

IN RETROSPECT

The Malay Archipelago

David Quammen re-enters the ‘Milky Way of land masses’ evoked by Alfred Russel Wallace’s masterpiece of biogeography.

ERNST MAYR LIB., MUS. COMPARATIVE ZOOLOGY, HARVARD UNIV.



Birds of paradise, as illustrated in *The Malay Archipelago* in 1869.

The Malay Archipelago: the land of the orang-utan, and the bird of paradise; a narrative of travel, with studies of man and nature

ALFRED RUSSEL WALLACE

Macmillan/Harper Brothers: first published 1869.

Alfred Russel Wallace was arguably the greatest field biologist of the nineteenth century. He played a leading part in the founding of both evolutionary theory and biogeography (see page 162). He was also, at times, a fine writer. The best of his literary side is on show in his 1869 classic, *The Malay Archipelago*, a wondrous book of travel and adventure that wears its deeper significance lightly.

The Malay Archipelago is the vast chain of islands stretching eastward from Sumatra for more than 6,000 kilometres. Most of it now falls within the sovereignties of Malaysia and Indonesia. In Wallace’s time, it was a world apart, a great Milky Way of land masses and seas and straits, little explored by Europeans, sparsely populated by peoples of diverse cultures, and harbouring countless species of unknown plant and animal in dense tropical forests. Some parts, such as the Aru group of islands, just off the coast of New Guinea, were almost legendary for their remoteness and biological riches. Wallace’s journeys throughout this region, sometimes by mail packet ship, sometimes in a trading vessel or a small outrigger canoe, were driven by a purpose: to collect animal specimens that might help to answer a scientific question. That question was: if species evolve, what is the mechanism?

“Wallace paid his expenses by selling specimens. So he collected series, not just samples.”

His Malay expedition began in 1854, five years before the publication of Charles Darwin’s *On the Origin of Species*. As he pursued his goal, oblivious to Darwin’s slow, secret theorizing, Wallace suffered one disadvantage that, paradoxically, proved advantageous: a need to earn money. He had no family funds such as those that had eased Darwin’s way. He paid his expenses by selling select specimens — pretty beetles and butterflies, stuffed birds and occasionally mammal skins — to museums and amateur collectors through an agent in London. So he had reason to take multiple individuals of the more striking species. He collected series, not just samples. From a riverside in southern Celebes he brought away “six good specimens” of *Papilio androcles*, a rare and beautiful swallowtail butterfly. And in Waigiou he harvested “twenty-four fine specimens” of *Paradisaea rubra*, the red bird of paradise. One effect of this redundant, commercial collecting was that he saw intraspecific variation laid out before him. ▶

▶ At a time when essentialist thinking (every species fixed, an ideal type) and the idea of special creations (each species shaped in a particular way by God) prevailed, even among most scientific naturalists, to see intraspecific variation was to take the first step towards an evolutionary theory.

Wallace travelled continuously among the islands for eight years (until 1862), and his mishaps, his hardships, his long weeks of illness and loneliness, his near-death scrapes, as well as his collecting and observing, are recounted in this vivid, encompassing book. From his jumping-off point in Singapore he proceeded to Borneo, then to Bali, then across the narrow, deep strait separating Bali from Lombok. From there he looped up to the Dutch entrepôt of Macassar, on the southwest tip of Celebes, which became his hub for further voyages around the archipelago.

He zigzagged from island to island — out east to the Aru cluster, back west to Sumatra and Java, up into the northern Moluccas and down south to Timor — depending

on his own whims and interests, but also on which way the monsoon winds were blowing or the next mail ship was heading. He took note of people and cultures, and recorded what intrigued him in his affable, bemused voice. Those parts of his book are as chatty and engaging as Samuel Pepys's diaries or James Boswell's *Life of Johnson*. And he continued his biological collections: an exquisite birdwing butterfly here, a giant beetle there, and, whenever possible, multiple specimens of each.

When the time came to combine his collecting notes with his anthropological observations and the diary of his many adventures and misadventures, Wallace recast the zig-zaggy muddle into a linear geographical structure. *The Malay Archipelago* proceeds from west to east — Singapore to Aru.

This structure serves well the implicit subjects of the book: biogeography and

evolution. Comparing the fauna of Bali with that of Lombok, Wallace laid the groundwork for one of his major contributions: the concept of Wallace's Line, the deep-water division between two major faunal realms, the Indo-Malay and the Australasian. Moving from Sumatra to Java to Borneo, he suggested (correctly) that those islands must once have been connected to the Asian mainland, resting high and dry on what we now call the Sunda Shelf. And he supported Darwinian theory, subtly rather than explicitly, by showing, in case after case, that only dispersal of ancestral forms, followed by isolation and evolutionary change, could explain the patterns of faunal distribution that his long labours had revealed. Special creations and essentialism just couldn't make sense of the Malay Archipelago.

But this book is not just a subtle compilation of data and argument. It is also a joyride through one of the wildest, most exotic, most remote regions of Earth that any Victorian explorer ever visited. There are deft descriptions of sensory experience, such as this on the taste of durian, the East's most notorious fruit: "A rich butter-like custard highly flavoured with almonds ... but intermingled with it come wafts of flavour that call to mind cream-cheese, onion-sauce, brown sherry, and other incongruities." There are moments of breathless drama, mostly involving not physical danger (about which Wallace tends to be matter-of-fact) but small, important triumphs, such as when he caught a new species of birdwing butterfly: "On taking it out of my net and opening the glorious wings, my heart began to beat violently, the blood rushed to my head, and I felt much more like fainting than I have done when in apprehension of immediate death." From his high points of discovery to his low points of misery, and through all the miles and days in between, Wallace is a companionable narrator with a dry wit, a keen eye, an inexhaustible curiosity and not a trace of self-pity. What more can readers ask of a literary and scientific traveller?

One thing is curiously absent from this long, brimming book. In recounting his stay on Ternate, in 1858, Wallace declined even to mention the world-altering idea — his theory of evolution — that he put to paper there. He could have crowed: in this place, at this time, I co-discovered natural selection. Evidently he didn't feel the need. By 1869, when *The Malay Archipelago* appeared, Wallace himself was a Darwinist. ■

David Quammen is a writer. His most recent book is *Spillover: Animal Infections and the Next Human Pandemic*. His 1996 book, *The Song of the Dodo: Island Biogeography in an Age of Extinctions*, features Wallace.
e-mail: quammen@imt.net



Some of the tropical butterflies that Wallace collected during his travels.

1855

**ON THE LAW WHICH HAS REGULATED THE
INTRODUCTION OF NEW SPECIES**

BY ALFRED RUSSEL WALLACE

FEBRUARY, 1855

(ALSO KNOWN AS THE SARAWACK LAW. ED.)

Wallace, A. R. 1855. On the law which has regulated the introduction of new species. *Annals and Magazine of Natural History*, 2nd Series, 16:184–196.

© 1998, Electronic Scholarly Publishing Project

<http://www.esp.org>

This electronic edition is made freely available for scholarly or educational purposes, provided that this copyright notice is included. The manuscript may not be reprinted or redistributed for commercial purposes without permission.

Wallace, A. R. 1855. On the law which has regulated the introduction of new species. *Annals and Magazine of Natural History*, 2nd Series, 16:184–196.

1855

**ON THE LAW WHICH HAS REGULATED THE
INTRODUCTION OF NEW SPECIES**

BY ALFRED RUSSEL WALLACE

FEBRUARY, 1855

(ALSO KNOWN AS THE SARAWACK LAW. ED.)

**Geographical Distribution dependent
on Geologic Changes.**

EVERY NATURALIST who has directed his attention to the subject of the geographical distribution of animals and plants, must have been interested in the singular facts which it presents. Many of these facts are quite different from what would have been anticipated, and have hitherto been considered as highly curious, but quite inexplicable. None of the explanations attempted from the time of Linnaeus are now considered at all satisfactory; none of them have given a cause sufficient to account for the facts known at the time, or comprehensive enough to include all the new facts which have since been, and are daily being added. Of late years, however, a great light has been thrown upon the subject by geological investigations, which have shown that the present state of the earth and of the organisms now inhabiting it, is but the last stage of a long and uninterrupted series of changes which it has undergone, and consequently, that to endeavour to explain and account for its present condition without any

© 1998, Electronic Scholarly Publishing Project

<http://www.esp.org>

This electronic edition is made freely available for scholarly or educational purposes, provided that this copyright notice is included. The manuscript may not be reprinted or redistributed for commercial purposes without permission.

reference to those changes (as has frequently been done) must lead to very imperfect and erroneous conclusions.

The facts proved by geology are briefly these: — That during an immense, but unknown period, the surface of the earth has undergone successive changes; land has sunk beneath the ocean, while fresh land has risen up from it; mountain chains have been elevated; islands have been formed into continents, and continents submerged till they have become islands; and these changes have taken place, not once merely, but perhaps hundreds, perhaps thousands of times: — That all these operations have been more or less continuous, but unequal in their progress, and during the whole series the organic life of the earth has undergone a corresponding alteration. This alteration also has been gradual, but complete; after a certain interval not a single species existing which had lived at the commencement of the period. This complete renewal of the forms of life also appears to have occurred several times: — That from the last of the geological epochs to the present or historical epoch, the change of organic life has been gradual: the first appearance of animals now existing can in many cases be traced, their numbers gradually increasing in the more recent formations, while other species continually die out and disappear, so that the present condition of the organic world is clearly derived by a natural process of gradual extinction and creation of species from that of the latest geological periods. We may therefore safely infer a like gradation and natural sequence from one geological epoch to another.

Now, taking this as a fair statement of the results of geological inquiry, we see that the present geographical distribution of life upon the earth must be the result of all the previous changes, both of the surface of the earth itself and of its inhabitants. Many causes, no doubt, have operated of which we must ever remain in ignorance, and we may, therefore, expect to find many details very difficult of explanation, and in attempting to give one, must allow ourselves to call into our service geological changes which it is highly probable may have occurred, though we have no direct evidence of their individual operation.

The great increase of our knowledge within the last twenty years, both of the present and past history of the organic world, has accumulated a body of facts which should afford a sufficient foundation for a comprehensive law embracing and explaining them all, and giving a direction to new researches. It is about ten years since the idea of such a law suggested itself to the writer of this essay, and he has since taken every opportunity of testing it by all the newly-ascertained facts with which he has become acquainted, or has

been able to observe himself. These have all served to convince him of the correctness of his hypothesis. Fully to enter into such a subject would occupy much space, and it is only in consequence of some views having been lately promulgated, he believes, in a wrong direction, that he now ventures to present his ideas to the public, with only such obvious illustrations of the arguments and results as occur to him in a place far removed from all means of reference and exact information.

A Law deduced from well-known Geographical and Geological Facts.

The following propositions in Organic Geography and Geology give the main facts on which the hypothesis is founded.

Geography

1. Large groups, such as classes and orders, are generally spread over the whole earth, while smaller ones, such as families and genera, are frequently confined to one portion, often to a very limited district.
2. In widely distributed families the genera are often limited in range; in widely distributed genera, well marked groups of species are peculiar to each geographical district.
3. When a group is confined to one district, and is rich in species, it is almost invariably the case that the most closely allied species are found in the same locality or in closely adjoining localities, and that therefore the natural sequence of the species by affinity is also geographical.
4. In countries of a similar climate, but separated by a wide sea or lofty mountains, the families, genera and species of the one are often represented by closely allied families, genera and species peculiar to the other.

Geology

5. The distribution of the organic world in time is very similar to its present distribution in space.
6. Most of the larger and some small groups extend through several geological periods.
7. In each period, however, there are peculiar groups, found nowhere else, and extending through one or several formations.

8. Species of one genus, or genera of one family occurring in the same geological time, are more closely allied than those separated in time.
9. As generally in geography no species or genus occurs in two very distant localities without being also found in intermediate places, so in geology the life of a species or genus has not been interrupted. In other words, no group or species has come into existence twice.
10. The following law may be deduced from these facts: — *Every species has come into existence coincident both in space and time with a pre-existing closely allied species.*

This law agrees with, explains and illustrates all the facts connected with the following branches of the subject: — 1st. The system of natural affinities. 2nd. The distribution of animals and plants in space. 3rd. The same in time, including all the phaenomena of representative groups, and those which Professor Forbes supposed to manifest polarity. 4th. The phaenomena of rudimentary organs. We will briefly endeavour to show its bearing upon each of these.

**The Form of a true system of
Classification determined by this Law.**

If the law above enunciated be true, it follows that the natural series of affinities will also represent the order in which the several species came into existence, each one having had for its immediate antitype a closely allied species existing at the time of its origin. It is evidently possible that two or three distinct species may have had a common antitype, and that each of these may again have become the antitypes from which other closely allied species were created. The effect of this would be, that so long as each species has had but one new species formed on its model, the line of affinities will be simple, and may be represented by placing the several species in direct succession in a straight line. But if two or more species have been independently formed on the plan of a common antitype, then the series of affinities will be compound, and can only be represented by a forked or many branched line. Now, all attempts at a Natural classification and arrangement of organic beings show, that both these plans have obtained in creation. Sometimes the series of affinities can be well represented for a space by a direct progression from species to species or from group to group, but it is generally found impossible so to continue. There constantly occur two or more modifications of an

organ or modifications of two distinct organs, leading us on to two distinct series of species, which at length differ so much from each other as to form distinct genera or families. These are the parallel series or representative groups of naturalists, and they often occur in different countries, or are found fossil in different formations. They are said to have an analogy to each other when they are so far removed from their common antitype as to differ in many important points of structure, while they still preserve a family resemblance. We thus see how difficult it is to determine in every case whether a given relation is an analogy or an affinity, for it is evident that as we go back along the parallel or divergent series, towards the common antitype, the analogy which existed between the two groups becomes an affinity. We are also made aware of the difficulty of arriving at a true classification, even in a small and perfect group; — in the actual state of nature it is almost impossible, the species being so numerous and the modifications of form and structure so varied, arising probably from the immense number of species which have served as antitype for the existing species, and thus produced a complicated branching of the lines of affinity, as intricate as the twigs of a gnarled oak or the vascular system of the human body. Again, if we consider that we have only fragments of this vast system, the stem and main branches being represented by extinct species of which we have no knowledge, while a vast mass of limbs and boughs and minute twigs and scattered leaves is what we have to place in order, and determine the true position each originally occupied with regard to the others, the whole difficulty of the true Natural System of classification becomes apparent to us.

We shall thus find ourselves obliged to reject all those systems of classification which arrange species or groups in circles, as well as those which fix a definite number for the divisions of each group. The latter class have been very generally rejected by naturalists, as contrary to nature, notwithstanding the ability with which they have been advocated; but the circular system of affinities seems to have obtained a deeper hold, many eminent naturalists having to some extent adopted it. We have, however, never been able to find a case in which the circle has been closed by a direct and close affinity. In most cases a palpable analogy has been substituted, in others the affinity is very obscure or altogether doubtful. The complicated branching of the lines of affinities in extensive groups must also afford great facilities for giving a show of probability to any such purely artificial arrangements. Their death-blow was given by the admirable paper of the lamented Mr. Strickland, published in the “Annals of Natural

History," in which he so cleverly showed the true synthetical method of discovering the Natural System.

Geographical Distribution of Organisms.

If we now consider the geographical distribution of animals and plants upon the earth, we shall find all the facts beautifully in accordance with, and readily explained by, the present hypothesis. A country having species, genera, and whole families peculiar to it, will be the necessary result of its having been isolated for a long period, sufficient for many series of species to have been created on the type of pre-existing ones, which, as well as many of the earlier-formed species, have become extinct, and thus made the groups appear isolated. If in any case the antitype had an extensive range, two or more groups of species might have been formed, each varying from it in a different manner, and thus producing several representative or analogous groups. The Sylviadae of Europe and the Sylvicolidae of North America, the Heliconidae of South America and the Euploeas of the East, the group of Trogons inhabiting Asia, and that peculiar to South America, are examples that may be accounted for in this manner.

Such phaenomena as are exhibited by the Galapagos Islands, which contain little groups of plants and animals peculiar to themselves, but most nearly allied to those of South America, have not hitherto received any, even a conjectural explanation. The Galapagos are a volcanic group of high antiquity, and have probably never been more closely connected with the continent than they are at present. They must have been first peopled, like other newly-formed islands, by the action of winds and currents, and at a period sufficiently remote to have had the original species die out, and the modified prototypes only remain. In the same way we can account for the separate islands having each their peculiar species, either on the supposition that the same original emigration peopled the whole of the islands with the same species from which differently modified prototypes were created, or that the islands were successively peopled from each other, but that new species have been created in each on the plan of the pre-existing ones. St. Helena is a similar case of a very ancient island having obtained an entirely peculiar, though limited, flora. On the other hand, no example is known of an island which can be proved geologically to be of very recent origin (late in the Tertiary, for instance), and yet possess generic or family groups, or even many species peculiar to itself.

When a range of mountains has attained a great elevation, and has so remained during a long geological period, the species of the two sides at and near their bases will be often very different, representative species of some genera occurring, and even whole genera being peculiar to one side, as is remarkably seen in the case of the Andes and Rocky Mountains. A similar phaenomena occurs when an island has been separated from a continent at a very early period. The shallow sea between the Peninsula of Malacca, Java, Sumatra and Borneo was probably a continent or large island at an early epoch, and may have become submerged as the volcanic ranges of Java and Sumatra were elevated. The organic results we see in the very considerable number of species of animals common to some or all of these countries, while at the same time a number of closely allied representative species exist peculiar to each, showing that a considerable period has elapsed since their separation. The facts of geographical distribution and of geology may thus mutually explain each other in doubtful cases, should the principles here advocated be clearly established.

In all those cases in which an island has been separated from a continent, or raised by volcanic or coralline action from the sea, or in which a mountain-chain has been elevated in a recent geological epoch, the phaenomena of peculiar groups or even of single representative species will not exist. Our own island is an example of this, its separation from the continent being geologically very recent, and we have consequently scarcely a species which is peculiar to it; while the Alpine range, one of the most recent mountain elevations, separates faunas and floras which scarcely differ more than may be due to climate and latitude alone.

The series of facts alluded to in Proposition (3), of closely allied species in rich groups being found geographically near each other, is most striking and important. Mr. Lovell Reeve has well exemplified it in his able and interesting paper on the Distribution of the *Bulimi*. It is also seen in the Hummingbirds and Toucans, little groups of two or three closely allied species being often found in the same or closely adjoining districts, as we have had the good fortune of personally verifying. Fishes give evidence of a similar kind: each great river has its peculiar genera, and in more extensive genera its groups of closely allied species. But it is the same throughout Nature; every class and order of animals will contribute similar facts. Hitherto no attempt has been made to explain these singular phaenomena, or to show how they have arisen. Why are the genera of Palms and of Orchids in almost every case confined to one hemisphere? Why are the closely allied

species of brownbacked Trogons all found in the East, and the green-backed in the West? Why are the Macaws and the Cockatoos similarly restricted? Insects furnish a countless number of analogous examples; — the Goliath of Africa, the Ornithopterae of the Indian Islands, the Heliconidae of South America, the Danaidae of the East, and in all, the most closely allied species found in geographical proximity. The question forces itself upon every thinking mind, — why are these things so? They could not be as they are had no law regulated their creation and dispersion. The law here enunciated not merely explains, but necessitates the facts we see to exist, while the vast and long-continued geological changes of the earth readily account for the exceptions and apparent discrepancies that here and there occur. The writer's object in putting forward his views in the present imperfect manner is to submit them to the test of other minds, and to be made aware of all the facts supposed to be inconsistent with them. As his hypothesis is one which claims acceptance solely as explaining and connecting facts which exist in nature, he expects facts alone to be brought to disprove it, not *à priori* arguments against its probability.

Geological Distribution of the Forms of Life.

The phaenomena of geological distribution are exactly analogous to those of geography. Closely allied species are found associated in the same beds, and the change from species to species appears to have been as gradual in time as in space. Geology, however, furnishes us with positive proof of the extinction and production of species, though it does not inform us how either has taken place. The extinction of species, however, offers but little difficulty, and the *modus operandi* has been well illustrated by Sir C. Lyell in his admirable "principles." Geological changes, however gradual, must occasionally have modified external conditions to such an extent as to have rendered the existence of certain species impossible. The extinction would in most cases be effected by a gradual dying-out, but in some instances there might have been a sudden destruction of a species of limited range. To discover how the extinct species have from time to time been replaced by new ones down to the very latest geological period, is the most difficult, and at the same time the most interesting problem in the natural history of the earth. The present inquiry, which seeks to eliminate from known facts a law which has determined, to a certain degree, what species could and did appear at a given epoch, may, it is hoped, be considered as one step in the right direction towards a complete solution of it.

**High Organization of very ancient
Animals consistent with this Law.**

Much discussion has of late years taken place on the question, whether the succession of life upon the globe has been from a lower to a higher degree of organization. The admitted facts seem to show that there has been a general, but not a detailed progression. Mollusca and Radiata existed before Vertebrata, and the progression from Fishes to Reptiles and Mammalia, and also from the lower mammals to the higher, is indisputable. On the other hand, it is said that the Mollusca and Radiata of the very earliest periods were more highly organized than the great mass of those now existing, and that the very first fishes that have been discovered are by no means the lowest organised of the class. Now it is believed the present hypothesis will harmonize with all these facts, and in a great measure serve to explain them; for though it may appear to some readers essentially a theory of progression, it is in reality only one of gradual change. It is, however, by no means difficult to show that a real progression in the scale of organization is perfectly consistent with all the appearances, and even with apparent retrogression, should such occur.

Returning to the analogy of a branching tree, as the best mode of representing the natural arrangement of species and their successive creation, let us suppose that at an early geological epoch any group (say a class of the Mollusca) has attained to a great richness of species and a high organization. Now let this great branch of allied species, by geological mutations, be completely or partially destroyed. Subsequently a new branch springs from the same trunk, that is to say, new species are successively created, having for their antitypes the same lower organized species which had served as the antitypes for the former group, but which have survived the modified conditions which destroyed it. This new group being subject to these altered conditions, has modifications of structure and organization given to it, and becomes the representative group of the former one in another geological formation. It may, however, happen, that though later in time, the new series of species may never attain to so high a degree of organization as those preceding it, but in its turn become extinct, and give place to yet another modification from the same root, which may be of higher or lower organization, more or less numerous in species, and more or less varied in form and structure than either of those which preceded it. Again, each of these groups may not have become totally extinct, but may have left a few species, the modified prototypes of which have existed in each succeeding period, a faint

memorial of their former grandeur and luxuriance. Thus every case of apparent retrogression may be in reality a progress, though an interrupted one: when some monarch of the forest loses a limb, it may be replaced by a feeble and sickly substitute. The foregoing remarks appear to apply to the case of the Mollusca, which, at a very early period, had reached a high organization and a great development of forms and species in the testaceous Cephalopoda. In each succeeding age modified species and genera replaced the former ones which had become extinct, and as we approach the present aera, but few and small representatives of the group remain, while the Gasteropods and Bivalves have acquired an immense preponderance. In the long series of changes the earth has undergone, the process of peopling it with organic beings has been continually going on, and whenever any of the higher groups have become nearly or quite extinct, the lower forms which have better resisted the modified physical conditions have served as the antitypes on which to found the new races. In this manner alone, it is believed, can the representative groups at successive periods, and the rising and fallings in the scale of organization, be in every case explained.

Objections to Forbes' Theory of Polarity.

The hypothesis of polarity, recently put forward by Professor Edward Forbes to account for the abundance of generic forms at a very early period and at present, while in the intermediate epochs there is a gradual diminution and impoverishment, till the minimum occurred at the confines of the Palaeozoic and Secondary epochs, appears to us quite unnecessary, as the facts may be readily accounted for on the principles already laid down. Between the Palaeozoic and Neozoic periods of Professor Forbes, there is scarcely a species in common, and the greater part of the genera and families also disappear to be replaced by new ones. It is almost universally admitted that such a change in the organic world must have occupied a vast period of time. Of this interval we have no record; probably because the whole area of the early formations now exposed to our researches was elevated at the end of the Palaeozoic period, and remained so through the interval required for the organic changes which resulted in the fauna and flora of the Secondary period. The records of this interval are buried beneath the ocean which covers three-fourths of the globe. Now it appears highly probable that a long period of quiescence or stability in the physical conditions of a district would be most favourable to the existence of organic life in the

greatest abundance, both as regards individuals and also as to variety of species and generic group, just as we now find that the places best adapted to the rapid growth and increase of individuals also contain the greatest profusion of species and the greatest variety of forms, — the tropics in comparison with the temperate and arctic regions. On the other hand, it seems no less probable that a change in the physical conditions of a district, even small in amount if rapid, or even gradual if to a great amount, would be highly unfavourable to the existence of individuals, might cause the extinction of many species, and would probably be equally unfavourable to the creation of new ones. In this too we may find an analogy with the present state of our earth, for it has been shown to be the violent extremes and rapid changes of physical conditions, rather than the actual mean state in the temperate and frigid zones, which renders them less prolific than the tropical regions, as exemplified by the great distance beyond the tropics to which tropical forms penetrate when the climate is equable, and also by the richness in species and forms of tropical mountain regions which principally differ from the temperate zone in the uniformity of their climate. However this may be, it seems a fair assumption that during a period of geological repose the new species which we know to have been created would have appeared, that the creations would then exceed in number the extinctions, and therefore the number of species would increase. In a period of geological activity, on the other hand, it seems probable that the extinctions might exceed the creations, and the number of species consequently diminish. That such effects did take place in connexion with the causes to which we have imputed them, is shown in the case of the Coal formation, the faults and contortions of which show a period of great activity and violent convulsions, and it is in the formation immediately succeeding this that the poverty of forms of life is most apparent. We have then only to suppose a long period of somewhat similar action during the vast unknown interval at the termination of the Palaeozoic period, and then a decreasing violence or rapidity through the Secondary period, to allow for the gradual repopulation of the earth with varied forms, and the whole of the facts are explained. We thus have a clue to the increase of the forms of life during certain periods, and their decrease during others, without recourse to any causes but these we know to have existed, and to effects fairly deducible from them. The precise manner in which the geological changes of the early formations were effected is so extremely obscure, that when we can explain important facts by a retardation at one time and an acceleration at another of a process which we know from its nature and from observation to have

been unequal, — a cause so simple may surely be preferred to one so obscure and hypothetical as polarity.

I would also venture to suggest some reasons against the very nature of the theory of Professor Forbes. Our knowledge of the organic world during any geological epoch is necessarily very imperfect. Looking at the vast numbers of species and groups that have been discovered by geologists, this may be doubted; but we should compare their numbers not merely with those that now exist upon the earth, but with a far larger amount. We have no reason for believing that the number of species on the earth at any former period was much less than at present; at all events the aquatic portion, with which geologists have most acquaintance, was probably often as great or greater. Now we know that there have been many complete changes of species; new sets of organisms have many times been introduced in place of old ones which have become extinct, so that the total amount which have existed on the earth from the earliest geological period must have borne about the same proportion to those now living, as the whole human race who have lived and died upon the earth, to the population at the present time. Again, at each epoch, the whole earth was no doubt, as now, more or less the theatre of life, and as the successive generations of each species died, their exuviae and preservable parts would be deposited over every portion of the then existing seas and oceans, which we have reason for supposing to have been more, rather than less, extensive than at present. In order then to understand our possible knowledge of the early world and its inhabitants, we must compare, not the area of the whole field of our geological researches with the earth's surface, but the area of the examined portion of each formation separately with the whole earth. For example, during the Silurian period all the earth was Silurian, and animals were living and dying, and depositing their remains more or less over the whole area of the globe, and they were probably (the species at least) nearly as varied in different latitudes and longitudes as at present. What proportion do the Silurian districts bear to the whole surface of the globe, land and sea (for far more extensive Silurian districts probably exist beneath the ocean than above it), and what portion of the known Silurian districts has been actually examined for fossils? Would the area of rock actually laid open to the eye be the thousandth or the ten-thousandth part of the earth's surface? Ask the same question with regard to the Oolite or the Chalk, or even to particular beds of these when they differ considerably in their fossils, and you may then get some notion of how small a portion of the whole we know.

But yet more important is the probability, nay almost the certainty, that whole formations containing the records of vast geological periods are entirely buried beneath the ocean, and for ever beyond our reach. Most of the gaps in the geological series may thus be filled up, and vast numbers of unknown and unimaginable animals, which might help to elucidate the affinities of the numerous isolated groups which are a perpetual puzzle to the zoologist, may there be buried, until future revolutions may raise them in their turn above the waters, to afford materials for the study of whatever race of intelligent beings may then have succeeded us. These considerations must lead us to the conclusion, that our knowledge of the whole series of the former inhabitants of the earth is necessarily most imperfect and fragmentary, — as much so as our knowledge of the present organic world would be, were we forced to make our collections and observations only in spots equally limited in area and in number with those actually laid open for the collection of fossils. Now, the hypothesis of Professor Forbes is essentially one that assumes to a great extent the completeness of our knowledge of the whole series of organic beings which have existed on the earth. This appears to be a fatal objection to it, independently of all other considerations. It may be said that the same objections exist against every theory on such a subject, but this is not necessarily the case. The hypothesis put forward in this paper depends in no degree upon the completeness of our knowledge of the former condition of the organic world, but takes what facts we have as fragments of a vast whole, and deduces from them something of the nature and proportions of that whole which we can never know in detail. It is founded upon isolated groups of facts, recognizes their isolation, and endeavours to deduce from them the nature of the intervening portions.

Rudimentary Organs.

Another important series of facts, quite in accordance with, and even necessary deductions from, the law now developed, are those of rudimentary organs. That these really do exist, and in most cases have no special function in the animal economy, is admitted by the first authorities in comparative anatomy. The minute limbs hidden beneath the skin in many of the snake-like lizards, the anal hooks of the boa constrictor, the complete series of jointed finger-bones in the paddle of the Manatus and whale, are a few of the most familiar instances. In botany a similar class of facts has long been recognised. Abortive stamens, rudimentary floral envelopes and undeveloped carpels, are of

the most frequent occurrence. To every thoughtful naturalist the question must arise, What are these for? What have they to do with the great laws of creation? Do they not teach us something of the system of Nature? If each species has been created independently, and without any necessary relations with pre-existing species, what do these rudiments, these apparent imperfections mean? There must be a cause for them; they must be the necessary results of some great natural law. Now, if, as it has been endeavoured to be shown, the great law which has regulated the peopling of the earth with animal and vegetable life is, that every change shall be gradual; that no new creature shall be formed widely differing from anything before existing; that in this, as in everything else in Nature, there shall be gradation and harmony, — then these rudimentary organs are necessary, and are an essential part of the system of Nature. Ere the higher Vertebrata were formed, for instance, many steps were required, and many organs had to undergo modifications from the rudimental condition in which only they had as yet existed. We still see remaining an antitypal sketch of a wing adapted for flight in the scaly flapper of the penguin, and limbs first concealed beneath the skin, and then weakly protruding from it, were the necessary gradations before others should be formed fully adapted for locomotion. Many more of these modifications should we behold, and more complete series of them, had we a view of all the forms which have ceased to live. The great gaps that exist between fishes, reptiles, birds, and mammals would then, no doubt, be softened down by intermediate groups, and the whole organic world would be seen to be an unbroken and harmonious system.

Conclusion.

It has now been shown, though most briefly and imperfectly, how the law that “*Every species has come into existence coincident both in time and space with a pre-existing closely allied species,*” connects together and renders intelligible a vast number of independent and hitherto unexplained facts. The natural system of arrangement of organic beings, their geographical distribution, their geological sequence, the phaenomena of representative and substituted groups in all their modifications, and the most singular peculiarities of anatomical structure, are all explained and illustrated by it, in perfect accordance with the vast mass of facts which the researches of modern naturalists have brought together, and, it is believed, not materially opposed to any of them. It also claims a superiority over

previous hypotheses, on the ground that it not merely explains, but necessitates what exists. Granted the law, and many of the most important facts in Nature could not have been otherwise, but are almost as necessary deductions from it, as are the elliptic orbits of the planets from the law of gravitation.

— THE END —

ON THE TENDENCY OF VARIETIES TO DEPART INDEFINITELY FROM THE ORIGINAL TYPE

by Alfred Russel Wallace written at Ternate, February, 1858

Instability of Varieties supposed to prove the permanent distinctness of Species

[Paper presented along with Darwin's "Abstract" on natural selection (actually the first edition of "The Origin of Species" at the meeting of the Linnaean Society which established Darwin's primacy to the theory of natural selection)].

ONE of the strongest arguments which have been adduced to prove the original and permanent distinctness of species is, that varieties produced in a state of domesticity are more or less unstable, and often have a tendency, if left to themselves, to return to the normal form of the parent species; and this instability is considered to be a distinctive peculiarity of all varieties, even of those occurring among wild animals in a state of nature, and to constitute a provision for preserving unchanged the originally created distinct species.

In the absence of scarcity of facts and observations as to varieties occurring among wild animals, this argument has had great weight with naturalists, and has led to a very general and somewhat prejudiced belief in the stability of species. Equally general, however, is the belief in what are called "permanent or true varieties,"- races of animals which continually propagate their like, but which differ so slightly (although constantly) from some other race, that the one is considered to be a variety of the other. Which is the variety and which the original species, there is generally no means of determining, except in those rare cases in which the one race has been known to produce an offspring unlike itself and resembling the other. This, however, would seem quite incompatible with the "permanent invariability of species," but the difficulty is overcome by assuming that such varieties have strict limits, and can never again vary further from the original type, although they may return to it, which, from the analogy of the domesticated animals, is considered to be highly probable, if not certainly proved.

It will be observed that this argument rests entirely on the assumption, that varieties occurring in a state of nature are in all respects analogous to or even identical with those of domestic animals, and are governed by the same laws as regards their permanence or further variation. But it is the object of the present paper to show that this assumption is altogether false, that there is a general principle in nature which will cause many varieties to survive the parent species, and to give rise to successive variations departing further and further from the original type, and which also produces, in domesticated animals, the tendency of varieties to return to the parent form.

The Struggle for Existence.

The life of wild animals is a struggle for existence. The full exertion of all their faculties and all their energies is required to preserve their own existence and provide for that of their infant offspring. The possibility of procuring food during the least favourable seasons, and of escaping the attacks of their most dangerous enemies, are the primary conditions which determine the existence both of individuals and of entire species. These conditions will also determine the population of a species; and by a careful consideration of all the circumstances we may be enabled to comprehend, and in some degree to explain, what at first sight appears so inexplicable- the excessive abundance of some species, while others closely allied to them are very rare.

The Law of Population of Species.

The general proportion that must obtain between certain groups of animals is readily seen. Large animals cannot be so abundant as small ones; the carnivora must be less numerous than the herbivora; eagles and lions can never be so plentiful as pigeons and antelopes; the wild asses of the Tartarian deserts cannot equal in numbers the horses of the more luxuriant prairies and pampas of America. The greater or less fecundity of an animal is often considered to be one of the chief causes of its abundance or scarcity; but a consideration of the facts will show us that it really has little or nothing to do with the matter. Even the least prolific of animals would increase rapidly if unchecked, whereas it is evident that the animal population of the globe must be stationary, or perhaps, through the influence of man, decreasing. Fluctuations there may be; but permanent increase, except in restricted localities, is almost impossible. For example, our own observation must convince us that birds do not go on increasing every year in a geometrical ratio, as they would do, were there not some powerful check to their natural increase. Very few birds produce less than two young ones each year, while many have six, eight, or ten; four will certainly be below the average; and if we suppose that each pair produce young only four times in their life, that will also be below the average, supposing them not to die either by violence or want of food. Yet at this rate how tremendous would be the increase in a few years from a single pair! A simple calculation will show that in fifteen years each pair of birds would have increased to nearly ten millions! whereas we have no reason to believe that the number of the birds of any country increases at all in fifteen or in one hundred and fifty years. With such powers of increase the population must have reached its limits, and have become stationary, in a very few years after the origin of each species. It is evident, therefore, that each year an immense number of birds must perish- as many in fact as are born; and as on the lowest calculation the progeny are each year twice as numerous as their parents, it follows that, whatever

be the average number of individuals existing in any given country, twice that number must perish annually,- a striking result, but one which seems at least highly probable, and is perhaps under rather than over the truth. It would therefore appear that, as far as the continuance of the species and the keeping up the average number of individuals are concerned, large broods are superfluous. On the average all above one become food for hawks and kites, wild cats and weasels, or perish of cold and hunger as winter comes on. This is strikingly proved by the case of particular species; for we find that their abundance in individuals bears no relation whatever to their fertility in producing offspring. Perhaps the most remarkable instance of an immense bird population is that of the passenger pigeon of the United States, which lays only one, or at most two eggs, and is said to rear generally but one young one. Why is this bird so extraordinarily abundant, while others producing two or three times as many young are much less plentiful? The explanation is not difficult. The food most congenial to this species, and on which it thrives best, is abundantly distributed over a very extensive region, offering such difference of soil and climate, that in one part or another of the area the supply never fails. The bird is capable of a very rapid and long-continued flight, so that it can pass without fatigue over the whole of the district it inhabits, and as soon as the supply of food begins to fail in one place is able to discover a fresh feeding-ground. This example strikingly shows us that the procuring a constant supply of wholesome food is almost the sole condition requisite for ensuring the rapid increase of a given species, since neither the limited fecundity, nor the unrestrained attacks of birds of prey and of man are here sufficient to check it. In no other birds are these peculiar circumstances so strikingly combined. Either their food is more liable to failure, or they have not sufficient power of wing to search for it over an extensive area, or during some season of the year it becomes very scarce, and less wholesome substitutes have to be found; and thus, though more fertile in offspring, they can never increase beyond the supply of food in the least favourable seasons. Many birds can only exist by migrating, when their food becomes scarce, to regions possessing a milder, or at least a different climate, though, as these migrating birds are seldom excessively abundant, it is evident that the countries they visit are still deficient in a constant and abundant supply of wholesome food. Those whose organization does not permit them to migrate when their food becomes periodically scarce, can never attain a large population. This is probably the reason why woodpeckers are scarce with us, while in the tropics they are among the most abundant of solitary birds. Thus the house sparrow is more abundant than the redbreast, because its food is more constant and plentiful,- seeds of grasses being preserved during the winter, and our farm-yards and stubble-fields furnishing an almost inexhaustible supply. Why, as a general rule, are aquatic, and especially sea birds, very numerous in individuals? Not because they are more prolific than others, generally the contrary; but because their food never fails, the sea-shores and river-banks daily swarming with a

fresh supply of small mollusca and crustacea. Exactly the same laws will apply to mammals. Wild cats are prolific and have few enemies; why then are they never as abundant as rabbits? The only intelligible answer is, that their supply of food is more precarious. It appears evident, therefore, that so long as a country remains physically unchanged, the numbers of its animal population cannot materially increase. If one species does so, some others requiring the same kind of food much diminish in proportion. The numbers that die annually must be immense; and as the individual existence of each animal depends upon itself, those that die must be the weakest- the very young, the aged, and the diseased,- while those that prolong their existence can only be the most perfect in health and vigour- those who are best able to obtain food regularly, and avoid their numerous enemies. It is, as we commenced by remarking, "a struggle for existence," in which the weakest and least perfectly organized must always succumb.

The Abundance or Rarity of a Species dependent upon its more or less perfect Adaptation to the Conditions of Existence.

It seems evident that what takes place among the individuals of a species must also occur among the several allied species of a group,- viz., that those which are best adapted to obtain a regular supply of food, and to defend themselves against the attacks of their enemies and the vicissitudes of the seasons, must necessarily obtain and preserve a superiority in population; while those species which from some defect of power or organization are the least capable of counteracting the vicissitudes of food, supply, &c., must diminish in numbers, and, in extreme cases, become altogether extinct. Between these extremes the species will present various degrees of capacity for ensuring the means of preserving life; and it is thus we account for the abundance or rarity of species. Our ignorance will generally prevent us from accurately tracing the effects to their causes; but could we become perfectly acquainted with the organization and habits of the various species of animals, and could we measure the capacity of each for performing the different acts necessary to its safety and existence under all the varying circumstances by which it is surrounded, we might be able even to calculate the proportionate abundance of individuals which is the necessary result.

If now we have succeeded in establishing these two points- 1st, that the animal population of a country is generally stationary, being kept down by a periodical deficiency of food, and other checks; and, 2nd, that the comparative abundance or scarcity of the individuals of the several species is entirely due to their organization and resulting habits, which, rendering it more difficult to procure a regular supply of food and to provide for their personal safety in some cases than in others, can only be balanced by a difference in the population which have to exist in a given area- we shall be in a condition to proceed to the consideration of varieties, to which the

preceding remarks have a direct and very important application.

Useful Variations will tend to Increase; useless or hurtful Variations to Diminish.

Most or perhaps all the variations from the typical form of a species must have some definite effect, however slight, on the habits or capacities of the individuals. Even a change of colour might, by rendering them more or less distinguishable, affect their safety; a greater or less development of hair might modify their habits. More important changes, such as an increase in the power or dimensions of the limbs or any of the external organs, would more or less affect their mode of procuring food or the range of country which they inhabit. It is also evident that most changes would affect, either favourably or adversely, the powers of prolonging existence. An antelope with shorter or weaker legs must necessarily suffer more from the attacks of the feline carnivora; the passenger pigeon with less powerful wings would sooner or later be affected in its powers of procuring a regular supply of food; and in both cases the result must necessarily be a diminution of the population of the modified species. If, on the other hand, any species should produce a variety having slightly increased powers of preserving existence, that variety must inevitably in time acquire a superiority in numbers. These results must follow as surely as old age, intemperance, or scarcity of food produce an increased mortality. In both cases there may be many individual exceptions; but on the average the rule will invariably be found to hold good. All varieties will therefore fall into two classes- those which under the same conditions would never reach the population of the parent species, and those which would in time obtain and keep a numerical superiority. Now, let some alteration of physical conditions occur in the district- a long period of drought, a destruction of vegetation by locusts, the irruption of some new carnivorous animal seeking "pastures new"- any change in fact tending to render existence more difficult to the species in question, and taxing its utmost powers to avoid complete extermination; it is evident that, of all the individuals composing the species, those forming the least numerous and most feebly organized variety would suffer first, and, were the pressure severe, must soon become extinct. The same causes continuing in action, the parent species would next suffer, would gradually diminish in numbers, and with a recurrence of similar unfavourable conditions might also become extinct. The superior variety would then alone remain, and on a return to favourable circumstances would rapidly increase in numbers and occupy the place of the extinct species and variety.

Superior Varieties will ultimately Extirpate the original Species.

The variety would now have replaced the species, of which it would be a more perfectly developed and more highly organized form. It would be in all

respects better adapted to secure its safety, and to prolong its individual existence and that of the race. Such a variety could not return to the original form; for that form is an inferior one, and could never compete with it for existence. Granted, therefore, a "tendency" to reproduce the original type of the species, still the variety must ever remain preponderant in numbers, and under adverse physical conditions again alone survive. But this new, improved, and populous race might itself, in course of time, give rise to new varieties, exhibiting several diverging modifications of form, any of which, tending to increase the facilities for preserving existence, must by the same general law, in their turn become predominant. Here, then, we have progression and continued divergence deduced from the general laws which regulate the existence of animals in a state of nature, and from the undisputed fact that varieties do frequently occur. It is not, however, contended that this result would be invariable; a change of physical conditions in the district might at times materially modify it, rendering the race which had been the most capable of supporting existence under the former conditions now the least so, and even causing the extinction of the newer and, for a time, superior race, while the old or parent species and its first inferior varieties continued to flourish. Variations in unimportant parts might also occur, having no perceptible effect on the life-preserving powers; and the varieties so furnished might run a course parallel with the parent species, either giving rise to further variations or returning to the former type. All we argue for is, that certain varieties have a tendency to maintain their existence longer than the original species, and this tendency must make itself felt; for though the doctrine of chances or averages can never be trusted to on a limited scale, yet, if applied to high numbers, the results come nearer to what theory demands, and, as we approach to an infinity of examples, become strictly accurate. Now the scale on which nature works is so vast- the numbers of individuals and periods of time with which she deals approach so near to infinity, that any cause, however slight, and however liable to be veiled and counteracted by accidental circumstances, must in the end produce its full legitimate results.

The Partial Reversion of Domesticated Varieties explained.

Let us now turn to domesticated animals, and inquire how varieties produced among them are affected by the principles here enunciated. The essential difference in the condition of wild and domestic animals is this,- that among the former, their well-being and very existence depend upon the full exercise and healthy condition of all their senses and physical powers, whereas, among the latter, these are only partially exercised, and in some cases are absolutely unused. A wild animal has to search, and often to labour, for every mouthful of food- to exercise sight, hearing, and smell in seeking it, and in avoiding dangers, in procuring shelter from the inclemency of the seasons, and in providing for the subsistence and safety

of its offspring. There is no muscle of its body that is not called into daily and hourly activity; there is no sense or faculty that is not strengthened by continual exercise. The domestic animal, on the other hand, has food provided for it, is sheltered, and often confined, to guard it against the vicissitudes of the seasons, is carefully secured from the attacks of its natural enemies, and seldom even rears its young without human assistance. Half of its senses and faculties are quite useless; and the other half are but occasionally called into feeble exercise, while even its muscular system is only irregularly called into action.

Now when a variety of such an animal occurs, having increased power or capacity in any organ or sense, such increase is totally useless, is never called into action, and may even exist without the animal ever becoming aware of it. In the wild animal, on the contrary, all its faculties and powers being brought into full action for the necessities of existence, any increase becomes immediately available, is strengthened by exercise, and must even slightly modify the food, the habits, and the whole economy of the race. It creates as it were a new animal, one of superior powers, and which will necessarily increase in numbers and outlive those inferior to it.

Again, in the domesticated animal all variations have an equal chance of continuance; and those which would decidedly render a wild animal unable to compete with its fellows and continue its existence are no disadvantage whatever in a state of domesticity. Our quickly fattening pigs, short-legged sheep, pouter pigeons, and poodle dogs could never have come into existence in a state of nature, because the very first step towards such inferior forms would have led to the rapid extinction of the race; still less could they now exist in competition with their wild allies. The great speed but slight endurance of the race horse, the unwielding strength of the ploughman's team, would both be useless in a state of nature. If turned wild on the pampas, such animals would probably soon become extinct, or under favorable circumstances might each lose those extreme qualities which would never be called into action, and in a few generations would revert to a common type, which must be that in which the various powers and faculties are so proportioned to each other as to be best adapted to procure food and secure safety,- that in which by the full exercise of every part of his organization the animal can alone continue to live. Domestic varieties, when turned wild, must return to something near the type of the original wild stock, or become altogether extinct.*

*That is, they will vary, and the variations which tend to adapt them to the wild state, and therefore approximate them to wild animals, will be preserved. Those individuals which do not vary sufficiently will perish.

Lamarck's Hypothesis very different from that now advanced.

We see, then, that no inferences as to varieties in a state of nature can be deduced from the observation of those occurring among domestic animals. The two are so much opposed to each other in every circumstance of their existence, that what applies to the one is almost sure not to apply to the other. Domestic animals are abnormal, irregular, artificial; they are subject to varieties which never occur and never can occur in a state of nature; their very existence depends altogether on human care: so far are many of them removed from that just proportion of faculties, that true balance of organization, by means of which alone an animal left to its own resources can preserve its existence and continue its race.

The hypothesis of Lamarck- that progressive changes in species have been produced by the attempts of animals to increase the development of their own organs, and thus modify their structure and habits- has been repeatedly and easily refuted by all writers on the subject of varieties and species, and it seems to have been considered that when this was done the whole question has been finally settled; but the view here developed renders such an hypothesis quite unnecessary, by showing that similar results must be produced by the action of principles constantly at work in nature. The powerful retractile talons of the falcon- and the cat-tribes have not been produced or increased by the volition of those animals; but among the different varieties which occurred in the earlier and less highly organized forms of these groups, those always survived longest which had the greatest facilities for seizing their prey. Neither did the giraffe acquire its long neck by desiring to reach the foliage of the more lofty shrubs, and constantly stretching its neck for the purpose, but because any varieties which occurred among its antitypes with a longer neck than usual at once secured a fresh range of pasture over the same ground as their shorter-necked companions, and on the first scarcity of food were thereby enabled to outlive them. Even the peculiar colours of many animals, especially insects, so closely resembling the soil or the leaves or the trunks on which they habitually reside, are explained on the same principle; for though in the course of ages varieties of many tints may have occurred, yet those races having colours best adapted to concealment from their enemies would inevitably survive the longest. We have also here an acting cause to account for that balance so often observed in nature,- a deficiency in one set of organs always being compensated by an increased development of some others- powerful wings accompanying weak feet, or great velocity making up for the absence of defensive weapons; for it has been shown that all varieties in which an unbalanced deficiency occurred could not long continue their existence. The action of this principle is exactly like that of the centrifugal governor of the steam engine, which checks and corrects any irregularities almost before they become evident; and in like manner no unbalanced deficiency in the animal kingdom can ever reach any conspicuous

magnitude, because it would make itself felt at the very first step, by rendering existence difficult and extinction almost sure to follow. An origin such as is here advocated will also agree with the peculiar character of the modifications of form and structure which obtain in organized beings- the many lines of divergence from a central type, the increasing efficiency and power of a particular organ through a succession of allied species, and the remarkable persistence of unimportant parts such as colour, texture of plumage and hair, form of horns or crests, through a series of species differing considerably in more essential characters. It also furnishes us with a reason for that "more specialized structure" which Professor Owen states to be a characteristic of recent compared with extinct forms, and which would evidently be the result of the progressive modification of any organ applied to a special purpose in the animal economy.

Conclusion.

We believe we have now shown that there is a tendency in nature to the continued progression of certain classes of varieties further and further from the original type- a progression to which there appears no reason to assign any definite limits- and that the same principle which produces this result in a state of nature will also explain why domestic varieties have a tendency to revert to the original type. This progression, by minute steps, in various directions, but always checked and balanced by the necessary conditions, subject to which alone existence can be preserved, may, it is believed, be followed out so as to agree with all the phenomena presented by organized beings, their extinction and succession in past ages, and all the extraordinary modifications of form, instinct, and habits which they exhibit.

Written at Ternate, February, 1858

CHAPTER I.

PHYSICAL GEOGRAPHY.

From a look at a globe or a map of the Eastern hemisphere, we shall perceive between Asia and Australia a number of large and small islands forming a connected group distinct from those great masses of land, and having little connection with either of them. Situated upon the Equator, and bathed by the tepid water of the great tropical oceans, this region enjoys a climate more uniformly hot and moist than almost any other part of the globe, and teems with natural productions which are elsewhere unknown. The richest of fruits and the most precious of spices are Indigenous here. It produces the giant flowers of the *Rafflesia*, the great green-winged *Ornithoptera* (princes among the butterfly tribes), the man-like Orangutan, and the gorgeous Birds of Paradise. It is inhabited by a peculiar and interesting race of mankind--the Malay, found nowhere beyond the limits of this insular tract, which has hence been named the Malay Archipelago.

To the ordinary Englishman this is perhaps the least known part of the globe. Our possessions in it are few and scanty; scarcely any of our travellers go to explore it; and in many collections of maps it is almost ignored, being divided between Asia and the Pacific Islands. It thus happens that few persons realize that, as a whole, it is comparable with the primary divisions of the globe, and that some of its separate islands are larger than France or the Austrian Empire. The traveller, however, soon acquires different ideas. He sails for days or even weeks along the shores of one of these great islands, often so great that its inhabitants believe it to be a vast continent. He finds that voyages among these islands are commonly reckoned by weeks and months, and that their several inhabitants are often as little known to each other as are the native races of the northern to those of the southern continent of America. He soon comes to look upon this region as one apart from the rest of the world, with its own races of men and its own aspects of nature; with its own ideas, feelings, customs, and modes of speech, and with a climate, vegetation, and animated life altogether peculiar to itself.

From many points of view these islands form one compact geographical whole, and as such they have always been treated by travellers and men of science; but, a more careful and detailed study of them under various aspects reveals the unexpected fact that they are divisible into two portions nearly equal in extent which differ widely in their natural products, and really form two parts of the primary divisions of the earth. I have been able to prove this in considerable detail by my observations on the natural history of the various parts of the Archipelago; and, as in the description of my travels and residence in the several islands I shall have to refer continually to this view, and adduce facts in support of it, I have thought it advisable to commence with a general sketch of the main features of the Malayan region as will render the facts hereafter brought forward more interesting, and their bearing upon the general question more easily understood. I proceed, therefore, to sketch the limits and extent of the Archipelago, and to point out the more striking features of its geology, physical geography, vegetation, and animal life.

Definition and Boundaries.--For reasons which depend mainly on the distribution of animal life, I consider the Malay Archipelago to include the Malay Peninsula as far as Tenasserim and the Nicobar Islands on the west, the Philippines on the north, and the Solomon Islands, beyond New Guinea, on the east. All the great islands included within these limits are connected together by innumerable smaller ones, so that no one of them seems to be distinctly separated from the rest. With but few exceptions all enjoy an uniform and very similar climate, and are covered with a luxuriant forest vegetation. Whether we study their form and distribution on maps, or actually travel from island to island, our first impression will be that they form a connected whole, all the parts of which are intimately related to each other.

Extent of the Archipelago and Islands.--The Malay Archipelago extends for more than 4,000 miles in length from east to west, and is about 1,300 in breadth from north to south. It would stretch over an expanse equal to that of all Europe from the extreme west far into Central Asia, or would cover the widest parts of South America, and extend far beyond the land into the Pacific and Atlantic oceans. It includes three islands larger than Great Britain; and in one of them, Borneo, the whole of the British Isles might be set down, and would be

surrounded by a sea of forests. New Guinea, though less compact in shape, is probably larger than Borneo. Sumatra is about equal in extent to Great Britain; Java, Luzon, and Celebes are each about the size of Ireland. Eighteen more islands are, on the average, as large as Jamaica; more than a hundred are as large as the Isle of Wight; while the isles and islets of smaller size are innumerable.

The absolute extent of land in the Archipelago is not greater than that contained by Western Europe from Hungary to Spain; but, owing to the manner in which the land is broken up and divided, the variety of its productions is rather in proportion to the immense surface over which the islands are spread, than to the quantity of land which they contain.

Geological Contrasts.--One of the chief volcanic belts upon the globe passes through the Archipelago, and produces a striking contrast in the scenery of the volcanic and non-volcanic islands. A curving line, marked out by scores of active, and hundreds of extinct, volcanoes may be traced through the whole length of Sumatra and Java, and thence by the islands of Bali, Lombok, Sumbawa, Flores, the Serwatty Islands, Banda, Amboyna, Batchian, Makian, Tidore, Ternate, and Gilolo, to Morty Island. Here there is a slight but well-marked break, or shift, of about 200 miles to the westward, where the volcanic belt begins again in North Celebes, and passes by Sian and Sanguir to the Philippine Islands along the eastern side of which it continues, in a curving line, to their northern extremity. From the extreme eastern bend of this belt at Banda, we pass onwards for 1,000 miles over a non-volcanic district to the volcanoes observed by Dampier, in 1699, on the north-eastern coast of New Guinea, and can there trace another volcanic belt through New Britain, New Ireland, and the Solomon Islands, to the eastern limits of the Archipelago.

In the whole region occupied by this vast line of volcanoes, and for a considerable breadth on each side of it, earthquakes are of continual recurrence, slight shocks being felt at intervals of every few weeks or months, while more severe ones, shaking down whole villages, and doing more or less injury to life and property, are sure to happen, in one part or another of this district, almost every year. On many of the islands the years of the great earthquakes form the chronological epochs of the native inhabitants, by the aid of which the ages of their children are remembered, and the dates of many important events are determined.

I can only briefly allude to the many fearful eruptions that have taken place in this region. In the amount of injury to life and property, and in the magnitude of their effects, they have not been surpassed by any upon record. Forty villages were destroyed by the eruption of Papandayang in Java, in 1772, when the whole mountain was blown up by repeated explosions, and a large lake left in its place. By the great eruption of Tomboro in Sumbawa, in 1815, 12,000 people were destroyed, and the ashes darkened the air and fell thickly upon the earth and sea for 300 miles around. Even quite recently, since I left the country, a mountain which had been quiescent for more than 200 years suddenly burst into activity. The island of Makian, one of the Moluccas, was rent open in 1646 by a violent eruption which left a huge chasm on one side, extending into the heart of the mountain. It was, when I last visited it in 1860, clothed with vegetation to the summit, and contained twelve populous Malay villages. On the 29th of December, 1862, after 215 years of perfect inaction, it again suddenly burst forth, blowing up and completely altering the appearance of the mountain, destroying the greater part of the inhabitants, and sending forth such volumes of ashes as to darken the air at Ternate, forty miles off, and to almost entirely destroy the growing crops on that and the surrounding islands.

The island of Java contains more volcanoes, active and extinct, than any other known district of equal extent. They are about forty-five in number, and many of them exhibit most beautiful examples of the volcanic cone on a large scale, single or double, with entire or truncated summits, and averaging 10,000 feet high.

It is now well ascertained that almost all volcanoes have been slowly built up by the accumulation of matter--mud, ashes, and lava--ejected by themselves. The openings or craters, however, frequently shift their position, so that a country may be covered with a more or less irregular series of hills in chains and masses, only here and there rising into lofty cones, and yet the whole may be produced by true volcanic action. In this manner the greater part of Java has been formed. There has been some elevation, especially on the south

coast, where extensive cliffs of coral limestone are found; and there may be a substratum of older stratified rocks; but still essentially Java is volcanic, and that noble and fertile island--the very garden of the East, and perhaps upon the whole the richest, the best cultivated, and the best governed tropical island in the world--owes its very existence to the same intense volcanic activity which still occasionally devastates its surface.

The great island of Sumatra exhibits, in proportion to its extent, a much smaller number of volcanoes, and a considerable portion of it has probably a non-volcanic origin.

To the eastward, the long string of islands from Java, passing by the north of Timor and away to Panda, are probably all due to volcanic action. Timor itself consists of ancient stratified rocks, but is said to have one volcano near its centre.

Going northward, Amboyna, a part of Bouru, and the west end of Ceram, the north part of Gilolo, and all the small islands around it, the northern extremity of Celebes, and the islands of Sian and Sang-air, are wholly volcanic. The Philippine Archipelago contains many active and extinct volcanoes, and has probably been reduced to its present fragmentary condition by subsidences attending on volcanic action.

All along this great line of volcanoes are to be found more or less palpable signs of upheaval and depression of land. The range of islands south of Sumatra, a part of the south coast of Java and of the islands east of it, the west and east end of Timor, portions of all the Moluccas, the Ke and Aru Islands, Waigiou, and the whole south and east of Gilolo, consist in a great measure of upraised coral-rock, exactly corresponding to that now forming in the adjacent seas. In many places I have observed the unaltered surfaces of the elevated reefs, with great masses of coral standing up in their natural position, and hundreds of shells so fresh-looking that it was hard to believe that they had been more than a few years out of the water; and, in fact, it is very probable that such changes have occurred within a few centuries.

The united lengths of these volcanic belts is about ninety degrees, or one-fourth of the entire circumference of the globe. Their width is about fifty miles; but, for a space of two hundred miles on each side of them, evidences of subterranean action are to be found in recently elevated coral-rock, or in barrier coral-reefs, indicating recent submergence. In the very centre or focus of the great curve of volcanoes is placed the large island of Borneo, in which no sign of recent volcanic action has yet been observed, and where earthquakes, so characteristic of the surrounding regions, are entirely unknown. The equally large island of New Guinea occupies another quiescent area, on which no sign of volcanic action has yet been discovered. With the exception of the eastern end of its northern peninsula, the large and curiously-shaped island of Celebes is also entirely free from volcanoes; and there is some reason to believe that the volcanic portion has once formed a separate island. The Malay Peninsula is also non-volcanic.

The first and most obvious division of the Archipelago would therefore be into quiescent and volcanic regions, and it might, perhaps, be expected that such a division would correspond to some differences in the character of the vegetation and the forms of life. This is the case, however, to a very limited extent; and we shall presently see that, although this development of subterranean fires is on so vast a scale--has piled up chains of mountains ten or twelve thousand feet high--has broken up continents and raised up islands from the ocean--yet it has all the character of a recent action which has not yet succeeded in obliterating the traces of a more ancient distribution of land and water.

Contrasts of Vegetation.--Placed immediately upon the Equator and surrounded by extensive oceans, it is not surprising that the various islands of the Archipelago should be almost always clothed with a forest vegetation from the level of the sea to the summits of the loftiest mountains. This is the general rule. Sumatra, New Guinea, Borneo, the Philippines and the Moluccas, and the uncultivated parts of Java and Celebes, are all forest countries, except a few small and unimportant tracts, due perhaps, in some cases, to ancient cultivation or accidental fires. To this, however, there is one important exception in the island of Timor and all the

smaller islands around it, in which there is absolutely no forest such as exists in the other islands, and this character extends in a lesser degree to Flores, Sumbawa, Lombock, and Bali.

In Timor the most common trees are Eucalypti of several species, also characteristic of Australia, with sandalwood, acacia, and other sorts in less abundance. These are scattered over the country more or less thickly, but, never so as to deserve the name of a forest. Coarse and scanty grasses grow beneath them on the more barren hills, and a luxuriant herbage in the moister localities. In the islands between Timor and Java there is often a more thickly wooded country abounding in thorny and prickly trees. These seldom reach any great height, and during the force of the dry season they almost completely lose their leaves, allowing the ground beneath them to be parched up, and contrasting strongly with the damp, gloomy, ever-verdant forests of the other islands. This peculiar character, which extends in a less degree to the southern peninsula of Celebes and the east end of Java, is most probably owing to the proximity of Australia. The south-east monsoon, which lasts for about two-thirds of the year (from March to November), blowing over the northern parts of that country, produces a degree of heat and dryness which assimilates the vegetation and physical aspect of the adjacent islands to its own. A little further eastward in Timor and the Ke Islands, a moister climate prevails; the southeast winds blowing from the Pacific through Torres Straits and over the damp forests of New Guinea, and as a consequence, every rocky islet is clothed with verdure to its very summit. Further west again, as the same dry winds blow over a wider and wider extent of ocean, they have time to absorb fresh moisture, and we accordingly find the island of Java possessing a less and less arid climate, until in the extreme west near Batavia, rain occurs more or less all the year round, and the mountains are everywhere clothed with forests of unexampled luxuriance.

Contrasts in Depth of Sea.--It was first pointed out by Mr. George Windsor Earl, in a paper read before the Royal Geographical Society in 1845, and subsequently in a pamphlet "On the Physical Geography of South-Eastern Asia and Australia", dated 1855, that a shallow sea connected the great islands of Sumatra, Java, and Borneo with the Asiatic continent, with which their natural productions generally agreed; while a similar shallow sea connected New Guinea and some of the adjacent islands to Australia, all being characterised by the presence of marsupials.

We have here a clue to the most radical contrast in the Archipelago, and by following it out in detail I have arrived at the conclusion that we can draw a line among the islands, which shall so divide them that one-half shall truly belong to Asia, while the other shall no less certainly be allied to Australia. I term these respectively the Indo-Malayan and the Austro-Malayan divisions of the Archipelago.

On referring to pages 12, 13, and 36 of Mr. Earl's pamphlet, it will be seen that he maintains the former connection of Asia and Australia as an important part of his view; whereas, I dwell mainly on their long continued separation. Notwithstanding this and other important differences between us, to him undoubtedly belongs the merit of first indicating the division of the Archipelago into an Australian and an Asiatic region, which it has been my good fortune to establish by more detailed observations.

Contrasts in Natural Productions.--To understand the importance of this class of facts, and its bearing upon the former distribution of land and sea, it is necessary to consider the results arrived at by geologists and naturalists in other parts of the world.

It is now generally admitted that the present distribution of living things on the surface of the earth is mainly the result of the last series of changes that it has undergone. Geology teaches us that the surface of the land, and the distribution of land and water, is everywhere slowly changing. It further teaches us that the forms of life which inhabit that surface have, during every period of which we possess any record, been also slowly changing.

It is not now necessary to say anything about how either of those changes took place; as to that, opinions may differ; but as to the fact that the changes themselves have occurred, from the earliest geological ages down to

the present day, and are still going on, there is no difference of opinion. Every successive stratum of sedimentary rock, sand, or gravel, is a proof that changes of level have taken place; and the different species of animals and plants, whose remains are found in these deposits, prove that corresponding changes did occur in the organic world.

Taking, therefore, these two series of changes for granted, most of the present peculiarities and anomalies in the distribution of species may be directly traced to them. In our own islands, with a very few trifling exceptions, every quadruped, bird, reptile, insect, and plant, is found also on the adjacent continent. In the small islands of Sardinia and Corsica, there are some quadrupeds and insects, and many plants, quite peculiar. In Ceylon, more closely connected to India than Britain is to Europe, many animals and plants are different from those found in India, and peculiar to the island. In the Galapagos Islands, almost every indigenous living thing is peculiar to them, though closely resembling other kinds found in the nearest parts of the American continent.

Most naturalists now admit that these facts can only be explained by the greater or less lapse of time since the islands were upraised from beneath the ocean, or were separated from the nearest land; and this will be generally (though not always) indicated by the depth of the intervening sea. The enormous thickness of many marine deposits through wide areas shows that subsidence has often continued (with intermitting periods of repose) during epochs of immense duration. The depth of sea produced by such subsidence will therefore generally be a measure of time; and in like manner, the change which organic forms have undergone is a measure of time. When we make proper allowance for the continued introduction of new animals and plants from surrounding countries by those natural means of dispersal which have been so well explained by Sir Charles Lyell and Mr. Darwin, it is remarkable how closely these two measures correspond. Britain is separated from the continent by a very shallow sea, and only in a very few cases have our animals or plants begun to show a difference from the corresponding continental species. Corsica and Sardinia, divided from Italy by a much deeper sea, present a much greater difference in their organic forms. Cuba, separated from Yucatan by a wider and deeper strait, differs more markedly, so that most of its productions are of distinct and peculiar species; while Madagascar, divided from Africa by a deep channel three hundred miles wide, possesses so many peculiar features as to indicate separation at a very remote antiquity, or even to render it doubtful whether the two countries have ever been absolutely united.

Returning now to the Malay Archipelago, we find that all the wide expanse of sea which divides Java, Sumatra, and Borneo from each other, and from Malacca and Siam, is so shallow that ships can anchor in any part of it, since it rarely exceeds forty fathoms in depth; and if we go as far as the line of a hundred fathoms, we shall include the Philippine Islands and Bali, east of Java. If, therefore, these islands have been separated from each other and the continent by subsidence of the intervening tracts of land, we should conclude that the separation has been comparatively recent, since the depth to which the land has subsided is so small. It is also to be remarked that the great chain of active volcanoes in Sumatra and Java furnishes us with a sufficient cause for such subsidence, since the enormous masses of matter they have thrown out would take away the foundations of the surrounding district; and this may be the true explanation of the often-noticed fact that volcanoes and volcanic chains are always near the sea. The subsidence they produce around them will, in time, make a sea, if one does not already exist.

But, it is when we examine the zoology of these countries that we find what we most require--evidence of a very striking character that these great islands must have once formed a part of the continent, and could only have been separated at a very recent geological epoch. The elephant and tapir of Sumatra and Borneo, the rhinoceros of Sumatra and the allied species of Java, the wild cattle of Borneo and the kind long supposed to be peculiar to Java, are now all known to inhabit some part or other of Southern Asia. None of these large animals could possibly have passed over the arms of the sea which now separate these countries, and their presence plainly indicates that a land communication must have existed since the origin of the species. Among the smaller mammals, a considerable portion are common to each island and the continent; but the vast physical changes that must have occurred during the breaking up and subsidence of such extensive regions

have led to the extinction of some in one or more of the islands, and in some cases there seems also to have been time for a change of species to have taken place. Birds and insects illustrate the same view, for every family and almost every genus of these groups found in any of the islands occurs also on the Asiatic continent, and in a great number of cases the species are exactly identical. Birds offer us one of the best means of determining the law of distribution; for though at first sight it would appear that the watery boundaries which keep out the land quadrupeds could be easily passed over by birds, yet practically it is not so; for if we leave out the aquatic tribes which are preeminently wanderers, it is found that the others (and especially the Passeres, or true perching-birds, which form the vast majority) are generally as strictly limited by straits and arms of the sea as are quadrupeds themselves. As an instance, among the islands of which I am now speaking, it is a remarkable fact that Java possesses numerous birds which never pass over to Sumatra, though they are separated by a strait only fifteen miles wide, and with islands in mid-channel. Java, in fact, possesses more birds and insects peculiar to itself than either Sumatra or Borneo, and this would indicate that it was earliest separated from the continent; next in organic individuality is Borneo, while Sumatra is so nearly identical in all its animal forms with the peninsula of Malacca, that we may safely conclude it to have been the most recently dismembered island.

The general result therefore, at which we arrive, is that the great islands of Java, Sumatra, and Borneo resemble in their natural productions the adjacent parts of the continent, almost as much as such widely-separated districts could be expected to do even if they still formed a part of Asia; and this close resemblance, joined with the fact of the wide extent of sea which separates them being so uniformly and remarkably shallow, and lastly, the existence of the extensive range of volcanoes in Sumatra and Java, which have poured out vast quantities of subterranean matter and have built up extensive plateaux and lofty mountain ranges, thus furnishing a *vera causa* for a parallel line of subsidence--all lead irresistibly to the conclusion that at a very recent geological epoch, the continent of Asia extended far beyond its present limits in a south- easterly direction, including the islands of Java, Sumatra, and Borneo, and probably reaching as far as the present 100-fathom line of soundings.

The Philippine Islands agree in many respects with Asia and the other islands, but present some anomalies, which seem to indicate that they were separated at an earlier period, and have since been subject to many revolutions in their physical geography.

Turning our attention now to the remaining portion of the Archipelago, we shall find that all the islands from Celebes and Lombok eastward exhibit almost as close a resemblance to Australia and New Guinea as the Western Islands do to Asia. It is well known that the natural productions of Australia differ from those of Asia more than those of any of the four ancient quarters of the world differ from each other. Australia, in fact, stands alone: it possesses no apes or monkeys, no cats or tigers, wolves, bears, or hyenas; no deer or antelopes, sheep or oxen; no elephant, horse, squirrel, or rabbit; none, in short, of those familiar types of quadruped which are met with in every other part of the world. Instead of these, it has Marsupials only: kangaroos and opossums; wombats and the duckbilled Platypus. In birds it is almost as peculiar. It has no woodpeckers and no pheasants--families which exist in every other part of the world; but instead of them it has the mound-making brush-turkeys, the honeysuckers, the cockatoos, and the brush-tongued lorries, which are found nowhere else upon the globe. All these striking peculiarities are found also in those islands which form the Austro-Malayan division of the Archipelago.

The great contrast between the two divisions of the Archipelago is nowhere so abruptly exhibited as on passing from the island of Bali to that of Lombok, where the two regions are in closest proximity. In Bali we have barbets, fruit-thrushes, and woodpeckers; on passing over to Lombok these are seen no more, but we have abundance of cockatoos, honeysuckers, and brush- turkeys, which are equally unknown in Bali, or any island further west. [I was informed, however, that there were a few cockatoos at one spot on the west of Bali, showing that the intermingling of the productions of these islands is now going on.] The strait is here fifteen miles wide, so that we may pass in two hours from one great division of the earth to another, differing as essentially in their animal life as Europe does from America. If we travel from Java or Borneo to Celebes or

the Moluccas, the difference is still more striking. In the first, the forests abound in monkeys of many kinds, wild cats, deer, civets, and otters, and numerous varieties of squirrels are constantly met with. In the latter none of these occur; but the prehensile-tailed Cuscus is almost the only terrestrial mammal seen, except wild pigs, which are found in all the islands, and deer (which have probably been recently introduced) in Celebes and the Moluccas. The birds which are most abundant in the Western Islands are woodpeckers, barbets, trogons, fruit-thrushes, and leaf-thrushes; they are seen daily, and form the great ornithological features of the country. In the Eastern Islands these are absolutely unknown, honeysuckers and small lorises being the most common birds, so that the naturalist feels himself in a new world, and can hardly realize that he has passed from the one region to the other in a few days, without ever being out of sight of land.

The inference that we must draw from these facts is, undoubtedly, that the whole of the islands eastwards beyond Java and Borneo do essentially form a part of a former Australian or Pacific continent, although some of them may never have been actually joined to it. This continent must have been broken up not only before the Western Islands were separated from Asia, but probably before the extreme southeastern portion of Asia was raised above the waters of the ocean; for a great part of the land of Borneo and Java is known to be geologically of quite recent formation, while the very great difference of species, and in many cases of genera also, between the productions of the Eastern Malay Islands and Australia, as well as the great depth of the sea now separating them, all point to a comparatively long period of isolation.

It is interesting to observe among the islands themselves how a shallow sea always intimates a recent land connexion. The Aru Islands, Mysol, and Waigiou, as well as Jobie, agree with New Guinea in their species of mammalia and birds much more closely than they do with the Moluccas, and we find that they are all united to New Guinea by a shallow sea. In fact, the 100-fathom line round New Guinea marks out accurately the range of the true Paradise birds.

It is further to be noted--and this is a very interesting point in connection with theories of the dependence of special forms of life on external conditions--that this division of the Archipelago into two regions characterised by a striking diversity in their natural productions does not in any way correspond to the main physical or climatal divisions of the surface. The great volcanic chain runs through both parts, and appears to produce no effect in assimilating their productions. Borneo closely resembles New Guinea not only in its vast size and its freedom from volcanoes, but in its variety of geological structure, its uniformity of climate, and the general aspect of the forest vegetation that clothes its surface. The Moluccas are the counterpart of the Philippines in their volcanic structure, their extreme fertility, their luxuriant forests, and their frequent earthquakes; and Bali with the east end of Java has a climate almost as dry and a soil almost as arid as that of Timor. Yet between these corresponding groups of islands, constructed as it were after the same pattern, subjected to the same climate, and bathed by the same oceans, there exists the greatest possible contrast when we compare their animal productions. Nowhere does the ancient doctrine--that differences or similarities in the various forms of life that inhabit different countries are due to corresponding physical differences or similarities in the countries themselves--meet with so direct and palpable a contradiction. Borneo and New Guinea, as alike physically as two distinct countries can be, are zoologically wide as the poles asunder; while Australia, with its dry winds, its open plains, its stony deserts, and its temperate climate, yet produces birds and quadrupeds which are closely related to those inhabiting the hot damp luxuriant forests, which everywhere clothe the plains and mountains of New Guinea.

In order to illustrate more clearly the means by which I suppose this great contrast has been brought about, let us consider what would occur if two strongly contrasted divisions of the earth were, by natural means, brought into proximity. No two parts of the world differ so radically in their productions as Asia and Australia, but the difference between Africa and South America is also very great, and these two regions will well serve to illustrate the question we are considering. On the one side we have baboons, lions, elephants, buffaloes, and giraffes; on the other spider-monkeys, pumas, tapirs, anteaters, and sloths; while among birds, the hornbills, turacos, orioles, and honeysuckers of Africa contrast strongly with the toucans, macaws, chatterers, and hummingbirds of America.

Now let us endeavour to imagine (what it is very probable may occur in future ages) that a slow upheaval of the bed of the Atlantic should take place, while at the same time earthquake- shocks and volcanic action on the land should cause increased volumes of sediment to be poured down by the rivers, so that the two continents should gradually spread out by the addition of newly-formed lands, and thus reduce the Atlantic which now separates them, to an arm of the sea a few hundred miles wide. At the same time we may suppose islands to be upheaved in mid- channel; and, as the subterranean forces varied in intensity, and shifted their points of greatest action, these islands would sometimes become connected with the land on one side or other of the strait, and at other times again be separated from it. Several islands would at one time be joined together, at another would be broken up again, until at last, after many long ages of such intermittent action, we might have an irregular archipelago of islands filling up the ocean channel of the Atlantic, in whose appearance and arrangement we could discover nothing to tell us which had been connected with Africa and which with America. The animals and plants inhabiting these islands would, however, certainly reveal this portion of their former history. On those islands which had ever formed a part of the South American continent, we should be sure to find such common birds as chatterers and toucans and hummingbirds, and some of the peculiar American quadrupeds; while on those which had been separated from Africa, hornbills, orioles, and honeysuckers would as certainly be found. Some portion of the upraised land might at different times have had a temporary connection with both continents, and would then contain a certain amount of mixture in its living inhabitants. Such seems to have been the case with the islands of Celebes and the Philippines. Other islands, again, though in such close proximity as Bali and Lombeck, might each exhibit an almost unmixed sample of the productions of the continents of which they had directly or indirectly once formed a part.

In the Malay Archipelago we have, I believe, a case exactly parallel to that which I have here supposed. We have indications of a vast continent, with a peculiar fauna and flora having been gradually and irregularly broken up; the island of Celebes probably marking its furthest westward extension, beyond which was a wide ocean. At the same time Asia appears to have been extending its limits in a southeast direction, first in an unbroken mass, then separated into islands as we now see it, and almost coming into actual contact with the scattered fragments of the great southern land.

From this outline of the subject, it will be evident how important an adjunct Natural History is to Geology; not only in interpreting the fragments of extinct animals found in the earth's crust, but in determining past changes in the surface which have left no geological record. It is certainly a wonderful and unexpected fact that an accurate knowledge of the distribution of birds and insects should enable us to map out lands and continents which disappeared beneath the ocean long before the earliest traditions of the human race. Wherever the geologist can explore the earth's surface, he can read much of its past history, and can determine approximately its latest movements above and below the sea-level; but wherever oceans and seas now extend, he can do nothing but speculate on the very limited data afforded by the depth of the waters. Here the naturalist steps in, and enables him to fill up this great gap in the past history of the earth.

One of the chief objects of my travels was to obtain evidence of this nature; and my search after such evidence has been rewarded by great success, so that I have been able to trace out with some probability the past changes which one of the most interesting parts of the earth has undergone. It may be thought that the facts and generalizations here given would have been more appropriately placed at the end rather than at the beginning of a narrative of the travels which supplied the facts. In some cases this might be so, but I have found it impossible to give such an account as I desire of the natural history of the numerous islands and groups of islands in the Archipelago, without constant reference to these generalizations which add so much to their interest. Having given this general sketch of the subject, I shall be able to show how the same principles can be applied to the individual islands of a group, as to the whole Archipelago; and thereby make my account of the many new and curious animals which inhabit them both, more interesting and more instructive than if treated as mere isolated facts.

Contrasts of Races.--Before I had arrived at the conviction that the eastern and western halves of the

Archipelago belonged to distinct primary regions of the earth, I had been led to group the natives of the Archipelago under two radically distinct races. In this I differed from most ethnologists who had before written on the subject; for it had been the almost universal custom to follow William von Humboldt and Pritchard, in classing all the Oceanic races as modifications of one type. Observation soon showed me, however, that Malays and Papuans differed radically in every physical, mental, and moral character; and more detailed research, continued for eight years, satisfied me that under these two forms, as types, the whole of the peoples of the Malay Archipelago and Polynesia could be classified. On drawing the line which separates these races, it is found to come near to that which divides the zoological regions, but somewhat eastward of it; a circumstance which appears to me very significant of the same causes having influenced the distribution of mankind that have determined the range of other animal forms.

The reason why exactly the same line does not limit both is sufficiently intelligible. Man has means of traversing the sea which animals do not possess; and a superior race has power to press out or assimilate an inferior one. The maritime enterprise and higher civilization of the Malay races have enabled them to overrun a portion of the adjacent region, in which they have entirely supplanted the indigenous inhabitants if it ever possessed any; and to spread much of their language, their domestic animals, and their customs far over the Pacific, into islands where they have but slightly, or not at all, modified the physical or moral characteristics of the people.

I believe, therefore, that all the peoples of the various islands can be grouped either with the Malays or the Papuans; and that these two have no traceable affinity to each other. I believe, further, that all the races east of the line I have drawn have more affinity for each other than they have for any of the races west of that line; that, in fact, the Asiatic races include the Malays, and all have a continental origin, while the Pacific races, including all to the east of the former (except perhaps some in the Northern Pacific), are derived, not from any existing continent, but from lands which now exist or have recently existed in the Pacific Ocean. These preliminary observations will enable the reader better to apprehend the importance I attach to the details of physical form or moral character, which I shall give in describing the inhabitants of many of the islands.



Wallace's Line in the Light of Recent Zoogeographic Studies

Author(s): Ernst Mayr

Source: *The Quarterly Review of Biology*, Vol. 19, No. 1 (Mar., 1944), pp. 1-14

Published by: [The University of Chicago Press](#)

Stable URL: <http://www.jstor.org/stable/2808563>

Accessed: 17/03/2014 23:44

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



The University of Chicago Press is collaborating with JSTOR to digitize, preserve and extend access to *The Quarterly Review of Biology*.

<http://www.jstor.org>

THE QUARTERLY REVIEW of BIOLOGY



WALLACE'S LINE IN THE LIGHT OF RECENT ZOOGEOGRAPHIC STUDIES

BY ERNST MAYR

American Museum of Natural History, New York

ZOOGEOGRAPHY has had a fate very much like taxonomy. It was flourishing during the descriptive period of biological sciences. Its prestige, however, declined rapidly when experimental biology began to come to the foreground. Again as with taxonomy, a new interest in zoogeography has been noticeable in recent years. It seems to me that this revival has had two causes. One is the interest of the student of geographic speciation in the findings of the zoogeographer. A study of past and present distributions yields much information on isolation of populations and on the dispersal of species. It is in this connection that I became interested in zoogeography.

The other reason is the introduction of new methods. The intensive exploration of all corners of the globe during the past fifty years has led to an accumulation of sufficient faunistic data to permit the application of statistical methods. Furthermore, the science of ecology has reached a level of maturity at which it is beginning to affect profoundly zoogeographic methods and principles. It seemed worth while to me to study the controversial and still wide open subject of the borderline between the Australian and Oriental Regions with the help of such modern methods.

A. R. Wallace, who is generally considered the foremost representative of classical zoogeography, states in his famous essay *On the zoological geography of the Malay Archipelago* (1860): "The western and eastern islands of the archipelago belong to regions more distinct and contrasted than any other of the great zoological divisions of the globe. South America and Africa, separated

by the Atlantic, do not differ so widely as Asia and Australia." There is much truth in this statement. Except for bats and a few rodents, the only native mammals of Australia are marsupials and monotremes. These same two groups are entirely lacking in Asia and are replaced by a wide variety of placental mammals, such as monkeys, shrews, squirrels, ungulates, and so forth. An equally pronounced faunal difference exists among birds, insects, and other groups of animals of the two regions.

Australia and Asia are connected by a belt of islands, the Malay Archipelago, and the question naturally comes up as to where in this island region the borderline is to be drawn between these two fundamentally different faunas. After reviewing the zoological evidence known to him, Wallace (*l.c.*) comes to the following conclusion: "We may consider it established that the Strait of Lombok [between Bali and Lombok] (only 15 miles wide) marks the limit and abruptly separates two of the great zoological regions of the globe." With these words he drew a zoogeographic boundary which was destined to gain fame under the name of its author: "Wallace's Line," a term first used by Huxley (1868) (Fig. 1). It runs between Bali and Lombok in the south, then through Makassar Strait between Borneo and Celebes, and finally turns into the open Pacific between Mindanao (Philippines) and the Sanghir Islands. This convenient borderline found quick acceptance in the zoological literature and was without hesitation adopted by nearly all the zoogeographers publishing between 1860 and 1890. Sarasin (1901) and Pelseneer (1904) should be

consulted for a historical survey of the earlier literature. The echo in the popular literature of this period was even more enthusiastic. A mysterious line, only 15 miles wide, that separates marsupials from tigers, and honey eaters and cockatoos from barbets and trogons, could not fail to appeal to the imagination of the layman.

facts became better known. Wallace himself was much less positive in his later writings. Since then many writers have insisted that Wallace's Line was entirely imaginary (Weber (1902), Pelseneer (1904), Mertens (1930), Brongersma (1936), and others). Van Kampen (1909), for example, asserted: "Such a sharp boundary as

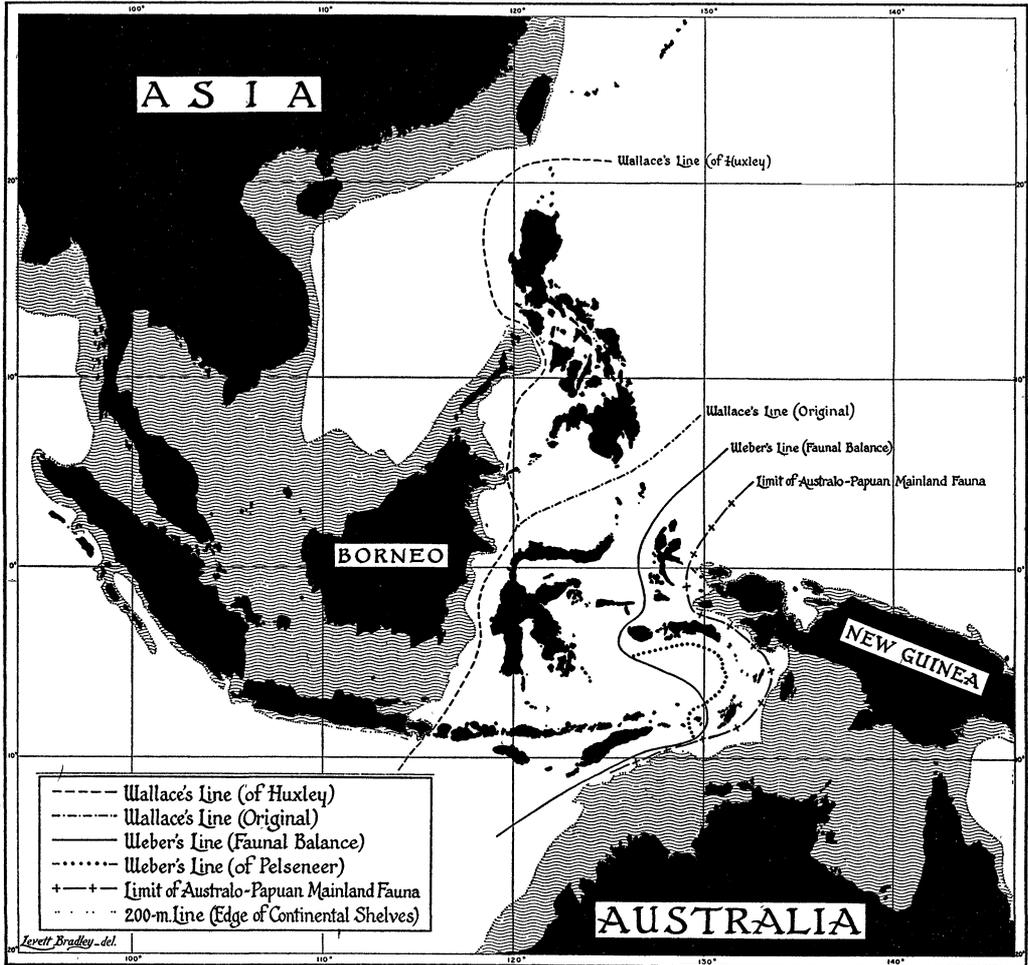


FIG. 1. ZOOGEOGRAPHIC BORDERLINES IN THE MALAY ARCHIPELAGO
The shaded areas are the continental shelves.

E. Haeckel (1893) outdid all his contemporaries by asserting: "Crossing the narrow but deep Lombok Strait we go with a single step from the Present Era to the Mesozoicum."

Statements of such exaggeration call for refutation and shortly after 1890 doubts were expressed more and more frequently as to the validity of Wallace's Line, particularly after the distributional

Wallace drew it does not exist. Not only is there none where he drew it, but no such line exists anywhere in the archipelago." On the other hand, Wallace's Line has been vigorously defended by such serious authors as Dickerson *et al.* (1928), Raven (1935), and Rensch (1936). Curiously enough most of the writers on this subject seem to be definitely in one or the other camp, either

they are for Wallace's Line or they are against it, and they tend to present their data accordingly. Others treat one aspect only of this diversified problem. An impartial study of the situation is still lacking at the present time.

Actually, a whole complex of questions is involved, of which the following seem to be the most important ones:

(1) Is Wallace's Line the borderline between the Oriental and the Australian Regions, and if not, where is this borderline?

(2) Does Wallace's Line represent the line of a major faunal break, and if this is true, how did such a break develop?

IS WALLACE'S LINE THE BOUNDARY BETWEEN THE ORIENTAL AND THE AUSTRALIAN REGIONS?

The fauna of the Malay Archipelago was rather poorly known in Wallace's days. Where he knew 20 species of birds, we now know 120; where he knew 5 species of reptiles, we know 40, and so forth. This lack of information caused Wallace to single out what he considered typical representatives of the respective faunas, and to use the borderline of their ranges as zoogeographic boundaries. The tiger, the squirrels and other mammals go as far east as Bali, but are absent from Lombok. Among birds the barbets (Capitonidae) and many other Oriental groups are abruptly brought to a halt by Lombok Strait. The Australian honeyeaters (genera *Philemon* and *Meliphaga*) and the cockatoo (*Cacatua*) reach Lombok, but not Bali. The faunal difference on either side of Makassar Strait is even more striking: A rich Oriental fauna on Borneo and a marsupial (*Phalanger*) on Celebes. It was on the basis of such data that Wallace came to the conclusion that Lombok and Makassar Straits form the boundary between the Oriental and the Australian Regions.

An analysis of the now available extensive faunal lists does not bear out Wallace's conclusion. After eliminating a few widespread species, the fauna of each of the islands of the Malay Archipelago can be divided readily into two groups: One consists of western species, that is, species which are derived from the Oriental Fauna, the other of eastern species, that is, such which are derived from the Australian Fauna. In a few species it is apparent that the genus or the family to which they belong was originally of western origin, but that the particular species arrived in the

island belt from the east as a descendant of a group of species that was isolated in Australia at an early date. Such secondarily eastern elements, as *Merops ornatus* among the birds, are included with the eastern group. The classification of a few species will always remain open to doubt, but a different decision in these cases would change the percentages only slightly and would not basically affect the following figures. A specialist of a given group usually has no difficulties in deciding which species are Indo-Malayan and which Australian.

Celebes. Weber (1902), the Sarasins (1901), de Beaufort (1926), Stresemann (1939) and other recent authors agree that at least three fourths of the Celebes animals are of western origin. According to Rensch (1936: 252) the figures are: Reptiles at least 88 per cent, Amphibia 80 per cent

TABLE 1

Percentage of western and eastern species on Lesser Sunda Islands

	REPTILES AND AMPHIBIANS		BIRDS		CHANGE OF PERCENTAGE IN BIRDS
	Western	Western	Eastern		
	<i>Per cent</i>	<i>Per cent</i>	<i>Per cent</i>		
Bali.....	94	87.0	13.0	14.5	
Lombok.....	85	72.5	27.5	4.5	
Sumbawa.....	87	68.0	32.0	5.0	
Flores.....	78	63.0	37.0	5.5	
Alor Group.....	—	57.5	42.5		

and butterflies 86 per cent. In birds the figure is slightly lower. Among 74 species of Passerine birds 67.6 per cent are western. The percentage for the old endemics (genera and good species) and for the more recent immigrants is quite similar. There is no doubt, Celebes must be included with the Oriental Region.

Lesser Sunda Islands. Table 1 shows the ratio of the western and the eastern elements on a number of islands (the data of reptiles and amphibians are from Mertens, 1930; the data on birds are original). Rensch's (1936) careful analysis shows that the Indo-Malayan element prevails numerically as far east as the islands of the Timor group. This is equally true for flying animals (birds and butterflies) and for flightless groups (mammals, land snails).

The figures in Table 1 permit only a single conclusion: Wallace's Line is not the borderline

between the Australian and the Oriental Regions. The first of the questions asked above is thus answered in favor of Wallace's opponents.

DOES WALLACE'S LINE INDICATE A MAJOR
FAUNAL BREAK?

The fact that Wallace's Line is not the border between the Oriental and the Australo-Papuan Regions is not the complete answer to our problem. A line which has been defended so vigorously by so many zoogeographers must have some significance. It is worthy of notice that its staunchest defenders were those naturalists who actually studied and collected the animal life on both sides of the line, like Dickerson and his associates in the Philippines, like Raven who repeatedly crossed Makassar Strait in a sail boat from Borneo to Celebes and back, and like Wallace and Rensch who crossed back and forth between Bali and Lombok. The actual impressions of these workers are vividly depicted in a quotation from one of Rensch's books. Arriving on Bali after a prolonged exploration of Lombok, Sumbawa and Flores, he asks himself:

"What about the animal life? Is it really as different from that of Lombok, as has been claimed by so many other travellers? Is the small strait between the two islands actually a sharp faunal division? A strait, which even the smallest bird could cross without any difficulties? . . . And the difference is indeed quite extraordinary! Much more conspicuous than I would have ever imagined. As soon as I entered the woods on a small native trail a whole chorus of strange bird songs greets me—in fact, among the real songsters there is not a single one with which I was familiar [from the islands east of Wallace's Line]. . . . One surprise follows the other. The very species that are most common on Bali, are absent on the islands to the east. The most characteristic bird of these woods is a green barbet . . . it belongs to the family Capitonidae which is entirely absent on Lombok! The woodpeckers also, which are represented on the islands farther east by a single species only, are found on Bali in five different species. On the other hand I missed a whole number of species of birds which are characteristic for the islands visited previously . . ." (Rensch 1930).

