An X-Linked Haplotype of Neandertal Origin Is Present Among All Non-African Populations

Vania Yotova,[1] Jean-Francois Lefebvre,[1] Claudia Moreau,[1] Elias Gbeha,[1] Kristine Hovhannesyan,[1,2] Stephane Bourgeois,†[,1] Sandra Bédarida,‡[,1] Luisa Azevedo,[3] Antonio Amorim,[3,4] Tamara Sarkisian,[2] Patrice Hodonou Avogbe,[5] Nicodeme Chabi,[5] Mamoudou Hama Dicko,[6] Emile Sabiba Kou' Santa Amouzou,[7] Ambaliou Sanni,[5] June Roberts-Thomson,§ Barry Boettcher,[8] Rodney J. Scott,[8] and Damian Labuda*[,1,9]

[1]Research Center, CHU Sainte-Justine, Université de Montréal, Montréal, Québec, Canada

[2]Center of Medical Genetics and Primary Health Care, Yerevan, Armenia

[3]IPATIMUP-Institute of Molecular Pathology and Immunology of the University of Porto, Porto, Portugal

[4]Faculty of Sciences of the University of Porto, Porto, Portugal

[5]Laboratoire de Biochimie et de Biologie Moléculaire, Université d'Abomey-Calavi, Cotonou, Bénin

[6]Université de Ouagadougou, Ouagadougou, Burkina Faso

[7]Département de Biochimie/Nutrition, Université de Lome, Lome, Togo

[8]Medical Genetics, School of Biomedical Sciences and Pharmacy, Faculty of Health, University of Newcastle, Australia

[9]Département de Pédiatrie, Université de Montréal, Montreal, Canada

†Present address: Wellcome Trust Sanger Institute, Cambridge, United Kingdom.

‡Present address: UMR 6578 CNRS, Équipe Émergence et Co-évolution Virale, Université de la Méditerranée, Marseille, France.

§Deceased.

***Corresponding author:** E-mail: damian.labuda@umontreal.ca.

Associate editor: Daniel Falush

Abstract

Recent work on the Neandertal genome has raised the possibility of admixture between Neandertals and the expanding population of *Homo sapiens* who left Africa between 80 and 50 Kya (thousand years ago) to colonize the rest of the world. Here, we provide evidence of a notable presence (9% overall) of a Neandertal-derived X chromosome segment among all contemporary human populations outside Africa. Our analysis of 6,092 X-chromosomes from all inhabited continents supports earlier contentions that a mosaic of lineages of different time depths and different geographic provenance could have contributed to the genetic constitution of modern humans. It indicates a very early admixture between expanding African migrants and Neandertals prior to or very early on the route of the out-of-Africa expansion that led to the successful colonization of the planet.

Key words: human evolution, archaic lineages, Neandertal admixture, out-of-Africa migration, genetic diversity, X-linked lineage.

Introduction

The issue of admixture between Neandertals and modern humans has been contentious. The arguments for and against such admixture were raised in several paleontological studies (Tattersall and Schwartz 1999; Klein 2003; Zilhao 2006; Trinkaus 2007). Neandertal mitochondrial DNA sequencing provided no evidence of Neandertal matrilineal contribution to contemporary humans (reviewed by, Hodgson and Disotell 2008). In contrast, a draft sequence of the Neandertal nuclear DNA revealed the presence of a number of bona fide Neandertal segments in non-African genomes (Green et al. 2010). Yet, although rather unlikely given all the precautions taken to eliminate such a possibility (Green et al. 2010), the admixture results could be due to contamination of ancient DNA with modern human genomic fragments (Wall and Kim 2007). In this paper, we provide further evidence of Neandertal admixture. Our laboratory did not participate in the determination of the Neandertal sequence. The admixed X-linked haplotype of Neandertal origin we report was suspected to descend from a non-African lineage well before the advent of the Neandertal genome (Zietkiewicz et al. 2003).

