

Understanding Evolutionary Trees

T. Ryan Gregory

Published online: 12 February 2008
© Springer Science + Business Media, LLC 2008

Abstract Charles Darwin sketched his first evolutionary tree in 1837, and trees have remained a central metaphor in evolutionary biology up to the present. Today, phylogenetics—the science of constructing and evaluating hypotheses about historical patterns of descent in the form of evolutionary trees—has become pervasive within and increasingly outside evolutionary biology. Fostering skills in “tree thinking” is therefore a critical component of biological education. Conversely, misconceptions about evolutionary trees can be very detrimental to one’s understanding of the patterns and processes that have occurred in the history of life. This paper provides a basic introduction to evolutionary trees, including some guidelines for how and how not to read them. Ten of the most common misconceptions about evolutionary trees and their implications for understanding evolution are addressed.

Keywords Branch · Clade · Common ancestor · Evolution · Node · Phylogeny · Sister taxa · Topology · Trend

Introduction: The Importance of Tree Thinking

In a flourish indicative of both his literary style and perceptive understanding of nature, Darwin (1859) offered the following arboreal metaphor to describe the diversification and extinction of species:

As buds give rise by growth to fresh buds, and these, if vigorous, branch out and overtop on all sides many a feebler branch, so by generation I believe it has been

with the great Tree of Life, which fills with its dead and broken branches the crust of the earth, and covers the surface with its ever-branching and beautiful ramifications.

Darwin clearly considered this Tree of Life as an important organizing principle in understanding the concept of “descent with modification” (what we now call evolution), having used a branching diagram of relatedness early in his exploration of the question (Fig. 1) and including a tree-like diagram as the only illustration in *On the Origin of Species* (Darwin 1859). Indeed, the depiction of historical relationships among living groups as a pattern of branching predates Darwin; Lamarck (1809), for example, used a similar type of illustration (see Gould 1999).

Today, evolutionary trees are the subject of detailed, rigorous analysis that seeks to reconstruct the patterns of branching that have led to the diversity of life as we know it (e.g., Cracraft and Donoghue 2004; Hodkinson and Parnell 2007; Lecointre and Le Guyader 2007; Maddison and Schultz 2007). An entire discipline known as phylogenetics (Gr. *phyle*, tribe + *genesis*, birth) has emerged, complete with professional societies, dedicated scientific journals, and a complex technical literature that can be impenetrable to many nonspecialists. The output of this profession has become prodigious: It has been suggested that phylogeneticists as a group publish an average of 15 new evolutionary trees per day (Rokas 2006). Little surprise, then, that it has been argued that evolutionary biology as a whole has undergone a shift to “tree thinking” (O’Hara 1988), akin to the earlier movement toward “population thinking” that helped to shape the Neo-Darwinian synthesis around the mid-twentieth century (Mayr and Provine 1980).

Whereas tree thinking has permeated much of professional evolutionary biology, it has yet to exert its full

T. R. Gregory (✉)
Department of Integrative Biology, University of Guelph,
Guelph, ON N1G 2W1, Canada
e-mail: rgregory@uoguelph.ca

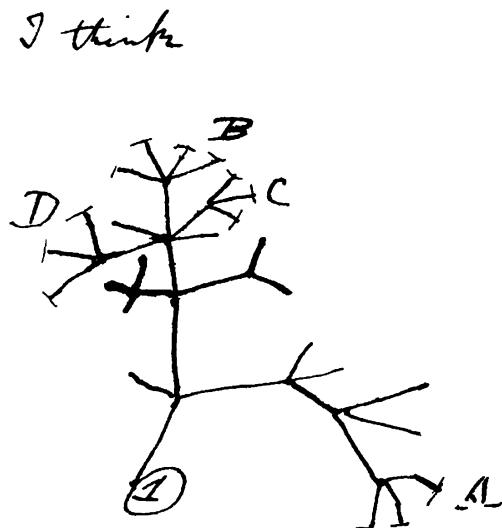


Fig. 1 The first evolutionary tree sketched by Darwin (1837) in one of his notebooks. It is also of note that the only illustration in *On the Origin of Species* (Darwin 1859) was an evolutionary tree. Other early evolutionists before and after Darwin, including Lamarck (1809), also drew branching diagrams to indicate relatedness (see Gould 1999)

influence among nonscientists. As Baum et al. (2005) recently pointed out, “Phylogenetic trees are the most direct representation of the principle of common ancestry—the very core of evolutionary theory—and thus they must find a more prominent place in the general public’s understanding of evolution.” In this regard, it is not so much the technical aspects of phylogenetic analysis¹ that are of interest but a more practical understanding of what evolutionary trees represent and, at least as important, what they do not represent. As Baum et al. (2005) continued,

Tree thinking does not necessarily entail knowing how phylogenies are inferred by practicing systematists. Anyone who has looked into phylogenetics from outside the field of evolutionary biology knows that it is complex and rapidly changing, replete with a dense statistical literature, impassioned philosophical debates, and an abundance of highly technical computer programs. Fortunately, one can interpret trees and use them for organizing knowledge of biodiversity without knowing the details of phylogenetic inference.

Unfortunately, it is becoming clear that many readers lack a sufficient level of phylogenetic literacy to properly interpret evolutionary patterns and processes. For example, a recent study of undergraduate students who had received at least introductory instruction in evolutionary science

revealed a range of common misconceptions about phylogenetic trees that represent “fundamental barriers to understanding how evolution operates” (Meir et al. 2007).² Early correction of these misconceptions would be of obvious benefit, and it has been suggested that the importance for biology students of learning how to interpret evolutionary trees is on par with that of geography students being taught how to read maps (O’Hara 1997). Given the growing significance of phylogenetic analyses in forensic, medical, and other applications (e.g., Vogel 1997; Rambaut et al. 2001; Mace et al. 2003; Mace and Holden 2005) in addition to their pervasive influence in evolutionary studies, this claim does not appear to be overstated.

This paper aims to provide a brief introduction to evolutionary trees and some basic details on how they should and should not be read and interpreted. This is followed by a discussion of ten of the most common misconceptions about evolutionary trees, many of which are held simultaneously and any of which can severely impede one’s understanding of evolution.

The Basics of Phylogenetic Literacy

What is an Evolutionary Tree?

In the most general terms, an evolutionary tree—also known as a *phylogeny*³—is a diagrammatic depiction of biological entities that are connected through common descent, such as species or higher-level taxonomic groupings. An overwhelming body of evidence supports the conclusion that every organism alive today and all those who have ever lived are members of a shared heritage that extends back to the origin of life some 3.8 billion years ago. One might therefore expect it to be possible, at least in principle, to reconstruct the Tree of Life, branch by branch and bough by bough, from the current diversity residing at the outermost twigs to a universally shared root. However, this proposition remains controversial—not because there is any scientific doubt about the historical relatedness of species (i.e., the fact of evolution; Gregory 2008) but because of the complex nature of evolutionary processes.

For a start, relatedness among species is a concept that depends on genetics as well as history, and there is ample evidence that even distantly diverged lineages have, at times, experienced significant gene sharing (a process known as lateral or horizontal gene transfer, in contrast to

¹ A discussion of phylogenetic methods is well beyond the scope of this article. Introductions to the technical aspects of phylogenetic analysis are provided by Hillis et al. (1996), Page and Holmes (1998), Nei and Kumar (2000), Felsenstein (2003), Salemi and Vandamme (2003), and Hall (2007).

² The quiz used by Meir et al. (2007) is available to instructors upon request by email (info@simbio.com). See also the “Tree Thinking Challenge” supplemental quiz by Baum et al. (2005).

³ For the purposes of this discussion and regardless of whether this will annoy some specialists, “evolutionary tree,” “phylogenetic tree,” and “phylogeny” are used interchangeably.

the more typical “vertical” transmission of genes from parent down to offspring). Some authors argue that this was sufficiently rampant in the earliest period of life’s history, and has been common enough throughout the more recent past, to create a “Web of Life” lacking any single root, rather than a strictly bifurcating tree in which branches, once split, remain separate forever (e.g., Doolittle 2000; Doolittle and Bapteste 2007). At the very least, it must be noted that in light of processes such as lateral gene transfer and gene duplication, the history of individual genes may not follow the same historical paths as those of the species in which they reside. In many cases, “gene trees” and “species trees” may not be equivalent, a fact that complicates (but does not preclude) the reconstruction of phylogenies using molecular information (e.g., Wolf et al. 2002; Rokas 2006).

These issues aside, living organisms do have a history, and this does include universal relatedness of one sort or another, be it analogous to a simple tree, a more complex web, or something else. Moreover, there is no fundamental principle that prevents the pattern of ancestry from differing both temporally and taxonomically: it is possible (but by no means confirmed) that a straightforward tree metaphor is inappropriate for, say, ancient (or perhaps even modern) bacteria but is accurate when applied to eukaryotes. In the case of the latter, at least, there may be a “true” phylogeny that accurately depicts the historical patterns of ancestry connecting eukaryote branches to their common root, but the shape of the tree is far from resolved (Baldauf 2003). In fact, except in rare instances where the pattern of evolutionary branching is created in the laboratory and observed directly as it occurs (e.g., Hillis et al. 1992; Sanson et al. 2002), it is impossible to know with certainty that any given phylogeny is historically accurate. As a result, any reconstructed phylogenetic tree is a hypothesis about relationships and patterns of branching and thus is subject to further testing and revision with the analysis of additional data. Fully resolved and uncontroversial phylogenies are rare, and as such, the generation, testing, and updating of phylogenetic hypotheses remain an active and sometimes hotly debated area of research.

Anatomy of a Phylogeny

The old cliché contends that an undue focus on individual trees can prevent one from appreciating the grandeur of a forest. The reverse applies with regard to evolutionary trees, in that their collective importance is obvious, but many people are unfamiliar with the basic features of individual phylogenies. Whether they illustrate relationships among a few species or thousands (e.g., Bininda-Emonds et al. 2007) or of larger groupings of species (genera, families,

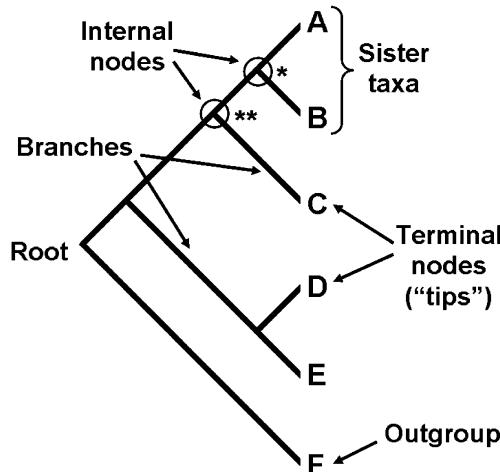


Fig. 2 The anatomy of a phylogeny. An evolutionary tree includes several components. At the *right* (in this case; see Figs. 4 and 5 for alternatives) are the terminal nodes or “tips” of the tree. These typically represent individual species or larger taxonomic groups, and all are contemporaries of one another (for example, all may be living at present, in which case *A* through *F* would all represent modern species). The terminal nodes are connected to one another through branches that join at “internal nodes.” Internal nodes represent inferred lineage splitting (speciation) events that give rise to descendant sister groups—in other words, they represent the common ancestors from which two or more related lineages are descended. In this figure, the *node marked with an asterisk* represents the most recent common ancestor of species *A* and *B*, and the *one marked with two asterisks* is the most recent common ancestor of species *A*, *B*, and *C*. The pattern of branching—known as the “topology” of the tree—indicates evolutionary relatedness. For example, species *A* and *B* share a recent common ancestor that was not shared by the other species and are therefore called “sister taxa.” Similarly, species *D* and *E* are sister taxa. Species *F* is the most distantly related of the sample of species and is known as the “outgroup.” Outgroup species are necessary to “root” an evolutionary tree—that is, to indicate the last common ancestor (i.e., the deepest internal node) shared by the entire group of species. The term “basal lineage” is sometimes used to describe the branch leading to the outgroup, but this is not recommended as it is often incorrectly taken to imply that it has undergone less change and is therefore more “primitive” or “ancestral” than the other lineages (Fig. 15; see also Crisp and Cook 2005)

phyla,⁴ etc.), all evolutionary trees provide the same basic information: a historical pattern of ancestry, divergence, and descent. They do so by depicting a series of branches that merge at points representing common ancestors, which themselves are connected through more distant ancestors.

The general anatomy of a phylogeny is summarized in Fig. 2. This tree shows the relationships among six species

⁴ Students (including many graduate students) sometimes exhibit confusion regarding the singular and plural forms of terms such as these. “Species” is both the singular and the plural (“specie” is not a biological term—it refers to coins). “Genus” is the singular, whereas “genera” is the plural. “Phylum” is the singular and “phyla” is the plural. Other terms of interest include “taxon” (singular) and “taxa” (plural) and the widely misused “data,” which is the plural form of “datum.” While on the topic, it bears mentioning that one human is still referred to as *Homo sapiens*, which means “wise man” and does not represent the plural of “Homo sapien.”

(labeled A through F), all of which are alive at present. Each species represents one “terminal node” or “tip” on the tree. In this tree, the far right where these terminal nodes are located represents the present day, with evolutionary time extending deeper into the past as one moves from right to left on the phylogeny. These living species are located at the ends of “branches” that join one another in the past at “internal nodes.” Each internal node is taken to represent an ancestor shared by two lineages, and each branch reflects the independent evolution of the lineages that has occurred after their divergence from a given common ancestor. Ultimately, all five species share a single common ancestor at the deepest internal node, also known as the “root” of the tree. Overall, the shape of the tree and therefore the pattern of branching that it hypothesizes are known as its “topology.”

By definition, the more common ancestors that two species share to the exclusion of other species, the more closely related they are. For example, in Fig. 2, from the terminal nodes to the root, species A and B share four common ancestors, species A and D share two common ancestors, and species F shares only one ancestor (the root itself) with any of the other five species. Species A and B are linked through a recent common ancestor that is not shared by any other taxa on the tree and are therefore known as “sister taxa.” The next closest relative of species A and B is species C, with whom they share an ancestor to the exclusion of species D, E, and F. Species D and E are sister taxa and are the next closest relatives of A+B+C. Species F, by contrast, is not linked to any of the other species beyond a single distant ancestor and is known as the “outgroup.” An outgroup is necessary to root a tree (unrooted trees also can be drawn, but these are less informative and are not covered here).

How to Read Evolutionary Trees

Phylogenies as Family Trees

Although the technical jargon of phylogenetics may be confusing on first pass, achieving a basic understanding of evolutionary trees need not be daunting. Notably, humans in all cultures are skilled at recognizing and understanding relatedness in other contexts, and many of these abilities apply equally well to phylogenies. There are some similarities between species phylogenies and human family pedigrees, and thinking of an evolutionary tree as a “family tree” can be helpful.⁵ This

⁵ Of course, one must not take this analogy too far. Human offspring have two parents, four grandparents, and so on, whereas each species in a phylogenetic tree is usually considered to have descended from a single parental species through a branching event (speciation). In this way, a more appropriate analogy would be to a pedigree showing only the males or only the females of a family or to the family tree of individual organisms that reproduce either through asexual fission or budding.

analogy is made explicit in Fig. 3a, which shows some easily recognized relationships among human family members. In this case, the reader is depicted as one of four terminal nodes, each of which represents a contemporary person. The reader’s closest cogenerational relative is a sibling, with whom he or she shares a common ancestor (the parent) to the exclusion of other individuals. The parent herself has a sibling, both of whom are descended from the reader’s grandparent; if this tree had been drawn one generation earlier, the grandparent would have been labeled as parent, and the aunt and parent would have been labeled as siblings. Descendants of the reader’s aunt (i.e., of the parent’s sister) are the reader’s cousins. The very same basic relationships apply to the species shown in Fig. 3b.

This simple comparison between phylogenies and family pedigrees highlights some other important points regarding the interpretation of evolutionary trees. First, contemporary entities (whether individual family members, species, or larger groupings) are related through common ancestors—they are not themselves ancestors of one another. Thus, the reader is not descended from a sibling; rather, both are descended from a shared parent. Likewise, the reader is not descended from a cousin, but they share a more distant common ancestor, namely their grandparent. Second, not only individual relatedness but also the relatedness of nested and increasingly inclusive groups is indicated on a tree. The reader, his or her sibling, and their shared parent represent an “immediate family,” whereas adding the cousins, the aunt, and the grandparent would also produce a coherent grouping that could be labeled more generally as a “family.” The analogous groups in phylogenetic terms, ones that include an ancestor and all of its descendants, are called “clades” (Fig. 3c). As O’Hara (1994) explained, “If you were to grab hold of the tree at any point, and cut immediately below your grip—below in the sense of toward the root—the chunk of the tree in your hand would by definition be a clade.” In other words, clades are branches that include all the twigs that have sprouted from them. Third, all members of an immediate family are equally related to individuals outside of their immediate family but with whom they share a more distant ancestor. For example, in Fig. 3a, both the reader and his or her sibling are equally related to both cousins. In like fashion, species Y and S in Fig. 3b are equally related to species C and to species K. Indeed, no matter how many descendants a parent and an aunt have, all siblings will be equally related to all of their first cousins. The same is true of species.

Types of Trees

In the most general terms, tree diagrams are known as “dendograms” (after the Greek for tree), whereas phylo-

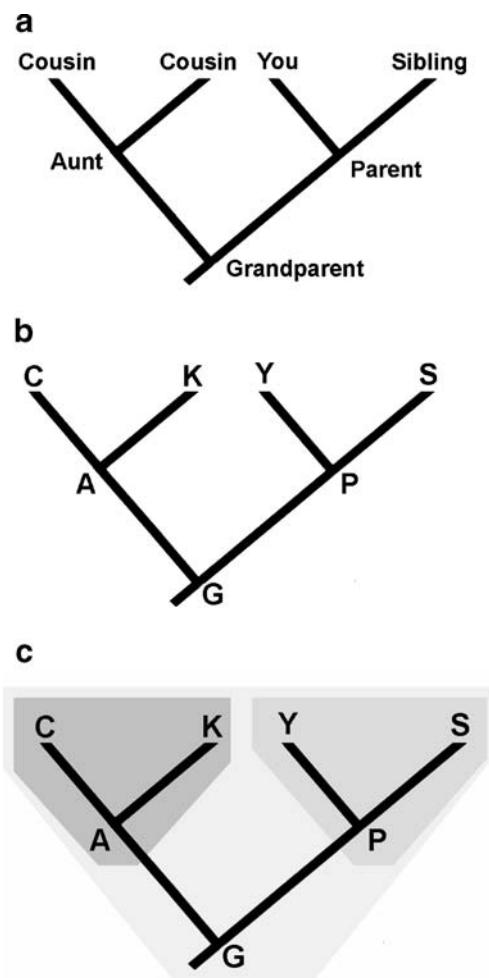


Fig. 3 Phylogenies indicate both relatedness and historical descent. As a rough analogy, phylogenies can be compared to pedigrees of human families. **a** shows a simple “phylogeny” of human relationships. You, the reader, are indicated as one terminal node, along with a sibling with whom you share a recent common ancestor (your parent). Also depicted are two of your cousins, each of whom is equally related to you and to your sibling and with whom you share a more distant common ancestor (your grandparent). You are not descended from your sibling or from your cousin (nor are they descended from you)—rather, your relatedness is determined by the pattern of descent and ancestry going back in time. **b** shows a phylogeny of the contemporary species *C*, *K*, *Y*, and *S*. *Y* is not descended from *S* (nor vice versa), but rather these are sister taxa that are both descended from their common ancestor, *P*. Similarly, *A* is the common ancestor of the sister taxa *C* and *K*. All of these species share a more distant common ancestor, *G*. The ancestors *A*, *P*, and *G* are extinct and may be known only from fossils or may be inferred to have existed using genetic or morphological data. **c** shows the groupings of related species and their shared ancestors, which are known as “clades.” The clades in this tree include the groupings $[C+K+A]$ and $[Y+S+P]$, which are nested within a larger clade $[C+K+A+Y+S+P+G]$. Both *A* and *P* are descended from their common ancestor *G*, and their descendant clades (i.e., $[C+K+A]$ and $[Y+S+P]$, respectively) are “sister groups” of one another (*A* and *P* were themselves sister taxa). Groups that do not include the most recent shared ancestor or that include only some of an ancestor’s descendants are not clades

genies that depict only branching order are known more specifically as “cladograms.” Cladograms can be drawn in many ways. Which one is used is largely insignificant: what matters is the order of branching. Figure 4 provides examples of six common ways of drawing cladograms, each of which is exactly equivalent in terms of topology. The overall orientation of cladograms is similarly irrelevant to their accuracy, although most conventional representations place the root either at the bottom or to the left (Fig. 5). The arrangement of the tips is unimportant as well, so long as the branching patterns are maintained. In this respect, phylogenies are like a baby’s mobile: every internal node can be rotated without any implications for the pattern in which the branches are connected (Fig. 6). As will be seen, grasping the concept of freely rotating nodes can, by itself, help to correct several major misinterpretations of phylogenetic relationships and thus of the evolutionary process that generates them.

As an example of the importance of rotating nodes, consider the relationships shown in Fig. 7. In this case, the tree is “unbalanced,” meaning that one of the branches

arising from a deep common ancestor contains many more species than the other. This can reflect either a real difference in “species richness” between two lineages, or it may simply result from sampling bias in which the full diversity of one lineage (or both) is not included in the tree (e.g., Harcourt-Brown et al. 2001). Panels a and b of this figure show a “ladderized” cladogram, in which the most diverse branches are consistently positioned to the right (panel a) or left (panel b) of each internal node. However, because each node can be rotated, the tree in panel c, although perhaps somewhat more cluttered, is equally accurate.

Branching pattern is the only piece of information that can be reliably gleaned from a cladogram, regardless of how it is presented. As with the order of the terminal nodes, the lengths of individual branches on a cladogram do not convey any information whatsoever (Fig. 8a). By contrast, trees known as “phylogenograms” present branch lengths as being proportional to some measure of divergence between species, for example based on comparisons of deoxyribonucleic acid sequences or morphological features. These trees typically include a scale bar to indicate the degree of divergence represented by a given length of branch (Fig. 8b). Topology remains important, but in these trees, the tips are not aligned at one end of the tree, although the species they represent are no less contemporary than in a cladogram. To both align the tips and present branch lengths as being proportional to divergence or time, one of several transformation algorithms can be used to “ultrametricize” a phylogram (Fig. 8c).

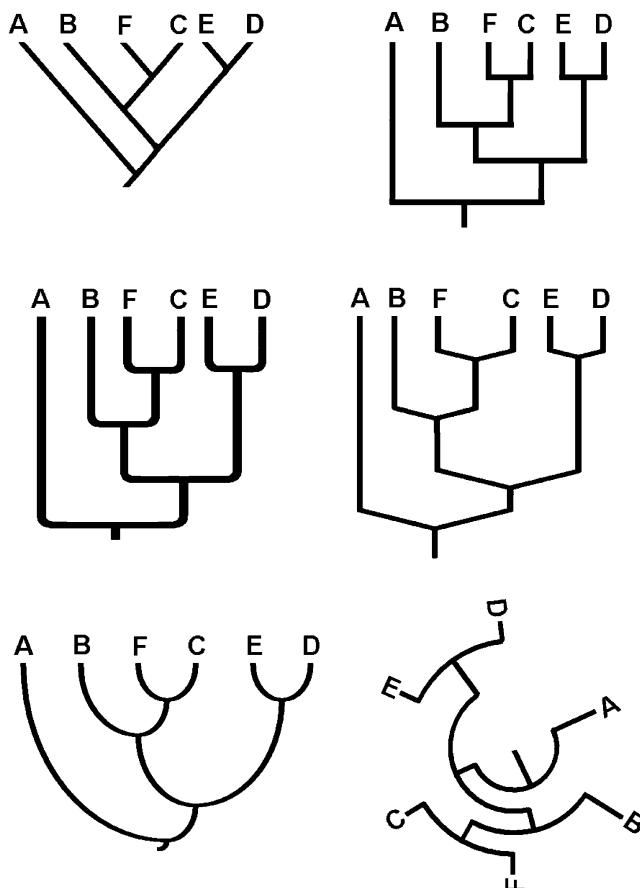


Fig. 4 Evolutionary trees can be presented in a variety of ways. This figure shows six common approaches to depicting evolutionary relationships using rooted trees. The two most common types, diagonal and rectangular, are shown at the top, but any of these may be encountered in the scientific literature or textbooks. In all six trees, the pattern of branching and relatedness is identical—for example, $F+C$ and $E+D$ are pairs of sister taxa, whereas A is the outgroup—and therefore all six are exactly equivalent. The reader is encouraged to confirm that the pattern of branching is the same in all six trees

Finally, some evolutionary trees are designed to provide information not only on branching order and time but also regarding features such as relative species diversity, geographical distribution, or ecological characteristics (Fig. 9). These are neither cladograms nor phylogenograms, and they may not provide any explicit information on the branching order of individual species within larger groups. Instead, they usually represent large-scale evolutionary trees that are drawn using fossil data and other information to provide a general overview of the history of a lineage.

How Not to Read Evolutionary Trees

Misunderstandings of evolutionary trees are pervasive among students, in the media, and among other non-specialists. Even more alarming, they also surface frequently

in the peer-reviewed scientific literature, often with significant implications for the conclusions drawn from comparative analyses (see Crisp and Cook 2005 for several examples). The following sections describe and seek to correct ten of the most commonly encountered misconceptions about evolutionary trees. Several of these are interrelated and therefore overlap to an extent, but each can be illustrated using distinct examples. Learning (and teaching students) to avoid these misunderstandings represents a key step toward the development of adequate tree thinking skills.

Misconception #1: Higher and Lower

Notions of a “Great Chain of Being” or *scala naturae* (scales of nature), in which living species (and, in some cases, nonliving matter and/or the divine) are ranked from lowest to highest and extend back at least as far as Aristotle. Although Darwin (1837) himself noted early on

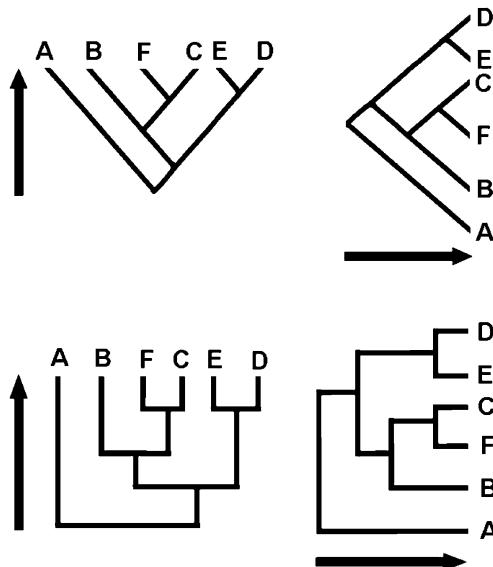


Fig. 5 Temporal directionality on a phylogeny. Regardless of the type of tree depicted (Fig. 4), a rooted tree can be read as indicating the earliest ancestor at the root, from which are descended the internal nodes and, more recently, the terminal nodes. That is to say, evolutionary trees indicate the passage of time beginning from the root (oldest) to the terminal nodes (youngest). Time cannot be read in any other direction on the tree (for example, across the tips), because all terminal nodes represent contemporary species (see Figs. 11 and 16). On all four trees shown here, the arrow indicates the direction of time from earliest ancestor (at the root) to modern species (at the tips). Trees are most commonly oriented to face up or right, but this is convention only, and downward or leftward trees would be equally accurate. Note that trees such as these do not imply specific amounts of time per branch, nor do they indicate when particular branching events occurred; they merely indicate the historical order of branchings within lineages. For example, this tree indicates that the split between the lineage leading to species D and E occurred sometime after the split of lineages from the common ancestor of $D+E+C+F+B$. By contrast, it does not indicate that the $D+E$ and $C+F$ splits occurred at the same time

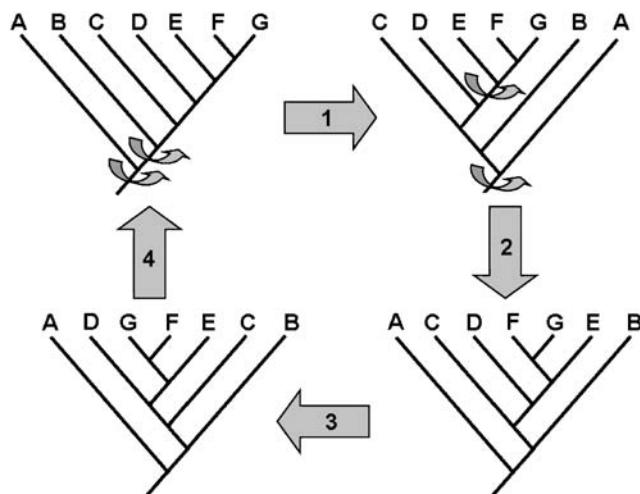


Fig. 6 All internal nodes can be rotated without changing the topology. Although they may look quite different, the four trees shown here are exactly equivalent to one another. This is because it is the order of branching events—the topology—that is relevant. Each internal node can be rotated without affecting the topology because this does not alter the groupings of species. For example, all four trees retain $F+G$ as sister taxa, with the next closest relative of these species being E , then D , and so on. The reader is encouraged to confirm that the topologies of these trees are identical. To convert from the tree at the top left into the one at the top right (arrow 1), one need only rotate the node joining B to $[C+D+E+F+G]$ and the node joining A to $[B+C+D+E+F+G]$, as indicated in the figure. To convert this second tree to the one at the bottom right (arrow 2), one must rotate the node joining $[E+F+G]$ and the node joining A to $[C+D+E+F+G+B]$. The reader is invited to identify the node rotations necessary to convert this third tree to the one at the bottom left (arrow 3) and thence back to the original tree at the top left (arrow 4)

that “It is absurd to talk of one animal being higher than another,” in many respects, his contribution merely shifted the explanation for the perceived rankings, replacing the scales of nature with an “evolutionary scale” or “evolutionary ladder” (Ruse 1996). Talk of “higher” and “lower” organisms, made in reference to contemporaneous species, persists in both public and professional scientific discourse. Not surprisingly, humans typically are (self-)designated as the “highest” organisms, with other living species ranked as higher or lower on the “evolutionary scale” according to how similar they are to this particular terminal node on the phylogeny of animals.

As many prominent authors have noted, there is no scientifically defensible basis on which to rank living species in this way, regardless of how interesting or unique some aspect of their biology may be to human observers (e.g., Dawkins 1992; Gould 1994, 1996). This error does not so much reflect a specific misunderstanding of phylogenetic diagrams *per se* but a failure to grasp the very concept of common descent. Therefore, the adjustment to be made in this case is from imagining evolution as a linear, progressive process that generates ladder-like ranks

to one of branching and diversification of which trees are the result (e.g., O’Hara 1992, 1997; Nee 2005).

Misconception #2: Main Line and Side Tracks

Although it is clearly a critical first step, recognizing evolution as tree-like does not in itself eliminate progres-

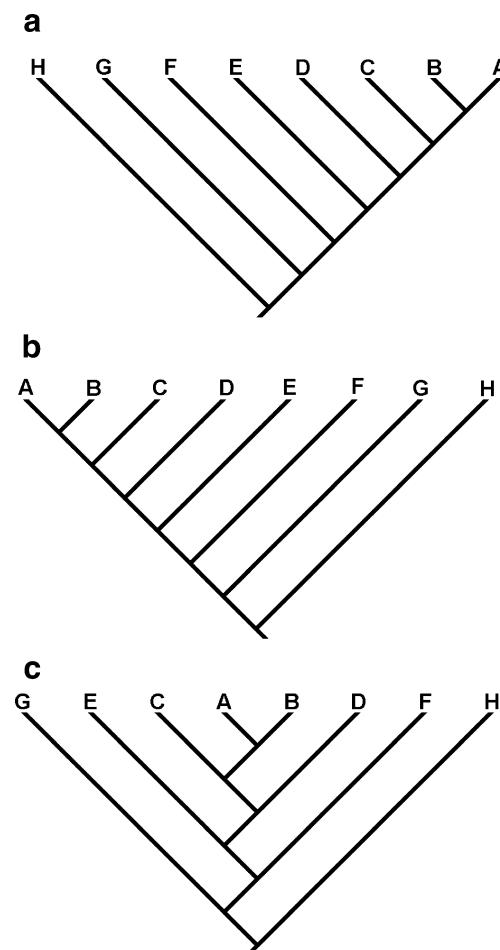


Fig. 7 Unbalanced trees can be ladderized or nonladderized without changing the topology. Many evolutionary trees are “unbalanced,” meaning that not all of the sister groups contain the same number of species. This can be due to real differences in diversity among groups or to incomplete sampling in which not all contemporary species are included in the tree. Figure 3b shows a balanced tree, but the trees depicted here are unbalanced because the major branches do not contain an equal number of species (i.e., one branch from the root contains only one species, H , whereas the other branch includes species G, F, E, D, C, B , and A). In **a** and **b**, the trees are “ladderized,” which means that they are presented with the most diverse sister group on the same side of every internal node. **c** shows the same tree, with an identical topology, in a nonladderized format. This was done simply by rotating several nodes (Fig. 6) so that more diverse groups do not always appear to the right (**a**) or left (**b**) of the internal nodes. Although ladderized trees appear less cluttered, they are no more accurate than nonladderized ones, and in fact they may cause readers to falsely interpret the information provided in the tree (Figs. 11 and 14; see also Crisp and Cook 2005)

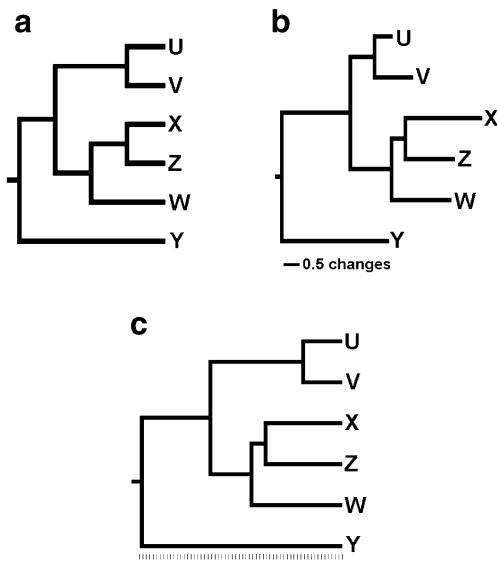


Fig. 8 Information other than topology requires different trees. The tree shown in **a** is known as a “cladogram” and is the same as the others in the paper in that the information contained within it is limited to branching order; the lengths of the branches in such trees do not convey any information. However, other types of trees can be drawn to indicate additional information, such as the degree of genetic or morphological divergence between species. **b** shows a special kind of phylogeny known as a “phylogram,” in which branch length is proportional to some measure of divergence. Phylogenograms typically include a scale bar to indicate how much change is reflected in the lengths of the branches. Total divergence between two species is determined on a phylogram by adding up the entire length of branches separating them: from one species to the common ancestor and then from the ancestor to the second species. In this tree, the lineage leading to species *U* has undergone less change than the lineage leading to species *V* since these lineages split from a common ancestor. Conversely, the lineage leading to *X* has undergone more change than the lineage leading to *W* since these lineages diverged from their most recent shared ancestor (which includes another split that led to *Z*—recall that neither *Z* nor *W* is an ancestor of *X*). It is important to note that, as with the cladogram in **a**, all species *U*–*Z* in **b** are contemporary species. To make this clearer, trees such as those in **b** are sometimes “ultrametricized” as in **c**, meaning that the terminal nodes are aligned with each other and the internal branch lengths are scaled to show the degree of divergence among sister groups rather than among individual species. Alternatively, a tree like that in **c** can be scaled to time (e.g., in millions of years before the present), for example if fossil or “molecular clock” data are available for calibrating the specific timing of branching events (Benton and Ayala 2003; see also Fig. 9)

sionist interpretations of life’s history. Even those who acknowledge the branching nature of evolutionary change may continue to interpret it as a progressive process in which a “main line” has led to a distinct endpoint (namely *Homo sapiens*). In this narrative, all other modern species are derivatives of “side tracks,” anomalous offshoots of the main line to humans that all went astray for one reason or another. Even Huxley (1880) fell prey to this line of thinking when he suggested that the teleost fishes “appear to me to be off the main line of evolution—to

represent, as it were, side tracks starting from certain points of that line.”

Figure 10a provides an illustration of how Huxley could reach such a false conclusion while still accepting the basic concept of tree-like branching. This represents an unbalanced, right-laddered tree with representatives of several vertebrate lineages, including the cartilaginous fishes, teleost fishes, amphibians, birds, and the mammalian lineage as represented by humans. For many, an intuitive interpretation of this tree is that humans represent an endpoint of a “main line,” with all other lineages branching off from this line at some time in the past.

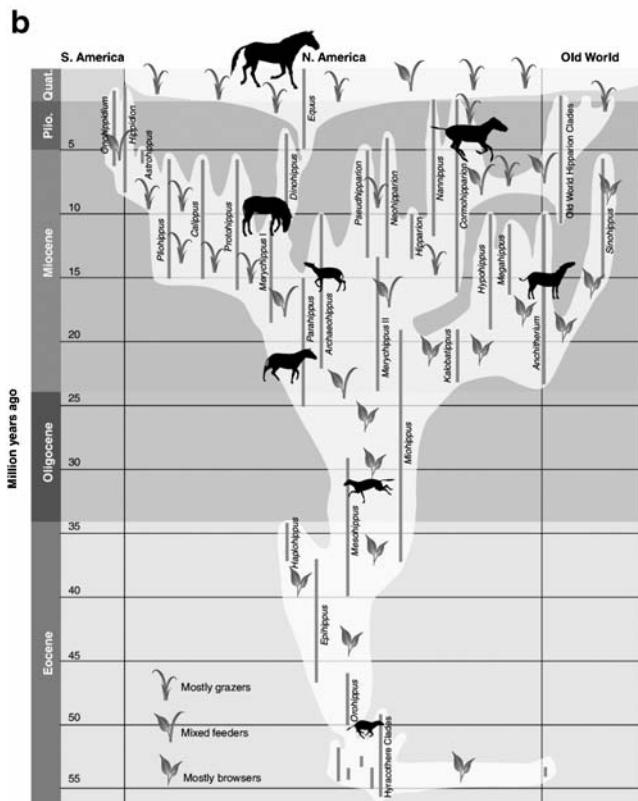
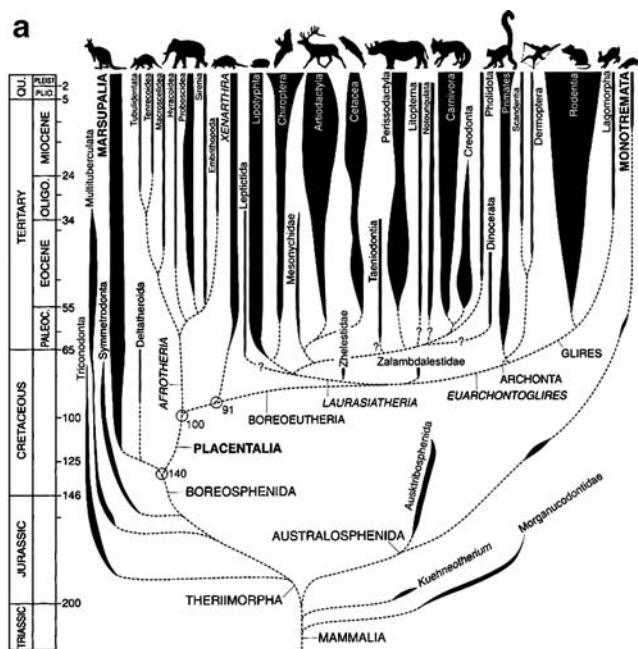
Two points can help to correct this misconception. First, recall from Fig. 3 that all siblings are equally related to their shared cousins. In this tree, all members of the clade that includes frogs, birds, and humans (tetrapods) are equally related to all members of the clade that includes goldfish and trout (teleost fishes). Second, a simple rotation of a few internal nodes or adding a better representation of some of the most diverse groups, as reflected in Fig. 10b, completely changes this perception without any effect on the accuracy of the tree. Few readers would interpret Fig. 10b as implying that a main line exists with perch as the endpoint and humans and other mammals as a minor, early branching offshoot, even given the fact that teleost fishes make up roughly 50% of all vertebrate diversity, whereas mammals represent about 10%. Still, the logic behind such an obviously silly reading of this tree is no more fallacious than of the intuitive “main line” interpretation of Fig. 10a. To quote O’Hara (1992),

When we come to realize that even among the vertebrates there are 50,000 different ‘vertebrate stories’, each one with a different ending and each one with a different narrative landscape; when we truly think in terms of the diverging tree, instead of the line; when we understand that it is absurd to talk of one animal being higher than another; only then will we see the full grandeur of the historical view of life.

As a matter of fact, it is most likely that evolutionary history will be misconstrued as representing a progressive “main line” when there is only one obvious endpoint available. In what he called “life’s little joke,” Gould (1991) noted that only *unsuccessful* lineages with very few living representatives are taken as endpoints of a supposed main line.

Misconception #3: Reading across the Tips

Referring to a cladogram similar to the one shown in Fig. 11a, Baum et al. (2005) asked readers to consider the following question on the basis of their reading of the tree: “Is the frog more closely related to the fish or the human?” The expectation, which has been reinforced by additional studies



(Meir et al. 2007), is that many people intuitively interpret the tree as indicating a closer relatedness between frog and fish than between frog and human. They do so because they incorrectly read meaning into the order of the terminal nodes, rather than assessing the pattern of branching that links these contemporary tips to one another historically.

Fig. 9 Some evolutionary trees include information about time and diversity. **a** shows an example of an evolutionary tree that includes not only information about topology but also time as given in the *axis at the left* and relative species diversity as indicated by the width of branches. Note that not all branches are of equal length because the tree is scaled to geological time periods and includes lineages that are extinct and therefore do not extend to the present. From Benton (2005), reproduced by permission of Blackwell. **b** shows an example of an evolutionary tree that provides information about geographical distribution and feeding ecology but that provides only a general indication of evolutionary relationships rather than explicit links between individual species (but see MacFadden 1992). From MacFadden (2005), illustration by P. Huey. Reprinted by permission of the American Association for the Advancement of Science

As a means of correcting this misinterpretation, one may take the time to identify the clades depicted in the tree (Baum et al. 2005). Humans, cats, and their common mammalian ancestor represent one clade, as do birds, lizards, and their common ancestor. These lineages together

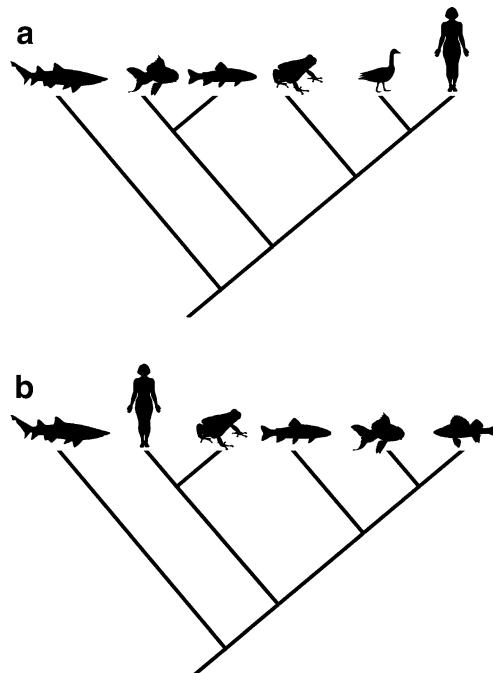


Fig. 10 There are no “main lines” or “side tracks” in evolution. Undoubtedly, many readers will consider the tree depicted in **a** to reflect a main line of evolutionary progress from a primitive ancestor to an “advanced” species like humans, with other groups such as cartilaginous or bony fishes appearing as side tracks off that line, despite the fact that roughly half of all vertebrate species are teleost fishes (and only 10% are mammals). Notably, the tree in **b** is equally valid and by the same false logic would have perch as the endpoint of an assumed main line and all terrestrial vertebrates, including humans, as an apparent side track. It is important that the positions of terminal nodes, all of which represent contemporary species, not be mistaken as having some significance, because they do not (see also Fig. 11). Note also that humans are more closely related to bony fishes than either is to sharks. Phylogenetically speaking, “fish” is an invalid category resulting from different rates of morphological change among lineages and does not reflect real relationships

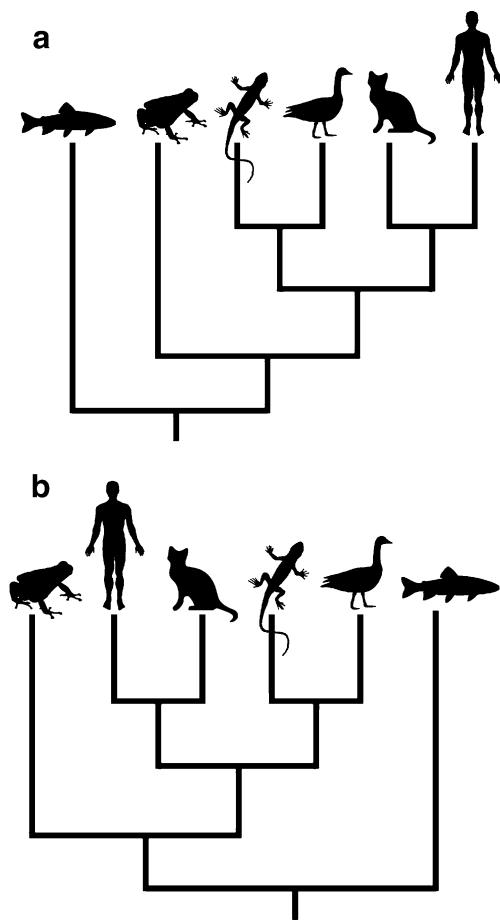


Fig. 11 The order of terminal nodes is meaningless. One of the most common misconceptions about evolutionary trees is that the order of the terminal nodes provides information about their relatedness. Only branching order (i.e., the sequence of internal nodes) provides this information; because all internal nodes can be rotated without affecting the topology of the tree, the order of the tips is meaningless. Nevertheless, there is a strong tendency for readers to take the tree in **a** as indicating that frogs are more closely related to fishes than humans are. They are not: both frogs and humans (and birds and lizards and cats) are equally closely related to fishes because as tetrapods they share a common ancestor to the exclusion of bony fishes. On the other hand, humans and cats are more closely related to each other than either is to any of the other species depicted because they share a recent common ancestor to the exclusion of the other species. The tree in **b** exhibits an identical topology to the one in **a** and is therefore equally valid. In this case, the same misinterpretation of “reading across the tips” would lead to the erroneous conclusion that birds are more closely related to fishes than cats are or that humans are more closely related to frogs than to lizards and birds. Because they share a common ancestor as amniotes, birds, cats, lizards, and humans are all equally related to frogs. It is good practice to rotate a few internal nodes mentally when first examining a tree to dispel misinterpretations based on reading the order of tips

with their shared ancestor represent a clade (amniotes) in which the first two clades are nested. Adding frogs and the ancestor linking them to the aforementioned species creates a yet larger clade (tetrapods). Adding fishes and the common ancestor of all species on this tree creates the final and largest clade (vertebrates). Because frogs can be

included in a clade with humans before fishes can—in other words, because frogs and humans share a common ancestor that is not shared with fishes—frogs are more closely related to humans than to fishes. Indeed, frogs and humans are exactly equally related to fishes through this common ancestor (recall that two cousins are equally related to a third, more distant relative).

A more rapid approach is to mentally rotate a few internal nodes with no effect on the topology of the tree, as shown in Fig. 11b. In this modified tree, humans are still sister to cats and birds are sister to lizards, frogs are then sister to amniotes, and fishes are the outgroup to the tetrapods. This second tree is identical in topology and is therefore equally accurate as the first tree. However, it

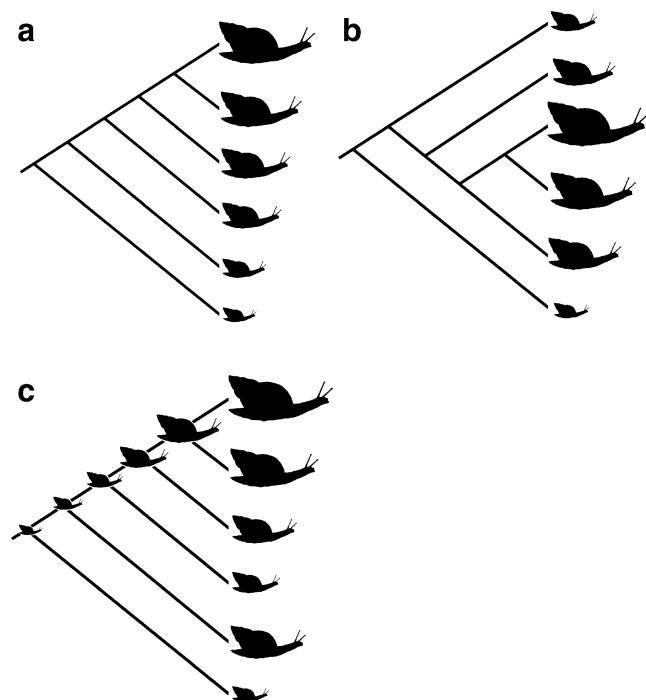


Fig. 12 Evolutionary trends cannot be identified by reading across the tips. In addition to resulting in incorrect interpretations of relatedness (Fig. 11), reading across the tips can engender a false impression of evolutionary trends. For example, many readers confronted with the tree in **a** might be tempted to infer an evolutionary trend toward increased body size in snail species over time (or, in Fig. 11a, an increase in complexity or intelligence over time). Unfortunately, misinterpretations such as this can be found even in the primary scientific literature. Once again, this can be corrected simply by rotating a few internal nodes, as has been done in **b**, in which the topology is the same but where the supposed trend is no longer apparent. **c** shows evidence of a real evolutionary trend toward increased body size. The important consideration is internal branching: In this case, there is information about ancestral states (e.g., from fossils), and it is evident that in every branching event, the two descendant species have been larger than their shared ancestor. Despite this being a clear evolutionary trend, there is no pattern evident across the terminal nodes. Thus, reading across the tips can create apparent trends where there are none and can mask real trends that are strongly supported by historical information

should be obvious that humans are not suddenly more closely related to frogs than to reptiles and birds.

Reading across the tips is not just problematic when interpreting relatedness. It can also lead readers and even authors of scientific publications to incorrectly intuit the existence of evolutionary trends where none exist or to overlook them where they do. For example, the phylogeny depicted in Fig. 12a may seem to show a trend toward increased body size in this snail clade. However, a simple rotation of a few internal nodes to produce an equivalent but nonladderized tree destroys this illusion (Fig. 12b). Conversely, although a reading across the tips in Fig. 12c would provide no indication of the fact, this tree indicates strong evidence for an evolutionary trend toward larger body size. In this case, information is available about the common ancestors, and it is clear that both descendants have been larger than their shared ancestor following every branching event. Only historical data or statistically rigorous inferences about history, and not a simple comparison of living species, can provide convincing support for claims of an evolutionary trend.

Misconception #4: Similarity versus Relatedness

The modern science of taxonomy is built upon the foundation laid by Carolus Linnaeus in the mid-eighteenth century. His system, which long predated the widespread scientific acceptance of common descent inspired by Darwin, categorized organisms on the basis of physical similarity. Notably, in the first edition of his *Systema Naturae* of 1735, whales were grouped with fishes—an oversight that he corrected in the tenth edition in 1758 by placing them with the other mammals. Today, the primary criterion for scientific classification is evolutionary relatedness, whereas differences in the degree of physical similarity across lineages are often a confounding variable. This can be so for two major reasons: First, as with whales and fishes, adaptation to similar environments can lead to a superficial convergence of physical appearance. Second, the rates of morphological change can vary considerably among lineages, with some remaining similar to a common ancestor and/or to more distantly related contemporary lineages and others becoming markedly different over the same time span (Baum et al. 2005).

By way of example, consider the phylogeny presented in Fig. 13. This tree shows one of the more prominent hypotheses regarding the relationships of major groups of nonmammalian tetrapods. Frogs are given as the outgroup in this tree, with turtles being the next most distantly related lineage to the others. Snakes are the sister group to lizards, and in fact, both modern lizards and snakes may be descended from a more ancestral lizard lineage. Most of these considerations are reasonably intuitive, but many

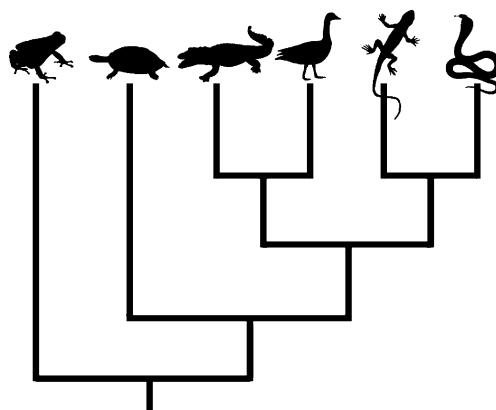
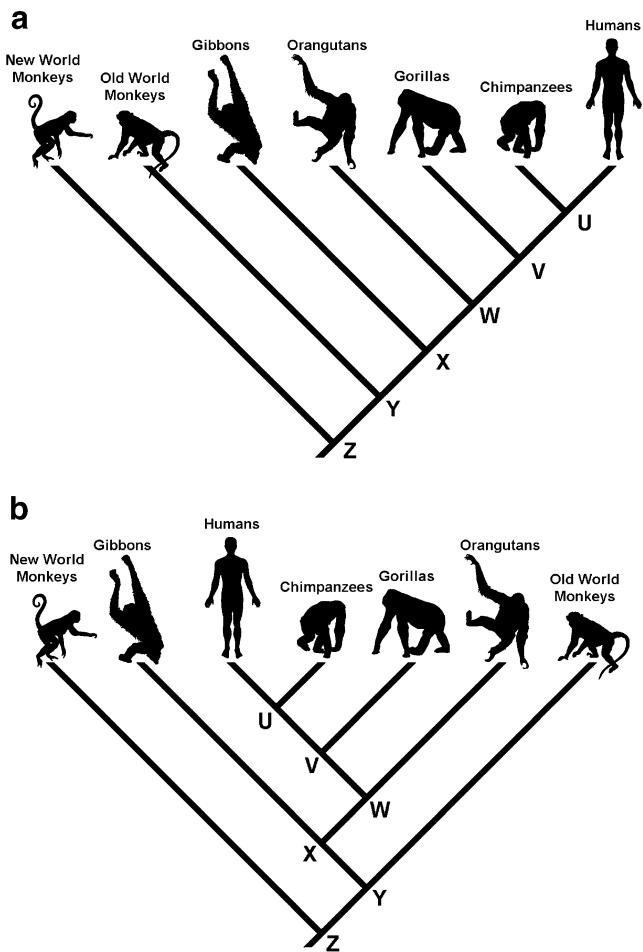


Fig. 13 Evolutionary relatedness and physical similarity are not necessarily linked. The rates at which physical features change can differ among lineages (Fig. 8), and superficially similar morphologies can evolve independently in more than one lineage. As a result, close relatives may look different from one another or distant relatives may look misleadingly similar. This tree presents evolutionary relationships among “reptiles,” with frogs as the outgroup. Although they look very different, birds and crocodiles are actually more closely related to each other than either is to any other group of reptiles. This particular phylogenetic hypothesis shows birds, crocodiles, lizards, and snakes to all be equally related to turtles (a detail that remains a subject of debate) and birds and all “reptiles” to be equally related to frogs (which is well accepted). The similarities between birds and mammals (e.g., four-chambered hearts, homeothermy) evolved independently in the two lineages well after their split from a distant reptilian ancestor. As with “fishes,” the category of “reptiles” is phylogenetically inconsistent

people find it surprising that phylogenetically birds are located within the “reptiles” and represent the sister group to crocodilians. Although physical similarities would seem to suggest otherwise, crocodiles are more closely related to birds than they are to lizards. The reason for this is that the bird lineage has experienced significant modification, whereas crocodilians have remained largely unchanged for tens of millions of years. It is important to note that birds and crocodiles represent each other’s closest *living* relatives but that birds are not descended from crocodiles nor vice versa—as taxonomic groups, crocodilians and birds both arose long after their respective lineages diverged from a common ancestor well over 200 million years ago. Birds are, in fact, descended from a lineage of theropod dinosaurs, making *Tyrannosaurus rex* far more similar to the last nonavian ancestor of modern birds than anything resembling a crocodile (see Prothero 2007).

Misconception #5: Sibling versus Ancestor

Mistaken assumptions that the ancestor of two modern groups must have been very similar to, or perhaps even was, one of the modern groups extend well beyond the case of crocodiles and birds. Any claim that two species represent each other’s closest living relative should not be construed as implying that one of the modern groups itself



is an ancestor of the other nor even that the common ancestor looked anything like either of the two groups. For example, the hypothesis that whales and hippopotamuses are sister groups (e.g., Boisserie et al. 2005) does not imply that the ancestor of whales was a hippo nor that it would even have been thought of as being similar to a hippo were it encountered when it was alive. Not surprisingly, the fossil record of whales, which is becoming increasingly extensive, shows that the early ancestors of whales (e.g., *Pakicetus*, *Ambulocetus*) bore no substantial resemblance to modern hippos at all (Thewissen and Bajpai 2001; Thewissen and Williams 2002).

Nowhere is this misconception more pronounced than in discussions of human evolution. One often hears it expressed in the rhetorical challenges offered by those who exhibit the poorest comprehension of evolutionary concepts: “If humans are descended from chimps,” so the question goes, “then why are there still chimps?” “If humans are descended from monkeys, then why has no one observed a monkey giving birth to a human baby?” The answer is simple because the premise is flawed: Humans are not descended from chimpanzees or monkeys, and no sane biologist suggests otherwise.

Fig. 14 Cousins are not ancestors, and humans are not descended from chimpanzees. **a** shows an evolutionary tree of anthropoid primates as it is often depicted, namely as an unbalanced, right-ladderized tree with humans at the extreme end. Viewed in this way, several of the most common fallacies in interpreting trees can arise: for example, that humans are the endpoint of a “main line,” that there is a trend toward “human-ness” from left to right, that the human lineage includes a monkey ancestor, or that there has been no branching in the lineages leading to the other modern species of primates. All are absolutely false. This becomes clearer if a few internal nodes are rotated, as in **b**, which is an equally accurate depiction of primate relationships. Humans and chimpanzees are more closely related to each other than either is to gorillas, orangutans, or any other living primates. However, note that “chimpanzees,” although depicted as one terminal node here, includes both the common chimpanzee (*Pan troglodytes*) and the bonobo (*Pan paniscus*), and if this tree was drawn as recently as 30,000 years ago, it would also include Neanderthals (*Homo neanderthalensis*) as a sister species to humans. Humans are not descended from chimpanzees any more than chimpanzees are descended from humans; rather, the two share a common ancestor (*U*) that lived some 5–7 million years ago and that was neither a human nor a chimpanzee. “Monkeys” are divided into Old World and New World lineages. Old World monkeys share a more recent ancestor with apes (*Y*) than either does with New World monkeys (*Z*), which means that apes (including humans) and Old World monkeys are equally related to New World monkeys. Monkeys are not ancestral to humans: The two lineages are related as distant cousins, not as grandparents and grandchildren

Figure 14a shows a ladderized phylogeny of the anthropoid primates. Humans and chimpanzees are sister taxa whose next (equally) close relatives are the gorillas, then the orangutans. Humans and chimpanzees share a common ancestor that lived around 5–7 million years ago. This ancestor was neither chimpanzee nor human, and as with whales, the increasingly detailed fossil record of the hominin lineage shows the extensive changes that have taken place since this divergence. Although the fossil record of chimpanzee ancestors is currently sparse, it can be presumed that a great deal of change characterized the evolutionary history of that branch as well.

The notion that other primates should have disappeared now that humans have evolved is based on a false understanding of species formation. Specifically, it assumes a process in which one species gradually transforms as a whole into another (called “anagenesis”). The reality of species diversification is that it most often proceeds by “cladogenesis,” the branching of new species from common ancestral populations. Chimps continue to exist because they are part of a separate branch that formed through cladogenesis when an ancestral population of a species, which was neither chimp nor human, split into independent lineages. Being confused about the coexistence of humans and chimpanzees is akin to being puzzled by the coexistence of Canada and Australia. Once again, rotating some internal nodes (Fig. 14b) can help to correct the misconception that other living primates are ancestors of humans or offshoots of a main line leading to humans or of

incorrectly assuming that the left- or bottom-most tip represents an ancestor to those at the terminal nodes of the other branches.

Misconception #6: Long Branch Implies no Change (or “Less Diverse Equals Basal Equals Ancestral”)

When viewing unbalanced trees such as those presented as Figs. 10a, 11a, 13, and 14a, there is a tendency among many people to misinterpret the long branch leading to the lone outgroup taxon in two ways. First, it is sometimes assumed that this species, although actually a contemporary of all others on the tree, is ancestral to the other lineages or at least is more similar to the root ancestor than any of the other species included in the tree (Crisp and Cook 2005). Second, this long branch is often taken to imply that no further branching has occurred along this lineage.

Figure 15 exposes the fallacy of both interpretations. In this case, humans are accurately included as the outgroup—the so-called basal lineage—to the echinoderms. It should go without saying that the branch leading from the common ancestor of echinoderms and vertebrates to modern mammals such as humans has not been devoid of additional divergence. In actuality, there have been hundreds of thousands, if not millions, of branching events along that lineage. The corollary of this observation, that humans do not resemble the ancestral echinoderm, should be even

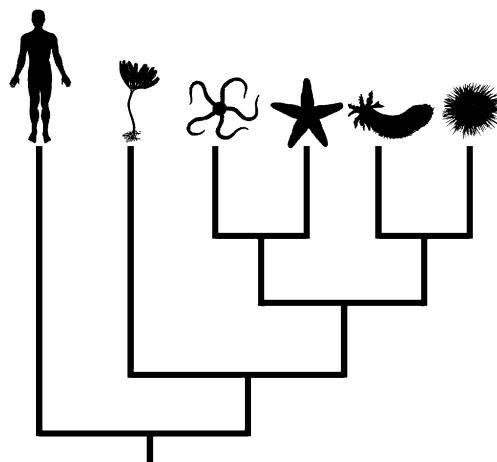


Fig. 15 A straight line does not mean that no change has occurred. This tree provides a simple illustration of the fact that the outgroup lineage cannot be assumed to be “basal,” “primitive,” or “ancestral” to the other species included on the tree. In this case, humans are accurately used as the outgroup to the echinoderms, which includes sea lillies, brittle stars, sea stars, sea cucumbers, and sea urchins. Of course, humans do not resemble the common ancestor of echinoderms, and there has been an enormous amount of branching among vertebrates since the very distant split of these two lineages from their common ancestor. It is most commonly argued that the sister group of the echinoderms is the hemichordates, both of which are more closely related to vertebrates than to any other group of animals, thus making the category “invertebrate” phylogenetically invalid

more obvious. Nonetheless, equally false interpretations of “basal” lineages are not uncommon, even in the scientific literature (Crisp and Cook 2005).

