

Global Genetic Variation at *OAS1* Provides Evidence of Archaic Admixture in Melanesian Populations

Fernando L. Mendez,¹ Joseph C. Watkins,² and Michael F. Hammer^{*,1,3}

¹Department of Ecology and Evolutionary Biology, University of Arizona

²Department of Mathematics, University of Arizona

³ARL Division of Biotechnology, University of Arizona

*Corresponding author: E-mail: mfh@u.arizona.edu.

Associate editor: Noah Rosenberg

Abstract

Recent analysis of DNA extracted from two Eurasian forms of archaic human shows that more genetic variants are shared with humans currently living in Eurasia than with anatomically modern humans in sub-Saharan Africa. Although these genome-wide average measures of genetic similarity are consistent with the hypothesis of archaic admixture in Eurasia, analyses of individual loci exhibiting the signal of archaic introgression are needed to test alternative hypotheses and investigate the admixture process. Here, we provide a detailed sequence analysis of the innate immune gene *OAS1*, a locus with a divergent Melanesian haplotype that is very similar to the Denisova sequence from the Altai region of Siberia. We resequenced a 7-kb region encompassing the *OAS1* gene in 88 individuals from six Old World populations (San, Biaka, Mandenka, French Basque, Han Chinese, and Papua New Guineans) and discovered previously unknown and ancient genetic variation. The 5' region of this gene has unusual patterns of diversity, including 1) higher levels of nucleotide diversity in Papuans than in sub-Saharan Africans, 2) very deep ancestry with an estimated time to the most recent common ancestor of >3 myr, and 3) a basal branching pattern with Papuan individuals on either side of the rooted network. A global geographic survey of >1,500 individuals showed that the divergent Papuan haplotype is nearly restricted to populations from eastern Indonesia and Melanesia. Polymorphic sites within this haplotype are shared with the draft Denisova genome over a span of ~90 kb and are associated with an extended block of linkage disequilibrium, supporting the hypothesis that this haplotype introgressed from an archaic source that likely lived in Eurasia.

Key words: human origins, hybridization, Denisova, TMRCA, linkage disequilibrium.

Introduction

Recent publications of the draft genomes of two archaic forms of *Homo sapiens* have reinvigorated the debate on the structure of ancient human populations. Although it is broadly accepted that the majority of extant human genome diversity traces to anatomically modern populations that lived in Africa ~100 thousand years ago (kya), the question of the extent of admixture with populations of archaic humans is still open. Full genome comparisons of the Neandertal and Denisova draft genomes with modern human sequences have revealed different amounts of shared ancestry between each of these archaic forms and anatomically modern human (AMH) populations from different geographic regions. For example, a higher proportion of single nucleotide polymorphisms (SNPs) was shared between non-African and Neandertal and between Melanesian and the Denisova genomes, than between either Neandertal or Denisova and extant African genomes (Green et al. 2010; Reich et al. 2010). An intriguing possibility is that these patterns result from introgression of archaic genes into AMH populations in Eurasia. However, this SNP sharing pattern could also be explained by ancestral population structure in Africa (i.e., without the need to posit introgression). For example, if non-Africans and the ancestors of Neandertals descend from the same deme in

a subdivided African population, and this structure persisted with low levels of gene flow among African residents until the ancestors of non-Africans migrated into Eurasia, then we would expect more SNP sharing between non-Africans and Neandertals (Durand et al. 2011).

As average measures of genetic similarity across the genome are compatible with several hypotheses (Reich et al. 2010; Durand et al. 2011), detailed studies of genetic regions that have introgressed are needed to distinguish among alternatives and increase our understanding of the admixture process. Unlike demographic processes that have similar effects across the genome, archaic introgression is not expected to impact all loci equally. After interbreeding occurs, the processes of recombination and genetic drift will leave a patchwork of loci carrying the signal of introgression. Analysis of the distribution of these loci and their patterns of variation may provide increased power to characterize the introgression process.

Here, we identify an *OAS1* haplotype that exhibits the signature of archaic introgression. The *OAS1* gene encodes for an enzyme (2'-5' oligoadenylate synthetase 1) with an important antiviral function that is involved in the innate immune response. Association studies have implicated variants at *OAS1* with susceptibility to infectious diseases (e.g., West Nile Fever and severe acute respiratory syndrome

