Denisovan Ancestry in East Eurasian and Native American Populations

Pengfei Qin[1] and Mark Stoneking*[,1]

[1]Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

*Corresponding author: E-mail: stoneking@eva.mpg.de.

Associate editor: Su Bing

Abstract

Although initial studies suggested that Denisovan ancestry was found only in modern human populations from island Southeast Asia and Oceania, more recent studies have suggested that Denisovan ancestry may be more widespread. However, the geographic extent of Denisovan ancestry has not been determined, and moreover the relationship between the Denisovan ancestry in Oceania and that elsewhere has not been studied. Here we analyze genome-wide single nucleotide polymorphism data from 2,493 individuals from 221 worldwide populations, and show that there is a widespread signal of a very low level of Denisovan ancestry across Eastern Eurasian and Native American (EE/NA) populations. We also verify a higher level of Denisovan ancestry in Oceania than that in EE/NA; the Denisovan ancestry in Oceania is correlated with the amount of New Guinea ancestry, but not the amount of Australian ancestry, indicating that recent gene flow from New Guinea likely accounts for signals of Denisovan ancestry across Oceania. However, Denisovan ancestry in EE/NA populations is equally correlated with their New Guinea or their Australian ancestry, suggesting a common source for the Denisovan ancestry in EE/NA and Oceanian populations. Our results suggest that Denisovan ancestry in EE/NA is derived either from common ancestry with, or gene flow from, the common ancestor of New Guineans and Australians, indicating a more complex history involving East Eurasians and Oceanians than previously suspected.

Key words: **Denisovans, Neanderthals, archaic admixture, modern humans.**

Introduction

Following the initial description and analysis of a genome sequence from an archaic human fossil from Denisova Cave in southern Siberia (Reich et al. 2010), Denisovan admixture was subsequently found to be limited to populations from eastern Indonesia, the Philippines, and Near and Remote Oceania (Reich et al. 2010, 2011; Meyer et al. 2012). This finding was quite surprising, given that the Denisova Cave site is located some 7,000 km away from the populations that currently exhibit Denisovan ancestry, and was interpreted as suggesting that Denisovan admixture occurred somewhere in the vicinity of island Southeast Asia (Reich et al. 2011). However, further studies have indicated that Denisovan ancestry may be more widespread than initially thought (Skoglund and Jakobsson 2011). In particular, a recent study inferred low levels of Denisovan ancestry of about 0.2% in three genome sequences, from a Dai and a Han Chinese from East Asia and a Karitiana from South America (Prüfer et al. 2014). Moreover, it appears that a genetic adaptation to high altitude in the Tibetan Plateau occurred through introgression from a Denisovan-related population into ancestral Tibetans (Huerta-Sánchez et al. 2014). Overall, these results suggest that Denisovan ancestry is not limited to populations from island Southeast Asia and Oceania, as originally thought. However, only a few populations have been systematically evaluated for signals of Denisovan ancestry; it is not clear whether a very low level of Denisovan ancestry is geographically widespread, or

rather limited to only a few populations outside of island Southeast Asia and Oceania. An additional question of interest is whether the Denisovan ancestry in these other populations reflects the same admixture event that contributed Denisovan ancestry to island Southeast Asian and Oceanian populations, or a different admixture event.

To address these and other questions related to Denisovan ancestry in human populations, we here present a systematic investigation of Denisovan introgression in Eastern Eurasia (defined here to include South Asia, East Asia, Southeast Asia, Central Asia, Siberia, and Oceania) and in Native American populations, hereafter abbreviated as EE/NA. Genome-wide data were collected from worldwide populations and analyzed along with high coverage genome sequences from a Denisovan (Meyer et al. 2012) and a Neanderthal (Prüfer et al. 2014). The analyses we report provide more details concerning the admixture history of modern humans with Denisovans.

Results

### Relationship of Present-Day Humans and Archaic Hominins

We assembled a data set which, after quality filtering, consisted of 2,493 individuals from 221 populations (supplementary tables S1 and S2, Supplementary Material online), all genotyped on the Affymetrix Human Origins Array (Patterson et al. 2012). After merging the human data with the chimpanzee,

Denisovan, and Neanderthal genome sequences, there were nearly 600,000 single nucleotide polymorphisms (SNPs) for analysis. To investigate the relationship of the diverse present-day human populations relative to archaic hominins, we carried out Principal Component Analysis (PCA) (Price et al. 2006) on the chimpanzee, Neanderthal, and Denisovan data, and projected the modern human samples onto the plane defined by the top two eigenvectors. The human samples all appear at the center of the plot (supplementary fig. S1A, Supplementary Material online); magnification of the central portion of the plot shows that humans separate into three clusters relative to archaic hominins and chimpanzees: African, Oceanian, and other Non-African (supplementary fig. S1B, Supplementary Material online). To more clearly visualize the patterns, we plotted the mean of eigenvectors 1 and 2 for each of the 221 modern human populations (fig. 1A). The first eigenvector separates the Africans from non-Africans and shows that the non-Africans are clearly closer to archaic hominins than are the Africans. The second eigenvector suggests closer genetic affinity between Oceanians and Denisovan than between other populations and Denisovan. There is a clear cline of Denisovan-related ancestry in Oceanians with Australians and New Guineans having the most Denisovan ancestry (supplementary fig. S2, Supplementary Material online). The Mamanwa, from the Philippines, are also involved in this cline, which is consistent with previous findings that the Mamanwa are related to Australians and New Guineans, and Denisovan admixture occurred in a common ancestral population of the Mamanwa, Australians, and New Guineans (Reich et al. 2011).

