Chapter 2

The Genetic Environment of Melanesia: Clines, Clusters and Contact

Murray P. Cox
Arizona Research Laboratories, University of Arizona, Tucson, Arizona, USA
Santa Fe Institute, Santa Fe, New Mexico, USA

Abstract

Current reconstructions of prehistory mostly reject early attempts to force artificial classifications on the Indo-Pacific region, such as the tripartite division of Melanesia, Micronesia and Polynesia historically defined by Dumont d’Urville. Instead, the modern anthropological community embraces a more fluid view of human society, in which communities change through time and space, experience internal developments, and interact with surrounding groups. However, such a pliant conception provides a poor model system for many biological, linguistic and cultural clusters observed in the Indo-Pacific region today. Natural features – acting as conduits and barriers to human mobility – have also been influential regulators of population demography, and contemporary molecular research reveals numerous biological markers that reflect this geography. Some genetic characters show signs of regulation by natural factors; others even have surprisingly accurate concordance with the western boundary of Melanesia. A reconsideration of Melanesia and its biological environs from genetic, linguistic and anthropological perspectives emphasizes the temporal and spatial complexity of the Indo-Pacific world, but also reveals a mosaic prehistory dominated by population clines and clusters, and most importantly, strong evidence for population contact.

1. Arizona Research Laboratories, 1041 East Lowell Street, Biological Sciences West 246B, University of Arizona, Tucson, AZ 85721, USA. Email: mpcox@u.arizona.edu
Introduction

The French explorer Dumont d’Urville (1831, 2003) first coined the term ‘Melanesia,’ defining the region as New Guinea and the Bismarck Archipelago, the Solomon and Vanuatu Archipelagos, New Caledonia, Fiji, and parts of Australia. With the exception of Australia, whose inclusion has never met with widespread approval (reviewed by Clark 2003a), and Fiji, whose placement has been routinely debated (for instance, Buxton 1926), anthropologists and biologists have used the concept of Melanesia (Figure 1) in their research for over 170 years. Attempting to classify the Indo-Pacific region further, d’Urville defined three additional geographical zones: Polynesia, the islands of the eastern Pacific Ocean; Micronesia, the islands of the northern Pacific; and Malaysia (modern ‘Island Southeast Asia’), a catch-all grouping of islands to the west, now primarily represented by the modern nations of Indonesia, Malaysia and the Philippines. The Maluku archipelago and the lesser Sundas in eastern Indonesia have an especially ambiguous placement in this setting. D’Urville (2003) did not define these islands as Melanesian, and some authors consider them a new group, Wallacea, an island zone between the Sunda (“Asian”) and Sahul (“Australian/New Guinea”) continental plates. However, these archipelagos have the same geographic affinity to western New Guinea as the Bismarck archipelago (New Britain and New Ireland) has to eastern New Guinea (neither was connected to the Sahul continent during periods of low sea level; O’Connell and Allen 2004). I therefore consider both groups as having Melanesian connections for the following discussion.

![Figure 1. The traditional boundaries of Melanesia with biogeographical, linguistic and anthropological divisions in the Indo-Pacific region.](image)

As anthropologists have come to recognize that cultural, linguistic and biological characters tend to intergrade from place to place, the validity of classifying the Indo-Pacific
The Genetic Environment of Melanesia

world into such discrete units has come under attack (although note that d’Urville did not envisage his divisions as being rigid and without interaction; D’Arcy 2003:223). Broad geocultural terms such as Melanesia are now deemed to be of limited utility for many anthropological purposes (Green 1991a, 2001); a view that has become standard enough even for inclusion in introductory textbooks (Jobling, Hurles and Tyler-Smith 2004:356). Here, however, I reassess the concept of Melanesia from a modern biological perspective, and show that a reconsideration of genetic data from this standpoint raises important new insights about demographic processes in the prehistoric Indo-Pacific region.

Importantly, even the earliest European explorers perceived Melanesia as more than just a geographical region. The term Melanesia is itself a biological descriptor: derived from the Greek words for “black islands,” it refers to indigenous peoples’ dark skin pigmentation (Harvey 1985, Norton et al. 2006), and was never purely a geographical definition. Early physical studies of Indo-Pacific peoples invariably supported this view of Melanesia as a region containing groups of people with relatively distinctive biological characteristics. Summarizing the early twentieth century’s biometrical research, Howells (1977:171) concluded that “there is nothing subjective about this assessment, which sets off an Australo-Melanesian population from anything to the immediate east, north or west, all of which seems to belong to a broad and varied Mongoloid population complex.” This conclusion was not limited to Howells, but is mirrored in other biometric works; for instance, Coon (1966), and more recently, Pietrusewsky (1994) and Hanihara (1996). Pietrusewsky, in particular, states that such large morphological differences imply distinct genetic roots. In concordance with this supposition, the distributions of classical genetic markers (systems such as the blood group antigen, ABO) have generally supported the assessment of Melanesia as a region containing population groups distinguishable from communities in Polynesia, Micronesia and Island Southeast Asia. Despite the relatively low limits of resolution inherent in classical genetic systems, Kirk (1976) demonstrated clear differences in the allele frequencies of Gm, Gc, Transferrin, and Diego red cell antigens between populations in Melanesia and Polynesia (see also Kirk’s 1992 review article). And summarizing seventy years of classical genetic research in the Indo-Pacific region, Cavalli-Sforza, Menozzi and Piazza (1994:383 ff.) produced a dataset suggesting that populations within the traditional boundaries of Melanesia carry at least eight blood group alleles at higher frequencies than Polynesians, Micronesians and Island Southeast Asians. In other words, populations within the traditional boundaries of Melanesia share certain classical genetic traits more commonly among themselves than with people in surrounding geographical regions (Figure 2).

Yet the anthropological community has become increasingly reluctant to recognize Melanesia as a region characterized by biologically related groups. For instance, Spriggs (1997:1) proposed that Melanesians are defined not so much by shared genetic inheritance than by their habitation of a common geographical region; and Green (2001:113, my italics) declares Melanesia a meaningless concept:

2 Eight blood protein alleles occur at higher frequencies in population groups from Melanesia than those from Polynesia, Micronesia or Island Southeast Asia: Esterase D, ESD*1; Duffy blood group, FY*A; Glutamic-Pyruvate Transaminase, GPT*1; Human Leukocyte Antigen B, HLAB*13; MNS blood group, MNS*N; Transferrin, TF*D; and Rhesus blood group, RH*CDe and RH*cde (see summary in Cox 2003, published data in Mourant, Kopeć and Domaniwska-Sobczak 1976, Cavalli-Sforza, Menozzi and Piazza 1994).
...there is no biologically well-marked population that can be termed “average” Melanesian, even if deemed to be of diverse and polyphyletic origin... “Melanesian” defines an unacceptable biological category.

Instead, many researchers prefer alternative conceptions of the Indo-Pacific region, which differ radically from d’Urville’s tetrapartite classification. For instance, some view biogeographical boundaries – primarily developed to explain distributions of plants and animals (Cox 2001) – as barriers to human mobility and interaction. To the west of Melanesia, Oppenheimer and Richards (2002:295) consider Wallace’s biogeographical line (Figure 1) as defining “a clear discontinuity in both maternal and paternal [human] lineages. The principal markers defining the recent Polynesian expansions are all derived from east of this line....” Mourant, Kopéć and Domaniewska-Sobczak (1976), and Handoko et al. (2001), similarly view Wallace’s Line as a barrier to gene flow. Conversely, to the east of Melanesia,
Green (1991a) has proposed the sizable sea gap between the main chain of the Solomon Islands and the Reefs-Santa Cruz group as an important line defining two regions known as Near and Remote Oceania. This biogeographical boundary was a likely barrier to human movements further eastward until technological advancements during the late Holocene allowed first settlement of the remote Pacific.

Yet other researchers consider that biogeographical boundaries played little role in human prehistory, viewing the Indo-Pacific as a largely panmictic region tempered by voyaging corridors along which human mobility has ebbed and flowed (Terrell 2004). This view seems similar to modern population genetic models showing that human variation often grades slowly from one region to the next (Serre and Pääbo 2004). Applying this view to the Indo-Pacific region, the major determinant of genetic similarity between any two individuals should simply reflect their geographic proximity (but see caveats regarding this model by Lum and Cann 1998).

All these arguments revolve in part around the existence (or otherwise) of genetic clusters in the Indo-Pacific region, which would require some form of migrationist explanation. One divide between the biological sciences and anthropology relates to the ability to detect migration. Directional human mobility is clear in biological data\(^3\) and can readily be approached using well-established analytical frameworks. However, in anthropology, the “apparent unpredictability and the difficulty of identifying [migration] archeologically combine to make migration an explanatory construct of limited utility” (Anthony 1990). Yet because the movements of people are structured processes with predictable outcomes, they are easily amenable to biological, if not archaeological analysis (see the contact archaeology of Zilhão 2001 for an elegant exception). However, although determination of causal factors remains a difficult (and contentious) task, migration need not be seen – as it sometimes is – as a chaotic, isotropic process with indefinable outcomes (see well-defined anthropological model systems described by Anthony 1990).