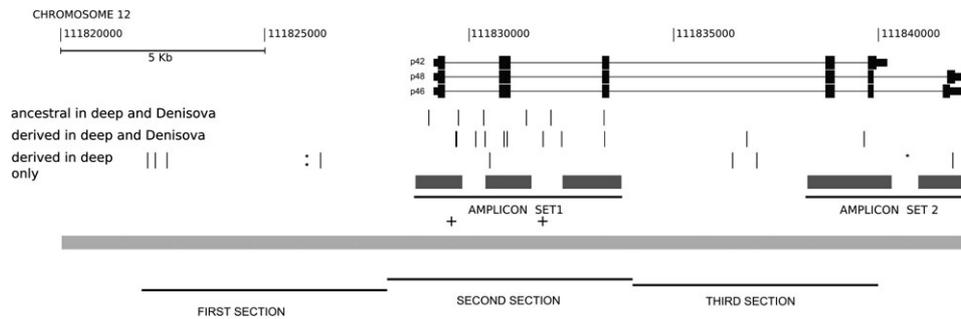


Fig. 1. Sequenced region on chromosome 12. Schematic representation of three splicing variants for isoforms p42, p48, and p46 of the gene *OAS1* (top), amplicons in sets 1 and 2 (gray boxes), and the region covered by extended resequencing (gray bar). Exons are represented by black boxes, joined by thin horizontal lines representing introns. The approximate positions of two genotyped SNPs are indicated with a plus sign “+.” Coordinates correspond to the hg18 build of the human genome. Three rows of thin vertical lines indicate sites ancestral only in the deep lineage (first row), derived mutations common only to the deep lineage and the Denisova sequence (second row), and mutations present only in the deep lineage (third row). Polymorphisms with no sequence coverage in Denisova are indicated with asterisks “*” (third row). Amplicon sets 1 and 2 are analyzed in the main text and [supplementary material](#) ([Supplementary Material](#) online), respectively. The three sections for the extended resequencing are 6 kb each and are compared on the values of divergence between the deep lineage and the reference (see Results).

[SARS]) (Hamano et al. 2005; Lim et al. 2009), severity of Hepatitis C infections (Knapp et al. 2003), and susceptibility to the autoimmune disorder multiple sclerosis (Fedetz et al. 2006). These findings make *OAS1* a good candidate for studies of local adaptation and differentiation among populations. We resequence a DNA segment of >7 kb encompassing this gene in six human populations from Africa, Europe, Asia, and Melanesia. In the 5' region of the *OAS1* gene, we find the highest levels of nucleotide diversity in samples from Melanesia and identify a divergent Melanesian haplotype that shares SNPs with the Denisova *OAS1* sequence. We also perform extended resequencing on chromosomes carrying the unusual haplotype, survey diagnostic SNPs in a large global panel of samples, measure levels of linkage disequilibrium (LD), and estimate divergence times for the branching between the unusual lineage and the human reference sequence. These results are consistent with a model of recent introgression of a segment containing the *OAS1* gene from an archaic source that likely lived in Eurasia.

Materials and Methods

Samples

Three panels of samples were used in this work. The first panel (referred to as the “diversity” panel), previously described (Wall et al. 2008; Hammer et al. 2010), consists of 88 individuals from three sub-Saharan African populations (10 San from Tsumkwe in Namibia, 16 Mandenka from Senegal, 15 Biaka Pygmy from the Central African Republic), one European population (16 Basque from France), one East Asian population (16 Han from China), and one Oceanian population (15 Papuans from the highlands of New Guinea). The second panel, referred to as the “Melanesian” panel, consists of five Melanesian individuals (four Papua New Guineans and one from Vanuatu) that are homozygous for an unusually divergent haplotype (see Results and Discussion). The diversity and Melanesian panels were used for resequencing surveys, whereas a third panel (the “genotyping” panel) was used for genotyping particular SNPs

within the *OAS1* region. The genotyping panel includes 1,572 samples from globally distributed populations (Karafet et al. 1999, 2001; Hammer et al. 2001; Cox et al. 2008).

Resequencing and Genotyping

In what follows, all positions are based on the 2006 build of the human genome (hg18). Five amplicons covering 7,048 bp of the *OAS1* gene were sequenced in all individuals in the diversity panel as well as in chimpanzee, bonobo, and gorilla. The sequence, which spans bases 111828677–111842221, covers the first six exons of the gene and contains all the information for the isoforms p42, p46, and p48 (fig. 1; [supplementary table S1](#), see [supplementary table S2a](#) for primer sequences, [supplementary material](#), [Supplementary Material](#) online). For simplicity, we abbreviate positions on the chromosome by referring only to the last five digits, that is, by dropping the “1118” throughout the main text. The Melanesian panel was resequenced for an extended region that includes positions 20038–42221 (for additional primer sequences, see [supplementary table S2b](#), [Supplementary Material](#) online). The genotyping panel was typed using restriction digests for SNPs diagnostic of newly described lineages ([supplementary table S2c](#), [Supplementary Material](#) online). The inferred geographic occurrences of the deep lineage (described in Results) were plotted using Generic Mapping Tools (Wessel and Smith 1998).

Sequence Analysis

The ancestral state at each polymorphic site in humans was inferred using sequences of the great apes. The polymorphic point mutations in each population were used to construct the corresponding frequency spectra. Estimates of the population mutation rate (Watterson 1975; Tajima 1983) and Tajima’s *D* values (Tajima 1989) were obtained for each individual amplicon ([supplementary material](#), [Supplementary Material](#) online) and for two sets of amplicons: amplicon set 1 (amplicons 1–3) and amplicon

set 2 (amplicons 4–5) (fig. 1). These two amplicon sets, which consist of a similar number of nucleotides (3,733 and 3,315 bp, respectively) and encompass exons 1–3 and 4–6, respectively, are separated by a noncoding region of ~4.5 kb.