To the admixed haplotypes from the bulk of genomic diversity, we assume that modern humans left Africa somewhere between 80–50 thousand years ago (Kya) and that at that time Neandertals occupied western Eurasia (Stringer 2002; Klein 2003; Mellars 2006; Oppenheimer 2009; Petraglia et al. 2010). Because these populations diverged 400–800 Kya, a number of derived alleles present in the Neandertal DNA can be expected to segregate in all human populations including sub-Saharan Africans. In contrast, segments in human DNA that were subsequently admixed outside Africa

Letter

Table 1. Twelve Major dys44 Haplotypes and the Neandertal Sequence.

Haplotypes	Polymorphic Sites																																	Sub-Saharan Africa (%) n=1420	Non-Africans (%) n=4672
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35		
Ancestral	T	A	A	A	G	T	8N	A	A	T	G	T	A	A	T	A	C	A	T	TGA	T	ACA	A	A	G	T	A	A	A	T	A	T	C	T	G	
Neandertal	C	.	.	.	.	.	na	na	na	.	na	na	na	na	na	.	.	na	na	.	na	.	.	na	.	.	.	.	.	.	G	.	na	na	na		
B001	C	G	G	.	.	.	.	.	.	C	.	.	C	T	.	G	.	C	.	.	C	.	.	G	.	.	.	.	G	.	.	.	.	G	.	12.5	35.8
B002	.	.	.	.	.	.	.	.	G	.	.	.	.	T	.	.	.	.	.	.	.	.	.	G	.	.	.	.	G	.	.	G	.	.	T	18.8	13.2
B003	C	G	G	.	.	.	.	.	.	C	T	.	C	T	.	G	.	C	.	.	C	.	.	G	.	.	.	.	G	.	.	.	.	G	.	4.4	15.6
B004	C	G	G	.	.	.	.	.	G	.	.	.	.	T	.	.	.	.	.	.	.	.	.	G	.	.	.	.	G	.	.	G	.	.	T	0.1	4.7
B005	C	G	G	.	.	.	.	.	.	.	T	.	C	T	.	G	.	.	C	.	C	.	.	G	.	C	.	.	G	.	.	.	.	.	.	9.7	5.3
B006	C	.	.	.	.	.	dupl	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	G	.	.	.	.	0.4	8.7
B007	.	.	.	.	.	.	.	T	.	.	.	.	C	T	.	.	.	.	.	.	C	.	.	.	.	C	.	.	G	C	.	.	.	.	.	7.1	0.0
B008	C	G	G	.	.	.	.	.	.	C	.	.	C	T	C	G	.	.	.	.	C	.	.	G	.	C	.	.	G	.	.	.	.	.	.	3.5	3.9
B009	C	G	G	.	.	.	.	.	.	C	T	.	C	T	.	G	.	C	.	.	C	.	.	G	.	C	.	.	G	.	.	.	.	G	.	2.7	0.1
B010	C	G	G	.	.	.	.	.	.	.	.	.	C	T	.	.	.	.	.	del	.	.	.	G	.	C	.	.	G	.	.	.	.	G	.	4.6	0.0
B011	C	G	G	.	.	.	.	.	.	.	T	.	C	T	.	G	.	C	C	.	C	.	.	G	.	C	.	.	G	.	.	.	.	G	.	0.4	1.8
B012	C	G	.	.	.	.	.	T	.	.	.	.	C	T	.	.	.	.	.	.	C	.	.	.	C	.	G	.	G	.	.	.	.	.	.	2.5	0.1
																																	Total %		66.8	89.2

1: rs6631517*, 2: rs6418638, 3: rs1456740*, 8: rs5972510, 9: rs6628685*, 10: rs331370*, 11: rs2854965*, 16: rs6653863, 18: rs331369*, 19: rs331368*, 20: rs72468607, 21: rs3747400*, 22: rs72468609, 23: rs73458098, 24: rs16990172*, 26: rs331367*, 28: rs58400027, 31 rs11795471*, 32 : rs67635204, 34: rs331366*, 35: rs6631518