Although Melanesia is now seldom treated as an informative analytical entity for anthropological research (but see D’Arcy 2003:218), there is little consensus regarding alternative population models (cf. Oppenheimer 1999, Bellwood 2005). Taking a model of simple clines of genetic variation across the region as a null hypothesis, this paper considers the genetic environment around Melanesia. By examining a wide assemblage of genetic data from both sex-specific and biparentally-inherited loci, I show that there is little support for simple clines of genetic variation across the Indo-Pacific region. However, there are broad-scale geographical patterns, including at least one genetic discontinuity. These ‘breaks,’ which divide genetically more homogeneous population clusters, do not occur at known biogeographical boundaries, such as Wallace’s Line or the Near/Remote Oceania division. Consequently, I suggest a re-evaluation of what biogeographical zones and boundaries may be acting to moderate the spread of human populations. While I am certainly not promoting

\(^3\) Discovery of the same genetic haplotype in two geographical locations presupposes either recurrent mutation or at least one migration event between the two localities. Given multiple genetic markers (e.g., linked Y-chromosome microsatellites) and extremely low mutation rates (e.g., on the order of \(2\times10^{-7}\) per generation; Gusmão et al. 2005), the likelihood of recurrent mutation creating exactly the same genetic haplotype becomes vanishingly small. (This basic principle underpins modern forensic genetics). Consequently, migration is frequently the only feasible explanation for identical, but geographically dispersed, genetic variants.
the resurrection of d’Urville’s simplistic classification system for what is an incredibly complex region (for instance, see Szabó and O’Connor 2004), re-examination of biological data from a ‘Melanesian’ perspective raises interesting new questions about the prehistoric processes that underlie human contact and mobility across this part of the Indo-Pacific region.

**Biogeography and Sociogeography**

Early in the twentieth century, most researchers believed in the existence of homogeneous, culturally distinctive and diachronically stable races of people (cf. the ‘Kava People’ of Rivers 1914, vol 2:240). Modern anthropology has correctly demolished any concept of ‘static races,’ replacing them with cultures that change over time, make innovations and interact with surrounding groups of people (Green 1991b). Human relations are therefore increasingly viewed as the sharing of sociogeographical regions, in which “aggregate[s] of people [have] more essential features in common and closer ties *inter se* than they have with groups in surrounding regions” (Read 1954:42, author’s italics). This important concept is central to modern population genetic research, where theoretical models often assume, as a preliminary thesis, that the genetic similarity of populations increases with geographic propinquity, and that populations need not have geographical or temporal permanence (Hey and Machado 2003). Such a dynamic definition of human society emphasizes interregional contact and interaction – with, of course, the important exception of population groups separated by natural geographical barriers, such as the Himalayas or the Saharan desert (Rosenberg *et al.* 2002, 2005). It is possible that two natural boundaries in the Indo-Pacific region may sometimes have taken this role: Wallace’s biogeographical line and the Near/Remote Oceania division (Figure 1).

Alfred Wallace defined his eponymous biogeographical line in the mid-nineteenth century. He based the concept on a natural division between Oriental and Australo-Papuan biota that he observed during his travels through the Indonesian Archipelago (Wallace 1869). Wallace’s biogeographical line marks an ocean boundary that has divided the Asian mainland permanently from the Pacific world since at least the Pliocene (Bird *et al.* 2005). Animals crossed this boundary only with great difficulty, and some researchers consider it a likely barrier zone for human mobility as well (Mourant, Kopeć and Domaniewska-Sobczak 1976, Handoko *et al.* 2001, Oppenheimer and Richards 2002).

The Near/Remote Oceania boundary divides the main Solomon Islands Archipelago from the Reefs-Santa Cruz group: sea distances increase appreciably between islands, the next island cannot be seen from the last, and the adoption of new voyaging practices was probably required for people to advance further eastward (Irwin 1992). Green (1991a) most clearly described the importance of this division in his landmark paper: “Near and Remote Oceania – Disestablishing ‘Melanesia’ in Culture History.” The Near/Remote Oceania boundary provides a terminus to the natural distribution of terrestrial mammals, and it likely proved a considerable barrier to the movement of people in their colonization of the greater Pacific. Although modern humans reached as far as the northern Solomon Islands at least 28,000 years ago (Wickler and Spriggs 1988, O’Connell and Allen 2004), they did not traverse the Near/Remote Oceania boundary until the late Holocene (Bedford, Spriggs and Wilson 1998).
The first settlement of Remote Oceania is defined by the Lapita cultural complex, characterized in particular by a distinct ceramic tradition, stone adze kit and shell ornaments, together with the first evidence for a full agricultural economy (Spriggs 1997:88 ff.). While the origin of the Lapita peoples is still hotly debated (Oppenheimer 1999, Terrell, Kelly and Rainbird 2001, Bellwood 2005, and see following sections), novel technological advances were almost certainly adopted or developed in order to breach the Near/Remote Oceania divide (Spriggs 1997:41 ff.), thus leading to the settlement of the greater Pacific region from around 3,200 years before present (BP).

**Disputed Histories of Settlement**

There have been many competing models for the settlement of the Indo-Pacific region, some of which have more biological support than others. For instance, Heyerdahl (1952) proposed that Remote Oceanic communities have an American origin, but the small number of Native American genetic markers in Polynesia today (Ohkura *et al.* 1999, Hurles *et al.* 2003) almost certainly result from population movements following European contact. Instead, most anthropologists and population geneticists believe that Melanesia and the wider Pacific region were settled from the west. This may be the extent of their agreement.

Views on Pacific settlement tend to fall towards one of two opposing camps, which I review here only briefly. (See citations for additional readings; Jobling, Hurles and Tyler-Smith 2004:354-370 present an especially readable argument from an alternative viewpoint to mine).

The first model discussed here (Bellwood 1978, 1997, 2005) proposes that Australo-Melanesian populations settled Near Oceania from 50,000 ± 10,000 BP (O’Connell and Allen 2004). During the mid-Holocene, peoples from Taiwan, but with ultimate connections to Mainland Asia, underwent an expansion driven by their adoption of an agricultural economy (Bellwood 2002, Diamond and Bellwood 2003). Their descendents moved through Island Southeast Asia, interacting with and adopting from indigenous populations they encountered along the way (*cf.* the intrusion/innovation/integration model of Green 1991b). A common misconception is that this demographic dispersal all but replaced pre-existing communities in the Indo-Pacific (Oppenheimer 2004), although biological data have long refuted this assertion (see the collection of papers in Hill and Serjeantson 1989). During a relatively short timeframe (Spriggs 2000, 2003), these admixed population groups moved along the northern coast of New Guinea and into the Bismarck Archipelago, where their arrival is postulated to have caused the sudden appearance of the Lapita cultural complex. With Neolithic technological advances allowing them to breach the Near/Remote Oceania boundary (Bedford, Spriggs and Wilson 1998), these partially admixed groups spread rapidly from the Solomons Archipelago to Vanuatu, ultimately settling the remote Pacific and becoming the ancestors of modern Polynesians (Cox *et al.* 2007). This model is bolstered by linguistic data, which places the basal lineages of the Austronesian language tree in Taiwan (Blust 1995, 1996, Pawley 2002); and it is supported by some genetic evidence, which has identified markers with clear Asian connections along coastal New Guinea (but not the New Guinea
The second model discussed here (Oppenheimer 1999, Oppenheimer and Richards 2002) also accepts that Near Oceania was settled around 50,000 ± 10,000 BP (O’Connell and Allen 2004), but postulates that subsequent developments in the Indo-Pacific region did not result from external influences (genetic, linguistic or cultural). Under this model, indigenous populations from the island-dominated region of East Indonesia (i.e., Wallacea) underwent an expansion during the late Pleistocene/early Holocene (Richards, Oppenheimer and Sykes 1998, but see criticism based on new data by Cox 2005, Trejaut et al. 2005, Pierson et al. 2006), radiating northwest into Island Southeast Asia and east along coastal New Guinea towards Island Melanesia (Oppenheimer 1999, Oppenheimer and Richards 2002). Under this model, it was the descendents of people from modern eastern Indonesian populations who ultimately settled remote Oceania. Other researchers have incorporated a voyaging corridor along northern New Guinea into this model, postulating periods of punctuated mobility driven by a complex interaction of climatic and cultural events (Terrell 2004).