As with several of the other misconceptions discussed here, the problem of “basal equals primitive” is most likely to emerge when the tree under consideration is unbalanced and ladderized. It must be borne in mind that even if the unbalanced nature of a phylogeny reflects real differences in species diversity (which it often does not, as most trees include an incomplete sample of species), the relative diversity of major lineages can change over time, with one being the most diverse now and the other having been so in the past (Crisp and Cook 2005).

Finally, it must be pointed out that the relevant comparison is not between “primitive” (as in the sense of “poorly developed”) versus “advanced.” The only legitimate comparison is between “primitive” in the technical sense, meaning more like the last common ancestor (also called “ancestral”), versus “derived” (i.e., different from the ancestor). Any other interpretation runs the risk of invoking the fallacy of a progressive evolutionary scale. Moreover, as Crisp and Cook (2005) put it,

Once two lineages have separated, each evolves new characters independently of the other and, with time, each will show a mixture of plesiomorphic [inherited largely unchanged from the ancestor] and apomorphic [newly evolved and thus not possessed by the ancestor] character states. Therefore, extant species in both lineages resemble, to varying degrees, their common ancestor. Consequently, whereas character states can be relatively ancestral (plesiomorphic) or derived (apomorphic), these concepts are nonsensical when applied to whole organisms.

Misconception #7: Different Lineage Ages for Modern Species

Groups of species recognized as taxonomically distinct on the basis of particular characteristics (say, “flowering plants” or “beetles”) may have appeared at different times in the history of life and thus may be of different ages. However, the overall lineage leading to any modern species is of exactly the same age as that leading to any other modern species with whom an ancestor is shared (Fig. 16). This is a fundamental consequence of the principle of common descent, but there nevertheless can be a tendency to conflate taxon age with lineage age. For example, the group identified as teleost fishes is thought to be older—that is, to have appeared as a recognizable taxonomic group earlier—than mammals. Similarly, the first organisms that would be recognized as rainbow trout (*Oncorhynchus mykiss*) probably lived and died before the first individuals

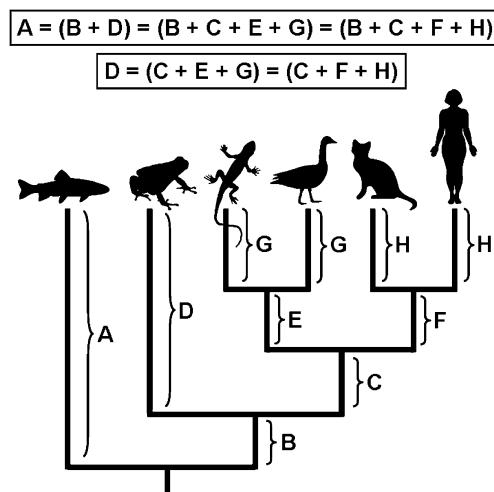


Fig. 16 The lineages leading to contemporary species have all been evolving for exactly the same amount of time. Rates of morphological change may vary among lineages, but the amount of time that separates two living lineages from their common ancestor does not. This figure shows the relationships among a sample of vertebrate lineages, all of which have been evolving for exactly the same amount of time, even if some lineages have undergone more change or more branching than others or if some taxonomically identifiable subsets of those lineages (e.g., teleost fishes) arose earlier than others (e.g., mammals). It is therefore a fallacy to describe one modern species as “more evolved” than another. Note, however, that this is a cladogram rather than an ultrametric tree, such that one cannot assume that any or all of *G*, *H*, *E*, *F*, *C*, and *B* are equal, only that the total amount of time between root and tip is the same along each of the lineages

that would have been classified as *Homo sapiens* were born. However, rainbow trout and humans are contemporary species, meaning that the lineages of which they are currently terminal nodes have been evolving for exactly the same amount of time since their divergence from a distant common ancestor. As a result, any notion that one of these lineages is “more evolved” or that one has had more time to accumulate differences is flawed.

Misconception #8: Backwards Time Axes

Among the common misconceptions identified by Meir et al. (2007) was the tendency for many students to misread the time axis on evolutionary trees. Many students interpreted the location of the terminal nodes as indicating time, for example by reading from left to right or from the leftmost tip to the root. In Fig. 17a, for example, many students read time as proceeding from birds (oldest) to the root *W* (youngest) or from birds (oldest) to kangaroos (youngest). Neither is correct, as time extends from the root to the terminal nodes, all of which are contemporary. This misinterpretation may have been exacerbated by the fact that the tree used in the quiz placed mammals—which many students assume to be the most “advanced” and hence most recent group—alone on the less diverse branch at the far right of an unbalanced, ladderized tree (unfortunately, a

tendency to place humans or some other preferred taxon at the top or right of every tree appears to be an unshakable habit among many phylogeneticists, although there is no objective reason for doing so). As indicated in Fig. 5, even on cladograms, in which the lengths of the branches are not

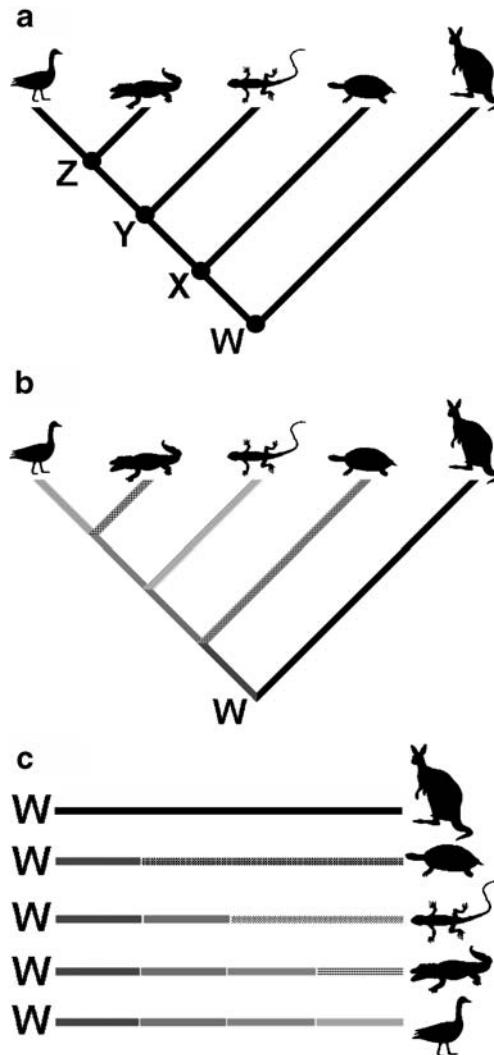


Fig. 17 The number of intervening nodes does not indicate overall relatedness between lineages. The tree in **a** is the same in topology as the one used in the study of Meir et al. (2007), which showed that many readers have a tendency to misread the directionality of time on phylogenies and to count nodes when asked to determine evolutionary relatedness among species. Confusion may arise in this particular case because many people maintain the erroneous assumption that mammals are the most “advanced” and therefore must be the youngest group. More generally, because the tree is unbalanced, students may tend to consider birds and mammals (separated by four internal nodes on this tree, *Z*, *Y*, *X*, and *W*) as more distantly related than turtles and mammals (separated by two internal nodes, *X* and *W*). However, this is simply an artifact of the species chosen for inclusion on the tree. All species descended from ancestor *X* are equally related to kangaroos, with which they all share the same last common ancestor, *W*. To demonstrate this, **b** illustrates the same tree with different patterns for each branch, which are then spliced together in **c** to reveal the identical total distance from the common ancestor *W* to all of the terminal nodes

scaled to time, the historical order in which ancestors lived extends from the most distant past at the root, through a series of increasingly recent branching events, to the tips representing contemporary species.

Misconception #9: More Intervening Nodes Equals More Distantly Related

In the study by Meir et al. (2007), many students demonstrated a tendency to assess relatedness in a phylogeny like the one depicted in Fig. 17a by “counting nodes.” For example, because birds on this tree are separated from mammals by four internal nodes (Z, Y, X, W), whereas the separation of turtles and mammals consists of only two internal nodes (X, W), many students incorrectly concluded that birds must be more distantly related to mammals than are turtles. The important point in calculating relatedness is not the number of intervening nodes along a given branch but the number of shared ancestors.

In Fig. 17a, both turtles and birds share one ancestor with mammals (node W), making them equally closely related to mammals. By contrast, birds share three common ancestors with crocodilians (nodes Z, Y, and X) but only two with turtles (nodes X and W), which makes birds and crocodilians more closely related to one another than either is to turtles. To illustrate the basic notion that all modern species in a tree are equally distant from their common ancestor, one can plot the same phylogeny as in Fig. 17a with different patterns for each branch (Fig. 17b) and then splice those branches together to show that the total distance from the root (node W) to any of the terminal nodes is exactly equal (Fig. 17c). The only difference is the number of branching events that occurred within the lineages, whereas the relatedness of the lineages themselves is not affected by this. Misconceptions about relatedness based on node counting also could be countered by balancing the tree, for example by deleting all but one species of birds/reptiles, resulting in a symmetrical V-shaped tree, regardless of which species remains along with mammals, or by adding an equal number of mammals to the sample to even out the diversity along the major branches.

Misconception #10: Change Only at Nodes

There is a legitimate debate among professional evolutionary biologists regarding the patterns of species formation, such as whether it occurs comparatively rapidly (in a geological sense) or is more gradual. Proponents of the punctuated equilibrium model of speciation argue that species remain largely unchanged morphologically for the duration of their existence, with most physical diversification occurring concomitant with species formation events (Eldredge and Gould 1972; Gould 2002; Eldredge 2008). If

punctuated equilibrium were established conclusively to represent the exclusive mode of species formation in a clade and an accurate and complete phylogenetic tree were available for that clade that included all living and extinct species, then one could reasonably interpret the internal nodes as the points at which most morphological divergence took place among species. As Meir et al. (2007) noted, many students do draw such a conclusion, although of course this is not because they possess the requisite knowledge on which to base it.

The fact is that one should not assume that an internal node indicates the exact moment (again, geologically speaking) when particular physical changes came about, any more than one should interpret a long, node-free branch as indicating that no change has occurred. More accurately, an internal node represents the time at which a formerly cohesive population diverged into two genetically isolated descendant populations, with morphological change possible both at this time and long afterward (Baum et al. 2005).

Finally, one must bear in mind that terminal nodes can also be misinterpreted if the diversity that they sometimes represent is neglected. For example, the tree in Fig. 11 shows only a single fish, a frog, a lizard, a bird, a cat, and a human, but in actuality, these six terminal nodes together represent more than 50,000 species of living vertebrates and an untold number of ancestors. The important point is that any given node, whether internal or at the tips, represents a diverse assemblage of organisms with a complex evolutionary history.

Looking Ahead to Better Understanding the Past

Two points are abundantly clear when it comes to phylogenetic literacy: (1) It is crucial for an understanding of modern evolutionary concepts, and (2) it is insufficiently common. Misconceptions abound regarding evolutionary trees—sometimes because of, and sometimes creating, incorrect preconceptions about how evolution operates. Many are holdovers of progressionist or even pre-evolutionary thinking about life’s diversity. Some, along with widespread misunderstandings of evolutionary mechanisms such as natural selection, undoubtedly contribute to the staggeringly low public acceptance of the principle of common descent in North America (Alters and Nelson 2002; Miller et al. 2006).

The way forward on this issue is unambiguous. Students, members of the public, and other nonspecialists must be better educated about the information that evolutionary trees do and do not convey. To this end, several teaching plans and software exercises for constructing and/or using phylogenetic hypotheses have become available (e.g., Bilardello and Valdes 1998; Gendron 2000; Singer et al.

2001; Goldsmith 2003; Meir et al. 2005). In addition, freely accessible online resources are making it possible for individuals to learn about and interact with evolutionary trees (see *Appendix*).

More generally, lessons at the high school and undergraduate level should de-emphasize the technical aspects of phylogeny reconstruction in favor of a focus on the concepts underlying tree thinking. In this regard, identifying, confronting, and clarifying misconceptions is perhaps the most important strategy. After all, a misconception corrected is a concept better understood. In few cases is this more relevant or more important than with Darwin's preferred metaphor of the Tree of Life.

Acknowledgments I thank Sarah Adamowicz, Alex Ardila Garcia, Martin Brummell, Niles Eldredge, Bruce Lieberman, Mark Pagel, Andy Purvis, Jillian Smith, Phillip Spinks, and Jonathan Witt for feedback on an early draft of the paper.

Appendix. Online resources

Understanding evolution

Understanding phylogenies:

http://evolution.berkeley.edu/evolibrary/article/evo_05
Phylogenetic systematics, a.k.a. evolutionary trees:
http://evolution.berkeley.edu/evolibrary/article/0_0_0/phylogenetics_01

Tree-Thinking Group

<http://www.tree-thinking.org>

The Tree-Thinking Challenge Supplemental Quiz

www.sciencemag.org/cgi/content/full/310/5750/979/DC1

Tree of Life Web Project

<http://tolweb.org/tree/phylogeny.html>

Interactive Tree of Life (iTOL)

<http://itol.embl.de/index.shtml>

TreeBASE

<http://www.treebase.org/treebase/index.html>

Encyclopedia of Life

<http://www.eol.org/>

Caminalcules teaching activity

<http://nsm1.nsm.iup.edu/rgendron/Caminalcules.shtml>

References

- Alters BJ, Nelson CE. Teaching evolution in higher education. *Evolution* 2002;56:1891–901.
- Baldau SL. The deep roots of eukaryotes. *Science* 2003;300:1703–6.
- Baum DA, DeWitt Smith S, Donovan SSS. The tree-thinking challenge. *Science* 2005;310:979–80.
- Benton MJ. Vertebrate palaeontology. 3rd ed. Malden, MA: Blackwell; 2005.
- Benton MJ, Ayala FJ. Dating the tree of life. *Science* 2003;300:1698–700.
- Bilardello N, Valdes L. Constructing phylogenies. *Am Biol Teach* 1998;60:369–73.
- Bininda-Emonds ORP, Cardillo M, Jones KE, MacPhee RDE, Beck RMD, Grenyer R, et al. The delayed rise of present-day mammals. *Nature* 2007;446:507–12.
- Boisserie J-R, Lihoreau F, Brunet M. The position of Hippopotamidae within Cetartiodactyla. *Proc Natl Acad Sci USA* 2005;102:1537–41.
- Cracraft J, Donoghue MJ, editors. Assembling the tree of life. Oxford, UK: Oxford University Press; 2004.
- Crisp MD, Cook LG. Do early branching lineages signify ancestral traits? *Trends Ecol Evol* 2005;20:122–8.
- Darwin C. Notebook B. 1837; p. 36, 74.
- Darwin C. On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. London, UK: Murray; 1859.
- Dawkins R. Progress. In: Keller EF, Lloyd EA, editors. *Keywords in evolutionary biology*. Cambridge, MA: Harvard University Press; 1992. p. 263–72.
- Doolittle WF. Uprooting the tree of life. *Sci Am* 2000;282:90–5. (Feb.).
- Doolittle WF, Bapteste E. Pattern pluralism and the tree of life hypothesis. *Proc Natl Acad Sci USA* 2007;104:2043–9.
- Eldredge N, Gould SJ. Punctuated equilibria: an alternative to phyletic gradualism. In: Schopf TJM, editor. *Models in paleobiology*. San Francisco, CA: Freeman Cooper; 1972. p. 82–115.
- Eldredge N. The early “evolution” of “punctuated equilibria”. *Evo Edu Outreach* 2008;2. DOI 10.1007/s12052-008-0032-0.
- Felsenstein J. Inferring phylogenies. Sunderland, MA: Sinauer; 2003.
- Gendron RP. The classification and evolution of Caminalcules. *Am Biol Teach* 2000;62:570–6.
- Goldsmith DW. The great clade race: Presenting cladistic thinking to biology majors and general science students. *Am Biol Teach* 2003;65:679–82.
- Gould SJ. *Bully for Brontosaurus*. New York: W.W. Norton; 1991. p. 168–81.
- Gould SJ. The evolution of life on the earth. *Sci Am* 1994;271:85–91. (Oct.).
- Gould SJ. *Full house*. New York: Harmony; 1996.
- Gould SJ. Branching through a wormhole. *Nat Hist* 1999;108:24–7, 84–9. (March).
- Gould SJ. The structure of evolutionary theory. Cambridge, MA: Harvard University Press; 2002.
- Gregory TR. Evolution as fact, theory, and path. *Evo Edu Outreach* 2008;1:46–52.
- Hall BG. Phylogenetic trees made easy. 3rd ed. Sunderland, MA: Sinauer; 2007.
- Harcourt-Brown KG, Pearson PN, Wilkinson M. The imbalance of paleontological trees. *Paleobiology* 2001;27:188–204.

- Hillis DM, Bull JJ, White ME, Badgett MR, Molineux IJ. Experimental phylogenetics: generation of a known phylogeny. *Science* 1992;255:589–92.
- Hillis DM, Moritz C, Mable BK, editors. Molecular systematics. 2nd ed. Sunderland, MA: Sinauer; 1996.
- Hodkinson TR, Parnell JAN, editors. Reconstructing the tree of life: taxonomy and systematics of species rich taxa. Boca Raton, FL: CRC; 2007.
- Huxley TH. On the application of the laws of evolution to the arrangement of the Vertebrata and more particularly of the Mammalia. *Proc Zool Soc Lond* 1880;43:659–62.
- Lamarck JB. Philosophie zoologique. Paris; 1809.
- Lecointre G, Le Guyader H. The tree of life: a phylogenetic classification. Cambridge, MA: Belknap; 2007.
- Mace GM, Gittleman JL, Purvis A. Preserving the tree of life. *Science* 2003;300:1707–9.
- Mace R, Holden CJ. A phylogenetic approach to cultural evolution. *Trends Ecol Evol* 2005;20:116–21.
- MacFadden BJ. Fossil horses: systematics, paleobiology, and evolution of the family Equidae. Cambridge, UK: Cambridge University Press; 1992.
- MacFadden BJ. Fossil horses—evidence for evolution. *Science* 2005;307:1728–9.
- Maddison DR, Schulz K-S, editors. The tree of life web project. <http://tolweb.org> (2007).
- Mayr E, Provine WB, editors. The evolutionary synthesis. Cambridge, MA: Harvard University Press; 1980.
- Meir E, Herron JC, Maruca S, Stal D, Kingsolver J. *EvoBeaker 1.0*. Ithaca, NY: SimBiotic Software. <http://www.simbio.com> (2005).
- Meir E, Perry J, Herron JC, Kingsolver J. College students' misconceptions about evolutionary trees. *Am Biol Teach* 2007; 69:71–6.
- Miller JD, Scott EC, Okamoto S. Public acceptance of evolution. *Science* 2006;313:765–6.
- Nee S. The great chain of being. *Nature* 2005;435:429.
- Nei M, Kumar S. Molecular phylogenetics and evolution. Oxford, UK: Oxford University Press; 2000.
- O'Hara RJ. Homage to Clio, or, toward an historical philosophy for evolutionary biology. *Syst Zool* 1988;37:142–55.
- O'Hara RJ. Telling the tree: narrative representation and the study of evolutionary history. *Biol Philos* 1992;7:135–60.
- O'Hara RJ. Evolutionary history and the species problem. *Am Zool* 1994;34:12–22.
- O'Hara RJ. Population thinking and tree thinking in systematics. *Zool Scr* 1997;26:323–9.
- Page RDM, Holmes EC. Molecular evolution: A phylogenetic approach. Malden, MA: Blackwell; 1998.
- Prothero DR. Evolution: What the fossils say and why it matters. New York: Columbia University Press; 2007.
- Rambaut A, Robertson DL, Pybus OG, Peeters M, Holmes EC. Phylogeny and the origin of HIV-1. *Nature* 2001;410:1047–8.
- Rokas A. Genomics and the tree of life. *Science* 2006;313:1897–8.
- Ruse M. Monad to man. Cambridge, MA: Harvard University Press; 1996.
- Salemi M, Vandamme A-M, editors. The phylogenetic handbook. Cambridge, UK: Cambridge University Press; 2003.
- Sanson GFO, Kawashita SY, Brunstein A, Briones MRS. Experimental phylogeny of neutrally evolving DNA sequences generated as a bifurcate series of nested polymerase chain reactions. *Mol Biol Evol* 2002;19:170–8.
- Singer F, Hagen JB, Sheehy RR. The comparative method, hypothesis testing and phylogenetic analysis. *Am Biol Teach* 2001;63: 518–23.
- Thewissen JGM, Bajpai S. Whale origins as a poster child for macroevolution. *BioScience* 2001;51:1037–49.
- Thewissen JGM, Williams EM. The early radiations of Cetacea (Mammalia): evolutionary pattern and developmental correlations. *Ann Rev Ecol Syst* 2002;33:73–90.
- Vogel G. Phylogenetic analysis: getting its day in court. *Science* 1997;275:1559–60.
- Wolf YI, Rogozin IB, Grishin NV, Koonin EV. Genome trees and the tree of life. *Trends Genet* 2002;18:472–9.

Mid-Pleistocene divergence of Cuban and North American ivory-billed woodpeckers

Robert C. Fleischer^{1,5,*}, Jeremy J. Kirchman²,
John P. Dumbacher³, Louis Bevier⁴, Carla Dove⁵,
Nancy C. Rotzel⁶, Scott V. Edwards⁶,
Martjan Lammertink⁷, Kathleen J. Miglia⁸
and William S. Moore⁸

¹Genetics Program, National Museum of Natural History and National Zoological Park, Smithsonian Institution, 3001 Connecticut Avenue, Northwest, Washington, DC 20008, USA

²Department of Zoology, University of Florida, Gainesville, FL 32611, USA

³California Academy of Sciences, San Francisco, CA 94103, USA

⁴Department of Biology, Colby College, 5720 Mayflower Hill Drive, Waterville, ME 04901, USA

⁵Smithsonian Ornithology and Division of Birds, National Museum of Natural History, Smithsonian Institution, Washington, DC 20526, USA

⁶Museum of Comparative Zoology, Harvard University, Cambridge, MA 02138, USA

⁷Cornell Laboratory of Ornithology, Cornell University, 159 Sapsucker Woods Road, Ithaca, NY 14850, USA

⁸Department of Biological Sciences, Wayne State University, Detroit, MI 48202, USA

*Author for correspondence (fleischr@si.edu).

We used ancient DNA analysis of seven museum specimens of the endangered North American ivory-billed woodpecker (*Campephilus principalis*) and three specimens of the species from Cuba to document their degree of differentiation and their relationships to other *Campephilus* woodpeckers. Analysis of these mtDNA sequences reveals that the Cuban and North American ivory bills, along with the imperial woodpecker (*Campephilus imperialis*) of Mexico, are a monophyletic group and are roughly equidistant genetically, suggesting each lineage may be a separate species. Application of both internal and external rate calibrations indicates that the three lineages split more than one million years ago, in the Mid-Pleistocene. We thus can exclude the hypothesis that Native Americans introduced North American ivory-billed woodpeckers to Cuba. Our sequences of all three woodpeckers also provide an important DNA barcoding resource for identification of non-invasive samples or remains of these critically endangered and charismatic woodpeckers.

Keywords: *Campephilus principalis*; ivory-billed woodpecker; Pleistocene divergence; phylogenetics; DNA barcoding

1. INTRODUCTION

The ivory-billed woodpecker (*Campephilus principalis*) is a spectacular bird that was thought to be extinct until publication of recent reports (Fitzpatrick *et al.* 2005). Ivory-billed woodpeckers occurred in two disjunct

The electronic supplementary material is available at <http://dx.doi.org/10.1098/rsbl.2006.0490> or via <http://www.journals.royalsoc.ac.uk>.

regions, with *C. p. principalis* found in primary hardwood forest throughout much of the southeastern United States and *C. p. bairdii* on the island of Cuba in mature lowland hardwood and hill pine forests. Populations in Cuba seriously declined over the past century and may now be extinct (Stattersfield & Capper 2000; Jackson 2002). The related imperial woodpecker (*Campephilus imperialis*) also underwent a major population decline in the highlands of Mexico and may also be extinct (Stattersfield & Capper 2000).

Cuban and mainland ivory-billed woodpeckers differ only slightly in plumage and size (Jackson 2002), and, although originally described as separate species by Cassin (1863), have been considered conspecific since at least 1976 (AOU 1976). In fact, Jackson (2002) suggested that ivory-billed woodpeckers in Cuba are a recent human-assisted introduction, citing evidence of extensive trade in birds between Native Americans in Cuba and Florida, a high value placed on ivory-billed woodpecker artefacts by Native Americans, and a presumed low propensity for dispersal by a forest-dwelling woodpecker over open water.

Here, we compare mitochondrial DNA sequences from museum specimens to estimate the level of genetic divergence between the two forms, and between each and the related imperial woodpecker (*C. imperialis*) and other members of their genus. We evaluate their taxonomic status and test whether the Cuban ivory-billed woodpecker is a recent introduction. Our sequences also provide an important DNA barcoding resource for the ivory-billed woodpecker and relatives. We find that the Cuban ivory-billed woodpecker is mitochondrial distinct from the North American form, a result that has implications for the conservation of the species in both regions.

2. MATERIAL AND METHODS

(a) Samples

We used scalpels to pare small samples of skin from the toe pads of museum specimens collected between 1861 and 1923 of seven North American and three Cuban ivory-billed woodpeckers, and three imperial woodpeckers. These museum specimens of *C. p. principalis*, *C. p. bairdii* and *C. imperialis* are listed in the electronic supplementary material table 1, along with accession number, locale and date of collection, sex, amount of sequence obtained and GenBank (NCBI) accession number. We also obtained frozen tissues of seven South and Central American species of *Campephilus*, along with two species each of *Dryocopus* and *Picumnus* woodpeckers as outgroups. Details concerning tissue specimens of *Campephilus* and other outgroups analysed at Wayne State University are shown in the electronic supplementary material table 2.

(b) Laboratory analysis

All pre-PCR ancient DNA labwork was completed in isolated ancient DNA laboratories at the Smithsonian, University of Florida, California Academy of Sciences and Harvard Museum of Comparative Zoology, following stringent protocols to avoid and detect potential contamination. DNA was isolated from museum specimen toe pads using phenol-chloroform extraction and centrifugal dialysis following the methods outlined in Fleischer *et al.* (2000). Short fragments of specific regions of mitochondrial DNA (*COI*, cytochrome *b*, *ND2*, *ATP6/8*, 12S rDNA) were amplified from the museum specimen extracts using PCR and specific primers. Amplification of these small fragments used primers listed in the electronic supplementary material table 3, developed in the case of *COI* from existing sequence for related species (DeFilippis & Moore 2000). Up to 1730 nucleotide sites were obtained, and at least 1079 bp were sequenced for at least two individuals of each taxon; smaller amounts of sequence were obtained for additional individuals. Nearly half of the sequences were at least partially and independently replicated by more than one laboratory (electronic supplementary material table 1). Analysis of

DNA from outgroups was conducted at Wayne State University, and limited to *COI*, cytochrome *b* and 12S rDNA sequences, with somewhat different primers and methods. Detailed methods and primers for amplification and sequencing of these outgroup taxa are available for cytochrome *b* in Moore & DeFilippis (1997), for *COI* in DeFilippis & Moore (2000) and for 12S in Webb & Moore (2005).

(c) Phylogenetic analyses

Phylogenies were reconstructed using three approaches: maximum likelihood (ML), maximum parsimony (MP) and Bayesian. We conducted an unweighted analysis with a heuristic search and 10 random addition repetitions under the MP criterion in PAUP* (Swofford 2002), and a subsequent bootstrap analysis with 1000 replications. The Akaike information criterion (AIC) in MODELTEST (Posada & Crandall 1998) was used to select the most appropriate model of sequence evolution for subsequent ML and Bayesian approaches. The AIC selected the GTR+I+G model, with a gamma shape parameter $\alpha=5.55$ and the proportion of invariable sites $I=0.657$. The ML analysis involved a heuristic search, and a bootstrap of 100 heuristic repetitions to indicate nodal support in PAUP*. Bayesian analyses were run in MrBAYES (Huelskenbeck & Ronquist 2001). We did a chain of 1 000 000 generations MCMC with the parameters as selected by MODELTEST (above). We excluded the first 500 000 generations as burn-in, and computed a consensus from 50 000 trees in the remainder. *Campephilus* taxa were rooted to *Picumnus* and *Dryocopus* outgroups.

(d) Dating of nodes

A likelihood-ratio test comparing a molecular clock-constrained and an unconstrained tree in PAUP* (Swofford 2002) indicated no significant departure from a molecular clock for this dataset ($\chi^2_{17} = 25.23$, $p > 0.05$). Dating of key nodes was estimated with MEGA using an ultrametric tree (clock-like) approach (Kumar *et al.* 2004), and with R8S (Sanderson 2003) using variable clock approaches (although the lack of significant among-lineage rate variation suggests the latter corrections are not necessary). For both of these methods, we used two calibrations. First, we used the rate of $2.0 \pm 0.5\%$ sequence divergence per million years (Myr) calibrated for cytochrome *b* in woodpeckers from Moore *et al.* (1999), based on a fossil *Colaptes* woodpecker. This rate is similar to those calibrated for many other avian taxa (e.g. Fleischer *et al.* 1998; Paxinos *et al.* 2002). Second, because of the basal position of *Campephilus haematogaster* in our phylogeny, we make an assumption that the ancestor of *Campephilus* clades A and B (see figure 1) occurred in South America, and was able to colonize North America only after the formation of the Isthmus of Panama about 3.1 Myr ago (Coates & Obando 1996). Once this colonization occurred, the two lineages could diverge, and the range of *Campephilus guatemalensis* would represent a secondary range expansion from South America into Central America.

3. RESULTS AND DISCUSSION

Phylogenetic analyses of these mtDNA sequences show that the Cuban and mainland North American ivory-billed woodpeckers are clearly distinct lineages. The analysis is ambiguous with regard to their status as sister taxa (figure 1), statistically a polytomy containing both ivory-billed woodpecker taxa and the imperial woodpecker. Our MP search resulted in three trees, the consensus of which matched in topology the ML and Bayesian trees, as did the consensus tree resulting from a 1000-repetition bootstrap with an MP search. The MP and ML trees have high bootstrap support values and the Bayesian tree has high posterior probabilities for many nodes, including for monophyly of the northern *Campephilus* taxa (*C. imperialis*, *C. p. principalis* and *C. p. bairdii*; clade A, figure 1). Clade A is sister to a clade containing all other *Campephilus* woodpeckers (clade B), except for *C. haematogaster*.

Genetic divergence between the North American and Cuban ivory-billed woodpeckers was calculated in MEGA (Kumar *et al.* 2004), and averaged 0.021 ± 0.007 substitutions per site. This value is similar to

that between each ivory-billed woodpecker taxon and the imperial woodpecker (0.017 ± 0.008 and 0.017 ± 0.004 , respectively). Applying the woodpecker calibrated rate of 2.0% sequence divergence per Myr for cytochrome *b* sequence (Moore *et al.* 1999) to the cytochrome *b* distance between the Cuban and North American ivory-billed woodpeckers results in an estimated split date of 1.0 Myr. Notably, the sequence divergence rate for cytochrome *b* calculated from our dataset using the date of the separation of the Isthmus of Panama was 1.9% per Myr.

Assuming clades A and B began to diverge 3.1 Myr ago, our estimates either with or without a molecular clock resulted in rates for our entire mtDNA dataset of between 0.011 and 0.014 substitutions per site per Myr. The predicted dates of divergence by penalized likelihood (PL, TN model) in R8S (Sanderson 2003) were 1.15 Myr ago for the split between *C. p. principalis* and (*C. p. bairdii*, *C. imperialis*), and 0.91 Myr ago for the split between *C. p. bairdii* and *C. imperialis*; by non-parametric rate smoothing (Powell model, Sanderson 2003): 1.57 and 1.27 Myr; by Langley-Fitch (Sanderson 2003): 1.05 and 0.83 Myr; and by MEGA (Kumar *et al.* 2004): 1.37 and 1.14 Myr. Thus, applying these different methods to the divergence of the three North American *Campephilus* lineages indicates that their split occurred sometime between about 0.8 and 1.6 Myr ago, although coalescent effects could push the species split closer to the present by as much as 200 000 years (Moore 1995; Edwards & Beerli 2000).

Our data suggest that, like *C. imperialis*, both *C. p. principalis* and *C. p. bairdii* should also be considered separate species. This contention is supported by the effectively unresolved trichotomy among the three northern taxa, and the degree of genetic divergence among them. Such a classification would appear to be supported under either a traditional biological species concept (i.e. allopatric, divergent lineages that appear reproductively isolated) or a phylogenetic species concept (i.e. allopatric, diagnosable lineages).

Our dating analyses reveal that the Cuban and North American ivory-billed woodpeckers and the imperial woodpecker diverged sometime in the Mid-Pleistocene. This period corresponds to the Mid-Pleistocene revolution (when global temperature oscillations changed from a 41 ky period to a 100 ky period), and was a period of global cooling and lowered sea-levels (Maasch 1988; Raymo *et al.* 1997). Reported sea-level differences of more than 30 m (Wright & Flower 2002) at this time would have increased the size of the Yucatan Peninsula and reduced the *ca* 176 km current distance between the Yucatan Peninsula and Cuba, and thus may have favoured colonization of Cuba by a woodpecker presumably averse to flying over water. Our data clearly do not support the hypothesis that Native Americans recently introduced the ivory-billed woodpecker from North America into Cuba (Jackson 2002).

Our discovery of three genetically divergent lineages of northern *Campephilus* that each may be a different species increases the already urgent need for

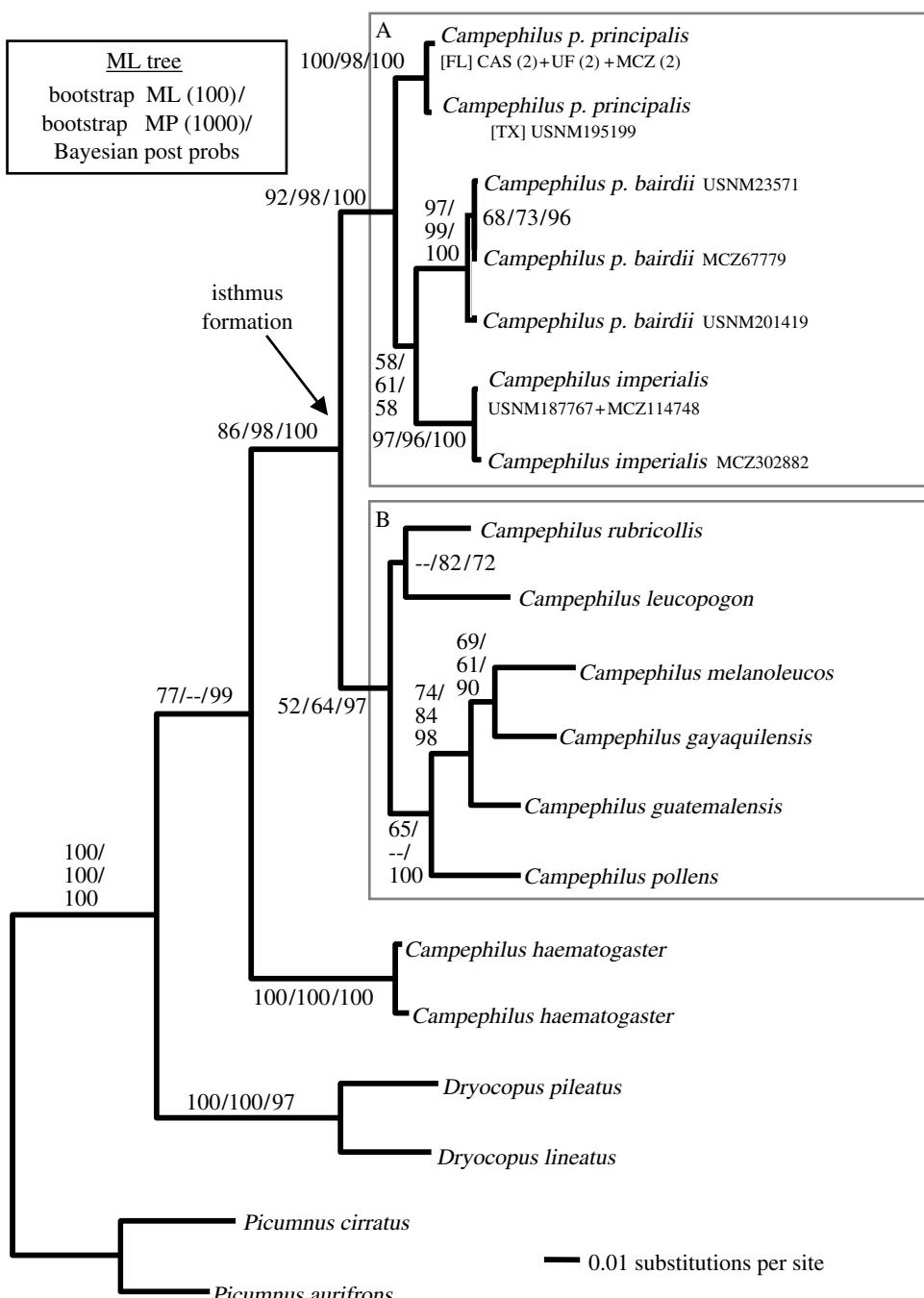


Figure 1. Phylogram of *Campephilus* species and outgroups produced by heuristic search with maximum-likelihood (ML) criterion. Topologies recovered from maximum-parsimony (MP) and Bayesian approaches were identical other than minor rearrangements of internal nodes in the three MP trees recovered. Numbers at nodes are ML (100-repetition) and MP (1000-repetition) bootstrap percentages and Bayesian posterior probabilities. Tree was rooted with *Dryocopus* and *Picumnus* outgroups. Note that the boxed clade A contains northern *Campephilus* taxa (i.e. ivory-billed and imperial woodpeckers), while the boxed clade B contains southern *Campephilus* taxa (with putative secondary northern expansion by *C. guatemalensis*). The basal position of *C. haematogaster* suggests a South American origin of *Campephilus* and colonization of North America after closure of the Isthmus of Panama. See text for details of phylogenetic and dating analyses.

rediscovery and conservation of this critical branch of the woodpecker tree. Our results will also provide an important DNA barcoding resource that could facilitate the discovery of living ivory-billed and imperial woodpeckers by identification of shed materials, such as feathers and faeces sampled from the wild.

We thank David Steadman and Gary Graves for permission to sample specimens under their care, and Rebecca Kimball, Kristy L. Deiner, Jeremiah Trimble, Carl E. McIntosh, Jennifer Reed and Dawn Reding for facilitating or assisting

in laboratory work. We appreciate insights from Storrs Olson, Jim Hanken, Jerome Jackson and Per Ericson. The Smithsonian Institution and NSF DEB#0452732 provided funding.

American Ornithologists' Union 1976 Check-list of North American birds. *Auk* **93**, 875.

Cassin, J. 1863 Notes on the Picidae. *Proc. Acad. Nat. Sci. Phila.* **15**, 322–328.

Coates, A. G. & Obando, J. A. 1996 Geological evolution of the Central American Isthmus. In *Evolution and*

- environment in tropical America (ed. J. B. C. Jackson, A. G. Coates & A. Budd), pp. 21–56. Chicago, IL: University of Chicago Press.
- DeFilippis, V. R. & Moore, W. S. 2000 Resolution of phylogenetic relationships among recently evolved species as a function of amount of DNA sequence: an empirical study based on woodpeckers (Aves: Picidae). *Mol. Phylogenet. Evol.* **16**, 143–160. (doi:10.1006/mpev.2000.0780)
- Edwards, S. V. & Beerli, P. 2000 Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution* **54**, 1839–1854.
- Fitzpatrick, J. W. et al. 2005 Ivory-billed woodpecker (*Campephilus principalis*) persists in continental North America. *Science* **308**, 1460–1462. (doi:10.1126/science.1114103)
- Fleischer, R. C., McIntosh, C. E. & Tarr, C. L. 1998 Evolution on a volcanic conveyor belt: using phylogeographic reconstructions and K–Ar based ages of the Hawaiian Islands to estimate molecular evolutionary rates. *Mol. Ecol.* **7**, 533–545. (doi:10.1046/j.1365-294x.1998.00364.x)
- Fleischer, R. C., Olson, S., James, H. A. & Cooper, A. C. 2000 The identity of the extinct Hawaiian eagle (*Haliaeetus*) as determined by mitochondrial DNA sequence. *Auk* **117**, 1051–1056.
- Huelsenbeck, J. P. & Ronquist, F. 2001 MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* **17**, 754–755. (doi:10.1093/bioinformatics/17.8.754)
- Jackson, J. A. 2002 Ivory-billed woodpecker (*Campephilus principalis*). In *The birds of North America*, no. 711 (ed. A. Poole & F. Gill). Philadelphia, PA: The Academy of Natural Sciences.
- Kumar, S., Tamura, K. & Nei, M. 2004 MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief. Bioinform.* **5**, 150–163. (doi:10.1093/bib/5.2.150)
- Maasch, K. A. 1988 Statistical detection of the mid-Pleistocene transition. *Clim. Dyn.* **2**, 133–143. (doi:10.1007/BF01053471)
- Moore, W. S. 1995 Inferring phylogenies from mtDNA variation: mitochondrial gene trees versus nuclear gene trees. *Evolution* **49**, 718–726.
- Moore, W. S. & DeFilippis, V. R. 1997 The window of taxonomic resolution for avian phylogenies based on mitochondrial cytochrome *b* DNA sequences. In *Avian molecular evolution and systematics* (ed. D. P. Mindell), pp. 83–119. San Diego, CA: Academic Press.
- Moore, W. S., Smith, S. M. & Prychitko, T. 1999 Nuclear gene introns versus mitochondrial genes as molecular clocks. In *Proc. 22nd Int. Ornithological Congress* (ed. N. Adams & R. Slotow), pp. 745–753. Durban, South Africa: University of Natal.
- Paxinos, E. E., James, H. F., Olson, S. L., Sorenson, M. D., Jackson, J. & Fleischer, R. C. 2002 mtDNA from fossils reveals a radiation of Hawaiian geese recently derived from the Canada goose (*Branta canadensis*). *Proc. Natl Acad. Sci. USA* **99**, 1399–1404. (doi:10.1073/pnas.032166399)
- Posada, D. & Crandall, K. A. 1998 Modeltest: testing the model of DNA substitution. *Bioinformatics* **14**, 817–818. (doi:10.1093/bioinformatics/14.9.817)
- Raymo, M. E., Oppo, D. W. & Curry, W. 1997 The mid-Pleistocene climate transition: a deep sea carbon isotopic perspective. *Paleoceanography* **12**, 546–559. (doi:10.1029/97PA01019)
- Sanderson, M. J. 2003 r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinform. Appl. Note* **19**, 301–302. (doi:10.1093/bioinformatics/19.2.301)
- Stattersfield, A. J. & Capper, D. R. (eds) 2000 *Threatened birds of the world*. Cambridge, UK: Lynx Edicions and BirdLife International.
- Swofford, D. L. 2002 PAUP*. *Phylogenetic analysis using parsimony (*and other methods)*, version 4.0. Sunderland, MA: Sinauer Associates.
- Webb, D. M. & Moore, W. S. 2005 A phylogenetic analysis of woodpeckers and their allies using 12S, Cyt *b*, and COI nucleotide sequences (class Aves; order Piciformes). *Mol. Phylogenet. Evol.* **36**, 233–248. (doi:10.1016/j.ympev.2005.03.015)
- Wright, A. K. & Flower, B. P. 2002 Surface and deep ocean circulation in the subpolar North Atlantic during the mid-Pleistocene revolution. *Paleoceanography* **17**, 1068. (doi:10.1029/2002PA000782)



Short Communication

The flight of the Passenger Pigeon: Phylogenetics and biogeographic history of an extinct species

Kevin P. Johnson ^{a,*}, Dale H. Clayton ^b, John P. Dumbacher ^{c,d}, Robert C. Fleischer ^d^a Illinois Natural History Survey, University of Illinois, Champaign, IL 61820, USA^b Department of Biology, University of Utah, Salt Lake City, UT 84112, USA^c California Academy of Sciences, San Francisco, CA 94103, USA^d Center for Conservation and Evolutionary Genetics, Smithsonian Conservation Biology Institute, National Zoological Park, Washington, DC 20008, USA

ARTICLE INFO

Article history:

Received 3 September 2009

Revised 7 May 2010

Accepted 11 May 2010

Available online 15 May 2010

Keywords:

Ectopistes migratorius

Molecular systematics

Ancient DNA

Columbiformes

Doves

ABSTRACT

The human-caused extinction of the Passenger Pigeon (*Ectopistes migratorius*) is one of the best known and documented of any bird. This event was particularly alarming because the Passenger Pigeon went from being one of the most numerous avian species in the world to extinct in a period of decades, when the last individual died in captivity in a Cincinnati Zoo in 1914. While a great deal of information exists on the likely direct and indirect causes of its demise, as well as information on life-history, the phylogenetic relationships of this species have been subject to considerable speculation. Here we use DNA sequences obtained from museum specimens to resolve the phylogenetic position of this species with respect to other pigeons and doves (Columbiformes). We show that the Passenger Pigeon is not related to the New World mourning doves (*Zenaida*) as many authors have suggested, but is the sister taxon of all other New World pigeons (*Patagioenas*). Biogeographic analysis suggests the Passenger Pigeon lineage may have colonized North America from Asia, and subsequently dispersed into South America, leading to a more extensive radiation of New World pigeons.

© 2010 Published by Elsevier Inc.

1. Introduction

The Passenger Pigeon (*Ectopistes migratorius*) was once the most numerous species of bird in North America but, due to human over-exploitation combined with other factors, became extinct in a matter of decades. This dramatic decline is perhaps one of the best known and documented extinctions for any species, and has served as a poignant reminder of the impact humans can have on nature. Unlike many cases, where uncertainty exists regarding the exact timing of extinction, considerable documentation of the demise of the Passenger Pigeon exists (Schorger, 1955; Blockstein, 2002); the last bird died in captivity in a Cincinnati Zoo in 1914. Information on the natural history of the Passenger Pigeon has also been relatively well documented (Blockstein, 2002). However, the evolutionary relationships of the Passenger Pigeon have been the subject of speculation, with little or no rigorous phylogenetic testing.

The phylogenetic relationships of the Passenger Pigeon are generally unclear, and it has typically been placed in the monotypic genus *Ectopistes*. Most authors suggest that the Passenger Pigeon is closely related to the New World mourning doves (*Zenaida*),

which include representatives in North and South America (Goodwin, 1983; Gibbs et al., 2001; Blockstein, 2002). Although the Passenger Pigeon shares a few superficially similar plumage patterns and a long tail with some members of *Zenaida* (Goodwin, 1983), it differs by its considerably larger size, sexual plumage dimorphism, and lack of a facial stripe. Because extant pigeons and doves (Columbiformes) in the New World have undergone three distinct radiations (Johnson and Clayton, 2000; Pereira et al., 2007), understanding the phylogenetic position of the Passenger Pigeon is important to more fully understand the origin of these radiations. Prior studies with limited taxon and gene sampling (Shapiro et al., 2002; Pereira et al., 2007) indicate that the Passenger Pigeon may not be closely related to *Zenaida*, as previously hypothesized. The goals of our study were to use additional DNA sequences obtained from museum specimens of the Passenger Pigeon and denser taxon sampling to identify its phylogenetic position within Columbiformes. We used this information to reconstruct biogeographic patterns in New World pigeons and doves.

2. Materials and methods

Sequences of the mitochondrial cytochrome *b* and ATPase8, as well as the nuclear beta-fibrinogen intron 7 (totaling 2401 aligned base pairs), are available for 78 species of pigeons and doves from

* Corresponding author. Address: Illinois Natural History Survey, 1816 South Oak Street, Champaign, IL 61820, USA. Fax: +1 217 244 0729.

E-mail address: kjohnson@inhs.uiuc.edu (K.P. Johnson).

around the world (Johnson, 2004). For the present study we also extracted DNA from the extinct Passenger Pigeon (*E. migratorius*) from museum skins. Because DNA from museum skin material is highly fragmented, only short sequences of the mitochondrial cytochrome *b* (130 bp) and ATPase8 (169 bp) genes could be obtained for the Passenger Pigeon, and both genes could only be sequenced from a single specimen USNM293588 (GenBank Accession Nos. HM195099–HM195100). Sequences from other individuals for one gene or the other were identical or differed by only one or two base pairs from these sequences. These sequences were obtained using methods described elsewhere (Sorenson et al., 1999; Dumbacher and Fleischer, 2001). We also compared our sequences of the cytochrome *b* gene to that of Shapiro et al. (2002), and only one substitution difference was detected. Thus, we used the longer 1044 bp cytochrome *b* sequence from this previous study in our analyses. As in previous studies (Johnson and Clayton, 2000; Johnson, 2004) the tree was rooted on a swiftlet.

First we used parsimony analyses to reconstruct a tree from all genes combined, with the Passenger Pigeon coded as missing data for the beta-fibrinogen intron 7 gene (10 random addition replicates). However, the phylogenetic placement of a taxon for which a large number of base pairs are missing (1188 of 2401 bp missing, i.e., the nuclear beta-fibrinogen intron 7) may be subject to bias. To account for this we also constructed a parsimony tree for only the 78 taxa for which we had complete sequences. This tree was then used as a constraint tree to evaluate the phylogenetic position of the Passenger Pigeon. We used this constraint in a second parsimony analysis where we analyzed only the two mitochondrial genes that were available for the Passenger Pigeon. This technique should lead to an unbiased placement of the Passenger Pigeon within the broader tree of pigeons and doves, while at the same time allowing the overall tree structure to be determined by the maximum number of available sequences.

For the purposes of comparison, we also conducted a Bayesian analysis using MrBayes (Huelsenbeck, 2001), with the GTR + I + G model. Posterior probabilities were calculated by sampling trees every 1000 generations from a 10 million generation chain. Inspection of likelihood scores indicated that they stabilized by 100,000 generations, so the first 100 trees were discarded as burn-in. This analysis was conducted with all sequences simultaneously and the Passenger Pigeon was coded with missing data for the beta-fibrinogen intron 7 gene.

Much of the radiation of pigeons and doves appears to have been the result of dispersal and this occurred after the breakup of Gondwana (Pereira et al., 2007). Thus, historical changes in biogeographic distribution between continental regions are likely due to dispersal. Brooks (1990) parsimony method is the most appropriate method for biogeographic reconstruction where changes are the result of dispersal (Johnson and Sorenson, 1999; DaCosta and Klicka, 2008). Because we were most interested in broad scale changes between major regions, with particular focus in the New World, we coded biogeographic areas as New World and Old World (including the Australasian region). No species has a native distribution in both the New and Old World so this coding is appropriate for examining major changes in distribution.

3. Results

Trees from both parsimony and Bayesian analyses were generally well resolved and in broad agreement about phylogenetic relationships among pigeons and doves (Fig. 1). Phylogenetic analyses suggest that the Passenger Pigeon is closely related to the clade containing a western North American pigeon, the Band-tailed Pigeon (*Patagioenas fasciata*), and other large-bodied New World pi-

geons (*Patagioenas*, Fig. 1). The Passenger Pigeon was distributed throughout eastern North America, and was likely the eastern counterpart of the western Band-tailed Pigeon.

Based on Brooks (1990) method of biogeographic reconstruction, pigeons and doves invaded the New World three times from the Old World (Fig. 1). The large-bodied New World pigeon clade, including the Passenger Pigeon, is embedded within a clade of Old World Columbiformes; within this clade, they are sister to a clade of cuckoo-doves (*Macropygia* and *Reinwardtoena*) from south-east Asia.

4. Discussion

Previous authors have suggested that the extinct Passenger Pigeon is a close relative of the New World genus *Zenaida* (Goodwin, 1983; Blockstein, 2002). However, phylogenetic analysis of DNA sequences reveals that the Passenger Pigeon (*E. migratorius*) is closely related to other New World pigeons in the genus *Patagioenas* (Fig. 1). This group falls in a larger clade containing Old World pigeons (*Columba*), turtle-doves (*Streptopelia*), and cuckoo-doves (*Macropygia* and *Reinwardtoena*). This large clade has 97% bootstrap support and 100% Bayesian posterior probability, a result that is consistent with previous limited molecular systematic studies of Columbiformes (Shapiro et al., 2002; Pereira et al., 2007). There is one inferred colonization event of the New World in this clade, possibly from south-east Asia (the distribution of cuckoo-doves). Both cuckoo-doves and the Passenger Pigeon have long tails, and the juvenile Passenger Pigeon is superficially similar to some cuckoo-doves, with scaling on the neck and ruddy coloration in the wings. Furthermore, cuckoo-doves and the Passenger Pigeon both lack facial stripes, unlike *Zenaida*, which possess them. Many species of cuckoo-doves also exhibit sexual dimorphism, with males having iridescent plumage in the neck region, similar to the Passenger Pigeon, while species of *Zenaida* are sexually monomorphic.

Evidence from the distribution and phylogeny of ectoparasitic lice also supports these relationships for the Passenger Pigeon. A species of louse, *Columbicola extinctus*, has been described from the Passenger Pigeon (Malcomson, 1937). Lice are generally quite host-specific, so this louse was thought to be extinct along with the Passenger Pigeon (Stork and Lyal, 1993). However, recent revisions of the lice of pigeons and doves (Clayton and Price, 1999; Bush et al., 2009) have revealed that *C. extinctus* is also found on the Band-tailed Pigeon (*P. fasciata*), here shown to be a close relative of the Passenger Pigeon. The *extinctus* group of lice is closely related to the *Columbicola angustus* species group (Johnson et al., 2007), found mainly on cuckoo-doves (*Macropygia* and *Reinwardtoena*) of south-east Asia, which also appear to be the closest relatives of the New World pigeons (Fig. 1). Thus, the host distribution of these groups of lice broadly mirrors the host phylogeny.

If indeed North American pigeons were derived by colonization of a lineage from south-east Asia, this would be one of the few such long-distance dispersal events known for birds. A few other lineages of New World birds are believed to have their closest relatives in south-east Asia. These include the Wrentit (*Chamaea fasciata*) from the west coast of North America, believed to be closely related to babblers (Timaliidae) from south-east Asia (Barboum and Burns, 2002) and New World vireos (Vireonidae), which are closely related to *Erpornis* from south-east Asia (Barker et al., 2004). It may also be possible that, prior to the Pleistocene, ancestors of these lineages lived further north than their current distributions, and the Beringian region served as a dispersal corridor between the Old and New World during the Miocene (Barker et al., 2004). Many species of large-bodied pigeons are exceptionally strong fliers and have colonized islands across much of the

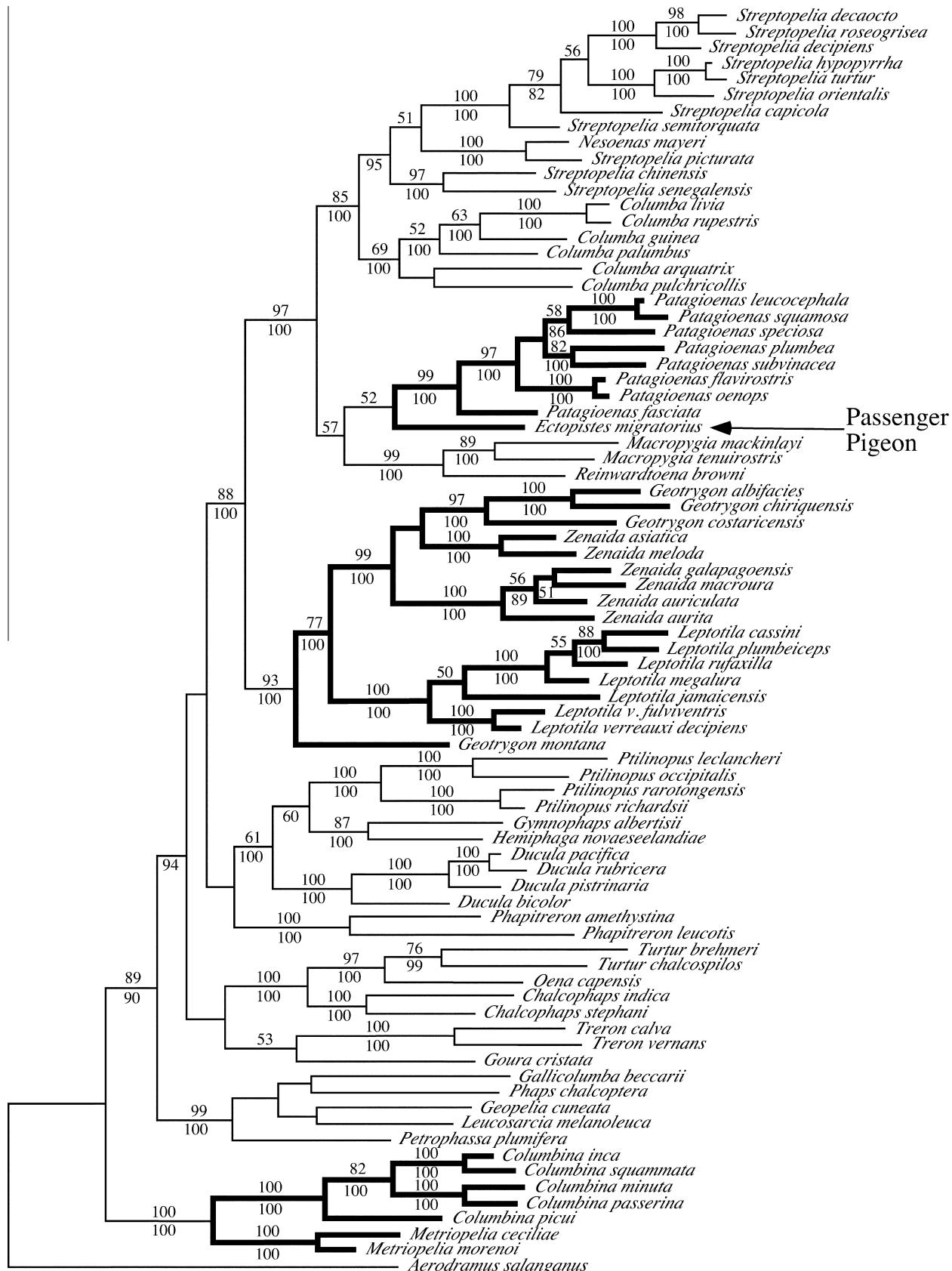


Fig. 1. Phylogeny for 79 taxa of pigeons and doves (Columbidae) reconstructed using parsimony analysis of nuclear (1188 aligned bp of beta-fibrinogen intron 7) and mitochondrial (1044 bp cytochrome *b*, 169 bp ATPase 8) sequences. GenBank Accession Nos. AF182648–AF182713. An identical placement of the Passenger Pigeon resulted from parsimony analyses that first excluded the Passenger Pigeon and then used this tree as a constraint in searches including only the mitochondrial genes with the Passenger Pigeon included. Numbers above branches are support from 1000 parsimony bootstrap replicates, and numbers below branches are posterior probabilities from Bayesian MCMC analysis. Bold branches are lineages that occur in the New World, while thin branches are Old World taxa (Eurasia, Africa, and Australasia).

south Pacific Ocean. The Passenger Pigeon was one of the fastest flying pigeons (ca. 100 km/h, Schorger, 1955), undergoing rapid long distance movements in large flocks. Thus, the ancestors of the Passenger Pigeon may have engaged in such an extraordinary flight.

Acknowledgments

The authors thank the Field Museum of Natural History, US National Museum of Natural History, and University of Kansas Museum of Natural History for providing samples of Passenger Pigeon skins. This work was supported by National Science Foundation Grants DEB-0107891 and DEB-0612938 to K.P.J. and DEB-0816877 to D.H.C.

References

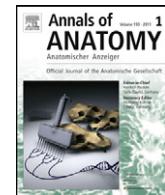
- Barboum, D.N., Burns, K.J., 2002. Phylogenetic relationships of the Wrentit based on mitochondrial cytochrome *b* sequences. *Condor* 104, 740–749.
- Barker, F.K., Cibois, A., Schikler, P., Feinstein, J., Cracraft, J., 2004. Phylogeny and diversification of the largest avian radiation. *PNAS* 101, 11040–11045.
- Blockstein, D.E., 2002. Passenger Pigeon (*Ectopistes migratorius*). In: Poole, A., Gill, F. (Eds.), *The Birds of North America*. No. 611. Philadelphia, PA.
- Brooks, D.R., 1990. Parsimony analysis in historical biogeography and coevolution: methodological and theoretical update. *Syst. Zool.* 39, 14–30.
- Bush, S.E., Price, R.D., Clayton, D.H., 2009. Description of eight species of *Columbicola* (Phthiraptera: Philopteridae), with a comprehensive world checklist. *J. Parasitol.* 95, 286–294.
- Clayton, D.H., Price, R.D., 1999. Taxonomy of New World *Columbicola* (Phthiraptera: Philopteridae) from the Columbiformes (Aves), with descriptions of five new species. *Ann. Entomol. Soc. Am.* 92, 675–685.
- DaCosta, J.M., Klicka, J., 2008. The Great American Interchange in birds: a phylogenetic perspective with the genus *Trogon*. *Mol. Ecol.* 17, 1328–1343.
- Dumbacher, J.P., Fleischer, R.C., 2001. Phylogenetic evidence for colour-pattern convergence in toxic pitohuis: Mullerian mimicry in birds? *Proc. R. Soc. Lond. B* 268, 1971–1976.
- Gibbs, D., Barnes, E., Cox, J., 2001. *Pigeons and Doves: A Guide to the Pigeons and Doves of the World*. Yale University Press, New Haven, CT.
- Goodwin, D., 1983. *Pigeons and Doves of the World*. Cornell University Press, Ithaca, NY.
- Huelsenbeck, J.P., 2001. MrBayes 1.10 (Bayesian Analysis of Phylogeny). University of Rochester, Rochester, NY.
- Johnson, K.P., 2004. Deletion bias in avian introns over evolutionary timescales. *Mol. Biol. Evol.* 21, 599–602.
- Johnson, K.P., Clayton, D.H., 2000. Nuclear and mitochondrial genes contain similar phylogenetic signal for pigeons and doves (Aves: Columbiformes). *Mol. Phylogenet. Evol.* 14, 141–151.
- Johnson, K.P., Reed, D.L., Hammond Parker, S.L., Kim, D., Clayton, D.H., 2007. Phylogenetic analysis of nuclear and mitochondrial genes supports species groups for *Columbicola* (Insecta: Phthiraptera). *Mol. Phylogenet. Evol.* 45, 506–518.
- Johnson, K.P., Sorenson, M.D., 1999. Phylogeny and biogeography of dabbling ducks (Genus: *Anas*): a comparison of molecular and morphological evidence. *Auk* 116, 792–805.
- Malcomson, R.O., 1937. Two new Mallophaga. *Annu. Rev. Entomol. Soc. Am.* 30, 53–56.
- Pereira, S.L., Johnson, K.P., Clayton, D.H., Baker, A.J., 2007. Mitochondrial and nuclear DNA sequences support a Cretaceous origin of Columbiformes and dispersal driven radiation in the Paleogene. *Syst. Biol.* 56, 656–672.
- Schorger, A.W., 1955. *The Passenger Pigeon: Its Natural History and Extinction*. University of Wisconsin Press, Madison, WI.
- Shapiro, B., Sibthorpe, D., Rambaut, A., Austin, J., Wragg, G.M., Bininda-Emonds, O.R.P., Lee, P.I.M., Cooper, A., 2002. Flight of the dodo. *Science* 295, 1683.
- Sorenson, M.D., Cooper, A., Paxinos, E.A., Quinn, T.W., James, H.F., Olson, S.L., Fleischer, R.C., 1999. Relationships of the extinct moa-nalos, flightless Hawaiian waterfowl, based on ancient DNA. *Proc. R. Soc. Lond. B* 266, 2187–2193.
- Stork, N.E., Lyal, C.H.C., 1993. Extinction or “co-extinction” rates? *Nature* 366, 307.