1: 32139960, 2: 32140361, 3: 32140683, 4: 32140819, 5: 32140848, 6: 32141227, 7: 32141804, 8: 32142291, 9: 32142304, 10: 32142364, 11: 32142804, 12: 32145805, 13: 32143100, 14: 32143103, 15: 32143141, 16: 32143469, 17: 32143713, 18: 32144154, 19: 32144398, 20: 32144834, 21: 32145216, 22: 32145280, 23: 32145700, 24: 32146159, 24: 32146159, 25: 32146199, 26: 32146414, 27: 32146531, 28: 32146861, 29: 32147134, 30: 32147392, 31: 32147542, 32: 32147779, 33: 32147878, 34: 32147882, 35: 32147883

Note.—Of 35 polymorphic sites that define dys44 haplotypes (Zietkiewicz et al. 1997, 2003), 26 of their derived alleles are found in the 12 major haplotypes listed here. All sites identifiable in the Neandertal (Green et al. 2010) are defined or otherwise marked as non available—na. Note that for the sake of compatibility with HapMap, and contrary to our previous reports (e.g., Zietkiewicz et al. 2003), these haplotypes are here defined on the DNA strand complementary to the dystrophin transcript (sequence numbering is according to March 2006 Assembly).

* Denotes 12 sites present in the HapMap3 database.

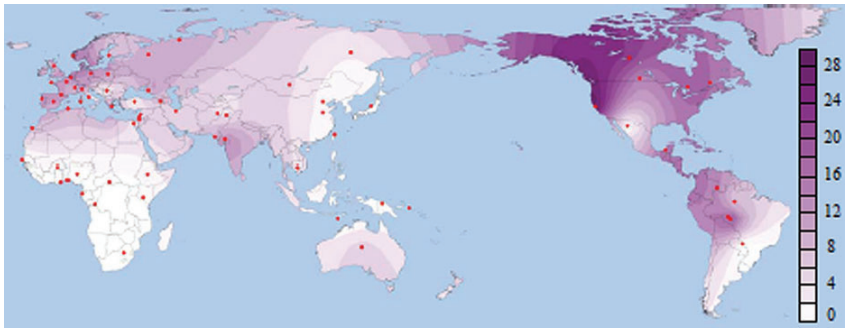

FIG. 1. Worldwide distribution of B006 haplotype based on a sample of 6,092 X-chromosomes. Samples are listed in supplementary table S1, Supplementary Material online. Certain subpopulations were merged, when justified by their geographic proximity.

are expected to carry additional derived alleles shared with Neandertals but absent in sub-Saharan Africans. The admixed haplotypes are also expected to be younger, that is, less diversified by recombination than the average African haplotypes at the same loci. Ideally, the allelic structure of these haplotypes would differ from the bulk of the common haplotypes reflecting their origin along a separately evolving lineage.

An X-linked haplotype that fulfills these characteristics occurs in an 8-Kb intronic segment spanning exon 44 of the dystrophin gene, referred to as dys44 (Zietkiewicz et al. 1997). We analyzed dys44 polymorphisms in a sample of 6,092 X-chromosomes representing populations from all habitable continents (supplementary table S1, Supplementary Material online), including previously published data (Labuda, Labuda, et al. 2000; Zietkiewicz et al. 2003; Xiao et al. 2004; Lovell et al. 2005; Bourgeois et al. 2009). We focus on the 12 major dys44 haplotypes that explain 89% of genetic diversity in non-Africans and 67% in sub-Saharan Africans (table 1). Of these, haplotype B006 is structurally distinct; with only four derived alleles, it is the closest to the ancestral one. Common outside Africa and virtually absent in sub-Saharan Africa (fig. 1), B006 was earlier proposed to represent an unknown non-African contribution to the human gene pool (Zietkiewicz et al. 2003). Of 1,420 sub-Saharan chromosomes, only one copy of B006 was observed in Ethiopia, and five in Burkina Faso, one among the Rimaibe and four among the Fulani and Tuareg, nomad-pastoralists known for having contacts with northern populations (supplementary table S1, Supplementary Material online). B006 only occurrence at the northern and northeastern outskirts of sub-Saharan Africa is thus likely to be a result of gene flow from a non-African source.