Both settlement models make specific predictions about the distribution of biological characteristics in the Indo-Pacific today. From a simplistic genetic perspective, the presence of Asian markers in Melanesia and the wider Pacific would best seem to fit a partial Asian expansion (our first model), whereas the near exclusive distribution of Melanesian lineages would best fit an eastern Indonesian or Melanesian expansion (our second model). Furthermore, any model must explain all the available data – including that from fields beyond biology (Hunley et al. 2007). This author (Cox 2003, 2005, 2006a, Cox and Lahr 2006, Cox et al. 2007, Lansing et al. 2007, Downey et al. 2008) believes that the totality of genetic, linguistic and archaeological evidence favors a partial Asian settlement model (i.e., genetic evidence reviewed by Cann and Lum 2004 and later sections, linguistic evidence reviewed by Bayard 1996, archaeological evidence reviewed by Bellwood 1997, Spriggs 1997, 2000). However, regardless of the settlement model adopted, contemporary molecular research indicates that there are different ‘sets’ of genetically distinguishable markers among Indo-Pacific peoples today, which must to some extent reflect different geographical origins. Unfortunately, dating genetic lineages and determining geographical ‘origins’ are not the strongest facets of modern molecular research (Cox 2005). Conversely, the field of human population genetics has well developed and powerful statistical tools to analyze demographic processes among and between populations. The results of these methods are not always considered when evaluating models of Pacific settlement, but they can provide new insight into the biological prehistory of the Indo-Pacific region (Cox and Lahr 2006). The demographic basis of genetic marker distributions forms a major focus of discussion here.

As an important aside, records of prehistoric processes are invariably distorted by the subsequent shared history of populations (Bellwood 2001). Major genetic distributions are necessarily the product of major demographic developments, but the possibility of confounding effects from recent or local demographic processes should always be kept in mind (for instance, see Davidson 1978 and Campbell 1995 for the effects of an interaction sphere centred on Fiji, Tonga and Samoa). Recent history can have big impacts on the genetic profiles of populations, but does so primarily when they involve directional movements of very large numbers of people (for instance, the African slave trade to Brazil, Alves-Silva et
al. 2000, Carvalho-Silva et al. 2001; or the Moorish settlement of Iberia, Corte-Real et al. 1996, Bosch et al. 2001). Genetic profiles can also vary radically within modern societal structures with no prehistoric counterpart (for instance, large urban centres). Genetic studies consciously adopt sampling strategies that avoid these obvious confounds, but not all problems of sampling can be similarly avoided.

One prevalent view is that observed genetic discontinuities, as opposed to clinal zones, are the product of population processes during the Neolithic (Bellwood 1997:71). However, the historic period has also seen increasing encroachment of external social and demographic influences on peoples in the Indo-Pacific region. Indonesia, with its close geographical ties to the Asian mainland, provides the best evidence for long-term contacts, including China (Zainu’ddin 1974:27-28), India (Wheatley 1961:181-182, Çoedes and Damais 1992), the Middle East (Tibbetts 1979:27-28, 206) and even Europe (Casson 1989). However, such contact was restricted largely to the far west of Island Southeast Asia with little impact further east. Genetic diversity on the island of Bali in central Indonesia indicates traces of historical migration from India as early as the late first millennium BCE (Lansing et al. 2004, Karafet et al. 2005), but intermarriage “between people from different islands and districts” in the east of the Indonesian archipelago was uncommon as recently as the last fifty years (Zainu’ddin 1974:15). Marriage within ethnic groups and village communities was apparently the norm for much of the Indo-Pacific region well into its modern history and continues in some parts of the region today.

There is less evidence for major, directional demographic processes in the broader Pacific region during the historical era. Non-indigenous genetic markers are rare and sporadic in modern Oceania, although they have been documented for the Bismarck Archipelago (Scrimgeour 1983), Vanuatu (Cox 2003, 2006a) and Tokelau (Cheer, Allen and Huntsman 2000). The major demographic disruption of modern times was the too often repeated collapse of population sizes caused by introduced diseases, such as measles, during the late nineteenth century AD (Buxton 1926, Deacon 1934). The census population of some island groups fell by over 90% – for instance, from ca. 500,000 people in Vanuatu in 1830 to less than 45,000 in 1945 (Spriggs 1997:261). This trend has been strongly reversed during the second half of the twentieth century. The effect of this population collapse on the genetic profiles of Indo-Pacific populations has not been examined and is seldom considered. Importantly, however, population sizes were still large in absolute terms. While such demographic changes were devastating from social, cultural and moral perspectives, they may have had little effect on the genetic profiles of these populations. Size fluctuations – especially over only a few generations – do not usually affect allele frequencies or genetic diversity to any great extent (Futuyma 1997:135, Amos and Harwood 1998, Zenger et al. 2003, Ehrich and Jorde 2005), although small population sizes over longer timeframes has substantial genetic repurcussions (Cox 2006a). Although historical and modern demographic processes are recorded in the genetic profiles of modern populations, prehistoric demographic processes appear to be the major cause of broad-scale genetic patterns observed in the Indo-Pacific region today.
Recent Genetic Research

Human population genetic research over the last twenty years has been based primarily on two genetic systems: mitochondrial DNA (mtDNA) and the Y-chromosome. Mitochondrial DNA is a circular DNA strand ca. 16,500 base pairs long (Anderson et al. 1981), which exists as approximately a thousand copies per cell (Robin and Wong 1988). MtDNA has an almost exclusively maternal mode of inheritance, the sole known exception being found in association with a pathological condition (Schwartz and Vissing 2002, Kraytsberg et al. 2004). While noting the possible confounding factor of limited natural selection (Liu et al. 2005, Kivisild et al. 2006), mtDNA variation can inform us about the relationships and historical mobility of women. Conversely, the Y-chromosome consists of a linear DNA strand ca. sixty million base pairs long, which is carried as a single copy in male cells. Because the non-recombining portion of the Y-chromosome is inherited only paternally (Y Chromosome Consortium 2001), variation in its DNA sequence allows us to reconstruct the historical relationships and movements of men.

MtDNA and the Y-chromosome are more informative for many population genetic questions than the classical genetic markers studied previously. Firstly, many more character states are available for the mtDNA and Y-chromosome systems, and this allows us to infer past relationships in greater detail. Secondly, because most mtDNA and Y-chromosome markers have no phenotypic effect on their carriers, natural or sexual selection can have played little role in their dispersal (but see Kivisild et al. 2006 for evidence of purifying selection on the mtDNA locus). This ensures that the current distribution of mtDNA and Y-chromosome lineages was determined primarily by the actions of people over time (with the confounding effects of genetic drift). In comparison, many classical genetic markers have phenotypic effects that correlate strongly with environmental determinants (for instance, frequencies of many α-globin alleles vary closely with the extent of malarial endemicity; Flint et al. 1986). Thirdly, the mtDNA and Y-chromosome systems are inherited as a single effective copy via one parent, whereas autosomal markers (including most classical genetic systems) are inherited as two copies from both parents. The mtDNA and Y-chromosome records are therefore more sensitive to very recent demographic processes, and they provide

4 For instance, consider a segment of the mtDNA control region containing (b =) 400 base pairs. If we assume an infinite sites model (only one mutation per nucleotide position), a resulting genealogy can contain as many as 401 unique sequence states (n = b + 1). Relaxing this model to allow for multiple mutations (homoplasy) for the four chemical bases (adenine, cytosine, guanine, and thymine), as many as 6.7×10^{240} unique sequence states (n = 4^b) could (theoretically) be observed. (Importantly, mtDNA has a relatively high mutation rate. Conversely, four unique phenotypes (A, B, O, and AB) are possible with the ABO blood group system, which are formed by only three different alleles: A, B, and O.

5 Autosomal genes are inherited biparentally and carried as two copies per cell (diploid); mtDNA and the Y-chromosome are inherited uniparentally and carried as one effective copy per cell (haploid). This difference in the number and inheritance of these loci controls the number of gene copies present in any given population at any given time. It therefore affects the information content of these genetic systems because genetic diversity (θ) is a function of the effective population size (N_e), which is four times smaller for mtDNA and the Y-chromosome than for autosomal loci. For our purposes, drift occurs four times more rapidly in mtDNA and the Y-chromosome, and recent population demography is recorded more quickly by these haploid systems than by autosomal genomic regions (also see Harris and Hey 1999).
more informative records about very recent evolutionary history. However, these records are inevitably biased towards processes acting specifically on women or men (Hage and Marck 2003), and demographic reconstructions inferred from these systems have greater uncertainty due to higher levels of stochastic genetic drift (Hudson and Turelli 2003, Cox 2006a). Nevertheless, these unique characteristics of mtDNA and the Y-chromosome collectively make genetic data from the past twenty years more generally informative for historical reconstructions than classical genetic data from the sixty-five years before. Movements away from mtDNA and the Y-chromosome towards large-scale, multilocus autosomal resequencing are continuing this trend towards more informative genetic data (Garrigan et al. 2007).

