

Article

Demographic History of Oceania Inferred from Genome-wide Data

Andreas Wollstein,^{1,2} Oscar Lao,² Christian Becker,¹ Silke Brauer,² Ronald J. Trent,³ Peter Nürnberg,¹ Mark Stoneking,⁴ and Manfred Kayser^{2,*}

¹Cologne Center for Genomics, University of Cologne, D-50931 Cologne, Germany

²Department of Forensic Molecular Biology, Erasmus MC University Medical Center Rotterdam, 3000 CA Rotterdam, The Netherlands

³Department of Molecular and Clinical Genetics, Royal Prince Alfred Hospital and Central Clinical School, The University of Sydney, Camperdown, NSW 2050 Sydney, Australia

⁴Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany

Summary

Background: The human history of Oceania comprises two extremes: the initial colonizations of Near Oceania, one of the oldest out-of-Africa migrations, and of Remote Oceania, the most recent expansion into unoccupied territories. Genetic studies, mostly using uniparentally inherited DNA, have shed some light on human origins in Oceania, particularly indicating that Polynesians are of mixed East Asian and Near Oceanian ancestry. Here, we use ~1 million single nucleotide polymorphisms (SNPs) to investigate the demographic history of Oceania in a more detailed manner.

Results: We developed a new approach to account for SNP ascertainment bias, used approximate Bayesian computation simulations to choose the best-fitting model of population history, and estimated demographic parameters. We find that the ancestors of Near Oceanians diverged from ancestral Eurasians ~27 thousand years ago (kya), suggesting separate initial occupations of both territories. The genetic admixture in Polynesian history between East Asians (~87%) and Near Oceanians (~13%) occurred ~3 kya, prior to the colonization of Polynesia. Fijians are of Polynesian (~65%) and additional Near Oceanian (~35%) ancestry not found in Polynesians, with this admixture occurring considerably after the initial settlement of Remote Oceania. Our data support a greater contribution of East Asian women than men in the admixture history of Remote Oceania and highlight population substructure in Polynesia and New Guinea.

Conclusions: Despite the inherent ascertainment bias, genome-wide SNP data provide new insights into the genetic history of Oceania. Our approach to correct for ascertainment bias and obtain reliable inferences concerning demographic history should prove useful in other such studies.

Introduction

The history of modern humans in Oceania comprises two extremes in terms of migrations. First, the initial colonization of Oceania is generally thought to reflect one of the oldest migrations of modern humans, namely an early out-of-Africa

exodus that reached Near Oceania (i.e., New Guinea and adjacent islands eastward, including the Solomon Islands up to Santa Cruz and the Reef Islands) and Australia at least 40 thousand years ago (kya) but did not extend into Remote Oceania (i.e., all islands from Santa Cruz and the Reef Islands eastward, including New Caledonia, Vanuatu, Fiji, and Polynesia) [1–5]. Descendants of that early occupation are thought to be Papuan-speaking (non-Austronesian languages) communities that nowadays mostly live in New Guinea, including the highlands and some coastal regions, mainly along the south coast, and some adjacent islands in the Bismarck Archipelago and the Solomon Islands [6, 7]. The second migration includes one of the most recent expansions of modern humans to settle unoccupied territories, namely the initial colonization of Remote Oceania that brought pottery-making farmers, seafarers, and fishermen, who most likely spoke Austronesian languages, into Near and subsequently Remote Oceania [8–10]. Austronesian-speaking groups of Oceania nowadays are found along the coast of mainland New Guinea, as well as on other islands of Near Oceania, and they comprise all groups in Remote Oceania, including Polynesians [11]. According to archaeological and linguistic evidence, the Austronesian expansion most likely started about 5.5 kya in Taiwan and continued through the Philippines and other parts of Island Southeast Asia, reaching the Bismarck Archipelago of northern Island Melanesia about 3.4 kya [12–14]. Here the typical elements of the Lapita cultural complex and the proto-Oceanic language developed and entered Remote Oceania about 3.2 kya, with a rapid spread eastward, leading to the initial occupation of all of Polynesia by about 1 kya [8, 11, 15].

This simplified two-stage scenario for the human occupation of Oceania is generally supported by uniparentally inherited Y-chromosomal (NRY) and mitochondrial (mt) DNA data [16–19]. For instance, Papuan-speaking groups of Near Oceania who do not live near Austronesian-speaking groups carry high frequencies of NRY/mtDNA haplogroups that are not found outside Near Oceania [20–23], except in parts of Eastern Indonesia and Remote Oceania, where they were brought from Near Oceania, in agreement with the assumption that they are descended from an early occupation of the region. Austronesian-speaking groups in Near and Remote Oceania usually carry these distinctive haplogroups, but they also carry other NRY/mtDNA haplogroups found in East and Southeast Asia [19, 24, 25]. Furthermore, there is a striking imbalance in the origin of Y chromosomes and mtDNAs in Polynesia: ~94% of Polynesian mtDNAs are of East Asian origin, whereas ~66% of Polynesian Y chromosomes are of Near Oceanian origin [26]. Higher frequencies of East Asian mtDNA and Near Oceanian NRY haplogroups were also observed in Austronesian-speaking groups from the Admiralty Islands of northern Island Melanesia [24]. These findings imply that the initial Austronesian arrival in Near Oceania was followed by sex-biased genetic admixture involving mostly Austronesian women and Near Oceanic men [24, 26], perhaps as a result of the matrilinear structure and matrilocality pattern of the pre-Polynesian societies [27, 28]. Furthermore, specific NRY/mtDNA haplogroups support a Taiwanese origin of the Austronesian expansion [16, 24] and highlight the important

*Correspondence: m.kayser@erasmusmc.nl

role of northern Island Melanesia in the population history of Remote Oceanians [24], as also suggested based on archaeological and linguistic data [9, 11].

However, uniparental data are not necessarily representative of the entire genome, nor do they permit detailed inferences concerning demographic history. Previous analyses of multilocus autosomal data sets do support the view that Polynesians are of admixed East Asian and Near Oceanian ancestry [29–31], but only one of these attempted any demographic inferences, and that was with a very simple demographic model focusing on the Near Oceanic admixture in Polynesians [29]. Hence, other approaches for analyzing genome-wide data are needed to investigate the demographic history of Near and Remote Oceanian populations in more detail. Moreover, there are still many open questions concerning the history of modern humans in Oceania that have not been addressed previously with genetic analyses (see overview in [32]). For example, although a southern dispersal route out of Africa into Near Oceania (and Australia) has been hypothesized [33, 34], rigorous tests of this hypothesis have so far not been carried out. Another question concerns the position of Fiji in the human history of Oceania. Fiji was probably initially occupied by Austronesian-speaking migrants coming from Near Oceania, which is supported by dates of Lapita sites that are older in Fiji than in Western Polynesia [5]. Fijians share cultural and phenotypic features with New Guineans [35], whereas their Austronesian language is closely related to Polynesian languages [36]. The NRY and mtDNA data indicate a closer relationship between Fiji and New Guinea than between Fiji and Polynesia [26], and some archaeological findings suggest that Fiji may have received subsequent migration from Near Oceania [5]. However, this secondary admixture scenario of Fijians has not been addressed with genetic analyses.

To investigate in more detail the human demographic history of Near and Remote Oceania, we generated and analyzed densely spaced genome-wide single nucleotide polymorphism (SNP) data obtained via Affymetrix 6.0 SNP microarrays for two population samples from Remote Oceania (Polynesia, including various islands, and Fiji), one from Near Oceania (Highlands of Papua New Guinea), and one from southern Borneo. We also included data for the same SNPs for Han Chinese from Beijing, Japanese from Tokyo, Yorubans from Ibadan, Nigeria, and U.S. European Americans from Utah with Northern and Western European ancestry, all from the International HapMap project [37]. We developed and applied a new approach to correct for the SNP ascertainment and performed approximate Bayesian computation for inferring demographic parameters reliably from such SNP microarray data. This approach, together with other analyses of the genome-wide data, provided new insights into the demographic history of Oceania and should also prove useful in analyses of similar data from other populations.