Phasing and Networks

PHASE v2.1 (Li and Stephens 2003; Stephens and Donnelly 2003) was used to perform phasing of the resequence data as well as to infer recombination rates from Yoruba HapMap phase II data (www.hapmap.org, supplementary material, Supplementary Material online). Median-joining networks were generated using Network 4.5.1.6 (<http://www.fluxus-engineering.com/sharenet.htm>) with equal weight for each mutation. Only SNPs were used to generate the network. A recombinant Papuan individual as well as a potentially gene converted polymorphism at position 29579 were removed from the network of amplicon set 1 (supplementary material, Supplementary Material online).

Divergence Time and TMRCA

To estimate the divergence time for a pair of human sequences, we assumed the divergence time between human and chimpanzee to be 6 Ma. First, we compared the sequences of two human lineages and a single chimpanzee and estimated the ratio between the human–human and human–chimpanzee divergence times. We then multiplied this ratio by the divergence time between human and chimpanzee lineages to obtain an estimate of the divergence time for the pair of human lineages. The ratio of the human–human and human–chimpanzee divergence times was estimated by first considering the point mutations derived on the human lineages (we ignore mutations in the chimpanzee lineage). Given the total number of mutations and assuming a constant mutation rate, the number of mutations shared between the human lineages, or private to each lineage, is distributed according to a multinomial distribution (Mendez et al. 2011). The parameters for the multinomial distribution are the length of the branches relative to that of the whole genealogy. When both human sequences were available (modern human lineages), the mutations private to each lineage were pooled. In comparisons involving the incomplete Denisova sequence, only the mutations in the modern human lineage were used (supplementary material, Supplementary Material online). This approach is necessary because the sampling time for the Denisova sequence is unknown, and areas of low sequence coverage may result in a higher SNP error rate.

A suite of demographic histories was considered to analyze the intraallelic variation in the deep lineage (supplementary material, Supplementary Material online). The distribution of time to the most recent common ancestor (TMRCA) values for perfectly matching frequency spectra was used to estimate median values and confidence intervals. The estimate including all considered demographic models is reported, but the results are not particularly sensitive to the model used (supplementary material, Supplementary Material online).

Table 1. Summary Statistics for Amplicon Sets 1 and 2 of OAS1.

Population	Amplicon Set 1(3,733 bp)					Amplicon Set 2(3,315 bp)				
	<i>n</i> ^a	<i>S</i>	θ_w (%) ^b	π (%) ^b	<i>D</i>	<i>S</i>	θ_w (%) ^b	π (%) ^b	<i>D</i>	
Biaka	30	10	0.068	0.083	0.69	6	0.046	0.027	−1.18	
Mandenka	32	13	0.087	0.089	0.07	18	0.136	0.046	−2.26	
San	20	8	0.061	0.057	−0.26	22	0.187	0.149	−0.84	
Papuans	30	18	0.123	0.173	1.41	12	0.092	0.137	1.61	
Han	32	6	0.040	0.033	−0.49	7	0.053	0.056	0.16	
Basque	32	6	0.040	0.035	−0.38	5	0.038	0.040	0.15	

^a Number of chromosomes in the sample.

^b Calculated per base.

Results

Five amplicons sequenced in 88 individuals (supplementary table S3, Supplementary Material online) were grouped in two amplicon sets (1 and 2) as described above. In the following sections we analyze sequence variability in the 5′ region of OAS1, which exhibits unusual patterns of diversity. For a description of variation in the 3′ region, consult the supplementary material (Supplementary Material online).

Sequence Diversity

Levels of nucleotide diversity are unusually high for Papuans in amplicon set 1 (i.e., exons 1–3) (0.17% per base) (table 1), despite the fact that sequence divergence with chimpanzee in amplicon set 1 (1%) is below the genome average (1.3%). Sequence diversity in Papuans is higher within this segment of the OAS1 gene than at 61 intergenic autosomal loci when estimated using the following three measures: 1) nucleotide diversity per base in Papuans ($P = 0.006$), 2) nucleotide diversity corrected by sequence divergence with chimpanzee in Papuans ($P = 0.0002$, assuming a normal distribution, supplementary fig. S1, Supplementary Material online), and 3) the ratio of nucleotide diversity in Papuans to that in each of the other populations (P varies from 0.014 to 0.001, assuming a log-normal distribution). Although ancestral alleles are rarely private to Papuans, this region has five SNPs with the ancestral allele observed only in Papuans. In addition, another seven SNPs that are private to Papuans are observed more than once (supplementary table S3, Supplementary Material online).