Contents lists available at ScienceDirect

Annals of Anatomy

journal homepage: www.elsevier.de/aanat



Research article

Ancient DNA from marine mammals: Studying long-lived species over ecological and evolutionary timescales

Andrew D. Foote^{a,*}, Michael Hofreiter^b, Phillip A. Morin^c

^a Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, ØsterVoldgade 5-7, DK-1350 Copenhagen, Denmark

^b Department of Biology, University of York, Wentworth Way, Heslington, York, YO10 5DD, UK

^c Protected Resources Division, Southwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 3333 N. Torrey Pines Ct., La Jolla, CA 92037, USA

ARTICLE INFO

Article history:

Received 14 January 2011

Received in revised form 14 April 2011

Accepted 15 April 2011

Keywords:

Ancient DNA

Distribution

DNA substitution rate

Ecology

Evolution

Marine mammals

Stable isotopes

SUMMARY

Marine mammals have long generation times and broad, difficult to sample distributions, which makes inferring evolutionary and demographic changes using field studies of extant populations challenging. However, molecular analyses from sub-fossil or historical materials of marine mammals such as bone, tooth, baleen, skin, fur, whiskers and scrimshaw using ancient DNA (aDNA) approaches provide an opportunity for investigating such changes over evolutionary and ecological timescales. Here, we review the application of aDNA techniques to the study of marine mammals. Most of the studies have focused on detecting changes in genetic diversity following periods of exploitation and environmental change. To date, these studies have shown that even small sample sizes can provide useful information on historical genetic diversity. Ancient DNA has also been used in investigations of changes in distribution and range of marine mammal species; we review these studies and discuss the limitations of such 'presence only' studies. Combining aDNA data with stable isotopes can provide further insights into changes in ecology and we review past studies and suggest future potential applications. We also discuss studies reconstructing inter- and intra-specific phylogenies from aDNA sequences and discuss how aDNA sequences could be used to estimate mutation rates. Finally, we highlight some of the problems of aDNA studies on marine mammals, such as obtaining sufficient sample sizes and calibrating for the marine reservoir effect when radiocarbon-dating such wide-ranging species.

© 2011 Published by Elsevier GmbH.

1. Introduction

Marine mammals can live for several decades (some even a century or more, [George and Bockstoce, 2008](#)) and have long generation times, a trait that can make studies on changes in genetic diversity through time in response to anthropogenic impacts (such as whaling) and environmental change challenging. Typically, such studies infer historical bottlenecks, founding events, population expansions or range shifts from the genetic signature found in extant populations (see [Hoelzel et al., 2002](#)). Ancient DNA (aDNA) studies, on the other hand, offer an opportunity for real-time studies of long-lived species over ecological and evolutionary timescales ([O'Corry-Crowe, 2008](#)). The genetic data from aDNA sequences allow tracking a variety of population parameters such as changes in genetic diversity ([de Bruyn et al., 2009](#)), past connectivity by using nuclear markers to look at changes in gene flow ([Valentine et al., 2008](#); [de Bruyn et al., 2009](#)) and past ecology by

combining stable isotope and molecular analysis (e.g. [Newsome et al., 2007](#); [Lindqvist et al., 2010](#)).

2. Sources of marine mammal aDNA

In this review we follow the example of [Navascués et al. \(2010\)](#) and define ancient DNA as DNA recovered post-mortem from non-ideal biological material. To date, aDNA studies have been heavily biased towards terrestrial species and in particular the abundant terrestrial megafauna species of the Pleistocene (e.g. [Hofreiter et al., 2004](#); [Shapiro et al., 2004](#); [Gilbert et al., 2008](#); [Campos et al., 2010](#)). This bias in the taxa being studied may reflect the relative availability of terrestrial versus marine samples, with sub-fossil remains of terrestrial mammals typically being far more accessible than sub-fossil remains of marine species. However, in locations where large numbers of marine mammal sub-fossil samples have been recovered they are often under-utilized for aDNA work. We suspect that this may arise in part from geneticists working on contemporary marine mammal species being unaware either of the availability of such remains or the potential insights gained by ancient DNA analyses. However, aDNA remains a relatively expensive and risky

* Corresponding author.

E-mail address: FooteAD@gmail.com (A.D. Foote).



Fig. 1. Mummified seal near Lake Bonney, Antarctica.

Photograph: Randy Stieletter.

line of research, especially when the data quality assurance and limitations to interpreting the data are considered.

Several thousand sub-fossil bones of terrestrial and marine mammals from the late Pleistocene and early Holocene have been recovered by trawling fishermen and suction-dredging from the Southern Bight of the North Sea (Mol et al., 2006). The marine specimens are thought to originate from inter-glacial periods when sea-level was high (Mol et al., 2006). For example, large numbers of samples of ice-associated or ice-dependent species such as walrus (*Odobenus rosmarus*), beluga (*Delphinapterus leucas*), and harp seal (*Pagophilus groenlandica*) dated to between 25,000 and 45,000 BP have been recovered (Post and Kompanje, 1995; Post, 1999; Mol et al., 2006). Similar findings have been made in southern Scandinavia (Aaris-Sørensen et al., 2010). Although the Southern Bight of the North Sea is exceptional in its sub-fossil record, there are sites throughout the sub-Arctic and Arctic with valuable zooarchaeological data that could also be used in aDNA studies (Dyke et al., 1996; Harington, 2003; Murray, 2008) including more recent middens (a mound or deposit of bones, shells and other refuse) from prehistoric aboriginal human populations (Etnier, 2004; Storå and Ericson, 2004; Newsome et al., 2007; Sommer et al., 2008) as well as animal middens (Magnanou et al., 2009). Additionally, bones found at former whaling stations and aboriginal whaling sites in both the Northern and Southern Hemispheres are a rich data source for aDNA studies (e.g. Rastogi et al., 2004; Lindqvist et al., 2009; Sremba et al., 2010).

Similar finds have also been made in Antarctica and the sub-Antarctic. Early explorers of the Antarctic continent were surprised to find the mummified remains of seals many kilometres inland and sometimes at high altitudes (Scott, 1905; Wilson, 1907; Fig. 1). Recent explorations have subsequently seen live seals travelling over 100 km inland and at an elevation close to 1000 m (Stirling and Rudolph, 1968). The mummified remains of Antarctic seals, which include the crabeater (*Lobodon carcinophagus*), leopard (*Hydrurga leptonyx*), southern elephant (*Mirounga leonina*) and Weddell (*Lepotnychotes weddellii*) seals, vary from twisted remains of skin to near perfect specimens (Péwé et al., 1959; Barwick and Balham, 1967; Dort, 1971; Gordon and Harkness, 1992; Banks et al., 2010). Although to date only one aDNA study of this material has been conducted to our knowledge (de Bruyn et al., 2009), the large number of samples available of the crabeater seal in particular (Péwé et al., 1959; Barwick and Balham, 1967) would be sufficient for several lines of investigation using aDNA studies.

The cold temperature of these high-latitude sites is key to the preservation of aDNA in samples such as sub-fossil tooth and bone or mummified skin over extended time scales. At low-latitude sites with warmer climates, DNA is expected to undergo more rapid post-mortem degradation and samples are therefore usually less suitable for aDNA studies (Hofreiter et al., 2001; Smith et al., 2003; Pääbo et al., 2004).

There is also great scope for applying aDNA methods to more recent samples due to the greater availability of historical samples, which are hundreds rather than thousands of years old. This allows investigation of the demographic and evolutionary history of taxa over the past few hundred years. The interest of marine mammals to taxonomists past and present has meant that many beach-stranded or bycaught cetaceans have been collected and maintained in collections at universities and museums. Several centuries of exploitation of marine mammals including cetaceans, pinnipeds, sirenians, sea otters and polar bears have also provided potential sources of DNA over a time series within which it is possible to address a range of demographic, ecological and evolutionary questions.

DNA has been extracted from historical samples of hair, teeth and bones of several taxa including the pelts of sea otters (Larson et al., 2002; Valentine et al., 2008), whale baleen (Rosenbaum et al., 1997) and mummified skin (Weber et al., 2000). Ethnographic material, which is found in large quantities in the anthropology collections of many museums such as engraved sperm whale (*Physeter macrocephalus*) teeth known as scrimshaw, or baleen used in aboriginal fishing line, nets and other utensils can also yield DNA (Pichler et al., 2001; Foote, unpublished data). The small amount of material needed for the extraction of sufficient DNA using improved extraction methods (e.g. Yang et al., 1998) has greatly reduced the destructiveness of this procedure on valuable museum specimens (Rosenbaum et al., 1997; Pichler et al., 2001; Cipriano and Pastene, 2009).

3. Markers for marine mammal aDNA studies

Mitochondrial DNA is the most widely used marker in ancient DNA studies, including those on marine mammals (see Table 1). As each cell contains hundreds to thousands of copies of the mitochondrial genome, compared to one diploid copy of the nuclear genome, the chances of amplification by the polymerase chain reaction are greatly increased (Binladen et al., 2006; Morin et al., 2007). The age and tissue type can have a dramatic effect on the proportion of mitochondrial to nuclear DNA preserved; e.g. for historical whale baleen the proportion is 40–60-fold higher than in historical bone and tooth samples or fresh preserved soft tissue (Morin et al., 2007).

Mitochondrial DNA is generally assumed to evolve under neutrality, and as such has been used for demographic inference. The diversity of neutrally evolving markers is expected to be proportional to the effective population size, and changes in diversity are therefore assumed to reflect demographic change (e.g. Harrison, 1989; Roman and Palumbi, 2003; Shapiro et al., 2004). However, recent studies have suggested that selective sweeps on the mtDNA could reduce genetic diversity and violate the assumption of neutrality (Bazin et al., 2006; Nabholz et al., 2009). There is evidence that changes in the protein coding genes of the mitochondrial genome of marine mammal species have evolved under positive selection (McClellan et al., 2005; da Fonseca et al., 2008; Foote et al., 2011). Individuals with these changes could be more successful and out-compete those without these changes, leading to a 'selective sweep' during which the variation across mitochondrial genomes without the adaptive mutation is eventually lost or driven to low frequency within the population. This process can result in

Table 1

Summary of ancient DNA studies on marine mammals.

Order	Species	Maximum age of samples (YBP)	Markers	References
Cetacea	Bowhead whale	51,000	mtDNA	Rastogi et al. (2004); Borge et al. (2007); McLeod et al. (2008, 2010)
	Bottlenose dolphin	1,400	mtDNA microsatellite	Nichols et al. (2007)
	Killer whale	<200	mtDNA	Morin et al. (2007); Foote et al. (2009)
	Longman's beaked whale	<100	mtDNA	Dalebout et al. (2003)
	North Atlantic right whale	<400	mtDNA	Rosenbaum et al. (1997, 2000); Rastogi et al. (2004); McLeod et al. (2008, 2010)
	Northern bottlenose whale	<100	SRY	Gowans et al. (2000)
Sirenia	Sperm whale	<100	mtDNA	Pichler et al. (2001)
	Steller's sea cow	>300	mtDNA	Rainey et al. (1984); Ozawa et al. (1997)
Carnivora				
Suborder Pinnipedia	Northern fur seal	2,425	mtDNA	Pinsky et al. (2010)
	Northern elephant seal	1,000	mtDNA	Weber et al. (2000)
	Guadalupe fur seal	>300	mtDNA	Weber et al. (2004)
	Southern elephant seal	7,087	mtDNA	de Bruyn et al. (2009)
	Walrus	<100	mtDNA	Lindqvist et al. (2009)
	Sea otter	<450	mtDNA, microsatellite	Larson et al. (2002); Valentine et al. (2008)
Suborder Fissipedia	Polar bear	>100,000	mtDNA	Lindqvist et al. (2010)

a genetic signature similar to a population bottleneck even if there has been no change in effective population size. It is therefore advisable to undertake statistical tests for neutrality of mutations against selective sweeps (e.g. Fu, 1997), or use two non-linked markers, when making inferences about demography using genetic diversity as a proxy (Ballard and Whitlock, 2004; Alter and Palumbi, 2009).

One of the most commonly used types of nuclear DNA marker in marine mammal population genetics of contemporary populations is the short-tandem-repeat or microsatellite. These are repetitive sequences of nucleotide pairs or triplets that vary in length depending upon the number of repeats. This variability has made them the marker of choice in contemporary population genetics. However, the length, low copy number, and difference in length between alleles of most microsatellite loci can make them difficult to amplify reliably. In particular, attempts to amplify microsatellite loci from degraded DNA samples can result in 'allelic dropout'. Multiple single-tube PCRs for the same loci for each sample can greatly reduce the bias from allelic drop out and misidentification of heterozygotes as homozygotes in degraded DNA samples (Taberlet et al., 1996; Allentoft et al., 2011). However, this greatly increases both the time and financial costs of the project and may explain why few marine mammal aDNA studies have utilized microsatellite loci. An alternative type of nuclear DNA marker to microsatellites are single nucleotide polymorphisms (SNPs) (Morin et al., 2004). The amplified fragment can therefore be much shorter than for microsatellite loci, reducing PCR failure rate, and each fragment can be the same length for all alleles, reducing the probability of allelic dropout (Morin and McCarthy, 2007). This principle also applies to other markers, such as Y-chromosomal markers, which are used to molecularly identify the sex of a specimen. Gowans et al. (2000) found that 1100 bp long ZFY-ZFX regions of the sex chromosomes were unsuitable for use on degraded samples of northern bottlenose whale (*Hyperoodon ampullatus*). They found that a short fragment (147 bp) of the Y-chromosome-specific SRY loci was more reliable and amplified more successfully in degraded and historic samples. Similarly, a real-time PCR assay targeting a 105 bp region of ZFY-ZFX has been shown to be effective for sex identification in most cetaceans and has high success rates for modern and historical samples (Morin et al., 2005, unpublished data).

4. Phylogenetic studies of marine mammals using aDNA

One of the earliest uses of aDNA sequences was for phylogenetic studies, whereby sequences of extinct species are included in molecular phylogenies to test taxonomic relationships that had previously been hypothesized based upon morphological characteristics alone. Although the first proliferation of sequences from extinct species in the 1980s and 1990s originated predominantly from terrestrial species such as the quagga (*Equus quagga quagga*; Higuchi et al., 1984), marsupial wolf (*Thylacinus cynocephalus*; Thomas et al., 1989), mammoth (*Mammuthus primigenius*; Höss et al., 1994; Hagelberg et al., 1995) and cave bear (*Ursus spelaeus*; Hänni et al., 1994), with Steller's sea cow they did include at least one marine mammal species. Steller's sea cow (*Hydrodamalis gigas*) was a giant Sireian first discovered in 1741 in the North Pacific, but following two decades of unsustainable harvest, is thought to have become extinct by 1768 (Turvey and Risley, 2006). The phylogenetic position of the Steller's sea cow within the Sirenia and the Tethytheria was determined first with immunological data (Rainey et al., 1984) and then by sequencing of the mitochondrial *cytochrome b* gene (Ozawa et al., 1997). These studies revealed that the divergence of the Steller's sea cow from the dugong (*Dugong dugon*) was almost as ancient (22 MYA) as the split between the manatees (*Trichechus* spp.) and the dugong. Prior to this study, it was thought that the sea cow and dugong had split more recently than the dugong and the manatees.

Ancient-DNA protocols have also been used on historical museum samples to identify the species of the sample and confirm taxonomic identification and status of extant species. For example, Longman's beaked whale (*Indopacetus pacificus*) had until recently rarely been observed in the wild due to its distribution being limited to a remote pelagic habitat. Prior to a study by Dalebout et al. (2003), only two specimens had been described. Phylogenetic analyses of short (130–409 bp) mitochondrial DNA (mtDNA) control region and *cytochrome b* fragments showed that the sequences obtained from four additional museum specimens grouped with the holotype of this species, thereby tripling the known number of specimens of this species. In cases where there are putative but cryptic subspecies/subtypes based on morphology or other phenotype data, molecular identification can be used to assign these museum specimens to subtype (see examples in the following

sections). In some cases putative cryptic subspecies have been shown to be geographic outliers rather than distinct taxonomic units. For example, the Laptev Sea walrus *O. rosmarus laptevi* is one of three described walrus subspecies. However, aDNA analysis combined with morphometrics did not support this taxonomic status, instead suggesting that the Laptev walrus was the westernmost population of the Pacific walrus *O. rosmarus divergens* (Linqvist et al., 2008).

There are important considerations when applying phylogenetic or population genetic analyses to aDNA datasets such as using appropriate methodology for utilizing heterochronous data (see Navascués et al., 2010 for a review) and assessing if apparent mutations are artefacts arising from post-mortem damage (see Hofreiter et al., 2001). These caveats are generally applicable to both terrestrial and marine datasets. However, a consideration relevant specifically to studies of marine species is that radiocarbon dating of marine-derived carbon is complicated by the so-called Marine Reservoir Effect (MRE). Marine-derived ^{14}C carbon has an extended mean residence time relative to terrestrial ^{14}C , meaning that the oceans are depleted in ^{14}C , which leads to an offset of ^{14}C age between marine- and terrestrial-derived carbon (Ascough et al., 2006). Additionally, local climatic and oceanographic variables lead to local variation in ^{14}C age within the marine environment (Ascough et al., 2006, 2009) and changes in the offset over time (Bondevik et al., 2006). Therefore, although samples may be collected from a localized area, the large ranges of many marine mammal species mean they are likely to have integrated the reservoir age from a number of water masses (Mangerud et al., 2007). The only study that we could find which considered these factors in detail used whales caught at known dates to calibrate pelagic ^{14}C reservoir ages, assuming the measured value from the whales to be a mean for the waters within their range (Mangerud et al., 2007). This problem becomes important when samples used for ancient DNA studies are dated to put them in a temporal context and even more so if the ages of the sequences are used to estimate the nucleotide substitution rate of the genetic locus investigated (e.g. Ho et al., 2007, 2011).

5. Estimating changes in genetic diversity due to anthropogenic exploitation

Using museum collections for aDNA analysis has the potential to be informative for species conservation and management (Etnier, 2004; Leonard, 2008). Ancient DNA studies can help establish baselines of pre-exploitation or pre-environmental degradation of genetic diversity and geographic range. Given sufficient samples along a timeline and across a geographic range, aDNA studies allow reconstructing responses to past environmental change through the investigation of changes in genetic variation across space and time (Roy et al., 1994; Leonard, 2008). In this section we review case studies that applied ancient DNA techniques to obtain data that were informative for the management of extant marine mammal populations.

Several studies on marine mammal species have investigated changes in genetic diversity using historical or ancient samples. A number of these studies found changes in genetic diversity during periods of high exploitation. For example, Weber et al. (2004) found, based on historical samples, that the Guadalupe fur seal (*Arctocephalus townsendi*) shows a loss of genetic diversity in mtDNA genotypes associated with 18th and 19th century commercial sealing, consistent with a signature of a genetic bottleneck previously found in modern samples from extant populations (Bernardi et al., 1998). Similarly, the northern elephant seal (*Mirounga angustirostris*) is known from census data to have undergone a severe bottleneck due to 18th and 19th century commercial sealing, but

subsequently rebounded from a low of approximately 20 individuals in 1892 to over 175,000 individuals today (Hoelzel, 1999). Impacts of this population bottleneck include low variation at all genetic markers analysed to date and a loss of fitness as inferred by increased fluctuating asymmetry and variability of quantitative traits in post-bottleneck skulls compared with pre-bottleneck ones (Hoelzel, 1999). Two separate studies found only two mtDNA genotypes in samples from over 150 post-bottleneck individuals, collected from colonies in central and southern California (Hoelzel et al., 1993; Weber et al., 2000). In contrast, four mtDNA genotypes were found in just five pre-bottleneck samples, indicating much greater genetic diversity before commercial harvesting (Weber et al., 2000). The aDNA analysis of just a few samples in these case studies provided confirmation of the loss of mtDNA diversity due to a population bottleneck, which had previously only been inferred from the genetic signature from modern populations.

Another pinniped species, the northern fur seal (*Callorhinus ursinus*), appears to have been more robust against a loss of genetic diversity over a prolonged period of exploitation despite losing most of its former breeding range (Pinsky et al., 2010). Using sub-fossil samples of harvested specimens collected from archaeological sites along the entire US Pacific coast, Newsome et al. (2007) used stable isotope analysis to show that both northern and southern sites contained unweaned pups and therefore were likely breeding colonies. Further, they found differences in life history strategies, based on isotopic values, between the northern and southern sites, with pups being nursed for up to three times as long in northern regions (Newsome et al., 2007). The contraction of the breeding range to the more northerly sites was concurrent with the loss of the alternative southern life history strategy. It could thus be expected to have been concurrent with a loss of genetic diversity. However, genetic diversity appears to have been maintained despite this loss of breeding range, and the life history variation likely resulted from behavioural plasticity (Pinsky et al., 2010), although the study could not exclude a potential loss of allelic variation at nuclear loci that may be responsible for the original behavioural variation. A lack of population structure and high dispersal rate, plus the use of Arctic refugia are all hypothesized to have led to the resilience in this species against loss of genetic diversity, at least with regard to mitochondrial DNA (Pinsky et al., 2010).

The sea otter (*Enhydra lutris*), once widely spread around the North Pacific rim, was also commercially hunted during the 18th, 19th and 20th centuries until a ban on hunting in 1911. Analyses of both mtDNA and nuclear microsatellites of one pre-fur-trade and five modern populations found substantial loss of diversity in the microsatellite alleles (Larson et al., 2002). The authors of this study suggested that this loss of genetic diversity could result in inbreeding depression. Following the extinction of numerous otter populations, there have been attempts to reintroduce sea otters from the Aleutian Islands to Alaska, British Columbia, Washington, Oregon and California during the past century. While the northern reintroductions have been successful, the Oregon reintroductions did not become established (Valentine et al., 2008). Analysis of the mtDNA control region of otter samples dated to between 175 and 2000 years old from two locations in Oregon indicated that the pre-fur-trade otter populations along the Oregon coast grouped genetically with the native California populations (subspecies *E. lutris nereis*) rather than the Alaskan populations (subspecies *E. lutris kenyoni*) that had been used to restock the area (Valentine et al., 2008). The historical Oregon sea otters were also morphologically more similar to the southern subspecies (Lyman, 1988) and these morphological differences between the two subspecies may, at least partially, be adaptive. Valentine et al. (2008) concluded that re-stocking from the California populations, which may be more adapted to the Oregon environment, could be more

successful than previous relocations of otters from Alaska to Oregon.

Similar studies to investigate the loss of genetic diversity due to whaling have been conducted on historical samples of two of the most heavily exploited baleen whale species, the bowhead whale (*Balaena mysticetus*) and the North Atlantic right whale (*Eubalaena glacialis*). Both species were thought to have been equally targeted by Basque whalers during the 16th and 17th centuries. However, two recent studies that conducted molecular analyses on bones collected from a comprehensive search of 16th and 17th century Basque whaling ports found that all but one bone, including those previously identified as from right whales based on osteology, were molecularly identified as being from bowhead whales (Rastogi et al., 2004; McLeod et al., 2008). The authors concluded that the bowhead whale, rather than the North Atlantic right whale, was the main target of Basque whalers. Therefore, it is expected that, if at all, bowhead whales but not northern right whales may have undergone a loss of genetic diversity from Basque whaling (McLeod et al., 2008). In line with these results, genotyping of the single right whale bone identified by McLeod et al. (2008) using 27 microsatellite loci and comparison with the alleles for these loci in the extant population also suggests little or no loss of diversity (McLeod et al., 2010). All the alleles present in the historical sample were found in the extant population, and heterozygosity was similar for the historical bone and the modern samples (McLeod et al., 2010). Comparison of mtDNA diversity of late 19th and early 20th century museum specimens with samples from 84% of the extant population also suggests that loss of genetic diversity due to modern 20th century whaling has been modest (Rosenbaum et al., 2000). Although bowhead whales appear to have been the primary target of Basque whalers, Borge et al. (2007) found little change in levels of genetic diversity of the mtDNA throughout the Holocene. As in the case of the northern fur seal, the lack of population structure and the high dispersal rate in bowhead whales (McLeod, 2008) may have helped reduce the loss of diversity.

6. Investigating genetic change during periods of environmental change

Loss of genetic diversity can also be caused by habitat loss and environmental change. Ancient-DNA studies can be applied to look at past responses to climatic variation and to help understand how different taxa may respond to ongoing directional climate change (e.g. Shapiro et al., 2004; Drummond et al., 2005; Campos et al., 2010). A recent study used DNA extracted from mummified southern elephant seal skins found on a beach in the Antarctic to reconstruct the demographic history of this former breeding colony (de Bruyn et al., 2009). Comparing the mtDNA haplotypes with those for other locations suggested that this Antarctic breeding colony had been founded by ancestors from the sub-Antarctic island of Macquarie during the retreat of glaciers and sea ice at the Ross Sea Embayment 7500–8000 YBP. Following this founding event, changes in diversity suggest that the newly established breeding colony rapidly expanded and became independent from other colonies. Interestingly, the authors detected a signal of one-way migration back from the Antarctic colony to the sub-Antarctic colony at Macquarie Island some 1300 YBP, when the glaciers and sea ice subsequently started to re-encroach on this breeding colony. However, despite this back migration, much of the new genetic diversity generated by the founding event and subsequent expansion was lost (de Bruyn et al., 2009). In addition to the demographic insights, this case study also provides a useful insight into how mobile species with large ranges may be able to adapt to current changes in climate by habitat tracking (see next section).

In terrestrial mammals there is little evidence of habitat tracking during past episodes of climate change. Instead, phylogeographic studies tracking changes in haplotype distribution over time using aDNA and comparisons of pre- and post-glacial genetic diversity suggest that only a few relict populations of formerly widespread species survived in refugia during glaciations (Hofreiter et al., 2004; Dalén et al., 2007; Hofreiter and Barnes, 2010). However, in contrast to terrestrial habitats, the low cost of movement (Williams, 1999) and few geographic barriers may facilitate habitat tracking in the marine environment during environmental change. Ancient-DNA studies using samples from across large spatial and temporal scales, such as those from the North Sea, which was covered by seasonal sea ice during the last ice age, offer the opportunity to investigate the influence of past glaciations on the phylogeography of extant populations of marine species. Unfortunately, despite samples being available, with the exception of the study by de Bruyn et al. (2009), no data on this question have so far been published to the best of our knowledge. However, such studies could provide useful insights into the potential for marine mammal species to respond to future climate change.

Studies have also used sequences from historical samples to infer more recent range shifts, contractions or extinction of local populations. Stranding data for marine species offer a long-term measure of relative occurrence over time (Hart et al., 2006), although such presence-only data cannot conclusively demonstrate the absence of a species (Tingley and Beissinger, 2009). However, stranding data combined with molecular techniques can provide important insights into historical ranges. For example, Morin et al. (2006) sequenced a diagnostic fragment of the mtDNA control region from killer whale (*Orcinus orca*) museum specimens that had stranded or were harpooned along the Pacific coast of North America during the past two centuries. As mtDNA haplotypes are fixed in many killer whale populations in this region, the sequence was to some extent informative of the population of origin. The distribution of haplotypes was mostly consistent with the range of current populations. However, a haplotype associated with a population from British Columbia was found in California (Morin et al., 2006).

Nichols et al. (2007) analysed bottlenose dolphin (*Tursiops truncatus*) bones recovered from a Saxon settlement on the Humber River estuary in England and compared a short mtDNA sequence and microsatellite alleles with those from current populations. Both the mtDNA and microsatellite data suggested that the Humber samples were from a population distinct from the extant ones that had been sampled in waters around the UK or neighbouring waters. The authors concluded that this result was best explained by local population extinction during the past few hundred years (Nichols et al., 2007). They argued that these findings may indicate a more general decline of a bottlenose dolphin meta-population. The excellent collections at UK and other European museums offer the possibility to measure the past genetic diversity in this species to further investigate this possible decline.

7. Investigating ecology and evolution

Further ecological and evolutionary inferences can be made when aDNA studies are combined with those of stable isotopes and/or morphometric measurements (e.g. Leonard et al., 2007; Richards et al., 2008). Phylogenetic reconstruction of sequences obtained from historical museum specimens using aDNA protocols demonstrated that killer whale (*O. orca*) specimens from the Northeast Atlantic belong to two distinct clades (Foote et al., 2009). The two lineages also differed in body length, tooth count, niche width based on Nitrogen N¹⁵ isotopic ratios, and tooth-wear (Fig. 2). Foote et al. (2009) suggested that the types were genetically,

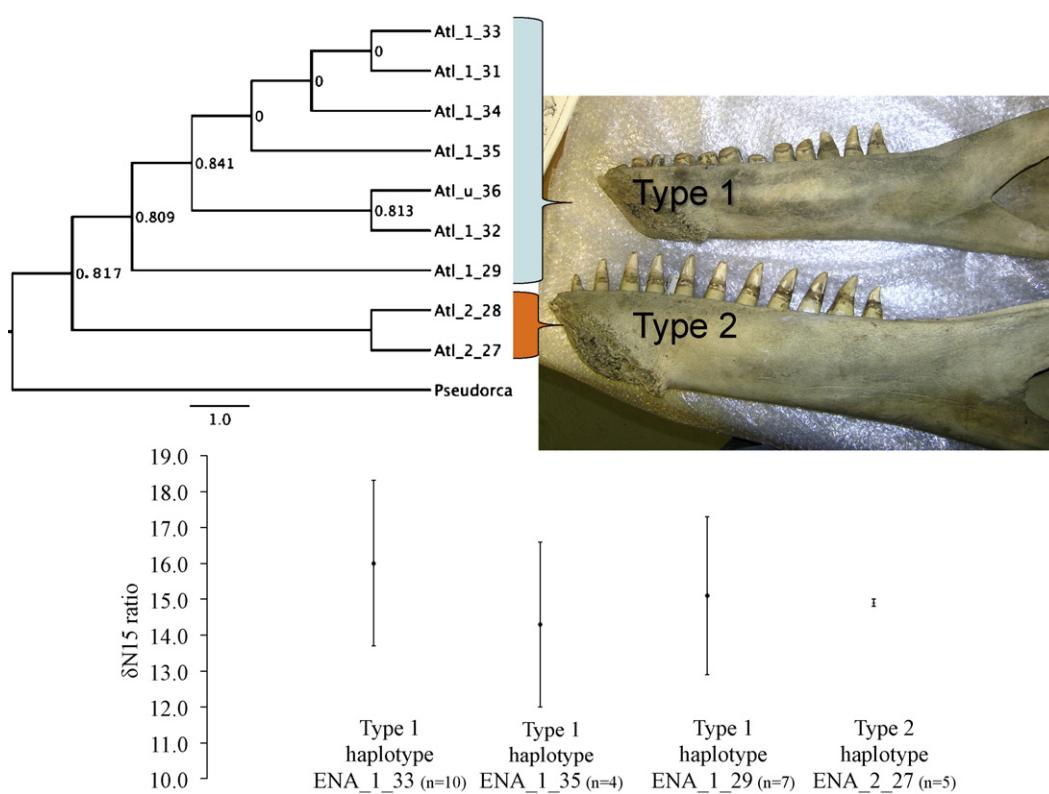


Fig. 2. Differences in apical tooth wear, tooth count, and niche width based on $\delta^{15}\text{N}$ stable isotope values between two sympatric, but genetically divergent lineages (type 1 and type 2) of North Atlantic killer whale (adapted from Foote et al., 2009).

morphologically and ecologically disparate enough that they should be considered separate 'Evolutionary Significant Units' (ESUs; see Ryder, 1986).

Using high-throughput sequencing techniques to generate the complete mitogenome of a polar bear jaw dated to between 110,000 and 130,000 years old, Lindqvist et al. (2010) were able to show that this sample was phylogenetically positioned at the branching point between polar bears and brown bears. By incorporating stable isotope measurements to infer diet, they could furthermore show that polar bears were hunting on the Arctic sea ice already at this early stage of their evolution. These examples show that there is indeed added scientific value in combining ancient DNA with stable isotope studies.

Molecular evolution can also be investigated using aDNA datasets. This has been applied to at least one marine mammal species. Evolutionary rates can be estimated using phylogenies, pedigree data and from aDNA sequences of C^{14} -dated samples. However, recent studies have found order-of-magnitude differences between short-term (<1 MY) mutation rates estimated using pedigree-based or aDNA data and long-term substitution rates estimated using species-level phylogeny-based data (Parsons et al., 1997; Lambert et al., 2002; Santos et al., 2005; Ho et al., 2007). For example, the short-term mutation rate of the mtDNA control region of the bowhead whale estimated using the aDNA dataset of Borge et al. (2007) is approximately 2×10^{-7} substitutions per site per year (Ho et al., 2007), compared with a long-term phylogeny-based substitution rate of approximately 2×10^{-8} (Rooney et al., 2001). However, subsequent data randomization analyses suggest there may not be sufficient temporal information to support rates estimated from the bowhead whale dataset (Ho et al., 2011). Given the relatively large number and the temporal range of samples sequenced in the bowhead whale aDNA study by Borge et al. (2007), it may be that very large aDNA datasets, spanning timescales of

10,000s of years are needed to reliably estimate evolutionary rates of species with slow molecular clocks such as baleen whales.

Ho et al. (2005) suggested that this time dependency of mutation rates is due to short-term transient mutations that will in the long-term eventually be removed by processes such as purifying selection. Therefore, genetic studies need to use an appropriate substitution rate for the timescales over which an analysis is conducted (Ho et al., 2005; Henn et al., 2009). For example, recent estimates of historical pre-whaling population sizes of minke (*Balaenoptera acutorostrata*), fin (*Balaenoptera physalus*) and humpback whale (*Megaptera novaeangliae*) based on current genetic diversity were an order-of-magnitude greater than previously estimated based on whaling catch records (Roman and Palumbi, 2003). However, these analyses used phylogenetic-based estimates of the mutation rate, which, as noted above, are likely to be an order-of-magnitude lower than mutation rates over the timescales for which they were estimating demographic change. Mutation rates estimated from moderate sized datasets of aDNA sequences (e.g. Ho et al., 2007) would be more appropriate for such studies investigating change over short timescales. This demonstrates both another important use for aDNA sequences and an important consideration for those using aDNA sequences to investigate phylogenetics or demographic change.

Genomic studies on marine mammals are just beginning to emerge. For example, recent studies have utilized mitogenomics of a global dataset of killer whales generated by high-throughput sequencing methods to test different evolutionary questions on divergences and adaptation (Morin et al., 2010; Foote et al., 2011). The first complete genome of a marine mammal, the bottlenose dolphin, has been sequenced, albeit currently at low coverage. Even so, the field of contemporary marine mammal genomics has a lot of catching up to do with the burgeoning field of palaeogenomics (Hofreiter, 2008). Complete mitogenome sequences of ancient terrestrial species are becoming relatively common, and

mitogenome phylogenies have been produced for several Pleistocene taxa (Gilbert et al., 2008; Willerslev et al., 2009; Stiller et al., 2009; Briggs et al., 2009), as have several nuclear genomic investigations (Noonan et al., 2005; Poinar et al., 2006; Green et al., 2006; Miller et al., 2008). Recently the first complete nuclear genome sequences of ancient hominids were published (Green et al., 2010; Rasmussen et al., 2010; Reich et al., 2010). With new protocols and ever-greater sequencing power becoming available, palaeogenomics may allow real-time tracking of evolutionary changes due to selection at the molecular level for marine mammals. However, the long generation times of many marine mammal species mean that for evolutionary change to be detectable, very strong selection would have had to have led to rapid change over relatively few generations. It would also only be possible for species for which large sample sizes and time series of specimens were available.

8. Summary

The purpose of this review was to highlight the variety and importance of aDNA studies that have been and could in future be applied to marine mammals. We have shown that there are sample sets available that would allow the type of population genetics investigations that have become increasingly common for terrestrial species. A number of studies have gained great insights by sequencing the mtDNA from only a small number of samples, and recently the first larger scale studies have started to appear. We predict that these studies will become increasingly common in the near future, leading to greater insights into past migrations, demographics and generally population developments of marine mammals, as testified by the few studies already published. These studies have proven particularly valuable when combined with comparisons with contemporary populations. Ancient-DNA analyses can provide a means of validating hypotheses based on the genetic 'footprint' of past demographic or evolutionary events in contemporary populations by investigating real time changes in populations from the time period under investigation. The rapid evolution of high-throughput sequencing technology is also facilitating the first wave of paleogenomic studies, which we feel certain will soon include marine mammal species.

Acknowledgments

AF was supported by the Lundbeck Foundation, a Marie Curie Intra-European Fellowship and the Centre for GeoGenetics, Natural History Museum, University of Copenhagen.

PAM was supported by the Southwest Fisheries Science Center.

MH was supported by the University of York.

The authors are grateful to Bill Perrin and Barb Taylor and 2 anonymous reviewers for helpful comments on the manuscript.

References

- Araris-Sørensen, K., Rasmussen, K.L., Kinze, C., Petersen, K.S., 2010. Late Pleistocene and Holocene whale remains (Cetacea) from Denmark and adjacent countries: species, distribution, chronology, and trace element concentrations. *Mar. Mamm. Sci.* 26, 253–281.
- Allentoft, M.E., Oskam, C., Houston, J., Hale, M.L., Gilbert, M.T.P., Rasmussen, M., Spencer, P., Jacob, C., Willerslev, E., Holdaway, R.N., Bunce, M., 2011. Profiling the dead: generating microsatellite data from fossil bones of extinct megafauna—protocols, problems, and prospects. *PLoS One* 6, e16670.
- Alter, S.E., Palumbi, S.R., 2009. Comparing evolutionary patterns and variability in the mitochondrial control region and *cytochrome b* in three species of baleen whales. *J. Mol. Evol.* 68, 97–111.
- Ascough, P.L., Cook, G.T., Church, M.J., Dugmore, A.J., Arge, S.V., McGovern, T.H., 2006. Variability in North Atlantic marine radiocarbon reservoir effects at c. AD 1000–Holocene 16, 131–136.
- Ascough, P.L., Cook, G.T., Dugmore, A.J., 2009. North Atlantic marine ^{14}C reservoir effects: Implications for late-Holocene chronological studies. *Quat. Geochronol.* 4, 171–180.
- Ballard, J.W.O., Whitlock, M.C., 2004. The incomplete natural history of mitochondria. *Mol. Ecol.* 13, 729–744.
- Banks, J.C., Ross, P.M., Smith, T.E., 2010. Report of a mummified leopard seal carcass in the southern Dry Valleys, McMurdo Sound, Antarctica. *Antarct. Sci.* 22, 43–44.
- Barwick, R.E., Balham, R.W., 1967. Mummified seal carcasses in a deglaciated region of South Victoria Land, Antarctica. *Tuatara* 15, 165–180.
- Bazin, E., Glemin, S., Galtier, N., 2006. Population size does not influence mitochondrial genetic diversity in animals. *Science* 312, 570–572.
- Bernardi, G., Fain, S.R., Gallo-Reynoso, J.P., Figueiro-Carranza, A.I., Le Boeuf, B.J., 1998. Genetic variability in Guadalupe fur seals. *J. Hered.* 89, 301–305.
- Binladen, J., Wiuf, C., Gilbert, M.T.P., Bunce, M., Barnett, R., Larson, G., Greenwood, A.D., Haile, J., Ho, S.Y.W., Hansen, A.J., Willerslev, E., 2006. Assessing the fidelity of ancient DNA sequences amplified from nuclear genes. *Genetics* 172, 733–741.
- Bondevik, S., Mangerud, J., Birks, H.H., Gulliksen, S., Reimann, P., 2006. Changes in North Atlantic radiocarbon reservoir ages during the Allerød and younger Dryas. *Science* 312, 1514–1517.
- Borge, T., Bachman, L., Bjørnstad, G., Wiig, Ø., 2007. Genetic variation in Holocene bowhead whales from Svalbard. *Mol. Ecol.* 16, 2223–2235.
- Briggs, A.W., Good, J.M., Green, R.E., Krause, J., Maricic, T., Stenzel, U., Lalueza-Fox, C., Rudan, P., Brajkovic, D., Kucan, Z., Gusic, I., Schmitz, R., Doronichev, V.B., Golovanova, L.V., de la Rasilla, M., Fortea, J., Rosas, A., Pääbo, S., 2009. Targeted retrieval and analysis of five Neanderthal mtDNA genomes. *Science* 325, 318–321.
- de Bruyn, M., Hall, B.L., Chauke, L.F., Baroni, C., Koch, P.L., Hoelzel, A.R., 2009. Rapid response of a marine mammal species to Holocene climate and habitat change. *PLoS Genet.* 5, e1000554.
- Campos, P.F., et al., 2010. Ancient DNA analyses exclude humans as the driving force behind late Pleistocene musk ox (*Ovibos moschatus*) population dynamics. *Proc. Natl. Acad. Sci. U.S.A.* 107, 5675–5680.
- Cipriano, F., Pastene, L., 2009. A review of current knowledge of techniques to extract and amplify DNA from 'difficult' whale samples. Report to Scientific Committee of the International Whaling Commission SC/61/SD2.
- Dalebout, M.L., Ross, G.J.B., Baker, C.S., Anderson, R.C., Best, P.B., Cockcroft, V.G., Hinsz, H.L., Peddermors, V., Pitman, R.L., 2003. Appearance, distribution, and genetic distinctiveness of Longman's beaked whale, *Indopacetus pacificus*. *Mar. Mamm. Sci.* 19, 421–461.
- Dalén, L., Nyström, V., Valdiosera, C., Germonpré, M., Sablin, M., Turner, E., Angerbjörn, A., Arsaga, J.-L., Götherstrom, A., 2007. Ancient DNA reveals lack of postglacial habitat tracking in the arctic fox. *Proc. Natl. Acad. Sci. U.S.A.* 104, 6726–6729.
- Dort, W., 1971. Mummified seals of southern Victoria Land. *Antarct. J. U.S.* 6, 210–211.
- Drummond, A.J., Rambaut, A., Shapiro, B., Pybus, O.G., 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol. Biol. Evol.* 22, 1185–1192.
- Dyke, A.S., Hooper, J., Savelle, J.M., 1996. A history of sea ice in the Canadian Arctic archipelago based on postglacial remains of the bowhead whale (*Balaena mysticetus*). *Arctic* 49, 235–255.
- Etnier, M.A., 2004. The potential of zooarchaeological data to guide pinnipeds management decisions in the eastern North Pacific. In: Lyman, R.L., Cannon, K.P. (Eds.), *Zooarchaeology and Conservation Biology*. University of Utah Press, Salt Lake City, pp. 88–102.
- Foote, A.D., Morin, P.A., Durban, J.W., Pitman, R.L., Wade, P., Willerslev, E., Gilbert, M.T.P., da Fonseca, R.R., 2011. Positive selection on the killer whale mitogenome. *Biol. Lett.* 7, 116–118.
- Foote, A.D., Newton, J., Piertney, S.B., Willerslev, E., Gilbert, M.T.P., 2009. Ecological, morphological and genetic divergence of sympatric North Atlantic killer whale populations. *Mol. Ecol.* 18, 5207–5217.
- da Fonseca, R., Johnson, W., O'Brien, S., Ramos, M., Antunes, A., 2008. The adaptive evolution of the mammalian mitochondrial genome. *BMC Genomics* 9, 119.
- Fu, Y.X., 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147, 915–925.
- George, J.C., Bockstoce, J.R., 2008. Two historical weapon fragments as an aid to estimating the longevity and movement of bowhead whales. *Polar Biol.* 31, 751–754.
- Gilbert, M.T.P., et al., 2008. Intraspecific phylogenetic analysis of Siberian woolly mammoths using complete mitochondrial genomes. *Proc. Natl. Acad. Sci. U.S.A.* 105, 8327–8332.
- Gordon, J.E., Harkness, D.D., 1992. Magnitude and geographic variation of the radiocarbon content in Antarctic marine life: implications for reservoir corrections in radiocarbon dating. *Quaternary Sci. Rev.* 7–8, 697–708.
- Gowans, S., Dalebout, M.L., Hooker, S.K., Whitehead, H., 2000. Reliability of photographic and molecular techniques for sexing northern bottlenose whales (*Hyperoodon ampullatus*). *Can. J. Zool.* 78, 1224–1229.
- Green, R.E., Krause, J., Ptak, S.E., Briggs, A.W., Ronan, M.T., Simons, J.F., Du, L., Egholm, M., Rothberg, J.M., Paunovic, M., Pääbo, S., 2006. Analysis of one million base pairs of Neanderthal DNA. *Nature* 444, 330–336.
- Green, R.E., et al., 2010. A draft sequence of the Neanderthal genome. *Science* 328, 710–722.
- Hagelberg, E., Thomas, M.G., Cook, C.E.J., Sher, A.V., Baryshnikov, G.F., Lister, A.M., 1995. DNA from ancient mammal bones. *Nature* 370, 333–334.
- Hänni, C., Laudet, V., Stehelin, D., Taberlet, P., 1994. Tracking the origins of the cave bear (*Ursus spelaeus*) by mitochondrial DNA sequencing. *Proc. Natl. Acad. Sci. U.S.A.* 91, 12336–12340.
- Harington, C.R., 2003. *Annotated Bibliography of the Quaternary Vertebrates of Northern North America*. University of Toronto Press.

- Harrison, R.G., 1989. Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. *Trends Ecol. Evol.* 4, 6–11.
- Hart, K.M., Moorside, P., Crowder, L.B., 2006. Interpreting the spatio-temporal patterns of sea turtle strandings: going with the flow. *Biol. Cons.* 129, 283–290.
- Henn, B.M., Gignoux, C.R., Feldman, M.W., Mountain, J.L., 2009. Characterizing the time dependency of human mitochondrial DNA mutation rate estimates. *Mol. Evol. Biol.* 26, 217–230.
- Higuchi, R., Bowman, B., Freiberger, M., Ryder, O.A., Wilson, A.C., 1984. DNA sequences from the quagga, an extinct member of the horse family. *Nature* 312, 282–284.
- Ho, S.Y.W., Kolokotronis, S.-O., Allaby, R.G., 2007. Elevated substitution rates estimated from ancient DNA sequences. *Biol. Lett.* 3, 702–705.
- Ho, S.Y.W., Lanfear, R., Phillips, M.J., Barnes, I., Thomas, J.A., Kolokotronis, S.-O., Shapiro, B., 2011. Bayesian estimation of substitution rates from ancient DNA sequences with low information content. *Syst. Biol.* doi:10.1093/sysbio/syq099.
- Ho, S.Y.M., Phillips, M.J., Cooper, A., Drummond, A.J., 2005. Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Mol. Biol. Evol.* 22, 1561–1568.
- Hoelzel, A.R., 1999. Impact of population bottlenecks on genetic variation and the importance of life-history: a case study of the northern elephant seal. *Biol. J. Linnean Soc.* 68, 23–39.
- Hoelzel, A.R., Goldsworthy, S.D., Fleischer, R.C., 2002. Population genetic structure. In: Hoelzel, A.R. (Ed.), *Marine Mammal Biology: an Evolutionary Approach*. Blackwell Science, Oxford, pp. 325–352.
- Hoelzel, A.R., Hally, J., O'Brien, S.J., Campagna, C., Arnborn, T., Leboeuf, B., Ralls, K., Dover, G.A., 1993. Elephant seal genetic-variation and the use of simulation-models to investigate historical population bottlenecks. *J. Hered.* 84, 443–449.
- Hofreiter, M., 2008. Palaeogenomics. *C. R. Palevol.* 7, 113–124.
- Hofreiter, M., Barnes, I., 2010. Diversity lost: are all Holarctic large mammal species just relict populations? *BMC Biol.* 8, 46.
- Hofreiter, M., Serre, D., Poinar, H.N., Kuch, M., Pääbo, S., 2001. Ancient DNA. *Nat. Rev.* 2, 353–359.
- Hofreiter, M., Serre, D., Rohland, N., Rabeder, G., Nagel, D., Conard, N., Münzel, S., Pääbo, S., 2004. Lack of phylogeography in European mammals before the last glaciation. *Proc. Natl. Acad. Sci. U.S.A.* 101, 12963–12968.
- Höss, M., Pääbo, S., Vereschagin, N.K., 1994. Mammoth DNA sequences. *Nature* 370, 333.
- Lambert, D.M., Ritchie, P.A., Millar, C.D., Holland, B., Drummond, A.J., Baroni, C., 2002. Rates of evolution in ancient DNA from Adélie penguins. *Science* 295, 2270–2273.
- Larson, S., Jameson, R., Etnier, M., Flemings, M., Bentzen, P., 2002. Loss of genetic diversity in sea otters (*Enhydra lutris*) associated with the fur trade of the 18th and 19th centuries. *Mol. Ecol.* 11, 1899–1903.
- Leonard, J.A., 2008. Ancient DNA applications for wildlife conservation. *Mol. Ecol.* 17, 4186–4196.
- Leonard, J.A., Vilà, C., Fox-Dobbs, K., Koch, P.L., Wayne, R.K., van Valkenburgh, B., 2007. Megafaunal extinctions and the disappearance of a specialized wolf ecomorph. *Curr. Biol.* 17, 1146–1150.
- Linqvist, C., Bachmann, L., Andersen, L.W., Born, E.W., Arnason, U., Kovacs, K.M., Lydersen, C., Abramov, A.V., Wiig, Ø., 2008. The Laptev Sea walrus *Odobenus rosmarus laptevi*: an enigma revisited. *Zoolog. Scripta* 38, 113–127.
- Lindqvist, C., et al., 2010. Complete mitochondrial genome of a Pleistocene jawbone unveils the origin of the polar bear. *Proc. Natl. Acad. Sci. U.S.A.* 107, 5053–5057.
- Lindqvist, C., Probst, A., Martin, A.R., Wiig, Ø., Bachmann, L., 2009. Molecular species identification of historical whale remains from South Georgia. *Mar. Mamm. Sci.* 25, 229–238.
- Lyman, R.L., 1988. Zoogeography of Oregon coast marine mammals: the last 3,000 years. *Mar. Mamm. Sci.* 4, 247–264.
- Magnanou, E., Malenke, J.R., Dearing, M.D., 2009. Expression of biotransformation genes in woodrat (*Neotoma*) herbivores on novel and ancestral diets: identification of candidate genes responsible for dietary shifts. *Mol. Ecol.* 18, 2401–2414.
- Mangerud, J., Bondevik, S., Gulliksen, S., Hufthammer, A.K., Hoisaeter, T., 2007. Marine C-14 reservoir ages for 19th century whales and molluscs from the North Atlantic. *Quaternary Sci. Rev.* 25, 3228–3245.
- McClellan, D.A., Palfreyman, E.J., Smith, M.J., Moss, J.L., Christensen, R.G., Sailsbury, J.K., 2005. Physicochemical evolution and molecular adaptation of the cetacean and artiodactyl *cytochrome b* proteins. *Mol. Biol. Evol.* 22, 437–455.
- McLeod, B.A., 2008. Historic Levels of Diversity in the North Atlantic Right (*Eubalaena glacialis*) and Bowhead Whale (*Balaena mysticetus*). PhD Trent University, Peterborough, Ontario, Canada, pp. 185.
- McLeod, B.A., Brown, M.W., Moore, M.J., Stevens, W., Barkham, S.H., White, B.N., 2008. Bowhead whales, and not right whales, were the primary target of 16th–17th-century Basque whalers in the western North Atlantic. *Arctic* 61, 61–75.
- McLeod, B.A., Brown, M.W., Frasier, T.R., White, B.N., 2010. DNA profile of a sixteenth century western North Atlantic right whale (*Eubalaena glacialis*). *Cons. Genet.* 11, 339–345.
- Miller, W., Drautz, D.I., Ratan, A., Pusey, B., Qi, J., Lesk, A.M., Tomsho, L.P., Packard, M.D., Zhao, F., Sher, A., Tikhonov, A., Raney, B., Patterson, N., Lindblad-Toh, K., Lander, E.S., Knight, J.R., Irzyk, G.P., Fredrikson, K.M., Harkins, T.T., Sheridan, S., Pringle, T., Schuster, S.C., 2008. Sequencing the nuclear genome of the extinct woolly mammoth. *Nature* 456, 387–390.
- Mol, D., Post, K., Reumer, J.W.F., van der Plicht, J., de Vos, J., van Geel, B., van Reenan, G., Pals, J.P., Glimmer veen, J., 2006. The Eurogeul-first report of the palaeonto-logical, palynological and archaeological investigations of this part of the North Sea. *Quaternary Int.* 142–143, 178–185.
- Morin, P.A., et al., 2010. Complete mitochondrial genome phylogeographic analysis of killer whales (*Orcinus orca*) indicates multiple species. *Genome Res.* 20, 908–916.
- Morin, P.A., Hedrick, N.M., Robertson, K.M., LeDuc, C.A., 2007. Comparative mitochondrial and nuclear quantitative PCR of historical marine mammal tissue, bone, baleen, and tooth samples. *Mol. Ecol. Notes* 7, 404–411.
- Morin, P.A., Leduc, R.G., Robertson, K.M., Hedrick, N.M., Perrin, W.F., Etnier, M., Wade, P., Taylor, B.L., 2006. Genetic analysis of killer whale (*Orcinus orca*) historical bone and tooth samples to identify western U.S. ecotypes. *Mar. Mamm. Sci.* 22, 897–909.
- Morin, P.A., Luikart, G., Wayne, R.K., and the SNP Workshop Group, 2004. SNPs in ecology, evolution and conservation. *Trends Ecol. Evol.* 19, 208–216.
- Morin, P.A., McCarthy, M., 2007. Highly accurate SNP genotyping from historical and low-quality samples. *Mol. Ecol. Res.* 7, 937–946.
- Morin, P.A., Nestler, A., Rubio-Cisneros, N.T., Robertson, K.M., Mesnick, S.L., 2005. Interfamilial characterization of a region of the ZFX and ZFY genes facilitates sex determination in cetaceans and other mammals. *Mol. Ecol.* 14, 3275–3286.
- Murray, M.S., 2008. Zooarchaeology and arctic marine mammal biogeography, conservation, and management. *Ecol. Appl.* 18, S41–S55.
- Nabholz, B., Glémén, S., Glater, N., 2009. The erratic mitochondrial clock: variations of mutation rate, not population size, affect mtDNA diversity across birds and mammals. *BMC Evol. Biol.* 9, 54.
- Navascués, M., Depaulis, F., Emerson, B.C., 2010. Combining contemporary and ancient DNA in population genetic and phylogeographic studies. *Mol. Ecol. Res.* 10, 760–772.
- Newsome, S.D., Etnier, M.A., Gifford-Gonzalez, D., Phillips, D.L., van Tuinen, M., Hadly, E.A., Costa, D.P., Kennett, D.J., Guilderson, T.P., Koch, P.L., 2007. The shifting baseline of northern fur seal ecology in the northeast Pacific Ocean. *Proc. Natl. Acad. Sci. U.S.A.* 104, 9709–9714.
- Nichols, C., Herman, J., Gaggiotti, O.E., Dobney, K.M., Parsons, K., Hoelzel, A.R., 2007. Genetic isolation for a now extinct population of bottlenose dolphins (*Tursiops truncatus*). *Proc. R. Soc. B* 274, 1611–1616.
- Noonan, J.P., Hofreiter, M., Smith, D., Priest, J.R., Rohland, N., Rabeder, G., Krause, J., Detter, J.C., Pääbo, S., Rubin, E.M., 2005. Genomic sequencing of Pleistocene cave bears. *Science* 309, 597–599.
- O'Corry-Crowe, G., 2008. Climate change and the molecular ecology of Arctic marine mammals. *Ecol. Appl.* 18, S56–S76.
- Ozawa, T., Hayashi, S., Mikhelson, V.M., 1997. Phylogenetic position of mammoth and Steller's sea cow within Tethytheria demonstrated by mitochondrial sequences. *J. Mol. Evol.* 44, 406–413.
- Pääbo, S., Poinar, H., Serre, D., Jaenic-Després, V., Hebler, J., Rohland, N., Kuch, M., Krause, J., Vigilant, L., Hofreiter, M., 2004. Genetic analyses from ancient DNA. *Annu. Rev. Genet.* 38, 645–679.
- Parsons, T.J., Muniec, D.S., Sullivan, K., Woodyatt, N., Alliston-Greiner, R., Wilson, M.R., Berry, D.L., Holland, K.A., Weedn, V.W., Gill, P., Holland, M.M., 1997. A high observed substitution rate in the human mitochondrial DNA control region. *Nat. Gen.* 15, 363–368.
- Péwé, T.L., Rivard, N.R., Llano, G.A., 1959. Mummified seal carcasses in the McMurdo Sound Region, Antarctica. *Science* 130, 716.
- Pichler, F.B., Dalebout, M.L., Baker, C.S., 2001. Nondestructive DNA extraction from sperm whale teeth and scrimshaw. *Mol. Ecol. Notes* 1, 106–109.
- Pinsky, M.L., Newsome, S.D., Dickerson, B.R., Fang, Y., van Tuinen, M., Kennet, D.J., Ream, R.R., Hadly, E.A., 2010. Dispersal provided resilience to range collapse in a marine mammal: Insights from the past to inform conservation biology. *Mol. Ecol.* 19, 2418–2429.
- Poinar, H.N., et al., 2006. Metagenomics to paleogenomics: large-scale sequencing of mammoth DNA. *Science* 311, 392–394.
- Post, K., 1999. Laat Pleistocene zeezoogdieren van de Nederlandse kustwateren. *Grondboor&Hamer* 53, 126–130.
- Post, K., Kompanje, E.J.O., 1995. Late Pleistocene white whales *Delphinapterus leucas* from Dutch coastal waters. *Lutra* 38, 67–76.
- Rainey, W.E., Lowenstein, J.M., Sarich, V.M., Magor, D.M., 1984. Sirenian molecular systematics—including the extinct Steller's sea cow (*Hydrodamalis gigas*). *Naturwissenschaften* 71, 586–588.
- Rasmussen, M., et al., 2010. Ancient human genome sequence of an extinct Palaeo-eskimo. *Nature* 463, 757–762.
- Rastogi, T., Brown, M.W., McLeod, B.A., Frasier, T.R., Grenier, R., Cumbaa, S.L., Nadarajah, J., White, B.N., 2004. Genetic analysis of 16th century whale bones prompts a revision of the impact of Basque whaling on right and bowhead whales in the western North Atlantic. *Can. J. Zool.* 82, 1647–1654.
- Reich, D., et al., 2010. Genetic history of an archaic hominin group from Denisova cave in Siberia. *Nature* 468, 1053–1060.
- Richards, M.P., Pacher, M., Stiller, M., Quilès, J., Hofreiter, M., Constantin, S., Zilhão, J., Trinkaus, E., 2008. Isotopic evidence for omnivory among European cave bears: late Pleistocene *Ursus spelaeus* from the Pestera cu Oase, Romania. *Proc. Natl. Acad. Sci. U.S.A.* 105, 600–604.
- Roman, J., Palumbi, S., 2003. Whales before whaling in the north Atlantic. *Science* 301, 508–510.
- Rooney, A.P., Honeycutt, R.L., Derr, J.N., 2001. Historical population size change of bowhead whales inferred from DNA sequence polymorphism data. *Evolution* 55, 1678–1685.
- Rosenbaum, H.C., Egan, M.G., Clapham, P.J., Brownell Jr., R.L., Desalle, R., 1997. An effective method for isolating DNA from historical specimens of baleen. *Mol. Ecol.* 6, 677–681.