An unemotional statistical analysis of the faunal data tends to support Rensch's assertions. The most striking feature of Wallace's Line is that it separates a zone with a rich animal life from a badly impoverished one. Borneo has about 420 species of breeding birds, Celebes only 220. Java has about 340 breeding species, Lombok only

120. It is even more true for freshwater fish: Borneo has 162 species of the carp family Cyprinidae, Celebes has none; Java has 55 species, Lombok has apparently only a single one. Raven (1935) shows that the Mammalian fauna is equally impoverished. The same is true for the Philippines, their fauna is badly depleted, as compared to that of Borneo and Palawan (Dickerson *et al.*, 1928).

THE GEOLOGY OF THE MALAY ARCHIPELAGO

Why the islands Sumatra, Java, Borneo, and Palawan should have a rich animal life, whereas the Philippines, Celebes, and the Lesser Sunda Islands have a poor one, cannot be understood without a study of the geological conditions. The British geologist Earle pointed out, as early as 1845, that geologically the Malay Archipelago consists of three parts, a western one comprising the greater Sunda Islands and the adjoining parts of Asia, which was very stable during the Tertiary, an eastern one consisting of New Guinea and Australia, which was also stable, and an unstable island belt in between. The unstable area, comprising the Philippines, Celebes, the Moluccas, and the Lesser Sunda Islands, has a most complicated geological structure. Deep sea basins, grabens, geosynclines and geanticlines are scrambled together in a bewildering manner. Geologists are still far from agreement in regard to the interpretation of these structures. So much, however, is clear—that this area is highly unstable and that it has seen many and violent changes in the recent past.

Originally, that is in late Mesozoic times, Celebes, the Moluccas, Misol, and western New Guinea, seem to have been situated on the same continental shelf. The fossil marine faunas of the mentioned regions, as well as tectonic features prove this close relationship. In fact, most geologists consider it as well established that Asia and Australia were in broad continental connection up to the very end of the Mesozoic. The Tertiary was a period of very active orogenesis. Part of the Philippines and of northern Celebes seem to have been folded up first. There is some evidence for the existence of additional islands during Eocene and Oligocene, as, for example, in the Timor region, but the exact position, size and chronology of such islands is unknown. In early Miocene, or according to other authors in very late Oligocene, the crust of the

earth seems to have buckled down in a gigantic manner along a line, which is roughly indicated by the west Sumatran Islands, Timor, Kei, Seran, and Halmahera. The very strong negative anomalies of the gravimetric measurements along this line are according to Vening Meinesz good evidence for the occurrence of such an event. The folding was so violent that it resulted in the widespread overthrusting of older strata over younger ones. The so-called outer Banda arc, consisting of the islands Sumba, Timor, Babber, Timorlaut, Kei, Seran, and Buru was formed along part of this fold. All of these islands are geologically very similar. Slightly later, but still in the Miocene, a second fold was formed consisting of parts of Sumatra and Java, as well as of the so-called inner Banda arc (Bali, Lombok, Sumbawa, Flores, Alor, Wetar, Dammer, and Banda). Most of this fold remained, however, at first submerged under the ocean. In fact, some of the islands may not have emerged until well in the Pleistocene. Later in the Tertiary, particularly in the Pliocene and Pleistocene, extensive fault lines developed which led to the lifting of large blocks (marine terraces in Timor rose 1280 m.!) and the corresponding sinking of other blocks to form flat-bottomed deep sea basins. The geological data indicate that periods of violent tectonic activity have alternated with periods of relative quietness, and the frequency of earthquakes and the continuous volcanic activity in this region reveal that the orogenic movements have not yet completely died down. The writings of Molengraaf (1922), Umbgrove (1932, 1934), and Kuenen (1935) should be consulted for further details concerning the geology of this region. Three facts of zoogeographic significance seem to stand out among the geological data: (1) There is no evidence whatsoever for any continental connection between Borneo and Celebes. In fact, the distance between the two islands was, up to the Pleistocene, greater than it is today; (2) Java, Bali, Lombok, and the other islands of the inner Banda arc are situated on the same geanticline; and (3) there is no geological evidence for any cross connections between inner and outer Banda arcs, except possibly between Sumba and Flores.

The first of these three conclusions shows that Makassar Strait is an ancient ocean barrier and that at least this particular part of Wallace's Line is geologically well-founded. Geologists and zoogeographers are in full agreement on this point.

Tertiary geology supplies, however, no explanation for a faunal difference between Bali and Lombok, a difference which seems to be due to events of a more recent geological past. A considerable quantity of ocean water accumulated in the polar ice caps during the Pleistocene glaciations. It has been calculated that this resulted in a lowering of the sea level of tropical ocean by at least 70 m., but more probably by 150 m. This caused the drying up of all shallow seas and resulted in a considerable extension of land on Sunda and Sahul shelves (see Fig. 1). Sumatra, Java, and Borneo united with the Malay Peninsula in the formation of "Sundaland," an extension of the Asiatic mainland and Bali became attached to this continent. Lombok, however, which is separated from Bali by a strait of a depth of 312 m., remained separated, even though it was fused temporarily with Sumbawa.

The geological background of Wallace's Line is thus as follows: In its central part, between Borneo and Celebes, it follows the edge of the continental Sunda shelf, in the south between Bali and Lombok (and the same is true in the north between Borneo-Palawan and the Philippines) it indicates the eastern edge of the Pleistocene Sundaland. The faunal break, which I have shown to exist along Wallace's Line, appears now in a new light. It is due to the fact that the line separates, on the whole, a continental from an insular fauna. This separation is clear cut in Makassar Strait, but it is rather obscured along the Sunda arc, where the geanticline of the inner Banda arc protrudes from Sundaland like a peninsula. Faunal breaks along this chain of islands occur not only on Lombok Strait, but also on all the other inter-island straits. A number of authors, among whom Mertens (1930) is foremost, have contended, that some of the other straits, as that between Java and Bali, or the one between Sumbawa and Flores, are even more efficient distribution barriers than Lombok Strait. This assertion is in conflict with the above-given findings of Pleistocene geology and it becomes therefore necessary to examine the relative efficiency of these water barriers in more detail.

THE EFFICIENCY OF THE WATER BARRIERS BETWEEN THE LESSER SUNDA ISLANDS

The faunal change between Borneo and Celebes is abrupt, but it is much more gradual along the west of Wallace's Line. The number of species

of birds on this island chain is as follows: Sumatra about 440; Java, 340; Bali, 166; Lombok, 119; Sumbawa, 123; Flores, 143; and Timor, 137. In the freshwater fish family Cyprinidae, Sumatra has 115 species; Java, 55; and Lombok only a single one. Of butterflies Sumatra has 334 species; Java, 270 species. Of reptiles Sumatra has 193 species; Java, 136 species (Rensch 1936). It is obvious from these figures that the animal life of Java is considerably impoverished as compared with that of Sumatra (or Borneo). The reasons for this are not entirely clear, but two factors seem to be most important. One is the heavy activity of the Javanese volcanoes, particularly during the Pleistocene, which covered a good part of the island with lava and ashes and may have exterminated a number of localized species. The second and more important reason is that Java is less humid

of Bali Strait as a zoogeographic barrier. It would be entirely misleading to subtract the number of Bali species from the total number of Java species and state that the difference comprises the species that are unable to cross Bali Strait. This method was actually applied by Mertens (1930) and Brongersma (1936), who arrived thereby at the erroneous conclusion that Bali Strait was the most important barrier along the Sunda chain.

A faunal change between the western and the eastern end occurs probably not only on Java, but on all elongated islands of the Sunda chain such as Sumbawa, Flores, and Timor. This fact invalidates to some extent the figures on the subsequent calculations, but it is fortunately of minor importance in respect to the small and rather round islands of Bali and Lombok.

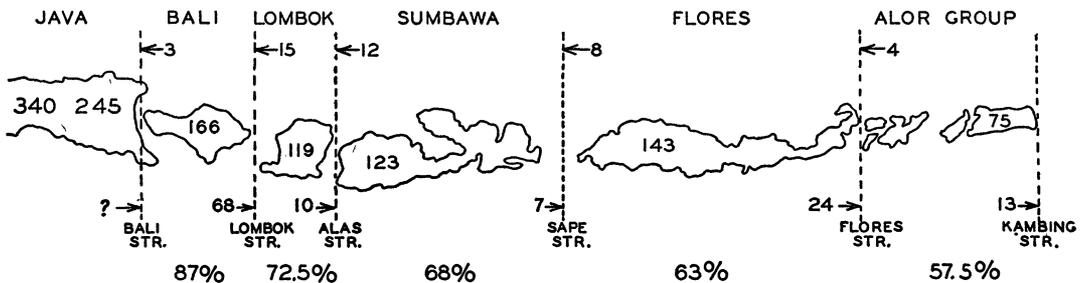


FIG. 2. INTER-ISLAND STRAITS IN THE LESSER SUNDA ISLANDS AND THEIR EFFICIENCY AS DISTRIBUTIONAL BARRIERS FOR BIRDS

Explanation of figures in text.

and poorer in habitats than Sumatra, also more peripheral and thus less accessible to colonists from the Asiatic mainland. The climatic deterioration, which is already indicated in western Java, accelerates rapidly in the eastern part of the island where in the lowlands true tropical rain forest seems to be largely replaced by monsoon forest. The result is that many of the most characteristic Java elements (including nearly all of the well-known endemics) are restricted to western Java. Of the 340 species of Java birds only 245 are found in the eastern half of the island and it is reasonable to believe that some 70 of these species drop out before the eastern tip of Java is reached, leaving only about 170 species for the eastern tip of the island. No natural history survey has ever been made of this section of Java. This is unfortunate because the fauna of a small area of easternmost Java, equivalent in size to Bali, must be compared with the Bali fauna, if one wants to test the signif-

This is by no means the only difficulty that is encountered in the attempt to determine the relative efficiency of the various straits in the Sunda chain. It happens that there is a gradual but steady change of climate and plant cover from west to east. Each more easterly island is somewhat more arid than its western neighbor and one after the other of the humidity-loving species drops out because the habitat becomes unsuitable and not necessarily because it can not cross the water barrier separating it from the next island.

The effect of six inter-island straits on the distribution of birds is illustrated in Fig. 2. The top line of figures records the number of eastern species that find the western limit of their ranges on the inter-island straits. Lombok Strait shows the highest figure with 15 species, but, on the whole, the difference between the various straits is rather slight. This is not surprising, since all the eastern species have the ability to jump water barriers

and it is probable that the ecological factors on the islands have as much or more to do with the limits of the ranges than age or width of the straits between them. The second row of figures gives the number of breeding species known from each island (the second figure on Java gives the number of species on the eastern half of Java). The third row of figures, and this is the most important one of all, gives the number of western species that are halted by the various straits. The significance of Lombok Strait becomes at once apparent. It prevents the passage of 68 (41 per cent) of the 166 Bali species. No other strait approximates this figure. The last row of figures gives the percentage of western species on each of the islands.

The relative efficiency of Lombok, Alas, Sape, and Flores Straits can be expressed by calculating what percentage of the species occurring on either side are stopped by the straits. Lombok Strait, for example, is a barrier for 83 (= 68 + 15) species

are, however, found on Wetar, Timor, or other more easterly islands, which proves that Flores Strait is not the eastern limit of their range.

The data presented in Fig. 2 and in Table 2 can be summarized as follows: Each of the straits in the Lesser Sunda Islands is a zoogeographic barrier. Lombok Strait, however, is more effective than any of the others. This is apparently due to the fact that this strait persisted throughout the Pleistocene, whereas Bali Strait and Alas Strait dried up at the height of the Pleistocene glaciation. Rensch's data (Table 2) indicate that reptiles, amphibia, butterflies, and land mollusks show similar conditions as birds and it is possible that a more thorough exploration of Bali, Flores, and Alor would make the two sets of data even more similar. Mertens (1930), for example, who denies that Lombok Strait is a more important barrier than Bali or Sape Straits, bases this claim on a study of the very insufficiently known herpeto-

TABLE 2
Relative efficiency of straits in Lesser Sunda Islands

	BIRDS (ORIG.)	REPTILES, AMPHIBIA BUTTERFLIES AND LAND MOLLUSKS (RENSCH)	RANK (FOR BIRDS)
Lombok Strait (Bali-Lombok)	83 of 285 = 29 per cent ¹	84 of 377 = 22 per cent	1
Alas Strait (Lombok-Sumbawa)	22 of 242 = 9 per cent	34 of 367 = 9 per cent	3
Sape Strait (Sumbawa-Flores)	15 of 266 = 5.5 per cent	52 of 364 = 15 per cent	4
Flores Strait (Flores-Alor)	28 of 218 = 13 per cent		2

¹ This figure indicates what percentage of the sum of the species of the two islands on either side of the strait have not crossed the strait.

of a total of 285 (166 + 119), that is, 29.1 per cent. The percentages for the other straits and for a number of other animal groups are given in Table 2.

The figures of Table 2 have, of course, only a relative value since the species totals include many species twice, once east and once west of the straits. Still they are valid as indicators of the relative efficiency of these straits and of their rank. Many of the smaller islands (Penida, Komodo, Sangeang, Rintja, etc.) are insufficiently explored and have therefore been omitted from the tabulation. Solor, Adonara, Pantar, and Alor have been united as Alor group. In the tabulation of the borders of western species only the easternmost occurrence has been used. This explains a seeming discrepancy of the figures. Flores Strait, for example, stops only 24 of the 143 species on Flores. One would expect the Alor group to have 119 species (143 less 24), but it actually has only 75. The "missing" 44 (119 less 75) species

fauna of Bali. De Beaufort (1926) and other students of freshwater fish are also handicapped in a discussion of this subject, since with few exceptions there are no true freshwater fish east of Lombok Strait. Even Bali has a badly depleted fish fauna. That this island is so poor in freshwater fish is not really surprising, since it is a volcanic island and has not a single large and only one medium-sized stream. It would be dangerous to base too many conclusions on the distributional data derived from a single group which is so exacting in its ecological requirements as are freshwater fish.

Freshwater fish are useful as negative zoogeographic indicators. The fact that primary freshwater fish (see Myers, 1938, for a definition of this term) are absent from Seran and Kei indicates, for example, that these islands have had no continental connection with New Guinea. The presence of four species of freshwater fish in the Lesser

Sunda Islands—*Rasbora elberti* on Lombok and Sumbawa; *Clarias batrachus* on Bali, Lombok, and Sumbawa; *Aplocheilichthys javanicus* on Lombok; and *A. celebensis* on Timor—does not necessarily prove continental connections for these, but it casts doubt on the means of dispersal of these species. The slight, or absent, differentiation of these species demands that these islands had a recent continental connection. However, if such had existed one would expect a much richer fish fauna. The transport of fish by water spouts is well substantiated and it is also possible that Lombok Strait had occasionally a surface sheet of freshwater while it was the outlet for the large streams of Pleistocene Sundaland. It would be dangerous to go too far in such speculations of possible chance dispersal but it is even more dangerous to base sweeping zoogeographic conclusions on the presence of a few species of so-called freshwater fish.

THE EASTERN COUNTERPART OF WALLACE'S LINE

It is obvious that there must be a line at the eastern edge of the island belt which corresponds to Wallace's Line in the west. Such a line would separate the zone of a more or less pure Australo-Papuan mainland fauna from the islands to the west with an impoverished Papuan fauna and an Indo-Malayan admixture. This line has been vaguely referred to by Lydekker and other nineteenth century writers, but I believe de Beaufort (1913) was the first to point out its true significance. It is not difficult to trace since it follows, except for a short stretch in the north, the 100 m. depth line, that is, the edge of that part of the Sahul shelf that was dry land at the height of the Pleistocene glaciation (Fig. 1). It passes between the Aru Islands, which have a pure Papuan fauna, and the Kei Islands with an impoverished fauna with Oriental elements. Of birds, for example, 166 species are known from the Aru Islands, including birds of paradise and many other typical Papuan types, while only 84 species are known from the Kei Islands, including some western elements. The line then passes between the mainland of New Guinea and Seran Island. There are 115 species of birds (about 30 per cent western) known from Seran as against more than 300 species from the Vogelkop, the neighboring part of New Guinea. The line that separates the Papuan mainland fauna from the island fauna swings from Seran north and passes through the Gilolo passage separating the western Papuan Islands (Waigeu, Batanta,

Salwati, and Misol) from the Northern Moluccas. In this section the line does not follow entirely the 100 m. contour, which would exclude Koffiau, Gebe, Batanta, and Waigeu. However, all these islands are so purely Papuan and form such a well-defined faunistic unit that it seems justified to be slightly inconsistent. It might be worth while to emphasize that the line, as just drawn, gives a better defined delimitation of the "Papuan mainland" and "Papuan island" fauna than does Wallace's Line in the west for the Indo-Malayan fauna. Its validity is particularly apparent for all groups with a limited dispersal faculty, for example freshwater fish. De Beaufort's map (1926, p. 103) of the range of the subfamily *Melanotaeniinae* illustrates it quite graphically. This Australian group extends westward as far as the Aru Islands and Waigeu, but is absent from the Kei Islands and from the Northern and Southern Moluccas.

The significance of this eastern line has been emphasized by a number of authors. It indicates, like Wallace's Line, a major faunal break; it separates, like its western counterpart, a continental from an island zone, as well as a zone with a more or less undiluted Papuan fauna from a mixed Papuan-Oriental fauna, a contrast which is least apparent in the north. It is for all these reasons that this line must be considered a major zoogeographic boundary.

SHOULD AN INTERMEDIATE ZOOGEOGRAPHIC REGION BE RECOGNIZED?

A gradual transition between the Oriental and the Australian faunas takes place in the island belt between Wallace's Line and its eastern counterpart. This was realized quite clearly by Salomon Müller (1846), the earliest zoogeographer of the Malay Archipelago. He lists correctly "Celebes, Flores, Timor, Gilolo and perhaps Mindanao" as islands on which a mixture of Indian and Australian elements is found. Wallace also, in his later publications, admitted the intermediate position of this region and stated of Celebes that it "hardly belongs to either [Oriental or Australian] region." Pelsener (1904:1007) lists a whole group of workers who recognized the transitional character of this region.

There are other factors, in addition to the lack of continental connection, which contribute toward the poverty of the fauna of this island belt. Salomon Müller (1846) very ably pointed out some

of the reasons, such as the small size of most of the islands, their low elevation and their aridity. There is a more or less arid corridor extending from the Philippines and Celebes to Buru and to the Sunda Islands from eastern Java to Timorlaut. This zone has acted as a barrier to many humidity-loving forms and has prevented their passage from Sundaland to the Papuan Region or vice versa. Additional reasons for the faunal poverty of this zone are the young geological age of many of the islands, which limits the number of chance colonizations, and the heavy volcanic activity over part of the region. There are three lines of volcanoes in this transition zone, one extending from Sumatra through Java to the inner Banda arc, a second one following the western edge of the northern Moluccas, and a third one reaching from north Celebes through the Sanghir Islands to the Philippines. The volcanic activity is thus strictly localized, but where it occurs it may be a very serious factor indeed. As mentioned, it seems to be one of the reasons why Java's animal life is so much poorer than that of Borneo or Sumatra (Rensch, 1936). There are not only 59 young volcanoes of more than 2000 m. altitude on Java, but also many extinct late Tertiary ones. This factor is even more evident on Lombok where heavy Pleistocene eruptions of Mount Rindjani seem to have destroyed much of the mountain fauna. The same is true for the volcano on Ternate Island (Stresemann, 1939:381).

All the mentioned factors combine to give the fauna of the transition zone a peculiar character. This has impressed some of the authors to such an extent that they have proposed to give formal recognition to this fauna and elevate the island belt to the rank of a separate zoogeographic region or subregion.

Dickerson *et al.* (1928), who coined the term *Wallacea* for this region, and Rensch (1936), who simply calls it *Zwischengebiet* (region of intermedicity), are the two most recent champions of such an arrangement. This region would include four different groups of islands, (1) the Lesser Sunda Islands from Lombok eastward; (2) the Moluccas and other outliers of the Papuan Region (Tenimber, Kei); (3) the Celebes group (with Sula and Talaut); and (4) the Philippines. Two reasons are usually quoted in favor of recognizing such a transition region. One is, that many endemic species and genera are confined to it. The other reason is, that all of the islands, which are included in this transition zone, are populated by a mixture

of Indo-Malayan and Australo-Papuan elements. As against these points which would favor the recognition of a transition region there are some very strong objections. Pelseneer (1904) has stated them clearly. He points out that it is only natural that

a zoogeographic border is not a line without width and that by necessity there is a mixture of faunal elements along the border of two zoogeographic regions, caused by a reciprocal penetration.

But if one would admit for this reason a special 'transition region' or a 'region of intermedicity,' one would obviously double the difficulties of delimitation. For now it would be necessary to trace both of the border lines which separate the transition region from either of the two adjoining zoogeographic regions.

These difficulties of delimitation are fully confirmed by the two most recently proposed transition regions. Dickerson *et al.* (1928:297) define theirs as follows: "Wallacea is outlined sharply by Wallace's Line (as modified) on the west and Weber's Line upon the east." It thus includes the Philippines, but it excludes the Moluccas, Timorlaut, and Kei Islands. Rensch (1936:265), however, includes in his *Zwischengebiet* "Celebes, the Lesser Sunda Islands, Timorlaut (perhaps also Kei), and the Moluccas (at least the southern Moluccas)." He definitely excludes the Philippines. Celebes and the Lesser Sunda Islands are, thus, the only two districts which the two transition regions have in common.

The "degree of intermedicity" of the various sections of the transition region is very uneven. It seems, for example, that the percentage of Australo-Papuan species in the Philippines (which are included in Wallacea by Dickerson and Merrill) is smaller than the percentage of Oriental species in New Guinea or Australia. Still, nobody would want to suggest including Australo-Papua in the transition zone.

Stresemann (1939:403) adds another weighty objection. He points out that the transition zone comprises four separate districts which have much less in common with one another than each one has with some outside region: The Moluccas are faunistically closest to New Guinea, and Celebes to the Philippines, but the Philippines are closer to Malaysia than to Celebes. The Lesser Sunda Islands, finally, have a close faunal relationship with Java and Australo-Papua, but only a very slight and recent one with Celebes. To unite four such heterogeneous districts in a single "region"

violates all principles of regional zoogeography. After all, if a zoogeographic region means anything, it means the home of a more or less homogeneous characteristic fauna. "Wallacea", however, is the home of four different faunas. It is self-evident that the formal recognition of a zoogeographic region of such heterogeneity is neither practical nor scientifically defensible. The term transition zone is justified only if applied informally as a descriptive attribute.

WEBER'S LINE

It is apparent from the preceding discussion that neither Wallace's Line nor the formal acceptance of a transition zone are satisfactory attempts of delimiting the Oriental against the Australian Region. This leaves, to my mind, only one other alternative solution, namely, the recognition of a line east of Wallace's Line. Before attempting to draft the best possible course of such a line, a few words must be said about the validity of any zoogeographic borderline.

A zoogeographic region is usually defined as a geographic subdivision of the earth that is the home of a peculiar fauna. Such a region is characterized by the presence of many endemic genera and families and by the absence of the characteristic genera and families of other zoogeographic regions. Its border should be drawn along the line where this specific fauna is replaced by a different fauna. This procedure is logical and presents no difficulties in all the cases where an efficient barrier separates the two regions, such as is formed by the South Atlantic between Africa and South America. However, an intermingling of the two faunas takes place in a border zone whenever two such regions come into direct contact. This is exactly what has happened in the island belt between Asia and Australia. Both the Indo-Malayan and the Australo-Papuan mainland faunas have spilled over into the intermediate island belt and it might seem impossible to delimit in such a mixed region one fauna from the other one. However, as Pelseneer (1904) says correctly, "it is evident that there must be a line . . . within the region of mixture, on one side of which the faunal elements of one region prevail and on the other side those of the second region. This line can serve usefully to mark the borderline between the two biogeographic regions."

On the basis of these considerations Pelseneer established a borderline between the Oriental and the Australian Regions, which he called "Weber's

Line." Pelseneer drew the course of this line on the basis of non-zoological data, that is, primarily on the soundings and other oceanographic results of the Siboga Expedition, many of which are no longer valid today. However, Weber's Line actually separates the islands with a more than 50 per cent Indo-Malayan fauna from the islands with a more than 50 per cent Papuan fauna, as is evident from Rensch's (1936) careful data and from all the other zoogeographic work of the region. With insignificant modification the line suggested by Pelseneer is still acceptable as the best possible borderline between the Oriental and the Australo-Papuan Regions.

The course of Weber's Line (Fig. 1) is as follows: In the north it begins between Talaut and Celebes in the west and the northern Moluccas in the east. In this section the line is extremely well defined, since the fauna of the northern Moluccas consists of about 80-90 per cent and that of Celebes of about 20-40 per cent Papuan elements. The line continues from here between the Sula Islands in the west and Obi in the east and then swings around Buru. The fauna of the Sula Islands is insufficiently known, but it is close to that of Celebes except much poorer and with a stronger Moluccan element. Still the Papuan component is probably less than 40 per cent, while it is about 63 per cent on Buru and even higher on Obi. It is difficult to trace Weber's Line from Buru on. Pelseneer attempted to follow the contour of the ocean bottom and this caused him to run the line between Banda ("Indo-malayan") and Seran ("Papuan") and between Sermatta ("Indo-malayan") and Babber ("Papuan"). The much more detailed information on the fauna of these islands, which is now available, indicates that a different course might be preferable. The young volcanic Banda Islands have a fauna which almost completely lacks endemic elements, and which is very close to that of Ambon, Seran, Seranlaut, etc. There is no doubt that the Banda Islands must be included in the southern Moluccas. Babber, on the other hand, has a fauna which is closer to that of Dammer and Sermatta, than to that of Timorlaut. It is, therefore, preferable to place the line between Babber and Timorlaut. Rensch (1926:206) has already pointed out the impossibility of separating Babber from the closely related Sermatta and Dammer. The fauna of Timorlaut is about 62.5 per cent Australo-Papuan. South West Islands, from Roma and Kisar to Dammer and Babber are a faunistic unit, but the progressive decrease of Indo-

Malayan elements which started on Java and Bali continues on these islands. It is possible that a future analysis may show that the eastern element on Babber and Dammer is already slightly more than 50 per cent of the total fauna of these islands. Even then I would be inclined to retain them in the Oriental Region rather than to draw a line through the middle of the South West Islands.

One glance at the map shows that Weber's Line is situated much closer to the Australo-Papuan than to the Asiatic shelf. The reason for this is twofold, faunal pressure and accessibility. The faunal pressure of the Indo-Malayan fauna is greater than that of the Papuan fauna because it is much richer in species and families. The sphere of influence of this rich fauna will, therefore, extend farther into the island belt than that of the poorer Papuan fauna. The second reason is that the chain of the Lesser Sunda Islands, forming practically a peninsula of Sundaland, was infinitely more easily accessible to colonists from the west than to those from the east, which had to jump the wide gap either from Australia to Timor or from New Guinea (and Aru) to the islands of Banda Sea. The preponderance of Oriental species in the Lesser Sunda Islands would be even more pronounced if ecological factors (aridity) had not favored colonization by Australian elements. These various factors explain the present course of the line of faunal balance, Weber's Line. Wallace's argument that Celebes should be included in the Australian Region because it had so few Oriental species as compared to Borneo, is beside the point. Every true island has, of course, a much impoverished fauna, but its zoogeographic position is determined by an analysis of its existing fauna and not by the elements it lacks. With an 80 per cent Oriental fauna Celebes can not be included in the Australian Region!

Weber's Line has found curiously few adherents among zoogeographers; Boden Kloss (1929) is one of the exceptions. There is nothing spectacular about this line and by crossing it one encounters a smaller faunal change than is found between Borneo and Celebes, or between New Guinea and Seran, or in general between the "mainland" and the "island" faunas (Fig. 1). The difference between the faunas of Sula and Buru and of Babber and Timorlaut is, indeed, rather small. Weber's Line is not acceptable to those who look for a strikingly conspicuous borderline between the Oriental and Australian Regions (Rensch, 1926:265).

Other objections have been raised against Weber's Line. Some authors, for example, have objected to Weber's Line because it separates islands which lie on the same submarine ridges. Thus it cuts between Babber and Timorlaut, between Dammer and Banda, and between Sula and Obi, each of these three pairs of islands lying on the same submarine ridge. It seems to me that this argument is another instance of confusing zoogeographic and geological interpretations, exactly as in the case of continental versus oceanic islands (Mayr, 1941). The geology of an island, and particularly of an oceanic island, is of no concern whatsoever, when we are attempting to classify its fauna. If the fauna of Seran and Kei is pre-vaillingly Papuan, I shall classify these islands with the Papuan region. The fact that Timor and Sumba with a prevaillingly Indo-Malayan fauna lie on the same tectonic arc has absolutely no bearing on this decision. In fact there is no evidence that any of these arcs were ever raised to the extent that they were exposed for their full length, and it is obvious that the undersea geology can have no influence over the distribution of forms that are dispersed across the water.

De Beaufort (1926:184) also rejects Weber's Line for a purely geological reason, because "the Moluccas are not the remains of a former greater land mass." This argument is entirely irrelevant, not only since Pelseneer nowhere makes such a claim, but also because the former geological history has a bearing on zoogeographic classification only to the extent to which it influences present day distribution.

I know of only a single valid argument against the adoption of Weber's Line as the boundary between the Australian and the Oriental Regions. It is the objection against dividing arbitrarily any continuous series of values at the halfway point between the extremes. In the case of Weber's Line the situation is aggravated by the fact that the 50:50 balance between the Indo-Malayan and the Australian elements is not always the same in the various taxonomic groups. The bird fauna of Wetar Island, for example, is more than 50 per cent Australo-Papuan, while in other groups the Oriental element seems to prevail. On Celebes about 67 per cent of the birds are of western origin, while among mammals, butterflies, reptiles, amphibians, and land snails it is more than 80 per cent. On the whole it seems as if among reptiles and butterflies the western element pushes farther eastward than among birds and snails. However, taking the

fauna as a whole, Weber's Line seems to separate rather neatly the islands with a prevailing Oriental fauna in the west from the islands with a prevailing Australo-Papuan fauna in the east. As stated above, the easternmost of the South West Islands (Dammer and Babber) possibly have slightly more than 50 per cent Australo-Papuan elements, but it is inadvisable to separate them from the larger group of islands of which they are an integral part. Lines of 50:50 balance face even more difficulties on continents than in archipelagos. The line in North America on which the Palearctic and the Nearctic elements balance, would be entirely unsuitable as a zoogeographic boundary. A 50:50 line is, thus, admittedly a more or less arbitrary boundary and may have to be modified in special cases. But it is no more arbitrary than to accept March 21st as the first day of spring (regardless of the weather!), or the 21st birthday as the day on which an adolescent reaches seniority. Such rigid divisions are of practical usefulness not only in human affairs, but frequently also in scientific matters. Different faunal regions are generally indicated on zoogeographic maps by different colors. It is obvious that the 50:50 line is the most convenient place where to replace one color by another. It is in this sense that Weber's Line (as modified above) may be accepted as the boundary between the region with a prevailing Oriental and the region with a prevailing Australo-Papuan fauna.

UNSOLVED PROBLEMS OF INDO-AUSTRALIAN ZOOGEOGRAPHY

The conclusions at which I arrived in the present analysis are not final. Many of the islands are insufficiently explored and it is certain that future exploration will add a good deal to our knowledge. A further refinement in the zoogeographic methods is also expected to yield increased results. Salomon Müller, P. L. Sclater, A. R. Wallace, and other early representatives of the classical school of zoogeography selected arbitrarily a number of indicator species and based the outlines of the zoogeographic regions and subregions on the distribution of these species. The preferred technique of the present paper is to calculate in percent the proportion of faunal element in the total number of species of certain localities. All the percentages in Fig. 2 and Table 2 are derived by this method.

In the matter of faunal composition an even superior method might be to determine the faunal

relationship of the dominant species of each habitat. It seems, for example, to judge by Rensch's description (see above) that the differences between the dominant species of birds of Bali and Lombok is even more striking than is apparent from a statistical analysis of the total faunas. Such a comparison of the dominant types of local faunas must be based on accurate census data gathered in the field and such data are not yet available. To gather them would be a worthwhile task of future explorers of the Malay archipelago.

The combination of ecological and zoogeographic methods promises to yield data of considerable interest. It seems, for example, that the faunal composition of each habitat is different. Of the eleven species of birds that are restricted to the mountain forest of Timor (above 4000 feet) only a single one is Papuan, the other ten are Indo-Malayan. The ratio is even, if not reversed, among the birds of the tree savanna of Timor. Lack of exact ecological data prevents a more accurate analysis at the present time. Steenis and other botanists have shown that a similar difference of floristic composition exists between different plant associations. Here is a practically untouched field for future investigators.

The delimitation of biogeographic regions depends to a considerable extent on the dispersal faculties and on the nature of the speciation processes of the organisms of which the distribution is studied. It has become evident in recent years that there is much difference between phytozoogeographic and zoogeographic classifications. The major floristic regions coincide fairly well with the major climatic regions. The major zoogeographic regions, on the other hand, indicate primarily the extent of formerly (or currently) isolated land areas. The biogeographic classification of New Guinea is a good illustration for this. New Guinea is, for the phytozoographer, a part of the Malayan region, but faunistically it is at least as close or even closer to Australia. A comparison of phytozoogeographic and zoogeographic maps indicates that it is impractical at the present time to construct biogeographic maps, that is, maps that intend to illustrate simultaneously the distribution of plants and of animals.

This is equally true, although to a lesser extent, for animal groups with different dispersal faculties. I have already mentioned above the differences between birds and reptiles in regard to the faunal composition of some of the islands. Much more accurate data are needed. It is possible that some

of the invertebrates show a distributional pattern that is much more similar to that of plants than to that of mammals or birds. Progress in this field depends largely on a more thorough faunistic exploration of the Indo-Australian Region.

SUMMARY

(1) Wallace's Line is not the boundary between the Indo-Malayan and the Australian Regions, but rather it indicates the edge of the area (Sunda shelf) that was dry at the height of the Pleistocene glaciations.

(2) The equivalent line along the edge of the Sahul Shelf separates New Guinea and the Aru Islands from the Moluccas and Kei Islands.

(3) Weber's Line separates the islands in the west on which the Indo-Malayan element is predominant from the islands in the east on which the Australo-Papuan element has a numerical superiority.

POSTSCRIPT

The results of an important symposium on Wallace's Line and on the zoogeography of the

Indo-Australian archipelago (Scrivenor, *et al.*, 1943) have been published after the completion of the present work. These papers contain nothing that would require a major modification of the conclusions at which I have arrived above. However, they contribute a considerable amount of interesting factual data and raise a number of questions which I have not treated. Corbet (*op. cit.*) shows that Weber's Line, at least in its northern part between Celebes and the northern Moluccas, constitutes a more pronounced faunal division in several families and genera of butterflies than Wallace's Line in its most effective section (between Borneo and Celebes). Malcolm Smith (*op. cit.*) comes to the conclusion, on the basis of the distribution of vertebrates, that Weber's Line is preferable to Wallace's Line, if a single borderline between the Oriental and the Australian Regions is to be found. The botanical contributors emphasize the discrepancy between the classification of biogeographic regions of the zoologists and of the botanists. This disagreement is much less striking in regard to the minor divisions.

LIST OF LITERATURE

- BEAUFORT, L. F. DE. 1913. Fishes of the eastern part of the Indo-Australian Archipelago, with remarks on its zoogeography. *Bijdr. tot de Dierk.*, 19: 95-164.
- . 1926. Zoögeographie van den Indischen Archipel. Volksuniversiteits Bibliotheek. Vol. 35, 202 pp. *Haarlem*.
- BRONGERSMA, L. D. 1936. Some comments upon H. C. Raven's paper: "Wallace's Line and the distribution of Indo-Australian mammals." *Arch. Néerl. zool.*, 2: 240-256.
- DICKERSON, R. E., *et al.* 1928. Distribution of life in the Philippines. *Bur. Sci. Monogr.*, 21, 322 pp. *Manila*.
- HAECKEL, E. 1893. Zur Phylogenie der australischen Fauna. *Denkschr. Medic. Nat. Ges. Jena*, 4: 5.
- HUXLEY, T. H. 1868. On the classification and distribution of the Alectoromorphae and Heteromorphae. *Proc. Zool. Soc. London*, pp. 294-319.
- KAMPEN, P. N. VAN. 1909. De zoögeographie van den Indische Archipel. *Bijbl. 3 en 4, Natuurk. Tijdschr. Ned. Indië*.
- KLOSS, C. B. 1929. The zoo-geographical boundaries between Asia and Australia and some oriental subregions. *Bull. Raff. Mus.*, 2: 1-10.
- KUENEN, PH. H. 1935. Geological interpretation of the bathymetrical results. *In: The Snellius Expedition in the eastern part of the Netherlands East-Indies 1929-1930*. Vol. 5, pt. 1, p. 1-124.
- MAYR, E. 1941. The origin and the history of the bird fauna of Polynesia. *Proc. Sixth Pac. Sci. Congr.* (1939), 4: 197-216.
- MERTENS, R. 1930. Die Amphibien und Reptilien der Inseln Bali, Lombok, Sumbawa and Flores. *Abh. Senckenb. Naturf. Ges.*, 42: 115-344.
- MOLENGRAEFF, G. A. F. 1922. Geologie. *In: De Zeen van Nederlandsch Oost-Indië*, pp. 272-357. *Leiden*.
- MÜLLER, SALOMON. 1846. Über den Charakter der Thierwelt auf den Inseln des indischen Archipels, etc. *Arch. Naturg.*, 12: 109-128.
- MYERS, G. S. 1938. Fresh-water fishes and West Indian Zoogeography. *Smithson. Rep. for 1937*: 339-364.
- PELSENEER, P. 1904. La "Ligne de Weber," limite zoologique de l'Asie et de l'Australie. *Bull. Classe Sci. Acad. R. Belgique*, 1904: 1001-1022.
- RAVEN, H. C. 1935. Wallace's Line and the distribution of Indo-Australian mammals. *Bull. Amer. Mus. Nat. Hist.*, 68: 177-293.
- RENSCH, B. 1930. Eine biologische Reise nach den Kleinen Sunda Inseln. P. 209. *Berlin*.
- . 1936. Die Geschichte des Sundabogens. Eine tiergeographische Untersuchung. Pp. 1-318 (with an exhaustive bibliography of 326 titles on the zoogeography of the Malay archipelago). *Berlin*.
- SARASIN, P., AND F. 1901. Über die geologische Geschichte der Insel Celebes auf Grund der Tierverbreitung. 169 pp. *Wiesbaden*.

- SCRIVENOR, J. B., T. H. BURKILL, MALCOLM A. SMITH, A. ST. CORBET, H. K. AIRY SHAW, P. W. RICHARDS, and F. E. ZEUNER. 1943. A discussion of the biogeographic division of the Indo-Australian Archipelago, with criticism of the Wallace and Weber lines and of any other dividing lines and with an attempt to obtain uniformity in the names used for the divisions. *Proc. Linn. Soc. London*, 154th Sess., pp. 120-165.
- STRESEMANN, E. 1939. Die Vögel von Celebes. Zoogeographie. *J. Ornith.*, 87: 312-425.
- UMBROGROVE, J. H. F. 1932. Het Neogen in den Indischen Archipel. *Tijdsch. Kon. Nederl. Aardrijksk. Genootsch.*, (2) 49: 769-833.
- , et al. 1934. Gravity, geology and morphology of the East Indian archipelago. In: Gravity Expeditions at Sea, 1923-1932. Vol. 2. *Publ. Netherl. Geod. Commission*.
- WALLACE, A. R. 1860. On the zoological geography of the Malay Archipelago. *Journ. Linn. Soc. London*, 4: 172-184.
- WEBER, M. 1902. Der Indo-australische Archipel und die Geschichte seiner Tierwelt. 46 pp. *Jena*.



Southeast Asian biodiversity: an impending disaster

Navjot S. Sodhi¹, Lian Pin Koh^{1,2}, Barry W. Brook³ and Peter K.L. Ng¹

¹Department of Biological Sciences, National University of Singapore, 14 Science Drive 4, Singapore 117543, Republic of Singapore

²Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544-1003, USA

³Key Centre for Tropical Wildlife Management, Charles Darwin University, Darwin, NT 0909, Australia

Southeast Asia has the highest relative rate of deforestation of any major tropical region, and could lose three quarters of its original forests by 2100 and up to 42% of its biodiversity. Here, we report on the current state of its biota and highlight the primary drivers of the threat of extinction now faced by much of the unique and rich fauna and flora of the region. Furthermore, the known impacts on the biodiversity of Southeast Asia are likely to be just the tip of the iceberg, owing to the paucity of research data. The looming Southeast Asian biodiversity disaster demands immediate and definitive actions, yet such measures continue to be constrained by socioeconomic factors, including poverty and lack of infrastructure. Any realistic solution will need to involve a multidisciplinary strategy, including political, socioeconomic and scientific input, in which all major stakeholders (government, non-government, national and international organizations) must participate.

Tropical ecosystems are exceptionally rich and exclusive reservoirs of much of the biodiversity on Earth. However, the rapid and extensive destruction of tropical habitats has become a serious threat to their native biota [1]. Deforestation is particularly severe in Southeast Asia, where natural habitats, such as lowland rain forests, are being destroyed at relative rates that are higher than those of other tropical regions [2]. If present levels of deforestation were to continue unabated, Southeast Asia will lose almost three-quarters of its original forest cover by the turn of the next century [2], resulting in massive species declines and extinctions [3]. More importantly, this biodiversity crisis is likely to develop into a full-fledged disaster, as the region is home to one of the highest concentrations of endemic species [4].

Here, we discuss the contribution of the unique geological history of Southeast Asia to its high species richness and endemism. We report on the current state of its terrestrial biota and highlight the primary drivers, such as forest conversion, that are responsible for the threat to the unique and rich biodiversity of the region. Finally, we discuss the major conservation challenges faced by this region.

The unique geological history and biodiversity of Southeast Asia

In a study reported in 2000, Myers *et al.* identified 25 'biodiversity hotspots' in the world as those areas containing high concentrations of endemic species and undergoing immense habitat loss (Figure 1) [4]. Southeast Asia overlaps with four of these hotspots, each of which has a unique geological history that has contributed to its rich and often unique biota [5]. During the Pleistocene glacial episodes, some temperate species from northern Asia expanded their ranges southwards into Indo-Burma and retained their presence thereafter [6]. Fluctuating sea levels periodically converted mountains into geographically isolated islands, creating conditions that were ideal for speciation. The episodic sea-level changes also repeatedly connected the islands of Sundaland (covering the western half of the Indo-Malayan archipelago) to the Asian mainland, enabling biotic migrations from the mainland to the archipelago [7]. As the sea level rose, the isolation of these islands also facilitated speciation. The presence of rain forest refugia in parts of Sundaland during the Pleistocene also enabled the persistence of its forest biota [8]. Although it was never connected to the Asian mainland, Wallacea (covering the central islands of Indonesia to the east of Java, Bali and Borneo, and west of the province of Papua on the island of New Guinea) is one of the most geologically complex regions in the world, because its islands originated from land fragments that rifted from Gondwanaland at different geological time periods [9]. This unique geological history, together with its stable tropical climate and numerous insular biotas, enabled Wallacea to evolve highly endemic biotas of its own. The other geologically unique region of Southeast Asia, the Philippines, consists of ~7000 islands, containing multiple centres of endemism [5]. The colonization of newly formed oceanic islands, followed by genetic differentiation and long-term persistence, has resulted in the extraordinarily high species richness and endemism of the Philippines [10]. As a result of the unique geological history Southeast Asia, the region ranks as one of the highest in the world in terms of species richness and endemism [5]. Furthermore, it features unique ecological processes, such as the strong synchrony of fruiting of trees (mast events) from the Dipterocarpaceae [11], which have major implications for forest ecology and conservation.

Corresponding author: Navjot S. Sodhi (dbsns@nus.edu.sg).

Available online 16 September 2004

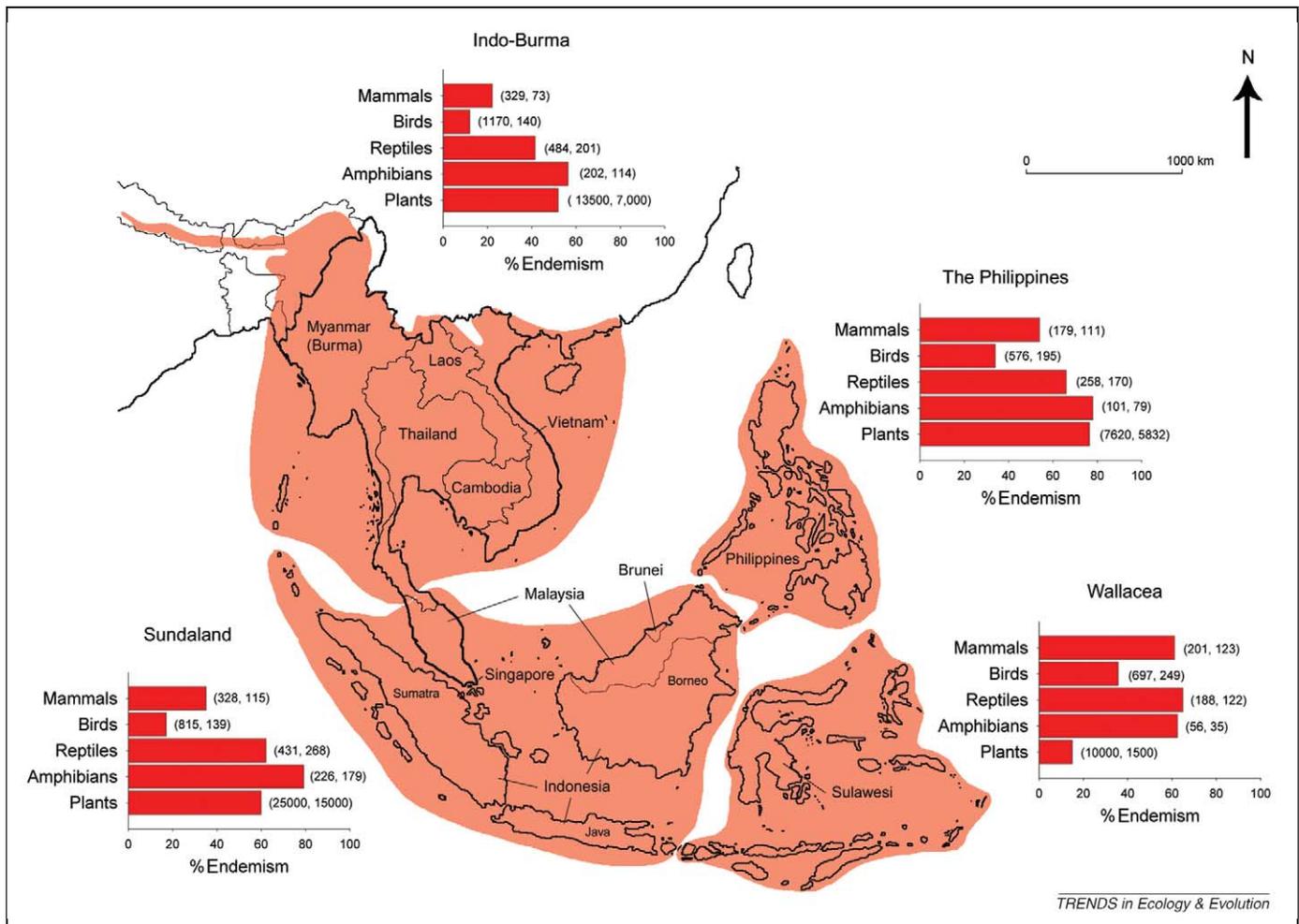


Figure 1. Species richness and endemism in Southeast Asia. The four biodiversity hotspots overlapping Southeast Asia are highlighted in red. Bars represent the percentage of species endemic to the respective hotspot. Numbers in parentheses represent total and endemic species known to science, respectively. The island of Borneo includes the political divisions of Brunei, Malaysia and Indonesia. The Indo-Burma hotspot includes part of Bhutan, Nepal, eastern India, southern China, as well as islands such as Hainan and the Andamans. Details of biodiversity hotspot boundaries, and numbers of total and endemic species within each hotspot were taken from Conservation International [74].

The current state of biodiversity

Three plant and eight animal species have been listed as 'extinct' in Southeast Asia by the International Union for the Conservation of Nature and Natural Resources (IUCN) [12]. Because Southeast Asia has a fairly recent history of large-scale deforestation (i.e. over the past two centuries), many of the native species of the region, such as rare long-lived trees, might be persisting as 'living dead' and are doomed to extinction owing to isolation caused by the fragmentation of habitats [3]. Therefore, although the actual number of extinct species in the region is not presently alarming, the level of endangerment of extant species reveals the seriousness of threats, such as deforestation, that are faced by the regional biota. The number of threatened species in Southeast Asia, including those in the IUCN categories of 'critically endangered' (CE), 'endangered' (EN) and 'vulnerable' (VU) ranges from 20 (CE) to 686 (VU) species for vascular plants, six to 91 species for fish, zero to 23 species for amphibians, four to 28 species for reptiles, seven to 116 species for birds, and five to 147 species for mammals (see Online Supplementary Material) [12]. The loss of many of these regional populations is likely to result in global extinctions because

of the high proportion of endemic species (Figure 1; see Online Supplementary Material) [13]. For example, 59.6% of the 29 375 vascular plant species in Indonesia do not occur anywhere else (see Online Supplementary Material) [13].

Drivers of biodiversity threat

Here, we discuss the drivers of biodiversity loss and endangerment in Southeast Asia, which might act either alone or in concert.

Forest conversion

The conversion of natural habitat to other land uses is the major driving force behind worldwide biodiversity loss [14–17]. Most of Southeast Asia was under forest cover 8000 years ago (see Online Supplementary Material) [18], but large-scale deforestation in the region began during the 1800s as a result of agricultural expansion that was needed to meet the increasing local and global demand for rice *Oryza sativa* [19]. The planting of perennial export crops, including rubber *Hevea brasiliensis*, oil palm *Elaeis guineensis* and coconut *Cocos nucifera*, also accounted for 20–30% of the total cultivated area of the region [19]. After

1950, increasing demand for Asian timber led to the proliferation of commercial logging activities [19]. Southeast Asian rain forests are particularly valuable to the logging industry for domestic consumption and export because its diverse dipterocarp species can be grouped for sale into just a few end-use classes [20]. Between 1880 and 1980, Southeast Asia experienced an average annual forest loss of 0.3% [19]. During the past decade, the loss of 'natural forest' in the region has continued at a rate of 1.4% y^{-1} (see Online Supplementary Material) [13,21], which was higher than the deforestation rates of other speciose tropical regions, such as Central America and the Caribbean (1.2%), and South America (0.5%) [13]. Currently, less than half (41.3–44.2%) of the original forests of Southeast Asia remain (see Online Supplementary Material) [21,22].

Deforestation in Southeast Asia has resulted in collateral impacts on its rich and unique biodiversity, with Singapore being the most heavily affected country to date (Box 1; see Online Supplementary Material). The most immediate impact of logging activities is the alteration of the unique multilayered and closed tropical forest canopy [20]. Reductions in canopy height, surface area and the crown size of selectively logged Malaysian forests are still evident after four decades of regeneration [23]. In Borneo, the species richness of trees was shown to be negatively associated with the intensity of logging activities [24]. Seedlings and saplings in these logged sites were also species poor [24], suggesting that logged forests require long periods of time to recover their original plant richness.

Conversion of land to agricultural use has even more detrimental impacts, owing to the depletion of soil nutrients and erosion following intense agricultural activities [25]. Even after a century of succession, the plant richness of abandoned agricultural areas in Singapore was only 60% of that of primary forests [26]. The selective logging of dipterocarps is also likely to decrease their reproductive success by reducing the extent and intensity of interspecific mast fruiting, which has been suggested to be a reproductive strategy that has evolved to satiate seed predators [11]. Logging has led to recruitment failure of dipterocarps within a national park in Borneo, which is now surrounded by logged forest, and logging has also exacerbated local El Niño events [11]. Recent studies reveal a trend of declining species richness and population density with increasing forest disturbance through logging activities, agriculture or urbanization across a range of Southeast Asian taxa, including termites [27], dung beetles [28], ants [29], bees [30], butterflies [31], moths [32], birds [33,34] and mammals [35–40]. Furthermore, the loss of animals that produce ecological services, such as seed dispersal, nutrient recycling and pollination, might further impede forest regeneration in the disturbed areas.

Forest fires

Although forest fires have always occurred in Southeast Asia, a combination of factors, such as poor land conversion practices, logging and more-intense El Niño events, act in concert to increase the chances of catastrophic fires

in the region [39,40]. For example, logging can increase the vulnerability of forests to fire by opening up the canopy, creating piles of flammable wooden debris and facilitating access to forests for people. In 1997–1998, up to 5 million ha of forests in Indonesia (Sumatra and Kalimantan) were consumed by forest fires [41]. It was estimated that 4.6% of the canopy trees, as well as 70–100% of seedlings and 25–70% of saplings, were destroyed in Sumatra (Barisan Selatan National Park). Owing to the loss of fruiting trees, many avian frugivores, such as the helmeted hornbill *Buceros vigil*, experienced population declines of up to 50% [42]. Immediately following the fires, other animals, such as the siamang *Holobates syndactylus*, the largest of the gibbons, disappeared completely from the burnt areas [39]. Siamang groups whose territory included forest that was either burned or adjacent to burned areas were significantly smaller and experienced higher infant and juvenile mortality than did groups not affected by fire [43].

Hunting for bushmeat

Humans have been hunting wildlife in Southeast Asia for at least 40 000 years [44,45]. However, hunting pressure has increased immensely with increasing human densities and declining forest areas in the region. The situation is further exacerbated by factors such as logging, creation of roads, better hunting equipment and ineffective wildlife protection [46–48]. Wildlife is currently being extracted from tropical forests at more than six times the sustainable rate [49]. In Sarawak, Malaysia, an estimated 2.6 million animals were shot and 23 500 tons of wildlife meat are consumed on annual basis [49,50]. A similar trend is reported from Sabah, Malaysia, where an estimated 108 million animals are hunted for bushmeat each year [50]. Such figures are alarming, especially given that intense hunting can depress animal densities [50].

Wildlife trade

Southeast Asia is a major hub of wildlife trade (see Online Supplementary Material) [51]. In 2000, the net legal export of lizard and snakeskins from Indonesia was 29.4% and 28.2% of global exports, respectively. Even a highly urbanized country, such as Singapore, was an active trader of wildlife and wildlife products, with a total net import of 7093 live animals and a total net export of 301 905 animal skins in 2000 alone (see Online Supplementary Material). The alarming numbers of wildlife resources that are legally traded in Southeast Asia are, in all likelihood, a gross underestimation of the actual volume of wildlife traded in the region. Trade unaccounted for includes illegal and unrecorded wildlife traffic, as well as many animals that died before reaching their intended destination or were sold in local markets. The lucrative pet trade in Southeast Asia is the major driving force behind its wildlife trade [52,53]. For example, rampant trapping for the pet trade, coupled with habitat conversion, has driven the endemic Bali starling *Leucopsar rothschildi* to the verge of extinction over the past three decades [54]. Currently, there are <20 wild individuals of this species, which are restricted to the Bali Barat National Park [54]. In spite of being listed on Appendix 1

Box 1. Singapore case study: a recipe for disaster

From both a scientific and management perspective, it would be informative if we could excise hypothetically a representative Southeast Asian site, allow it to fulfil its economic potential, and document the consequent losses of natural habitats and biodiversity, all within a greatly accelerated time frame. Perhaps it is both depressing and fortunate that Singapore is one such ecological 'worst case scenario'.

Singapore has experienced exponential population growth, from ~150 subsistence-economy villagers in the early 1800s to four million people in 2002 [13,72]. In particular, Singapore has transformed itself from a developing country of squatters and slums to a developed metropolis of economic prosperity within the past few decades and, thus, has been widely regarded by the leaders of regional developing countries as the ideal economic model.

However, the success of Singapore came with a hefty price, paid for by its biodiversity. The island has suffered massive

deforestation (>95%), initially from the cultivation of short-term cash crops (e.g. gambier *Uncaria gambir* and rubber *Hevea brasiliensis*) and subsequently from urbanization and industrialization [72]. A recent study by Brook *et al.* [3] showed substantial rates of documented (observed) and inferred (based on what could have occurred in Singapore before habitat loss) extinctions, with most extinct taxa (34–87%) being species of butterflies, fish, birds and mammals (Figure 1). Similar environmental scenarios are already unfolding on a much larger scale in other Southeast Asian countries, such as Indonesia [73]. Extrapolations based on the species-area model calibrated to the biodiversity losses in Singapore indicate that the current rate of habitat destruction in Southeast Asia will result in the loss of 13–42% of regional populations of all species by 2100, at least half of which could represent global species extinctions (Figure 1) [3].

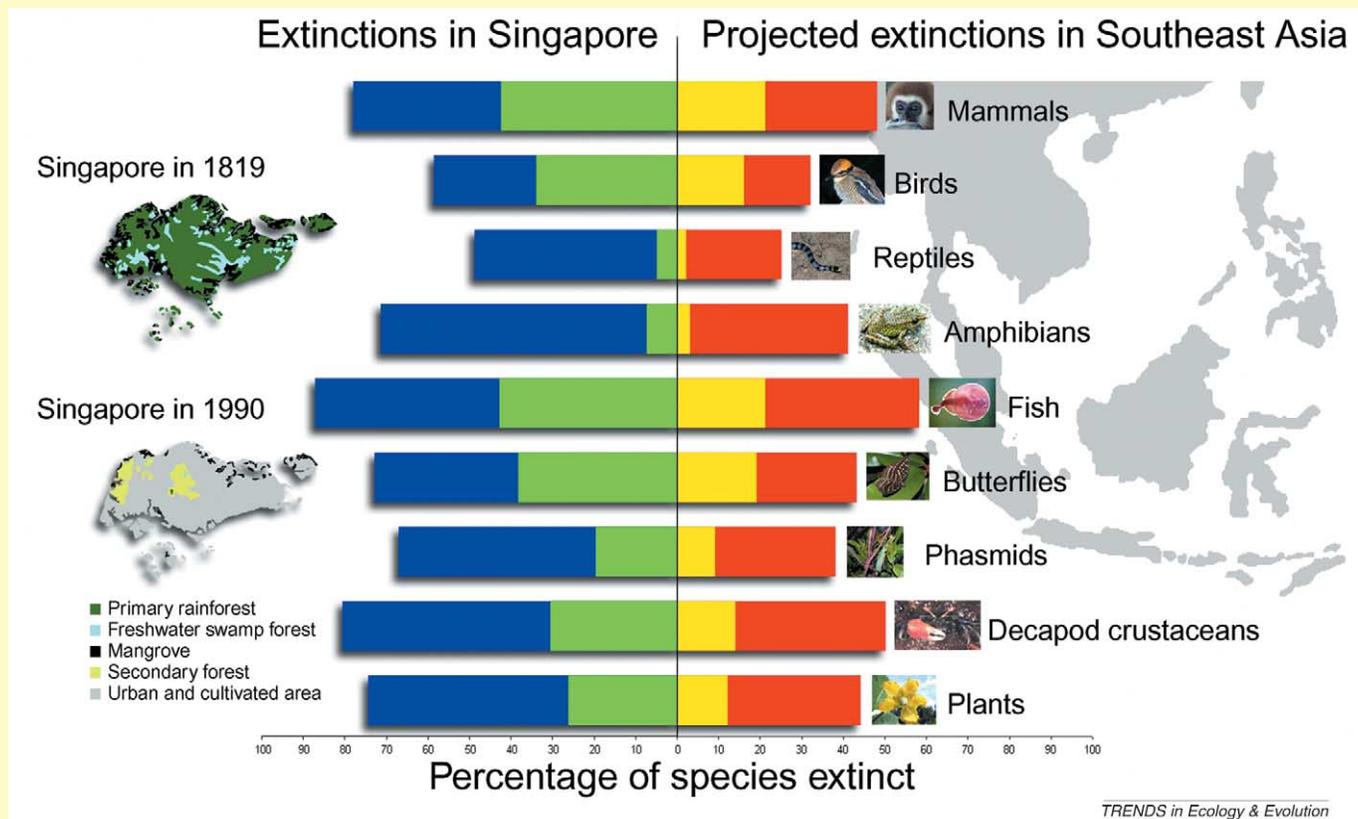


Figure 1. Population extinctions in Singapore and Southeast Asia. Green and blue bars represent recorded and inferred extinctions in Singapore, respectively. Yellow and red bars represent minimum and maximum projected extinctions in Southeast Asia, respectively.

of the Convention on International Trade in Endangered Species (CITES) in 1970 and having been protected in Indonesia since 1971, 19 individuals were observed being sold illicitly in shops in Singapore in 1979 and 16 individuals were observed in cages in Denpasar (Bali, Indonesia) in 1982 [54].

Many animal and plant products are used in traditional Chinese medicine, which dates back 5000 years. Trade in the raw materials of traditional Chinese medicine has a detrimental impact on many vertebrates in Southeast Asia, including tigers, bears, rhinos, turtles, snakes, tokay geckos, pangolins, monkeys and swiftlets [5]. This is exemplified by the Sumatran tiger *Panthera tigris sumatrae*, from which body parts such as bones and penises are used in traditional medicine. Between 1975

and 1992, South Korea imported 6128 kg of tiger bones (340 kg y^{-1}), of which 60% originated from Indonesia [55]. The demand for the Sumatran tiger continues to threaten the remaining wild population of an estimated 500 individuals [56].

Other potential drivers

Sala *et al.* [14] showed that, relative to the overwhelming impact of forest conversion, other drivers, such as climate change, nitrogen deposition, invasive species and atmospheric CO_2 change, are not expected to have significant effects on biodiversity in tropical terrestrial ecosystems. Climate change was shown to have the largest proportional effect on biodiversity in extreme environments, such as the arctic and boreal zones, and the least effect in

the tropics, with only montane areas being substantially affected [14]. Nitrogen deposition is expected to have the largest impact on biodiversity in areas that are most nitrogen limited, such as temperate forests [57]. Invasive species are also not expected to impact tropical ecosystems greatly because various abiotic and biotic factors, such as high species diversity, minimize the probability of successful establishment by invaders in undisturbed communities in the tropics [58]. The increased atmospheric concentration of CO₂ is expected to have a large impact on biodiversity mainly in areas where plant growth is most limited by water availability and where there is a mixture of C₃ and C₄ species, such as grassland and savannas. This is because of known species differences in the effect of CO₂ on water-use efficiency [59]. Nevertheless, in spite of these findings, the relative importance and long-term implications of these drivers of biodiversity threats and, more critically, their synergistic impacts on the biodiversity of Southeast Asia, remain poorly understood.

Conservation challenges

The outlook for the biodiversity of Southeast Asia appears bleak, owing to several key social, scientific and logistical conservation challenges faced by the region. The major challenges in mitigating the imminent threats to its biodiversity are primarily socioeconomic in origin, including population growth, poverty, chronic shortage of conservation resources (both expertise and funding) and corrupt national institutions. As regional societies strive to match the living standards of developed nations, environmental issues are inevitably marginalized. Increasing human population density is a primary socioeconomic driver of forest loss in Asia [60] and, in Southeast Asia, both this and economic growth are positively associated with forest loss (Figure 2).

Research on Southeast Asian biodiversity over the past 20 years has also been neglected in comparison to other tropical regions. We compared all of the internationally peer-reviewed research articles about biodiversity from

the Biological Abstracts data base (excluding exclusively marine studies) that were published between 1983 and 2003. The number of scientific publications about Southeast Asian biodiversity was fewer than was expected for its forest area compared with other tropical regions, including Central America and the Caribbean, Sub-Saharan Africa, and South America. Furthermore, the distribution of research effort in Southeast Asia was also taxonomically biased. For example, there were more research papers on mammals than would be expected if studies were distributed according to the relative species richness of taxonomic groups; however, there remains a dearth of research on other important taxa, such as vascular plants, invertebrates and fish. To remedy the paucity of biodiversity studies in Southeast Asia, particularly in poorly studied taxonomic groups, collaborative research efforts of regional and international expertise on Southeast Asian biotas are urgently needed [61].

The diversity of habitat types in Southeast Asia is another major challenge for the conservation of its biodiversity [62] and there are numerous habitats that remain poorly studied. For example, recent studies reveal that the highly acidic blackwaters of peat swamps have much higher biodiversity and productivity than was previously thought [63], and contain many hitherto unrecognized rare and endangered fish species [64]. Furthermore, recent studies of limestone formations in Southeast Asia revealed the presence of many endemic and highly specialized taxa [65]. Additionally, large tracts of rain forests and mangroves in Southeast Asia have yet to be surveyed by professional biologists [66]. The lack of biodiversity studies in such neglected habitats is a serious impediment to the conservation of Southeast Asian biodiversity because sound biological knowledge is needed to prioritize conservation areas and habitat, and to model sustainable use of resources, such as timber and bushmeat. Currently, such biological understanding in Southeast Asia lags behind those of other regions.

Protected areas (both existing and future) are the main hope for preserving the biodiversity of Southeast Asia.

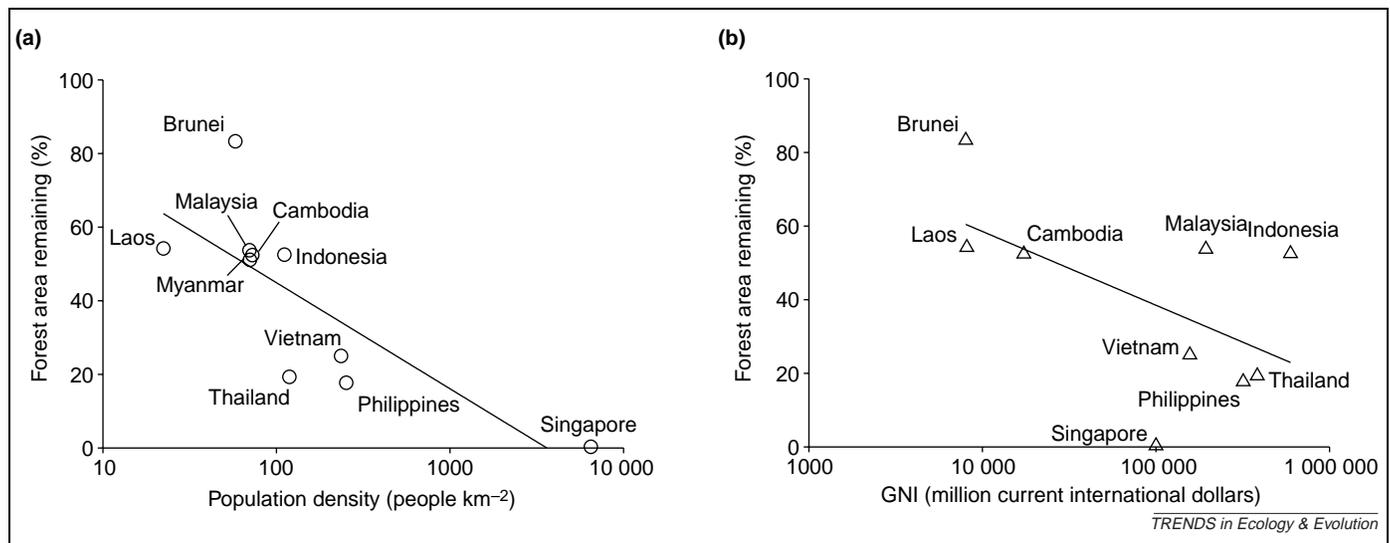


Figure 2. Socioeconomic correlates of forest loss. The proportion of forest area remaining in Southeast Asian countries correlated with (a) population density in 2000 ($r = -0.78$, $P = 0.008$); and (b) Gross National Income (GNI) at current international dollars in 2000 ($r = -0.57$, $P = 0.111$) [13]. GNI of Brunei was taken in 1998.

Currently, there are 2262 protected areas in the region, which cover a total land area of 58 million ha (13.4%) (see Online Supplementary Material [13,67]. More than half of the total cover of these areas is in Indonesia (24 million ha) and Malaysia (10 million ha). Recent studies show that, in spite of their 'protected' status, some of these areas have become increasingly isolated and deforested [68]. For example, from 1985 to 2001, the forest cover of lowland protected areas in West Kalimantan, Indonesia, was reduced by 1.85 million ha (63%). In addition, of the 64 remaining forest fragments, only 16 were considered large enough (> 10 000 ha) to support intact vertebrate fauna [68–70]. Protected areas also vary considerably in their degree of effectiveness [71]. Bruner *et al.* showed that the effectiveness of such areas for biodiversity conservation was correlated most strongly with the density of guards [71]. Therefore, the enforcement and management of parks is crucial to the success of protected areas in conserving the native habitats and biodiversity of Southeast Asia.

Conclusion

Massive anthropogenic habitat modifications, forest fires and the overexploitation of wildlife in Southeast Asia are clear-and-present dangers to its biodiversity. In spite of the pessimistic outlook, there are ways to conserve at least some of the regional natural resources. Given that many of the drivers of biodiversity loss (e.g. international demand for rain forest timber and elevated global CO₂ levels) are issues that transcend national boundaries, any realistic solution will need to involve a multi-national and multi-disciplinary strategy, including political, socioeconomic and scientific input, in which all major stakeholders (governmental, non-governmental, national and international organizations) must partake. Key solutions should include enhancing public environmental awareness, delineating adequately protected reserves and providing economic incentives for conservation.

Acknowledgements

We thank three anonymous reviewers for their comments about the article. We also thank Tom Brooks for help. This study was supported by the National University of Singapore (R-154-000-210-112).

Supplementary data

Supplementary data associated with this article can be found at [doi:10.1016/j.tree.2004.09.006](https://doi.org/10.1016/j.tree.2004.09.006)

References

- Laurance, W.F. (1999) Reflections on the tropical deforestation crisis. *Biol. Conserv.* 91, 109–117
- Achard, F. *et al.* (2002) Determination of deforestation rates of the world's humid tropical forests. *Science* 297, 999–1002
- Brook, B.W. *et al.* (2003) Catastrophic extinctions follow deforestation in Singapore. *Nature* 424, 420–423
- Myers, N. *et al.* (2000) Biodiversity hotspots for conservation priorities. *Nature* 403, 853–858
- Mittermeier, R.A. *et al.* (1999) *Hotspots: Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions*, Cemex, Conservation International and Agrupacion Sierra Madre
- Jablonski, N.G. (1993) Quaternary environments and the evolution of primates in East Asia, with notes on two new specimens of fossil Cercopithecidae from China. *Folia Primatol.* 60, 118–132
- Meijaard, E. (2004) Biogeographic history of the Javan leopard *Panthera pardus* based on a craniometric analysis. *J. Mammal.* 85, 302–310
- Gathorne-Hardy, F.J. *et al.* (2002) Quaternary rainforest refugia in south-east Asia: using termites (Isoptera) as indicators. *Biol. J. Linn. Soc.* 75, 453–466
- Audley-Charles, M.G. (1983) Reconstruction of eastern Gondwanaland. *Nature* 306, 48–50
- Steppan, S.J. *et al.* (2003) Molecular phylogeny of the endemic Philippine rodent *Apomys* (Muridae) and the dynamics of diversification in an oceanic archipelago. *Biol. J. Linn. Soc.* 80, 699–715
- Curran, L.M. *et al.* (1999) Impact of El Niño and logging on canopy tree recruitment in Borneo. *Science* 286, 2184–2188
- IUCN (2003) *2003 IUCN Red List of Threatened Species*, (<http://www.redlist.org>)
- WRI (2003) *World Resources 2002–2004: Decisions for the Earth: Balance, Voice, and Power*, United Nations Development Programme, United Nations Environment Programme, World Bank and World Resources Institute (<http://www.wri.org>)
- Sala, O.E. *et al.* (2000) Global biodiversity scenarios for the year 2100. *Science* 287, 1770–1774
- Brooks, T.M. *et al.* (1999) Threat from deforestation to montane and lowland birds and mammals in insular South-east Asia. *J. Anim. Ecol.* 68, 1061–1078
- Brooks, T.M. *et al.* (2002) Habitat loss and extinction in the hotspots of biodiversity. *Conserv. Biol.* 16, 909–923
- Brooks, T.M. *et al.* (1997) Deforestation predicts the number of threatened birds in insular Southeast Asia. *Conserv. Biol.* 11, 382–394
- Billington, C. *et al.* (1996) *Estimated Original Forest Cover Map – A First Attempt*, World Conservation Monitoring Centre (<http://www.unep-wcmc.org/forest/original.htm>)
- Flint, E.P. (1994) Changes in land use in South and Southeast Asia from 1880 to 1980: a data base prepared as part of a coordinated research program on carbon fluxes in the tropics. *Chemosphere* 29, 1015–1062
- Whitmore, T.C. (1998) *Tropical Rain Forests*, Oxford University Press
- FAO (2001) *Global Forest Resources Assessment 2000: Main Report. FAO Forestry Paper No. 140*, FAO (<http://www.fao.org>)
- Iremonger, S. *et al.* (1997) A statistical analysis of global forest conservation. In *A Global Overview of Forest Conservation* (Iremonger, S. *et al.*, eds), Center for International Forestry Research (CIFOR) and World Conservation Monitoring Centre (<http://www.unep-wcmc.org/forest/data/cdrom2>)
- Okuda, T. *et al.* (2003) Effect of selective logging on canopy and stand structure and tree species composition in a lowland dipterocarp forest in Peninsular Malaysia. *For. Ecol. Manage.* 175, 297–320
- Foody, G.M. and Cutler, M.E.J. (2003) Tree biodiversity in protected and logged Bornean tropical rain forests and its measurement by satellite remote sensing. *J. Biogeogr.* 30, 1053–1066
- Grubb, P.J. *et al.* (1994) Mineral nutrient status of coastal hill dipterocarp forest and adinandra belukar in Singapore: analysis of soil, leaves and litter. *J. Trop. Ecol.* 10, 559–577
- Turner, I.M. *et al.* (1997) Tree species richness in primary and old secondary tropical forest in Singapore. *Biodiv. Conserv.* 6, 537–543
- Gathorne-Hardy, F.J. *et al.* (2002) A regional perspective on the effects of human disturbance on the termites of Sundaland. *Biodiv. Conserv.* 11, 1991–2006
- Davis, A.J. *et al.* (2001) Dung beetles as indicators of change in the forests of northern Borneo. *J. Appl. Ecol.* 38, 593–616
- Bruehl, C.A. *et al.* (2003) Size does matter: effects of tropical rainforest fragmentation on the leaf litter ant community in Sabah, Malaysia. *Biodiv. Conserv.* 12, 1371–1389
- Liow, L.H. *et al.* (2001) Bee diversity along a disturbance gradient in tropical lowland forests of south-east Asia. *J. Appl. Ecol.* 38, 180–192
- Koh, L.P. and Sodhi, N.S. Importance of reserves, fragments and parks for butterfly conservation in a tropical urban landscape. *Ecol. Appl.* (in press)
- Beck, J. *et al.* (2002) From forest to farmland: diversity of geometrid moths along two habitat gradients on Borneo. *J. Trop. Ecol.* 18, 33–51
- Sodhi, N.S. (2002) A comparison of bird communities of two fragmented and two continuous Southeast Asian rainforests. *Biodiv. Conserv.* 11, 1105–1119

- 34 Thiollay, J.-M. (1995) The role of traditional agroforests in the conservation of rain and forest bird diversity in Sumatra. *Conserv. Biol.* 9, 335–353
- 35 Robertson, J.M.Y. and van Schaik, C.P. (2001) Causal factors underlying the dramatic decline of the Sumatran orang-utan. *Oryx* 35, 26–38
- 36 Heydon, M.J. and Bulloh, P. (1997) Mousedeer densities in a tropical rainforest: the impact of selective logging. *J. Appl. Ecol.* 34, 484–496
- 37 O'Brien, T.G. *et al.* (2003) Crouching tigers, hidden prey: Sumatran tiger and prey populations in a tropical forest landscape. *Anim. Conserv.* 6, 131–139
- 38 Meijaard, E. and Nijman, V. (2000) The local extinction of the proboscis monkey *Nasalis larvatus* in Pulau Kaget Nature Reserve, Indonesia. *Oryx* 34, 66–70
- 39 O'Brien, T.G. *et al.* (1998) *Effects of the 1997 Fires on the Forest and Wildlife of the Bukit Barisan Selatan National Park, Sumatra. Wildlife Conservation Society Working Paper No. 13*, Wildlife Conservation Society
- 40 Taylor, D. *et al.* (1999) Environmental change and rain forests on the Sunda shelf of Southeast Asia: drought, fire and the biological cooling of biodiversity hotspots. *Biodiv. Conserv.* 8, 1159–1177
- 41 Schweithelm, J. (1998) *The Fire this Time, an Overview of Indonesia's Forest Fires in 1997/1998*, World Wide Fund for Nature
- 42 Kinnaid, M.F. and O'Brien, T.G. (1998) Ecological effects of wildfire on lowland rainforest in Sumatra. *Conserv. Biol.* 12, 954–956
- 43 O'Brien, T.G. *et al.* (2003) Fire, demography and the persistence of siamang (*Symphalangus syndactylus*: Hylobatidae) in a Sumatran rainforest. *Anim. Conserv.* 6, 115–121
- 44 Zuraina, M. (1982) The West Mouth, Niah, in the prehistory of Southeast Asia. *Sarawak Mus. J.* 31, 1–20
- 45 Milner-Gulland, E.J. and Bennett, E.L. (2003) Wild meat: the bigger picture. *Trends Ecol. Evol.* 18, 351–357
- 46 Ling, S. *et al.* (2002) No new recipes for bushmeat. *Oryx* 36, 330–330
- 47 Bulte, E.H. and Horan, R.D. (2002) Does human population growth increase wildlife harvesting? An economic assessment. *J. Wildl. Manage.* 66, 574–580
- 48 Smith, R.J. *et al.* (2003) Governance and the loss of biodiversity. *Nature* 426, 67–70
- 49 Bennett, E.L. (2002) Is there a link between wild meat and food security? *Conserv. Biol.* 16, 590–592
- 50 Bennett, E.L. *et al.* (2000) Saving Borneo's bacon: the sustainability of hunting in Sarawak and Sabah. In *Hunting for Sustainability in Tropical Forests* (Robinson, J.B. and Bennett, E.L., eds), pp. 305–324, Columbia University Press
- 51 WCMC (2002) *Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) Annual Report Data*, World Conservation Monitoring Centre, CITES Trade Database (<http://www.cites.org>)
- 52 Robinson, J.M. (2001) The dynamics of avicultural markets. *Environ. Conserv.* 28, 76–85
- 53 Beissinger, S.R. (2001) Trade of live birds: potential, principles and practices of sustainable use. In *Conservation of Exploited Species* (Reynolds, J.D. *et al.*, eds), pp. 182–202, Cambridge University Press
- 54 van Balen, S.B. *et al.* (2000) Status and distribution of the endemic Bali starling *Leucopsar rothschildi*. *Oryx* 34, 188–197
- 55 Mills, J.A. (1993) Tiger bone trade in South Korea. *Cat News* 19, 13–16
- 56 Seidensticker, J. *et al.* (1999) *Riding the Tiger: Tiger Conservation in Human-dominated Landscapes*, Cambridge University Press
- 57 Tilman, D. (1993) Species richness of experimental productivity gradients: how important is colonization limitation? *Ecology* 74, 2179–2191
- 58 Rejmanek, M. (1996) Species richness and resistance to invasion. In *Biodiversity and Ecosystem Processes in Tropical Forests* (Orlans, G.H. *et al.*, eds), pp. 153–172, Springer-Verlag
- 59 Jackson, R.B. *et al.* (1994) CO₂ alters water use, carbon gain, and yield for the dominant species in a natural grassland. *Oecologia* 98, 257–262
- 60 Bawa, K.S. and Dayanandan, S. (1997) Socioeconomic factors and tropical deforestation. *Nature* 386, 562–563
- 61 Sodhi, N.S. and Liow, L.H. (2000) Improving conservation biology research in Southeast Asia. *Conserv. Biol.* 14, 1211–1212
- 62 Dennis, C. and Aldhous, P. (2004) A tragedy with many players. *Nature* 430, 396–398
- 63 Ng, P.K.L. *et al.* (1992) *The Conservation of the Fish and Other Aquatic Fauna of the North Selangor Peat Swamp Forest and Adjacent Areas*, Asian Wetland Bureau, Institute of Advanced Studies University of Malaya, World Wide Fund for Nature and National University of Singapore
- 64 Ng, P.K.L. *et al.* (1994) Diversity and conservation of blackwater fishes in Peninsular Malaysia, particularly in the north Selangor peat swamp forest. *Hydrobiologia* 285, 203–218
- 65 Jubertie, C. and Decu, V., eds (2001) *Encyclopaedia Biospeologica, Tome III*. Société Internationale de Biospéologie
- 66 Ng, H.H. and Rachmatika, I. (1999) The catfishes (Teleostei: Siluriformes) of Bentuang Karimun National Park, West Kalimantan, Indonesia. *Raffles B. Zool.* 47, 167–183
- 67 WCMC (2003) *World Database on Protected Areas (WDPA) Version 6*, World Database on Protected Areas Consortium (http://www.unep-wcmc.org/protected_areas)
- 68 Curran, L.M. *et al.* (2004) Lowland forest loss in protected areas of Indonesian Borneo. *Science* 303, 1000–1003
- 69 MacKinnon, K.S. *et al.* (1996) *The Ecology of Kaimantan*, Periplus Editions
- 70 Lambert, F.R. and Collar, N.J. (2002) The future for Sundaic lowland forest birds: long-term effects of commercial logging and fragmentation. *Forktail* 18, 127–146
- 71 Bruner, A.G. *et al.* (2001) Effectiveness of parks in protecting tropical biodiversity. *Science* 291, 125–128
- 72 Corlett, R.T. (1992) The ecological transformation of Singapore 1819–1990. *J. Biogeogr.* 19, 411–420
- 73 Jepson, P. *et al.* (2001) The end for Indonesia's lowland forests? *Science* 292, 859–861
- 74 Conservation International (2004) *Biodiversity Hotspots*, Conservation International (<http://www.biodiversityhotspots.org>)



Trends in Ecology and Evolution backfiles

Would you like to know what was in the first issue of TREE ?

Ever wondered how you could get an electronic version of articles from the early issues of the journal?

Issues of TREE, from the very first one in July 1986 to the current issue, are now available online at

<http://www.sciencedirect.com>





A marine fish follows Wallace's Line: the phylogeography of the three-spot seahorse (*Hippocampus trimaculatus*, Syngnathidae, Teleostei) in Southeast Asia

Sara A. Lourie* and Amanda C. J. Vincent

Project Seahorse, Department of Biology,
McGill University, Montréal, Québec, Canada

ABSTRACT

Aim To test the potential of two contrasting biogeographical hypotheses ('Indian/Pacific Ocean Basin' vs. 'Wallace's Line') to explain the distribution of genetic diversity among populations of a marine fish in Southeast Asia.

Location The marine waters of Asia and Southeast Asia: from India to Japan, and east to the Indonesian islands of Sulawesi and Flores.

Methods We sequenced a 696 base pair fragment of cytochrome *b* DNA of 100 individuals of *Hippocampus trimaculatus* Leach 1814 (three-spot seahorse), obtained from across its range. We tested our hypotheses using phylogenetic reconstructions and analyses of molecular variance.

Results Significant genetic divergence was observed among the specimens. Two distinct lineages emerged that diverged by an average of 2.9%. The genetic split was geographically associated, but surprisingly it indicated a major east–west division similar to the terrestrial Wallace's Line ($\Phi_{ST} = 0.662$, $P < 0.001$) rather than one consistent with an Indian-Pacific ocean basin separation hypothesis ($\Phi_{ST} = 0.023$, $P = 0.153$). Samples from east of Wallace's Line, when analysed separately, however, were consistent with an Indian/Pacific Ocean separation ($\Phi_{ST} = 0.461$, $P = 0.005$). The degree of genetic and geographical structure within each lineage also varied. Lineage A, to the west, was evolutionarily shallow (star-like), and the haplotypes it contained often occurred over a wide area. Lineage B to the east had greater genetic structure, and there was also some evidence of geographical localization of sublineages within B.

Main conclusions Our results indicate that the genetic diversity of marine organisms in Southeast Asia may reflect a more complex history than the simple division between two major ocean basins that has been proposed by previous authors. In particular, the east–west genetic division observed here is novel among marine organisms examined to date. The high haplotype, but low nucleotide diversity to the west of Wallace's Line is consistent with post-glacial colonization of the Sunda Shelf. Additional data are needed to test the generality of these patterns.

Keywords

Marine biogeography, Sunda Shelf, AMOVA, cytochrome *b*, Pleistocene, sea level, Syngnathidae, mitochondrial DNA.

*Correspondence: Sara A. Lourie, Project Seahorse, Redpath Museum, McGill University, 859 Sherbrooke St. W., Montréal, Québec, H3A 2KG, Canada.

E-mail: sara.lourie@mail.mcgill.ca

Present address: Amanda C. J. Vincent, Project Seahorse, Fisheries Centre, University of British Columbia, 2204 Main Mall, Vancouver, British Columbia V6T 1Z4, Canada.

INTRODUCTION

Uncovering genetic patterns among populations is the first step towards disentangling the effects of historical, vs. present-day

forces in structuring biodiversity. Marine patterns are still poorly known, especially within highly diverse Southeast Asia. In the terrestrial realm, Wallace's Line separates terrestrial flora and fauna into Asian (west) and Australasian (east) elements

(Van Oosterzee, 1997). Marine organisms, by contrast, are expected to show a north-east to south-west division perpendicular to Wallace's Line, reflecting Pleistocene separation of Indian Ocean vs. Pacific Ocean populations (Benzie, 1998).

Past geological and climatic events (e.g. plate tectonic movement or glacial episodes) have undoubtedly played a major role in terrestrial biogeography, but a similar role for historical geological events has yet to be comprehensively explored for marine species. Exposed land-bridges that enabled terrestrial organisms to intermingle (e.g. the Isthmus of Panama) would have resulted in vicariance of marine populations (Knowlton & Weigt, 1998), and former oceanic basins that separated terrestrial continents (e.g. Tethys Seaway) would have enabled marine taxa to extend their ranges (McKenzie, 1991).

Geologically, Southeast Asia is one of the most complex areas in the world (Hall, 2002). Such geological complexity has most likely contributed to the tremendous biological diversity of the region (Michaux, 1991). It may also have led to distinctive, congruent distribution patterns across species; the most famous terrestrial example being Wallace's Line (Huxley, 1868; Whitmore, 1981) (Fig. 1). Two potential geological

causes for Wallace's Line are (1) the collision of Asian and Australasian tectonic elements, between c. 15 and 3 Ma, which brought terrestrial organisms from different realms into close contact (Hall, 2002), and (2) more recently, periods of lowered sea-levels (up to at least 120 m below present levels) associated with Pleistocene glaciations (2.4 Ma–10,000 years ago) which resulted in the formation of land-bridges between mainland Asia, Borneo and other western Indonesian islands, and among Philippine islands (Voris, 2000).

Marine passage between the Pacific (north/east) and Indian (south/west) Oceans would have been drastically reduced when the Asian and Australasian elements collided in the Miocene, and almost completely interrupted when sea levels were lowered in the Pleistocene (Hall, 1998; Voris, 2000). This has led to the proposal that such barriers would have been sufficient to result in biogeographical discontinuities in marine taxa between the two ocean basins at the species, subspecies or population level (Springer & Williams, 1990; Randall, 1998).

Historical biogeographical patterns are often believed to be obscured by extensive dispersal and high contemporary gene-flow in marine species (Bohonak, 1999). However, recent

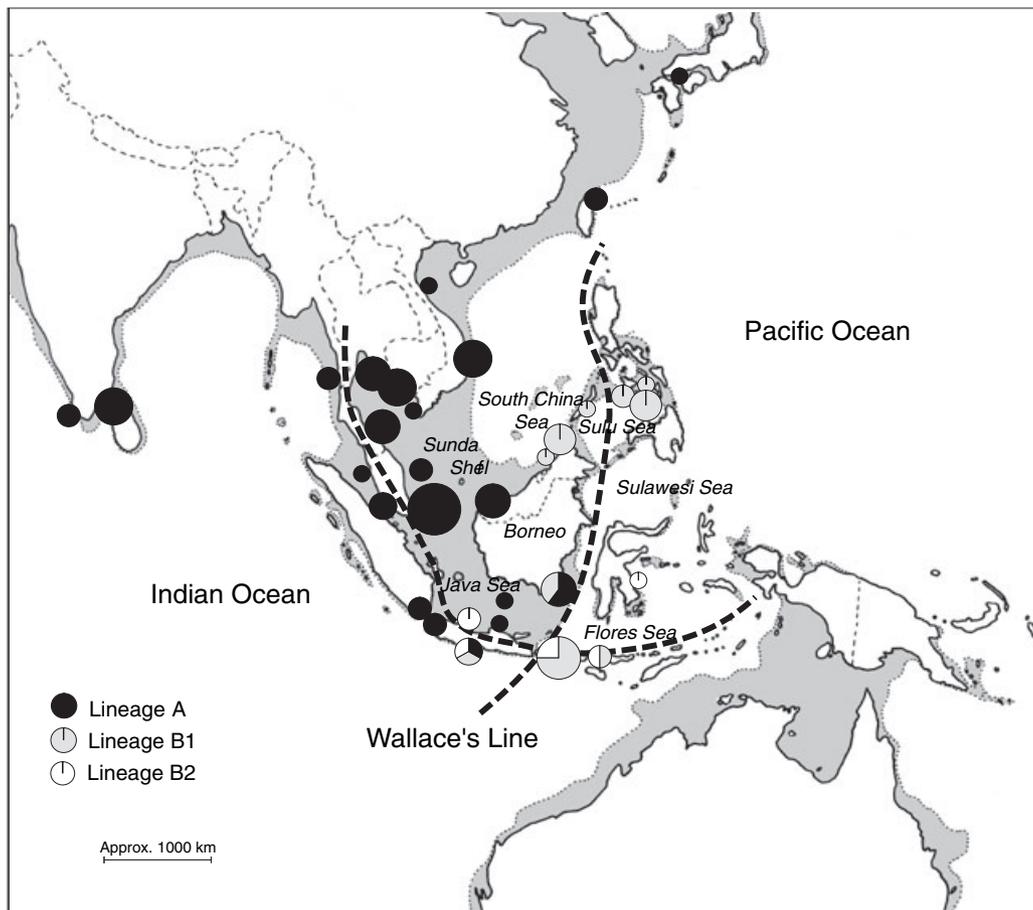


Figure 1 Sample locations and geographical distribution of lineages of *Hippocampus trimaculatus* Leach. The location of Wallace's Line and the division between the Indian and Pacific Oceans as defined in the text are shown. The size of the pies is proportional to the number of individuals sampled from each location. Shaded sea areas are continental shelves that would have been exposed to the air during periods of low sea-level.

studies that show significant phylogeographical structure among marine populations, are beginning to question the paradigm of widespread effective dispersal over ecological and evolutionary time-frames (Barber *et al.*, 2002). Particular life histories, behaviours and the vagaries of oceanic currents may, furthermore, lead some marine organisms to retain greater historical structure than previously assumed. Near-shore species that brood their young, have benthic larvae, or experience direct development, are likely to exhibit lower dispersal than pelagic species with broadcast spawning and young that spend weeks in the plankton (Doherty *et al.*, 1995). For example, damselfish populations have been shown to experience significant self-recruitment (Jones *et al.*, 1999; Swearer *et al.*, 1999), large numbers of gastropod larvae can be entrained in gyres and returned to their natal reef (Scheltema *et al.*, 1996), and larval behaviour can strongly affect realized marine fish dispersal (Leis & Carson-Ewart, 1998).

Based on their biology, seahorses are likely to be at the lower end of the marine fish dispersal continuum, and likely to retain historical patterns. All seahorses have internal fertilization, brood their offspring throughout larval development, are fully pigmented and independent when they are released from the male's pouch, and remain largely sedentary as adults (Foster & Vincent, 2004).

We assessed the phylogeographical pattern exhibited by the three-spot seahorse (*Hippocampus trimaculatus*) and tested the hypothesis that lowered sea-levels during the Pleistocene period (2.4 Ma–10,000 years ago) would have resulted in divergent populations in the Indian vs. the Pacific Oceans ('Ocean-Basin' hypothesis) (Springer & Williams, 1990). We tested this against an alternative that is consistent with the terrestrial biogeographical division known as Wallace's Line ('Wallace's Line' hypothesis) (Fig. 1).