Results

Genetic Diversity and Population Substructure Heterozygosity, Population Differentiation, and Linkage Disequilibrium

Average heterozygosity in the four Asian and Oceanian populations for which Affymetrix 6.0 SNP data were generated (Figure 1) was highest in Borneons and lowest in New Guinea Highlanders, with Polynesians and Fijians both having highly similar, intermediate values (see Table S1 available online).

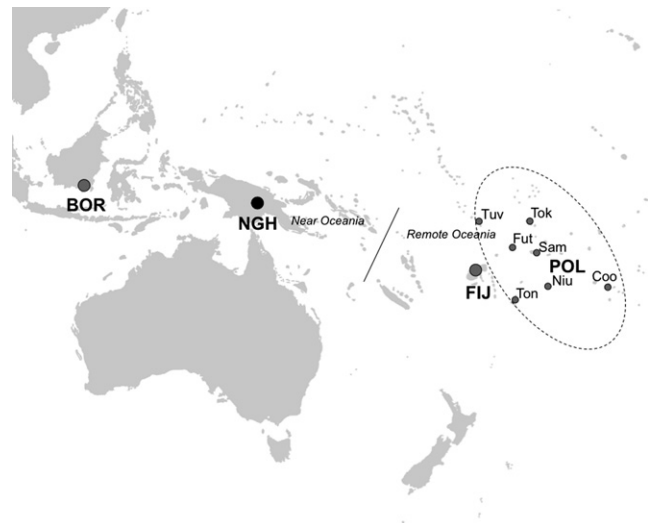


Figure 1. Geographic Map of Asia and Oceania with Study Sample Sites Indicated for Which Genome-wide Affymetrix SNP Microarray Data Were Generated

The following abbreviations are used: BOR, Borneo (southern part, Barito river area); NGH, New Guinea Highlands (Southern Highlands of Papua New Guinea); FIJ, Fiji; POL, Polynesia, including seven islands: Tuvalu (Tuv), Samoa (Sam), Tonga (Ton), Futuna (Fut), Niue (Niu), Tokelau (Tok), and Cook (Coo) Islands. Grey dots denote Austronesian-speaking groups, black dot denotes non-Austronesian (Papuan)-speaking group.

Comparing these data with those from the four worldwide HapMap populations revealed highest heterozygosity in the Yorubans from Africa, but also higher values in Europeans, whereas heterozygosity values in Chinese and Japanese were lower than in Borneons but higher than in all three Oceanian populations. The same trends were seen in the number of polymorphic sites. Mutual F_{st} , as overall measure for population differentiation, was second highest in New Guinea Highlanders after Yorubans, and higher in Polynesians and Fijians than in Chinese and Borneons (Table S1). Pairwise F_{st} values were, among the four Asian and Oceanian populations, highest between New Guinea Highlanders and the other three groups (Table S2). Furthermore, we detected a strong and highly significant correlation between pairwise F_{st} values and geographic distance for all eight population samples included ($R^2 = 0.387$; Mantel test with 1 million permutations $p = 0.00688$). We also analyzed for all population samples the decay of mean linkage disequilibrium over genomic distance (data not shown). Most notably, at genomic distances of 500 kb from the target SNP, we observed higher mean linkage disequilibrium in Polynesians and Fijians compared to all other populations analyzed (Table S1).

Frappe and Principal Component Analyses

We applied the Frappe analysis to an increasing number of clusters (K) and obtained consistent results over independent runs for $K = 2$ to $K = 6$, as depicted in Figure 2 (see Figure S1 for results of multiple independent runs). For $K = 2$, all individuals clustered almost perfectly as Africans versus non-Africans, except the Europeans expressed an appreciable amount of the African ancestry component, which, however, disappeared at higher K values. For $K = 3$, individuals were clustered into three groups, Africans, Europeans, and Asians/Oceanians, with very little sharing of ancestry components between them. At $K = 4$, the New Guinea Highlanders appeared as an additional cluster, and this signal also occurred at a frequency

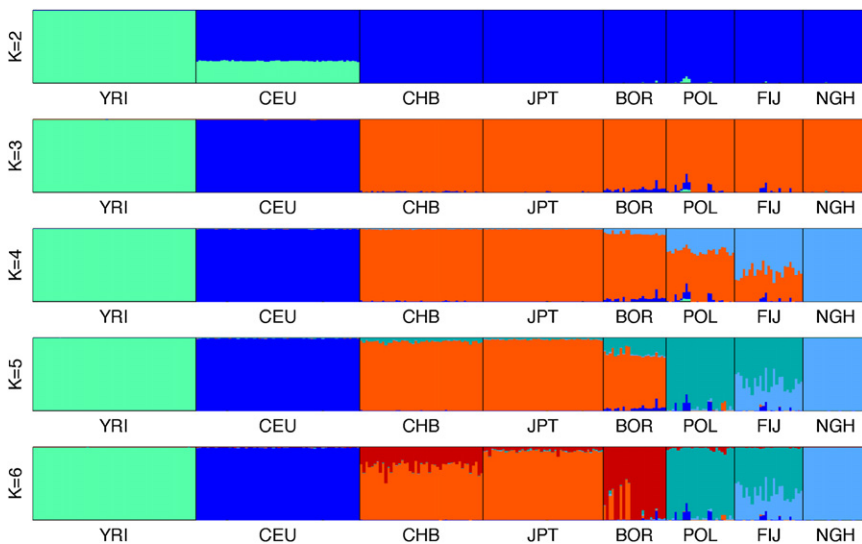


Figure 2. Frappe Analysis Based on Genome-wide SNP Data for the Four Asia and Oceania Populations Together with the Four HapMap Populations

Analysis is based on 765,708 autosomal SNPs intersecting between the Affymetrix 6.0 data generated for the four Asia and Oceania populations and the HapMap data set. Each bar represents a single individual, with abbreviations as follows: YRI, Yorubans from Africa; CEU, Europeans; CHB, Han Chinese; JPT, Japanese; BOR, Borneans; POL, Polynesians; FIJ, Fijians; NGH, New Guinea Highlanders. For results from multiple independent Frappe runs, see Figure S1. For Frappe analyses to estimate admixture of Polynesians and Fijians, and for Frappe analysis per chromosome, see Figure S3.

of about 58% in Fijians, 30% in Polynesians, and 6% in Borneans, but nowhere else. At $K = 5$, Polynesians appeared as an additional separate cluster, and this signal also appeared at a frequency of about 60% in Fijians and 20% in Borneans. At $K = 6$, Borneans appeared as an additional separate cluster; this signal was also seen at a frequency of about 26% in Chinese and rarely in Japanese (~4%) or Polynesians (~2%). Notably, from $K = 3$, we consistently observed a small European signal in a few individuals each in Polynesia, Fiji, and Borneo, but not in the New Guinea Highlanders, although with $K = 6$ this signal disappeared from Borneans (except for one sample with 13%). In particular, with $K = 6$, six Polynesian samples (four Cooks and two Tokelau) showed on average 13% of the European signal not observed in the remaining 19 Polynesians (Figure 2), and 5 of the 25 Fijian samples showed on average 7% of the European component (Figure 2). In order to avoid any influence of potential recent European admixture in the demographic parameter estimations, we excluded these 12 individuals from the demographic analyses.