Additional Archaic Ancestry in EE/NA Populations

PCA is a descriptive analysis that is useful for indicating potential admixture events, but cannot be used to prove that admixture occurred. We therefore applied formal tests to document potential admixture between archaic hominins and modern humans. As EE/NA populations have on average inherited more archaic ancestry than West Eurasian populations (Wall et al. 2013), we computed f_4 statistics (Reich et al. 2009; Patterson et al. 2012) of the form $f_4(Yoruba, Archaic; French, X)$, in which X is an EE/NA population. A significantly positive statistic (Z-score >2) is evidence that EE/NA possesses more Archaic (either Neanderthal or Denisovan) alleles than does the French population. Significantly positive statistics (Z-score >6) are obtained for all Oceanian populations, which are much higher than those for other EE/NA populations, indicating more archaic ancestry in Oceanians (fig. 1B and supplementary table S3, Supplementary Material online). Moreover, there are more Denisovan than Neanderthal alleles shared with the Oceanian and Mamanwa populations (fig. 1B), although Neanderthal ancestry is also elevated, probably because signals of Denisovan and Neanderthal ancestry are difficult to distinguish in this analysis. In addition to the Oceanian populations, many additional EE/NA populations exhibit significant Z-scores (>2), indicating that they have more archaic alleles than the French population has. However, unlike the Oceanian populations, the inferred amounts of Denisovan and Neanderthal alleles are approximately the same in these EE/NA populations (fig. 1B). It is thus not clear from this analysis whether the additional archaic ancestry in these EE/NA populations reflects Neanderthal ancestry, Denisovan ancestry, or both. In order to increase the power of these tests, we combined the data from the East Asian, Siberian, and native American populations, and obtained significantly higher signals of archaic ancestry (Z-score = 3.12 for Neanderthal and 3.64 for Denisovan, compared with the average single population Z-scores of 2.79 ± 0.07 for Neanderthal and 3.17 ± 0.08 for Denisovan).
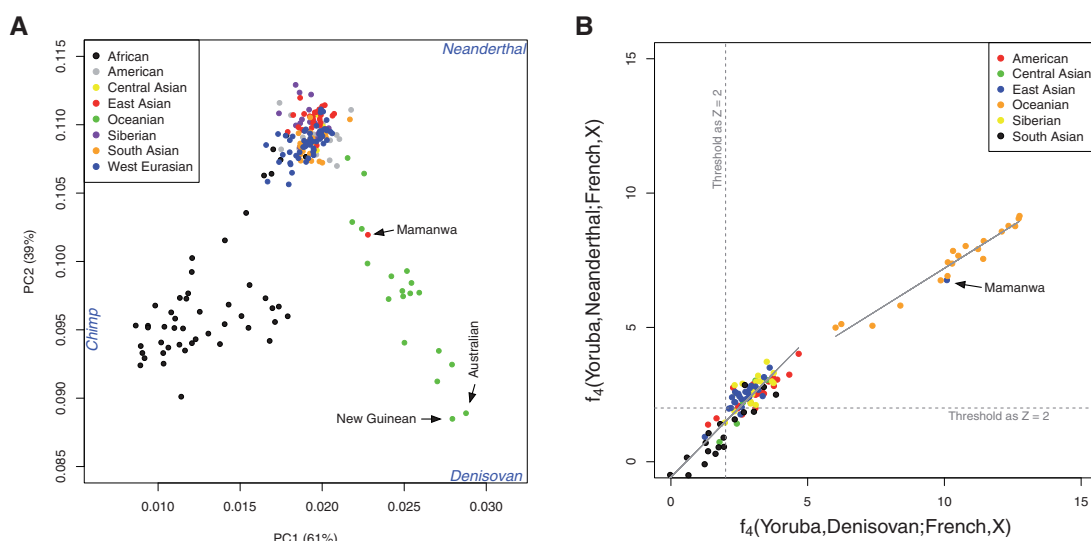
Fig. 1. The relationships among modern human populations relative to archaic humans. (A) PCA of 221 populations projected onto the top two eigenvectors defined by Neanderthal, Denisovan, and chimpanzee. The mean values of eigenvectors 1 and 2 are plotted for each population. (B) Formal admixture tests suggest a substantial number of EE/NA and Oceanian populations inherited significantly more archaic ancestry than West Eurasian populations. We used f_4 statistics of the form $f_4(Yoruba, Archaic; French, X)$ to test admixture between archaic humans and modern human populations. A Z-score larger than 2 was set as the threshold for determining whether admixture is significantly greater than zero.

Although gene flow between the modern human reference populations could inflate the f_4 statistics, the magnitude of the inflation is expected to be small relative to any archaic ancestry (see supplementary text S1, Supplementary Material online). To ensure that our results are not significantly influenced by the choice of African (Yoruba) and European (French) reference populations used in the f_4 analyses, we repeated the analysis with different reference populations and found no significant differences in the results (supplementary fig. S3, Supplementary Material online). Moreover, values of f_4 statistics are perfectly correlated with each other using different pairs of reference populations (supplementary fig. S4, Supplementary Material online).

We further computed the statistic $f_4(Yoruba, X; Neanderthal, Denisovan)$ which compares the genetic affinity of present-day non-Africans with different archaic hominins (fig. 2). Positive values of this f_4 statistic indicate excess sharing of Denisovan alleles (relative to Neanderthals); negative values indicate excess sharing of Neanderthal alleles, and values near zero indicate equivalent amounts of alleles shared with Denisovans and Neanderthals. We obtain larger values in Oceanians than in other non-Africans, which is consistent with the observations based on the PCA and on formal tests of admixture. The largest values are observed in New Guineans, Australians, and some populations from Remote Oceania (supplementary table S4, Supplementary Material online), consistent with previous results (Reich et al. 2011). Negative values are observed in most non-African populations, indicating sharing of more Neanderthal than Denisovan alleles in these populations. Moreover, larger values are observed in East Asians than in West Eurasians,

suggesting that most East Asians share more alleles with Neanderthal than West Eurasians do. Native Americans are generally more similar to West Eurasians in the patterns of allele sharing, suggesting that either Native Americans have the same amount of allele sharing with Neanderthal as West Eurasians do, or Native Americans shared more Neanderthal alleles than West Eurasians and additionally share a small amount of Denisovan alleles. This procedure was repeated using different African populations in the f_4 statistics and similar results were obtained (supplementary fig. S5, Supplementary Material online).

Denisovan Ancestry in Oceanians

Denisovan ancestry in Oceanians has been documented previously (Reich et al. 2010, 2011; Meyer et al. 2012). To verify and extend these previous results, we analyzed a larger set of Oceanian populations that were genotyped on a different platform, and also utilized the high-coverage archaic genomes (Meyer et al. 2012; Prüfer et al. 2014).