MATERIALS AND METHODS

In order to test whether a north-east–south-west 'Indian/Pacific Ocean-Basin' hypothesis vs. an east–west 'Wallace's Line' hypothesis could account for genetic variation in the three-spot seahorse (*H. trimaculatus* Leach 1814) we first defined our area of analysis. We took the boundary between the ocean basins as following the spine of Peninsular Malaysia, then the north coasts of Sumatra, Java, Bali, Lombok, Flores, etc. (Wyrarki, 1961), and Wallace's Line as following the edge of the Sunda Shelf (Huxley, 1868) (Fig. 1). Specimens ($n = 100$) of *H. trimaculatus* were obtained from either side of Wallace's Line, and from both the Indian and Pacific Ocean basins (see Appendix 1). Dried specimens were obtained from as close to their source as possible in order to be confident of their locality of origin. Seventy per cent of the specimens came directly from fishers, and a further 12% from first-level buyers or researchers. Seventeen per cent were from traditional medicine shops where, through interviews, we had good confidence of their origin. *Hippocampus trimaculatus* lives at depths of > 15 m, on areas of gravel, sand and dead bivalves (Lourie *et al.*, 1999a), possibly in association with octocorals and sea-whips (Choo &

Liew, 2003). Most specimens would have been captured by small trawl boats. No seahorses were specifically killed for this project, and issues of seahorse conservation were raised, wherever possible, with fishers or traders.

Genetic methods

We used a small piece of tail muscle (*c.* 0.005–0.05 g dry weight) for DNA extraction, and a standard proteinase-K/phenol-chloroform protocol except that no salt was added at the ethanol precipitation step. A section of the cytochrome *b* gene was amplified in a 25 μ L reaction using a Perkin-Elmer 9600 Cycle Sequencer, and the following PCR reaction mix: 19.0 μ L H₂O, 2.5 μ L RedTaq buffer (10X), 0.5 μ L dNTP (10 mM), 1.0 μ L RedTaq, 0.5 μ L of each primer (10 μ M), 1.0 μ L (containing *c.* 5–50 ng) DNA, under the following PCR conditions: 94 °C 2:30 min (94 °C 0:30, 50 °C 0:30, 72 °C 1:15) 35 times, 72 °C 5:00. We designed and used the following seahorse-specific primers: forward shf 5'-CTACCTGCACCATCAAATATTTTC-3' or shf2 5'-TTGCAACCGCATTTTCTTCAG-3' and reverse shr2 5'-CGGAAGGTGAGTCCTCGTTG-3'.

PCR products were cleaned using QIAQuick columns (Qiagen, Mississauga, Ontario, Canada) or Millipore filters according to the manufacturer's instructions and sequenced in both directions in 10 μ L reactions using the following reaction mix: 4.5 μ L H₂O, 1.5 μ L buffer (5X), 0.5 μ L DMSO, 1.0 μ L Big Dye Terminator v.2.0 (ABI, Foster City, CA), 0.5 μ L primer (20 μ M), 2 μ L (containing *c.* 40 ng) DNA, under the following PCR conditions (96 °C 0:30, 50 °C 0:15, 60 °C 4:00) 40 times. After adding 20 μ L sterile miliQ H₂O, precipitating with 95% ethanol (68 μ L) and sodium acetate (3 μ L, 5 M), and adding high-dye formamide (10 μ L), the samples were sequenced with an ABI 3730xl DNA analyser.

Analysis

The primers shf and shr2 amplified a fragment 895 bp long, and shf2 and shr2 a fragment 780 bp long. A total of 692 bases between bases 219 and 910 of the cytochrome *b* gene were unambiguously edited using Sequencher version 3.0.1 (Gene Codes Corporation, Ann Arbor, MI) and BBEdit Lite version 3.0 (Felciano, 1994), aligned using Clustal X (Thompson *et al.*, 1997) and SeqPup version 0.6f (Gilbert, 1996), and checked by eye. Polymorphic sites were rechecked with the original sequence trace files. Six sequences included in the data set came from S. Casey (GenBank accession numbers: AF192698–AF192703). We altered the identity of two bases (from G to A) in Casey's sequences (positions 812 and 856 in reference to the entire *cyt b* sequences). These base positions were invariable in our remaining 94 samples suggesting that an error may have been made in Casey's editing. Identical sequences were assigned the same haplotype identity, and only a single example of each was used in the phylogenetic reconstructions assuming that identical haplotypes shared the same evolutionary origin. Haplotype definitions have been submitted to GenBank (accession numbers AY322433–AY322476).

Distance-based, parsimony and maximum-likelihood (ML) phylogenetic reconstructions were carried out using PAUP* version 4.0b10 (Swofford, 2002) with settings for the ML analyses obtained from Modeltest version 3.0 (Posada & Crandall, 1998). The major phylogenetic divisions observed were identical in all analyses. A 95% likely statistical parsimony network of haplotypes was estimated using TCS version 1.13 (Clement *et al.*, 2000). Calculations of molecular diversity and analysis of molecular variance (AMOVA) results were obtained using Arlequin version 2.0 (Excoffier *et al.*, 1992; Schneider *et al.*, 2000) with Kimura's two-parameter (K2P) genetic distance (Kimura, 1980).

To test our hypotheses, we split the samples into 'populations' according to their ocean basins of origin (Pacific or Indian) and their position with respect to Wallace's Line (west or east) and used these groupings to test the geographical distributions of the lineages using an exact contingency test (lineage membership × 'population') as executed by GEODIS version 2.0 (Posada *et al.*, 2000) with 1000 permutations. Results from the AMOVA supported our results from the chi-square analysis. Divergence times (for K2P genetic distance) were estimated to be 1.4% per million years based on a molecular clock hypothesis for the cytochrome *b* gene of seahorses from either side of the Isthmus of Panama (Casey, 1999).

RESULTS AND DISCUSSION

Our results from a total of 100 individuals surveyed for mtDNA cytochrome *b* sequence variation indicated the presence of two distinct lineages, separated by an average $d = 2.90\%$ sequence divergence (K2P distance, corrected for

within-lineage diversity) (Fig. 2, Table 1). Within-group sequence diversity was $\pi = 0.29\%$ for lineage A, and $\pi = 0.79\%$ for lineage B. Lineages A and B were separated geographically into east and west groups reminiscent of the terrestrial biogeographical division, rather than north-east-south-west as predicted by the ocean-basin hypothesis (Figs 1 and 3). This east-west division of the major lineages is consistent, although not precisely congruent, with Wallace's Line ($\chi^2 = 51.22, P < 0.001; \Phi_{ST} = 0.662, P < 0.001$). It is not consistent with the ocean-basin hypothesis ($\chi^2 = 1.62, P = 0.251; \Phi_{ST} = 0.023, P = 0.153$). Lineage B, found primarily to the east of Wallace's Line, was further split into two clusters (B1 and B2) separated by an average divergence of $d = 1.1\%$. The geographical structure between B1 and B2 was consistent with an Indian-Pacific ocean-basin hypothesis ($\chi^2 = 10.0, P = 0.004; \Phi_{ST} = 0.461, P = 0.005$), although of course, consistence does not necessarily imply causality.

Differences in their molecular diversity and phylogenetic structure indicate that the two primary lineages have been affected by different historical events (Table 1). Lineage A contains 32 haplotypes, haplotype diversity is high ($h = 0.90$), but nucleotide diversity is low because the haplotypes are only separated by a maximum of five nucleotide differences (average within-lineage pairwise difference = 2.04 mutations). By contrast, the 15 haplotypes that make up lineage B are widely separated, have a lower overall haplotype diversity ($h = 0.83$), but a higher nucleotide diversity because they are separated by up to 18 nucleotide differences (average pairwise difference = 5.40). Thus, lineage B has much deeper phylogenetic structure, despite the smaller geographical distances involved.

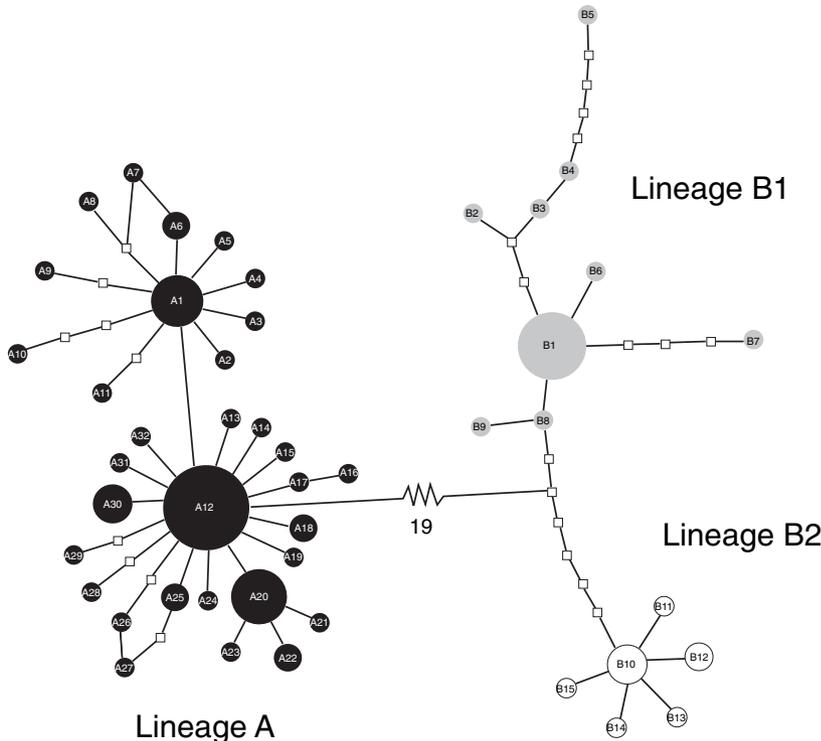


Figure 2 Statistical parsimony network showing phylogenetic relationships among haplotypes of *Hippocampus trimaculatus*. Each branch represents a single nucleotide mutation except for the central one, which represents 19 mutations. The area of each circle is proportional to the number of specimens sharing that haplotype. Small open squares represent hypothesized intermediate haplotypes not observed in our sample. Loops reflect alternative connections in the 95% likely network.

Table 1 Summary of molecular diversity and AMOVA results for *Hippocampus trimaculatus*. Numbers of individuals (n), numbers of haplotypes (k), average pairwise differences among individuals ($p \pm$ standard deviation), haplotype diversity ($h \pm$ standard deviation), nucleotide diversity ($\pi \pm$ standard deviation) for each grouping of samples. Average sequence divergence between groups (d_{xy}), corrected sequence divergence ($d_{xy(corr)}$), Φ_{ST} and P -values for group comparisons (lineage A vs. B, west vs. east of Wallace's Line and Indian vs. Pacific Ocean basins)

	n	k	p	h	π	d_{xy}	$d_{xy(corr)}$	Φ_{ST} (P -value)
All	100	47	11.93 \pm 5.44	0.94 \pm 0.014	0.017 \pm 0.0087	n/a	n/a	n/a
Lineages								
A	70	32	2.04 \pm 1.16	0.90 \pm 0.026	0.0029 \pm 0.0019	3.44%	2.90%	0.869 (< 0.0001)
B	30	15	5.46 \pm 2.70	0.83 \pm 0.064	0.0079 \pm 0.0043			
Wallace's Line hypothesis								
West	82	40	7.89 \pm 3.71	0.93 \pm 0.020	11.4 \pm 5.93	3.13%	2.14%	0.662 (< 0.0001)
East	18	9	5.89 \pm 2.95	0.82 \pm 0.075	8.51 \pm 4.77			
Ocean Basin hypothesis								
Indian	31	19	13.02 \pm 6.04	0.94 \pm 0.026	18.8 \pm 9.69	1.79%	0.004%	0.023 (0.153)
Pacific	69	35	11.28 \pm 5.18	0.94 \pm 0.017	16.3 \pm 8.30			

Low sequence diversity, yet high haplotype diversity, on the continental shelf (west of Wallace's Line) across huge geographical distances (from India to Japan > 10,000 km by coastal routes) suggests that any previous phylogeographical structure in *H. trimaculatus* in this region has been erased by more recent range-expansion events (Avice, 2000). Range expansion is perhaps most clearly demonstrated by the distribution of the central and abundant haplotype A12 which extends from India to Vietnam, and south to Borneo and Java (Appendix 1). Because free marine passage across the Sunda Shelf would only have been possible since the end of the last glacial maximum (*c.* 10,000 years ago) successful dispersal events resulting in a range expansion of the order of > 1 km year⁻¹ must have occurred in order to account for this present distribution. The lack of phylogeographical structure to the west of Wallace's Line may reflect the recency of this range expansion and insufficient time to attain migration-drift equilibrium (Slatkin, 1993).

Given the evidence for relatively rapid dispersal across the continental shelf regions by lineage A, the lack of lineage A to the east of Wallace's Line is surprising. Ocean currents in this region are seasonally reversing. However, they flow from west to east through the Java Sea for most of the year (Wyrski, 1961) apparently favouring the movement of lineage A haplotypes from the west to the east. The narrow zone of overlap of lineages A and B on the continental margin (north and south Borneo and Java) indicates that the observed genetic disjunction is not absolute and some unidirectional (east-west) dispersal has occurred onto the shelf. Lack of further westward expansion by lineage B is again inconsistent with the apparent wide dispersal of lineage A.

The maintenance of the observed phylogeographical break could be due to competitive interactions or assortative matings between the two lineages (Planes & Doherty, 1997), the result of preferences for different marine ecosystems (oceanic vs. continental shelf) (Rocha *et al.*, 2002), or different evolution-

ary tendencies for dispersal by oceanic vs. continental populations. Alternatively, the break may not be actively maintained; the widespread nature of lineage A may reflect ecological release and founding of new populations in previously unoccupied habitats following the last ice-age, whereas the more limited extent of movement of lineage B haplotypes onto the Sunda Shelf, and the higher level of genetic structure to the east of Wallace's Line, may reflect more limited effective dispersal of individuals moving into already saturated habitats.

The deeper genetic structure within lineage B, and the geographical associations of its haplotypes suggests a role for fragmentation and isolation of populations. Haplotype B1 is widespread in both the north and the south, whereas the rest of lineage B1 (except haplotype B9) is restricted to the Sulu Sea and the whole of lineage B2 is restricted to the south. It is possible that these patterns reflect the predicted separation of the Indian/Pacific Ocean basins or it may reflect isolation of smaller ocean basins (Sulu/Sulawesi/Flores Seas). Further data from the east are needed to critically assess these hypotheses.

Assuming a molecular clock, calibrated at 1.4% K2P sequence divergence/Ma (Casey, 1999) *H. trimaculatus* separated from its sister taxon, 'split-spot trimaculatus' (= *Hippocampus biocellatus* of Lourie *et al.*, 1999b; Kuitert, 2001) in Australia *c.* 9.5 Ma (S.L., unpubl. data). This is approximately the time at which open water (a possible barrier to a coastal-shelf species) between the two continental shelves was at its minimum (Hall, 2002) and successful dispersal might be most likely.

Using the same clock rate, lineages A and B last shared a common ancestor *c.* 2.1 Ma. This date supports an early Pleistocene event, but the location of the geographical disjunction between these two lineages is not consistent with vicariance between separated Indian and Pacific Ocean populations. An alternative explanation, still consistent with

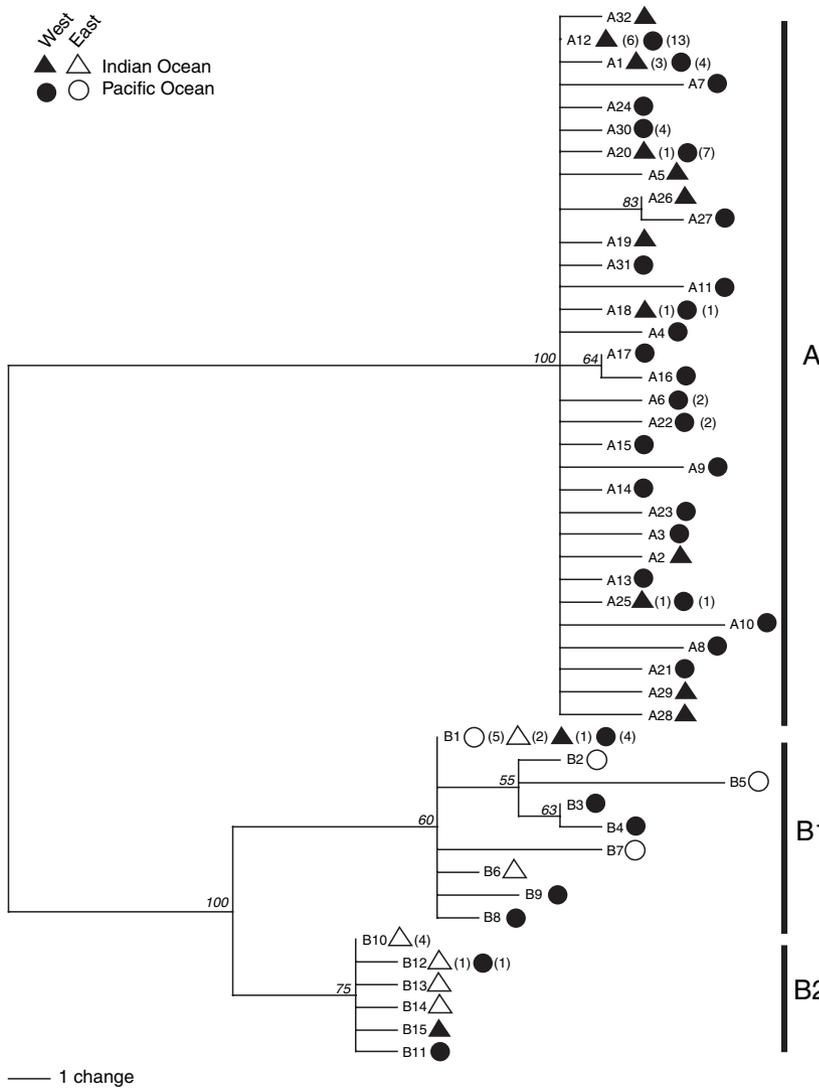


Figure 3 Maximum-likelihood (ML) mid-point rooted bootstrap consensus tree among haplotypes of *Hippocampus trimaculatus*. ML settings: model HKY + I with transition/transversion ratio = 17.8597, nucleotide frequencies: A = 0.2878, C = 0.2573, G = 0.1474, T = 0.3075, proportion of invariable sites = 0.4. Symbols reflect geographical origins of specimens (filled = west of Wallace's Line, empty = east of Wallace's Line, triangles = Indian Ocean, circles = Pacific Ocean). Haplotype identities are shown and numbers in parentheses indicate the number of individuals sharing a haplotype ($n = 1$ unless indicated otherwise). Figures above branches indicate bootstrap support (100 replicates). Partitions with < 50% support are not shown.

Pleistocene vicariance resulting from lowered sea-levels, is that the populations were separated by a land bridge between Borneo and the Philippines (Heaney, 1985). In this case, populations in the South China Sea (proto-lineage A) would have been separated from those in the Sulu/Sulawesi/Banda/Flores Seas (i.e. the Philippines and eastern Indonesian islands) (proto-lineage B). Following the final inundation of the shallow water Sunda Shelf, lineage A may have rapidly expanded to its present range. The lack of lineage A in northern Borneo contradicts this hypothesis, but with only four specimens from this area only limited conclusions can be drawn. Further sampling, especially on the northern and western edges of the Sunda Shelf, and a higher resolution genetic marker would be needed to test the hypothesis of a South China Sea origin for lineage A (as opposed to one in the Indian Ocean).

The 1.1% sequence divergence within lineage B is equivalent to *c.* 780,000 years of independent evolution, possibly the result of a second vicariant or dispersal event during the Pleistocene. The higher genetic diversity in lineage B implies

that it is divided into fewer, smaller populations, isolated from one another and subject to drift. Certainly specimens were much harder to obtain from east of Wallace's Line than west. Smaller population sizes, and greater isolation may reflect lack of suitable habitat as extensive shelf-margins are absent from much of eastern Indonesia. It may also reflect the stability (i.e. resistance to extinction) of these populations over evolutionary time (Cunningham & Collins, 1998).

Other phylogeographical surveys within Southeast Asia (more limited in geographical scope) indicate that our major phylogeographical division is novel. In mantis shrimps (*Haptosquilla pulchella*) a deep divergence exists in central Indonesia (cytochrome *b*, 6.6%, $\Phi_{ST} = 0.821$, $P < 0.001$) but the major separation is north-south, rather than east-west (Barber *et al.*, 2000). The authors dubbed their division a 'marine Wallace's Line' with its orientation perpendicular to the terrestrial Wallace's Line. A survey of cytochrome *b* sequences in false clownfish (*Amphiprion ocellaris*) also revealed significant phylogeographical structure ($\Phi_{ST} = 0.560$ among all populations, $P < 0.001$), but the

primary lineages were separated by only 0.48% and there was extensive sharing of haplotypes among sites (Nelson *et al.*, 2000). One of these lineages occurred in the Indian Ocean, well to the west of Wallace's Line, the other encompassed areas in the Pacific Ocean both west and east of Wallace's Line. Although *A. ocellaris* is considered 'low-dispersing' with a pelagic larval phase of 8–12 days, its genetic structure suggests that it has greater dispersal potential than *H. trimaculatus* does across Wallace's Line, but not within the Sunda Shelf.

Our study, plus those cited above are among the first to focus on marine genetic connections within Southeast Asia. Although their details differ, they suggest that significant genetic structure does exist among marine populations in Southeast Asian waters. The patterns imply retention of signatures of historical events, and that knowledge of contemporary ocean currents may not necessarily be a good predictor of population connections over evolutionary time. The precise driving forces behind the patterns are still unclear (particularly those that mimic a terrestrial 'Wallace's Line' pattern) and further studies are warranted. The existence of sharp phylogeographical breaks also has implications for regional-scale conservation planning: (1) restocking programmes should proceed with great caution as populations may not be evolutionarily equivalent, (2) populations may be vulnerable to local extirpation as input from propagules from far a field may be minimal, (3) establishing marine protected areas on the basis of ocean currents may not necessarily maintain species distributions, (4) concepts of connectedness in marine conservation must be grounded in an good understanding of empirically recognized connections among populations, and (5) as low dispersal organisms retain historical signatures they should be a focus for designing marine protected areas at the most detailed spatial scale.

ACKNOWLEDGMENTS

This work is a contribution from Project Seahorse. Our thanks to R. Burton and his group (Scripps Institution of Oceanography, San Diego), T. Bureau and his group (McGill University, Montréal), P. Lepage and his group (Québec Genome Centre) for use of laboratory facilities, sequencing and advice. Also to V. Khalil, N. Dun and S. Elgazzar for technical assistance. Discussions with R. Atkinson, B. Bowen, S. Casey, J. Curtis, J. Dodson, D. Green, A. Hendry, F. Lecomte, A. Smith, G. Stirling, J. Turgeon, T. Wilson, Y. Wong and input from two anonymous reviewers greatly improved this paper. We would also like to thank many people in Southeast Asia who provided specimens or logistic support (particularly M. Chaiyapu, C.-K. Choo, M. Erdmann, T.H., Haq, H. Nusantari and A. Perry). Financial support was provided by a Leverhulme Foundation Study Abroad Studentship, McGill University (Canada) Trust and Commonwealth Scholarships to SL and a Dawson Award to AV (McGill University). LIPI (Indonesian Institute of Sciences) are thanked for granting a research permit (no. 1485/I/KS/2001).

REFERENCES

- Avice, J.C. (2000) *Phylogeography: the history and formation of species*. Harvard University Press, Cambridge, MA.
- Barber, P.H., Palumbi, S.R., Erdmann, M.V. & Moosa, M.K. (2000) A marine Wallace's line? *Nature*, **406**, 692–693.
- Barber, P.H., Palumbi, S.R., Erdmann, M.V. & Moosa, K.M. (2002) Sharp genetic breaks among populations of *Haptosquilla pulchella* (Stomatopoda) indicate limits to larval transport: patterns, causes, and consequences. *Molecular Ecology*, **11**, 659–674.
- Benzie, J.A.H. (1998) Genetic structure of marine organisms and Southeast Asian biogeography. *Biogeography and geological evolution of Southeast Asia* (ed. by R. Hall and J.D. Holloway), pp. 197–209. Backhuys Publishers, Leiden.
- Bohonak, A.J. (1999) Dispersal, gene flow, and population structure. *Quarterly Review of Biology*, **74**, 21–45.
- Casey, S.P. (1999) *Conservation genetics of seahorses (Hippocampus species)*. PhD Thesis, University of London, London.
- Choo, C.L. & Liew, H.C. (2003) Spatial distribution, substrate assemblages and size composition of sea horses (Family Syngnathidae) in the coastal waters of Peninsular Malaysia. *Journal of the Marine Biological Association of the United Kingdom*, **83**, 271–276.
- Clement, M., Posada, D. & Crandall, K.A. (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.
- Cunningham, C.W. & Collins, T.M. (1998) Beyond area relationships: extinction and recolonization in molecular marine biogeography. *Molecular approaches to ecology and evolution* (ed. by R. DeSalle and B. Schierwater), pp. 297–321. Birkhauser Verlag, Basel.
- Doherty, P.J., Planes, S. & Mather, P. (1995) Gene flow and larval duration in seven species of fish from the Great Barrier Reef. *Ecology*, **76**, 2373–2391.
- Excoffier, L., Smouse, P.E. & Quattro, J.M. (1992) Analysis of molecular variance inferred from metric distances among mtDNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.
- Felciano, R. (1994) *BBEdit Lite*. Bare Bones Software, Bedford, MA.
- Foster, S.J. & Vincent, A.C.J. (2004) The life history and ecology of seahorses, *Hippocampus* spp.: implications for conservation and management. *Journal of Fish Biology*, **65**(1), 1–61.
- Gilbert, D.G. (1996) *SeqPup: a biosequence editor and analysis application*. Indiana University, Bloomington, IN.
- Hall, R. (1998) The plate tectonics of Cenozoic SE Asia and the distribution of land and sea. *Biogeography and geological evolution of SE Asia* (ed. by R. Hall and J.D. Holloway), pp. 99–131. Backhuys Publishers, Leiden.
- Hall, R. (2002) Cenozoic geological and plate tectonic evolution of SE Asia and the SW Pacific: computer-based reconstructions and animations. *Journal of Asian Earth Sciences*, **20**, 353–434.

- Heaney, L.R. (1985) Zoogeographic evidence for middle and late Pleistocene land bridges to the Philippine islands. *Modern Quaternary Research in South East Asia*, **9**, 127–143.
- Huxley, T.H. (1868) On the classification and distribution of the Alectoromorphae and Heteromorphae. *Proceedings of the Zoological Society of London*, **1868**, 294–319.
- Jones, G.P., Milicich, M.J., Emslie, M.J. & Lunow, C. (1999) Self-recruitment in a coral reef fish population. *Nature*, **402**, 802–803.
- Kimura, M. (1980) A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, **16**, 111–120.
- Knowlton, N.R. & Weigt, L.A. (1998) New dates and new rates for divergence across the Isthmus of Panama. *Proceedings of the Royal Society of London, Series B*, **165**, 2257–2263.
- Kuiter, R.H. (2001) Revision of the Australian seahorses of the genus *Hippocampus* (Syngnathiformes: Syngnathidae) with descriptions of nine new species. *Records of the Australian Museum*, **53**, 293–340.
- Leis, J.M. & Carson-Ewart, B.M. (1998) Complex behaviour by coral-reef fish larvae in open-water and near-reef pelagic environments. *Environmental Biology of Fishes*, **53**, 259–266.
- Lourie, S.A., Pritchard, J.C., Casey, S.P., Truong, S.K., Hall, H.J. & Vincent, A.C.J. (1999a) The taxonomy of Vietnam's exploited seahorses. *Biological Journal of the Linnean Society*, **66**, 231–256.
- Lourie, S.A., Vincent, A.C.J. & Hall, H.J. (1999b) *Seahorses: an identification guide to the world's species and their conservation*. Project Seahorse, London.
- McKenzie, K.G. (1991) Implications of shallow Tethys and the origin on modern oceans. *Australian Systematic Botany*, **4**, 37–40.
- Michaux, B. (1991) Distributional patterns and tectonic development in Indonesia: Wallace reinterpreted. *Australian Systematic Botany*, **4**, 25–36.
- Nelson, J.S., Hoddell, R.J., Chou, L.M., Chan, W.K. & Phang, V.P.E. (2000) Phylogeographic structure of the false clownfish, *Amphiprion ocellaris*, explained by sea level changes on the Sunda shelf. *Marine Biology*, **137**, 727–736.
- Planes, S. & Doherty, P.J. (1997) Genetic and color interactions at a contact zone of *Acanthochromis polyacanthus*: a marine fish lacking pelagic larvae. *Evolution*, **51**, 1232–1243.
- Posada, D. & Crandall, K.A. (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.
- Posada, D., Crandall, K.A. & Templeton, A.R. (2000) GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology*, **9**, 487–488.
- Randall, J.E. (1998) Zoogeography of shore fishes of the Indo-Pacific region. *Zoological Studies*, **37**, 227–268.
- Rocha, L.A., Bass, A.L., Robertson, D.R. & Bowen, B.W. (2002) Adult habitat preferences, larval dispersal, and the comparative phylogeography of three Atlantic surgeonfishes (Teleostei: Acanthuridae). *Molecular Ecology*, **11**, 243–252.
- Scheltema, R.S., Williams, I.P. & Lobel, P.S. (1996) Retention around and long-distance dispersal between oceanic islands by planktonic larvae of benthic gastropod Mollusca. *American Malacological Bulletin*, **12**, 67–75.
- Schneider, S., Roessli, D. & Excoffier, L. (2000) *Arlequin ver. 2.000: a software for population genetics data analysis*. Genetics and Biometry Laboratory, University of Geneva, Switzerland.
- Slatkin, M. (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, **47**, 264–279.
- Springer, V.G. & Williams, J.T. (1990) Widely distributed Pacific plate endemics and lowered sea-level. *Bulletin of Marine Science*, **47**, 631–640.
- Swearer, S.E., Caselle, J.E., Lea, D.W. & Warner, R.R. (1999) Larval retention and recruitment in an island population of a coral-reef fish. *Nature*, **402**, 799–802.
- Swofford, D.L. (2002) *PAUP*. Phylogenetic analysis using parsimony (*and other methods)*, Version 4. Sinauer Associates, Sunderland, MA.
- Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. & Higgins, D.G. (1997) The ClustalX interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, **24**, 4876–4882.
- Van Oosterzee, P. (1997) *Where worlds collide: the Wallace Line*. Cornell University Press, New York.
- Voris, H.K. (2000) Maps of Pleistocene sea-levels in South East Asia: shorelines, river systems, time durations. *Journal of Biogeography*, **27**, 1153–1167.
- Whitmore, T.C. (ed.) (1981) *Wallace's Line and plate tectonics*. Oxford University Press, Oxford.
- Wyrski, K. (1961) Scientific results of marine investigations of the South China Sea and the Gulf of Thailand. *Naga Report*, **2**, 195.

BIOSKETCHES

Sara Lourie's research includes morphological and molecular systematics, phylogeography and the use of biogeography in setting marine conservation priorities. Her current work focuses on seahorses in Southeast Asia.

Amanda Vincent holds the Canada Research Chair in Marine Conservation. Her early research on seahorse ecology has led to broad initiatives in conservation science, management and policy, often featuring seahorses.

Appendix 1 Detailed list of specimens of *Hippocampus trimaculatus*, including museum numbers, haplotypes and GenBank accession numbers

Catalogue number	Location	Latitude	Longitude	Source	Haplotype	GenBank number
West of Wallace's Line/Indian Ocean (<i>n</i> = 21)						
RM2522a	Kerala, India	c. 10°00' N	c. 076°30' E	Buyer	A19	AY322451
RM2522b	Kerala, India	c. 10°00' N	c. 076°30' E	Buyer	A32	AY322433
RM2116	Mandapam, Tamil Nadu, India	c. 09°25' N	c. 079°20' E	Researcher	A02	AF192698
RM2636a	Mandapam, Tamil Nadu, India	c. 09°25' N	c. 079°20' E	Researcher	A01	AF192699
RM2281b	Palk Bay, Tamil Nadu, India	c. 09°25' N	c. 079°20' E	Fisher	A12	AY322435
RM2281f	Palk Bay, Tamil Nadu, India	c. 09°25' N	c. 079°20' E	Fisher	A12	AY322435
RM2281g	Palk Bay, Tamil Nadu, India	c. 09°25' N	c. 079°20' E	Fisher	A01	AF192699
RM2281h	Palk Bay, Tamil Nadu, India	c. 09°25' N	c. 079°20' E	Fisher	A28	AY322460
RM2420	Myanmar	Exact location unknown	Exact location unknown	Buyer	A12	AY322435
RM2423a	Border area between Thailand and Myanmar	Exact location unknown	Exact location unknown	Buyer	A26	AY322458
RM2393	Bulon Island, Thailand	c. 06°50' N	c. 099°30' E	Buyer	A12	AY322435
RM2352a	Near Pulau Pangkor, Perak, Malaysia	c. 04°15' N	c. 100°34' E	Fisher	A20	AY322452
RM2352c	Near Pulau Pangkor, Perak, Malaysia	c. 04°15' N	c. 100°34' E	Fisher	A18	AY322450
RM2352b	Near Pulau Pangkor, Perak, Malaysia	c. 04°15' N	c. 100°34' E	Fisher	A12	AY322435
RM2476	Bandar Lampung, Sumatra, Indonesia	c. 05°32' S	c. 105°17' E	Fisher	A05	AY322438
RM2466	Bandar Lampung, Sumatra, Indonesia	c. 05°32' S	c. 105°17' E	Fisher	A12	AY322435
RM2712a	Pulau Sangiang or P. Panaitan, Anyer, West Java, Indonesia	c. 06°00' or 06°30' S	c. 105°50' or 105°15' E	Fisher	A29	AY322461
RM2712c	Pulau Sangiang or P. Panaitan, Anyer, West Java, Indonesia	c. 06°00' or 06°30' S	c. 105°50' or 105°15' E	Fisher	A01	AF192699
P 10	Pangandaran, Ciamis, West Java, Indonesia	c. 07°41' S	c. 108°40' E	Buyer	B15	AY322476
RM2715a	Pangandaran, Ciamis, West Java, Indonesia	c. 07°41' S	c. 108°40' E	Fisher	A25	AY322457
RM2715b	Pangandaran, Ciamis, West Java, Indonesia	c. 07°41' S	c. 108°40' E	Fisher	B01	AY322434
West of Wallace's Line/Pacific Ocean (<i>n</i> = 61)						
RM2791	Rampa, Kota Baru, Pulau Laut, Kalimantan Selatan, Indonesia	c. 03°14' S	c. 116°12' E	Fisher	B09	AY322470
RM2785a	Tanjung Dewa, Kalimantan Selatan, Indonesia	c. 03°10' S	c. 116°20' E	Fisher	B01	AY322434
RM2785b	Tanjung Dewa, Kalimantan Selatan, Indonesia	c. 03°10' S	c. 116°20' E	Fisher	A10	AY322443
RM2787a	Pagatan, Kalimantan Selatan, Indonesia	c. 03°36' S	c. 115°58' E	Fisher	A12	AY322435

Appendix 1 *continued*

Catalogue number	Location	Latitude	Longitude	Source	Haplotype	GenBank number
RM2787b	Pagatan, Kalimantan Selatan, Indonesia	c. 03°36' S	c. 115°58' E	Fisher	A20	AY322452
RM2721	Karimunjawa, Central Java, Indonesia	c. 05°53' S	c. 110°26' E	Fisher	A20	AY322452
RM2713a	Jepara, Central Java, Indonesia	c. 06°30' S	c. 110°30' E	Fisher	A25	AY322457
RM2736d	Indramayu, West Java, Indonesia	c. 06°15' S	c. 108°30' E	Fisher	B11	AY322472
RM2736b	Indramayu, West Java, Indonesia	c. 06°15' S	c. 108°30' E	Fisher	B12	AY322473
RM2686b	Mersing, Johor, Malaysia	c. 02°25' N	c. 103°50' E	TCM shop	A04	AY322437
RM2686c	Mersing, Johor, Malaysia	c. 02°25' N	c. 103°50' E	TCM shop	A17	AY322449
RM2686d	Mersing, Johor, Malaysia	c. 02°25' N	c. 103°50' E	TCM shop	A12	AY322435
RM2686e	Mersing, Johor, Malaysia	c. 02°25' N	c. 103°50' E	TCM shop	A12	AY322435
RM2686f	Mersing, Johor, Malaysia	c. 02°25' N	c. 103°50' E	TCM shop	A06	AY322439
RM2686g	Mersing, Johor, Malaysia	c. 02°25' N	c. 103°50' E	TCM shop	A12	AY322435
RM2686h	Mersing, Johor, Malaysia	c. 02°25' N	c. 103°50' E	TCM shop	A30	AY322462
RM2686i	Mersing, Johor, Malaysia	c. 02°25' N	c. 103°50' E	TCM shop	A22	AY322454
RM2686j	Mersing, Johor, Malaysia	c. 02°25' N	c. 103°50' E	TCM shop	A20	AY322452
RM2686k	Mersing, Johor, Malaysia	c. 02°25' N	c. 103°50' E	TCM shop	A15	AY322447
RM2686m	Mersing, Johor, Malaysia	c. 02°25' N	c. 103°50' E	TCM shop	A09	AY322442
RM2686n	Mersing, Johor, Malaysia	c. 02°25' N	c. 103°50' E	TCM shop	A14	AY322446
RM2356a	Near Pulau Kapas and P. Tenggol, Terengganu, Malaysia	c. 04°45' N	c. 103°40' E	Fisher	A12	AY322435
RM2356b	Near Pulau Kapas and P. Tenggol, Terengganu, Malaysia	c. 04°45' N	c. 103°40' E	Fisher	A24	AY322456
RM2405	Pattani, Thailand	c. 07° N	c. 101° E	Fisher	A01	AY322436
RM2409a	Pattani, Thailand	c. 07° N	c. 101° E	TCM shop	A30	AY322462
RM2409b	Pattani, Thailand	c. 07° N	c. 101° E	TCM shop	A12	AY322435
RM2409c	Pattani, Thailand	c. 07° N	c. 101° E	TCM shop	A12	AY322435
RM2409e	Pattani, Thailand	c. 07° N	c. 101° E	TCM shop	A07	AY322440
MC 12	Ban Koh Prerd, Laem Sing, Chanthaburi, Thailand	c. 12°10' N	c. 102°10' E	Fisher	A06	AY322439
MC 37	Paknam, Samut Prakan, Thailand	c. 13°25' N	c. 100°36' E	Fisher	A01	AF192699
MC 38	Paknam, Samut Prakan, Thailand	c. 13°25' N	c. 100°36' E	Fisher	A23	AY322455
MC 15	Ban Koh Prerd, Laem Sing, Chanthaburi, Thailand	c. 12°10' N	c. 102°10' E	Fisher	A12	AY322435
RM2115	Chonburi, Thailand	c. 13°15' N	c. 100°40' E	Researcher	A12	AF192702
RM2850a	Kampong Som, Cambodia	10.60355° N	103.48604° E	Fisher	A20	AY322452
RM2850b	Kampong Som, Cambodia	10.60355° N	103.48604° E	Fisher	A12	AY322435
RM2850e	Kampong Som, Cambodia	10.60355° N	103.48604° E	Fisher	A16	AY322448
RM2850g	Kampong Som, Cambodia	10.60355° N	103.48604° E	Fisher	A27	AY322459
RM2878	Kep, Cambodia	10.48060° N	104.32182° E	Fisher	A21	AY322453
RM2833	Lob, Cambodia	10.43218° N	104.43185° E	Fisher	A22	AY322454
RM2121b	Song-Doc, Thang, Minh Hai, Vietnam	c. 09°00' N	c. 104°45' E	Fisher?	A13	AY322445
RM2124b	Nha Trang, Khanh Hoa, Vietnam	c. 12°15' N	c. 109°10' E	Fisher?	A31	AY322463
RM2129b	Nha Trang, Khanh Hoa, Vietnam	c. 12°15' N	c. 109°10' E	Fisher?	A11	AY322444
RM2131a	Nha Trang, Khanh Hoa, Vietnam	c. 12°15' N	c. 109°10' E	Fisher?	A12	AY322435

Appendix 1 *continued*

Catalogue number	Location	Latitude	Longitude	Source	Haplotype	GenBank number
RM2131b	Nha Trang, Khanh Hoa, Vietnam	c. 12°15' N	c. 109°10' E	Fisher?	A30	AY322462
RM2130a	Nha Trang, Khanh Hoa, Vietnam	c. 12°15' N	c. 109°10' E	Fisher?	A18	AY322450
RM2130d	Nha Trang, Khanh Hoa, Vietnam	c. 12°15' N	c. 109°10' E	Fisher?	A20	AY322452
RM2132	Thuan An, Vietnam	Exact location unknown	Exact location unknown	Fisher?	A03	AF192703
SW001	Kuching, Sarawak, Malaysia	c. 01°50' N	c. 110°15' E	Fisher	A08	AY322441
SW002	Kuching, Sarawak, Malaysia	c. 01°50' N	c. 110°15' E	Fisher	A12	AY322435
SW003	Kuching, Sarawak, Malaysia	c. 01°50' N	c. 110°15' E	Fisher	A20	AY322452
SW015	Kuching, Sarawak, Malaysia	c. 01°50' N	c. 110°15' E	Fisher	A12	AY322435
SW017	Kuching, Sarawak, Malaysia	c. 01°50' N	c. 110°15' E	Fisher	A20	AY322452
SB116	Pulau Tigabu, Sabah, Malaysia	c. 07°00' N	c. 117°20' E	Fisher	B03	AY322465
SB117	Pulau Tigabu, Sabah, Malaysia	c. 07°00' N	c. 117°20' E	Fisher	B08	AY322469
SB118	Pulau Tigabu, Sabah, Malaysia	c. 07°00' N	c. 117°20' E	Fisher	B01	AY322434
SB119	Pulau Tigabu, Sabah, Malaysia	c. 07°00' N	c. 117°20' E	Fisher	B04	AY322466
RM2467	Pulau Tiga, Sabah, Malaysia	c. 05°42' N	c. 115°38' E	Fisher	B01	AY322434
RM2659b	Dumaran, Palawan, Philippines	c. 10°32' N	c. 119°48' E	Fisher	B01	AY322434
RM2746	I-Lan, Taiwan	c. 24°46' N	c. 121°45' E	TCM shop	A01	AF192699
RM2831	Keelung Island, Taiwan	c. 25°15' N	c. 121°40' E	Fisher	A30	AY322462
RM2123	Japan	Exact location unknown		Aquarium	A01	AF192700
East of Wallace's Line/Indian Ocean (<i>n</i> = 10)						
RM2206b	Batu Nampar, Lombok, Indonesia	c. 08°52' S	c. 116°24' E	Fisher	B10	AY322471
RM2206c	Batu Nampar, Lombok, Indonesia	c. 08°52' S	c. 116°24' E	Fisher	B12	AY322473
RM2206e	Batu Nampar, Lombok, Indonesia	c. 08°52' S	c. 116°24' E	Fisher	B10	AY322471
RM2206g	Batu Nampar, Lombok, Indonesia	c. 08°52' S	c. 116°24' E	Fisher	B10	AY322471
RM2206h	Batu Nampar, Lombok, Indonesia	c. 08°52' S	c. 116°24' E	Fisher	B10	AY322471
RM2206i	Batu Nampar, Lombok, Indonesia	c. 08°52' S	c. 116°24' E	Fisher	B06	AY322467
RM2206k	Batu Nampar, Lombok, Indonesia	c. 08°52' S	c. 116°24' E	Fisher	B13	AY322474
RM2206a	Batu Nampar, Lombok, Indonesia	c. 08°52' S	c. 116°24' E	Fisher	B01	AY322434
RM2703a	Labuan Bajo, Flores, Nusa Tenggara, Indonesia	c. 08°29' S	c. 119°53' E	Fisher	B01	AY322434
RM2703b	Labuan Bajo, Flores, Nusa Tenggara, Indonesia	c. 08°29' S	c. 119°53' E	Fisher	B14	AY322475
East of Wallace's Line/Pacific Ocean (<i>n</i> = 8)						
RM2346	Kendari, Sulawesi, Indonesia	c. 04°00' S	c. 123°00' E	Boy in market	B01	AY322434
RM2153	Maumauan Island, Bohol, Philippines	c. 10°08' N	c. 124°08' E	Fisher	B05	AF192701
RM2023	Nasingin Island, Bohol, Philippines	c. 10°08' N	c. 124°08' E	Fisher	B01	AY322434
RM2676	Daram Island, West Samar, Philippines	c. 11°38' N	c. 124°47' E	Fisher	B01	AY322434
RM2681a	Suwangan, Bantayan Island, Cebu, Philippines	c. 11°12' N	c. 123°45' E	Fisher	B07	AY322468
RM2681b	Suwangan, Bantayan Island, Cebu, Philippines	c. 11°12' N	c. 123°45' E	Fisher	B01	AY322434
RM2667a	Carles, Iloilo, Philippines	Exact location unknown	Exact location unknown	Buyer	B02	AY322464
RM2667b	Carles, Iloilo, Philippines	Exact location unknown	Exact location unknown	Buyer	B01	AY322434

RM = Redpath Museum, McGill University, 859 Sherbrooke St. W., Montréal, Québec H3A 2K6, Canada. P = Tissue Sample in Redpath Museum. MC = Collection of Monruedee Chaiyapu, Department of Fisheries, Kasetsart University, Bangkok, Thailand. SW and SB = Collection of Choo Chee Kuang, University College of Science and Technology, Kuala Terengganu, Malaysia. TCM = Traditional Chinese Medicine. *Note:* Haplotype A01 appears three times in GenBank.

Molecular phylogenetics of the butterflyfishes (Chaetodontidae): Taxonomy and biogeography of a global coral reef fish family

Jennifer L. Fessler^{a,b,*}, Mark W. Westneat^b

^a Department of Biological Sciences, University of Illinois at Chicago, 845 W. Taylor Street, Chicago, IL 60607, USA

^b Department of Zoology, Field Museum of Natural History, 1400 S. Lakeshore Drive, Chicago, IL 60605, USA

Received 6 November 2006; revised 22 May 2007; accepted 24 May 2007

Available online 9 June 2007

Abstract

Marine butterflyfishes (10 genera, 114 species) are conspicuously beautiful and abundant animals found on coral reefs worldwide, and are well studied due to their ecological importance and commercial value. Several phylogenies based on morphological and molecular data exist, yet a well-supported molecular phylogeny at the species level for a wide range of taxa remains to be resolved. Here we present a molecular phylogeny of the butterflyfishes, including representatives of all genera (except *Parachaetodon*) and at least one representative of all commonly cited subgenera of *Chaetodon* (except *Roa sensu Blum, 1988*). Genetic data were collected for 71 ingroup and 13 outgroup taxa, using two nuclear and three mitochondrial genes that total 3332 nucleotides. Bayesian inference, parsimony, and maximum likelihood methods produced a well-supported phylogeny with strong support for a monophyletic Chaetodontidae. The *Chaetodon* subgenera *Exornator* and *Chaetodon* were found to be polyphyletic, and the genus *Amphichaetodon* was not the basal sister group to the rest of the family as had been previously proposed. Molecular phylogenetic analysis of data from 5 genes resolved some clades in agreement with previous phylogenetic studies, however the topology of relationships among major butterflyfish groups differed significantly from previous hypotheses. The analysis recovered a clade containing *Amphichaetodon*, *Coradion*, *Chelmonops*, *Chelmon*, *Forcipiger*, *Hemitaenichthys*, *Johnrandallia*, and *Heniochus*. *Prognathodes* was resolved as the sister to all *Chaetodon*, as in previous hypotheses, although the topology of subgeneric clades differed significantly from hypotheses based on morphology. We use the species-level phylogeny for the butterflyfishes to resolve long-standing questions regarding the use of subgenera in *Chaetodon*, to reconstruct molecular rates and estimated dates of diversification of major butterflyfish clades, and to examine global biogeographic patterns.

© 2007 Elsevier Inc. All rights reserved.

Keywords: Biodiversity; *Chaetodon*; Molecular evolution; Phylogeny; Reef fishes

1. Introduction

Butterflyfishes are diverse and abundant marine perciform fishes that have spectacular coloration, high levels of ecological and morphological diversity, and are found on reefs throughout shallow, tropical seas. Because of these features, they have attracted the enthusiasm and inquiry of divers, scientists, and coral reef conservationists. Research-

ers have an enduring legacy of fascination with the butterflyfishes because they are a prominent focus for studies of feeding specialization (Ferry-Graham et al., 2001a; Motta, 1985, 1988, 1989; Sazima and Sazima, 2001), feeding ecology and resource partitioning (Cox, 1994; Pitts, 1991; Pratchett, 2005; Zekeria et al., 2002), mating behavior and pair bonding (Driscoll and Driscoll, 1988; Fricke, 1986; Hourigan, 1989; Tricas and Hiramoto, 1989; Yabuta, 1999, 2002), biogeography (Bellwood et al., 2004; Blum, 1989; Findley and Findley, 1989, 2001; Righton et al., 1996), color patterns (McMillan et al., 1999), and the curious, armored, tholichthys larval stage (Leis, 1989). Butterflyfishes comprise a biogeographical mixture of species, in

* Corresponding author. Address: Department of Zoology, Field Museum of Natural History, 1400 S. Lakeshore Drive, Chicago, IL 60605, USA. Fax: +1 312 665 7391.

E-mail address: jfessler@fieldmuseum.org (J.L. Fessler).

which some are highly restricted endemics whereas others exhibit nearly circumtropical distributions. Although recent research has generated phylogenetic hypotheses using both morphological data (Ferry-Graham et al., 2001b; Smith et al., 2003) and molecular data (Hsu et al., 2007; Littlewood et al., 2004; McMillan and Palumbi, 1995) the exploration of evolutionary, ecological and biogeographic patterns remains limited by lack of a well-resolved molecular phylogeny involving species from each of the major chaetodontid clades. The present study provides a phylogenetic analysis of butterflyfishes with a wide range of taxonomic sampling, enabling the examination of evolutionary and biogeographic patterns among species groups within the family.

The most comprehensive previous study of chaetodontid phylogenetics was based on morphology (Blum, 1988). Blum's coding and analysis of butterflyfish morphological characters persists as the most relied-upon taxonomic and phylogenetic guide. This phylogeny was constructed using composite taxa as representatives of subgenera, and a combination of features of outgroups distilled to a hypothetical ancestor. Ferry-Graham et al. (2001b) published a morphologically-based phylogeny using several character state revisions of Blum's (1988) dissertation work. Smith et al. (2003) updated, refined, and expanded this morphological database, including new characters based on soft tissue anatomy. The resultant phylogeny strongly supported the monophyly of the Chaetodontidae, identified the major clades within the family, and proposed a topology for the branching pattern of *Chaetodon* subgenera.

An active area of evolutionary research on the family is the phylogenetic placement of the long-jawed butterflyfishes. This question is of interest in order to understand the evolution of several specializations of skull biomechanics in the genera *Forcipiger* and *Chelmon*. The mechanism of feeding in the long-jawed butterflyfishes has been explored from the perspectives of quantitative kinematics, morphology, and the trends in feeding that are apparent when integrated with a phylogeny (Ferry-Graham et al., 2001b). This study used the morphological phylogeny to propose a scenario for the evolution of specialized jaw protrusion in the group. Molecular evidence illuminating the evolution of this clade, however, remains an important goal for the continued exploration of ecological and evolutionary trends in this family.

Several molecular phylogenetic studies have been conducted on butterflyfishes, including analysis of the cytochrome *b* gene on a group of 13 Pacific species (McMillan and Palumbi, 1997). This study questioned the monophyly of the *Chaetodon* (*Exornator*) clade, finding that *C. argentatus* (typically placed in *Chaetodon* subgenus *Exornator*) grouped with members of the *Roapons* “*tinkeri*” complex. Further, they found that *Chaetodon* subgenera were genetically structured by the Pacific and Indian Ocean basins rather than by subgeneric assignment. A molecular phylogeny focusing on the genus *Chaetodon* (Littlewood et al., 2004) and based on the Genbank data of Nelson et al. (unpublished) also conflicted in part with previous

hypotheses. The results of this analysis placed *Parachaetodon* as sister to *C. trifascialis* of the *C.* (*Megaprotodon*) subgenus and found monophyly of some *Chaetodon* subgenera. These studies represent the first steps toward resolution of relationships within *Chaetodon* using molecular data, but were not able to address broad-scale questions of intrafamily relationships or species-level chaetodontid phylogeny.

The phylogenetic composition and taxonomic utility of many of the proposed generic and subgeneric groupings within the Chaetodontidae remain to be explored. Some butterflyfish systematists (Allen et al., 1998; Blum, 1988) and the authors of popular aquarist books (Kuitert, 2002) vary in the species comprising various subgenera and even genera. The current study contains representatives of all currently recognized genera and *Chaetodon* subgenera, with the exception of the monotypic *Parachaetodon ocellatus* and *Roapons* (*sensu* Blum, 1988). Phylogenetic resolution of the family will allow assessment of taxonomic subdivisions within the family, and provide a framework for reconstructing the patterns of diversification of these charismatic reef inhabitants. Our goals in this contribution are to (1) generate a phylogeny that tests species-level relationships within the family Chaetodontidae, (2) assess the phylogenetic validity of the *Chaetodon* subgenera, and (3) examine the relative rates of molecular evolution, dates of divergence of major groups, and biogeographic history of butterflyfishes.

2. Materials and methods

2.1. Taxon sampling and genes sequenced

We analyzed data from 71 butterflyfishes and 13 perciform outgroups for a total of 84 taxa. Ingroup taxa were selected to maximize sampling of generic and subgeneric ranking within the Chaetodontidae. Outgroups were selected based on availability and consideration of numerous past proposals of relational proximity to the butterflyfishes. Outgroups consist of the Zaclidae, Drepaneidae, Ephippidae, Kyphosidae (including *Microcanthus* and *Atypichthys*), Scatophagidae and members of the Pomacanthidae. The majority of past work has placed the Pomacanthidae as a sister family to the Chaetodontidae, and recent research has raised the question of whether the Scatophagidae might be the closest butterflyfish relatives (Bellwood et al., 2004). We therefore used a wide selection of outgroups to provide resolution among several of the possible closest relatives to the Chaetodontidae.

We sequenced fragments of 5 genes for this study, including 3 mitochondrial regions (12S, 16S, and ND3) and 2 nuclear loci. The mitochondrial genes are all relatively widely used markers for phylogenetics. Tmo-4C4 is a single-copy nuclear DNA locus first isolated by Streelman and Karl (1997) for use in perciform fish phylogenetic analysis, and thought to be similar to titin, a large protein that contributes to muscle assembly and resting tension.

Also from the nuclear genome, recombination activating gene-2 (RAG2) works with RAG1 to synergistically activate V(D)J recombination, the process by which developing lymphocytes generate immune system diversity (Oettinger et al., 1990). Both of these nuclear loci have shown their utility in family-level phylogenetic studies of fishes (e.g., Sparks and Smith, 2004; Westneat and Alfaro, 2005).

2.2. DNA extraction and PCR amplification

Muscle and gill samples preserved in ethanol were cut into approximately 2 mm³ sections and processed using a Puregene DNA isolation kit (Gentra Systems, Minneapolis, MN). To increase our yield, we opted to include glycogen in our DNA recovery step of the extraction. DNA was stored at 4 °C, in the DNA hydration solution included in the Puregene kit. To amplify extracted DNA, we mixed a PCR cocktail solution that contained 2.5 µl 10× reaction buffer (100 mM Tris-HCl, 500 mM KCl, 15 mM MgCl₂, pH 8.3; Roche, Mannheim, Germany), 1.5 µl of 8 mM pre-mixed deoxynucleotide triphosphates, 0.1 µl of 5 U ml⁻¹ Taq polymerase (Roche, Mannheim, Germany), and 1 µl of each oligonucleotide primer, each at 10 µM concentration. To every 25 µl aliquot of the PCR cocktail, we added 1 µl of DNA stock (diluted 1:10 in some cases). Primers used for amplification and sequencing are given in Table 1. Some samples that initially amplified weakly were redone using double the amount of DNA in a standard PCR with the annealing temperature lowered by 5 °C. Each PCR

used an annealing temperature of 60 °C and a cycle duration of 30 s. PCR reactions were performed using an MJ Research PTC-200 Peltier Thermal Cycler (MJ Research Inc., Watertown, MA).

All PCR products were loaded and run on agarose gels, then bands visually assessed for intensity, cut out, and digested with GELase™ Agarose Gel-Digesting Preparation (Epicentre, Madison, WI). PCR products were cycle sequenced using an ABI PRISM® Big Dye™ Terminator Cycle Sequencing Ready Reaction Kit, version 1.0 (Applied Biosystems, Foster City, CA). Each sequencing reaction of 10 µl total volume contained 0.5 µl of Big Dye, 0.5 µl of DNA, and 0.25 µl of each amplification primer. Reactions were processed using 32 cycles at an annealing temperature of 50 °C for 15 s, and extension at 60 °C for 4 min. Forward and reverse labeled fragments were analyzed on an ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA). Sequences of both strands were determined for at least one sample per species. Random re-extraction and re-sequencing of taxa was performed throughout the project for error checking.

2.3. Sequence alignment and partition congruence

Contiguous sequences of each forward and reverse strand were joined using Sequencher 3.0 (Gene Codes Corporation). 12S and 16S sequences were aligned to a secondary structure model for perciform fishes (Westneat and Alfaro, 2005). The model and our sequences were imported into Se-AL (Rambaut, 1996) and aligned manually. We excluded ambiguously alignable regions and loop regions from further analysis. Sequences for protein coding nuclear genes were aligned automatically using Se-AL or MacClade 4.0 and visually edited for proper gap placement in regions of insertion or deletion.

Genes were assessed for incongruence via separate analysis of each partition. We first conducted separate maximum likelihood and Bayesian analyses for each data partition and surveyed corresponding nodes for support and congruence. Nodes that differed between methods but possessed low support values were treated as congruent. No strongly supported nodes were found in a partition tree that disagreed with strongly supported nodes from other partition trees, so all gene fragments were combined in subsequent analyses.

2.4. Phylogenetic analysis: likelihood and parsimony

Phylogenetic analyses were conducted using a beta version of PAUP* (Swofford, 2000). We used ModelTest Version 3.06 (Posada and Crandall, 1998) to eliminate the least likely models from a limited set. We adopted the recommended model for this analysis, which was (GTR+I+G) selected by AIC in Modeltest (Iset base = (0.2751, 0.2665, and 0.2118), nst = 6, rmat = (1.9785, 12.5647, 2.3986, 0.8385, and 17.0951), rates = gamma, shape = 0.4423,

Table 1
Primers and primer sequences used for amplification and sequencing of the five genes employed for phylogenetic reconstruction of the Chaetodontidae

Gene	Primer name	Primer sequence
12s rDNA	12s53F	CAC AAA GGC TTG GTC CTG ACT TT
	12s489F	CTG GGA TTA GAT ACC CCA CTA TGC
	12s613R	TCG GTT CTA GAA CAG GCT CCT CTA G
	12s991R	GGT ACA CTT ACC ATG TTA CGA CT
Tmo 4C4	Tmo-f1-5	CCT CCG GCC TTC CTA AAA CCT CTC
	Tmo-f1-6	GAA AAG AGT GTT TGA AAA TGA
	Tmo-r1-3	CAT CGT GCT CCT GGG TGA CAA AGT
RAG2	Andy-f	CTG GCC AAA ACG CTC ATG TCC AAC
	RAG2-f1	GAG GGC CAT CTC CTT CTC CAA
	Isben-f	TGG AAC AGT GTM RTT GAC TGT CC
	RAG2-f2	GAC TGT CCT CCT CAG GTG TTC
	RAG2-r2	GTC TGT AGA GTC TCA CAG GAG AGC A
	Fisk-r	CCA CCA AAC CAT GTA CGA CTA TGG
RAG2-r3	JohnD-r	GAT GGC CTT CCC TCT GTG GGT AC
	JohnD-r	GTG GAC TCC TGG CTG CAK CCC TG
16s	AR	CGC CTG TTT ATC AAA AAC AT
	BR	CCG GTC TGA ACT CAG ATC ACG T
ND3	F280	GAT GAG GMT CWT AAT CTT TCT AGT A
	R750	TTG ATT TCG RCTC AAA ARW TTR TGG

pinvar = 0.4492). A heuristic likelihood search with 10 random sequence additions was used to find the optimal tree. We used the ML tree as a framework to compare sequence divergence for each of the gene fragments. We also calculated likelihood scores in PAUP for each data partition using the ML tree with and without a molecular clock enforced.

For parsimony analyses, heuristic searches to find the most parsimonious tree(s) were performed using tree bisection–reconnection (TBR) branch swapping. One thousand random sequence addition replicates were used to minimize the chance of finding only locally optimal trees (Maddison, 1991). All sites were equally weighted and gaps treated as missing characters. We used nonparametric bootstrapping (Felsenstein, 1985) to measure support of clades with 1000 total pseudoreplicates and TBR branch swapping with 2 random sequence addition replicates per pseudoreplicate.

2.5. Phylogenetic analysis: Bayesian inference

We performed a Bayesian analysis to calculate posterior probabilities of clades using MrBayes 3.1.2 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). Data was partitioned by gene in all analyses and we chose the HKY model (Nst = 2), adopting a conservative approach that did not over-parameterize the analysis in response to advised caution against selecting a model more complex than necessary (Huelsenbeck et al., 2004). Results based on a more complex GTR model (Nst = 6) vs. a simpler model, however, were tested in analyses using MrBayes runs of ten million generations and we determined that the model choice had little effect on topology and support values in this analysis.

A Markov chain Monte Carlo simulation was run for 10 million generations, with trees sampled and saved every 1000 generations (10,000 trees saved per run). We ran four analyses, each with six chains, using the Illinois BioGrid courtesy of DePaul University. Default settings were used for all other parameters and included a flat prior, ts/tv rate ratio of Beta 1,1, branch length of Exp 10, alpha parameter of the gamma distribution of rate heterogeneity as uniform 0.1–50, proportion of invariant sites as uniform 0–1, base frequencies of Dirichlet 1,1,1,1, and tree topology parameters as uniform. All trees prior to both stationarity and convergence of the runs were discarded and a majority rule consensus was computed for those remaining. We assigned the beginning region of the plateau of stable log likelihood scores as the point at which stationarity was reached (427,000 generations). Convergence was reached when the value for the standard deviation of split frequencies remained below 0.01 (6.2 million generations). We therefore retained trees after 6.2 million generations. All post-burn in tree files were concatenated into a single file (totaling 15,200 trees), which was used to compute a majority rule consensus tree and posterior probabilities in PAUP*.

2.6. Hypothesis testing

To test the fit of various phylogenetic hypotheses to our data set, we used the SH Test of Shimodaira and Hasegawa (1999), which corrects for multiple comparisons, and tests the hypothesis that each tree in a set of trees is an equally good explanation of the data. We applied the RELL method of resampling with 1000 bootstrap replicates and $p < 0.05$. The test generated an average log likelihood for the collection of trees being considered, re-centered the distribution, and corrected for the bias of finding the ML tree and comparing suboptimal trees to it. Methods of hypothesis testing differ between traditional frequentist statistical approaches and the Bayesian school of hypothesis testing (Ghosh and Heo, 2003). We chose a post hoc frequentist approach to sampling the population of tree topologies derived from the Bayesian analysis. We constructed four filters, which consisted of several trees corresponding to systematic hypotheses of interest: (1) our Bayesian consensus solution similar to that in Fig. 1, without outgroups, (2) the Bayesian consensus solution at the clade level of resolution without outgroups, (3) taxa included in our analysis rearranged to match the Smith et al. (2003) hypothesis without outgroups, and (4) taxa included in our analysis rearranged to match the Littlewood et al. hypothesis without outgroups. These four filters were loaded as constraints in PAUP* and used to determine the number of matching trees from the Bayesian set of trees.

2.7. Rates, dates and evolutionary history

To investigate the possible dates of origin and diversification of major butterflyfish groups, we employed the penalized likelihood (PL) variable molecular clock module of the r8s 1.70 software of Sanderson (1997, 2002). The Bayesian MPP tree topology was used with the following parameters: algorithm = tn, rates = gamma, shape = 0.53, and smoothing = 10 (determined by cross validation). We set the root age of the outgroup node at 65 million years old, the transition between Mesozoic and Cenozoic marine fish faunas, and the probable earliest origins of many reef fish lineages (Bellwood and Wainwright, 2002). Using this date, we would expect to see origins of the family Chaetodontidae and Scatophagidae at or before 50 mya, marking the earliest known fossil evidence of most reef fish families (Bellwood et al., 2004). Few fossil Chaetodontidae are available for calibration, with the exception of *Chaetodon fischeuri* (Carnevale, 2004) a *Chaetodon* of uncertain phylogenetic position from approximately 7 mya. Thus, we set a minimum age of 7 mya on the root node of the genus *Chaetodon*.

3. Results

Phylogenetic analysis resulted in a well-resolved hypothesis of relationships at the species level (Fig. 1), and the tree topology supported monophyly of the Chaetodontidae, monophyly of all butterflyfish genera, and monophyly of

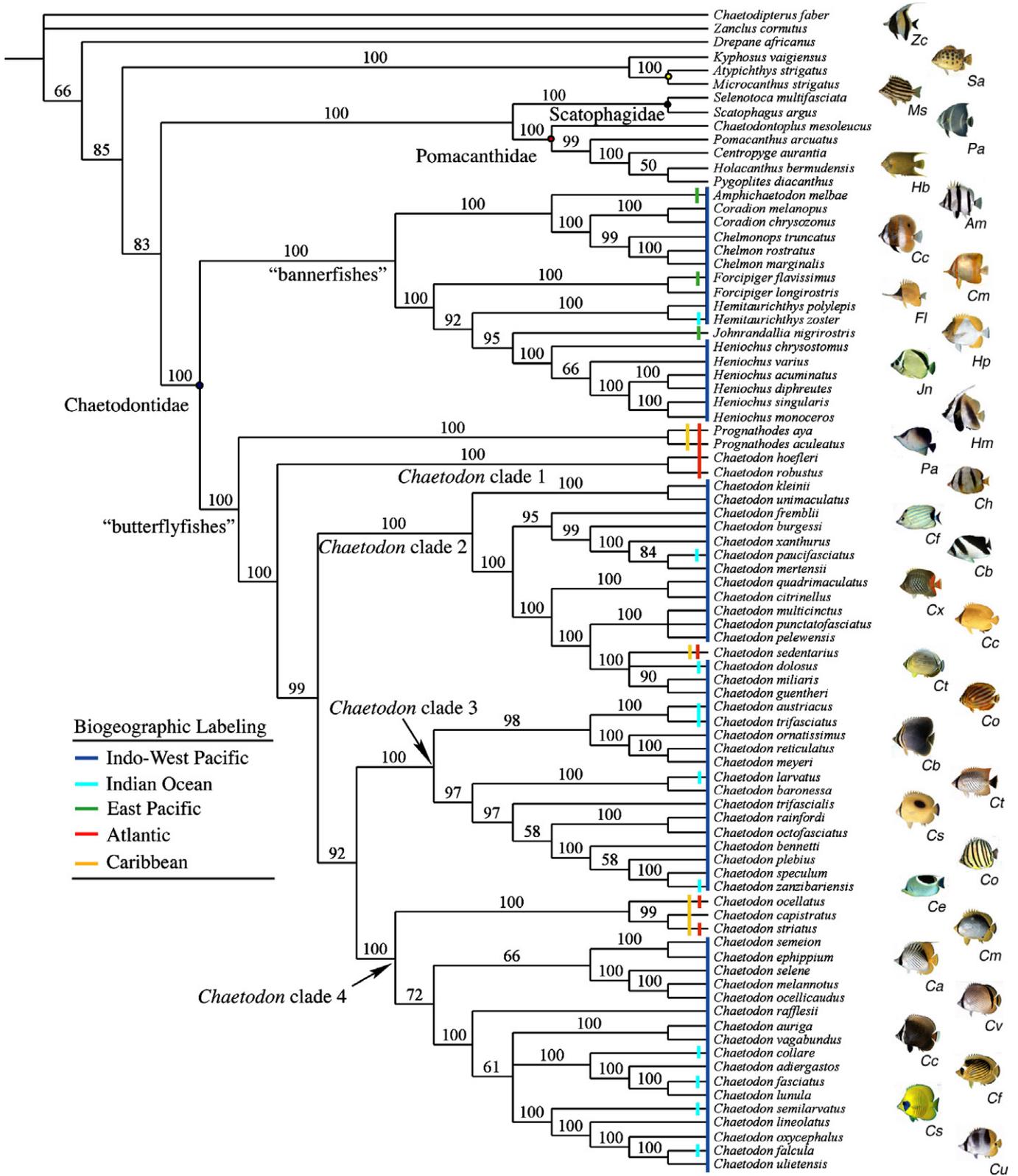


Fig. 1. Phylogeny of the Chaetodontidae resulting from Bayesian analysis of DNA sequences from 12s, 16s, ND3, Tmo-4c4, and RAG2 genes. The tree is the majority-rule consensus of 4 synchronous 10 million generation MCMC runs resulting in 36,000 sampled post-burn in trees. Posterior probabilities are indicated as the support for each node. Nodes corresponding to the families Chaetodontidae, Pomacanthidae, and Scatophagidae are labeled. Color coding at the tips indicate biogeographic provinces. Color photos are identified by the first letters of genus and species in proximity to the photo (courtesy of J.E. Randall and Fishbase.org).

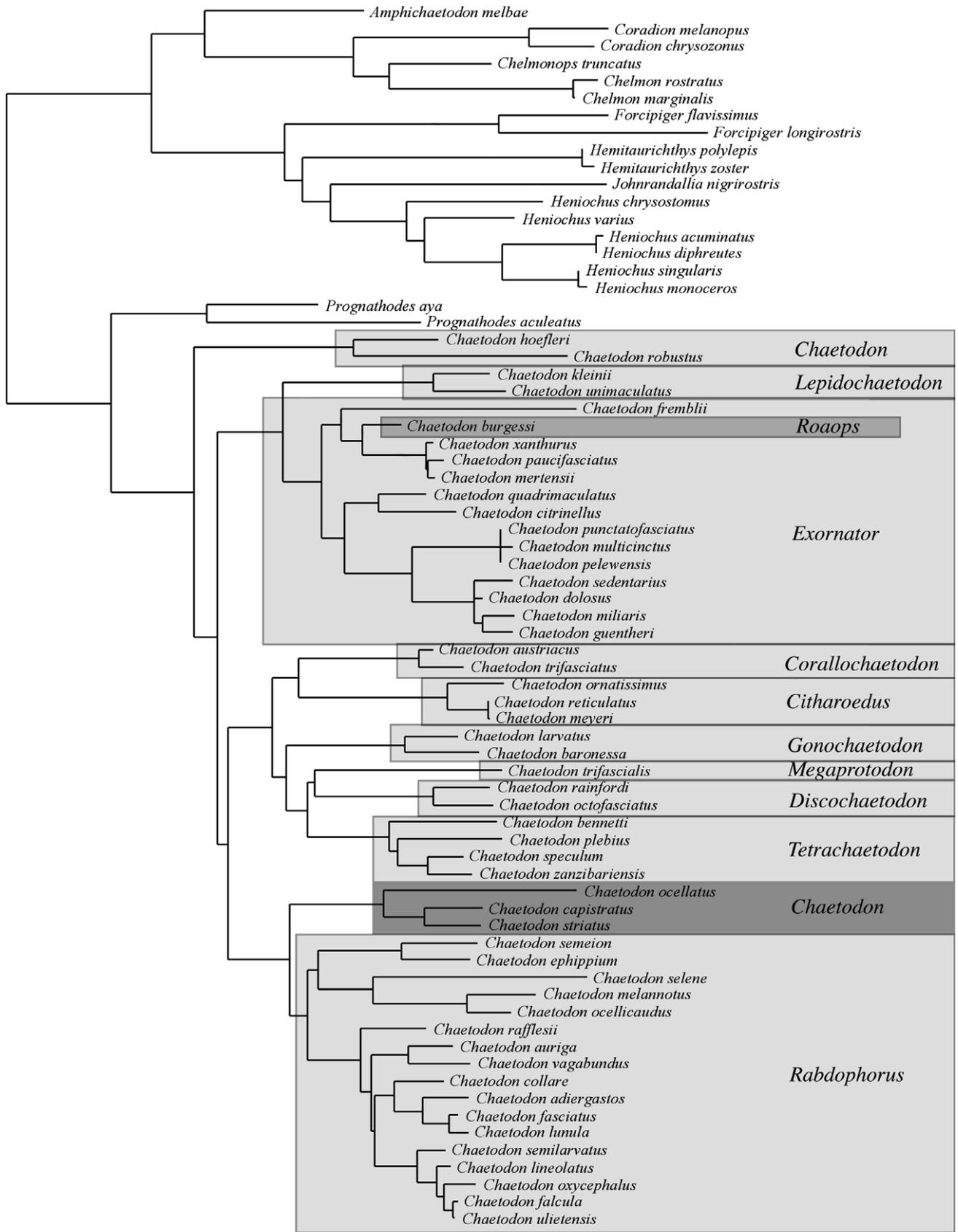


Fig. 2. Maximim likelihood phylogeny of the Chaetodontidae resulting from a heuristic search using the GTR+I+G model. Traditional subgenera of *Chaetodon* are indicated by gray boxes.

most but not all of the *Chaetodon* subgenera (Fig. 2). *C. burgessi*, which had been categorized as a member of subgenus *Roaops*, grouped within *C. (Exornator)* in our study, rendering it polyphyletic. *Chaetodon* subgenus *Chaetodon* (*sensu* Blum, 1988) was split into two non-sister groups. Other clades of butterflyfishes were congruent with previous hypotheses, although hypothesis testing of our phylogenetic branching patterns showed that they differed significantly from both the Blum (1988) morphological and Littlewood et al. (2004) molecular hypotheses of relationships between genera and subgenera. The Chaetodontidae originated approximately 37 mya, the origin of *Chaetodon* was about 20 mya, and several of the species complexes and sister pairs appear to have diverged within the past few million years.

3.1. Phylogenetic analyses

The Bayesian inference (BI) analysis resulted in a 95% credible set of 19,893 out of 21,393 unique trees sampled. Fig. 1 shows the resulting topology with posterior probabilities listed at each node. The maximum likelihood (ML) analysis conducted in PAUP resulted in a ML topology with a negative log likelihood of 28414.94 (Fig. 2). The BI and ML topologies were nearly identical, and differed only in the placement of taxa with very low support. Overall, we found five major divisions in butterflyfishes: (1) a basal clade formed by *Amphichaetodon*, *Chelmon* plus *Chelmonops*, and *Coradion*, (2) *Hemitaurichthys* and *Johnrandallia* with *Forcipiger* and *Heniochus*, (3) *Prognathodes* (4) *Chaetodon hoefleri* plus *C. robustus*, and (5) the rest of *Chaetodon*, including three major divisions within *Chaetodon*. We found strong support for monophyly of the genus *Chaetodon* and tentatively support monophyly of most of Blum's (1988) subgenera with a few exceptions (Table 3). All analysis methods split subgenus *Chaetodon sensu* Blum (1988) into two groups that are now widely separated on the BI tree. Our placement of *C. striatus* is with the Caribbean taxa portion of subgenus *C. (Chaetodon)*. This taxon was previously stated to be of uncertain position, and listed separately in the Smith et al. (2003) analysis, who placed it near, to but separate from, a monophyletic subgenus *C. Chaetodon* and basal to *C. (Rabdophorus)*. A portion of *C. (Chaetodon)* remains the most basal sister to the rest of *Chaetodon*. Excluding *C. (Rabdophorus)*, the next most basal split is similar to the Smith et al. (2003) hypothesis:

one node containing *Chaetodon* subgenera *Lepidochaetodon*, *Roaops*, and *Exornator*; the other containing subgenera *Citharoedus*, *Corallochaetodon*, *Gonochaetodon*, *Tetrachaetodon*, *Megaprotodon*, *Discochaetodon*, the Caribbean section of *C. (Chaetodon)*, and *Rabdophorus* (the last two being where our hypothesis differs at this level in the hierarchy). After that node, however, there are few other comparable patterns except that *C. (Exornator)* remains grouped with *C. (Lepidochaetodon)*. In addition, *C. burgessi* (subgenus *Roaops sensu* Blum, 1988) was placed within *C. (Exornator)* in our analysis.

3.2. Data partitions, congruence, and the utility of molecular characters

GenBank accession numbers for the nucleotide data are the following: 12s (EF616820–EF616903), 16s (EF616904–EF616985), ND3 (EF616986–EF617067), RAG2 (EF617068–EF617150), and Tmo-4C4 (EF617151–EF617232). Combined analysis provided good resolution at multiple levels of the phylogeny, and gene partitions displayed noticeably different base composition and T_i/T_v ratios (Table 2). We obtained (after trimming primers) 981 bp of 12S, 511 bp of Tmo-4C4, 803 bp of RAG2, 640 bp of 16s, and 397 bp of ND3, for a total of 3332 bp. After exclusion of short ends and ambiguous regions, we used 790 bp of 12s, 460 bp of TMO4c4, 765 bp of RAG2, 460 bp of 16s, and 376 bp of ND3 for a total of 2851 characters.

Likelihood ratio testing (LRT) resulted in the rejection of clock-like rates of evolution for all data partitions. Examination of the rates of each gene mapped onto the ML topology (Figs. 3 and 4) illustrates the pace of nucleotide substitution and the contribution of each gene to particular regions of the tree. Mitochondrial 12s contributed characters that help to resolve relationships both basally and among species, with particularly useful sequence divergence among the bannerfishes and coralfishes and among *Chaetodon* species-level topology such as the *Chaetodon* clades 2 and 4 (Fig. 3a). The 16s sequence (not illustrated) was most useful among the outgroups and at deeper nodes in the tree. A faster locus with more substitution along tip branches, ND3 was informative regarding relationships between closely-related taxa, such as species pairs within *Chaetodon* and within *Coradion* and *Chelmon* (Fig. 3b). The two nuclear genes contributed useful information to resolve clades at multiple levels in

Table 2

Base composition, maximum percentages of divergence among Chaetodontidae and among all taxa, including outgroups, and transition to transversion ratios for each gene partition

Partition	Nucleotides included	%A	%C	%G	%T	Max % ingroup divergence	Max % divergence to outgroup	T_i/T_v
12s	790	30.1	25.3	22.9	21.6	10.93	18.18	3.27
4c4	460	28.1	20.5	25.3	26.1	7.18	20.87	4.25
RAG2	765	21.9	26.1	26.4	25.6	6.96	24.66	2.87
16s	460	27.1	24.2	25.2	23.5	9.47	24.96	3.52
ND3	376	22.5	30.5	16.2	30.7	24.66	58.45	3.78

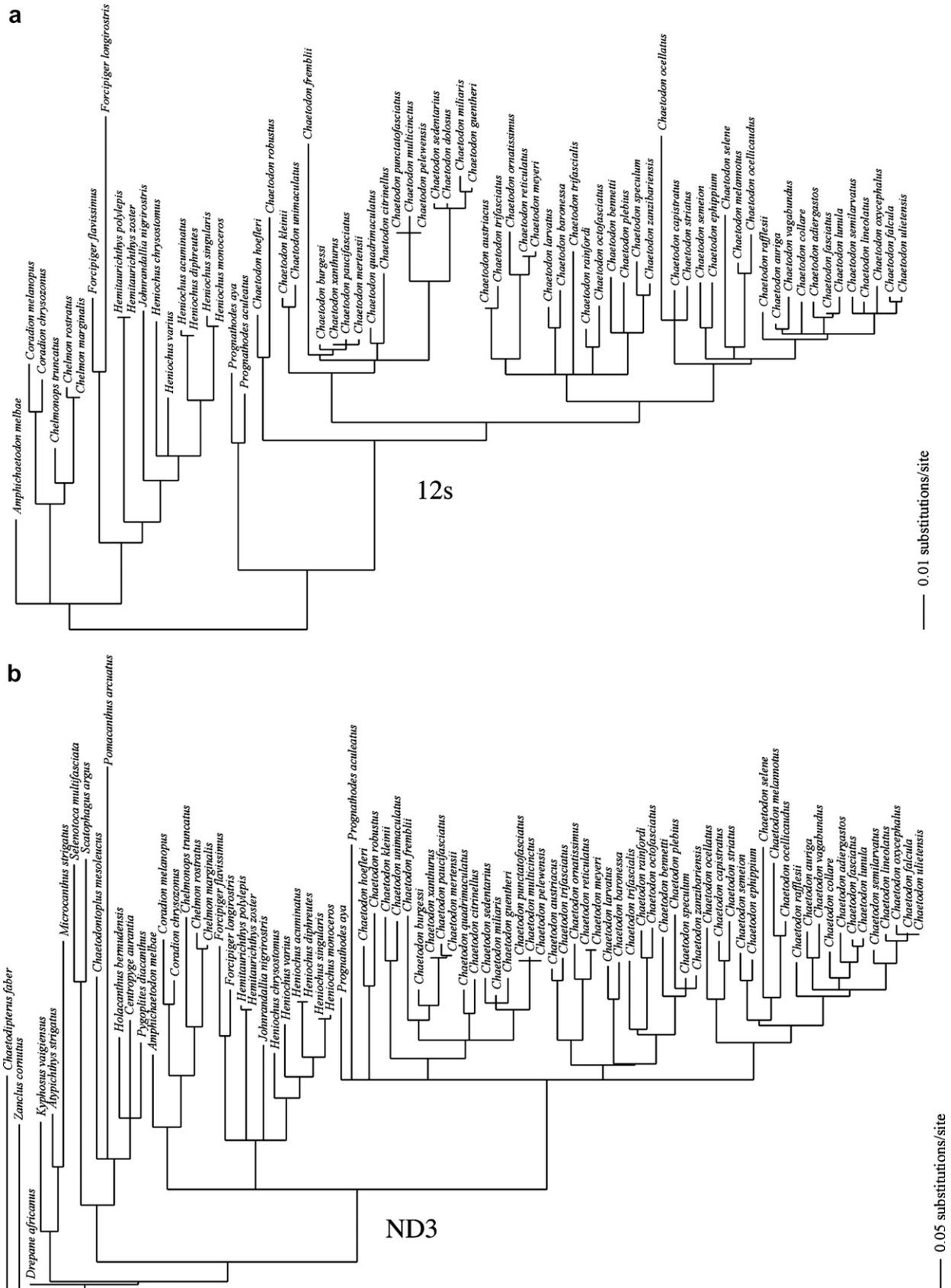


Fig. 3. Evolutionary map of gene sequence change in two mitochondrial partitions in the Chaetodontidae. Substitution rates are illustrated on the maximum likelihood topology for (a) the mitochondrial 12s and (b) mitochondrial ND3 gene. Note that these are character optimizations on the topology of Fig. 2, not gene trees. Outgroups were included in the optimization of characters on the ML topology, but removed for the illustration.

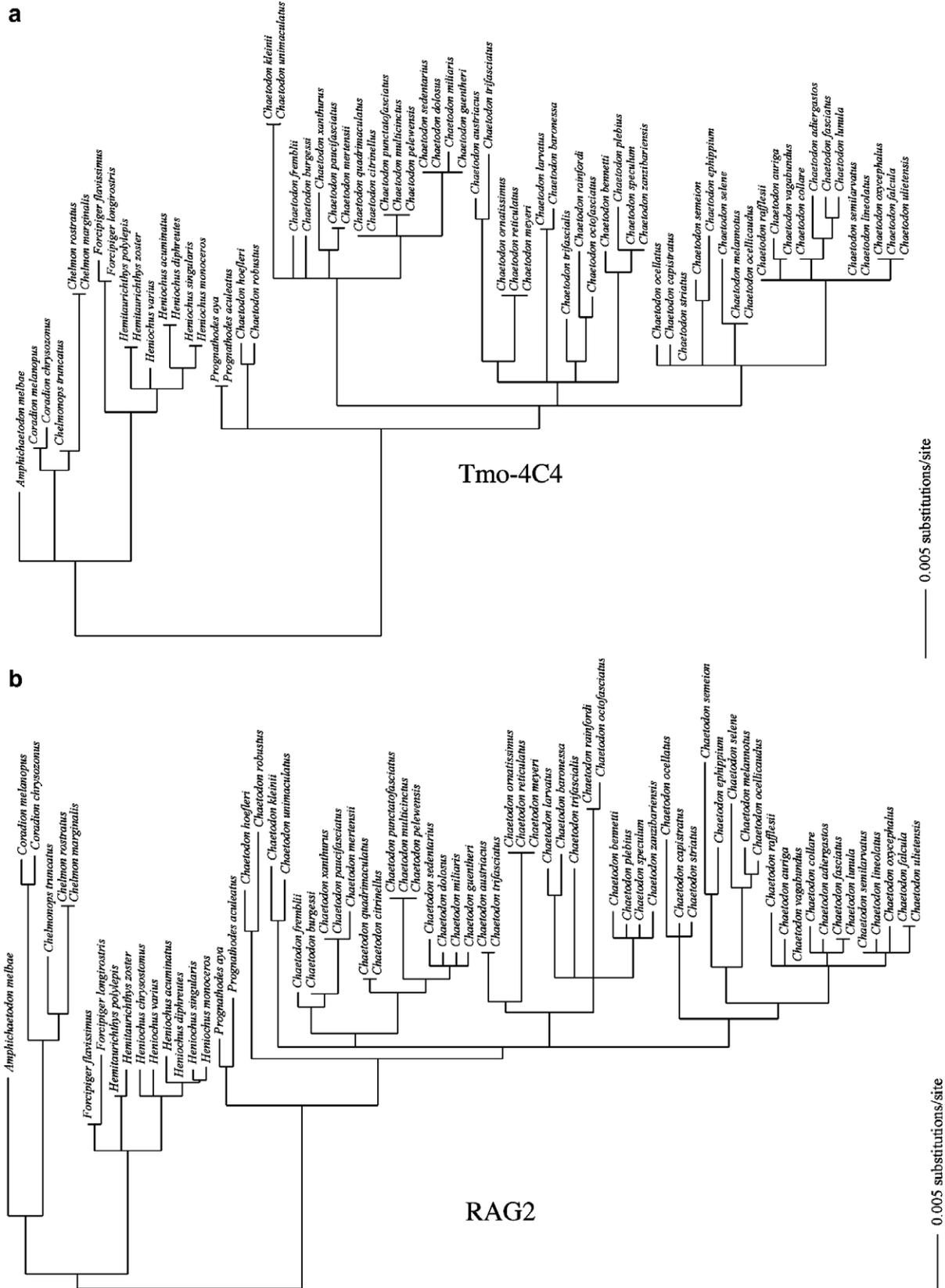


Fig. 4. Evolutionary map of gene sequence change in two nuclear partitions in the Chaetodontidae. Substitution rates are illustrated on the maximum likelihood topology for (a) the nuclear TMO4c4 gene and (b) the nuclear RAG2 gene. Note that these are character optimizations on the topology in Fig. 2, not gene trees. Outgroups were included in the optimization of characters on the ML topology, but removed for the illustration.

the tree, and were nicely complementary in their information content (Fig. 4). For example, Tmo-4C4 (Fig. 4a) was informative in the resolution of the *Forcipiger* clade plus closely-related bannerfishes, whereas RAG2 helped to resolve some of the species groups within *Chaetodon* such as *Chaetodon* clade 4 (Figs. 1 and 4b).

3.3. Hypothesis testing of phylogenetic relationships

Four filters were used to search the total set of 15,200 Bayesian inference (BI) post-burn topologies. Results are expressed as the number of trees found to be in agreement and retained using each filter. The currently proposed Bayesian inference (BI) solution without outgroups (Fig. 1) filter retained a match of 2942 trees, or 18.3% of the sampled trees. When our Bayesian tree was pruned to include just the major clades of butterflyfishes, then 9171 matching trees are retained, or 26% of the sampled trees. We found that zero trees in our Bayesian set matched either the Smith et al. (2003) or Littlewood et al. (2004) hypotheses. Our molecular phylogeny thus disagrees with the previously proposed pattern of branching among several of the subgeneric clades within the genus *Chaetodon* in several areas of the phylogeny (Fig. 5).

SH test results on the same four tree topologies assigned the most favorable rank to the Bayesian consensus tree, with a likelihood score of -29917.7 , second to the Bayesian consensus tree at the clade resolution, with a likelihood score of -31951.2 , third to the Smith et al. (2003) morphology-based tree with a likelihood score of -32875.7 , and the least likely was the Littlewood et al. (2004) hypothesis with a likelihood score of -39134.0 .

3.4. Origins, divergence times, and biogeography

Penalized likelihood (Sanderson, 2002) estimates using an initial root age estimate of 65 mya yielded an age estimate for the Chaetodontidae + Scatophagidae + Pomacanthidae at about 54 mya, with the root of the family Chaetodontidae at about 37 mya (Fig. 6 and Table 4), the date of the major split between the ancestors of the two major butterflyfish clades. The root of the first major clade, including *Amphichaetodon*, *Coradion*, *Chelmon*, *Forcipiger*, *Johnrandallia*, and *Heniochus* was estimated to be 33 mya, with some of the major splits between genera in that clade ranging from 20 to 28 mya (Table 4). The second major chaetodontid clade, *Prognathodes* + *Chaetodon*, showed a root age of about 24 mya, with the estimated initial divergences among the sampled species from the four major clades of *Chaetodon* occurring from about 9 to 14 mya (Fig. 6). Many of the species sister pairs and small multi-species clades were estimated to have radiated in just the past few million years (Table 4).

The vast majority of butterflyfishes are located in the Indo-West Pacific, including members from all basal clades. Biogeographic mapping of species locality onto the phylogeny (Fig. 1) suggests that there have been three

invasions of clades from the Pacific into the Atlantic. Two clades of butterflyfishes that live in the Atlantic are adjacent to each other in the molecular tree (genus *Prognathodes* and part of subgenus *C. Chaetodon*). A second clade split from the now polyphyletic subgenus *C. (Chaetodon)*, and also the Atlantic *C. sedentarius* of the *C. (Exornator)* clade, are both Atlantic inhabitants. There are many examples of clades that show a split into separate, closely-related Pacific and Indian Ocean relatives or sister-species pairs (Fig. 1). Finally, *Johnrandallia*, *Forcipiger*, and one species of *Amphichaetodon* are all members of the diverse sister-group to *Chaetodon* that are located in the Eastern Pacific Ocean. Although not included in the current analysis, members of *Prognathodes* and subgenus *C. (Chaetodon)* also occur in the Eastern Pacific.

4. Discussion

Phylogenetic analysis of molecular data for the Chaetodontidae resulted in a robust phylogeny that agrees with previous morphological hypotheses at the level of most major clades and some of the relationships among butterflyfish genera, but disagrees with earlier work regarding the topology of relationships within *Chaetodon*. We use the present molecular data set to statistically test previous phylogenetic hypotheses in a Bayesian framework and to reevaluate taxonomic groupings within the family. Our data support most genus-level nodes of the Blum (1988) and Smith et al. (2003) data sets but support different relationships in other regions of the tree. We found strong support for four major clades within *Chaetodon*, and did not support the *Chaetodon* subgenus *C. (Exornator)*, which contains the subgenus *C. (Roaops)*, or subgenus *C. (Chaetodon)* which was split into two non-sister groups. This finding is consistent with the McMillan and Palumbi (1995) result that *C. argentatus*, a member of *Chaetodon* subgenus *Exornator* (*sensu* Blum, 1988) was instead forming the basal section of the *C. (Roaops)* clade. The molecular phylogeny suggests a Pacific origin of the family Chaetodontidae, with a likely date for genesis of the family of about 37 mya at the end of the Eocene, somewhat more recent than the origin of many other major reef fish groups at 50 mya or earlier. Biogeographic analysis of the family supports three invasions of the Atlantic/Caribbean region, an early history of butterflyfish presence and diversification in the Eastern Pacific, and a repeated pattern of Pacific/Indian Ocean splits between sister-species pairs.