Furthermore, we carried out principal component analyses. Figure 3A shows the plot of the first versus the second principal component for all individuals studied (see Figure S2 for higher-order components). As evident, the first principal component separated Africans, Europeans, and the Asian/Oceanian samples, whereas the second principal component differentiated Chinese/Japanese/Borneans from New Guinea Highlanders, with Polynesians and Fijians placed in between. When applying the principal component analyses only to the six populations from Asia/Oceania (Figure 3B; Figure S2 for higher-order components), we still observed a clustering of Chinese/Japanese/Borneans separated from New Guinea Highlanders and from Polynesians/Fijians. Notably, in both Figure 3A (with principal components 1 and 2) and Figure 3B (with principal component 1), Polynesian and Fijian samples appeared between Chinese/Japanese/Borneans and New Guinea Highlanders, and Fijians appeared between Polynesians and New Guinea Highlanders, but closer to Polynesians than to New Guinea Highlanders.

We also investigated via separated principal component and Frappe analyses the fine-scale population substructure within Polynesians and within New Guinea Highlanders, respectively. The principal component analysis of just the Polynesian

samples (Figure 3C) resulted in a separation of the Cook Islanders from all other Polynesians along principal component 1, whereas component 2 largely separated non-Cook Polynesians according to their island of origin. The Frappe analysis for $K = 2$ showed the same separation of Cook Islanders from other Polynesians (data not shown). The New Guinea Highlanders samples come from two different groups, Huli speakers and Angal-Kewa speakers, and these are completely separated in the principal component analysis carried out for New Guinea Highlanders alone (Figure 3D) and in the corresponding Frappe analysis for $K = 2$ (data not shown).

Furthermore, dedicated principal component and Frappe with $K = 2$ analyses were used to estimate the amount of admixture in Polynesians and Fijians. Different combinations of assumed ancestral populations were tested, such as Chinese/New Guinea Highlanders and Borneans/New Guinea Highlanders for Polynesians, as well as Chinese/New Guinea Highlanders, Borneans/New Guinea Highlanders, and Polynesians/New Guinea Highlanders for Fijians. The admixture estimates from the principal component analysis were almost identical to those obtained from Frappe for the same population trios (Table S3; Figure S3). We estimated for Polynesians about 85% Bornean and 15% New Guinean admixture and for Fijians about 63% Polynesian and 37% New Guinean admixture (as shown below from simulation analyses, these are the most likely ancestral populations for Polynesians and Fijians).

We additionally applied principal component analysis to SNPs in linkage equilibrium ($R^2 < 0.8$) and found a high and significant correlation of principal components 1 and 2 for the reduced versus the full data set (all populations: $R^2 = 0.9888$, 234619/765708 SNPs; all Asians: $R^2 = 0.9930$, 301463/767764 SNPs; Polynesians: $R^2 = 0.9958$, 566685/816283 SNPs; New Guineans: $R^2 = 0.9651$, 606289/818967 SNPs; $p < 1e^{-6}$ in all cases using a Mantel test). Hence, we conclude that data thinning based on linkage disequilibrium does not significantly influence our results for the first and second principal components.

We also applied Frappe to SNPs from the X chromosome and each autosomal chromosome separately in order to estimate the admixture of Polynesians and Fijians using Borneans/New Guinea Highlanders and Polynesians/New Guinea Highlanders, respectively, as proxies of the parental populations.

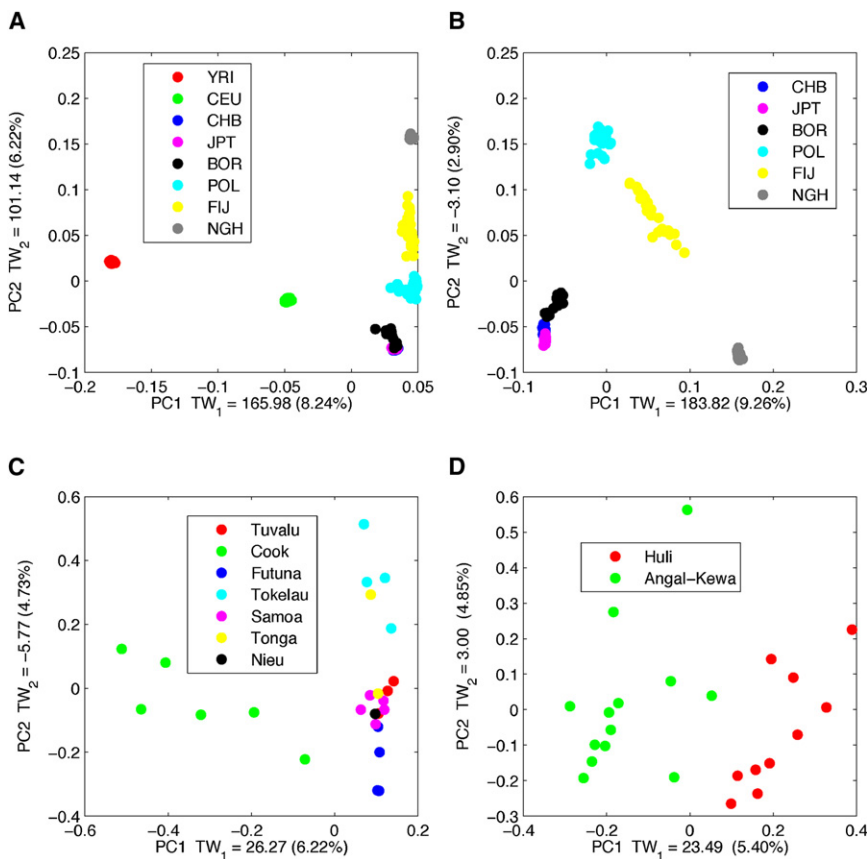


Figure 3. Principal Component Analysis from Genome-wide SNP Data using Different Subsets of the Four Asia and Oceania Populations and the Four HapMap Populations

Each dot represents a single individual; the same numbers of individuals per group were analyzed. Tracy-Widom (TW) statistics and the amount of explained variance are given for the corresponding components. According to the TW distribution, a component is significant with $p < 0.05$ for $TW > 1.2$.

(A) Principal component 1 versus 2 for all populations analyzed (765,708 autosomal SNPs).

(B) Principal component 1 versus 2 for only the Asian (including Chinese, Japanese, and Borneon) and Oceanian (including New Guinea Highlander, Fijian, and Polynesian) populations (767,764 autosomal SNPs).

(C) Principal component 1 versus 2 for only the Polynesians (816,283 autosomal SNPs).

(D) Principal component 1 versus 2 for only the New Guinea Highlanders (818,967 autosomal SNPs).

The following abbreviations are used: YRI, Yorubans from Africa; CEU, Europeans; CHB, Han Chinese; JPT, Japanese; BOR, Borneons; POL, Polynesians; FIJ, Fijians; NGH, New Guinea Highlanders. For higher-order principal components, see Figure S2. For principal component analysis to estimate admixture of Polynesians and Fijians, see Figure S3.

The X-chromosomal admixture estimates were 97.6% Borneon ancestry for Polynesians and 95.9% Polynesian ancestry for Fijians, which was considerably higher than the corresponding admixture estimates for the autosomes (Figure S3).

Demographic Inferences

Affymetrix SNP Ascertainment Bias Correction

To obtain reliable demographic inferences from SNP microarray data, we developed and implemented in the demographic model a novel procedure to account for ascertainment bias in the SNPs represented on the Affymetrix 6.0 chip. The approximate posterior distributions of the discovery depth for the different populations originally used in the SNP ascertainment procedure are depicted in Figure S4A. Based on the median values of this analysis, we conclude that two Yorubans, one European, and two Chinese chromosomes are the most likely values to explain the differences between the observed summary statistics for the ENCODE sequence data and the Affymetrix SNP array genotype data. We then applied the approximate posterior distributions of the discovery depth as a prior in the subsequent Bayesian estimations using the Affymetrix 6.0 SNP genotypes. As a proof of concept, we first estimated demographic parameters for a simple out-of-Africa scenario (model 1 in Figure 4) using the ascertainment bias-corrected Affymetrix 6.0 SNP data and compared them to estimates obtained from ENCODE sequence data not prone to any SNP ascertainment bias. We observed results similar to the ENCODE regions for effective population size estimates, as well as somewhat higher estimates of the divergence times (Table S4; Figures S4B and S4C), indicating that our approach

is suitable to infer demographic parameters for Oceania using the ascertainment bias-corrected SNP array data.