Following previous methods (Moorjani et al. 2011; Patterson et al. 2012), we used a ratio of f_4 statistics to estimate the admixture proportion of Denisovans ($P_D(X)$) in Oceanians (see Materials and Methods). As Oceanians retain both Denisovan and Neanderthal ancestry, we used Han Chinese to control for the Neanderthal ancestry in Oceanians, under the assumption that Han and Oceanians share a similar number of Neanderthal alleles. To evaluate the validity of this assumption, we examined the $f_4(Yoruba, Han; Neanderthal, X)$ and $f_4(Yoruba, Han; Denisovan, X)$ statistics for each Oceanian population X. If Han and the X possess similar amounts of Neanderthal alleles, then any changes in the two statistics will be driven by the varying
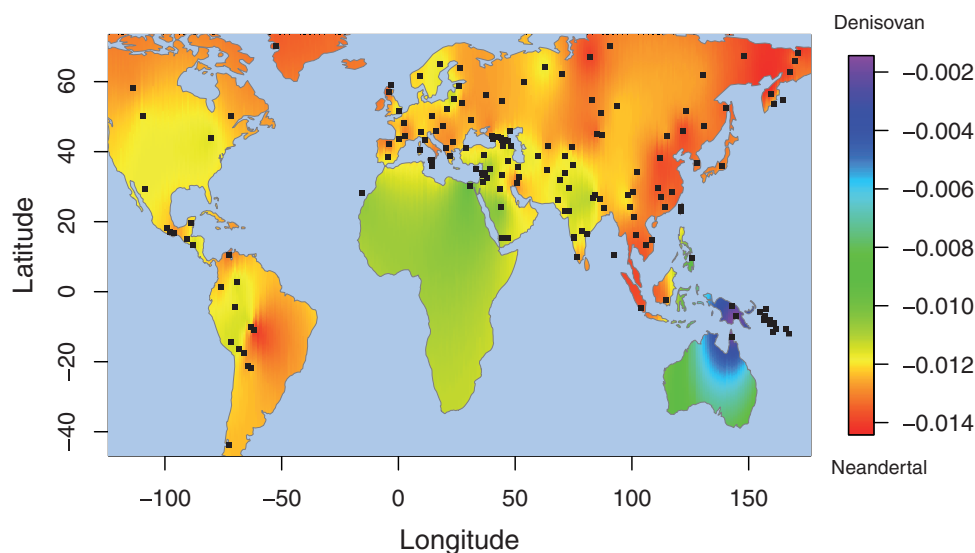
FIG. 2. Archaic introgression in modern humans is prevalent and varies across different geographic regions. The sharing of Neanderthal and Denisovan alleles with each non-African population was measured by f_4 statistics of the form $f_4(Yoruba, X; Neanderthal, Denisovan)$. An excess of allele sharing with Denisovan yields positive values, whereas an excess with Neanderthal yields negative values. The heat plot values indicated on the map are valid only for regions covered by our samples.

amount of their Denisovan alleles. The two statistics will thus have a linear relationship with an intercept close to (0,0). However, if the amount of Neanderthal alleles differs in the two populations, then the linear model will not cross at the point of origin. Empirically, these two sets of f_4 statistics are indeed linearly correlated ($R^2 = 0.99$; supplementary fig. S6, Supplementary Material online) and the intercept for the linear model fitting the data is near (0,0), indicating that Han and Oceanians are similarly close to Neanderthal (see Materials and Methods).

The highest $P_D(X)$ value is observed in Australians and New Guineans (0.034 ± 0.002 and 0.034 ± 0.005, respectively) (fig. 3A), which is consistent with previous results (Reich et al. 2011) and suggests Denisovan introgression into the common ancestor of Australians and New Guineans. We also observed high amounts (>3%) of Denisovan ancestry in Bougainville, as observed previously (Reich et al. 2011), and in Santa Cruz, a population from Remote Oceania which was not analyzed previously. This latter result is in keeping with previous observations of extraordinarily high frequencies and diversity of mtDNA and Y-chromosome haplogroups of New Guinean origin in Santa Cruz (Delfin et al. 2012; Duggan et al. 2014), which suggest high amounts of New Guinean ancestry (and thereby Denisovan ancestry) in Santa Cruz. All other Oceanian populations have Denisovan ancestry ranging from 0.9% to 3%.

It was shown previously that Denisovan ancestry in Oceanian groups (other than Australia and New Guinea), as well as in eastern Indonesia, was likely to be an indirect consequence of admixture with New Guineans, as the Denisovan ancestry in these other Oceanian and eastern Indonesian groups is proportional to their New Guinea ancestry (Reich et al. 2011). We observe a similar correlation for the 17 Oceanian populations (excluding Australia and New Guinea) in this study (fig. 3B). However, given that Australia and New Guinea share common ancestry, it is possible that

the Denisovan ancestry in these Oceanian populations was contributed by admixture from Australia or from the ancestral Australia–New Guinea population, rather than directly from New Guinea; these possibilities were not examined in previous studies. We evaluated these alternative possibilities by the statistic $f_4(Yoruba, X; NewGuinea, Australia)$, and found that Oceanians share significantly more alleles with New Guineans than with Australians (supplementary table S5, Supplementary Material online). These results suggest that Denisovan ancestry in Oceanians is likely to derive from recent admixture with New Guineans, rather than admixture with Australians or the common ancestor of Australians and New Guineans.

Denisovan Introgression in East/South Asian, Siberian, and Native American Populations

To detect Denisovan introgression in EE/NA populations, we computed the ratio of two f_4 statistics ($R_D(X)$): $f_4(Yoruba, Denisovan; French, X)$ and $f_4(Yoruba, Neanderthal; French, X)$, for each EE/NA population. Populations with $R_D(X) > 1$ are likely to have Denisovan ancestry (see Materials and Methods). Before we applied the approach to empirical data, we evaluated the performance of this statistic through simulations (see Materials and Methods and supplementary fig. S7, Supplementary Material online). For simulated populations with or without Denisovan ancestry, we applied a strict threshold of $R_D(X) > 1$, which yields low false positive rates and underestimates true Denisovan ancestry (supplementary fig. S8A and B, Supplementary Material online), and hence is a conservative indicator of Denisovan ancestry. Moreover, gene flow between African and European populations will not inflate the false positive rate and has no obvious impact on the performance of this indicator (supplementary fig. S8C–F, Supplementary Material online). We then computed $R_D(X)$ for all EE/NA populations exhibiting significant admixture signals with Denisovans and/or Neanderthals in
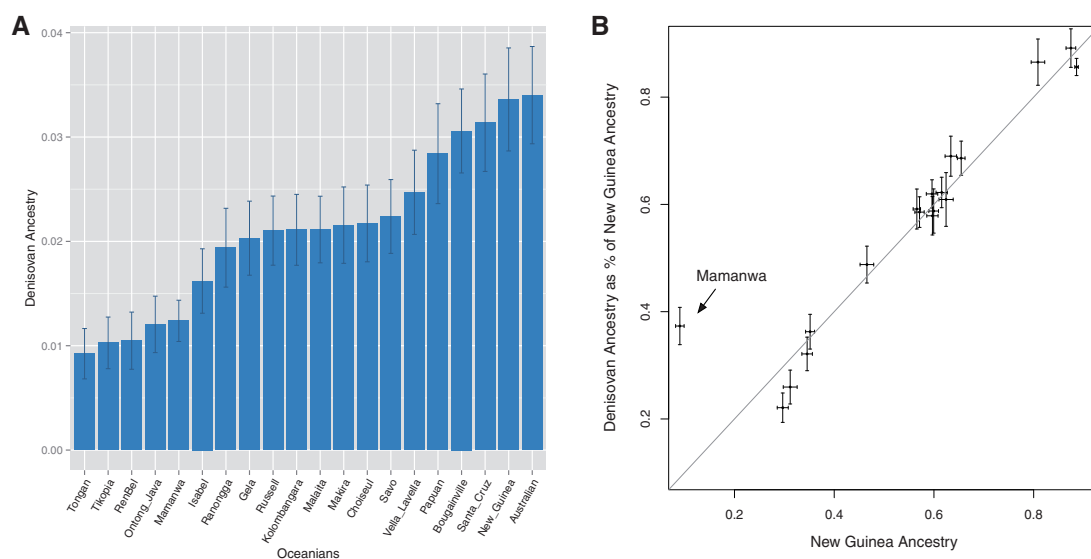