In the available Neandertal sequence (Green et al. 2010), there is information on 20 of 35 dys44 polymorphic sites. These represent 18 ancestral and 2 derived alleles, fully matching the corresponding sites of B006 (table 1). One of the derived alleles, C of rs6631517, is also shared with other dys44 haplotypes, whereas the second one, G of rs11795471, is unique to B006 (the information on two remaining B006-polymorphisms is not available). Figure 2 illustrates plausible historical pathways leading to the three observed categories of the dys44 haplotypes. Haplotypes, such as B007, B010, and B012 in table 1 are specific to sub-Saharan Africa. They carry common (sites of type 1 and 3) and African-specific polymorphisms (sites 2, 4, and 6). The remaining haplotypes, except B006, are cosmopolitan and are found both inside and outside Africa. The third category, absent from Africa, is represented by B006, which carries two types of derived alleles that are shared with Neandertal DNA. The mutation at site 1 (as rs6631517 above) was presumably fixed in Neandertals and segregated along the ancestral allele in the human lineage. Derived alleles acquired through recent admixture (site 5) are expected to be only found in the progeny of the admixed individuals as in the case of rs11795471. Moreover, the Neandertal origin of B006 is consistent with the allelic status of the remaining 19 of the 20 Neandertal sites of known identity (table 1) and confirmed by analysis of the B006 flanking segments in HapMap3 (Altshuler et al. 2010).

Twelve of the 35 dys44 polymorphisms available in the HapMap3 database are sufficient to identify the 12 major dys44 haplotypes presented in table 1. In the considered HapMap3 populations (supplementary table S1, Supplementary Material online), these haplotypes represent 369 of 523 (70.6%) sub-Saharan chromosomes and 815 of 891 (91.5%) non-African chromosomes that include 77 copies of B006. We extended the dys44 8-kb region by 28 additional polymorphisms on the left and 49 on the right (108 kb in total) to include sites that by inspection appeared to maintain some degree of linkage disequilibrium with the B006 haplotype (fig. 3). In the Neandertal sequence, no information is available for 28 of these sites, 36 sites represent ancestral alleles and 13 derived alleles. Importantly, three of the derived alleles (rs17243319, rs1456729, and rs11796299 in supplementary table S2, Supplementary Material online) are absent from the African chromosomes, as in the case of the derived G of rs11795471 from within B006. Moreover, all derived alleles shared with Neandertals occur at high frequencies (0.75 and more) on a background of the extended B006 haplotype (fig. 3 and supplementary table S2, Supplementary Material online) as expected in a segment of recent Neandertal origin.