Modeling Population Splits and Demographic Parameter Estimation

We first tested competing hypotheses (Figure 4, models 2a–2c) on the split of New Guineans (1) directly from Africans (model 2c), (2) from the joined ancestral population of Europeans and Chinese (i.e., Eurasia, model 2b), or (3) from Chinese (i.e., East Asia, model 2a). Model 2b, a split of New Guineans from Eurasia, received the highest support, with an estimated posterior probability of 0.74 (Table 1), whereas the posterior probability of a New Guinea split from East Asians was only 0.24, and there was practically no support for a direct split of New Guineans from Africans ($p = 0.02$). Next, we tested various hypotheses for the origins of Polynesians (Figure 4, models 3a–3c): (1) divergence of Polynesians from Chinese (model 3a), (2) divergence of Polynesians from New Guineans (model 3b), or (3) Polynesians arising from admixture between Chinese and New Guineans, incorporating the admixture rate as an additional model parameter (model 3c). Model 3c, the admixture model of Polynesians, received the highest probability ($p = 0.75$). Considering this finding, we then applied two types of model comparisons to further explore the origin of Borneons (Figure 4, models 4a–4c) and Fijians (Figure 4, models 5a–5c). For Borneons, we obtained a very high posterior probability of 0.93 for a split from Chinese (Table 1). For Fijians (assuming no admixture), we obtained the highest posterior probability for a split from Polynesians at 0.75, with a probability of 0.10 for a split from Chinese and of 0.14 for a split from New Guineans (Table 1). We then tested specifically

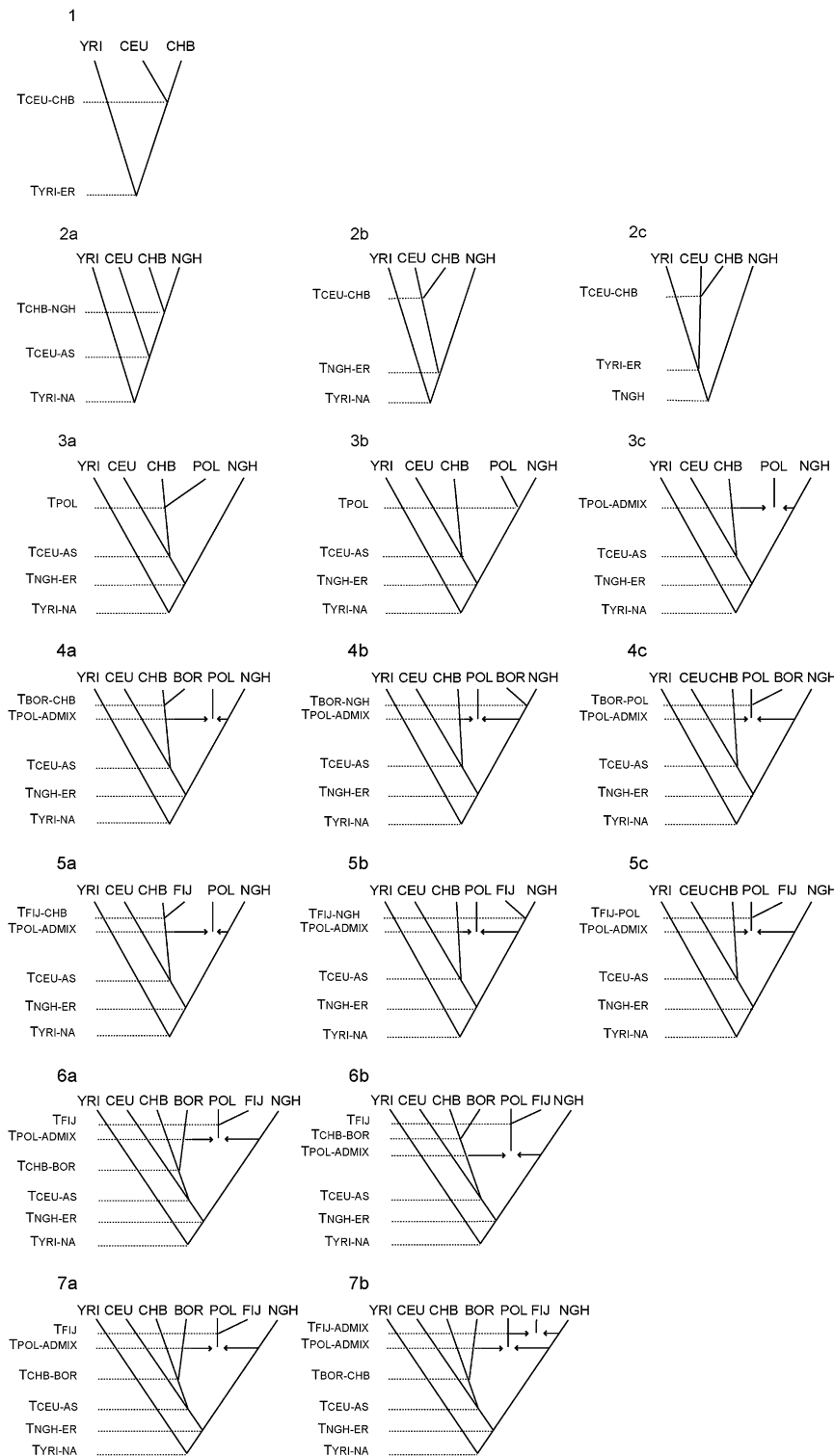


Figure 4. Investigated Demographic Models, Ordered by Increasing Complexity, using Genome-wide SNP Data after Correction for Affymetrix SNP Ascertainment Bias

Model 1 depicts the simplified out-of-Africa model for the three populations: sub-Saharan Africans (YRI), Europeans (CEU), and East Asians (Chinese, CHB). Models 2a–2c are based on model 1 and describe competing split scenarios for New Guinea Highlanders (NGH). Models 3a–3c describe alternative scenarios for the origin of Polynesians (POL). Models 4a–4c define alternative scenarios for the split of Borneans (BOR), and models 5a–5c define alternative scenarios for the split of Fijians (FIJ); both sets of models assume Polynesians to be an admixed population between Asians (Chinese or Borneans) and New Guineans and a split of New Guineans from Eurasia. Models 6a and 6b provide alternative scenarios for the Asian admixture of Polynesians. Models 7a and 7b describe alternative scenarios for the origin of Fijians. Model 7b depicts the most-likely tree topology, as derived from the previous model comparisons, and was used to derive the demographic parameters shown in Table 2. The following abbreviations are used: NA, non-Africans; AS, Asians; ER, Eurasians; ADMX, admixture. See Figures S4B and S4C for the posterior distributions of the demographic parameters for model 1, and see Figure S5 for the posterior distributions of the demographic parameters for model 7b.

from Polynesians (Figure 4, model 7a) versus Fijians representing an admixture of Polynesians and New Guineans (Figure 4, model 7b). We found the admixture of Fijians between Polynesians and New Guineans as parental populations (0.6) to be preferred against a split of Fijians from Polynesians (0.4). Overall, we can conclude that model 7b (Figure 4) provides the most likely scenario for the human history of Oceania, based on the population samples included in this study.