FIG. 3. Denisovan ancestry in Oceanians. (A) Estimated Denisovan ancestry in Oceanian populations. (B) Denisovan ancestry in Oceanians is highly correlated with their New Guinean ancestry.

formal admixture tests (fig. 1B and supplementary table S3, Supplementary Material online). As expected, we observed large ratios ($R_D(X) > 1$) in all Oceanian populations (fig. 4). Variable results were obtained for the other EE/NA populations, with $R_D(X) > 1$ observed in several populations, including most Native American populations. Overall, this analysis indicates that there are EE/NA populations outside Oceania with a clear signal of Denisovan ancestry. Similar results are obtained with the use of different reference populations (supplementary table S6, Supplementary Material online).

Is this presumptive Denisovan ancestry in EE/NA populations from the same admixture event that contributed Denisovan ancestry to Oceanian populations, or does it rather represent a separate admixture event (or events) between modern humans and Denisovans? We postulated that if it reflects the same event that contributed Denisovan ancestry to Oceanians, then the amount of Denisovan ancestry in EE/NA populations should be correlated with the amount of New Guinean ancestry. Because the estimated amount of Denisovan ancestry is quite small in EE/NA populations, and difficult to distinguish from Neanderthal ancestry, we instead compared the overall amount of archaic admixture in EE/NA populations (as a fraction of that in New Guinean), which is calculated by the ratio of $f_4(Yoruba, Denisovan; French, X)$ and $f_4(Yoruba, Denisovan; French, NewGuinean)$, to the statistic $f_4(Yoruba, NewGuinean; French, X)$. These two values are significantly correlated (fig. 5A, Pearson $R^2 = 0.23$, $P = 2.5 \times 10^{-3}$). However, this analysis could be confounded by admixture between East Asians and New

Guineans, which has occurred as a consequence of the Austronesian expansion (Wollstein et al. 2010; Duggan and Stoneking 2014) and perhaps other population movements. We therefore removed the East Asian populations and repeated the analysis just for Siberians and Native Americans, and obtained an even higher correlation (fig. 5B, Pearson $R^2 = 0.54$, $P = 3.8 \times 10^{-6}$). Thus, archaic ancestry in EE/NA populations is significantly correlated with their New Guinea ancestry, suggesting that the Denisovan ancestry in EE/NA and Oceanian populations reflects the same admixture event.