Network

Computationally phased sequences were used to generate a median-joining network of haplotypes. The network shows a long branch separating a fraction of Papuan haplotypes from a cluster with all remaining haplotypes (fig. 2). When the network is rooted through comparison with great apes, Papuan lineages fall on either side of the root. We refer to the basal branch containing only Papuans as the “deep lineage.”

Divergence Time

To estimate the time of the deepest branching point, we aligned the reference sequences of the human, chimpanzee,

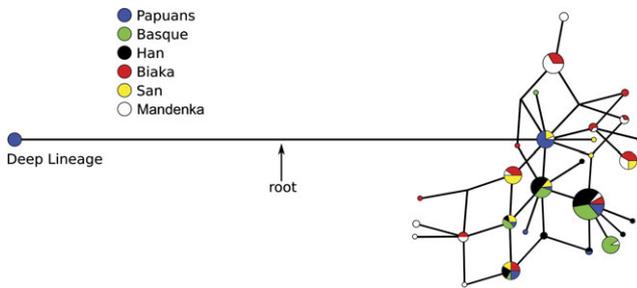


Fig. 2. Median-joining network of phased haplotypes for the first three amplicons. Branch lengths are proportional to the number of mutations along each branch, and the area of the circle is proportional to haplotype frequency.

and gorilla, with the sequence from an individual carrying two copies of the deep lineage, and compared the number of mutations segregating between the two humans with the number of mutations shared by the human sequences since the most recent common ancestor with chimpanzee. Assuming that the human and chimpanzee sequences diverged 6 Ma, the estimated time of the branching event is 3.7 Ma (2–5.2 Ma, 95% confidence interval [CI]) (table 2).

Comparisons with the Denisova Sequence

We then compared the deep lineage sequence with that of the recently reported sequence of the hominin from Denisova in the sequence matching amplicon set 1. The two sequences are identical at nearly all the sites at which the Denisova sequence has $2\times$ or higher coverage, with the exception of one site (position 30504), at which the extant human carries the derived C and the Denisova specimen carries the ancestral T (supplementary material, Supplementary Material online). The estimated divergence time between the two sequences is 0.6 Ma (0.015–2.7 Ma, 95% CI) (table 2).

Extended Resequencing within and 5' of OAS1

We performed extended resequencing on a >18 -kb region that begins ~ 9.1 kb upstream of *OAS1* and encompasses the entire gene in the Melanesian panel (supplementary table S4, Supplementary Material online) to 1) estimate the length of the unusual haplotype characterizing the deep lineage; 2) infer branching times of the deep lineage, the human reference, and the Denisova sequence; and 3) analyze intraallelic variation in the deep lineage. The deep lineage and Denisova sequence share SNPs over a >12 -kb segment. For example, the deep lineage and Denisova share the ancestral T and C at positions 29013 and 41363, respectively, whereas Eurasians and Africans share the derived C and T at these positions (supplementary

table S3, Supplementary Material online). Haplotype homozygosity in the deep lineage was explored 5' from the gene, extending up to position 20638 (supplementary table S4, Supplementary Material online).

Three contiguous sections of ~ 6 kb (22001–28000, 28001–34000, and 34001–40000) were analyzed separately and in various combinations (fig. 1). Analysis of the second of these three sections (i.e., 28001–34000) yields an estimated time for the branching between the deep lineage and the human reference sequence of 3.3 Ma (2.0–4.5 Ma, 95% CI) (table 2). This compares with the corresponding divergence time estimates of 0.6 Ma (0.2–1.2 Ma, 95% CI) and 1.1 Ma (0.4–2.2 Ma, 95% CI) for the first (22001–28000) and third (34001–40000) sections, respectively. Fisher's exact test shows that the divergence time of the second section is significantly greater than those of the first section ($P < 10^{-4}$) and of the third section ($P < 0.01$).

Extent of LD Associated with Deep Lineage

To characterize the extent of LD associated with the deep lineage, we examined publicly available Illumina 650 array data and identified a distinctive ~ 90 kb haplotype in an HGDP (Human Genome Diversity Project) individual who is known to be homozygous for the deep lineage (supplementary material, Supplementary Material online). Notably, for regions throughout this 90-kb block that have good sequence coverage in Denisova, there is complete agreement between deep lineage and Denisova SNP calls. To estimate the time that this extent of haplotype agreement between the Denisova sequence and the deep lineage is expected to persist in a single population, we performed the following analysis. We used PHASE v2.1 on HapMap phase II Yoruban data, assuming an effective population size of 10,000, and the method described in supplementary material (Supplementary Material online), to estimate the genetic map length for this block and obtained a value of 0.22 cM. If we assume that all recombination events are detectable and that there are 25 years per generation, the time of persistence of a block of this genetic length is exponentially distributed with parameter 0.0022 per generation. Thus, a block of 0.22 cM would have a 50%, 95%, and 99% probability of decaying in 8, 34, and 52 ky, respectively. Given the markers and their allelic frequency, not all recombination events are detectable. For example, if only half of the recombination events were detectable, these times would be doubled.