Mitochondrial DNA – A history of women. Research during the 1990s identified three major mtDNA lineages that are carried by most Oceanic individuals (Hagelberg and Clegg 1993, Lum et al. 1994). These lineages – B4a, P, and Q – were found to vary in frequency across the Indo-Pacific region (Figure 3). An early study of mtDNA variation in Oceanic populations (Hertzberg et al. 1989) showed that lineage B4a (or rather, the more general lineage B form that these researchers screened) is absent from the New Guinea Highlands, but occurs at moderate frequencies in coastal New Guinea, Island Melanesia and Fiji (see also Merriwether et al. 1999). A specific variant called the ‘Polynesian motif’ was identified in all of the study’s Niuean, Samoan and New Zealand Maori samples.

Stoneking and Wilson (1989) confirmed the presence of B4a along coastal New Guinea, and its complete absence in the New Guinea Highlands. With the possible exception of one Australian individual, the lineage is also absent from West New Guinea (Tommaseo-Ponzetta et al. 2002) and Australia (Betty et al. 1996). B4a has since become associated with the proposed spread of Asian-derived Austronesian-speaking migrants into Oceania during the mid-Holocene for three main reasons: its high frequencies in Polynesia (the eastern tail end of the Austronesian expansion), its close association with the distribution of the Austronesian language family (including Madagascar; Cox 2003, Hurles et al. 2005), and particularly, its absence from the New Guinea Highlands where non-Austronesian (or Papuan) languages alone are spoken.

There is increasing evidence for substantial mitochondrial DNA diversity in the Indo-Pacific, particularly in western and central Island Southeast Asia (Hill et al. 2007) and the Bismarck Archipelago (Friedlaender et al. 2007b). However, two mtDNA lineages predominate near New Guinea. Q and P are common in the New Guinea Highlands (Redd and Stoneking 1999), have never been detected on mainland Asia, and are not closely related to other lineages there (Kivisild et al. 2002). Because Q and P are restricted geographically to the traditional boundaries of Melanesia (with minor exceptions detailed below) and have no close relationships to extant Asian lineages (Friedlaender et al. 2005), Forster et al. (2001) proposed that Q and P probably evolved in situ in the vicinity of Melanesia. Genetic dating indicates that these lineages may have been isolated in the New Guinea region for upwards of 40,000 years, and the most parsimonious explanation remains that Q and P descend from genetic lineages carried by Pleistocene populations in Melanesia and eastern Island Southeast Asia (Forster et al. 2001, Cox 2003, Friedlaender et al. 2005). (Friedlaender et al. 2007 note that two new lineages, M28 and M29, may be conflated with lineage Q in earlier studies. However, M28 and M29 screening currently has insufficient geographical coverage to make use of these new lineages here).
Figure 3. Distribution of three major mtDNA lineages in the Indo-Pacific region. Lineage B4a, P and Q frequencies from Sykes et al. (1995), Betty et al. (1996), Lum et al. (1998), Hagelberg et al. (1999), Kivisild et al. (2002), Tommaseo-Ponzetta et al. (2002), Cox (2003) and Tajima et al. (2003).

Although individuals with lineage Q have been identified beyond Melanesia—in the Philippines (n = 2), Marshall Islands (n = 1) and sporadically across Polynesia (Sykes et al. 1995, Hurles et al. 2003)—the marker never exceeds low frequencies in these areas. Furthermore, the distributional patterns of B4a versus P and Q differ statistically between Coastal and Highland New Guinea populations, and between population groups with Austronesian and non-Austronesian linguistic affiliations (Stoneking et al. 1990). Interestingly, Giles, Ogan and Steinberg (1965) identified a similar association with γ-immunoglobulin (GM) variation, thus stimulating Terrell and Fagan (1975) to begin this very debate about the utility of d’Urville’s classification system.

Y-chromosome — A history of men. Our knowledge of Y-chromosome variation stems largely from the last ten years (Underhill et al. 1997), during which the young field of Y-chromosome analysis has changed rapidly: many papers’ conclusions were outdated months after they were published (although their data remain sound), and several systems of nomenclature coupled with the rapid discovery of new markers made for a fluid and confusing field. More recently, a universal nomenclature system has been adopted (Y-Chromosome Consortium 2001, Jobling and Tyler-Smith 2003), and the suite of markers generally typed in Indo-Pacific samples (Figure 4) has been an increasingly informative panel (Cox 2006b). Most Indo-Pacific Y-chromosomes can be assigned to one of five lineages: C2, K*, K5, M or O3. One of these, K*, is a Y-chromosome paragroup; a genetic grouping that is not defined by a specific marker, but by the absence of mutations for other lineages. This situation likely reflects the ascertainment bias inherent in Y-chromosome lineage detection (for a related Oceanic case, see Cox et al. 2007) rather than capturing any clear signal of
prehistoric demography. Although $K^*$ is frequent in Oceania, its paragroup status limits interpretation of its distribution, and it is not discussed further here.

Figure 4. Distribution of Y-chromosome haplogroups in the Indo-Pacific region (summarized by Cox and Lahr 2006). Abbreviations: AUS, Australia (Aborigines); BRN, Northern Borneo; CHN, China; CKI, Cook Islands; FIJ, Fiji; FPL, French Polynesia; JAV, Java; KOR, Korea; MLY, Malaysia; MOL, Moluccas; MSI, Malaita Province of Solomon Islands; NBR, New Britain; NIR, New Ireland; NTG, Nusa Tenggara; PHL, Philippines; PNG, Papua New Guinea; SBR, Southern Borneo; SLW, Sulawesi; TON, Tonga; TRB, Trobriand Islands; TWN, Taiwan (Aborigines); VAN, Vanuatu; VTN, Viet Nam; WNG, West New Guinea; and WSM, Western Samoa.

O3 occurs at variable frequencies in Polynesia, but is uncommon in Island Melanesia. Except for one individual, O3 is also absent from genetic samples collected in the West and East New Guinea Highlands, and Australia (Kayser et al. 2001b, 2003, Mona et al. 2007). The widespread distribution of O3 throughout mainland Asia (Su et al. 1999b), a demographic expansion of O3 individuals dated using genetic techniques to the mid Holocene (Kayser et al. 2000, 2001a), and the marker’s almost complete absence from the New Guinea Highlands, have led researchers to associate lineage O3 tentatively with the expansion of Asian-derived, Austronesian-speaking peoples into Island Southeast Asia and Oceania during the mid-Holocene (Kayser et al. 2000, 2001a, Capelli et al. 2001, Underhill et al. 2001).

The three remaining markers follow different distributions, but none have been found on the Asian mainland. Lineage M is common within the traditional boundaries of Melanesia, but occurs only infrequently in central Indonesia and Polynesia (Karafet et al. 2005). Lineage M is also found on the islands immediately west of New Guinea, i.e., the Maluku islands and Timor (Kayser et al. 2001b). To the east, Su et al. (2000) detected lineage M in every one of their Nasiol sample (Bougainville Island), although the marker is less frequent in the Tolais of New Britain, the East New Guinea Highlands and along the New Guinea coast (Kayser et al. 2001b). Eight percent of Tongan individuals represent the only Polynesians identified as carriers of lineage M to date.
Lineage K5 shows a similarly clear Melanesian connection. The marker attains its highest frequency in the East New Guinea Highlands (exceeding 50%), but reaches only lower frequencies in coastal and lowland New Guinea, East Indonesia, New Britain, Solomon Islands and Vanuatu (Cox 2003, Kayser et al. 2003, Cox and Lahr 2006). K5 has not been detected in Polynesia or the Asian mainland, and its distribution is restricted to eastern Indonesia and Island Melanesia.

Lineage C2 presents a more ambiguous picture (Cox et al. 2007). It occurs sporadically in East Indonesia, where the mutation may have arisen (Underhill et al. 2001), but has not been found on the Asian mainland. C2 – defined either by the presence of the point mutation M38, or the closely associated microsatellite deletion, DYS390.3del (Kayser et al. 2003) – is frequent in Melanesia (Kayser et al. 2003), and descendent lineages (such as C-P33) reach moderate to high frequency in Polynesia; from 27% in Tongans (Hurles et al. 2002) to 82% in Cook Islanders (Kayser et al. 2000, 2003). The presumed origin of this marker either in East Indonesia (Oppenheimer and Richards 2002) or Melanesia (Kayser et al. 2000), coupled with its prevalence in Polynesia, have been taken as support for an origin of Remote Oceanic communities from East Indonesia (Oppenheimer and Richards 2002). However, this model fails to explain the additional presence of Asian lineage O3 at moderate frequencies in the Pacific, as well as the almost complete absence of lineages M and K5 in Polynesia despite high frequencies in East Indonesia and Melanesia.