However, there are (at least) two potential alternate scenarios that could explain these results. First, Denisovan admixture could have occurred in a population that was ancestral to both EE/NA and Oceanian populations; second, admixture could have occurred in a population that was ancestral specifically to Mamanwa, Australians, and New Guineans (as suggested previously; Reich et al. 2011), followed by a back-migration from New Guinea to mainland East Asia. This putative back-migration would then have spread both New Guinea and Denisovan ancestry throughout East Asia and Siberia, and ultimately to the Americas. To distinguish between these two scenarios, we repeated the previous analysis but substituted Australians for New Guineans, comparing the archaic admixture in EE/NA populations (as a fraction of that in Australians) to the statistic $f_4(Yoruba, Australian; French, X)$. The results are virtually identical to those obtained with New Guineans as the comparison (fig. 5C and D). Moreover, although Oceanian populations are more closely related to New Guineans than to
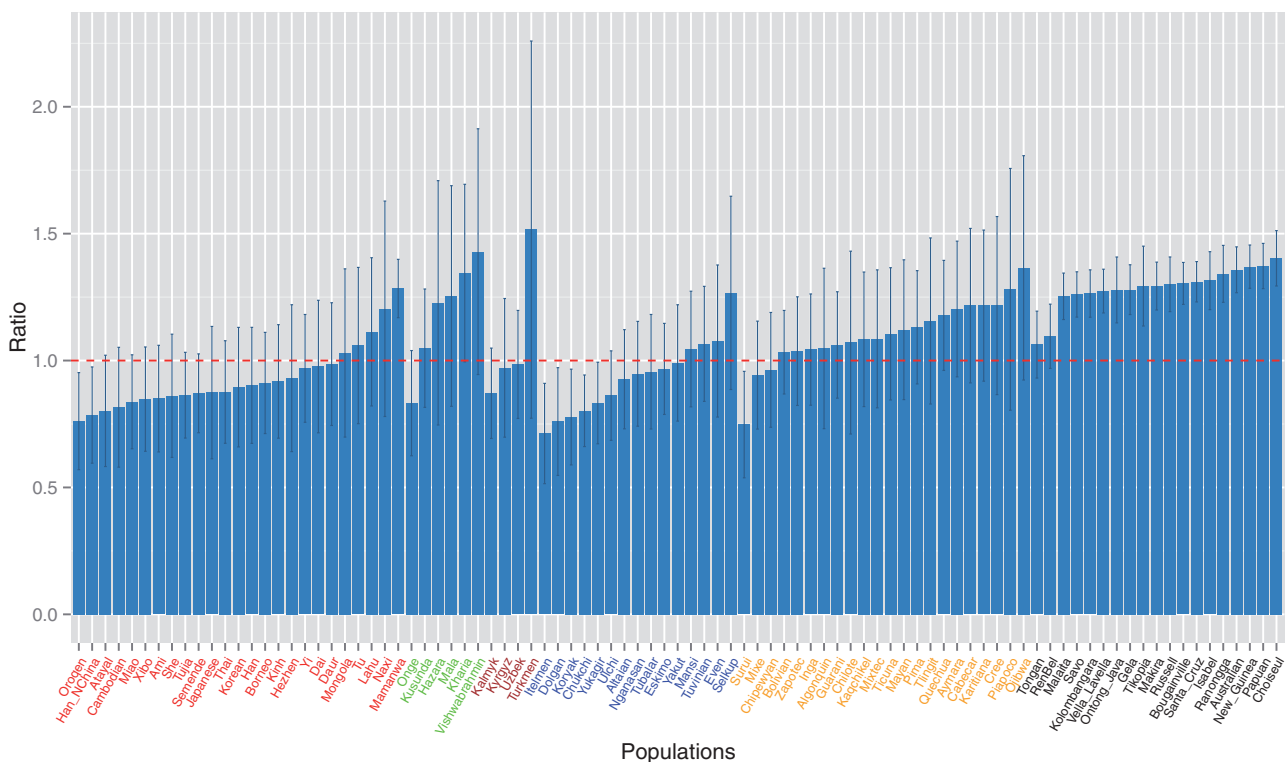
Fig. 4. Widespread Denisovan ancestry in EE/NA populations. Values of the $R_D(X)$ ratio are plotted for all EE/NA populations which give significant signals of admixture with Neanderthal or Denisovan in formal tests; values greater than 1 (dashed line) are indicative of Denisovan ancestry.
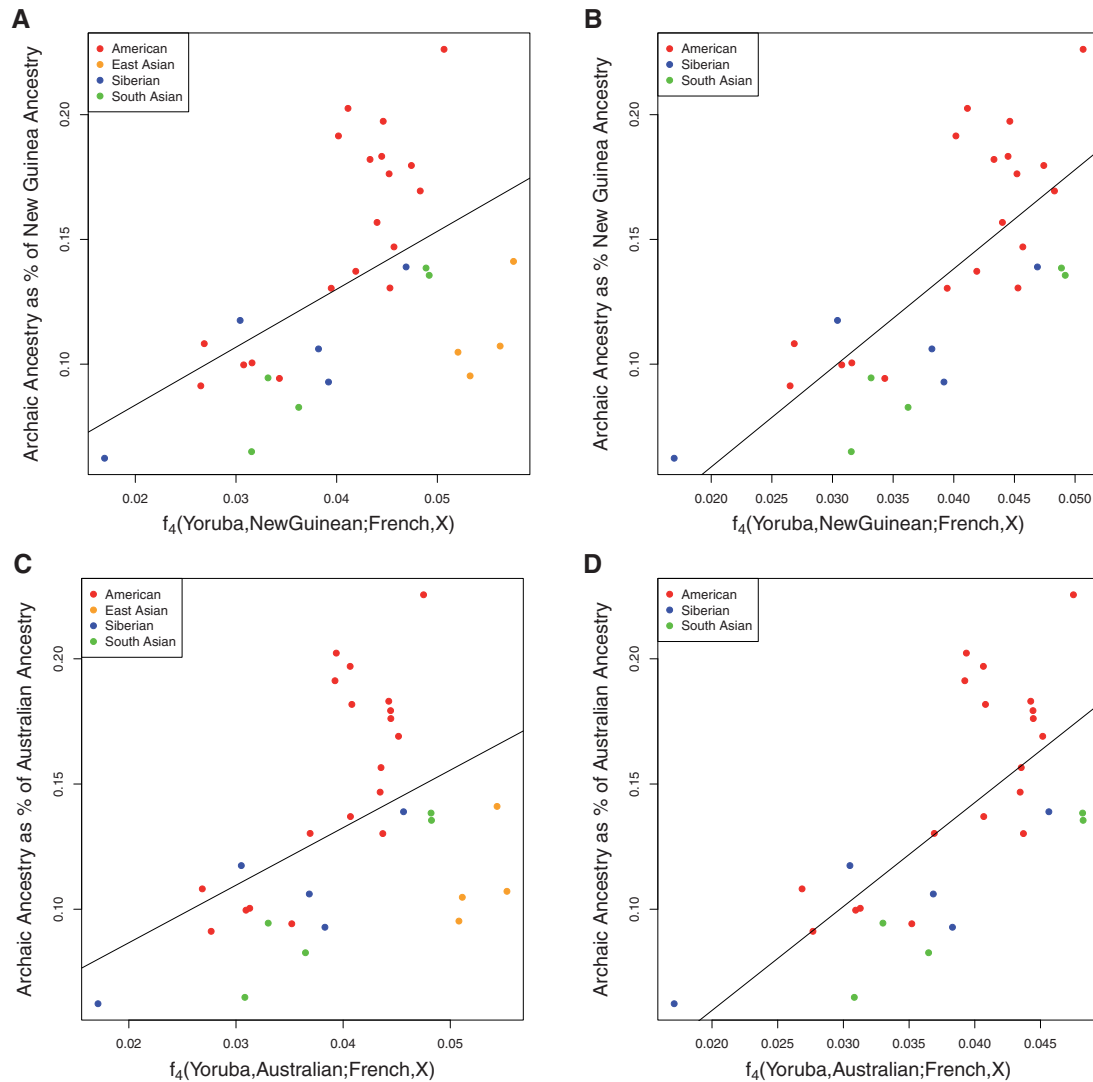
FIG. 5. Denisovan introgression in EE/NA populations is correlated with their genetic affinity with Oceanians. Genetic affinity is measured by $f_4(Yoruba, NewGuinean; French, X)$. We observed a significant correlation between Denisovan ancestry and New Guinea ancestry with (A) $R^2 = 0.23$ ($P = 2.5 \times 10^{-3}$) for all EE/NA populations with $R_D(X) > 1$, and (B) $R^2 = 0.54$ ($P = 3.8 \times 10^{-6}$) when we remove East Asians. Significant correlations were also observed when New Guinean was replaced by Australian with (C) $R^2 = 0.19$ ($P = 6.7 \times 10^{-3}$) and (D) $R^2 = 0.48$ ($P = 2.0 \times 10^{-5}$) when we remove East Asians.