Geographic Distribution of the Deep Lineage

Diagnostic polymorphisms for the deep lineage were used to assess its geographic distribution (fig. 3; supplementary table S5a, Supplementary Material online). This lineage

Table 2. Divergence Time Estimates (Ma).

Lineages Compared	Amplicon Set 1 ^a	Time/(95% CI)			
		First Section	Second Section	Third Section	Second and Third sections.
Deep–Reference	3.7 (2.0–5.2)	0.6 (0.2–1.2)	3.3 (2.0–4.5)	1.1 (0.4–2.2)	
Deep–Denisova	0.6 (0.015–2.7)				0.7 (0.15–1.8)

^a Amplicon set 1 is a subset of the second section (see text).

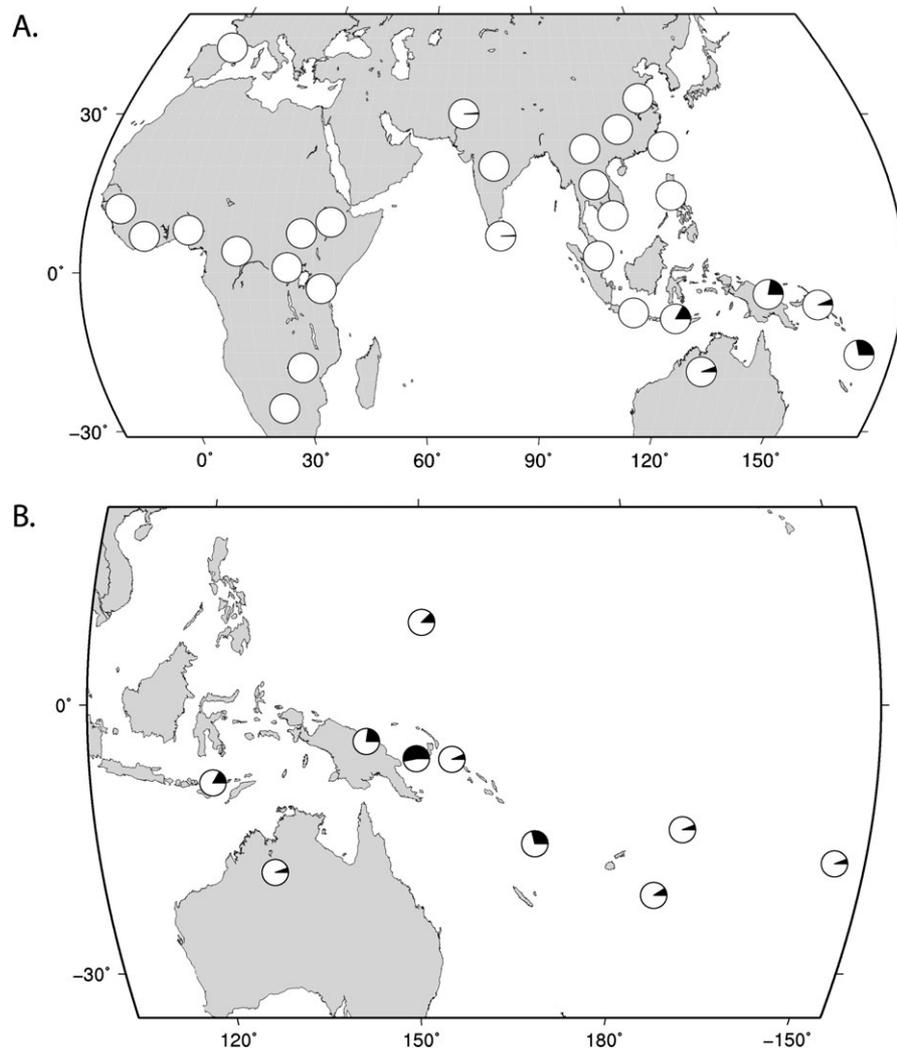


Fig. 3. Geographic distribution of the deep lineage in (A) Old World populations and (B) South East Asians and Oceanians.

is mostly restricted to and widely distributed within Melanesia. Exceptions include the finding of the diagnostic deep lineage SNPs in two Pakistanis and a single Sri Lankan (supplementary material, Supplementary Material online) in our sample of 323 South Asians.

Functional Variation

We identified 12 nonsynonymous, one frameshift, and one splice site polymorphisms in our data set (supplementary table S3, Supplementary Material online). Five of these nonsynonymous SNPs are in exonic regions that are spliced out of some versions of the gene product. By analyzing the protein sequences (supplementary material, Supplementary Material online), we identified four amino acid replacements with a likely functional effect, and three of them (R104G, P129R, and E183D) are found only in the deep lineage and Denisova sequence. The remaining replacement change (T69N) was observed on a single Papuan chromosome.