4.1. Phylogeny and taxonomy of the Chaetodontidae

Using a range of squamipinne outgroups, we found that the monophyletic Chaetodontidae are most closely related to a sister-family pair of Scatophagidae + Pomacanthidae (Fig. 1). This particular relationship has not previously been proposed, although Bellwood et al. (2004) identified this closely-related triumvirate and placed the scats as sister to the butterflyfishes. A more complete sampling at the

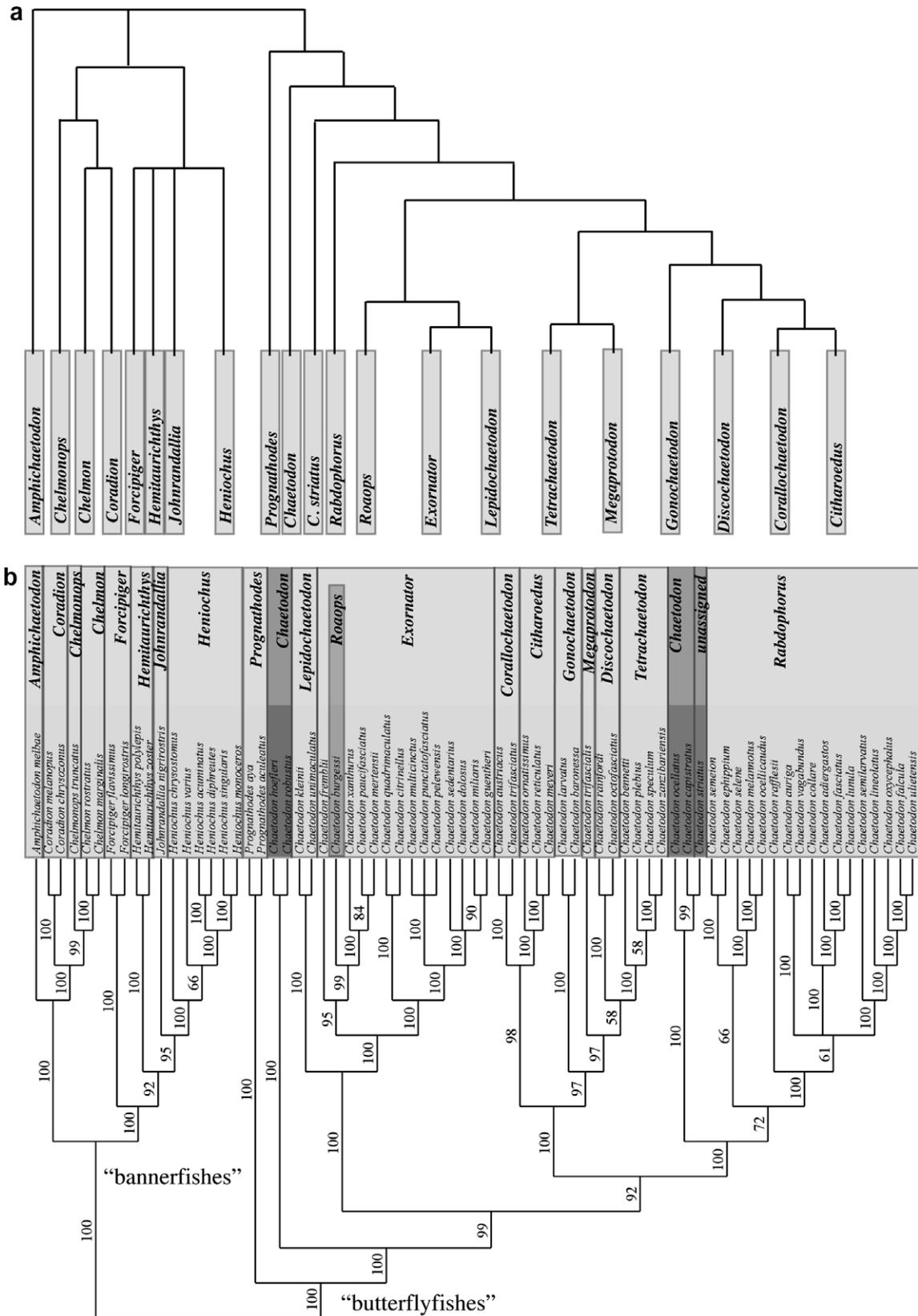


Fig. 5. A comparison of (a) the morphological topology of Smith et al. (2003) to (b) the present molecular hypothesis for the Chaetodontidae. Major features of the higher-level tree are similar, with the exception of the position of *Amphichaeatodon*, however phylogenetic relationships among subgroups (often defined as subgenera) within the genus *Chaetodon* (light gray boxes) show significant topological differences. *Chaetodon* subgenera for which monophyly is rejected are indicated by dark gray boxes.

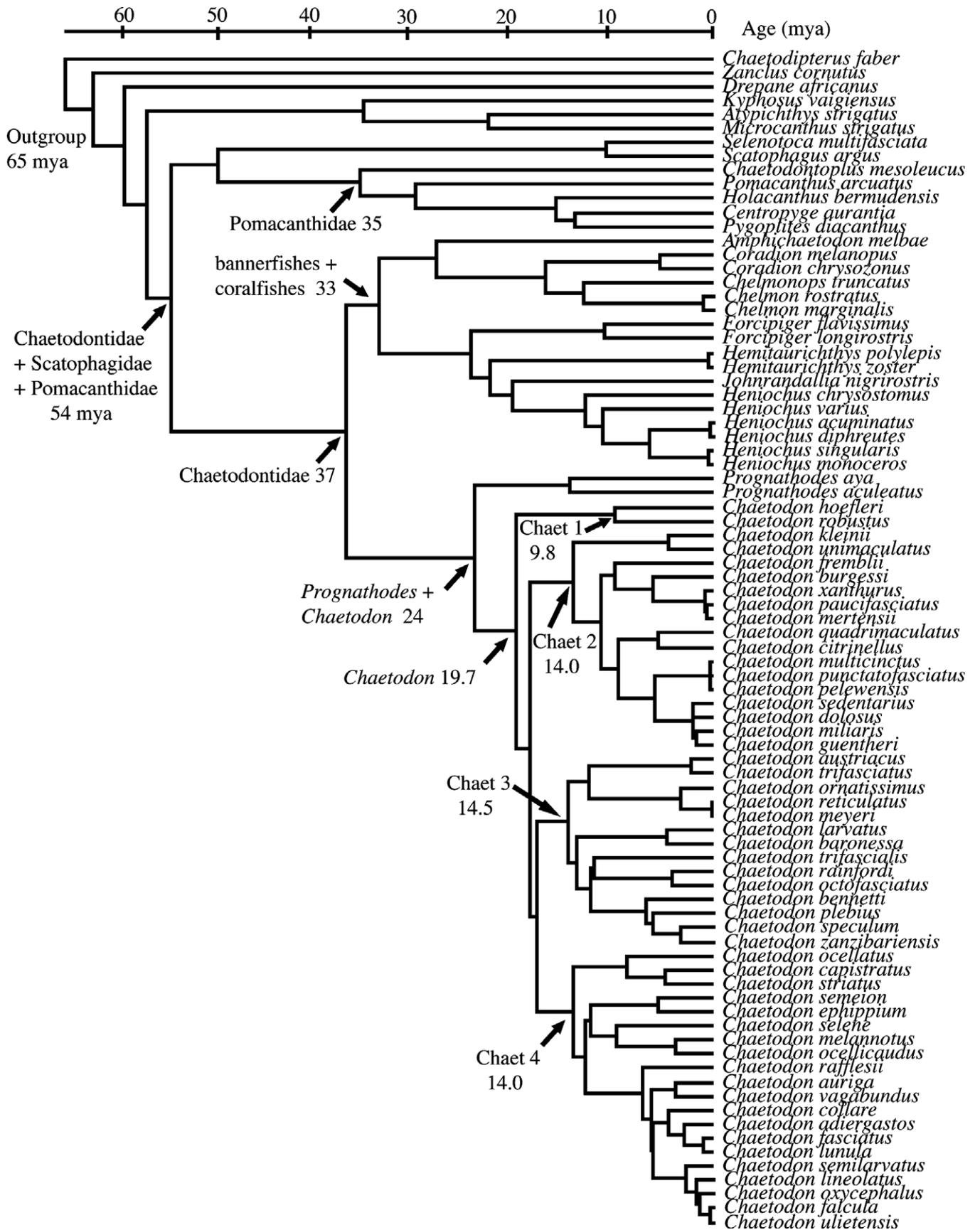


Fig. 6. A chronogram for the family Chaetodontidae and close relatives, based on penalized likelihood analysis (r8s software, Sanderson, 2002). Major clades of interest are labeled with estimated ages in millions of years.

species level among the various squamipinne families will be required to assess how molecular phylogenetics can contribute to the higher-level question of interfamily relationships.

The Chaetodontidae are split into two major lineages, the “bannerfish + coralfish” lineage that also includes *Amphichaetodon* and the long-nosed *Forcipiger*, and the “butterflyfish” lineage that includes *Prognathodes* and *Chaetodon* (Fig. 1). This topology with two large sister-clades has not been previously proposed. Previous morphological studies, primarily based on the dataset of Blum (1988), identified three primary clades, consisting of (1) a clade of *Amphichaetodon*, (2) a clade of *Chelmonops*, *Chelmon*, *Coradion*, *Forcipiger*, *Hemitaurichthys*, *Heniochus*, and *Johnrandallia*, and (3) a clade of *Chaetodon*, *C. (Roa)*, and *Prognathodes*. Our molecular phylogeny agrees with the higher-level structure of these hypotheses, with some exceptions (Fig. 5). Our study did not find *Amphichaetodon* to be the sister-group to all other butterflyfishes, as in previous studies, however it was placed with strong support at the base of the clade also containing *Coradion*, *Chelmonops* and *Chelmon* (Fig. 5). Other differences at varying taxonomic levels are worth noting. For example, we found high support for (*Chelmonops* + *Chelmon*) with *Coradion* as sister to them, rather than the Smith et al. (2003) result of ((*Chelmon* + *Coradion*) + *Chelmonops*). Finally, our analysis resolved the previous polytomy presented by Blum (1988) of (*Forcipiger* + *Hemitaurichthys* + *Heniochus* + *Johnrandallia*).

Also in agreement with morphological topologies, *Prognathodes* remained the sister-group to the genus *Chaetodon*. Within *Chaetodon*, however, relationships among the subgenera bear less likeness to previous hypotheses (Fig. 5). In our study, subgenus *Rabdophorus* was more deeply nested than as reported in the morphologically-based tree. Smith et al. (2003) placed *C. (Rabdophorus)* basal to the clade containing *C. (Exornator)* + *C. (Lepidochaetodon)*. We found *C. (Exornator)* grouped with *C. (Lepidochaetodon)* and immediately basal to *C. (Rabdophorus)*. Smith et al. (2003) also found *C. (Discochaetodon)* with *C. (Corallochaetodon)* + *C. (Citharoedus)*; we found *C. (Discochaetodon)* + *C. (Tetrachaetodon)* (with low support) but also our node uniting all of *C. (Discochaetodon)* + *C. (Tetrachaetodon)* with *C. (Megaprotodon)* as sister emerged with high support. The position of *C. (Gonochaetodon)* differed as well. Smith et al. (2003) placed *C. (Gonochaetodon)* as sister to (*C. (Discochaetodon)* + (*C. (Corallochaetodon)* + *C. (Citharoedus)*)), whereas our results show a well-supported node with *C. (Gonochaetodon)* the sister to ((*C. (Discochaetodon)* + *C. (Tetrachaetodon)*) + *C. (Megaprotodon)*).

Most discussions of butterflyfish taxonomy have largely followed Blum (1988) and Smith et al. (2003). However, a number of other taxonomies are also widely used, and this has created some confusion in this literature (Allen et al., 1998; Burgess, 1978; Kuitert, 2002) due to the practice of retaining the same subgenus name for different sets of taxa

in popular aquarium books without formal revision. Here we summarize the current sets of species included in *Chaetodon* subgenera (Table 4) in order to illustrate this diversity of taxonomies. Blum (1988) modified membership in many of the subgenera of Burgess (1978), elevated *Roa* and *Prognathodes* from subgenus to genus status, and demoted *Parachaetodon* from genus to subgenus level. Blum (1988) also recognized *C. (Roaops)* and *C. (Exornator)*, but eliminated the use of *C. (Rhombochaetodon)*, *C. (Heterochaetodon)*, *C. (Chaetodontops)*, and *C. (Pseudochaetodon)*. Kuitert (2002) largely followed these conventions while modifying membership of *C. (Lepidochaetodon)*, *C. (Roaops)*, *C. (Exornator)*, and *C. (Rabdophorus)*, recognizing *C. (Rhombochaetodon)* and *C. (Heterochaetodon)* which had been eliminated by Blum (1988), and following the use of *Parachaetodon* as a genus. Regarding these membership modifications, one interesting point is that the species *Chaetodon reticulatus* and *C. collare* are strikingly similar in appearance. Blum (1988), however, placed *C. reticulatus* some distance away from *C. collare* and with both *C. ornatissimus* and *C. meyeri* (as part of subgenus *C. Chaetodon*), a decision which Kuitert (2002) reversed. While the current study did not find monophyly of the entire subgenus *C. (Chaetodon)*, our data agreed with the Blum (1988) conclusion and did not support a close relationship between *C. collare* and *C. reticulatus*. Allen et al. (1998) largely followed the grouping of Burgess (1978) but recognized *Prognathodes* as a genus and also eliminated the use of *Pseudochaetodon*. The result is a taxonomy for species within the single genus *Chaetodon* upon which there is significant disagreement.

We evaluated the membership of *Chaetodon* subgroups by assessing their validity strictly in terms of monophyly in our phylogenetic hypothesis (grayscale coding of Table 3) without proposing any taxonomic changes. Our current study included enough taxa to evaluate over half of these groupings. We considered a hypothesis of monophyly partially testable if we were able to acquire at least two of the taxa, although we often had many species from a group. We conclude that there is a strong genetic basis for subdividing the genus *Chaetodon* into four major clades (using Blum (1988) subgeneric definitions): (1) *C. (Chaetodon)*, (2) *C. (Lepidochaetodon)* + *C. (Roaops)* + *C. (Exornator)*, (3) *C. (Corallochaetodon)* + *C. (Citharoedus)* + *C. (Gonochaetodon)* + *C. (Tetrachaetodon)* + *C. (Megaprotodon)* + *C. (Discochaetodon)*, and (4) *C. (Rabdophorus)*. Each of these clades received phylogenetic support with long branch lengths (Figs. 1 and 2) and seem likely to withstand additional species and genetic sampling. The node corresponding to the monophyletic *C. (Rabdophorus)* clade received lower support than the other clades (Fig. 1). However, in an exploratory Bayesian analysis excluding the ND3 partition, the root node of *C. (Rabdophorus)* had much higher posterior support (100), although it showed a corresponding loss of resolution within the clade.

If the basal *Chaetodon* genus stands, then the other 3 major cladistic groups would require elevation to genus

Table 3

Taxa belonging to each of the prevailing taxonomic hypotheses (identified by author and date) for subgenera of *Chaetodon*

	monophyly supported	monophyly rejected	not testable in current study	
Author <i>Chaetodon</i> subgenus	Blum, 1988	Burgess, 1978	Allen & Steene, 1998	Kuiter, 2002
<i>Citharoedus</i> subgenus Kaup, 1860	<i>C. meyeri</i> ✓ <i>C. ornatissimus</i> ✓ <i>C. reticulatus</i> ✓	<i>C. meyeri</i> ✓ <i>C. ornatissimus</i> ✓ <i>C. reticulatus</i> ✓	<i>C. meyeri</i> ✓ <i>C. ornatissimus</i> ✓ *	<i>C. meyeri</i> ✓ <i>C. ornatissimus</i> ✓ *
<i>Corallochaetodon</i> subgenus Burgess, 1978	<i>C. austriacus</i> ✓ * <i>C. melapterus</i> <i>C. trifasciatus</i> ✓	<i>C. austriacus</i> ✓ * <i>C. melapterus</i> <i>C. trifasciatus</i> ✓	<i>C. austriacus</i> ✓ <i>C. lunulatus</i> <i>C. melapterus</i> <i>C. trifasciatus</i> ✓	<i>C. austriacus</i> ✓ <i>C. lunulatus</i> <i>C. melapterus</i> <i>C. trifasciatus</i> ✓
<i>Tetrachaetodon</i> subgenus Weber & de Beaufort, 1936	<i>C. bennetti</i> ✓ <i>C. plebius</i> ✓ <i>C. speculum</i> ✓ <i>C. zanzibariensis</i> ✓ *	<i>C. bennetti</i> ✓ <i>C. plebius</i> ✓ <i>C. speculum</i> ✓ <i>C. zanzibariensis</i> ✓ *	<i>C. bennetti</i> ✓ <i>C. plebius</i> ✓ <i>C. speculum</i> ✓ <i>C. zanzibariensis</i> ✓ *	<i>C. bennetti</i> ✓ <i>C. plebius</i> ✓ <i>C. speculum</i> ✓ <i>C. zanzibariensis</i> ✓ <i>C. andamanensis</i>
<i>Lepidochaetodon</i> subgenus	* * <i>C. kleinii</i> ✓ * * * <i>C. trichrous</i> <i>C. unimaculatus</i> ✓	* * * * * * <i>C. unimaculatus</i> ✓	* * * * * * <i>C. unimaculatus</i> ✓	<i>C. daedalma</i> <i>C. interruptus</i> <i>C. kleinii</i> ✓ <i>C. litus</i> <i>C. nippon</i> <i>C. smithi</i> <i>C. trichrous</i> <i>C. unimaculatus</i> ✓
<i>Roa</i> genus, Jordan 1923	as a genus * * <i>R. excelsa</i> * <i>R. jakari</i> * <i>R. modestus</i> * *	as a subgenus <i>C. burgessi</i> ✓ <i>C. declevis</i> <i>C. excelsa</i> <i>C. flavocoronatus</i> <i>C. jakari</i> <i>C. mitratus</i> <i>C. modestus</i> <i>C. nippon</i> <i>C. tinkeri</i>	as a subgenus <i>C. burgessi</i> ✓ <i>C. declevis</i> <i>C. excelsa</i> <i>C. flavocoronatus</i> * <i>C. mitratus</i> <i>C. modestus</i> <i>C. nippon</i> <i>C. tinkeri</i>	as a genus * * <i>R. excelsa</i> * <i>R. jakari</i> * <i>R. modestus</i> * *
<i>Rooops</i> subgenus Mauge & Bauchot, 1984	as a subgenus <i>C. burgessi</i> ✓ <i>C. declevis</i> * <i>C. flavocoronatus</i> <i>C. mitratus</i> <i>C. nippon</i> <i>C. tinkeri</i>	not recognized	not recognized	as a subgenus <i>C. burgessi</i> ✓ <i>C. declevis</i> <i>P. guyotensis</i> <i>C. flavocoronatus</i> <i>C. mitratus</i> * <i>C. tinkeri</i>
<i>Rhombochaetodon</i> subgenus Burgess, 1978	Not recognized: Blum combined with <i>Exornator</i> , and called <i>Exornator</i>	<i>C. argentatus</i> * * <i>C. madagascariensis</i> <i>C. mertensii</i> ✓ <i>C. paucifasciatus</i> ✓ <i>C. xanthurus</i> ✓	<i>C. argentatus</i> * * * <i>C. mertensii</i> ✓ <i>C. paucifasciatus</i> ✓ <i>C. xanthurus</i> ✓	<i>C. argentatus</i> <i>C. blackburnii</i> <i>C. fremblii</i> ✓ <i>C. madagascariensis</i> <i>C. mertensii</i> ✓ <i>C. paucifasciatus</i> ✓ <i>C. xanthurus</i> ✓
<i>Exornator</i> subgenus Nalbant 1971	<i>C. argentatus</i> <i>C. assarius</i> <i>C. blackburnii</i> <i>C. citrinellus</i> ✓ <i>C. daedalma</i> <i>C. dolossus</i> ✓ <i>C. fremblii</i> ✓ <i>C. guentheri</i> ✓ <i>C. guttatissimus</i> <i>C. litus</i> <i>C. madagascariensis</i> <i>C. mertensii</i> ✓ <i>C. miliaris</i> ✓ <i>C. multicinctus</i> ✓ <i>C. paucifasciatus</i> ✓ <i>C. pelewensis</i> ✓ <i>C. punctatofasciatus</i> ✓ <i>C. quadrimaculatus</i> ✓ <i>C. sanctahelenae</i> <i>C. sedentarius</i> ✓ <i>C. smithi</i> <i>C. xanthurus</i> ✓	not recognized	not recognized	* * * <i>C. citrinellus</i> ✓ * * * * <i>C. guttatissimus</i> * * * * <i>C. multicinctus</i> ✓ * <i>C. pelewensis</i> ✓ <i>C. punctatofasciatus</i> ✓ <i>C. quadrimaculatus</i> ✓ * * *

(continued on next page)

Table 3 (continued)

<i>Heterochaetodon</i> subgenus Mauge & Bauchot, 1984	not recognized	not recognized	not recognized	<i>C. assarius</i> <i>C. dolosus</i> ✓ <i>C. guentheri</i> ✓ <i>C. miliaris</i> ✓ <i>C. sanctahelenae</i> <i>C. sedentarius</i> ✓
<i>Chaetodon</i> subgenus Linnaeus, 1758	* * <i>C. capistratus</i> ✓ * * * * * * * <i>C. hoefleri</i> ✓ <i>C. humeralis</i> * * <i>C. marleyi</i> * * * <i>C. ocellatus</i> ✓ * * * <i>C. robustus</i> ✓ * * * * *	<i>C. assarius</i> <i>C. blackburni</i> <i>C. capistratus</i> ✓ <i>C. citrinellus</i> ✓ <i>C. daedalma</i> * <i>C. dolosus</i> ✓ <i>C. fremblii</i> ✓ <i>C. guentheri</i> ✓ <i>C. guttatissimus</i> <i>C. hoefleri</i> ✓ <i>C. humeralis</i> <i>C. kleinii</i> ✓ <i>C. litus</i> <i>C. marleyi</i> <i>C. melannotus</i> ✓ <i>C. miliaris</i> ✓ <i>C. multicinctus</i> ✓ <i>C. ocellatus</i> ✓ <i>C. ocellicaudus</i> ✓ <i>C. pelewensis</i> ✓ <i>C.</i> <i>punctatofasciatus</i> ✓ <i>C. quadrimaculatus</i> ✓ <i>C. robustus</i> ✓ <i>C. sanctahelenae</i> <i>C. sedentarius</i> ✓ <i>C. smithi</i> <i>C. striatus</i> ✓ <i>C. trichrous</i>	<i>C. assarius</i> <i>C. blackburni</i> <i>C. capistratus</i> ✓ <i>C. citrinellus</i> ✓ <i>C. daedalma</i> <i>C. dialeucos</i> <i>C. dolosus</i> ✓ <i>C. fremblii</i> ✓ <i>C. guentheri</i> ✓ <i>C. guttatissimus</i> <i>C. hoefleri</i> ✓ <i>C. humeralis</i> <i>C. kleinii</i> ✓ <i>C. litus</i> <i>C. marleyi</i> <i>C. melannotus</i> ✓ <i>C. miliaris</i> ✓ <i>C. multicinctus</i> ✓ <i>C. ocellatus</i> ✓ <i>C. ocellicaudus</i> ✓ <i>C. pelewensis</i> ✓ <i>C.</i> <i>punctatofasciatus</i> ✓ <i>C. quadrimaculatus</i> ✓ <i>C. robustus</i> ✓ <i>C. sanctahelenae</i> <i>C. sedentarius</i> ✓ <i>C. smithi</i> <i>C. striatus</i> ✓ <i>C. trichrous</i>	* * <i>C. capistratus</i> ✓ * * * * * * <i>C. hoefleri</i> ✓ <i>C. humeralis</i> * * <i>C. marleyi</i> * * * <i>C. ocellatus</i> ✓ * * * <i>C. robustus</i> ✓ * * * * *
Uncertain position	<i>C. striatus</i> ✓	*	*	*
<i>Prognathodes</i> subgenus and genus Gill, 1862	as a genus	as a subgenus <i>C. aculeatus</i> ✓ <i>C. aya</i> ✓ <i>C. dichrous</i> <i>C. falcifer</i> <i>C. guezi</i> <i>C. guyanensis</i> <i>C. guyotensis</i> <i>C. marcellae</i> <i>C. obliquus</i>	as a genus	as a genus
<i>Rabdophorus</i> (<i>Radophorus</i>) subgenus Swainson, 1839	<i>C. adiergastos</i> ✓ <i>C. auriga</i> ✓ <i>C. auripes</i> <i>C. collare</i> ✓ <i>C. decassatus</i> <i>C. ephippium</i> ✓ <i>C. falcula</i> ✓ <i>C. fasciatus</i> ✓ <i>C. flavirostris</i> <i>C. gardineri</i> <i>C. leucopleura</i> <i>C. lineolatus</i> ✓ <i>C. lunula</i> ✓ <i>C. melannotus</i> ✓ <i>C. mesoleucos</i> <i>C. nigropunctatus</i> <i>C. ocellicaudus</i> ✓ <i>C. oxycephalus</i> ✓ <i>C. rafflesi</i> ✓ <i>C. selene</i> ✓ <i>C. semeion</i> ✓ <i>C. semilarvatus</i> ✓ <i>C. ulietensis</i> ✓ <i>C. vagabundus</i> ✓ <i>C. wiebeli</i> <i>C. xanthocephalus</i>	* <i>C. auriga</i> ✓ * * <i>C. decassatus</i> <i>C. ephippium</i> ✓ <i>C. falcula</i> ✓ * * <i>C. gardineri</i> <i>C. leucopleura</i> <i>C. lineolatus</i> ✓ * * <i>C. mesoleucos</i> <i>C. nigropunctatus</i> * <i>C. oxycephalus</i> ✓ <i>C. rafflesi</i> ✓ <i>C. selene</i> ✓ <i>C. semeion</i> ✓ * <i>C. ulietensis</i> ✓ <i>C. vagabundus</i> ✓ * <i>C. xanthocephalus</i>	* <i>C. auriga</i> ✓ * * <i>C. decassatus</i> <i>C. ephippium</i> ✓ <i>C. falcula</i> ✓ * * <i>C. gardineri</i> <i>C. leucopleura</i> <i>C. lineolatus</i> ✓ * * <i>C. mesoleucos</i> <i>C. nigropunctatus</i> * <i>C. oxycephalus</i> ✓ <i>C. rafflesi</i> ✓ <i>C. selene</i> ✓ <i>C. semeion</i> ✓ * <i>C. ulietensis</i> ✓ <i>C. vagabundus</i> ✓ * <i>C. xanthocephalus</i>	* <i>C. auriga</i> ✓ * * <i>C. decassatus</i> <i>C. ephippium</i> ✓ <i>C. falcula</i> ✓ * * <i>C. gardineri</i> <i>C. leucopleura</i> <i>C. lineolatus</i> ✓ * * <i>C. mesoleucos</i> <i>C. nigropunctatus</i> * <i>C. oxycephalus</i> ✓ <i>C. rafflesi</i> ✓ <i>C. selene</i> ✓ <i>C. semeion</i> ✓ * <i>C. ulietensis</i> ✓ <i>C. pictus</i> (<i>vagabundus</i>) ✓ * <i>C. xanthocephalus</i>
<i>Megaprotodon</i> subgenus Guichenot, 1848	<i>C. trifascialis</i> ✓	<i>C. trifascialis</i> ✓	<i>C. trifascialis</i> ✓	<i>C. trifascialis</i> ✓

Table 3 (continued)

<i>Chaetodontops</i> subgenus Bleeker, 1876	Not recognized (Blum, 1988 combined with <i>Rabdophorus</i> and called <i>Rabdophorus</i>)	<i>C. adiergastos</i> ✓ <i>C. auripes</i> <i>C. collare</i> ✓ <i>C. fasciatus</i> ✓ <i>C. flavirostris</i> <i>C. lunula</i> ✓ * * <i>C. semilarvatus</i> ✓ <i>C. wiebeli</i>	<i>C. adiergastos</i> ✓ <i>C. auripes</i> <i>C. collare</i> ✓ <i>C. fasciatus</i> ✓ <i>C. flavirostris</i> <i>C. lunula</i> ✓ * <i>C. reticulatus</i> ✓ <i>C. semilarvatus</i> ✓ <i>C. wiebeli</i>	<i>C. adiergastos</i> ✓ <i>C. auripes</i> <i>C. collare</i> ✓ <i>C. fasciatus</i> ✓ <i>C. flavirostris</i> <i>C. lunula</i> ✓ <i>C. nigropunctatus</i> <i>C. reticulatus</i> ✓ <i>C. semilarvatus</i> ✓ <i>C. wiebeli</i> ✓
<i>Gonochaetodon</i> subgenus Bleeker, 1876	<i>C. baronessa</i> ✓ <i>C. larvatus</i> ✓ <i>C. triangulum</i>	<i>C. baronessa</i> ✓ <i>C. larvatus</i> ✓ <i>C. triangulum</i>	<i>C. baronessa</i> ✓ <i>C. larvatus</i> ✓ <i>C. triangulum</i>	<i>C. baronessa</i> ✓ <i>C. larvatus</i> ✓ <i>C. triangulum</i>
<i>Discochaetodon</i> subgenus Nalbant, 1971	<i>C. aureofasciatus</i> <i>C. octofasciatus</i> ✓ <i>C. rainfordi</i> ✓ <i>C. tricinctus</i>	<i>C. aureofasciatus</i> <i>C. octofasciatus</i> ✓ <i>C. rainfordi</i> ✓ <i>C. tricinctus</i>	<i>C. aureofasciatus</i> <i>C. octofasciatus</i> ✓ <i>C. rainfordi</i> ✓ <i>C. tricinctus</i>	<i>C. aureofasciatus</i> <i>C. octofasciatus</i> ✓ <i>C. rainfordi</i> ✓ <i>C. tricinctus</i>
<i>Parachaetodon</i> genus Bleeker, 1974	as a subgenus <i>C. oligacanthus</i>	as a genus <i>P. ocellatus</i>	as a genus <i>P. ocellatus</i>	as a genus <i>P. ocellatus</i>
<i>Pseudochaetodon</i> Burgess, 1978	not recognized	as a genus <i>Johnrandallia</i>	not recognized	not recognized

Asterisks (*) denote taxa that are excluded from a particular hypothesis, which in some instances was prior to the discovery of that taxon. Taxa analyzed in the current study are denoted by ✓ after the name.

level, thus having the highly desirable effect of eliminating the proliferation of subgenera. Within clade 2, our phylogeny rejects the Blum (1988) *C. (Exornator)* subgenus as polyphyletic. Within clade 3, our phylogeny tentatively supports the monophyly of all six *Chaetodon* subgenera (*Corallochaetodon*, *Citharoedus*, *Gonochaetodon*, *Tetrachaetodon*, *Megaprotodon*, and *Discochaetodon*), but we question the utility of subgeneric groups each containing few members (some with 1–5 species each) when the group of 19 species forms a monophyletic group that might receive a single taxonomic designation. The current phylogeny also rejects the Kuitert (2002) membership of *C. (Rhombochaetodon)*, *C. (Exornator)*, *C. (Rabdophorus)*, and *C. (Chaetodontops)*, and the Allen et al. (1998) definition of *C. (Chaetodon)*, *C. (Rabdophorus)*, and *C. (Chaetodontops)*. We predict that the four main clades of *Chaetodon* will stand up to further taxonomic sampling, and that additional sequencing of species in the subgenus *C. (Roaops)* will result in unification of *C. (Roaops)* and *C. (Exornator)*.

Overall, we found that the many *Chaetodon* subgenera do not clarify taxonomy or phylogeny. To the contrary, the variable subgenera with alternative definitions have primarily made the discussion of butterflyfish taxonomy more difficult. We conclude that complete species sampling of the entire *Chaetodon* group, additional molecular data, and a synthesis of molecular and morphological approaches will be required before researchers should propose additional formal classification schemes at the genus and subgenus level for the butterflyfishes.

4.2. Evolutionary tempo and biogeography of the butterflyfishes

The butterflyfishes appear to have shared a common ancestor with the Scatophagidae and Pomacanthidae

about 54 million years ago, and the first divergence within the Chaetodontidae was reconstructed to have occurred approximately 37 mya (Fig. 6). Thus, the early diversification of the family Chaetodontidae occurred in the early Oligocene, during the increasing subdivision of the broad Tethys Sea that has been implicated in the origin and diversification of many reef fish clades (Bellwood and Wainwright, 2002). Most chaetodontids are distributed in the Indo-Pacific ocean basins, although a concentration of Eastern Pacific and Atlantic/Caribbean taxa provide some interesting biogeographic implications of the phylogeny. *Amphichaetodon* and *Forcipiger* have both Eastern Pacific and Indo-West Pacific distributions, with no members in the Atlantic. These two genera split off from their coralfish and bannerfish relatives between 20 and 30 mya (Fig. 6). If, from an Indo-West Pacific origin, dispersal was via the ancient Tethys Sea and across the Atlantic into the Eastern Pacific prior to the Terminal Tethyan event (closing of the Middle Eastern land bridge), it appears that *Amphichaetodon* and *Forcipiger* left no Atlantic descendants. Alternatively, dispersal could have been eastward across the East Pacific Barrier, a vast expanse of land-free ocean spanning 5000 km. The presence of such a barrier is both disputed and supported in various studies on a wide assortment of taxa (Briggs, 1995; Grigg and Hey, 1992; Lessios et al., 1998). Our data contain examples from different clades that may support several alternative hypotheses of dispersal routes. The broad distribution of *F. flavissimus* suggests that trans-Pacific dispersal should be considered for this group that includes the Eastern Pacific *Johnrandallia* and the widespread bannerfishes of the genus *Heniochus*. In contrast, the genus *Prognathodes* and subgenus *C. (Chaetodon)* have members in the Eastern Pacific, Indo-West Pacific and the Atlantic, which is consistent

with dispersal to the Atlantic either through the Panamanian isthmus or around the southern tip of Africa.

The primary biogeographic patterns of *Prognathodes* and *Chaetodon* are of several independent occurrences in the Atlantic basin dating to early divergence times, and a more recent pattern of species pairs split into Pacific and Indian Ocean inhabitants. *Prognathodes* and *Chaetodon* both have members in the Atlantic, and because *C. sedentarius* is now confirmed to be a member of the *Chaetodon* clade 2, there are three *Chaetodon* clades with species in the Atlantic. This is a similar pattern to the multiple lineages of labrid fishes that have Atlantic and Caribbean representatives (Westneat and Alfaro, 2005). The present sampling and distribution pattern (Fig. 1) raise the possibility that the main butterflyfish lineage originated in the Atlantic and subsequently diversified into the Indo-Pacific. Ideally, to address these questions of origin and dispersal, one would need a complete sampling of *Prognathodes* and subgenus *Chaetodon*, as the partial species sample in this basal region of the tree limits our ability to optimize biogeographic characters at the root of this main butterflyfish clade. However, the third example, that of *C. sedentarius*, is a much more recent invasion of the Atlantic, dated here at about 1.6 mya. This may be too recent to be a split due to closure of the Panamanian isthmus, and may be more likely to be attributed to dispersal around the southern tip of Africa from the Indian Ocean, which has been implicated in the dispersal of angelfishes (Bowen et al., 2006).

Molecular rates resulted in the estimate that *Prognathodes* and *Chaetodon* split approximately 24 mya and the subsequent divergences into the four main clades of *Chaetodon* occurred from 20 to about 14 mya (Fig. 6). This is a

period of diverse coral reef assemblages in the Tethys Sea (Bellwood and Wainwright, 2002), suggesting that the widespread ecological niche of consuming coral polyps by butterflyfishes (Pratchett, 2005) may have originated during this period. Interestingly, the three most species-rich clades of *Chaetodon* all originated about 14 mya, about the time of closure of the Tethys and final formation of separate Atlantic and Pacific basins between 12 and 18 mya. These origins considerably predate the earliest fossil evidence of *Chaetodon*, the Miocene butterflyfish *C. fischeuri* (Carnevale, 2006) of the lower Messinian, at approximately 7 mya. However, this fossil fits in nicely with the present phylogeny as it was tentatively placed by Carnevale (2006) as close to *Chaetodon kleinii*, forming part of the basal sister clade to the rest of *Chaetodon* clade 2 which diverged into currently living species during the last 10 million years (Fig. 6).

At finer levels of resolution within the genus *Chaetodon*, there are multiple species distribution splits between the Pacific and Indian Oceans (Fig. 1). For example, in *Chaetodon* clades 3 and 4 (Fig. 1), there are a total of seven sampled Indian Ocean basin species that have close relatives with disjunct distributions in the Pacific. Most of these divergences have occurred within the past 5 million years, and some of them (e.g., *C. falcula* + *C. ulietensis*) have apparently occurred within the past few hundred thousand years (Table 4). The Pliocene and Pleistocene were a turbulent time for shallow, reef associated animals, with extensive sea level changes resulting in the drying of reefs (Haq et al., 1987), and frequent cooling events that were responsible for faunal turnovers in many reef regions. Although the biogeographic provinces are different, the

Table 4

Lineage rate/time analysis for the butterflyfish phylogeny, showing the estimated ages for major clades and estimates of local substitution rates using penalized likelihood (r8s software, Sanderson, 2002)

Node	PL age (mya)	Local rate
Root age (fixed at 65.0)	65.00	
Chaetodontidae + Scatophagidae + Pomacanthidae	54.29	0.0047
Pomacanthidae	35.34	0.0028
Chaetodontidae	36.72	0.0032
“Bannerfishes” (<i>Amphichaetodon</i> to <i>Heniochus</i>)	33.49	0.0025
<i>Amphichaetodon</i> to <i>Chelmon</i>	27.80	0.0019
<i>Forcipiger</i> to <i>Heniochus</i>	24.19	0.0033
<i>Johnrandallia</i> + <i>Heniochus</i>	20.15	0.0027
<i>Heniochus</i>	12.80	0.0023
<i>Forcipiger</i>	10.90	0.0036
<i>Prognathodes</i> + <i>Chaetodon</i>	23.93	0.0036
<i>Chaetodon</i> (minimum age 7.0)	19.67	0.0036
<i>Chaetodon</i> clade 1 (<i>hoeferi</i> + <i>robustus</i>)	9.75	0.0037
<i>Chaetodon</i> clade 2 (<i>kleinii</i> to <i>guentheri</i>)	14.04	0.0035
<i>C. sedentarius</i> + <i>miliaris</i> + <i>guentheri</i> + <i>dolosus</i>	1.64	0.0051
<i>C. paucifasciatus</i> + <i>mertensii</i> + <i>xanthurus</i>	0.87	0.0028
<i>Chaetodon</i> clade 3 (<i>austriacus</i> to <i>zanzibariensis</i>)	14.51	0.0032
<i>C. ornatissimus</i> + <i>reticulatus</i> + <i>meyeri</i>	3.22	0.0033
<i>C. rainfordi</i> + <i>octofasciatus</i>	4.01	0.0039
<i>Chaetodon</i> clade 4 (<i>ocellatus</i> to <i>ulietensis</i>)	13.96	0.0039
<i>C. capistratus</i> + <i>striatus</i> + <i>ocellatus</i>	8.65	0.0040
<i>C. melannotus</i> + <i>ocellicaudus</i>	3.72	0.0031
<i>C. falcula</i> + <i>ulietensis</i>	0.26	0.0021

timing of divergences are similar to the pattern of rapid speciation in the parrotfish genus *Sparisoma* that took place 2.8–5.6 million years ago (Bernardi et al., 2000; Robertson et al., 2006), corresponding to periods of sealevel fluctuation and extinction of corals in the Caribbean 1–4 million years ago (Johnson et al., 1995).

An important emerging trend in coral reef fish biogeography is the relatively recent and fairly rapid pace of cladogenesis on branches that may extend back into the Oligocene or Eocene, resulting in the radiation of species richness in many reef-fish clades that we see in the oceans today. Robust, well-supported molecular phylogenies for coral reef fish lineages are the key to recent advances in the evolutionary biology and biogeography of marine fishes. Increased genetic and genomic data integrated with morphological information, and the ability to set confidence limits on historical divergence times of clades with fossil information and flexible molecular clock approaches will continue to clarify the fascinating biogeographic history of coral reef fishes.

Acknowledgments

Field work for tissue collection was aided by the Lizard Island Research Station, Mark McGrouther, Jeff Leis, Randy Mooi, Tom Trinski, Jeff Williams, Kent Carpenter, Luz Regis, Keith Lapuos, Jeff Janovetz, Brad Wright, Aaron Rice, Jim Cooper, and others on expeditions to the northern Great Barrier Reef, Vanuatu, Solomon Islands, and the Philippines. Thanks also to Overnight Fish and Apet, Inc. for assistance obtaining a few taxa that were otherwise difficult to acquire. DNA sequencing was aided by Lydia Smith. Thanks to Aaron N. Rice and William J. Cooper for helpful discussions and comments on the manuscript, and to Rob Vogelbacher at DePaul CTI for access to computational time on Illinois BioGrid. This research was funded by NSF Grant DEB 0235307 to M. Westneat.

References

- Allen, G.R., Steene, R., Allen, M., 1998. *A Guide to Angelfishes and Butterflyfishes*. Odyssey Publishing, Vanguard Press, Perth, WA.
- Bellwood, D.R., van Herwerden, L., Konow, N., 2004. Evolution and biogeography of marine angelfishes (Pisces: Pomacanthidae). *Mol. Phylogenet. Evol.* 33, 140–155.
- Bellwood, D.R., Wainwright, P.C., 2002. The history and biogeography of fishes on coral reefs. In: Sale, P.F. (Ed.), *Coral Reef Fishes: Dynamics and Diversity in a Complex Ecosystem*. Academic Press, San Diego, pp. 5–32.
- Bernardi, G., Robertson, D.R., Clifton, K.E., Azzurro, E., 2000. Molecular systematics, zoogeography, and evolutionary ecology of the Atlantic parrotfish genus *Sparisoma*. *Mol. Phylogenet. Evol.* 15, 292–300.
- Blum, S.D., 1988. *Osteology and Phylogeny of the Chaetodontidae* (Pisces: Perciformes). University of Hawaii, Honolulu.
- Blum, S.D., 1989. Biogeography of the Chaetodontidae: an analysis of allopatry among closely related species. *Environ. Biol. Fish.* 25, 9–31.
- Bowen, B.W., Muss, A., Rocha, L.A., Grant, W.S., 2006. Shallow mtDNA coalescence in Atlantic pygmy angelfishes (genus *Centropyge*) indicates a recent invasion from the Indian Ocean. *Journal of Heredity* 97, 1–12.
- Grigg, J.C., 1995. *Global Biogeography*. Elsevier, New York.
- Burgess, W.E., 1978. *Butterflyfishes of the World*. Tropical Fish Hobbyist, Neptune City, NJ.
- Carnevale, G., 2006. Morphology and biology of the Miocene butterflyfish *Chaetodon ficheri* (Teleostei: Chaetodontidae). *Zool. J. Linn. Soc.* 146, 251–267.
- Cox, E.F., 1994. Resource use by corallivorous butterflyfishes (Family Chaetodontidae) in Hawaii. *Bull. Mar. Sci.* 54, 535–545.
- Driscoll, J.W., Driscoll, J.L., 1988. Pair behavior and spacing in butterflyfishes (Chaetodontidae). *Environ. Biol. Fish.* 22, 29–37.
- Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39, 783–791.
- Ferry-Graham, L.A., Wainwright, P.C., Bellwood, D.R., 2001a. Prey capture in long-jawed butterflyfishes (Chaetodontidae): the functional basis of novel feeding habits. *J. Exp. Mar. Biol. Ecol.* 256, 167–184.
- Ferry-Graham, L.A., Wainwright, P.C., Hulsey, C.D., 2001b. Evolution and mechanics of long jaws in butterflyfishes (family Chaetodontidae). *J. Morphol.* 248, 120–143.
- Findley, J.S., Findley, M.T., 1989. Circumtropical patterns in butterflyfish communities. *Environ. Biol. Fish.* 25, 32–46.
- Findley, J.S., Findley, M.T., 2001. Global, regional, and local patterns in species richness and abundance of butterflyfishes. *Ecol. Monogr.* 71, 69–91.
- Fricke, H.W., 1986. Pair swimming and mutual partner guarding in monogamous butterflyfish (Pisces, Chaetodontidae): a joint advertisement for territory. *Ethology* 73, 307–333.
- Ghosh, M., Heo, J., 2003. Noninformative priors, credible sets and Bayesian hypothesis testing for the intraclass model. *J. Stat. Plann. Infer.* 112, 133–146.
- Grigg, R.W., Hey, R., 1992. Paleooceanography of the tropical eastern Pacific ocean. *Science* 255, 172–178.
- Haq, B.U., Hardenbol, J., Vail, P.R., 1987. Chronology of fluctuating sealevels since the Triassic. *Science* 235, 1156–1167.
- Hourigan, T.F., 1989. Environmental determinants of butterflyfish social systems. *Environ. Biol. Fish.* 25, 61–78.
- Huelsenbeck, J.P., Larget, B., Alfaro, M.E., 2004. Bayesian phylogenetic model selection using reversible jump Markov chain Monte Carlo. *Mol. Biol. Evol.* 6, 1123–1133.
- Huelsenbeck, J.P., Ronquist, F., 2001. MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* 17, 754–755.
- Hsu, K.C., Chen, J.P., Shao, K.T., 2007. Molecular phylogeny of *Chaetodon* (Teleostei: Chaetodontidae) in the Indo-West Pacific: evolution in geminate species pairs and species groups. *Raffles Bull. Zool. Suppl.* 14, 77–86.
- Johnson, K.G., Budd, A.F., Stemmann, T.A., 1995. Extinction selectivity and ecology of Neogene Caribbean corals. *Paleobiology* 21, 52–73.
- Kuiter, R.H., 2002. *Butterflyfishes, Bannerfishes and their Relatives: A Comprehensive Guide to Chaetodontidae and Microcanthidae*. TMC Publishing, Chorleywood, UK.
- Leis, J.M., 1989. Larval biology of butterflyfishes (Pisces, Chaetodontidae): What do we really know? *Environ. Biol. Fish.* 25, 87–100.
- Lessios, H.A., Kessing, B.D., Robertson, D.R., 1998. Massive gene flow across the world's most potent marine biogeographic barrier. *Proc. Biol. Sci.* 256, 583–588.
- Littlewood, D.T.J., McDonald, S.M., Gill, A.C., 2004. Molecular phylogenetics of *Chaetodon* and the Chaetodontidae (Teleostei: Perciformes) with reference to morphology. *Zootaxa* 779, 1–20.
- Maddison, W.P., 1991. Squared-change parsimony reconstructions of ancestral states for continuous-valued characters on a phylogenetic tree. *Syst. Zool.* 40, 304–314.
- McMillan, O.W., Palumbi, S.R., 1995. Concordant evolutionary patterns among Indo-West Pacific butterflyfishes. *Proc. R. Soc. Lond. B Biol. Sci.* 260, 229–236.
- McMillan, O.W., Palumbi, S.R., 1997. Rapid rate of control-region evolution in Pacific butterflyfishes (Chaetodontidae). *J. Mol. Evol.* 45, 473–484.

- McMillan, W.O., Weigt, L.A., Palumbi, S.R., 1999. Color pattern evolution, assortative mating, and genetic differentiation in brightly colored butterflyfishes (Chaetodontidae). *Evolution* 53, 247–260.
- Motta, P.J., 1985. Functional morphology of the head of Hawaiian and mid-Pacific butterflyfishes (Perciformes, Chaetodontidae). *Environ. Biol. Fish.* 13, 253–276.
- Motta, P.J., 1988. Functional morphology of the feeding apparatus of ten species of Pacific butterflyfishes (Perciformes, Chaetodontidae): an ecomorphological approach. *Environ. Biol. Fish.* 22, 39–67.
- Motta, P.J., 1989. Dentition patterns among Pacific and Western Atlantic butterflyfishes (Perciformes, Chaetodontidae): relationships to feeding ecology and evolutionary history. *Environ. Biol. Fish.* 25, 159–170.
- Oettinger, M.A., Schatz, D.G., Gorka, C., Baltimore, D., 1990. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. *Science* 248, 1517–1523.
- Pitts, P.A., 1991. Comparative use of food and space by three Bahamian butterflyfishes. *Bull. Mar. Sci.* 48, 749–756.
- Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics* 14, 817–818.
- Pratchett, M.S., 2005. Dietary overlap among coral-feeding butterflyfishes (Chaetodontidae) at Lizard Island, northern Great Barrier Reef. *Mar. Biol.* 148, 373–382.
- Rambaut, A., 1996. Se-AI: Sequence Alignment Editor. Freeware available at <ftp://evolve.zo.ox.ac.uk/packages/Se-AI>.
- Righton, D., Kemp, J., Ormond, R., 1996. Biogeography community structure and diversity of Red Sea and western Indian Ocean butterflyfishes. *J. Mar. Biol. Assoc. UK* 76, 223–228.
- Robertson, D.R., Karg, F., de Moura, R.L., Victor, B.C., Bernardi, G., 2006. Mechanisms of faunal enrichment in Atlantic parrotfishes. *Mol. Phylogenet. Evol.* 40, 795–807.
- Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574.
- Sanderson, M.J., 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. *Mol. Biol. Evol.* 14, 1218–1231.
- Sanderson, M.J., 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Mol. Biol. Evol.* 19, 101–109.
- Sazima, C., Sazima, I., 2001. Plankton-feeding aggregation and occasional cleaning by adult butterflyfish, *Chaetodon striatus* (Chaetodontidae), in Southwestern Atlantic. *Cybius* 25, 145–151.
- Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* 16, 1114–1116.
- Smith, W.L., Webb, J.F., Blum, S.D., 2003. The evolution of the laterophysic connection with a revised phylogeny and taxonomy of butterflyfishes (Teleostei: Chaetodontidae). *Cladistics* 19, 287–306.
- Sparks, J.S., Smith, W.L., 2004. Phylogeny and biogeography of the Malagasy and Australasian rainbowfishes (Teleostei: Melanotaeniidae): Gondwanan vicariance and evolution in freshwater. *Mol. Phylogenet. Evol.* 33, 719–734.
- Streelman, T.J., Karl, S.A., 1997. Reconstructing labroid evolution with single-copy nuclear DNA. *Proc. R. Soc. Lond. B* 264, 1011–1020.
- Swofford, D.L., 2000. PAUP* 4.00: Phylogenetic Analysis Using Parsimony (*and other methods). Sinauer Associates, MA.
- Tricas, T.C., Hiramoto, J.T., 1989. Sexual differentiation, gonad development, and spawning seasonality of the Hawaiian butterflyfish *Chaetodon multicinctus*. *Environ. Biol. Fish.* 25, 111–124.
- Westneat, M.W., Alfaro, M.E., 2005. Phylogenetic relationships and evolutionary history of the reef fish family Labridae. *Mol. Phylogenet. Evol.* 36, 370–390.
- Yabuta, S., 1999. Behavioral rules and tail-up display in extra- and intra-pair interactions of the butterflyfish, *Chaetodon lunulatus*. *J. Ethol.* 17, 79–86.
- Yabuta, S., 2002. Uncertainty in partner recognition and the tail-up display in a monogamous butterflyfish. *Anim. Behav.* 63, 165–173.
- Zekeria, Z.A., Dawit, Y., Ghebremedhin, S., 2002. Resource partitioning among four butterflyfish species in the Red Sea. *Mar. Freshwat. Res.* 53, 163–168.

SYNTHESIS



The historical biogeography of coral reef fishes: global patterns of origination and dispersal

Peter F. Cowman* and David R. Bellwood

School of Marine and Tropical Biology, and
Australian Research Council Centre of
Excellence for Coral Reef Studies, James Cook
University, Townsville, QLD, 4811, Australia

ABSTRACT

Aim To use recently published phylogenies of three major reef fish families to explore global patterns of species origin and dispersal over the past 65 million years. The key questions are: when and where did reef fishes arise, and how has this shaped current biodiversity patterns?

Location Biogeographic reconstructions were performed on globally distributed reef fish lineages. Patterns of lineage origination and dispersal were explored in five major biogeographic regions: the East Pacific, the Atlantic, the Indian Ocean, the Indo-Australian Archipelago hotspot, and the Central Pacific.

Methods A dispersal, extinction and cladogenesis (DEC) model implemented in LAGRANGE was used to infer the most likely biogeographic scenarios at nodes on chronograms of three diverse reef fish families (Labridae, Pomacentridae, Chaetodontidae). For the terminal branches ANOVA was used to compare patterns of origination on a regional and global scale. Patterns of origination and dispersal were examined within discrete time periods for the five biogeographic regions.

Results Temporal examination of hypothetical ancestral lineages reveal a pattern of increasing isolation of the East Pacific and Atlantic regions from the Eocene, and the changing role of the Indo-Australian Archipelago from a location of accumulating ranges in the Palaeo/Eocene, a centre of origination in the Miocene, to extensive expansion of lineages into adjacent regions from the Pliocene to Recent.

Main conclusions While the East Pacific and Atlantic have a history of isolation, the Indo-Australian Archipelago has a history of connectivity. It has sequentially and then simultaneously acted as a centre of accumulation (Palaeocene/Eocene onwards), survival (Eocene/Oligocene onwards), origin (Miocene onwards), and export (Pliocene/Recent) for reef fishes. The model suggests that it was the proliferation and expansion of lineages in the Indo-Australian Archipelago (the Coral Triangle) during the Miocene that underpinned the current biodiversity in the Indian and Pacific Oceans.

Keywords

Biodiversity hotspot, centre of origin, coral reef, Coral Triangle, DEC model, dispersal, evolution, Indo-Australian Archipelago, Lagrange, survival.

*Correspondence: Peter F. Cowman, School of Marine and Tropical Biology, James Cook University, Townsville, QLD 4811, Australia.
E-mail: peter.cowman@my.jcu.edu.au

INTRODUCTION

Understanding the historical forces shaping biodiversity is an important aspect of marine biogeography. More than 5000 species of fishes are associated with coral reefs globally (Bell-

wood *et al.*, 2010). A maximum diversity is reached on reefs within the central Indo-Australian Archipelago (IAA), the largest marine biodiversity hotspot. This ‘bull’s eye’ pattern has both latitudinal and longitudinal gradients in diversity. The latitudinal gradients in species diversity away from the

IAA hotspot have been recognized for some time (Ekman, 1935; Rosen, 1981), and reflect the temperate–tropical gradients seen in terrestrial systems. However, the longitudinal decline in marine species diversity has inspired much debate in the literature (Palumbi, 1997; Briggs, 1999; Hoeksema, 2007). Central to the debate over the last 30 years have been three cornerstone ‘centre of’ hypotheses describing the origin and maintenance of faunal diversity within the IAA hotspot: whether it is a centre of origin, overlap, and/or accumulation. As in terrestrial biogeography, rates of origination have been used as the primary basis for evaluating these various hypotheses. These hypotheses examine processes maintaining biodiversity in the IAA hotspot and are framed within the context of the Indo-Pacific region. Yet, the Indo-Pacific does not exist in isolation. There has been little attempt to explore global patterns of origination and dispersal between regions to directly compare the major marine biogeographic provinces. Particularly, how do rates of origination within the IAA compare with those of adjacent regions in the Indo-Pacific (Indian Ocean, Central Pacific) and other more distant regions (Atlantic, East Pacific)? There are challenges to addressing these questions. The lack of physical barriers in the marine environment and connectivity through larval dispersal has blurred the biogeographic origins of marine species (Bellwood & Meyer, 2009a). Even on shallow time-scales it is often unclear whether speciation has occurred in allopatry through vicariance, or in sympatry (Barracough & Vogler, 2000; Losos & Glor, 2003; Rocha *et al.*, 2005; Quenouille *et al.*, 2011). However, large-scale patterns are becoming increasingly open to quantitative evaluation (e.g. Hughes *et al.*, 2002; Floeter *et al.*, 2008; Rocha & Bowen, 2008; Williams & Duda, 2008). Nevertheless, a detailed appraisal of global patterns in origination and dispersal among the major biogeographic regions for coral reef fishes remains to be seen.

Since their initial formulation by Potts (1985) in the context of reef building corals, the three ‘centre of’ hypotheses have been repeatedly modified and expanded (reviewed by Bellwood *et al.*, 2012) in an attempt to explain the extensive overlapping and widespread ranges seen in multiple coral reef taxa (Hughes *et al.*, 2002; Connolly *et al.*, 2003). Furthermore, a fourth ‘centre of survival’ (Heck & McCoy, 1978; Barber & Bellwood, 2005) has been added which seeks to explain why most of the taxa remain in the IAA regardless of the location of origin of the taxa. This hypothesis allows multiple sources of biodiversity, widespread ranges and post-speciation range expansion. Recent study has shown that coral reefs may provide the mechanism for this survival in the IAA, allowing both higher rates of diversification and reduced vulnerability to extinction for associated lineages (Cowman & Bellwood, 2011). Several phylogeographic studies of reef fish evolution centred in the IAA have invoked one or more of the ‘centre of’ hypotheses to explain current biogeographic patterns (Bernardi *et al.*, 2002; McCafferty *et al.*, 2002; Timm *et al.*, 2008). However, there has been no family-level study for reef fishes exploring congruence in global patterns of biogeographic origination, dispersal and geographic range inheritance in a temporal framework. Central questions regarding rates and

locations of origination, direction and amplitude of dispersal over time, and the extent of range inheritance along lineages need to be answered before we can begin to understand the drivers of coral reef biodiversity. To answer these questions a global perspective is needed, for multiple groups, across all of the major marine biogeographic realms.

The circum-tropical belt can be divided into three major realms: the Indo-Pacific, Atlantic, and East Pacific. These realms are distinguished by a taxonomic makeup resulting from a long history of vicariance associated with both hard and soft barrier formation dating as far back as the Eocene (Bellwood & Wainwright, 2002; Floeter *et al.*, 2008). The formation of three separate barriers has acted to divide these regions: (1) the East Pacific Barrier (EPB), an open expansion of ocean which separates the Indo-Pacific from the East Pacific; (2) the Terminal Tethian event (TTE), which largely cuts off dispersal between the Indo-Pacific and the Atlantic (but see Bowen *et al.*, 2006); and (3) the Isthmus of Panama (IOP), which separates the Atlantic from the East Pacific (reviewed by Lessios, 2008). Within the Indo-Pacific, the lack of hard barriers (land bridges) has allowed many taxa to maintain widespread geographic ranges spanning from the east coast of Africa to islands in the central Pacific, or in some cases to the Pacific coast of the Americas (Hughes *et al.*, 2002; Reece *et al.*, 2011). However, the combination of tectonic activity and several semi-permeable hydrological barriers (Barber *et al.*, 2000, 2002; Santini & Winterbottom, 2002) has resulted in a complex distribution of taxa including both provincial endemics and widespread species that characterize the Indian Ocean, the IAA hotspot, and the Central Pacific island arcs (Bellwood & Wainwright, 2002; Jones *et al.*, 2002; Connolly *et al.*, 2003; Hoeksema, 2007). These conditions make it extremely difficult to identify origination and directionality of dispersal between regions. For both endemic taxa and those that are widespread across the entire Indo-Pacific (Indian Ocean, IAA, Central Pacific) we need to answer several questions: (1) In what region did a lineage first arise? (2) To what extent has within-region and between-region origination contributed to patterns of diversity? (3) How has dispersal affected their distribution? In a temporal framework, the answers to these questions will illuminate both the evolutionary history of the IAA hotspot, and its relationship with adjacent and distant geographic regions.

Within the IAA, it may be difficult to detect the overall contribution of vicariant speciation through isolation due to the transient nature of barriers and rapid dispersal potential of marine fishes (Quenouille *et al.*, 2011). However, on a larger taxonomic and geographic scale, vicariance between biogeographic regions may be clearly observed in deeper lineages (Losos & Glor, 2003). Likewise, at a regional scale it may be beneficial to quantify the extent of lineage origination within a biogeographic region (regional origination) versus between regions. In considering dispersal between regions, it is important to quantify the frequency with which dispersal is followed by vicariance and whether it reflects the palaeogeographic record.

To address these issues, we examine inferred patterns of origination and dispersal in the reef fish families Labridae, Pomacentridae and Chaetodontidae under a dispersal, extinction and cladogenesis (DEC) model. All three families are amongst the most diverse and abundant families distributed on coral reefs globally (Bellwood & Wainwright, 2002). They also inhabit other tropical reef habitats as well as more temperate latitudes within the five biogeographic regions. Chronograms for these families have recently been reconstructed that include the highest taxon sampling to date (45% labrids, 46% pomacentrids, 75% chaetodontids), with relatively even representation in all regions and habitats (Cowman & Bellwood, 2011; see Table S1, Appendix S1 in the Supporting Information). These chronologies show remarkable congruence in the tempo of diversification (Cowman & Bellwood, 2011), with previous work highlighting temporal concordance in trophic innovation (Cowman *et al.*, 2009; Bellwood *et al.*, 2010). The three families contain endemic species in most major regions, as well as widespread species (Moura & Sazima, 2000; Jones *et al.*, 2002; Kuiter, 2002; Floeter *et al.*, 2008; Craig *et al.*, 2010). Previous studies have explored biogeographic patterns of various taxa within each family (Floeter *et al.*, 2001, 2008; McCafferty *et al.*, 2002; Barber & Bellwood, 2005; Rocha *et al.*, 2005; Westneat & Alfaro, 2005; Beldade *et al.*, 2009; Hodge *et al.*, 2012). However, there has been no biogeographic reconstruction of ancestral ages at the family level. These three families, therefore, present ideal candidates for ancestral range reconstruction and the exploration of patterns of origination on coral reefs. Using recently developed software for biogeographic reconstruction (Ree & Smith, 2008) hypothetical inheritance scenarios along the molecular lineages can be inferred from extant ranges. This can provide us with a framework in which we can assess patterns of origination within each region, and dispersal between them.

The aim of this study, therefore, is to examine patterns of origination and dispersal in the biogeographic histories of the Labridae, Pomacentridae and Chaetodontidae. In a global context this will provide an opportunity to evaluate possible sources of current biodiversity, directionality of dispersal, and the relationship between the major biogeographic realms over the past 65 Myr. The specific questions to be answered are:

1. Do the three focal families of coral reef fishes display congruent patterns of origination on a global scale?
2. How have post-speciation dispersal and range inheritance contributed to current patterns of biodiversity of coral reef fishes?
3. Has the role of the IAA hotspot changed through time?

MATERIALS AND METHODS

Chronologies and geographic range data

Recently constructed chronograms for the families Labridae (wrasses), Pomacentridae (damselfishes) and Chaetodontidae

(butterflyfishes) were used in the ancestral range reconstruction analysis. These chronograms were obtained using Bayesian inference and fossil data (see Cowman & Bellwood, 2011). The chronograms focused on relative ages within the three families. Given the uncertainty in stem lineage ages some older origins may be possible for some taxa, especially the Chaetodontidae (Santini *et al.*, 2009). The geographic ranges of all nominal species in each family were assessed using published sources (Randall *et al.*, 1990; Allen, 1991; Allen *et al.*, 1998; Kuiter, 2002; Randall, 2005) and FishBase (Froese & Pauly, 2011). Geographic ranges were allocated into five separate regions: (1) Indian Ocean, (2) IAA, (3) Central Pacific, (4) East Pacific, and (5) Atlantic (Table S1, Appendix S1; areas incorporated in each region provided in Table S2, Appendix S1). Presence within a geographic region required a record of one location within the region; there was no limit to the number or order of regions occupied. Temperate lineages are included in each region for each of the three families to maintain taxonomic completeness of clades where possible. For taxa sampled in the chronograms, the presence or absence of a species in each region was coded as a character state to be used in the ancestral range reconstruction (Table S3, Appendix S1). Although not complete, all major lineages and tribes contain high and relatively even taxon sampling across all regions (Table S1, Appendix S1).

Ancestral range reconstruction

Reconstruction of ancestral ranges based on the time-calibrated phylogenies was implemented in the program LAGRANGE 2.01 (likelihood analysis of geographic range evolution; Ree & Smith, 2008). LAGRANGE implements a maximum likelihood approach based on a stochastic model of geographic range evolution involving dispersal, extinction and cladogenesis (DEC model). The DEC model implemented by the LAGRANGE program has been used in several recent studies of various taxa: plants (Smith, 2009), snakes (Kelly *et al.*, 2009), insects (Ramírez *et al.*, 2010), and gastropods (Göbbeler & Klussmann-Kolb, 2010). In marine fishes, it has been used to investigate biogeographic patterns in the smelt *Hypomesus* (Ilves & Taylor, 2007). Buerki *et al.* (2011) show, using palaeogeographic evidence, that it can lead to more accurate reconstructions than previous approaches, provided care is taken with defining the initial model.

In the DEC model, anagenetic (internode) range evolution is governed by a Q matrix of instantaneous transition rates that infer dispersal (range expansion) between geographic regions, or local extinction (range contraction) within a region along phylogenetic branches that can be calibrated to time (Ree & Smith, 2008). The model also allows global rates of dispersal and local extinction to be estimated by maximum likelihood (Ree & Smith, 2008). Dispersal was restricted to only occur between adjacent regions.

Cladogenetic (node speciation) change is modelled under three alternative inheritance scenarios (Ree *et al.*, 2005): (1)

vicariance, where a widespread ancestor diverges across a regional boundary with descendants present in adjacent regions (Fig. S1a, Appendix S2); (2) within-region origination, where the ancestor and both descendants are present in the same region (Fig. S1b, Appendix S2); and (3) peripheral cladogenesis, where an ancestral lineage maintains a range across more than one of the defined regions, while one descendant originates within one of the regions (Fig. S1c, Appendix S2). The third range inheritance scenario allows a widespread ancestral range to be inherited by a single descendant lineage. It is this scenario that will be most useful in modelling range evolution within these fish families, which is not implemented in traditional range reconstruction software such as DIVA (Ronquist, 1997). For each node, range inheritance scenarios are ranked according to the fractional likelihood they received by the analysis. Where other likely scenarios exist for range inheritance, LAGRANGE ranks these scenarios within two log-likelihood units of the optimal scenario.

LAGRANGE also allows constraints to be placed on the DEC model to reflect past geological events (e.g. formation of barriers) and fossil information. Each family was exposed to three separated models: M0, an unconstrained model allowing equal probability of dispersal between adjacent areas at any time; M1, a constrained model reflecting formation of biogeographic barriers; and M2, a constrained model with an added fossil constraint at the root of each tree reflecting the fossil record for each family (Fig. S2, Appendix S2). The constrained model, M1, reduced the probability of dispersal from the Central Pacific to the East Pacific to 0.05 for the entire duration of the chronogram (i.e. from root to tip) reflecting the East Pacific Barrier. The probability of dispersal from the Indian Ocean to the Atlantic Ocean was reduced to 0.05 from 18 Ma onwards, reflecting the closure of the Tethys seaway (Steininger & Rögl, 1979), but allowing the possibility of dispersal around the Horn of Africa (Bowen *et al.*, 2006). Dispersal from the Atlantic to the East Pacific was not allowed from 3.1 Ma to present, reflecting the closure of the Isthmus of Panama (Lessios, 2008).

The fossil constrained model, M2, forced the range of the root node to include the Atlantic (Tethys) region, based on fossils from Monte Bolca (Bellwood, 1996; Fig. S2, Appendix S2). The three models were compared using log-likelihood scores, and range inheritance scenarios were compared across the tree. The best model was used to map origination, dispersal and extinction events on to the family chronograms.

Exploring patterns

Patterns of origination and dispersal among the Labridae, Pomacentridae and Chaetodontidae were assessed based on the DEC model with the best likelihood score. By mapping the range inheritance scenarios reconstructed by the DEC model, the estimated timing and location of origination was recorded for each of the lineages on each of the chronologies. Inferred dispersal events were recorded along each

branch. The timing for each dispersal event was taken as the midway point along the branch on which it occurred. This procedure was employed for the tips of the trees (i.e. the extant taxa) and across the entire chronology (i.e. hypothetical ancestral molecular lineages).

Patterns of origination were compared as a proportion of 'sampled regional diversity' (extant lineages sampled in the trees present in each region) and as a proportion of 'sampled family diversity' (extant lineages sampled in the trees) for each family. This was to identify regions where origination of lineages through regional cladogenesis has been a major contribution to diversity of the region, and to identify which regions have been a major contributor to the diversity of each family. To investigate if taxon sampling has had an effect on the patterns observed we also compare regional origination as a proportion of 'total regional diversity' (all nominal species in the region for each family) and 'total family diversity' (all nominal species in the family) for each family. All proportional values were arcsin transformed prior to analyses.

In order to investigate the pattern of origination and dispersal throughout the evolutionary history of the three families, the number of lineages present in each region at four time intervals were calculated from the ancestral reconstruction: the Eocene/Oligocene boundary (33 Ma), the Oligocene/Miocene boundary (25 Ma), the Miocene/Pliocene boundary (5 Ma), and the present (0). This gives an estimate of hypothetical relative 'palaeodiversity' from the molecular history (N.B. this is only relative and makes no allowances for subsequent extinction). The palaeodiversity in each region is the result of lineage origination within the region, and lineage dispersal into the region during the previous epoch. In cases where vicariance has occurred after dispersal, the sum of origination and dispersal into the region will be greater than the total palaeodiversity. Where a lineage has inherited a range, origination and dispersal will be less than palaeodiversity. Proportions of origination, dispersal, and inheritance were estimated in each family for each epoch (Palaeocene/Eocene, Oligocene, Miocene, Pliocene/Recent).

RESULTS

Range reconstruction

The constrained models, M1 (barrier constrained) and M2 (fossil constrained), were favoured over the unconstrained model M0, with significantly better log-likelihood scores (over 30 log-likelihood units) for all three families (Table S4, Appendix S1). In all cases M1 had a better likelihood score than M2; however, this was negligible for the Labridae (0.3 units) and Chaetodontidae (1 unit). For the Pomacentridae, the addition of the fossil calibration gave a lower log-likelihood score by 2.7 units. Optimal reconstructed scenarios showed consistently higher likelihood under M1. Global rates of dispersal were slightly higher under constrained models while extinction rates were similar under all three

models (Table S4, Appendix 1). Node reconstructions of model M1 are therefore used to explore patterns of origination and dispersal. Node reconstructions were overlaid onto family chronograms (Figs S3–S5, Appendix S2).

Origination and dispersal of extant taxa

For sampled regional diversity of each family, the relative contribution of origination, dispersal and range inheritance to regional biodiversity was calculated from the LAGRANGE M1 reconstruction (Table 1). For all three families, the East Pacific, Atlantic and IAA show high rates of within-region origination (75–100%; Table 1). Dispersal into these regions is low (1–6%). The Atlantic region appears to be isolated except for two putative dispersal events into the region around the Cape of Good Hope: *Halichoeres maculipinna*; *Anampses caeruleopunctatus* (Fig. S3, Appendix S2). Both the Indian Ocean and Central Pacific are characterized by low origination (20–40% Indian Ocean; 10–16% Central Pacific), but high dispersal of lineages into the region, from the IAA (Fig. 1a, Table 1). Dispersal from the IAA into the Indian Ocean and the Central Pacific appears to be of a similar magnitude (Fig. 1a). The Central Pacific, while having comparable species richness to the Indian Ocean (Fig. 1a), has lower origination (7–16%). For all three families ancestral inheritance in the Central Pacific appears to be higher than within-region origination (Table 1).

In order to investigate whether patterns of origination (both as a proportion of sampled regional diversity and sampled family diversity) varied among regions and families, a two-way ANOVA was used, with regions and families treated as fixed effects. There was no significant family effect on regional origination as a proportion of sampled regional diversity ($F_{2,8} = 0.29$, $P = 0.76$) i.e. patterns of origination within each region were consistent among families. However, there was a significant region effect on origination as a proportion of sampled diversity in each region ($F_{4,8} = 32.7$, $P < 0.0001$), i.e. there were significant regional differences in the contribution of within-region origination to sampled regional diversity. Tukey–Kramer post-hoc comparisons

showed that lineage originations in the IAA, Atlantic, and East Pacific were significantly higher than in the Indian Ocean and Central Pacific (Fig. 1b).

Regional origination as a proportion of sampled family diversity likewise showed no significant family effect ($F_{2,8} = 0.001$, $P = 0.99$). There was, again a significant region effect on regional origination as a proportion of sampled family diversity ($F_{2,8} = 56.7$, $P < 0.0001$). Tukey–Kramer post-hoc comparisons identified the IAA as having a significantly higher proportion of origination contributing to the sampled diversity of each of the three families, compared with the other four regions (Fig. 1c).

Regional origination as a proportion of total regional diversity and total family diversity showed similar results. There was no significant family effect found for regional origination as a proportion of total regional diversity ($F_{2,8} = 3.2$, $P = 0.09$) or total family diversity ($F_{2,8} = 2.4$, $P = 0.15$) i.e. patterns of regional origination both as a proportion of total regional diversity and total family diversity did not differ significantly among families. As with sampled diversity, there was a significant region effect on the proportion of regional origination of both total regional diversity ($F_{2,8} = 10.22$, $P < 0.005$), and total family diversity ($F_{2,8} = 20.3$, $P < 0.001$), i.e. the contribution of regional origination to both the total biodiversity of a region, and the total biodiversity of each family, differed significantly between regions. Tukey–Kramer post-hoc comparisons identified the same groups as reported for sampled regional diversity (EP, Atl, IAA versus In, CP) and for sampled family diversity (IAA versus all other regions)

Origination and dispersal through time

The ancestral reconstruction implemented in LAGRANGE allowed patterns of anagenetic and cladogenetic change to be investigated in different epochs. As with the extant tips of the trees, this allowed the contribution of origination and dispersal to the palaeodiversity (of molecular lineages) in each biogeographic region to be estimated for the Labridae (Fig. 2a–d), Pomacentridae (Fig. 2e–h), and the Chae-

Table 1 Relative contribution of origination (Origin), dispersal (Disp.) and inheritance (Inh.) per capita regional biodiversity of the Labridae, Pomacentridae, and Chaetodontidae in five major biogeographic regions. Origination is the proportion of species that arose within that region; Dispersal is the proportion of species within each region that dispersed to that region and maintain it as part of its range; Inheritance is the proportion of species that inherited that region as a part of an ancestral lineage range that arose in another region.

Region	Labridae			Pomacentridae			Chaetodontidae		
	Origin	Disp.	Inh.	Origin	Disp.	Inh.	Origin	Disp.	Inh.
EP	0.76	0.06	0.18	0.90	0.00	0.10	0.67	0.00	0.33
Atl.	0.99	0.01	0.00	1.00	0.00	0.00	1.00	0.00	0.00
In	0.21	0.64	0.15	0.41	0.49	0.10	0.40	0.47	0.14
IAA	0.96	0.03	0.01	0.88	0.11	0.01	0.95	0.05	0.00
CP	0.10	0.78	0.12	0.07	0.72	0.21	0.16	0.53	0.30

EP, East Pacific; Atl, Atlantic; In, Indian; IAA, Indo-Australian Archipelago; CP, Central Pacific.

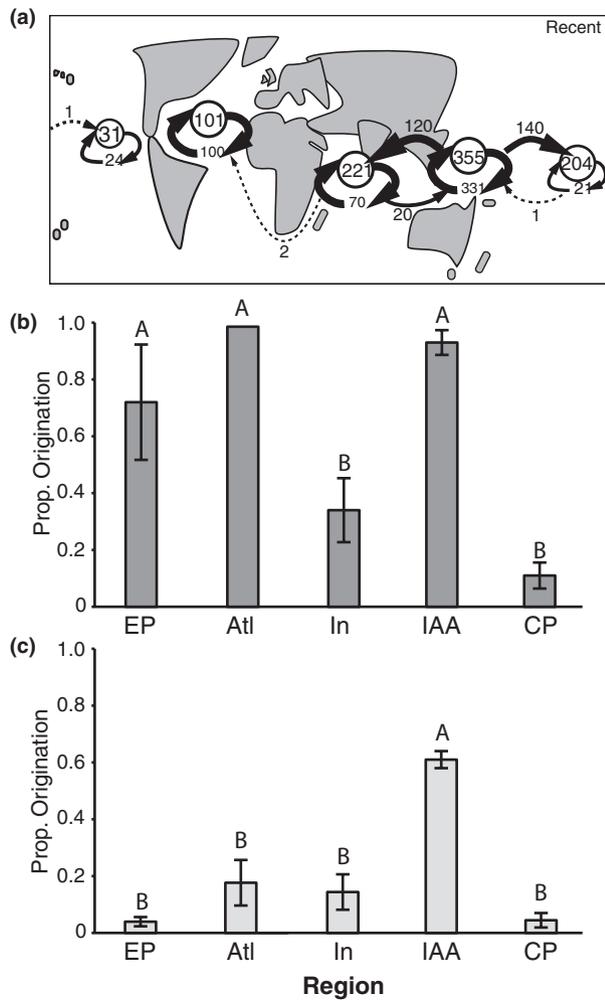


Figure 1 (a) Schematic drawing of the globe displaying regional patterns of combined diversity, origination and dispersal of extant labrids, pomacentrids and chaetodontids (the tips of the trees) in each region. EP = East Pacific, Atl = Atlantic, In = Indian, IAA = Indo-Australian Archipelago, CP = Central Pacific. Numbers in circles show extant diversity of sampled fishes in each region. Numbers below circles indicate within-region originations. Numbers next to arrows between adjacent regions indicate inferred dispersal events. Line thickness of arrows reflects the number of lineage originations in each region, and number of dispersal events between regions based on the reconstruction. For example, of the 355 lineages in the IAA, 331 arose within the region; 21 dispersed in and 260 out (see Appendix S2 for family reconstructions). (b) Proportion (\pm SE) of lineage originations in each region (regional origination/sampled regional richness; $n = 3$ families). (c) Proportion of lineage origination (\pm SE) in each region (regional origination/sampled family diversity; $n = 3$ families). A and B above bars mark non-significant groups in Tukey's post-hoc comparisons ($P < 0.001$).

odontidae (Fig. 3). Within the Palaeo/Eocene, Oligocene, Miocene, and Pliocene/Recent epochs the reconstructions yielded congruent patterns among the three families. These epochs will be considered separately below.

Palaeo/Eocene (65–33 Ma)

Both the Labridae and Pomacentridae have highest palaeodiversity in the IAA with relatively low dispersal of lineages from the area (Fig. 2a,e). The Chaetodontidae have yet to diverge at this stage, but the reconstruction places the origin of the family in the IAA (Fig. 3a; Fig. S5, Appendix S2). Dispersal is potentially possible between all five regions but the number of lineages that potentially dispersed between regions is low. The East Pacific, Atlantic and IAA regions contain lineages that have origins within the region, and lineages that have dispersed from adjacent regions. However, within the Indo-Pacific, the Indian Ocean and Central Pacific contain only lineages that have dispersed in to the regions. The panmixia inferred from the reconstruction suggests that lineages of the hypsigenyines and scarines may have maintained widespread ranges. By 33 Ma we have highest lineage diversity in the IAA, of which the majority remain restricted there (Fig. 2a,e). The most recent common ancestor (MRCA) of widespread lineages overlap with lineages restricted to the Atlantic (Labrini), and the IAA (Cheilini, Julidini, Odacini, *Bodianus*, *Choerodon*; Figs S3 & S4, Appendix S2). For the Pomacentridae, the root is placed within the IAA, but by the end of the Eocene three major lineages representing precursors to the Stegastinae, Chrominae and Abudedefdufinae are widespread, overlapping with the basal lineages of the Pomacentrinae that are restricted to the IAA (Fig. S4, Appendix S2). Overall, there are several widespread lineages that have resulted from initial connectivity between regions from the beginning of the Palaeocene, and dispersal throughout the Eocene. However, the accumulation of widespread lineages with those restricted to the IAA marked the region as an early centre for palaeodiversity of molecular lineages (Fig. 2a,e).

Oligocene (33–23 Ma)

The Oligocene sees an increase in diversity within the IAA, but again dispersal out of the region is limited (Figs 2b,f & 3b). Cladogenesis appears to be suppressed in all regions except the in the IAA. For both the Labridae and the Pomacentridae, further diversification of lineages restricted to the IAA more than doubles the inferred palaeodiversity from the previous epoch (Fig. 2b,f). By comparison, the adjacent regions in the Indo-Pacific (Indian Ocean, Central Pacific) show little or no within-region origination. Similarly, the East Pacific and Atlantic show limited support for diversification of lineages during this epoch. The Atlantic region becomes increasingly isolated from the Indo-Pacific during this period. There is no connection maintained either side of the Atlantic region through lineage inheritance, or through dispersal of labrid lineages (Fig. 2b). For the duration of the Oligocene, the Atlantic is severed from the Indo-Pacific, with the exception of a single dispersal event of a damselfish lineage (*Chromis*) from the Atlantic to the Indian Ocean, and a

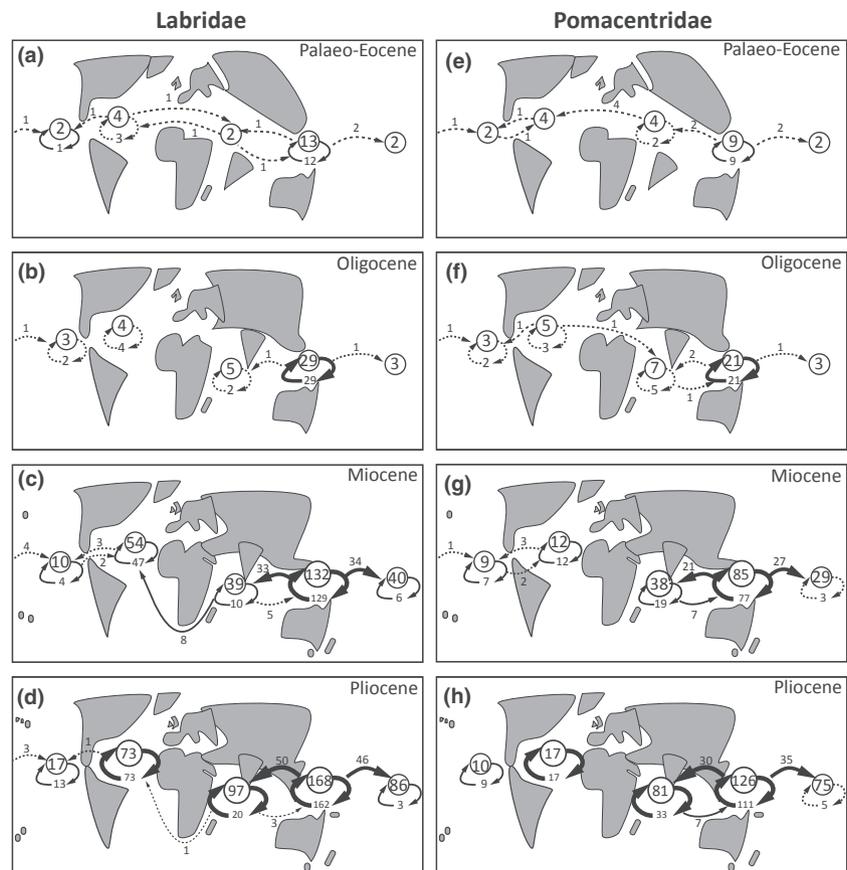


Figure 2 Schematic drawing of global palaeomaps for four time periods: Palaeo/Eocene (65–33 Ma); Oligocene (33–23 Ma); Miocene (23–5 Ma); Pliocene/Recent (5–0 Ma). Numbers show total hypothetical palaeodiversity (in circle), origination (below circle), and dispersal events (next to arrow) for each region in each period for the Labridae (a–d) and the Pomacentridae (e–h). Where dispersal in + origination is less than regional richness, the difference is range inheritance from the previous epoch. Where it is greater than regional richness, the difference is origination through vicariance.

continuous range of the *Abudefduf* lineage between these two regions (Fig. 2f; Fig. S4, Appendix S2). Vicariance events split lineages between the Atlantic and the Indian Ocean in the Labridae (*Calotomus/Sparisoma* clade) and the Pomacentridae (Chrominae, Stegastinae) (Fig. 2b,f; Figs S3 & S4, Appendix S2). In contrast, the Chaetodontidae reconstruction shows the initial divergence between the two major lineages; the bannerfish lineage restricted to the IAA, while the crown *Chaetodon* lineage disperses out to the Central Pacific, and through the Indian Ocean into the Atlantic, maintaining this widespread range into the Miocene (Fig. 3b).

Miocene (23–5 Ma)

In all three families, in all five regions, the Miocene is characterized by a leap in the palaeodiversity of molecular lineages (Figs 2c,g & 3c). Within the Atlantic and the IAA, this increase in palaeodiversity is characterized by high rates of within-region origination, which accounts for over 90% of the diversity in each of the two regions. However, the rate of origination and overall diversity in the IAA is far higher than in the Atlantic. This increased diversity is followed by numerous dispersal events from the IAA to the Indian Ocean and the Central Pacific, increasing overall diversity in those two adjacent regions. There is no dispersal between the Indian Ocean and the Atlantic for the Pomacentridae; however, within the Abudefdufinae the Atlantic is retained in its widespread range

until the early Pliocene. Both labrids and pomacentrids display a similar pattern of lineage origination and dispersal. Lineage diversity is highest in the IAA and dispersal from the IAA to adjacent regions of the Indian Ocean and the Central Pacific is of a similar magnitude (i.e. c. 25% of labrids and 25–30% of pomacentrid lineages arising in the IAA disperse out to either the Indian Ocean or Central Pacific; Fig. 2c,g). This pattern of lineage expansion from the IAA is also seen in the chaetodontid lineages (25–30% to adjacent regions; Fig. 3c). In addition, the reconstruction infers dispersal of chaetodontid lineages from the Indian to the Atlantic (MRCA *Chaetodon* C2 & C3 & C4); across the EPB (*Chaetodon* C4, *Amphichaetodon*, MRCA *Johnrandallia/Heniochus*); and from the East Pacific to the Atlantic (*Chaetodon* C4) (Fig. 3c; Fig. S5, Appendix S2). Dispersal from the Indian Ocean to the Atlantic also occurs along several labrid lineages, both before (*Bodianus*, *Thalassoma*, *Coris*, *Novaculines*) and after (*Bodianus*, *Scarus*) the TTE (c. 12–18 Ma; Fig. 2c,g; Fig. S3, Appendix S2). Labrid lineages restricted to the Atlantic appear to have undergone more cladogenesis than the pomacentrid or chaetodontid lineages there. Several vicariance events occur between Indian and IAA regions in all three families.

Pliocene/Recent (5 Ma)

Exceptionally high origination within the IAA continues in the Pliocene to Recent (Figs 2d,h & 3d) and dispersal from

Chaetodontidae

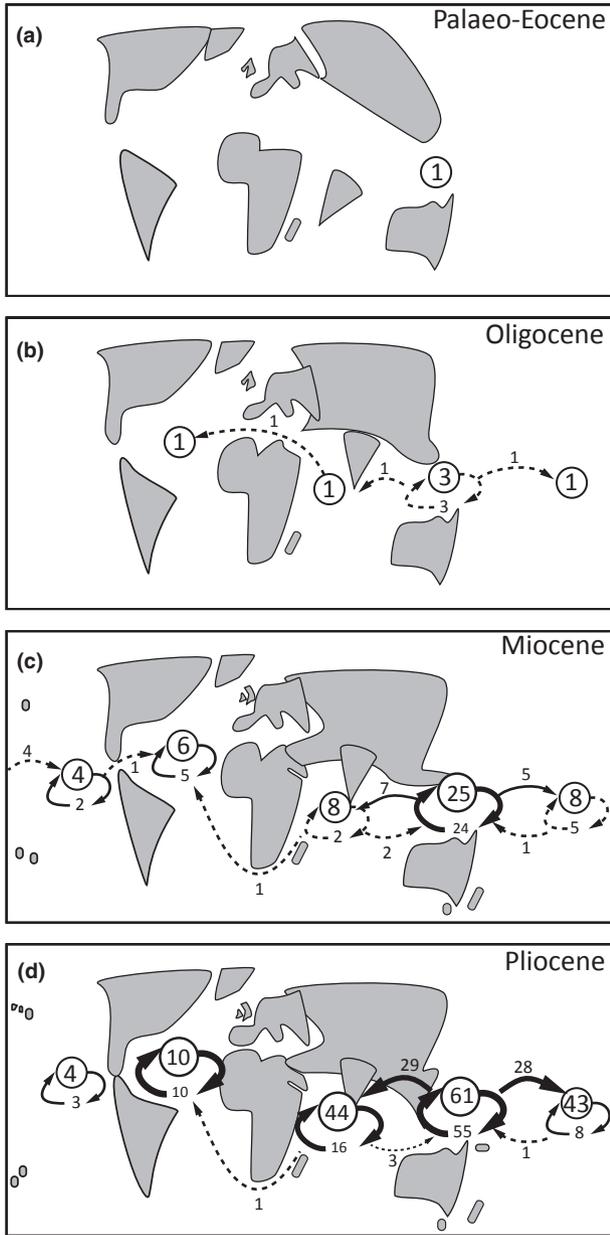


Figure 3 Schematic drawing of global palaeomaps for four time periods: Palaeo/Eocene (65–33 Ma); Oligocene (33–23 Ma); Miocene (23–5 Ma); Pliocene/Recent (5–0 Ma). Numbers show total hypothetical palaeodiversity (in circle), origination (below circle), and dispersal events (next to arrow) for each region in each period for the Chaetodontidae. See Fig. 2 for detailed legend.

the region continues to be the main source of diversity in adjacent regions. Again, the reconstruction infers a similar magnitude of lineage expansion into the Indian Ocean and/or the Central Pacific, from the IAA (27–30% of labrids and pomacentrids, 50% of chaetodontids). Within the Indian Ocean and Central Pacific regions, we now see evidence of range inheritance from the Miocene. This is particularly apparent in the Central Pacific for labrid and pomacentrid

lineages, where regional origination and dispersal from the IAA during the Pliocene only accounts for c. 50% of the biodiversity present (Fig. 2d,h). The East Pacific and Atlantic become separated by the IOP, and this is reflected by several vicariance events (Figs S3–S5, Appendix S2). No dispersal is apparent out of the Atlantic for any lineage. Lineages dispersing into the Atlantic region from the Indian Ocean quickly get separated by vicariance (Figs S3–S5, Appendix S2), with the exception of the recent movement of *Anampses caeruleopunctatus* [recorded on FishBase (Froese & Pauly, 2011) in the south-east Atlantic]. In the Pomacentridae and Chaetodontidae, no dispersal is evident across the EPB, while the Labridae maintains five separate widespread ranges across the barrier (*Calotomus carolinus*, *Scarus rubroviolaceus*, *Scarus ghobban*, *Novaculichthys taeniourus*, *Stethojulis bandanensis*; Fig. S3, Appendix S2).

DISCUSSION

The reconstruction provides a clear global overview of the last 65 Myr, with evidence of regional isolation, connection and cladogenesis reflecting the major tectonic and biogeographic events. The East Pacific and Atlantic have been largely independent regions, a pattern that reflects the increasing isolation of these regions in the Tertiary. High extinction within these regions appears to have resulted in a restricted cladogenic history. The Indo-Pacific, in contrast, is characterized by extensive origination and connectivity between adjacent regions. The reconstruction suggests that the hotspot in the IAA is, at least in part, a result of prolonged origination over the past c. 30 Myr and that diversity in the Indian Ocean and Central Pacific is largely a result of dispersal from the IAA in the last 5 Myr. Survival of ancestral lineages in the IAA laid the foundation for the subsequent cladogenesis and range expansion that has led to present-day patterns of biodiversity across the entire Indo-Pacific.

Our study is the first to apply the DEC model to the exploration of globally distributed coral reef fish taxa. The results of the reconstruction presented herein show striking parallels with the patterns found in other taxa. Most notable is the congruence with global patterns of dispersal with terrestrial plants, where Australia and the IAA support numerous lineages with extensive evidence of dispersal from the Miocene onwards (Buerki *et al.*, 2011). The use of the DEC model highlights the origins and progression of biodiversity in the marine tropics, with disparate rates of extinction and survival playing key roles in shaping the global distribution of reef associated fishes. The changing role of the IAA hotspot in driving the current biodiversity across the entire Indo-Pacific is illustrated from inferred pattern of palaeodiversity of molecular lineages.

Origination and biodiversity through time and space

Each of the five regions can be grouped reflecting three basic patterns of origination and diversity:

1. Independent and isolated (East Pacific and Atlantic).
2. A macroevolutionary source of diversity (IAA).
3. Diversity recipients, or macroevolutionary sinks (Indian Ocean and Central Pacific).

Each will be examined separately below, reflecting these associations.

East Pacific and Atlantic

The history of the East Pacific and Atlantic biota, before the closing of the IOP, are closely linked. Both regions share a subset of the families present in the Indo-Pacific, although there are a few absent families or subfamilies in the Atlantic and East Pacific that are present in the Indo-Pacific (e.g. Caesionidae, Nasinae, Siganidae) (Bellwood & Wainwright, 2002). The origination of extant lineages within the East Pacific and Atlantic regions are comparable with other areas in terms of percentage regional origination (Fig. 1b). However, there appears to have been less origination as a percentage of all extant lineages in each family (Fig. 1c). This may reflect reduced speciation capacity within the two regions, or increased rate of lineage extinction. Historically, both regions have experienced periods of increased faunal turnover (McCoy & Heck, 1976; Budd, 2000; Bellwood & Wainwright, 2002; O'Dea *et al.*, 2007). Increased extinction may result in fewer extant species despite the region having potentially high origination. The lack of connection to the Indo-Pacific has reduced the potential for lineage replenishment. Our results offer support for the recent works of Rocha *et al.* (2005), Floeter *et al.* (2008) and Joyeux *et al.* (2008), which have suggested that these areas are largely independent and have been isolated from the Indo-Pacific for some time. However, the reconstruction highlights the early development of this isolation in the Oligocene following broader connectivity in the Palaeocene.