These large differences in Y-chromosome allele frequencies throughout Oceania are almost certainly due in part to genetic drift resulting from population bottlenecks, founder events and small population sizes (Cox and Lahr 2006), and possibly to sex-specific patterns of mobility (Hage and Marck 2003). A more complete explanation for the distribution of Y-chromosome lineages in the Oceanic world (Figure 4) suggests a model whereby Austronesian-speaking populations carrying haplogroup O from mainland Asia absorbed a large proportion of indigenous individuals with C Y-chromosomes as they passed by New Guinea (cf. the cultural intrusion/innovation/integration model of Green 1991b). This scenario helps explain the change of Y-chromosome lineage variants observed across Island Southeast Asia (Kayser et al. 2000). Statistical analyses of Y-chromosome diversity indicate that an admixture model is necessary to explain the observed data (Cox and Lahr 2006, Lansing et al. 2007), whereas models that invoke a completely indigenous Melanesian or East Indonesian origin do not provide good matches to the genetic data.

It should be emphasized that much more detailed patterns of mitochondrial DNA and Y-chromosome diversity are apparent at smaller scales than are discussed here. Important advances are currently being made on community-level diversity, especially for Island Melanesia (e.g., Scheinfeldt et al. 2006, Friedlaender et al. 2007a, b) and eastern Indonesia (Lansing et al. 2007). However, haploid genetic diversity at small scales is very susceptible to the stochastic effects of genetic drift (Cox 2006a), and a strong case can be made that viewing the Oceanic world from a broader perspective is more useful for detecting general geographical trends in genetic data. This is the approach taken in following sections.

It is also worth noting that mtDNA and the Y-chromosome are currently the only two genetic systems with sufficiently dense geographical sampling to address questions related to broad demographic processes in the Indo-Pacific region. However, considerable research from autosomal and obligate parasitic genetic systems also support some distinction between
Melanesia and surrounding regions (although these may be affected to an unknown extent by natural selection), and some autosomal loci also seem to fit a model of recent movements of Asian derived peoples into the New Guinea region with subsequent contact leading to the partially admixed populations of today. Although seldom argued from this perspective and lacking any statistical confirmation, aspects of the patterns recorded by mtDNA and the Y-chromosome are also evident in molecular data from ABO blood group genes (Ohashi et al. 2004), the albumin-vitamin D-binding protein gene (Chen et al. 1990), the α(1,2)-fucosyltransferase gene (Chang et al. 2002), the band 3 gene involved in Southeast Asian/Melanesian ovalocytosis (Kimura et al. 1998, 2003), the immunoglobulin λ genomic region (Robledo et al. 2003), the HIV-1 resistance gene SDF1 (Su et al. 1999a, Kimura et al. 2002), the Human Leukocyte Antigen system (Yoshida et al. 1995, Bugawan et al. 1999, Zimdahl et al. 1999, Main et al. 2001), and autosomal non-genic STR variation (Parra et al. 1999). Similar demographic models have been adopted to explain the phylogeographic distribution of Indo-Pacific variation in obligate human viruses: human T-lymphotropic virus type 1 (Gessian et al. 1991, 1993, Nerurkar et al. 1993, 1994a, 1994b, Yanagihara et al. 1995), hepatitis B (Mulyanto et al. 1997), and human polyomavirus JC (Chang et al. 1999, Yanagihara et al. 2002, Miranda et al. 2003). Consequently, I expect that the models discussed here from mtDNA and Y-chromosome evidence likely have a much broader genetic foundation.

**Clines and Clusters**

The mtDNA and Y-chromosome distribution maps presented in Figures 3 and 4 are necessarily broad-scale, and are intended only to depict major genetic trends in the Indo-Pacific region for population samples with full genetic typing (thus excluding otherwise applicable studies like Redd et al. 1995 and Souto et al. 2005). Clearly, more detailed genetic and community-level analyses reveal smaller-scale patterns in these lineages (Friedlaender et al. 2005, Merriwether et al. 2005, Lansing et al. 2007, Wilder and Hammer 2007), but this does not invalidate the larger picture. Different patterns are observable at different scales (Levin 1992, D’Arcy 2003), and it is a valid exercise to view genetic, linguistic and anthropological evidence from a broad Indo-Pacific perspective. Here, the mitochondrial and Y-chromosome maps show that some genetic markers – most notably mtDNA lineages P and Q, and Y-chromosome lineage M – occur frequently among peoples inhabiting the region traditionally defined as Melanesia. Similar patterns are reflected in the frequency maps of Kayser et al. (2006, Figure 1). These markers intergrade with surrounding populations to some degree; for instance, carriers of mtDNA lineage Q occur sporadically in the Philippines and Marshall Islands. However, these three Oceanic markers have not been detected even as far west as the Asian mainland, and they certainly occur nowhere else in the world. With the vast amount of genetic data now available for a representative range of global human populations, these distributions can no longer be viewed simply as an artifact of genetic sampling strategies. Some lineages have a strong connection to the geographical region traditionally defined as Melanesia.
Importantly, genetic diversity does not cline smoothly across this region. Taking Y-chromosome variation as an exemplar, the genetic diversity of populations does not change linearly in relation to the geographical distances between them, as would be expected if these populations had been interacting with each other over a two-dimensional landscape for a moderate period of time (Rousset 1997). In Figure 5, the leftmost pane shows the association between genetic distance ($F_{ST}$) and geographical distance westward of New Guinea; the rightmost pane shows the same relationship eastward of New Guinea. These figures illustrate near textbook examples of non-linearity, where the data best fit a nonlinear regression (solid lines). To further confirm the poor fit of a linear model to these data, a linear fit can be computed for both datasets (dotted lines), and compared with the fit to one example of a curvilinear relationship (West: linear $r^2 = 0.34$, quadratic $r^2 = 0.50$; East: linear $r^2 = 0.21$, quadratic $r^2 = 0.24$). While a higher order polynomial function probably does not provide the best theoretical fit with the data (which likely depends on many unknown underlying factors), the quadratic models provide far more adequate matches than do the linear models (West: linear $F_{1, 103} = 29.5, P << 0.001$; East: $F_{2, 102} = 17.3, P << 0.001$). Indeed, the fit of the linear model is so inaccurate that any suggestion of a linear association between geographical distance and genetic distance can be rejected, especially so for the geographical region to the west of New Guinea. (Note, however, that isolation-by-distance may be reasonable at much smaller scales; e.g., Kirk 1992:178 ff.).

These patterns may fit a multifactorial threshold model, in which many underlying cultural, demographic and environmental causes interacted until their combined effect increased to the point where easy mobility from one geographical region to another was prevented. Non-equilibrium migration models, where rates of mobility increased rapidly at a specific time (such as a transition to agriculture followed by demographic dispersal) can also produce non-linear genetic/geographic distance plots like those observed here (Wilkins and Marlowe 2006). Alternately, these graphs may record the signal of a demographic expansion funneled along an essentially one-dimensional migration route (see Figure 4 in Rousset 1997). Similar patterns are observed when isolation-by-distance occurs over a strongly elongated spatial range (i.e., one-dimensional habitat ranges, such as populations connected along a river). For the geographical region west of New Guinea, this explanation would require that population contact was channeled through an interaction zone on the order of 10-

---

6 Analyses use Y-chromosome data from Indo-Pacific populations ($n = 21$) summarized by Cox and Lahr (2006). The West of Melanesia dataset includes population samples from Taiwanese aboriginals, the Philippines, Java, Southern Borneo, the Maluku Islands, Nusa Tenggara, Lowland and Highland West New Guinea, and Coastal and Highland Papua New Guinea; the East of Melanesia dataset includes population samples from Lowland and Highland West New Guinea, Coastal and Highland Papua New Guinea, the Trobriand Islands, New Britain, New Ireland, Western Province of Solomon Islands, Malaita Province of Solomon Islands, Vanuatu, Fiji, Tonga, Western Samoa, French Polynesia and the Cook Islands. Geographical distances between sampling locations were $\ln$ transformed to compensate for two-dimensional spatial analysis, and pairwise $F_{ST}$ values (i.e., genetic distances between population samples) were transformed by $F_{ST}/(1-F_{ST})$ to generate values varying from 0 to $\infty$ in direct comparison with geographical distance (Rousset 1997). The statistical package R (R Project 2005) was used to generate figures, evaluate the fit of the data to linear and quadratic regression models, and determine statistical significance via an analysis of variance (ANOVA) approach.

7 It is worth noting that increasing deviation from a perfect two-dimensional grid of populations, as might be expected for populations sampled randomly from an island setting, would increase the variance of sample points about the local regression. However, this regression should still be approximately linear. A strong non-linear regression, as shown here, necessitates alternative causes.
km wide by 1000-km long. Although a recent and strongly directional Austronesian expansion might explain aspects of this pattern, either through increased rates of migration as agricultural populations spread, or through the use of a narrow migration corridor, other factors (cultural, demographic or environmental) may be equally reasonable alternatives.