We then applied model 7b (Figure 4) to estimate demographic parameters using 3.7 million simulations (Table 2; Figure S5A). For Africans, Europeans, and Chinese, the effective population sizes and divergence times were similar to those estimated for model 1 (Table S4). Using a generation time of 30 years [38], the divergence time of New Guineans from Eurasia was ~27 kya, the time of admixture of Polynesians between Borneans and New Guineans

whether Borneans or Chinese better represent the Asian parental population of the admixed Polynesians/Fijians, considering New Guineans as the second parental population (Figure 4, models 6a and 6b). We found that Polynesians are more likely to be an admixture of Borneans and New Guineans (0.67) than of Chinese and New Guineans (0.33) (Table 1). Finally, we estimated the posterior probability of a Fijian split

was ~3 kya, the admixture time of Fijians between Polynesians and New Guineans was ~0.5 kya, and the effective population sizes were ~1100 individuals for Polynesians and ~2000 for New Guinea Highlanders. The admixture proportions for Polynesians were ~87% Bornean and ~13% New Guinean ancestry, and for Fijians they were ~65% Polynesian and ~35% New Guinean ancestry (Table 2).

Table 1. Results from Demographic Model Testing using Genome-wide SNP Data after Correction for Affymetrix SNP Ascertainment Bias

Model ^a	Description	Probability	Power ^b
2a	New Guinea split from East Asia	0.237	0.801
2b	New Guinea split from Eurasia	0.743	0.582
2c	New Guinea split from Africa	0.020	0.523
3a	Polynesia split from East Asia	0.177	0.643
3b	Polynesia split from New Guinea	0.069	0.832
3c	Polynesian admixture between East Asia and New Guinea	0.753	0.401
4a	Borneo split from East Asia	0.928	0.904
4b	Borneo split from New Guinea	0.000	0.851
4c	Borneo split from Polynesia	0.072	0.732
5a	Fiji split from East Asia	0.104	0.904
5b	Fiji split from New Guinea	0.142	0.851
5c	Fiji split from Polynesia	0.754	0.732
6a	Admixture of Polynesian/Fijian ancestors between Borneo and New Guinea	0.673	0.635
6b	Admixture of Polynesian/Fijian ancestors between East Asia and New Guinea	0.327	0.746
7a	Fiji split from Polynesia	0.396	0.907
7b	Fijian admixture between Polynesia and New Guinea	0.604	0.534

^a For tree topologies, see Figure 4, with competing models summarized by the same numbers used here.

^b Estimated from 250 simulations from random samples of the prior of a model; a power of 0.9, for example, means that in 90% of the simulations, we could recover the true model.

As a further check on the performance of our method, we used the inferred demographic parameters to generate a simulated data set and then used the approximate Bayesian computation framework to infer the demographic parameters. Almost all parameters are perfectly recovered by the median posterior distribution, and the true parameter value always lies within the 95% confidence range of the posterior distribution (Figure S5B).

Discussion

Descriptive analyses on the genetic diversity of the four newly analyzed population samples from Asia and Oceania and the four HapMap populations fit general expectations. Most notably, the Yorubans from Africa have the largest amount of genetic diversity and the lowest linkage disequilibrium, and we see in our data a general increase of linkage disequilibrium with increased geographic distance from Africa, as observed previously [37, 39], both usually interpreted as genetic evidence in favor of an out-of-Africa scenario of modern human origins. The lower diversity for New Guinea Highlanders may be explained by the absence of New Guinean samples in the Affymetrix SNP ascertainment procedure, whereas the geographic regions of other populations studied here either were directly involved in the SNP ascertainment (Africans, East Asians, and Europeans) or are assumed to be related to populations included in the ascertainment (e.g., Borneans, Fijians, and Polynesians) [40]. Alternatively, the low diversity of New Guinea Highlanders may be explained by a smaller effective population size and a correspondingly stronger impact of genetic drift. We also detect higher mean linkage disequilibrium in Polynesians and in Fijians compared to all other populations analyzed, which could reflect either recent admixture [41] or bottlenecks or founder events [42], or both.

Table 2. Prior and Mean Approximate Distributions of Demographic Parameters for the Best-Fitting Model 7b of Figure 4, Based on 3.7 Million Simulations from Genome-wide SNP Data after Correction for Affymetrix SNP Ascertainment Bias

Parameter	Prior Distribution	Median Approximate Posterior (95% CI)
N_{NGH}	$U(100, N_{YRI})$	2026 (1261, 3081)
N_{POL}	$U(100, N_{YRI})$	1134 (395, 3146)
N_{BOR}	$U(100, N_{YRI})$	4034 (1993, 6497)
N_{FIJ}	$U(100, N_{YRI})$	1678 (437, 4697)
T_{YRI-NA}	Posterior of model 1	1847 (1488, 2317)
T_{CEU-AS}	Posterior of model 1	619 (437, 826)
T_{NGH-ER}	$U(100, T_{CEU-AS})$	904 (614, 1345)
$T_{BOR-CHB}$	$U(10, T_{CEU-AS})$	138 (54, 257)
$T_{POL-ADMX}$	$U(10, \min[T_{NGH-ER}, T_{BOR-CHB}])$	99 (19, 267)
r_{POL}	$U(0, 1)$	0.866 (0.645, 0.974)
$T_{FIJ-ADMX}$	$U(10, T_{POL-ADMX})$	17 (3, 59)
r_{FIJ}	$U(0, 1)$	0.655 (0.238, 0.939)

The following abbreviations are used: CI, credible interval; N, effective population size in number of individuals; T, divergence time in generations; r, admixture rate; NA, non-Africans; ER, Eurasians; AS, East Asians (Chinese); CEU, Europeans; YRI, Yorubans from Africa; POL, Polynesians; BOR, Borneans; FIJ, Fijians; NGH, New Guinea Highlanders; ADMX, admixture; U(a,b), the uniform prior in the range [a,b]. For posterior distributions, see Figure S5A.

The main goal of this study, however, was to use dense genome-wide SNP data, obtained with the Affymetrix 6.0 platform, to reconstruct the human population history of Near and Remote Oceania. As a consequence of the discovery procedure of SNPs included on commercially available microarrays, the marker set is enriched by common variants [43], which influences the distribution of summary statistics [40] and therefore biases the values of the inferred demographic parameters. The approach we adopted to correct for such ascertainment bias was to model the discovery depth per population and then incorporate this information in the Bayesian framework to estimate demographic parameters. From the posterior distribution of the discovery depth, it can be seen that it is unlikely that no Yoruban and no Chinese samples were involved in the SNP discovery process. We then demonstrated that demographic inferences for the four HapMap populations based on the genotype data after ascertainment bias correction yielded similar results as those obtained from ENCODE sequence data not affected by any SNP ascertainment bias and that the estimated population parameters were similar to previous estimates for the same populations [44, 45]. These findings provide confidence that our approach is valid for inferring demographic parameters in Oceania.

Settlement History of Near Oceania

Among the three demographic models examined for the peopling of Near Oceania (Figure 4, models 2a–2c), the model receiving the highest support involves a split of New Guineans from a common European–East Asian (i.e., Eurasian) ancestor population. This finding does not support the southern-dispersal hypothesis of separate human migrations from Africa to Near Oceania and to East Asia [33, 34]. The existence of a single ancestral population for all present-day non-Africans is supported, among other genetic evidence, by recent data from the Neandertal genome sequence, indicating that all present-day non-African genome sequences studied (including one from a Papua New Guinean) have equivalent amounts of Neandertal admixture [46]. However, the model suggested by our data does imply an earlier migration of the

Near Oceanian ancestors from this ancestral non-African population, followed by a later migration of the East Asian ancestors from this same ancestral population. The estimated ~27 kya for the split time of New Guineans from Eurasians is in broad agreement with the earliest archeological dates for the human occupation of Near Oceania at 35–40 kya [1–4] and is considerably older than the estimated split time of East Asians from Europeans (18 kya). Still, there also is appreciable support in our data for a split of New Guineans from East Asians, and there is large overlap in the 95% credible intervals for the divergence time of New Guineans from Eurasians versus New Guineans from East Asians (Table 1). This could reflect migration between Near Oceania and East Asia after the initial colonization, or it could indicate that there was just one major migration that led to the settlement of both New Guinea and East Asia, as suggested by another study of genome-wide SNP data, albeit using very little data from Near Oceania [47]. Genetic data from more populations are therefore needed to establish whether Near Oceania and East Asia were indeed colonized via separate migration waves, or instead via the same migration wave.