- Rosenbaum, H.C., Egan, M.G., Clapham, P.J., Brownell Jr., R.L., Malik, S., Brown, M.W., White, B.N., Walsh, P., Desalle, R., 2000. Utility of North Atlantic right whale museum specimens for assessing changes in genetic diversity. *Cons. Biol.* 14, 1837–1842.
- Roy, M.S., Girmar, D.J., Taylor, A.C., Wayne, R.K., 1994. The use of museum specimens to reconstruct the genetic variability and relationships of extinct populations. *Experientia* 50, 551–557.
- Ryder, O.A., 1986. Species conservation and systematics: the dilemma of subspecies. *Trends Ecol. Evol.* 1, 9–10.
- Santos, C., Montiel, R., Sierra, B., Bettencourt, C., Fernandez, E., Alvarez, L., Lima, M., Abade, A., Aluja, M.P., 2005. Understanding differences between phylogenetic and pedigree-derived mtDNA mutation rate: A model using families from the Azores Islands (Portugal). *Mol. Evol. Biol.* 22, 1490–1505.
- Scott, R.F., 1905. *The Voyage of the 'Discovery'*. Smith, Elder & Co., London.
- Shapiro, B., et al., 2004. Rise and fall of the Beringian steppe bison. *Science* 306, 1561–1565.
- Smith, C.I., Chamberlain, A.T., Riley, M.S., Stringer, C., Collins, M.J., 2003. The thermal history of human fossils and the likelihood of successful DNA amplification. *J. Hum. Evol.* 45, 203–217.
- Sommer, R.S., Pasold, J., Schmölk, U., 2008. Post-glacial immigration of the harbour porpoise (*Phocoena phocoena*) into the Baltic Sea. *Boreas* 37, 458–464.
- Sremba, A., Martin, A.R., Baker, C.S., 2010. Genetic approach to species identification of whale bones from South Georgia Island whaling stations. Report to Scientific Committee of the International Whaling Commission SC/62/SH19.
- Stiller, M., Knapp, M., Stenzel, U., Hofreiter, M., Meyer, M., 2009. Direct multiplex sequencing (DMPS)—a novel method for targeted high-throughput sequencing of ancient and highly degraded DNA. *Genome Res.* 19, 1843–1848.
- Stirling, I., Rudolph, E.D., 1968. Inland record of a live crabeater seal in Antarctica. *J. Mammal.* 49, 161–162.
- Storå, J., Ericson, P.G.P., 2004. A prehistoric breeding population of harp seals (*Phoca groenlandica*) in the Baltic Sea. *Mar. Mamm. Sci.* 20, 115–133.
- Taberlet, P., Griffin, S., Goossens, B., Questiau, S., Manceau, V., Escaravage, N., Waits, L.P., Bouvet, J., 1996. Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Res.* 24, 3189–3194.
- Thomas, W.K., Schaffner, W., Wilson, A.C., Pääbo, S., 1989. DNA phylogeny of the extinct marsupial wolf. *Nature* 340, 465–467.
- Tingley, M.W., Beissinger, S.R., 2009. Detecting range shifts from historical species occurrences: new perspectives on old data. *Trends Ecol. Evol.* 24, 625–633.
- Turvey, S.T., Risley, C.L., 2006. Modelling the extinction of Steller's sea cow. *Biol. Lett.* 22, 94–97.
- Valentine, K., Duffield, D.A., Patrick, L.E., Hatch, D.R., Butler, V.L., Hall, R.L., Lehman, N., 2008. Ancient DNA reveals genotypic relationships among Oregon populations of the sea otter (*Enhydra lutris*). *Cons. Genet.* 9, 933–938.
- Weber, D.S., Stewart, B.S., Garza, J.C., Lehman, N., 2000. An empirical genetic assessment of the severity of the northern elephant seal population bottleneck. *Curr. Biol.* 10, 1287–1290.
- Weber, D.S., Stewart, B.S., Lehman, N., 2004. Genetic consequences of a severe population bottleneck in the Guadalupe fur seal (*Arctocephalus townsendi*). *J. Hered.* 95, 144–153.
- Willerslev, E., et al., 2009. Analysis of complete mitochondrial genomes from extinct and extant rhinoceroses reveals lack of phylogenetic resolution. *BMC Evol. Biol.* 9, 95.
- Williams, T.M., 1999. The evolution of cost efficient swimming in marine mammals: limits to energetic optimization. *Phil. Trans. R. Soc. B* 354, 193–201.
- Wilson, E.A., 1907. *Mammalia (whales and seals)*. National Antarctic Expedition, 1901–4, Natural History, vol. 2. British Museum, London, pp. 1–66.
- Yang, D.Y., Eng, B., Wayne, J.S., Duder, J.C., Saunders, S.R., 1998. Improved DNA extraction from ancient bones using silica-based spin columns. *Am. J. Phys. Anthropol.* 105, 539–543.

Taxonomic revision of the olingos (*Bassaricyon*), with description of a new species, the Olinguito

Kristofer M. Helgen^{1,†}, C. Miguel Pinto^{2,3,4,5,‡}, Roland Kays^{6,7,8,§},
Lauren E. Helgen^{1,¶}, Mirian T. N. Tsuchiya^{1,9,10,¶¶}, Aleta Quinn^{1,11,¶},
Don E. Wilson^{1,††}, Jesús E. Maldonado^{1,10,¶¶}

1 Division of Mammals, National Museum of Natural History, NHB 390, MRC 108, Smithsonian Institution, P.O. Box 37012, Washington, DC 20013-7012, USA **2** Centro de Investigación en Enfermedades Infecciosas, Escuela de Ciencias Biológicas, Pontificia Universidad Católica del Ecuador, Av. 12 de Octubre y Roca, Quito, Ecuador **3** Department of Mammalogy, and Sackler Institute for Comparative Genomics, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024, USA **4** The Graduate Center, City University of New York, 365 Fifth Ave., New York, NY, 10016 USA **5** Department of Biological Sciences and the Museum, Texas Tech University, Lubbock, Texas 79409-3131, USA **6** North Carolina Museum of Natural Sciences, 11 West Jones Street, Raleigh, NC, 27601, USA **7** Fisheries, Wildlife & Conservation Program, North Carolina State University, Raleigh, NC, 27695, USA **8** Smithsonian Tropical Research Institute, Balboa Ancón, Republic of Panamá **9** Department of Environmental Science & Policy, George Mason University, Fairfax, VA, 22030 USA **10** Center for Conservation and Evolutionary Genetics, Smithsonian Conservation Biology Institute, National Zoological Park, Washington, DC 20008, USA **11** Department of History and Philosophy of Science, University of Pittsburgh, Pittsburgh, PA, 15260, USA

† <http://zoobank.org/66786588-E2AF-4A97-B188-8A31367C7975>

‡ <http://zoobank.org/473217A8-3902-4B95-89D2-28FC4452DCD1>

§ <http://zoobank.org/12018BFE-6422-418F-8923-4EB4B9CFB9C9>

¶ <http://zoobank.org/D78E4F09-4ED6-46F7-AF36-7C28B0329385>

¶¶ <http://zoobank.org/80A3E87F-4C7A-4B66-AEE4-1E4DC405BC7A>

¶¶¶ <http://zoobank.org/E2F31DAB-24A2-4F53-B6AE-8A8B3CF5328A>

†† <http://zoobank.org/C3AB3E54-C11A-4A68-BD26-4439C82B956F>

††† <http://zoobank.org/0CBF5D72-8C0A-489C-A148-B262884B60FC>

Corresponding author: Kristofer M. Helgen (helgen@si.edu)

Academic editor: W. Bogdanowicz | Received 23 June 2013 | Accepted 12 August 2013 | Published 15 August 2013

<http://zoobank.org/A5747BD0-FC61-42C9-A1A8-FA1C9D27CD22>

Citation: Helgen KM, Pinto CM, Kays R, Helgen LE, Tsuchiya MTN, Quinn A, Wilson DE, Maldonado JE (2013) Taxonomic revision of the olingos (*Bassaricyon*), with description of a new species, the Olinguito. ZooKeys 324: 1–83. doi: 10.3897/zookeys.324.5827

Abstract

We present the first comprehensive taxonomic revision and review the biology of the olings, the endemic Neotropical procyonid genus *Bassaricyon*, based on most specimens available in museums, and with data derived from anatomy, morphometrics, mitochondrial and nuclear DNA, field observations, and geographic range modeling. Species of *Bassaricyon* are primarily forest-living, arboreal, nocturnal, frugivorous, and solitary, and have one young at a time. We demonstrate that four olingo species can be recognized, including a Central American species (*B. gabbii*), lowland species with eastern, cis-Andean (*B. alleni*) and western, trans-Andean (*B. medius*) distributions, and a species endemic to cloud forests in the Andes. The oldest evolutionary divergence in the genus is between this last species, endemic to the Andes of Colombia and Ecuador, and all other species, which occur in lower elevation habitats. Surprisingly, this Andean endemic species, which we call the Olinguito, has never been previously described; it represents a new species in the order Carnivora and is the smallest living member of the family Procyonidae. We report on the biology of this new species based on information from museum specimens, niche modeling, and fieldwork in western Ecuador, and describe four Olinguito subspecies based on morphological distinctions across different regions of the Northern Andes.

Keywords

Andes, *Bassaricyon*, biogeography, Neotropics, new species, olingo, Olinguito

Contents

Introduction.....	3
Materials and methods	4
Museum specimens and comparisons	4
DNA extraction	8
DNA Sequencing	8
Molecular analysis	9
Bioclimatic range modeling.....	11
Results.....	12
Phylogenetics	12
Biogeography	15
Morphology and morphometrics.....	15
Bioclimatic range modeling.....	19
Systematics	28
<i>Bassaricyon neblina</i> sp. n.	28
<i>Bassaricyon gabbii</i> J. A. Allen, 1876	41
<i>Bassaricyon alleni</i> Thomas, 1880	44
<i>Bassaricyon medius</i> Thomas, 1909	49
Discussion	52
Carnivore taxonomy.....	52
Conservation.....	53
Biogeography	54
Additional Olinguito study priorities	56
Appendix 1.....	78
Appendix 2.....	81

Introduction

“New Carnivores of any sort are always few and far between...”
Oldfield Thomas (1894:524)

Olingos (genus *Bassaricyon* J.A. Allen, 1876) are small to medium-sized (0.7 to 2 kg) arboreal procyonids found in the forests of Central America and northern South America. No comprehensive systematic revision of the genus has ever been undertaken, such that species boundaries in *Bassaricyon* are entirely unclear, and probably more poorly resolved than in any other extant carnivoran genus. There are various reasons for limited knowledge of *Bassaricyon*. For such a widespread genus of Carnivora, olingos were discovered surprisingly late (first described from Central America in 1876 and from South America in 1880; Allen 1876, Thomas 1880); they were long known by few specimens in museum collections; they are often overlooked in the field because they are regularly confused with another better known procyonid, the kinkajou, *Potos flavus* (Schreber, 1774) (e.g., Thomas 1880, Manville 1956, Ford and Hoffmann 1988, Sampaio et al. 2011); and they are often largely or entirely omitted from both authoritative and popular references on Neotropical wildlife and natural history (e.g., Janzen 1983, Henderson 2002, Lord 2007). In the absence of a detailed systematic review, five species of *Bassaricyon* are tentatively recognized in most recent taxonomic references, including three species in Central America (*B. gabbii* Allen, 1876; *B. lasius* Harris, 1932; *B. pauli* Enders, 1936) and three species in South America (with *B. gabbii* recognized as occurring west of the Andes, and *B. alleni* Thomas, 1880 and *B. beddardi* Pocock, 1921a east of the Andes), but most authors have explicitly identified a longstanding need for a detailed taxonomic overview to clarify species diversity and distributions in this genus (Cabrera 1958, Decker and Wozencraft 1991, Eisenberg 1989, Eisenberg and Redford 1999, Eizirik 2012, Emmons 1990, 1997, Ewer 1973, Glatston 1994, González-Maya et al. 2011, Hall 1981, Hall and Kelson 1959, Helgen et al. 2008c, Hunter 2011, Kays 2009, Kays and Russell 2001, Nowak 1999, Poglayen-Neuwall and Poglayen-Neuwall 1965, Prange and Prange 2009, Reid 1997, 2009, Reid and Helgen 2008a, 2008b, 2008c, Russell 1984, Samudio et al. 2008, Stains 1967, Wozencraft 1989, 1993, 2005, Zeveloff 2002).

Here we review the taxonomic standing of all named forms of *Bassaricyon* based on morphological, morphometric, and molecular comparisons of voucher specimens in museums; we clarify the distribution and conservation status of each valid taxon; and, as far as possible, we enable information from published literature on olingo anatomy (e.g., Beddard 1900, Mivart 1885, 1886, Pocock 1921a, 1921b, Segall 1943, Stains 1973, Story 1951), ecology and behavior (e.g., Aquino and Encarnación 1986, Emmons 1990, 1991, Glanz 1990, Goldman 1920, Hunter and Caro 2008, Janson and Emmons 1990, Kays 2000, Loyola et al. 2008, Patton et al. 1982, Peres 2001, Poglayen-Neuwall 1966, 1973, 1976, 1989, Poglayen-Neuwall and Poglayen-Neuwall 1965, Prange and Prange 2009, Reid 1997, Rodríguez and Amanzo 2001, Wainwright 2002), and parasites and disease (e.g., Grimaldi and Tesh 1993, Hendricks

1977, Herrer and Christensen 1975, Jewell et al. 1972, Magalhães-Pinto et al. 2009) to be associated with particular olingo taxa now recognized as valid.

All previously described olingo taxa occur in lower to middle-elevation tropical or subtropical forests (≤ 2000 meters in elevation). Remarkably, our morphological, morphometric, molecular, and field studies document the existence of an undescribed species in the genus, endemic to higher-elevation cloud forests (1500 to 2750 meters) in the Western and Central Andes of Colombia and Ecuador, which we describe here as a new species. (This species has been discussed preliminarily, in advance of its formal description, by Kays [2009] and Hunter [2011].) This species, upon which we bestow the common name of Olinguito (oh-ling-GHEE'-toh), is the sister taxon to a lineage comprising all previously described species of *Bassaricyon*; is the smallest living procyonid; and is the first new species of American carnivore described since the discovery of the Colombian weasel (*Mustela felipei*) in similar habitats in the same region of the Andes more than three decades ago (Izor and de la Torre 1978). We discuss what is known to date of the biology of this remarkable new procyonid, the Olinguito.

Materials and methods

Museum specimens and comparisons

We examined all *Bassaricyon* specimens in the collections of the American Museum of Natural History, New York, USA (AMNH); Academy of Natural Sciences, Philadelphia, USA (ANSP); Natural History Museum, London, UK (BMNH); Museo de Zoología, Universidad Politecnica, Quito, Ecuador (EPN); Field Museum of Natural History, Chicago, USA (FMNH); Biodiversity Institute, University of Kansas, Lawrence, USA (KU); Los Angeles County Natural History Museum, Los Angeles, USA (LACM); Museum of Comparative Zoology, Harvard University, Cambridge, USA (MCZ); Museo Ecuatoriano de Ciencias Naturales, Quito, Ecuador (MECN); Museum of Vertebrate Zoology, University of California, Berkeley, USA (MVZ); Naturhistoriska Riksmuseet, Stockholm, Sweden (NMS); Museo de Zoología, Pontificia Universidad Católica del Ecuador, Quito, Ecuador (QCAZ); Royal Ontario Museum, Toronto, Canada (ROM); Biodiversity Research and Teaching Collections, Texas A&M University, College Station, USA (TCWC); Museum of Zoology, University of Michigan, Ann Arbor, USA (UMMZ); National Museum of Natural History, Smithsonian Institution, Washington, D.C., USA (USNM); Peabody Museum of Natural History, Yale University, New Haven, USA (YPM); and Museum für Naturkunde, Humboldt Universität, Berlin, Germany (ZMB). These holdings include all type specimens in the genus and represent the great majority (well over 95%) of olingo specimens in world museums. We also had access to published information on a few additional specimens in museum collections in Colombia and Bolivia (Saavedra-Rodríguez and Velandia-Perilla 2011, Anderson 1997). Tissue sam-

ples are stored in the frozen tissue collections of the MVZ, ROM, USNM (including specimens to be accessioned at QCAZ), the New York State Museum, Albany, New York, USA (NYSM), and the Museum of Texas Tech University, Lubbock, Texas, USA (TTU) (Table 1).

Table 1. List of samples (and associated information) used in phylogenetic analysis. Boldfaced entries represent samples newly sequenced in this study.

SPECIES	Identifier in Figure 1	Specific locality	Source (catalog reference)	Genbank Accession Numbers	
				Cytochrome <i>b</i>	<i>CHRNA1</i>
<i>Bassaricyon medius orinomus</i>	Panama	Limbo plot	NYSM ZT105	EF107703	KC773757
<i>Bassaricyon medius orinomus</i>	Panama	Rio Juan Grande	NYSM ZT106	EF107704	KC773758
<i>Bassaricyon medius orinomus</i>	Panama	Limbo plot	Koepfli et al. (2007)	DQ660300	DQ660210
<i>Bassaricyon medius medius</i>	Ecuador	Las Pampas	QCAZ 8659; tk149097	EF107706	KC773759
<i>Bassaricyon medius medius</i>	Ecuador	Las Pampas	QCAZ 8658; tk149094	EF107707	KC773760
<i>Bassaricyon alleni</i>	Guyana	Iwokrama	ROM 107380	EF107710	KC773763
<i>Bassaricyon alleni</i>	Peru	Rio Cenapa	MVZ 155219; Koepfli et al. (2007)	DQ660299	DQ660209
<i>Bassaricyon gabbi</i>	Costa Rica	Monteverde	KU 165554	JX948744	---
<i>Bassaricyon neblina neblina</i>	Ecuador	La Cantera	QCAZ 8662; tk149108	EF107708	KC773761
<i>Bassaricyon neblina neblina</i>	Ecuador	Otonga Reserve	QCAZ 8661; tk149001	EF107709	KC773762
<i>Bassaricyon neblina osborni</i>	Colombia	Vicinity of Cali	Genbank	X94931	DQ533950
<i>Potos flavus</i>	<i>Potos flavus</i>	Costa Rica	Koepfli et al. (2007)	DQ660304	DQ660214
<i>Procyon cancrivorus</i>	<i>Procyon cancrivorus</i>	Paraguay	Koepfli et al. (2007)	DQ660305	DQ660215
<i>Procyon lotor</i>	<i>Procyon lotor</i>	Montana, USA	Koepfli et al. (2007)	DQ660306	AF498152
<i>Bassariscus astutus</i>	<i>Bassariscus astutus</i>	Arizona, USA	Koepfli et al. (2007)	AF498159	AF498151
<i>Bassariscus sumichrasti</i>	<i>Bassariscus sumichrasti</i>	Mexico	Koepfli et al. (2007)	DQ660301	DQ660211
<i>Nasua nasua</i>	<i>Nasua nasua</i>	Bolivia	Koepfli et al. (2007)	DQ660303	DQ660213
<i>Nasua narica</i>	<i>Nasua narica</i>	Panama	Koepfli et al. (2007)	DQ660302	DQ660212
<i>Enhydra lutris</i>	Mustelidae	Attu Island, Alaska, USA	Koepfli et al. (2007)	AF057120	AF498131
<i>Eira barbara</i>	Mustelidae	Bolivia	Koepfli et al. (2007)	AF498154	AF498144
<i>Taxidea taxus</i>	Mustelidae	New Mexico, USA	Koepfli et al. (2007)	AF057132	AF498148
<i>Neovison vison</i>	Mustelidae	Texas, USA	Koepfli et al. (2007)	AF057129	AF498140

SPECIES	Identifier in Figure 1	Specific locality	Source (catalog reference)	Genbank Accession Numbers	
				Cytochrome <i>b</i>	<i>CHRNA1</i>
<i>Martes americana</i>	Mustelidae	Rocky Mtn Research Station, USA	Koepfli et al. (2007)	AF057130.1	AF498141
<i>Lontra longicaudis</i>	Mustelidae	Kagka, Peru	Koepfli et al. (2007)	AF057123.1	AF498134
<i>Ictonyx libyca</i>	Mustelidae	Brookfield Zoo	Genbank	EF987739.1	EF987699
<i>Meles meles</i>	Mustelidae	No voucher infromation	Koepfli et al. (2007)	AM711900.1	AF498147
<i>Mephitis mephitis</i>	Mephitidae	San Diego Zoo	Eizirik et al. (2010), Yu et al. (2011)	HM106332.1	GU931029.1
<i>Spilogale putorius</i>	Mephitidae		Arnason et al. (2007), Eizirik et al. (2010)	NC_010497.1	GU931030.1
<i>Ailurus fulgens</i>	Ailuridae		Arnason et al. (2007), Eizirik et al. (2010)	AM711897.1	GU931037.1
<i>Arctocephalus australis</i>	Otariidae		Davis et al. (2004), Fulton and Strobeck (2006)	AY377329.1	DQ205738.1
<i>Odobenus rosmarus</i>	Odobenidae		Bardeleben et al. (2005), Fulton and Strobeck (2010)	GU174611.1	DQ093076.1
<i>Phoca fasciata</i>	Phocidae		Fulton and Strobeck (2010)	GU167294.1	GU167764.1
<i>Mirounga angustirostris</i>	Phocidae		Bardeleben et al. (2005), Peng et al. (2007)	AY377325.1	DQ093075.1
<i>Canis lupus</i>	<i>Canis lupus</i>		Delisle and Strobeck (2005), Fulton and Strobeck (2006)	AY598499	DQ205757
<i>Nyctereutes procyonoides</i>	other Canidae		Eizirik et al. (2010), Chen and Zhang (2012)	GU256221	GU931027.1
<i>Urocyon cinereoargenteus</i>	other Canidae		Eizirik et al. (2010), Naidu et al. (2012)	JF489121.1	GU931028.1
<i>Ailuropoda melanoleuca</i>	Ursidae		Bardeleben et al. (2005), Peng et al. (2007)	NC_009492	DQ093074.1
<i>Ursus americanus</i>	Ursidae		Delisle and Strobeck (2002), Fulton and Strobeck (2006)	NC_003426.1	DQ205726.1

Values from external measurements of 95 specimens are presented to provide an appreciation of general body size and lengths and proportions of appendages. Values (in mm) for total length and length of tail are those recorded by collectors on labels attached to skins; subtracting length of tail (abbreviated TV) from total length produced

a value for length of head and body (HB). Values for length of hind foot (HF), which includes claws, were either obtained from skin labels or from our measurements of dry study skins; those for length of external ear (E), or pinna, come from collector's measurements recorded on skin labels or in field journals (we assume, but are not certain for all specimens, that ear-length measurements represent the greatest length from the notch to the distal margin of the pinna).

Morphological terminology follows Evans (1993) and Ahrens (2012). Craniodental variables were measured by the first author with digital calipers to the nearest 0.1 mm. Single-tooth measurements are measured on the crown. All measurements of length are in millimeters, and measurements of mass are given in grams. Only fully adult, wild-collected specimens that are sufficiently intact were included in our morphometric analyses. A total of 115 specimens were included (51 male, 64 female). The classification of "adult" was applied generally only to skulls in which the full dentition is completely erupted, and in which the basilar (basioccipital-basisphenoid) suture (synchondrosis) in particular is obliterated via ossification. Variables measured include maximum crown widths (W) of premolars (p1, p2, p3, p4, P2, P3, P4, with lower case designating lower teeth and uppercase designating upper teeth) and molars (m1, m2, M1, and M2); maximum crown lengths (L) of the larger premolars and molars (P4, M1, M2, m1, and m2); condylobasal length (CBL), zygomatic width (ZYG), breadth of braincase (BBC), external width across the canines (CC), and length of the maxillary toothrow, C-M2 (MTR), all as defined by Kennedy and Lindsay (1984); and four posterior skull measurements: greatest width across the postdental palatal shelf (WPP), length of the postdental palate behind an imaginary line delineated by the back of the second molars (LPP), anteroposterior length of the auditory bullae including the eustachian tube (LAB), and the dorsoventral diameter inside the external auditory meatus (EAM). Unless explicitly noted, all reported metrics (and resulting statistical and multivariate comparisons) refer only to fully mature (adult) specimens, as judged by direct examination of skulls. Because some olingo taxa demonstrate significant sexual dimorphism in cranial measurements, patterns of morphometric variation in males and females were compiled and analyzed separately. Principal Component Analysis (PCA) and Discriminant Function Analyses (DFA) were undertaken using a combination of cranial and dental measurements indicated in tables and in the text, selected to sample craniodental size and shape, and to maximize sample size. All measurement values were transformed to natural logarithms prior to multivariate analysis. Principal components were extracted from the covariance matrix. The software program Statistica 8.0 (Statsoft Inc., Tulsa, Oklahoma, USA) was used for all multivariate analyses.

The taxa and sequences included in our analysis are listed in Table 1. Our choice of taxa outside of *Bassaricyon* was guided by the findings of Koepfli et al. (2007), Fulton and Strobeck (2007), and Eizirik et al. (2010). These studies provide strong statistical support for relationships and divergence dates within Procyonidae and Carnivora based on >6,000 bases of DNA and fossil evidence. We chose one mitochondrial marker and one nuclear marker used in these and many other mammalian

studies, in order to capture the evolutionary histories of these distinct genetic systems. Although deeper relationships within the order Carnivora cannot be resolved solely by using these two genes, we are confident that they provide the appropriate level of support to resolve species-level relationships within this group of procyonids (Koepfli et al. 2007). As our specific goal was to estimate the timing of divergence within *Bassaricyon*, and our reduced dataset did not provide enough support to resolve deeper nodes in Caniformia, we decided to use highly supported divergence date estimates from Koepfli et al. (2007) and Eizirik et al. (2010) as priors in our analysis.

DNA extraction

Tissues from fresh and frozen specimens were processed using a Qiagen DNeasy kit (QIAGEN, Valencia, CA, USA) to obtain genomic DNA. The sample from the skull of KU 165554, a museum specimen of *B. gabbii*, was taken from the turbinate bones and extracted following the method of Wisely et al. (2004). Including this turbinate sample of *B. gabbii*, we successfully extracted DNA from eight individuals of *Bassaricyon* (four *B. medius*, one *B. alleni*, and two *B. neblina* sp. n.). All pre-PCR protocols were conducted in an isolated ancient DNA laboratory facility located in a separate building from the one containing the primary DNA laboratory.

DNA Sequencing

Mitochondrial gene, cytochrome *b*: For cytochrome *b* (1140 bp), polymerase chain reaction (PCR) and sequencing reactions were carried out with primers LGL 765 and LGL 766 from Bickham et al. (2004) and using a thermal cycler (MJ Research, Waltham, MA, USA) under the following conditions, repeated for 35 cycles: denaturation at 92°C for 1 min, annealing at 50°C for 1 min, extension at 72°C for 1 min. The PCR reagents in a 25 µL reaction were 0.2 µL AmpliTaq (5 units µL-1, Applied Biosystems, Foster City, CA, USA), 1 µL per primer (10 µM), 2.5 µL dNTP (2 µM), 2 µL MgCl₂ (25 mM), 2.5 µL AmpliTaq Buffer (Applied Biosystems), 2 µL BSA (0.01 mg/µL), 1 µL gDNA and 12.8 µL sterile water. To amplify DNA from turbinate samples, PCR and sequencing were carried out with internal primers designed for this study using sequences generated from the tissue samples; the reverse primer, H151949Pro (5' CTCCTCAAAAGGATTTGYCCTCA 3'), located at 14611 – 14636 on the *Nasua nasua* mitochondrial genome, was used with LGL 765. A new forward primer, BAS420F (5' TCAGACAAAATCCCCTTCCA 3'), position 14825 - 14845 on the *N. nasua* mitochondrial genome, was used with LGL 766. The thermal cycle regime was modified to 50 cycles; reagents were as above.

Nuclear intron, Cholinergic Receptor Nicotinic Alpha Polypeptide 1 precursor (*CHRNA1*): For *CHRNA1* (347 bp), we used the primers described by Lyons

et al. (1997) and the thermocycling conditions consisted of an initial denaturation (95°C for 10 min), followed by 30 cycles of denaturation at 95°C for 30 s, annealing at 54°C for 30 s and extension for 72°C for 45 s, and final extension of 72°C for 5 min. Reagent volumes were the same as for cytochrome *b* amplification (above), except 2 µL of gDNA was added for *CHRNA1* amplification, decreasing sterile water volume to 10.8 µL. We were unable to sequence the nuclear intron from the turbinate bone sample.

Each PCR was conducted with negative and positive controls to minimize risk of spurious results from contamination or failure of the reaction. A 2 µL sample of the PCR product was stained with ethidium bromide and run on an agarose gel with a 1 kb ladder. The gel was placed under UV light to visualize the PCR products. Polymerase chain reaction products were amplified for sequencing using a 10 µL reaction mixture of 2 µL of PCR product, 0.8 µL of primer (10 µM), 1.5 µL Big Dye 5 x Buffer (Applied Biosystems), 1 µL Big Dye version 3 (Applied Biosystems), and 4.7 µL sterile water. The reaction was run using a thermal cycler (MJ Research) with denaturation at 96°C for 10 s, annealing at 50°C for 10 s and extension at 60°C for 4 min: this was repeated for 25 cycles. The product was cleaned using sephadex filtration method and sequences for both strands were run on a 50 cm array using the ABI PRISM 3100 Genetic Analyzer (Applied Biosystems).

Molecular analysis

Sequences were aligned and edited in Sequencher version 4.1.2 using the implemented Clustal algorithm and the default gap penalty parameters (Gene Codes Corporation, Ann Arbor, MI, USA, <http://www.genecodes.com>).

For *Bassaricyon*, we included all newly sequenced and previously available sequences for cytochrome *b* and *CHRNA1* (for cytochrome *b* this included five individuals of *B. medius*, two *B. alleni*, one *B. gabbi* and three *B. neblina* sp. n.; for *CHRNA1* this included five individuals of *B. medius*, two *B. alleni*, and three *B. neblina* sp. n.) (Table 1).

Maximum Parsimony, Maximum Likelihood and Bayesian analyses were performed for each gene and a concatenation of the two genes to check for any incongruence in structure and support of the *Bassaricyon* clade. All Bayesian and Maximum Likelihood phylogenetic inferences were carried out using the Cipres Portal (Miller et al. 2009). Indels were treated as missing data or non-informative data for all of the analyses as in previous molecular phylogenetic studies of procyonids (Koepfli et al. 2007; Eizirik et al. 2010).

Pairwise distances for cytochrome *b* were generated using the Kimura 2-parameter model using MEGA4 (Tamura et al. 2007).

The branch and bound search method implemented in the software package TNT (Goloboff et al. 2008) was used for the maximum parsimony analyses. Parsimony bootstrap support was generated using the heuristic search method with 100 random stepwise additions for 1000 replicates.

Maximum Likelihood analysis was conducted using the software package GARLI 0.96b (Zwickl 2006). The genetic-like algorithm was used to simultaneously find the topology, branch lengths and substitution model parameters with the greatest log-likelihood (lnL) for each dataset. Bootstrap support was generated with 1000 replicates and two independent searches per replicate.

jModeltest version 0.1.1 (Posada 2008) was used to find the best model of sequence evolution. We chose to partition the cytochrome *b* data in order to minimize the number of parameters and to account for differences in base composition and substitution rates among the different codon positions (Corse et al. 2013). The software PartitionFinder (Lanfear et al. 2012) was used to determine the best partitioning scheme, and for the cytochrome *b*, the scheme with 1st, 2nd, and 3rd codon positions partitioned separately was selected. The best fit model under the Bayesian information criterion (BIC) for cytochrome *b* for the first and second codon position partitions was HKY + G + I (Hasegawa et al. 1985), and for the third codon position, the best model under BIC was TrN + I + G (Tamura and Nei 1993). The model chosen for *CHRNA1* was K80 + G (Kimura 1980). The parameters were then applied in MrBayes version 3.1p (Huelsenbeck and Ronquist 2001). The model parameters were set to nst = 2 using a gamma distribution for *CHRNA1*, nst = 2 and the rate parameter invariant with a gamma distribution for the cytochrome *b* 1st and 2nd codon partitions, and nst = 6 with a gamma distribution and rate parameter invariant for cytochrome *b* 3rd codon partition. Since this version of MrBayes did not include the specific model selected for the cytochrome *b* 3rd codon position partition, we opted for using a more complex model (nst = 6) following the results of Huelsenbeck and Rannala (2004). The Bayesian analysis was run using 5,000,000 generations along four chains with 2 replicates at a temperature of 0.05. The convergence between the two replicates run in MrBayes was assessed by the average standard deviation of split frequencies (ASDSF) between runs. After 5,000,000 generations, the ASDSF was 0.003. Sample frequency was set to 1000 with a burn-in of 1,250.

Molecular divergence estimates were generated in BEAST (Drummond and Rambaut 2007). The following calibration nodes were included based on Eizirik et al. (2010): *Nasua* – *Bassaricyon* truncated normal mean 7.2 million years ago (mya) (± 1.7 s.d.); *-Potos* truncated normal mean 16.2 mya (± 2.5 s.d.); Procyonidae normal mean 20.7 mya (± 4.0 s.d.); Procyonidae-Mustelidae-Ailuridae-Mephitidae normal mean 30 mya (± 7.0 s.d.); *Phoca-Mirounga* normal mean 20 mya (± 6 s.d.); Caniformia normal mean 48 mya (± 6.5 s.d.). The molecular clock was estimated using the uncorrelated lognormal setting, operators were left to the default setting, and trees were searched using the Yule process. The substitution and clock models were left unlinked, partition tree model was linked, and the models for the two gene partitions were: cytochrome *b* (1), (2), and (3) => TN93 + I + G (all parameters unlinked) and *CHRNA1* K80 + G (HKY + G). In order to evaluate the effects of the priors on the divergence time estimates, we carried out a run using an empty alignment but with the same settings and

compared it to our results, with the outcome indicating that the priors are not having an especially strong effect on the estimated divergence times (Drummond et al. 2006).

To infer geographical range evolution of procyonids we used the Maximum Likelihood model of dispersal-extinction-cladogenesis (DEC) implemented in Lagrange v. 20130526 (Ree and Smith 2008). The BEAST chronogram tree was trimmed to keep one representative per procyonid species, and two additional lineages, one representing Mustelidae and one representing Mephitidae. Six general geographic areas were used to characterize the distribution ranges: Eurasia, North America, Central America, Chocó, Andes, and Amazonia. The branches of the mustelid and mephitid lineages were treated as belonging to the ancestors of the families and their hypothesized distributions are according to previous ancestral range estimations (Koepfli et al. 2008, Sato et al. 2009). Reconstruction of potential ancestral area combinations and dispersal scenarios took into account realistic dispersal routes (e.g., allowing Eurasia to connect only with North America) and the geological history of the region (e.g., formation of the Panama Isthmus during the late Miocene and Pliocene; Weyl 1980, Almendra and Rogers 2012).

Bioclimatic range modeling

Vouchered localities of occurrence for *Bassaricyon* used in our analyses were extracted from museum specimen labels, often as clarified by associated field notes and journals, and from definitive published accounts. Gazetteers published by Paynter (1982, 1993, 1997), Stephens and Taylor (1983, 1985), Fairchild and Handley (1966), Handley (1976), Voss (1988), and Voss et al. (2002) were especially helpful in georeferencing Neotropical expedition and collecting localities represented in museum collections.

We used Maximum Entropy Modeling (Maxent) (Phillips et al. 2006) to predict the geographic range of the geographic range of the four *Bassaricyon* species at broad scales based on vouchered localities (Appendix 2) and 20 environmental variables representing potential vegetation and climate. For potential vegetation we used the 15 major habitat types classified as ecological biomes (Olson et al. 2001). For climate we used 19 BIOCLIM variables representing annual trends, seasonality, and extremes in temperature and precipitation across Central and South America (derived from Hijmans et al. [2005] as described at <http://www.worldclim.org/bioclim.htm>). We used all vouchered specimen localities to train the final model (excluding published records based only on visual observations). We also tested overall performance by running 10 model iterations while randomly withholding 20% of the points as test locations. To produce geographic ranges showing presence/absence of a species we used the average equal training sensitivity and specificity for the 10 test models as our probability cutoff value (Phillips et al. 2006).

Results

Phylogenetics

With the largest molecular sampling effort to date, we show that *Bassaricyon* is well resolved as a monophyletic genus (cf. Nyakatura and Bininda-Emonds 2012) within the family Procyonidae. All of our analyses resolve *Bassaricyon* as a clade with bootstrap and probability values of 100%. The sister genus to *Bassaricyon* is *Nasua*, a relationship consistently recovered in our analyses with 100% support. The divergence between *Bassaricyon* and *Nasua* was estimated at 10.2 million years old (mya) (95% Confidence Interval [CI] = 7.6 – 12.7 mya), consistent with previously published results (Koepfli et al. 2007, Eizirik et al. 2010).

The family Procyonidae is well resolved as monophyletic (100% bootstrap and probability values) with a divergence date of 21.4 mya (CI 18.1 – 25.0 mya), in agreement with the divergence estimate of 22.6 mya (CI 19.4 – 25.5 mya) by Eizirik et al. (2010). Eizirik et al. (2010) had a more constrained confidence interval on the age of this divergence, due to the incorporation of genes that are more informative at deeper nodes in the tree. We chose *CHRNA1* and cytochrome *b* with a focus toward resolving relationships within *Bassaricyon*; these markers are far more useful for determining relationships in recent radiations within Procyonidae than the deeper relationships within Carnivora. The only part of the Procyonidae where *CHRNA1* and cytochrome *b* did not provide sufficient resolution to re-construct recently published multi-gene topologies (Koepfli et al. 2007, Eizirik et al. 2010) was the divergence between the two species of *Bassariscus*, and *Procyon*. In our BEAST chronogram the divergence for *Bassariscus* is 7.6 mya (CI 4.8 – 10.6 mya) but the branch leading to their divergence has no support, and therefore is collapsed in the phylogeny (Figure 1; see also Koepfli et al. 2007, Eizirik et al. 2010). The other procyonid genera are well-supported monophyletic groups; according to our chronogram *Procyon lotor* and *P. cancrivorus* diverged 4.2 mya (CI 2.3 – 6.5 mya) and *Nasua narica* and *N. nasua* diverged 5.6 mya (CI 3.5 – 7.9 mya).

The concordance of our recovered topology and estimates of genetic divergence with previous phylogenetic studies of the Procyonidae suggests that data from cytochrome *b* and *CHRNA1* across sampled taxa have provided a well-supported framework in which the species relationships and divergence dates within *Bassaricyon* can be reliably assessed. Previous molecular phylogenetic studies have included either only one species (e.g., Ledje and Arnason 1996a, 1996b, and further studies using the same sequences, see below), identified as “*B. gabii*” (Genbank identifier X94931), but actually representing *B. neblina* sp. n.; or, two species (Koepfli et al. 2007), identified as *B. alleni* (correctly, sample from Amazonian Peru) and “*B. gabii*” (actually *B. medius orinomus*, from Panama). Koepfli et al. (2007) gave the divergence estimate for these latter two taxa (i.e. *B. alleni* and *B. medius orinomus*) as 2.5–2.8 mya (CI 1.2–5.0 mya). Our results indicate that the earliest divergence within *Bassaricyon*, corresponding to the split between the ancestors of *B. neblina* sp. n. and other *Bassaricyon*, occurred 3.5 mya (CI = 2.1 – 5.2 mya). Sequence divergence in cytochrome *b* between *B. neblina* sp. n. and other *Bassaricyon* (including specimens of *B. m. medius* collected in regional

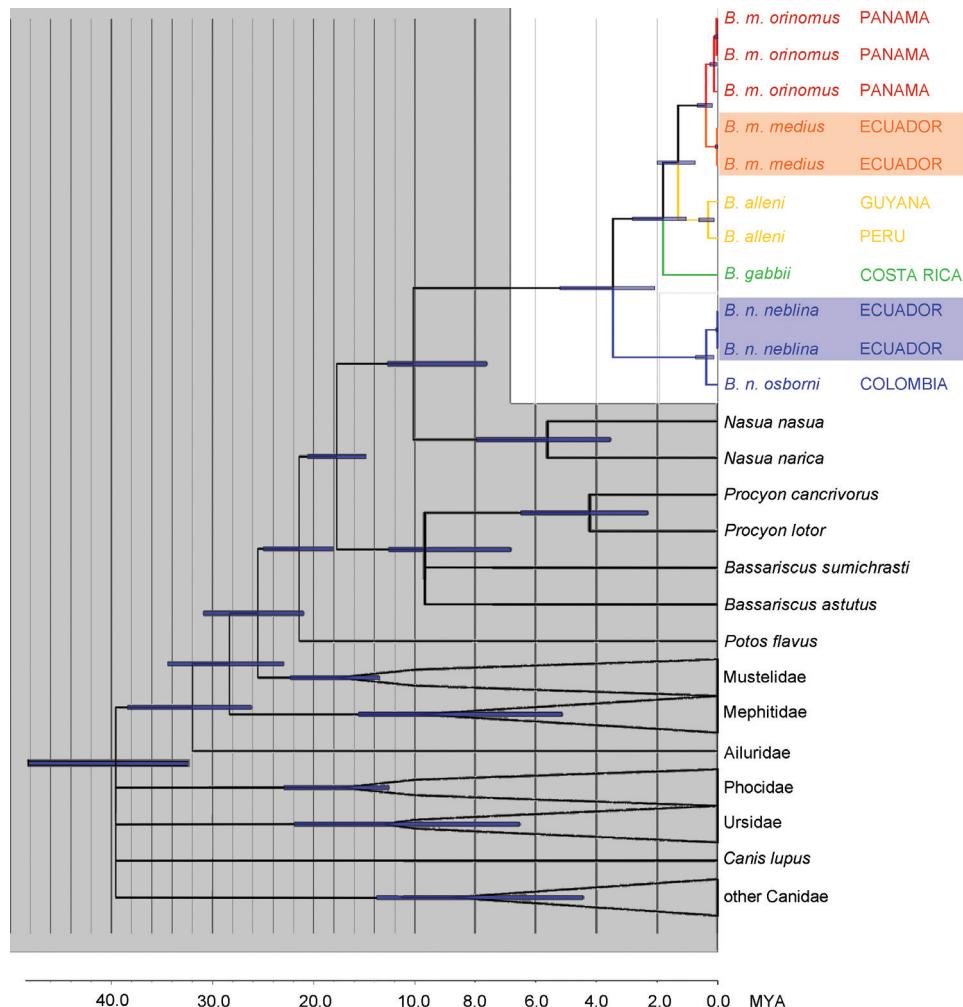


Figure 1. Phylogeny of the genus *Bassaricyon*. Phylogeny generated from the concatenated *CHRNA1* and cytochrome *b* sequences. All analyses consistently recovered the same relationships with high support. Divergence dating was generated in BEAST; bars show the 95% confidence interval at each node. Branches without support are collapsed and outgroup clades have been collapsed, leaving monophyletic groupings with 100% support. Data for *CHRNA1* are missing for *B. gabbii*, for which DNA was extracted from a museum skull. All nodes in *Bassaricyon* have 1.00 Bayesian posterior probability, except the split between *B. gabbii* and *B. allenii/B. medius* (0.97 Bayesian posterior probability). Non-focal and outgroup taxa are shaded in gray, *Bassaricyon* species and subspecies are color coded, samples of *B. m. medius* and *B. n. neblina* that were collected within 5 km of each other in Ecuador are shaded.

sympatry with *B. neblina* sp. n. in the Western Andes of Ecuador) is 9.6-11.3% (Table 2). Cytochrome *b* sequence divergences between *B. gabbii*, *B. medius*, and *B. allenii* are 6-7% (Table 2). Subspecific distances (see Systematics, below, for discussion of subspecies boundaries) are 1.6-2.0% within *B. medius* (between *B. m. medius* and *B. m.*

Table 2. Percentage sequence divergence in cytochrome *b* sequences (Kimura 2-Parameter) among specimens of *Bassaricyon* (numbers 1-11) and other Procyonidae (numbers 12-18) in our analyses (see Table 1, Figure 1). Numbers across the top row match numbered samples in the vertical column.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. <i>B. m. orinomus</i> (Panama)																	
2. <i>B. m. orinomus</i> (Panama)	0.2																
3. <i>B. m. orinomus</i> (Panama)	0.3	0.4															
4. <i>B. m. medius</i> (Ecuador)	1.9	1.9	1.6														
5. <i>B. m. medius</i> (Ecuador)	1.9	2.0	1.6	0.1													
6. <i>B. alleni</i> (Guyana)	6.9	7.0	6.6	7.2	7.4												
7. <i>B. alleni</i> (Peru)	6.3	6.4	6.0	6.3	6.5	1.3											
8. <i>B. gabbii</i> (Costa Rica)	7.3	7.1	7.0	6.9	6.7	6.3	6.6										
9. <i>B. n. neblina</i> (Ecuador)	10.1	10.1	9.8	10.4	10.6	11.3	11.0	9.9									
10. <i>B. n. neblina</i> (Ecuador)	10.1	10.1	9.8	10.5	10.6	11.3	11.0	9.9	0.0								
11. <i>B. n. osborni</i> (Colombia)	10.0	9.9	9.6	10.3	10.4	11.2	10.6	10.4	1.6	1.6							
12. <i>Potos flavus</i>	28.7	28.9	28.7	29.5	29.5	29.8	29.0	28.1	29.8	29.9	28.9						
13. <i>Procyon lotor</i>	34.8	34.3	34.3	35.2	34.9	35.6	34.9	33.0	33.8	33.7	32.7	27.3					
14. <i>Procyon cancrivorus</i>	31.9	31.2	31.2	32.2	32.0	32.1	29.9	31.9	32.0	31.8	30.4	29.4	13.1				
15. <i>Bassariscus astutus</i>	30.7	30.5	30.0	29.8	30.0	30.8	30.0	29.4	29.3	29.1	29.5	29.6	20.7	17.8			
16. <i>Bassariscus sumichrasti</i>	28.1	27.4	27.7	27.7	27.9	27.7	25.7	28.3	26.2	26.1	25.6	26.8	17.1	18.3	15.8		
17. <i>Nasua nasua</i>	26.8	26.7	26.7	28.1	28.4	25.4	24.1	25.7	25.0	24.8	24.1	35.6	35.8	30.3	30.5	29.1	
18. <i>Nasua narica</i>	30.3	29.7	30.0	30.2	30.0	29.0	29.2	28.8	25.1	25.1	24.2	31.3	29.7	26.4	27.3	26.3	20.4

orinomus) and 1.6% within *B. neblina* sp. n. (between *B. n. neblina* subsp. n. and *B. n. osborni* subsp. n., the two subspecies for which we have molecular data).

We obtained the highest bootstrap and posterior probability support values (100% and 1.0 respectively) for relationships within *Bassaricyon* with every method of phylogenetic inference that was used in this study. The single exception was that the topology that recovered the node uniting *B. alleni* and *B. medius* to the exclusion of *B.*

gabbi was assigned a slightly lower Bayesian posterior probability value of 0.97, but all other methods lent full support to this topology (*B. gabbi*, (*B. medius*, *B. alleni*)). These results were also well-supported by our comparisons of morphological characters and together lend strong support for this scenario as being an accurate representation of the evolutionary history of diversification within *Bassaricyon*.

Biogeography

The historical biogeographic reconstruction for the Procyonidae using the DEC model sets Central America as the likely center of origin of crown-group procyonids (Figure 2) (though we note that the family has many extinct, Eocene to Miocene representatives in North America and Europe). Major splits within the family appear to have occurred in Central America previous to the formation of the Panamanian Isthmus, and all the dispersal events resulting in the extant species have occurred within the last 10 million years. All those dispersal events involving southward movements seem to have occurred up to *circa* 6 mya, coinciding with the initial uplift of the Panamanian Isthmus, and, presumably once it was consolidated, with the Great American Biotic Interchange (GABI) (Figures 1–2). The clade containing all olingo species is likely to have originated directly as a result of the formation of the Panamanian Isthmus, and provides evidence of a complex pattern of dispersal events out of Central America (Figure 2).

Morphology and morphometrics

Our study of *Bassaricyon* taxonomy originally began with close examination of craniodental traits of museum specimens, which quickly revealed to the first author the existence of *B. neblina* sp. n., which is highly distinctive morphologically. Close examinations of skins and skulls revealed clear differences in qualitative traits, and in external and craniodental measurements and proportions, between the four principal *Bassaricyon* lineages identified in this paper (which we recognize taxonomically as *B. neblina* sp. n., *B. gabbi*, *B. alleni*, and *B. medius*; Figures 3–5). Externally, these especially include differences in body size, pelage coloration, pelage length, relative length of the tail, and relative size of the ears (Figure 3, Table 5). Cranioidentally, these especially include differences in skull size, relative size of the premolars and molars, configuration of molar cusps, relative size of the auditory bullae and external auditory meati, and the shape of the postdental palatal shelf (Figures 4–5, Tables 3–4). These and other differences are discussed in greater detail in the species accounts provided later in the paper.

Principal component analyses of cranial and dental measurements support our molecular results in clearly identifying a fundamental morphometric separation between the Olinguito (*B. neblina* sp. n.) and all other *Bassaricyon* taxa, in separate comparisons involving both males and females (Figures 6–7, Appendix 1). When first and

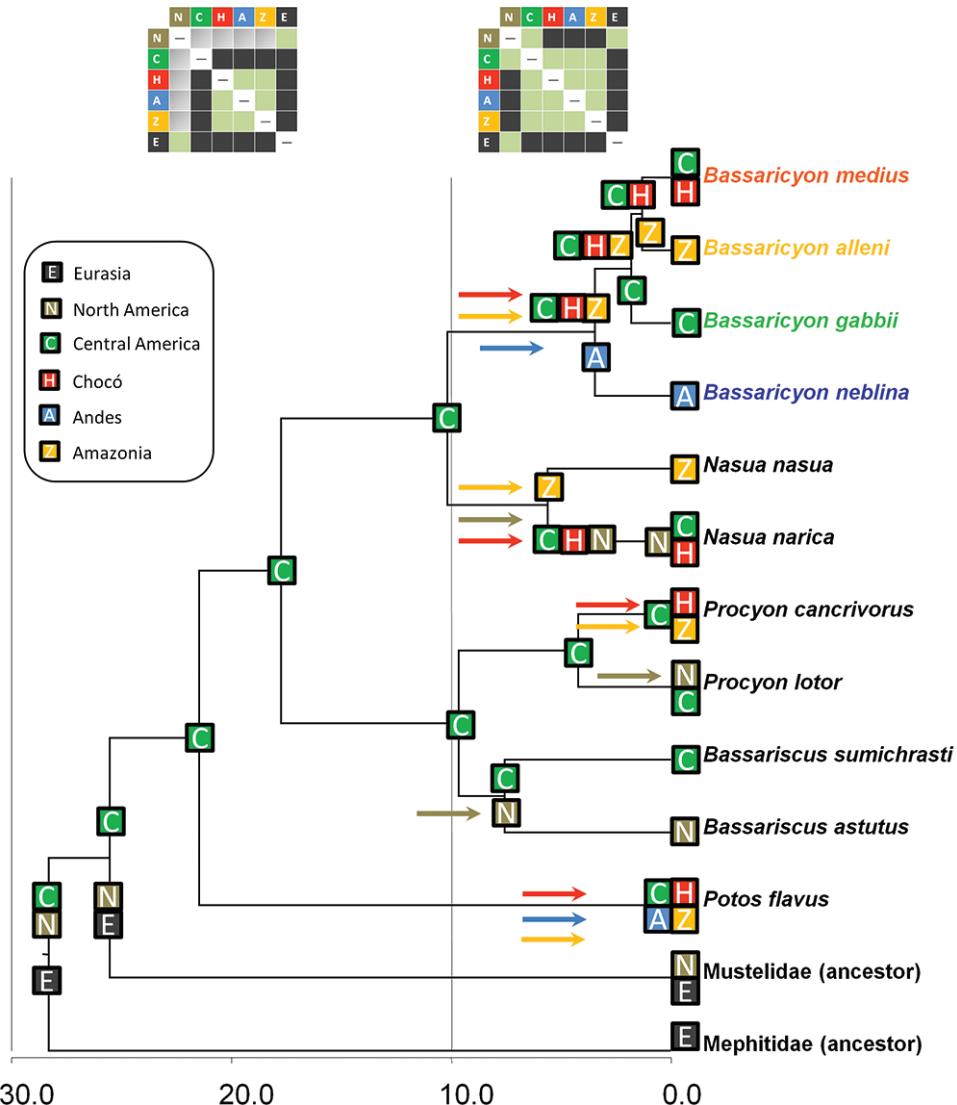


Figure 2. Historical biogeography of procyonids. Reconstructed under the DEC model implemented in Lagrange. See legend for geographical areas used in the analysis. Colored squares at the tip of the branches reflect the distribution of taxa, and previously inferred distributions of the ancestors of mustelids and mephitids. Colored squares at the nodes represent the geographic ranges with the highest probabilities in the DEC model inherited by each descendant branch. Colored arrows reflect dispersal events between ancestral and derived areas, with colors matching with recipient areas. Upper boxes: different dispersal constraints at time intervals 0–10 mya and 10–30 mya, the former to simulate the effect of the land bridge formation between Central and South America, the latter restricted dispersal due to the absence of the land bridge; the cells in green indicate no restriction to dispersal, cells in gray indicate a reduction by half in dispersal capability, and cells in black do not allow dispersal. Timescale in millions of years before present (mya).



Figure 3. Illustrations of the species of *Bassaricyon*. From top to bottom, *Bassaricyon neblina* sp. n. (*B. n. ruber* subsp. n. of the western slopes of the Western Andes of Colombia), *Bassaricyon medius* (*B. m. orinomus* of eastern Panama), *B. alleni* (Peru), and *B. gabbii* (Costa Rica, showing relative tail length longer than average). Artwork by Nancy Halliday.

second principal components are juxtaposed in a bivariate plot, Olinguito specimens demonstrate clear morphometric separation from all other *Bassaricyon*, despite overlap between these clusters in body size (as indicated by overlap in the first principal component, on which all or most variables in the analysis show positive [males] or negative [females] loadings). Despite smaller average body size compared to other *Bassaricyon*, the morphometric distinctness of Olinguito specimens is reflected not especially in small size but rather primarily by separation along the second principal component, indicating trenchant differences in overall shape and proportion, especially reflecting consistent differences in the molars, auditory bullae, external auditory meati, and pal-

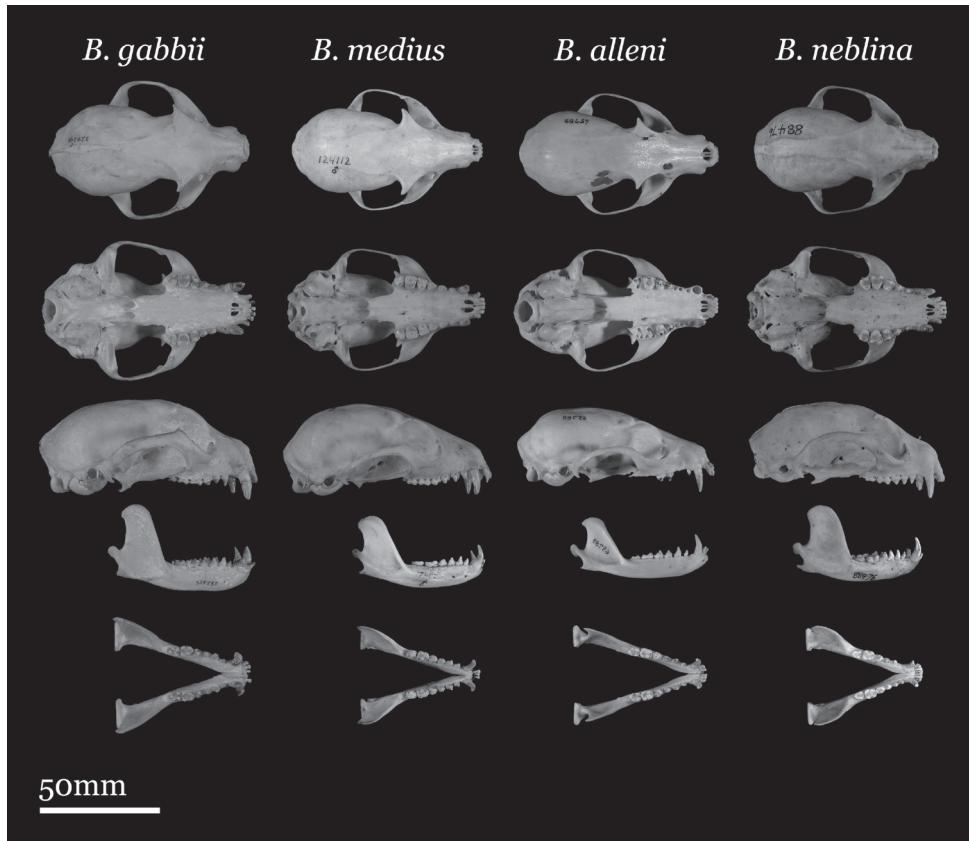


Figure 4. Skulls of adult male *Bassaricyon*. From left to right: *Bassaricyon gabbii* (USNM 324293, Cerro Punta, 1700 m, Chiriquí Mountains, Panama); *B. medius medius* (MVZ 124112, Dagua, 1800 m, Colombia); *B. allenii* (FMNH 65789, Chanchamayo, 1200 m, Junin, Peru); *B. neblina osborni* (FMNH 88476, Munchique, 2000 m, Cauca, Colombia). Scale bar = 50 mm.

ate, in which the Olinguito differs strongly and consistently from other *Bassaricyon* (Figures 6–7, Tables 3–4, Appendix 1).

The lower elevation olingo taxa *B. gabbii*, *B. medius*, and *B. allenii* are not separable in most principal component analyses of craniodental measurements (e.g., Figures 6–7), but discriminant function analyses of craniodental measurements (e.g., Figure 8, showing separation of male skulls) separates them into discrete clusterings with few misclassifications, and identifies some of the more important craniodental traits that help to distinguish between them (Appendix 1). These (and other, qualitative) craniodental distinctions are complemented by differences in pelage features and genetic divergences that we discuss below.

Because of marked and consistent differences in body size between the two regional populations of *B. medius* (one distributed in western South America, the other primarily distributed in Panama), we choose to recognize these two as separate subspecies (*B. m. medius* and *B. m. orinomus*, respectively, Tables 6–7). The Olinguito likewise

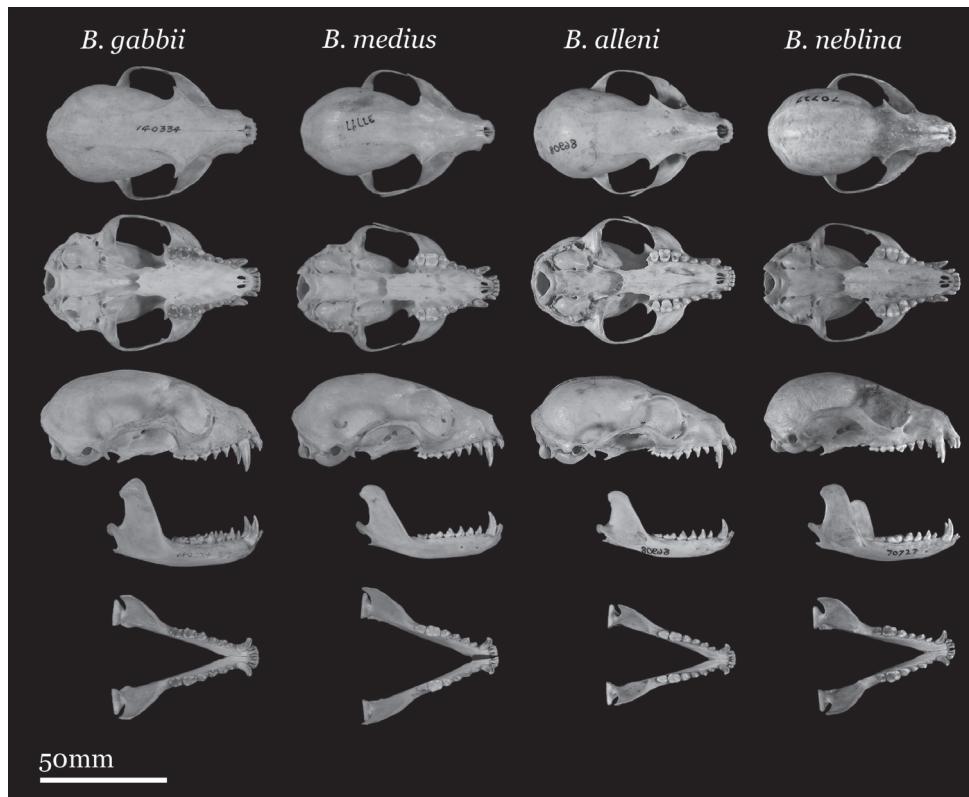


Figure 5. Skulls of adult female *Bassaricyon*. From left to right: *Bassaricyon gabbii* (AMNH 140334, Lajas Villa, Costa Rica); *B. medius orinomus* (AMNH 37797, Puerta Valdivia, Antioquia District, Colombia); *B. alleni* (FMNH 86908, Santa Rita, Rio Nanay, Maynas, Loreto Region, Peru); *B. neblina hershkovitzii* (FMNH 70727, San Antonio, Agustín, Huila District, Colombia). Scale bar = 50 mm.

shows a clear pattern of geographic variation, with different regional populations in the Northern Andes showing consistent differences in craniodental size and morphology (Figures 9–10, Table 8, Appendix 1), as well as pelage coloration and length. We recognize four distinctive subspecies of the Olinguito throughout its recorded distribution, as discussed in the description of *B. neblina* sp. n., below. Two of these subspecies are included in our genetic comparisons; genetic comparisons involving the remaining two subspecies remain a goal for the future.

Bioclimatic range modeling

Distribution models for all species are judged to have performed well based on their high values for ‘area under the curve’ (AUC) and unregularized test gain (Table 9), as well as their fit of the final prediction to the locality data (Figures 11–12). There was relatively low impact of withholding test data from these models, as indicated by the

Table 3. Cranial measurements for olingo species (compiled separately by sex). For each measurement, means are provided, \pm standard deviation, with ranges in parentheses.