Figure 5. Nonlinear associations between Y-chromosome haplogroup distances ($F_{ST}$) and geographical distances for Indo-Pacific populations to the west (left) and east (right) of New Guinea. Local nonlinear regressions (solid curves) provide better fits than linear alternatives (dotted lines).

Unfortunately, none of these models can readily be tested without far more detailed spatial sampling of genetic data and a much better knowledge of underlying demographic parameters (e.g., prehistoric hunter/gatherer carrying capacities in Island Southeast Asia). Such data are not currently available. Yet the view that Y-chromosome variability grades smoothly between New Guinea and the populations of Island Southeast Asia to the west or the populations of Near/Remote Oceania to the east (i.e., simple genetic isolation-by-distance) can be rejected with high statistical confidence. A nonlinear trend in the genetic data relative to geographical distance (i.e., a discontinuity that separates two more internally distinctive genetic clusters) is a significantly better model for this region. A nonlinear trend also holds for the region from Melanesia out into the Pacific, but is much less strongly pronounced. With some provisos, a close association between genetic similarity and geographic propinquity may provide a passable proxy for island communities eastward of New Guinea.

Of course, the observance of this genetic discontinuity should not be taken to imply that genetic marker distributions have perfectly discrete boundaries. Genetic variation is commonly conceived as an array of overlapping clines of individual genetic markers (Serre and Pääbo 2004), where similarity between individuals is dictated more by geographic proximity than other demographic factors. Contemporary genetics usually assumes spatial interaction of populations, and therefore, clinal variation in genetic diversity. However, global clines are sometimes fractured into genetic clusters at discontinuities – places where genetic
frequencies change rapidly over relatively short geographical distances (Rosenberg et al. 2002, 2005). Physical barriers (e.g., the Himalayas or the Saharan desert) can often explain such discontinuities as obvious barriers to human mobility, but where simple geological reasons are insufficient (as it seems here), discontinuities (and their corresponding genetic clusters) may signal the action of important demographic or selective factors during prehistory. Furthermore, similar trends at multiple genetic loci require a demographic explanation, because natural and sexual selection act on individual genetic loci (Bamshad and Wooding 2003). Importantly, relatively discrete genetic clusters cannot be explained by demographic models that emphasize undirected and stochastic (i.e., isotropic) mobility of individuals (Terrell 2004). Evidence of genetic clusters in the Indo-Pacific points strongly towards directional (i.e., anisotropic) demographic models as their underlying cause.

Unfortunately, the magnitude of genetic discontinuities and the amount of population admixture observed today says little about the timeframe over which processes causing these patterns may have acted. Kayser et al. (2000) showed high rates of admixture in the vicinity of New Guinea, and consequently proposed a ‘slow boat’ settlement scenario, in which migration past New Guinea slowed and contact between communities increased. However, admixture rates need not be related directly to time. Highly admixed populations in East Indonesia, New Guinea and Island Melanesia (Cox and Lahr 2006, Ohashi et al. 2006) may result from a slowing of the migration rate, which led to population contact over a longer period of time; or a constant migration rate, coupled with a cultural acceleration of the contact/admixture process. Although genetic research is well suited to determining the proportion of admixture in contemporary Indo-Pacific populations, the rate of the spread of communities will remain an archaeological question unless genetic dating methods improve significantly. Currently, archaeological analyses of the Neolithic spread across the Indo-Pacific region (Spriggs 2000, 2003) favor uniformly fast rates of mobility, thereby suggesting that high admixture proportions around New Guinea may instead result from increased rates of population contact, or alternately, a longer period of population interaction (“shared history”) near New Guinea following initial contact (Cox and Lahr 2006).

It is perhaps also important to emphasize my emphasis on ‘discrete’ clusters of human genetic variation does not mark a return to race-based thinking (Rosenberg et al. 2005, but see concerns in Foster and Sharp 2004). Related studies that detect regional discontinuities in genomic DNA variation stress the overall similarity of humans from global and genomic perspectives. Most autosomal genetic variants occur all around the world (International HapMap Consortium 2005), and less than 8% of genetic variants, many of them extremely rare (Rosenberg et al. 2002), are restricted to broad geographical regions (e.g., Africa, the Middle East, India). The genetic characters considered here are ‘neutral’ genetic markers (i.e., they have no phenotypic effects). These markers just happened to mutate during, and therefore record, particular demographic processes, the odds of which are high with nearly four billion genomic characters to choose from. Genetic data clearly do not support historical views of race, but rather favor modern anthropological concepts of continuous interaction between people. Consequently, discussion by biologists about discontinuities in the genetic data should not be conflated with any concept of racially defined or “hermetically sealed” groups of people (see rebuttal of this criticism by Bellwood 2001:107, 2002). Instead,
geographic discontinuities in genetic relationships simply reflect barriers to gene flow that affect many, but not all, human participants within a particular geographical region.

Importantly, although mitochondrial DNA and the Y-chromosome lineages have relatively discrete distributions in the Indo-Pacific region, these do not match closely with either Wallace’s biogeographical boundary or the Near/Remote Oceania division (Figure 1). Their geographical spread does, however, coincide surprisingly well with the traditional western boundary of Melanesia. There is no similarly clear discontinuity in the east, where isolation-by-distance may provide a reasonable working hypothesis instead.

The Near-Remote Oceania Division

The poor association between haploid genetic data and the Near-Remote Oceania division probably results from relatively recent demographic processes. The distributions of many Melanesian lineages (e.g., mitochondrial haplogroups P and Q, Y-chromosome haplogroups C, M and K5) are shifted eastward relative to the Near/Remote Oceania division: these lineages remain common in Vanuatu and Fiji, but they drop in frequency towards Polynesia. Why does this distribution have an eastward skew?

The Near/Remote Oceania boundary is archaeologically well supported as a barrier to the eastward spread of human populations during the Pleistocene and early Holocene (Spriggs 1997:41 ff.). Technologically pre-Neolithic peoples have lived west of this line since the late Pleistocene (O’Connell and Allen 2004), but first settlement to the east dates only from ca. 3,200 BP (Bedford et al. 1998). The initial settlement must have involved communities with a large Asian-derived component in order to explain the frequencies of these lineages in the greater Pacific (although drift effects should not be underestimated; Cox 2006a). However, the high frequencies of Melanesian lineages in parts of Remote Oceania, particularly the Vanuatu Archipelago (Cox 2003), indicate that large numbers of individuals carrying these markers must have crossed the Near/Remote Oceania boundary, presumably from more northerly parts of Island Melanesia. An early spread of these Melanesian lineages is suggested by their dominance in Vanuatu today (Cox 2003, 2006a, Cox and Lahr 2006), and this is also consistent with Melanesian mtDNA lineages being detected in archaeological remains from late Lapita skeletons in Near and Remote Oceania (Hagelberg and Clegg 1993). Individuals from the early Lapita cemetery of Teouma on Éfaté, Vanuatu (M. Spriggs, pers. comm. 2006, Bentley et al. 2007) might well be expected to carry a near modern range of mtDNA and Y-chromosome lineages – including a high proportion of Melanesian markers. Nevertheless, the discrete distribution of some mtDNA and Y-chromosome lineages also supports later movements from northern Island Melanesia to at least the Reefs-Santa Cruz group (Friedlaender et al. 2002) and the northern parts of Vanuatu (Cox 2003:140). Melanesian lineages may have dispersed to Fiji at a comparatively late date as well (Campbell 1995:4-5).

The spread of Melanesian markers cannot have been an entirely indigenous development, because an indigenous process does not explain the presence of Asian mtDNA lineages (such as mtDNA haplogroup B4a) in Remote Oceania. A more parsimonious explanation is that this spread was facilitated by novel Neolithic technologies introduced by Lapita peoples with ultimate Asian ancestry spreading eastward from Island Southeast Asia (Spriggs 1997). If
colonists derived from Pleistocene-era populations in Melanesia took part in this population dispersal, the homogeneous material culture of the earliest settlements in Remote Oceania implies that they adopted their new cultural lifestyle quickly. The first settlers crossing into Remote Oceania were probably already partially admixed (for instance, see parallels in Devlin et al. 2001), probably in Island Melanesia rather than further west (Hill and Serjeantson 1989), with subsequent migratory contributions trending eastward in later centuries (Friedlaender et al. 2002, Cox 2003).