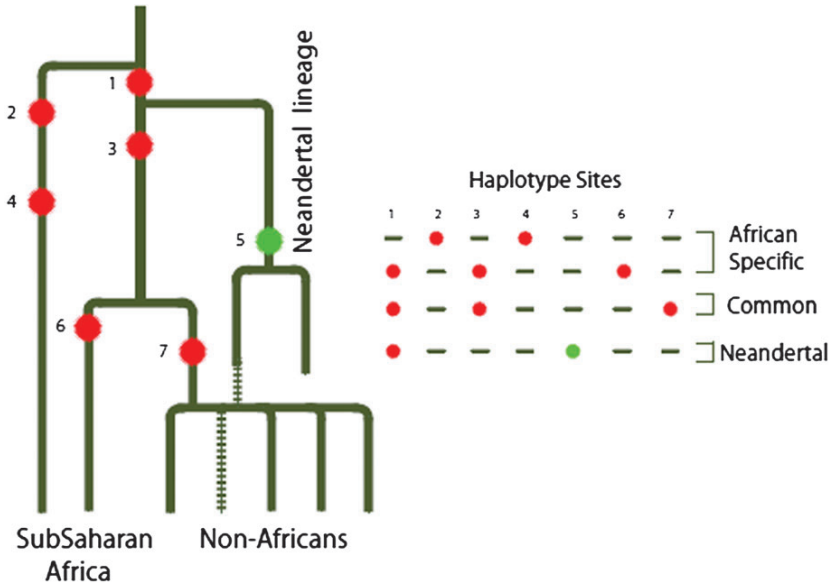

FIG. 2. Scheme of evolutionary pathways leading to three categories of haplotypes in the dys44 segment.

Outside Africa, B006 is found in all habitable continents including Australia, as determined from a remote community of isolated indigenous Australians living in Central Australia (fig. 1). The ubiquity of B006 lineage reflects a worldwide contribution of Neandertal lineages to non-African genomes. It indicates very early Neandertal admixture prior to successful range expansion of the population ancestral to virtually all contemporary non-African populations and confirms earlier contention of very early admixture based on the analysis of Neandertal segments in European, Han, and Papuan genomes (Green et al. 2010). Hodgson et al. 2010 proposed such admixture through early Levantine contacts of modern humans and Neandertals, prior to the most recent out-of-Africa expansion, and suggested that traces of such admixture should be still detectable in sub-Saharan populations of Northeastern Africa. In contrast, rather than considering admixture, an alternative explanation posits that the ancestral population of present-day non-Africans was more closely related to Neandertals than the ancestral population of present-day Africans (Green et al. 2010). Indeed, the evidence is accumulating on deep subdivisions within the ancestral population in Africa (Labuda, Zietkiewicz, Yotova 2000; Falush et al. 2003; Cohen et al. 2007; Yotova et al. 2007; Behar et al. 2008; Gunz et al. 2009). As in the case of admixture through Levantine contacts, the latter explanation also implies significant sharing of Neandertal haplotypes between all non-Africans and Northeastern Africans. On the other hand, no evidence for these scenarios is found in our data, whereas the oldest lineages tend to be found in South rather than Northeastern Africa (Behar et al. 2008; Campbell and Tishkoff 2010). Paleontological findings (Shea 2008; Petraglia et al. 2010) point to the occupation of the Levant by Neandertals and early *H. sapiens* at nonoverlapping time periods making their early contacts unlikely. Interestingly, some of the HapMap3 haplotypes from the segments proposed by Green et al. 2010 and fulfilling our criteria of Neandertal admixture, also turn out in Maasai, where, however, their occurrence can be due to recent back-to-Africa migration (Sikora et al. 2011).

In contrast to other candidate haplotypes, such as these from "ASPM" and "MCPH1" that seem to have failed the

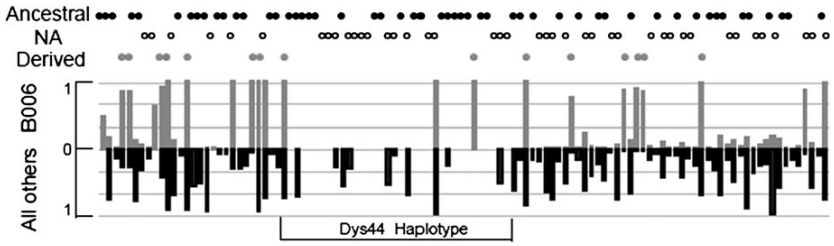

FIG. 3. Derived allele frequencies in the HapMap3-based extended dys44 haplotype. Superimposed histograms compare derived allele frequencies seen on the background of the 77 extended B006 haplotypes (upper) with frequencies of the same alleles on the remaining 1,337 haplotypes from this data set (supplementary table S2, Supplementary Material online). The status of the Neandertal alleles is marked on the top as ancestral, derived, or not available, NA.