		<i>B. gabbii</i> <i>n</i> = 11 ♂♂, 11 ♀♀	<i>B. medius</i> <i>n</i> = 18 ♂♂, 27 ♀♀	<i>B. allenii</i> <i>n</i> = 12 ♂♂, 17 ♀♀	<i>B. neblina</i> <i>n</i> = 10 ♂♂, 9 ♀♀
CBL	♂♂	80.8 \pm 1.50 (78.1 - 83.0)	79.4 \pm 2.67 (74.5 - 85.1)	79.4 \pm 1.81 (76.5 - 82.8)	74.5 \pm 3.26 (70.1 - 79.5)
	♀♀	78.2 \pm 1.75 (75.0 - 80.2)	77.3 \pm 2.70 (70.8 - 82.3)	77.0 \pm 2.24 (73.1 - 80.5)	75.1 \pm 1.49 (72.9 - 77.9)
ZYG	♂♂	55.2 \pm 2.76 (49.5 - 58.7)	52.0 \pm 2.66 (48.3 - 56.7)	51.6 \pm 1.02 (49.0 - 52.8)	50.1 \pm 3.02 (46.2 - 54.4)
	♀♀	51.3 \pm 1.90 (48.1 - 54.4)	50.0 \pm 2.50 (44.4 - 54.0)	50.2 \pm 0.99 (48.6 - 52.2)	49.0 \pm 2.69 (44.6 - 53.0)
BBC	♂♂	36.1 \pm 0.86 (34.3 - 37.6)	35.1 \pm 1.16 (32.9 - 37.5)	35.4 \pm 0.80 (34.2 - 36.8)	34.6 \pm 1.62 (32.4 - 37.5)
	♀♀	35.7 \pm 1.34 (33.1 - 37.5)	34.6 \pm 1.20 (32.0 - 37.2)	34.9 \pm 0.91 (33.3 - 36.8)	34.2 \pm 1.62 (31.0 - 36.6)
HBC	♂♂	28.7 \pm 0.88 (26.4 - 29.7)	27.6 \pm 0.84 (26.5 - 29.3)	27.4 \pm 0.73 (26.2 - 28.5)	27.4 \pm 0.61 (26.5 - 28.3)
	♀♀	27.9 \pm 0.74 (26.9 - 28.8)	26.9 \pm 0.90 (25.4 - 28.5)	26.9 \pm 0.63 (26.0 - 28.1)	26.5 \pm 0.93 (24.9 - 27.8)
MTR	♂♂	28.5 \pm 0.50 (27.8 - 29.3)	28.6 \pm 0.87 (27.0 - 30.4)	28.4 \pm 0.83 (26.5 - 29.5)	26.5 \pm 1.38 (24.5 - 28.7)
	♀♀	27.3 \pm 1.02 (26.0 - 29.0)	27.7 \pm 0.90 (25.6 - 29.1)	27.3 \pm 0.69 (26.1 - 28.5)	26.9 \pm 0.88 (25.8 - 28.3)
CC	♂♂	18.7 \pm 1.12 (17.2 - 20.4)	16.4 \pm 0.92 (15.0 - 17.9)	16.8 \pm 0.51 (15.8 - 17.6)	15.9 \pm 0.94 (14.7 - 17.1)
	♀♀	16.9 \pm 0.76 (15.6 - 17.9)	15.7 \pm 0.80 (14.5 - 17.2)	15.9 \pm 0.55 (14.8 - 16.8)	15.7 \pm 0.47 (14.9 - 16.4)
WPP	♂♂	11.3 \pm 1.27 (9.0 - 12.9)	10.3 \pm 0.95 (8.4 - 12.1)	10.4 \pm 0.82 (8.7 - 11.8)	11.7 \pm 1.05 (10.6 - 14.0)
	♀♀	10.7 \pm 0.99 (9.3 - 12.7)	10.3 \pm 0.90 (9.0 - 13.0)	9.9 \pm 0.89 (8.2 - 11.7)	11.6 \pm 0.87 (10.5 - 12.7)
LPP	♂♂	12.3 \pm 0.99 (10.7 - 14.0)	10.2 \pm 0.88 (7.9 - 11.7)	10.8 \pm 1.21 (9.3 - 12.9)	11.2 \pm 1.24 (9.2 - 12.7)
	♀♀	10.8 \pm 0.77 (9.7 - 12.0)	10.1 \pm 0.90 (8.1 - 11.8)	10.4 \pm 0.67 (8.7 - 11.6)	11.1 \pm 0.82 (9.7 - 12.3)
LAB	♂♂	13.8 \pm 0.63 (12.9 - 14.7)	14.0 \pm 0.81 (12.8 - 15.6)	15.1 \pm 0.76 (14.1 - 16.8)	11.8 \pm 0.76 (10.9 - 13.3)
	♀♀	13.8 \pm 0.67 (12.9 - 14.8)	14.0 \pm 0.80 (12.6 - 15.2)	14.4 \pm 0.81 (13.0 - 15.6)	12.2 \pm 0.51 (11.0 - 12.7)
EAM	♂♂	3.6 \pm 0.47 (2.6 - 4.2)	3.9 \pm 0.33 (3.4 - 4.5)	3.8 \pm 0.40 (3.2 - 4.5)	2.9 \pm 0.22 (2.5 - 3.1)
	♀♀	3.6 \pm 0.39 (3.0 - 4.2)	3.9 \pm 0.30 (3.5 - 4.7)	3.8 \pm 0.36 (3.2 - 4.4)	3.2 \pm 0.33 (2.6 - 3.5)

low Mean Test AUC values. These values are lowest for *B. allenii*, probably reflecting its larger distribution relative to the variation of environmental data (Phillips et al. 2006). The strongest environmental predictors for *B. neblina* sp. n. were seasonal variation

Table 4. Selected dental measurements of olingo species. For each measurement, means are provided, \pm standard deviation, with ranges in parentheses.

	<i>B. gabbii</i> <i>n</i> = 22	<i>B. medius</i> <i>n</i> = 45	<i>B. alleni</i> <i>n</i> = 34	<i>B. neblina</i> <i>n</i> = 19
p1 width	1.7 \pm 0.17 (1.4 - 2.1)	1.7 \pm 0.13 (1.4 - 2.0)	1.7 \pm 0.12 (1.5 - 1.9)	1.6 \pm 0.13 (1.4 - 1.8)
p2 width	2.4 \pm 0.24 (2.0 - 2.8)	2.2 \pm 0.18 (1.8 - 2.6)	2.2 \pm 0.15 (1.9 - 2.5)	2.1 \pm 0.17 (1.9 - 2.5)
p3 width	2.7 \pm 0.21 (2.3 - 3.0)	2.5 \pm 0.18 (2.2 - 2.9)	2.6 \pm 0.16 (2.2 - 2.9)	2.4 \pm 0.22 (2.1 - 2.9)
p ⁴ width	3.4 \pm 0.27 (3.0 - 3.9)	3.2 \pm 0.18 (2.8 - 3.6)	3.4 \pm 0.21 (2.8 - 3.7)	3.3 \pm 0.15 (3.0 - 3.7)
P2 width	2.4 \pm 0.24 (2.1 - 2.9)	2.3 \pm 0.19 (1.9 - 2.8)	2.2 \pm 0.17 (1.9 - 2.7)	2.1 \pm 0.19 (1.8 - 2.5)
P3 width	2.9 \pm 0.22 (2.5 - 3.3)	3.0 \pm 0.29 (2.5 - 3.6)	3.0 \pm 0.22 (2.6 - 3.5)	2.9 \pm 0.21 (2.6 - 3.4)
P4 length	4.4 \pm 0.24 (3.9 - 4.8)	4.2 \pm 0.27 (3.6 - 4.9)	4.2 \pm 0.20 (3.8 - 4.6)	4.5 \pm 0.24 (4.1 - 4.9)
P4 width	5.1 \pm 0.35 (4.5 - 5.6)	4.7 \pm 0.26 (4.2 - 5.4)	4.8 \pm 0.23 (4.4 - 5.6)	5.0 \pm 0.40 (4.5 - 5.9)
M1 length	5.0 \pm 0.27 (4.4 - 5.4)	5.0 \pm 0.29 (4.3 - 5.6)	5.1 \pm 0.21 (4.6 - 5.5)	5.3 \pm 0.35 (4.8 - 6.1)
M1 width	5.5 \pm 0.30 (4.7 - 5.9)	5.3 \pm 0.32 (4.7 - 5.9)	5.5 \pm 0.28 (4.9 - 6.0)	5.8 \pm 0.31 (5.4 - 6.4)
M2 length	3.7 \pm 0.32 (2.8 - 4.1)	4.0 \pm 0.25 (3.2 - 4.4)	3.8 \pm 0.27 (3.3 - 4.4)	3.8 \pm 0.35 (3.3 - 4.4)
M2 width	4.6 \pm 0.38 (4.0 - 5.3)	4.7 \pm 0.27 (4.1 - 5.2)	4.7 \pm 0.28 (4.0 - 5.2)	4.8 \pm 0.24 (4.4 - 5.4)
m1 length	5.6 \pm 0.31 (5.0 - 6.3)	5.7 \pm 0.26 (4.9 - 6.2)	5.6 \pm 0.22 (5.2 - 6.0)	5.8 \pm 0.29 (5.4 - 6.3)
m1 width	4.3 \pm 0.29 (3.8 - 4.9)	4.3 \pm 0.21 (3.9 - 4.7)	4.3 \pm 0.23 (3.7 - 4.8)	4.8 \pm 0.22 (4.5 - 5.3)
m2 length	4.8 \pm 0.25 (4.4 - 5.3)	5.1 \pm 0.36 (4.2 - 5.7)	4.8 \pm 0.25 (4.4 - 5.4)	5.0 \pm 0.35 (4.4 - 5.6)
m2 width	3.8 \pm 0.24 (3.3 - 4.2)	3.7 \pm 0.24 (3.2 - 4.2)	3.7 \pm 0.19 (3.3 - 4.0)	3.8 \pm 0.17 (3.5 - 4.1)

in temperature (suitability declines with higher variation, after sharp threshold) and the temperature of the wettest quarter (negative relationship). The annual range of temperatures was the most important predictor for the *B. gabbii* and *B. medius* distributions (both sharp negative relationships). *Bassaricyon alleni* was the only one of the four species to have an ecological biome ranked as one of the top predictors (Tropical Moist Broadleaf Forests as highly suitable).

The full Maxent distribution models predict the suitability of habitat across South and Central America (Figure 11). To make the binary prediction maps (Figure 12) we excluded areas with high probability that were disjunct from areas where specimens have been recorded (e.g., western Venezuela excluded from the map for *B. neblina*

Table 5. External measurements of olingo species. For each measurement, means are provided, \pm standard deviation, with ranges in parentheses.

	<i>B. gabbi</i> <i>n</i> = 13	<i>B. medius</i> <i>n</i> = 36	<i>B. allen</i> <i>n</i> = 27	<i>B. neblina</i> <i>n</i> = 19
TL	873 \pm 54.8 (785 - 970)	819 \pm 60.5 (680 - 905)	842 \pm 50.6 (705 - 985)	745 \pm 33.7 (660 - 820)
Tail	445 \pm 40.3 (400 - 521)	441 \pm 44.6 (350 - 520)	450 \pm 28.8 (401 - 530)	390 \pm 21 (335 - 424)
HF	84 \pm 8.7 (65 - 100)	81 \pm 7.3 (58 - 92)	81 \pm 5.8 (70 - 92)	76 \pm 6.9 (60 - 86)
Ear	36 \pm 4.7 (25 - 44)	37 \pm 5.4 (25 - 44)	37 \pm 3.4 (30 - 43)	34 \pm 4.3 (25 - 39)
Mass (g)	1382 \pm 165 (1136 - 1580)	1076 \pm 71.6 (915 - 1200)	1336 \pm 152 (1100 - 1500)	872 \pm 169 (750 - 1065)
HB	428 \pm 27.9 (373 - 470)	379 \pm 23.2 (310 - 415)	391 \pm 29.3 (304 - 455)	355 \pm 21.1 (325 - 400)
Tail/HB	1.04 \pm 0.1 (0.9 - 1.2)	1.16 \pm 0.1 (1.0 - 1.4)	1.15 \pm 0.08 (1.0 - 1.3)	1.10 \pm 0.08 (1.0 - 1.2)

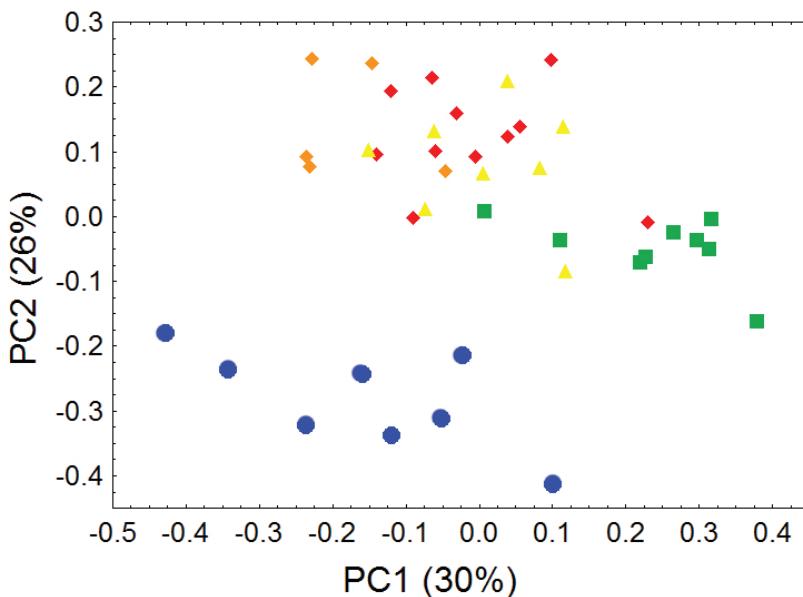


Figure 6. Morphometric distinction between Olinguitos and other *Bassaricyon*, males. Morphometric dispersion (first two components of a principal component analysis) of 41 adult male *Bassaricyon* skulls based on 21 craniodental measurements (see Appendix 1, Table A1). The most notable morphometric distinction is between the Olinguito (blue circles) and all other *Bassaricyon* taxa. The plot also demonstrates substantial morphometric variability across geographic populations of the Olinguito, which we characterize with the description of four subspecies across different Andean regions. Symbols: blue circles (*B. neblina*), green squares (*B. gabbi*), yellow triangles (*B. allen*), orange diamonds (*B. medius medius*), red diamonds (*B. medius orinomus*).

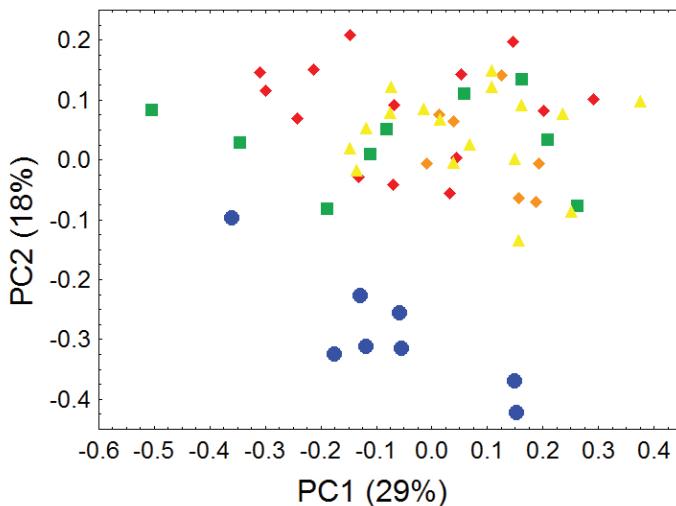


Figure 7. Morphometric distinction between Olinguitos and other *Bassaricyon*, females. Morphometric dispersion (first two components of a principal component analysis) of 55 adult female *Bassaricyon* skulls based on 24 craniodental measurements (see Appendix 1, Table A2). The most notable morphometric distinction is between the Olinguito (blue circles) and all other *Bassaricyon* taxa. The plot also demonstrates substantial morphometric variability across geographic populations of the Olinguito, which we characterize with the description of four subspecies across different Andean regions. Symbols: blue circles (*B. neblina*), green squares (*B. gabbi*), yellow triangles (*B. alleni*), orange diamonds (*B. medius medius*), red diamonds (*B. medius orinomus*).

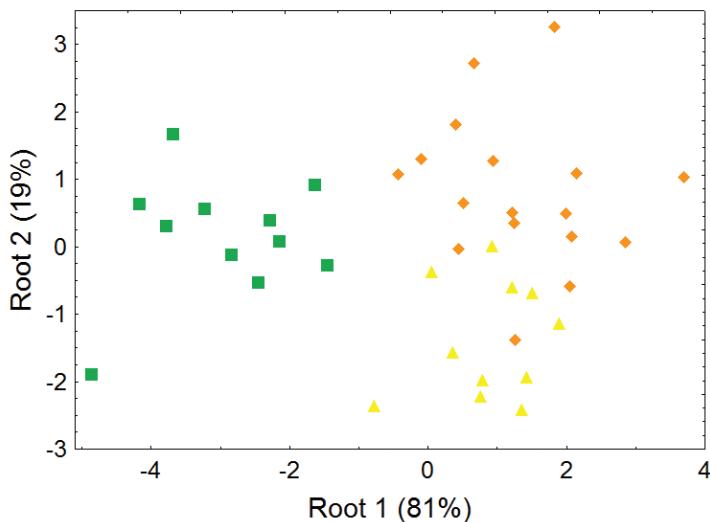


Figure 8. Morphometric distinction between species of *Bassaricyon*, excluding the Olinguito, adult males. Morphometric dispersion (first two variates of a discriminant function analysis) of 39 adult male *Bassaricyon* skulls based on 8 craniodental measurements (see Appendix 1, Table A3). Symbols: green squares (*B. gabbi*), yellow triangles (*B. alleni*), orange diamonds (*B. medius*).

Table 6. Cranial measurements for the two subspecies of *Bassaricyon medius*. For each measurement, means are provided, \pm standard deviation, with ranges in parentheses.

		<i>B. medius medius</i>	<i>B. medius orinomus</i>
		W Colombia, W Ecuador	C Panama to N Colombia
		n = 5 ♂♂, 7 ♀♀	n = 12 ♂♂, 17 ♀♀
CBL	♂♂	77.2 \pm 1.81 (74.5 - 78.8)	80.3 \pm 2.50 (76.2 - 85.1)
	♀♀	75.4 \pm 1.65 (72.4 - 76.7)	78.8 \pm 1.72 (75.5 - 82.3)
ZYG	♂♂	50.2 \pm 1.14 (48.9 - 51.2)	53.0 \pm 2.57 (48.9 - 56.7)
	♀♀	48.5 \pm 1.69 (46.5 - 51.0)	51.2 \pm 1.98 (47.4 - 54.0)
BBC	♂♂	34.0 \pm 0.80 (32.9 - 34.8)	35.6 \pm 0.98 (34.0 - 37.5)
	♀♀	34.4 \pm 0.41 (33.7 - 35.0)	35.0 \pm 1.15 (32.8 - 37.2)
HBC	♂♂	28.2 \pm 1.06 (27.1 - 29.3)	27.4 \pm 0.62 (26.6 - 28.3)
	♀♀	26.8 \pm 0.89 (26.1 - 28.5)	27.0 \pm 0.89 (25.4 - 28.5)
MTR	♂♂	28.5 \pm 0.97 (27.3 - 29.8)	28.7 \pm 0.90 (27.0 - 30.4)
	♀♀	27.1 \pm 0.78 (25.6 - 27.9)	28.0 \pm 0.77 (26.4 - 29.1)
CC	♂♂	15.9 \pm 0.69 (15.1 - 17.0)	16.7 \pm 0.94 (15.0 - 17.9)
	♀♀	15.0 \pm 0.46 (14.5 - 15.8)	16.1 \pm 0.71 (14.6 - 17.2)
WPP	♂♂	9.7 \pm 0.95 (8.4 - 10.8)	10.6 \pm 0.91 (8.6 - 12.1)
	♀♀	10.0 \pm 0.57 (9.1 - 10.6)	10.3 \pm 1.04 (9.0 - 13.0)
LPP	♂♂	9.4 \pm 1.03 (7.9 - 10.6)	10.5 \pm 0.64 (9.8 - 11.7)
	♀♀	9.8 \pm 0.84 (8.9 - 11.3)	10.2 \pm 1.01 (8.1 - 11.8)
LAB	♂♂	13.6 \pm 0.72 (12.8 - 14.6)	14.2 \pm 0.84 (13.1 - 15.6)
	♀♀	13.4 \pm 0.45 (12.6 - 13.9)	14.3 \pm 0.73 (12.8 - 15.2)
EAM	♂♂	3.9 \pm 0.47 (3.4 - 4.5)	3.9 \pm 0.27 (3.5 - 4.4)
	♀♀	3.9 \pm 0.34 (3.5 - 4.4)	3.9 \pm 0.28 (3.6 - 4.7)

Table 7. External measurements for the two subspecies of *Bassaricyon medius*. For each measurement, means are provided, \pm standard deviation, with ranges in parentheses.

		<i>B. m. medius</i> W Colombia, W Ecuador n = 12	<i>B. m. orinomus</i> C Panama to N Colombia n = 24
TL		754 \pm 49.7 (680 - 819)	844 \pm 42.9 (770 - 905)
Tail		392 \pm 29.1 (350 - 435)	460 \pm 33.6 (400 - 520)
HF		73 \pm 5.4 (58 - 79)	85 \pm 3.5 (77 - 92)
Ear		32 \pm 4.8 (25 - 40)	39 \pm 4 (30 - 44)
Mass (g)		1058 \pm 146 (915 - 1200)	1090 \pm 19.2 (1050 - 1100)
HB		362 \pm 29.5 (310 - 415)	385 \pm 17.2 (355 - 410)
Tail/HB		1.1 \pm 0.09 (0.97 - 1.24)	1.2 \pm 0.08 (1.04 - 1.35)

sp. n., central and eastern Brazil excluded from the *B. allenii* map, northern Central America excluded from the *B. medius* map, South America excluded from the *B. gabbi* map). For *B. neblina* sp. n. we excluded areas of high probability from the Eastern Cordillera of Colombia and the Andes of southern Ecuador and northern Peru because of the lack of specimens. Likewise, predicted suitable habitat for *B. gabbi* in northern

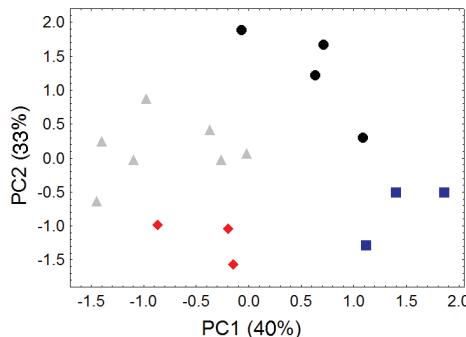


Figure 9. Morphometric distinction between Olinguito subspecies. Both sexes combined. Morphometric dispersion (first two components of a principal component analysis) of 17 adult skulls based on 13 cranial measurements (see Appendix 1, Table A4). (Dental measurements also discretely partition these subspecies in a separate principal component analysis, not shown.) Black dots = *B. neblina neblina*; gray triangles = *B. n. osborni*; red diamonds = *B. n. ruber*; blue squares = *B. n. hershkovitzi*.

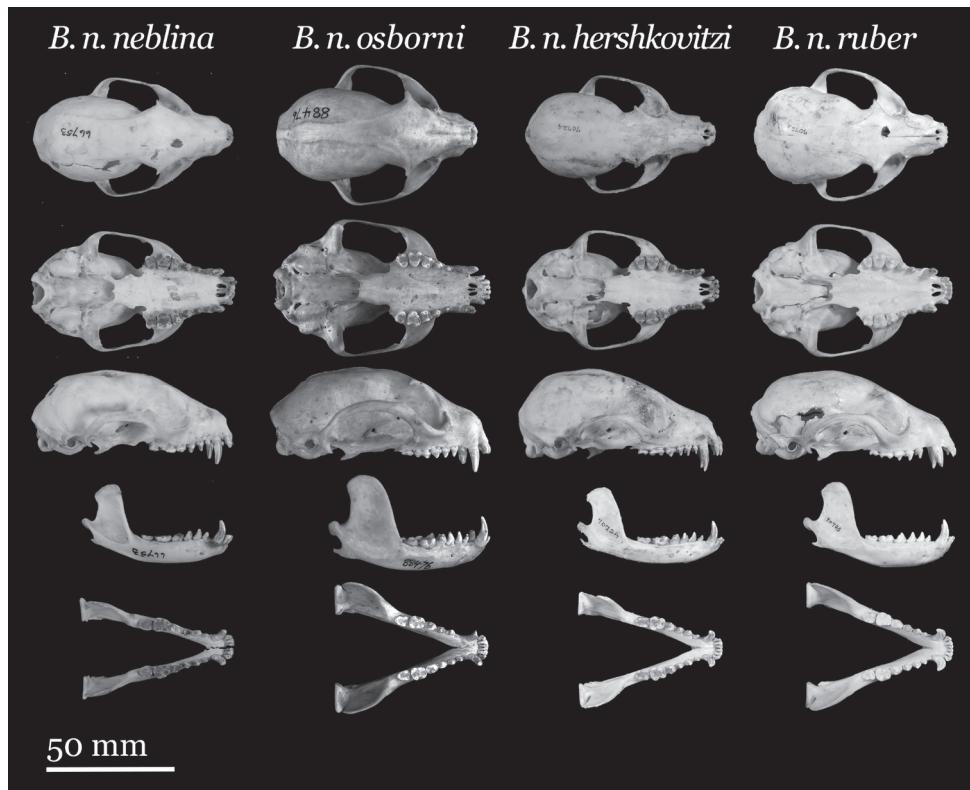


Figure 10. Skulls of Olinguito subspecies. From left to right: *Bassaricyon neblina neblina* (AMNH 66753, holotype, old adult female, Las Maquinas, Ecuador); *B. n. osborni* (FMNH 88476, holotype, adult male, Munchique, 2000 m, Cauca Department, Colombia); *B. n. hershkovitzi* (FMNH 70724, paratype, adult male, San Antonio, Agustin, Huila District, Colombia); *B. neblina ruber* (FMNH 70723, paratype, adult male, Guapantel, 2200 m, Urrao, Antioquia Department, Colombia). Scale bar = 50 mm.

Table 8. Dental and cranial measurements of Olinguito (*Bassaricyon neblina*) subspecies. For each measurement, means are provided, \pm standard deviation, with ranges in parentheses.

	<i>B. n. ruber</i> <i>n</i> = 3	<i>B. n. bershkovitzi</i> <i>n</i> = 4	<i>B. n. osborni</i> <i>n</i> = 8	<i>B. n. neblina</i> <i>n</i> = 4
p1 width	1.4 \pm 0.06 (1.4 - 1.5)	1.5 \pm 0.12 (1.4 - 1.6)	1.6 \pm 0.09 (1.6 - 1.8)	1.7 \pm 0.11 (1.5 - 1.8)
p2 width	2.1 \pm 0.14 (1.9 - 2.2)	1.9 \pm 0.06 (1.9 - 2.0)	2.2 \pm 0.15 (2.0 - 2.5)	2.2 \pm 0.17 (2.1 - 2.4)
p3 width	2.4 \pm 0.08 (2.3 - 2.5)	2.2 \pm 0.06 (2.1 - 2.2)	2.5 \pm 0.16 (2.4 - 2.8)	2.4 \pm 0.32 (2.2 - 2.9)
p4 width	3.3 \pm 0.11 (3.2 - 3.4)	3.1 \pm 0.12 (3.0 - 3.3)	3.4 \pm 0.13 (3.2 - 3.7)	3.4 \pm 0.09 (3.3 - 3.5)
P2 width	2.0 (2.0 - 2.0)	1.9 \pm 0.05 (1.8 - 2.0)	2.2 \pm 0.17 (2.1 - 2.5)	2.3 \pm 0.15 (2.2 - 2.5)
P3 width	2.9 \pm 0.17 (2.7 - 3.1)	2.7 \pm 0.10 (2.6 - 2.8)	3.0 \pm 0.19 (2.8 - 3.4)	3.1 \pm 0.15 (2.9 - 3.3)
P4 length	4.3 \pm 0.21 (4.1 - 4.5)	4.2 \pm 0.13 (4.1 - 4.3)	4.5 \pm 0.17 (4.3 - 4.8)	4.7 \pm 0.17 (4.5 - 4.9)
P4 width	4.6 \pm 0.14 (4.5 - 4.8)	5.0 \pm 0.23 (4.8 - 5.3)	4.9 \pm 0.20 (4.6 - 5.1)	5.7 \pm 0.13 (5.6 - 5.9)
M1 length	5.0 \pm 0.12 (5.0 - 5.2)	5.0 \pm 0.25 (4.8 - 5.4)	5.3 \pm 0.23 (5.0 - 5.6)	5.7 \pm 0.4 (5.2 - 6.1)
M1 width	5.5 \pm 0.14 (5.4 - 5.6)	5.5 \pm 0.10 (5.4 - 5.6)	5.8 \pm 0.20 (5.5 - 6.1)	6.2 \pm 0.13 (6.1 - 6.4)
M2 length	3.6 \pm 0.22 (3.5 - 3.9)	3.5 \pm 0.16 (3.3 - 3.7)	4.1 \pm 0.29 (3.6 - 4.4)	3.9 \pm 0.4 (3.3 - 4.2)
M2 width	4.5 \pm 0.13 (4.4 - 4.6)	4.7 \pm 0.03 (4.7 - 4.8)	4.8 \pm 0.20 (4.6 - 5.2)	4.9 \pm 0.3 (4.7 - 5.4)
m1 length	5.5 \pm 0.05 (5.4 - 5.5)	5.8 \pm 0.21 (5.6 - 6.0)	5.8 \pm 0.18 (5.6 - 6.0)	6.2 \pm 0.03 (6.2 - 6.3)
m1 width	4.7 \pm 0.12 (4.6 - 4.8)	4.8 \pm 0.17 (4.7 - 5.0)	4.8 \pm 0.26 (4.5 - 5.3)	5.0 \pm 0.22 (4.7 - 5.2)
m2 length	4.7 \pm 0.39 (4.4 - 5.1)	5.0 \pm 0.37 (4.5 - 5.2)	5.2 \pm 0.26 (4.9 - 5.6)	4.8 \pm 0.22 (4.5 - 5.1)
m2 width	3.7 \pm 0.09 (3.6 - 3.8)	3.7 \pm 0.19 (3.5 - 3.9)	3.9 \pm 0.10 (3.7 - 4.0)	3.9 \pm 0.16 (3.7 - 4.1)
CBL	73.0 \pm 0.58 (72.4 - 73.5)	71.4 \pm 1.13 (70.1 - 72.9)	76.6 \pm 1.64 (75.1 - 79.5)	75.9 \pm 1.4 (74.6 - 77.9)
ZYG	51.1 \pm 2.28 (48.9 - 53.4)	46.7 \pm 0.60 (46.2 - 47.5)	51.7 \pm 1.73 (49.1 - 54.4)	46.9 \pm 1.59 (44.6 - 48)
BBC	36.0 \pm 1.44 (34.7 - 37.5)	32.9 \pm 0.54 (32.4 - 33.6)	35.1 \pm 0.90 (33.9 - 36.6)	33.2 \pm 1.62 (31.0 - 34.9)
HBC	27.7 \pm 0.55 (27.2 - 28.3)	27.6 \pm 0.38 (27.1 - 27.9)	27.2 \pm 0.58 (26.5 - 28.2)	25.8 \pm 0.63 (24.9 - 26.2)
MTR	25.9 \pm 0.22 (25.7 - 26.1)	25.1 \pm 0.56 (24.5 - 25.8)	27.4 \pm 0.78 (26.0 - 28.7)	27.5 \pm 0.56 (27 - 28.3)
CC	15.7 \pm 0.52 (15.4 - 16.3)	14.9 \pm 0.15 (14.7 - 15.0)	16.4 \pm 0.54 (15.5 - 17.1)	15.6 \pm 0.25 (15.4 - 15.9)
WPP	12.1 \pm 0.25 (11.8 - 12.3)	11.8 \pm 1.54 (10.6 - 14.0)	11.8 \pm 0.74 (10.8 - 12.8)	10.9 \pm 0.8 (10.5 - 12.1)
LPP	10.9 \pm 0.54 (10.3 - 11.4)	9.7 \pm 0.34 (9.2 - 9.9)	11.9 \pm 0.56 (11.0 - 12.7)	11.2 \pm 1.05 (9.7 - 12.3)
LAB	11.7 \pm 0.38 (11.4 - 12.1)	11.2 \pm 0.40 (10.9 - 11.8)	12.3 \pm 0.60 (11.2 - 13.3)	12.5 \pm 0.18 (12.3 - 12.7)
EAM	2.7 (2.7 - 2.7)	3.2 \pm 0.16 (3.1 - 3.4)	2.9 \pm 0.29 (2.5 - 3.3)	3.4 \pm 0.05 (3.4 - 3.5)

Table 9. Performance of bioclimatic distribution models for four *Bassaricyon* species using voucherized specimen localities. Mean values are averages of 10 models run, each withholding 20% of data as test localities, while the Full Model AUC used all available data. The mean value for equal training sensitivity and specificity was used as a logistic threshold to create a range map predicting presence/absence.

	Localities	Mean Test AUC (stdev)	Full Model AUC	Mean Unregularized Training Gain	Mean equal training sensitivity and specificity (logistic threshold)
<i>B. alleni</i>	43	0.901 (0.036)	0.939	1.85	0.302
<i>B. gabbi</i>	18	0.977 (0.012)	0.993	4.09	0.222
<i>B. medius</i>	31	0.952 (0.028)	0.988	3.76	0.119
<i>B. neblina</i>	16	0.996 (0.002)	0.998	4.77	0.160

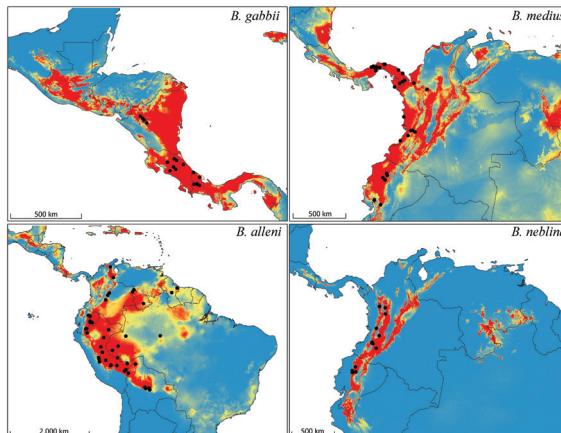


Figure 11. Bioclimatic distribution models and localities for *Bassaricyon* species. Models from MAXENT using all voucherized occurrence records, 19 bioclimatic variables, and one potential habitat variable.

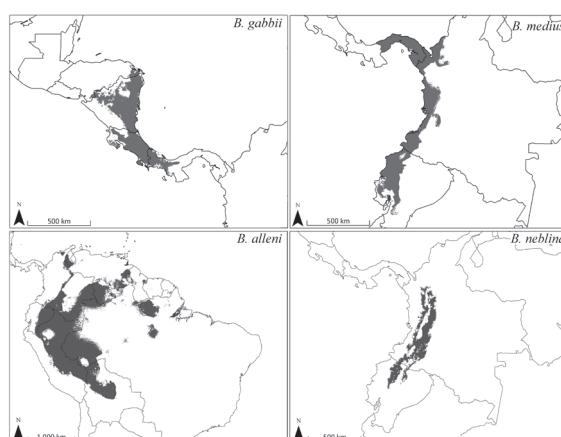


Figure 12. Predicted distribution for *Bassaricyon* species based on bioclimatic models. To create these binary maps we used the average minimum training presence for 10 test models as our cutoff. In addition, we excluded areas of high probability that were outside of the known range of the species if they were separated by unsuitable habitat.

Central America (Honduras, Guatemala) remains unverified by specimen data. Although there are two recent unconfirmed records in the region (Ordóñez Garza et al. 1999–2000), the specific locations of these sightings did not fall in areas predicted as suitable habitat by our models. Finally, the exact area of transition between *B. gabbi* and *B. medi* in Panama remains unclear. All of these regions should be considered high priority areas for future surveys, especially areas identified as potential *B. neblina* sp. n. habitat (see Discussion, below).

The range of *B. neblina* sp. n. is typical of many Andean species in being restricted to wet cloud forest habitats, which are limited in area and also under heavy development pressure. In comparing recent land use (Eva et al. 2004) of suitable historical *B. neblina* habitat, we found that 42% of suitable habitats have been converted to agriculture or urban areas, and 21% remain in natural but largely unforested conditions. Thus we predict that only 37% (40,760 km²) of appropriate Olinguito habitats remain forested.

Systematics

Bassaricyon neblina sp. n.

<http://zoobank.org/94DDB038-2111-44D1-A940-766BF8F15E51>

http://species-id.net/wiki/Bassaricyon_neblina

Holotype: We designate as the holotype of *neblina* specimen number 66753 in the mammalogy collection of the American Museum of Natural History, New York, a skin and complete skull of an old adult female, from Las Máquinas (= Las Machinas [see Voss 1988:474], *circa* 00°32'S, 78°39'W, 2130 m), Pichincha Province, Ecuador, collected 21 September 1923 by G.H.H. Tate.

Referred specimens: QCAZ 0159, partial skin, Otonga Reserve, 1800 m, Cotopaxi Province, Ecuador; MECN 2177, adult female, skin and skull, La Cantera 2300 m, Cotopaxi Province, Ecuador; QCAZ 8661, young adult female, skin, skull, and postcranial skeleton, Otonga Reserve, 2100 m, Cotopaxi Province, Ecuador (collected by K. Helgen et al., August 2006); QCAZ 8662, young adult female, skin, skull, and postcranial skeleton, ["forested gully near"] La Cantera, 2260 m, Cotopaxi Province, Ecuador (collected by M. Pinto et al., August 2006). We have also seen photographs of this species from Tandayapa, 2350 m, Pichincha Province (Figure 13).

Below, we identify additional referred specimens when we describe three additional subspecies of *B. neblina* from the cordilleras of Colombia (Figures 9–10, 13–16).

Diagnosis: *Bassaricyon neblina* can be easily identified on the basis of both external and craniodental characteristics (Figures 3–7, Tables 3–5). It differs from other *Bassaricyon* in its smaller body and cranial size; longer, denser, and more richly coloured dorsal pelage (black-tipped, tan to strikingly orange- to reddish-brown); indistinctly banded, bushier, and proportionally shorter tail (at least compared to the lowland olingos, *B. alleni* and *B. medi*, Table 5); (externally) more rounded face with a blunter, less tapering muzzle; smaller and more heavily furred external ears, and considerably



Figure 13. The Olinguito, *Bassaricyon neblina neblina*, in life, in the wild. Taken at Tandayapa Bird Lodge, Ecuador (for mammalogical background of Tandayapa, see Lee et al. 2006). Photograph by Mark Gurney.



Figure 14. Olinguito skins from different regions of the Colombian Andes. Left, *B. n. ruber*, of the western slopes of the Western Andes of Colombia (FMNH 70722, adult male); Middle, *B. n. hershkovitzii*, of the eastern slopes of the Central Andes of Colombia (FMNH 70727, adult female); Right, *B. n. osborni*, of the eastern slopes of the Western Andes and eastern slopes of the Central Andes of Colombia (FMNH 90052, adult female).

reduced auditory bullae, with a markedly smaller external auditory meatus; broadened and more elongate postdental palate ('palatal shelf'), bearing more prominent lateral 'flanges' (sometimes developed to the point where it nearly closes off the "palatal notch" *sensu* Asher 2007); and proportionally much larger first molars (M1 and m1), achieved especially by the development of more massive and bulbous principal molar cusps (protocone, paracone, metacone, hypocone) in M1, and for m1 by the widening of the talonid with the expansion in particular of the entoconid and hypoconid. The m1 paraconid is reduced relative to other *Bassaricyon*.

Where *B. medius* and *B. neblina* occur in regional sympatry on the western slopes of the Andes, *B. neblina* is smaller and more richly rufous and/or blackish in coloration, and is distinguished by all of the characteristics noted above. Externally, *B. neblina* can only be confused with the highest elevation populations of *B. alleni*, from forests above 1000 m on the eastern slopes of the Andes (specimens from Pozuzo and Chanchamayo in Peru), which, like *B. neblina*, also have long, black-tipped dorsal pelage (though not so strongly rufous as in *B. neblina*), ears that are especially furry (though not so small as in *B. neblina*), and tails averaging slightly shorter than in lowland populations of *B. alleni* (but not as short as in *B. neblina*). The craniodontal characteristics of *B. neblina* (especially of the palate, bullae, and molars) are unmistakable.

Etymology: The specific epithet *neblina* (Spanish, "fog or mist"), a noun in apposition, references the cloud forest habitat of the Olinguito.

Distribution: The recorded distribution of *B. neblina* comprises humid montane rainforests ("cloud forests") from 1500 m to 2750 m in the Northern Andes, spe-



Figure 15. The Olinguito, *Bassaricyon neblina osborni*, in life. Photograph taken in captivity, at the Louisville Zoo (see Poglayen-Neuwall 1976). This animal, named "Ringer", was received as an adult in 1967 from the mountains of Colombia near Cali, and exhibited in various zoos, including the National Zoo in Washington, D.C. (see text). Photographs by I. Poglayen-Neuwall, previously unpublished (additional photographs published by Poglayen-Neuwall 1976).

cifically along the western and eastern slopes of the Western Andes of Colombia and Ecuador, and along the western and eastern slopes of the Central Andes of Colombia (Figure 16). *Bassaricyon neblina* occurs in regional sympatry with *B. medius medius* on the western slopes of the Ecuadorian Andes, where we have encountered the two species at localities less than 5 km apart. On the basis of our museum and field research, we document *B. neblina* from 16 localities (representing 19 elevational records) in the Western Andes of Ecuador and the Western and Central Andes of Colombia. All sites are situated between 1500 and 2750 m (mean 2100 m, median 2130 m, \pm 280 s.d.) and are associated with humid montane forest (“cloud forest”, Churchill et al. 1995). We used bioclimatic modeling to predict the global geographic distribution

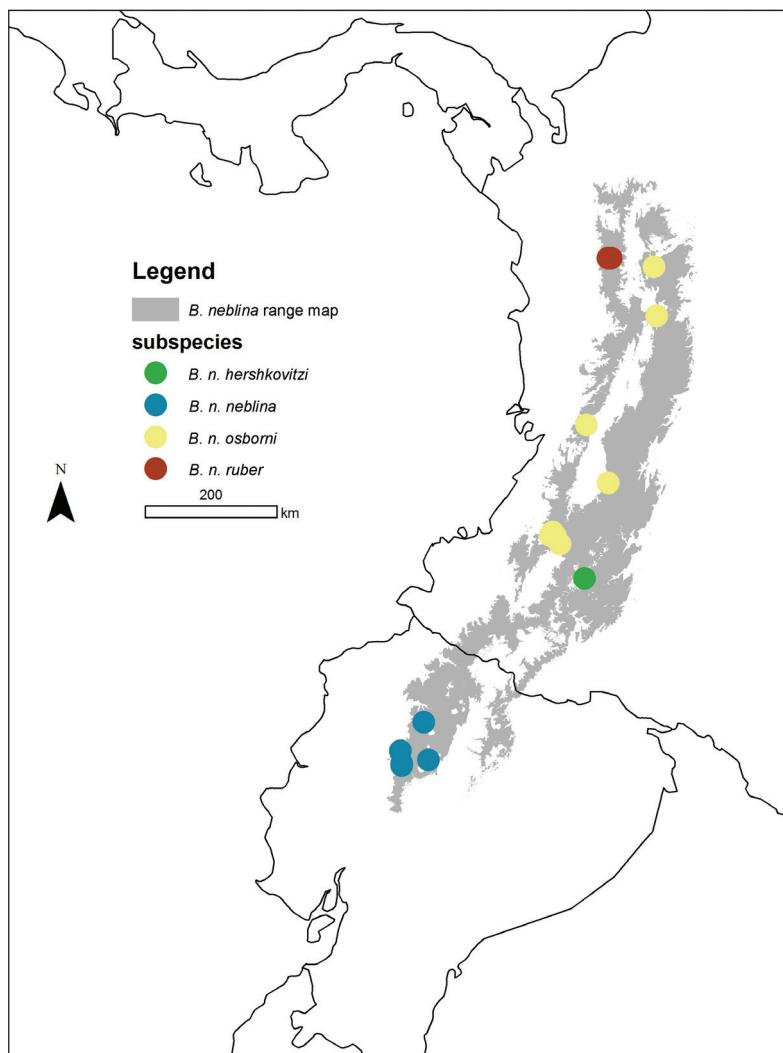


Figure 16. Distributions (localities) of the four Olinguito subspecies in the Andes of Colombia and Ecuador.

of *B. neblina*, which comprises wet, forested ecoregions typical of the habitats where Olinguitos have been recorded (Figures 11–12). As noted above, of the entire land area predicted to be suitable for Olinguito occurrence, 42% has been converted to agriculture or urban areas and 21% comprises other unforested landscapes; only 37% (40,760 km²) of this land area is currently forested.

Geographic variation: Geographic variation in the Olinguito is remarkable, reflecting consistent regional differences in color, size, and craniodental features associated with differential distributions in disjunct areas of the Andes. This is unsurprising given that the montane forests of the Central and Western Cordilleras of the Northern Andes are a region where major evolutionary differentiation has unfolded in many endemic Andean vertebrate groups (e.g., Benham 2012, Graham et al. 2010, Voss et al. 2002, Velasco et al. 2010). Below we diagnose four distinctive subspecies of *Bassaricyon neblina* and describe their geographic ranges as so far understood.

Subspecies of *Bassaricyon neblina*

Bassaricyon neblina neblina subsp. n.

http://species-id.net/wiki/Bassaricyon_neblina_neblina
(western slopes of Western Andes of Ecuador)

Diagnosis: This subspecies is (in skull length) smaller than *B. n. osborni* subsp. n., but larger than *B. n. hershkovitzi* subsp. n. and *B. n. ruber* subsp. n. (though *B. n. ruber* subsp. n. is more robust cranially, with a wider skull). It has proportionally very large teeth, especially P4 and the first molars, and a narrow skull, with a narrow and low-domed braincase (Figures 9–10, Table 8). In color it most closely resembles *B. n. osborni* subsp. n., but is the least rufous of the subspecies, usually with the greatest preponderance of black tipping to the fur (e.g., Figure 13).

Distribution: The nominate subspecies is endemic to Ecuador, where it is recorded from the western slopes of the Andes, in Pichincha and Cotopaxi Provinces, in forests at elevations from 1800 to 2300 m (Figure 16).

Referred specimens: As listed for *B. neblina*, above.

Bassaricyon neblina osborni subsp. n.

http://species-id.net/wiki/Bassaricyon_neblina_osborni
(eastern slopes of Western Andes and western slopes of Central Andes of Colombia)

Diagnosis: This is the largest subspecies of *B. neblina*, with a short rostrum, widely splayed zygomatica, wide rostrum and braincase, and very large molars and posterior premolars; the dorsal pelage is of moderate length, tan to orangish-brown in overall color, with prominent black and gold tipping, with a more grayish face and limbs, with the limbs bearing relatively short fur, and a tail usually grizzled with golden-brown fur tipping.

Distribution: This is the representative of *B. neblina* on the eastern slopes of the Western Andes of Colombia (e.g., Castilla Mountains [AMNH]; Sabanetas [FMNH]; El Tambo [NMS]; the vicinity of Cali [Poglayen-Neuwall 1976]; El Duende [Saavedra-Rodríguez and Velandia-Perilla 2011]; Gallera: “western slope of most eastern ridge of southern Western Andes” [AMNH, Paynter 1997:222]) and the western slopes of the Central Andes of Colombia (Cerro Munchique [FMNH]). One specimen (AMNH 42351, from Santa Elena, Antioquia Department) derives from the eastern slopes of the Central Andes in northern Colombia (habitat described as “deforested, grassy, and bushy (Chapman 1917:61)”; Paynter 1997:403); this shows that this subspecies also crosses to the eastern slopes of the Central Andes in Antioquia. Further south, in the department of Huila, the smaller subspecies *B. n. hershkovitzi* subsp. n. (see below) occurs on the eastern slopes of the Central Andes.

Records to date of *B. n. osborni* are from 1500 to at least 2750 m elevation in Cauca, Valle del Cauca, and Antioquia Departments of Colombia (Figure 16). *Bassaricyon medius medius* is also recorded from the Cauca Valley (east slopes of Western Andes and western slopes of Central Andes) at elevations up to at least 725 m (UV-3774: Saavedra-Rodríguez and Velandia-Perilla 2011; see account of *B. medius* below), so these two taxa (*B. m. medius* and *B. n. osborni*) are presumably regionally sympatric (and probably elevationally stratified) across the range of this Olinguito subspecies on the slopes of the Western and Central Andes.

Etymology: The name honors Henry Fairfield Osborn (1857–1935), paleontologist, faculty of Princeton and Columbia Universities, and Curator of Vertebrate Paleontology (1891–1909) and President (1909–1933) of the American Museum of Natural History (Gregory 1937, Colbert 1996). “*Bassaricyon osborni*” is a manuscript name (never formally published) associated with a specimen of this taxon (AMNH 32609, with “Type” written on the skull), demonstrating a century-old intention, later discarded (probably by J.A. Allen or H.E. Anthony, see below), to name this taxon after Osborn. Here we validate this unpublished name as a newly described subspecies of *B. neblina*, but we choose a more complete specimen than AMNH 32609, which has a damaged mandible and various broken teeth, as holotype.

Holotype: FMNH 88476, adult male, skin and skull, Munchique, 2000 m, Cauca Department, Colombia (collected by K. von Sneidern, 3 June 1957).

Paratypes: AMNH 32608, adult female, skin and skull, and AMNH 32609, adult male, skin and skull, Gallera (Chapman 1912:155; = “La Gallera” of Paynter 1997:222), 5000 feet (=1524 m), Cauca Department, Colombia (both collected by L. Miller, 13 July 1911); NMS A59-5083, adult female, skin and skull, El Tambo, 1700 m, Cauca Department, Colombia (collected by K. von Sneidern); FMNH 85818, adult male, skin and skull, Munchique, 2000 m, Cauca Department, Colombia (collected by K. von Sneidern, 19 January 1956); FMNH 89220, adult female, skin and skull, Sabanetas, 2000 m, Cauca Department, Colombia (collected by K. von Sneidern, 26 September 1957); FMNH 90052, adult female, skin and skull, Sabanetas, 1900 m, Cauca Department, Colombia (collected by K. von Sneidern, 12 February 1959).

Referred specimens: AMNH 14185, skin (skull not found), adult male, Castilla Mountains (“La Castilla” of Paynter 1997), Valle del Cauca Department (collected by J.H. Batty, 9 June 1898); AMNH 42351, adult male, skin and skull, Santa Elena, apparently at 9000 feet (= 2750 m), Antioquia Department, Colombia (collected by H. Niceforo Maria, 10 January 1919) (Paynter 1997:403); USNM 598996, adult male, skin, skull, and postcranial skeleton, from Colombia, specific locality unknown (received from Tulane University).

***Bassaricyon neblina hershkovitzi* subsp. n.**

http://species-id.net/wiki/Bassaricyon_neblina_hershkovitzi
(eastern slopes of Central Andes of Colombia)

Diagnosis: This is the smallest subspecies of *B. neblina*, with the fur of the dorsum and tail very long, and richly orange-brown (brown with strong golden and black tipping) in coloration, and more golden brown face and limbs, with the limbs well-furred. The



Figure 17. Area of sympatric occurrence between *Bassaricyon* species in western Ecuador. Farmland cutting into cloud forest habitat at Las Pampas, approximately 1800 m, on the western slopes of the Western Andes, Ecuador, along the boundary of Otonga, a protected forest reserve. It is at this elevational and environmental boundary that *B. medius medius* (lower elevations, including more anthropogenically disturbed habitats) and *B. n. neblina* (higher elevations, less disturbed forests) co-occur in regional sympatry on the western slopes of the Andes.



Figure 18. Type series of an Olinguito subspecies, *Bassaricyon neblina herskovi*zti, in the field. Two Olinguito specimens (FMNH 70726, paratype of *herskovi*zti, and FMNH 70727, holotype of *herskovi*zti, along with a Long-tailed weasel, *Mustela frenata*, FMNH 70998) brought in by a local hunter, 6 September 1951, at San Antonio, San Agustín, Huila District, Colombia. Photo by P. Hershkovitz, courtesy of the Field Museum of Natural History.

skull, braincase, and rostrum are especially narrowed, the posterior palatal shelf is extremely broad, and the molars are proportionally very large.

Distribution: This is the representative of *B. neblina* on the eastern slopes of the Central Andes of southern Colombia (Figure 16). Records to date are from 2300 to 2400 m elevation in the vicinity of San Antonio (Huila Department), a forested locality “on eastern slope of Central Andes at headwaters of Rio Magdalena, near San Agustin” (Paynter 1997:380) (see Kattan et al. 1994).

Etymology: The name honors American mammalogist Philip Hershkovitz (1909–1997), collector of the type series, Curator of Mammals at the Field Museum of Natural History (1947–1974; Emeritus Curator until 1997), and authority on South American mammals (Patterson 1987, 1997).

Holotype: FMNH 70727, adult female, skin, skull, and postcranial skeleton, San Antonio, 2300 m, San Agustin, Huila Department, Colombia (collected by P. Hershkovitz, 6 September 1951) (see Figure 18).

Paratypes: FMNH 70724, adult male, skin, skull, and postcranial skeleton, San Antonio, 2400 m, San Agustin, Huila Department, Colombia (collected by P. Hershkovitz, 20 August 1951); FMNH 70725, adult male, skin, skull, and postcranial skeleton, San Antonio, 2400 m, San Agustin, Huila Department, Colombia (collected by P. Hershkovitz, 25 August 1951); FMNH 70726, adult male, skin, skull, and postcranial skeleton, San Antonio, 2300 m, San Agustin, Huila Department, Colombia (collected by P. Hershkovitz, 6 September 1951).

Bassaricyon neblina ruber subsp. n.

http://species-id.net/wiki/Bassaricyon_neblina_ruber

(Urrao District, western slope of Western Andes of Colombia)

Diagnosis: This subspecies is markedly smaller (at least in skull length) than *B. n. neblina* and *B. n. osborni*, with the fur longest and most strikingly reddish of all the Olinguito populations (reddish with golden and black tipping), and more golden brown face and reddish brown limbs, with the limbs well-furred. Though similar in overall skull length to *B. n. hershkovitzi*, the skull is especially wide for its size (Table 8), with broad zygomatica, braincase, and rostrum compared to that subspecies.

Distribution: This subspecies is recorded from the Urrao District of Colombia (2200–2400 m in Huila and Antioquia Departments), on the western slope of the Western Andes, where it is documented by specimens collected in 1951 by Philip Hershkovitz.

Etymology: The name refers to the rich reddish-brown pelage of this subspecies (Figures 3, 14).

Holotype: FMNH 70722, adult male, skin, skull, and postcranial skeleton, Rio Urrao, 2400 m, Urrao, Huila Department, Colombia (collected by P. Hershkovitz, 24 April 1951).

Paratypes: FMNH 70721, adult female, skin, skull, and postcranial skeleton, Rio Ana, 2200 m, Urrao, Huila Department, Colombia (collected by P. Hershkovitz, 19 April 1951); FMNH 70723, adult male, skin, skull, and postcranial skeleton, Guapantel, 2200 m, Urrao, Antioquia Department, Colombia (collected by P. Hershkovitz, 28 April 1951).

Reproductive isolation and genetic divergence of *B. neblina*

Information from sympatric occurrences and captive breeding demonstrates that the Olinguito, *B. neblina*, is reproductively isolated from other species of *Bassaricyon* and clearly constitutes a distinct “biological species” (i.e., *sensu* Mayr 1940, 1942).

In Ecuador we documented the Olinguito (*B. neblina neblina*) in regional sympatry with the Western lowland olingo, *B. medius medius*; we recorded the two species at localities less than 5 km apart (i.e., at Otonga and San Francisco de las Pampas) during fieldwork in August 2006. The ecogeographic relationship between the two species is probably one of elevational parapatry or limited elevational overlap along the western slopes of the Andes. *Bassaricyon medius medius* extends into the elevational range of *B. neblina*, perhaps especially in areas where cloud forests have been cleared for human settlement, agriculture, and pastoralism (Figure 17).

Ingeborg Poglayen-Neuwall (pers comm. to R. Kays, 2006) informed us that an adult female zoo animal named “Ringerl” (Figure 15; also figured by Poglayen-Neuwall 1976), which we can now identify as an Olinguito (*B. n. osborni*), was moved among several zoos during the 1970s because it would not successfully breed with other captive olingos (i.e., not *B. neblina*), most of which were apparently *B. alleni* (see Poglayen-Neuwall 1976).

The Olinguito differs from congeners (*B. alleni*, *B. medius*, and *B. gabbi*) by 9.6–11.3% in base-pair composition of the (mitochondrial) cytochrome *b* gene (Table 2), a level of divergence consistent with that separating biological species in many groups of mammals, including carnivores (Baker and Bradley 2006). For comparison with other procyonids, this level of genetic distinction is equivalent to the 10–11% divergence between *Procyon lotor* and *P. cancrivorus*, sympatrically-occurring raccoons traditionally classified in separate subgenera (Goldman 1950, Helgen and Wilson 2005), and comparable to the 9–13% divergence between *Nasua narica* and *Nasuella olivacea* (Helgen et al. 2009), coatis traditionally classified in separate genera (Hollister 1915, Decker and Wozencraft 1991, Wozencraft 1993, 2005).

Karyotype: The karyotype of an adult female Olinguito (*B. n. osborni*, then identified as “*B. gabbi*”, with $2n = 38$, as in all procyonids) was reported and discussed (but not described in detail) by Wurster-Hill and Gray (1975), and figured by Nash (2006). This was based on a captive animal originally captured from mountains in the vicinity of Cali in Colombia (Figure 15).

Description: The Olinguito is the smallest species of *Bassaricyon*, both in skull and body size (Tables 3, 5), and is thus, on average, the smallest living procyonid

(matched only by small individuals of the Ringtail, *Bassariscus astutus*). The tail averages 10% longer than the head-body length (Table 5). The pinnae are proportionally much smaller in *B. neblina* than in other *Bassaricyon*, appearing shorter and rounder, and standing out less conspicuously on the head; they are also more heavily furred and usually fringed with a paler, contrasting border of buffy or golden fur. The dorsal fur is dense, long, and luxurious, with the longer hairs measuring 30–40 mm in length (usually much shorter in other *Bassaricyon*, at least in the predominantly lowland taxa *B. mediis* and *B. alleni*, but reaching 25 mm in the highest-elevation populations of *B. alleni* on the eastern versant of the Andes). The hairs of the dorsum, crown, upper limbs, and tail are golden-orange, with grey bases and dark red-brown or blackish-brown tips, generating a distinctly dark, often red-brown appearance, more striking than the relatively drab fur colors (more tan or yellowish-brown to grayish-brown) of other *Bassaricyon* (Figure 3). The fur of the cheeks, chin, venter, and underside of the limbs is yellow to the bases, often washed with orange. The fur of the face in front of the eyes is shorter and gray or buff with black tipping, sometimes with a pale cream ring around the eyes. The hairs of the tail are strongly tipped with gold, or with both golden and blackish-brown tipping. In contrast to specimens of other *Bassaricyon*, the tail is not conspicuously banded, though when viewed in the right light, a banding pattern of alternating golden and brown hues is weakly apparent in some specimens. A white terminal tail tip is present in a minority of individuals.

Like other *Bassaricyon*, the cranium of *B. neblina* is long relative to its width, with a moderately long and broad rostrum, an elongate and somewhat globose braincase with a smooth dorsal surface, and moderately developed postorbital processes.

In *B. neblina*, the temporal ridges do not meet to form a sagittal crest, even in older animals. The postdental palate is usually flared laterally, but is smoothly parallel-sided, tapers posteriorly, or bears only weaker bony flaring in other *Bassaricyon* (Figures 4–7, 19). At its more extreme development (e.g., in FMNH 70726), the portion of the bony palate sitting behind M2 is almost continuously joined to the postdental palate by a continuous shelf of bone, rather than bearing a deep excavation separating the molar-bearing portion of the bony palate from the postdental shelf (Figure 19). The auditory bullae are very small in the Olinguito relative to other *Bassaricyon*, both in length and vertical inflation, and the external auditory meatus is considerably narrower in diameter, on average (Figures 4–7). The median septal foramen of the anterior palate (Steno's Foramen), between the paired incisive (or anterior palatal) foramina, is usually well-developed. The mandible is proportionally less elongate than in other *Bassaricyon*, with a proportionally larger and more vertically-oriented coronoid process (Figures 4–5). The first two upper premolars are caniniform, similar in size and shape to those of other *Bassaricyon*. P3 is usually relatively smaller in *B. neblina* than in other *Bassaricyon*. P4 is similar in structure to congeners but is relatively larger with a more bulbous protocone and more prominent metacone. M1 and M2 are proportionally lengthened and considerably more massive in appearance, especially relative to skull

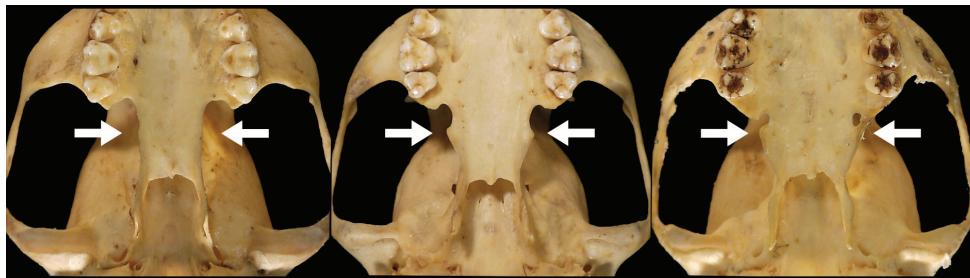


Figure 19. Lateral flare of the postpalatal shelf in *Bassaricyon*. Lateral extension of the postpalatal shelf (shown by white arrows) is usually absent or little-developed in other *Bassaricyon* (e.g., left, *B. alleni*, FMNH 41501), but is well-developed in *B. neblina* (e.g., center, *B. n. ruber*, FMNH 70721, and right, *B. n. hershkovitzii*, FMNH 70726).

size, than in other *Bassaricyon*. *p4* is variable in size among *B. neblina* subspecies, generally smaller than other *Bassaricyon* in *B. n. ruber* and *B. n. hershkovitzii*, but proportionally quite large in *B. n. neblina*. *m1* is relatively much larger in *B. neblina* than in other *Bassaricyon*; each of the four major cusps that define the subrectangular shape of this tooth are massive and bulbous, and the posterior portion is especially broadened, with the metaconid and hypoconid particularly large and laterally expanded relative to congeners. *m2* is also often expanded in size in *B. neblina* relative to other *Bassaricyon*.

Natural history: Our field observations document that *B. neblina* is nocturnal, arboreal, frugivorous, and probably largely solitary (compiled during July and August 2006 at Otonga Forest Reserve in Ecuador: 00°41'S, 79°00'W; for faunal and floral context see Freiberg and Freiberg 2000, Nieder and Barthlott 2001, Jarrín-V 2001). It occupies cloud forest canopies and is an adept leaper. It has a single pair of mammae and probably raises one young at a time. Notes associated with AMNH 14185, the first specimen to arrive in a museum, mention that it was “shot at 2 pm [an error for 2 am?] in high trees while coming down mountain to feed on guavas; strictly nocturnal.”