During the Palaeocene/Eocene, the East Pacific and Atlantic were linked to the Indo-Pacific by dispersal through the Tethys Sea. Evidence of this can be seen in the fossil record of Monte Bolca in Italy with the majority of the fish families represented in the deposits being present in both the Atlantic and Indo-Pacific (reviewed by Bellwood & Wainwright, 2002). This pantropical connection between the Atlantic and Indo-Pacific is evident from the reconstruction, with several inferred dispersal events between all tropical marine regions for the Labridae and Pomacentridae (Fig. 2a,e). While panmixia is possible during the Eocene, we begin to see incipient provinciality with restricted lineages in each region, while the inferred palaeodiversity of molecular lineages in the Atlantic and East Pacific is lower than in the IAA. This lower palaeodiversity maybe a consequence of high extinction throughout the epoch, or a period of high extinction at the end of the Eocene, which has been hypothesized for both these fishes (Cowman & Bellwood, 2011) and for flowering plants (Antonelli & Sanmartín, 2011).

The isolation of the Atlantic and East Pacific regions from the Indo-Pacific in the Eocene escalates in the Oligocene,

with few dispersal events inferred between the Indian Ocean and Atlantic, or the Central Pacific and East Pacific regions. Dispersal between the East Pacific and Atlantic is also reduced. These earlier divisions may reflect earlier, temporary barriers within the Atlantic (Hallam, 1973), and East Pacific (Lessios *et al.*, 1995). However, by the end of the Miocene, the Atlantic and the East Pacific are largely isolated from the Indo-Pacific and from each other.

The Miocene epoch represents the time period with the highest lineage diversification in the East Pacific and Atlantic especially within the Labridae, reflecting patterns previously described in the labrid genus *Halichoeres* (Barber & Bellwood, 2005) and other reef genera (Floeter *et al.*, 2008). Circum-African lineages within the Labridae and Chaetodontidae appear to allow some Indo-Pacific lineage expansion into the Atlantic. This is not possible for pomacentrid lineages (with the exception of *Abudefduf*), perhaps lacking ability to survive in the more temperate climes of the Southwest African coast (cf. Floeter *et al.*, 2008). A temporary reconnection before the final closure of the IOP is implied by several bidirectional dispersal events between the East Pacific and Atlantic in each of the three families (Figs S3–S5, Appendix S2). After the closure of the IOP the regions are completely separated from each other and the Atlantic is largely isolated from the Indo-Pacific. While the influence of the Indo-Pacific biota on the Atlantic fauna has been recorded recently (Robertson *et al.*, 2004; Rocha *et al.*, 2005; Bowen *et al.*, 2006), it has made little overall impact to the extant biodiversity of the Atlantic and East Pacific regions (Fig. 1a; Floeter *et al.*, 2008). Regional origination does continue in the Atlantic and East Pacific during the Pliocene/Recent epoch, but little significant increase in regional richness is observed (with the possible exception of Atlantic wrasses; Barber & Bellwood, 2005).

Indo-Australian archipelago

For the Labridae, Pomacentridae, and Chaetodontidae, the reconstruction suggests that the IAA has been a centre for lineage survival, origination, and range expansion at different stages in the history of the region (Fig. 4). In terms of global diversity for each family, the IAA stands out as a significant source of diversity in terms of both origination within the region and the expansion of lineages into adjacent regions. Based on the reconstruction, the origination of extant lineages within the IAA accounts for c. 60% of total global biodiversity for each family (Fig. 2c). This is relatively direct evidence for the IAA as a centre of origin in contrast to inferred origination based on an assumed preponderance of endemic taxa (e.g. Briggs, 2003; Mora *et al.*, 2003; Roberts *et al.*, 2002; but see Bellwood & Meyer, 2009a,b). Patterns of dispersal inferred from the reconstruction also show the connectivity within Indo-Pacific, with the IAA acting as a source of biodiversity for the Indian Ocean and Central Pacific (Table 1; Fig. 4a). However, a temporal perspective highlights the progression of high biodiversity in the IAA from

overlapping ranges in the Eocene, survival in the Oligocene, origination in the Miocene, and expansion in the Pliocene and Recent (Fig. 4).

During the Palaeocene/Eocene, the IAA is reconstructed by the model as a centre of palaeodiversity of molecular lineages, with both localized cladogenesis and overlapping widespread ranges (Labridae, Pomacentridae). A similar West Pacific origin for basal julidine lineages has been suggested previously (Westneat & Alfaro, 2005). However, fossil evidence from the middle Eocene (42–39 Ma) highlights the Tethys regions as the centre for palaeodiversity of reef associated taxa (Renema *et al.*, 2008). Indeed, fossils from the Eocene of Monte Bolca mark the earliest record of many extant reef fish forms (Bellwood, 1996), with fewer fossils found in the Indo-Pacific or IAA region. The conflicting biogeographic patterns from the fossil record and the ancestral reconstruction of molecular lineages presented here may reflect the influence of extinction in other regions (Budd, 2000; Bellwood & Wainwright, 2002). The greater palaeodiversity of ancestral molecular lineages in the IAA from the early Oligocene onwards may highlight also the ability of the IAA to support ancestral lineages, i.e. it is a centre of survival of ancestral lineages (Barber & Bellwood, 2005; Cowman & Bellwood, 2011). In this, it appears that the key to variation is in the ability to support taxa. As mentioned by Rabosky (2009), ecology may overshadow temporal patterns of diversity. In this respect the expansion of coral reef area in the IAA may have underpinned the observed pattern. Indeed, total reef area is the primary variable explaining variation in

reef fish diversity (Bellwood & Hughes, 2001; Bellwood *et al.*, 2005). Thus, the expansion in the IAA is probably not just age and cladogenesis but the capacity of the area to maintain and support species, even in the Oligocene. By the beginning of the Oligocene, the IAA is already emerging as a centre for diversity, however, this may be the result of higher rates of survival within the region rather than exceptional rates of origination.

The IAA hosts the initial divergences of the Chaetodontidae in the Oligocene and the origination of labrid and pomacentrid lineages continues. There is little dispersal into the Indian Ocean, possibly highlighting adverse conditions there (Hallam, 1984). However, it is in the Miocene that the diversity synonymous with the IAA begins to emerge, coinciding with the expansion of coral-dominated reefs and a mosaic of reef habitats (Renema *et al.*, 2008). Lineages that have previously been highlighted as containing exceptional diversity, have inferred ancestral origins in the IAA during the Miocene and many show strong links to coral reefs (Cowman & Bellwood, 2011). It is in the Miocene that lineages from the IAA now begin to recolonize the Indian Ocean. Origination continues in the Pliocene/Recent epochs primarily in *Chaetodon* lineages. The extensive dispersal from the IAA to the Indian Ocean and Central Pacific regions highlight the role of the IAA as a source of diversity across the entire Indo-Pacific during this period (Fig. 4).

It appears that the survival of ancient lineages in the IAA resulted in more subsequent cladogenesis, which in turn permitted extensive range expansion and the export of lineages

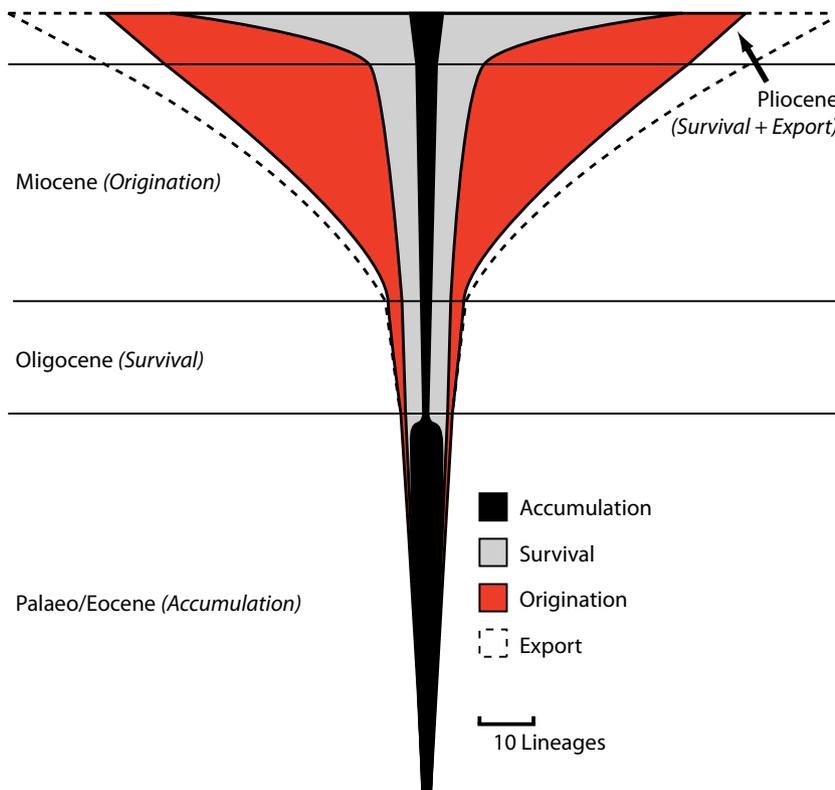


Figure 4 Schematic diagram illustrating the changing role of the Indo-Australian Archipelago (IAA) hotspot in the origins of Indo-Pacific reef fish biodiversity, inferred from the reconstruction. In each epoch, the dominant roles of the IAA hotspot changed from accumulation (Palaeo/Eocene), survival (Oligocene), origination (Miocene) and export (Pliocene) of lineages. Accumulation = lineages acquired from outside the IAA; Survival = lineages retained from the previous epoch (= Accumulation + Origination from previous epoch); Origination = lineages arising in the IAA during that epoch; Export = lineages in the IAA that were exported/expanded to adjacent regions during that epoch. Each process has been dominant at one time, but most act simultaneously through time. The relative width of the four main roles, accumulation (black), origination (red, or dark grey in greyscale), survival (pale grey) and export (open) reflect the relative contribution to palaeodiversity in the labrid reconstruction (a–d).

into adjacent regions. In summary, the IAA has sequentially and then simultaneously acted as a centre of accumulation (Palaeocene/Eocene onwards), survival (Eocene/Oligocene onwards), origin (Miocene onwards), and export/expansion (Pliocene/Recent) for reef fishes in the three families (Fig. 4). In this way, it reflects the general patterns of the 'hopping hotspot' discussed by Renema *et al.* (2008). However, this is the first clear evidence of the IAA being a centre for lineage export/expansion over an extended period of time.

Indian Ocean and Central Pacific

The Indian Ocean and Central Pacific regions are both recipients, or macroevolutionary sinks, for biodiversity from the IAA. Regional origination accounts for little of the diversity in both regions (Table 1). While both regions appear to have relatively low regional origination since the Palaeocene, the influence of expanding IAA lineages has escalated from the Miocene to Recent. The timing of these events is examined below for each region.

The current biodiversity of the Indian Ocean is the result of moderate cladogenesis in the region, with the majority of species being of IAA origin that maintain the Indian Ocean as part of a wider Indo-Pacific range (cf. Hughes *et al.*, 2002). It has been a sink for lineages from the IAA since the Miocene. Excluding the Chaetodontidae, the majority of lineages that have arisen in the Indian Ocean are of late Miocene/early Pliocene age. Invasion of Indian Ocean species into the IAA appears to be much less frequent. However, they may make up almost half of the total origination in the Indian Ocean due to the low origination of lineages there. Dispersal from the IAA to the Indian Ocean and to the Central Pacific seems to be of a similar magnitude in both directions. Starting in the early Miocene and continuing into the Pliocene, approximately 25–30% of lineages arising in the IAA disperse to adjacent regions.

The pattern in the Indian Ocean reflects the ancient history of the region with links to declining or receding hotspots (Renema *et al.*, 2008). By the end of the Eocene the Arabian hotspot would have been in place in the western Indian Ocean (Renema *et al.*, 2008). Tectonic rearrangement and global cooling in the Oligocene/Miocene would have had a dramatic effect on the shallow water reefs in this area (Hallam, 1994). With the restriction of the tropics and decline in reef area in the Arabian hotspot, faunal loss would have occurred in the Indian Ocean throughout the Oligocene/early Miocene. Those lineages that maintained an extended range into the IAA would presumably have been able to survive best at this time. This pattern appears to hold for mangrove, foraminifera, and other reef associated organisms (Renema *et al.*, 2008), with the resistance to extinction in the IAA (Cowman & Bellwood, 2011) being based on the unique features of the area (Hoeksema, 2007) and the diversity of reef habitats (Rosen, 1984). This loss is counteracted by invasion from the IAA to maintain palaeodiversity.

In contrast to the Indian Ocean, the island arcs that make up the Central Pacific region have been closely linked to the IAA, with gene flow between the two regions allowing the Central Pacific to easily inherit ancestral widespread ranges (and vice versa; see Craig *et al.*, 2010; Gaither *et al.*, 2011). The EPB has prevented the invasion of East Pacific/Atlantic species (but see Lessios & Robertson, 2006), resulting in most of the diversity seen in the Central Pacific being of IAA origin. The Central Pacific has comparable species richness to the Indian Ocean; however, its low regional origination appears to be a function of a higher occurrence of range inheritance and speciation through peripheral isolation (cf. Planes & Fauvelot, 2002; Hodge *et al.*, 2012). The reconstruction suggests that while there is moderate range expansion of lineages originating in the Indian Ocean to the IAA, expansion appears very limited from the Central Pacific to the IAA (Fig. 1a). Overall, the Central Pacific remains largely a recipient of lineages from the IAA.

Recent dispersal from the Indian Ocean to the IAA has been highlighted in previous work where populations form temporal rather than geographic clades (Horne *et al.*, 2008; Gaither *et al.*, 2011). From the IAA to the Central Pacific region there is evidence of geographic clade structure of both populations (Planes & Fauvelot, 2002; but see Horne *et al.*, 2008) and species (Bernardi *et al.*, 2002). However, these are all relatively young associations. The reconstruction herein identifies prolonged movement from the IAA to the Indian and Pacific Oceans beginning in the Miocene and escalating during the Pliocene to Recent. While the Indian Ocean reflects a history of extinction and faunal recovery from lineages arising in the IAA, the Central Pacific is characterized by widespread ancestrally inherited ranges, with a smaller number of peripheral endemics.

Caveats and considerations

The combined use of dated phylogenetic trees and the DEC model implemented in LAGRANGE has allowed the evolutionary history of the Labridae, Pomacentridae and Chaetodontidae to be explored in a biogeographic context. There are, however, several sources of potential bias. This applies to the phylogenetic trees used for this reconstruction (discussed in Cowman & Bellwood, 2011), and with the ancestral reconstruction itself. Five major caveats need to be considered when interpreting the results:

1. This is a hypothetical reconstruction, based on present-day distributions of taxa and as such, there are limitations to the interpretation of a lineage's geographic origin (Losos & Glor, 2003). Post-speciation expansion is fast, giving allopatric species the ability to become completely sympatric within 4 Myr (Quenouille *et al.*, 2011), (although the reverse may also be true). For older lineages this has the potential to obscure geographic origins, especially on a smaller within-region scale. For the broad, regional patterns described herein the effect will be relatively limited. Discussion was therefore limited to patterns between ocean basins known to have distinct faunal assemblages. Discussions of patterns

within regions are restricted to lineage origination with no implications for the mode of speciation.

2. Incomplete taxon sampling of extant lineages will also affect the biogeographic reconstruction. For example, the *Prognathodes* lineage of the Chaetodontidae was reconstructed to be restricted to the Atlantic; however, species not sampled in the phylogenies have been recorded in the Indo-Pacific (*P. guezei*) and the East Pacific (*P. falcifer*) (Kuitert, 2002). Inclusion of these species has the potential to change the geographic origins of the clade (Fig. S5, Appendix S2). This problem may be exacerbated by unsampled cryptic speciation. Recent phylogeographic studies in the Atlantic have identified cryptic species in *Halichoeres* lineages (Rocha *et al.*, 2005). Such studies have been limited in the Indo-Pacific (Rocha & Bowen, 2008) and those that have been undertaken have not found the level of species division seen in the Atlantic. Nevertheless, further studies may alter presently accepted taxonomy and species distributions. In turn, this may affect the mean ages of lineages within regions, and may reveal more recent cladogenic events in the Indo-Pacific.

The relatively even sampling of taxa across regions and among clades within families (Table S1, Appendix S1), and near complete generic sampling (Cowman & Bellwood, 2011), allows for some confidence in the interpretation of patterns of historical biodiversity of lineages. By examining the origination of lineages, rather than species, at the larger scale of the five major biogeographic regions, interpretations for global patterns should be relatively robust. Furthermore, analysis using total species counts rather than just species sampled within clades resulted in no change in the observed patterns. However, a more detailed evaluation of the effect of sampling would be valuable, especially if analyses are to be informative at a finer within-region scale.

3. The ancestral reconstruction does not allow for the global extinction of a lineage. While the contraction of a lineage's range through local extinction can be inferred from the reconstruction, if large numbers of taxa have gone extinct in a region this will translate into a perceived lower palaeodiversity of molecular lineages within that region. Palaeodiversities may therefore be underestimated. There is known variation in extinction rates among the five regions. Extinction of reef-associated fauna has been described in previous studies for the East Pacific and Atlantic (McCoy & Heck, 1976; Budd, 2000; Bellwood & Wainwright, 2002; O'Dea *et al.*, 2007), the Indian Ocean (Renema *et al.*, 2008), and may have been high in the Central Pacific during periods of sea-level change (Paulay, 1997; Fauvelot *et al.*, 2003). The only region where high past extinction rates have not been recorded on reefs is in the IAA (but see Springer & Williams, 1994; Williams & Duda, 2008). This may result in the IAA being inferred as a centre of origin for deeper lineages, when in fact they were probably peripheral to a widespread ancestral range.

The reconstruction of the widespread ancestral root of the Labridae reflects the widespread records of fossil labrids in the West Tethys (D.R.B., unpublished data). This offers some confidence in the palaeo-reconstructed distribution (Fig. S3, Appen-

dix S2). However, the pomacentrids may provide some evidence of the effect that extinction has had on biogeographic interpretations. The reconstruction of the pomacentrid root origin in the IAA is contrary to the fossil record with two pomacentrid representatives in the Eocene of West Tethys (Bellwood & Sorbini, 1996; Bellwood, 1999), which indicates the presence but not, necessarily, the origin of pomacentrids in the West Tethys. Therefore, while the reconstruction of the Labridae provides support for the shift from a widespread Tethys palaeodistribution to the IAA, the conflicting pomacentrid reconstruction is potential evidence for extinction in the West Tethys.

4. There is also a strong likelihood that clade diversity within these families is not related to the age of the clade itself. In the absence of a relationship between clade age and clade richness, clade diversity can be limited by other ecological factors (Rabosky, 2009). If the regional diversity of a lineage is limited by ecological factors, the regional capacity for diversification will be influenced by changes in ecological limits rather than by length of time a clade has had to diversify. Habitat area is one such factor that can influence the carrying capacity of region. While lineages in different regions may experience the same rate of speciation and/or extinction, the capacity for cladogenesis may be regulated by habitat area. While this may not alter the inferred ancestral patterns, it does limit the interpretation. Reef area has been shown to be a major factor in explaining observed variation in the richness of coral reef species in the IAA (Bellwood & Hughes, 2001; Bellwood *et al.*, 2005). However, on an evolutionary scale reef area may also be a factor regulating diversity, by giving clades an extended growth phase before carrying capacity is reached (see Rabosky, 2009), rather than underpinning exceptionally high rates of diversification. If so, this extended growth phase appears to extend throughout the Miocene (Fig. 4), where several significantly diverse clades arise (Cowman & Bellwood, 2011). In this respect, the diversification of reef fishes in the IAA during the Miocene may be the result of expanding reef area elevating the regions capacity to support biodiversity.

5. The final caveat, and potentially the most important, is 'the pull of the present' (Pybus & Harvey, 2000). This has been discussed as a bias arising from lineages arising in the recent past being less likely to go extinct than lineages in the distant past. In a historical biogeographic context it can be seen as a bias in the likelihood of an inferred lineage arising within a specific location. Ancestral origination within the IAA may therefore arise purely from it being a location of overlapping extant ranges. In this way factors that have resulted in the accumulation, survival or maintenance of species in the IAA will result in the reconstruction of origination in the IAA. The conflict between maintenance and origination is an old one with little resolution. However, the hopping hotspot theory (Renema *et al.*, 2008) does provide some support, with fossil evidence of a shift in invertebrate fossil diversity from the Tethys to the IAA. However, the lack of independent evidence of origination of fishes in the IAA (endemics or fossils; Bellwood & Meyer, 2009a) is a concern

and therefore conclusions from the reconstruction must be taken as an indication only. Solid independent evidence, preferably fossil, is urgently required.

CONCLUSIONS

For the first time, temporally congruent patterns in origination and dispersal of molecular lineages highlight global patterns of historical biodiversity of reef fishes within five major tropical marine regions. The Atlantic and East Pacific regions have been largely independent and isolated from the Indo-Pacific from the Oligocene, but have had a minor influx of Indo-Pacific biota. Within the Indo-Pacific, the IAA has played important roles as a centre of lineage accumulation, survival and origination both sequentially and simultaneously. Those lineages that survived in the IAA while others went extinct elsewhere are the driving forces behind current biodiversity in the Indo-Pacific. While ancestral reconstruction requires careful interpretation, it has allowed insights into the changing role of the largest marine biodiversity hotspot. Diversity of reef fishes in the IAA hotspot potentially began as a location within a widespread range that gradually accumulated ancestral lineages, for which it provided a refuge during a period of extinction in adjacent ancestral habitat. Surviving lineages in the IAA proliferated and formed the basis for extensive recolonization of the Indian and Pacific Ocean realms.

ACKNOWLEDGEMENTS

We would like to thank L. van Herwerden, C. Goatley and A. Hoey for helpful discussions; J. Tanner, M. Alfaro, F. Santini and an anonymous referee for valuable comments on earlier drafts. This work was supported by The Australian Research Council (D.R.B.).

REFERENCES

- Allen, G.R. (1991) *Damselfishes of the world*. Aquarium Systems, Melle, Germany.
- Allen, G.R., Steene, R. & Allen, M. (1998) *A guide to angelfishes and butterflyfishes*. Odyssey Publishing/Tropical Reef Research, Perth.
- Antonelli, A. & Sanmartín, I. (2011) Mass extinction, gradual cooling, or rapid radiation? Reconstructing the spatiotemporal evolution of the ancient angiosperm genus *Hedyosmum* (Chloranthaceae) using empirical and simulated approaches. *Systematic Biology*, **60**, 596–615.
- Barber, P.H. & Bellwood, D.R. (2005) Biodiversity hotspots: evolutionary origins of biodiversity in wrasses (*Halichoeres*: Labridae) in the Indo-Pacific and New World tropics. *Molecular Phylogenetics and Evolution*, **35**, 235–253.
- Barber, P.H., Palumbi, S.R., Erdmann, M.V. & Moosa, M.K. (2000) A marine Wallace's line? *Nature*, **406**, 692–693.
- Barber, P.H., Palumbi, S.R., Erdmann, M.V. & Moosa, M.K. (2002) Sharp genetic breaks among populations of a ben-
thic marine crustacean indicate limited oceanic larval transport: patterns, causes, and consequences. *Molecular Ecology*, **11**, 659–674.
- Barracough, T. & Vogler, A. (2000) Detecting the geographical pattern of speciation from species-level phylogenies. *The American Naturalist*, **155**, 419–434.
- Beldade, R., Heiser, J.B., Robertson, D.R., Gasparini, J.L., Floeter, S.R. & Bernardi, G. (2009) Historical biogeography and speciation in the Creole wrasses (Labridae, *Clepticus*). *Marine Biology*, **156**, 679–687.
- Bellwood, D.R. (1996) The Eocene fishes of Monte Bolca: the earliest coral reef fish assemblage. *Coral Reefs*, **15**, 11–19.
- Bellwood, D.R. (1999) Fossil pharyngognath fishes from Monte Bolca, Italy, with a description of a new pomacentrid genus and species. *Studi e Ricerche sui Giacimenti Terziari di Bolca, Miscellanea Palaeontologica, Museo Civico di Storia Naturale di Verona*, **8**, 207–217.
- Bellwood, D.R. & Hughes, T.P. (2001) Regional-scale assembly rules and biodiversity of coral reefs. *Science*, **292**, 1532–1535.
- Bellwood, D.R. & Meyer, C.P. (2009a) Searching for heat in a marine biodiversity hotspot. *Journal of Biogeography*, **36**, 569–576.
- Bellwood, D.R. & Meyer, C.P. (2009b) Endemism and evolution in the Coral Triangle: a call for clarity. *Journal of Biogeography*, **36**, 2010–2012.
- Bellwood, D.R. & Sorbini, L. (1996) A review of the fossil record of the Pomacentridae (Teleostei: Labroidae) with a description of a new genus and species from the Eocene of Monte Bolca, Italy. *Zoological Journal of the Linnean Society*, **117**, 159–174.
- Bellwood, D.R. & Wainwright, P.W. (2002) The history and biogeography of fishes on coral reefs. *Coral reef fishes: dynamics and diversity in a complex ecosystem* (ed. by P.F. Sale), pp. 5–32. Academic Press, London.
- Bellwood, D.R., Hughes, T.P., Connolly, S.R. & Tanner, J. (2005) Environmental and geometric constraints on Indo-Pacific coral reef biodiversity. *Ecology Letters*, **8**, 643–651.
- Bellwood, D.R., Klanten, S., Cowman, P.F., Pratchett, M.S., Konow, N. & van Herwerden, L. (2010) Evolutionary history of the butterflyfishes (f: Chaetodontidae) and the rise of coral feeding fishes. *Journal of Evolutionary Biology*, **23**, 335–349.
- Bellwood, D.R., Renema, W. & Rosen, B.R. (2012) Biodiversity hotspots, evolution and coral reef biogeography: a review. *Biotic evolution and environmental change in Southeast Asia* (ed. by D.J. Gower, K.G. Johnson, J.E. Richardson, B.R. Rosen, L. Rüber and S.T. Williams), pp. 216–245. Cambridge University Press, Cambridge.
- Bernardi, G., Holbrook, S.J., Schmitt, R.J., Crane, N.L. & DeMartini, E. (2002) Species boundaries, populations and colour morphs in the coral reef three-spot damselfish (*Dascyllus trimaculatus*) species complex. *Proceedings of the Royal Society B: Biological Sciences*, **269**, 599–605.
- Bowen, B.W., Muss, A., Rocha, L.A. & Grant, W.S. (2006) Shallow mtDNA coalescence in Atlantic pygmy angelfishes

- (genus *Centropyge*) indicates a recent invasion from the Indian Ocean. *Journal of Heredity*, **97**, 1–12.
- Briggs, J.C. (1999) Coincident biogeographic patterns: Indo-West Pacific Ocean. *Evolution*, **53**, 326–335.
- Briggs, J.C. (2003) Marine centres of origin as evolutionary engines. *Journal of Biogeography*, **30**, 1–18.
- Budd, A.F. (2000) Diversity and extinction in the Cenozoic history of Caribbean reefs. *Coral Reefs*, **19**, 25–35.
- Buerki, S., Forest, F., Alvarez, N., Nylander, J.A.A., Arrigo, N. & Sanmartín, I. (2011) An evaluation of new parsimony-based versus parametric inference methods in biogeography: a case study using the globally distributed plant family Sapindaceae. *Journal of Biogeography*, **38**, 531–550.
- Connolly, S., Bellwood, D.R. & Hughes, T.P. (2003) Indo-Pacific biodiversity of coral reefs: deviations from a mid-domain model. *Ecology*, **84**, 2178–2190.
- Cowman, P.F. & Bellwood, D.R. (2011) Coral reefs as drivers of cladogenesis: expanding coral reefs, cryptic extinction events, and the development of biodiversity hotspots. *Journal of Evolutionary Biology*, **24**, 2543–2562.
- Cowman, P.F., Bellwood, D.R. & van Herwerden, L. (2009) Dating the evolutionary origins of wrasse lineages (Labridae) and the rise of trophic novelty on coral reefs. *Molecular Phylogenetics and Evolution*, **52**, 621–631.
- Craig, M.T., Eble, J.A. & Bowen, B.W. (2010) Origins, ages and population histories: comparative phylogeography of endemic Hawaiian butterflyfishes (genus *Chaetodon*). *Journal of Biogeography*, **37**, 2125–2136.
- Ekman, S. (1935) *Tiergeographie des Meeres*. Akademische Verlagsgesellschaft, Leipzig.
- Fauvelot, C., Bernardi, G. & Planes, S. (2003) Reductions in the mitochondrial DNA diversity of coral reef fish provide evidence of population bottlenecks resulting from Holocene sea-level change. *Evolution*, **57**, 1571–1583.
- Floeter, S.R., Guimaraes, R.Z.P., Rocha, L.A., Ferreira, C.E.L., Rangel, C.A. & Gasparini, J.L. (2001) Geographic variation in reef-fish assemblages along the Brazilian coast. *Global Ecology and Biogeography*, **10**, 423–431.
- Floeter, S.R., Rocha, L.A., Robertson, D.R., Joyeux, J.C., Smith-Vaniz, W.F., Wirtz, P., Edwards, A.J., Barreiros, J.P., Ferreira, C.E.L., Gasparini, J.L., Brito, A., Falcón, J.M., Bowen, B.W. & Bernardi, G. (2008) Atlantic reef fish biogeography and evolution. *Journal of Biogeography*, **35**, 22–47.
- Froese, F. & Pauly, D. (2011) *FishBase*. Available at: <http://www.fishbase.org> (accessed 10 May 2012).
- Gaither, M.R., Bowen, B.W., Bordenave, T.R., Rocha, L.A., Newman, S.J., Gomez, J.A., van Herwerden, L. & Craig, M.T. (2011) Phylogeography of the reef fish *Cephalopholis argus* (Epinephelidae) indicates Pleistocene isolation across the Indo-Pacific Barrier with contemporary overlap in the Coral Triangle. *BMC Evolutionary Biology*, **11**, 189.
- Göbbeler, K. & Klusmann-Kolb, A. (2010) Out of Antarctica? – New insights into the phylogeny and biogeography of the Pleurobranchomorpha (Mollusca, Gastropoda). *Molecular Phylogenetics and Evolution*, **55**, 996–1007.
- Hallam, A. (1973) Distributional patterns in contemporary terrestrial and marine animal. *Organisms and continents through time* (ed. by N.F. Hughes), pp. 93–105. Paleontological Association, London.
- Hallam, A. (1984) Pre-Quaternary sea-level changes. *Annual Review of Earth Planetary Sciences*, **12**, 205–243.
- Hallam, A. (1994) *An outline of Phanerozoic biogeography*. Oxford University Press, Oxford.
- Heck, K.L. & McCoy, E.D. (1978) Long-distance dispersal and the reef-building corals of the Eastern Pacific. *Marine Biology*, **48**, 349–356.
- Hodge, J.R., Read, C.I., van Herwerden, L. & Bellwood, D.R. (2012) The role of peripheral endemism in species diversification: evidence from the coral reef fish genus *Anampses* (Family: Labridae). *Molecular Phylogenetics and Evolution*, **62**, 653–663.
- Hoeksema, B. (2007) Delineation of the Indo-Malayan centre of maximum marine biodiversity: the Coral Triangle. *Biogeography, time, and place: distributions, barriers, and islands* (ed. by W. Renema), pp. 117–178. Springer, Dordrecht, The Netherlands.
- Horne, J.B., van Herwerden, L., Choat, J.H. & Robertson, D.R. (2008) High population connectivity across the Indo-Pacific: congruent lack of phylogeographic structure in three reef fish congeners. *Molecular Phylogenetics and Evolution*, **49**, 629–638.
- Hughes, T.P., Bellwood, D.R. & Connolly, S.R. (2002) Biodiversity hotspots, centres of endemism, and the conservation of coral reefs. *Ecology Letters*, **5**, 775–784.
- Ilves, K.L. & Taylor, E.B. (2007) Evolutionary and biogeographical patterns within the smelt genus *Hypomesus* in the North Pacific Ocean. *Journal of Biogeography*, **35**, 48–64.
- Jones, G.P., Caley, M.J. & Munday, P.L. (2002) Rarity in coral reef fish communities. *Coral reef fishes, dynamics and diversity in a complex ecosystem* (ed. by P.F. Sale), pp. 81–101. Academic Press, San Diego, CA.
- Joyeux, J.C., Floeter, S.R., Ferreira, C.E.L. & Gasparini, J.L. (2008) Biogeography of tropical reef fishes: the South Atlantic puzzle. *Journal of Biogeography*, **28**, 831–841.
- Kelly, C.M.R., Barker, N.P., Villet, M.H. & Broadley, D.G. (2009) Phylogeny, biogeography and classification of the snake superfamily Elapoidea: a rapid radiation in the late Eocene. *Cladistics*, **25**, 38–63.
- Kuiter, R.H. (2002) *Butterflyfishes, bannerfishes and their relatives: a comprehensive guide to Chaetodontidae and Microcanthidae*. TMC Publishing, Chorleywood, UK.
- Lessios, H.A. (2008) The great American schism: divergence of marine organisms after the rise of the Central American Isthmus. *Annual Review of Ecology, Evolution and Systematics*, **39**, 63–91.
- Lessios, H.A. & Robertson, D.R. (2006) Crossing the impassable: genetic connections in 20 reef fishes across the eastern Pacific barrier. *Proceedings of the Royal Society B: Biological Sciences*, **273**, 2201–2208.

- Lessios, H.A., Allen, G.R., Wellington, G. & Bermingham, E. (1995) Genetic and morphological evidence that the eastern Pacific damselfish *Abudefduf declivifrons* is distinct from *A. concolor* (Pomacentridae). *Copeia*, **1995**, 277–288.
- Losos, J.B. & Glor, R.E. (2003) Phylogenetic comparative methods and the geography of speciation. *Trends in Ecology and Evolution*, **18**, 220–227.
- McCafferty, S., Bermingham, E., Quenouille, B., Planes, S., Hoelzer, G. & Asoh, K. (2002) Historical biogeography and molecular systematics of the Indo-Pacific genus *Dascyllus* (Teleostei: Pomacentridae). *Molecular Ecology*, **11**, 1377–1392.
- McCoy, E.D. & Heck, K.L., Jr (1976) Biogeography of corals, seagrasses and mangroves: an alternative to the center of origin concept. *Systematic Zoology*, **25**, 201–210.
- Mora, C., Chittaro, P.M., Sale, P.F., Kritzer, J.P. & Ludsins, S.A. (2003) Patterns and processes in reef fish diversity. *Nature*, **421**, 933–936.
- de Moura, R. & Sazima, I. (2000) Species richness and endemism levels of the Southwestern Atlantic reef fish fauna. *Proceedings of the 9th International Coral Reef Symposium*, **1**, 1–6.
- O’Dea, A., Jackson, J.B.C., Fortunato, H., Smith, J.T., D’Croz, L., Johnson, K.G. & Todd, J.A. (2007) Environmental change preceded Caribbean extinction by 2 million years. *Proceedings of the National Academy of Sciences USA*, **104**, 5501–5506.
- Palumbi, S.R. (1997) Molecular biogeography of the Pacific. *Coral Reefs*, **16**, S47–S52.
- Paulay, G. (1997) Diversity and distribution of reef organisms. *Life and death of coral reefs* (ed. by C.M. Birkeland), pp. 298–353. Chapman and Hall, New York.
- Planes, S. & Fauvelot, C. (2002) Isolation by distance and vicariance drive genetic structure of a coral reef fish in the Pacific Ocean. *Evolution*, **56**, 378–399.
- Potts, D.C. (1985) Sea level fluctuations and speciation in Scleractina. *Proceedings of the 5th International Coral Reef Congress*, **4**, 127–132.
- Pybus, O.G. & Harvey, P.H. (2000) Testing macro-evolutionary models using incomplete molecular phylogenies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **267**, 2267–2272.
- Quenouille, B., Hubert, N., Bermingham, E. & Planes, S. (2011) Speciation in tropical seas: allopatry followed by range change. *Molecular Phylogenetics and Evolution*, **58**, 546–552.
- Rabosky, D.L. (2009) Ecological limits and diversification rate: alternative paradigms to explain the variation in species richness among clades and regions. *Ecology Letters*, **12**, 735–743.
- Ramírez, S.R., Roubik, D.W., Skov, C. & Pierce, N.E. (2010) Phylogeny, diversification patterns and historical biogeography of euglossine orchid bees (Hymenoptera: Apidae). *Biological Journal of the Linnean Society*, **100**, 552–572.
- Randall, J.E. (2005) *Reef and shore fishes of the South Pacific*. University of Hawaii Press, Honolulu.
- Randall, J.E., Allen, G.R. & Steene, R.C. (1990) *Fishes of the Great Barrier Reef and Coral Sea*. Crawford House, Bathurst, NSW.
- Ree, R.H. & Smith, S.A. (2008) Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology*, **57**, 4–14.
- Ree, R.H., Moore, B.R., Webb, C.O. & Donoghue, M.J. (2005) A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution*, **59**, 2299–2311.
- Reece, J., Bowen, B., Smith, D. & Larson, A. (2011) Comparative phylogeography of four Indo-Pacific moray eel species (Muraenidae) reveals comparable ocean-wide genetic connectivity despite five-fold differences in available adult habitat. *Marine Ecology Progress Series*, **437**, 269–277.
- Renema, W., Bellwood, D.R., Braga, J.C., Bromfield, K., Hall, R., Johnson, K.G., Lunt, P., Meyer, C.P., McMonagle, L., Morley, R.J., O’Dea, A., Todd, J.A., Wesselingh, F.P., Wilson, M.E.J. & Pandolfi, J.M. (2008) Hopping hotspots: global shifts in marine biodiversity. *Science*, **321**, 654–657.
- Roberts, C.M., McClean, C.J., Veron, J.E.N., Hawkins, J.P., Allen, G.R., McAllister, D.E., Mittermeier, C.G., Schueler, F.W., Spalding, M., Wells, F., Vynne, C. & Werner, T.B. (2002) Marine biodiversity hotspots and conservation priorities for tropical reefs. *Science*, **295**, 1280–1284.
- Robertson, D.R., Grove, J.S. & McCosker, J.E. (2004) Tropical transpacific shore fishes. *Pacific Science*, **58**, 507–565.
- Rocha, L.A. & Bowen, B.W. (2008) Speciation in coral-reef fishes. *Journal of Fish Biology*, **72**, 1101–1121.
- Rocha, L.A., Robertson, D.R., Roman, J. & Bowen, B.W. (2005) Ecological speciation in tropical reef fishes. *Proceedings of the Royal Society B: Biological Sciences*, **272**, 573–9.
- Ronquist, F. (1997) Dispersal–vicariance analysis: a new approach to the quantification of historical biogeography. *Systematic Biology*, **46**, 195–203.
- Rosen, B.R. (1981) The tropical high diversity enigma – the corals’-eye view. *Chance, change and challenge: the evolving biosphere* (ed. by P.L. Forey), pp. 103–129. British Museum (Natural History), London, and Cambridge University Press, Cambridge.
- Rosen, B.R. (1984) Reef coral biogeography and climate through the late Cainozoic: just islands in the sun or a critical pattern of islands? *Fossil and climate* (ed. by P. Brenchley), pp. 201–262. Geological Journal Special Issue, No. 11. John Wiley and Sons, Chichester.
- Santini, F. & Winterbottom, R. (2002) Historical biogeography of Indo-western Pacific coral reef biota: is the Indonesian region a centre of origin? *Journal of Biogeography*, **29**, 189–205.
- Santini, F., Harmon, L.J., Carnevale, G. & Alfaro, M.E. (2009) Did genome duplication drive the origin of teleosts? A comparative study of diversification in ray-finned fishes. *BMC Evolutionary Biology*, **9**, 194.
- Smith, S.A. (2009) Taking into account phylogenetic and divergence-time uncertainty in a parametric biogeography-

- ical analysis of the Northern Hemisphere plant clade Caprifoliaceae. *Journal of Biogeography*, **36**, 2324–2337.
- Springer, V.G. & Williams, J.T. (1994) The Indo-West Pacific blennioid fish genus *Istiblennius* reappraised: a revision of *Istiblennius*, *Blenniella*, and *Paralticus*, new genus. *Smithsonian Contributions to Zoology*, **565**, 1–193.
- Steininger, F.F. & Rögl, F. (1979) The Paratethys history. A contribution towards the Neogene geodynamics of the alpine orogene. *Annales Geologiques des Pays Helleniques*, **3**, 1153–1165.
- Timm, J., Figiel, M. & Kochzius, M. (2008) Contrasting patterns in species boundaries and evolution of anemonefishes (Amphiprioninae, Pomacentridae) in the centre of marine biodiversity. *Molecular Phylogenetics and Evolution*, **49**, 268–76.
- Westneat, M.W. & Alfaro, M.E. (2005) Phylogenetic relationships and evolutionary history of the reef fish family Labridae. *Molecular Phylogenetics and Evolution*, **36**, 370–390.
- Williams, S.T. & Duda, T.F. (2008) Did tectonic activity stimulate Oligo-Miocene speciation in the Indo-West Pacific? *Evolution*, **62**, 1618–1634.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Supplemental tables. Table S1 shows species richness for each family by clade and by region with percentage of taxa sampled in the present study. Table S2 shows areas incorporated in five major biogeographic regions. Table S3 shows presences/absences of fish species in major marine biogeographic regions. Table S4 gives an overview of DEC model results for likelihood comparisons.

Appendix S2 Supplemental figures. Figure S1 illustrates regional inheritance scenarios considered in the DEC model implemented in LAGRANGE. Figure S2 illustrates DEC model constraints used in LAGRANGE analyses. Figures S3–S5 show resulting biogeographic reconstructions overlaid on chronograms of the Labridae, Pomacentridae and Chaetodontidae (modified from Cowman & Bellwood, 2011).

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

BIOSKETCHES

Peter Cowman is a researcher at James Cook University. His interests include vertebrate evolution and genomic research. His current research focuses on the use of bioinformatics methods to study the evolution of reef fishes and associated biodiversity patterns.

David Bellwood is a Professor of Marine Biology at James Cook University. His broad interests include the ecology and evolution of reef fishes. He is particularly interested in the role of fishes in ecosystem function and the impact of history and biodiversity on reef processes. His current research programme examines the origins, patterns and consequences of variation in global coral reef biodiversity.

Editor: Luiz A. Rocha

Evolutionary history of the butterflyfishes (f: Chaetodontidae) and the rise of coral feeding fishes

D. R. BELLWOOD*†, S. KLANTEN*‡, P. F. COWMAN*†, M. S. PRATCHETT†, N. KONOW*§ & L. VAN HERWERDEN*‡

*School of Marine and Tropical Biology, James Cook University, Townsville, Qld, Australia

†Australian Research Council Centre of Excellence for Coral Reef Studies, James Cook University, Townsville, Qld, Australia

‡Molecular Evolution and Ecology Laboratory, James Cook University, Townsville, Qld, Australia

§Ecology and Evolutionary Biology, Brown University, Providence, RI, USA

Keywords:

biogeography;
chronogram;
coral reef;
innovation;
molecular phylogeny;
trophic novelty.

Abstract

Of the 5000 fish species on coral reefs, corals dominate the diet of just 41 species. Most (61%) belong to a single family, the butterflyfishes (Chaetodontidae). We examine the evolutionary origins of chaetodontid corallivory using a new molecular phylogeny incorporating all 11 genera. A 1759-bp sequence of nuclear (S7I1 and ETS2) and mitochondrial (cytochrome *b*) data yielded a fully resolved tree with strong support for all major nodes. A chronogram, constructed using Bayesian inference with multiple parametric priors, and recent ecological data reveal that corallivory has arisen at least five times over a period of 12 Ma, from 15.7 to 3 Ma. A move onto coral reefs in the Miocene foreshadowed rapid cladogenesis within *Chaetodon* and the origins of corallivory, coinciding with a global reorganization of coral reefs and the expansion of fast-growing corals. This historical association underpins the sensitivity of specific butterflyfish clades to global coral decline.

Introduction

Coral reef fishes are a highly diverse group, with an evolutionary history extending back more than 50 Myr (Bellwood & Wainwright, 2002). From the fossil record, it appears that scleractinian-dominated coral reefs and modern coral reef fish families first appeared and then diversified at approximately the same time, in the early Cenozoic (Bellwood, 1996; Bellwood & Wainwright, 2002; Wallace & Rosen, 2006). This suggests that the origins of modern coral reefs and their associated fish families may be closely linked. However, it is remarkable that of the 5000 or more fish species recorded from coral reefs today only 128 eat corals (Cole *et al.*, 2008; Rotjan & Lewis, 2008) and just 41 are believed to feed directly on scleractinian corals as their primary source of nutrition. Moreover, 61% (25 of 41) belong to a single family, the

butterflyfishes (f. Chaetodontidae); of the remainder most (eight) are in the Labridae. Why so few species have been able to exploit such a widespread resource remains a mystery. It also highlights the exceptional abilities of the few corallivores that have managed to subsist on corals, and the remarkable status of butterflyfishes. Despite being one of the most intensively studied families of reef fishes, the evolution of this highly specialized feeding mode remains poorly understood. In particular, how many times has corallivory arisen within the group and when did this unusual feeding mode first arise? Did corallivory arise along with major coral groups in the early Eocene?

Butterflyfishes are conspicuous and iconic inhabitants of coral reef environments. The family contains over 130 species with representatives in all coral reef regions (Allen *et al.*, 1998; Kuitert, 2002). Their colourful patterns, and ease of identification and observation have ensured that the behavioural, ecological, morphological and biogeographic characteristics of butterflyfishes have been extensively studied (e.g. Motta, 1988; Ferry-Graham *et al.*, 2001a; Findley & Findley, 2001; Pratchett,

Correspondence: David R. Bellwood, School of Marine and Tropical Biology, James Cook University, Townsville, Qld 4811, Australia. Tel.: +61 7 4781 4447, fax: +61 7 4725 1570; e-mail: david.bellwood@jcu.edu.au

2005). Indeed, they have been regularly identified as indicator species of reef health (Reese, 1975; Roberts *et al.*, 1988; Kulbicki *et al.*, 2005). It is this close association with corals and coral reefs that stands as one of the most important features of this family. With almost a quarter of the species feeding on corals, they have what is arguably the closest association of any fish group with coral reefs. The key to understanding the history of this relationship, however, is to obtain well-supported phylogenies based on multiple genes, and to use robust molecular dating methodologies, informed by reliable fossil data, to provide a temporal framework in which to interpret recent ecological evidence.

Fossil evidence of increasing reef–fish interactions points to a major change between the late Mesozoic and the beginning of the Tertiary (Bellwood, 2003). Although there is a diverse range of acanthomorph fishes in the late Cretaceous, the earliest evidence of the vast majority of extant reef fish families, for which there is a fossil record, is from the Eocene. Most of these families are first recorded from the 50-Myr-old deposits of Monte Bolca, in northern Italy (e.g. Blot, 1980; Bellwood, 1996; although it should be noted that a few molecular studies have suggested that some lineages may predate the Cretaceous/Tertiary (K/T) boundary, e.g. Streebman *et al.*, 2002; Alfaro *et al.*, 2007; Azuma *et al.*, 2008). The Monte Bolca deposits mark the first modern coral reef fish assemblage in terms of both its taxonomic composition and the functional attributes of the component taxa (Bellwood, 1996), and it is here that we observe the first evidence of increased interactions between fishes and the benthos, with the appearance of several lineages of fishes that were almost certainly grazing herbivores (Bellwood, 2003). Although there have been many Eocene fossils ascribed to the Chaetodontidae, a recent evaluation of this material has rejected all of these taxa; there is no reliable record of the family from the Eocene (Bannikov, 2004). The oldest reliable fossil evidence for the Chaetodontidae is of Miocene age (Carnevale, 2006). However, with a robust molecular phylogeny we can build on the fossil record and place recent ecological advances in an evolutionary framework.

Recent ecological research has provided a new perspective on the nature of corallivory in reef fishes, especially butterflyfishes. Building on existing information (e.g. Birkeland & Neudecker, 1981; Blum, 1988; Motta, 1988) recent studies have provided a detailed understanding of the family in terms of feeding morphology and kinematics (Ferry-Graham *et al.*, 2001a,b; Konow *et al.*, 2008; Konow & Ferry-Graham, in press), feeding strategies and behavioural interactions (Zekeria *et al.*, 2002; Gregson *et al.*, 2008) and, most importantly, the nature and extent of corallivory (Pratchett *et al.*, 2004; Berumen *et al.*, 2005; Pratchett, 2005, 2007; Cole *et al.*, 2008; Rotjan & Lewis, 2008). It is now known that butterflyfishes exhibit considerable diversity in the nature of corallivory, underpinned by both ecomorpho-

logical and behavioural variation. This includes at least two different modes of coral feeding, exploitation of both soft and hard corals, and a spectrum of coral feeding specializations ranging from highly specialized obligate coral feeders that primarily target just one coral species to facultative and generalist coral feeders that can feed on a wide range of coral species (Pratchett *et al.*, 2004; Berumen *et al.*, 2005).

The evolutionary history of these traits is poorly understood. Existing morphological and molecular phylogenies have produced discordant tree topologies, opening questions about the nature of character evolution. Most molecular analyses of the family have focused on relationships between species pairs or species complexes (e.g. McMillan & Palumbi, 1995; Hsu *et al.*, 2007) or have included butterflyfishes as part of broader studies of putative sister taxa (Bellwood *et al.*, 2004). The most comprehensive analysis to date using 3332 bp of mitochondrial and nuclear DNA yielded a well-supported phylogeny, which differed significantly from all previous topologies, and underpinned a thorough evaluation of the systematics and taxonomy of the family (Fessler & Westneat, 2007).

The evolutionary and ecological ramifications of these relationships for corallivory, however, have yet to be fully explored. We address four critical questions: (1) How many times has corallivory arisen? (2) When did it first arise? (3) Did corallivory and/or a move onto coral reefs underpin diversification within the family? and (4) What are the broader implications for the evolution of coral reefs? In this present study, we use a comprehensive new molecular phylogeny to explore the evolutionary history of the family. For the first time, we use molecular methods to examine the relationships between all 10 genera and all respective subgenera (with the exception of *Roa Roa sensu*, Blum, 1988; = *Roa*, Kuitert, 2002). We examined 56 species using two nuclear (ETS2 and S711) and a mitochondrial marker (*cyt b*); the nDNA markers have not been used previously in this family. We also apply, for the first time, current Bayesian analyses for chronogram construction with multiple fossil calibrations. Using this chronogram, ecological evidence and diversification statistics, we explore the evolutionary origins of corallivory and the timing of this highly unusual fish–coral interaction. We then consider the basis for these trophic innovations in the context of the evolution of coral reefs.

Materials and methods

Taxon sampling

In total, 56 butterflyfish species were examined, with multiple representatives of all 11 identified butterflyfish genera and 12 subgenera (authorities given in Allen *et al.*, 1998). Specimens were obtained using spears or nets with additional material obtained from the

ornamental fish trade. A further eight species from the Pomacanthidae, Ehippidae, Kyphosidae and Scatophagidae were included as putative outgroup species. Based on Smith & Wheeler (2006), the Pomacanthidae then the Ehippidae are the putative successive sister taxa to the Chaetodontidae; the more distant kyphosid and microcanthid species were used to root the entire phylogeny. Our taxon sampling specifically included 18 of the 26 species known to be obligate coral feeders (Appendix S1; Table 1). Full taxon sampling in several lineages (*Forcipiger*, *Chelmon* and *Chelmonops*) also enables us to explore the temporal and biogeographic patterns of species origins.

Laboratory procedures

Total DNA was extracted from tissues using standard salt-chloroform and proteinase K digestion extraction procedures (Sambrook & Russell, 2001). Two nuclear genes, ETS2 (ETS is a transcription factor important in cell proliferation; Dwyer *et al.*, 2007; Lyons *et al.*, 1997); S7 Intron 1 (S7 is a ribosomal protein required for assembling 16s RNA; Maguire & Zimmermann, 2001; Chow & Hazama, 1998) and the mitochondrial protein-coding region, cytochrome *b* (which participates in electron transport; Kocher *et al.*, 1989; Irwin *et al.*, 1991; McMillan & Palumbi, 1995) were used to explore the evolutionary relationships among the butterflyfishes (Appendix S1; Table 2). An average of two specimens were sequenced for each species. Each 20 μ L polymerase chain reaction (PCR) volume contained 2.5 mM Tris-Cl (pH 8.7), 5 mM KCl(NH₄)₂SO₄, 200 μ M each dNTP, MgCl₂ ranging from 1.5 to 4 mM, 10 μ M each primer, 1 U of Taq Polymerase (Qiagen, Doncaster, Victoria, Australia) and 10 ng template DNA. Amplifications followed the same basic cycling protocol: an initial

denaturing step of 2 min at 94 °C, followed by 35 cycles, with the first five cycles at 94 °C for 30 s, 30 s at primer-specific annealing temperatures (T_a) (Appendix S1; Table 2), followed by 1 min 30 s extensions at 72 °C and the remaining 30 cycles were performed as before, but at $T_a - 2$ °C. PCR products were purified by isopropanol precipitation (cyt *b* and S7I1) or gel-purification on 2% agarose gels, as two bands appeared routinely (ETS2). This was also the case for S7I1 amplified fragments of some species. A 500-bp fragment was retained for ETS2 whilst a 700-bp fragment was retained for S7I1. Gel-excised fragments were purified in a column following manufacturer's protocols (Qiagen). Purified templates were quantified by UV-Vis absorbance (ND-1000 Spectrophotometer, NanoDrop®, Wilmington, NC, USA) and sent to Macrogen Inc. (Seoul, South Korea) for direct sequencing in both directions.

Analytical procedures

Data compilation

The consensus sequence of the multiple specimens sequenced was used to represent each taxon. Sequences were edited using Sequencher 4.5 (Gene codes corporation, Ann Arbor, MI, USA), and automatically aligned using CLUSTALX (Thompson *et al.*, 1997) and finally manually corrected using SE-AL version 2.0 available at <http://evolve.zoo.ox.ac.uk> (Rambaut, 1996). Sequences of this study are available at GenBank accession numbers (ETS2: FJ167730–FJ167792; S7I1: FJ167793–FJ167846, FJ167848–FJ167856 and cyt *b*: FJ167682–FJ167709, FJ167711–FJ167719, FJ167721–FJ167729). Several sequences of cytochrome *b* were used from GenBank (*Chaetodon*: *C. argentatus* AF108580, *C. citrinellus* AF108585, *C. kleinii* AF108591, *C. lineolatus* AF108593, *C. lunula*

Table 1 Departure of chaetodontid lineages from global diversification rate estimated of the family Chaetodontidae.

Clade name	Age	Total	$\epsilon = 0$	$\epsilon = 0.3$	$\epsilon = 0.5$	$\epsilon = 0.6$	$\epsilon = 0.8$	$\epsilon = 0.9$
Crown group (CT)	32.8	130(56)	$r_G = 0.1274$	$r_G = 0.1246$	$r_G = 0.1188$	$r_G = 0.1141$	$r_G = 0.0971$	$r_G = 0.0787$
CF	26.1	27(16)	7.61E-01	6.81E-01	6.61E-01	6.60E-01	6.76E-01	7.00E-01
CP	23.9	103(40)	4.23E-02	7.27E-02	9.32E-02	1.05E-01	1.42E-01	1.81E-01
AC	22.9	12(8)	8.84E-01	7.95E-01	7.76E-01	7.75E-01	7.94E-01	8.18E-01
C1*	17.8	3(1)	8.04E-01	8.49E-01	8.76E-01	8.89E-01	9.19E-01	9.38E-01
CH	17.8	93(38)	5.03E-04*	2.89E-03*	7.39E-03	1.16E-02	3.13E-02	6.29E-02
FH	14.2	16(8)	2.69E-01	3.01E-01	3.29E-01	3.50E-01	4.31E-01	5.19E-01
C3	13.4	21(13)	9.94E-02	1.44E-01	1.80E-01	2.04E-01	2.88E-01	3.83E-01
C2	12.6	37(14)	3.15E-03	1.14E-02	2.39E-02	3.46E-02	8.15E-02	1.49E-01
C4	11.3	31(11)	3.10E-03	1.12E-02	2.39E-02	3.50E-02	8.43E-02	1.57E-01
Ho	7.8	8(3)	2.01E-01	2.46E-01	2.81E-01	3.05E-01	3.96E-01	4.99E-01
PR	5.5	10(2)	2.07E-02	4.37E-02	7.08E-02	9.13E-02	1.69E-01	2.67E-01
C2 + C3 + C4	16.6	89(38)	1.61E-04*	1.26E-03*	3.90E-03	6.71E-03	2.21E-02	5.00E-02
C3 + C4	15.7	52(24)	5.44E-03	1.69E-02	3.14E-02	4.28E-02	8.81E-02	1.49E-01

Bold *P*-values highlight significantly higher species richness in subtending clade than expected under the global rate of cladogenesis (*significance after Bonferroni correction). ϵ is the extinction rate, r_G is the estimated global Chaetodontidae speciation rate conditional on the extinction rate. Clade names and ages are taken from the node labels and mean node heights of Fig 1, and table 1 in Appendix S2.

AF108594, *C. meyeri* AF108597, *C. milliaris* U23606, *C. multicinctus* U23588, *C. ornatissimus* AF108600, *C. plebius* AF108602, *C. quadrimaculatus* AJ748302, *C. unimaculatus* AJ748304; *Chelmon rostratus* AF108612, *Coradion altivelis* AF108613, *Coradion chrysozonus* AF108614, *Forcipiger flavissimus* AF108615, *Hemitaurichthys polylepis* AF108616, *Heniochus acuminatus* AF108618 and *Parachaetodon ocellatus* AF108622 as per Littlewood *et al.*, 2004; McMillan & Palumbi, 1995; Nelson *et al.*, unpublished GenBank submission). Each locus was first examined for saturation using DAMBE version 5.010 (Xia & Xie, 2001). This method employs an entropy-based index of substitution saturation (Iss); if Iss is significantly larger than the critical Iss (i.e. Iss.c), sequences have experienced substitution saturation (Xia *et al.*, 2003). Prior to concatenating gene regions, each gene was partitioned based on its function or structure. Coding genes [*cyt b* and a short exon region (99 bp) of *ETS2*] were partitioned according to codon positions (1st and 2nd combined as conserved region, and 3rd codon separately). Nuclear introns (*ETS2* and *S711*) were partitioned into putative stem (conserved) and loop (hypervariable) regions. Eight separate gene partitions were identified in total.

Phylogenetic analyses

Maximum parsimony (MP) analyses were implemented in PAUP* 4.0b10 (Swofford, 1998) using heuristic search methods with 1000 pseudo-replicate bootstraps, tree-bisection-reconnection branch swapping and random addition of taxa. Two separate heuristic MP runs were performed. First, all sites were treated equally and second, sites were weighted according to gene partitions [3rd codon, loop regions = 1; conserved (1st and 2nd codon) and stem regions = 2]. A 50% majority rule consensus tree was generated from all shortest trees obtained. Bayesian inference (BI) analyses were implemented in MrBayes version 3.1.2 (Huelsenbeck & Ronquist, 2001) using James Cook University's HPC GridSphere system (<https://ngportal.hpc.jcu.edu.au/gridsphere/>). The analysis of the combined data used a partition mix model method (pMM) according to gene partitions with locus-specific substitution models, using MRMODELTEST version 2.2 (Nylander, 2004) and Akaike information criterion (AIC) (Nylander *et al.*, 2004). Two Bayesian pMM analyses were performed using Markov chain Monte Carlo (MCMC) simulations with four chains of 2 000 000 generations each, sampling trees every 100 generations. Stationarity was reached after 10 000 generations and a 50% majority rule consensus tree was computed using the best 16 000 post-burn-in trees from each run. Six putative sister taxa were included in the analyses, three pomacanthids (*Pomacanthus annularis*, *P. rhomboids* and *P. sexstriatus*), an ephippid (*Platax orbicularis*), and two scats (*Scatophagus argus* and *Selonotoca multifasciata*) and, in addition, two distant outgroups, a kyphosid (*Kyphosus vaigiensis*) and a microcanthid (*Tilodon sexfasciatum*), which were used to root resulting trees. The single best tree was selected for molecular dating.

Maximum likelihood (ML) analysis was performed using Garli version 0.95 (Zwickl, 2006). Ten independent runs were performed using the best substitution model (as per AIC) for the combined data (not partitioned) implemented with MODELTEST version 3.7 (Posada & Crandall, 1998). The best trees from the individual runs were compared with ensure they did not differ in topology and that the ML search was not arriving in a suboptimal area of tree space. In addition, a ML analysis with 100 bootstrap replicates was performed to show support of individual clades in the tree.

Molecular dating

Age estimation of the chaetodontid lineages was performed in the program BEAST v1.4.8 (Drummond & Rambaut, 2007). BEAST implements BI and a MCMC analysis to simultaneously estimate branch lengths, topology, substitution model parameters and dates based on fossil calibrations. It also does not assume substitution rates are autocorrelated across lineages, allowing the user to estimate rates independently from an uncorrelated exponential distribution or lognormal distribution (UCLD). Many empirical data sets have been shown not to demonstrate autocorrelation of rates and times (Drummond *et al.*, 2006; Alfaro *et al.*, 2007; Brown *et al.*, 2008). An initial ultrametric tree was constructed in r8s 1.71 (Sanderson, 2004) from the topology and branch lengths of the best Bayesian tree recovered from phylogenetic analyses, using a penalized likelihood (PL) method (Sanderson, 2002). This topology was used for calibration purposes by using parametric priors implemented in BEAST to make assumptions *a priori* based on fossil and biogeographic data (see below).

The vast majority of reported fossil chaetodontids are demonstrably erroneous (Bannikov, 2004), a pattern common in many other reef groups (e.g. scarids, Bellwood & Schultz, 1991; pomacentrids, Bellwood & Sorbini, 1996). Fossil selection is critical, and fossils were only used if placed in a family based on reliable morphological criteria. Fossil calibrations were therefore restricted to fossils from two families recorded from the Eocene (50 Ma) deposits of Monte Bolca: *Eoplatax papilio* (Ephippidae) (Blot, 1969) and *Eoscatophagus frontalis* (Scatophagidae) (Tyler & Sorbini, 1999). With both putative ephippid and scatophagid fossils having a minimum age of 50 Ma, an exponential prior was placed on the *Platax* node (PL) with a hard lower bound age of 50 Ma and a 95% soft upper bound of 65 Ma. The exponential prior reflects the decreasing probability of a lineage being older than its oldest fossil (Yang & Rannala, 2006; Ho, 2007). The soft upper bound of 65 Ma representing the transition of fish faunas at the K/T boundary (following Bellwood & Wainwright, 2002; Bellwood *et al.*, 2004; Fessler & Westneat, 2007) beyond which there is no fossil record of modern reef fish families.

BEAST MCMC runs of 10×10^6 generations were performed assuming the UCLD model with eight

unlinked data partitions and unlinked substitution models specified by MRMODELTEST v2.2 (Nylander, 2004). Ten independent analyses were run sampling every 500th generation. Resulting log files were examined using Tracer v1.4 (Rambaut & Drummond, 2007) to ensure all analyses were converging on the same area in tree space. Tree files (approx. –10% burn-in) were then combined using LogCombiner (Rambaut & Drummond, 2007) and compiled into a maximum clade credibility chronogram to display mean node ages and highest posterior density (HPD) intervals at 95% (upper and lower) for each node.

Optimizing ecological traits

Two ecological traits (corallivory and habitat use) were mapped to the best phylogenetic tree that was the basis for dating diversification in the chaetodontids, using Mesquite v 2.6 (Maddison & Maddison, 2007). Habitat use traits were scored as 0 = not on reefs, 1 = rocky reefs and 2 = coral reefs. Diet traits were scored as 0 = non-corallivore, 1 = omnivores (that have < 1% coral in the diet and feed on other invertebrates and/or algae), 2 = facultative or occasional coral feeders [which include some (1–80%) hard coral in the diet] and 3 = corallivores (in which the diet is dominated by i.e. > 80% hard or soft corals). Ecological character states were drawn from the published literature (Appendix S1; Table 1). Within the corallivores, species are further identified as obligate hard or soft corallivores when they feed exclusively on a specific coral type.

Diversification rates

All diversification statistics were performed in R version 2.7.2 (<http://www.Rproject.org>) (Ihaka & Gentleman, 1996) using functions written for GEIGER (Harmon *et al.*, 2008), LASER (Rabosky, 2006) and associated packages. The constant rates (CR) test of Pybus & Harvey (2000) was used to investigate the rates of cladogenesis of the chaetodontid crown group. This test estimated the gamma statistic of the BEAST generated chronogram. Significantly negative gamma values (< –1.645, one-tailed test) indicate a decrease in the rates of cladogenesis over time. This implies that internal nodes of the tree are distributed closer to the root than would be expected under a Yule (pure birth) process. To account for incomplete taxon sampling (which increases Type 1 error of the CR test; Pybus & Harvey, 2000) a Markov chain CR (MCCR) test (Pybus & Harvey, 2000) was used to compare the observed gamma to that of the null distribution created from 10 000 randomly subsampled, simulated (full) topologies under a Yule process. The relative cladogenesis statistic (Nee *et al.*, 1992) was used to identify lineages with significantly faster/slower rate of cladogenesis. These methods have previously been used to investigate diversification rates in tetraodontiform lineages (Alfaro *et al.*, 2007).

Methods implemented in the tetraodontiform study (Alfaro *et al.*, 2007) were used to calculate the global diversification rate (r_G) of the chaetodontids across extinction rates (ϵ) in increments of 0.1 from 0 to 0.9 (see Magallon & Sanderson, 2001). Using functions in GEIGER (based on the method of moments estimator of Magallon & Sanderson, 2001) the probabilities of the observed species richness in each of the major chaetodontid lineages were calculated using crown group ages and the global estimates of diversification rate (r_G) for each increment of extinction. In case of the *C. robustus* lineage, no other reported taxa in this clade were included in this study and thus the stem group age estimator was used (equation 10a, Magallon & Sanderson, 2001; see Alfaro *et al.*, 2007) to calculate the above probability.

Results

Sequence variability

We examined 1759 bp of sequence of which approximately 50% was parsimony-informative. The two nuclear markers, ETS2 and S7I1 had 647 and 655 bp respectively, cytochrome *b* contributed a further 426 bp with (47%, 65% and 47% parsimony-informative sites respectively). None of the individual gene regions were saturated (Iss 0.43 < 0.8 Iss.c, Iss 0.3 < Iss.c 0.8 and Iss 0.3 < Iss.c 0.78 respectively), neither was the concatenated data (Iss 0.4 < 0.8 Iss.c).

Model selection

The gene-specific models (AIC) for each of the eight gene partitions used for Bayesian analysis were as follows: Cyt *b* conserved region (gene 1, 1st and 2nd codons) required a GTR + G model (gamma shape parameter = 0.2397) with substitution Nst = 6, for its 3rd codon region (gene 2) a GTR + I + G model (invariable sites = 0.02, γ = 3.9890) with substitution Nst = 6. ETS2 coding region required for its conserved (gene 3) region a K80 model with substitution Nst = 2 (ti/tv ratio = 2.5419), and its variable (gene 4, 3rd codon) region a HKY model with substitution Nst = 2 (ti/tv ratio = 1.5633). Both ETS2 stem (gene 5) and loop (gene 6) required a GTR + G model (γ = 0.9426 and 1.2595 respectively) with Nst = 6 substitution classes. S7I1 stem (gene 7) region required a HKY + G model (γ = 1.5603) with substitution Nst = 2 (ti/tv ratio = 1.3058) and its loop regions (gene 8) a GTR + G model (γ = 5.4651) with substitution Nst = 6. The model selections for pMM BI only requires a general 'form' of the model (Nylander, 2004), as the Markov chain integrates uncertainties of the parameter values. Therefore, seven of the eight gene partitions had a base frequency = dirichlet (1,1,1,1) (i.e. unequal) while the eight gene (gene 3, ETS2 coding, conserved region) base frequency was set to = fixed (equal).

Maximum likelihood analysis (GARLI) required an overall GTR + G model ($\gamma = 0.5120$) for all regions, substitution $Nst = 6$ with substitution rates fixed (0.7160, 2.3497, 0.8239, 0.6913 and 3.8464), and base frequencies fixed (0.26, 0.2336, 0.2115 and 0.2949).

Tree inference

Stationarity of the Bayesian analyses was reached after much fewer than 10 000 generations in both runs, (visualized in Tracer version 1.4; Rambaut & Drummond, 2007) and the 50% majority rule consensus tree topology was no different from the best trees of each run ($-\ln L = -22\ 255.116$ 1st run and $-\ln L = -22\ 254.149$ 2nd run) with very high posterior probabilities (Fig. 1). Both MP and ML analyses inferred the same tree topology as per Bayesian analysis. We therefore included only the support for each retrieved node (Fig. 1). Four major clades of *Chaetodon* were retrieved, resembling closely the four clades retrieved in a previous molecular study by Fessler & Westneat (2007). Although identical species were not analysed in the two studies, the placement of species common to both studies was identical, despite the use of different loci in the two studies. For clarity, in *Chaetodon* we follow the four clades of Fessler & Westneat (2007). Old taxonomic groupings were found to be of limited utility (only one remains intact and retains its traditional boundaries (*Radophorus* in clade 4) and we will not consider them further within *Chaetodon*. Fessler & Westneat (2007) provide a thorough evaluation of the taxonomy of the family. The only additional detail from our study is that *Parachaetodon* would make *Chaetodon* (and the subgenus *Discochaetodon*) paraphyletic and *Parachaetodon* is probably best placed within *Chaetodon* (as a junior synonym).

Molecular dating

The best Bayesian topology and branch lengths received by phylogenetic analyses (Fig. 1) was used as the initial starting tree with an exponential prior used to calibrate the PL node (see Materials and methods). BEAST log files analysed in Tracer showed convergence between independent runs in tree space. High effective sample size scores of individual parameters indicated valid estimates based on independent samples from the posterior distribution of the MCMC. A maximum clade credibility chronogram was compiled in Tree Annotator from 180 000 post-burn-in trees (9×10^7 generations from 10 BEAST MCMC runs). The chronogram displays mean node heights received at each node by BEAST MCMC with bars representing 95% HPD (Appendix S2, Fig. 1). The family Chaetodontidae dates back to the early Eocene where it split from pomacanthids with a mean age of the most recent common ancestor (MRCA) of 50.1 Ma (41.5–60.7, 95% HPD). Estimated ages indicate the origin of the butterfly fish and bannerfish clades at a

mean age of 32.8 Ma (24.9–40.9, 95% HPD) after which they rapidly diversified, with the four major *Chaetodon* lineages in place by the mid Miocene (Appendix S2; Table 1). Also during the early Miocene we see the origins of the three major divisions within the bannerfish clade.

Optimization of ecological traits

Based on the species examined it is clear that corallivory has arisen on at least five separate occasions (Fig. 2). Corallivory has been reported in 25 chaetodontid species (Appendix S1; Table 1). All are in the reef-butterflyfishes clade and are restricted to a single monophyletic genus, *Chaetodon*. Of these 25 corallivores, 17 are included in the current phylogeny. The remaining eight species are easily included in the four main *Chaetodon* clades based on previous phylogenetic and taxonomic evidence (Appendix S1; Table 1).

Chaetodon clade 1 contains only three species, all are restricted to West African coastal waters, with no record of coral feeding. *Chaetodon* clade 2 (37 species) contains three distinct lineages of coral feeders (estimated MRCA to omnivorous sister taxa in parentheses): *C. quadrimaculatus* (3.2 Ma), the *C. multicinctus* clade (4.9 Ma) and the *C. unimaculatus-interruptus* clade (4.3 Ma). The first two lineages are hard coral feeders and probably have an obligate dependence on corals. *Chaetodon unimaculatus* and *C. interruptus* feed on soft and hard corals. *Chaetodon* clade 3 is predominantly corallivorous, with 19 of the 21 species being obligate corallivores. Of the remaining species, the diet of *C. tricinctus* is unknown, leaving one noncorallivore, *Parachaetodon ocellatus*. The chronogram places the origins of corallivory at about 15.7 Ma. Of the 31 species in *Chaetodon* clade 4 only two are corallivores: *C. melannotus* and *C. ocellicaudus*. These sister species are both obligate soft coral feeders (Appendix S1; Table 1). They separated from their omnivorous sister at about 9.8 Ma.

Diversification rates

The relative cladogenesis statistics identified the *Chaetodon* lineage as having a significantly different rate of cladogenesis than its sister lineage. Both CR and MCCR tests showed no evidence for a slowdown in the rate of cladogenesis through time for the family Chaetodontidae ($\gamma = -1.248$, MCCR adjusted $P = 0.55$). As noted by Magallon & Sanderson (2001) the estimates of r_G decreased with increasing extinction rates (Table 1). The *Chaetodon* clade (CH) with all subtending lineages showed significantly higher species diversity than expected given the global diversification rate up to 90% extinction rate (adjusted $P = 0.03$; $\alpha = 0.9$). Clades 2 and 4 (C2 and C4) showed significantly higher species diversity than expected (up to 80% extinction), however, the corallivorous clade 3 (C3) is not significantly more diverse than expected given the crown diversification

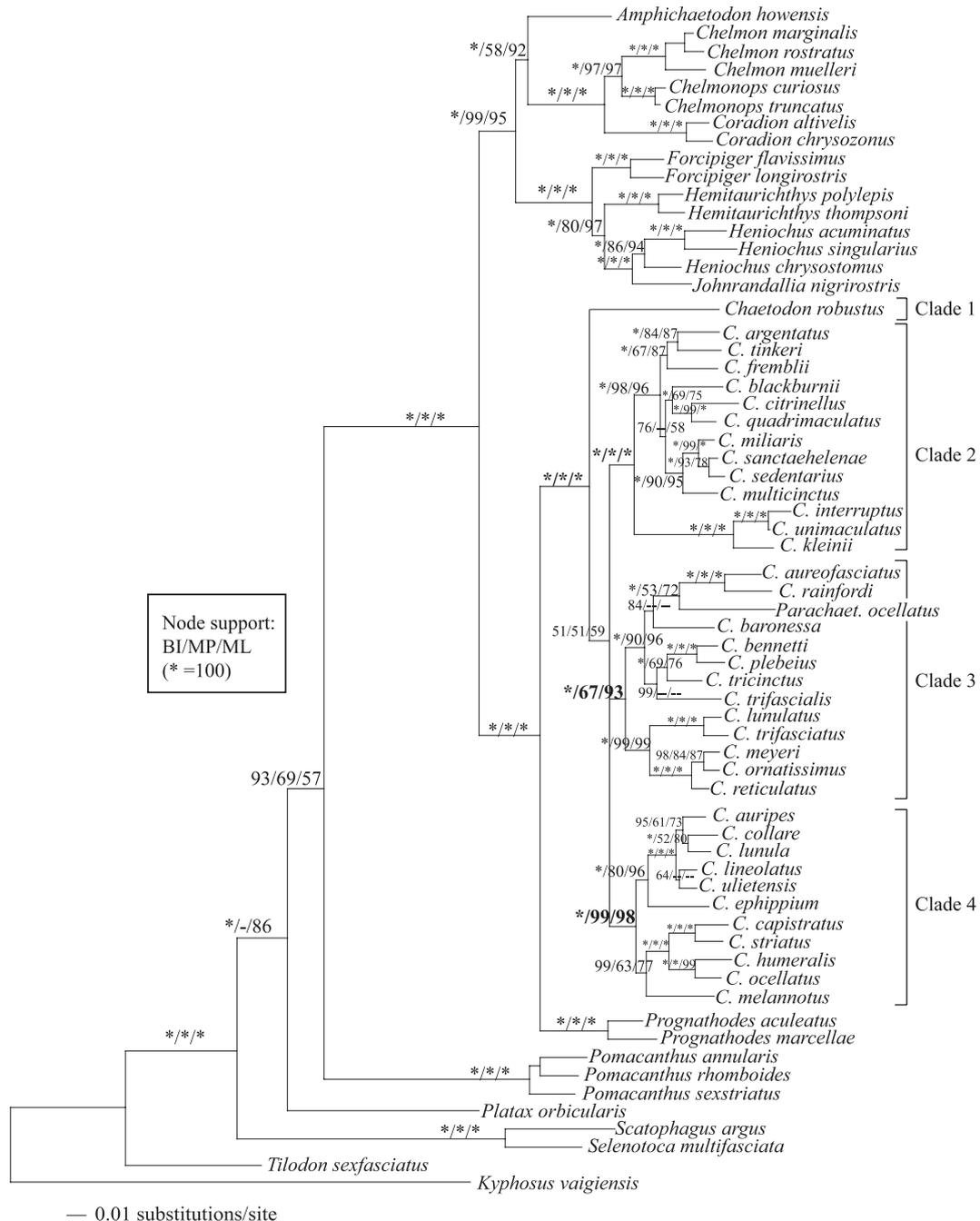


Fig. 1 Inferred phylogeny of the butterflyfish and bannerfish (f. Chaetodontidae), based on 56 species with representatives from all 11 genera and 12 subgenera, obtained by Bayesian, maximum parsimony (MP) and maximum likelihood analyses for three loci (ETS2, S711 and *cyt b*). The topology shows the best bayesian tree with posterior probabilities (consensus of 32 000 trees) and bootstrap support (> 50%) of MP and ML (1000 and 100 bootstrap replicates respectively). (*) 100% support. (–) no bootstrap support. The tree was rooted with *Kyphosus vaigiensis* and *Tilodon sexfasciatus*.

rate even in the absence of extinction. The *Prognathodes* lineage also shows significantly higher diversification for up to 30% extinction ($P = 0.044$). If using a Bonferroni

correction (adjusted $P = 0.0038$) the Chaetodon (CH and C2 + 3 + 4) clade still remains significant at low extinction rates.

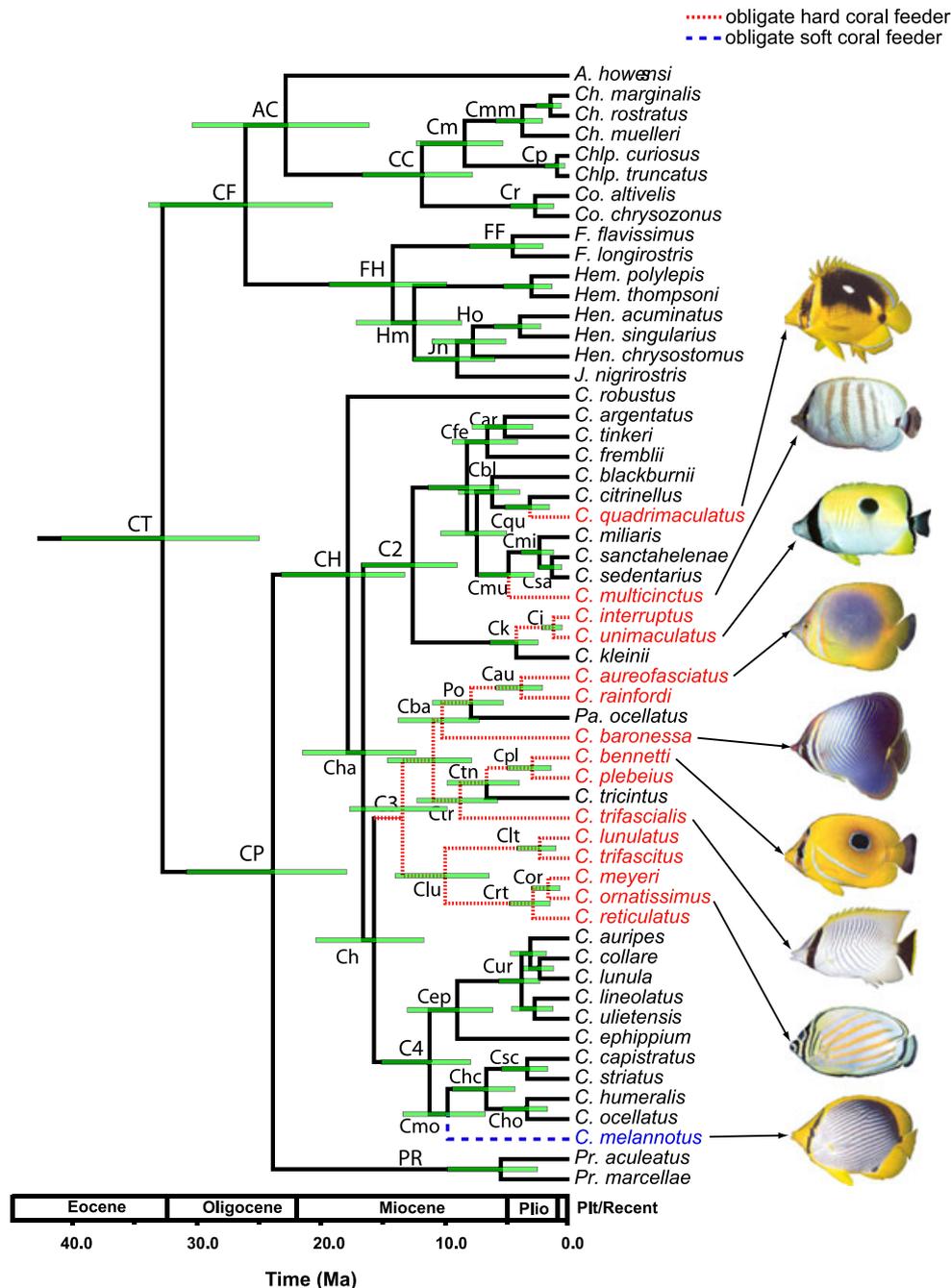


Fig. 2 A chronogram of the Chaetodontidae with optimized trophic modes reveals five independent origins of corallivory over the last 15.7–3.2 Ma. Red/dotted branches indicate obligate hard coral feeders and blue/dashed branches obligate soft coral feeders. The estimated ages are in Ma (see Fig. 1 and table 1 in Appendix S2 and for confidence intervals of the mrca age estimates). The butterflyfish illustrations exemplify some of the corallivores in each of the independent clades in which corallivory has arisen (images from Kuiter, 2002).

Discussion

Systematics of the Chaetodontidae

We present a comprehensive evaluation of the Chaetodontidae, with representatives from all described

genera and currently recognized subgenera. There was an extremely high degree of congruence among gene regions and among methods (Likelihood, Parsimony and Bayesian). Using independent models for each gene partition the resultant phylogeny had strong support for all major nodes. In all analyses, the phylogeny strongly

supports the monophyly of the family with a basal split into two clades: the long-snouted bannerfishes and the reef butterflyfishes.

Our phylogeny identified three major divisions in the long-snouted bannerfish clade and four divisions in the reef butterflyfish clade. A comparable pattern, for nine of the 10 genera, was reported by Fessler & Westneat (2007). Our data support this earlier study in placing *Amphichaetodon* within the bannerfish clade, rather than as a sister taxon to all remaining species in the family, as suggested by most morphological phylogenies (Smith *et al.*, 2003). The degree of agreement between the two molecular studies is noteworthy. Despite using different markers and different representative species, the topology of the resultant trees were almost identical. This provides excellent independent corroboration of our tree. With a robust, well supported phylogenetic reconstruction for this family we are now able to explore the evolutionary history of corallivory.

Divergence times within the Chaetodontidae

Given a well-supported cladogram, with independent support for the topology, we endeavoured to provide robust molecular age estimates within a chronogram. The use of the program BEAST allowed more precise age calibrations than previous approaches. Furthermore, the use of exponential priors accommodates both the influence of the faunal transition at the K/T boundary and the stronger influence of the 50 Ma calibration based on the fossil ephippid (*Eoplatax*). Our age estimates in the resultant chronogram are supported by several independent lines of evidence.

Firstly, our estimated ages agree well with the available fossil record. We used the two best fossil dates for calibration [i.e. (1) The K/T boundary, marking the transition between Mesozoic and Cenozoic faunas (Patterson, 1993; Bellwood & Wainwright, 2002) and, (2) the calibration 50 Ma, marking the earliest fossil record of numerous reef fish families (Bellwood, 1996)]. However, there is a third piece of fossil evidence: a fully articulated Miocene fossil chaetodontid (Carnevale, 2006). It is morphologically extremely similar to extant taxa in clade 4, and at 7 Ma old lies shortly after the age estimates for this clade with a mean age of 11.3 and 95% HPD of 7.9–15.2 Ma.

Secondly, as an independent check, we can compare our estimated ages with major biogeographic events. These again compare favourably. Firstly, the terminal Tethyan event (TTE) marking the final closure of the Red Sea land bridge is dated between 12 and 18 Ma (Steininger & Rögl, 1984). These ages approximate the minimum age of the initial division between the major clades within *Chaetodon* at 17.8 Ma (13.3–23.2 Ma, 95% HPD). Of the major clades, two (2 and 4) lie on either side of the land bridge, whereas clade 1 is restricted to the

Atlantic and clade 3 is restricted to the Indo-Pacific. Secondly, the ages of lineages that appear to have been separated by the rising of the Isthmus of Panama (IOP) i.e. *Chaetodon humeralis*–*C. ocellatus* at 3.4 Ma (1.8–5.4 Ma, 95% HPD) are again extremely close to the estimated final closure of this land bridge at 3.1 Ma (Coates *et al.*, 1992), with the 95% density distribution encompassing the geological dates (Lessios, 2008). Finally, divisions between the Indian Ocean and Pacific Ocean pairs (e.g. *C. unimaculatus*–*interruptus* at 1.3 Ma and *C. trifasciatus* at 2.4 Ma) closely match the estimated ages of other Indian Ocean–Pacific Ocean divisions (McCafferty *et al.*, 2002; Read *et al.*, 2006). The separation of *C. sedentarius* and *C. sanctaehelenae* from their closest known sister lineage in the Indo-Pacific may be a further example of an invasion of the Atlantic via the Cape of Good Hope (reviewed in Floeter *et al.*, 2008). Based on the first two biogeographic divisions (TTE and IOP), our estimated ages with HPD intervals closely approximate these two well dated biogeographic events.

Finally, a comparison of our age estimates with those in previous studies, using a range of calibration methods, suggest that the estimated ages of our terminal taxa are comparable with those of other reef fishes (e.g. Fauvelot *et al.*, 2003; Bernardi *et al.*, 2004; Klanten *et al.*, 2004; Barber & Bellwood, 2005; Read *et al.*, 2006; Cowman *et al.*, 2009) and other reef organisms (Palumbi *et al.*, 1997; Lessios *et al.*, 1999; Renema *et al.*, 2008). The closest study to the present work is by Fessler & Westneat (2007) which yielded a very similar phylogeny and broadly comparable ages, even though they used a single model in tree construction and an additive PL method for age estimation, while we used a partitioned mixed model and Bayesian MCMC analyses with informative prior calibrations. These differences will not necessarily change the tree topology but can change relative branch lengths, while the BEAST analyses take into account uncertainty in topology, sequence dataset and model parameters. Overall, fossil, biogeographic and comparative data provide strong support for our chronogram. This provides a relatively robust platform for evaluating the evolution of corallivory on coral reefs.

Evolutionary and biogeographic patterns within the Chaetodontidae

In the Chaetodontidae, a move onto reefs was associated with a significant increase in species richness. Interestingly, there was no increase associated with a switch to corallivory and the exploitation of this widely available reef resource. The Chaetodontidae can be effectively divided into two ecologically and morphologically distinct clades that should be represented as sub-families: the bannerfishes and the butterflyfishes. The bannerfish clade is characterized by a distinctive long-snout morphology and it is within this clade that we see a novel suspensorial protrusion mechanism

(Ferry-Graham *et al.*, 2001a). Despite the morphological variation and innovation within the bannerfish clade, however, the standing species richness of the bannerfish lineage is not significantly different than expected given the global rate of cladogenesis (even at high extinction rates) for the crown Chaetodontidae (Table 1). Biogeographically, the bannerfish clade has close links with Australia and temperate or sub-tropical waters, and a subtropical Australian origin for this clade remains a distinct possibility. Although habitat optimization in the bannerfishes is uncertain (Fig. S1), three of the eight lineages are found on temperate subtropical rocky reefs and many species in the other lineages are found in rocky or coastal waters, supporting these temperate associations.

In contrast to the bannerfishes, the butterflyfishes exhibit limited morphological variation. Indeed, they appear to be relatively uniform (Motta, 1988) with relatively simple oral jaw mechanics and kinematics (Ferry-Graham *et al.*, 2001a). Only with respect to intramandibular flexion does there appear to be any clear morphological variation (Konow *et al.*, 2008). Depending on the definition of a coral reef and a coral reef fish (cf. Bellwood & Wainwright, 2002), it appears that there have been multiple invasions of coral reefs by chaetodontids. The butterflyfish clade contains 103 species, approximately 80% of species within the family, and is strongly associated with coral reefs. As in parrotfishes (Streelman *et al.*, 2002), wrasses (Westneat & Alfaro, 2005) and tetraodontoids (Alfaro *et al.*, 2007), the reef dwelling clades are exceptionally species rich. The *Chaetodon* clade, in particular, exhibits far higher numbers of species than expected even at high extinction rates ($P = 0.03$, $\alpha = 0.8$). When considering just the reef-based clades 2, 3 and 4 there is a greater significant difference ($P = 0.021$, $\alpha = 0.8$) from expected. It thus appears that a move to reefs did indeed underpin diversification in *Chaetodon*, as previously reported in the tetraodontiformes (Alfaro *et al.*, 2007). This pattern may be expected in a number of reef fish groups (Bellwood & Wainwright, 2002). However, it is noteworthy that clade 3 does not demonstrate higher species richness than expected, even though this obligate reef fish clade contains the largest number of corallivores found in any teleost taxon. It appears that a move onto reefs, not a switch to corallivory, underpinned diversification within the family.

The rise of corallivory

The Chaetodontidae contains more corallivores than any other fish family; however, this did not arise as a result of a single exceptional event. Corallivory has arisen at least 5 times, with representatives in almost every major butterflyfish clade. Furthermore, it appears to have arisen relatively recently (15.7–3.2 Ma) and in a number of markedly different ways.

The oldest estimated record of corallivory is in *Chaetodon* clade 3 at 15.7 Ma (the MRCA with an omnivorous sister lineage; Fig. S2). Of the 13 species examined in this clade, 12 are corallivorous (the exception is *Parachaetodon ocellatus*). Nine additional species can be placed in this clade based on phylogenetic (*austriacus*, *larvatus*, *octofasciatus*, *speculum* and *zanzibarensis*) and taxonomic (*melapterus*, *lunulatus*, *andamanensis* and *triangulum*) evidence (Fessler & Westneat, 2007; Hsu *et al.*, 2007). All these taxa are obligate corallivores. This is the oldest record of corallivory in the family and it is in clade 3 that we see the strongest reef associations and the tightest links between fishes and corals. Several species feed on just one or two coral species and may be incapable of switching prey species (Berumen & Pratchett, 2008), while others specialize by ingesting specific parts of the coral or just mucous (Cole *et al.*, 2008). These species have relatively long intestines and appear to represent an extreme level of coral feeding specialization (Elliott & Bellwood, 2003; Konow & Ferry-Graham, in press). Given this level of specialization, it is no surprise that it is species within this clade that exhibit the most extreme negative response to the decline in coral cover as a result of anthropogenic disturbances and climate change (Pratchett *et al.*, 2006, 2008; Wilson *et al.*, 2006).

Given this long association with corallivory, the monotypic *Parachaetodon* was a striking inclusion in clade 3. *Parachaetodon ocellatus* is not a corallivore and often lives in sheltered sediment rich areas (Allen *et al.*, 1998). Given its position in the tree, this appears to be the first recorded reversal from corallivory to omnivory. The evolutionary scenario that may have triggered such a change is unclear. The explanation may be biogeographic, with a dietary switch following the loss of corals in an isolated marine basin.

The second oldest record of corallivory is in *Chaetodon* clade 4 at about 9.8 Ma, in *C. melannotus* and its sister species *C. ocellicaudus* (cf. Fessler & Westneat, 2007; Hsu *et al.*, 2007). These species are again strongly reef associated and highly specialized obligate coral feeders. However, these taxa are restricted exclusively to soft corals. Their relationship with other members of the clade is not well resolved and a sister group relationship with the omnivore *C. selene* suggested by Fessler & Westneat (2007) would imply that the origins of corallivory in the *melannotus*–*ocellicaudus* clade are younger than our estimate. Nevertheless, this represents an independent, and highly distinctive, obligate soft coral feeding lineage.