Settlement and Admixture History of Remote Oceania

All analyses we performed on our genome-wide SNP data clearly indicate that Polynesians represent a genetically admixed population carrying ancestry components from both East Asia and Near Oceania. The approximate Bayesian computation simulations in particular provide general support for the East Asian ancestry component of Polynesians, with Borneans more likely to represent this ancestry component than Han Chinese, although the latter were previously used for estimating Asian admixture in Polynesians from multilocus autosomal data [29, 31]. This finding is in good agreement with linguistic data, because the people of Borneo speak an Austronesian language (as Polynesians do), whereas Han Chinese speak a Sino-Tibetan language (<http://www.ethnologue.com>). Our findings could be taken as support for the Slow Boat to Melanesia model of Polynesian origins [48], which suggests that Indonesia represents the homeland of Polynesians in Paleolithic times [48]; however, there is little genetic or other data supporting this hypothesis [13, 26, 29–31, 49–51]. Other models, including the Slow Boat from East Asia model of Polynesian origins that we proposed earlier [52], place the homeland for the Austronesian expansion in Taiwan or coastal China in Neolithic times, which is supported by various genetic, archaeological, and linguistic data [12–14, 26, 29–31, 49, 53–56]. Under any of these scenarios, Borneo is on the direct route of the Austronesian expansion, and hence our results cannot clearly distinguish among them. However, the estimated Polynesian admixture time between Borneans and New Guineans of about 3 kya makes a Paleolithic origin of Polynesians unlikely.

Our approximate Bayesian computation simulations also indicate that Polynesians are of about 87% Bornean and 13% New Guinean admixed ancestry, which closely agrees with the estimates we obtained from dedicated principal component and Frappe analyses for estimating Polynesian admixture (even though the latter analyses do not incorporate a biological model, nor do they take into account the ascertainment bias). Overall, the Polynesian admixture estimates from this study agree quite closely with previous estimates from multilocus autosomal data (although they were achieved either from cluster algorithms only [31] or from simulations employing a much more simple demographic model [29]). The estimated

Polynesian admixture time of about 3 kya suggests that Austronesian migrants (arriving from East Asia) mixed with local Near Oceanians, most likely in Near Oceania prior to the occupation of Remote Oceania. Notably, our genetic admixture time estimate is in good agreement with a date of 3.7 kya estimated recently from linguistic data for the arrival of Austronesian languages in Near Oceania [13], as well as with archaeological evidence for the advent of the Lapita cultural complex in Near Oceania about 3.4 kya [9, 55, 57], which subsequently spread to (western) Polynesia. Based on uniparental genetic data, we previously proposed that this admixture involved primarily Austronesian women and primarily Near Oceanian men [24, 26, 52]. The new X chromosomal and autosomal data reported here provide further support for this Slow Boat (from East Asia) model [52], because the estimated East Asian ancestry in Polynesians and Fijians is significantly higher for X-chromosomal SNPs than for genome-wide autosomal SNPs. This finding implies that the number of East Asian women involved in the admixture history of Remote Oceanians was larger than that of East Asian men.

The genetic history of Fiji is analyzed here in detail for the first time. If Polynesia was colonized via Fiji, as strongly indicated by archaeological, linguistic, and NRY/mtDNA data [5, 13, 26], then in the absence of any subsequent contact between Fiji and Near Oceania, we would expect Fijians to be nearly identical to Polynesians genetically. The only genetic differences should then reflect bottleneck effects and genetic drift during the subsequent colonization of the rest of Remote Oceania. However, our model comparisons indicate that Fijians carry about 2 times more Near Oceanian ancestry than do Polynesians (in agreement with admixture estimates obtained with principal component and Frappe analyses), thereby suggesting substantial contact between Fiji and Near Oceania that did not extend to Polynesia. The Fijian admixture time estimate of about 0.5 kya may not be reliable, because this reflects admixture involving the recently admixed Polynesians as one of the parental groups, making reliable estimates more difficult. However, this time estimate does suggest that the additional Near Oceanian genetic admixture of Fijians happened well after the initial occupation of Remote Oceania, which, according to archaeological evidence, occurred between 3.2 and 2.1 kya for most islands (for overview, see [5]). Phenotypic, cultural, and linguistic data have previously indicated more similarities between Fijians and Near Oceanians than for other groups of Remote Oceania [5, 35]. Our data provide the first genetic evidence that after the initial settlement of Fiji, and subsequent migration from Fiji throughout other parts of Remote Oceania, there was additional, subsequent contact between Fiji and Near Oceania, thereby explaining such similarities. Detailed analyses of additional populations located between New Guinea and Fiji (such as the Solomons, New Caledonia, and Vanuatu) would provide further insights into the source and timing of this additional genetic input into Fijians.

Although the specific modern populations used in this study to represent the parental populations of Polynesians and Fijians may not be the exact (real) ancestral populations, they are likely to be closely related to them. Notably, we chose Papuan-speaking Highlanders from Papua New Guinea to represent the Near Oceanian ancestral group because linguistic, archaeological, and uniparental genetic data suggest that the Austronesian expansion never reached the Highlands of New Guinea [7, 18, 20, 57]. Choosing the best modern population proxy to represent the Asian ancestral

group of Polynesians may indeed be more complex. We show here that using Han Chinese and Borneans provides slightly different admixture proportions for Polynesians and that Borneans were preferred over Chinese as one of the ancestral groups of Polynesians in our simulation analyses; additional candidate populations (such as Aboriginal groups from Taiwan) should be investigated in the future.

Fine-Scale Population Structure within Near and Remote Oceania

Another question of interest is the extent to which genome-wide SNP array data can reveal fine-scale population structure. Previous studies have found fine-scale structure in SNP microarray data from European populations [58, 59], but to what extent can such data reveal fine-scale structure in populations that are only remotely (or not at all) related to the populations used to ascertain the microarray SNPs? We indeed find indications of fine-scale structure in the Affymetrix 6.0 SNP data in both the Highlanders of Papua New Guinea and in Polynesians. In the case of New Guinea Highlanders, most of the samples came from two language groups, Huli and Angal-Kewa, both belonging to the Engan branch of the Trans-New Guinea languages (<http://www.ethnologue.com>), and the villages from which the sampled individuals came are separated by only 70–100 km. And yet, the principal component and Frappe analyses clearly separated Huli speakers from Angal-Kewa speakers. Concerning Polynesians, the first principal component, as well as Frappe, separated the Cook Islanders from the other six Polynesian islands, and there was some separation of these Islanders with the second principal component. Differential European ancestry could, in principle, play a role in the genetic separation of the Cook Islanders from the other Polynesians, because previous studies showed a somewhat larger frequency of European Y chromosomes in Cook Islanders relative to other Polynesians (except New Zealand Maori) [19, 26, 51, 60]. However, none of the Cook Islander or other Polynesian samples used in our study carried a European Y chromosome (data not shown). Not only did a Frappe analysis of our genome-wide SNP data identify six Polynesians (including four Cook Islanders) carrying a small fraction of the major European component, but we also found a significant correlation between this Frappe component and the first principal component in the Polynesians ($R^2 = 0.51$, $p = 6.15e^{-05}$). However, repeating the Polynesian principal component analysis without these six samples provided essentially the same results: the Cook Islanders were largely separated from other Polynesians by component 1, with additional separation among Polynesian Islanders by component 2 (data not shown). Hence, overall, differential European ancestry does not explain the fine-scale population structure seen in Polynesia (or in the New Guinea Highlands), thereby indicating that dense SNP array data can detect fine-scale structure, even with ascertainment bias.