An adult female Olinguito (an animal named “Ringerl”, *B. n. osborni*, Figure 15) that lived at the Louisville Zoological Park and the National Zoological Park in Washington during 1967–1974 made vocalizations different from those of other *Bassaricyon* according to Poglayen-Neuwall (1976). Poglayen-Neuwall (1976) figured a picture of this animal in characteristic estrus behavior and in various other circumstances (see below for more discussion of this captive Olinguito).

Previous identifications and references:

Though described taxonomically for the first time in this paper, the Olinguito (heretofore misidentified as other species of *Bassaricyon*) has been represented in museum collections for more than a century, has been exhibited in zoos, has had its karyotype published, and has been included in published molecular phylogenetic studies.

Olinguito museum specimens previously reported in the literature include specimens from Gallera, Colombia, mentioned by Allen (1912, 1916) (AMNH 32608 and

32609, as “*B. medioides*”); a specimen from Santa Elena, Colombia, reported by Anthony (1923) (AMNH 42351, as “*B. medioides*”), specimens from “El Duende Regional Reserve” (2200 m asl; 04°02'55.6"N, 76°27'28.4"W) and “Los Alpes, Florida, 2250 m asl” in Valle del Cauca Department, Colombia (mammal collection of the Universidad del Valle, Cali, Colombia, specimen numbers 12736, 13700) discussed by Saavedra-Rodríguez and Velandia-Perilla (2011) (as “*B. gabbii*”); and a skull from San Antonio, Huila Department, Colombia (FMNH 70727) figured by Prange and Prange (2009) (as “*B. gabbii*”, designated above as the holotype of *B. n. hershkovitzi*). One Olinguito specimen, AMNH 32609, bears an unpublished scientific name, “*Bassaricyon osborni*”, written on the skull and on the tags, apparently during the early twentieth century—correctly reflecting an understanding that the specimen represented an undescribed species. This appellation (a “manuscript name”) is likely attributable to J.A. Allen or H.E. Anthony (seemingly too early to be G.H.H. Tate). In any case, the name was never published, and by 1923, Allen had passed away (in 1921) and Anthony had decided that the specimen in question was best referable to *Bassaricyon medioides* (see Anthony 1923). We have chosen to validate this name under our own authorship, above, as a subspecies of *B. neblina*.

Mejía Correa (2009) reported camera-trap photos of a species of *Bassaricyon* at Munchique in Colombia; these records presumably represent *B. n. osborni*, the only *Bassaricyon* recorded at Munchique.

Molecular data for *B. neblina* from a cell line were first generated and used in a phylogenetic study of carnivore relationships by Ledje and Arnason (1996a, 1996b), apparently the same animal whose karyotype was reported and discussed by Wurster-Hill and Gray (1975) (also Nash 2006). DNA sequence data (12S rRNA, cytochrome *b*) from this sample, available on Genbank, have been used in various other published studies (e.g., Flynn and Nedbal 1998, Koepfli and Wayne 1998, Emerson et al. 1999, Flynn et al. 2000, Gaubert et al. 2004, Marmi et al. 2004, Flynn et al. 2005, Fulton and Strobeck 2007, Yonezawa et al. 2007, Wolsan and Sato 2010, Nyakatura and Bininda-Emonds 2012, but not in some important studies, e.g., Koepfli et al. 2007, Agnarsson et al. 2010). This cell line apparently originated from the zoo animal “Ringerl” (discussed by Poglayen-Neuwall [1976]), an adult female Olinguito (*B. n. osborni*, originally from mountains near Cali, Colombia), apparently exhibited at the Louisville Zoo, National Zoo, Tucson Zoo, Bronx Zoo, and possibly Salt Lake City Zoo during the late 1960s and 1970s (Ingeborg Poglayen-Neuwall, pers. comm. to R. Kays, 2006). Ivo Poglayen-Neuwall (in litt. to C.O. Handley, Jr, 6 November 1964) mentioned another *Bassaricyon*, a young adult male at the Louisville Zoo, also from Cali, received in 1964, that seems also to have been an Olinguito (“shows the following unusual physical features: (1) strikingly round head... (2) very short, round ears! (3) rather short tail (no amputation!)”). This latter animal seems not to be discussed in Poglayen-Neuwall’s various publications on olingos, and it is unclear what became of it.

Relevant field notes associated with *B. medius* include: “shot at dusk in high tree in forest” (FMNH 29180); “shot at 8 pm, 40 feet up in large tree, active and agile, but curious, eyes shine brightly” (USNM 305748); “shot at 8:30 pm in avocado plantation” (USNM 305749); “shot near banana plantation (at night), stomach with banana” (USNM 305750); “shot at 8:30 pm in large tree in cafetal [coffee plantation], stomach with soft fruit with tomato-like seed” (USNM 305751); “shot at 8 pm in forest” (USNM 307037); “lactating” and pregnant with “1 embryo”, “stomach: fruit pulp” (USNM 310666); “shot in tree at night” (USNM 335767, 338348); “shot at night in tree in forest” (USNM 335769); “shot at night in tree in cocoa grove” (USNM 335770); “shot in small tree in plantain patch at night” (USNM 335771); “one embryo” in a pregnant female “shot in forest” (USNM 363342); “shot in banana tree” (USNM 363343).

Specimens examined:

B. m. medius

Colombia: BMNH 9.7.17.10 (holotype of *medius*), 9.7.17.11, FMNH 29180, 86852, 90049, 90051, MVZ 124112, USNM 598997. **Ecuador:** AMNH 66752, BMNH 34.9.10.81, 34.9.10.82, EPN 841, 900, MECN DAP37, NMS A59-5081, A59-5082, QCAZ 8758, 8659.

B. m. orinomus

Panama: USNM 171138, 179053, 179157 (holotype of *orinomus*), 179158, 179779, 179917, 206123, 284773, 284903, 284933, 284934, 284935, 305748, 305749, 305750, 305751, 305752, 305753, 305754, 307035, 307036, 307037, 310666, 310667, 310668, 324295, 324296, 335767, 335768, 335769, 335770, 335771, 338348, 338894, 363342, 363343, 363344. **Colombia** (tentatively attributed): AMNH 37797, FMNH 69578.

Discussion

Carnivore taxonomy

Descriptions of new species of carnivores are especially rare, and the order Carnivora is generally considered one of the most completely characterized groups across the entire tree of life (Collen et al. 2004, Reeder et al. 2007). *Bassaricyon neblina* is a deeply divergent lineage within its genus, a very morphologically distinctive member of the family Procyonidae, and even shows signs of evolutionary diversification across its geographic range. It thus adds significantly to current understanding of taxonomic, phylogenetic, and ecomorphological evolution in the family Procyonidae. It has presumably been overlooked by taxonomists for several reasons—principally the lack of close taxonomic attention paid to Neotropical procyonids for nearly a century (Helgen and Wilson 2003, Helgen et al. 2009), but probably also because of its nocturnal and arboreal habits, relatively limited geographic distribution, and the small number of specimens scattered across various museum collections (see Patterson 1994, 2000).

The description of the Olinguito highlights how incompletely known the taxonomy of almost all kinds of mammals remains, including the Carnivora (Gutiérrez and Helgen 2013). Our study of olingo taxonomy is part of a series of studies that have better clarified species diversity in insufficiently studied genera of Carnivora, especially in Neotropical small carnivores (e.g., *Procyon*: Helgen and Wilson 2002, 2003, 2005; Helgen et al. 2008b; *Nasuella*: Helgen et al. 2009; *Galictis*: Bornholdt et al. 2013), but also in other little-known genera (*Arctonyx*: Helgen et al. 2008a; *Eupleres*: Goodman and Helgen 2010), often revealing considerable overlooked biodiversity in poorly studied groups. Many additional carnivore genera have not been the subject of modern integrative systematic reviews, especially in the Neotropics (e.g., *Potos*, *Nasua*, *Conepatus*). Detailed reviews of these groups are likely to reveal additional overlooked diversity.

Conservation

The rapid and ongoing discovery of endemic mammals and birds in northern Andean cloud forests (e.g., Robbins and Stiles 1999, Anderson and Jarrín-V 2002, Cuervo et al. 2001, 2005, Lara et al. 2012, Ojala-Barbour et al. in press) reaffirms the evolutionary importance of these unique habitats and betrays how incompletely inventoried this biota remains. Though a center of diversity and endemism for many groups (e.g., Young et al. 2002, Brehm et al. 2005, Mittermeier et al. 2005, Hughes and Eastwood 2006, Patterson et al. 2012), northern Andean cloud forests are among the most threatened ecosystems in the Neotropics (Young 1994, Myers et al. 2000, Mittermeier et al. 2005, Schipper et al. 2008). Drawing on the criteria used by the International Union for the Conservation of Nature (IUCN; Schipper et al. 2008; in this case, based on inferred population declines due to habitat declines over last three generations), we suggest classifying the Olinguito under the IUCN category of “Near Threatened.” Given that Olinguitos are directly dependent on cloud forest for habitat and food, deforestation appears to be the primary threat to Olinguito populations, and this IUCN categorization reflects our concerns about habitat destruction across its relatively restricted geographic range. Based on our distribution model, it appears that 42% of potential Olinguito habitat in Colombia and Ecuador has already been converted to agriculture or urban environments. Remaining habitat is highly fragmented and faces increasing threats from farming, grazing, deforestation for drug cultivation, logging, and climate change (Kattan et al. 1994, Myers et al. 2000, Brooks et al. 2002, Sarmiento 2002, Armenteras et al. 2003). The long-term survival of *B. neblina* will depend on the preservation of those upland forest fragments that remain, and restoration of degraded habitat to maintain connectivity between populations. Its discovery introduces a novel flagship species around which to rally conservation initiatives in the region. Preserving cloud forests in this region would benefit the long-term conservation of the Olinguito, and many other Northern Andean cloud forest endemics.

Based on their relatively expansive distributional ranges, all of which include various protected areas (Figures 11, 12), we suggest IUCN Red List rankings of “Least Concern” for *B. alleni*, *B. medius*, and *B. gabbi*, for the present.

Biogeography

A well-resolved taxonomy for olingos has never been available, such that biogeographic patterns within the genus, and their origins, have never before been critically reviewed (Eizirik 2012). Our overview of *Bassaricyon* allows us to glimpse these patterns for the first time, unveiling both anticipated and unexpected biogeographic patterns.

Previous overviews of procyonid biogeography have focused especially on the important potential role of the Great American Biotic Interchange (GABI) in the diversification of the family (Marshall et al. 1979, Koepfli et al. 2007, Eizirik 2012). We complement this focus by suggesting that northern Andean uplift, proceeding in greatest part since the middle Miocene (Gregory-Wodzicki et al. 2000, Ollier 2006, Weir 2006), has played an almost equally important role in procyonid diversification.

The most detailed previous phylogenetic comparisons of olingos (Koepfli et al. 2007) highlighted the genetic divergence between taxa originating from North America and from South America (*B. medius* from Panama [then called “*B. gabbi*” by Koepfli et al. 2007] and *B. alleni* from Peru), finding that this split apparently postdated the GABI. This comparison was undertaken prior to the discovery of the Olinguito lineage, the deepest split in the genus, and could not resolve the question of whether the radiation of crown group olingos unfolded first in North or in South America. Our phylogenetic comparisons indicate that *B. neblina*, an Andean cloud forest endemic, is the sister taxon to all other *Bassaricyon* and last shared a common ancestor with congeners 3–4 million years ago, a timescale concordant with the timing of both the GABI and Northern Andean mountain-building. That *Bassaricyon* mainly occurs in South America, with only one species, *B. gabbi*, endemic to Central America, and that the earliest divergence in *Bassaricyon* is between *B. neblina* and the other three species allows us to suggest that the most important events in the diversification of crown group *Bassaricyon* occurred in northwestern South America (as suggested by Poglayen-Neuwall 1973) (see Velazco and Patterson [2013] for particularly clear example of this same biogeographic pattern). That the two earliest divergences within the genus involve what are today montane (*B. neblina* of the Andes) or mostly montane (*B. gabbi* of the Costa Rican, Nicaraguan, and western Panama highlands) taxa provides an indication that the isolation of upland Neotropical habitats was likely important in the early diversification of the genus. Uplift of the Andes simultaneously created a barrier to dispersal that is ultimately reflected in the speciation event between the allopatric pair of lowland olingos, *B. alleni* (eastern, cis-Andean) and *B. medius* (western, trans-Andean) (cis- and trans-Andean *sensu* Haffer 1967). In addition to promoting evolutionary diversification within *Bassaricyon*, Northern Andean uplift has fostered the evolution of other endemic montane procyonids (the Mountain coatis *Nasuella olivacea* and *N. meridensis* [Helgen et al. 2009] as well as currently unrecognized montane species of *Nasua*,

synonymized uncritically with *N. nasua* under current taxonomic checklists [following Decker 1991]). These mountains also served as a key barrier to dispersal of presumed recent North American procyonid immigrants (*Procyon lotor* and *Nasua narica*) to South America, which penetrate South America only west of the Andes, primarily in western Colombia (Marín et al. 2012), with *N. narica* perhaps extending also to western Ecuador (Decker 1991) and *P. lotor* perhaps also to western Venezuela (Helgen and Wilson 2005).

The phylogenetic topology seen in *Bassaricyon*, with an Andean species sister to a clade of lowland congeners, is unusual among mammals, but seen in some groups with lowland representatives restricted to the Amazon. For example, the echimyid rodent genera *Dactylomys* and *Isothrix* present this pattern, with *Isothrix barbarabrownae* and *Dactylomys peruanus* restricted to the Andes and their congeners to the Amazon lowlands (Patterson and Velazco 2008, Lim 2012, Patterson et al. 2012). In olings, the time estimates for this diversification are broadly equivalent with the estimated Pliocene divergence timing (2–5 mya) proposed between *Isothrix barbarabrownae* and its lowland congeners (Upham and Patterson 2012). A similar pattern of inferred colonization from the Andes to the Amazonian lowlands was proposed for dendrobatid frogs, but occurred earlier, during the late Miocene (11.2 – 5.3 mya), when the Andes were considerably lower in elevation (Santos et al. 2009).

One species of olingo, *B. alleni*, is endemic to habitats east of the Andes, especially the Amazon. The Amazon is arguably the most diverse region of the planet (e.g. Bass et al. 2010, Malhado et al. 2013; but see Solari et al. 2012), and it has been postulated that its high current diversity is a result of an accumulation of lineages for a prolonged period of time, covering mostly the Pliocene and Miocene, with subsequent local divergences (e.g., Hoorn et al. 2010; Leite and Rogers 2013). However, *Bassaricyon alleni* appears to be a considerably more recent immigrant to this region, likely arriving in the Pleistocene, during the past 1–2 mya (Figures 1–2), well after the last major uplift of the Andes, which occurred until *circa* 3.0 mya (Gregory-Wodzicki 2000). Thus, it is likely that a dispersal event across the North Andes is responsible for the cis-Andean distribution of *B. alleni*. This supports the idea that the Andes and the trans-Andean Neotropics (the western side of South America, and Central America) serve as continuous pumps of diversity into the Amazon, as proposed in other vertebrate groups such as tanagers and woodcreepers (e.g. Sedano and Burns 2010, Weir and Price 2011). The western boundary of the Amazon with the Andes and close proximity to the Chocó and Central America contribute to an influx of species from these regions into the Amazon and this influx seems to be a principal driver of the high diversity of the western Amazon and the eastern slopes of the Andes (Patterson et al. 2012).

One species of olingo, *B. medius*, is distributed in the Chocó forests to the west of the Western Andes of Colombia and Ecuador, as well as in tropical forests of eastern Panama in Central America (Figure 12). For vertebrates, this is a common pattern: the Chocó has closer biogeographic affinities with Central America than with other areas of South America (Ron 2000). Mammalian examples of a Chocó + Central America distributional pattern include many medium-sized species in the region, including *Nasua narica*, *Procyon lotor*, *Coendou rothschildi*, *Tamandua mexicana*, *Caluromys derbianus*,

and *Philander opossum* (Eisenberg 1989, Brown 2004, Voss 2011, Voss et al. 2013, Marín et al. 2012). That these various distributions result from multiple biogeographic events is evidenced by the dissimilar evolutionary divergence timings involved, but the GABI and Northern Andean uplift no doubt are key events that collaborated to generate these co-distributions. The divergence between the two subspecies of *B. medius* is recent (*circa* 1.0 mya, Figure 1), but considering both subspecies are recorded in Colombia, it seems possible that *B. medius* entered Panama quite recently, perhaps penetrating the North American continent as far as the distribution of the Mesoamerican endemic taxon *B. gabii*. The location of the geographic boundary between *B. medius* and *B. gabii* in Panama is not yet clear, and the nature of interaction between these species, if any, at this boundary, will be a very interesting subject for further investigation.

The last species of olingo to consider, *B. gabii*, is a Mesoamerican endemic, distributed from Nicaragua to western Panama and recorded primarily in montane contexts: the Nicaraguan highlands, Costa Rican cordilleras, and Chiriquí Mountains. The elucidation of the phylogenetic relationship, depth of divergence (we estimate a *circa* 2.0 mya divergence between *B. gabii* and the lowland species-pair *B. medius/B. alleni*) and the distinctive morphological features of *B. gabii* allow us to recognize it as the only carnivore species endemic to this region of Central America, although many vertebrate species, especially birds, reptiles and amphibians, and small mammals, are endemic to this same region (Savage 1966, 1982, Slud 1964, Stiles and Skutch 1989, Carleton and Musser 1995). As noted by Carleton and Musser (1995:357–358), “some have attributed the high endemism to the possible isolation of the Talamanca-Chiriquí region as an island, or a series of islands, within the Panamanian portal prior to complete closure and late-Pliocene formation of the landbridge” (citing McPherson 1985, 1986, among others). This vision of insular or archipelagic diversification in *Bassaricyon* during the GABI may provide insight into the early splits in the genus that ultimately gave rise to the principal modern lineages so far identified in the genus: *B. neblina* in the Andes of northwestern South America, *B. gabii* in the Nicaragua-W Panama highlands, and *B. medius/B. alleni* in the Neotropical lowlands primarily in South America (southward from eastern Panama). Additional geographic surveys, specimen collecting, and specimen-based comparisons are needed to better understand the nature of differentiation in *B. gabii* across different Central American cordilleras, and the true easternmost extent of its distribution, where it may co-occur with or abut the range of *B. medius*.

Additional Olinguito study priorities

Our studies of Olinguito specimens in museums reveal a remarkable pattern of geographic variation, allowing for the delineation of four distinctive subspecific taxa distributed in separate biogeographic regions of the Andes of Colombia and Ecuador. Additional study is needed to more fully evaluate the level of genetic divergence between different Olinguito subspecies, especially for *B. n. ruber*, perhaps the most isolated and distinctive of the four (Figures 3, 9–10, 13–16).

Our bioclimatic analyses (Figures 11–12) also identify a number of high-priority candidate regions where further exploration is needed to assess whether additional populations of the Olinguito, or other distinctive high-elevation *Bassaricyon* populations, are present (Figure 24). One of these is the Colombian Eastern Andes, or Cordillera Oriental, the eastern branch of the Andes in Colombia. Olinguitos are recorded from the Western and Central Andes of Colombia, but not yet from the Eastern Andes, an area of substantial montane biotic endemism, where only *B. alleni* is known to occur. Another survey priority is the Quijos region of Ecuador, a county and river situated on the eastern side of the Andes, which comprises relevant cloud forest habitats (Quijos is an old, pre-Spaniard name for the indigenous community in the area). This region deserves greater attention and contains the important Papalacta region discussed by Voss (2003). The Pallatanga-Sangay region in the Central Andes of Ecuador is another important priority study area; Pallatanga is an important mammal type locality (Tomes 1860), and Sangay is a national park with peculiar cloud forest mammal representation (Fonseca et al. 2003, Lee et al. 2011). Finally the Loja-Huancabamba, a low elevation region of the Andes in southern Ecuador and northern Peru has potential as Olinguito habitat. Though situated on the eastern side of the Andes, this region was recognized as biogeographically important by Chapman (1926). Patterson et al. (1992) showed little differentiation between *Artibeus* from the

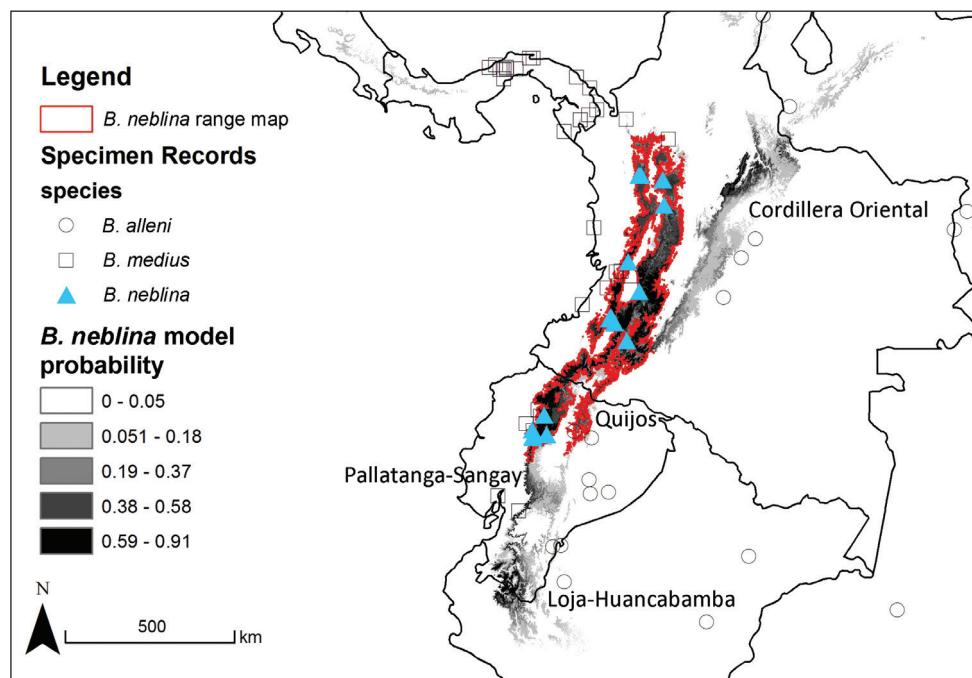


Figure 24. Selected priority areas to search for Olinguitos. Areas mentioned in the text with appropriate cloud forest habitats. A) Cordillera Oriental, the eastern branch of the Andes in Colombia. B) Quijos, a county on the eastern side of the Andes in Ecuador. C) Pallatanga-Sangay in central Ecuador. D) The Loja-Huancabamba region of the Andes in southern Ecuador and northern Peru.

western slope of the Andes and the Marañón valley in this area of northern Peru, and Pinto (2009) inferred potential cases of east-west dispersal in vampire bats across both slopes of the Andes in this region of southern Ecuador, suggesting this could well be an area where the Olinguito could cross from the western to the eastern versant of the Andes.

Much remains to be learned about the Olinguito, including its distribution. The taxonomic description of this species is the first step toward further studies of its biology, and we look forward to future reports of additional discoveries from Andean cloud forests regarding this beautiful procyonid.

Acknowledgments

For assistance in Ecuador we thank E. Tapia, C. Tapia, G. Onore, P. Asimbaya, L. Velásquez, P. Rosero, M. Grijalva, G. Acosta, P. Pinto, M. Gallo, A. Padilla, M. Vargas, P. Vargas, I. Rodriguez, C. Byrne, and Fundación Otonga. For access to specimens under their care, loan of tissues, and other assistance during museum visits we are grateful to D. Lunde, E. Westwig, R. Voss, W. Stanley, B. Patterson, L. Heaney, R. Banasiak, S. Burneo, R. Timm, J. Patton, C. Conroy, M. Engstrom, B. Lim, J. Eger, P. Jenkins, R. Portela Miguez, D. Hills, J. Chupasko, N. Gilmore, J. Dines, J. Light, W. Murphy, R. Asher, I. Thomas, L. Albuja-V., O. Groenwall, R. Anderson, E. Sargis, K. Zyskowski, P. Myers, S. Jansa, R. Baker, K. Koepfli, R. Fleischer, R. Arcos, C. Handley, Jr., L. Gordon, and H. Kafka. For assistance in the registrar's office at the National Zoo, we thank L. Morse. For helpful reviews, helpful discussion, and other assistance, we thank W. Bogdanowicz, B. Patterson, R. Sampaio, E. Gutiérrez, G. Slater, K. Koepfli, R. Anderson, I. Poglayen-Neuwall, J. Gibbons, P. Engelbrektsson, F. Reid, J. Pino, R. Samudio, M. Bass, T. Flannery, E. Eizirik, R. Baker, H. Mantilla, R. Ree, R. Voss, and L. Emmons. This work was partially funded by the Smithsonian Institution. M.P. was supported in part by a grant from the Texas Tech University Association of Biologists. We thank Nancy Halliday for original artwork. For photographic contributions we thank P. Asimbaya and L. Velásquez of *Finding Species*, and M. Gurney, I. Poglayen-Neuwall, T. Gregory (Smithsonian Conservation Biology Institute), C. Aguilar, G. Basco, M. Guerra, R. Banasiak, R. Bornholdt, and S. Burke.

References

- Agnarsson I, Kuntner M, May-Collado LJ (2010) Dogs, cats, and kin: a molecular species-level phylogeny of Carnivora. *Molecular Phylogenetics and Evolution* 54: 726–745. doi: 10.1016/j.ympev.2009.10.033
- Ahrens HE (2012) Craniodental characters and the relationships of Procyonidae (Mammalia: Carnivora). *Zoological Journal of the Linnean Society* 164: 669–713. doi: 10.1111/j.1096-3642.2011.00778.x

Back to the future: museum specimens in population genetics

Peter Wandeler, Paquita E.A. Hoeck and Lukas F. Keller

Zoological Museum, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

Museums and other natural history collections (NHC) worldwide house millions of specimens. With the advent of molecular genetic approaches these collections have become the source of many fascinating population studies in conservation genetics that contrast historical with present-day genetic diversity. Recent developments in molecular genetics and genomics and the associated statistical tools have opened up the further possibility of studying evolutionary change directly. As we discuss here, we believe that NHC specimens provide a largely underutilized resource for such investigations. However, because DNA extracted from NHC samples is degraded, analyses of such samples are technically demanding and many potential pitfalls exist. Thus, we propose a set of guidelines that outline the steps necessary to begin genetic investigations using specimens from NHC.

Introduction

Given that evolution is change over time, documenting and understanding temporal patterns has long been at the heart of evolutionary studies. In disciplines such as palaeontology, inferences about evolutionary processes are drawn from the analyses of temporal patterns in the fossil record. Similarly, our understanding of microevolutionary processes (i.e. changes in gene frequencies over time) has often involved the analyses of records taken over several years; Dobzhansky's [1] early studies of microevolution among *Drosophila* used this approach, a tradition that continues among students of this model organism today [2]. However, such microevolutionary studies were often limited to certain taxa and questions because the time available to document temporal changes was often limited to a few generations.

How can these limitations be overcome? Long term studies, running over several decades, are one possibility and they are yielding fascinating insights, for example, into the role of reinforcement and character displacement in adaptive radiation and speciation [3,4]. Another approach, which gives longer time series, is to extend the data back in time using well preserved fossil samples or specimens from natural history collections (NHC). Here, we review the use of specimens from NHC for the study of evolutionary change. We aim to increase awareness of both the methodological limitations involved in using molecular methods with NHC specimens and their future potential.

We focus on studies of evolutionary change rather than the widespread use of NHC specimens in phylogenetics and phylogeography (e.g. Ref [5]) or pathogen origin and dynamics (e.g. Ref [6]). Similarly, we limit ourselves to studies of NHC specimens and do not consider studies of ancient DNA (Box 1; see [7–8] for excellent reviews on the latter). Although ancient DNA studies have yielded spectacular results [9,10], they will remain restricted to a relatively small set of species because the samples required for such work are rare and difficult to obtain. By contrast, NHC specimens generally cover a broader taxonomic range and are more easily obtained, thus enabling a wider range of questions and taxa to be studied.

NHC samples in conservation genetics

A large proportion of empirical studies of NHC samples published to date contrast past and recent genetic diversity in threatened and endangered populations or species (Table 1). Many now endangered or extinct populations and species became so within the past two centuries [11], a time period that coincides with the establishment of the majority of NHC (but see Ref [12]). As a result, specimens stored in NHC often represent the genetic diversity of populations shortly before significant anthropogenic influence. By inferring temporal changes in neutral genetic diversity, biologists can obtain estimates of the magnitude of anthropogenic influences on population sizes and connectivity (i.e. gene flow between populations) [13] and they can detect cryptic introductions or genetic introgression. Such insights could guide future conservation action.

Population declines and loss of genetic diversity

Low genetic variation in endangered populations is of conservation concern because genetic variation is the raw material required for future adaptive evolution. Low levels of genetic variation can be the consequence of recent population declines, or it can represent an ancestral state. Differentiating these two causes is therefore an important task in conservation biology and it can be achieved by comparing levels of genetic variability among NHC samples collected before a genetic bottleneck with those found in current populations.

One of the earliest studies to use such an approach was an investigation of an endangered population of the greater prairie chicken (*Tympanuchus cupido*) [14]. This study provided direct evidence for a human-induced decrease in genetic diversity over time: specific alleles present in the NHC samples were no longer present in the current population. Such 'ghost alleles' have subsequently

Corresponding author: Wandeler, P. (peter.wandeler@access.uzh.ch).
Available online 7 November 2007.

Box 1. Ancient DNA versus DNA from NHC

NHC specimens have often been included in the definition of ancient DNA [7]. However, there are several differences between ancient DNA and that extracted from NHC. For example, ascertaining authenticity is crucial in ancient DNA work whereas in NHC-based studies authenticity is more easily ascertained by comparisons with results from high-quality DNA samples taken from extant populations. We therefore propose to make a distinction based on the criteria in

Table I, even if there is not always a sharp contrast. Based on our criteria, working with DNA from NHC specimens is feasible when dedicated laboratory facilities are available for working with low-quality DNA and when the guidelines outlined in Boxes 3 and 4 can be fulfilled. By contrast, working with ancient DNA is methodologically considerably more demanding and, therefore, remains the domain of specialist groups and laboratories.

Table I. Differences between DNA extracted from NHC and ancient samples

	DNA from NHC	Ancient DNA
Sample origin	Museum, private and archive collections, herbarium	Archaeological and palaeontological sites, museum
Sample age	≤200 years	~100 000 years ^a
Sample size	Often large enough for population sample	Generally small ^b
Tissue types	Several (see Box 2)	Mainly hard tissue
Hominid samples	Seldom	Common
Main causes of DNA decay	Preservation methods, storage condition	Physical factors at sampling site, storage conditions
Level of DNA decay	High	Very high ^c
Variance of DNA quality among samples	Very high	High
Microsatellite amplification	Common	Seldom
Problems with authenticity	Moderate	Very high
Risk of contamination	High	Very high

^aSample age of one million years might be possible.

^bThe number of investigations with larger sample size is currently increasing.

^cAlthough some ancient samples can yield higher DNA quality.

been reported in several bottlenecked animal populations. By contrast, other studies have reported stable genetic diversity despite declines in population size [24–26]. For example, no obvious loss of genetic diversity was detected among Canadian peregrine falcons (*Falco peregrinus*) despite a DDT induced bottleneck [25].

Inferring the genetically effective population size

The genetically effective population size (N_e) is a key parameter that determines the rate of loss of genetic variation owing to genetic drift. Measuring N_e is thus of interest in conservation biology, but is notoriously difficult [27]. However, several methods based on temporal samples have been developed, including forward probabilistic approaches and backwards coalescence approaches (reviewed in Ref [27]). Using microsatellites amplified from DNA extracted from historical samples, temporal estimates of N_e have been obtained in several different fish species [16,28–30], grizzly bears (*Ursus arctos*) [31], greater prairie chicken [14] and leopard frog (*Rana pipiens*) [32]. Some of these N_e estimates were lower than anticipated whereas others were higher. This pattern highlights the fact that including NHC samples in such analyses can prevent us from reaching misleading conclusions.

Changes in connectivity

Habitat fragmentation often leads to changes in population connectivity. With increasing population fragmentation, migration among populations and, thus, gene flow is expected to decrease. This leads to a reduction in genetic diversity within populations and an elevation of genetic differentiation among populations [18,33]. Such an increase in spatial genetic differentiation over a time period of six to seven generations has been reported for the endangered Spanish imperial eagle (*Aquila adalberti*) [33]. This finding suggests that future management

policies should attempt to restore the ancestral panmictic situation through habitat restoration or translocations.

Detecting introductions and introgression

Specimens from NHC have been used also to detect new introductions and to assess the rate of genetic introgression into indigenous populations. For example, the cryptic invasion of a non-native genotype of the common reed (*Phragmites australis*) in North America during the past century was detected by sequencing samples collected worldwide at two noncoding regions of the chloroplast [34]. Similarly, archival fish scales have been used to quantify the genetic consequences for wild fish populations of stocking or of aquaculture escapees [35,36]. Stocking of Spanish Atlantic salmon (*Salmo salar*) since the 1970s appears to have increased the mitochondrial DNA diversity of four endangered populations. However, it failed to halt the ongoing population decline [35].

NHC samples in evolutionary biology

Fisher and Ford [37] provided an early example of the use of NHC specimens to study evolutionary change directly. Their 1947 study of the spread of the *medionigra* gene among the scarlet tiger moth (*Callimorpha dominula*) provided clear evidence for gene frequency change owing to natural selection. Since then, surprisingly few studies have taken advantage of the evolutionary history preserved in NHC samples to investigate the molecular footprint of selection. One such study investigated the evolutionary genetic mechanisms that underlie the rapid evolution of insecticide resistance in the blowfly (*Lucilia cuprina*) in Australia [38]. Sequence data from 16 pinned NHC specimens collected before the first use of organophosphate insecticides revealed that the rapid evolution of insecticide resistance was because of the pre-existence of mutant alleles in the historical gene pool and that the

Table 1. A selection of studies contrasting past and recent genetic diversity using specimens from NHC^{a,b}

Species	Marker ^c	Time span ^d	Sample ^e N ^f	Summary	Refs
Plants					
Common reed	1400 bp	Before 1910; 5/2	62/283	Cryptic invasion of non-native genotype	[34]
<i>Phragmites australis</i>	cpDNA	after 1960			
Marsh orchid	250–280 bp	1832–recent	5/4	Genetic variation in time and space	[72]
<i>Anacamptis palustris</i>	plastid DNA				
Insects					
Sheep blowfly	287 bp	1930–1990	-/2	16/35 Evolution of insecticide resistance	[38]
<i>Lucilia cuprina/sericata</i>	<i>LcxE7</i>				
Adonis blue <i>Polyommatus bellargus</i>	4	1897–1999	1/2	20/74 High genetic drift in an insect population	[73]
Fish					
Northern pike <i>Esox lucius</i>	7	1961–1993	1/3	196/72 Temporal N_e estimation of introduced population	[29]
Atlantic salmon	6	1913–1994	7/4	228/90 Genetic diversity between and within extant and extinct populations	[74]
<i>Salmo salar</i>					
Newfoundland cod	6	1964–1994	5/3	574/570 Stability of genetic structure despite population size decline	[24]
<i>Gadus morhua</i>					
Brown trout <i>Salmo trutta</i>	8	1910–1992	5/3	191/311 Estimation of N_e and temporal stability of genetic structure	[28]
	9	1945–2000	2/7	146/397 Genetic effects of stocking domestic trout into wild populations	[36]
New Zealand snapper	7	1950–1998	2/5	372/96 Genetic diversity and low N_e in two overexploited populations	[16]
<i>Pagrus auratus</i>					
Steelhead trout	7	1958–1998	3/4	180/90 Temporal diversity and N_e for three populations	[30]
<i>Oncorhynchus mykiss</i>					
Brown trout <i>S. trutta</i>	6; 1 <i>Satr-UBA</i> ^g	1958–1995	1/5	232/50 Temporal variation of MHC class I gene during aquaculture activities	[68]
Atlantic salmon <i>S. salar</i>	1409 bp RFLP	1948–2002	4/5	592/125 Effects of stocking on mitochondrial diversity	[35]
Amphibians					
Leopard frog <i>Rana pipiens</i>	7	1971–2001	5/2	204/188 N_e and temporal stability of genetic structure	[32]
Birds					
Greater prairie chicken	6	1930s–recent	4/2	15/127 Genetic diversity of pre- and post-bottleneck populations	[14]
<i>Tympanuchus cupido</i>					
Mauritius kestrel	12	1830–recent	7/2	52/250 Genetic monitoring of a pesticide-induced bottleneck	[17]
<i>Falco punctatus</i>					
Greater prairie chicken	384 bp; 6	1951–2000	4/2	125/81 N_e and temporal genetic variation in bottlenecked populations	[18]
<i>Tympanuchus cupido</i>					
White-headed duck	192 bp	1861–2003	-/2	67/46 Loss of genetic diversity	[21]
<i>Oxyura leucocephala</i>					
Peregrine falcon	405 bp; 11	1885–2004	-/2	95/184 Pesticide-induced bottleneck in Canadian falcons	[25]
<i>Falco peregrinus</i>					
Spanish imperial eagle	345 bp; 10	1860–recent	1/2	34/79 Effects of fragmentation on spatiotemporal genetic structure	[33]
<i>Aquila adalberti</i>					
Mammals					
Kangaroo rat	225 bp	1911–1988	3/2	49/63 Continuity of spatial and temporal mtDNA diversity	[75]
<i>Dipodomys panamintinus</i>					
Northern hairy-nosed wombat	90 bp; 8	1883–1994	3/2	5/29 Genetic variation of extinct and extant populations	[76]
<i>Lasiorhinus krefftii</i>					
Hector's dolphin	206 bp	1870–1998	2/4	55/108 Loss of genetic diversity owing to fishery-related mortality	[20]
<i>Cephalorhynchus hectori</i>					
European otter <i>Lutra lutra</i>	9	1883–1993	3/3	67/58 Genetic consequences of population decline	[26]
Elephant seal	116 bp; 4	1500s–1990s	1/2	22/185 Effects of bottleneck on genetic diversity and on symmetry of bilateral characters	[19]
<i>Mirounga angustirostris</i>					
Grizzly bear <i>Ursus arctos</i>	8	1912–1999	1/3	110/136 N_e of Yellowstone grizzly	[31]
Common hamster	240 bp; <i>DRB</i> ^h	1924–2000	2/3	20/31 Loss of MHC diversity in the <i>DRB</i> exon 2	[67]
<i>Cricetus cricetus</i>					
Grey wolf <i>Canis lupus</i>	229 bp; 15+ ⁱ	1829–recent	4/2	33/22 Genetic variability and migration during population decline	[77]
	425 bp	1856–recent	-/2	32/399 Genetic variability and population size of extirpated US wolf populations	[23]
Arctic fox <i>Alopex lagopus</i>	292 bp; 5	1831–2004	2/2	21/41 Demographic bottleneck	[15]
Red fox <i>Vulpes vulpes</i>	354 bp	1911–2002	5/2	29/35 Genetic evidence for the persistence of Sierra Nevada red fox	[22]

^aWithin taxonomic groups, the order is by year of publication.^bAbbreviations: cpDNA, chloroplast DNA; MHC, major histocompatibility complex; RFLP, restricted fragment length polymorphisms.^cUnless otherwise noted, sequence length in base pairs (bp) stands for mitochondrial DNA and the single digit for the total number of microsatellites loci applied.^dThe oldest and the most recent year of sampling.^eNumber of populations and (/) number of temporal samples.^fSample size of historical and (/) most recent samples.^gA microsatellite locus embedded in the MHC class I locus of brown trout.^hGene that presents extracellular proteins to T lymphocytes.ⁱRefers to Y chromosome microsatellites.

Box 2. Tissue samples from natural history collections for genetic studies

Several types of tissue have been applied to obtain genetic data from NHC specimens. Here, we provide an overview of the most common tissue types used and their advantages and disadvantages in terms of potential damage to the specimens and the expected DNA quality.

Hides and skins

Epithelium tissue is a preferred source of DNA because damage is often insignificant and sampling is straightforward. Different preservation methods can, however, create substantial DNA degradation and PCR inhibition [48], which causes a high variance in DNA quality among samples. Superior DNA can be extracted from claws on hides of mammalian specimens [49] or from toe pads of birds [78].

Bones

DNA in bones is generally better protected than in epithelium tissue. However, DNA extraction is laborious because the material needs to be ground with a drill or a mortar. The use of maxilloturbinal bone material (i.e. thin bones inside the nasal cavities) might minimize damage to specimens [45], except where these bones are important for morphological analyses.

Teeth

DNA from teeth is generally well preserved given the hard tissue protecting it. Extraction protocols for teeth [51,77] are often identical to bones, although the efficiency can vary among different protocols [79]. To prevent the destruction of the whole tooth, only the root tip can be removed or the material following drilling inside the root cavity can be recovered. Nondestructive extraction methods have been developed [44,46].

Feathers

DNA has been extracted from the base of the feather calamus or the blood clot from the superior umbilicus. The latter is thought to yield higher amounts of DNA [80]. Careful selection of feathers results in little damage to the appearance of the specimen and removed feathers can be reattached [81].

Fish scales

Initially collected for use in age determination of individuals, several extraction protocols have been developed for scales [35,40]. Because they are plucked from fresh material, dermic and epidermic cells remain attached to the scale and dry up sufficiently fast to prevent DNA degradation [74].

Pinned insects

DNA can be extracted from pinned specimens using the whole insect or body parts [73]. Protocols conferring no external damage have recently been developed [82]. Different killing methods and storage conditions can affect the recovered DNA [83].

Fluid specimens

Many NHC specimens are stored in aqueous formalin. DNA extraction and PCR amplification from formalin-fixed specimens, however, is particularly difficult [50,54]. Formalin storage can cause frequent nucleotide misincorporations [50,55]. Particular care is needed to detect potential cross-contamination among samples because several specimens are often stored within the same fluid or because fluid among different storage containers could have been exchanged.

Herbarium specimens

Leaf tissue [34,84] and seeds [85] from plants dried and stored under controlled conditions and not preserved with chemicals can provide useful amounts of nuclear and chloroplast DNA [84].

Ethnographic artefacts

Animal and plant parts are also preserved in artefacts of ethnographic collections. Such collections can contain specimens that were collected before the major collection activities of NHC and thus might yield even longer time series (J. Groombridge, personal communication).

associated selective sweep has led to a significant loss of genetic variability [38]. This study illustrates the power of harnessing modern molecular genetic approaches with long time series data, either through the use of NHC samples or through long term studies [2].

Pitfalls and precautions

NHC hold an unchallenged wealth of specimens that reflect past and current biodiversity of our planet. However, molecular studies based on historical samples are challenging because genotype and sequence data obtained by PCR are often error prone. Consequently, precautions are needed to guarantee reliable genetic data.

Incomplete specimen records and small sample size

Unsurprisingly, most specimens of NHC collections were not assembled with the aim of carrying out genetic studies. As described by Fisher and Ford in their scarlet tiger moth study [37], this can considerably limit the utility of NHC collections for genetic studies. Collectors were likely to be biased in their sampling effort, particularly in relation to colour morphs and other phenotypic varieties. This might be less problematic when neutral genetic markers rather than functional genes are studied, but sampling localities, age and sex of the animals can still be biased. In addition, records for specimens from NHC are frequently incomplete, imprecise, missing or incorrect [39].

In a few instances, archival collections are reasonably comprehensive. For example, commercially and recreationally important fish species are often represented in systematic and continuous collections of scales or otolithes, sampled since the beginning of the 19th century to determine the age of individuals [40]. Few other taxa are represented by such comprehensive collections. Given that the number of historical specimens is limited, sample sizes can only be increased by searching longer for appropriate material, which is likely to exceed the time and effort spent on sampling modern material.

A related problem is that some specimens might not be suitable for molecular studies, because DNA needs to be physically extracted from tissue, which causes potential damage to the specimen (Box 2). Fifteen years ago, this problem led to a debate over the benefits and misuse of molecular studies based on NHC [41,42] and prompted several museums to establish guidelines for the use of specimens from their collections (Box 3). However, given the small amount of tissue now needed for extracting sufficient amounts of DNA for PCR amplification and the development of less or nondestructive DNA extraction methods [43–46], we believe that few NHC specimens should be out of bounds for molecular studies.

Molecular work with low-quality DNA

Numerous biological, physical and chemical factors affect the DNA quality of specimens from NHC. Most of these factors have been extensively described for ancient DNA work (reviewed in Refs [7,8]) and are also relevant for NHC, such as endogenous nuclease activity and hydrolytic damage. As a consequence, DNA extracted from historical material can be expected to be highly degraded and hence highly diluted similar to DNA derived from noninvasive

Box 3. Guidelines for genetic studies using natural history collections: study design

These guidelines are intended to outline the steps necessary to begin genetic investigations using specimens from NHC. They are divided into two sections covering the general (this Box) and the laboratory aspects of such projects (Box 4).

Collaboration with natural history museums

Unless specified otherwise by museums, researchers should provide a short project proposal that includes justification for the required material and evidence of the experience with such work. An agreement between the researchers and the museum is advantageous to clarify storage of and access to surplus extracted DNA and the future use of the genetic data. Museums should be regularly informed, acknowledged in publications and supplied with reprints of publications at the end of the study.

Sample selection

Records of specimens need to be verified to account for potential identification errors (e.g. imprecision or error in the location of a record [39]). Depending on the geographical and temporal scale of the intended study, specimens with imprecise records should be excluded. Only small tissue samples should be taken, causing the least amount of damage while providing a good likelihood of extracting sufficiently preserved DNA (Box 2).

Pilot study

A pilot study is invaluable to evaluate PCR amplification success and, where possible, to quantify the copy number of the target DNA. In addition, the frequency of genotyping errors in microsatellites and single base pair errors in sequence data should be assessed (Box 4). These results will clarify whether the scientific goals are realistic given the number of samples available, the sample quality and the molecular methods applied and will help to justify damage to additional specimens.

Sceptical attitude to own results

As records of specimens might be incorrect, results from genetic analyses should be evaluated carefully. Any samples for which specimen records and genetic data do not match or do not make biological sense should be treated with particular care. Moreover, results that might indicate methodological artefacts rather than biological findings have to be interpreted carefully.

sampling [47]. In addition, different preservation methods can negatively affect the ability to extract, amplify and sequence DNA [48–50]. PCR amplification of historical DNA is, therefore, generally restricted to short amplicons (<200 bp) and is further vulnerable to contamination by recent DNA and PCR products from the study species. Because preservation methods can vary considerably, the variance expected in DNA quality among samples of similar age can be large and the risk of cross contamination is considerable. Consequently, sample age is not the only important factor that affects DNA quality ([44], but see Ref [51]). Molecular work with DNA from NHC specimens requires special precautions, including an isolated and dedicated laboratory environment (Box 4).

The cumulative damage to the DNA can also cause incorrect bases to be inserted during enzymatic amplification. The main source for these alterations are single nucleotide misincorporations ([7,52], and references therein). C to T transitions are the main type of such alterations that occur when an erroneous DNA strand is replicated during the first cycle of a PCR. Initially thought to be limited to ancient DNA, nucleotide misincorporations have recently been reported in studies based on specimens from NHC

Box 4. Guidelines for genetic studies using natural history collections: laboratory work

The main laboratory criteria relevant for working with specimens from NHC are based, in part, on those originally proposed for ancient DNA work [7,8]. However, there are some significant differences between ancient DNA and DNA from NHC (Box 1), which lead to the following guidelines.

Choosing appropriate genetic markers

Because DNA extracted from NHC specimens is highly degraded, amplicons of sequence data and microsatellite loci ideally should be <200 bp in size. All primers should be specifically designed for the species of interest to yield the best possible PCR efficiency and should have been extensively tested in modern samples. Species specific primers often can be redesigned within the flanking region of microsatellite loci and PCR multiplexing can be advantageous to delay the depletion of template DNA. The amplification of overlapping variable sequences of mitochondrial DNA will further diminish the chance that the sequence derives from a nuclear insertion.

Isolation of laboratory work area

Cross-contamination and contamination with exogenous DNA is a key concern when working with historical DNA. Therefore, DNA extraction and PCR preparation should be carried out in a dedicated and isolated laboratory with one way movement of DNA out of the laboratory. Only reagents exclusively purchased and maintained for working within such an environment must be used and extensive decontamination (e.g. UV radiation) of the laboratory, working surfaces and equipment is essential. Because the variation in the quantity of recovered DNA among NHC specimens can be very large, the number of simultaneously extracted samples needs to be kept small. The interspersion of samples that are likely to yield different haplotypes can further improve the detection of cross-contamination.

Negative control for extraction and PCR

Negative extraction and PCR controls need to be included to detect potential contamination in reagents and cross-contamination between samples.

Appropriate molecular behaviour

PCR amplification intensity of historical DNA is inversely related to product size, and products >500 bp are unusual. Deviations from such appropriate behaviour should be cause for careful checking and repetitions.

Reproducibility of genetic data

Several biological and nonbiological processes can damage DNA causing single nucleotide misincorporations [52,53]. Independent PCR and sequencing of unknown haplotypes are therefore required and particularly important sequences should be verified by cloning. Genotyping errors in microsatellites such as allelic dropout and false alleles are widespread in historical samples [15,51] and can be more common among loci with longer fragment size and among older samples [51]. Several independent PCR replications are required to attain a sufficiently high genotyping reliability [47]. Alternatively, the template DNA concentration can be estimated by quantitative PCR to adjust the number of PCR replications [56].

[15,52,53]. Estimates so far range from 17% to 21% of sequences that show one or more errors [15,53]. Special cases are formalin-preserved specimens, in which sequence alterations can occur at even higher frequencies [50,54,55]. Misincorporations might look like new alleles or new sequences and can therefore lead to systematic overestimation of the genetic diversity of past populations [53]. Consequently, repeated sequencing of independent PCR

products and cloning of important sequences are needed to ensure reliable genetic data [52].

The highly degraded and, therefore, diluted nature of DNA extracted from NHC specimens can cause a significant rate of genotyping errors when biparental genetic markers, such as microsatellites, are amplified by PCR. Two types of genotyping errors are likely: allelic dropout and false alleles. Allelic dropout is the stochastic nonamplification of one of the two alleles present at a heterozygote locus [47]. False alleles are PCR amplification artefacts that occur by the slippage of the taq polymerase during the first cycles of the PCR [47]. False alleles are less frequent than allelic dropouts. Because the rate of both types of error is inversely correlated with the concentration of the extracted DNA [47,51,56], a high frequency of allelic dropout can be expected in samples from NHC with low DNA quality. Consequently the genetic diversity measured will be systematically underestimated in historical samples compared with modern samples. This contrasts with the overestimation in sequence data that can occur owing to misincorporations and owing to false alleles. Several approaches have been proposed to achieve higher genotyping reliability for microsatellites in samples with low DNA quantity, including a predefined number of repeated and independent PCR amplifications [47] and the initial quantification of the template DNA concentration [51,56] by quantitative PCR assays.

Appropriate standards for reliable genetic data

On the one hand, the findings of molecular studies based on NHC can be misleading when ignoring the effects of sequence alterations and genotyping errors. On the other hand, it might be difficult to fulfil stringent standards, for example, when only a minute amount of DNA from a precious specimen is available and, as a consequence, only a small number of PCR reactions can be performed. Thus, although we propose guidelines for molecular genetic studies based on NHC samples (Boxes 3 and 4), we caution against sweeping, global standards. Instead, and similar to the approach advocated by Gilbert *et al.* [57] for assessing the authenticity of ancient DNA, we recommend a flexible approach, cognizant of the problems particular to every study.

Investigations using historical material need to be carefully assessed on a case-by-case basis by scrutinizing the magnitude of potential errors in historical genetic data in relation to the general findings of the investigation. For example, a minor and overlooked rate of allelic dropout might have a small effect on estimates of allele frequencies but could significantly alter the outcome of relatedness analyses with NHC samples. In particular, a sceptical attitude [7] towards one's own results is needed when findings can equally well be explained by methodological artefacts and biological processes.

Prospects

In the near future, advances in molecular technologies will enable access to more and more genetic information from specimens archived in NHC. This progress will allow us to shift from neutral genetic markers to specific genes under selection.

Gaining more genetic information from NHC samples

Mitochondrial sequences and microsatellites have been the preferred genetic markers in studies using NHC, which reflects the present genetic tools of choice for non-model organisms in ecology and evolution. Recently, single nucleotide polymorphisms (SNPs) have been put forward as an alternative to microsatellites [58]. SNPs have several advantages, including a known mutation model and a higher genotyping efficiency. Furthermore, they are suitable for highly degraded DNA because genotyping requires only short target DNA sequences (<100 bp) and protocols for genotyping of SNPs have been developed for degraded DNA [59]. However, the discovery of SNPs, that is, the selection of representative samples used for finding polymorphic sites, is crucial and can significantly bias the estimates of genetic variation within and between populations ([58], and references therein). To avoid this ascertainment bias, samples from different time periods and therefore also from historical samples could be used for the discovery of SNPs. Obviously, the degraded nature of historical DNA and the potential of nucleotide misincorporations will make this difficult. At present, the trade-off between these methodological constraints is unresolved and new molecular methods are needed to deploy the full potential of SNPs in NHC based studies. One such method is parallel pyrosequencing with a 454 instrument [60], which has been used in ancient DNA studies [61]. Overall, we believe that SNPs could soon become the marker of choice in studies based on NHC.

Genotyping and sequencing can rapidly deplete the precious and often limited amounts of DNA extracted from NHC specimens. This process can partly be delayed by multiplexing several loci within one PCR reaction, an approach now regularly used in microsatellite and SNP genotyping [59]. PCR multiplex amplification has also been used to obtain the complete sequence of the mitochondrial genome from the woolly mammoth (*Mammuthus primigenius*) from a 200 mg bone sample [62]. However, a complete and representative amplification of the limited DNA is preferred, which would provide a nearly unlimited copy number of genomic DNA for future molecular work. This could be achieved by applying whole-genome amplification (WGA) techniques. Several different WGA protocols have been developed to amplify the genomic DNA from minute amounts of template DNA, that is, from a few cells [63]. Most of these protocols rely on high quality template DNA, although protocols tolerant of degraded DNA have recently been developed [64]. We hope that in the near future novel WGA protocols will be adopted for use with NHC specimens.

Conservation genetics

DNA extracted from NHC specimens adds an important temporal dimension to the genetic study of endangered species and will remain pivotal in conservation genetics to assess phylogenetic positions and evolutionary significant (or management) units and to infer changes in population size and structure. Furthermore, contrasting past and recent genetic variability could serve as a retrospective approach of genetically monitoring populations for conservation and management [65]. Currently, there is growing

interest in understanding gene flow in a heterogeneous landscape by combining landscape and ecological data with individual based genetic data [66]. Historical samples might prove particularly valuable in this context because they enable direct tests of hypotheses concerning human-induced changes in gene flow patterns over time.

Beyond neutral markers

The main body of population genetic research using samples from NHC to date has focussed on temporal changes in genetic diversity of neutral genetic markers. By contrast, only a few NHC-based studies have identified allele frequencies of coding DNA sequences [38,67] or of microsatellites closely linked to sequences potentially under selection [68]. The relative ease with which one can now assay candidate genes and an increasingly large number of genetic markers makes it possible to study (natural) selection directly using temporal samples. Comparing the genetic variation from different time periods can be a powerful tool to detect molecular signatures of selection or environmental causes of selection. This has been exemplified by documenting insecticide resistance [38], evolutionary response to climate change [2] and human-induced allelic selection in maize [69].

Nevertheless, these approaches are not without caveats. Temporal variation of allele frequencies of a gene suspected to be under selection (or of a marker linked to it) occur because of selection, genetic drift or a combination of both. Consequently, statistical methods that separate the effects of selection from drift are required and these are often sensitive to assumptions about the demography of the populations [70]. Furthermore, different selection regimes can yield the same allele frequency dynamics [71], thus limiting the details that can be inferred from allele frequency dynamics data alone. Statistical analyses of selection are, however, a very active research field [70] and future methodological developments are likely to open up more possibilities.

Conclusions

Analyses of DNA from NHC samples have played an important role in conservation genetics by identifying processes that have shaped current levels of genetic diversity. A strong taxonomic bias is apparent among the studies to date (Table 1). Vertebrates and particularly fish predominate, whereas studies on plants and invertebrates are surprisingly rare. This bias might, in part, have arisen from biases in sample availability, conservation interests and methodological constraints, but the lack of more plant studies remains puzzling.

As NHC-based genetic work is coming of age, studies might fruitfully shift from investigating neutral genetic variation to studying the interplay of selection and drift. Although pioneered by Fisher and Ford 60 years ago [37], such direct studies of evolutionary change through the use of NHC samples are only beginning. Although technically demanding, such studies will be worth all the effort. Fisher and Ford [37] put it nicely: 'The spread of a gene in natural conditions is an event which repays detailed study since it provides an opportunity for examining evolution in progress. It has long been apparent to us that a careful watch

should be kept for this occurrence and that whenever found it should, if possible, be analysed from two distinct points of view, ecological and genetic – a technique which has so far received much less attention than it deserves.' Combined with recent samples and, where possible, with ecological data that represent both current and historical time periods [39], NHC specimens provide a largely untapped resource for such investigations.

Acknowledgements

We thank Mark Beaumont, Love Dalén, Felix Gugerli, Michael Hofreiter and Phil Morin for helpful discussions and comments. Homayoun Bagheri, Jim Groombridge, Ken Petren and three anonymous referees made helpful comments on an earlier version of this article.

