Genetics and Molecular Biology, 35, 4 (suppl), 904-911 (2012)

www.sbg.org.br

Research Article

Homo sapiens, Homo neanderthalensis and the Denisova specimen: New insights on their evolutionary histories using whole-genome comparisons

Vanessa Rodrigues Paixão-Côrtes, Lucas Henrique Viscardi, Francisco Mauro Salzano, Tábita Hünemeier and Maria Cátira Bortolini

Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil.

Abstract

After a brief review of the most recent findings in the study of human evolution, an extensive comparison of the complete genomes of our nearest relative, the chimpanzee (*Pan troglodytes*), of extant *Homo sapiens*, archaic *Homo neanderthalensis* and the Denisova specimen were made. The focus was on non-synonymous mutations, which consequently had an impact on protein levels and these changes were classified according to degree of effect. A total of 10,447 non-synonymous substitutions were found in which the derived allele is fixed or nearly fixed in humans as compared to chimpanzee. Their most frequent location was on chromosome 21. Their presence was then searched in the two archaic genomes. Mutations in 381 genes would imply radical amino acid changes, with a fraction of these related to olfaction and other important physiological processes. Eight new alleles were identified in the Neanderthal and/or Denisova genetic pools. Four others, possibly affecting cognition, occured both in the *sapiens* and two other archaic genomes. The selective sweep that gave rise to *Homo sapiens* could, therefore, have initiated before the modern/archaic human divergence.

Keywords: human evolution, comparative genomics, positive selection, Neanderthal, Denisova.

Introduction

Until recently it was believed that the first hominid genus (or hominin, primates basically characterized by erect posture, bipedal locomotion and relatively large brains; Johanson and Edgar, 1996) was *Australopithecus*, whose fossil record is relatively broad and convincing in showing the conditions described above. More recent discoveries, however, have brought up the possibility of change to this traditional view, since they describe at least three new species of hominids whose existence dates back to much more remote times (~4-7 million years ago or BP): *Sahelanthropus tchadensis* (Brunet *et al*., 2002), *Orrorin tugenensis* (Haile-Selassie, 2001) and *Ardipithecus ramidus* (Suwa *et al*., 2009; White *et al.*, 2009). Although there are controversies regarding the hominid phylogeny and its nomenclature (recent discussion in González-José *et al.*, 2008; Endicott *et al*., 2010; Schwartz and Tattersall, 2010), some paleoanthropologists have postulated that from *Ardipithecus ramidus* would have emerged the first species of the genus *Australopithecus, Australopithecus anamensis* (~4 million years BP), which in turn gave rise to *Australophitecus afarensis* (~3.5 years million BP), one of the best documented extinct hominid species (the famous skeleton of a female named Lucy, which is part of the collection that helped define the characteristics of the species (Johanson and Edgar, 1996: Leakey *et al.*, 1998). It is likely that *Australopithecus afarensis* was the ancestor of several other species currently identified as belonging to the *Paranthropus* genus (earlier identified as robust australopith lineages), as well as to others classified in the genus *Homo* (Johanson and Edgar, 1996; Kimbel and Delezene, 2009). At around 2 million years BP individuals belonging to at least three *Homo* species (*Homo habilis, Homo ergaster* and *Homo rudolfensis*) inhabited the area around Lake Turkana, although paleoanthropologists do not have the slightest idea about whether or how these apparent relatives may have interacted (Tattersall, 1997). Additionally, this temporal overlap of early *Homo* species indicates that ancestor-descendant relationships are far from straightforward (Johanson and Edgar, 1996). On the other hand, there is a consensus that until that moment the history of hominids was restricted to Africa. This changed around ~1.8 million years BP when *Homo* hominins colonized Europe, Asia and Oceania, where their probable descendants survived until recently. A noteworthy example is *Homo floresiensis,* a short-statured hominin whose remains were

Send correspondence to Maria Cátira Bortolini. Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Caixa Postal 15053, 91501-970 Porto Alegre, RS, Brazil. E-mail: maria.bortolini@ufrgs.br.

found in the island of Flores, Indonesia and who remained there until at least 17,000 years BP (Brown *et al.*, 2004). The typical morphology of *Homo neanderthalensis*, on the other hand, appeared first in Europe about 400,000 years ago, and probably evolved from some *H. erectus* branch that left Africa in that first round of migrations. Other distinctive Neanderthal forms subsequently evolved until 30,000-40,000 BP, when the species became extinct (Dodge, 2012).