test (Green et al. 2010; Lari et al. 2010), the evidence for Neandertal origin of B006 appears very strong. This provides additional verification of the findings by Green et al. (2010) that Neandertals contributed to the genetic makeup of modern human populations outside Africa. Our data indicate that Neandertal admixture occurred very early or prior to their worldwide expansion. Considering such early encounter of *H. sapiens* with Neandertals, a question may be raised: was this encounter coincidental and without important evolutionary consequences or (either through genetic or cultural exchanges, or both, Premo and Hublin 2009) did it facilitate adaptations to novel environmental conditions that actually contributed to the successful expansion of human migrants from Africa to other continents?

Materials and Methods

DNA samples, listed in supplementary table S1, Supplementary Material online, were genotyped for dys44 polymorphisms as previously reported (Zietkiewicz et al. 1997). All samples were non-nominative, originating either from existing collections or peripheral blood samples donated by consenting informed adults, following the protocol approved by the institutional review boards of the Sainte-Justine University Hospital Center in Montreal. The study of samples from a community of isolated indigenous Australians was approved by the University of Newcastle Human Research Ethics Committee and Hunter New England Health Research Ethics Committee (written approval was additionally obtained from the traditional land owners of the community from which the samples were originally collected). The extended dys44 haplotypes were analyzed using a subset of HapMap3 data (Altshuler et al. 2010) described in supplementary table S1, Supplementary Material online. Contour maps were obtained from haplotype frequencies using Surfer 8.02 from Golden Software.

Supplementary Material

Supplementary tables S1 and S2 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

We are grateful to Luis Barreiro, Youssef Idaghdour, Alon Keinan, and David Reich for their comments on earlier versions of this paper and wish to thank all individuals who kindly contributed their genetic material for our studies. L.A. is supported by the Fundação para a Ciência e a Tecnologia (FCT) Ciência 2007 and by the European Social Fund. Institute of Molecular Pathology and Immunology of the University of Porto (IPATIMUP) is an Associate Laboratory of the Portuguese Ministry of Science, Technology and Higher Education and is partially supported by FCT. This study was supported by grants from the Canadian Institutes for Health Research (MOP-67150 and IGI-94494).

References

Altshuler DM, Gibbs RA, Peltonen L, et al. (66 co-authors). 2010. Integrating common and rare genetic variation in diverse human populations. *Nature* 467:52–58.

Behar DM, Villems R, Soodyall H, et al. (12 co-authors). 2008. The dawn of human matrilineal diversity. *Am J Hum Genet.* 82:1130–1140.

Bourgeois S, Yotova V, Wang S, et al. (8 co-authors). 2009. X-chromosome lineages and the settlement of the Americas. *Am J Phys Anthropol.* 140:417–428.

Campbell MC, Tishkoff SA. 2010. The evolution of human genetic and phenotypic variation in Africa. *Curr Biol.* 20:R166–R173.

Cohen AS, Stone JR, Beuning KR, et al. (9 co-authors). 2007. Ecological consequences of early Late Pleistocene megadroughts in tropical Africa. *Proc Natl Acad Sci U S A.* 104:16422–16427.

Falush D, Wirth T, Linz B, et al. (15 co-authors). 2003. Traces of human migrations in Helicobacter pylori populations. *Science* 299:1582–1585.

Green RE, Krause J, Briggs AW, et al. (53 co-authors). 2010. A draft sequence of the Neandertal genome. *Science* 328:710–722.

Gunz P, Bookstein FL, Mitteroecker P, Stadlmayr A, Seidler H, Weber GW. 2009. Early modern human diversity suggests subdivided population structure and a complex out-of-Africa scenario. *Proc Natl Acad Sci U S A.* 106:6094–6098.

Hodgson JA, Bergey CM, Disotell TR. 2010. Neandertal genome: the ins and outs of African genetic diversity. *Curr Biol.* 20:R517–R519.