The most recent examples of corallivory are found in *Chaetodon* clade 2. This clade contains a large number of species that occasionally graze on live corals, but only four obligate corallivores. Here, corallivory arose as a result of three independent events: *C. multinctus* clade (inc. *pelewensis* and *punctatofasciatus*) at about 4.9 Ma, *C. quadrimaculatus* at about 3.2 Ma and the *C. unimaculatus*–*interruptus* clade at about 4.3 Ma. These ages are not

well-established as incomplete taxon sampling precludes robust estimates. Nevertheless, all the three stand as relatively recent independent events, a pattern that is unlikely to be altered by further taxon sampling. There are two different feeding modes. The first two lineages contain obligate corallivores and in both cases the preferred coral prey appear to be *Pocillopora* spp. (Berumen & Pratchett, 2006). The latter clade consists of two sister taxa that feed exclusively on corals; *C. unimaculatus* in the Pacific and *C. interruptus* in the Indian Ocean. *Chaetodon unimaculatus* appears to be unique in that it feeds on both soft and hard corals (hard in French Polynesia and Hawaii vs. soft on the GBR and in Guam; Motta, 1988; Wylie & Paul, 1989; Konow & Ferry-Graham, in press; M.S. Pratchett, unpublished). It is also the only butterflyfish to take large bites from corals that remove both the polyp and the surrounding tissues. In this, the bite is more reminiscent of excavating parrotfishes which leave distinctive scars at the feeding site (Bellwood & Choat, 1990). This robust feeding mode is reflected by an unusually robust jaw morphology in this lineage (Motta, 1988; Konow *et al.*, 2008).

Despite the clear patterns, care is needed when interpreting evolutionary history from phylogenies. The ages of origination refer to the approximate ages at which extant lineages are hypothesized to have commenced corallivory. The ages of these taxa are comparable with those recorded from other reef fish families such as the Pomacentridae (e.g. McCafferty *et al.*, 2002), Labridae (e.g. Read *et al.*, 2006; Cowman *et al.*, 2009) and Acanthuridae (Klanten *et al.*, 2004). Yet in each of these three families the Eocene fossil record yields several extinct fossil taxa that are the functional equivalents of extant taxa (Bannikov & Sorbini, 1990; Bellwood & Sorbini, 1996; Tyler & Sorbini, 1999). One can not, therefore, discount the possibility that corallivory predated the origins of extant lineages and our minimum age estimates. However, our estimates do provide a clear indication of the minimum age of this feeding mode and evidence of an increasing diversity of corallivores, in terms of both feeding modes and number of lineages, during the Miocene and Pliocene (15.7–3 Ma).

Corallivory and its implications for reef–fish interactions and the evolution of coral reefs

Our chronogram clearly suggests that corallivory did not arise with the origins of the major coral groups in the Eocene. Rather, it ties in with a major expansion and reorganization of reefs in the Miocene, and coincides with the initial formation of the biodiversity hotspot in the Indo-Australian Archipelago.

Even given that our estimates are minimum ages, 15.7–3 Ma still represents a relatively recent origination for such a derived feeding mode as corallivory. Scleractinian corals have been a significant component of shallow

carbonate reefs since the early Tertiary, with most of the major *Acropora* clades (the coral genus targeted by most modern corallivores) already represented in the Eocene at 49–37 Ma (Wallace & Rosen, 2006). In contrast, other major coral reef benthic feeding modes, e.g. grazing herbivory and crushing with pharyngeal jaws, have been present for at least 50 Ma (Bellwood & Sorbini, 1996; Bellwood, 1999, 2003; Cowman *et al.*, 2009). The problem of minimum estimates notwithstanding, this relatively recent rise of corallivory raises two questions: are chaetodontids one of the most recent taxa to switch to corallivory and does this switch reflect a broader change in the nature of reef–fish interactions?

In terms of the evolution of corallivory, the evidence is scarce but all the indications are that the timing of corallivory in chaetodontids is comparable to that of the only other major group with significant numbers of corallivores, the labrids. Based on the most recent labrid phylogeny (Cowman *et al.*, 2009) corallivory appears to be derived, to have arisen only once (in the *Labropsis–Labrichthys* clade) and to have arisen relatively recently, although considerably earlier than in the Chaetodontidae (at 29 Ma). In the parrotfishes (i.e. *Bolbometopon muricatum* and *Sparisoma viride*) coral feeding probably arose prior to the late Miocene (12 and 10 Ma respectively; Robertson *et al.*, 2006; Cowman *et al.*, 2009). Overall, it appears that the chaetodontids are only exceptional in terms of the number of corallivorous species within the family. Their dietary shift appears to have coincided with a general rise in corallivory in a range of reef fish families.

In terms of the broader changes in the nature of reef–fish interactions, the rise of corallivory in the Miocene is consistent with several other lines of evidence. We see a progressive increase in detritivory in the Miocene (Harmelin-Vivien, 2002) and a number of novel specialist groups e.g. specialist foraminifera feeders and fish cleaners (*Macropharyngodon*; Read *et al.*, 2006; Cowman *et al.*, 2009). The origins of corallivory, therefore, fit in a broader context in which the Miocene exhibits a new level of reef–fish interactions with more specialized reef-associated taxa. This may be associated with the rise of *Acropora* and *Pocillopora* as the dominant coral groups during this period (Johnson *et al.*, 2008; B. Rosen, personal communication). The vast majority of corallivores and all obligate specialists feed only on these coral genera.

The number of independent origins of corallivory and the lack of morphological modifications to the feeding apparatus suggest that there are few morphological restrictions to corallivory, although the elongation of the intestine suggests that the difficulty, if any, may lie in processing rather than procuring coral tissues. Extant corallivores are often highly selective feeders, exploiting specific coral species or even specific sites on a coral (e.g. damaged tissues) (McIlwain & Jones, 1997). The rise of corallivory may therefore have been dependent on corals reaching sufficient densities to permit the selective feeding necessary to adequately process the coral tissues; the

increased access to *Acropora* and *Pocillopora* colonies triggering the expansion of corallivory in the Miocene. The rapid expansion of coral bearing carbonate platforms in the Indo-Australian Archipelago in the early-mid Miocene (Wilson, 2008) may therefore have acted as the trigger for not only the rapid expansion of numerous fish groups but the origins of trophic novelty, including corallivory (Renema *et al.*, 2008; Cowman *et al.*, 2009). In this context, it is interesting to note that on modern coral reefs the number of corallivores declines swiftly in response to the loss of coral cover (Pratchett *et al.*, 2006, 2008).

Coral reefs have been exposed to escalating predation pressure for millennia (Vermeij, 1977; Bellwood, 2003). For corals, predation by fishes certainly appears to have increased over the last 15 Ma. We now have, for the first time, an understanding of the origins of corallivory in fishes. Of all corallivorous fishes 63% are found within a single family, the Chaetodontidae. Yet, surprisingly, within this family this derived feeding mode has arisen at least five times over the last 3–15.7 Ma, with specialists on both soft and hard corals. This unusual feeding mode appears to reflect an exceptionally close association between this family and coral reefs. An understanding of this history offers a new perspective on the nature of the relationships between fishes and coral reefs in a changing world.

Acknowledgments

The authors wish to thank C. Fulton, A. Hoey, P.C. Wainwright, F. Walsh and P. Wirtz for tissue samples; S. Wismer for graphics assistance; the staff of Carrie Bow Cay, Orpheus Island, Lizard Island and Moorea CRIOBE Research Stations for invaluable field support, colleagues in the Centre of Excellence for Coral Reef Studies for helpful discussions and several reviewers for constructive comments. This work was financially supported by James Cook University and the Australian Research Council.

References

- Alfaro, M.E., Brock, C.D. & Santini, F. 2007. Do reefs drive diversification in marine teleosts? Evidence from the pufferfishes and their allies (Tetraodontiformes: Acanthomorpha). *Evolution* **61**: 2104–2126.
- Allen, G.R., Steene, R. & Allen, M. 1998. *A Guide to Angelfishes and Butterflyfishes*. Odyssey Publishing-Tropical Reef Research, Perth.
- Azuma, Y., Kumazawa, Y., Miya, M., Mabuchi, K. & Nishida, M. 2008. Mitogenomic evaluation of the historical biogeography of cichlids toward reliable dating of teleostean divergences. *BMC Evol. Biol.* **8**: 215.
- Bannikov, A.F. 2004. Fishes from the Eocene of Bolca, northern Italy, previously classified with the Chaetodontidae (Perciformes). *Studie Ricerche sui Giacimenti Terziari di Bolca* **10**: 55–74.
- Bannikov, A.F. & Sorbini, L. 1990. *Eocoris bloti*, a new genus and species of labrid fish (Perciformes, Labroidae) from the Eocene of Monte Bolca, Italy. *Studie Ricerche sui Giacimenti Terziari di Bolca* **6**: 133–148.
- Barber, P.H. & Bellwood, D.R. 2005. Biodiversity hotspots: evolutionary origins of biodiversity in wrasses (*Halichoeres*: Labridae) in the Indo-Pacific and New World tropics. *Mol. Phylogenet. Evol.* **35**: 235–253.
- Bellwood, D.R. 1996. The Eocene fishes of Monte Bolca: the earliest coral reef fish assemblage. *Coral Reefs* **15**: 11–19.
- Bellwood, D.R. 1999. Fossil Pharyngognath fishes from Monte Bolca, Italy, with a description of a new pomacentrid genus and species. *Studie Ricerche sui Giacimenti Terziari di Bolca* **8**: 207–217.
- Bellwood, D.R. 2003. Origins and escalation of herbivory in fishes: a functional perspective. *Paleobiology* **29**: 71–83.
- Bellwood, D.R. & Choat, J.H. 1990. A functional analysis of grazing in parrotfishes (family Scaridae): the ecological implications. *Environ. Biol. Fish* **28**: 189–214.
- Bellwood, D.R. & Schultz, O. 1991. A review of the fossil record of the parrotfishes (Labroidae: Scaridae) with a description of a new *Calotomus* species from the Middle Miocene (Badenian) of Austria. *Ann. Nat. Mus. Wien.* **92**: 55–71.
- Bellwood, D.R. & Sorbini, L. 1996. A review of the fossil record of the Pomacentridae (Teleostei: Labroidae) with a description of a new genus and species from the Eocene of Monte Bolca, Italy. *Zool. J. Linn. Soc.* **117**: 159–174.
- Bellwood, D.R. & Wainwright, P.W. 2002. The history and biogeography of fishes on coral reefs. In: *Coral Reef Fishes: Dynamics and Diversity in a Complex Ecosystem* (P.F. Sale, ed.), pp. 5–32. Academic Press, London.
- Bellwood, D.R., van Herwerden, L. & Konow, N. 2004. Evolution and biogeography of marine angelfishes (Pisces: Pomacanthidae). *Mol. Phylogenet. Evol.* **33**: 140–155.
- Bernardi, G., Bucciarelli, G., Costagliola, D., Robertson, D.R. & Heiser, J.B. 2004. Evolution of coral reef fish *Thalassoma* spp. (Labridae): I. Molecular phylogeny and biogeography. *Mar. Biol.* **144**: 369–375.
- Berumen, M.L. & Pratchett, M.S. 2006. Persistent disturbance and long-term shifts in the structure of fish and coral communities at Tiahura Reef, Moorea. *Coral Reefs* **25**: 647–653.
- Berumen, M.L. & Pratchett, M.S. 2008. Trade-offs associated with dietary specialization in corallivorous butterflyfishes (Chaetodontidae: *Chaetodon*). *Behav. Ecol. Sociobiol.* **62**: 989–994.
- Berumen, M.L., Pratchett, M.S. & McCormick, M.I. 2005. Within reef variation in the diet and condition of two coral feeding butterflyfish (Pisces: Chaetodontidae). *Mar. Ecol. Prog. Ser.* **287**: 217–227.
- Birkeland, C. & Neudecker, S. 1981. Foraging behavior of two Caribbean chaetodontids: *Chaetodon capistratus* and *C. aculeatus*. *Copeia* **1**: 169–178.
- Blot, J. 1969. Les poissons fossiles du Monte Bolca. Classe's jusqu'ici dans les familles des Carangidae, Menidae, Ephippidae, Scatophagidae. *Mem. Mus. Civ. Stor. Nat. Verona* **1**: 1–525.
- Blot, J. 1980. La faune ichthyologique des gisements du Monte Bolca (Province de Verone, Italie): catalogue systématique présentant l'état actuel des recherches concernant cette faune. *Bull. Mus. Nat. Hist. Nat.* **2**: 339–396.
- Blum, S.D. 1988. Biogeography of the Chaetodontidae: an analysis of allopatry among closely related species. *Environ. Biol. Fish* **25**: 9–31.
- Brown, J.W., Rest, J.S., García-Moreno, J., Sorenson, M.D. & Mindell, D.P. 2008. Strong mitochondrial DNA support for a Cretaceous origin of modern avian lineages. *BMC Biol.* **6**: 6.

- Carnevale, G. 2006. Morphology and biology of the Miocene butterflyfish *Chaetodon fischeuri* (Teleostei: Chaetodontidae). *Zool. J. Linn. Soc.* **146**: 251–267.
- Chow, S. & Hazama, K. 1998. Universal PCR primers for S7 ribosomal protein gene introns in fish. *Mol. Ecol.* **7**: 1255–1256.
- Coates, A.G., Jackson, J.B.C., Collins, L.S., Cronin, T.M., Dowsett, H.J., Bybell, L.M., Jung, P.J. & Obando, J.A. 1992. Closure of the Isthmus of Panama: the nearshore record of Costa Rica and western Panama. *Geol. Soc. Am. Bull.* **104**: 814–828.
- Cole, A.J., Pratchett, M.S. & Jones, G.P. 2008. Diversity and functional importance of coral-feeding fishes on tropical coral reefs. *Fish Fish.* **9**: 286–307.
- Cowman, P.F., Bellwood, D.R. & van Herwerden, L. 2009. Dating the evolutionary origins of wrasse lineages and the rise of trophic novelty on coral reefs. *Mol. Phylogenet. Evol.* **52**: 621–631.
- Drummond, A.J. & Rambaut, A. 2007. Bayesian evolution analysis by sampling trees. *BMC Evol. Biol.* **7**: 214.
- Drummond, A.J., Ho, S.Y.W., Phillips, M.J. & Rambaut, A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* **4**: 699–710.
- Dwyer, J., Li, H., Xu, D. & Liu, J.P. 2007. Transcriptional regulation of telomerase activity: roles of the Ets transcription factor family. *Ann. N. Y. Acad. Sci.* **1114**: 36–47.
- Elliott, J. & Bellwood, D.R. 2003. Alimentary tract morphology and diet in three coral reef fish families. *J. Fish Biol.* **63**: 1598–1609.
- Fauvelot, C., Bernardi, G. & Planes, S. 2003. Reductions in the mitochondrial DNA diversity of coral reef fish provide evidence of population bottlenecks resulting from Holocene sea-level change. *Evolution* **57**: 1571–1583.
- Ferry-Graham, L.A., Wainwright, P.C. & Bellwood, D.R. 2001a. Prey capture in long-jawed butterflyfishes (Chaetodontidae): the functional basis of novel feeding habits. *J. Exp. Mar. Biol. Ecol.* **256**: 167–184.
- Ferry-Graham, L.A., Wainwright, P.C., Hulsey, C.D. & Bellwood, D.R. 2001b. Evolution and mechanics of long jaws in butterflyfishes (Family Chaetodontidae). *J. Morphol.* **248**: 120–143.
- Fessler, J.L. & Westneat, M.W. 2007. Molecular phylogenetics of the butterflyfishes (Chaetodontidae): taxonomy and biogeography of a global coral reef fish family. *Mol. Phylogenet. Evol.* **45**: 50–68.
- Findley, J.S. & Findley, M.T. 2001. Global, regional, and local patterns in species richness and abundance of butterflyfishes. *Ecol. Monogr.* **71**: 69–91.
- Floeter, S.R., Rocha, L.A., Robertson, D.R., Joyeux, J.C., Smith-Vaniz, W.F., Wirtz, P., Edwards, A.J., Barreiros, J.P., Ferreira, C.E.L., Gasparini, J.L., Brito, A., Falcón, J.M., Bowen, B.W. & Bernardi, G. 2008. Atlantic reef fish biogeography and evolution. *J. Biogeogr.* **35**: 22–47.
- Gregson, M., Berumen, M.L. & Pratchett, M.S. 2008. Bite rates of butterflyfishes (Chaetodontidae) correlate with proportional consumption of corals. *Coral Reefs* **27**: 583–591.
- Harmelin-Vivien, M.L. 2002. Energetics and fish diversity on coral reefs. In: *Coral Reef Fishes: Dynamics and Diversity in a Complex Ecosystem* (P.F. Sale, ed.), pp. 265–274. Academic Press, New York.
- Harmon, L.J., Weir, J.T., Brock, C.D., Glor, R.E. & Challenger, W. 2008. GEIGER: investigating evolutionary radiations. *Bioinformatics* **24**: 129–131.
- Ho, S.Y.W. 2007. Calibrating molecular estimates of substitution rates and divergence times in birds. *J. Avian Biol.* **38**: 404–414.
- Hsu, K., Chen, J. & Shao, K. 2007. Molecular phylogeny of *Chaetodon* (Teleostei: Chaetodontidae) in the Indo-West Pacific: evolution in geminate species pairs and specific groups. *Raffles Bull. Zool.* **14**: 77–86.
- Huelsenbeck, J.P. & Ronquist, F. 2001. MR. BAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**: 754–755.
- Ihaka, R. & Gentleman, R. 1996. R: a language for data analysis and graphics. *J. Comput. Graph. Stat.* **5**: 299–314.
- Irwin, D.M., Kocher, T.D. & Wilson, A.C. 1991. Evolution of the cytochrome b gene of mammals. *J. Mol. Evol.* **32**: 128–144.
- Johnson, K.G., Jackson, J.B.C. & Budd, A.F. 2008. Caribbean reef development was independent of coral diversity over 28 million years. *Science* **319**: 1521–1523.
- Klanten, S.O., van Herwerden, L., Choat, J.H. & Blair, D. 2004. Patterns of lineage diversification in the genus *Naso* (Acanthuridae). *Mol. Phylogenet. Evol.* **32**: 221–235.
- Kocher, T.D., Thomas, W.K., Meyer, A., Edwards, S.V., Paabo, S., Villablanca, F.X. & Wilson, C. 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proc. Natl Acad. Sci.* **86**: 6196–6200.
- Konow, N. & Ferry-Graham, L. (in press). Functional morphology of the butterflyfishes, In: *The Biology of Butterflyfishes*, Chapter 2 (M.S. Pratchett, M.L. Berumen & B.G. Kapoor, eds). Science Publishers Inc, Enfield, USA.
- Konow, N., Bellwood, D.R., Wainwright, P.C. & Kerr, A.M. 2008. Evolution of novel jaw joints promote trophic diversity in coral reef fishes. *Biol. J. Linn. Soc.* **93**: 545–555.
- Kuiter, R.H. 2002. *Butterflyfishes, Bannerfishes and Their Relatives: A Comprehensive Guide to Chaetodontidae and Microcanthidae*. TMC Publishing, Chorleywood.
- Kulbicki, M., Bozec, Y.M. & Green, A. 2005. Implications of biogeography in the use of butterflyfishes (Chaetodontidae) as indicators for Western and Central Pacific areas. *Aquat. Conserv. Mar. Freshw. Ecosys.* **15**: S109–S126.
- Lessios, H.A. 2008. The great American schism: divergence of marine organisms after the rise of the Central American Isthmus. *Annu. Rev. Ecol. Evol. Syst.* **39**: 63–91.
- Lessios, H.A., Kessing, B.D., Robertson, D.R. & Paulay, G. 1999. Phylogeography of the pantropical sea urchin *Eucidaris* in relation to land barriers and ocean currents. *Evolution* **53**: 806–817.
- Littlewood, D.T.J., McDonald, S.M., Gill, A.C. & Cribb, T.H. 2004. Molecular phylogenetics of *Chaetodon* and the Chaetodontidae (Teleostei: Perciformes) with reference to morphology. *Zootaxa* **779**: 1–20.
- Lyons, L.A., Laughlin, T.F., Copeland, N.G., Jenkins, N.A., Womack, J.E. & O'Brien, S.J. 1997. Comparative anchor tagged sequences (CATS) for integrative mapping of mammalian genomes. *Nat. Genet.* **15**: 47–56.
- Maddison, W. & Maddison, D. 2007. Mesquite v 2.01. URL <http://mesquiteproject.org>.
- Magallon, S. & Sanderson, M.J. 2001. Absolute diversification rates in angiosperm clades. *Evolution* **55**: 1762–1780.
- Maguire, B.A. & Zimmermann, R.A. 2001. The ribosome in focus. *Cell* **104**: 813–816.
- McCafferty, S., Bermingham, E., Quenouille, B., Planes, S., Hoelzer, G. & Asoh, K. 2002. Historical biogeography and molecular systematics of the Indo-Pacific genus *Dascyllus* (Teleostei: Pomacentridae). *Mol. Ecol.* **11**: 1377–1392.
- McIlwain, J.L. & Jones, G.P. 1997. Prey selection by an obligate coral-feeding wrasse and its response to small-scale disturbance. *Mar. Ecol. Prog. Ser.* **155**: 189–198.

- McMillan, W.O. & Palumbi, S.R. 1995. Concordant evolutionary patterns among Indo-West Pacific butterflyfishes. *Proc. R. Soc. Lond. B* **260**: 229–236.
- Motta, P.J. 1988. Functional morphology of the feeding apparatus of ten species of Pacific butterflyfishes (Perciformes, Chaetodontidae): an ecomorphological approach. *Environ. Biol. Fish* **22**: 39–67.
- Nee, S., Mooers, O. & Harvey, P.H. 1992. Tempo and mode of evolution revealed from molecular phylogenies. *Proc. Natl Acad. Sci. USA* **89**: 8322–8326.
- Nylander, J.A.A. 2004. *MrModeltest v.2*. Evolutionary Biology Centre, Uppsala University, Uppsala.
- Nylander, J.A.A., Ronquist, F., Huelsenbeck, J.P. & Nieves-Aldrey, J.L. 2004. Bayesian phylogenetic analysis of combined data. *Syst. Biol.* **53**: 47–67.
- Palumbi, S.R., Grabowsky, G., Duda, T., Geyer, L. & Tachino, N. 1997. Speciation and population genetic structure in tropical Pacific sea urchins. *Evolution* **51**: 1506–1517.
- Patterson, C. 1993. Osteichthyes: Teleostei. In: *The Fossil Record 2* (M.J. Benton, ed.), pp. 621–656. Chapman and Hall, London.
- Posada, D. & Crandall, K.A. 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics* **14**: 814–818.
- Pratchett, M.S. 2005. Dietary overlap among coral-feeding butterflyfishes (Chaetodontidae) at Lizard Island, northern Great Barrier Reef. *Mar. Biol.* **148**: 373–382.
- Pratchett, M.S. 2007. Dietary selection by coral-feeding butterflyfishes (Chaetodontidae) on the Great Barrier Reef, Australia. *Raffles Bull. Zool.* **14**: 155–160.
- Pratchett, M.S., Wilson, S.K., Berumen, M.L. & McCormick, M.I. 2004. Sub-lethal effects of coral bleaching on an obligate coral feeding butterflyfish. *Coral Reefs* **23**: 352–356.
- Pratchett, M.S., Wilson, S.K. & Baird, A.H. 2006. Declines in the abundance of *Chaetodon* butterflyfishes (Chaetodontidae) following extensive coral depletion. *J. Fish Biol.* **69**: 1269–1280.
- Pratchett, M.S., Munday, P.L., Wilson, S.K., Graham, N.A.J., Cinner, J.E., Bellwood, D.R., Jones, G.P., Polunin, N.V.C. & McClanahan, T.R. 2008. Effects of climate-induced coral bleaching on coral-reef fishes: ecological and economic consequences. *Oceanogr. Mar. Biol. Ann. Rev.* **46**: 251–296.
- Pybus, O.G. & Harvey, P.H. 2000. Testing macro-evolutionary models using incomplete molecular phylogenies. *Philos. Trans. R. Soc. (Lond.) B* **267**: 2267–2272.
- Rabosky, D.L. 2006. LASER: a maximum likelihood toolkit for detecting temporal shifts in diversification rates from molecular phylogenies. *Evol. Bioinform. Online* **2**: 247–250.
- Rambaut, A. 1996. Se-Al: sequence alignment editor. URL <http://evolve.zoo.ox.ac.uk>.
- Rambaut, A. & Drummond, A.J. 2007. Tracer v1.4. URL <http://beast.bio.ed.ac.uk/Tracer>.
- Read, C.I., Bellwood, D.R. & van Herwerden, L. 2006. Ancient origins of Indo-Pacific coral reef fish biodiversity: a case study of the leopard wrasses (Labridae: *Macropharyngodon*). *Mol. Phylogenet. Evol.* **38**: 808–819.
- Reese, E.S. 1975. A comparative field study of the social behavior and related ecology of reef fishes of the family Chaetodontidae. *Z. Tierpsychol.* **37**: 37–61.
- Renema, W., Bellwood, D.R., Braga, J.C., Bromfield, K., Hall, R., Johnson, K.G., Lunt, P., Meyer, C.P., McMonagle, L., Morley, R.J., O’dea, A., Todd, J.A., Wesselingsh, F.P., Wilson, M.E.J. & Pandolfi, J.M. 2008. Hopping hotspots: global shifts in marine biodiversity. *Science* **321**: 654–657.
- Roberts, C.M., Ormond, R.F.G. & Shepherd, A.R.D. 1988. The usefulness of butterflyfishes as environmental indicators on coral reefs. In: *Proceedings of the 6th International Coral Reef Symposium 2* (J.H. Choat, ed.), pp. 331–336. Townsville, Australia.
- Robertson, D.R., Krag, F., de Moura, R.L., Victor, B.C. & Bernardi, G. 2006. Mechanisms of speciation and faunal enrichment in Atlantic parrotfishes. *Mol. Phylogenet. Evol.* **40**: 795–807.
- Rotjan, R.D. & Lewis, S.M. 2008. Impact of coral predators on tropical reefs. *Mar. Ecol. Prog. Ser.* **367**: 73–91.
- Sambrook, J. & Russell, D.W. 2001. *Molecular Cloning, a Laboratory Manual*. Cold Spring Harbour Laboratory Press, New York.
- Sanderson, M.J. 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Mol. Biol. Evol.* **19**: 101–109.
- Sanderson, M.J. 2004. *r8s, Version 1.70: User’s Manual*. Section of Evolution and Ecology, University of California, Davis, CA.
- Smith, W.L. & Wheeler, W.C. 2006. Venom evolution widespread in fishes: a phylogenetic road map for the bioprospecting of piscine venoms. *J. Hered.* **97**: 206–217.
- Smith, W.L., Webb, J.F. & Blum, S.D. 2003. The evolution of the laterophysic connection with a revised phylogeny and taxonomy of butterflyfishes (Teleostei: Chaetodontidae). *Cladistics* **19**: 287–306.
- Steininger, F.F. & Rögl, F. 1984. Paleogeography and palinspastic reconstruction of the Neogene of the Mediterranean and Paratethys. In: *The Geological Evolution of the Eastern Mediterranean* (J.E. Dixon & A.H. Robertson, eds), pp. 659–668. Blackwell, Oxford.
- Streelman, J.T., Alfaro, M., Westneat, M.W., Bellwood, D.R. & Karl, S.A. 2002. Evolutionary history of the parrotfishes: biogeography, ecomorphology, and comparative diversity. *Evolution* **56**: 961–971.
- Swofford, D.L. 1998. *PAUP* Phylogenetic Analysis Using Parsimony (* and Other Methods)*. Sinauer Associates, Sunderland, MA.
- Thompson, J.D., Gibson, T.J., Plewniak, F.J. & Higgins, D.G. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **24**: 4876–4882.
- Tyler, J.C. & Sorbini, C. 1999. Phylogeny of the fossil and recent genera of fishes of the family Scatophagidae (Squamipinnes). *Boll. Mus. Civ. Stor. Nat. Verona* **23**: 353–393.
- Vermeij, G.J. 1977. The mesozoic marine revolution: evidence from snails, predators and grazers. *Paleobiology* **3**: 245–258.
- Wallace, C.C. & Rosen, B. 2006. Diverse staghorn corals (*Acropora*) in high-latitude Eocene assemblages: implications for the evolution of modern diversity patterns of reef corals. *Proc. Biol. Sci.* **273**: 975–982.
- Westneat, M.W. & Alfaro, M. 2005. Phylogenetic relationships and evolutionary history of the reef fish family Labridae. *Mol. Phylogenet. Evol.* **36**: 370–390.
- Wilson, M.E.J. 2008. Global and regional influences on equatorial shallow-marine carbonates during the Cenozoic. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **265**: 262–274.
- Wilson, S.K., Graham, N.A.J., Pratchett, M.S., Jones, G.P. & Polunin, N.V.C. 2006. Multiple disturbances and the global degradation of coral reefs: are reef fishes at risk or resilient? *Glob. Chang. Biol.* **12**: 2220–2234.
- Wylie, C.R. & Paul, V.J. 1989. Chemical defenses in three species of *Simularia* (Coelenterata, Alcyonacea): effects against generalist predators and the butterflyfish *Chaetodon unimaculatus* Bloch. *J. Exp. Mar. Biol. Ecol.* **129**: 141–160.

- Xia, X.H. & Xie, Z. 2001. DAMBE: data analysis in molecular biology and evolution. *J. Hered.* **92**: 371–373.
- Xia, X.H., Xie, Z., Salemi, M., Chen, L. & Wang, Y. 2003. An index of substitution saturation and its application. *Mol. Phylogenet. Evol.* **26**: 1–7.
- Yang, Z. & Rannala, B. 2006. Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Mol. Biol. Evol.* **23**: 212–226.
- Zekeria, Z.A., Dawit, Y., Ghebremedhin, S., Naser, M. & Videler, J.J. 2002. Resource partitioning among four butterflyfish species in the Red Sea. *Mar. Freshw. Res.* **53**: 63–168.
- Zwickl, D.J. 2006. Garli – genetic algorithm for rapid likelihood inference. URL <http://www.bio.utexas.edu/faculty/antisense/garli/Garli.html>.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Table 1: List of species in the Chaetodontidae; their taxonomy, diet and habitat associations. Table 2: Primer sequences.

Appendix S2 Estimated ages of nodes with HPD distributions.

Figure S1 Coral reef use as habitat preference and the likelihood (ML) of the presence of this trait in the most recent common ancestor (mrca) of each clade, as indicated by the pie charts at each node, when mapped to the best phylogenetic reconstruction of the group.

Figure S2 Coravllivory and the likelihood (ML) of the presence of this trait in the most recent common ancestor (mrca) of each clade, as indicated by the pie charts at each node, when mapped to the best phylogenetic reconstruction of the group.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

Received 20 October 2009; accepted 23 October 2009

The anthropogenic environment lessens the intensity and prevalence of gastrointestinal parasites in Balinese long-tailed macaques (*Macaca fascicularis*)

Kelly E. Lane · Concerta Holley · Hope Hollocher · Agustin Fuentes

Received: 19 August 2010 / Accepted: 22 November 2010 / Published online: 17 December 2010
© Japan Monkey Centre and Springer 2010

Abstract The distribution of wildlife parasites in a landscape is intimately tied to the spatial distribution of hosts. In parasite species, including many gastrointestinal parasites, with obligate or common environmental life stages, the dynamics of the parasite can also be strongly affected by geophysical components of the environment. This is especially salient in host species, for example humans and macaques, which thrive across a wide variety of habitat types and quality and so are exposed to a wealth of environmentally resilient parasites. Here, we examine the effect of environmental and anthropogenic components of the landscape on the prevalence, intensity, and species diversity of gastrointestinal parasites across a metapopulation of long-tailed macaques on the island of Bali, Indonesia. Using principal-components analysis, we identified significant interaction effects between specific environmental and anthropogenic components of the landscape, parsing the Balinese landscape into anthropogenic (PC1), mixed environment (PC2), and non-anthropogenic (PC3) components. Further, we determined that the anthropogenic environment can mitigate the prevalence and intensity of specific gut parasites and the intensity of the overall community of gut parasites, but that non-anthropogenically driven landscape components have no significant effect in increasing or reducing the intensity or prevalence of the community of gut parasites in Balinese macaques.

Keywords Macaques · Gastrointestinal parasites · Intensity · Parasite species diversity · Ecosystem

Introduction

Ecological factors are often described as significant drivers of primate sociality and intra-group interactions, including social organization, mating systems, and social dominance (van Schaik 1996; Lee 1999; Hartwig 2007; Thierry 2007; Fuentes 2007). Given the shared evolutionary history of humans and non-human primates, the current levels of human-driven habitat disturbance, and the low levels of host specificity for many parasites of primates, the risk of disease emergence or re-emergence from non-human primates into human populations is high (Schrag and Weiner 1995; Wolfe et al. 1998; Jones-Engel et al. 2004). While it is challenging to identify bi-directional transmission between humans and other primates, the risk of shared parasites may be great in areas with high human and non-human primate densities and where resource sharing occurs (Jones-Engel et al. 2004; Pederson and Fenton 2006; Gasser et al. 2009; Gurarie and Seto 2008). The increase in overlap of humans and wildlife populations, and the significant effect of parasite-mediated host mortality, fertility, and fecundity, necessitates the further understanding of environmental components of host–parasite interactions (Baudoin 1975; Dobson 1988; Combes 1996; Stirnadel and Ebert 1997; Hurd 2001; Daszak et al. 2001; Keesing et al. 2006). Further, investigation of wild primate populations can provide insight into the interaction between the environment and the population dynamics and social structure of both the host and parasite (Wolfe et al. 1998; Ostfeld et al. 2005; Wolfe and Switzer 2009; but see Gasser et al. 2009).

K. E. Lane (✉) · C. Holley · H. Hollocher
004 Galvin Life Sciences, Department of Biological Sciences,
University of Notre Dame, Notre Dame, IN 46556, USA
e-mail: klane3@nd.edu

A. Fuentes
Department of Anthropology, University of Notre Dame,
Notre Dame, IN 46556, USA

Although the effect of the environment on the transmission dynamics of gastrointestinal parasites is evident (Haufstater and Meade 1982; Jaffee et al. 1992; Roepstorff and Murrell 1997; Slifko et al. 2000; Patz et al. 2000), many studies compare host populations that differ drastically in only one important environmental component (i.e. water availability or habitat) (Anderson and Beaudoin 1966; Grutter 1998; Baird 2003; Ferrari et al. 2004; Trejo-Macias et al. 2007). Although this is important in determining the role of specific facets of the environment and eliminating confounding factors, this approach minimizes the ability to identify and understand the interactive effects, and thus the emergent properties, of the landscape (Fuentes 2009). Few studies have examined how these environmental factors have affected a community of parasites across a metapopulation of non-human primate populations thriving in a heterogeneous, anthropogenic landscape. For example, Trejo-Macias et al. (2007) examined parasite species prevalence and diversity in populations of howler monkeys (*Alouatta palliata mexicana* and *A. pigra*) and found habitat fragmentation affects the diversity of infection of coccidia and strongylids. This study emphasizes the importance of examining the effect of an environmental component on a suite of parasites across host populations; however, its fine-scale focus potentially limits the ability to detect larger scale effects of parasitism across the landscape.

The Indonesian island of Bali is an ideal model system for evaluating the effect of a complex anthropogenic landscape on gastrointestinal parasites. Long-tailed macaques (*Macaca fascicularis*) live in temples associated with forest patches, or monkey forests, across the island. Individual monkey forests are closely associated with human populations and are often provisioned (Fuentes et al. 2005; Lane et al. 2010). The habitat of the island is wide ranging, and in addition to a large urban center in the south of the island, includes both tropical forest patches, bamboo forests, wet rice agriculture, and dry scrub habitat, with water availability reflecting the varied habitats (Southern 2002). The existence of thriving populations of long-tailed macaques across these habitats demonstrates the robustness of this species to environmental fluctuations and perturbations. However, the ability of macaques to access food and water has had a substantial impact on the growth of these populations, with a small number of populations undergoing recent, explosive increases in population density. Although all macaque populations are in close spatial proximity to human populations, a small number of populations are located in the center of high-density human populations and often act as tourist destinations for international travelers (Fuentes et al. 2005). This immediacy to dense human populations and interactions with international tourists can be a source of novel infections within

macaque populations, potentially exposing macaques to human pathogens from across the globe and risking human infection throughout Bali (Jones-Engel et al. 2004).

Gastrointestinal parasites that largely utilize the fecal-oral route of transmission, including those helminths and protozoa examined in this study, are potentially susceptible to changing environmental conditions (Thieltges et al. 2008). Several of these parasites are water-borne or soil-borne parasites and have environmentally resilient infective stages that can survive outside the host (Niezen et al. 2001; Mouritsen and Poulin 2003; Thieltges et al. 2008). However, the success of these infective stages is largely dependent upon the environment in which they occur and is commonly susceptible to temperature fluctuations and desiccation (Anderson and Beaudoin 1966). In Bali, the substantial ranges in water availability, habitat type, elevation, host population size, food availability and provisioning rate, and proximity to high-density human populations should facilitate predictable differences in gut parasite species intensity, diversity, and prevalence. Namely, gastrointestinal parasite prevalence, intensity, and diversity should increase with water availability—given their transmissibility through this medium—and with macaque population size and nearness of high-density human populations, because of the increased numbers of interactions and larger sources of parasites under these conditions. In contrast, high levels of provisioning and readily available sources of year-round food among monkey forests should result in a decrease in parasite prevalence, intensity, and diversity, because of the combined effects of better nutrition and minimization of time spent foraging, reducing potential exposure to parasite infective stages. We also predict that specific elements of the landscape: amount of forest, rice agriculture, and elevation—those landscape features significant in landscape connectedness—will result in an increase in parasite prevalence, intensity, and diversity by enhancing connections between macaque populations and thus, parasite transmission.

Here, we report the prevalence, intensity, and diversity of the gastrointestinal parasites of long-tailed macaques across Bali, and the effect of specific environmental and anthropogenic landscape components and their interactions on these patterns of parasitism. Preliminary genetic evidence comparing mitochondrial gene genealogies across populations (Lane et al. 2010) reveals that the matrilineal nature of macaques has resulted in long-term female stability, exhibited as genetic clustering at this locus by population. Thus, we hypothesize complex geophysical features will have a real effect on the prevalence, intensity, and diversity of gastrointestinal parasites among these macaque populations. Specifically, we evaluate how the complex anthropogenic landscape facets of habitat type, water availability, proximity to humans, provisioning rate

or food availability, elevation, host population size, and the interactions between these variables affect the patterns of gastrointestinal parasites.

Methods

We collected samples of feces ($n = 468$) from 13 populations of long-tailed macaques on the Indonesian island of Bali during the summer of 2007. Populations sampled here have been monitored through multiple site visits from 2000 to 2007, and during these visits, population size estimates, habitat assessments, and interviews with locals and tourists were collected (Fuentes et al. 2005; Loudon et al. 2006; Lane et al. 2010, 2011). These populations occur in habitats dominated by rice agriculture, scrub-type habitat, bamboo forest, and both wet and dry forest. Although most populations are visited only infrequently and by Balinese and other Indonesian tourists, a small number are visited daily and by large numbers of international tourists, resulting in a pattern of provisioning heavily skewed towards these large populations (Fuentes et al. 2005; Lane et al. 2010, 2011). The influx of international tourists at these sites has enabled a level of management at these few temples, further skewing the level of provisioning to the scale of heavy and routine daily feeding.

A mean of 35 samples of feces (ranging from 16 samples at Angseri to 115 samples at Ubud) was collected per population. In this system, macaque population sizes range from 25 to approximately 400 individuals, and collections of feces ranged from 15% (Bedugal) to 64% (Batu Pageh) of the estimated macaque population size, with the average collection representing 37% of the estimated population size. A linear regression comparing prevalence against sampling effort, to confirm no effect of over or under-sampling of an individual population, found no significant effect ($R^2 = 0.086$, $p = 0.843$). Only fresh, non-dry samples were collected and, whenever possible, when defecation was observed. Collections were completed during a 1 day collection period for each site in order to minimize duplication of sampling from individual macaques and to minimize the risk of pseudoreplication (Paterson and Lello 2003). Samples were partitioned for immediate examination for helminths (fresh samples) and for examination of protozoa stages (preserved). Samples were preserved in 10% formalin before being shipped to the United States for further analysis (Garcia 1999; Chapman et al. 2005a, b; Trejo-Macias et al. 2007). No necropsies of any macaques were performed to determine the number of individual parasites present.

Quantification of helminth infective stages was completed in the field immediately after daily collections. Fecal flotations were performed using a Sheather's solution with

a specific gravity of 1.23. Because of the limitations of equipment accessible while in the field, in lieu of centrifugation, samples were mixed by inversion and left for a minimum of 30 min before being analyzed (Zajac and Conboy 2006). Parasites were identified by color, shape, symmetry, and size (Foreyt 2001); identification is tentative below the genus level. Measurements to the nearest 0.1 μm were made with a field micrometer. Egg counts were measured as the number of eggs or infective stages per gram of feces (Table 1). Quantification of protozoa infective stages was performed on samples preserved in formalin using Wheatley's modified trichrome stain. For each sample, approximately 500 fields per sample were examined, with counts averaged across five slides (Garcia 1999).

Prevalence and mean intensity of infection of each parasite identified are reported for both the entire island and for each individual macaque population. Further, because of interest in the effect of the environment on overall macaque "health", prevalence and mean intensity were also reported for the community of helminth parasites, the community of protozoa parasites, and the gut parasite community overall, again for the entire island and for each individual population. Prevalence was reported as the number of infected macaques per sampled subset of the population (Margolis et al. 1982; Rozsa et al. 2000; Chapman et al. 2005a, b; Greiner and McIntosh 2009). Mean intensity was defined as the average number of infective stages, as described in Table 1, among the infected members of each population. Parasite species diversity was measured using the Shannon–Weiner diversity index (SDI), a standard ecological measure of species diversity, which combines species richness and abundance (Spellerberg and Fedor 2003; Trejo-Macias et al. 2007). This was accomplished using the "vegan" package in R (R Core Development Team 2009). Table 1 lists the stage identified and counted for each parasite in this analysis. Quantification of diagnostic stage for each parasite was used, in lieu of the gold standard of necropsy and quantification of individual parasites, because of the large-scale census of wild non-human primates as hosts in this research. Preliminary genotyping of approximately 60% of our samples across 13 microsatellite loci supports our assumption of single samples of feces representing individual macaques (data not presented).

Measures of environmental variables were made directly or using a previously collected GIS dataset (Southern 2002). More specifically, measurements of forest, rice agriculture, and urban area were measured as the amount (m^2) of contiguous forest, rice, or urban area surrounding each macaque population. Elevation was measured directly in 2007, using GPS, in the central area of each temple associated with each macaque population. Population size

Table 1 Stages and methods used for parasite identification

Parasite	Stage identified	Method	Citation
Hookworm	Eggs	Sugar flotation	Varadharajan and Pythal (1999), Stuart et al. (1998)
<i>Ascaris</i> spp.	Fertilized egg	Sugar flotation	Varadharajan and Pythal (1999), Stuart et al. (1993)
<i>Taenia</i> spp.	Egg, only in presence of proglottids	Sugar flotation	Garcia (1999)
<i>Trichostrongylus</i> spp.	Egg	Sugar flotation	Ekanayake et al. (2006), Varadharajan and Pythal (1999)
<i>Enterobius</i> spp.	Eggs	Sugar flotation	Ekanayake et al. (2006)
<i>Trichuris</i> spp.	Eggs	Sugar flotation	Singh et al. (2009)
<i>Paragonimus</i> spp.	Eggs	Sugar flotation	Varadharajan and Pythal (1999), Stuart et al. (1998)
<i>Alaria</i> spp.	Eggs	Sugar flotation	Mohl et al. (2009)
<i>Entamoeba</i> spp.	Cysts, but noted trophozoites when found	Trichrome stain	Kumagai et al. (2001), Ekanayake et al. (2006), Stuart et al. (1998)
<i>Giardia</i> spp.	Cysts, but noted trophozoites when found	Trichrome stain	Stuart et al. (1993), Salzer et al. (2007)
<i>Endolimax</i> spp.	Cysts, but noted trophozoites when found	Trichrome stain	Ekanayake et al. (2006)
<i>Isoospora</i> spp.	Sporulated cysts	Trichrome stain	Stuart et al. (1993)
<i>Iodamoeba</i> spp.	Cysts, but noted trophozoites when found	Trichrome stain	Ekanayake et al. (2006)
<i>Cryptosporidia</i> spp.	Sporulated oocysts	Lugol's iodine	Ekanayake et al. (2006), Salzer et al. (2007)
<i>Balantidium</i> spp.	Cysts	Trichrome stain	Ekanayake et al. (2006), Stuart et al. (1998)
<i>Blastocystis</i> spp.	Cysts	Trichrome stain	Sheehan et al. (1986)
<i>Trichomonas</i> spp.	Cysts, but noted trophozoites when found	Trichrome stain	Stuart et al. (1993)
<i>Retortomonas</i> spp.	Cysts, but noted trophozoites when found	Trichrome stain	Stuart et al. (1993)

was determined by repeated censuses taken throughout 2002–2005 (Fuentes et al. 2005) and again, in 2007, before and after collection of feces. Water days is a measure of the number of days per year that water is readily available for drinking. This was determined by direct observation of water sources, interviewing native Balinese (Lane et al. 2011), and from published datasets of annual rainfall (Kripalani and Kulkarni 1998). Tourism days and offering weight were also determined by direct observation and by interviewing native Balinese, including temple staff (Lane et al. 2011). Specific measurements made at each population are listed in Table 2.

To identify any significant effect of the environment on levels of macaque parasitism, principle-components analysis (PCA) was performed, using R, on eight measures of the macaque environment: the number of days water is available (water days), a measure of food provisioning (offering weight), the number of days of tourism (tourism days), elevation, population size, and the number of m² of rice agriculture, urban area, and forest patch in and around a 100 ha buffer zone for each macaque population (Table 2). Linear regression analysis was then performed to compare principal components with standard deviations greater than 1 (Table 3) with parasite prevalence, mean

intensity, and parasite species diversity across populations. Regression analysis was also performed in R to compare parasite community prevalence, mean intensity, and parasite species diversity, across populations, to principal components (R Core Development Team 2009).

Results

Descriptive statistics

Prevalence of helminths across the island ranged from 0.012 (*Trichuris* spp.) to 0.137 (*Taenia* spp.), with the prevalence of the community of helminths on the island equaling 0.312 (Table 4). Prevalence of protozoa across the island ranged from 0.017 (*Retortomonas* spp.) to 0.632 (*Entamoeba* spp.), with the prevalence of the community of protozoa on the island equaling 0.786 (Table 4). Island-wide infection intensity peaked at 73.2 cysts (*Giardia* spp. in Angseri (AS)). Across the island, helminth infection intensity ranged from 3.5 eggs (*Trichostrongylus* spp.) to 12.5 eggs (*Taenia* spp.) and protozoa infection intensity ranged from 2.7 cysts (*Balantidium* spp.) to 35.6 cysts (*Giardia* spp.). Protozoa community intensity across the

Table 2 Environmental measurements, for each macaque population, that contributed to the principal-components analysis

Population	Water days	Tourism days	Offering weight (kg/day)	Forest (m ²)	Rice (m ²)	Urban (m ²)	Population size	Elevation (m)
Alas Nengahn (AN)	244	365	75	510781	724907	32233	50	109
Angseri (AS)	244	91	0.4	536319	123203	102191	40	730
Bedugal (BD)	244	244	10	536319	0	9473	200	1387
Bukit Gumang (BG)	91	244	10	0	131	0	100	213
Batu Pageh (BP)	91	91	0.4	0	0	6546	45	270
Batur (BT)	91	244	33.5	0	0	0	25	1717
Cekik (CK)	91	91	0.8	977641	0	0	50	647
Lempuyang (LM)	365	91	2.5	180053	0	127074	60	754
Mekori (MK)	365	91	1.3	566899	298983	87969	60	630
Pulaki (PL)	244	244	40	106198	0	34488	200	679
Ubud (PU)	365	365	100	36002	807051	72013	400	62
Tejakula (TK)	91	91	5	417946	114393	23372	75	70
Uluwatu (U)	91	365	60	0	0	23372	300	80

Table 3 PCA results: components 1–3 explain almost 85% of the variation in the environment, and, as such, were used in the remaining analyses

	Component 1	Component 2	Component 3	Component 4
Standard deviation	1.8468642	1.4676169	1.0953147	0.76643758
Prop. of variance	42.63	26.92	14.99	7.34
Cumulative variance	42.63	69.56	84.56	91.89
Loading scores				
Forest	0.236	0.329	0.509	0.402
Rice	0.344	0.313	0.333	0.443
City		-0.589		0.310
Water days	-0.161	-0.609		0.310
Tourism days	-0.453	0.240	-0.352	
Offering weight	-0.506		-0.223	
Elevation	0.286		-0.620	0.642
Population size	-0.498		-0.241	-0.229

Component 1 is significantly affected by “anthropogenic effects” or number of tourism days, offering weight, and macaque population size. Component 2 is affected by a combination of anthropogenic and non-anthropogenic effects, largely comprising the amount of city surrounding each temple and the number of days with available water. And, finally, Component 3 is driven largely by non-anthropogenic components of the environment—the amount of forest surrounding macaque populations and elevation

island was 67.5 cysts, with the protozoa community dominating infection of macaques across the island, resulting in over five times as many protozoa stages found than helminth stages (Table 5).

Principal-components analysis

The strength of the environmental variables and their interaction effects on the strength of parasitic infection in the long-tailed macaques of Bali was determined by PCA. Three principle components were found to play a significant role in partitioning environmental variation in this system (Table 3). The first component (PC1) is largely

determined by the strength and interaction effects of the amount of tourism, food offering weights, macaque population size, and the interaction effects of these variables. These are largely anthropogenic effects, with macaque population size being heavily affected by the amount of food provisioning the population receives. PC1 is negative in direction, suggesting that any change in variables related to this effect will be related to the decreasing amount of overall anthropogenic effect of the landscape. Or, more simply, any significant increase in parasitism related to an increase in PC1 is suggestive of a decrease in anthropogenic landscape effects. The second component (PC2) largely comprises the effects of the area of urbanization

Table 4 Parasite prevalence: prevalence of each parasite identified and for the community of parasites across the island and within individual macaque populations

	Island (n = 468)	AN (n = 26)	AS (n = 16)	BD (n = 36)	BG (n = 44)	BP (n = 29)	BT (n = 13)	CK (n = 19)	LM (n = 16)	MK (n = 34)	PL (n = 47)	PU (n = 115)	TK (n = 25)	U (n = 49)
Helminth														
<i>Ancylostoma</i>	0.058	0	0.69	0.03	0.05	0.03	0	0.05	0.13	0	0.04	0.04	0	0.04
<i>Ascaris</i>	0.06	0	0.44	0	0.09	0.03	0.08	0.05	0.25	0	0.04	0.02	0.04	0.06
<i>Taenia</i>	0.137	0.04	0.75	0.08	0.05	0	0.31	1	0.19	0.06	0.04	0.14	0.08	0
<i>Trichostrongylus</i>	0.035	0.12	0	0.03	0	0	0.08	0.16	0.06	0.03	0.09	0.02	0	0
<i>Enterobius</i>	0.064	0.15	0	0.03	0.14	0.1	0.08	0.11	0.19	0.06	0.02	0.04	0	0.08
<i>Trichuris</i>	0.012	0	0	0	0	0.03	0	0	0	0.03	0	0.008	0.12	0
<i>Paragonimus</i>	0.056	0	0	0	0.02	0	0.38	0	0.31	0	0.04	0.02	0.48	0
<i>Alaria</i>	0.023	0	0	0	0	0	0.38	0	0.25	0	0	0	0.08	0
Protozoa														
<i>Entamoeba</i>	0.632	0.62	0.94	0.47	0.55	0.79	0.77	0.89	0	0.26	0.7	0.68	1	0.65
<i>Giardia</i>	0.609	0.5	0.81	0.47	0.52	0.76	0.62	1	0.81	0.5	0.47	0.67	0.96	0.34
<i>Isospora</i>	0.137	0	0	0.03	0.02	0.14	0.62	0	0.13	0.21	0.32	0.17	0.36	0
<i>Endolimax</i>	0.301	0	0.25	0.22	0.36	0.52	0.62	0.58	0.88	0.03	0.34	0.33	0.6	0
<i>Iodamoeba</i>	0.096	0	0	0.06	0	0.14	0.38	0.37	0.38	0.03	0.06	0.11	0.2	0
<i>Cryptosporidia</i>	0.249	0.04	0.75	0	0	0.41	0.92	0	0.25	0.53	0.45	0.34	0	0
<i>Balantidium</i>	0.044	0	0	0.06	0	0.03	0.15	0.05	0	0	0.11	0.07	0.08	0
<i>Blastocystis</i>	0.031	0	0	0.03	0	0	0	0.05	0	0	0.11	0.06	0.04	0
<i>Trichomonas</i>	0.056	0.04	0.69	0.03	0.02	0.03	0	0	0	0	0.02	0.08	0.04	0
<i>Retortamonas</i>	0.017	0	0	0.06	0	0	0	0	0	0.03	0.06	0.02	0	0
Helminth community	0.312	0.231	0.813	0.139	0.227	0.172	1	0.947	0.686	0.176	0.213	0.235	0.56	0.163
Protozoa community	0.786	0.769	1	0.667	0.682	0.793	1	1	0.938	0.559	0.723	0.869	1	0.673
Gut parasite community	0.831	0.885	1	0.694	0.773	0.862	1	1	0.938	0.676	0.787	0.869	1	0.755

Table 5 Parasite mean intensity for each parasite identified and for the community of parasites reported island-wide and for individual macaque populations

Island (n = 468)	AN (n = 26)	AS (n = 16)	BD (n = 36)	BG (n = 44)	BP (n = 29)	BT (n = 13)	CK (n = 19)	LM (n = 16)	MK (n = 34)	PL (n = 47)	PU (n = 115)	TK (n = 25)	U (n = 49)
Helminth													
<i>Ancylostoma</i>	10.1	16.8	6	7.5	4		15	1		2.5	3.4		11
<i>Ascaris</i>	7.8	15.3		6.8	2	3	6	6.5		7	2.5	3	3
<i>Taenia</i>	12.5	11.7	4.3	5.5		7.8	24.2	6.3	5	3.5	5.5	20.3	
<i>Trichostrongylus</i>	3.5	4	7			4	4.7	2	3	2.25	2.5		
<i>Enterobius</i>	3.8	3.3	2	3.5	4.7	7	3	2.3	3.5	3	3.6		3.5
<i>Trichuris</i>	4				2				4		2	5.3	
<i>Paragonimus</i>	10.5			2		6.8		5.8		4.5	2.5	17.08	
<i>Alaria</i>	5.7					7		3.8				6.5	
Protozoa													
<i>Entamoeba</i>	26.8	19.8	12.6	19.7	44.5	18.9	33.3		25.7	21.64	30	36.9	24.56
<i>Giardia</i>	35.6	15.8	11.4	30.6	29.4	24	55.8	84.7	23.3	18.41	39.4	66.8	23.59
<i>Isospora</i>	6.3		2	3	3.3	4.4		4	5.9	7.87	8.3	4	
<i>Endolimax</i>	18.3	4.8	15.8	10.6	10.6	16.8	19.5	39	9	32.69	13.1	19.8	
<i>Iodamoeba</i>	6.9		5		6.8	3.6	5.7	8.7	18	3.67	9.9	2.6	
<i>Cryptosporidia</i>	26.9	17		29.7	20	52.9		27.3	21.1	20.95	22.4		
<i>Balantidium</i>	2.7		4		1	4	8			2	2.3	2.5	
<i>Blastocystis</i>	3.3		2				2			3.2	3.9	2	
<i>Trichomonas</i>	6.5	7	1	2	3					1	4.3	1	
<i>Retortamonas</i>	4.6		11.5						3	1.67	3		
Helminth community	12.5	4.8	5.6	7.6	4.4	8.8	27.8	9.1	4	4.7	5.2	19.7	5.6
Protozoa community	67.5	27.4	30.5	45.1	91.9	93.2	96.5	121.1	69.3	65.9	71.2	115.2	39.5
Gut parasite community	68.7	25.1	30.4	42	85.4	101.9	122.9	127.7	58.2	61.9	72.6	126.2	36.5

(m²), the amount of days of available water around the monkey forest, and their interaction effects. This component is, thus, dominated by a mixed anthropogenic and non-anthropogenic ecological effect. Finally, the third significant component (PC3) comprises the area of surrounding forest (m²), population elevation, and the interaction effects between these variables. This final component is therefore dominated by non-anthropogenic ecological effects. Both PC2 and PC3 are positive in direction, unlike PC1. Therefore, any significant increase in parasitism related to PC2 or PC3 is related to an increase in mixed or non-anthropogenic landscape effects, respectively.

Environmental effects on the parasite community

By using linear regression analysis we were able to identify how the three principal components (i.e. PC1: the anthropogenic component, PC2: the mixed environment component, and PC3: the non-anthropogenic ecological component) affect each individual parasite's prevalence and intensity (Table 6) and how they affect the gastrointestinal parasite community's prevalence, intensity, and diversity (Table 7). It is important to remember when interpreting these results that because the PC1 trend is negative, a positive relationship involving PC1 is representative of the negative effect of that particular component.

Among the 18 parasites identified, the prevalence of only three was shown to be significantly related to any environmental component. Specifically, the prevalence of *Giardia* spp. and *Endolimax* spp. were positively related to the anthropogenic effects, or PC1 ($p = 0.008$ and 0.037 , respectively), and the prevalence of *Entamoeba* spp. was positively related to the mixed landscape, or PC2 ($p = 0.049$). Thus, the prevalence of *Giardia* spp. and *Endolimax* spp. are lessened with increased anthropogenic alteration to the landscape whereas the prevalence of *Entamoeba* spp. is enhanced in a mixed landscape. Unlike prevalence, however, the intensity of eight of the eighteen parasites identified showed a significant relationship with PC1, PC2, or PC3. Four were positively related to PC1: the intensity of *Entamoeba* spp. ($p = 0.007$), *Giardia* spp. ($p = 0.041$), *Isospora* spp. ($p = 0.022$), and *Blastocystis* spp. ($p = 0.017$) was lessened by increasing effects of anthropogenic alteration of the landscape. *Iodamoeba* spp. and *Trichomonas* spp. were both positively related to a mixed landscape, or PC2 ($p = 0.028$ and 0.013 , respectively). And, finally, *Cryptosporidia* spp. and *Retortomonas* spp. ($p = 0.049$ and 0.001 , respectively) had a positive relationship with PC3, or the non-anthropogenic effects of the landscape. Interestingly, these results demonstrate that for 15 out of the 18 gastrointestinal parasites

Table 6 Individual parasite to principal-component regression analysis: below are the R^2 and p values from linear regression analysis comparing individual parasite prevalence and mean intensity to PC1 (the anthropogenic effect), PC2 (the mixed effect), and PC3 (the non-anthropogenic effect)

	PC1		PC2		PC3	
	R^2	p	R^2	p	R^2	p
Individual parasite prevalence						
<i>Ancylostoma</i> spp.	0.019	0.622	0.161	0.138	0.070	0.340
<i>Ascaris</i> spp.	0.0001	0.999	0.047	0.439	0.041	0.472
<i>Taenia</i> spp.	0.192	0.102	0.004	0.817	0.070	0.341
<i>Trichostonglyus</i> spp.	0.002	0.882	0.006	0.792	0.0001	0.970
<i>Enterobius</i> spp.	0.010	0.718	0.0001	0.965	0.051	0.418
<i>Trichuris</i> spp.	0.038	0.487	0.017	0.644	0.202	0.092
<i>Paragonimus</i> spp.	0.093	0.270	0.025	0.575	0.002	0.868
<i>Alaria</i> spp.	0.091	0.273	0.013	0.690	0.069	0.346
<i>Entamoeba</i> spp.	0.054	0.404	0.267	0.049	0.114	0.218
<i>Giardia</i> spp.	0.432	0.008	0.001	0.935	0.042	0.463
<i>Isospora</i> spp.	0.040	0.473	0.081	0.302	0.0001	0.981
<i>Endolimax</i> spp.	0.293	0.037	0.031	0.531	0.001	0.934
<i>Iodamoeba</i> spp.	0.144	0.163	0.021	0.609	0.040	0.475
<i>Cryptosporidia</i> spp.	0.082	0.301	0.008	0.751	0.028	0.554
<i>Balantidium</i> spp.	0.015	0.661	0.200	0.094	0.024	0.585
<i>Blastocystis</i> spp.	0.014	0.670	0.012	0.700	0.0001	0.940
<i>Trichomonas</i> spp.	0.014	0.669	0.088	0.284	0.011	0.707
<i>Retortomonas</i> spp.	0.0001	0.956	0.066	0.783	0.019	0.230
Individual parasite intensity						
<i>Ancylostoma</i> spp.	0.113	0.343	0.023	0.675	0.020	0.700
<i>Ascaris</i> spp.	0.050	0.472	0.300	0.081	0.132	0.272
<i>Taenia</i> spp.	0.235	0.110	0.062	0.435	0.010	0.755
<i>Trichostonglyus</i> spp.	0.157	0.291	0.157	0.291	0.197	0.231
<i>Enterobius</i> spp.	0.009	0.773	0.211	0.133	0.005	0.832
<i>Trichuris</i> spp.	0.240	0.510	0.002	0.958	0.067	0.741
<i>Paragonimus</i> spp.	0.190	0.388	0.019	0.796	0.211	0.360
<i>Alaria</i> spp.	0.457	0.527	0.990	0.063	0.050	0.857
<i>Entamoeba</i> spp.	0.694	0.007	0.006	0.790	0.028	0.566
<i>Giardia</i> spp.	0.305	0.041	0.069	0.366	0.017	0.656
<i>Isospora</i> spp.	0.553	0.022	0.122	0.358	0.029	0.659
<i>Endolimax</i> spp.	0.0001	0.994	0.022	0.662	0.007	0.803
<i>Iodamoeba</i> spp.	0.041	0.574	0.474	0.028	0.0001	0.970
<i>Cryptosporidia</i> spp.	0.089	0.472	0.243	0.214	0.504	0.049
<i>Balantidium</i> spp.	0.182	0.340	0.004	0.895	0.237	0.268
<i>Blastocystis</i> spp.	0.884	0.017	0.721	0.069	0.035	0.762
<i>Trichomonas</i> spp.	0.018	0.729	0.609	0.013	0.007	0.833
<i>Retortomonas</i> spp.	0.180	0.576	0.147	0.617	0.998	0.001

Significant values are highlighted

identified, there is no significant environmental effect on prevalence in the macaques across the island. Only for *Giardia* spp. was there a significant relationship—both prevalence and intensity increase as populations have less

Table 7 Community parasitism regression analysis: below, we report the R^2 and p values from linear regression analysis performed to compare helminth, protozoa, and total gut parasite community prevalence, intensity, and diversity to PC1 (the anthropogenic effect), PC2 (the mixed effect), and PC3 (the non-anthropogenic effect)

	PC1		PC2		PC3	
	R^2	p	R^2	p	R^2	p
Community prevalence						
Helminth	0.211	0.085	0.001	0.926	0.067	0.350
Protozoa	0.197	0.097	0.025	0.574	0.016	0.648
Gut parasite	0.246	0.060	0.013	0.685	0.057	0.391
Community intensity						
Helminth	0.135	0.179	0.003	0.857	0.009	0.733
Protozoa	0.328	0.026+	0.006	0.787	0.002	0.876
Gut parasite	0.307	0.032+	0.006	0.782	0.001	0.910
Community diversity						
Helminth	0.008	0.757	0.003	0.847	0.005	0.800
Protozoa	0.103	0.244	0.029	0.547	0.012	0.701
Gut parasite	0.041	0.468	0.037	0.491	0.074	0.327

Significant values are highlighted

tourism, less provisioning, and smaller population sizes. For no individual helminth was there a significant relationship of either prevalence or intensity with any component of the landscape.

At the total parasite community level, prevalence or parasite species diversity was not significantly related with any component of the landscape. However, protozoa community intensity and total gut parasite community intensity were both positively related to PC1, or the anthropogenic effect of the landscape ($p = 0.026$ and 0.032 , respectively), suggesting that tourism, provisioning, and macaque population size can act to mitigate the intensity of gastrointestinal parasite infections. No pattern emerged relating PC2 or PC3 to community level intensity.

Discussion

In wild systems, no single component of the environment acts truly independently. However, recognizing how seemingly independent elements interact is crucial in understanding how anthropogenically altered landscapes affect the dynamics of the wildlife existing in that environment. Here, we used principal-components analysis to determine the strength of the interaction effects of the environmental elements in this system. With this, we were able to identify how specific elements of the environment and their interactions affect the prevalence, intensity, and diversity of the gastrointestinal parasites of the long-tailed macaques of Bali.

The success of directly transmitted parasites in primates can be positively affected by high host population size, density, and group size, and evidence demonstrates that the larger the host population size, the stronger and more diverse the infection (Freeland 1979; Stuart et al. 1993; Cote and Poulin 1995; Arneberg et al. 1998; Morand 2000; Roberts et al. 2002; Poulin and Morand 2004; Bagge et al. 2004; Chapman et al. 2005a, b; but see Snaith et al. 2008). This trend is often linked to the highly social nature of many primate species (Fuentes 2007) with the epidemiology of gut parasites enabling rapid individual to individual transmission within the population (Brown and Brown 1986; Moller 1993). In Bali, however, macaque population growth is often tied to provisioning rate in individual monkey forests. Previous work has demonstrated that increases in human contact leads to increased exposure to parasites, especially gastrointestinal parasites (Sleeman et al. 2000; Nunn and Altizer 2006). However, in Bali, the increase in human contact is largely with a small number of populations, with recent increases in Australian, European, and American tourism. It is, therefore, unlikely to be the source of macaque gut parasite transmission. Instead of an increasing parasite burden with a high population size, our results demonstrate that macaques are in high contact with tourists, and so are heavily provisioned, and are, as a result, thriving in large populations, with lessened prevalence and intensity of gastrointestinal parasites.

Variability in access to water can also increase the risk of exposure to parasites (Rendtorff and Holt 1954; Sogandaresbernal 1955; Combes 2001). Contact with slow-moving or standing water sources often increases exposure to many infective stages of parasites associated with human and animal feces (Wang and Dunlop 1954; Combes 2001; Roberts et al. 2002; Nunn and Altizer 2006). Host populations living in areas near multiple water sources should be exposed to a higher number and variety of parasites relative to their host counterparts with little access to year-round water sources. However, our results demonstrate water availability, in conjunction with the amount of urban area surrounding a monkey forest, does not directly limit or enhance the prevalence or intensity of the majority of gastrointestinal parasites, with only *Entamoeba* spp. linked with an increase in prevalence and *Iodamoeba* spp. and *Trichomonas* spp. linked with an increase in intensity in populations with high water availability and urban contact. Further, at the community level, no pattern emerged when comparing water availability and proximity to high-density human populations and gut parasite prevalence, intensity, or diversity.

Primates living in complex habitats have been shown to be infected by a greater intensity and diversity of parasites; however, this trend is confounded by the great variety in diet of these primates (Poulin 1997; Nunn et al. 2003; Nunn

and Altizer 2006). In parasites with multi-stage life cycles, or with infective stages transmitted through the environment, the habitat of the host–parasite interaction plays a more direct role in the transmission of that parasite (Combes 1996; Grutter 1998; Nunn and Altizer 2006). Poor quality habitat and disturbed habitats are much more likely to harbor a greater intensity and diversity of parasites relative to host populations in optimum habitats (Poulin 1997; Morand 2000; Nunn and Altizer 2006). In a complex anthropogenic landscape, such as the island of Bali, Indonesia, the interaction between specific elements of the environment and system dynamic components, such as tourism and provisioning, can act in concert to minimize the prevalence and intensity of parasitic infection across populations.

In the context of long-term, anthropogenic changes elemental to the Balinese landscape, the patterns emerging suggest a more refined interpretation of the interface between host, parasite, and environment. In anthropogenic landscapes, heavily provisioned host populations have reduced exposure to the infective stages of many parasites. Traditionally, non-human primates with diverse diets have the greatest parasite prevalence, intensity, and diversity; however, the large-scale provisioning afforded the Balinese macaques, especially in areas of high tourism, reduces the need to forage and, ultimately, the likelihood of exposure. Thus, despite human populations being a recognized source of pathogen exchange (Jones-Engel et al. 2004), they can also minimize parasite infection. Moreover, this human provisioning can provide enough nutritional benefit to wild populations to enable substantial population growth. How this macaque–human interaction has limited or enhanced gene flow opportunities through this landscape, and thus parasite transmission, remains unanswered. Our data imply overall lessening of several measures of gut parasitism in macaque populations that are in regular, close-contact with often high-density human populations. Nevertheless, the complexity of host–parasite interactions warrants additional investigation. Seasonality in parasite infection, for example, is a potentially confounding variable, and following one or more of these populations over the course of the year would help us identify the importance of these anthropogenic and environmental effects on the long-term success of the parasites identified here (Legesse and Erko 2004). Beyond this, egg counts have been found to be potentially unreliable measures of the level of infection (McKenna 1981). However, egg counts have been used reliably both as an estimate of burden and as a reliable measure of parasite presence, especially when examining wild, non-human primate populations (Trejo-Macias et al. 2007; Ekanayake et al. 2006; Ezenwa 2006; Salzer et al. 2007).

Our results demonstrate that an increasingly anthropogenic environment can lead to an overall lessening of the

intensity and diversity of gut parasites in wild macaque populations. This contradicts previous findings linking increases in parasitism with increases in human contact and interactions (Nunn and Altizer 2006). It is likely that these results are due to the type of human contact—in Bali, the high interaction rate results in a benefit to macaques via provisioning. Within this context, it will be important to consider the anthropogenic environment when analyzing primate–parasite dynamics. The anthropogenic landscape continues to change these dynamics, and indeed our expectations, of the opportunities for transmission of parasites of non-human primates, especially those parasites reliant on the environment for transmission success.

Acknowledgments We would like to thank Udayana Universitas and the Pusat Kajian Primata. We would also like to thank Rachel Polando, Natalie Griffiths, Sean Hoban, Janine Reugg, and three anonymous reviewers for their helpful comments and insights. This work was supported by the National Science Foundation (BSC-0629787), the University of Notre Dame’s Institute for Scholarship in the Liberal Arts, and the Leakey Foundation. All collections were approved by the University of Notre Dame IACUC (protocol 07-001) and the Indonesian Institute of Science (permit number 662.02/1090.DIII).

References

- Anderson DR, Beaudoin RL (1966) Host habitat and age as factors in the prevalence of intestinal parasites of the muskrat. *Bull Wildl Dis Assoc* 2:70–77
- Arneberg P, Skorpung A, Grenfell B, Read AF (1998) Host densities as determinants of abundance in parasite communities: comparative analyses of directly transmitted nematodes of mammals. *Ecography* 25:88–94
- Bagge AM, Poulin R, Valtonen ET (2004) Fish population size, and not density, as the determining factor of parasitic infection: a case study. *Parasitology* 128:305–313
- Baird JK (2003) Host age as a determinant of naturally acquired immunity in *Plasmodium falciparum*. *Parasitol Today* 11:105–111
- Baudoin M (1975) Host castration as a parasitic strategy. *Evolution* 29(2):335–352
- Brown CR, Brown MB (1986) Ectoparasitism as a cost of coloniality in cliff swallows (*Hirundo pyrrhonota*). *Ecology* 67:1206–1218
- Chapman CA, Gillespie TR, Goldberg TL (2005a) Primates and the ecology of their infectious diseases: how will anthropogenic change affect host–parasite interactions? *Evol Anthropol* 14:134–144
- Chapman CA, Gillespie TR, Speirs ML (2005b) Parasite prevalence and richness in sympatric colobines: gastrointestinal parasites from forest edge and interior primate groups. *Am J Primatol* 68:1–12
- Combes C (1996) Parasites, biodiversity and ecosystem stability. *Biodivers Conserv* 5:953–962
- Combes C (2001) Parasitism: the ecology and evolution of intimate interactions. University of Chicago Press, Chicago
- Cote IM, Poulin R (1995) Parasitism and group size in social mammals: a meta-analysis. *Behav Ecol* 6:159–165
- Daszak P, Cunningham AA, Hyatt AD (2001) Anthropogenic environmental change and the emergence of infectious diseases in wildlife. *Acta Trop* 78:103–116

- Dobson AP (1988) The population biology of parasite-induced changes in host behavior. *Q Rev Biol* 63:139–165
- Ekanayake DK, Arulkanthan A, Horadagoda NU, Sanjeevani GKM, Kieft R, Gunatilake S, Dittus WPJ (2006) Prevalence of *Cryptosporidium* and other enteric parasites among wild non-human primates in Polonnaruwa, Sri Lanka. *Am J Trop Med* 74:322–329
- Ezenwa VO (2006) Host social behavior and parasitic infection: a multifactorial approach. *Behav Ecol* 15:446–454
- Ferrari N, Cattadori IM, Nespereira J, Rizzoli A, Hudson PJ (2004) The role of host sex in parasite dynamics: experiments on the yellow-necked mouse *Apodemus flavicollis*. *Ecol Lett* 7:88–94
- Foreyt WJ (2001) *Veterinary parasitology: reference manual*, 5th edn. Blackwell, Oxford
- Freeland WJ (1979) Primate social groups as biological islands. *Ecology* 61:1297–1303
- Fuentes A (2007) Social organization: social systems and the complexities in understanding the evolution of primate behavior. In: Campbell CJ, Fuentes A, MacKinnon KC, Panger M, Bearder SK (eds) *Primates in perspective*. Oxford University Press, Oxford, pp 609–621
- Fuentes A (2009) A new synthesis: resituating approaches to the evolution of human behavior. *Anthropol Today* 25:12–17
- Fuentes A, Southern M, Suaryana KG (2005) Monkey forests and human landscapes: is extensive sympatry sustainable for *Homo sapiens* and *Macaca fascicularis* in Bali? In: Patterson J (ed) *Commensalism and conflict: the primate–human interface*. American Society of Primatology Publications, Norman
- Garcia LS (1999) *Practical guide to diagnostic parasitology*. American Society of Microbiology Press, Washington DC
- Gasser RB, de Grijter JM, Polderman AM (2009) The utility of molecular methods for elucidating primate–pathogen relationships—the *Oesophagostomum bifurcum* example. In: Huffman MA, Chapman CA (eds) *Primate parasite ecology: the dynamics and study of host–parasite relationships*. Cambridge University Press, Cambridge, pp 47–63
- Greiner EC, McIntosh A (2009) Collection methods and diagnostic procedures for primate parasitology. In: Huffman MA, Chapman CA (eds) *Primate parasite ecology: the dynamics and study of host–parasite relationships*. Cambridge University Press, Cambridge, pp 3–28
- Grutter AS (1998) Habitat-related differences in the abundance of parasites from a coral reef fish: an indication of the movement patterns of *Hemigymnus melapterus*. *J Fish Biol* 53:49–57
- Gurarie D, Seto EYW (2008) Connectivity sustains disease transmission in environments with low potential for endemicity. *J R Soc Interface* 6(35):495–508
- Hartwig W (2007) Primate evolution. In: Campbell CJ, Fuentes A, MacKinnon KC, Panger M, Bearder SK (eds) *Primates in perspective*. Oxford University Press, Oxford, pp 11–22
- Haufstater G, Meade BJ (1982) Alteration of sleeping groves by yellow baboons (*Papio cynocephalus*) as a strategy for parasite avoidance. *Primates* 23:287–297
- Hurd H (2001) Host fecundity reduction: a strategy for damage limitation? *Trends Parasitol* 17:363–368
- Jaffee B, Phillips R, Muldoon A, Mangel M (1992) Density-dependent host–pathogen dynamics in soil microcosms. *Ecology* 73:495–506
- Jones-Engel L, Engel GA, Schillaci MA, Kyes K, Froelich J, Papatungan U, Kyes RC (2004) Prevalence of enteric parasites in pet macaques in Sulawesi, Indonesia. *Am J Primatol* 62:71–82
- Keesing F, Holt RD, Ostfeld RS (2006) Effects of species diversity on disease risk. *Ecol Lett* 9:485–498
- Kripalani RH, Kulkarni A (1998) Rainfall variability over South-east Asia—connections with Indian monsoon and ENSO extremes: new perspectives. *Int J Climatol* 17:1155–1168
- Kumagai M, Koayashi S, Okita T, Ohtoma H (2001) Modifications of Kohn's chlorazol black E staining and Wheatley's trichrome staining for temporary wet mount and permanent preparation of *Entamoeba histolytica*. *J Parasitol* 87:701–704
- Lane KE, Lute M, Rompis A, Wandia IN, Arta Putra IG, Hollocher H, Fuentes A (2010) Pests, pestilence, and people: the long-tailed macaque and its role in the cultural complexities of Bali. In: Gursky-Doyen S, Supriatna J (eds) *Indonesian primates, developments in primatology: progress and prospects*. Springer, Berlin, pp 235–248
- Lane KE, Putra IGA, Wandia IN, Rompis A, Hollocher H, Fuentes A (2011) Balinese perceptions of pathogen exposure in human communities surrounding long-tailed macaque populations. *Ecol Environ Anthropol* (in press)
- Lee P (1999) *Comparative primate socioecology*. Cambridge University Press, Cambridge
- Legesse M, Erko B (2004) Zoonotic intestinal parasites in *Papio anubis* (baboon) and *Cercopithecus aethiops* (vervet) from four localities in Ethiopia. *Acta Tropica* 90:231–236
- Loudon J, Howell M, Fuentes A (2006) The importance of integrative anthropology: a preliminary investigation employing primatological and cultural anthropological data-collection methods in assessing human–monkey co-existence in Bali, Indonesia. *Ecol Environ Anthropol* 2(1):2–13
- Margolis L, Esch GW, Holmes JC, Kuris AM, Schad GA (1982) The use of ecological terms in parasitology (report of an ad hoc committee of the American Society of Parasitologists) *J Parasitol* 68:131–133
- McKenna P (1981) The diagnostic value and interpretation of fecal egg counts in sheep. *N Z Vet J* 29:129–132
- Mohl K, Grosse K, Hamedy A, Wuste T, Kabelitz P, Lucker E (2009) Biology of *Alaria* spp. and human exposition risk to *Alaria mesocercariae*—a review. *Parasitol Res* 105:1–15
- Moller AP (1993) A fungus infecting domestic flies manipulates sexual behavior of its host. *Behav Ecol Sociobiol* 33:403–407
- Morand S (2000) Wormy world: comparative tests of theoretical hypotheses on parasite species richness. In: Poulin R, Morand S, Skorping A (eds) *Evolutionary biology of host–parasite relationships: theory meets reality*. Elsevier, Amsterdam
- Mouritsen KN, Poulin R (2003) The mud flat anemone–cockle association: mutualism in the intertidal zone? *Oecologia* 135:131–137
- Niezen JH, Charleston WAG, Hodgson J, Miller CM, Waghorn TS, Robertson HA (2001) Effect of plant species on the larvae of gastrointestinal nematodes which parasitize sheep. *Int J Parasitol* 28:791–803
- Nunn CL, Altizer S (2006) *Infectious diseases in primates*. Oxford University Press, Oxford
- Nunn CL, Altizer S, Jones KE, Sechrest W (2003) Comparative tests of parasite species richness in primates. *Am Nat* 162:597–614
- Ostfeld RS, Glass GE, Keesing F (2005) Spatial epidemiology: an emerging (or reemerging) discipline. *Trends Ecol Evol* 20:328–336
- Paterson S, Lello J (2003) Mixed models: getting the best use of parasitological data. *Trends Ecol Evol* 19:370–375
- Patz JA, Graczyk TK, Geller N, Vittor AY (2000) Effects of environmental change of emerging parasitic diseases. *Int J Parasitol* 30:1395–1405
- Pederson AB, Fenton A (2006) Emphasizing the ecology in parasite community ecology. *Trends Ecol Evol* 22:133–139
- Poulin R (1997) Species richness of parasite assemblages: evolution and patterns. *Annu Rev Ecol Syst* 28:341–358
- Poulin R, Morand S (2004) *Parasite biodiversity*. Smithsonian Institution Press, Washington, DC
- R Development Core Team (2009) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. <http://www.R-project.org>

- Rendtorff RC, Holt CJ (1954) The experimental transmission of human intestinal protozoan parasites. 4. Attempts to transmit *Entamoeba coli* and *Giardia lamblia* cysts by water. *Am J Hyg* 60:327–338
- Roberts MG, Dobson AP, Arneberg P, de Leo GA, Kreczek RC, Manfredi MT, Lanfranchi P, Zaffaroni E (2002) Parasite community ecology and biodiversity. In: Hudson PJ, Rizzoli A, Grenfell BT, Hesterbeek H, Dobson AP (eds) *The ecology of wildlife diseases*. Oxford University Press, Oxford, pp 63–82
- Roepstorff A, Murrell KD (1997) Transmission dynamics of helminth parasites of pigs on continuous pasture: *Ascaris suum* and *Trichuris suis*. *Int J Parasitol* 27:563–572
- Rozsa L, Reiczigel J, Majoros G (2000) Quantifying parasites in samples of hosts. *J Parasitol* 86:228–232
- Salzer JS, Rwego IB, Goldberg TL, Kuhlenschmidt MS, Gillespie TR (2007) *Giardia* sp. and *Cryptosporidium* sp. infections in primates in fragmented and undisturbed forest in western Uganda. *J Parasitol* 93:439–440
- Schrag SJ, Weiner P (1995) Emerging infectious disease: what are the relative roles of ecology and evolution? *Trends Ecol Evol* 10:319–324
- Singh S, Shrivastav AB, Sharma RK (2009) The epidemiology of gastrointestinal parasitism and body condition in free-ranging herbivores. *Journal Threatened Taxa* 1:535–537
- Sleeman JM, Meader LL, Mudakikwa AB, Foster JW, Patton S (2000) Gastrointestinal parasites of mountain gorillas (*Gorilla gorilla beringei*) in the Parc National des Volcans, Rwanda. *J Zoo Wildl Med* 31:322–328
- Slifko TR, Smith HV, Rose JB (2000) Emerging parasite zoonoses associated with water and food. *Int J Parasitol* 30:1379–1393
- Snaith TV, Chapman CA, Rothman JM, Wasserman MD (2008) Bigger groups have fewer parasites and similar cortisol levels: a multi-group analysis in red colobus monkeys. *Am J Primatol* 70:1–9
- Sogandaresbernal F (1955) Some helminth parasites of fresh and brackish water fishes from Louisiana and Panama. *J Parasitol* 41:582–587
- Southern MW (2002) An assessment of potential habitat corridors and landscape ecology for long-tailed macaques (*Macaca fascicularis*) on Bali, Indonesia [dissertation]. Central Washington University, Ellensburg
- Spellerberg IF, Fedor PJ (2003) A tribute to Claude Shannon (1916–2001) and a plea for more rigorous use of species richness, species diversity, and the ‘Shannon–Wiener’ Index. *Glob Ecol Biogeogr* 12:177–179
- Stirnadel HA, Ebert D (1997) Prevalence, host specificity and impact on host fecundity of microparasites and epibionts in three sympatric *Daphnia* species. *J Anim Ecol* 66:212–222
- Stuart MD, Strier KB, Peirberg SM (1993) A coprological survey of parasites of wild miquis, *Brachyteles arachnoides*, and brown howling monkeys, *Alouatta fusca*. *J Helminthol Soc WA* 60:111–115
- Stuart M, Pendergast V, Rumpfelt S, Pierberg S, Greenspan L, Glander K, Clarke M (1998) Parasites of wild howlers (*Alouatta* spp.). *Int J Primatol* 19:493–512
- Thieltges DW, Jensen KT, Poulin R (2008) The role of biotic factors in the transmission of free-living endohelminth stages. *Parasitology* 135:407–426
- Thierry B (2007) The macaques: a double-layered social organization. In: Campbell CJ, Fuentes A, MacKinnon KC, Panger M, Bearder SK (eds) *Primates in perspective*. Oxford University Press, Oxford, pp 224–240
- Trejo-Macias G, Estrada A, Cabrera MAM (2007) Survey of helminth parasites in populations of *Alouatta palliata mexicana* and *A. pigra* in continuous and in fragmented habitat in southern Mexico. *Int J Primatol* 28:931–945
- van Schaik CP (1996) Social evolution in primates: the role of ecological factors and male behaviour. *Proc Br Acad* 88:9–31
- Varadharajan A, Pythal C (1999) Parasites of wildlife—I: a preliminary investigation on the parasites of wild animals at the Zoological Garden, Thiruvananthapuram, Kerala. *Zoos’ Print Journal* 14:159–164
- Wang WLL, Dunlop SG (1954) Animal parasites in sewage and irrigation water. *Sew Ind Wastes* 26:1020–1032
- Wolfe ND, Switzer WM (2009) Primate exposure and the emergence of novel retroviruses. In: Huffman MA, Chapman CA (eds) *Primate parasite ecology: the dynamics and study of host–parasite relationships*. Cambridge University Press, Cambridge, pp 353–370
- Wolfe ND, Escalante AA, Karesh WB, Kilbourn A, Spielman A, Lal AA (1998) Wild primate populations in emerging infectious disease research: the missing link? *EID* 4:149–158
- Zajac AM, Conboy GA (2006) *Veterinary clinical parasitology*, 7th edn. Blackwell–Wiley, Oxford

COMPARATIVE PHYLOGEOGRAPHY OF THREE CODISTRIBUTED STOMATOPODS: ORIGINS AND TIMING OF REGIONAL LINEAGE DIVERSIFICATION IN THE CORAL TRIANGLE

PAUL H. BARBER,^{1,2} MARK V. ERDMANN,^{3,4} AND STEPHEN R. PALUMBI^{5,6}

¹*Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138*

³*Center for Applied Biodiversity Science, Conservation International, 1919 M Street NW, Washington, DC 20036*

⁴*E-mail: m.erdmann@conservation.org*

⁵*Department of Biological Sciences, Stanford University, Hopkins Marine Station, Pacific Grove, California 93950*

⁶*E-mail: spalumbi@stanford.edu*

Abstract.—The Indonesian-Australian Archipelago is the center of the world's marine biodiversity. Although many biogeographers have suggested that this region is a "center of origin," criticism of this theory has focused on the absence of processes promoting lineage diversification in the center. In this study we compare patterns of phylogeographic structure and gene flow in three codistributed, ecologically similar Indo-West Pacific stomatopod (mantis shrimp) species. All three taxa show evidence for limited gene flow across the Maluku Sea with deep genetic breaks between populations from Papua and Northern Indonesia, suggesting that limited water transport across the Maluku Sea may limit larval dispersal and gene flow across this region. All three taxa also show moderate to strong genetic structure between populations from Northern and Southern Indonesia, indicating limited gene flow across the Flores and Java Seas. Despite the similarities in phylogeographic structure, results indicate varied ages of the genetic discontinuities, ranging from the middle Pleistocene to the Pliocene. Concordance of genetic structure across multiple taxa combined with temporal discordance suggests that regional genetic structures have arisen from the action of common physical processes operating over extended time periods. The presence in all three species of both intraspecific genetic structure as well as deeply divergent lineages that likely represent cryptic species suggests that these processes may promote lineage diversification within the Indonesian-Australian Archipelago, providing a potential mechanism for the center of origin. Efforts to conserve biodiversity in the Coral Triangle should work to preserve both existing biodiversity as well as the processes creating the biodiversity.

Key words.—Allopatric speciation, biodiversity, center of origin, Coral Triangle, larval dispersal, marine phylogeography, stomatopod.

Received October 24, 2005. Accepted May 24, 2006.

The oceans of Indonesia are home to the highest marine biodiversity in the world (Myers et al. 2000; Veron 2000; Hughes et al. 2002; Roberts et al. 2002). Although a number of theories attempt to explain this pattern (for a review see Barber and Bellwood 2005), recent studies (Briggs 2000; Mora et al. 2003) support the hypothesis that this region is a center of origin (Ekman 1953), with increased biodiversity arising from speciation within the waters of Indonesia, Philippines, and Papua New Guinea, a region collectively known as the Coral Triangle (sensu Green and Mous 2004). A center of origin requires that mechanisms must be acting to promote lineage diversification within the Coral Triangle, and critics of this hypothesis have highlighted the lack of a convincing mechanism (Jokiel and Martinelli 1992). Although allopatric differentiation is the most common model of speciation invoked for terrestrial organisms (Mayr 1942; Bush 1975; Lynch 1989; Turelli et al. 2001) the pelagic larval dispersal common in marine species results in few clear boundaries to dispersal and should limit opportunities for allopatric divergence (Palumbi 1992, 1994). Thus, high biodiversity in the Indonesian-Australian Archipelago (IAA) seems paradoxical.

The dynamic geologic history of the IAA, however, suggests ample opportunity for vicariance. During Pleistocene glacial maxima, sea levels dropped 130 m below present levels (Porter 1989; Voris 2000), exposing the Sunda and Sahul continental shelves and constricting the waterways of

the IAA, potentially leading to increased isolation and allopatric differentiation between populations spanning the multiple ocean basins of the IAA. Support for Pleistocene vicariance in the IAA comes from biogeographic studies demonstrating patterns consistent with Pleistocene vicariance (McManus 1985; Potts 1985; Fleminger 1986; Woodland 1986; Springer and Williams 1990) and from genetic studies that indicate Pleistocene speciation (e.g., McMillan and Palumbi 1995). Further support comes from studies that demonstrate intraspecific genetic breaks on broad spatial scales (>1000 km) among populations of fish (Lacson and Clark 1995), crabs (Lavery et al. 1996), prawns (Duda and Palumbi 1999), sea stars (Williams and Benzie 1998; Benzie 1999) and mangroves (Duke et al. 1998) that span the IAA. These phylogeographic studies emphasize the IAA as a barrier between two oceans, resulting in allopatric diversification during the Pleistocene. However, the broad-scale geographic sampling employed in these studies is too coarse to determine whether phylogeographic patterns across species are truly concordant, limiting our ability to rigorously examine the specific processes promoting lineage diversification in the IAA.

A robust test of the processes promoting lineage diversification in the IAA requires comparative population genetic studies of multiple taxa on fine spatial scales throughout the region. Such comparative studies have been performed in the Caribbean (Shulman and Bermingham 1995), Great Barrier Reef (Doherty et al. 1995), Western Atlantic (Avisé et al.

² Present address: Department of Biology, Boston University, Boston, Massachusetts 02215; E-mail: pbarber@bu.edu.

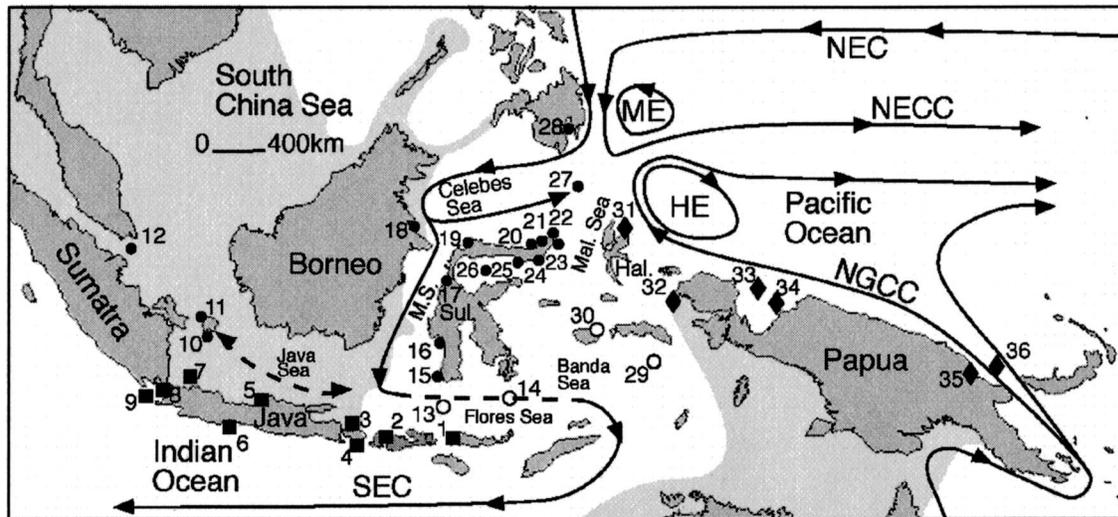


FIG. 1. Map of the Indo-West Pacific showing dominant (solid lines) and seasonally reversing currents (dashed lines), including the North Equatorial Current (NEC), North Equatorial Counter Current (NECC), New Guinea Coastal Current (NGCC), Mindanao Eddy (ME), Halmahera Eddy (HE), Makassar Strait (MS; the main passageway for the Indonesian Throughflow), and the South Equatorial Current (SEC). Squares indicate Southern Indonesia localities, closed circles indicate Northern Indonesia localities, and diamonds refer to Papua localities. Open circles indicate localities that could not be a priori assigned to one of the aforementioned regions. The islands of Halmahera and Sulawesi are identified by Hal. and Sul., respectively. Numbers identify localities described in Table 2. Light gray shading indicates coastal margins during Pleistocene low sea level stands, after Voris (2000).