References

- 1 Dobzhansky, T. (1943) Genetics of natural populations IX. Temporal changes in the composition of populations of *Drosophila pseudoobscura*. *Genetics* 28, 162–186
- 2 Balanya, J. *et al.* (2006) Global genetic change tracks global climate warming in *Drosophila subobscura*. *Science* 313, 1773–1775
- 3 Silvertown, J. *et al.* (2006) The Park Grass Experiment 1856–2006: Its contribution to ecology. *J. Ecol.* 94, 801–814
- 4 Grant, P.R. and Grant, B.R. (2006) Evolution of character displacement in Darwin's finches. *Science* 313, 224–226
- 5 Shapiro, B. *et al.* (2002) Flight of the dodo. *Science* 295, 1683
- 6 Bearchell, S.J. *et al.* (2005) Wheat archive links long-term fungal pathogen population dynamics to air pollution. *Proc. Natl. Acad. Sci. U. S. A.* 102, 5438–5442
- 7 Paabo, S. *et al.* (2004) Genetic analyses from ancient DNA. *Annu. Rev. Genet.* 38, 645–679
- 8 Hofreiter, M. *et al.* (2001) Ancient DNA. *Nat. Rev. Genet.* 2, 353–359
- 9 Hadly, E.A. *et al.* (2004) Genetic response to climatic change: insights from ancient DNA and phylochronology. *PLoS Biol.* 2, e290 (<http://biology.plosjournals.org>)
- 10 Shepherd, L.D. *et al.* (2005) Microevolution and mega-icebergs in the Antarctic. *Proc. Natl. Acad. Sci. U. S. A.* 102, 16717–16722
- 11 Ceballos, G. and Ehrlich, P.R. (2002) Mammal population losses and the extinction crisis. *Science* 296, 904–907
- 12 Paxinos, E.E. *et al.* (2002) Prehistoric decline of genetic diversity in the nene. *Science* 296, 1827
- 13 Roy, M.S. *et al.* (1994) The use of museum specimens to reconstruct the genetic-variability and relationships of extinct populations. *Experientia* 50, 551–557
- 14 Bouzat, J.L. *et al.* (1998) The ghost of genetic diversity past: historical DNA analysis of the greater prairie chicken. *Am. Nat.* 152, 1–6
- 15 Nyström, V. *et al.* (2006) Genetic consequences of a demographic bottleneck in the Scandinavian arctic fox. *Oikos* 114, 84–94
- 16 Hauser, L. *et al.* (2002) Loss of microsatellite diversity and low effective population size in an overexploited population of New Zealand snapper (*Pagrus auratus*). *Proc. Natl. Acad. Sci. U. S. A.* 99, 11742–11747
- 17 Groombridge, J.J. *et al.* (2000) 'Ghost' alleles of the Mauritius kestrel. *Nature* 403, 616
- 18 Johnson, J.A. *et al.* (2004) Temporal changes in allele frequencies and low effective population size in greater prairie-chickens. *Mol. Ecol.* 13, 2617–2630
- 19 Hoelzel, A.R. *et al.* (2002) Impact of a population bottleneck on symmetry and genetic diversity in the northern elephant seal. *J. Evol. Biol.* 15, 567–575
- 20 Pichler, F.B. and Baker, C.S. (2000) Loss of genetic diversity in the endemic Hector's dolphin due to fisheries-related mortality. *Proc. R. Soc. Lond. B Biol. Sci.* 267, 97–102
- 21 Munoz-Fuentes, V. *et al.* (2005) Population structure and loss of genetic diversity in the endangered white-headed duck, *Oxyura leucocephala*. *Conserv. Genet.* 6, 999–1015
- 22 Perrine, J.D. *et al.* (2007) Genetic evidence for the persistence of the critically endangered Sierra Nevada red fox in California. *Conserv. Genet.* 8, 1083–1095
- 23 Leonard, J.A. *et al.* (2005) Legacy lost: genetic variability and population size of extirpated US grey wolves (*Canis lupus*). *Mol. Ecol.* 14, 9–17

- 24 Ruzzante, D.E. *et al.* (2001) Stability in the historical pattern of genetic structure of Newfoundland cod (*Gadus morhua*) despite the catastrophic decline in population size from 1964 to 1994. *Conserv. Genet.* 2, 257–269
- 25 Brown, J.W. *et al.* (2007) Appraisal of the consequences of the DDT-induced bottleneck on the level and geographic distribution of neutral genetic variation in Canadian peregrine falcons, *Falco peregrinus*. *Mol. Ecol.* 16, 327–343
- 26 Pertoldi, C. *et al.* (2001) Genetic consequences of population decline in the European otter (*Lutra lutra*): an assessment of microsatellite DNA variation in Danish otters from 1883 to 1993. *Proc. R. Soc. Lond. B Biol. Sci.* 268, 1775–1781
- 27 Wang, J.L. (2005) Estimation of effective population sizes from data on genetic markers. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 1395–1409
- 28 Hansen, M.M. *et al.* (2002) Long-term effective population sizes, temporal stability of genetic composition and potential for local adaptation in anadromous brown trout (*Salmo trutta*) populations. *Mol. Ecol.* 11, 2523–2535
- 29 Miller, L.M. and Kapuscinski, A.R. (1997) Historical analysis of genetic variation reveals low effective population size in a northern pike (*Esox lucius*) population. *Genetics* 147, 1249–1258
- 30 Heath, D.D. *et al.* (2002) Temporal change in genetic structure and effective population size in steelhead trout (*Oncorhynchus mykiss*). *Mol. Ecol.* 11, 197–214
- 31 Miller, C.R. and Waits, L.P. (2003) The history of effective population size and genetic diversity in the Yellowstone grizzly (*Ursus arctos*): Implications for conservation. *Proc. Natl. Acad. Sci. U. S. A.* 100, 4334–4339
- 32 Hoffman, E.A. *et al.* (2004) Effective population sizes and temporal stability of genetic structure in *Rana pipiens*, the northern leopard frog. *Evolution Int. J. Org. Evolution* 58, 2536–2545
- 33 Martinez-Cruz, B. *et al.* (2007) Population fragmentation leads to spatial and temporal genetic structure in the endangered Spanish imperial eagle. *Mol. Ecol.* 16, 477–486
- 34 Saltonstall, K. (2002) Cryptic invasion by a non-native genotype of the common reed, *Phragmites australis*, into North America. *Proc. Natl. Acad. Sci. U. S. A.* 99, 2445–2449
- 35 Ciborowski, K.L. *et al.* (2007) Stocking may increase mitochondrial DNA diversity but fails to halt the decline of endangered Atlantic salmon populations. *Conserv. Genet.* DOI: 10.1007/s10592-007-9286-2 (<http://www.springerlink.com/content>)
- 36 Hansen, M.M. (2002) Estimating the long-term effects of stocking domesticated trout into wild brown trout (*Salmo trutta*) populations: an approach using microsatellite DNA analysis of historical and contemporary samples. *Mol. Ecol.* 11, 1003–1015
- 37 Fisher, R.A. and Ford, E.B. (1947) The spread of a gene in natural conditions in a colony of the moth *Panaxia dominula* L. *Heredity* 1, 143–174
- 38 Hartley, C.J. *et al.* (2006) Amplification of DNA from preserved specimens shows blowflies were preadapted for the rapid evolution of insecticide resistance. *Proc. Natl. Acad. Sci. U. S. A.* 103, 8757–8762
- 39 Graham, C.H. *et al.* (2004) New developments in museum-based informatics and applications in biodiversity analysis. *Trends Ecol. Evol.* 19, 497–503
- 40 Nielsen, E.E. *et al.* (1997) Analysis of microsatellite DNA from old scale samples of Atlantic salmon *Salmo salar*: A comparison of genetic composition over 60 years. *Mol. Ecol.* 6, 487–492
- 41 Graves, G.R. and Braun, M.J. (1992) Museums - storehouses of DNA. *Science* 255, 1335–1336
- 42 Schram, F.R. (1992) Museum collections - why are they there? *Science* 256, 1502
- 43 Hutchinson, W.F. *et al.* (1999) A nondestructive technique for the recovery of DNA from dried fish otoliths for subsequent molecular genetic analysis. *Mol. Ecol.* 8, 893–894
- 44 Rohland, N. *et al.* (2004) Nondestructive DNA extraction method for mitochondrial DNA analyses of museum specimens. *Biotechniques* 36, 814–821
- 45 Wisely, S.M. *et al.* (2004) A technique for sampling ancient DNA that minimizes damage to museum specimens. *Conserv. Genet.* 5, 105–107
- 46 Pichler, F.B. *et al.* (2001) Nondestructive DNA extraction from sperm whale teeth and scrimshaw. *Mol. Ecol. Notes* 1, 106–109
- 47 Taberlet, P. *et al.* (1999) Noninvasive genetic sampling: look before you leap. *Trends Ecol. Evol.* 14, 323–327
- 48 Hall, L.M. *et al.* (1997) Association of enzyme inhibition with methods of museum skin preparation. *Biotechniques* 22, 928–934
- 49 Hedmark, E. and Ellegren, H. (2005) Microsatellite genotyping of DNA isolated from claws left on tanned carnivore hides. *Int. J. Legal Med.* 119, 370–373
- 50 Tang, E.P.Y. (2006) *Path to Effective Recovering of DNA from Formalin-Fixed Biological Samples in Natural History Collections Workshop Summary*, The National Academies Press
- 51 Wandeler, P. *et al.* (2003) Patterns of nuclear DNA degeneration over time - a case study in historic teeth samples. *Mol. Ecol.* 12, 1087–1093
- 52 Stiller, M. *et al.* (2006) Patterns of nucleotide misincorporations during enzymatic amplification and direct large-scale sequencing of ancient DNA. *Proc. Natl. Acad. Sci. U. S. A.* 103, 13578–13584
- 53 Sefc, K.M. *et al.* (2007) Single base errors in PCR products from avian museum specimens and their effect on estimates of historical genetic diversity. *Conserv. Genet.* 8, 879–884
- 54 Stuart, B.L. *et al.* (2006) Extraction of nuclear DNA from bone of skeletonized and fluid-preserved museum specimens. *Syst. Biodivers.* 4, 133–136
- 55 Williams, C. *et al.* (1999) A high frequency of sequence alterations is due to formalin fixation of archival specimens. *Am. J. Pathol.* 155, 1467–1471
- 56 Morin, P.A. *et al.* (2001) Quantitative polymerase chain reaction analysis of DNA from noninvasive samples for accurate microsatellite genotyping of wild chimpanzees (*Pan troglodytes verus*). *Mol. Ecol.* 10, 1835–1844
- 57 Gilbert, M.T.P. *et al.* (2005) Assessing ancient DNA studies. *Trends Ecol. Evol.* 20, 541–544
- 58 Morin, P.A. *et al.* (2004) SNPs in ecology, evolution and conservation. *Trends Ecol. Evol.* 19, 208–216
- 59 Morin, P.A. and Mc Carthy, M. (2007) Highly accurate SNP genotyping from historical and low-quality samples. *Mol. Ecol. Notes* 7, 937–946 (<http://www.blackwell-synergy.com/loi/men>)
- 60 Margulies, M. *et al.* (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437, 376–380
- 61 Poinar, H.N. *et al.* (2006) Metagenomics to paleogenomics: large-scale sequencing of mammoth DNA. *Science* 311, 392–394
- 62 Krause, J. *et al.* (2006) Multiplex amplification of the mammoth mitochondrial genome and the evolution of Elephantidae. *Nature* 439, 724–727
- 63 Panelli, S. *et al.* (2006) Towards the analysis of the genomes of single cells: further characterisation of the multiple displacement amplification. *Gene* 372, 1–7
- 64 Wang, G. *et al.* (2004) DNA amplification method tolerant to sample degradation. *Genome Res.* 14, 2357–2366
- 65 Schwartz, M.K. *et al.* (2007) Genetic monitoring as a promising tool for conservation and management. *Trends Ecol. Evol.* 22, 25–33
- 66 Manel, S. *et al.* (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends Ecol. Evol.* 18, 189–197
- 67 Smulders, M.J.M. *et al.* (2003) Complete loss of MHC genetic diversity in the common hamster (*Cricetus cricetus*) population in The Netherlands. Consequences for conservation strategies. *Conserv. Genet.* 4, 441–451
- 68 Coughlan, J. *et al.* (2006) Temporal variation in an immune response gene (MHC I) in anadromous *Salmo trutta* in an Irish river before and during aquaculture activities. *ICES J. Mar. Sci.* 63, 1248–1255
- 69 Jaenicke-Despres, V. *et al.* (2003) Early allelic selection in maize as revealed by ancient DNA. *Science* 302, 1206–1208
- 70 Nielsen, R. (2005) Molecular signatures of natural selection. *Annu. Rev. Genet.* 39, 197–218
- 71 Denniston, C. and Crow, J.F. (1990) Alternative fitness models with the same allele frequency dynamics. *Genetics* 125, 201–205
- 72 Cozzolino, S. *et al.* (2007) Genetic variation in time and space: the use of herbarium specimens to reconstruct patterns of genetic variation in the endangered orchid *Anacamptis palustris*. *Conserv. Genet.* 8, 629–639
- 73 Harper, G.L. *et al.* (2006) Analysis of museum specimens suggests extreme genetic drift in the adonis blue butterfly (*Polyommatus bellargus*). *Biol. J. Linnean Soc.* 88, 447–452
- 74 Nielsen, E.E. *et al.* (1999) Analysis of DNA from old scale samples: technical aspects, applications and perspectives for conservation. *Hereditas* 130, 265–276
- 75 Thomas, W.K. *et al.* (1990) Spatial and temporal continuity of kangaroo rat populations shown by sequencing mitochondrial-DNA from museum specimens. *J. Mol. Evol.* 31, 101–112

- 76 Taylor, A.C. *et al.* (1994) Genetic variation of microsatellite loci in a bottlenecked species: the northern hairy-nosed wombat *Lasiorhinus krefftii*. *Mol. Ecol.* 3, 277–290
- 77 Flagstad, O. *et al.* (2003) Two centuries of the Scandinavian wolf population: patterns of genetic variability and migration during an era of dramatic decline. *Mol. Ecol.* 12, 869–880
- 78 Mundy, N.I. *et al.* (1997) Skin from feet of museum specimens as a non-destructive source of DNA for avian genotyping. *Auk* 114, 126–129
- 79 Rohland, N. and Hofreiter, M. (2007) Comparison and optimization of ancient DNA extraction. *Biotechniques* 42, 343–352
- 80 Horváth, M.B. *et al.* (2005) An overlooked DNA source for non-invasive genetic analysis in birds. *J. Avian Biol.* 36, 84–88
- 81 Payne, R.B. and Sorenson, M.D. (2002) Museum collections as sources of genetic data. *Bonn. Zool. Beitr.* 51, 97–104
- 82 Gilbert, M.T.P. *et al.* (2007) DNA extraction from dry museum beetles without conferring external morphological damage. *PLoS One* DOI: 10.1371/journal.pone.0000272 (<http://www.plosone.org>)
- 83 Dean, M.D. and Ballard, J.W.O. (2001) Factors affecting mitochondrial DNA quality from museum preserved *Drosophila simulans*. *Entomol. Exp. Appl.* 98, 279–283
- 84 Drabkova, L. *et al.* (2002) Comparison of seven DNA extraction and amplification protocols in historical herbarium specimens of Juncaceae. *Plant Mol. Biol. Rep.* 20, 161–175
- 85 Walters, C. *et al.* (2006) The utility of aged seeds in DNA banks. *Seed Sci. Res.* 16, 169–178

**Have you contributed to an Elsevier publication?
Did you know that you are entitled to a 30% discount on
books?**

A 30% discount is available to all Elsevier book and journal contributors when ordering books or stand-alone CD-ROMs directly from us.

To take advantage of your discount:

1. Choose your book(s) from www.elsevier.com or www.books.elsevier.com

2. Place your order

Americas:

Phone: +1 800 782 4927 for US customers

Phone: +1 800 460 3110 for Canada, South and Central America customers

Fax: +1 314 453 4898

author.contributor@elsevier.com

All other countries:

Phone: +44 (0)1865 474 010

Fax: +44 (0)1865 474 011

directorders@elsevier.com

You'll need to provide the name of the Elsevier book or journal to which you have contributed. Shipping is free on prepaid orders within the US.

If you are faxing your order, please enclose a copy of this page.

3. Make your payment

This discount is only available on prepaid orders. Please note that this offer does not apply to multi-volume reference works or Elsevier Health Sciences products.

For more information, visit www.books.elsevier.com

Inferring the ancient population structure of the vulnerable albatross *Phoebastria albatrus*, combining ancient DNA, stable isotope, and morphometric analyses of archaeological samples

Masaki Eda · Hiroko Koike · Masaki Kuro-o ·

Shozo Mihara · Hiroshi Hasegawa ·

Hiroyoshi Higuchi

Received: 30 March 2011 / Accepted: 2 September 2011 / Published online: 28 September 2011

© Springer Science+Business Media B.V. 2011

Abstract The history of population structure is a key to effective wildlife management and conservation. However, inferring the history of population structure using present genetic structures is problematic when the method is applied to species that have experienced severe population bottlenecks. Ancient DNA analysis seemed to be a promising, direct method for inferring ancient population structures. However, the usual methods for inferring modern population structure, i.e. the phylogeographic approach using mitochondrial DNA and the Bayesian approach using microsatellite DNA, are often unsuitable for ancient samples. In this study, we combined ancient DNA obtained from zooarchaeological bones with carbon/nitrogen stable isotope ratios and morphological variations to infer ancient population structure of the short-tailed

albatross *Phoebastria albatrus*. The results showed that the bird existed in two populations, between which the genetic distance was greater than that of distinct sister albatross species, although no subspecies of *P. albatrus* have been proposed. Our results suggest that the birds at the present two breeding regions (Torishima in the Izu Islands and two islets of the Senkaku Islands) are descended from these two ancient populations, and that reevaluation of the status and conservation strategy for the species is required. Our results also indicate that lineage breeding on the Senkaku Islands has drastically reduced genetic diversity, while that on Torishima has not. The approach proposed in this study would be useful for inferring ancient population structure, using samples of highly mobile animals and/or samples from archaeological sites, and the reconstructed ancient population structure would be useful for conservation and management recommendations.

Electronic supplementary material The online version of this article (doi:10.1007/s10592-011-0270-5) contains supplementary material, which is available to authorized users.

M. Eda (✉) · H. Higuchi
School of Agriculture and Life Sciences, University
of Tokyo, Yayoi 1-1-1, Bunkyo-ku, Tokyo 113-8657, Japan
e-mail: edamsk@med.tottori-u.ac.jp

Present Address:
M. Eda
Faculty of Medicine, Tottori University, Nishi-cho 86,
Yonago, Tottori 683-8503, Japan

H. Koike · S. Mihara
Graduate School of Social and Cultural Studies, Kyushu
University, Motooka 744, Nishi-ku, Fukuoka 819-0395, Japan

Present Address:
H. Koike
The Kyushu University Museum, Hakozaki 6-10-1,
Higashi-ku, Fukuoka 812-8581, Japan

M. Kuro-o
Department of Biology, Hirosaki University, Bunkyo-cho 3,
Hirosaki, Aomori 036-8561, Japan

Present Address:
S. Mihara
Faculty of Urban Liberal Arts, Tokyo Metropolitan University,
Minami-Osawa 1-1, Hachioji, Tokyo 192-0397, Japan

H. Hasegawa
Biology Department, Toho University, Miyama 2-2-1,
Funabashi, Chiba 274-8510, Japan

Keywords Ancient DNA · Archaeological materials · *Phoebastria albatrus* · Population structure · Stable isotope analysis · Short-tailed albatross

Introduction

Because all evolutionary forces, such as natural selection, genetic drift, mutation, and gene flow, are ultimately related to population structure, the history of population structure is an important element in understanding evolution (Avise 2000; Fox et al. 2001; Futuyma 1986). Present population structures are usually inferred by analysing genetic structure, i.e. the distribution patterns of neutral molecular markers among breeding sites (Avise 2000). Maternally inherited mitochondrial DNA (mtDNA) and bisexual inheritance microsatellite DNA are two of the molecular markers most widely used for this purpose, because of their fast mutation rates. Present genetic structures reflect both ongoing and past gene flow (Avise 2000; Avise et al. 1987; Templeton et al. 1995). Avise et al. (1987) argued that the history of population structure can be reconstructed from the current genetic structure by assuming that genetic drift has fixed different alleles among isolated populations. The history of population structure in many organisms has been estimated from this phylogeographic perspective and used in wildlife management and conservation.

However, inferring the history of population structure using present genetic structures is problematic when the method is applied to reconstruct the population histories of species that have experienced severe population bottlenecks. This is because allele frequencies may have been influenced by genetic drift and historical information on the population structure can be lost (Crandall et al. 2000; Excoffier and Schneider 1999; Knowles and Maddison 2002; Templeton and Georgiadis 1995). For example, Crandall et al. (2000) found that two recently divided small (e.g. bottlenecked) populations could be regarded as two long-term isolated populations, if divergent alleles in an ancestral population were fixed by genetic drift that has a more pronounced effect in small populations than in large ones. They suggested that in such the case, to manage the “two small populations” as different management units would result in decreased genetic diversity and increased risk of inbreeding depression (Crandall et al. 2000).

Ancient DNA, which is DNA extracted from historical, paleontological, and archaeological remains, has recently been applied to wildlife conservation (Leonard 2008; Ramakrishnan and Hadly 2009). Ancient DNA analysis is a promising, direct method for inferring ancient population structures, which can be studied using resident animals or

animal bones that have accumulated within colonies (Hadly et al. 1998, 1995; Lawrence et al. 2008). However, the usual methods for inferring modern population structure, i.e. the phylogeographic approach using mtDNA (Avise 2000) and the Bayesian approach using microsatellite DNA, or STRUCTURE analysis (Pritchard et al. 2000; Rosenberg et al. 2002), are often unsuitable for ancient samples. The phylogeographic approach requires information on the breeding site for each sample. However, knowing where the animals bred is impossible for samples of mobile animals and samples from archaeological material. STRUCTURE analysis for inferring the ancient population structure is an impractical solution for this problem, because low numbers of nuclear DNA copies remain in typical ancient samples and nuclear DNA analysis is difficult. For example, Ramakrishnan & Hadly (2009) reviewed ancient-DNA studies that investigated population-level genetic variation, only two of which dealt with nuclear DNA analysis: one examined penguin fossils from permafrost, in which DNA was unusually well preserved (Shepherd et al. 2005), and the other was a maize study that amplified 50- to 80-bp target genes (Jaenicke-Despres et al. 2003).

The short-tailed albatross *Phoebastria albatrus* is a vulnerable seabird species that is regarded as a management unit, according to IUCN criteria (BirdLife International 2008; U. S. Fish and Wildlife Service 2008). Whereas several millions of birds bred in more than 13 sites in the late 19th century, feather hunting drastically reduced their numbers and caused the extirpation of many breeding sites (Hasegawa 2003; Tickell 2000; U. S. Fish and Wildlife Service 2008). In the mid-twentieth century, only about 50 individuals and one breeding site, Torishima in the Izu Islands (Fig. 1), remained. However, since the cessation of commercial harvesting, the population has grown steadily, and current estimates of world population size are around 3,000 individuals, with about 2,500 birds on Torishima and 500 birds on two islets of the Senkaku Islands (Fig. 1).

In an earlier study, we analyzed mtDNA control region 2 (CR2) of modern short-tailed albatrosses from Torishima and the Senkaku Islands, and found two distinct haplotypic clades, one specific to Torishima and the other distributed across both regions (Eda et al. 2010; Kuro-o et al. 2010). This genetic structure is classified to Phylogeographic pattern II and can be explained by two population history scenarios (Avise 2000): two isolated populations with recent co-occupation on Torishima, or a population with a large evolutionarily effective population size and significant gene flow in which two separate, ancient lineages were retained by chance. Different conservation strategies follow from each scenario; in the first case, the species should be regarded as comprising two different evolutionarily significant units or management units (Moritz 1994), whereas in the second case, the species could be

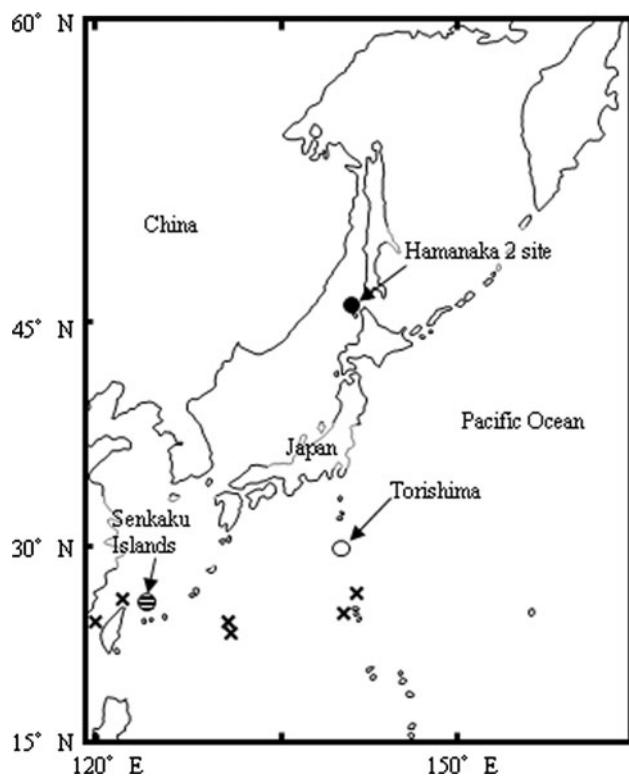


Fig. 1 Distribution of the breeding sites for the short-tailed albatross (Torishima and Senkaku Islands) and sampled archaeological site (Hamanaka 2 site). *X* Former breeding site

managed as a single genetically diverse entity. Therefore knowledge of ancient population structure is important for determining the conservation strategy for this species.

In the previous study, we identified short-tailed albatross bones from Hamanaka 2 site (HM2) on Rebun Island in northern Japan (Fig. 1), using ancient DNA analysis (Eda et al. 2006). Because the bones seemed to result from non-breeding and/or foraging birds harvested by ancient humans at sea (Eda et al. 2005), they were unsuitable for reconstructing ancient population structure via the conventional phylogeographic approach.

In order to investigate the ancient population structure, we obtained mtDNA CR2 sequence, carbon/nitrogen stable isotope ratios and morphological data from each of the archaeological short-tailed albatross samples and combined them in a new approach (see details in [materials and methods](#)). In the results, we suggested that short-tailed albatross existed in two populations, between which the genetic distance was greater than that of distinct sister albatross species. Our results suggest that the birds at the present two breeding sites would be descended from these two ancient populations, and that reevaluation of the status and conservation strategy for the species is required.

Materials and methods

Materials

Fifty-eight left-side Diomedeidae carpometacarpi that included the proximal articulate of Os metacarpale majus and minus were collected from HM2 (Eda et al. 2005; Maeda and Yamaura 1992). All of these bones were determined to be about 1,000 years old, because they were found with pottery and stone tools from the Okhotsk period, between the 7 and 12th centuries AD (Maeda and Yamaura 1992). Thirty-five of the 58 samples were also studied in Eda et al. (2006). Using morphological criteria (Eda et al. 2006), 23 of the 58 bones were labeled osteologically mature while the others were considered immature or unknown. In addition, one ulna collected from the Minami-kojima of the Senkaku Islands in May 1996 was analyzed to supplement the scarce sample size for the islands in the previous study (Kuro-o et al. 2010). In total, we analyzed 45 modern samples (41 from Torishima and four from the Senkaku Islands) in the following analysis.

DNA extraction and sequencing

Ancient DNA extractions with precautions to avoid contamination were conducted as described in Eda et al. (2006). The target sequences (CR2 domain I (139 bp) and cytochrome *b* (cyt *b*; 143 bp) from mtDNA) were amplified and sequenced independently for each sample, and authentic sequences were deduced as described in Hofreiter et al. (2001). The extraction was repeated for each sample that shared its CR2 sequence with another sample to be confirmed sample origin. The amplification of cyt *b* was achieved as described in Eda et al. (2006). For the amplification of CR2, semi-nested PCR was done using hot-start ExTaq polymerase (Takara) with the same program used for cyt *b* amplification and the following primers: 1st round, Lcon2.dio (Eda et al. 2010) and Hcon3.dio (5'-TAC CATTCAARRGTATGYTGCTCRAGC-3') and 2nd round, Lcon3.dio (5'-ATTAAATGCATGTACTACGCACATAS AC-3') or Lcon5.dio (5'-CATTAAATGCATGTACTACG CACATACACATA-3') and Hcon3.dio. Sequencing was done as described in Eda et al. (2006), using Lcon3.dio, Lcon5.dio and Hcon3.dio. The negative controls yielded no product. Whole DNA was extracted from an ulna collected from the Senkaku Islands as described in Kuro-o et al. (2010), and the CR2 sequences were determined. All sequences have been deposited in DDBJ/EMBL/GenBank under accession numbers AB551841–AB551868.

DNA data analyses

The haplotypes were aligned using ClustalX in MEGA 4.0 (Tamura et al. 2007). A statistical parsimony network for the CR2 haplotypes was generated using TCS version 1.2.1 (Clement et al. 2000) with the sequences from modern short-tailed albatrosses (Kuro-o et al. 2010). Because the nesting algorithm was terminated with the non-parsimonious union of two 4-step clades in the TCS, the least genetic distance between the two clades was calculated using MEGA. The resulting haplotype genealogy was converted to a nested design using a previously described procedure (Templeton and Sing 1993). In this procedure, briefly, haplotypes (“0-step clades”) separated by a single substitution were grouped together into “one-step clades” proceeding from the tips to the interior of the haplotype genealogy, then one-step clades by a single substitution were grouped together into “two-step clades”, and so on until the next level of nesting would encompass the entire haplotype genealogy. Nucleotide and haplotype diversity was calculated using DnaSP version 5.00.07 (Librado and Rozas 2009). MEGA was employed to calculate inter haplotype group sequence divergence.

Bone measurement and stable isotope analysis

We analysed bone measurements and stable nitrogen/carbon isotope ratios as intrinsic traits that were independent from neutral molecular marker. These traits primarily reflect food availability during the nestling period (e.g. Klicka et al. 1999; Rasmussen 1994), the food chain that the animal belongs to (Deniro and Epstein 1978), and the trophic levels within the food chain (Deniro and Epstein 1981). For bone length, the greatest length of carpometacarpi (Driesch 1976) was measured. For the carbon/nitrogen isotope ratio analyses, collagen was extracted from the samples and measured using ANCA-SL (Europa Scientific Ltd.) as described in Mihara et al. (2002). At least two aliquots (about 0.8 mg of collagen in tin capsules per sample) were combusted in ANCA-SL. The difference between the stable isotope measurements for the various aliquots was 0.08 ± 0.01 for carbon and 0.13 ± 0.01 for nitrogen. Glycine was used as the standard. The purity of the extracted collagen samples was determined by the carbon/nitrogen ratio, calculated from the wt% C by wt% N. The collagen carbon/nitrogen ratios were between 2.94 and 2.95, and fell within the range of “collagen” (DeNiro 1985).

Reconstruction of an ancient population structure

The logic behind the approach is as follows. A random-mating population is defined as all alleles, including neutral

and functional ones, distributed randomly among individuals and areas (Futuyma 1986; Hartl and Clark 1997). Intrinsic traits are determined by environmental and/or genetic factors (Futuyma 1986); if traits vary among clades classified by neutral molecular markers, environmental and/or genetic differences should exist between the clades. Differences that arose only through environmental variation imply a geographically structured clade distribution, those that arose through genetic differences imply a structured distribution of functional alleles in particular clades, and those that arose through a combination of environmental and genetic differences imply both a geographically structured clade distribution and structured distribution of functional alleles. In all cases, the assumption of random allele distribution is rejected; therefore, if traits differ among clades, individuals belonging to different clades may form different populations. The null hypothesis (no association between intrinsic markers and the nested group of clades) was examined with respect to the average excess ($a_{i(M)}$; Templeton et al. 1988) and analogous to the clade distance (Templeton 1993).

Let $a_{i(M)}$ be the average excess of the b -step clade, i , nested within $b + 1$ step clade, M . Then,

$$a_{i(M)} = \sum_{K=1}^n y_k h_{ik} / \sum_{K=1}^n h_{ik} - \sum_{K=1}^n y_k h_{Mk} / \sum_{K=1}^n h_{Mk}$$

where n refers to the number of individuals, y_k refers to the intrinsic marker value of individual k , and h_{ik} and h_{Mk} refer to the total number of, respectively, clade i and M haplotypes carried by individual k . Let $p_{i(M)}$ be the relative frequency of clade i within clade, M . That is,

$$p_{i(M)} = \sum_{K=1}^n h_{ik} / \sum_{K=1}^n h_{Mk}$$

Then, the weighted sum of squares of the b -step clades nested within the clade M , $S_{c(M)}$, is

$$S_{c(M)} = \sum_i p_{i(M)} a_{i(M)}^2$$

where i is the summed over all b -step clades contained within clade M . In the nested clade analysis, the clade distance is defined as the mean spatial distance of members of a clade from the geographical center of the clade (Templeton 1993). Similar to the distance, we calculated $V_{(i)}$, the variance of intrinsic markers for all individuals of each b step clade, i .

The significance was calculated by comparison with a null distribution constructed from 5,000 random permutations against the intrinsic marker values as described by Templeton et al. (1995) and Templeton & Sing (1993) using R 2.10.1 (R Development Core Team 2009). Null distributions were constructed by randomizing the data table for each $b + 1$ step clade, M , in which haplotype frequencies and sample

sizes were preserved, and estimating again the test statistics (i.e. $S_{c(M)}$ and $V_{(i)}$) for each randomized data set. These tests are asymptotically independent, both at the same clade level and across clade levels because of the nested design (Prum et al. 1990; Templeton and Sing 1993).

Results

The target sequence in CR2 was successfully isolated in 46 of 58 (79.3%) bones collected from HM2, and 32 polymorphic sites defining 28 haplotypes were found. A statistical parsimony network clearly showed two haplotypic groups (Fig. 2): group 1, which included 19 haplotypes from 38 chicks on Torishima and nine haplotypes from 15 archaeological bones, and group 2, which included two haplotypes from three chicks on Torishima, one haplotype from four samples on the Senkaku Islands and 18 haplotypes from 31 archaeological bones. Using a nested-clade approach, the entire haplotype genealogy was included in two 4-step clades. The target sequence in cyt *b* was successfully isolated in 42 bones (73.7%) in which all samples determining CR2 sequence were included. One polymorphic site defining two haplotypes was found, as previously described in Eda et al. (2006). Each of the 4-step clades in CR2 corresponded to the cyt *b* haplotypes without any inconsistency (Supplementary material).

For the entire data set, a null hypothesis of no association between intrinsic markers and clades between clades 4–1 and 4–2 was rejected. The carpometacarpal length and nitrogen isotope ratio ($\delta^{15}\text{N}$) were significantly different between the clades (for bone length and $\delta^{15}\text{N}$, $P < 0.05$). Within clade 4–2, the clade distances, or variances in bone length and $\delta^{15}\text{N}$, were significantly smaller than for the randomly extracted samples ($P < 0.01$). The null hypothesis was not rejected at any other clade level. The results remained unchanged if immature and age-unknown sample data were excluded, except for clade distance of bone length within clade 4–2 (not significant).

For the ancient samples, the haplotype and nucleotide diversities were, respectively, 0.857 (± 0.09 , SD) and 0.02556 (± 0.00462) for the clade 4–1 population and 0.933 (± 0.03) and 0.03964 (± 0.00304) for the clade 4–2 population (Fig. 3). For the modern samples, the haplotype and nucleotide diversities were, respectively, 0.925 (± 0.025) and 0.03180 (± 0.00202) for the haplotypes belonging to clades 4–1 and 0.667 (± 0.160) and 0.03357 (± 0.00895) for the haplotypes belonging to clades 4–2. Similar levels of nucleotide diversity were found for the ancient and modern sample, suggesting a lack of post-mortem genetic degradation causing artifactual mutations. Sequence divergence between the two haplotypic groups was 11.5%.

Discussion

Ancient population structure

CR2 and cyt *b* sequences clearly show two haplotypic clades, between which a null hypothesis of no association between intrinsic markers (nitrogen isotope ratio and bone length) and clades was rejected. These facts suggest that individuals having mtDNA belonging to the same haplotypic clade formed a different population about 1,000 years ago, and they strongly support a historical scenario explaining the modern genetic structure: the co-occupation of populations isolated for long periods of time. Judging from the current haplotype frequencies at the breeding regions, Torishima and the Senkaku Islands would have been mostly occupied by, respectively, clade 4–1 and 4–2 lineage birds. Unfortunately, there is no data that prove a relict population of albatrosses survived at the Senkaku Islands: from 1910 to the 1970s (a period almost equivalent to the maximum longevity), no chicks were observed on the Senkaku Islands, and furthermore, no individuals had been observed on or near the Islands from 1939 to 1971, despite several biological surveys during the breeding season (1939, 1950, 1952, 1963, 1970, 1971, 1979 and 1980) (Hasegawa 2003). However, the number of birds surviving on the Senkaku Islands is presumed to have been very small, even compared with Torishima where the population was reduced to around 50 birds (Hasegawa 2003; Tickell 2000; U. S. Fish and Wildlife Service 2008). Given the breeding location of the species (mid-point or tops of cliffs) and the short survey periods for each island, it is possible that this very small population was overlooked.

Genetic difference between two populations

Although no subspecies of *P. albatrus* have been proposed, the sequence divergence between the two clades (11.5%) is greater than that between recognized sister species of albatrosses analyzed for CR domain I: 1.0% for *Diomedea antipodensis* and *D. gibsoni*, 1.8% for *Thalassarche cauta* and *T. steadi*, 2.9% for *T. salvini* and *T. eremita*, 4.5% for *D. exulans* and *D. dabbenena*, 4.6% for *D. exulans* and *D. amsterdamensis*, and 7.2% for *T. melanophris* and *T. impavida* (Abbott and Double 2003; Burg and Croxall 2001, 2004; Rains et al. 2011). Since the degree of divergence between the two *P. albatrus* populations is larger than that between *T. melanophris* and *T. impavida*, the relationship between the two populations on Torishima is analogous to that observed between *T. melanophris* and *T. impavida* on Campbell Island, which includes incomplete assortative mating and outbreeding depression (Moore et al. 1997). Sympatric breeding revealed by mtDNA analysis does not

Fig. 2 The statistical parsimony network of the haplotype from archaeological (Hamanaka 2 site; $N = 46$) and living short-tailed albatrosses (Torishima, $N = 41$; Senkaku Islands, $N = 4$). The size of each circle is proportional to the number of albatrosses sharing that haplotype. Numbers in the network show haplotype name corresponding to supporting information

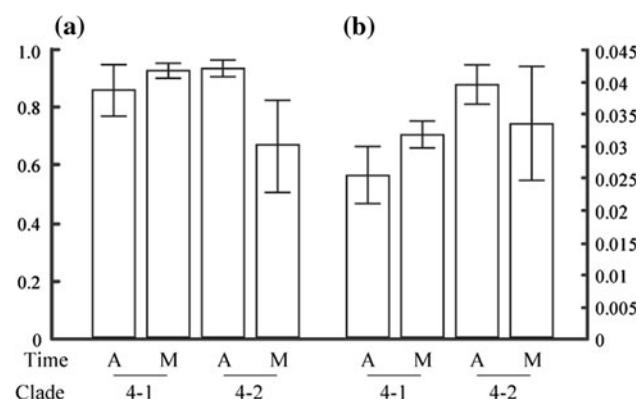
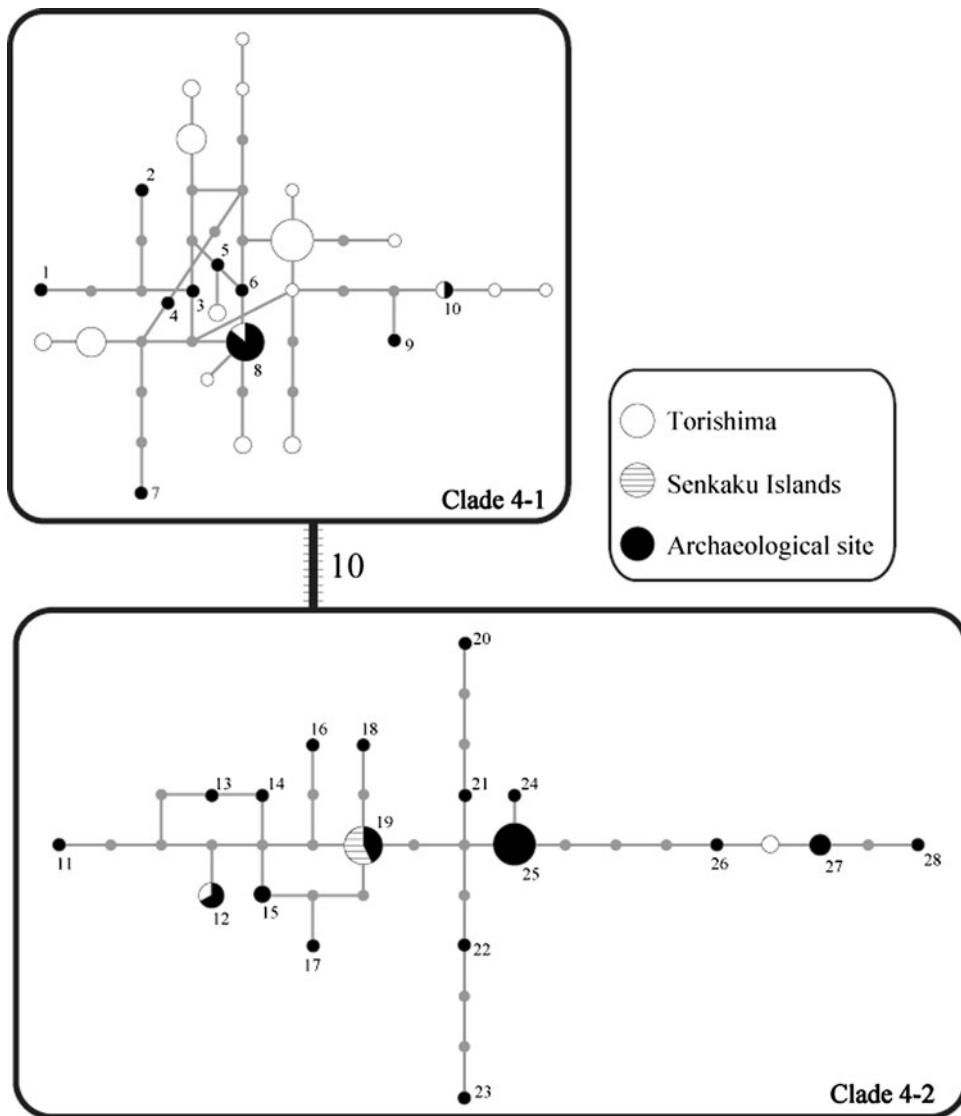


Fig. 3 Haplotype (a) and nucleotide diversity (b) for ancient (A) and modern (M) samples from clade 4-1 and 4-2 birds. Error bars show standard deviations

imply mating between individuals from the two populations. Pre-mating isolation is suggested by differences in courtship display between ringed (known to have fledged from Torishima) and un-ringed young birds (likely fledged from the Senkaku Islands) and the difference in breeding season (about 2 weeks earlier on the Senkaku Islands; H. Hasegawa, personal observation). On Torishima, one or two developmentally disabled chicks were observed each year beginning in the late 1990s (Hasegawa 2003), but they became rare in the late 2000s (H. Hasegawa, personal observation). Because the chicks were observed at a similar position in the colony (Hasegawa 2003), their disability may have been genetic in origin. Outbreeding depression rather than inbreeding depression would be a reasonable mechanism since currently the haplotype diversity on Torishima is high (Kuro-o et al. 2010) and the genetic distance between the lineages is greater than that between *T. melanophris* and *T. impavida*,

for which outbreeding depression has been suggested (Burg and Croxall 2001; Moore et al. 1997).

Difference in genetic diversities among modern and ancient populations

The ranges of haplotype diversity for the two ancient populations overlapped and were not significantly different. However, the ranges in nucleotide diversity did not overlap, indicating a significant difference between the two populations. These observations suggest that the populations had maintained stable numbers of individuals for a long time and that the origin of the clade 4–2 population predates the clade 4–1 population.

For the haplotypes belonging to clade 4–1, the haplotype diversity overlapped between the ancient and modern samples and was not significantly different. This similarity in haplotype diversity is consistent with population genetic simulations predicting a minor reduction in haplotype diversity among the birds on Torishima, given the short amount of time (relative to generation time) during which the population numbers were severely depressed (Kuro-o et al. 2010). In contrast, the haplotype diversities did not overlap between the ancient and modern samples for those haplotypes belonging to clade 4–2, suggesting a significant loss of genetic diversity in this lineage. This observation is concordant with the model in which a small number of birds survived the commercial harvest era on the Senkaku Islands. One would expect the loss of genetic diversity to have been greater than that shown in Fig. 3 since birds in that lineage do not seem to mate randomly; they breed on Torishima and the Senkaku Islands, and no evidence of emigration from Torishima to the Senkaku Islands has been observed (nearly all Torishima birds are fitted with metal leg bands as chicks since 1979, and no banded birds observed on the Senkaku Islands in 1980, 1988, 1991, 1992, 1996, 2001 and 2002; H. Hasegawa, personal observation).

Lessons for conservation strategies of the short-tailed albatross

Given our results, we conclude that two populations of the short-tailed albatross existed about 1,000 years ago, and that the level of genetic divergence exceeded that between sister species of other albatrosses. The exact point at which the two lineages began to co-distribute on Torishima is unknown; however, after the early 20th century (i.e. after massive population decline of the species) is a reasonable assumption. In the wandering albatross, the proportion of juveniles recruited to the natal population is suggested to be positively related to population size (Inchausti and Weimerskirch 2002). If this trend is true for the short-tailed albatross, the dispersal rate from the Senkaku Islands and

emigration rate to Torishima is expected to have increased during the early 20th century when few birds survived on the Senkaku Islands. Although during the late 19th and early 20th centuries the species was exterminated at most breeding colonies, descendants of each population survived at both locations. The species is considered to be vulnerable and a management unit. Biparental molecular marker analyses (e.g. microsatellite DNA) of the birds on Torishima are required to assess whether the two populations are mixing. In 2008, to create a third stable breeding station, the translocation of chicks from Torishima 350 km south to Mukojima of the Bonin Islands was begun as an international conservation project (U. S. Fish and Wildlife Service 2008). To exclude the risk of artificial population admixture and outbreeding depression, genetic monitoring of the chicks on Torishima using mtDNA or microsatellite DNA and careful selection of the individuals to be translocated are required. Furthermore, genetic diversity among the clade 4–2 lineage birds was drastically decreased through exploitation during the late 19th and early 20th centuries. Careful and extensive study of the Senkaku Islands birds, which has not been conducted since the 2001–2002 breeding season due to territorial disputes among the neighboring countries, is therefore necessary.

Future perspectives of the ancient population structure and conservation biology

The approach used in our study proved useful for reconstructing the ancient population structure of the short-tailed albatross. This approach may be useful for inferring the ancient population structures of highly mobile animals captured within or outside their breeding area, and in classifying recently mixed or sympatrically diverged populations. It is possible to apply this approach to other species for which there is no information on breeding-site location. For example, we can infer population structure from archaeological remains or the fossils of animals that were traded, such as large herbivores, and of animals that are highly mobile, like migratory birds, whales, and fishes. The reconstructed ancient population structure would be useful for conservation and management recommendations. In addition, when more than one population is detected, the temporal change of the genetic diversity can be compared among populations. The information will be useful to assign conservation priority to each population. In this study, we analysed bone measurements and stable nitrogen/carbon isotope ratios as intrinsic traits that were independent from neutral molecular marker. In addition to these traits, stable isotopes of hydrogen (δD ; Hobson and Wassenaar 1997; Chamberlain et al. 1997; Lott et al. 2003; Clegg et al. 2003), strontium ($\delta^{87}\text{Sr}$; Chamberlain et al. 1997), and sulphur ($\delta^{34}\text{S}$; Lott et al. 2003) have been used

to reveal the connectivity of breeding and non-breeding areas in migratory animals (reviewed in Hobson 1999; Webster et al. 2002). In particular, analyses of stable isotope ratios of hydrogen, which display an approximate latitudinal gradient (Hobson and Wassenaar 1997), and those of strontium, which correspond with those found in regional soil parent materials (Chamberlain et al. 1997), may improve the power to reconstruct ancient population structures in other animals. The most important lesson from zooarchaeological short-tailed albatross bones would be that the ancient population structure can be reconstructed from zooarchaeological remains and that the reconstructed population structure is useful for conservation and management recommendations.

Acknowledgements We thank Dr. Ushio Maeda for use of the albatross bones from HM2. We also thank Drs. Go Fujita, Tadashi Miyashita, Gregory Balogh, Haruki Tatsuta, Yoko Kunitake, Masaki Fujita, and Tatsuya Amano for their insightful comments on the manuscript and Dr. Van't Hof and Jiro Eda for improving our English. Helpful comments from two anonymous reviewers clarified the strengths and weaknesses of this study. This study was supported financially in part by a Grant-in-Aid for JSPS Fellows to ME from the Japan Society for the Promotion of Science (No. 16-6316).

References

- Abbott CL, Double MC (2003) Phylogeography of shy and white-capped albatrosses inferred from mitochondrial DNA sequences: implications for population history and taxonomy. *Mol Ecol* 12:2747–2758. doi:10.1046/j.1365-294X.2003.01944.x
- Avise JC (2000) Phylogeography: the history and formation of species. Harvard University Press, Cambridge
- Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC (1987) Intraspecific phylogeography—the mitochondrial-DNA bridge between population-genetics and systematics. *Annu Rev Ecol Syst* 18:489–522
- BirdLife International (2008) *Phoebastria albatrus*. www.iucnredlist.org. Accessed 22nd July 2009
- Burg TM, Croxall JP (2001) Global relationships amongst black-browed and grey-headed albatrosses: analysis of population structure using mitochondrial DNA and microsatellites. *Mol Ecol* 10:2647–2660
- Burg TM, Croxall JP (2004) Global population structure and taxonomy of the wandering albatross species complex. *Mol Ecol* 13:2345–2355. doi:10.1111/j.1365-294X.2004.02232.x
- Chamberlain CP, Blum JD, Holmes RT, Feng XH, Sherry TW, Graves GR (1997) The use of isotope tracers for identifying populations of migratory birds. *Oecologia* 109:132–141
- Clegg SM, Kelly JF, Kimura M, Smith TB (2003) Combining genetic markers and stable isotopes to reveal population connectivity and migration patterns in a Neotropical migrant, Wilson's warbler (*Wilsonia pusilla*). *Mol Ecol* 12:819–830
- Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Mol Ecol* 9:1657–1659
- Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. *Trends Ecol Evol* 15:290–295
- DeNiro MJ (1985) Postmortem preservation and alteration of in vivo bone collagen isotope ratios in relation to palaeodietary reconstruction. *Science* 317:806–809
- Deniro MJ, Epstein S (1978) Influence of diet on distribution of carbon isotopes in animals. *Geochim Cosmochim Acta* 42:495–506
- Deniro MJ, Epstein S (1981) Influence of diet on the distribution of nitrogen isotopes in animals. *Geochim Cosmochim Acta* 45: 341–351
- Development Core Team R (2009) R: a language and environment for statistical computing., 2.10.1 edn. R Foundation for Statistical Computing, Vienna
- Driesch A (1976) A guide to the measurements of animal bones from archaeological sites. *Peabody Museum Bull* 1:1–136
- Eda M, Koike H, Sato F, Higuchi H (2005) Why were so many albatross remains found in northern Japan? In: Grupe G, Peters J (eds) Feathers, grit and symbolism: birds and humans in the old and new worlds. Verlag Marie Leidorf GmbH, Westfalen, pp 131–140
- Eda M, Baba Y, Koike H, Higuchi H (2006) Do temporal size differences influence species identification of archaeological albatross remains when using modern reference samples? *J Archaeol Sci* 33:349–359. doi:10.1016/j.jas.2005.07.017
- Eda M, Kuro-o M, Higuchi H, Hasegawa H, Koike H (2010) Mosaic gene conversion after a tandem duplication of mtDNA sequence in Diomedaeidae (albatrosses). *Genes Genet Syst* 85:129–139
- Excoffier L, Schneider S (1999) Why hunter-gatherer populations do not show signs of Pleistocene demographic expansions. *Proc Natl Acad Sci USA* 96:10597–10602
- Fox CW, Roff DA, Fairbairn DJ (2001) Evolutionary ecology. Oxford University Press, Oxford
- Futuyma DJ (1986) Evolutionary Biology. Sinauer Association Inc., Sunderland
- Hadly EA, Kohn MH, Leonard JA, Wayne RK (1998) A genetic record of population isolation in pocket gophers during Holocene climatic change. *Proc Natl Acad Sci USA* 95:6893–6896
- Hardy C, Callou C, Vigne JD, Casane D, Dennebouy N, Mounolou JC, Monnerot M (1995) Rabbit mitochondrial DNA diversity from prehistoric to modern times. *J Mol Evol* 40:227–237
- Hartl DL, Clark AG (1997) Principles of population genetics. Sinauer Association Inc., Sunderland
- Hasegawa H (2003) From fifty to five thousands: for the restoration of the short-tailed albatross. Doubutsu-sha, Tokyo (in Japanese)
- Hobson KA (1999) Tracing origins and migration of wildlife using stable isotopes: a review. *Oecologia* 120:314–326
- Hobson KA, Wassenaar LI (1997) Linking breeding and wintering grounds of Neotropical migrant songbirds using stable hydrogen isotopic analysis of feathers. *Oecologia* 109:142–148
- Hofreiter M, Jaenicke V, Serre D, von Haeseler A, Paabo S (2001) DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. *Nucleic Acids Res* 29:4793–4799
- Inchausti P, Weimerskirch H (2002) Dispersal and metapopulation dynamics of an oceanic seabird, the wandering albatross, and its consequences for its response to long-line fisheries. *J Animal Ecol* 71:765–770
- Jaenicke-Despres V, Buckler ES, Smith BD, Gilbert MTP, Cooper A, Doebley J, Paabo S (2003) Early allelic selection in maize as revealed by ancient DNA. *Science* 302:1206–1208
- Klicka J, Zink RM, Barlow JC, McGillivray WB, Doyle TJ (1999) Evidence supporting the recent origin and species status of the Timberline Sparrow. *Condor* 101:577–588
- Knowles LL, Maddison WP (2002) Statistical phylogeography. *Mol Ecol* 11:2623–2635
- Kuro-o M, Yonekawa H, Saito S, Eda M, Higuchi H, Koike H, Hasegawa H (2010) Unexpectedly high genetic diversity of mtDNA control region through severe bottleneck in vulnerable albatross *Phoebastria albatrus*. *Conserv Genet* 11:127–137
- Lawrence HA, Scofield RP, Crockett DE, Millar CD, Lambert DM (2008) Ancient genetic variation in one of the world's rarest seabirds. *Heredity* 101:543–547. doi:10.1038/hdy.2008.85

- Leonard JA (2008) Ancient DNA applications for wildlife conservation. *Mol Ecol* 17:4186–4196. doi:[10.1111/j.1365-294X.2008.03891.x](https://doi.org/10.1111/j.1365-294X.2008.03891.x)
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451–1452. doi:[10.1093/bioinformatics/btp187](https://doi.org/10.1093/bioinformatics/btp187)
- Lott CA, Meehan TD, Heath JA (2003) Estimating the latitudinal origins of migratory birds using hydrogen and sulfur stable isotopes in feathers: influence of marine prey base. *Oecologia* 134:505–510. doi:[10.1007/s00442-002-1153-8](https://doi.org/10.1007/s00442-002-1153-8)
- Maeda U, Yamaura K (1992) Excavation research of Hamanaka 2 site. Maeda Insatsu, Tsukuba (in Japanese)
- Mihara S, Okuno M, Ogawa H, Tanaka K, Nakamura T, Koike H (2002) AMS 14-C dating and dietary analysis for Lal-lo shell midden sites, Philippines. *Summ Res Using AMS Nagoya Univ* 13:82–104 (in Japanese)
- Moore PJ, Taylor GA, Amey JM (1997) Interbreeding of black-browed albatross *Diomedea m. melanophris* and New Zealand black-browed albatross *D. m. impavida* on Campbell Island. *Emu* 97:322–324
- Moritz C (1994) Defining evolutionarily significant units for conservation. *Trends Ecol Evol* 9:373–375
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959
- Prum B, Guilloultbataille M, Clergetdarpoux F (1990) On the use of chi-2 tests for nested categorized data. *Ann Hum Genet* 54: 315–320
- Rains D, Weimerskirch H, Burg TM (2011) Piecing together the global population puzzle of wandering albatrosses: genetic analysis of the Amsterdam albatross *Diomedea amsterdamensis*. *J Avian Biol* 42:69–79. doi:[10.1111/j.1600-048X.2010.05295.x](https://doi.org/10.1111/j.1600-048X.2010.05295.x)
- Ramakrishnan U, Hadly EA (2009) Using phylochronology to reveal cryptic population histories: review and synthesis of 29 ancient DNA studies. *Mol Ecol* 18:1310–1330. doi:[10.1111/j.1365-294X.2009.04092.x](https://doi.org/10.1111/j.1365-294X.2009.04092.x)
- Rasmussen PC (1994) Geographic-variation in morphology and allozymes of south-american imperial shags. *Auk* 111(1): 143–161
- Rosenberg NA, Pritchard JK, Weber JL, Cann HM, Kidd KK, Zhivotovsky LA, Feldman MW (2002) Genetic structure of human populations. *Science* 298:2381–2385
- Shepherd LD, Millar CD, Ballard G, Ainley DG, Wilson PR, Haynes GD, Baroni C, Lambert DM (2005) Microevolution and mega-icebergs in the Antarctic. *Proc Natl Acad Sci USA* 102:16717–16722. doi:[10.1073/pnas.0502281102](https://doi.org/10.1073/pnas.0502281102)
- Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596–1599. doi:[10.1093/molbev/msm092](https://doi.org/10.1093/molbev/msm092)
- Templeton AR (1993) The “Eve” hypothesis: a genetic critique and reanalysis. *Am Anthropol* 95:51–72
- Templeton A, Georgiadis RNJ (1995) A landscape approach to conservation genetics: conserving evolutionary processes in the African Bovidae. In: Avise JC, Hamrick JL (eds) *Conservation genetics*. Chapman & Hall, New York, pp 398–430
- Templeton AR, Sing CF (1993) A cladistic-analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. IV. nested analyses with cladogram uncertainty and recombination. *Genetics* 134:659–669
- Templeton AR, Sing CF, Kessling A, Humphries S (1988) A cladistic analysis of phenotype associations with haplotypes inferred from restriction endonuclease mapping. II. The analysis of natural populations. *Genetics* 120:1145–1154
- Templeton AR, Routman E, Phillips CA (1995) Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics* 140: 767–782
- Tickell WLN (2000) *Albatrosses*. Yale University Press, New Haven
- U. S. Fish and Wildlife Service (2008) Short-tailed albatross recovery plan.1-105
- Webster MS, Marra PP, Haig SM, Bensch S, Holmes RT (2002) Links between worlds: unraveling migratory connectivity. *Trends Ecol Evol* 17:76–83

Seals Quickly Respond To Gain And Loss Of Habitat Under Climate Change

Date: July 12, 2009

Source: Durham University

Southern Elephant seals responded rapidly to climate and habitat change and established a new breeding site thousands of kilometres from existing breeding grounds, according to new research.

An international research team, including post-doctorate Dr Mark de Bruyn and collaborators from the US, South Africa and Italy, led by Professor Rus Hoelzel from the School of Biological and Biomedical Sciences, Durham University, found that when the Antarctic ice sheets of the Ross Sea Embayment retreated in the Holocene period 8,000 years ago, elephant seals, *Mirounga leonina*, adopted the emergent habitat and established a new population which flourished.



Elephant seal. In the past, southern elephant seals responded rapidly to climate and habitat change and established a new breeding site thousands of kilometres from existing breeding grounds.

Credit: iStockphoto/Alexander Hafemann

DNA sequences from the ancient remains of seals from the now extinct Antarctic colony showed high levels of genetic diversity, probably due to the very large population size sustained there. The study, published in the academic journal *PLoS Genetics*, and funded by the US National Science Foundation, shows how environmental change can drive the demographic and evolutionary processes that determine diversity within and among species.

Tracking these processes during periods of change reveals mechanisms for the establishment of populations, and provides predictive data on response to potential future impacts, including those caused by climate change.

Professor Rus Hoelzel said: "In general, this approach of looking to the past to understand what might happen in the future, has good potential for predicting the impact of environmental change in both marine and terrestrial systems.

"We've shown how a highly mobile marine species responded to the gain and loss of new breeding habitat. The new habitat was quickly adopted, probably because seals migrate annually into Antarctic waters to feed. However, when the ice returned and the habitat was lost, only a small proportion returned to the original source population. The Antarctic population crashed and much diversity was lost."

This habitat was released after the retreat of the grounded ice sheet in the Ross Sea Embayment 7,500-8,000 years ago, and is within the range of modern foraging excursions from the Macquarie Island colony. Using ancient mtDNA and evolutionary models, the research team tracked the population dynamics of the now extinct colony and the connectivity between this and modern breeding sites.

The team found clear signs of rapid expansion in the new colony 8,000 years ago. This was followed by directional migration away, coupled with a loss of diversity 1,000 years ago, when the sea ice is thought to have expanded. The data suggest that the new colony seals came initially from Macquarie Island, and that some returned there, but in much smaller numbers, when the new colony habitat was lost 7,000 years later.

In order to understand how biodiversity is generated and maintained over time, the team has set out to understand the process by which the seal populations formed and diverged. They analysed data from ancient DNA to show how elephant seal populations responded when new breeding habitat was gained and then lost over the course of approximately 7,000 years.

Professor Rus Hoelzel said: "Using ancient DNA, we were able to track the dynamics and diversity of a population from its foundation, through to its extinction, in the context of Holocene climate change. We learned that new habitat emerging within the species' migratory range could be quickly taken advantage of, but that the reverse was not true. The movement patterns of seals from this Antarctic breeding site would have been unlikely to take them near other potential breeding sites, and so when their breeding site was lost, their numbers crashed.

"The seals that discovered the new breeding site had things good, because food was abundant and nearby, however when the ice returned, the new colony collapsed and only a few seals made it back to their original home.

"This illustrates the importance of understanding the behaviour and life history of a species, in order to model how it may be able to respond to rapid change."

The key factors in the expansion of the new colony were likely to be the abundant local food resource and extensive physical habitat that allowed rapid expansion after the initial founder event, and a tendency for females to return to annual breeding sites in this species.

When the ice expanded again 1,000 years ago, the seals returned to their origins but in much smaller numbers.

Story Source:

The above story is based on materials provided by **Durham University**. *Note: Materials may be edited for content and length.*

null

Cite This Page:

MLA **APA** **Chicago**

Durham University. "Seals Quickly Respond To Gain And Loss Of Habitat Under Climate Change." ScienceDaily. ScienceDaily, 12 July 2009. <www.sciencedaily.com/releases/2009/07/090709201849.htm>.

Rapid Response of a Marine Mammal Species to Holocene Climate and Habitat Change

Mark de Bruyn¹✉*, Brenda L. Hall², Lucas F. Chauke³, Carlo Baroni⁴, Paul L. Koch⁵, A. Rus Hoelzel¹*

1 School of Biological Sciences, University of Durham, Durham, United Kingdom, **2** Climate Change Institute and Department of Earth Sciences, University of Maine, Orono, Maine, United States of America, **3** Mammal Research Institute, Department of Zoology and Entomology, University of Pretoria, Pretoria, South Africa, **4** Dipartimento di Scienze della Terra, Università di Pisa, Pisa, Italy, **5** Department of Earth and Planetary Sciences, University of California, Santa Cruz, California, United States of America

Abstract

Environmental change drives demographic and evolutionary processes that determine diversity within and among species. Tracking these processes during periods of change reveals mechanisms for the establishment of populations and provides predictive data on response to potential future impacts, including those caused by anthropogenic climate change. Here we show how a highly mobile marine species responded to the gain and loss of new breeding habitat. Southern elephant seal, *Mirounga leonina*, remains were found along the Victoria Land Coast (VLC) in the Ross Sea, Antarctica, 2,500 km from the nearest extant breeding site on Macquarie Island (MQ). This habitat was released after retreat of the grounded ice sheet in the Ross Sea Embayment 7,500–8,000 cal YBP, and is within the range of modern foraging excursions from the MQ colony. Using ancient mtDNA and coalescent models, we tracked the population dynamics of the now extinct VLC colony and the connectivity between this and extant breeding sites. We found a clear expansion signal in the VLC population ~8,000 YBP, followed by directional migration away from VLC and the loss of diversity at ~1,000 YBP, when sea ice is thought to have expanded. Our data suggest that VLC seals came initially from MQ and that some returned there once the VLC habitat was lost, ~7,000 years later. We track the founder-extinction dynamics of a population from inception to extinction in the context of Holocene climate change and present evidence that an unexpectedly diverse, differentiated breeding population was founded from a distant source population soon after habitat became available.

Citation: de Bruyn M, Hall BL, Chauke LF, Baroni C, Koch PL, et al. (2009) Rapid Response of a Marine Mammal Species to Holocene Climate and Habitat Change. PLoS Genet 5(7): e1000554. doi:10.1371/journal.pgen.1000554

Editor: Molly Przeworski, University of Chicago, United States of America

Received January 12, 2009; **Accepted** June 9, 2009; **Published** July 10, 2009

Copyright: © 2009 de Bruyn et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This research was funded by the Office of Polar Programs of the National Science Foundation and with the support of the Italian National Program on Antarctic Research (PNRA). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: markus.debruyn@gmail.com (MDB); a.r.hoelzel@dur.ac.uk (ARH)

✉ Current address: School of Biological Sciences, Bangor University, Bangor, United Kingdom

Introduction

Populations can respond to changing habitats by adapting (through natural selection or phenotypic plasticity), moving (to avoid habitat of reduced suitability, or take advantage of emerging habitat), by adjusting population size, or some combination of the above. Both natural selection and genetic drift can shape populations as they evolve in this context. It is unusual for there to be an opportunity to track population dynamics and genetics during extended periods of environmental change, but a now extinct population of elephant seals from the Victoria Land Coast (VLC) in the Ross Sea, Antarctica provides such an opportunity. Here we investigate the demographic and population genetic consequences of emerging breeding habitat suitable to this species in an environment rich in resources, and the impact of the eventual loss of that habitat.

Open beach along the VLC was probably released around 8,000–7,500 YBP, based on the grounding-line retreat of the West Antarctic ice sheet [1]. More recently, there is evidence that ice encroachment has taken place over the last 1,000 years at VLC, based on indications that Holocene raised beaches in several locations have been overrun by glacial expansion [2–4], and that these same beaches show evidence of being formed under wave-dominated conditions characterized by reduced sea ice [5]. Today

the region is mostly enclosed by year round land-fast sea ice, and is therefore unsuitable for southern elephant seal breeding. Elephant seals are not found along the VLC today.