Hodgson JA, Disotell TR. 2008. No evidence of a Neanderthal contribution to modern human diversity. *Genome Biol.* 9:206.

Klein RG. 2003. Paleoanthropology. Whither the Neanderthals? *Science* 299:1525–1527.

Labuda D, Zietkiewicz E, Yotova V. 2000. Archaic lineages in the history of modern humans. *Genetics* 156:799–808.

Labuda M, Labuda D, Miranda C, Poirier J, Soong BW, Barucha NE, Pandolfo M. 2000. Unique origin and specific ethnic distribution of the Friedreich ataxia GAA expansion. *Neurology* 54:2322–2324.

Lari M, Rizzi E, Milani L, et al. (13 co-authors). 2010. The microcephalin ancestral allele in a Neanderthal individual. *PLoS One.* 5:e10648.

Lovell A, Moreau C, Yotova V, Xiao F, Bourgeois S, Gehl D, Bertranpetit J, Schurr E, Labuda D. 2005. Ethiopia: between Sub-Saharan Africa and western Eurasia. *Ann Hum Genet.* 69:275–287.

Mellars P. 2006. Going east: new genetic and archaeological perspectives on the modern human colonization of Eurasia. *Science* 313:796–800.

Oppenheimer S. 2009. The great arc of dispersal of modern humans: Africa to Australia. *Quat Int.* 202:2–13.

Petraglia MD, Haslam M, Fuller DQ, Boivin N, Clarkson C. 2010. Out of Africa: new hypotheses and evidence for the dispersal of Homo sapiens along the Indian Ocean rim. *Ann Hum Biol.* 37:288–311.

Premo LS, Hublin JJ. 2009. Culture, population structure, and low genetic diversity in Pleistocene hominins. *Proc Natl Acad Sci U S A.* 106:33–37.

Shea JJ. 2008. Transitions or turnovers? Climatically-forced extinctions of *Homo sapiens* and Neanderthals in the east Mediterranean Levant. *Quat Sci Rev.* 27:2253–2270.

Sikora M, Laayouni H, Calafell F, Comas D, Bertranpetit J. 2011. A genomic analysis identifies a novel component in the genetic structure of sub-Saharan African populations. *Eur J Hum Genet.* 19:84–88.

Stringer C. 2002. Modern human origins: progress and prospects. *Philos Trans R Soc Lond B Biol Sci.* 357:563–579.

Tattersall I, Schwartz JH. 1999. Hominids and hybrids: the place of Neanderthals in human evolution. *Proc Natl Acad Sci U S A.* 96:7117–7119.

Trinkaus E. 2007. European early modern humans and the fate of the Neandertals. *Proc Natl Acad Sci U S A*. 104:7367–7372.
Wall JD, Kim SK. 2007. Inconsistencies in Neanderthal genomic DNA sequences. *PLoS Genet*. 3:1862–1866.
Xiao FX, Yotova V, Zietkiewicz E, et al. (8 co-authors). 2004. Human X-chromosomal lineages in Europe reveal Middle Eastern and Asiatic contacts. *Eur J Hum Genet*. 12:301–311.
Yotova V, Lefebvre JF, Kohany O, Jurka J, Michalski R, Modiano D, Utermann G, Williams SM, Labuda D. 2007. Tracing genetic history of modern humans using X-chromosome lineages. *Hum Genet*. 122:431–443.
Zietkiewicz E, Yotova V, Gehl D, et al. (10 co-authors). 2003. Haplotypes in the dystrophin DNA segment point to a mosaic origin of modern humans' diversity. *Am J Hum Genet*. 73:994–1015.
Zietkiewicz E, Yotova V, Jarnik M, et al. 1997. Nuclear DNA diversity in worldwide distributed human populations. *Gene* 205:161–171.
Zilhao J. 2006. Neandertals and moderns mixed, and it matters. *Evol Anthropol*. 18:183–195.