Following the initial settlement of Remote Oceania, there was considerable mobility between the Solomon and Vanuatu Archipelagos and Fiji. Spriggs (1997:154 ff.) presents evidence that these island chains later underwent trade contraction, adopted an ‘inward-looking’ approach, and demonstrably reduced their mobility near the end of the Lapita period around 2,000 BP. The rapid drop in Melanesian lineage frequencies near Fiji may simply represent an eastward limit to the drift of Melanesian populations at that time. Alternately, this ‘Melanesianisation’ may be relatively recent (Campbell 1995). Only these, or similarly complex scenarios, can account for the high frequency of Melanesian population markers in Vanuatu and Fiji – reaching nearly 100% in some populations eastward of the Near/Remote Oceania boundary (Cox 2003). The key point is that genetic profiles from Melanesia to Polynesia do not change particularly abruptly (Figure 5), perhaps due to significant population mobility during later periods. The Near/Remote Oceania division may be an incredibly useful concept for conceptualizing human demographic processes during the Pleistocene (and for explaining the distribution of some terrestrial animals; Austin 1999, Matisoo-Smith and Robins 2004), but it does not adequately reflect human diversity in this region today (Clark 2003b, D’Arcy 2003).

**Wallace’s Biogeographical Line**

The genetic data also produce a poor match to Wallace’s biogeographical line. The distributions of many Melanesian lineages are again shifted eastward from this line; they are uncommon on Sulawesi, but reach high frequencies in East Indonesia and West New Guinea. Furthermore, the genetics of this region do not suggest a smooth cline between New Guinea and Island Southeast Asia (Figure 5), but rather indicate a strong discontinuity, where the genetic profiles of individual populations change quickly over relatively short geographical distances.

Mourant, Kopeć and Domaniewska-Sobczak (1976:94 ff.) maintained that the MNS*M blood group allele drops rapidly from high to low frequency at Wallace’s biogeographical line, but their claim was based on few, widely-distributed samples, and has not been replicated independently. Similarly, Oppenheimer and Richards (2002: 295) believed that Wallace’s biogeographical line defined “a clear discontinuity in both maternal and paternal lineages [where] the principal markers defining the recent Polynesian expansions are all derived from east of this line…” However, although lineages found in Polynesia (such as the mtDNA ‘Polynesian Motif’) do occur east of Wallace’s line, in practice the discontinuity in these genetic distributions is actually located much farther east.
This is not surprising. Setting aside pre-modern humans (e.g., *Homo erectus, H. floresiensis*), for which there is no evidence of a haploid genetic contribution to anatomically modern humans (e.g., Capelli et al. 2001, Ke et al. 2001, Richards and Macaulay 2001), the Indo-Pacific region seems to have been settled in patchwork fashion from 50,000 ± 10,000 years ago. On the Asian side of Wallace’s biogeographical line, the earliest signs of modern humans occur at Niah Cave, Borneo just over 45,000 years ago (Kennedy 1977, Bellwood 1997:83-4, Barker 2005:94). Across Wallace’s biogeographical line in Wallacea and on the Sahul shelf, the most well supported dates fall in the range of 45,000-42,000 BP (reviewed by O’Connell and Allen 2004). Settlement dates up to 60,000 BP have been proposed (most famously for Lake Mungo in south-eastern Australia; Thorne et al. 1999), but remain contentious. In any case, an arrival of modern humans in Australia by 60,000 BP presupposes their prior movement through (at least parts of) Island Southeast Asia. Certainly at some stage around 60,000-40,000 years ago, the first groups of modern humans were making their way through Island Southeast Asia towards New Guinea and Australia.

As these peoples crossed Wallace’s biogeographical line, they would have encountered unfamiliar flora, and more especially, fauna (Cox 2001). Their survival required accommodation and adaptation to this new biota, but little factual evidence indicates that these environmental factors at Wallace’s biogeographical line hindered the spread of modern human populations for any length of time. The earliest dates for anatomically modern humans on both sides of Wallace’s line appear broadly contemporary (or even paradoxically younger in the east), and the effect of changing biodiversity across this boundary may not have been particularly demanding on mobile human populations. Even a relatively fast spread of hunter/gatherer groups (for instance, in the range of 6-10 km/year as suggested for the colonization of North America; Hazelwood and Steele 2004:677) would leave considerable time for encountering and adapting to new flora and fauna. The absence of a punctuated chronological gradient for the settlement of North America despite crossing a broad variety of habitat zones (for instance, see Steele, Adams and Sluckin 1998, Figures 1-4) suggests that a changing habitat and biota may have little effect on human mobility. Indeed, the biotic change that best mimics Wallace’s biogeographical boundary in terms of the extent of change in plant and animal types lies between North and South America (Cox 2001). Yet current estimates for the advance of humans through the Americas indicate that this division was crossed almost instantaneously from an archaeological perspective (Steele, Gamble and Sluckin 2000).

Despite its near mythical status, Wallace’s biogeographical line has probably never been a meaningful boundary in human terms. Indeed, recent research concludes that even an early hominin, *Homo floresiensis*, must have crossed this sea boundary (Brown et al. 2004; Morwood et al. 2004), a view that has historically been anathema. In spite of the elegant suggestion that this important biogeographical boundary should necessitate major changes in the genetic profiles of modern human populations, it was apparently not a major barrier to population spread in reality. From a genetic perspective, Wallace’s biogeographical line does not show any clear sign of being a meaningful division of humans today, or at any time in the genetically reconstructable past.
**Wallace’s Phenotypic Boundary**

However, the genetic data do fit another, little-known nineteenth century division through eastern Indonesia. Alfred Wallace developed his biogeographical line only as a boundary between Oriental and Australo-Papuan fauna; he never meant it as a significant division for humans. Instead, he defined a second line across Indonesia that classified indigenous groups on the basis of superficial physical criteria (Wallace 1869, vol. 2:439 ff.). Wallace (in his own terms) placed Malay-like peoples to the west of this phenotypic dividing line, and Papuan-like peoples to the east. Wallace’s phenotypic division has surprisingly accurate correspondence to the discontinuity observed in the genetic data (Figure 1). For instance, frequencies of mitochondrial haplogroups P and Q and Y-chromosome haplogroups M, C and K5 increase rapidly east of Wallace’s phenotypic line, but are uncommon to the west; Y-chromosome haplogroup O carriers are more frequent to the west, but drop rapidly to the east (Figures 3 and 4).

At a broad scale, this genetic distinction is not restricted solely to mtDNA and the Y-chromosome data. In one of the first studies to scan variation systematically across the human genome, Rosenberg et al. (2002) typed 1066 individuals from global populations for 377 microsatellites (small stretches of DNA that vary in length). Because this study is not limited to one or a few genetic loci, it should represent an unbiased and statistically robust estimate of human genomic (i.e., autosomal) variation (Rosenberg and Nordborg 2002, but see also comments by Serre and Pääbo 2004 and reply by Rosenberg et al. 2005). The genetic discontinuity between Melanesia and Southeast Asia does not appear to be an artifact of the mtDNA or Y-chromosome records, but is also reflected in genomic data (Rosenberg et al. 2002, Figures 1 and 2; Rosenberg et al. 2005, Figure 2). Rosenberg’s study cannot determine the exact location of this discontinuity due to the sampling limitations of their study, but it is not surprising to find the mtDNA or Y-chromosome discontinuity mirrored in genomic data. Rosenberg and colleagues’ (2002, 2005) markers are selectively neutral, but their study suggests that a large proportion of the genomic variation in each region is on average markedly dissimilar. Correspondingly, physical anthropologists researching human morphology have argued for a discontinuity in physical form between Melanesia and Southeast Asia since the early twentieth century (Coon 1966, Howells 1977, Pietrusewsky 1994). Because autosomal DNA contains the majority of human genes – the causative agents of physical variation, this suggests a likely link between the recent studies of broad-scale neutral genetic diversity and historical studies of physical variation.

Interestingly, Wallace’s phenotypic division also corresponds closely to a linguistic division (Figure 1). Austronesian languages of the Central Malayo-Polynesian (CMP) and South Halmahera-West New Guinea (SHWNG) groups occur only to the east of this line, whereas languages of the Western Malayo-Polynesian (WMP) group – effectively a residual category of non-CMP and non-SHWNG languages – are restricted to the west (Pawley 2002). Non-Austronesian (“Papuan”) languages also occur sporadically in East Indonesia (the Timor-Maluku region), and these too are only found eastward of Wallace’s phenotypic line (Downey et al. 2008). However, these non-Austronesian languages are often considered to be
late-Holocene introductions from New Guinea (Pawley 2002:267), and their current distribution may be misleading.