Australians, EE/NA populations are equally related to Australians and New Guineans (supplementary table S5, Supplementary Material online). These results indicate that the archaic ancestry in EE/NA populations is shared with the common ancestor of Australians and New Guineans, and hence reflects the same admixture event. A more detailed model for archaic introgression in modern humans, combining results from this and previous studies (Prüfer et al. 2014; Vernot and Akey 2015), is presented in figure 6.

Discussion

Our analyses demonstrate that, in addition to being prevalent in Oceanian populations, Denisovan introgression is present in East Eurasian and Native American populations, even though the amount of Denisovan alleles in these latter populations is relatively small. These results thus confirm and extend previous studies suggesting Denisovan ancestry

outside of Oceania (Skoglund and Jakobsson 2011; Huerta-Sánchez et al. 2014; Prüfer et al. 2014). In particular, as found previously (Reich et al. 2011), Denisovan ancestry in Oceania is highly correlated with New Guinea ancestry. This suggests that these populations have either shared ancestry or contact with New Guinea that is more recent than the Denisovan admixture event. However, previous studies did not exclude the possibility that more ancient shared ancestry with New Guinean (after the Denisovan admixture event but before the divergence between New Guinean and Australian) explains the correlated signals of Denisovan and New Guinean ancestry in Oceania. Another potential explanation would be migrations from Australian rather than New Guinean, which could still produce a significant correlation between Denisovan and New Guinean ancestry as a consequence of the genetic relationship of Australians and New Guineans. To test these other possibilities, we compared amounts of
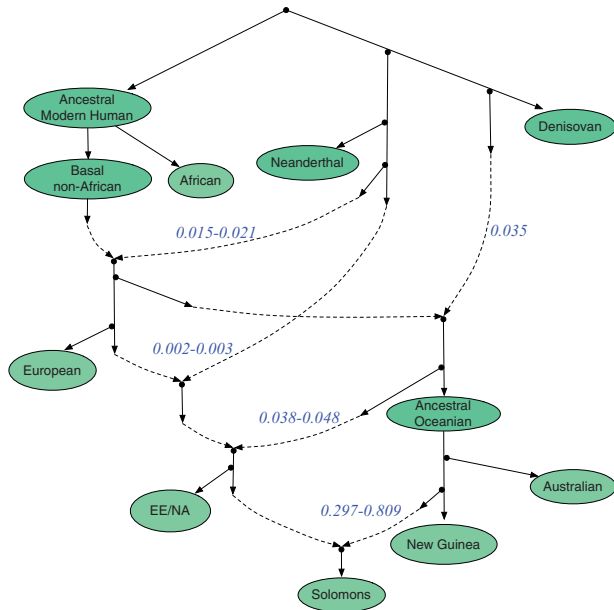
FIG. 6. A model of admixture events between archaic hominins and modern humans. Solid lines represent descent without admixture, and dashed lines represent admixture events. Admixture proportions relating the archaic introgression as estimated from both this study and previous studies are shown.

Denisovan and Australian ancestry in Oceanian populations, and found that New Guinean ancestry does indeed provide a better explanation for the Denisovan ancestry in these Oceanian populations than does Australian ancestry (supplementary table S5, Supplementary Material online).

Our results also show a consistent signal of a low-level of Denisovan ancestry outside of Oceania, in populations of East Eurasia and the Americas. Although this signal does not reach significance in all populations (fig. 4), given how widespread the signal is, it seems most reasonable to assume that all EE/NA populations probably do harbor some Denisovan ancestry. The actual amount of Denisovan ancestry is difficult to estimate directly as it is so low. As it was previously estimated that the Denisovan-related ancestry in EE/NA populations is 3.8–4.8% of that in Oceania (Prüfer et al. 2014), combining this information with our new estimate of approximately 3.5% Denisovan ancestry in New Guinea and Australian leads to estimates of Denisovan ancestry in EE/NA populations of 0.13–0.17%. We would expect full genome sequences to provide more accurate estimates.

As with Oceanian populations, the Denisovan ancestry in EE/NA populations is correlated with their New Guinean ancestry. However, unlike Oceanian populations, the Denisovan ancestry in EE/NA populations is equally correlated with their Australian ancestry, and moreover EE/NA populations are just as closely related to Australians as they are to New Guineans (fig. 5 and supplementary table S5, Supplementary Material online).

There are (at least) two potential scenarios that could account for these results. First, there was introgression from a population related to Denisovans into a modern human population that was ancestral to all EE/NA and Oceanian populations. After the separation of the ancestral EE/NA and Oceanian populations, subsequent migration(s) then brought other modern human ancestry into the ancestors of EE/NA (but not Australian or New Guinean) populations, thereby "diluting" Denisovan ancestry in EE/NA populations. This scenario has two important consequences. First, it means that the introgression between Denisovans and modern humans did not necessarily occur in island Southeast Asia as postulated previously (Reich et al. 2011), but instead could have occurred closer to the vicinity of Denisova Cave, in southern Siberia. Second, identifying the source(s) of the other modern human ancestry in EE/NA populations would be of great interest for further understanding the genetic history of human populations.

The second scenario that could explain the different amounts of Denisovan ancestry in Oceanians versus EE/NA populations would be that Denisovan introgression occurred specifically in a population ancestral to Australians, New Guineans, and the Mamanwa, as hypothesized previously (Reich et al. 2011). After the Denisovan admixture, but before the divergence of Australians and New Guineans, there would then have been a back-migration from Oceania to mainland East Asia, which would have contributed both Denisovan and shared Australian/ New Guinea ancestry to the ancestors of present day EE/NA populations. Although we are not aware of any previous suggestions of such a back-migration from archaeological, anthropological, or genetic evidence, we also are not aware of any evidence that would disprove such a back-migration.

It thus seems that at present these two scenarios are equally plausible explanations for our results. However, ancient DNA analyses could potentially distinguish between them. Current results indicate that modern humans colonizing Oceania split from ancestral Eurasian as early as 62,000–75,000 years before present (BP) (Rasmussen et al. 2011), whereas East Asians subsequently diverged from Europeans much later, around 23,000–38,000 years BP (Keinan et al. 2007; Gutenkunst et al. 2009; Gravel et al. 2011). If ancient DNA from East Asian fossils that predate the divergence between East Asians and Europeans contains high levels of Denisovan ancestry (related to the Denisovan ancestry in Oceania), this would support the first scenario. But if high levels of Denisovan ancestry remain restricted to Oceania, with only low levels found outside Oceania even in ancient DNA, then this would support the second scenario.