1987), and French Polynesia (Fauvelot and Planes 2002), and Lourie et al. (2005) examined patterns of gene flow in brooding sea horses on the Sunda Shelf. However, there have been no comparative population genetic studies to date on species with pelagic larval dispersal on fine spatial scales in the center of the Coral Triangle. Such studies are critical to provide a robust test of the processes promoting lineage diversification in the IAA, improving our understanding of the origins of marine biodiversity in the IAA and of marine speciation in general.

Although sea-level fluctuations may promote lineage diversification in the IAA, this isolation may be compounded or mitigated by the physical oceanography of this region because of the ability of oceanic currents to affect dispersal through the transport of larvae entrained therein. Indonesia's oceanography is dominated by the Indonesian Throughflow (Wyrki 1987), a movement of 20 million m^3/sec of water from the Pacific to the Indian Ocean (Godfrey 1996; Gordon and Fine 1996) through the waters of Indonesia (Fig. 1) with associated currents that exceed 1 m/sec (Wyrki 1961). The vast majority (~90%) of the Indonesian Throughflow water comes from the North Equatorial Current (Nof 1995; Morey et al. 1999), providing a dispersal corridor to potentially link populations throughout much of this region. The remaining fraction comes from the New Guinea Coastal Current that travels north and west along the northern coast of the island of Papua (Papua New Guinea and Indonesian Papua). The New Guinea Coastal Current is deflected in the proximity of the island of Halmahera, forming the Halmahera Eddy (Fig. 1) and greatly limiting transport of New Guinea Coastal Current waters into the Indonesian Throughflow (Nof 1995; Morey et al. 1999). A powerful graphical representation of this isolation is illustrated in Lukas et al. (1991, p. 7092) in which surface drifter buoys that enter the Halmahera Eddy do not

cross the Maluku Sea, but are instead retroflected and transported eastward. Thus, although marine larvae are often active swimmers (Stobutzki and Bellwood 1997; Leis and Carson-Ewart 1997; Wolanski et al. 1997; Armsworth 2001) and can engage in behaviors (Cronin and Forward 1986; Tankersley and Forward 1994; Forward et al. 1997) that may result in limited dispersal (Jones et al. 1999; Swearer et al. 1999; Thorrold et al. 2001), the physical oceanography of the IAA may have a profound impact on larval dispersal, which in turn will impact gene flow and genetic cohesion throughout this region.

Counter to predictions of genetic homogeneity resulting from larval transport within the Indonesian Throughflow, Barber et al. (2000) found pronounced patterns of genetic differentiation among populations of the mantis shrimp *Haptosquilla pulchella* (Crustacea: Stomatopoda) throughout the IAA. Sampling on a scale of hundreds of kilometers revealed that the genetic break between Pacific and Indian Ocean populations was more precisely a break between populations north and south of the Java and Flores Seas that are separated by only 300–400 km. The discovery of a similarly large genetic break between populations from Papua New Guinea and Indonesia led Barber et al. (2002a) to suggest that although Pleistocene vicariance may contribute to patterns of genetic structure in marine populations of the IAA, other processes such as limited water/larval transport across the Maluku Sea and ecologically limited larval dispersal across the Flores and Java Seas may be driving regional diversification in this taxon. However, the validity of these hypotheses can only be evaluated through a rigorous comparative study of multiple ecologically similar species spanning this region.

In this paper, we compare patterns of genetic structuring in three Indo-Pacific stomatopod species. The congeneric species *Haptosquilla pulchella* and *H. glyptocercus*, as well as

TABLE 1. Egg size and postlarval settlement size for *Haptosquilla pulchella*, *H. glyptocercus*, *Gonodactylellus viridis*, and the western Atlantic gonodactyloid *Neogonodactylus oerstedii*. We have previously estimated the larval period for *H. pulchella* at four to six weeks (Barber et al. 2000, 2002a) based on the established facts that stomatopod larval development is highly regular and hence size is an appropriate proxy for age, the larval duration of the related species *N. oerstedii* has been experimentally determined as 35 days, and the three species in question hatch at a slightly smaller size yet recruit at a slightly larger size than *N. oerstedii* and hence are expected to have a slightly longer larval period. Asterisk indicates estimated values.

Species	Egg size (mm)	Size at settlement (mm)	Larval period	Source
<i>H. pulchella</i>	0.73–0.85	9.5–10	4–6 weeks*	Erdmann 1997
<i>H. glyptocercus</i>	0.60–0.78	7–9.5	4–6 weeks*	Erdmann 1997
<i>G. viridis</i>	0.63–0.95	5–9	4–6 weeks*	Erdmann 1997
<i>N. oerstedii</i>	1.04–1.30	6	35 days	Manning 1963; Provenzano and Manning 1978

the more distantly related *Gonodactylellus viridis* (Barber and Erdmann 2000), are three of the most common coral reef flat stomatopods found in the IAA (Erdmann 1997). Larvae of each of these species hatch from similar-sized eggs and recruit as similar sized postlarvae (Table 1). Stomatopod larval development is highly regular such that larval size is good predictor of age (Hamano et al. 1995). Therefore, it can be assumed that the length of the larval period in these three species is roughly similar, previously estimated at four to six weeks (Barber et al. 2000, 2002a). Previous work has further confirmed that the larvae of both *H. pulchella* and *H. glyptocercus* are pelagic (Barber et al. 2002b). Ecologically, all three species inhabit the intertidal and subtidal zones of coral reef flats and are frequently found cohabitating in the same piece of coral rubble (P. H. Barber and M. V. Erdmann, pers. obs.). However, whereas *H. glyptocercus* and *G. viridis* occur in a wide range of habitats and are common inhabitants of both coastal and oceanic reefs, *H. pulchella* requires proximity to freshwater, excluding them from oceanic reefs systems, including those that span Flores Sea (Erdmann 1997; P. H. Barber and M. V. Erdmann, pers. obs.).

The comparison of these three taxa on fine spatial scales across the same geographic region permits an examination of how three ecologically similar species respond to the dynamic physical environment of the IAA, increasing our understanding of the processes promoting lineage diversification in this region. Because broadly acting physical processes produce similar patterns of regional genetic structure across taxa with similar ecologies and dispersal abilities (Avice et al. 1987; Avice 1992) this comparison will allow a robust test of previously hypothesized phylogeographic barriers. In particular, having similar larval durations, these three taxa should be similarly affected by previously hypothesized limited larval transport across the Maluku Sea between the islands of Papua and Indonesia (subsequently referred to as the Eastern Barrier), resulting in concordant genetic breaks across this region (Barber et al. 2002a). In contrast, we predict that the sharp genetic break seen in *H. pulchella* across the Flores and Java Seas (subsequently referred to as the Southern Barrier) should be absent or less pronounced in *H. glyptocercus* and *G. viridis*, given those two species' ability to inhabit oceanic reefs and hence use them as stepping-stones in dispersal.

MATERIALS AND METHODS

Sampling and Sequencing

From July 1994 until April 2005, over 800 juvenile and adult specimens of *H. pulchella*, *H. glyptocercus*, and *G. viridis* were collected from numerous localities throughout the IAA (Fig. 1, Table 2) and preserved in 70–95% ethanol. For both species of *Haptosquilla*, approximately 700 base pairs of the mitochondrial cytochrome *c* oxidase subunit-1 gene (CO1) was amplified from Chelex (Bio-rad, Hercules, CA) DNA extractions (Walsh et al. 1991) via polymerase chain reaction (PCR) using primers H1490 and L2198 (Folmer et al. 1994) following published protocols (Barber and Erdmann 2000). Primers internal to these were used for *G. viridis* (GV2LCO=GAGCATGATCAGGAATAGTAGG and GV5HCO=AAGTTGCGGTCTGTAAAGAAGTA) with thermocycling parameters of 38 cycles of 94°C/30 sec, 50°C/30 sec, 72°C/30 sec. Five microliters of double-stranded PCR fragments were cleaned by adding 0.5 units of shrimp alkaline phosphatase and 5 units of exonuclease, then incubating at 37°C for 30 min and 80°C for 15 min. Cleaned PCR fragments were sequenced on an ABI 377 automated sequencer using BigDye (Applied Biosystems, Foster City, CA) terminator chemistry. Forward and reverse sequences were proofread in Sequencher (Gene Codes Corp., Ann Arbor, MI) and subsequently aligned by eye. Protein translations were confirmed in MacClade version 4.05 (Maddison and Maddison 2002).

Phylogenetics, Population Structuring, and Gene Flow

Due to previously reported high levels of regional genetic variation (Barber et al. 2002a) and potentially polyphyletic lineages in stomatopods (Barber and Boyce 2006), sequence data from each individual species was initially screened to confirm that all sequences from a given species formed a monophyletic lineage and could therefore be reasonably assumed to be a single taxonomic unit. First, all sequences of *H. pulchella* and *H. glyptocercus* were compared to other Indo-Pacific Protosquillidae, and all *G. viridis* samples were compared to other Gonodactylidae from Barber and Erdmann (2000) using neighbor-joining as implemented in PAUP version 4.0b10 (Swofford 2002). Representative sequences from major clades identified through neighbor-joining analyses were then compared to all Protosquillidae and Gonodactylidae taxa from Barber and Erdmann (2000) via heuristic max-

TABLE 2. Collection localities and samples sizes of *Haptoquilla pulchella*, *H. glyptocercus*, and *Gonodactylellus viridis* examined in this study. Numbers correspond to Figure 1. South, North, and Papua refer to phylogeographic regions in Southern and Northern Indonesia and Papua New Guinea as discussed in the text. Flores and Banda populations could not be a priori assigned to a geographic region and were omitted from gene flow analyses.

Locality	Region	<i>H. pulchella</i>	<i>H. glyptocercus</i>	<i>G. viridis</i>
1 Komodo	South	6	9	12
2 Lombok	South	10	10	9
3 Lovina, Bali	South	48	5	
4 Sanur, Bali	South	16	11	
5 Jepara, Java	South			13
6 Pangandaran, Java	South		11	
7 Pulau Seribu	South	14	11	21
8 Karita, Java	South	16		
9 Krakatau	South	11	11	
10 Belitung, Batu Hitam	North	10		15
11 Belitung, Tanjung Pandan	North	10		
12 Riau	North	21		10
13 Postiljons	Flores		10	10
14 Tukang Besi/Wakatobi	Flores		10	
15 Spermonde, Sulawesi	North	76	11	10
16 Pare Pare, Sulawesi	North	10		5
17 Pantaloan, Sulawesi	North	10	11	10
18 Derawan, Kalimantan	North		10	12
19 Toli Toli, Sulawesi	North	13		
20 Kwandang, Sulawesi	North	10	11	
21 Pulau Tiga, Sulawesi	North	14		
22 Malalayang, Sulawesi	North	10	11	
23 Lembeh, Sulawesi	North	10		
24 Tanjung Dodepo, Sulawesi	North		10	
25 Gorontalo, Sulawesi	North		10	
26 Togians	North	21	11	12
27 Talaud	North	20	11	9
28 Davao, Philippines	North	10	10	
29 Hatta, Banda	Banda		11	11
30 Obi	Banda	9		
31 Halmahera	Papua		11	9
32 Sorong, W. Papua	Papua	9		
33 Biak, W. Papua	Papua		11	16
34 Yapen, W. Papua	Papua	10		10
35 Long Island, Papua New Guinea	Papua	10	11	10
36 Madang, Papua New Guinea	Papua	8		
Total		403	238	204

imum likelihood with model parameters determined by ModelTest 3.0 (Posada and Crandall 1998), and 100 maximum likelihood bootstrap replicates were performed to assess node support. Sequences that fell outside monophyletic species groupings were assumed to be heterospecific and were omitted from all subsequent analyses. Furthermore, because highly divergent clades could likely represent cryptic species, all separations greater than 4.8% corrected sequence divergence (based on the maximum likelihood parameters from ModelTest) were also assumed to be heterospecific and were omitted from analyses of gene flow. This value was chosen based on Knowlton and Weigt (1998) in which the lowest corrected pairwise sequence divergence for CO1 between transisthmian sister species of alpheid shrimp was 4.8%. These divergent clades are, however, included in the phylogeography figures to show the geographic positions of these potentially cryptic species. Although defining taxonomic boundaries based on sequence divergence is inherently arbitrary, in the absence of detailed morphological studies, such a strategy was required for consistency among the analyses.

For each taxon, a minimum-spanning tree based on pairwise sequence differences was created via the MINSNPNET

algorithm as employed in Arlequin 2.0 (Schneider et al. 2000). All alternative connections were examined to determine whether they resulted in fundamental changes in the tree topology. Clades separated by more than 10 mutational steps were color-coded, and the frequency of these divergent clades was plotted onto geography using pie diagrams.

Patterns of genetic structuring were estimated using the AMOVA algorithm as implemented in Arlequin 2.0 (Schneider et al. 2000). Analyses were run with regional subdivisions of Papua (populations east of the Maluku Sea and north of the Banda Sea), Northern Indonesia (populations west of the Maluku Sea and north of the Java and Flores Seas) and Southern Indonesia (populations south of the Java and Flores Sea) except for *H. pulchella* where additional groupings defined by phylogeographic structure were also included. Significance levels were determined through 10,000 randomization replicates. To examine levels of gene flow among defined regions, pairwise Φ_{ST} , an analog of F_{ST} (Michalakis and Excoffier 1996), among regions were calculated in AMOVA from both pairwise and Kimura two-parameter distance. For these analyses, populations from distinct phylogeographic regions were grouped into single regional populations (e.g.,

Papua, Northern Indonesia, Southern Indonesia) to be comparable with other migration estimates (see below). To standardize regional comparisons, samples from the Tukang Besi (Wakatobi), Banda, and Postiljons were omitted from gene flow analyses because their geographic positions in the Flores or Banda Seas precluded clear a priori assignment to Papua, Northern Indonesian, or Southern Indonesian populations. Gene flow was then estimated using the formula $Nm = 0.5 \times [(1/F_{ST}) - 1]$ (Hudson et al. 1992).

Gene flow and divergence time among geographic regions across the Southern and Eastern Barriers were simultaneously examined with Mdiv (Nielsen and Wakeley 2001), a program that can account for existing polymorphism prior to vicariance through simultaneously estimating θ , Nm , and divergence time, and can give effective estimates of these parameters with single nonrecombining loci such as mtDNA. Analyses were run using a finite-sites model, a Markov chain length of 2×10^6 , burn-in time of 50,000, maximum migration rate and divergence time of 10. Scaled divergence time was converted to years following Smith and Farrell (2005) where $T_{div} = T\theta/(2\mu)$, with T and θ being estimated from Mdiv, and μ being calculated by multiplying a 1.4%/million year rate of CO1 divergence (Knowlton and Weigt 1998) by the length of the CO1 sequence data for each of the three species. Because Mdiv reports results in terms of generations, the resulting value was converted into years by assuming a minimum generation time of two years in gonodactylid stomatopods (R. Caldwell, pers. comm.). As with AMOVA, analyses were run excluding clades in excess of 4.8% divergence. Although Nm values were estimated from mtDNA and hence only represent female dispersal, sex-biased dispersal of pelagic crustacean larvae is highly unlikely, suggesting that mtDNA derived dispersal estimates are likely representative of males as well.

RESULTS

A total of 625 bp of mitochondrial cytochrome oxidase I was sequenced from 403 individuals of *H. pulchella* and 227 individuals of *H. glyptocercus*, and 491 bp of CO1 was sequenced from 213 individuals of *G. viridis*. All sequences aligned easily and translated into protein. A total of 14 amino acid substitutions were recorded for *H. pulchella*, three for *H. glyptocercus*, and five for *G. viridis*; the remainder of nucleotide variation was silent.

Neighbor-joining analyses of all haplotypes of *H. pulchella* and *H. glyptocercus* with seven Protosquillid taxa from Barber and Erdmann (2000) showed these two taxa to be monophyletic. This result was confirmed with maximum likelihood analyses (Fig. 2) of representative sequences from the unique clades identified in neighbor-joining analyses. However, while the *H. glyptocercus* clades received 93% bootstrap support, there was no bootstrap support for the *H. pulchella* clades with the exception of the unity of Northern Indonesia and Yapen haplotypes receiving 74% bootstrap support. In contrast, phylogenetic analysis of *G. viridis* haplotypes with 20 other gonodactylids from Barber and Erdmann (2000) indicated two small clades that fell outside of the otherwise monophyletic *G. viridis* sequences. Maximum likelihood analysis confirmed (Fig. 2) that instead these sequences close-

ly allied (88% bootstrap support) with a recently described species *G. erdmanni* (Ahyong 2001) although their specific taxonomic status is uncertain (S. Ahyong, pers. comm.). As such, these samples were omitted from analyses of genetic structure and gene flow of *G. viridis*. However, the remainder of *G. viridis* *sensu stricto* received 99% bootstrap support.

Haplotypes of *H. pulchella* formed four distinct clades: a black clade separated by a genetic break of 18 mutational steps (2.9% corrected sequence divergence) from a red clade that itself was separated by 30 mutational steps from a white clade (5.5% corrected sequence divergence), which in turn was separated by 34 mutational steps (6.6% uncorrected sequence divergence) from a blue clade (Fig. 3A). As such, the blue and white clades were omitted from AMOVA and Mdiv analyses. Haplotypes from *H. glyptocercus* also fell into multiple groups: a black clade separated by 27 mutational steps (4.7% corrected sequence divergence) from a blue clade which was separated by 50 mutational steps (11.7% corrected sequence divergence) from a group of two rare haplotypes (white and red) that were 30 mutational steps (5.3% corrected sequence divergence) different from each other (Fig. 3B). The highly divergent red and white clades were omitted from AMOVA and Mdiv analyses, but the blue clade was included despite being close to the 4.8% cutoff. This decision was made because although the red and white clades could be morphologically distinguished from the blue and black, the latter could not be distinguished even with the aid of knowing their genetic identity (S. Ahyong, pers. comm.). Haplotypes of *G. viridis* had four divisions: a black clade separated by 17 mutational steps (3.53% corrected sequence divergence) from a white clade which was 15 mutational steps (3.0%) from a light blue clade which in turn was 38 mutational steps (10.6% corrected sequence divergence) from a dark blue clade (Fig. 3C). Because of the deep divergence, the dark blue clade was omitted from AMOVA and Mdiv analyses. However, the remaining clades were included because no clear morphological differences could be found among them, even when separated genetically (S. Ahyong, pers. comm.). Although all minimum-spanning trees had alternate connections, none affected the results reported herein.

Geographic distribution of the genetic variation was highly concordant with geography. Previous results from *H. pulchella* (Barber et al. 2002a) showed that the Papua New Guinea populations had only blue clade haplotypes, Northern Indonesia populations had only black clade haplotypes and the white clade was exclusive to Southern Indonesia (Fig. 4A). However, new results obtained by including additional populations in the analysis show black clade haplotypes spanning the Eastern Barrier, extending across the Maluku Sea to western Papua (population 32, Fig. 1), whereas a fourth less divergent clade was found only in northwestern Papua on the island of Yapen. Results from *H. glyptocercus* show that blue clade haplotypes dominate in Papua and decrease dramatically in frequency west of the Eastern Barrier. Black clade haplotypes span the Southern Barrier, dominating in both Northern and Southern Indonesia, although the two rare highly divergent white and red haplotypes are only encountered in Bali (Fig. 4B). Similarly, in *G. viridis*, light blue haplotypes are only seen in Papua, black clade haplotypes dominate Northern Indonesia, and white clade haplotypes are predom-

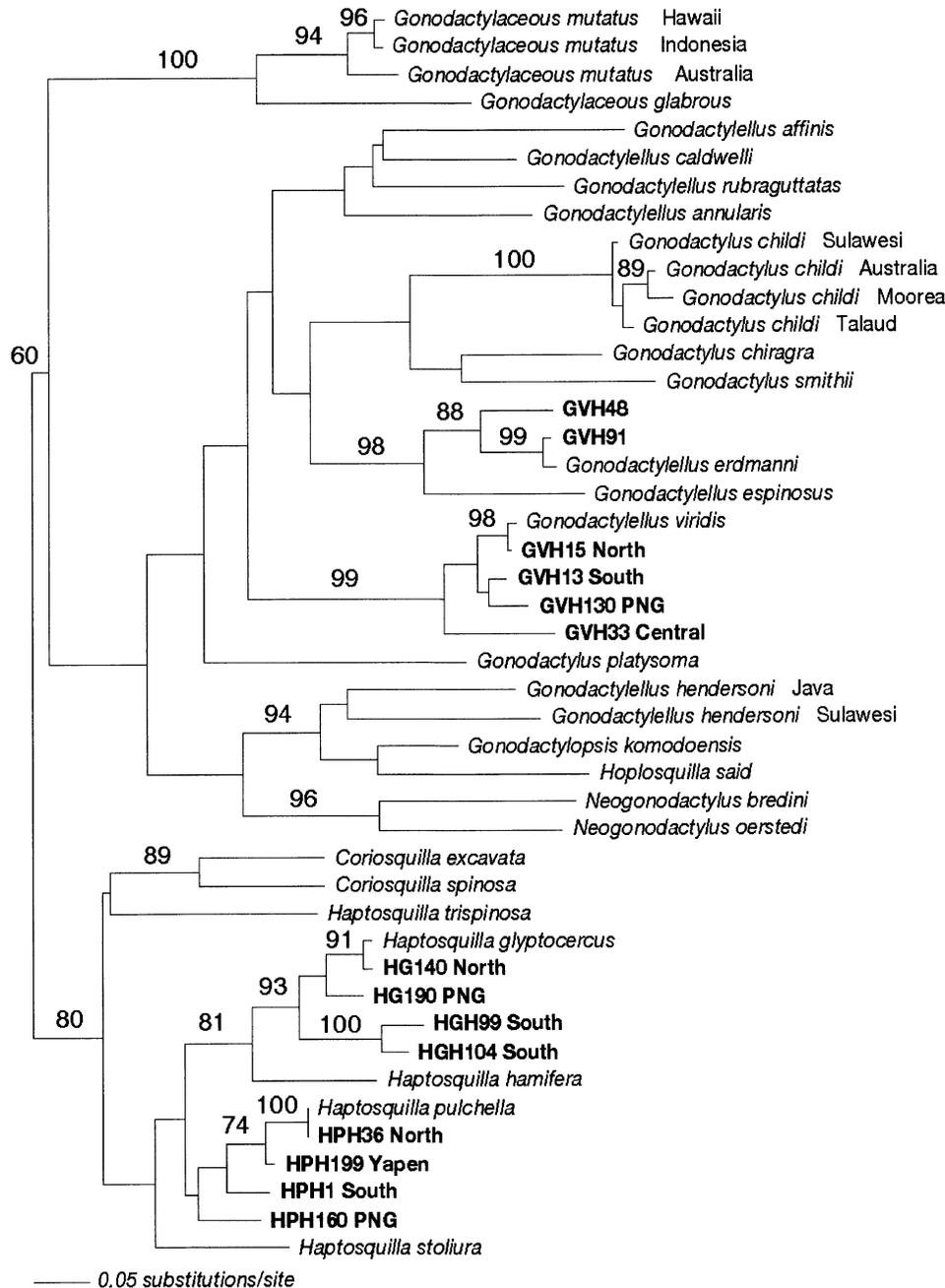


FIG. 2. Maximum likelihood tree of representative haplotypes from divergent clades of *Haptosquilla pulchella*, *H. glyptocercus*, and *Gonodactylellus viridis* (in bold) with gonodactylid and protosquillid stomatopods from Barber and Erdmann (2000). Numbers at nodes indicate support from 100 maximum likelihood bootstrap replicates.

inant in Southern Indonesia. Haplotypes from the dark blue clade were distributed sparsely throughout central Indonesia, and are subsequently referred to as the central clade (Fig. 4C).

AMOVA results calculated from pairwise distance indicated that populations of *H. pulchella* were highly structured. Excluding the highly divergent blue clade from Papua New Guinea and white clade from Southern Indonesia, $\Phi_{ST} = 0.800$ ($P < 0.00001$) and a total of 77.2% of the observed variance was the result of differences between Papua, South-

ern Indonesian, and Northern Indonesian populations. Variance among populations within groups and variance within populations accounted for 2.9% and 19.9%, respectively. Even excluding all but the black clade haplotypes, genetic structure was strong in *H. pulchella* ($\Phi_{ST} = 0.16$, $P < 0.00001$). Similarly, *H. glyptocercus* was highly structured, but less structured than *H. pulchella*, with $\Phi_{ST} = 0.527$ ($P < 0.00001$), excluding the highly divergent southern haplotypes. Again, even when only black clade haplotypes are considered, $\Phi_{ST} = 0.113$ ($P < 0.00001$) for *H. glyptocercus*.

TABLE 3. AMOVA results showing degrees of freedom (df), variance components (Var.), percent variation (% Var.) and ϕ -statistics for *Haptosquilla pulchella*, *H. glyptocercus*, and *Gonodactylellus viridis*. Runs included all populations as described in the text. Asterisk indicates significance at the $P < 0.00001$ level.

	Among regions				Among populations within regions				Within populations			
	df	Var.	% Var.	ϕ_{CT}	df	Var.	% Var.	ϕ_{SC}	df	Var.	% Var.	ϕ_{ST}
<i>H. pulchella</i>	2	4.8	77.2	0.772*	14	0.18	2.9	0.125*	259	1.24	19.9	0.800*
<i>H. glyptocercus</i>	2	4.3	47.9	0.479*	17	0.44	4.8	0.092*	187	4.33	47.3	0.527*
<i>G. viridis</i>	2	4.30	43.5	0.434*	13	2.36	23.9	0.423*	142	3.23	32.7	0.674*

and Northern Indonesian populations across the Eastern Barrier. Estimates of N_m from Φ_{ST} and MDiv were highly consistent with the exception of N_m values of *H. glyptocercus*. For *H. pulchella*, the presence of multiple phylogroups in Papua prevented a true Papua–Northern Indonesia comparison of gene flow. Instead, two separate comparisons were made. First, populations from Northern Indonesia and Sorong (western Papua) were compared to Yapen, the next sampled population of conspecifics to the East in Papua. The resulting Φ_{ST} between Northern Indonesia and Yapen was 0.89 ($P < 0.00001$), yielding an N_m estimate of 0.064 similar to the $N_m = 0.12$ obtained from Mdiv. To examine gene flow specifically across the Maluku Sea, a comparison between populations from Northern Indonesia and Sorong was done, yielding $\Phi_{ST} = 0.16$ ($P < 0.002$), even though this comparison only included populations containing black clade haplotypes. This comparison yielded an N_m estimate of 2.61, similar to $N_m = 3.8$ obtained from Mdiv. Populations of *H. glyptocercus* had a $\Phi_{ST} = 0.66$ ($P < 0.00001$) resulting in $N_m = 0.26$, but $N_m = 1.58$ was obtained from Mdiv. For *G. viridis*, $\Phi_{ST} = 0.73$ ($P < 0.00001$) resulting in $N_m = 0.20$, roughly similar to the $N_m = 0.68$ obtained from Mdiv. Although individual clades of haplotypes spanned the Maluku Sea in both *H. glyptocercus* and *G. viridis*, sample sizes were too limited to permit analysis of gene flow within individual clades as with *H. pulchella*.

Gene flow between Northern and Southern Indonesia

TABLE 4. Summary of gene flow (N_m) estimates across the Eastern and Southern Barriers in populations of *Haptosquilla pulchella*, *H. glyptocercus* and *Gonodactylellus viridis* as calculated from Φ_{ST} values from Arlequin (ARL) and directly from Mdiv. Because *H. pulchella* had multiple phylogeographic regions in Papua, comparisons were not made between Northern Indonesia and Papua populations. Instead, two separate comparisons were made: (1) Northern Indonesia populations (including Sorong, see text) were compared to Yapen, and (2) Northern Indonesia populations were compared to Sorong. Results include estimates that include divergent clades as well as estimates derived using only haplotypes from a single (black) clade, noted by double asterisks.

	<i>H. pulchella</i>	<i>H. glyptocercus</i>	<i>G. viridis</i>
Papua vs. North, ARL	n.a.	0.26	0.20
Papua vs. North, Mdiv	n.a.	1.58	.28
North vs. Yapen, ARL	.064	n.a.	n.a.
North vs. Yapen, Mdiv	0.12	n.a.	n.a.
North vs. Sorong, ARL**	2.61	n.a.	n.a.
North vs. Sorong, MDiv**	3.80	n.a.	n.a.
North vs. South, ARL	0.095	n.a.	0.94
North vs. South, MDiv	0.042	n.a.	2.06
North vs. South, ARL**	4.26	3.94	7.83
North vs. South, Mdiv**	4.02	4.02	5.14

across the Flores and Java Seas across the Southern Barrier was similar in all three species (Table 4). Including only black clade haplotypes, populations of *H. pulchella* had $\Phi_{ST} = 0.105$ ($P < 0.00001$) across this span, yielding $N_m = 4.26$, comparable to the $N_m = 4.02$ obtained from Mdiv. However, these values would drop to $N_m = 0.095$ and 0.042, respectively, were the black and white clades found to be conspecific. Nearly identical results were obtained from *H. glyptocercus* with $\Phi_{ST} = 0.113$ ($P < 0.00001$) across the Flores Sea, resulting in $N_m = 3.9$, similar to $N_m = 4.02$ obtained from Mdiv. Excluding only the dark blue clade, genetic structure in *G. viridis* was $\Phi_{ST} = 0.35$ ($P < 0.00001$), yielding $N_m = 0.94$, or $N_m = 2.06$ from Mdiv. Including only the black clade haplotypes, genetic structure was still significant, with $\Phi_{ST} = 0.06$ ($P < 0.0009$), yielding $N_m = 7.83$, or $N_m = 5.14$ as obtained from Mdiv.

Relative divergence times in Mdiv were scaled assuming a rate of CO1 sequence divergence of 1.4%/million years (MY; Knowlton and Weigt 1998) and a minimum generation time of two years. For *H. pulchella*, populations from Northern Indonesia and Yapen are estimated to have diverged 2.64 million years ago (MYA), whereas divergence within the black clade between Northern Indonesia and Sorong has been more recent at 0.47 MYA (Table 5). For *H. glyptocercus* and *G. viridis*, divergence between Northern Indonesia and Papua New Guinea populations were estimated at 3.34 MYA and 2.42 MYA, respectively. Divergence times between Northern and Southern Indonesia populations were 0.37 MYA for *H. pulchella*, 0.96 MYA for *H. glyptocercus*, and 0.81 MYA for

TABLE 5. Theta and divergence time estimates between populations of *Haptosquilla pulchella*, *H. glyptocercus*, and *Gonodactylellus viridis* from Northern Indonesia (North) and Papua and Northern and Southern Indonesia (South) as computed with Mdiv. Due to multiple clades of *H. pulchella* in Papua, multiple comparisons were required as described in the text. Results include estimates that include divergent clades as well as estimates derived using only haplotypes from a single (black) clade, noted by double asterisks.

	θ	Years
North-Papua		
<i>H. pulchella</i> (North-Yapen)	54.9	2.64
<i>H. pulchella</i> (North-Sorong)**	41.3	0.47
<i>H. glyptocercus</i>	54.1	3.34
<i>G. viridis</i>	24.4	2.41
North-South		
<i>H. pulchella</i>	48.0	4.8
<i>H. pulchella</i> **	39.9	0.37
<i>H. glyptocercus</i> **	59.7	0.96
<i>G. viridis</i>	18.7	0.81
<i>G. viridis</i> **	16.6	0.53

G. viridis. Even when only black clade haplotypes were examined, divergence of *G. viridis* populations across this span was similar at 0.53 MYA. θ values were similar for *H. pulchella* (39.9–54.9) and *H. glyptocercus* (54.1–59.7), but these were two to three times as large as θ values for *G. viridis* (18.6–24.4).

DISCUSSION

Regional Patterns of Isolation

Three ecologically similar mantis shrimp species show broad similarity of deep genetic structure throughout the Indonesian Archipelago. All three species include genetically distinct southern populations on the Indian Ocean side of the archipelago, distinct eastern populations along the coast of Papua and distinct central populations in between. Although Indonesia has long been described as an area of strong genetic breaks between the Pacific and Indian Oceans (Lacson and Clark 1995; Lavery et al. 1996; Duke et al. 1998; Williams and Benzie 1998; Benzie 1999; Duda and Palumbi 1999; Williams et al. 2002), our results show that these breaks are not the simple genetic clines seen across genetic breaks in other oceans (Avice 1996). Instead, the population patterns described here show a complex genetic mosaic within the complex oceanography and insular geography of the Indonesian Archipelago. Although the broad patterns among species are similar, the detailed edges of the internal genetic mosaics and the timing of population differentiation among mosaic patches vary slightly among species.

The commonalities of phylogeographic pattern and gene flow at the largest spatial scale observed in all three taxa suggests that the action of common regional scale physical process(es) have acted in their creation (Avice 2000). However, for patterns to be maximally concordant, they must also exhibit temporal concordance (Cunningham and Collins 1994). Yet each of the three species is unique in its response to the physical processes acting in this region with individualistic patterns, depths, and timing of regional genetic divergence. This result indicates that processes promoting regional genetic differentiation are strong, but that these processes may be occasionally swamped by gene flow during periods of altered dispersal. A nonexclusive alternative is that chance events or short-term environmental forcing affects some aspects of regional population-level differentiation within the Indonesian Archipelago. Such variable environmental impacts may be common in complex habitats such as Indonesia and during complex climate periods such as the Pleistocene.

Divergence across the Eastern Barrier

All three taxa show evidence of isolation in the region of the Eastern Barrier between populations on the island of Papua and the Indonesian islands west of the Maluku Sea (Fig. 3). In each case, major clades divergent by 15–38 nucleotide substitutions separate haplotypes on either side of this break, and for two of three species these divergent haplotypes form sister clades in the phylogenetic reconstruction (Fig. 2). The low estimates of gene flow and population divergence dating to the Pliocene suggests that this structure has evolved from

limited genetic exchange over an extended period of time. However, the recovery of pronounced genetic structure even within the black clade of *H. pulchella* indicates that this region continues to be a barrier to gene flow. Similar genetic discontinuities across this region are also seen in nautilus (Wray et al. 1995) as well as in human populations (Capelli et al. 2001), supporting the notion that this region is a biogeographic barrier.

The region is also host to two historically named biogeographic breaks, Weber's Line (through the Maluku Sea, see Van den Bergh et al. 2001) and Lydekker's Line (through the Halmahera Sea, see Van Oosterzee 1997). The former was defined as the middle of a zone of faunal mixing of Oriental and Australian fresh water fauna, whereas the latter marks the edge of the Sahul Continental shelf and the very far eastern reaches of Wallace's zoogeographic province. Both lines denote a strong west-east decline in terrestrial species of Oriental origin across this region, with different species showing slightly different biogeographic patterns. Likewise, although our three mantis shrimp species show genetic breaks from Papua New Guinea to the Indonesian Archipelago, the precise location of this break varies. For *H. glyptocercus*, the break occurs across the 400-km-wide Maluku Sea, whereas for *H. pulchella* and *G. viridis*, the break occurs farther to the east between Halmahera and Papua. Further evidence for isolation across the Maluku Sea comes from analysis of intraclade genetic structure in *H. pulchella*. In this species, black clade haplotypes extend across the Maluku Sea from Northern Indonesia to western Papua. However, genetic structure within the black clade between Northern Indonesia and western Papua was still pronounced ($\Phi_{ST} = 0.161$, $P < 0.005$) further indicating limited genetic exchange between these regions.

Based on Mdiv results, divergences between major clades in Northern Indonesia and Papua date to 3.34 MYA and 2.42 MYA, respectively, for *H. glyptocercus* and *G. viridis* and to 2.64 MYA for *H. pulchella*. These dates broadly mark the end of the Pliocene and a period of marked sea-level fluctuations that began 2.6 MYA (Pillans et al. 1998), a process that has long been hypothesized to have sparked marine speciation (Potts 1985). Increased upwelling during low sea-level stands (Fleminger 1986) may have acted to isolate both Indian and Pacific populations as well as populations in Indonesia and Papua that span the Eastern Barrier.

An alternative but not mutually exclusive explanation is that regional dispersal is limited by low water transport from Papua to Indonesia across the Maluku Sea even at high sea levels. The retroflexion of South Pacific waters by the Halmahera Eddy results in limited water exchange between Papua and the remainder of Indonesia. Although these are present-day surface currents, they result from processes that are highly stable. First, the proximity of Asia and Australia continents and their respective shallow continental shelves (see Fig. 1) constricts the waterways between the Pacific and Indian Oceans, a constriction that has been pronounced for at least the past 12 million years (Hall 1998). Because of this constriction, trade winds caused by convective heating in the tropics combined with the eastward rotation of the earth creates a sea surface height and pressure differential between the Pacific and Indian Ocean, driving the Indonesian

Throughflow (Wyrтки 1987) and shaping the associated Western Boundary Currents (e.g., Mindanao and Halmahera Eddies). Although there is variability in the intensity of the Indonesian Throughflow (Bray et al. 1996), particularly with respect to climate events such as the El Niño/Southern Oscillation (ENSO; Meyers 1996), many of the features that drive the physical oceanography of this region have been in place and operating for millions of years. The intensity of the Throughflow and associated currents may have diminished at times, but a wholesale reversal with Indian to Pacific transport is unlikely (R. Lukas, pers. comm.). Thus, while variation is expected, the currents associated with the Indonesian Throughflow have likely been a persistent feature of this oceanographic region since the middle Miocene. Limited water and larval transport between distinct geographic regions for large parts of the past several million years may promote lineage diversification *de novo*, and/or may act to reinforce Pleistocene vicariance, and may explain the genetic structure across this region even within the black clade of *H. pulchella*.

Divergence in Southern Indonesia

Regional genetic differentiation among populations spanning the Flores Sea was also seen in all three taxa. The strongest phylogeographic pattern was observed in *H. pulchella* (see also Barber et al. 2000, 2002a), in which very limited mixing was seen between populations dominated by divergent haplotypes on the southern versus northern shores of the Flores and Java Seas. Across the same area, *G. viridis* showed a less pronounced genetic break across the Flores Sea: instead, differentiation occurred along the southern shores of the Java Sea, particularly along the islands of Java and Belitung off the coast of Sumatra (Fig. 4). Just as in *H. pulchella*, a dramatic genetic shift occurs along a 200-km stretch of coast between these two islands. These deep phylogeographic breaks vary greatly in age, with divergence across the Flores Sea dating to 4.8 MYA in *H. pulchella* and only 0.81 MYA in *G. viridis*. Thus, they either originated from different processes separated by millions of years, or, perhaps more likely, by a physical process that has been present for an extended period of time.

The least genetic structure was observed in *H. glyptocercus*, in which a shift between major haplotypes did not occur. However, genetic structuring among closely related haplotypes between Northern and Southern Indonesian populations was observed ($\Phi_{ST} = 0.112$, $P < 0.00001$) with levels of gene flow much less than would be expected of a taxon with a four- to six-week pelagic larval dispersal phase in an environment where strong currents should facilitate dispersal. Similar conclusions are drawn from *H. pulchella* and *G. viridis*, in which comparable levels of genetic structure were seen across the Flores Sea, even when clades of deeply divergent haplotypes were excluded from the analysis. In all three species, divergence within individual clades spanning the Flores Sea dates to the mid- to late Pleistocene (0.37–0.96 MYA, Table 5). Given at least 10,000 years of modern oceanographic conditions since the end of the last glacial maxima, strong gene flow across the Flores Sea would be expected to erase any incomplete genetic boundaries sepa-

rating populations north and south of the Flores Sea, yet significant levels of genetic structure remain. The combination of deeply divergent phylogeographic lineages of highly varied temporal origins with continued evidence of limited genetic exchange across the Flores Sea suggests that isolation most likely results from a process with a prolonged effect. However, unlike divergence across the Maluku Sea, there is no obvious physical process to explain divergence across this span.

Strong genetic differentiation across the Flores Sea shows that individual ecological requirements of these three species do not strictly control dispersal. The tendency of *H. pulchella* to settle only on reefs in proximity to freshwater influx could prevent the use of small oceanic islands in the Flores Sea as stepping-stones, whereas the other two species are free of this constraint. However, the small degree of leakage of *H. pulchella* black clade haplotypes across the Maluku Sea combined with gene flow estimates that are similar in all three species across this span indicates that dispersal can occur in the absence of stepping-stones. This result suggests that stepping-stone islands do not strongly influence migration across the Flores Sea in these stomatopods.

Genetic Endemism and the Metapopulations of Indonesia

The islands of the IAA have complex geologic origins with a mixture of continental shelf and deepwater ocean origins. Habitat differences among islands could conceivably account for some of the phylogenetic breaks observed here. For example, Springer (1982) noted a distinct shift in shallow-water fish fauna on islands of adjacent tectonic plates. For cowries, an association between larval dispersal duration and ocean habitat type experienced by the larvae contributes to patterns of diversification (Meyer 2003). Genetic differences across the Halmahera Sea correspond geographically to the edge of the Sahul Shelf. However, habitat-based explanations are unlikely because we see similar genetic breaks in species with specialist and generalist habitat requirements. *Haptosquilla pulchella* is stenotopic in its habitat requirements, preferring only shallow reef flats in close proximity to freshwater sources. In over 13 years of collecting these animals throughout the Indo-Pacific, we have found *H. pulchella* only in these areas. In contrast, *G. viridis* is extremely tolerant of habitat quality. Even within a single observed phylogeographic region (e.g., black clade), this taxon can be found in oceanic reefs and coastal/continental reefs, as well as in pristine reefs and highly eutrophied reefs that are nearly dead (Erdmann 1997; P. H. Barber and M. V. Erdmann, pers. obs.). *Haptosquilla glyptocercus* is found across similarly divergent habitat types, although it is less tolerant of eutrophication than *G. viridis*. Thus, because habitat variation within phylogeographic regions exceeds that among regions, habitat variation is an unlikely explanation for the generation of phylogeographic structure.

In addition, the phylogeographic structure does not correspond in a simple way to the geologic origin of the islands they inhabit. For example, genetic homogeneity is observed along the shores of Sulawesi (Barber et al. 2002a), an island with multiple geologic origins (Hall 1998). In contrast, the islands of Sumatra and Java have similar geologic origins,

yet there are major genetic breaks at their juncture in two of three species. Explanations for sharp genetic breaks in taxa with high potential dispersal remain obscure (Barber et al. 2000; Taylor and Hellberg 2003). However, the interaction of extrinsic habitat variation and intrinsic forces such as swimming (Katz et al. 1994; Yoshimura et al. 1999), vertical migrations (Cronin and Forward 1986; Tankersley and Forward 1994; Forward et al. 1997), or resident adults interfering with larval recruitment (Woodin 1976; Morgan 1995) cannot be completely discounted.

The mosaic nature of genetic structure in Indonesian stomatopods is also seen in the geographic distribution of genetic endemics. Distinct mtDNA haplotypes or clades often occur in just one or two collection localities. Sometimes these haplotypes are just a few base pairs different from haplotypes elsewhere, for example in the Togian Islands (location 26, Fig. 1) in Sulawesi (see Barber et al. 2002b). Similarly, a distinct group of haplotypes in *H. pulchella* was found only at Yapen in West Papua (location 34, Fig. 1; red haplotypes in Figs. 3 and 4). In contrast, a highly divergent clade of *G. viridis* haplotypes was found in only a few disjunct locations but over a wide geographic area. Individuals with blue clade haplotypes were found near Bali and over 1000 km away on Banda and Kalimantan (locations 3, 13, 18, and 30, Fig. 1; dark blue haplotype in Figs. 3 and 4), yet no representatives of this clade were found in the numerous intervening populations.

Private haplotypes are common in populations that disperse poorly. In marine species, however, such haplotypes tend to be young ones because, in general, older haplotypes have had the opportunity for long-distance dispersal and tend to be widespread (Palumbi 1997). The occurrence of old private haplotypes in Indonesian stomatopods suggests a long history of restricted gene flow, or a metapopulation structure in which occasional long-distance gene flow and habitat colonization is followed by local population retention and periodic extinction.

In some cases, metapopulation models can also help explain the high F_{ST} values seen here and in other high dispersal marine species (for review see Hellberg et al. 2002). When colonization of a new habitat occurs from a small subsample of adjacent habitats, then extinction and recolonization events increase F_{ST} dramatically (the propagule-pool model of Slatkin 1977). In contrast, if recolonization derives from immigrants from the whole metapopulation (the migrant-pool model), extinction and recolonization decreases F_{ST} (Slatkin 1977; Wakeley 2000). For stomatopods, genetic data from recolonization of the reefs of Krakatau show that adjacent populations may have contributed more heavily to the colonists (Barber et al. 2002a), supporting the propagule-pool model for these species.

The impact of metapopulation dynamics on marine genetic diversity is only likely to be seen in cases where extinction rate per population per generation is greater than the migration rate per population per generation (for discussion see Pannell 2003). Shallow-water reef populations, especially near rivers, may have a high rate of local extinction from typhoons and extended periods of either low salinity due to torrential rains or extreme temperatures from ENSO events. If so, we might predict lower F_{ST} values among ecologically

similar deepwater stomatopods with lower exposure to local extinction.

Ideally, we would be able to estimate extinction and migration parameters from the kinds of population genetic data presented here. However, simulations show that single-locus genealogical data have low power to detect the dynamics of metapopulation processes (Pannell 2003), and such calculations require more extensive multilocus genealogical data.

Taxonomy and Speciation

Levels of genetic diversity were extremely high in all three taxa, with genetic divergences among clades similar to those seen in species of alpheid shrimp (Knowlton and Weigt 1998). Although phylogenetic analysis indicated that some of these lineages were clearly polyphyletic (e.g., *G. viridis* clades allied with *G. erdmanni*), other divergent clades within monophyletic taxa might be cryptic species based on depth of genetic divergence. The taxonomic status of these groups, however, can only be substantiated with thorough, detailed morphological study and/or genetic analysis of nuclear genes. Whether these clades are eventually determined to be morphologically distinct species or simply highly divergent conspecific lineages, the results of this study indicate that present levels of biodiversity in the IAA—the center of all marine biodiversity (Myers et al. 2000; Hughes et al. 2002; Roberts et al. 2002)—may be significantly underestimated (Knowlton 1993). Similar conclusions are drawn from Paulay and Meyer (2002), Meyer (2003), Williams and Reid (2004), Kirkendale and Meyer (2004), and Meyer et al. (2005), each of which show deep regional divergences within individual taxa of a variety of marine invertebrates throughout the Indo-Pacific, and from Barber and Boyce (2006) who discovered multiple new stomatopod taxa through DNA bar coding of stomatopod larvae.

More importantly, rather than merely indicating multiple cryptic lineages within these three taxa, the results suggest a mechanism for the creation of biodiversity in the IAA. All three taxa demonstrate varied degrees of genetic isolation across the Eastern and Southern Barriers. In both regions this isolation ranges from statistically detectable limits to gene flow within individual clades to nearly reciprocal monophyletic phylogeographic regions to putative cryptic species. Thus, these patterns suggest a progression of regional genetic diversification, indicating that these barriers to gene flow may both contribute to the creation of distinct phylogeographic lineages as well as ultimately drive cladogenesis and the creation of species diversity, providing a mechanism to support the center of origin hypothesis.

Particularly instructive is the case of *H. pulchella*, in which there is evidence of limited gene flow on multiple levels. Within populations with black clade haplotypes, there is moderately strong isolation ($\Phi_{ST} = 0.161$, $P < 0.005$) between Northern Indonesia and Sorong across the Eastern Barrier. Then there is increased isolation between Northern Indonesia/Sorong (black clade haplotypes) and Yapen (red clade haplotypes), the next sampled population about 400 km to the east. Although there may be an as yet unidentified process creating this phylogeographic break along the shores of Western Papua, a likely explanation is that this break arose because

of isolation across the Eastern Barrier and that there has been a subsequent breach of this barrier by populations containing black clade haplotypes leading to the present unity of Western Papua and Northern Indonesia. This may have occurred from rare dispersal across the Eastern Barrier or via stepping-stones across the islands south of the Maluku Sea, as suggested by black clade haplotypes on the island of Obi (Fig. 4A; population 30 in Table 2). On a deeper time scale, such a process may have also given rise to the blue clade of *H. pulchella* in eastern Papua, resulting in a type of founder speciation seen in peripheral habitats of the Indo-West Pacific (Meyer et al. 2005).

Similar patterns are seen in populations of all three taxa that span the Flores Sea. In each species, limited gene flow within individual clades parallels deeper phylogeographic breaks that may represent cryptic lineages. Particularly interesting is the contrast between *H. pulchella* and *G. viridis*. Including the deeply divergent black and white clades, results for *H. pulchella* indicate that these lineages have been diverging for nearly five million years. When the white clade is excluded, the history of limited genetic exchange is much more recent (0.37 MYA), a result that closely mirrors divergence of *H. pulchella* populations that span the Maluku Sea. In contrast, populations of *G. viridis* have a history of divergence across the Flores Sea that dates to 0.81 MYA when both black and white clades are included and 0.53 MYA when the divergent white clade is removed. These results highlight the temporal variability in lineage divergence as well as show that lineage divergence across the same geographic area can occur in relatively rapid succession. Similarly, they argue against discrete individual events as their cause.

Combined, these results suggest a gradual diversification due to limited gene flow in two regions of Indonesia, and support the notion of lineage diversification within Indonesia as predicted by the center of origin hypothesis (Briggs 2000). Although a center of origin is likely not the only explanation for high marine biodiversity in the IAA (Barber and Bellwood 2005; Meyer et al. 2005), the patterns seen in stomatopods appear to support lineage diversification within the Coral Triangle. By understanding the processes that create biodiversity, we can then improve our ability to conserve this biodiversity through appropriately directed conservation efforts (Moritz 1999; Moritz 2002).

Consequences for Conservation

The connectivity of marine populations is a critical gap in present scientific knowledge needed for large-scale marine conservation planning (Palumbi 2004; Sale et al. 2005). Many tropical marine conservation efforts are presently focused on the development of networks of marine protected areas (MPAs) that are connected through demographic exchange of adults or larvae (e.g., Roberts et al. 2003), implying that a thorough understanding of patterns of genetic and demographic connectivity among distant populations is a priority.

However, equally important to understanding connectivity is identifying regions where connectivity is lacking, as well as understanding the processes that limit that connectivity. The data in this study suggest that limited gene flow across

the Maluku and Flores/Java Seas leads to the creation of pronounced phylogeographic structure, and possibly cryptic speciation. This limited gene flow likely results from limited dispersal and connectivity among these geographic regions through interactions of individual larvae with their physical environment. Thus, our results suggest the importance of regional filters to dispersal and gene flow within the IAA, data that has been used to support current seascape-level initiatives by defining boundaries of regionally distinct biogeographic regions in which to create regional networks of genetically and ecologically connected MPAs that will both conserve the rich marine biodiversity of the Indo-West Pacific (e.g., Green and Mous 2004; MPA News 2004). Furthermore, the long-term maintenance of biodiversity requires the conservation of processes that create this biodiversity (Briggs 2005). Allopatric divergence across the Maluku and Flores Seas may be one important process generating biodiversity in the Coral Triangle.

It is not yet clear whether the Eastern and Southern Barriers in the regions of the Maluku and Flores Seas represent dispersal barriers only to stomatopods, or whether these are broadly acting barriers that may be shaping patterns of intraspecific genetic diversity across a wide range of taxa. Determination of the overall importance and general applicability of these barriers will require comparisons across a wide range of codistributed invertebrate and vertebrate taxa, and should be considered a top priority in the developing marine conservation strategies for the region.

ACKNOWLEDGMENTS

The authors are grateful to the Indonesian government, Indonesian Institute of Sciences (LIPI), and M. K. Moosa for permitting and sponsoring our work in Indonesia, and to the government of Papua New Guinea for permitting our collections there. We also thank the numerous villagers and fisherman that assisted us during field collections. J. K. Choi, M. Said, J. Kulongoski, L. Zeidberg, and A. Mehta assisted with field collections and logistics. Field collections were supported by a Putnam Grant to PHB and a National Science Foundation (NSF) Doctoral Dissertation Improvement Grant (9503060) and NSF International Postdoctoral Fellowship (INT-970416) to MVE. Additionally, PHB thanks C. Wallace and N. Lindquist for supporting collections in the Togian Islands and Papua New Guinea, respectively, and P. Crews for organizing permits in Papua New Guinea. A NSF Minority Postdoctoral Fellowship, Boston University, and NSF grant OCE-0349177 supported field and lab work by PHB during this study. Laboratory work was further supported by grants to SRP from NSF and the Andrew Mellon Foundation. This research was also supported in part by a Pew Marine Conservation Fellowship to MVE.

LITERATURE CITED

- Ahyong, S. T. 2001. Revision of the Australian Stomatopod Crustacea. *Rec. Aust. Mus. Suppl.* 26:1–326.
- Armstrong, P. R. 2001. Directed motion in the sea: efficient swimming by reef fish larvae. *J. Theor. Biol.* 210:81–91.
- Avise, J. C. 1992. Molecular population structure and the biogeography of history of a regional fauna: a case history with lessons for conservation biology. *Oikos* 63:62–67.

- . 2000. *Phylogeography: the history and formation of species*. Harvard Univ. Press, Cambridge, MA.
- Avise, J. C., J. Arnold, R. M. Ball, E. Bermingham, T. Lamb, J. E. Neigel, C. A. Reeb, and N. C. Saunders. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annu. Rev. Ecol. Syst.* 18: 489–522.
- Barber, P. H., and D. R. Bellwood. 2005. Biodiversity hotspots: evolutionary origins of biodiversity in wrasses (Halichoeres: Labridae) in the Indo-Pacific and new world tropics. *Mol. Phylogenet. Evol.* 35:235–253.
- Barber, P. H., and S. L. Boyce. 2006. Estimating diversity of Indo-Pacific coral reef stomatopods through DNA barcoding of stomatopod larvae. *Proc. R. Soc. Lond. B* 273:2053–2061.
- Barber, P. H., and M. V. Erdmann. 2000. Molecular systematics of the Gonodactylidae (Stomatopoda) using mitochondrial cytochrome oxidase C (subunit I) DNA sequence data. *J. Crustacean Biol.* 20(Spec. No. 2):20–36.
- Barber, P. H., S. R. Palumbi, M. V. Erdmann, and M. K. Moosa. 2000. A marine Wallace's line? *Nature* 406:692–693.
- Barber, P. H., S. R. Palumbi, M. V. Erdmann, and M. K. Moosa. 2002a. Sharp genetic breaks among populations of a benthic marine crustacean indicate limited oceanic larval transport: patterns, causes, and consequences. *Mol. Ecol.* 11:659–674.
- Barber, P. H., M. K. Moosa, and S. R. Palumbi. 2002b. Rapid recovery of genetic diversity of stomatopod populations on Krakatau: temporal and spatial scales of marine larval dispersal. *Proc. R. Soc. Lond. B* 269:1591–1597.
- Benzie, J. A. H. 1999. Major genetic differences between crown-of-thorns starfish (*Acanthaster planci*) populations in the Indian and Pacific Oceans. *Evolution* 53:1782–1795.
- Bray, N. A., S. Hautala, J. Chong, and J. Pariwono. 1996. Large-scale sea level, thermocline, and wind variations in the Indonesian Throughflow region. *J. Geophys. Res.* 101:12239–12254.
- Briggs, J. C. 2000. Centrifugal speciation and centres of origin. *J. Biogeogr.* 27:1183–1188.
- . 2005. Coral reefs: conserving the evolutionary sources. *Biol. Conserv.* 126:297–305.
- Bush, G. L. 1975. Modes of animal speciation. *Annu. Rev. Ecol. Syst.* 6:339–365.
- Capelli, C., J. F. Wilson, M. Richards, M. P. Stumpf, F. Gratrix, S. Oppenheimer, P. Underhill, V. L. Pascali, T. M. Ko, and D. B. Goldstein. 2001. A predominantly indigenous paternal heritage for the Austronesian-speaking peoples of insular Southeast Asia and Oceania. *Am. J. Hum. Genet.* 68:432–443.
- Cronin, T. W., and R. B. Forward. 1986. Vertical migration cycles of crab larvae and their role in larval dispersal. *Bull. Mar. Sci.* 39:192–201.
- Cunningham, C. W., and T. M. Collins. 1994. Developing model systems for molecular biogeography: vicariance and interchange in marine invertebrates. Pp. 405–434 in B. Schierwater, B. Streit, G. P. Wagner, and R. DeSalle, eds. *Molecular ecology and evolution: approaches and applications*. Birkhauser Verlag, Basel.
- Doherty, P. J., S. Planes, and P. Mather. 1995. Gene flow and larval duration in seven species of fish from the Great Barrier Reef. *Ecology* 76:2373–2391.
- Duda, T. F., and S. R. Palumbi. 1999. Population structure of the black tiger prawn, *Penaeus monodon*, among western Indian Ocean and western Pacific populations. *Mar. Biol.* 134:705–710.
- Duke, N. C., J. A. H. Benzie, J. A. Goodall, and E. R. Ballment. 1998. Genetic structure and evolution of species in the mangrove genus *Avicennia* (Avicenniaceae) in the Indo-West Pacific. *Evolution* 52:1612–1626.
- Ekman, S. 1953. *Zoogeography of the sea*. Sidgwick and Jackson, Ltd., London.
- Erdmann, M. V. N. 1997. The ecology, distribution, and bioindicator potential of Indonesian coral reef stomatopod communities. Pp. 1–290. *Integrative biology*. Ph.D. diss. University of California, Berkeley.
- Fauvelot, C., and S. Planes. 2002. Understanding origins of present-day genetic structure in marine fish: biologically or historically driven patterns? *Mar. Biol.* 141:773–788.
- Fleminger, A. 1986. The Pleistocene equatorial barrier between the Indian and Pacific Oceans and a likely cause for Wallace's Line. Pp. 84–97 in A. C. Pierrot-Bults, ed. *Pelagic biogeography: proceedings of an international conference*. May 29–June 5 1985, The Netherlands. UNESCO, Paris.
- Folmer, O., M. Black, W. Hoeh, R. Lutz, and R. Vrijenhoek. 1994. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit 1 from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotechnol.* 3:294–299.
- Forward, R. B., J. Swanson, R. A. Tankersely, and J. M. Welch. 1997. Endogenous swimming rhythms of blue crab, *Callinectes sapidus*, megalopae: effects of offshore and estuarine cues. *Mar. Biol.* 127:621–628.
- Godfrey, J. S. 1996. The effect of the Indonesian Throughflow on ocean circulation and heat exchange with the atmosphere: a review. *J. Geophys. Res.* 101:12217–12237.
- Gordon, A. L., and R. A. Fine. 1996. Pathways of water between the Pacific and Indian oceans in the Indonesian seas. *Nature* 379: 146–149.
- Green, A., and P. J. Mous. 2004. Delineating the Coral Triangle, its ecoregions and functional seascapes. The Nature Conservancy Expert Workshop, South East Asia Center for Marine Protected Areas, Bali, Indonesia. 24 pp.
- Hall, R. 1998. The plate tectonics of Cenozoic SE Asia and the distribution of land and sea. Pp. 99–132 in R. Hall and J. D. Holloway, eds. *Biogeography and geological evolution of SE Asia*. Backhuys Publishers, Leiden.
- Hamano, T., T. Kikkawa, S. Ueno, and K. I. Hayashi. 1995. Use of larval size, instead of larval stage, to study the ecology of a stomatopod crustacean *Oratosquilla oratoria*. *Fish. Sci.* 61: 165–166.
- Hellberg, M. E., R. S. Burton, J. E. Neigel, and S. R. Palumbi. 2002. Genetic assessment of connectivity among marine populations. *Bull. Mar. Sci.* 70:273–290.
- Hudson, R. R., M. Slatkin, and W. P. Maddison. 1992. Estimation of gene flow from DNA sequence data. *Genetics* 132:583–589.
- Hughes, T. P., D. R. Bellwood, and S. R. Connolly. 2002. Biodiversity hotspots, centres of endemicity, and the conservation of coral reefs. *Ecol. Lett.* 5:775–784.
- Jokiel, P., and F. J. Martinelli. 1992. The vortex model of coral reef biogeography. *J. Biogeogr.* 19:449–458.
- Jones, G. P., M. J. Milicich, M. J. Emslie, and C. Lunow. 1999. Self-recruitment in a coral reef fish population. *Nature* 402: 802–804.
- Katz, C. H., J. S. Cobb, and M. Spaulding. 1994. Larval behavior, hydrodynamic transport, and potential offshore-to-inshore recruitment in the American lobster *Homarus americanus*. *Mar. Ecol. Prog. Ser.* 103:265–273.
- Kirkendale, L. A., and C. P. Meyer. 2004. Phylogeography of the *Patelloida profunda* group (Gastropoda: Lottidae): diversification in a dispersal driven marine system. *Mol. Ecol.* 13: 2749–2762.
- Knowlton, N. 1993. Sibling species in the sea. *Annu. Rev. Ecol. Syst.* 24:189–216.
- Knowlton, N., and L. A. Weigt. 1998. New dates and new rates for divergence across the Isthmus of Panama. *Proc. R. Soc. Lond. B* 265:2257–2263.
- Lacson, J. M., and S. Clark. 1995. Genetic divergence of Maldivian and Micronesian demes of the damselfishes *Stegastes nigricans*, *Chrysiptera biocellata*, *C. glauca* and *C. leucopoma* (Pomacentridae). *Mar. Biol.* 121:585–590.
- Lavery, S., C. Moritz, and D. R. Fielder. 1996. Indo-Pacific population structure and evolutionary history of the coconut crab *Birgus latro*. *Mol. Ecol.* 5:557–570.
- Leis, J. M., and B. M. Carson-Ewart. 1997. In situ swimming speeds of the late pelagic larvae of some Indo-Pacific coral-reef fishes. *Mar. Ecol. Prog. Ser.* 159:165–174.
- Lourie, S. A., D. M. Green, and A. C. J. Vincent. 2005. Dispersal, habitat differences, and comparative phylogeography of Southeast Asian seahorses (Syngnathidae: Hippocampus). *Mol. Ecol.* 14:1073–1094.
- Lukas, R., E. Firing, P. Hacker, P. L. Richardson, C. A. Collins, R. Fine, and R. Gammon. 1991. Observations of the Mindanao

- Current during the Western Equatorial Pacific Ocean circulation study. *J. Geophys. Res.* 96:7089–7104.
- Lynch, J. D. 1989. The gauge of speciation: on the frequencies of modes of speciation. Pp. 527–556 in D. Otte and J. A. Endler, eds. *Speciation and its consequences*. Sinauer, Sunderland, MA.
- Maddison, W. P., and D. R. Maddison. 2002. *MacClade*. Sinauer, Sunderland, MA.
- Manning, R. B. 1963. Notes on the embryology of the stomatopod crustacean *Gonodactylus oerstedii* Hansen. *Bull. Mar. Sci. Gulf Caribb.* 13:422–432.
- Mayr, E. 1942. *Systematics and the origin of species*. Columbia Univ. Press, New York.
- McManus, J. W. 1985. Marine speciation, tectonics and sea-level changes in southeast Asia. *Proc. Fifth Int. Coral Reef Congr. Tahiti* 4:133–138.
- McMillan, W. O., and S. R. Palumbi. 1995. Concordant evolutionary patterns among Indo-West Pacific butterflyfishes. *Proc. R. Soc. Lond. B* 260:229–236.
- Meyer, C. P. 2003. Molecular systematics of cowries (Gastropoda: Cypraeidae) and diversification patterns in the tropics. *Biol. J. Linn. Soc.* 79:401–459.
- Meyer, C. P., J. B. Geller, and G. Paulay. 2005. Fine scale endemism on coral reefs: archipelagic differentiation in turbinid gastropods. *Evolution* 59:113–125.
- Meyers, G. 1996. Variation of Indonesian Throughflow and the El Niño Southern Oscillation. *J. Geophys. Res.* 101:12255–12263.
- Michalakis, Y., and L. Excoffier. 1996. A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics* 142:1061–1064.
- Mora, C., P. M. Chittaro, P. F. Sale, J. P. Kritzer, and S. A. Ludsin. 2003. Patterns and processes in reef fish diversity. *Nature* 421:933–936.
- Morey, S. L., J. F. Shriver, and J. J. O'Brien. 1999. The effects of Halmahera on the Indonesian Throughflow. *J. Geophys. Res.* 104:23281–23296.
- Morgan, S. A. 1995. Life and death in the plankton: larval mortality and adaptation. Pp. 279–322 in L. McEdwards, ed. *Ecology of marine invertebrate larvae*. CRC Press, New York.
- Moritz, C. 1999. Conservation units and translocations: strategies for conserving evolutionary processes. *Hereditas* 130:217–228.
- . 2002. Strategies to protect biological diversity and the evolutionary processes that sustain it. *Syst. Biol.* 51:238–254.
- MPA News. 2004. Australian parliament passes re-zoning bill for Great Barrier Reef, creating world's largest reserve system. Marine Affairs Research and Education, Seattle, WA. Available via <http://depts.washington.edu/mpanews>.
- Myers, N., R. A. Mittermeier, C. G. Mittermeier, G. A. B. da Fonseca, and J. Kent. 2000. Biodiversity hotspots for conservation priorities. *Nature* 403:853–858.
- Nielsen, R., and J. Wakeley. 2001. Distinguishing migration from isolation: a Markov chain Monte Carlo approach. *Genetics* 158:885–896.
- Nof, D. 1995. Choked flows from the Pacific to the Indian-Ocean. *J. Phys. Oceanogr.* 25:1369–1383.
- Palumbi, S. R. 1992. Marine speciation on a small planet. *Trends Ecol. Evol.* 7:114–118.
- . 1994. Genetic divergence, reproductive isolation, and marine speciation. *Annu. Rev. Ecol. Syst.* 25:547–572.
- . 1997. Molecular biogeography of the Pacific. *Coral Reefs* 16:S47–S52.
- . 2004. Marine reserves and ocean neighborhoods: the spatial scale of marine populations and their management. *Annu. Rev. Environ. Res.* 29:31–68.
- Pannell, J. R. 2003. Coalescence in a metapopulation with recurrent local extinction and recolonization. *Evolution* 57:949–961.
- Paulay, G., and C. Meyer. 2002. Diversification in the tropical Pacific: comparisons between marine and terrestrial systems and the importance of founder speciation. *Int. Comp. Biol.* 42:922–934.
- Pillans, B., J. Chappell, and T. R. Naish. 1998. A review of the Milankovitch climatic beat: template for Plio-Pleistocene sea-level changes and sequence stratigraphy. *Sediment. Geol.* 122:5–21.
- Porter, S. C. 1989. Some geological implications of average Quaternary glacial conditions. *Quat. Res.* 32:245–261.
- Posada, D., and K. A. Crandall. 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics* 14:817–818.
- Potts, D. C. 1985. Sea-level fluctuations and speciation in Scleractinia. *Proc. Fifth Int. Coral Reef Congr. Tahiti* 4:127–132.
- Provenzano, A. J., and R. B. Manning. 1978. Studies on development of stomatopod crustacea II. The later larval stages of *Gonodactylus oerstedii* Hansen reared in the laboratory. *Bull. Mar. Sci.* 28:297–315.
- Roberts, C. M., C. J. McClean, J. E. N. Veron, J. P. Hawkins, G. R. Allen, D. E. McAllister, C. G. Mittermeier, F. W. Schueler, M. Spalding, F. Wells, C. Vynne, and T. B. Werner. 2002. Marine biodiversity hotspots and conservation priorities for tropical reefs. *Science* 295:1280–1284.
- Roberts, C. M., G. Branch, R. H. Bustamante, J. C. Castilla, J. Dugan, B. S. Halpern, K. D. Lafferty, H. Leslie, J. Lubchenco, D. McArdle, M. Ruckelshaus, and R. R. Warner. 2003. Application of ecological criteria in selecting marine reserves and developing reserve networks. *Ecol. Appl.* 13:S215–S228.
- Sale, P. F., R. K. Cowen, B. S. Danilowicz, G. P. Jones, J. P. Kritzer, K. C. Lindeman, S. Planes, N. V. C. Polunin, G. R. Russ, Y. J. Sadovy, and R. S. Steneck. 2005. Critical science gaps impede use of no-take fishery reserves *Trends Ecol. Evol.* 20:74–80.
- Schneider, S., D. Roessli, and L. Excoffier. 2000. Arlequin. Ver. 2.000. A software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland.
- Shulman, M. J., and E. Bermingham. 1995. Early life histories, ocean currents, and the population genetics of Caribbean reef fishes. *Evolution* 49:897–910.
- Slatkin, M. 1977. Gene flow and genetic drift in a species subject to frequent local extinction. *Theor. Popul. Biol.* 12:253–262.
- Smith, C. I., and B. D. Farrell. 2005. Range expansions in the flightless longhorn cactus beetles, *Moneilema gigas* and *Moneilema armatum*, in response to Pleistocene climate changes. *Mol. Ecol.* 14:1025–1044.
- Springer, V. G. 1982. Pacific Plate biogeography, with special reference to shorefishes. *Smithson. Contrib. Zool.* 367:1–182.
- Springer, V. G., and J. T. Williams. 1990. Widely distributed Pacific plate endemics and lowered sea-level. *Bull. Mar. Sci.* 47:631–640.
- Stobutzki, I. C., and D. R. Bellwood. 1997. Sustained swimming abilities of the late pelagic stages of coral reef fishes. *Mar. Ecol. Prog. Ser.* 149:35–41.
- Swearer, S. E., J. E. Caselle, D. W. Lea, and R. R. Warner. 1999. Larval retention and recruitment in an island population of a coral-reef fish. *Nature* 402:799–802.
- Swofford, D. L. 2002. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Sinauer, Sunderland, MA.
- Tankersley, R. A., and R. B. Forward. 1994. Endogenous swimming rhythms in estuarine crab megalopae: implications for flood-tide transport. *Mar. Biol.* 118:415–423.
- Taylor, M. S., and M. E. Hellberg. 2003. Genetic evidence for local retention of pelagic larvae in a Caribbean reef fish. *Science* 299:107–109.
- Thorrold, S. R., C. Latkoczy, P. K. Swart, and C. M. Jones. 2001. Natal homing in a marine fish metapopulation. *Science* 291:297–299.
- Turelli, M., N. H. Barton, and J. A. Coyne. 2001. Theory and speciation. *Trends Ecol. Evol.* 16:330–343.
- Van den Bergh, G., J. de Vos, and P. Y. Sondaar. 2001. The Late Quaternary palaeogeography of mammal evolution in the Indonesian Archipelago. *Palaeogeog. Palaeoclimatol. Palaeoecol.* 171:385–408.
- Van Oosterzee, P. 1997. *Where worlds collide: the Wallace Line*. Cornell Univ. Press, Ithaca, NY.
- Veron, J. E. N. 2000. *Corals of the World*. AIMS, Townsville, Australia.
- Voris, H. K. 2000. Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *J. Biogeogr.* 27:1153–1167.
- Wakeley, J. 2000. The effects of subdivision on the genetic divergence of populations and species. *Evolution* 54:1092–1101.

- Walsh, P. S., D. A. Metzger, and R. Higuchi. 1991. Chelex-100 as a medium for simple extraction of DNA for PCR based typing from forensic material. *BioTechniques* 10:506–513.
- Williams, S. T., and J. A. H. Benzie. 1998. Evidence of a phylogeographic break between populations of a high-dispersal starfish: congruent regions within the Indo-West Pacific defined by color morphs, mtDNA and allozyme data. *Evolution* 52:87–99.
- Williams, S. T., J. Jara, E. Gomez, and N. Knowlton. 2002. The marine Indo-West Pacific break: Contrasting the resolving power of mitochondrial and nuclear genes. *Int. Comp. Biol.* 42: 941–952.
- Williams, S. T., and D. G. Reid. 2004. Speciation and diversity on tropical rocky shores: a global phylogeny of snails of the genus *Echinolittorina*. *Evolution* 58:2227–2251.
- Wolanski, E., P. Doherty, and J. Carleton. 1997. Directional swimming of fish larvae determines connectivity of fish populations on the Great Barrier Reef. *Naturwissenschaften* 84:262–268.
- Woodin, S. A. 1976. Adult-larval interactions in dense infaunal assemblages: patterns of abundance. *J. Mar. Res.* 34:25–41.
- Woodland, D. J. 1986. Wallace's Line and the distribution of marine inshore fishes. Pp. 453–460 in T. Uyeno, R. Arai, T. Taniuchi and K. Matsuura, eds. *Indo-Pacific fish biology: proceedings of the Second International Conference on Indo-Pacific Fishes*. Ichthyological Society of Japan, Tokyo.
- Wray, C. G., N. H. Landman, and J. Bonacum. 1995. Genetic divergence and geographic diversification in *Nautilus*. *Paleobiology* 21:220–228.
- Wyrтки, K. 1961. *Physical oceanography of the southeast Asian waters*. University of California, San Diego.
- . 1987. Indonesian Throughflow and the associated pressure gradient. *J. Geophys. Res.* 92:12941–12946.
- Yoshimura, T., H. Yamakawa, and E. Kozasa. 1999. Distribution of final stage phyllosoma larvae and free-swimming pueruli of *Panulirus japonicus* around the Kuroshio Current off southern Kyusyu, Japan. *Mar. Biol.* 133:293–306.

Corresponding Editor: D. Ayre

Recent speciation in the Indo–West Pacific: rapid evolution of gamete recognition and sperm morphology in cryptic species of sea urchin

C. Landry, L. B. Geyer, Y. Arakaki, T. Uehara and Stephen R. Palumbi

Proc. R. Soc. Lond. B 2003 **270**, doi: 10.1098/rspb.2003.2395, published 7 September 2003

References

Article cited in:

<http://rsps.royalsocietypublishing.org/content/270/1526/1839#related-urls>

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click [here](#)

Recent speciation in the Indo-West Pacific: rapid evolution of gamete recognition and sperm morphology in cryptic species of sea urchin

C. Landry¹, L. B. Geyer^{1†}, Y. Arakaki², T. Uehara³ and Stephen R. Palumbi^{1*†}

¹Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA

²Department of Tourism, Faculty of International Studies, Meio University, Nago, Okinawa, Japan

³Department of Biology, University of the Ryukyus, Nishihara, Okinawa, Japan

The rich species diversity of the marine Indo-West Pacific (IWP) has been explained largely on the basis of historical observation of large-scale diversity gradients. Careful study of divergence among closely related species can reveal important new information about the pace and mechanisms of their formation, and can illuminate the genesis of biogeographic patterns. Young species inhabiting the IWP include urchins of the genus *Echinometra*, which diverged over the past 1–5 Myr. Here, we report the most recent divergence of two cryptic species of *Echinometra* inhabiting this region. Mitochondrial cytochrome oxidase 1 (CO1) sequence data show that in *Echinometra oblonga*, species-level divergence in sperm morphology, gamete recognition proteins and gamete compatibility arose between central and western Pacific populations in the past 250 000 years. Divergence in sperm attachment proteins suggests rapid evolution of the fertilization system. Divergence of sperm morphology may be a common feature of free-spawning animals, and offers opportunities to simultaneously understand genetic divergence, changes in protein expression patterns and morphological evolution in traits directly related to reproductive isolation.

Keywords: speciation; binding; Okinawa

1. INTRODUCTION

The tropical shelf waters of the Indo-West Pacific (IWP) house a rich species diversity that has its zenith around the East Indies. Species richness of many taxa increases from the eastern Pacific and the Indian Ocean towards the Indonesian Archipelago (reviewed in Briggs 1999; Bellwood & Hughes 2001). Although previous investigations of the genesis of this diversity emphasized the slow accumulation of species over long periods of time (Springer & Williams 1990), recent genetic evidence suggests that many of the reef species in this diverse area are relatively new, having originated since the beginning of the Pleistocene (Palumbi 1996; McMillan *et al.* 1999; Barber *et al.* 2002). These recent speciation events provide a laboratory in which to understand the speciation process of marine taxa, and allow a combination of genetic, morphological, behavioural and biogeographic data to be used to answer questions about speciation mechanisms. Among others, allopatry, peripheral speciation and divergence after secondary range contact have been proposed as drivers of species diversity in the IWP (Palumbi 1997; Bellwood & Hughes 2001), but as yet there have been very few clear reconstructions of speciation history of marine taxa in this area.

Footprints of species formation are most likely to be identified when comparing recently diverged species, initial differentiation of which can be correlated with the different proposed speciation processes. Molecular tools have helped to uncover closely related and still diverging

species from a wide variety of taxa and have provided insight into the mechanisms, pace and geography of marine speciation (Knowlton 1993; Palumbi 1997; Lessios *et al.* 2001). Coupled with information about geological history of regions, faunistic affiliations of species found in particular regions, and the oceanic currents that may provide dispersal corridors, molecular data can greatly help document the geographic pattern and relative timing of species formation.

Among the closely related species that help contribute to the richness of the IWP are the sea urchins of the genus *Echinometra*. Pan-tropical sea urchins of the genus *Echinometra* diverged into two different evolutionary lineages 3.3–4.5 Myr ago when the Indo-Pacific (IP) lineage separated from the one leading to the neotropical species (McCartney *et al.* 2000). Since then, the IP species lineage has undergone several speciation events. Molecular phylogenetic analyses, morphological studies and fertilization experiments have been combined to map the boundaries of four closely related *Echinometra* species common throughout the tropical Pacific (see Mortensen 1943; Uehara *et al.* 1986; Matsuoka & Hatanaka 1991; Palumbi & Metz 1991; Metz *et al.* 1994; Palumbi 1996; Arakaki *et al.* 1998a,b).

Among these is *Echinometra oblonga*, a black species that occupies reefs from Mauritius in the Indian Ocean to Isla del Coco in the eastern Pacific (figure 1). Typically, this species inhabits burrows above the mean low water level, in microhabitats that are highly exposed to wave action (Keslo 1970; Nishihira *et al.* 1991). This habitat preference correlates with species-specific physiological adaptations providing resistance to extreme conditions such as high and low temperatures, salinity changes (Arakaki & Uehara 1991) and strong shearing action of waves on

* Author for correspondence (spalumbi@stanford.edu).

† Present address: Department of Biological Sciences, Stanford University, Hopkins Marine Station, Pacific Grove, CA 93950, USA.

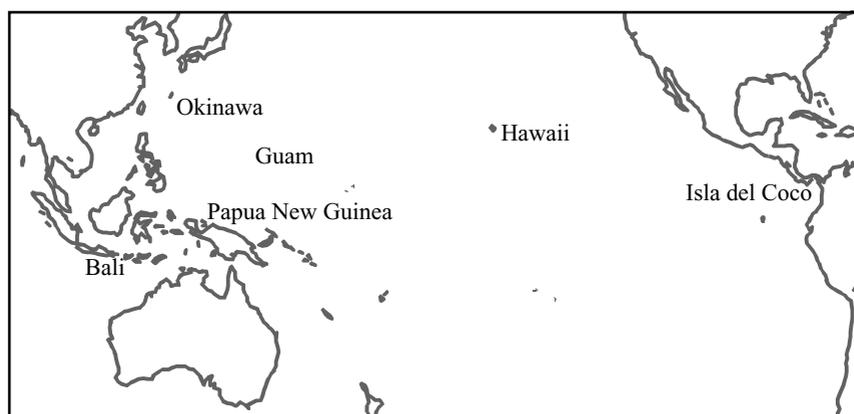


Figure 1. Schematic representation of collecting localities in the Pacific Ocean for *Echinometra oblonga* populations. Western Pacific: Bali (Indonesia), Okinawa (Japan) and Papua New Guinea; central Pacific: Guam (US), Hawaii (US) and Isla del Coco (Costa Rica).

gametes (Mead 1996). This species has been described from Japan and Hawaii on the basis of monophyletic mitochondrial DNA (mtDNA) lineages (Palumbi & Metz 1991), allozyme differences (Matsuoka & Hatanaka 1991), reproductive incompatibility with sister species (Uehara *et al.* 1990; Palumbi & Metz 1991; Metz *et al.* 1994) and large differences in the sperm protein bindin that attaches sperm to eggs before fertilization (Metz & Palumbi 1996). This species is among a complex of at least three other species that occur commonly throughout the IWP, and have been examined closely for population structure (Palumbi *et al.* 1997) and mechanisms of speciation (Palumbi 1999).

Careful analysis of data on *E. oblonga* from numerous localities throughout its range has shown marked polymorphism. For instance, individuals from Mauritius, Okinawa, Bonin, Guam and Hawaii differ slightly, although not diagnostically, in the number of pore pairs, as well as the shape of the spicules of the gonads and tube feet (Arakaki & Uehara 1999). Furthermore, these characteristics are not consistent with each other within *E. oblonga*: some characteristics cluster populations of Okinawa and Hawaii together and Indonesia and Guam together, whereas others differentiate the Hawaiian population from the rest of the Pacific (Arakaki & Uehara 1999). Another difference across the Pacific is sperm morphology: western Pacific populations have a sperm head that is twice as long as those from the central Pacific (figure 4; Arakaki & Uehara 1999). Recently, some central and western Pacific populations have been shown to be differentiated at bindin, a sperm–egg recognition protein (Geyer & Palumbi 2003). These results suggest that populations of *E. oblonga* may actually represent diverging species that differ in reproductive characteristics such as bindin sequence and sperm morphology. However, genetics, sperm morphology and gamete attachment features have never been examined in the same suite of populations, making delineation of biogeographic boundaries for these potential species difficult, and obscuring potential insights into the process of speciation.

In this study, our aims are to characterize populations of *E. oblonga* for two genetic loci and for sperm morphology to test if the differences observed reflect the reproductive isolation of these populations and to discuss

the history and the geography of this diversification. We investigate the phylogenetic relationship among *E. oblonga* populations from the western and central Pacific and compare genetic data with morphological differences. Because of their large population size and their recent divergence, *Echinometra* species are expected to be monophyletic at very few of their loci. However, mitochondrial genomes are expected to drift faster than nuclear genes and are then expected to show a signal of differentiation earlier in the speciation process (Palumbi *et al.* 2001). Previous work on mitochondrial sequences provided only slight evidence of genetic separation between *E. oblonga* populations based on the segregation of a single polymorphic site (Palumbi *et al.* 1997). Here, we increase the sequence length to provide more resolving power, and include additional populations that have been investigated for both sperm and bindin features. To put mtDNA diversity data within species into a larger perspective, we also include in our analysis other closely related species found in the IP (*E. mathaei*, *E. sp. C* and *E. sp. A*) and a sister species endemic to Easter Island (*E. insularis*). *Echinometra vanbrunti*, a congeneric species found in the eastern Pacific is used as an outgroup.

2. METHODS

Populations of *E. oblonga* were sampled from Isla del Coco ($n = 2$), Hawaii ($n = 5$), Guam ($n = 5$), Okinawa ($n = 11$), Bali ($n = 11$) and Papua New Guinea ($n = 3$). We included 15 samples of *E. mathaei*, three of *E. sp. A*, 14 of *E. sp. C* and 10 of *E. insularis*. DNA was extracted from tube feet, gonad or spine muscle tissue by proteinase K digestion and phenol/chloroform extraction, or by using NucleoSpin columns (BD, Biosciences, Clontech).

Two overlapping fragments of the cytochrome oxidase 1 (CO1) gene corresponding to the interval between sites 5851 to 7013 of *Strongylocentrotus purpuratus* mitochondrial genome were sequenced for each individual. Samples corresponded in part to the samples analysed for bindin by Geyer & Palumbi (2003). The 3' fragment was amplified and sequenced by using primers CO1f and CO1d according to Palumbi *et al.* (1997). The 5' fragment was obtained by using primers LCO1490 and HCO2198 (Folmer *et al.* 1994) by using the same chemistry and the following cycles: 94 °C 30 s; 50 °C 45 s; and 72 °C 1 min,

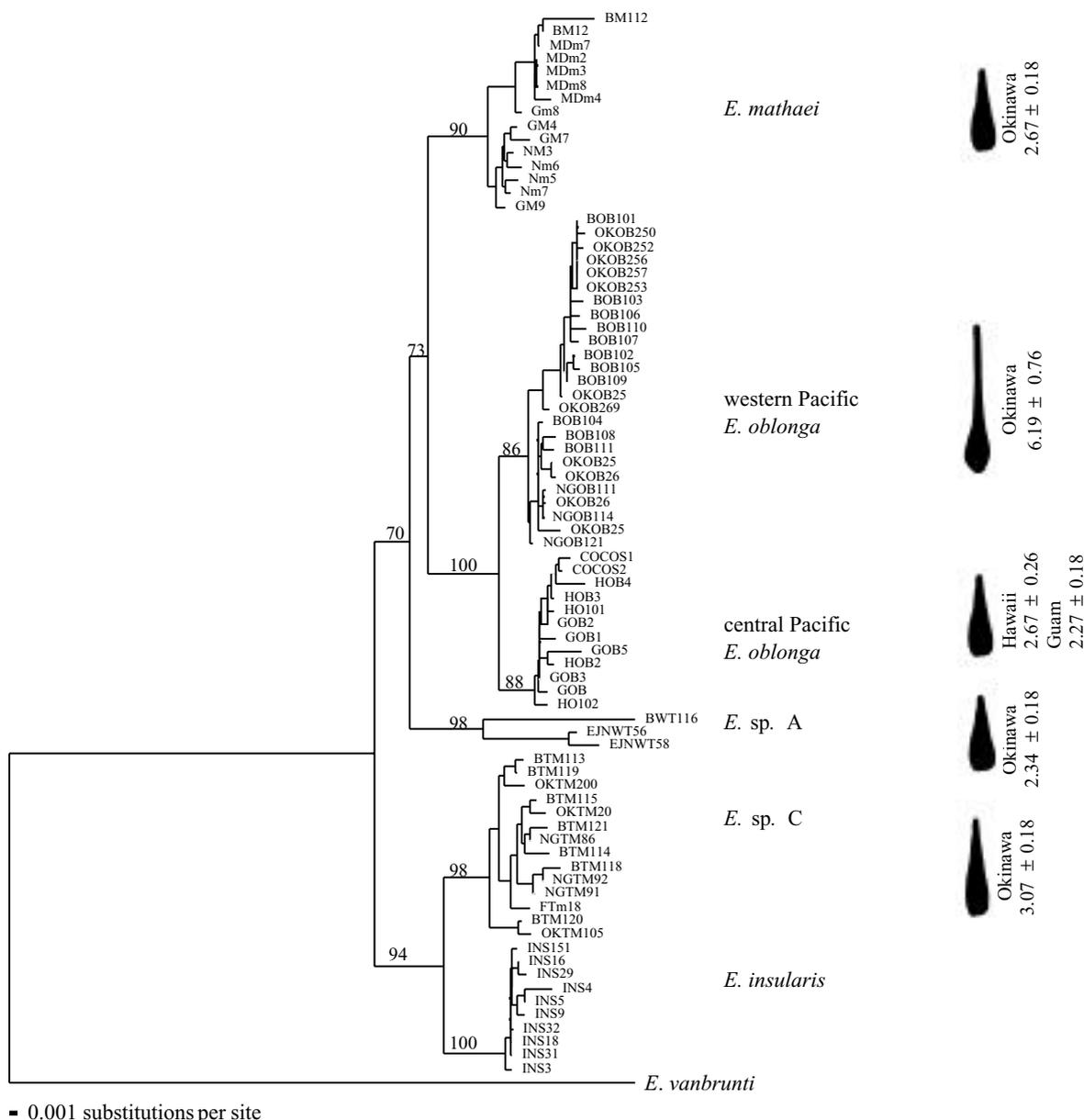


Figure 2. Rooted NJ tree of CO1 sequences (K_2 distance). Bootstraps were carried out by resampling all nucleotide positions 1000 times. The first one or two first letters of the sample names represent the sampling locality. (F, Fiji; OK, Okinawa; NG, Papua New Guinea; B, Bali; H, Hawaii; Co, Isla del Coco; M, Midway, Hawaii; N, Niue; EJJN, Japan; INS, Easter Island). Sperm morphologies are schematized and lengths (micrometres) are from Arakaki *et al.* (1998b) and Arakaki & Uehara (1999).

and a final elongation of 5 min at 72 °C. Including the primers, the two fragments overlap by 66 bp. Polymerase chain reaction (PCR) products were prepared for sequencing by using Exonuclease I and shrimp alkaline phosphatase digestion and were then sequenced on a 3100 ABI automated sequencer (Applied Biosystems, CA, USA). Sequences were aligned by hand. Within-species and among-species polymorphisms of CO1 sequences were estimated by using MEGA (Kumar *et al.* 2001). The net average distance corresponds to the average distance minus the average within-species polymorphism. The phylogenetic relationship of the samples was drawn by using the neighbour-joining (NJ) algorithm from PAUP v. 4.0 (Swofford 2001).

Bindin alleles were amplified, cloned and sequenced as previously described (Palumbi 1999; Geyer & Palumbi 2003) from individuals collected in Okinawa ($n = 4$), Bali ($n = 14$), Guam

($n = 6$), Hawaii ($n = 11$) and Niue ($n = 8$). A total of three to five clones from each individual were sequenced. When two alleles were encountered, both were included in the analysis. Otherwise, a single allele was analysed from each individual.

3. RESULTS

A portion of the mitochondrial CO1 gene corresponding to bases 5851 to *ca.* 6550 of the *S. purpuratus* complete mitochondrial genome was obtained for 79 individuals of five putative IWP species of *Echinometra* (GenBank accession numbers AY262861–AY262940). When added to sequences already available for these individuals, our dataset comprises 1164 bp of the CO1 gene.

Most of the substitutions within *Echinometra* are at silent sites in this coding sequence, and the

Table 1. Between-species average K_2 distance (below the diagonal) at all sites (pairwise deletion) and net average distance (above the diagonal). Average intraspecific nucleotide diversity (π) is shown on the diagonal (in italics).

	1	2	3	4	5	6	7
<i>Echinometra oblonga</i> western Pacific	<i>0.004</i>	0.009	0.033	0.027	0.022	0.034	0.132
<i>E. oblonga</i> central Pacific	0.013	<i>0.003</i>	0.034	0.027	0.019	0.036	0.133
<i>E. sp. C</i>	0.039	0.039	<i>0.007</i>	0.030	0.030	0.015	0.127
<i>E. sp. A</i>	0.040	0.040	0.046	<i>0.021</i>	0.021	0.029	0.127
<i>E. mathaei</i>	0.027	0.024	0.037	0.034	<i>0.006</i>	0.030	0.126
<i>E. insularis</i>	0.037	0.038	0.019	0.040	0.034	<i>0.002</i>	0.129
<i>E. vanbrunti</i>	0.135	0.134	0.131	0.138	0.129	0.130	0

transition/transversion ratio is about 11 : 1. We observed two amino acid changes in two sequences. Phylogenetic analysis of mtDNA sequences from 80 individuals revealed five major clades (figure 2) defined by high bootstrap values (90–100%). Four of these clades correspond to the four currently recognized IP species of *Echinometra* (Palumbi 1996). The clade including *E. oblonga* shows two additional, well-defined sub-clades, corresponding to *E. oblonga* from the central Pacific (Hawaii, Isla del Coco and Guam) and those from western Pacific populations (Indonesia, Okinawa and Papua New Guinea). These two clades are differentiated by seven fixed transition substitutions. The same organization among clades was obtained by using a parsimony analysis (C. Landry, unpublished data).

Using *E. vanbrunti* as an outgroup (McCartney *et al.* 2000), the IP *Echinometra* fall into two multi-species clades. One contains *E. sp. C*, a species found among few Pacific island archipelagoes (Palumbi 1996), and *E. insularis*, which is endemic to Easter Island. The other clades contain the widely distributed *E. mathaei*, *E. oblonga* and *E. sp. A*. Bootstrap support for the branching order of these three species is poor, but data from both morphology and bindin sequences suggest a sister-species relationship between *E. mathaei* and *E. sp. A*. Within this cluster, the two clades of *E. oblonga* are derived sister taxa (figure 2).

Average K_2 parameter distance (where K_2 is the silent rate at twofold degenerate nucleotide positions; Kumar *et al.* 2001) among IWP *Echinometra* species considering all the sites (pairwise deletion) ranges from 1.3% to 4.6%, which is very similar to the distances measured using only the CO1f-a gene fragment (Palumbi 1996; McCartney *et al.* 2000). Mean K_2 parameter distances to the eastern Pacific *E. vanbrunti* vary between 12.9% and 13.8%, which is slightly higher than was calculated by McCartney *et al.* (2000) for the smaller gene segment. Variation within species ranges from 0.2% to 2.1% for these samples (table 1).