The presence of a genetic discontinuity in this geographical location seems difficult to reconcile with Oppenheimer’s (1999) model for the settlement of the Oceanic world (our second model). This postulates that a loss of land caused by rising sea levels following the Last Glacial Maximum (~18,000 BP) forced populations in southern Island Southeast Asia (particularly Wallacea) to move into surrounding regions, an exodus ultimately resulting in the first human settlement of the Remote Pacific. More recently, the climatic cause of this demographic spread has been downplayed (Oppenheimer 2003). Regardless of the cause, however, it is difficult to envisage how ancient, related population groups radiating from a demographic center in eastern Indonesia (e.g., Oppenheimer 2004, Figure 1) would be consistent with a major genetic discontinuity running through their presumed geographical origin. Conversely, the genetic discontinuity fits well with a model of human populations spreading from northwest Island Southeast Asia, and interacting with (and being impeded by) indigenous, genetically differentiated populations already present in the southeast (see similar suggestions on archaeological grounds by Spriggs 2003). Notably, this contact zone, reflected in a correlation between genetics and linguistics, is observable in the region today even at very small scales, such as on the small island of Sumba in eastern Indonesia (Lansing et al. 2007). This signal of a demographic discontinuity, apparent at both large and small scales, is therefore unlikely to be an artifact.

**Possible Causes of the Discontinuity**

Why does the genetic discontinuity exist? Put simply, there is no clear answer to this question. However, the basis of an answer must lie in the region’s prehistoric demography, and the discontinuity is most readily explained using a model of recent Asian movements into the Indo-Pacific region (Bellwood 2005, Lansing et al. 2007). Under this model, Wallace’s phenotypic division may mark the location where indigenous groups, present since the Pleistocene, were living in sufficiently large numbers to resist incursive populations spreading into the region during the mid-Holocene. Perhaps this demographic resistance was driven by an indigenous agricultural tradition related to that found in New Guinea (as documented by Denham 2005). Only the East New Guinea highlands have provided direct evidence for an autochthonous development of agriculture (Denham et al. 2003), but indigenous agriculture may have been practiced more widely in this region than we currently have archaeological evidence for. A case for this has recently been argued from a genetic perspective (Mona et al. 2007).

However, it seems more likely that Wallace’s phenotypic division may better represent a change in the agricultural traditions of the Austronesian-speaking peoples who moved into, and through, the Indo-Pacific region during the mid-Holocene. Bellwood (1978:148) suggests rice agriculture as a presumptive trigger for the expansion of Austronesian-speaking populations into the Indo-Pacific, and shows that its prevalence decreases in frequency from west to east through Indonesia. Natural climatic variation across the archipelago likely underpins this change from rice agriculture to tuber-based economies (Spriggs 2000:54).
Curiously, Wallace’s phenotypic line marks the general location where the seasonal tropical climate of more northerly latitudes is transformed into the season-less monotony of the equatorial zone. Dewar (2003, Figure 1) places the current eastward limit of rice (*Oryza sativa* L.) almost exactly along Wallace’s phenotypic line, and presents convincing evidence that this limit is regulated by climatic factors, particularly changes in rainfall variability. Although Japanese millet (*Echinochloa frumentacea* Link) has a similar eastward limit, other grain crops follow notably different distributions.

A model of rice agriculture driving the Austronesian expansion into Island Southeast Asia (Diamond and Bellwood 2003:601) is compelling primarily because an extensive array of terms for rice and its cultivation can be reconstructed for proto-Austronesian (Blust 1976, Pawley 2002:264). Reflexes of these terms occur in languages throughout Island Southeast Asia, including East Indonesia, and provide key evidence that rice cultivation was a component of early Austronesian societies. However, no such reflexes exist in Oceanic languages, and one must conclude that rice agriculture declined in importance as the Austronesian expansion moved through and beyond Island Southeast Asia. Thus it remains an intriguing possibility that the discontinuity observed in the genetic data was caused not simply by a collision of incursive agricultural peoples with sizable Melanesian populations, but rather by the decreasing ability of rice agriculture to propel the expansion of Austronesian-speaking peoples in the presence of indigenous inhabitants. This region may well have been a ‘friction zone,’ a place where human or environmental factors hindered the Asian population dispersal from interacting with (or imposing itself on) pre-existing populations in the region (Bellwood 2002). Borneo (~40,000 BP; Kennedy 1977, Bellwood 1997:84), the Philippines (22-20,000 BP; Bellwood 1997:85) and Java (~4,000 BP; Bellwood 1978:76) likely hosted Australo-Melanesian populations from the late Pleistocene to the early Holocene, thus supporting Howells (1976) contention of an “Old Melanesia,” a swathe of Australo-Melanesian populations that once stretched across much of modern Island Southeast Asia. Yet individuals of Asian descent dominate this region today, and their ancestors must have replaced Australo-Melanesian populations across most of this region at some stage in the past. The failure of agricultural practices due to changing climatic conditions provides a compelling explanation for why population replacement changed to a more evenly balanced admixture process in contemporary eastern Indonesia.

Alternative explanations must necessarily invoke geographically widespread and demographically disruptive processes, none of which are immediately obvious in the prehistoric, historic or modern anthropological literature. There is a paucity of evidence for rice in most early Island Southeast Asian archaeological sites, although there are also very few relevant archaeological studies on this topic. Setting aside some Taiwanese sites, archaeological evidence for rice remains have been found sparsely in southern Island Southeast Asia (reviewed by Paz 2002). The earliest rice appears in Sarawak (Gua Sireh, ~3850 BP), Luzon (Andarayan, ~3400 BP) and South Sulawesi (Ulu Leang, early but not securely dated). Interestingly, there is no clear evidence for rice further east, which is a key consequence of the horticultural hypothesis discussed above. Diamond and Bellwood (2003:601) suggest that “rice of subtropical South Chinese origin was abandoned … when the [Austronesian] expansion entered the equatorial tropics,” and Paz (2002:282) indicates that rice cultivation would have been hindered by the changing climatic conditions encountered by agricultural
populations as they moved through Island Southeast Asia. It remains a striking possibility that the discontinuity seen in haploid genetic data represents the demographic echo of this change to the Austronesian agricultural economy.

**Conclusion**

There is little doubt that the peoples of Melanesia have diverse and polyphyletic origins (e.g., Friedlaender *et al.* 2007, Hill *et al.* 2007, Mona *et al.* 2007) – a complex history that is reflected in the sex-specific and autosomal genetic lineages they carry today, as well as their cultural and linguistic diversity. Importantly, however, this region also shows clear signals of shared prehistoric demographic processes, including extensive mobility, constraints to population movement, and contact with groups in surrounding regions. The actors moving on, off, and through this Melanesian stage are participants in complex societies who interacted with their environment and human neighbors in dynamic ways over geographical space and chronological time. However, it is not necessary to characterize the prehistory of this region solely by its complexity. Demographic changes in the Indo-Pacific – and the underlying processes that drive them – can be approached, at least in some small way, via biological data.

The Indo-Pacific region hosts at least one clear discontinuity in the genetic variation of humans today, and this must reflect some outcome of prehistoric demographic processes. The observed clusters of genetic similarity do not accurately match known biogeographical boundaries, such as Wallace’s biogeographical line or the Near/Remote Oceania division, but they do produce a reasonable match to Alfred Wallace’s phenotypic boundary, his conceptual representation of changing physical characteristics that subdivides modern-day eastern Indonesia. This division is also a surprisingly good fit with linguistic, climatic and horticultural boundaries, and the action of one or more influential prehistoric demographic causes seems necessary to explain the association among these multiple strands of evidence. The discontinuity to the west of Melanesia may have been driven by retardation of agricultural Austronesian-speaking populations as they expanded through Island Southeast Asia towards New Guinea. There, the Austronesian expansion experienced partial admixture with Pleistocene-era populations, before rapidly moving eastward across the Near/Remote Oceania boundary and out into the greater Pacific. It seems that Melanesia, as an analytical unit, retains some ability to spur new ways of thinking about the ancient Indo-Pacific.

**Acknowledgements**

I extend my sincerest appreciation to the Indo-Pacific individuals and communities whose voluntary participation in genetic studies over many years has facilitated this partial reconstruction of their extraordinary prehistory. I also thank the late Thor Heyerdahl, an early benefactor of this research. I would like to acknowledge James Green (University of Otago) for statistical discussions; Peter Bellwood (Australian National University), J. Stephen Lansing (University of Arizona/Santa Fe Institute), John Schoenfelder (University of
California at Los Angeles), Matthew Spriggs (Australian National University) and John Terrell (Field Museum of Natural History) for anthropological discussions; Takafumi Ishida (University of Tokyo), Heather Norton (University of Arizona) and Jon Wilkins (Santa Fe Institute) for genetic discussions; and several anonymous reviewers for their remarks on an earlier version of this manuscript. This work was supported by grants from the Foundation for Research, Science and Technology (New Zealand), Thor Heyerdahl and the Kon-Tiki Museum (Norway), and the Isaac Newton Trust (United Kingdom); and by the patronage of the University of Otago (New Zealand), the University of Oslo (Norway), and the University of Cambridge (United Kingdom).

References


Dispersal Hypothesis (pp. 17-28). Cambridge: McDonald Institute for Archaeological Research.


Taiwanese patients with MELAS or MERRF syndrome. *Annals of the New York Academy of Sciences*, 1042, 82-87.