Discussion

Analyses of the recently published Neandertal and Denisova sequences from Eurasia suggest that archaic

introgression may have played an important role in shaping genetic variation in non-Africans (Green et al. 2010; Reich et al. 2010). Although genome-wide comparisons detect more sequence agreement between non-African and Neandertal genomes and between Melanesian and Denisova genomes, the specific loci exhibiting these signals have not yet been identified. Furthermore, current analyses do not elucidate the relative roles of recent introgression versus long-term population structure in Africa in explaining these patterns. Another unanswered question concerns how gene flow occurred given the large geographic distance between the Denisova archeological site in the Altai Mountains and the likely route of migration of Melanesian ancestors. It is unclear whether individuals carrying Denisova-like sequences in the Altai region contributed directly or shared some genetic variation with the population that contributed to the ancestry of Melanesians. We address these issues in our study of a region encompassing the *OAS1* gene, a locus with a divergent Melanesian haplotype that is very similar to the Denisova sequence. Importantly, none of the three individuals

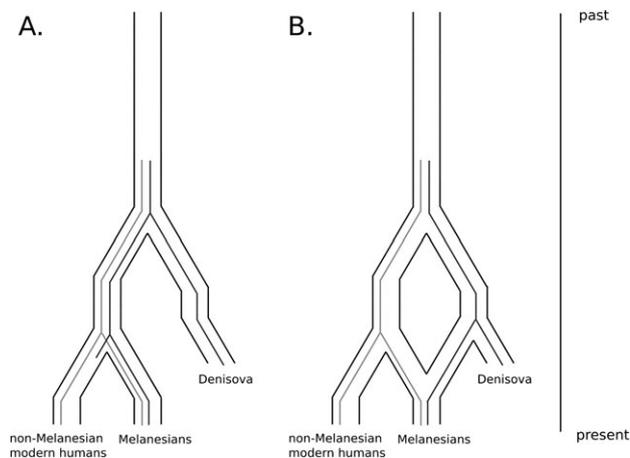


Fig. 4. Demographic models to explain nucleotide polymorphism observed at *OAS1*. (A) The deep lineage and the Denisova sequence are similar due to ancestral polymorphism maintained in Denisova and modern humans but subsequently lost in all non-Melanesians. (B) The deep lineage is observed in Melanesians due to admixture with Denisova-like humans.

(1 Sri Lankan and 2 Pakistanis) found to carry this haplotype in our survey of 1,212 non-Melanesians was an African ($n = 653$).

Evidence for Archaic Introgression at *OAS1*

Three unusual features of genetic diversity characterize the 5' region of *OAS1* (fig. 1): 1) sequence variation in Papuans is higher than that of 61 intergenic control loci, despite the fact that sequence divergence with chimpanzee in this region is below the genome average; 2) polymorphic sites with the ancestral state are observed only in Papuans; and 3) one of the two basal branches on a median-joining network (the “deep” lineage) is found only in Papuans. These patterns contrast with those of the overwhelming majority of loci in the genome, which show higher African diversity and basal lineages that are exclusive to sub-Saharan African populations (Takahata et al. 2001; Excoffier 2002; Garrigan and Hammer 2006; Henn et al. 2011). Moreover, the sequence obtained for the deep *OAS1* lineage is very similar to that of the hominin specimen from Denisova (fig. 1; supplementary table S6, Supplementary Material online).

The presence of a modern Melanesian haplotype that is similar to the Denisova sequence may be the result of ancestral shared polymorphism that persisted in the ancestors of Denisova and AMH and that was subsequently lost in all non-Melanesian AMH populations (fig. 4A). Alternatively, the similarity between the deep lineage and the Denisova sequence may be a consequence of recent admixture between a Denisova-like population and the ancestors of extant Melanesians (fig. 4B). We may be able to distinguish these alternatives because variants that recently introgressed are expected to be associated with longer blocks of LD (Wall 2000; Nordborg 2001; Garrigan et al. 2005). We found a block of ~ 90 kb in which there is complete agreement between deep lineage and Denisova SNP calls and then evaluated whether the observed extent of LD

associated with the deep lineage was likely under a model of ancestral polymorphism. In other words, we considered the probability that a block with the observed genetic length persisted in a single population since the estimated Denisova-AMH split time of 300 kya (Reich et al. 2010). We found that a block of length 0.22 cM would have a 99% probability of decaying in 52 ky, or more conservatively, within ~ 100 ky (see Results). This suggests that the deep lineage was isolated from most of the other *OAS1* lineages (e.g., those characterized by different Illumina SNP-based haplotypes, which are extremely abundant in Africa) as expected under a model of isolation and admixture (Nordborg 2001) (fig. 4B). A genome-wide comparison shows a greater average similarity between the sequences of Neandertal and Denisova than with those of modern humans (Reich et al. 2010). We made a preliminary comparison between the Neandertal sequence, the deep lineage, and extant human sequences and found that the former is very similar to the human reference sequence (data not shown). An in-depth analysis of the sequence of the *OAS1* region in Neandertal will be presented elsewhere. To ensure that our conclusions were not biased by comparisons with the reference sequence, we repeated our analysis replacing the reference sequence with other non-Melanesian sequences and obtained very similar results (data not shown).