Future Prospects

Overall, we demonstrated that our approach is able to obtain reliable inferences for some demographic parameters from ascertained, genome-wide SNP microarray data, and hence it should be widely applicable. However, with our current approach, we could not reliably infer additional demographic parameters of interest such as population expansion or bottleneck events (data not shown). In principle, such parameters may be inferable from the second-order moments of the site frequency spectrum. Our inference framework includes the

site frequency spectrum with the joint allele frequency spectrum, but ascertainment bias is known to reduce the variance of the heterozygosity; hence, information about population size change in the past would be lost, and therefore such parameters could not be considered in our approach. Further work is needed to allow reliable inference of parameters involving population size changes. Nevertheless, we expect our approach of correcting for ascertainment bias in Affymetrix SNP microarray data for more reliable inference of demographic parameters of human population history to prove useful in other such studies.

Experimental Procedures

For information on materials and methods, see [Supplemental Experimental Procedures](#).

Supplemental Information

Supplemental Information includes five figures, five tables, and Supplemental Experimental Procedures and can be found with this article online at [doi:10.1016/j.cub.2010.10.040](https://doi.org/10.1016/j.cub.2010.10.040).

Acknowledgments

We are grateful to all volunteers who provided samples and to the International HapMap Project for making their data available. We thank Mark Vermeulen and Ingelore Bäßmann for technical assistance and the compute-cluster Clio of the University of Cologne Computing Centre for technical support. This study was supported by the German Research Council (DFG) Sonderforschungsbereich (SFB) 680 “Molecular Basis of Evolutionary Innovations” to P.N. and M.K. M.S. was supported by the Max Planck Society, and M.K. was additionally supported by the Erasmus MC University Medical Center, Rotterdam.

Received: June 18, 2010

Revised: September 22, 2010

Accepted: October 16, 2010

Published online: November 11, 2010

References

1. Groube, L., Chappell, J., Muke, J., and Price, D. (1986). A 40,000 year-old human occupation site at Huon Peninsula, Papua New Guinea. *Nature* 324, 453–455.
2. Allen, J., Gosden, C., Jones, R., and White, J.P. (1988). Pleistocene dates for the human occupation of New Ireland, northern Melanesia. *Nature* 331, 707–709.
3. O’Connell, J.F., and Allen, J. (2004). Dating the colonization of Sahul (Pleistocene Australia-New Guinea): A review of recent research. *J. Archaeol. Sci.* 31, 835–853.
4. Pavlides, C., and Gosden, C. (1994). 35,000-year-old sites in the rainforests of West New Britain, Papua New Guinea. *Antiquity* 68, 604–610.
5. Kirch, P.V. (2000). *On the Road of the Winds: An Archaeological History of the Pacific Islands before European Contact* (London: University of California Press).
6. Wurm, S.A., and Hattori, S. (1981). *Language Atlas of the Pacific Area, Volume 66* (Canberra: Australian Academy of the Humanities, in Collaboration with the Japan Academy).
7. Foley, W.A. (2000). The languages of New Guinea. *Annu. Rev. Anthropol.* 29, 357–404.
8. Green, R.C. (1991). The Lapita cultural complex: Current evidence and proposed models. *Indo-Pacific Prehist. Assoc. Bull.* 11, 295–305.
9. Kirch, P.V. (1997). *The Lapita Peoples: Ancestors of the Oceanic World* (Oxford: Blackwell).
10. Bellwood, P., Fox, J., and Tryon, D. (1995). *The Austronesians: Historical and Comparative Perspectives* (Canberra, Australia: Australian National University).
11. Lynch, J., Ross, M., and Crowley, T. (2002). *The Oceanic Languages* (London: Curzon Press).