In any event, the inescapable conclusion is that Denisovan ancestry is more widespread in modern human populations than thought previously to be the case, and moreover human genetic history must consequently also be more complicated than previously believed. Mapping the segments of Denisovan ancestry in modern human populations, as has been done for Neanderthal ancestry (Prüfer et al. 2014; Sankararaman et al. 2014; Vernot and Akey 2014), should provide more

insights into the history and consequences of the interactions between Denisovans and modern humans.

Materials and Methods

Population Samples and Data

We report 168 new samples from 20 populations of Oceania and Southeast Asia (supplementary table S1, Supplementary Material online) genotyped on the Affymetrix Human Origins SNP Array, in which SNPs are cleanly ascertained (Patterson et al. 2012). We merged the new data with those previously released samples (Lazaridis et al. 2014) genotyped on the same array, resulting in 2,890 samples from 236 world-wide modern human populations, and 1 chimpanzee sample. The full data set was filtered to 2,493 individuals from 221 populations (supplementary table S2, Supplementary Material online) after removing outlier individuals or relatives, based on visual inspection of PCA plots and/or model-based clustering analysis as described previously (Lazaridis et al. 2014). We then merged high coverage sequences of archaic hominins into the data set: Altai Neanderthal (52×) (Prüfer et al. 2014) and Denisovan (31×) (Meyer et al. 2012). Problematic SNPs were filtered out following the approach described previously (Lazaridis et al. 2014), and analyses were carried out on a set of 594,903 autosomal SNPs.

Principal Component Analysis

PCA was performed with EIGENSOFT (Patterson et al. 2006) version 5.0.1. We performed PCA on a subset of individuals from which the top two eigenvectors could determine a plane and then projected others onto the plane. To explore the relationship between modern humans and archaic hominins, we carried out PCA on the chimpanzee, Neanderthal, and Denisovan only and projected present-day human samples onto the map defined by the top two eigenvectors of archaic hominins and chimpanzee.

Statistical Analysis to Detect Archaic Ancestry in Modern Humans

We applied f_4 statistics (Reich et al. 2009; Patterson et al. 2012) with the form $f_4(A, B; C, D) = \frac{\sum_{i=1}^{N}(p_A - p_B) \times (p_C - p_D)}{\sum_{i=1}^{N} p_{Outgroup} \times (1 - p_{Outgroup})}$, where p is the allele frequency for populations A, B, C, D or outgroup, to assess the correlation between the allele frequency differences of the two pairs of populations. If populations A and B are consistent with forming a clade in an unrooted tree with respect to populations C and D, the statistic is expected to be 0. We computed f_4 statistics with the form $f_4(African, Archaic; WestEurasian, EE/NA)$ to detect gene flow between archaic hominins and non-Africans. Additional gene flow from archaic hominins to EE/NA will yield significant positive values, whereas additional gene flow from archaic humans to West Eurasians will yield significant negative values. A Weighted Block Jackknife procedure (Kunsch 1989; Busing et al. 1999), which drops 5-centimorgan (cM) blocks of the genome in each run, was used to compute standard errors. These f_4

statistics have been shown to be robust to ascertainment bias, which will affect the observed magnitude of these statistics, but almost never causes a deviation from zero and hence do not cause significant signals of admixture (Patterson et al. 2012).

Estimating Denisovan Ancestry in Oceanians

Oceanians have been previously shown to contain both Denisovan and Neanderthal ancestry (Reich et al. 2010, 2011). We computed the proportion of Denisovan ancestry in Oceanians $P_D(X)$ by the following f_4 ratio :

$$P_D(X) = \frac{f_4(Yoruba, Neanderthal; Han, X)}{f_4(Yoruba, Neanderthal; Han, Denisovan)},$$

which assumes that Han and Oceanians have similar amounts of Neanderthal ancestry.

To evaluate whether Neanderthal ancestry in Oceanians is indeed similar to that in Han, we investigated the following f_4 statistics (Reich et al. 2009; Patterson et al. 2012) for each population of Oceania:

$$f_4(Yoruba, Neanderthal; Han, X) = W \times q + (W + M) \times r$$

$$f_4(Yoruba, Denisovan; Han, X) = (W + L) \times q + W \times r$$

in which W, M, and L are the quantitative measures of branch length in the phylogenetic model (supplementary fig. S9, Supplementary Material online), and q and r are the proportion of Denisovan ancestry and Neanderthal ancestry, respectively. If there is no Neanderthal ancestry ($r = 0$), then the values of these two f_4 statistics should be correlated and correspond to a linear model crossing the origin point (0, 0) with

$$Slope = \frac{f_4(Yoruba, Neanderthal; Han, X_i) - f_4(Yoruba, Neanderthal; Han, X_j)}{f_4(Yoruba, Denisovan; Han, X_i) - f_4(Yoruba, Denisovan; Han, X_j)}$$

$$= \frac{W \times q_i - W \times q_j}{(W + L) \times q_i - (W + L) \times q_j} = \frac{W}{W + L}.$$

Otherwise, Neanderthal ancestry ($r > 0$) will yield a linear model with the same slope

$$Slope = \frac{f_4(Yoruba, Neanderthal; Han, X_i) - f_4(Yoruba, Neanderthal; Han, X_j)}{f_4(Yoruba, Denisovan; Han, X_i) - f_4(Yoruba, Denisovan; Han, X_j)}$$

$$= \frac{W \times (q_i - q_j) + (r_i - r_j) \times (W + M)}{(W + L) \times (q_i - q_j) + (r_i - r_j) \times W} \approx \frac{W}{W + L}$$

and the same intercept (0,0) if $r_i \approx r_j$. If $r_i \neq r_j$, then the intercept will be shifted. With Han as the comparison to different Oceanian populations, the intercept does not differ from (0,0), whereas replacing Han with French results in an intercept that is significantly different from (0,0) (supplementary fig. S6, Supplementary Material online). Thus, this analysis indicates that the amount of Neanderthal ancestry is roughly the same in Han and Oceanians, but not in French and Oceanians.

Detecting Denisovan Introgression in East Eurasian and Native American Populations

To detect Denisovan introgression in EE/NA populations, we investigated the ratio of two f_4 statistics

$$R_D(X) = \frac{f_4(Yoruba, Denisovan; French, X)}{f_4(Yoruba, Neanderthal; French, X)}.$$

Given the null hypothesis that there is no Denisovan admixture in population X, the expected value of the numerator should be $W \times r$, which is smaller than the expected value of the denominator ($[W + M] \times r$), in which W and M are the branch lengths in the phylogenetic model (the same as that in supplementary fig. S9, Supplementary Material online, replacing Han with French), and r is the admixture proportion from Neanderthals. If population X has Denisovan ancestry, then the numerator could be larger than the denominator, depending on the admixture proportion q from Denisovans and the branch length L. Thus, populations with large ratios ($R_D(X) > 1$) are inferred to have Denisovan ancestry.