Southern elephant seals have a circumpolar distribution, breeding on sub-Antarctic (north of the Antarctic convergence) and Antarctic (south of the Antarctic convergence) islands (Figure 1). They are annual breeders, showing fidelity to a set of established breeding colonies, and spend the non-breeding seasons foraging over large distances at sea, and hauled out on beaches for moulting [6]. The same beaches used for breeding are often used for moulting, but other locations may also be used, especially by adult males, sometimes including the Antarctic mainland [6]. Temperate continental breeding habitat historically included South Africa and Tasmania, but is now restricted to Peninsula Valdés, Argentina [6]. Small colonies are found on open beaches at Palmer Station and in the Windmill Islands (probably originating from the Macquarie and Heard Island colonies [7]) close to the Antarctic continent, but the presence of southern elephant seals along the Antarctic coastline in modern times is primarily on foraging excursions that originate from sub-and/or Antarctic breeding colonies or haul-out sites [8]. There are no examples of elephant seal breeding colonies where it is necessary to cross ice to reach the site from open water [6], possibly because the



Author Summary

In order to understand how biodiversity is generated and maintained over time, we need to understand the process by which populations form and diverge. Natural variation within species is typically partitioned among populations, which sometimes forms the basis for speciation events. One mechanism for the establishment of novel variation at the population level is through a response to emerging habitat. Here we use data from ancient DNA to show how elephant seal populations responded when new breeding habitat was gained and then lost over the course of approximately 7,000 years. We show that the seals quickly took advantage of newly available breeding habitat, far from the nearest extant breeding site, and that a highly diverse and genetically differentiated population was established over a matter of generations. The key factors were likely the abundant local food resource and extensive physical habitat that allowed rapid expansion after the initial founder event and a tendency for females to return to annual breeding sites in this species. Tracking the founder-extinction dynamics of historical populations provides insight into the likely implications of future environmental change. This is an important tool in our efforts to mitigate the impact of human-induced change.

energetic expense would be too high for competing, polygynous seals that fast during the breeding period.

A large number of exceptionally well-preserved skin samples and mummified remains, including pups, were collected from raised Holocene beaches along the VLC, adjacent to the Ross Sea (Figure 1), from a period when the beaches were exposed to open water [1,9]. The vast quantity of seal remains and their extensive geographical coverage (Table S1) suggests that the VLC was not merely a moulting site, though breeding could have been a minor or temporary aspect of site usage. To investigate this further, and to understand the evolutionary history of this putative population, we used strict ancient DNA (aDNA) protocols [10] to amplify and sequence 325 bp of the most variable part of the mitochondrial DNA (mtDNA) control region. Direct radiocarbon dating was used to date 223 of these samples (Table S1).

We utilise these data to track the founder-extinction dynamics of an ephemeral population from inception to extinction in the context of Holocene climate change. We test the hypothesis that the release of new habitat in a highly productive environment can lead to the rapid founding and expansion of a new population, and investigate the consequences with respect to genetic diversity and gene flow among populations. We further investigate the consequences of the loss of this habitat on genetic diversity and connectivity.

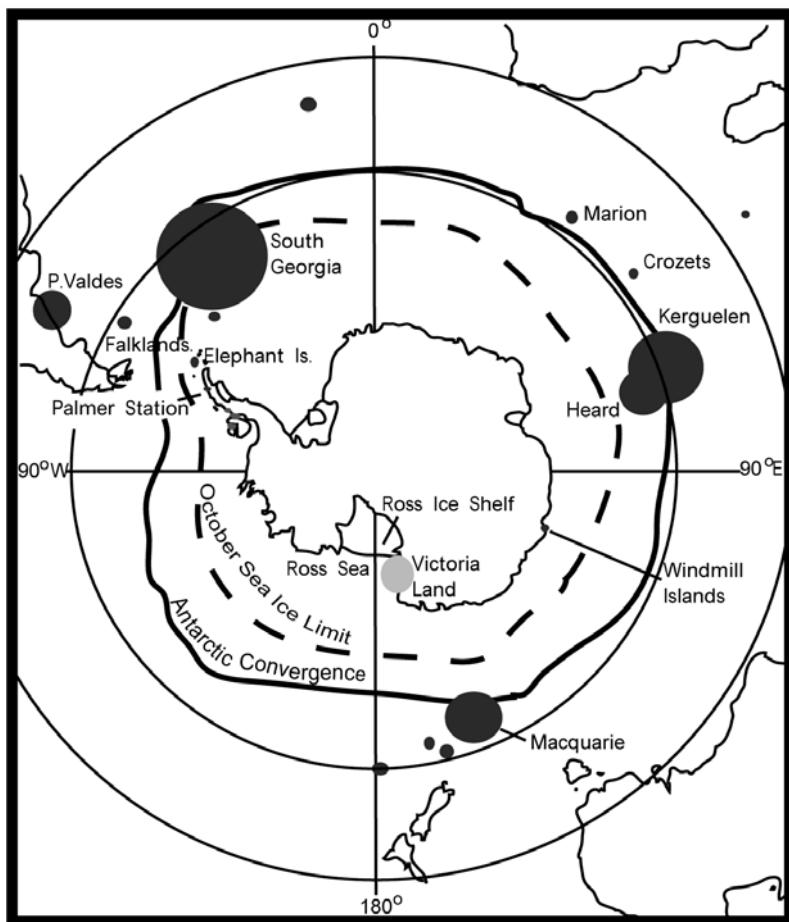


Figure 1. Map of study region showing major and minor southern elephant seal (*Mirounga leonina*) breeding colonies. With the exception of VLC (shown in gray), size of the circles indicates relative colony size.
doi:10.1371/journal.pgen.1000554.g001

Results

Ancient DNA extractions were successful for approximately 90% of samples trialled. From 223 successfully sequenced ancient VLC sample extractions, 177 haplotypes were identified, whereas from 49 MQ samples, 16 haplotypes were identified (Table 1; GenBank accession numbers: FJ168073–FJ168343). A single haplotype was shared between VLC and MQ. Summary details for all other extant colonies are presented elsewhere [11,12]. To determine relationships between the VLC seals and extant seal colonies, we used various phylogeny reconstruction methods [13,14]. While a large degree of reticulation was evident for the VLC samples (Figure 2), all methods unambiguously defined two major lineages; the ancient VLC samples together with extant samples from MQ, and a lineage including all other extant colonies (Figure 2; only median-joining network reconstructions are shown). Bayesian analyses of the ancient VLC data, incorporating an explicit post-mortem damage model [15], calibrated against the estimated calendar age of the ancient samples (Table S1) produced a mtDNA hypervariable region (HVR) rate estimate of 9.80×10^{-7} (Figure S1; 95% highest posterior density interval (HPDI) 1.67×10^{-9} – 2.06×10^{-6}) substitutions per site per year (s.s.yr $^{-1}$). This rate is in agreement with estimates obtained from various other aDNA datasets [16] including a close match with an HVR rate of 9.6×10^{-7} s.s.yr $^{-1}$ calculated over a similar timeframe (6,424 YBP) from Adelie penguin aDNA [17], and with the mean human HVR pedigree rate estimate derived from a meta-analysis (9.5×10^{-7} s.s.yr $^{-1}$) [18]. Although HPDIs were broad, the sampling distribution returned a well-resolved peak with strong bounds (Figure S1). After [19–21], uncertainty in the actual substitution rate estimate was not incorporated into further analyses, in order to isolate uncertainty in the coalescent process. In the context of the well-documented geochronology of the VLC [1–5,9,22], our results applied in further analyses (e.g. IM; mismatch distributions – see below) provide strong support for the HPD rate estimate, and for the similar molecular rates calculated from other species over Holocene timeframes and from pedigree data [16–18,23,24]. Although controversial [25,26], as they are an order of magnitude or more faster than ‘traditional’ substitution rate estimates derived from interspecific phylogenetic datasets or the fossil record [24], these high intraspecific estimates have been quite consistent and are now well supported [16–18,23,24,26].

Founder analysis identified three closely related haplotypes within the MQ/VLC clade as potential founders (Figure 2B), based on minimum ρ value estimates for that clade. Two of these were sampled from the VLC site, while one was sampled from MQ. Coalescent theory predicts that ancestral haplotypes will be basal and/or central within the haplogroup [27,28], as is the case for these haplotypes. Mapping the estimated calendar ages of the VLC haplotypes onto the MQ/VLC network supported these results, as the potential MQ founder haplotype’s nearest-neighbours date to early in the VLC colonies’ history, as do the potential founder haplotypes sampled from the VLC (Figure 2C). In addition, minimum inter-population ρ value estimates identify MQ as the most likely source of the VLC samples (Table 2).

Coalescent estimates [29] of the time of splitting for MQ and VLC ($t = 6,167$ YBP; 90% HPDI = 5,116–7,674 YBP; Figure S2; Table S2) corresponds closely with the mid-Holocene retreat of the Ross Sea ice sheet from the VLC (7,500–8,000 YBP) [1], and thus the consequent opening of newly available seal habitat, and coincides with the age of the oldest ancient sample (7,087 YBP; Table S1). On this basis, colonisation of the VLC likely began prior to 7,000 calendar YBP.

Bayesian demographic model selection and population genetic and coalescent-based statistics provide a signature of subsequent expansion in the VLC population, while MQ shows a signature of a long-term stable population (Table 1). Similarly, mismatch distributions [30] show a strong signal for expansion in the VLC population (with $\tau = 5.48$; mean expansion time estimate = 8,600 YBP (95% CI = 7,185–10,298 YBP)), while the mismatch distribution for MQ is multi-modal and therefore consistent with long-term stability (Figure 3). FLUCTUATE [31] analyses indicated strong growth in VLC samples from $\geq 3,000$ YBP ($g = 648 \pm 44$ sd), while 1,500–3,000 year old samples ($g = 254 \pm 33$) and the most recent samples ($< 1,500$ years old; $g = 119 \pm 14$) showed lower growth rates.

Recent back migration to MQ can be estimated by examining for the presence of MQ lineages that are most likely to have evolved at VLC (Figure 2), that is, MQ haplotypes that are derived from VLC haplotypes [27]. It is clear from the network that this is the case for a number of MQ haplotypes, and again mapping the ages of their nearest VLC neighbours suggests increasing recent (range of nearest-neighbour dates: 506–1,565 YBP) contact from VLC to MQ (Figure 2C). Consistent with this notion of some MQ haplotypes being derived from VLC haplotypes, ϕ_{ST} values appear to decrease over time between VLC and MQ ($> 3,000$ YBP: $\phi_{ST} = 0.208$; 1,500–3,000 YBP: $\phi_{ST} = 0.161$; $< 1,500$ YBP: $\phi_{ST} = 0.149$; *c.f.* Table S3). Comparing VLC to other modern populations, ϕ_{ST} showed the opposite trend ($> 3,000$ YBP: $\phi_{ST} = 0.458$; 1,500–3,000 YBP: $\phi_{ST} = 0.539$; $< 1,500$ YBP: $\phi_{ST} = 0.560$). One possible interpretation would be increasing contact between VLC and MQ. These significant ϕ_{ST} values also show that VLC must have become an independent breeding population, and was not merely or predominantly a geographically distant moulting haul-out site. Results using an isolation with migration model [29] provide additional support for post-founder unidirectional migration from VLC to MQ only (m_1 (into Macquarie) = 0.45, 90% HPDI = 0.07–1.79; m_2 (into VLC) = 0.005, 90% HPDI = 0.005–0.115), and an estimated time of migration of 1,365 YBP (90% HPDI = 565–4,325 YBP) (Figure S2, Table S2). Based on the age of the youngest seal remains recovered, final abandonment of the VLC site likely took place around 400 YBP (Table S1).

Bayesian skyline plots [32], which simultaneously estimate both the timing and magnitude of effective (female) population size change (\mathcal{N}_e), provided more equivocal results. There is only a slight indication of expansion when VLC is considered on its own (Figure 4A), and the 95% HPDIs are broad and overlapping. Considering MQ on its own suggests a stable population at an estimated \mathcal{N}_e that is substantially lower than for VLC (Figure 4B). We then combined VLC with extant populations, first with Marion Island where the census population size is known to be similar to that on MQ [6,33,34], and for which there is little evidence of connectivity with either MQ or VLC ([11]; see above). In this case there is a stronger indication of expansion, but no indication of a population decline (Figure 4C). A very similar result was found when we combined VLC and Elephant Island (BSP not presented), which also has a similar or smaller census size to MQ [6,35]. Finally, we combine VLC with MQ (Figure 4D). Here we see a trace that indicates both population expansion and decline, at dates roughly consistent with our previous estimates for these events.

We then assessed levels of diversity from two time periods for the VLC, and for the modern MQ, though our VLC samples from the relatively recent past come mostly from the beginning of the apparent decline period (VLC samples $> 1,000$ YBP: $\pi = 0.0243 \pm 0.0001$, $r = 44.7$, $n = 163$; VLC $< 1,000$ YBP:

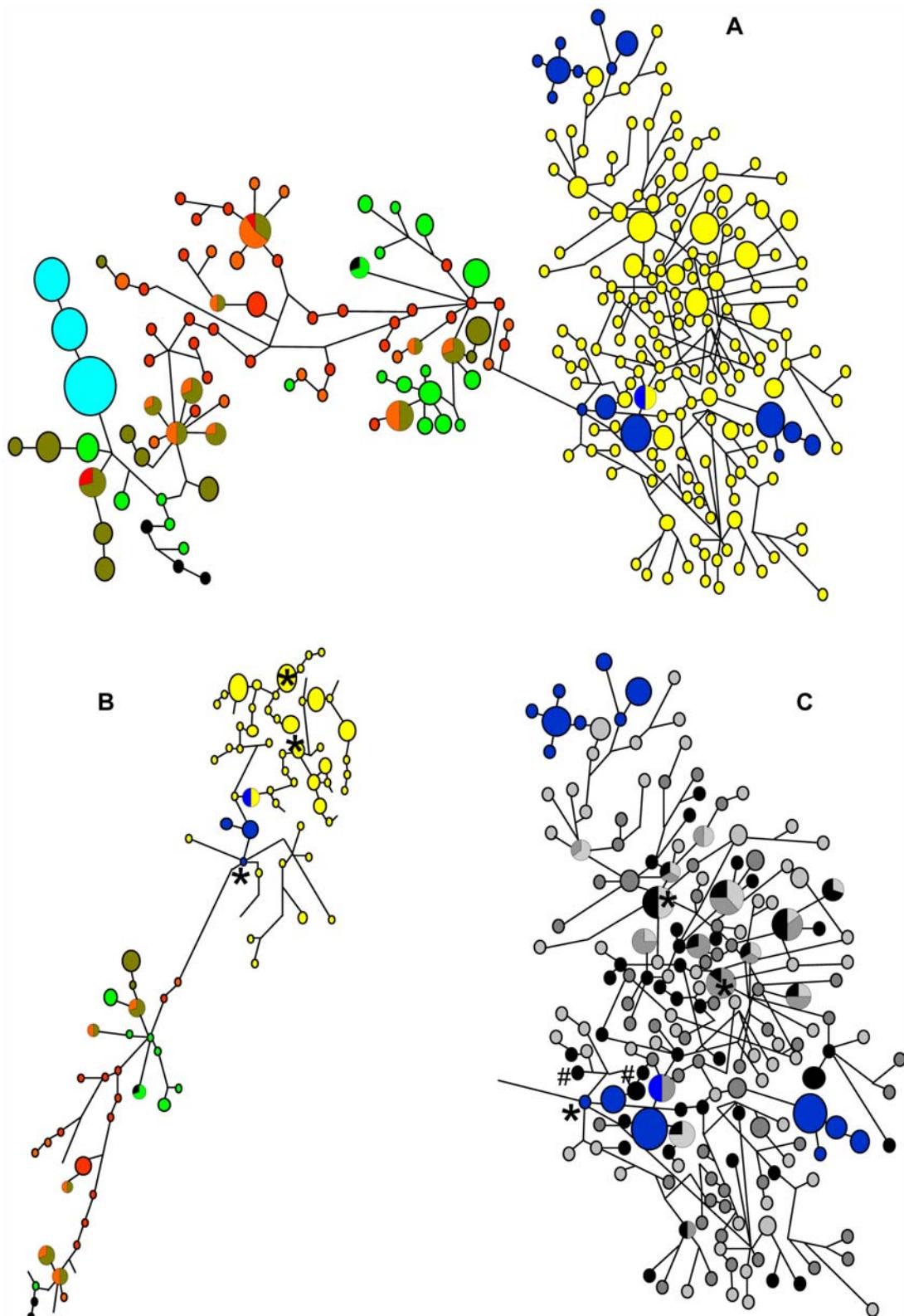


Figure 2. Networks of phylogenetic relationships among ancient and extant mitochondrial DNA haplotypes. (A) Median-joining network where the size of the circle indicates relative frequency of the haplotype. Yellow – Victoria Land Coast (VLC); dark blue – Macquarie Island (MQ); brown – Falklands; orange – Elephant Island; green – Marion Island; light blue – Peninsula Valdés; red – South Georgia; black – Heard Island. South Georgia and Heard Island relationships are based on 240 bp of mtDNA (see Materials and Methods). Networks were post-processed to reduce complexity using the Maximum Parsimony option in Network 4.5 [13]. (B) Network showing torso only, illustrating major connections among basal haplotypes. Image has been rotated to fit the page. Samples coloured as in part 'A'. Likely VLC founder haplotypes, based on rho (ρ) estimates, indicated by **. (C) MQ and VLC clade only, illustrating relationship between MQ haplotypes and the age of the VLC samples. Blue: MQ; light gray:

VLC<1500 YBP; dark gray: VLC 1501–3000; black: VLC 3000+. Likely VLC founder haplotypes, based on rho (ρ) estimates, indicated by **. Nearest VLC neighbours of the proposed VLC-founder haplotype (sampled from MQ) indicated by #. doi:10.1371/journal.pgen.1000554.g002

$\pi = 0.0224 \pm 0.0012$, $r = 41.7$, $n = 60$; MQ: $\pi = 0.0202 \pm 0.0013$, $r = 15.0$, $n = 49$; r = haplotype richness, corrected for sample size). Only the VLC>1,000 YBP and MQ samples had non-overlapping distributions for π according to the analysis in Dnasp (see Materials and Methods). The haplotype network (Figure 2) and relative haplotype richness values for VLC and MQ illustrate the large difference in diversity, which may in part be due to diversity being lost during 19th century sealing at MQ. However, there is no signal for a population bottleneck and subsequent expansion at MQ (Figure 3), and the level of diversity there is comparable to that seen at other islands [11].

Estimates of N_e derived using IM indicate a large effective population size at VLC (219,255; 90% HPDI = 187,888–267,582), and much lower estimates for MQ and the putative ancestor population (MQ: 6,727 (2,566–11,969); Ancestral: 13,263 (9,320–22,827)) (Figure S2, Table S2). This may suggest continuity between the ancestor and MQ populations. Note that un-sampled ('ghost') populations may inflate the apparent size of the ancestor population in this two population model [29]. The larger size estimate for VLC is consistent with the skyline plot results (Figure 4). Calculating the time to the most recent common ancestor (TMRCA) using Bayesian phylogenetic methods [14]

Table 1. Population genetic summary and demographic statistics for Victoria Land Coast (VLC) and Macquarie (MQ) populations.

	<i>N</i>	<i>S</i>	<i>h</i>	<i>H_d</i>	<i>k</i>	<i>S/k</i>	<i>F_s</i>	<i>D</i>	BF
VLC	223	91	177	0.996 (0.001)	7.73	11.77	−323.00	−1.66	exp*
							$P < 0.01$	$0.10 > P > 0.05$	
MQ	49	23	16	0.915 (0.019)	6.58	3.50	−0.11	0.88	con*
							$P = 0.13$	$P > 0.10$	

P values were obtained with 50,000 coalescent simulations. *N*: number of individuals; *S*: segregating sites, *h*: number of haplotypes; *H_d*: haplotype diversity (\pm s.d.); *k*: average number of nucleotide differences; *S/k*: ratio of *S* over *k*; *F_s*: Fu's *F_s*; *D*: Tajima's *D*; BF, approximate Bayes Factor selection of alternative demographic models (constant vs. exponential), using the difference in harmonic means of sampled marginal likelihoods implemented in BEAST [58]. * indicates $3 < BF(\pm s.e.) < 10$. doi:10.1371/journal.pgen.1000554.t001

Table 2. ρ distances (\pm s.d.) between the mtDNA pools of Victoria Land Coast (VLC) samples and those of potential source populations, representing all major extant southern elephant seal breeding colonies.

MQ	SG	HD	MA	EI	FL	PV
0.2 (0.2)	1.6 (0.14)	3.0 (0.25)	3.5 (0.5)	3.5 (0.5)	6.86 (0.86)	20.16 (1.67)

VLC: Victoria Land Coast; MQ: Macquarie Island; SG: South Georgia Island; HD: Heard Island; MA: Marion Island; EI: Elephant Island; FL: Falklands; PV: Peninsula Valdés. Calculations for SG and HD were based on 240 bp of sequence data (see Materials and Methods), and are likely underestimates. doi:10.1371/journal.pgen.1000554.t002

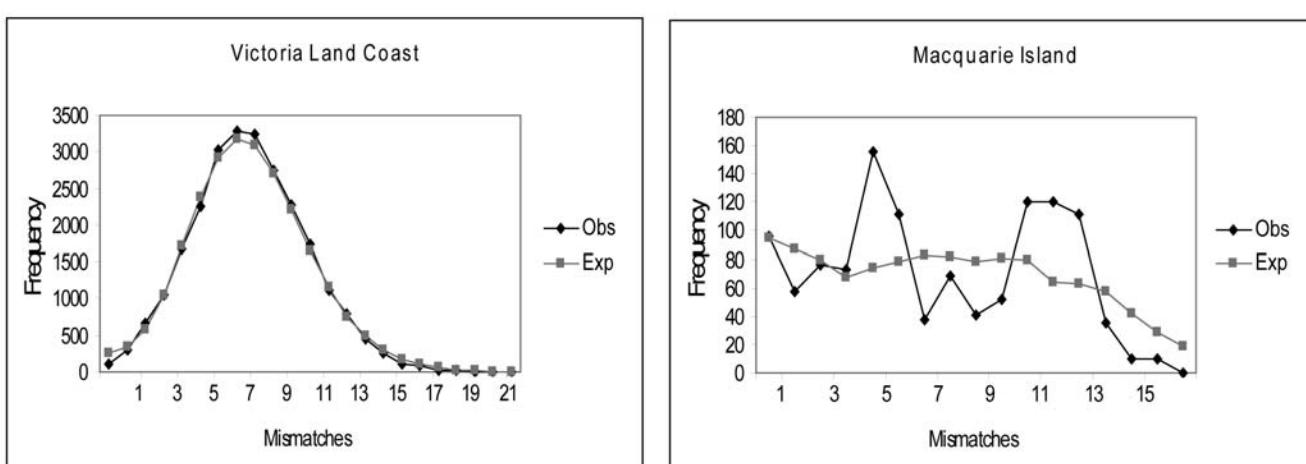


Figure 3. Mismatch distributions for mitochondrial DNA haplotypes sampled from the VLC and MQ populations.
doi:10.1371/journal.pgen.1000554.g003

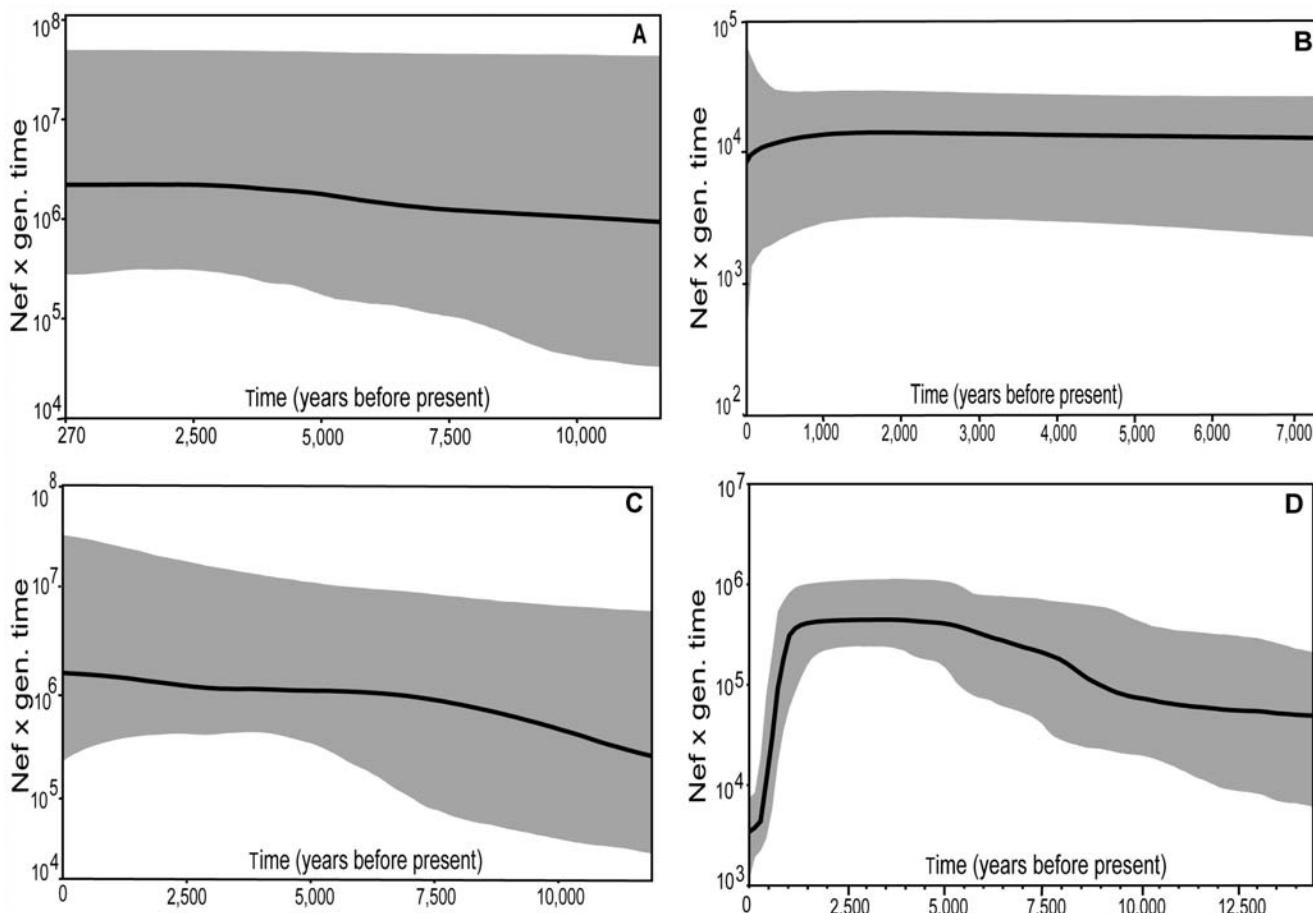


Figure 4. Bayesian skyline plots of effective population size change through time. The black line indicates the median posterior effective population size through time. The gray lines illustrate the 95% HPDI, taking into account coalescent model and phylogenetic uncertainty, for: (A) VLC alone; (B) MQ alone; (C) VLC and Marion Island combined; (D) VLC and MQ combined.

doi:10.1371/journal.pgen.1000554.g004

showed similar overlapping age estimates for the VLC and MQ samples (VLC: 11,740 YBP; 95% HPDI = 9,097–14,710; MQ: 11,090; 95% HPDI = 8,461–14,030), pre-dating the founding of VLC. All other modern populations considered both separately and pooled gave a younger TMRCA estimate, which approximated the time when the VLC site was founded ('Modern' pooled: 5,570 YBP (3,833–7,189); Figure S3).

Discussion

Effective management and mitigation of the impact of climate change on biodiversity requires an understanding of how populations are likely to respond to the consequent gain and loss of critical habitat. Predictive models [e.g. 36,37] typically base expectations on extant population distributions. However, investigating the dynamics of populations during past climate change provides a resource on which future inference can be more accurately based [38–40]. In this study the impact on habitat was driven by climate, as it involved the retreat and subsequent advance of both glacial and sea ice. The elephant seal, like many other large, highly mobile marine species, travels great distances during foraging excursions, whereas population structure is seen on a smaller geographic scale [11].

One of our key hypotheses was that a highly mobile species can respond to distant habitat gain and loss through migration.

All potential breeding habitat on the Antarctic mainland for this species was likely under ice during the last glacial period, so the VLC population must have been established sometime after the last glacial maximum. Geological evidence suggests that the timing of this would have been possible around 8,000 YBP [1–5,22]. Our results indicate that the VLC population is most closely related to MQ among the extant populations. This was supported by the phylogenetic reconstructions and inbreeding coefficients. Rho statistics also show MQ to be the most likely source for the founding population at VLC under this scenario. However, while we assume that VLC must have been founded from an extant sub-Antarctic population, there are aspects of the phylogenetic reconstructions that may suggest alternative interpretations. First, the level of diversity at VLC is higher than that of the putative source population, second, a number of the MQ haplotypes sampled appear to be derived from VLC haplotypes, and third, the TMRCA for MQ and VLC is older than the splitting time.

We therefore investigated this aspect of the relationship between VLC and MQ in considerable detail. One possible interpretation is that both VLC and MQ were derived historically from a third population. However, all major extant populations were sampled for this study, and none are possible candidates due to their genetic distance from the VLC. Therefore, this hypothetical population would have to be extinct today. The only known possible

candidate would be the population from the north-western coast of Tasmania, extinct since prehistoric times [41]. Our only extant example of a mainland population is depauperate of mtDNA variation (on Punta Norte, Peninsula Valdés, Argentina, see Figure 2), and Tasmania is 1,500 km further from VLC than MQ (raising questions as to why the VLC population would be founded by the Tasmanian rather than by the geographically closer MQ population). Neither of these considerations rule out this possibility, but other factors may.

If both VLC and MQ had been derived as founders from an unsampled historical population, both should show evidence for expansion, though only VLC does (based on Bayesian demographic model selection, coalescent assessments from Fluctuate analyses, neutrality statistics, and the mismatch distributions; see Table 1). Also, if they split separately from a common ancestor, why should an estimated splitting time between VLC and MQ coincide with the likely date that habitat was released along the VLC, when the ice sheet was retreating? These observations also count against the notion that MQ was founded from VLC, which is in any case made highly unlikely due to the lack of available habitat at VLC more than 7,000–8,000 years ago.

We propose instead that VLC was founded by MQ, and that some seals from a declining VLC population returned to MQ. These interpretations are supported by the following. If a population is founded from another, there should be haplotypes in the source population that are basal within a lineage that includes both populations. This is true for the VLC and MQ populations, though the most convincing case is for just one, infrequently sampled haplotype in MQ found at the division point between the VLC/MQ lineage and the rest of the populations (see Figure 2B). This haplotype also has a close relationship to the two VLC-sampled potential founder haplotypes. There is also a basal haplotype within the network that is shared between MQ and VLC (Figure 2B), suggesting a common root for the two populations. In further support of this is the illustration that relatively old VLC haplotypes within the network are closely allied with MQ basal haplotypes (see Figure 2C). Even so, it is clear that basal haplotypes are not very well sampled from the MQ population. One possible explanation could be the loss of lineages at MQ in the relatively recent past (perhaps as a consequence of sealing pressure) and other sampling effects.

There is a signal from the IM analyses indicating unidirectional migration from VLC to MQ at about 1,300 YBP. This is consistent with the idea that seals from VLC returned to MQ as the VLC population was starting to decline. A trend for decreasing genetic distance between these populations over time may also support this (though unlike the various non-equilibrium analyses used, the ϕ_{ST} data assume equilibrium conditions that probably don't apply). From the phylogenies, you would expect to see some MQ haplotypes as derived tips within the network, and this pattern is clearly evident. Furthermore, these are primarily derived from the more recent VLC samples (dated to 506–1,565 YBP; see Figure 2C). The IM estimates depend on an accurate substitution rate estimate, which from our calculations have a large error associated with them. However, the HPD rate estimate we derived is very similar to estimates derived from intraspecific analyses for a variety of taxa [16–18,23,24,26,38], and in all cases where analysed (as in this study), these rates provide good correspondence to the timing of known historical and geologic events (e.g. [38–40]). It is in that context that we apply this substitution rate.

A further consideration with IM is the assumption that the two populations being compared were sampled at the same point in time. While this is not the case for comparisons between VLC and MQ, the chronological results from this analysis are nevertheless

consistent with those obtained using other methods and with the geologic record. This also applies to other analyses that involve the incorporation of different chronologies into estimates of population differentiation (such as ϕ_{ST}) or population expansion (such as Tajima's D) [42]. We tested the robustness of these comparisons for VLC sample subsets involving timeframes that represent a small proportion of the TMRCA (<10%), and show that our conclusions remained consistent (Table S3).

Further support for the proposed pattern of demographic expansion and contractions at VLC comes from the BSP profile shown in Figure 4D. This BSP plot shows population expansion and contraction signals consistent with all other data from the genetic analyses, and with the expected timing based on geologic inference [1–5,22]. The problem is that the construction of this plot violates some key assumptions. First, there is an assumption of panmixia, and the ϕ_{ST} data (for example) show that MQ and VLC are differentiated. However, empirical data and simulation studies suggest skyline plots are robust to violations of assumptions of panmixia [19,20,32,38]. There is also the potential problem associated with combining populations that have very different effective sizes. However, we test this by combining VLC with other extant breeding populations (Marion Island; Elephant Island) with very similar or smaller census sizes as that for MQ [6,33–35]; see Figure 1 and Figure 4C). In this case there is no decline seen at 1,000 YBP, or at any other point. Further, it isn't clear why the combination of VLC and MQ should show a decline as old as 1,000 YBP, nor why the end point should be lower than the estimated effective population size on MQ (see Figure 4B and 4D). We suggest that this could be a result of VLC seals having returned to MQ, and taken with them lineages that retain the signal of the VLC population crash. The profile for VLC alone may not show this crash due to the paucity of very recent samples from this site (see Table S1). However, the various problems discussed mean that this interpretation should be treated cautiously, and only as consistent with the stronger inference gained from the various other analyses.

As mentioned previously, the greater diversity of the VLC site was unexpected for a founder population, and much of it is unique to that population. Two key factors may be important. First, the founded population could have established and grown quickly, if as now the forage resource was high. Second, high numbers were supported over time (estimated in this study as at least an order of magnitude higher than a long-term stable population size estimate for MQ). Again, this likely reflects a large (hundreds of kilometres of newly available habitat [22]) local environment rich in resources for the VLC population. The greater diversity at VLC compared to MQ may also reflect more rapid lineage extinctions at MQ, where N_e was lower (and perhaps impacted by sealing pressure on population size during the last century [43]). Given that the mtDNA genome includes numerous functional loci, selection needs to be a consideration when interpreting these data. However, in this case the high diversity and lack of evidence for increased frequency of specific haplotypes in the VLC mean that there is no clear evidence to support that interpretation.

Another factor is the relatively old TMRCA for MQ and VLC, however this is expected if they share common ancestry (as they would if one was derived from the other) and if some level of gene flow was maintained (as is suggested by the IM analysis). The more recent, shared TMRCA among the other populations is perhaps more unexpected (see Figure S3). One possible explanation is that the environmental changes that led to the founding of VLC also had a demographic impact on the other sub-Antarctic breeding sites, though this would need to be investigated further to be confirmed.

Taken together, our results indicate that new habitat was quickly exploited, and that the founder population became independent of the distant source population (given high and significant ϕ_{ST} values throughout the relevant time period, the distinct mismatch distribution profile, and IM results for time of splitting). The genetic data clearly distinguishes the VLC population from a site used exclusively for moulting, which could have represented the temporary assemblage of seals from one or more separate breeding populations. The implication is that discovery of emergent habitat was facilitated by long-distance foraging excursions, and that some level of continued dispersal between sites allowed the return of a proportion of the founder population to the source colony, when the new habitat was lost many generations later. This type of long distance networking may be possible for other species of conservation concern at times of rapid environmental change. In contrast, species with a more local distribution and range may be less able to cope. For example, Arctic foxes from midlatitude Europe went extinct during the Pleistocene when available habitat contracted [44]. Predictive modelling for the development of conservation strategies should therefore take these factors into account. In general, these data indicate the rapid establishment of a genetically differentiated population in emerging habitat, where diversity levels remained very high. Much of this novel diversity was later lost when the VLC breeding habitat was lost. This suggests that when conditions are optimal, the potential for future adaptive radiation over relatively short time frames is high, and that this potential can be lost again just as rapidly.

Materials and Methods

Sample collection, ancient DNA extraction, PCR, and sequencing

Southern elephant seal skin samples were collected from beaches along the Victoria Land Coast, Ross Sea, Antarctica (Table S1), and stringent sample collection procedures were followed to avoid anthropogenic- or cross-contamination of samples. Ancient DNA extractions and PCR setup were conducted in a dedicated aDNA laboratory at the University of Durham. Strict aDNA protocols were adhered to [10], and results carefully assessed in this context: 1) One-way movement of personnel/ equipment/ PCR set-up from aDNA lab to Molecular Ecology lab; 2) Regular bleach & UV treatment of all work surfaces; 3) Use of aerosol-resistant pipette filter-tips for aDNA work; 4) Regular autoclaving of all equipment, plasticware, and DNase & RNase-free deionized sterile water from a dedicated aDNA autoclave; 5) Use of multiple blank controls for both extraction and PCR procedures; 6) Replication, cloning and independent replication of a subset of aDNA samples [22]; 7) Confirmation that results make phylogeographical sense.

To check the accuracy of the extraction, PCR, and sequencing steps, approximately 30% of samples were replicated for all steps, while 10 samples were replicated 4 times. These steps exceed established replication protocols for the aDNA field [10,45]. Nine samples extracted independently in a different lab five years prior to the current study by another researcher for an earlier study [22] were re-extracted and sequenced (i.e. independent replication), and all sequences matched. Six PCR products were cloned (six clones from each product were sequenced) to evaluate template damage and to check for nuclear pseudogenes and contamination.

Skin samples were digested in 3 mL of extraction buffer, containing 0.5% SDS, 0.45 M EDTA (pH 8), 10 mM Tris-HCL, and 50 μ l Proteinase K, at 50°C overnight. Digests were centrifuged at 13,000 g for 1 min, the aqueous phase was

removed, and added to 5 volumes of PB Buffer (QIAgen) and vortexed. This solution was passed in 700 μ l quantities through QIAquick spin filter columns (QIAgen), centrifuged at 13,000 g for 1 min, and the flow-through discarded. DNA was then twice washed with 700 μ l of Buffer PE (QIAgen), centrifuged as before, and the flow through discarded. The filter column was placed in a sterile 1.5 ml tube, and DNA was eluted by applying 35 μ l of Buffer EB (QIAgen) to the column. This was then centrifuged at 13,000 g for 1 min. DNA samples were stored at -20°C until required. DNA extractions were performed in batches of 20 samples with 1–2 blanks (no tissue) per batch.

PCR reactions (25 μ l volumes) contained: 2 μ l extract, 1 U Platinum *Taq* Hi-Fidelity, 1×buffer, 1.5 mM MgCl₂ (Invitrogen Ltd., UK), 0.2 μ M each primer and 200 μ M each dNTP. BSA was found to be inhibitory, possibly due to the presence of hairs in the elephant seal skin samples [46]. Thermal cycles were as follows: 95°C for 2 min, 45 cycles of 94°C for 45 sec, 52–55°C for 45 sec, and 72°C for 45 sec, followed by a final extension at 68°C for 10 min. Negative controls were used for all PCR runs. Positive PCR amplicons of the correct size were purified using a QIAgen Purification Kit according to manufacturer's instructions. Overlapping fragments of DNA were amplified for a final target of 325 bp of the most variable portion of the mitochondrial control region (HVR1). Primers used: SESanc3f 5'-GCTGACATTC-TACTTAACT-3' & SESanc3r 5'-ATGTACATGCTTATGCAT-3'; SESmdbf 5'-AGCCCTATGTATATCCTG-CATT-3' & SESmdbr 5'-CAGTATAGAAACCCCCACATGA-3'. Purified PCR products were sequenced at Macrogen (S. Korea) and at the University of Durham. There was no variation among replicates for a given sample. During cloning and sequencing, no evidence of nuclear copies was detected, while one sample displayed evidence of damage and was removed from further analyses, consistent with the generally excellent preservation of these 'freeze-dried' Antarctic samples.

¹⁴C Radiocarbon dating of ancient tissue and calibration methods

Radiocarbon dating of 223 ancient VLC samples was by Accelerator Mass Spectrometry (AMS) at the NOSAMS laboratory at Woods Hole Oceanographic Institution and at the University of Arizona. All dates were calibrated using the CALIB [47] program and the Marine04 dataset, with a time-dependent Southern Ocean marine-reservoir effect (on average, $\Delta_{\text{LR}} = 791 \pm 121$ yrs), derived from paired radiocarbon and uranium-thorium dates of Holocene corals collected along the VLC (BL Hall et al. unpublished data). Radiocarbon dates were converted to calendar dates (Table S1).

Modern DNA extraction, PCR, and sequencing

Modern southern elephant seal samples were analysed after completion of all aDNA work (Macquarie Island $n=49$; Marion Island $n=48$) or data incorporated from earlier studies using Genbank sequences (Peninsula Valdez, $n=50$; Falklands, $n=56$; Elephant Island, $n=30$; South Georgia, $n=30$; Heard Island, $n=6$) [11,12]. All major extant southern elephant seal breeding colonies are represented in our analyses (Figure 1). Modern sequences were aligned to the aDNA sequences and trimmed to create a final combined HVR1 dataset of 269 modern and 223 aDNA sequences of 325 bp in length; except for the South Georgia and Heard samples, for which only 240 bp of homologous sequence data were available. Tissue digestions followed the protocol for aDNA samples, followed by standard phenol: chloroform extraction [48]. Blank controls were incorporated. PCR setup was the same as for aDNA, except that the

number of thermal cycles was reduced to 35, and annealing temperature was set at 55°C. Primers (SESM1F: 5'-TGACATC-CATCCCCCTTATT-3' and SESMR: 5'-TGTGTG AT-CATGGGCTGATT-3') were designed to amplify a 440 bp fragment encompassing the entire aDNA target fragment.

Phylogeographic and demographic analyses of ancient and modern sequences

Phylogenies [13,14] were constructed from the full dataset to visualise relationships among modern and ancient haplotypes. Median-joining networks were implemented in Network 4.5 (<http://www.fluxus-engineering.com>) [13] based on 325 bp sequences for the full dataset, and a second network was generated focussing on the MQ/VLC clade. Additional analyses were also conducted with a reduced dataset of 240 bp incorporating South Georgia and Heard Islands samples. No pre-processing was required, but Maximum Parsimony (MP) post-processing was implemented, to remove links not used by the shortest trees in the network, thereby reducing network complexity. Networks were analysed according to founders' analysis principles [27]. As a consequence of the close genetic relationship identified between MQ and VLC and initial founders' analysis results, we also adopted criteria $f0-2$ proposed by Richards et al. [27] to distinguish between candidate founders and back-migration. In brief (see [27] for details), potential founders are not permitted at the tips of the clade to minimise the possibility of recurrent mutations being designated founders, while back-migration can be identified as haplotypes derived from haplotypes intermediate to those of the founder. Ancestral haplotypes for the MQ/VLC clade were identified on the basis of minimum ρ value estimates [49] for that clade [24]. In addition, calendar ages (derived from radiometric ages) were mapped onto the MQ/VLC network to determine whether founder haplotypes were consistent with haplotype ages along their respective lineages. Minimum inter-population ρ value estimates were then calculated between ancestral VLC haplotypes and the seven modern population haplogroups, representing all major extant breeding colonies (potential source colonies). An implicit assumption in these analyses is that no alternate potential source population (and its' mitochondrial lineages) has gone to extinction, and thus remains unsampled in the current study.

Bayesian coalescent analysis using Markov Chain Monte Carlo (MCMC) was implemented in the program BEAST v. 1.4.8 and v. 1.5.b [14], using a substitution rate model (GTR+I+G) determined by MODELTEST [50]. Phylogenetic relationships among all samples were also investigated in BEAST through the construction of a Maximum Clade Credibility Tree. Version 1.5.b was provided by one of the BEAST program developers (Andrew Rambaut), and incorporates a post-mortem damage model [15], which takes into account the potential for sequence damage to influence the outcome of the aDNA analyses. Demographic inferences were essentially the same with or without the incorporation of the post-mortem damage model (details available from authors). BEAST analyses used MCMC to integrate over all credible genealogies, while simultaneously estimating both substitution rate (using corrected calendar ages as calibrations) and the demographic history of the sampled sequences. This was done under three demographic coalescent models: constant, exponential, and Bayesian Skyline Plot (BSP [32]). Serially sampled aDNA sequences provide the opportunity to calibrate substitution rate directly [17]. Rate estimates for further analyses were derived under the BSP model, from the VLC samples only (strict clock model), as genetic subdivision (i.e. population structure) within a dataset may bias rate estimates upwards [51]. A BSP coalescent

model is preferable for rate estimation as it 'averages out' the demographic history of the sample, which can be considered a nuisance parameter in this instance [32]; notwithstanding, substitution rate estimates have been shown to be consistent across alternate underlying demographic coalescent models [16].

To ensure the substitution rate estimate was not merely an artefact of the dates provided, i.e. that the aDNA sequences used in this study were rate informative, these analyses were repeated three times by randomly re-assigning calendar ages to sequences. A recent re-analysis [51] of exceptionally fast rates reported for tuatara [52] utilised this approach to show that the tuatara aDNA dataset was not rate informative. For our study, rates estimated from the re-assigned ages were in all cases highly skewed to zero (Figure S4), while the rate estimate derived from the 'true' calendar ages returned a bell curve estimate not encompassing zero (Figure S1). This test indicates that the sequences were informative. Three alignments were run in BEAST: 1. Full dataset - all samples; 2. Ancient samples only (VLC); 3. VLC and MQ clade (BSP model only). As a control measure to determine whether the MQ/VLC BSP would be influenced primarily by the smaller effective population size of MQ relative to that of the VLC, we generated a further two BSPs by including VLC with two modern populations, Elephant Island and Marion Island, respectively, each with similar or smaller population sizes to MQ (Figure 1) [6,33–35]. Although combining separate populations violates the assumption of panmixia, simulation and empirical studies have shown skyline plots to be robust to violations of this assumption [19,20,32,38]. Three independent MCMC samples per alignment were run for 20,000,000 generations, sampled every 2,000 generations, after the initial 10% were discarded as burn-in. These three independent samples were combined using LogCombiner and analyzed in Tracer v1.4 (both programs distributed with BEAST), to generate credibility intervals that represent coalescent model and phylogenetic uncertainty, and to produce the final BEAST results. The effective sample size (ESS) for all parameter estimates was at least 200, while the autocorrelation times of MCMC plots indicated that runs converged on the equilibrium distribution.

Summary and population genetic statistics were calculated in DNASP v. 10.4.9 [53]. The number of segregating sites S over the average number of pairwise nucleotide differences k describes the 'expansion coefficient' [54], with high values indicating an increase in population size through time, while low values should be indicative of relatively long-term constant population size. DNASP was also used to compare genetic diversity estimates for nucleotide diversity. We used 1,000 coalescent simulations to derive 95% confidence intervals for these estimates, in order to assess whether the distributions overlapped (with non-overlapping distributions providing a conservative estimate of significant differentiation). Tajima's D [55] and Fu's F_s [56] neutrality test statistics were also estimated. Under assumptions of neutrality, negative values indicate a signature of population expansion. Fu's F_s is less conservative than Tajima's D , and is expected to be more sensitive to large population expansions (indicated by highly significant negative values). Mismatch distributions [30] and ϕ_{ST} (using a substitution rate model determined by MODELTEST [50]) were estimated using ARLEQUIN v. 3.01 [57]. The validity of the estimated stepwise expansion model was tested using the sum of squared deviations (SSD) between the observed and the expected mismatch as a test statistic. Both the mismatch distribution and neutrality test statistics operate under infinite site assumptions which may not hold, particularly for the large number of haplotypes identified within the VLC sample. Thus, these estimates were derived for MQ and VLC to see whether results

were concordant with demographic model selection implemented in BEAST.

To understand changes in VLC population size through time, FLUCTUATE [31] analyses was implemented using 20 short chains of 1,000 steps each, and five long chains of 20,000 steps were used to determine parameters for the production runs. Production runs were implemented using 20 short chains of 8,000 steps each, and 10 long chains of 50,000 steps. The program was run multiple times to ensure concordance of parameter estimates. Haplotype richness for VLC was measured in relation to MQ ($n=49$) using Rarefactor (<http://www2.biology.ualberta.ca/jbrzusto/rarefact.php>). An Isolation-with-Migration model was implemented in IM [29] to estimate effective population sizes, time of splitting, and migration rates between the VLC and MQ populations. All IM analyses used a flat prior (0–10) for m (migration) and t (divergence time), and changes were allowed in population size incorporated into the IM model. To facilitate these analyses, the substitution rate estimated in BEAST was incorporated. This was converted to a rate per generation based on a 4-year generation time for N_e estimates (Table S2). Three independent replicates of each pairwise comparison were performed, with the Markov chains run for 100,000,000 generations, after discarding a 10% burn-in. To ensure convergence, simulations were run until the smallest ESS estimates were greater than 200, and update rates were greater than 20%. Although the basic IM model cannot account for the sizes of founding populations, the parameter ‘ s ’ controls for this [21], providing for the proportion of the ancestral population that founds the descendant populations. In this way, expansion or contraction can be modelled. A uniform prior distribution for s between 0.5 and 1.0 was set (after a similar analysis in [21]) to ensure that the estimates were robust to different patterns of expansion (data not shown). IM results were concordant with either uninformative (0–1) or hard (0.5–1) priors on s .

Supporting Information

Figure S1 BEAST output of southern elephant seal substitution rate estimated from 223 directly radiocarbon dated samples from Victoria Land Coast (VLC). Radiocarbon dates were converted to calendar years.

Found at: doi:10.1371/journal.pgen.1000554.s001 (1.49 MB TIF)

Figure S2 Isolation-with-migration (IM) output of population genetic parameters for the Macquarie Island (MQ), Victoria Land Coast (VLC) and Ancestral “populations.” See Table S2 for text version of IM parameter estimate distributions.

Found at: doi:10.1371/journal.pgen.1000554.s002 (8.75 MB TIF)

References

1. Conway H, Hall BL, Denton GH, Gades AM, Waddington ED (1999) Past and future grounding-line retreat of the West Antarctic ice sheet. *Science* 286: 280–283.
2. Baroni C, Orombelli G (1991) Holocene raised beaches at Terra Nova Bay, Victoria Land, Antarctica. *Quat Res* 36: 157–177.
3. Baroni C, Orombelli G (1994) Holocene glacier variations in the Terra Nova Bay area (Victoria Land, Antarctica). *Antarctic Sci* 6: 497–505.
4. Baroni C, Hall B (2004) A new Holocene relative sea-level curve for Terra Nova Bay, Victoria Land, Antarctica. *Jnl Quat Science* 19: 377–396.
5. Hall BL, Denton GH (2002) Holocene history of the Wilson Piedmont Glacier along the southern Scott Coast, Antarctica. *The Holocene* 12: 619–627.
6. LeBoeuf B, Laws R (1994) Elephant Seals: Population Ecology, Behaviour, and Physiology. Berkeley: University of California.
7. Van den Hoff J, Davies R, Burton H (2003) Origins, age composition and change in numbers of moulting southern elephant seals (*Mirounga leonina* L.) in the Windmill Islands, Vincennes Bay, east Antarctica, 1988–2001. *Wild Res* 30: 275–280.
8. Biuw M, Boehme L, Guinet C, Hindell M, Costa D, et al. (2007) Variations in behaviour and condition of a Southern Ocean top predator in relation to *in situ* oceanographic conditions. *Proc Nat Acad Sci USA* 104: 13705–13710.
9. Nichols R (1961) Characteristics of beaches formed in polar climates. *Am Jnl Sci* 259: 694–708.
10. Gilbert MTP, Bandelt HJ, Hofreiter M, Barnes I (2005) Assessing ancient DNA studies. *Trends Ecol Evol* 20: 541–544.
11. Fabiani A, Hoelzel AR, Galimberti F, Muelbert MMC (2003) Long range paternal gene flow in the southern elephant seal. *Science* 299: 676.
12. Hoelzel AR, Campagna C, Arnbom T (2001) Genetic and morphometric differentiation between island and mainland southern elephant seal populations. *Proc Roy Soc B* 268: 325–332.
13. Bandelt HJ, Forster P, Rohl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16: 37–48.
14. Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7: 214.

Figure S3 BEAST output of Time-to-Most-Recent-Common-Ancestor (TMRCA) estimates (YBP). For: all modern populations grouped (besides Macquarie Island) (blue), Macquarie Island (MQ - yellow), and Victoria Land Coast (VLC - red).
Found at: doi:10.1371/journal.pgen.1000554.s003 (1.49 MB TIF)

Figure S4 BEAST output of southern elephant seal substitution rate estimated from 223 randomized radiocarbon dated samples from Victoria Land Coast (VLC). Radiocarbon dates were converted to calendar years. Dates were randomly assigned to sequences to assess whether the non-randomized ages and sequences were rate informative. These analyses were repeated three times, and all randomized rate estimates were highly skewed to zero.

Found at: doi:10.1371/journal.pgen.1000554.s004 (1.99 MB TIF)

Table S1 Sample location and age.

Found at: doi:10.1371/journal.pgen.1000554.s005 (0.32 MB DOC)

Table S2 Isolation-with-migration (IM) output of population genetic parameters for the Macquarie Island (1), Victoria Land Coast (2) and Ancestral (a) “populations.” See Figure S2 for graphic depiction of IM parameter estimate distributions.

Found at: doi:10.1371/journal.pgen.1000554.s006 (0.04 MB DOC)

Table S3 Parameter comparisons against chronology subsets for VLC. Φ_{ST} = pairwise estimate between MQ and VLC “subset.” % TMRCA = subset age interval (%) relative to the time-to-most-recent-common-ancestor of the VLC sample (estimated at 11,740 YBP in BEAST).

Found at: doi:10.1371/journal.pgen.1000554.s007 (0.03 MB DOC)

Acknowledgments

We thank Joanne Stamford and Oliver Pye for lab assistance; Clive McMahon, Mark Hindell, and the Antarctic Wildlife Research Unit, University of Tasmania, for the provision of Macquarie Island samples; and Alexei Drummond, Simon Ho, Roger Thorpe, Burney Le Boeuf, Andrew Rambaut, and four anonymous reviewers and the Subject Editor who provided helpful advice and comments. A. Bamberg, B. Chase, A. Doughty, S. Newsome, B. Overturf, and M.C. Salvatore assisted with field collections in the Ross Sea.

Author Contributions

Conceived and designed the experiments: BLH ARH. Performed the experiments: MDB LFC. Analyzed the data: MDB BLH ARH. Contributed reagents/materials/analysis tools: BLH LFC CB PLK ARH. Wrote the paper: MDB ARH. Fieldwork: BLH CB PLK.

15. Rambaut A, Ho SYW, Drummond AJ, Shapiro B (2009) Accommodating the effects of ancient DNA damage on inferences of demographic histories. *Mol Biol Evol* 26: 245–248.
16. Ho SYW, Kolokotronis S-O, Allaby RG (2007) Elevated substitution rates estimated from ancient DNA sequences. *Biol Letters* 3: 702–705.
17. Lambert DM, Ritchie PA, Millar CD, Holland B, Drummond AJ, et al. (2002) Rates of evolution in ancient DNA from Adélie penguins. *Science* 295: 2270–2273.
18. Howell N, Smejkal CB, Mackey DA, Chinnery PF, Turnbull DM, et al. (2003) The pedigree rate of sequence divergence in the human mitochondrial genome: there is a difference between phylogenetic and pedigree rates. *Am J Hum Genet* 72: 659–670.
19. Atkinson QD, Gray RD, Drummond AJ (2008) mtDNA Variation predicts population size in humans and reveals a major southern Asian chapter in human prehistory. *Mol Biol Evol* 25: 468–474.
20. Atkinson QD, Gray RD, Drummond AJ (2009) Bayesian coalescent inference of major human mitochondrial DNA haplogroup expansions in Africa. *Proc Roy Soc B* 276: 367–373.
21. Hey J (2005) On the number of New World founders: A population genetic portrait of the peopling of the Americas. *PLoS Biology* 3: e193. doi:10.1371/journal.pbio.0030193.
22. Hall BL, Hoelzel AR, Baroni C, Denton GH, Le Boeuf BJ, et al. (2006) Holocene elephant seal distribution implies warmer-than-present climate in the Ross Sea. *Proc Nat Acad Sci USA* 103: 10213–10217.
23. Ho SYW, Phillips MJ, Cooper A, Drummond AJ (2005) Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Mol Biol Evol* 22: 1561–1568.
24. Henn BM, Gignoux CR, Feldman MW, Mountain JL (2009) Characterizing the time dependency of human mitochondrial DNA mutation rate estimates. *Mol Biol Evol* 26: 217–230.
25. Bandelt H-J (2008) Clock debate: when times are a-changin': Time dependency of molecular rate estimates: tempest in a teacup. *Heredity* 100: 1–2.
26. Howell N, Howell C, Elson JL (2008) Molecular clock debate: Time dependency of molecular rate estimates for mtDNA: this is not the time for wishful thinking. *Heredity* 101: 107–108.
27. Richards M, Macaulay V, Hickey E, Vega E, Sykes B, et al. (2000) Tracing European founder lineages in the near Eastern mtDNA pool. *Am J Hum Genet* 67: 1251–1276.
28. Castelloe J, Templeton AR (1994) Root probabilities for intraspecific gene trees under neutral coalescent theory. *Mol Phyl Evol* 3: 102–113.
29. Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* 167: 747–760.
30. Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Mol Biol Evol* 9: 552–569.
31. Kuhner MK, Yamato J, Felsenstein J (1998) Maximum likelihood estimation of population growth rates based on the coalescent. *Genetics* 149: 429–434.
32. Drummond AJ, Rambaut A, Shapiro B, Pybus OG (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol Biol Evol* 22: 1185–1192.
33. McMahon CR, Burton HR, Bester MN (2003) A demographic comparison of two southern elephant seal populations. *J Anim Ecol* 72: 61–74.
34. Pistorius PA, Taylor FE, Bester MN, Hofmeyr GJG, Kirkman SP (2008) Evidence for density dependent population regulation in southern elephant seals in the southern Indian Ocean. *Afr Zool* 43: 75–80.
35. Fabiani A, Galimberti F, Sanvito S, Hoelzel AR (2004) Extreme polygyny among southern elephant seals on Sea Lion Island, Falkland Islands. *Behav Ecol* 15: 961–969.
36. Araújo MB, Rahbek C (2006) Ecology: How does climate change affect biodiversity? *Science* 312: 1396–1397.
37. Kapelle M, Van Vuuren MMI, Baas P (1999) Effects of climate change on biodiversity: a review and identification of key research issues. *Biodiv Cons* 8: 1383–1397.
38. Shapiro B, Drummond AJ, Rambaut A, Wilson MC, Matheus PE, et al. (2004) Rise and fall of the Beringian steppe bison. *Science* 306: 1561–1565.
39. Ho SYW, Larson G, Edwards CJ, Heupink TH, Lakin KE, et al. (2008) Correlating Bayesian date estimates with climatic events and domestication using a bovine case study. *Biol Letters* 4: 370–374.
40. Ritchie PA, Millar CD, Gibb GC, Baroni C, Lambert DM (2004) Ancient DNA enables timing of the Pleistocene origin and Holocene expansion of two Adélie penguin lineages in Antarctica. *Mol Biol Evol* 21: 240–248.
41. Bryden MM, O'Connor S, Jones R (1999) Archaeological evidence for the extinction of a breeding population of elephant seals in Tasmania in prehistoric times. *Int J Osteoarchaeol* 9: 430–437.
42. Depaulis F, Orlando L, Hanni C (2009) Using classical population genetic tools with heterochronous data: Time matters. *PLoS One* 4: e5541. doi:10.1371/journal.pone.0005541.
43. Hindell MA, Burton HR (1988) The history of the elephant seal industry at Macquarie Island and estimates of the pre-sealing numbers. *Pap Proc Roy Soc Tas* 122: 159–176.
44. Dalen L, Nyström V, Valdiosera C, Germonpré M, Sablin M, et al. (2007) Ancient DNA reveals lack of postglacial habitat tracking in the arctic fox. *Proc Nat Acad Sci USA* 104: 6726–6729.
45. Willerslev E, Cooper A (2005) Ancient DNA. *Proc Roy Soc B* 272: 3–16.
46. Gilbert MTP, Wilson AS, Bunce M, Hansen AJ, Willerslev E, et al. (2004) Ancient mitochondrial DNA from hair. *Curr Biol* 14: R463–R464.
47. Stuiver M, Reimer PJ (1993) Extended ¹⁴C data base and revised CALIB 3.0 ¹⁴C age calibration program. *Radiocarbon* 35: 215–230.
48. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual. New York: Cold Spring Harbor Laboratory Press.
49. Forster P, Harding R, Torroni A, Bandelt HJ (1996) Origin and evolution of native American mtDNA variation: a reappraisal. *Am J Hum Genet* 59: 935–945.
50. Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics* 14: 817–818.
51. Miller HC, Moore JA, Allendorf FW, Daugherty CH (2009) The evolutionary rate of tuatara revisited. *Trends Genetics* 25: 13–15.
52. Hay JM, Subramanian S, Millar CD, Mohandesan E, Lambert DM (2008) Rapid molecular evolution in a living fossil. *Trends Genetics* 24: 106–109.
53. Rozas J, Sanchez DelBarrio JC, Messeguer X, Rozas R (2003) DNAsP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19: 2496–2497.
54. Peck DR, Congdon DC (2004) Reconciling historical population processes and population structure in the sooty tern *Sterna fuscata*. *J Avian Biol* 35: 327–335.
55. Tajima F (1989) The effect of change in population size on DNA polymorphism. *Genetics* 123: 597–601.
56. Fu Y-X (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147: 915–925.
57. Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evol Bioinf Online* 1: 47–50.
58. Suchard MA, Weiss RE, Sinsheimer JS (2001) Bayesian selection of continuous-time Markov chain evolutionary models. *Mol Biol Evol* 18: 1001–1013.