The average genetic distance between the two groups of *E. oblonga* (1.3%; net distance: 0.9%) is low compared with the 4.0–4.8% difference between *E. sp. A*, *E. sp. C* and *E. insularis*, but is comparable to the distance between *E. sp. C* and the endemic *E. insularis* (1.9%). Under the assumption of constant rates of CO1 evolution, it is possible to use divergence data between mitochondrial sequences to estimate the time of the split between *E. oblonga* clades. Using the average sequence divergence at CO1 between species of *Echinometra* found on both sides

of the Isthmus of Panama, McCartney *et al.* (2000) derived an estimate of 3.49% Kimura-corrected sequence divergence per million years, with a range that spans 24% of this mean. Using this rate, we estimated a divergence time between the two types of *E. oblonga* (net $K_2 = 0.9%$, s.e.m. 0.5%) of ca. 250 000 years.

(a) *Bindin* sequences

Central and western Pacific populations of *E. oblonga* are characterized by very divergent bindin alleles. Bindin sequence diversity among all *E. oblonga* populations combined is much higher (2.5%, average corrected p -distance) than in any of the other IP *Echinometra* species (1.6%, *E. mathaei*; 0.9% *E. sp. A*, 0.8% *E. sp. C*). Division of the data into central and western Pacific populations explains 55% of the variation observed. Central Pacific *E. oblonga* (Nieu, Guam, Hawaii) have bindin sequences that fall into three major clades (clades one, two and four; figure 3). By contrast, bindin from western Pacific *E. oblonga* (Indonesia, Okinawa and Papua New Guinea) fall in a distinct clade separated from the rest by unique insertion and deletion events, plus amino acid substitutions (clades three a and three b; figure 3). Clade three alleles are found almost exclusively in the west Pacific (Bali, Okinawa, New Guinea). The only exception was a single individual from Niue which was homozygous for this allele, but carried a central Pacific CO1 haplotype.

(b) *Sperm morphology*

Two different morphologies typify sperm from western and central Pacific populations. The western Pacific type has a slender sperm head, with a ratio of length over width of about six, whereas the central Pacific type has a shape more compact, with a ratio of length over width of about three (figure 4). The primary difference lies in the tip of the sperm head and the shape of the nucleus (Arakaki *et al.* 1998a,b). Populations of *E. oblonga* from Okinawa show the sperm morphology of western Pacific populations. Populations in Guam and Hawaii both have the shorter, stouter sperm morphology.

4. DISCUSSION

The species *E. oblonga* appears to have recently diverged into two species distinguishable by sequences of mitochondrial genes, genes for gamete recognition, and by sperm morphology. Data from CO1 show that the species *E. oblonga* contains two major monophyletic clades that differ from one another by 0.9% and that divide the

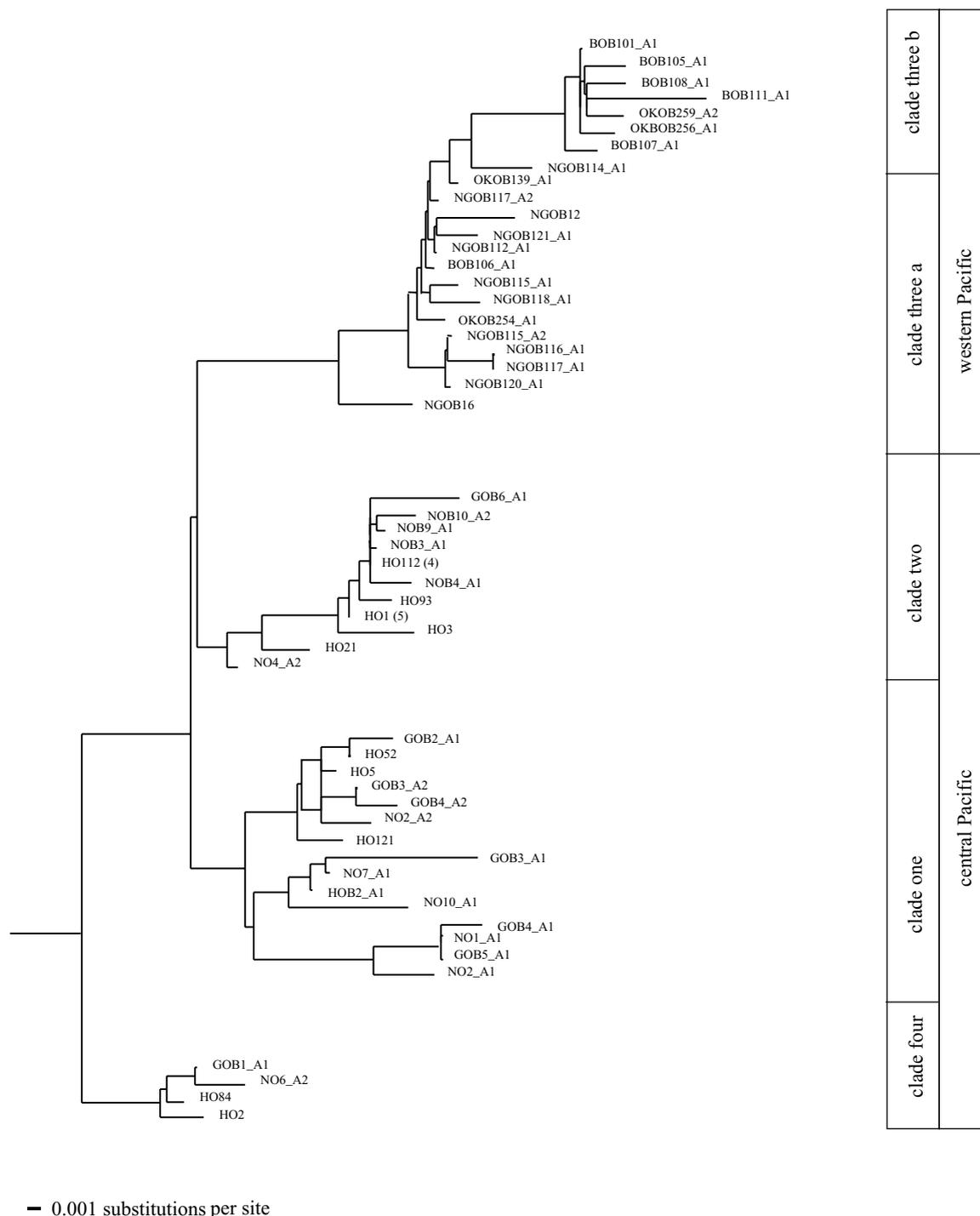


Figure 3. Divergence of bindin sequences in western and central Pacific populations of *Echinometra oblonga*. Sequences from the central Pacific fall into three major clades, all of which differ from the bindins seen in western Pacific populations. The branch leading to the western Pacific clades shows amino acid substitutions and insertions and deletions in the region of the protein implicated in functional sperm–egg attachment differences.

tropical Pacific into central and western areas. However, other sea urchin species in this genus harbour old mitochondrial lineages (Palumbi 1996), and so the mere existence of a significant split between central and western populations in CO1 does not guarantee species distinction. Instead, corroborative data are needed from other loci or organismal features before these data can be used to confirm species status of central versus western Pacific populations. Data from three other sources show agreement with this geographic pattern.

(a) *Bindin gene sequences and fertilization barriers*

The genetic differentiation seen in mtDNA is paralleled by differences at the gamete recognition locus bindin. Central Pacific populations differ strongly in bindin sequence from western Pacific populations, including amino acid and insertion/deletion changes in the protein coding region. The region of the bindin gene where these changes occur is the previously identified evolutionary hotspot where the action of positive diversifying selection

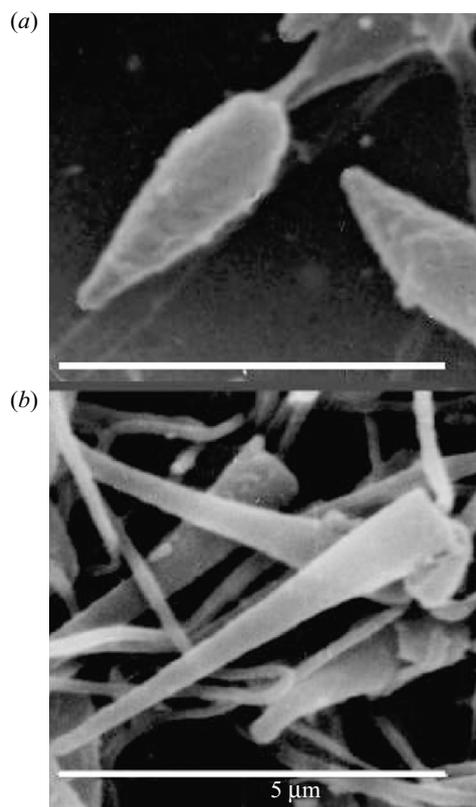


Figure 4. Electron micrograph of sperm of *Echinometra oblonga* from (a) Guam and (b) Okinawa showing the difference in sperm length and shape.

has been observed (Metz & Palumbi 1996; Palumbi 1999). Bindin is involved in the attachment of the sperm to the egg, before its entrance into the egg, and initiates the first steps of fertilization (Vacquier & Moy 1977). This protein has been shown to influence fertilization success within species and has been proposed to participate in reproductive isolation among species (Metz *et al.* 1994; Palumbi 1999). As a result, these bindin gene changes are likely to have functional significance and could change egg–sperm attachment properties.

In fact, a reproductive barrier was suggested by recently reported cross-fertilization experiments between *E. oblonga* from Okinawa versus Guam and Okinawa versus Hawaii. Fertilization rates were high in control crosses, but were near zero for crosses between central and western populations (Arakaki 2000). Experiments with Indonesian or Papua New Guinean individuals have not been performed, and more extensive cross-fertilization work between populations should be completed. However, within-species and between-species variation in fertilization success owing to bindin differences has been demonstrated in other *Echinometra* species (Palumbi 1999) and as the western and central Pacific do not share bindin alleles, the same mechanism could explain the fertilization differences between *E. oblonga* populations.

(b) Sperm morphology

Sperm morphology among most species of *Echinometra* is highly conserved, with sperm heads about twice as long as wide, and central Pacific *E. oblonga* (Hawaii and Guam) have a typical *Echinometra* sperm shape (figures 2 and 4).

In marked contrast, spermatozoa of western Pacific populations (Okinawa, Indonesia and Mauritius) have a long anterior process, and are up to six times longer than they are wide (figure 4; Arakaki & Uehara 1989; Arakaki 2000). As with other mass-spawning invertebrates, sea urchins are known to have little premating behaviour or communication (e.g. Lamare & Stewart 1998). Especially when habitats and spawning seasons may overlap such as in *Echinometra* (Arakaki & Uehara 1991), gametic incompatibility is likely to be an important prezygotic mechanism of isolation (Palumbi 1992). Accordingly, the properties of sperm and eggs may often diverge during species formation, or function as the most proximal characteristics, ensuring the integrity of the species boundaries after secondary contact following allopatric divergence (Geyer & Palumbi 2003).

Sperm size in sea urchins can be polymorphic among populations within species (as in *S. droebachiensis*; C. Biermann, personal communication) but variation in length more than twofold usually represents differences among species (Chia *et al.* 1975; Amy 1983; Raff *et al.* 1990). In the present case, the two divergent sperm morphologies occur across the same geographical boundaries as genetic differences in mtDNA and bindin, and have not been reported to be polymorphic within populations (Arakaki 2000). The unique sperm morphology of the western Pacific *E. oblonga* suggests that this population exhibits a derived state that evolved in this region.

(c) Gamete divergence and sympatry

These concordant results of mtDNA differences, bindin sequence divergence, sperm morphology and fertilization suggest that the two populations of *E. oblonga* in the central and western Pacific have diverged into two sibling species. Although they each deserve a species name, the origin of the *E. oblonga* type specimen is not clearly known and therefore attributing a new species name to either the central or western Pacific populations will require more investigation.

The state of the genetic and morphological traits of the western Pacific *E. oblonga* suggests that it is the derived species. Factors driving the divergence of western Pacific *E. oblonga* are as yet unclear. However, the occurrence of slender sperm morphology and bindin sequence differences are correlated with the presence of *E. sp. C*, a species that generally occupies similar habitats on the western Pacific reefs (Nishihira *et al.* 1991; Uehara *et al.* 1996; Geyer & Palumbi 2003). This pattern is clearest in Guam. Guam is closer to Okinawa than Hawaii (*ca.* 2200 km versus *ca.* 7500 km), and shares many marine species (such as acroporid corals) with Okinawa and the rest of the IP that are absent in Hawaii. Even the echinometrid *E. sp. A* is shared between Guam and Okinawa, but is absent in Hawaii. On this basis, the black urchin in Guam was expected to be the western Pacific *E. oblonga*, but genetic and sperm traits consistently place it within the central Pacific *E. oblonga*. Despite the overall biogeographic similarity of Guam and Okinawa, Guam and Hawaii also share at least one biogeographic feature: they both lack *E. sp. C*.

When they co-occur, *E. oblonga* has been reported to occupy stronger wave action areas than *E. sp. C* (Arakaki & Uehara 1991). The derived sperm morphology might then be an adaptation to this habitat of higher wave

action. This hypothesis is supported by the similar elongated sperm morphology encountered in *Colobocentrotus mertensii* (Arakaki & Uehara 1989), another member of the family Echinometridae inhabiting high wave-action areas. It is also interesting that the sperm head of Okinawan *Echinometra* species increases in length in the order of *E. sp. A* < *E. mathaei* < *E. sp. C* < western Pacific *E. oblonga* but decreases in width in the same order, which correlates with the relative exposure of these species to wave action (Nishihira *et al.* 1991; Arakaki *et al.* 1998*a,b*; Mita *et al.* 2002). Longer sperm may not directly perform better in higher wave action but could be better at fertilizing bigger eggs, as proposed by Raff *et al.* (1990), which in turn could be better and more resistant in higher wave action. Data on egg size have not been reported for all the species but *E. sp. A* has been reported to have a smaller egg diameter than *E. sp. C*, which has a smaller egg diameter than western Pacific *E. oblonga* ($66.9 \pm 1.3 \mu\text{m}$ versus $71.8 \pm 1.4 \mu\text{m}$ versus $74.0 \pm 0.9 \mu\text{m}$) (Aslan *et al.* 1999; Rahman *et al.* 2001).

(d) Speciation in *Echinometra*

Along with previous reports, our results allow us to propose a history of this recent speciation event. Based on the derived genetic state of western Pacific *E. oblonga*, combined with derived sperm type, it appears that *E. oblonga* evolved in allopatry in the central Pacific, diverging from an ancestral stock common in the western Pacific within the past 0.5–1.0 Myr. *Echinometra oblonga* subsequently reinvaded the western Pacific, where it encountered the ancestral species *E. sp. C*, causing rapid changes in sperm morphology and gamete binding proteins. These changes resulted in gametic differences and reproductive barriers between *E. oblonga* currently found in the central Pacific and the new species in the western Pacific. Molecular clock estimates of divergence suggest that this reinvansion and speciation has taken place within the past 250 000 years.

(e) Sperm and speciation

Sex and speciation have been bound intimately in the definition of the biological species concept. Evidence is now accumulating that sexual selection might be acting when species are diverging and that within-species sexual selection may drive the development of reproductive isolation (reviewed in Panhuis *et al.* (2001)). Furthermore, there is ample evidence that sex-related genes are evolving quickly in a broad range of organisms (e.g. *Drosophila*, Civetta & Singh 1995; *Chlamydomonas*, Ferris *et al.* 1997; primates, Wyckoff *et al.* 2000; *Arabidopsis*, Mayfield *et al.* 2001; and marine invertebrates, Metz & Palumbi 1996). These rapidly changing genes may be correlated with the rapid evolution of genitalia and gamete morphology reported in many species. Some evidence suggests that sperm and reproductive organs are indeed rapidly evolving. For instance, in insects, genitalia are often the first morphological characters to change, which is why they are often used in systematics for the differentiation of closely related species (reviewed in Shapiro & Porter 1989). Sperm shape is also known to be rapidly evolving (e.g. Joly *et al.* 1991). In many species, sperm size has been shown to be related to male fitness (see, for example, Radwan 1996) and to respond to selection (e.g.

Gomendio & Roldan 1991; LaMunyon & Ward 2002). Spermatozoa in broadcast spawners are simultaneously both gametes and unicellular genitalia and therefore are expected to evolve quickly as well. Direct attempts to test this hypothesis have seldom been done but indirect evidence suggests that sperm may be coevolving with egg size (Raff *et al.* 1990; Levitan 1996; Levintan & Irvine 2001). Other evidence of rapid evolution comes from the phylogenetic analysis of sperm characteristics, which often suggests rampant secondary loss or reappearance of sperm structures (Rouse 1999; Koufopanou *et al.* 1999; Keys & Healy 2000). Further work is needed to clarify the role of sperm morphology in speciation and how quickly these traits change evolutionarily. This approach may make it possible to investigate the role of specific genes in the evolution of sperm morphology, and thereby elucidate some of the genetic mechanisms by which species evolve.

We thank N. Aubin-Horth and D. Neafsey for comments on an earlier version of the manuscript. C. DeVargas, H. Lessios, D. G. Reid, R. Robertson and S. Vollmer provided some of the specimens used in this study. C.L. is supported by a doctoral fellowship from CRSNG (Conseil de Recherches en Sciences Naturelles et en Génie du Canada) and FCAR (Fond pour la Formation des Chercheurs et d'Aide à la Recherche du Québec) and the Frank Knox Memorial Funds. L.B.G. was supported by graduate research grants from the Department of Organismic and Evolutionary Biology. This research was supported by NSF grants to S.R.P.

REFERENCES

- Amy, R. L. 1983 Gamete sizes and developmental time tables of five tropical sea urchins. *Bull. Mar. Sci.* **33**, 173–176.
- Arakaki, Y. 2000 The black *Echinometra* distributed in the Indo-West Pacific: a species complex. In *Echinoderms 2000* (ed. M. F. Barker), pp. 407–412. Lisse, The Netherlands: Swets & Zeitlinger.
- Arakaki, Y. & Uehara, T. 1989 Comparative external morphology on the spermatozoa of four species, the family of Echinometridae, found on Okinawan reef flat. *Zool. Sci.* **6**, 1221.
- Arakaki, Y. & Uehara, T. 1991 Physiological adaptations and reproduction of the four types of *Echinometra mathaei* (Blainville). In *Biology of Echinodermata* (ed. T. Yanagisawa, I. Yasumasu, C. Oguro, N. Suzuki & T. Motokawa), pp. 105–112. Rotterdam: Balkema.
- Arakaki, Y. & Uehara, T. 1999 Morphological comparison of black *Echinometra* individuals among those in the Indo-West Pacific. *Zool. Sci.* **16**, 551–558.
- Arakaki, Y., Uehara, T. & Fagoonee, I. F. 1998*a* Comparison between *Echinometra* species from Mauritius and Okinawa. In *Proc. 9th Int. Echinoderm Conf.* (ed. R. Mooi & M. Telford), pp. 547–551. Rotterdam: Balkema.
- Arakaki, Y., Uehara, T. & Fagoonee, I. F. 1998*b* Comparative studies of the genus *Echinometra* from Okinawa and Mauritius. *Zool. Sci.* **15**, 159–168.
- Aslan, L. M., Rahman, M. & Uehara, T. 1999 Relationship between egg size and larval size of two closely related species of sea urchins genus *Echinometra* and the egg size effects to the larval size of their reciprocal hybrids. *Biol. Mag. Okinawa* **37**, 63–73.
- Barber, P. H., Palumbi, S. R., Erdmann, M. V. & Moosa, M. K. 2002 Sharp genetic breaks among populations of *Haptosquilla pulchella* (Stomatopoda) indicate limits to larval transport: patterns, causes, and consequences. *Mol. Ecol.* **11**, 659–674.

- Bellwood, D. R. & Hughes, T. P. 2001 Regional-scale assembly rules and biodiversity of coral reefs. *Science* **292**, 1532–1534.
- Briggs, J. C. 1999 Coincident biogeographic patterns: Indo-West Pacific Ocean. *Evolution* **52**, 326–335.
- Chia, F. S., Atwood, D. & Crawford, B. 1975 Comparative morphology of echinoderm sperm and possible phylogenetic implications. *Am. Zool.* **15**, 533–565.
- Civetta, A. & Singh, R. S. 1995 High divergence of reproductive tract proteins and their association with postzygotic reproductive isolation in *Drosophila melanogaster* and *Drosophila virilis* group species. *J. Mol. Evol.* **41**, 1085–1095.
- Ferris, P. J., Pavlovic, C., Fabry, S. & Goodenough, U. W. 1997 Rapid evolution of sex-related genes in *Chlamydomonas*. *Proc. Natl Acad. Sci. USA* **94**, 8634–8639.
- Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. C. 1994 DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotechnol.* **3**, 284–299.
- Geyer, L. B. & Palumbi, S. R. 2003 Reproductive character displacement and the genetics of gamete recognition in tropical sea urchins. *Evolution* **57**, 1049–1060.
- Gomendio, M. & Roldan, E. R. S. 1991 Sperm competition influence sperm size in mammals. *Proc. R. Soc. Lond. B* **243**, 181–185.
- Joly, D., Bressac, C., Devaux, J. & Lachaise, D. 1991 Sperm length diversity in Drosophilidae. *Dros. Inf. Serv.* **70**, 104–108.
- Kelso, D. 1970 A comparative morphological and ecological study of two species of sea urchins: genus *Echinometra* in Hawaii. PhD thesis, University of Hawaii, Honolulu, USA.
- Keys, J. L. & Healy, J. M. 2000 Relevance of sperm ultrastructure to the classification of giant clams (Mollusca, Cardioidea, Cardiidae, Tridacninae). In *The evolutionary biology of the Bivalvia* (ed. E. M. Harper, J. D. Taylor & J. A. Crame), pp. 191–205. The Geological Society of London.
- Knowlton, N. 1993 Sibling species in the sea. *A. Rev. Ecol. Syst.* **24**, 189–216.
- Koufopanou, V., Reid, D. G., Ridgway, S. A. & Thomas, R. H. 1999 A molecular phylogeny of the patellid limpets (Gastropoda: Patellidae) and its implications for the origins of their antitropical distribution. *Mol. Phylogenet. Evol.* **11**, 138–156.
- Kumar, S., Tamura, K., Jakobsen, I. & Nei, M. 2001 MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics* **17**, 1244–1245.
- Lamare, M. D. & Stewart, B. G. 1998 Mass spawning by the sea urchin *Evechinus chloroticus* (Echinodermata: Echinoidea) in a New Zealand fiord. *Mar. Biol.* **132**, 135–140.
- LaMunyon, C. W. & Ward, S. 2002 Evolution of larger sperm in response to experimentally increased sperm competition in *Caenorhabditis elegans*. *Proc. R. Soc. Lond. B* **269**, 1125–1128. (DOI 10.1098/rspb.2002.1996.)
- Lessios, H. A., Kessing, B. D. & Pearse, J. 2001 Population structure and speciation in tropical seas: global phylogeography of the sea urchin *Diadema*. *Evolution* **55**, 955–975.
- Levitan, D. R. 1996 Effects of gamete traits on fertilization in the sea and the evolution of sexual dimorphism. *Nature* **382**, 153–155.
- Levitan, D. R. & Irvine, S. D. 2001 Fertilization selection on egg and jelly-coat size in the sand dollar *Dendraster excentricus*. *Evolution* **55**, 2479–2483.
- McCartney, M. A., Keller, G. & Lessios, H. A. 2000 Dispersal barriers in tropical oceans and speciation in Atlantic and eastern Pacific sea urchins of the genus *Echinometra*. *Mol. Ecol.* **9**, 1391–1400.
- McMillan, W. O., Weigt, L. A. & Palumbi, S. R. 1999 Color pattern evolution, assortative mating, and genetic differentiation in brightly colored butterfly fishes (Chaetodontidae). *Evolution* **53**, 247–260.
- Matsuoka, N. & Hatanaka, T. 1991 Molecular evidence for the existence of four sibling species within the sea-urchin, *Echinometra mathaei* in Japanese waters and their evolutionary relationships. *Zool. Sci.* **8**, 121–134.
- Mayfield, J. A., Fiebig, A., Johnstone, S. E. & Preuss, D. 2001 Gene families from the *Arabidopsis thaliana* pollen coat proteome. *Science* **292**, 2482–2485.
- Mead, K. S. 1996 Sex in the surf zone: the effect of hydrodynamic shear stress on the fertilization and early development of free-spawning invertebrates. PhD thesis, Stanford University, CA, USA.
- Metz, E. C. & Palumbi, S. R. 1996 Positive selection and sequence rearrangements generate extensive polymorphism in the gamete recognition protein bindin. *Mol. Biol. Evol.* **13**, 397–406.
- Metz, E. C., Kane, R. E., Yanagimachi, H. & Palumbi, S. R. 1994 Fertilization between closely related sea urchins is blocked by incompatibilities during sperm-egg attachment and early stages of fusion. *Biol. Bull.* **187**, 23–34.
- Mita, M., Uehara, T. & Nakamura, M. 2002 Comparative studies on the energy metabolism in spermatozoa of four closely related species of sea urchins (genus *Echinometra*) in Okinawa. *Zool. Sci.* **19**, 419–427.
- Mortensen, T. H. 1943 *Monograph of the Echinoidea: Camarondonta*. Copenhagen: C. A. Reitzel.
- Nishihira, M., Sato, Y., Arakaki, Y. & Tsuchiya, M. 1991 Ecological distribution and habitat preference of four types of the sea urchin *Echinometra mathaei* on the Okinawan coral reefs. In *Biology of Echinodermata* (ed. T. Yanagisawa, I. Yasumasu, C. Oguro, N. Suzuki & T. Motokawa), pp. 91–104. Rotterdam: Balkema Press.
- Palumbi, S. R. 1992 Marine speciation on a small planet. *Trends Ecol. Evol.* **7**, 114–117.
- Palumbi, S. R. 1996 What can molecular genetics contribute to marine biogeography? An urchin's tale. *J. Exp. Mar. Biol. Ecol.* **203**, 75–92.
- Palumbi, S. R. 1997 Molecular biogeography of the Pacific. *Coral Reefs* **16**, S47–S52.
- Palumbi, S. R. 1999 All males are not created equal: fertility differences depend on gamete recognition polymorphisms in sea urchins. *Proc. Natl Acad. Sci. USA* **96**, 12 632–12 637.
- Palumbi, S. R. & Metz, E. C. 1991 Strong reproductive isolation between closely related tropical sea urchins (genus *Echinometra*). *Mol. Biol. Evol.* **8**, 227–239.
- Palumbi, S. R., Grabowsky, G., Duda, T., Geyer, L. & Tachino, N. 1997 Speciation and population genetic structure in tropical Pacific sea urchins. *Evolution* **51**, 1506–1517.
- Palumbi, S. R., Cipriano, F. & Hare, M. P. 2001 Predicting nuclear gene coalescence from mitochondrial data: the three-times rule. *Evolution* **55**, 859–868.
- Panhuis, T. M., Butlin, R., Zuk, M. & Tregenza, T. 2001 Sexual selection and speciation. *Trends Ecol. Evol.* **10**, 228–231.
- Radwan, J. 1996 Intraspecific variation in sperm competition success in the bulb mite: a role for sperm size. *Proc. R. Soc. Lond. B* **263**, 855–859.
- Raff, R. A., Herlands, L., Morris, V. B. & Healy, J. 1990 Evolutionary modification of echinoid sperm correlates with developmental mode. *Dev. Growth Differ.* **32**, 283–291.
- Rahman, M., Uehara, T. & Pearse, J. 2001 Hybrids of two closely related tropical sea urchins (genus *Echinometra*): evidence against postzygotic isolating mechanisms. *Biol. Bull.* **200**, 97–106.
- Rouse, G. W. 1999 Polychaete sperm: phylogenetic and functional considerations. *Hydrobiologia* **402**, 215–224.

- Shapiro, A. M. & Porter, A. H. 1989 The lock-and-key hypothesis: evolutionary and biosystematic interpretation of insect genitalia. *A. Rev. Entomol.* **34**, 231–245.
- Springer, V. G. & Williams, J. T. 1990 Widely distributed Pacific plate endemics and lowered sea-level. *Bull. Mar. Sci.* **47**, 631–640.
- Swofford, D. 2001 *PAUP*: phylogenetic analysis using parsimony and other methods*, v. 4. Sunderland, MA: Sinauer.
- Uehara, T., Shingaki, M. & Taira, K. 1986 Taxonomic studies in the sea urchin, genus *Echinometra*, from Okinawa and Hawaii. *Zool. Sci.* **3**, 1114.
- Uehara, T., Asakura, H. & Arakaki, Y. 1990 Fertilization blockage and hybridization among species of sea urchins. In *Advances in invertebrate reproduction*, vol. 5 (ed. M. Hoshi & O. Yamashita), pp. 305–310. Amsterdam: Elsevier Science.
- Uehara, T., Tsukahara, J. & Tsukashima, M. 1996 Distribution and characteristics of sea urchin genus *Echinometra* from Palau. Kagoshima University Research Centre South Pacific Occasional Papers, pp. 55–59.
- Vacquier, V. D. & Moy, G. W. 1977 Isolation of bindin: the protein responsible for adhesion of sperm to sea urchin eggs. *Proc. Natl Acad. Sci. USA* **74**, 2456–2460.
- Wyckoff, G. J., Wang, W. & Wu, C. I. 2000 Rapid evolution of male reproductive genes in the descent of man. *Nature* **403**, 304–309.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.



Contrasting patterns in species boundaries and evolution of anemonefishes (Amphiprioninae, Pomacentridae) in the centre of marine biodiversity

Janne Timm*, Malgorzata Figiel, Marc Kochzius

Biotechnology and Molecular Genetics, FB2-UFT, University of Bremen, Leobener Strasse UFT, 28359 Bremen, Germany

ARTICLE INFO

Article history:

Received 16 April 2008

Accepted 18 April 2008

Available online 26 April 2008

Keywords:

Amphiprion
Molecular Phylogeny
Phylogeography
Southeast Asia
Clownfishes

ABSTRACT

Many species of coral reef fishes are distinguished by their colour patterns, but genetic studies have shown these are not always good predictors of genetic isolation and species boundaries. The genus *Amphiprion* comprises several species that have very similar colouration. Additionally, morphological characters are so variable, that sibling species can show a considerable overlap, making it difficult to differentiate them. In this study, we investigated the species boundaries between the sibling species pair *A. ocellaris* and *A. percula* (Subgenus *Actinicola*) and three closely related species of the subgenus *Phalerebus* (*A. akallopisos*, *A. perideraion*, *A. sandaracinos*) by phylogenetic analysis of mitochondrial cytochrome *b* and control region sequences. These two subgenera show strong differences in their patterns of species boundaries. Within the *A. ocellaris/A. percula* complex, five clades were found representing different geographic regions. Two major divergences both with genetic distances of 4–7% in *cty b* and 17–19% in the *d*-loop region indicate the presence of three instead of two deep evolutionary lineages. The species of the subgenus *Phalerebus* show three monophyletic clades, independent of the geographical location of origin, but concordant to the morphological species classification. The genetic distances between the *Phalerebus* species were 2–5% in *cty b* and 10–12% in the control region.

© 2008 Elsevier Inc. All rights reserved.

1. Introduction

The Indo-Malay Archipelago, also called the “coral triangle”, contains the world’s richest marine shallow water biodiversity (Briggs, 2005; Allen and Werner, 2002) and is therefore a well-suited region to study divergence and speciation processes. If the high diversity in this area is caused by an overlap of the Pacific and Indian Ocean faunas (Woodland, 1983), or the result of an accumulation of species that evolved at the periphery (Jokiel and Martinelli, 1992), or if it is actually the “center of origin” where species evolve (Briggs, 2000, 2005) is discussed controversially.

In the genus *Amphiprion*, the latter theory was supported by Santini and Polacco (2006), who found the “center of origin” in an area reaching from the Philippines to the Great Barrier Reef and from Sumatra to Melanesia, which does not exclude speciation in peripheral remote areas (e.g. *A. trilineatus*, Marshall Islands endemic). Furthermore, a rather recent radiation was indicated in the Indian Ocean, because derived and endemic species are dominant (Santini and Polacco, 2006).

In order to evaluate biodiversity correctly it is important to clarify species boundaries, integrities, and phylogenetic relationships (Frankham et al., 2002). Many species of coral reef fishes are

distinguished by their colour patterns, but genetic studies have shown that these are not always sufficient indicators of genetic isolation and species boundaries (Bernardi et al., 2002). The genus *Amphiprion* comprises several species with very similar colouration. Additionally, morphological characters are so variable, that sibling species can show a considerable overlap, making it difficult to differentiate them.

There are closely related species that only show slight differences in their colour pattern in the subgenus *Phalerebus* (Allen, 1991), such as *A. akallopisos* and *A. sandaracinos*. *A. akallopisos* has a white caudal fin and an orange to pinkish body colour, whereas *A. sandaracinos* has an orange caudal fin and its body colour is usually clearly orange. The white stripe on the back is supposed to be slightly longer in *A. sandaracinos*, spanning from the upper lip to the caudal peduncle, whereas in *A. akallopisos* it begins more on the forehead. However, this character seems to vary especially in the latter species. Regarding our observations, the white stripe often started also at the upper lip in *A. akallopisos*, so it is an overlapping and therefore rather weak character for distinguishing these two species. A more stable character separating the two species is the differently shaped teeth, indicating slightly different ecological adaptations (Fautin and Allen, 1994). *A. perideraion* has a similar body colour like *A. akallopisos*, but shows an additional white stripe between head and trunk. *A. perideraion* and *A. sandaracinos* show a sympatric distribution with the latter having a more restricted range and higher host specificity, accepting only two

* Corresponding author.

E-mail addresses: janne.timm@uni-bremen.de, janneket@gmx.de (J. Timm).

anemone species (*Heteractis crista* and *Stichodactyla mertensii*) instead of four (*Heteractis crista*, *H. magnifica*, *Stichodactyla gigantea*, and *Macroactyla dorensis*) in *A. perideraion* (Fautin and Allen, 1994). *A. akallopisos* has a parapatric distribution with the former two species, overlapping around the upper Sunda Islands and is also associated with only two anemone species (*Heteractis magnifica* and *Stichodactyla mertensii*; Fautin and Allen, 1994; Fig. 1a).

The sibling species *A. ocellaris* and *A. percula* show more or less the same colour pattern, although *A. percula* is described to have larger black bands in its colouration (Fautin and Allen, 1994), which could not be confirmed by our observations. There is rather a large variation: some specimens showing no black bands in their colouration while others do. Morphologically, these two species are differentiated by the number of spines in the dorsal fin, but also this character is overlapping: *A. ocellaris* has 10–11 and *A. percula* 9–10 spines. The ecological requirements of both species seem to be identical; both of them prefer the same host anemone species (*Heteractis magnifica* and *Stichodactyla gigantea*). Regarding Fautin and Allen (1994) these siblings have an allopatric distribution (Fig. 1b), but Kuitert and Tonzuka (2004) reported both species in the Tomini Bay (Sulawesi), which indicates a parapatric distribution.

The high morphological similarity of the above mentioned species raises the question whether these species form distinct genetic clades within the subgenera. Additionally, their similar biology could lead to similar species boundaries patterns.

This study aims to reveal (1) species boundaries within the anemonefish genus *Amphiprion* in the Indo-Malay Archipelago, and (2) speciation processes in the hotspot of marine shallow water biodiversity.

2. Materials and methods

2.1. Sampling

A total of 86 tissue samples of five coral reef associated fish species of the genus *Amphiprion* were collected at different locations in the Indo-Malay Archipelago (Table 1 and Fig. 1). The fishes were caught with two aquarium nets. A fin clip from the caudal fin was taken and the fishes were released into their host anemones. It was therefore possible to obtain tissue samples without killing the animals. The samples were stored in 96% ethanol.

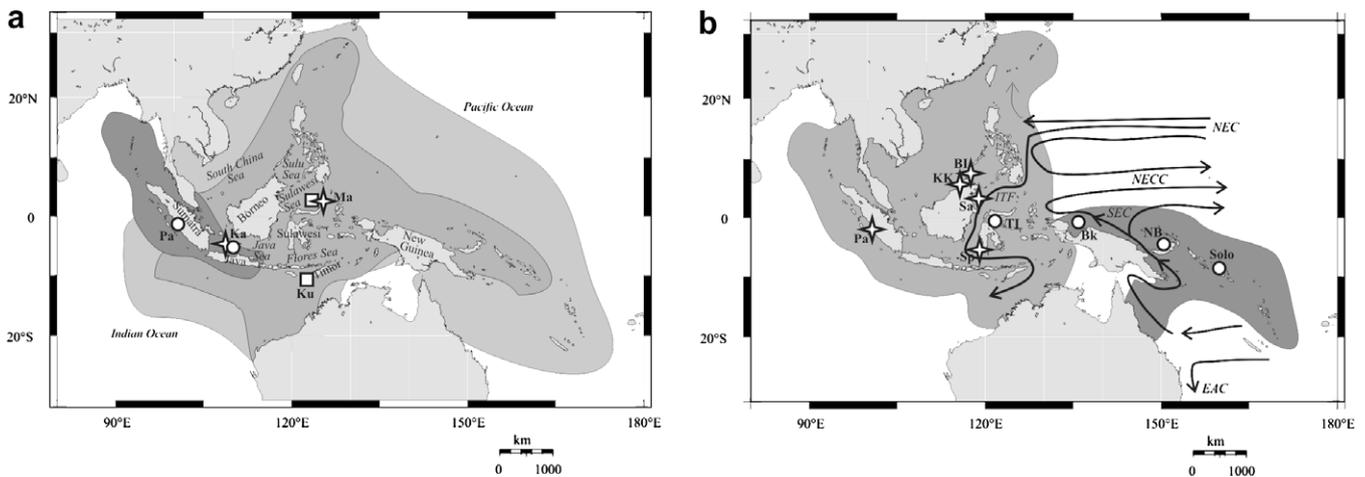


Fig. 1. (a) Distribution patterns and sample sites of the species *Amphiprion akallopisos* (dark grey, circles), *A. perideraion* (grey, stars) and *A. sandaracinos* (light grey, diamonds) in the Indo-Malay Archipelago (Fautin and Allen 1994). Names of the main ocean basins and islands are added. (b) Distribution patterns and sample sites of the species *Amphiprion ocellaris* (grey, stars) and *A. percula* (dark grey, circles) in the Indo-Malay Archipelago (Fautin and Allen 1994). Dominant currents are added (simplified after Godfrey, 1996).

Table 1

Number of sampled individuals from 5 species of the genus *Amphiprion*, abbreviations used (Abbr.), accession numbers (EMBL), and the corresponding sample locations in the Indo-Malay Archipelago

Species	Location	Abbr.	No. CR	Accession No. CR	No. cyt b	Accession No. cyt b
<i>A. ocellaris</i>	Spermonde, Sulawesi	AoS _p	5	AM747125–AM747129	3	AM942669–AM942671
	Banggi Islands, North Borneo	AoB _I	5	AM747144–AM747148	-	-
	Kota Kinabalu, Northwest Borneo	AoK _K	6	AM941144–AM941149	6	AM942678–AM942683
	Sangkalaki, Northeast Borneo	AoS _a	6	AM941138–AM941143	6	AM942672–AM942677
	Padang, Sumatra	AoS _p	6	AM747139–AM747143	6	AM942684–AM942689
<i>A. percula</i>	Tomini Bay, Sulawesi	Aper _T	9	AM747130–AM747138	9	AM942690–AM942695
	Biak, New Guinea	Aper _B	10	AM747149–AM747157	6	AM942696–AM942701
	New Britain, New Guinea	Aper _N	4	AM745732–AM745734	4	AM942702–AM942705
	Solomon Islands(Genebank)	Aper _S	1	DQ343939	1	DQ343958
<i>A. akallopisos</i>	Padang, Sumatra	AaP _a	5	AM747158–AM747162	2	AM942706–AM942707
	Karimun Jawa, Jawa Sea	AaK _a	5	AM747163–AM747167	3	AM942708–AM942710
<i>A. perideraion</i>	Manado, Sulawesi	ApM _a	5	AM747168–AM747171	3	AM942711–AM942713
	Karimun Jawa, Jawa Sea	ApK _a	5	AM747172–AM747174	3	AM942714–AM942716
<i>A. sandaracinos</i>	Manado, Sulawesi	AsM _a	5	AM747175–AM747178	3	AM942717–AM942719
	Kupang, Timor	AsK _u	5	AM747179–AM747183	2	AM942720–AM942721

The sequence from the Solomon Islands was obtained from Genebank. CR, control region; cyt b, cytochrome b.

Thirty-two samples of *A. ocellaris* were obtained from five different locations over a range of 1500 km. The sibling species *A. percula* was collected in Biak (New Guinea; 10 samples) and from New Britain (New Guinea; 4 samples). An additional number of 9 tissue samples have been collected from depths between 1 and 25 m from the Togian Islands (Tomini Bay, Sulawesi). Since *A. ocellaris* is distributed across the Indo-Malay Archipelago (Fautin and Allen, 1994), with its eastern border of distribution around the Molucca Islands and the most western tip of New Guinea, the samples from the Togian Islands have been classified as *A. ocellaris*. One sequence of *A. percula* from Solomon Islands was added for each marker from GeneBank (Table 1).

Within the subgenus *Phalerebus*, 10 specimens from two locations in the Indo-Malay Archipelago were analysed for each of the species *A. akallopisos*, *A. sandaracinos*, and *A. perideraion*.

As outgroup for the phylogenetic analyses *Chromis viridis* was used, which is a member of the same family (Pomacentridae).

2.2. DNA extraction and amplification

Genomic DNA was extracted with filter column based extraction kits from Qiagen and Macherey–Nagel, following the manufacturers' protocols.

A fragment with a maximum length of 420 bp of the mitochondrial control region (CR) was amplified by PCR with the primers CR-A (TTC CAC CTC TAA CTC CCA AAG CTA G) and CR-E (CCT GAA GTA GGA ACC AGA TG) (Lee et al., 1995). PCR was performed in a Perkin Elmer and Eppendorf Ep S Mastercycler with the following thermo-profile: 95 °C for 2 min, followed by 35 cycles of 95 °C for 30 s, 50 °C for 30 s and 72 °C for 60 s. The Terminal elongation was at 72 °C for 2 min. 25 µl reactions contained 2.5 µl 10× PCR buffer, 0.075 µmol Mg²⁺, 0.25 µmol dNTP mix, 10 pmol of each primer and 0.5 U Taq polymerase. Between 10 and 30 ng genomic DNA was used of each sample as template.

For a subset of samples of each species cytochrome *b* (*cyt b*) sequences were obtained (Table 1). This fragment of the mitochondrial genome is suitable for resolving phylogenetic patterns on intraspecific (Nelson et al., 2000) to intrageneric level (Kocher et al., 1989).

For the amplification of the *cyt b* fragment of around 400 bp length, the primers tRNAgluF (AAAACCACCGTTGTTATTCACACTACA; Nelson et al., 2000) and H15149 (AAACTGCAGCCCCTCAGAATGATATTTGTCCTCA; Kocher et al., 1989) were used. PCR was performed in Eppendorf Ep Mastercyclers in 25 µl reaction mix, containing 2.5 µl 10× PCR buffer, 0.0625 µmol Mg²⁺, 0.25 µmol dNTP mix, 10 pmol of each primer and 0.5 U Taq polymerase. Again 10–30 ng genomic DNA was used of each sample. The thermo-profile was 95 °C for 5 min, 35 cycles of 95 °C for 45 s, 58 °C for 45 s and 72 °C for 60 s. The final elongation was 72 °C for 5 min.

All PCR products were purified with the QIA-quick PCR Purification Kit (Qiagen). Sequencing of both strands was conducted with the PCR primers using the Big Dye Terminator Cycle Sequencing Kit (ver. 1.3 and ver. 3.1; Applied Bioscience) according to the manufacturer's recommendations and an ABI 310 and 3100 automated sequencer.

2.3. Phylogenetic analyses

Both strands were assembled and edited with help of the programme Seqman (Ver. 4.05, DNASTAR). Multiple sequence alignment was done using Clustal W (Thompson et al., 1994) as implemented in the software Bioedit (ver. 7.0.0.1, Hall, 1999).

The phylogenetic relationships of species were analysed with all sequences available for *cyt b* and a subset of sequences for the CR, adjusted to the *cyt b* dataset, by using maximum parsimony (MP), maximum likelihood (ML) and neighbour joining (NJ) algorithms

with the software Paup* (Ver. 4.0b10; Swofford, 1998). The statistical confidences were evaluated by 1000 non-parametric bootstrap replicates for NJ and MP analyses and by 100 for ML analysis. In order to test if the molecular clock will be rejected, the ML analysis was done with and without molecular clock enforced. The Shimodaira–Hasegawa (Goldman et al., 2000; Shimodaira and Hasegawa, 1999) and Kishino–Hasegawa (Kishino and Hasegawa, 1989) tests were used to verify if the topologies of the two ML trees are significantly different.

Haplotype networks were drawn including the CR sequences of all samples and *cyt b* sequences from a subset of samples, based on the results obtained from Arlequin (Schneider et al., 2000).

The program Modeltest (Ver. 3.06; Posada and Cranall, 1998) was used to determine the best-fit model of DNA evolution for the two datasets.

Sequence divergences between individuals were calculated with Paup* (Swofford, 1998) and the average within and between each group was given, as well as the genetic distances after correcting for within population diversity ($d_A = d_{XY} - (\pi_X + \pi_Y)/2$, Nei 1987 as cited in Campton et al., 2000). The corrected genetic distances were further used for comparison and molecular divergence time calculations, whereat the latter were done based on a mutation rate of 6.41% per million years for the CR, estimated in a study on the phylogeography of the coral reef fish *Pseudochromis fuscus* (Messmer et al., 2005). This species belongs to the same order and the same CR fragment was used, including a conserved and a hypervariable region (Alvarado et al., 1995). For *cyt b*, mutation rates of 1.0–2.8% per million years were utilised, as assumed for different fish species (Ortí et al., 1994; Martin and Bermingham, 1998; Perdices et al., 2002; Chenoweth et al., 2002; Banford et al., 2004; Casey et al., 2004).

3. Results

3.1. Phylogenetic trees

An alignment of 371 base pairs of the CR fragment, containing 55 sequences from five species, was obtained. The alignment included several gaps and the Ts/Tv ratio was 1.77. The best-fit model of evolution for the present dataset was the General Time Reversible model (GTR; Tavaré, 1986) with a proportion of invariable sites of 0.17, and a gamma distribution shape parameter of 1.34.

The 55 sequences of the *cyt b* fragment resulted in an alignment of 357 base pair length. The sequences represent the same species as the CR sequences. The best-fit model of evolution for the *cyt b* dataset was the Total Variation Metric model (TVM; Pond, 2007), with a proportion of invariable sites of 0.67. The latter model is a modified GTR model with equal substitution rates for A-G and C-T (Paraskeris et al., 2004). In the *cyt b* dataset the Ts/Tv ratio was 2.24. Of the observed 93 substitutions, 94% were at third codon positions, not changing the amino acid sequence of the fragment, whereas 4% were at first and 2% at second codon positions.

These evolutionary models were used for the NJ and ML analyses.

The phylogenetic analysis of the CR dataset is presented as a NJ cladogram with bootstrap values of NJ, MP and ML analysis (Fig. 2a), showing a grouping of all species in two main clades, both of them supported by high bootstrap values.

One clade contained the species *A. sandaracinos*, *A. akallopisos*, and *A. perideraion*, the other *A. ocellaris* and *A. percula*. The species of the subgenus *Phalerebus* (Allen, 1991) form monophyletic subclades in all analyses, well supported by bootstrap values between 97% and 100% for the MP and NJ analysis. The ML analysis gave weaker bootstrap support for many branches. With a low bootstrap support of 66% only in the NJ analysis *A. sandaracinos* and *A. peri-*

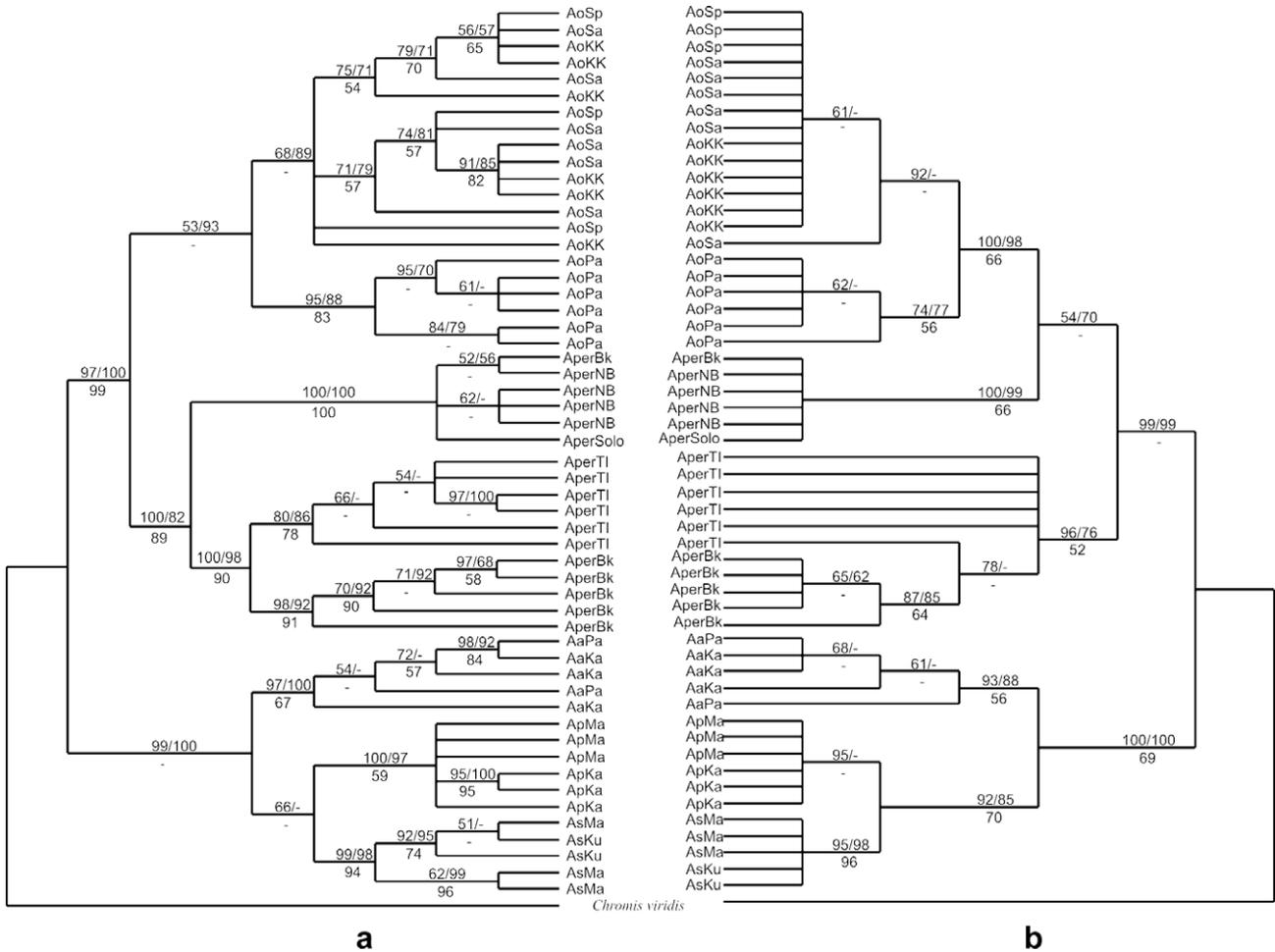


Fig. 2. Neighbour-Joining trees of CR (a) and *cyt b* (b) sequences of anemonefish species from the genus *Amphiprion* in the Indo-Malay Archipelago. Bootstrap values above branches are based on 1000 replicates for the NJ/MP analyses and below branches on 100 replicates for the ML analysis. For abbr. see Table 1.

perideraion formed sister clades, and *A. akallopisos* was basal to them (Fig. 2a).

The species *A. ocellaris* and *A. percula* formed one well-supported monophyletic clade, in all analyses, with a clear subdivision into an *A. ocellaris* and an *A. percula* sub-clade. The specimens from Togian Islands (Tomini Bay) were found within the *A. percula* clade. This clade showed a subdivision into three branches, corresponding to the geographical regions of Tomini Bay, Biak, and New Britain. The sequence from the Solomon Islands was included into the New Britain group. One sample from Biak was as well clearly found within the New Britain subdivision. A phylogeographic structure regarding the population from Padang (Indian Ocean) was found in *A. ocellaris* (Fig. 2a).

The phylogenetic analyses done for the *cyt b* dataset is represented by a NJ tree as well. In this case, the grouping of the *A. ocellaris* and *A. percula* clades was not resolved in the ML analysis. All algorithms showed the same major groupings as the CR dataset, but in the NJ and MP analyses the branch of *A. percula* from New Britain and the Solomon Islands was associated to *A. ocellaris* instead of *A. percula* from Tomini Bay and Biak (Fig. 2b). This inconsistency was not well supported (Bootstrap values of 54 for NJ and 70 for MP) and not shown in the ML analysis, which did not resolve the relationships among the *A. ocellaris*/*A. percula* species complex. In the NJ tree one sequence of *A. percula* from Tomini Bay was grouped basal to the Biak clade, but this was not supported by the other algorithms and the other *A. percula* sequences from Tomini Bay were not resolved at all. The same sample of *A. percula* from

Biak which appeared in the *A. percula* group from New Britain in the CR dataset was found in the New Britain clade in the *cyt b* dataset as well.

3.2. Haplotype networks

The parsimonious haplotype networks of the *A. akallopisos*/*perideraion*/*sandaracinos* complex revealed three clearly separated clades for both markers (Fig. 3). These clades were concordant to the morphologically defined species and no geographical pattern could be observed. For the CR network the numbers of mutational steps in the *A. akallopisos* clade were 6–13, and in the *A. perideraion* clade 1–21. Two sub-clades, separated by 19 mutations, were present in *A. sandaracinos*. Within these sub-clades the variation was between one and four steps. The mutational steps between species pairs were 59 (*A. akallopisos*–*A. perideraion*) and 53 (*A. perideraion*–*A. sandaracinos*). In the *cyt b* network the number of mutational steps between *A. akallopisos* and *A. perideraion* was 10 and between *A. perideraion* and *A. sandaracinos* 5. Within the species clades, there was one dominant haplotype in each, and one or two other haplotypes separated by only one or two mutational steps.

The haplotype network of the *A. ocellaris*/*percula* species complex based on the CR dataset showed that *A. ocellaris* is separated from *A. percula* collected in Tomini Bay by 80 substitutions (Fig. 4a). The genetic break between *A. ocellaris* individuals from Padang and its conspecifics from other sample sites was determined by 31 substitutions. Two to 25 mutational steps separated the haplotypes

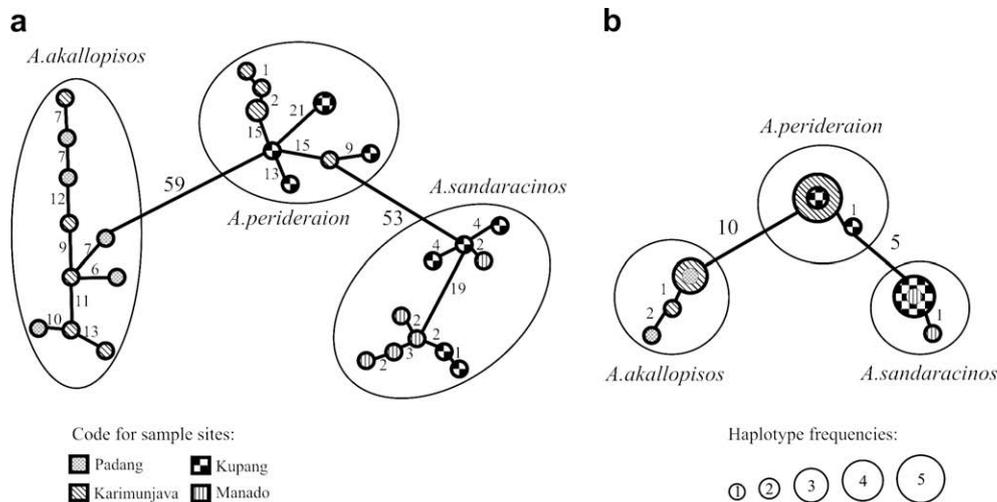


Fig. 3. Haplotype networks of (a) CR and (b) *cyt b* sequences of *Amphiprion akallopisos*, *A. perideraion* and *A. sandaracinos* from different locations in the Indo-Malay Archipelago. Pattern fillings of circles correspond to the different sample locations and numbers indicate mutational steps. The size of the circles indicate the haplotype frequencies.

within each clade of this species. The populations of *A. percula* in Tomini Bay and Biak were divided by 41 steps. The sequences of *A. percula* from New Britain were strongly separated to those from Biak by 82 substitutions. Within each population of *A. percula* we found 2–26 mutational steps. Especially among haplotypes from Tomini Bay a high variability could be observed.

The *cyt b* network showed the same division into five clades corresponding to geographical regions. The population of Padang was separated from the other *A. ocellaris* populations by three steps. The same separation was revealed between *A. percula* from Tomini Bay and Biak. The mutational steps between *A. ocellaris* and *A. percula* from Tomini Bay was with 15 noticeable higher. In this dataset, the clade of New Britain and Solomon Islands was connected to Tomini Bay, separated by 13 mutations, instead of Biak, as it was shown in the CR dataset. In both networks, one sequence, sampled in Biak, was included into the New Britain clade, as shown in the trees.

3.3. Sequence divergence

The sequence divergences for both markers within and among clades, as well as the genetic distances between clades after accounting for the diversity within them are shown in Table 2.

The genetic distances of 6.6% in CR and 1% in *cyt b* sequences between *A. percula* from Tomini Bay and Biak was within the same range as the distances between *A. ocellaris* from Padang and other populations of this species (5.8% and 0.9%, respectively). Most of these values are between the maximum value within clades of species (*A. percula*; Togian Islands = 6% in CR and 0.4% in *cyt b*) and the minimum value among species (*A. perideraion*–*A. sandaracinos*, CR = 10.1% and *cyt b* = 1.5%) observed in the datasets. The genetic distances between *A. percula* from Tomini Bay/Biak and New Britain/Solomon Islands were 16.7–18% in CR and 3.8–5% in *cyt b* sequences. In comparison, the genetic differences in the *A. akallopisos/sandaracinos/perideraion* complex did not exceed 12.3% in CR and 4.8% in *cyt b* (Table 2).

3.4. Molecular clock

The Shimodaira–Hasegawa and Kishino–Hasegawa tests, conducted to compare the ML trees reconstructed with and without the molecular clock enforced, showed no significant difference for both markers (CR: $p = 0.288$ and $p = 0.715$, respectively; *cyt b*:

$p = 0.238$ and $p = 0.847$, respectively). Therefore, the molecular clock was not rejected and the separation among the species' ancestors could be estimated. With the used mutation rate of 6.41% per million years for the CR, the genetic distance of 11.1% could be translated to approx. 1.7 million years divergence time between *A. akallopisos* and *A. perideraion*. This time frame fits into the divergence range assumed for the *cyt b* fragment (1.1–3.1 million years ago), comprising the geological border between Pliocene and Pleistocene (Table 3). The genetic distance of 12.3% between *A. akallopisos* and *A. sandaracinos* of the CR was translated to 1.9 million years and therefore revealed a slightly longer separation time between these species. The divergence range calculated for the *cyt b* fragment dated the split between *A. akallopisos* and *A. sandaracinos* as well into the Pleistocene - Pliocene border (1.7–4.8 million years ago). *A. sandaracinos* and *A. perideraion* revealed, with 1.6 million years for the CR and a range of 0.5–1.5 million years for *cyt b*, a more recent split in the Pleistocene (Table 3).

The genetic distances between the sibling species *A. percula* and *A. ocellaris* were between 16.6–18.5% in CR and 4.8–7.2% in *cyt b*, which indicated that the split between them is approx. 1.7–7.2 (2.6–2.9 for CR) million years old. The divergence time between the Indian Ocean clade of *A. ocellaris* and the other populations within the Indo-Malay Archipelago reaches back 300,000 to 900,000 years (900,000 years for CR). This is similar to the separation of the populations of *A. percula* from Tomini Bay and Biak, which is 400,000 years to 1 million years old (1 million years in CR). The divergence between Biak/Tomini Bay and New Britain/Solomon Islands is 1.4–5 million years old (2.6–2.8 million years in CR), similar to the split between *A. ocellaris* and *A. percula* (Table 3).

4. Discussion

Molecular phylogenetic analyses of closely related species provide insights into their relationships, allowing us to verify their morphological taxonomic classification. Sometimes, such studies indicate that the previously assumed classification is wrong or not sufficient. This was the case in a study on the *Dascyllus trimaculatus* species complex (Bernardi et al., 2001), in which the authors revealed inconsistency between morphological and colouration traits, as well as their molecular phylogenetic relationships. In contrast to that, research on coral reef fishes of the genus *Thalassoma* (Costagliola et al., 2004) confirmed the morphological species definition, even though the colouration pattern

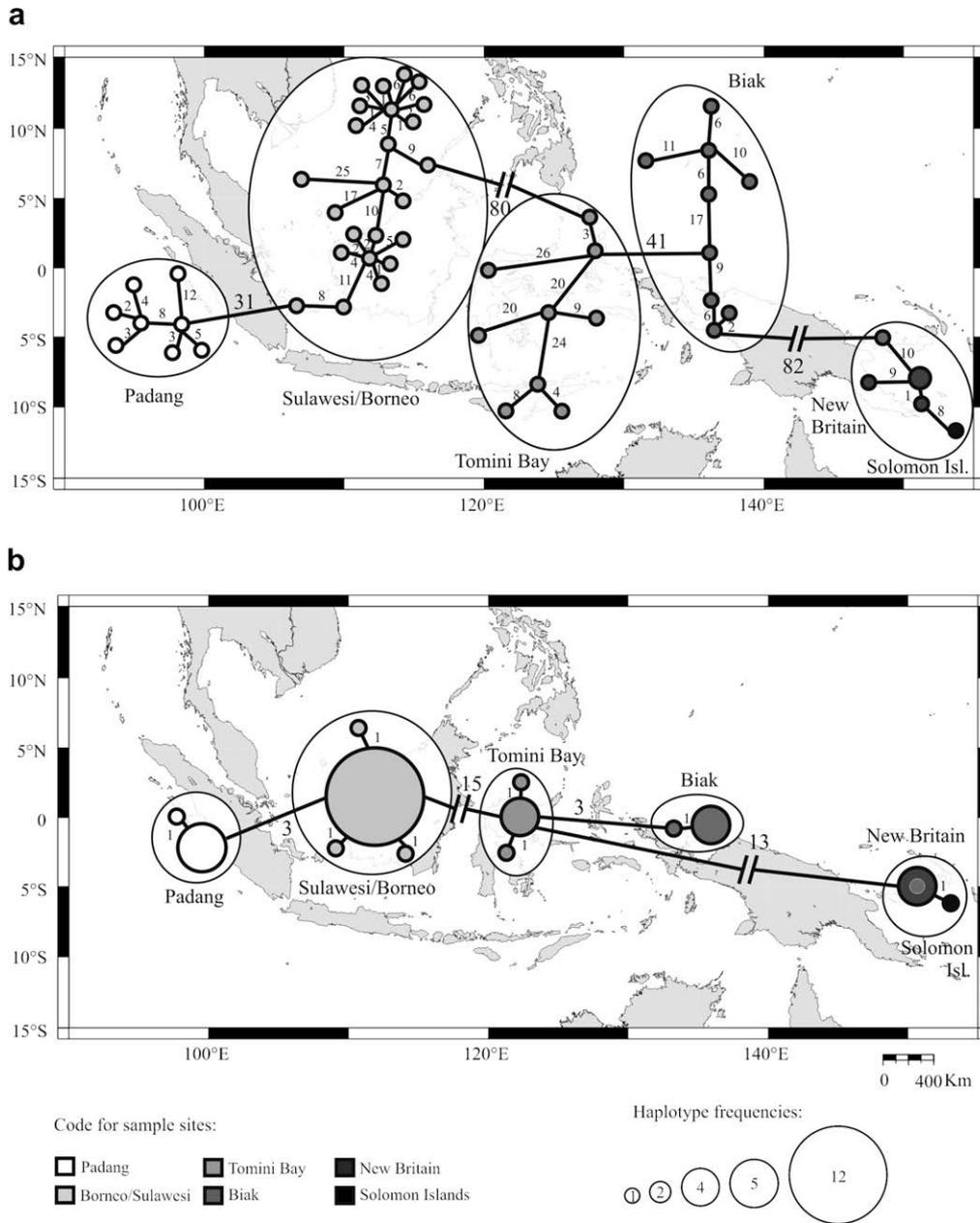


Fig. 4. Haplotype networks of (a) CR and (b) *cyt b* sequences of *Amphiprion ocellaris* and *A. percula* from different locations in the Indo-Malay Archipelago. Numbers indicate mutational steps. The size of the circles indicate the haplotype frequencies.

observed in that group of fishes questioned this. Additionally, such studies can provide insides into divergence and speciation processes, as well as geographic locations of these events.

Different mutation rates for the mitochondrial control region in fish have been estimated, ranging from 2–6.41% per million years (Faber and Stepien, 1998; Campton et al., 2000; Donaldson and Wilson, 1999; Waters et al., 2001; Messmer et al., 2005). In this study, we used a mutation rate of 6.41%, because it was obtained for the same CR fragment. The other studies used different fragments or the complete CR, resulting in lower mutation rates. The mutation rate for the *cyt b* region varies widely in literature, but for teleost fish species rates between 1.0% and 2.8% per million years were found (Ortí et al., 1994; Martin and Bermingham, 1998; Perdices et al., 2002; Chenoweth et al., 2002; Banford et al., 2004; Casey et al., 2004). CR divergence times were within the ranges calculated for *cyt b*. It shows a consistency for the divergence times calculations between the results of the two markers.

4.1. *Amphiprion akallopisos/perideraion/sandaracinos* species complex

The definition by morphological characters and colouration pattern of *A. akallopisos*, *A. perideraion*, and *A. sandaracinos* was supported by the molecular phylogenetic analysis in this study. Additionally, the comparably low sequence divergences in this group confirmed the close relatedness of these species.

Based on the colouration pattern, it was expected that *A. akallopisos* and *A. sandaracinos* are sister species, but *A. sandaracinos* and *A. perideraion* were observed to be more closely related in the phylogenetic trees as well as in the haplotype networks. This confirms the findings of another study on the molecular phylogeny of anemonefishes (Santini and Polacco, 2006). The authors of this study assume that the main characters of the *A. akallopisos/perideraion/sandaracinos* species complex (orange-pinkish colouration with a white band on the back, a slender body with a rounded caudal fin) evolved before the radiation into different species.

Table 2

Sequence divergences (calculated with Paup*) of (a) CR (considering the GTR model) and (b) *cyt b* sequences (considering the TVM model) between (below diagonal), within taxa (diagonal), and genetic distances between taxa after accounting for diversities within them (bold, above diagonal), of the genus *Amphiprion* (Ao = pooled sequences of *A. ocellaris* from Borneo and Sulawesi, Aa = all sequences of *A. akallopisos*, Ap = all sequences of *A. perideraion*, As = all sequences of *A. sandaracinos*) from different locations in the Indo-Malay Archipelago (Abbr. see Table 1)

	Ao	AoPa	AperTI	AperBk	AperNB	AperSolo	Aa	Ap	As
(a)									
Ao	0,043	0,058	0,174	0,185	0,169	0,184	0,197	0,209	0,203
AoPa	0,094	0,030	0,173	0,183	0,166	0,184	0,210	0,225	0,215
ArTI	0,225	0,218	0,060	0,066	0,167	0,179	0,238	0,248	0,224
ArBk	0,228	0,219	0,118	0,043	0,169	0,180	0,240	0,248	0,224
ArNB	0,198	0,189	0,205	0,198	0,016	0,020	0,251	0,261	0,248
ArSolo	0,205	0,199	0,209	0,201	0,028	0,000	0,258	0,252	0,264
Aa	0,233	0,240	0,283	0,276	0,274	0,273	0,030	0,111	0,123
Ap	0,251	0,261	0,299	0,290	0,290	0,273	0,147	0,042	0,101
As	0,236	0,241	0,265	0,257	0,267	0,275	0,149	0,133	0,023
(b)									
Ao	0,001	0,009	0,050	0,062	0,048	0,052	0,272	0,588	0,325
AoPa	0,010	0,001	0,060	0,072	0,058	0,062	0,242	0,343	0,285
ArTI	0,052	0,062	0,004	0,010	0,038	0,042	0,212	0,232	0,214
ArBk	0,063	0,073	0,012	0,001	0,046	0,050	0,237	0,252	0,231
ArNB	0,048	0,058	0,040	0,046	0,000	0,003	0,217	0,225	0,210
ArSolo	0,052	0,062	0,044	0,050	0,003	0,000	0,225	0,231	0,216
Aa	0,274	0,244	0,216	0,239	0,219	0,227	0,004	0,031	0,048
Ap	0,588	0,343	0,234	0,252	0,225	0,231	0,033	0,000	0,015
As	0,325	0,285	0,216	0,231	0,210	0,216	0,050	0,015	0,000

Table 3

Divergence times between clades of the genus *Amphiprion* (Ao, pooled sequences of *A. ocellaris* from Borneo and Sulawesi; Aa, all sequences of *A. akallopisos*; Ap, all sequences of *A. perideraion*; As, all sequences of *A. sandaracinos*; further abbr. see Table 1)

Pairs of taxa	Control region	Cytochrome <i>b</i>	Geological era
Ao–AoPa	0.9	0.3–0.9	Pleistocene
Ao/AoPa–AperTI	2.7	1.8–6.0	Pliocene/end of Miocene
Ao/AoPa–AperBk	2.9	2.2–7.2	Pliocene/end of Miocene
Ao/AoPa–AperNB/AperSolo	2.6–2.9	1.7–6.2	Pliocene/end of Miocene
AperTI–AperBk	1.0	0.4–1.0	Pleistocene
AperTI–AperNB/AperSolo	2.6–2.8	1.4–4.2	Pliocene
AperBk–AperNB/AperSolo	2.6–2.8	1.6–5.0	Pliocene
Aa–Ap	1.7	1.1–3.1	Pleistocene/Pliocene
Aa–As	1.9	1.7–4.8	Pliocene
Ap–As	1.6	0.5–1.5	Pleistocene

[Miocene: 23–5.33 MYA, Pliocene: 5.33–1.8 MYA, Pleistocene: 1.8 MY–11500 YA (Grabstein & Ogg 2004)].

Molecular clock estimates on the divergence time among the *Phalerebus* species indicated that the separation of *A. akallopisos* and the other two species took place by the end of the Pliocene and beginning of the Pleistocene 1.1–4.8 million years ago (1.7–1.9 in CR). The species *A. perideraion* and *A. sandaracinos* diverged later, in the Pleistocene, 500,000 years to 1.5 million years ago (1.6 million years in CR).

By the end of the Pliocene and during the Pleistocene the sea level dropped, following the increasing glaciations, which created barriers for migration between the ocean basins. The ancestral population inhabiting the Indian Ocean gave rise to *A. akallopisos*, while the ancestral population on the Pacific side went through a sympatric speciation or allopatric speciation in separate ocean basins, such as the South China Sea, the Sulu Sea, and the Sulawesi Sea during sea level low stands (Voris, 2000). This gave rise to *A. perideraion* and *A. sandaracinos*. The present distribution patterns of the species support these findings. *A. akallopisos* is distributed mainly in the Indian Ocean, but also present in the Java Sea and probably re-colonised the Sunda Shelf through the Sunda Strait

after the glacial times. *A. sandaracinos* and *A. perideraion* both inhabit almost the whole Indo-Malay Archipelago. However, *A. sandaracinos* has a smaller distribution area (Fautin and Allen, 1994). Although, there is an overlap in host anemone acceptance (both fish species can occur in *Heteractis crispa*), it seems to be that they avoid competition by specialisation on different host anemone species (Elliott and Mariscal, 2001). *A. sandaracinos* represents a specialised and *A. perideraion* a rather generalised behaviour accepting two and four anemone species as hosts, respectively (Fautin and Allen, 1994). The speciation of the latter two species took place around 100,000 to 300,000 years later than the split between *A. akallopisos* and the ancestor of *A. sandaracinos* and *A. perideraion*.

The separation of the Pacific and Indian Ocean by sea level low stands triggered divergence and speciation also in other coral reef dwelling animals (McMillan and Palumbi, 1995; Williams, 2000; Kochzius et al. 2003).

4.2. *Amphiprion ocellaris/percula* species complex

All specimens from the Togian Islands (Tomini Bay) were clearly included in the *A. percula* clade. This is contrary to the distribution pattern proposed by Allen (1991) and Fautin and Allen, (1994), assuming the presence of *A. ocellaris* in Tomini Bay. Regarding Kuitner and Tonozuka (2005), both species occur in Tomini Bay, inhabiting different depth. *A. percula* is supposed to live in shallow water close to the coastline, whereas *A. ocellaris* inhabits the deeper areas. We sampled at different locations and different depths varying between one and 25 m, but could not confirm these findings, because the phylogenetic analysis clearly showed that all specimens belong to *A. percula*.

The specimens of *A. percula* collected at three different locations formed distinct geographic subclades, supported, by high bootstrap values. The haplotype networks both indicated a strong separation of the samples from New Britain (New Guinea) including the sample from the Solomon Islands, reaching back around 1.9–10 million years. In contrast to the CR analysis, the clade from New Britain and Solomon Islands is not connected to Biak in the *cyt b* network, but to Togian Island. Obviously, the New Britain lineage is genetically very distinct and its connection to the other clades is not well resolved. This was as well shown in the inconsistencies and not resolved nodes in the phylogenetic trees. Possibly, there is a tendency of mutation saturation present for the strong diverged clades for the third codon positions of the *cyt b* gene (Farias et al. 2001) and/or in the highly variable CR sequences, as well indicated by the comparably small Ts/Tv ratios in both markers (3.9 in labroid fishes, Bernardi and Bucciarelli 1999, 3 in lionfishes, Kochzius et al. 2002, 3.93 in cichlid fishes, Farias et al., 2001). The divergence times between the Biak/Tomini Bay and the New Britain clade are similar to the separation between *A. ocellaris* and *A. percula*. Nelson et al., (2000) found the similar divergence time of 1.9–7.5 million years between *A. ocellaris* and *A. percula* using the same *cyt b* fragment as in the present study. Specimens of *A. percula* from Biak and from Tomini Bay are separated by a number of mutational steps similar to that one separating the clade of *A. ocellaris* from Padang (Indian Ocean) to their conspecifics. In the *cyt b* fragment the mutational steps of the abovementioned clades were even equal.

On the one hand, the genetic distances between some of the geographic groups of the *A. percula/ocellaris* complex were larger than between species in the subgenus *Phalerebus*. The clade of *A. percula* from New Britain and the Solomon Islands shows a divergence of 82 steps from its conspecifics, raising the question if they can be still regarded as one species.

On the other hand, the separation by 31 mutations between the *A. ocellaris* specimens from Padang (Indian Ocean) and the other *A. ocellaris* clade, was lower than between the *Phalerebus* species. The

same could be observed between *A. percula* samples from Tomini Bay and from Biak.

Both, the genetic distance of CR (6.6%) and *cyt b* (1%) between *A. percula* from Biak and *A. percula* from the Togian Islands could correspond to strongly diverged populations, whereas sequence divergences of 16.7–18% (CR) and 3.8–5% (*cyt b*) present between *A. percula* from New Britain/Solomon Islands and the other *A. percula* clades are rather at species level. This view is supported by comparison with the lower genetic distances among the clearly distinct species in the subgenus *Phalerebus* (Table 2). The analysis revealed a sharp genetic break, although the strong South Equatorial Current (SEC) along the northern coast of New Guinea could indicate high gene flow. Although, part of the SEC joins the North Equatorial Countercurrent (NECC) northwest of New Guinea, a noticeable part of it branches off before it reaches Biak (Fig. 1b), which might prevent a continuous mixing (Godfrey, 1996). Different current regimes might have also restricted connectivity along the northeastern coast of New Guinea during times of low sea level stands in the glacials. The New Britain group might be rather connected to populations in the Great Barrier Reef to the South than to the other clades in the Northwest. A strong connectivity of New Britain and the Great Barrier Reef could not be found in another coral reef fish species (*Pseudochromis fuscus*, Messmer et al., 2005), however, the close relatedness to the neighbouring Solomon Islands, although only represented by one sequence, could indicate connectivity in southeast direction.

The fact that two mitochondrial lineages were found in Biak (one specimen corresponding to the New Britain Clade) indicates that occasional migration takes place resulting either in hybridisation and introgression, or a coexistence of these diverged lineages without genetic exchange. To resolve the question of hybridisation and the presence of a cryptic species, it would be necessary to use a nuclear genetic marker for further analyses.

The basic differentiation pattern of the *A. ocellaris/A. percula* complex, with the incertitude of the genetic isolation of the New Britain Clade, can be traced back to the separation of ocean basins during late Pliocene sea level low stands (Van Andel, 1994). Due to climate oscillations in the late Pliocene, the sea level dropped by around 40–70 m during glacial periods (Van Andel, 1994). This caused intermediate frequent disturbances that are likely to increase the probability of divergence and speciation (Roxburgh et al., 2003; Shea et al., 2004). Even the divergence of the New Britain clade might have been influenced by possible different current patterns during the Pliocene.

In the following Pleistocene, glaciations increased and the sea level dropped up to 120 m (Voris, 2000). During sea level low stands the Sunda Shelf was exposed and the population of *A. ocellaris* from Padang (Indian Ocean) was separated from the ones on the other side of the Sunda shelf. This separation, most probably initiated also the speciation in the *Phalerebus*-species complex (see above). Such a divergence of marine animals in the Indo-Malay Archipelago was also shown for other species, e.g. the mantis shrimp *Haptosquilla pulchella* (Barber et al., 2000, 2002), starfish species (Benzie, 1999; Williams, 2000) and the giant clam *Tridacna crocea* (Nuryanto and Kochzius, 2006). In butterflyfishes, different species assemblages were found in the Indian and Pacific Ocean with genetic variations of around 2% between the species of the two different oceans. Comparably to the anemonefishes, butterflyfishes are also coral reef associated organisms, which seem to have undergone diversifications during sea level low stands in that area (McMillan and Palumbi, 1995). Additionally, the central part of the Indo-Malay Archipelago was separated from the Pacific coast of New Guinea, leading to a genetic break between populations. Similar genetic differentiation was revealed in other marine organisms, such as *Haptosquilla pulchella* (Barber et al., 2002) and *Tridacna crocea* (Nuryanto and Kochzius, 2006). The population of

A. percula from Tomini Bay is very isolated and obviously only connected to other populations of this species in the East by currents going through the Moluccan Islands. This is especially notably, because at the neighbouring coastal areas of Sulawesi (e.g. Manado in the North, Kendari in the South) populations of *A. ocellaris* were found (unpublished data). The population of mantis shrimps in Tomini Bay was shown to be also very divergent (Barber et al., 2002), supporting the distinctiveness of that area (Wallace et al., 2002). A study on scad mackerel found a connection between populations in Tomini Bay and the Moluccan Islands (Arnaud et al., 1999). It would be necessary to analyse populations of that area to proof this zone as a connection between Tomini Bay and the Pacific coast of New Guinea.

The results of the present study suggest that there are possibly three species present within the *A. ocellaris/percula* species complex, which is not concordant to the current taxonomy based on morphological characters and colour pattern. Within the *A. akallopisos perideraion/sandaracinos* species complex, genetic clades follow species boundaries defined by morphology and colour pattern. The latter results gave a helpful scaling within the genus *Amphiprion* for the interpretation of the *A. ocellaris/A. percula* species complex analyses. The species of the two subgenera (*Actinicola* and *Phalerebus*) are biologically very similar to each other, they are quite specialized in host selection (2–4 host anemones), share the same food sources, and behave and reproduce in a similar way. The species boundaries patterns revealed in this study were contrasting, though. Within the *A. ocellaris/A. percula* complex, there was a strong subdivision corresponding to geographical locations, but not following the species definitions and distributions assumed before. Whereas within the *Phalerebus* subgenus, species boundaries were concordant to their prior species classification but did not reveal geographical clades.

This study shows, that important factors, for divergences of lineages until speciation of the genus *Amphiprion* in the “coral triangle”, were sea level fluctuations since the Miocene (Van Andel, 1994). This supports on the one hand the “centre of origin” theory for explaining the high biodiversity present in that region (Briggs, 2000, 2005). On the other hand it shows, that through the separation between the ocean basins, divergences took place at each side of the barriers, which obviously also contributed to species richness in this marine biodiversity hotspot.

Acknowledgments

We thank the institutions and individuals that have made our study possible: German Federal Ministry of Education and Research (BMBF, Grant No. 03F0390B), which funded the project “Molecular Genetics as a Tool for the Management of Marine Ornamentals in Sulawesi (Indonesia)” in the framework of the SPICE project (Science for the Protection of Indonesian Coastal Marine Ecosystems); colleagues from the University of Bremen, especially A. Nuryanto for help during field work, and D. Blohm; S. Planes from the University of Perpignan (France) for providing samples from New Britain; U. John from the Alfred-Wegener-Institute Bremerhaven (Germany) for sequencing support; colleagues from Universitas Hasanuddin (Makassar, Indonesia) for logistical support during field work in Spermonde Archipelago, especially J. Jompa.

References

- Allen, G.R., 1991. Riffbarsche der Welt. Mergus Verlag, Melle, Germany. pp. 1–272.
- Allen, G.R., Werner, T.B., 2002. Coral reef fish assessment in the “coral triangle” of southeastern Asia. *Environ. Biol. Fishes* 65, 209–214.
- Alvarado, B.J.R., Baker, A.J., Mejuto, J., 1995. Mitochondrial DNA control region sequences indicate extensive mixing of swordfish (*Xiphias gladius*) populations in the Atlantic Ocean. *Can. J. Fish Aquat. Sci.* 52, 1720–1732.

- Arnaud, S., Bonhomme, F., Borsa, P., 1999. Mitochondrial DNA analysis of the genetic relationships among populations of scad mackerel (*Decapterus macarellus*, *D. macrosoma*, *D. russelli*) in South-East Asia. *Mar. Biol.* 135, 699–707.
- Banford, H.M., Bermingham, E., Collette, B.B., 2004. Molecular phylogenetics and biogeography of transisthmian and amphi-Atlantic needlefishes (Blennioidei: *Strongylura* and *Tylosurus*): perspectives on New World marine speciation. *Mol. Phylogenet. Evol.* 31, 833–851.
- Barber, P.H., Palumbi, S.R., Erdmann, M.V., Moosa, M.K., 2000. A marine Wallace's line? *Nature* 406, 692–693.
- Barber, P.H., Palumbi, S.R., Erdmann, M.V., Moosa, M.K., 2002. Sharp genetic breaks among populations of *Haptosquilla pulchella* (Stomatopoda) indicate limits to larval transport patterns, causes and consequences. *Mol. Ecol.* 11, 659–674.
- Benzie, J.A.H., 1999. Genetic structure of coral reef organisms: ghost of dispersal past. *Am. Zool.* 39, 131–145.
- Bernardi, G., Bucciarelli, G., 1999. Molecular phylogeny and speciation of the surperches (Embiotocidae, Perciformes). *Mol. Phylogenet. Evol.* 13, 77–81.
- Bernardi, G., Holbrook, S.J., Schmitt, R.J., 2001. Gene flow at three spatial scales in a coral reef fish, the three-spot dascyllus, *Dascyllus trimaculatus*. *Mar. Biol.* 138, 457–465.
- Bernardi, G., Holbrook, S.J., Schmitt, R.J., Crane, N.L., DeMartini, E., 2002. Species boundaries, populations and colour morphs in the coral reef three-spot damselfish (*Dascyllus trimaculatus*) species complex. *Proc. R Soc. Lond. B* 269, 599–605.
- Briggs, J.C., 2000. Centrifugal speciation and centres of origin. *J. Biogeogr.* 27, 1183–1188.
- Briggs, J.C., 2005. The marine East Indies: diversity and speciation. *J. Biogeogr.* 32, 1517–1522.
- Campton, D.E., Bass, A.L., Chapman, F.A., Bowen, B.W., 2000. Genetic distinction of pallid, shovelnose, and Alabama sturgeon: emerging species and the US Endangered Species Act. *Conserv. Genet.* 1, 17–32.
- Casey, S.P., Hall, H.J., Stanley, H.F., Vincent, A.C.J., 2004. The origin and evolution of seahorses (genus *Hippocampus*): a phylogenetic study using the Cytochrome *b* gene of mitochondrial DNA. *Mol. Phylogenet. Evol.* 30, 261–272.
- Chenoweth, S.F., Hughes, J.M., Connolly, R.C., 2002. Phylogeography of the pipefish, *Urocampus carinirostris*, suggests secondary intergradation of ancient lineages. *Mar. Biol.* 141, 541–547.
- Costagliola, D., Robertson, D.R., Guidetti, P., Stefanni, S., Wirtz, P., Heiser, J.B., Bernardi, G., 2004. Evolution of the coral reef fish *Thalassoma* spp. (Labridae). 2. Evolution of the eastern Atlantic species. *Mar. Biol.* 144, 377–383.
- Donaldson, K.A., Wilson, R.R., 1999. Amphi-panamic geminates of snook (Percoidae: Centropomidae) provide a calibration of the divergence rate in the mitochondrial DNA control region of fishes. *Mol. Phylogenet. Evol.* 13, 208–213.
- Elliott, J.K., Mariscal, R.N., 2001. Coexistence of nine anemonefish species: differential host and habitat utilization, size and recruitment. *Mar. Biol.* 138, 23–36.
- Faber, J.E., Stepien, C.A., 1998. Tandemly repeated sequences in the mitochondrial DNA control region and phylogeography of the pike-perches *Stizostedion*. *Mol. Phylogenet. Evol.* 10, 310–322.
- Farias, I.P., Orti, G., Sampaio, I., Schneider, H., Meyer, A., 2001. The cytochrome *b* gene as a phylogenetic marker: the limits of resolution for analysing relationships among cichlid fishes. *J. Mol. Evol.* 53, 89–103.
- Fautin, D.G., Allen, G.R., 1994. *Anemonenfische und ihre Wirte*. Tetra-Verlag, Melle. pp. 1–168.
- Frankham, R., Ballou, J.D., Briscoe, D.A., 2002. *Introduction to Conservation Genetics*. Cambridge Univ. Press, UK. pp. 1–617.
- Godfrey, J.S., 1996. The effect of the Indonesian throughflow on ocean circulation and heat exchange with the atmosphere: a review. *J. Geophys. Res.* 101, 12217–12237.
- Goldman, N., Anderson, J.P., Rodrigo, A.G., 2000. Likelihood-based tests of topologies in phylogenetics. *Syst. Biol.* 49, 652–670.
- Hall, T.A., 1999. BIOEDIT: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Series* 41, 95–98.
- Jokiel, P., Martinelli, F.J., 1992. The vortex model of coral reef biogeography. *J. Biogeogr.* 19, 449–458.
- Kishino, H., Hasegawa, M., 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in hominoidea. *J. Mol. Evol.* 29, 170–179.
- Kocher, T.D., Thomas, W.K., Meyer, A., Edwards, S.V., Pääbo, S.F., Villablanca, F.X., Wilson, A.C., 1989. Dynamics of mtDNA evolution in animals: amplification and sequencing with conserved primers. *Proc. Natl. Acad. Sci. USA* 86, 6196–6200.
- Kochzius, M., Soller, R., Khalaf, M.A., Blohm, D., 2003. Molecular phylogeny of the lionfish genera *Dendrochirus* and *Pterois* (Scorpaenidae, Pteroidae) based on mitochondrial DNA sequences. *Mol. Phylogenet. Evol.* 28, 396–403.
- Kuiter, H.K., Tonozuka, T., 2004. *Indonesian Reef Fishes. Part 2. PT Dive and Dives*, Bali, Indonesia. pp. 376–377.
- Lee, W.J., Howell, W.H., Kocher, T.D., 1995. Structure and evolution of teleost mitochondrial control regions. *J. Mol. Evol.* 41, 54–66.
- Martin, A.P., Bermingham, E., 1998. Systematics and evolution of lower Central American Cichlids inferred from analysis of Cytochrome *b* gene sequences. *Mol. Phylogenet. Evol.* 9, 192–203.
- McMillan, W.O., Palumbi, S.R., 1995. Concordant evolutionary patterns among Indo-West Pacific butterflyfishes. *Proc. R Soc. Lond.* 260, 229–236.
- Messmer, V., VanHerwerden, L., Munday, P.L., Jones, G.P., 2005. Phylogeography of colour polymorphism in the coral reef fish *Pseudochromis fuscus*, from Papua New Guinea and the Great Barrier Reef. *Coral Reefs* 24, 392–402.
- Nelson, J.S., Hoddell, R.J., Chou, L.M., Chan, W.K., Phang, V.P.E., 2000. Phylogeographic structure of false clownfish, *Amphiprion ocellaris*, explained by sea level changes on the Sunda shelf. *Mar. Biol.* 137, 727–736.
- Nuryanto, A., Kochzius, M., 2006. Limited gene flow in the giant clam *Tridacna crocea* across the Indo-Malay Archipelago. In: Abstracts of the 6th European meeting of the International Society for Reef Studies (ISRS), Bremen, Germany, pp. 19–22.
- Orti, G., Bell, M.A., Reimchen, T.E., Meyer, A., 1994. Global Survey of mitochondrial DNA sequences in the Threespine Stickleback: evidence for recent migrations. *Evolution* 48, 608–622.
- Paraskeris, D., Magiorkinis, E., Magiorkinis, G., Kiosses, V.G., Lenny, P., Vandamme, A.-M., Rambaut, A., Hatzakis, A., 2004. Phylogenetic reconstructions of a known HIV-1 CRF04_cpx transmission network using maximum likelihood and Bayesian methods. *J. Mol. Evol.* 59, 709–717.
- Perdices, A., Bermingham, E., Montilla, A., Doadrio, I., 2002. Evolutionary history of the genus *Rhombia* (Teleostei: Pimelododae) in Central America. *Mol. Phylogenet. Evol.* 25, 172–189.
- Pond, S.L.K., 2007. Evolutionary model selection with a genetic algorithm: a case study using stem RNA. *Mol. Biol. Evol.* 24, 159–170.
- Posada, D., Cranall, K., 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14, 817–818.
- Roxburgh, S.H., Shea, K., Wilson, J.B., 2004. The intermediate disturbance hypothesis: patch dynamics and mechanisms of species coexistence. *Ecology* 85, 359–371.
- Santini, S., Polacco, G., 2006. Finding Nemo: molecular phylogeny and evolution of the unusual life style of anemonefish. *Gene* 385, 19–27.
- Schneider, S., Roessli, D., Excoffier, L., 2000. ARLEQUIN version 2.000: A Software Program for Population Genetic Data Analysis. Genetic and Biometry Laboratory, Department of anthropology, University of Geneva, Switzerland.
- Shea, K., Roxburgh, S.H., Rauschert, E.S.J., 2004. Moving from pattern to process: coexistence mechanisms under intermediate disturbance regimes. *Ecol. Lett.* 7, 491–508.
- Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* 16, 1114–1116.
- Swofford, D.L., 1998. PAUP. Phylogenetic Analysis using Parsimony (and other Methods). Version 4. Sinauer, Sunderland.
- Tavaré, S., 1986. Some probabilistic and statistical problems on the analysis of DNA sequences. *Lect. Math. Live Sci.* 17, 57–86.
- Thompson, J.G., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignments through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673–4680.
- Van Andel, T.H., 1994. *New views on an Old Planet: A History of Global Change*. Cambridge Univ. Press, UK. pp. 1–439.
- Voris, H.K., 2000. Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and durations. *J. Biogeogr.* 27, 1153–1167.
- Wallace, C.C., Paulay, G., Hoeksema, B.W., 2002. Nature and origins of unique high diversity reef faunas in the Bay of Tomini, Central Sulawesi: The ultimate "centre of diversity"? *Proc. International Coral Reef Symp.*, Bali, Indonesia, pp. 185–192.
- Waters, J.M., Craw, D., Youngson, J.H., Wallis, G.P., 2001. Genes meet geology: fish phylogeographic pattern reflects ancient, rather than modern, drainage connections. *Evolution* 55, 1844–1851.
- Williams, S.T., 2000. Species boundaries in the starfish genus *Linckia*. *Mar. Biol.* 136, 137–148.
- Woodland, D.J., 1983. Zoogeography of the Siganidae (Pisces): an interpretation of distribution and richness patterns. *Bull. Mar. Sci.* 33, 713–717.