On the Possible Source(s) of Introgression

The sequence of the deep lineage has several properties that fit expectations of a haplotype that has introgressed from a distinct population that was isolated in Asia, including its differentiation from other haplotypes, extent of linkage disequilibrium, and close relationship with an archaic Asian sequence. Broadly distributed throughout Melanesia, the deep lineage exhibits very low intraallelic diversity (supplementary table S4, Supplementary Material online), with an estimated TMRCA of ~ 25 kya (supplementary material and table S7a and b, Supplementary Material online), suggesting that it may have diversified and spread with the expansion of modern humans in the region. To further test for the presence of the deep lineage outside of Melanesia, we used the 90-kb haplotype described above as a query in a database search, we identified five similar haplotypes from Oceania and two from Pakistan (supplementary material, Supplementary Material online) and confirmed that these individuals also carry the deep lineage. These two Pakistani individuals (not genotyped in our survey) share a ~ 170 kb block with a Melanesian individual. The presence of this single long haplotype, which is otherwise restricted to Melanesia, in combination with the absence of non-deep lineages that are common in Pakistan and higher deep lineage diversity in Melanesia, suggests that the deep lineage migrated to South Asia from Melanesia. Our deep lineage results are consistent with those of a recent study of the distribution of SNPs inferred to be shared with the Denisova genome, which showed that the admixture signal is nearly limited to aboriginal groups in eastern Indonesia and Oceania (Reich et al. 2011). In sum, these patterns are unexpected for sequences that evolved in

AMH in Africa and support the hypothesis that the deep lineage introgressed into the ancestors of Melanesians from an archaic human population outside of Africa.

Considering the large geographic distance between the Altai mountains and the likely southern out-of-Africa migration route of the ancestors of Melanesians, which is not thought to have occurred earlier than ~70 kya (Macaulay et al. 2005), and the much older divergence time of the human deep lineage and the Denisova sequence (~700 kya, table 2), we suggest that the archaic ancestor contributing the deep lineage to Melanesians and the specimen from Denisova were members of genetically differentiated populations. Interestingly, the coalescent time of the human reference and deep lineage of 3.3 Ma is much older than the population divergence time of ~0.3 Ma for humans and Denisova suggested by Reich et al. (2010). Although TMRCA are expected to predate population divergence events, the discrepancy observed here is substantial. Preliminary analysis that suggests simple models that invoke balancing selection do not sufficiently explain the pattern of such a deep TMRCA combined with the concurrent maintenance of LD extending ≥ 6 kb (supplementary material, Supplementary Material online), although balancing selection and local adaptation may have influenced the current distribution of the deep lineage. Furthermore, an exclusively selectionist explanation for the observed variation at this locus would require that after evolving and being maintained by selection for ~3 myr in Africa, the deep lineage is completely lost in Africans and Neandertals in <300 ky. These observations present the intriguing possibility that this deeply diverged region of *OAS1* may have introgressed into the common ancestor of Denisova and Melanesians via admixture with an unsampled hominin group, such as *Homo erectus*. In fact, the introgression of a more archaic form into the ancestors of Denisova was also considered by Reich et al. (2010) to explain some archaic morphological features of the Denisova molar. However, without DNA from this “ghost” population or evidence from other loci showing a similar pattern of extreme haplotype differentiation, such a scenario is speculative. Fortunately, complete genomes from several Melanesian and other populations will soon be available, which will provide an opportunity to identify additional introgressed haplotypes, if they indeed exist. Detailed analysis of such candidate loci holds the key to untangling the complex relationships between modern and archaic hominins suggested by this study of *OAS1*.

Supplementary Material

Supplementary material, tables S1–S9, and figures S1–S5 are available at *Molecular Biology and Evolution* online (<http://www.mbe.oxfordjournals.org/>).

Acknowledgments

We thank Jennifer Reid for technical assistance, Tatiana Karafet and August Woerner for helpful discussions, Latifa Jackson for support, and Krishna Veeramah for comments on the manuscript. This research was funded by a National

Science Foundation HOMINID grant (BCS-0423670) to M.F.H.