12. Bellwood, P., and Dizon, E. (2005). The Batanes archaeological project and the "Out of Taiwan" hypothesis for Austronesian dispersal. *J. Austronesian Studies* 1, 1–31.
13. Gray, R.D., Drummond, A.J., and Greenhill, S.J. (2009). Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science* 323, 479–483.
14. Blust, R. (1999). Subgrouping, circularity and extinction: Some issues in Austronesian comparative linguistics. *Symposium Series of the Institute of Linguistics Academia Sinica* 1, 31–94.
15. Green, R.C. (2003). The Lapita horizon and tradition: Signature for one set of Oceanic migrations. In *Pacific Archaeology: Assessments and Prospects (Proceedings of the International Conference for the 50th Anniversary of the First Lapita Excavation, Koné-Nouméa 2002)*, C. Sand, ed. (Nouméa, New Caledonia: Museum of New Caledonia), pp. 95–120.
16. Melton, T., Peterson, R., Redd, A.J., Saha, N., Sofro, A.S., Martinson, J., and Stoneking, M. (1995). Polynesian genetic affinities with Southeast Asian populations as identified by mtDNA analysis. *Am. J. Hum. Genet.* 57, 403–414.
17. Sykes, B., Leibold, A., Low-Beer, J., Tetzner, S., and Richards, M. (1995). The origins of the Polynesians: An interpretation from mitochondrial lineage analysis. *Am. J. Hum. Genet.* 57, 1463–1475.
18. Stoneking, M., Jorde, L.B., Bhatia, K., and Wilson, A.C. (1990). Geographic variation in human mitochondrial DNA from Papua New Guinea. *Genetics* 124, 717–733.
19. Capelli, C., Wilson, J.F., Richards, M., Stumpf, M.P., Gratrix, F., Oppenheimer, S., Underhill, P., Pascali, V.L., Ko, T.M., and Goldstein, D.B. (2001). A predominantly indigenous paternal heritage for the Austronesian-speaking peoples of insular Southeast Asia and Oceania. *Am. J. Hum. Genet.* 68, 432–443.
20. Kayser, M., Brauer, S., Weiss, G., Schiefenhövel, W., Underhill, P., Shen, P., Oefner, P., Tommaseo-Ponzetta, M., and Stoneking, M. (2003). Reduced Y-chromosome, but not mitochondrial DNA, diversity in human populations from West New Guinea. *Am. J. Hum. Genet.* 72, 281–302.
21. Mona, S., Tommaseo-Ponzetta, M., Brauer, S., Sudoyo, H., Marzuki, S., and Kayser, M. (2007). Patterns of Y-chromosome diversity intersect with the Trans-New Guinea hypothesis. *Mol. Biol. Evol.* 24, 2546–2555.
22. Scheinfeldt, L., Friedlaender, F., Friedlaender, J., Latham, K., Koki, G., Karafet, T., Hammer, M., and Lorenz, J. (2006). Unexpected NRY chromosome variation in Northern Island Melanesia. *Mol. Biol. Evol.* 23, 1628–1641.
23. Friedlaender, J.S., Friedlaender, F.R., Hodgson, J.A., Stoltz, M., Koki, G., Horvat, G., Zhadanov, S., Schurr, T.G., and Merriwether, D.A. (2007). Melanesian mtDNA complexity. *PLoS ONE* 2, e248.
24. Kayser, M., Choi, Y., van Oven, M., Mona, S., Brauer, S., Trent, R.J., Suarika, D., Schiefenhövel, W., and Stoneking, M. (2008). The impact of the Austronesian expansion: Evidence from mtDNA and Y chromosome diversity in the Admiralty Islands of Melanesia. *Mol. Biol. Evol.* 25, 1362–1374.
25. Kayser, M., Brauer, S., Weiss, G., Schiefenhövel, W., Underhill, P.A., and Stoneking, M. (2001). Independent histories of human Y chromosomes from Melanesia and Australia. *Am. J. Hum. Genet.* 68, 173–190.
26. Kayser, M., Brauer, S., Cordaux, R., Casto, A., Lao, O., Zhivotovskiy, L.A., Moyse-Faurie, C., Rutledge, R.B., Schiefenhövel, W., Gil, D., et al. (2006). Melanesian and Asian origins of Polynesians: mtDNA and Y chromosome gradients across the Pacific. *Mol. Biol. Evol.* 23, 2234–2244.
27. Hage, P., and Marck, J. (2003). Matrilineality and the Melanesian origin of Polynesian Y chromosomes. *Curr. Anthropol.* 44, 121–127.
28. Jordan, F.M., Gray, R.D., Greenhill, S.J., and Mace, R. (2009). Matrilocal residence is ancestral in Austronesian societies. *Proc. Biol. Sci.* 276, 1957–1964.
29. Kayser, M., Lao, O., Saar, K., Brauer, S., Wang, X., Nürnberg, P., Trent, R.J., and Stoneking, M. (2008). Genome-wide analysis indicates more Asian than Melanesian ancestry of Polynesians. *Am. J. Hum. Genet.* 82, 194–198.
30. Friedlaender, J.S., Friedlaender, F.R., Reed, F.A., Kidd, K.K., Kidd, J.R., Chambers, G.K., Lea, R.A., Loo, J.H., Koki, G., Hodgson, J.A., et al. (2008). The genetic structure of Pacific Islanders. *PLoS Genet.* 4, e19.
31. Kimura, R., Ohashi, J., Matsumura, Y., Nakazawa, M., Inaoka, T., Ohtsuka, R., Osawa, M., and Tokunaga, K. (2008). Gene flow and natural selection in oceanic human populations inferred from genome-wide SNP typing. *Mol. Biol. Evol.* 25, 1750–1761.
32. Kayser, M. (2010). The human genetic history of Oceania: Near and remote views of dispersal. *Curr. Biol.* 20, R194–R201.
33. Lahr, M.M., and Foley, R. (1994). Multiple dispersals and modern human origins. *Evol. Anthropol.* 3, 48–60.
34. Bulbeck, D. (2007). Where river meets sea: A parsimonious model for *Homo sapiens* colonization of the Indian Ocean rim and Sahul. *Curr. Anthropol.* 48, 315–321.
35. Frost, E.L. (1979). Fiji. In *The Prehistory of Polynesia*, J.D. Jennings, ed. (Canberra, Australia: Australian National University Press), pp. 61–81.
36. Ross, M., Pawley, A., and Osmond, M. (2003). *The Lexicon of Proto Oceanic. The Culture and Environment of Ancestral Oceanic Society, Volume 2* (Canberra, Australia: Pacific Linguistics).
37. Frazer, K.A., Ballinger, D.G., Cox, D.R., Hinds, D.A., Stuve, L.L., Gibbs, R.A., Belmont, J.W., Boudreau, A., Hardenbol, P., Leal, S.M., et al.; International HapMap Consortium. (2007). A second generation human haplotype map of over 3.1 million SNPs. *Nature* 449, 851–861.
38. Fenner, J.N. (2005). Cross-cultural estimation of the human generation interval for use in genetics-based population divergence studies. *Am. J. Phys. Anthropol.* 128, 415–423.
39. Jakobsson, M., Scholz, S.W., Scheet, P., Gibbs, J.R., VanLiere, J.M., Fung, H.C., Szpiech, Z.A., Degnan, J.H., Wang, K., Guerreiro, R., et al. (2008). Genotype, haplotype and copy-number variation in worldwide human populations. *Nature* 451, 998–1003.
40. Clark, A.G., Hubisz, M.J., Bustamante, C.D., Williamson, S.H., and Nielsen, R. (2005). Ascertainment bias in studies of human genome-wide polymorphism. *Genome Res.* 15, 1496–1502.
41. Chakraborty, R., and Weiss, K.M. (1988). Admixture as a tool for finding linked genes and detecting that difference from allelic association between loci. *Proc. Natl. Acad. Sci. USA* 85, 9119–9123.
42. Service, S., DeYoung, J., Karayiorgou, M., Roos, J.L., Pretorius, H., Bedoya, G., Ospina, J., Ruiz-Linares, A., Macedo, A., Palha, J.A., et al. (2006). Magnitude and distribution of linkage disequilibrium in population isolates and implications for genome-wide association studies. *Nat. Genet.* 38, 556–560.
43. Nielsen, R., and Signorovitch, J. (2003). Correcting for ascertainment biases when analyzing SNP data: Applications to the estimation of linkage disequilibrium. *Theor. Popul. Biol.* 63, 245–255.
44. Tenesa, A., Navarro, P., Hayes, B.J., Duffy, D.L., Clarke, G.M., Goddard, M.E., and Visscher, P.M. (2007). Recent human effective population size estimated from linkage disequilibrium. *Genome Res.* 17, 520–526.
45. Fagundes, N.J., Ray, N., Beaumont, M., Neuenschwander, S., Salzano, F.M., Bonatto, S.L., and Excoffier, L. (2007). Statistical evaluation of alternative models of human evolution. *Proc. Natl. Acad. Sci. USA* 104, 17614–17619.
46. Green, R.E., Krause, J., Briggs, A.W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M.H., et al. (2010). A draft sequence of the Neandertal genome. *Science* 328, 710–722.
47. Abdulla, M.A., Ahmed, I., Assawamakin, A., Bhak, J., Brahmachari, S.K., Calacal, G.C., Chaurasia, A., Chen, C.H., Chen, J., Chen, Y.T., et al. HUGO Pan-Asian SNP Consortium/Indian Genome Variation Consortium. (2009). Mapping human genetic diversity in Asia. *Science* 326, 1541–1545.
48. Oppenheimer, S.J., and Richards, M. (2001). Polynesian origins. Slow boat to Melanesia? *Nature* 410, 166–167.
49. Moodley, Y., Linz, B., Yamaoka, Y., Windsor, H.M., Breurec, S., Wu, J.Y., Maady, A., Bernhöft, S., Thiberge, J.M., Phuanukoonnon, S., et al. (2009). The peopling of the Pacific from a bacterial perspective. *Science* 323, 527–530.
50. Diamond, J. (2001). Reply to Oppenheimer and Richards. *Nature* 410, 167.
51. Hurles, M.E., Nicholson, J., Bosch, E., Renfrew, C., Sykes, B.C., and Jobling, M.A. (2002). Y chromosomal evidence for the origins of oceanic-speaking peoples. *Genetics* 160, 289–303.
52. Kayser, M., Brauer, S., Weiss, G., Underhill, P.A., Roewer, L., Schiefenhövel, W., and Stoneking, M. (2000). Melanesian origin of Polynesian Y chromosomes. *Curr. Biol.* 10, 1237–1246.
53. Diamond, J.M. (1988). Express train to Polynesia. *Nature* 336, 307–308.
54. Diamond, J.M. (2000). Taiwan's gift to the world. *Nature* 403, 709–710.
55. Spriggs, M. (2003). Chronology of the Neolithic transition in Southeast Asia and the Western Pacific: A view from 2003. *Rev. Archaeol.* 24, 57–80.
56. Cox, M.P. (2005). Indonesian mitochondrial DNA and its opposition to a Pleistocene era origin of proto-Polynesians in island southeast Asia. *Hum. Biol.* 77, 179–188.

57. Spriggs, M. (1997). *The Island Melanesians* (Oxford: Blackwell Publishers).
58. Lao, O., Lu, T.T., Nothnagel, M., Junge, O., Freitag-Wolf, S., Caliebe, A., Balasckova, M., Bertranpetit, J., Bindoff, L.A., Comas, D., et al. (2008). Correlation between genetic and geographic structure in Europe. *Curr. Biol.* *18*, 1241–1248.
59. Novembre, J., Johnson, T., Bryc, K., Kutalik, Z., Boyko, A.R., Auton, A., Indap, A., King, K.S., Bergmann, S., Nelson, M.R., et al. (2008). Genes mirror geography within Europe. *Nature* *456*, 98–101.
60. Hurles, M.E., Irlen, C., Nicholson, J., Taylor, P.G., Santos, F.R., Loughlin, J., Jobling, M.A., and Sykes, B.C. (1998). European Y-chromosomal lineages in Polynesians: A contrast to the population structure revealed by mtDNA. *Am. J. Hum. Genet.* *63*, 1793–1806.