Simulations for Evaluating $R_D(X)$

We applied coalescent simulations implemented in "ms" (Hudson 2002) to evaluate the performance of $R_D(X)$ in investigating Denisovan ancestry in EE/NA populations. We specified the population splitting time between modern humans and the common ancestor of Neanderthals and Denisovans to be 560 ka, and the splitting time between Neanderthals and Denisovans to be 380 ka (Prüfer et al. 2014). We assumed a generation time of 29 years (Fenner 2005) and mutation rate of 5×10^{-8} per base pair per year. We set the Neanderthal introgression into the common ancestor of both West and East Eurasian populations as 0.02. As it has been recently estimated that there was about 15% (with 95% confidence interval 6.8–26.6%) more Neanderthal introgression into the ancestors of East Asians (Vernot and Akey 2015), we simulated the archaic introgression in East Eurasian populations under two different scenarios: 1) Additional Neanderthal admixture (ranging from 0.002 to 0.006) but without Denisovan admixture, and 2) both Denisovan (ranging from 0.001 to 0.005) and additional Neanderthal (0.003) admixture. To evaluate the influence of genetic interactions between African and West Eurasian populations on our indicator $R_D(X)$, we also simulated different levels of gene flow between African and West Eurasian populations.

We sampled 20 haploid sequences from each simulated modern human population and 2 from each simulated archaic hominins. Around 600K markers were obtained for each simulation, to match the empirical data. For each set of parameters we repeated the simulations 1,000 times, using the demographic model shown in Supplementary figure S7, Supplementary Material online.

Supplementary Material

Supplementary text S1, figures S1–S10, and tables S1–S6 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

References

Busing F, Meijer E, Leeden RVD. 1999. Delete-m jackknife for unequal m. *Stat Comput.* 9:3–8.

Delfin F, Myles S, Choi Y, Hughes D, Illek R, van Oven M, Pakendorf B, Kayser M, Stoneking M. 2012. Bridging near and remote Oceania: mtDNA and NRY variation in the Solomon Islands. *Mol Biol Evol.* 29:545–564.

Duggan AT, Evans B, Friedlaender FR, Friedlaender JS, Koki G, Merriwether DA, Kayser M, Stoneking M. 2014. Maternal history of Oceania from complete mtDNA genomes: contrasting ancient diversity with recent homogenization due to the Austronesian expansion. *Am J Hum Genet.* 94:721–733.

Duggan AT, Stoneking M. 2014. Recent developments in the genetic history of East Asia and Oceania. *Curr Opin Genet Dev.* 29C:9–14.

Fenner JN. 2005. Cross-cultural estimation of the human generation interval for use in genetics-based population divergence studies. *Am J Phys Anthropol.* 128:415–423.

Gravel S, Henn BM, Gutenkunst RN, Indap AR, Marth GT, Clark AG, Yu F, Gibbs RA, Bustamante CD. 2011. Demographic history and rare allele sharing among human populations. *Proc Natl Acad Sci U S A.* 108:11983–11988.

Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD. 2009. Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genet.* 5:e1000695.

Hudson RR. 2002. Generating samples under a Wright-Fisher neutral model of genetic variation. *Bioinformatics* 18:337–338.

Huerta-Sánchez E, Jin X, Asan, Bianba Z, Peter BM, Vinckenbosch N, Liang Y, Yi X, He M, Somel M. 2014. Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature* 512:194–197.

Keinan A, Mullikin JC, Patterson N, Reich D. 2007. Measurement of the human allele frequency spectrum demonstrates greater genetic drift in East Asians than in Europeans. *Nat Genet.* 39:1251–1255.

Kunsch H. 1989. The jackknife and the bootstrap for general stationary observations. *Ann Stat.* 17:1217–1241.

Lazaridis I, Patterson N, Mittnik A, et al. 2014. Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* 513:409–413.

Meyer M, Kircher M, Gansauge M-T, et al. 2012. A high-coverage genome sequence from an archaic Denisovan individual. *Science* 338:222–226.

Moorjani P, Patterson N, Hirschhorn J. 2011. The history of African gene flow into Southern Europeans, Levantines, and Jews. *PLoS Genet.* 7:e1001373.

Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, Genschoreck T, Webster T, Reich D. 2012. Ancient admixture in human history. *Genetics* 192:1065–1093.

Patterson N, Price AL, Reich D. 2006. Population structure and eigenanalysis. *PLoS Genet.* 2:e190.

Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. 2006. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet.* 38:904–909.

Prüfer K, Racimo F, Patterson N, et al. 2014. The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* 505:43–49.

Rasmussen M, Guo X, Wang Y, et al. 2011. An aboriginal Australian genome reveals separate human dispersals into Asia. *Science* 334:94–98.

Reich D, Green RE, Kircher M, et al. 2010. Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature* 468:1053–1060.

Reich D, Patterson N, Kircher M, et al. 2011. Denisova admixture and the first modern human dispersals into Southeast Asia and Oceania. *Am J Hum Genet.* 89:516–528.

Reich D, Thangaraj K, Patterson N, Price AL, Singh L. 2009. Reconstructing Indian population history. *Nature* 461:489–494.

Sankararaman S, Mallick S, Dannemann M, Prüfer K, Kelso J, Pääbo S, Patterson N, Reich D. 2014. The genomic landscape of Neanderthal ancestry in present-day humans. *Nature* 507:354–357.

Skoglund P, Jakobsson M. 2011. Archaic human ancestry in East Asia. *Proc Natl Acad Sci U S A.* 108:18301–18306.

Vernot B, Akey JM. 2014. Resurrecting surviving Neandertal lineages from modern human genomes. *Science* 343:1017–1021.

Vernot B, Akey JM. 2015. Complex history of admixture between modern humans and Neandertals. *Am J Hum Genet.* 96:448–453.

Wall JD, Yang MA, Jay F, et al. 2013. Higher levels of neanderthal ancestry in East Asians than in Europeans. *Genetics* 194:199–209.

Wollstein A, Lao O, Becker C, Brauer S, Trent RJ, Nürnberg P, Stoneking M, Kayser M. 2010. Demographic history of Oceania inferred from genome-wide data. *Curr Biol.* 20:1983–1992.