References

- Cox MP, Mendez FL, Karafet TM, Pilkington MM, Kingan SB, Destro-Bisol G, Strassmann BI, Hammer MF. 2008. Testing for archaic hominin admixture on the X chromosome: model likelihoods for the modern human RRM2P4 region from summaries of genealogical topology under the structured coalescent. *Genetics* 178:427–437.
- Durand EY, Patterson N, Reich D, Slatkin M. 2011. Testing for ancient admixture between closely related populations. *Mol Biol Evol*. 28:2239–2252.
- Excoffier L. 2002. Human demographic history: refining the recent African origin model. *Curr Opin Genet Dev*. 12:675–682.
- Fedetz M, Matesanz F, Caro-Maldonado A, Fernandez O, Tamayo JA, Guerrero M, Delgado C, Lopez-Guerrero JA, Alcina A. 2006. *OAS1* gene haplotype confers susceptibility to multiple sclerosis. *Tissue Antigens* 68:446–449.
- Garrigan D, Hammer MF. 2006. Reconstructing human origins in the genomic era. *Nat Rev Genet*. 7:669–680.
- Garrigan D, Mobasher Z, Kingan SB, Wilder JA, Hammer MF. 2005. Deep haplotype divergence and long-range linkage disequilibrium at Xp21.1 provide evidence that humans descend from a structured ancestral population. *Genetics* 170:1849–1856.
- Green RE, Krause J, Briggs AW, et al. (56 co-authors). 2010. A draft sequence of the Neandertal genome. *Science* 328:710–722.
- Hamano E, Hijikata M, Itoyama S, et al. (15 co-authors). 2005. Polymorphisms of interferon-inducible genes *OAS-1* and *MxA* associated with SARS in the Vietnamese population. *Biochem Biophys Res Commun*. 329:1234–1239.
- Hammer MF, Karafet TM, Redd AJ, Jarjanazi H, Santachiara-Benerecetti S, Soodyall H, Zegura SL. 2001. Hierarchical patterns of global human Y-chromosome diversity. *Mol Biol Evol*. 18:1189–1203.
- Hammer MF, Woerner AE, Mendez FL, Watkins JC, Cox MP, Wall JD. 2010. The ratio of human X chromosome to autosome diversity is positively correlated with genetic distance from genes. *Nat Genet*. 42:830–831.
- Henn BM, Gignoux CR, Jobin M, et al. (19 co-authors). 2011. Hunter-gatherer genomic diversity suggests a southern African origin for modern humans. *Proc Natl Acad Sci U S A*. 108:5154–5162.
- Karafet T, Xu L, Du R, Wang W, Feng S, Wells RS, Redd AJ, Zegura SL, Hammer MF. 2001. Paternal population history of East Asia: sources, patterns, and microevolutionary processes. *Am J Hum Genet*. 69:615–628.
- Karafet TM, Zegura SL, Posukh O, et al. (14 co-authors). 1999. Ancestral Asian source(s) of new world Y-chromosome founder haplotypes. *Am J Hum Genet*. 64:817–831.
- Knapp S, Yee LJ, Frodsham AJ, et al. (12 co-authors). 2003. Polymorphisms in interferon-induced genes and the outcome of hepatitis C virus infection: roles of *MxA*, *OAS-1* and *PKR*. *Genes Immun*. 4:411–419.
- Li N, Stephens M. 2003. Modeling linkage disequilibrium and identifying recombination hotspots using single-nucleotide polymorphism data. *Genetics* 165:2213–2233.
- Lim JK, Lisco A, McDermott DH, et al. (14 co-authors). 2009. Genetic variation in *OAS1* is a risk factor for initial infection with West Nile virus in man. *PLoS Pathog*. 5:e1000321.
- Macaulay V, Hill C, Achilli A, et al. (21 co-authors). 2005. Single, rapid coastal settlement of Asia revealed by analysis of complete mitochondrial genomes. *Science* 308:1034–1036.
- Mendez FL, Karafet TM, Krahn T, Ostrer H, Soodyall H, Hammer MF. 2011. Increased resolution of Y chromosome haplogroup T defines relationships among populations of the Near East, Europe, and Africa. *Hum Biol*. 83:39–53.

- Nordborg M. 2001. On detecting ancient admixture. In: Donnelly P, Foley RA, editors. *Genes, fossils and behaviour: an integrated approach to human evolution*. Amsterdam: IOS Press.
- Reich D, Green RE, Kircher M, et al. (28 co-authors). 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. (15 co-authors). 2011. Denisova admixture and the first modern human dispersals into southeast Asia and Oceania. *Am J Hum Genet.* 89:1–13.
- Stephens M, Donnelly P. 2003. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *Am J Hum Genet.* 73:1162–1169.
- Tajima F. 1983. Evolutionary relationships of DNA sequences in finite populations. *Genetics* 105:437–460.
- Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585–595.
- Takahata N, Lee S, Satta Y. 2001. Testing multiregionality of modern human origins. *Mol Biol Evol.* 18:172–183.
- Wall JD. 2000. Detecting ancient admixture in humans using sequence polymorphism data. *Genetics* 154:1271–1279.
- Wall JD, Cox MP, Mendez FL, Woerner A, Severson T, Hammer MF. 2008. A novel DNA sequence database for analyzing human demographic history. *Genome Res.* 18:1354–1361.
- Watterson GA. 1975. On the number of segregating sites in genetical models without recombination. *Theor Popul Biol.* 7:256–276.
- Wessel P, Smith WHF. 1998. New, improved version of Generic Mapping Tools released. *EOS Trans AGU.* 79:579.