The discovery of a fossil of a probable hominin always triggers further discussions about hominid evolutionary history. It was no different with a recent discovery in Russia, a distal manual phalanx of a juvenile hominin, dated to 50,000 to 30,000 years BP, which was excavated at the Denisova Cave in the Altai Mountains of southern Siberia (Derevianko *et al.*, 2003). In the same layer, body ornaments of polished stone normally associated with modern humans, as well as other lithic artifacts connected to more ancient technology traditions were found. These conflicting cultural characteristics and the scarcity of more representative fossil bones made it difficult to define the exact taxonomic category of this specimen.

The finding of archaic humans in distinct regions and remote times raises a pertinent question: when, how and where did *Homo sapiens* appear?

The most recent discoveries of a fossil attributed to early anatomically modern *Homo sapiens* were made in Ethiopia, northeast Africa (White *et al.*, 2003; Haile Selassie, *et al.*, 2004). These and other findings suggest that modern humans emerged ~155,000 years ago from an archaic phase of *Homo sapiens* and the latter from *Homo erectus*, in successive evolutionary events which occurred in Africa, although this view is far from consensual (Gibbons, 2002; Schwartz and Tattersall, 2010).

It is likely that the first migration of anatomically modern *Homo sapiens* out of Africa occurred immediately before or during an interglacial period that occurred from 135,000 to 74,000 years BP (Armitage *et al*., 2011). Around 30,000-40,000 years ago, evidence for the presence of the anatomically modern *Homo sapiens* in Europe is striking (Dodge, 2012). The contemporaneity of modern and archaic humans in Europe, Asia and Oceania implies that they could have interacted, although the fossil and archeological records are controversial concerning the consequences of these probable contacts (Schwartz and Tattersall, 2010; Dodge, 2012).

The complete sequencing of plant and animal genomes has increased our ability to discover and understand many important biological phenomena, including those related to our own evolutionary history. In the case of *Homo sapiens*, only one complete genome was known in February 2001, when its draft was simultaneously published in Science and Nature by two separate research teams. Today, the complete genome of several individuals, including a paleo-Eskimo, sub-Saharan Africans, Asians and Europeans are known (Rasmussen *et al.*, 2010, Schuster *et al*., 2010; The 1000 Genomes Project Consortium, 2010); and the drafts of the Neanderthal and Denisova genomes were published, revealing for the first time details about the complete nuclear genome of other species of the *Homo* genus (Green *et al*., 2010, Reich *et al*., 2010). These nuclear data sets have provided a much more accurate view of our own evolutionary history (Gibbons, 2002). For instance, Green *et al.* (2010) found evidence that present day non-Africans have 1% to 4% of nuclear DNA of Neanderthal origin, while Reich *et al*. (2010) showed that Denisova populations must have shared a closer common ancestor with Neanderthals than with modern humans. They also suggested that Denisovans contributed 4%-6% genes to ancestors of present-day Melanesians from Papua New Guinea and the Bougainville Islands. These results indicate at least two crossbreeding events between modern and archaic humans, raising the question of whether *H. sapiens* is or is not a species distinct from the others (Gibbons, 2002). This possibility has an important implication since the complete replacement postulated by the "Out of Africa" model could be questioned, in favor of alternative models that admit some level of assimilation between local archaic and migrant modern hominins.

Comparative analyses of these data sets allowed researchers, for the first time, to identify common or taxonomically-restricted molecular toolkits of these *Homo* lineages. For instance, a recent study performed by our research team revealed that 194 individuals from different human populations presented 16 substitutions, without variability in a 546-base pair segment that acts as an enhancer of gene expression (*HACNS1),* distinguishing us from the chimpanzees. Equal lack of variability in this region was also found in the Neanderthal sequence, favoring the interpretation of past positive and present conservative selection, as would be expected in a region which influences traits as important as opposable thumbs, manual dexterity and bipedal walking (Hünemeier *et al.*, 2010). A particularly important result of Hünemeiers paper was the suggestion that the *HACNS1* mutant alleles had an origin that predated the emergence of *Homo sapiens* and that these variants could have had important roles in the evolution of *Homo* specific traits.

In the present study we compared the *Homo sapiens* genome with those of Neanderthal and Denisova to explore some issues on the nature of the differences and similarities of these modern and archaic hominin lineages, an essential approach to unravel the genetic components that make us human.

Material and Methods

The *Homo neanderthalensis* draft genome (Green *et al.*, 2010), reference human genome (NCBI Build 36/hg18), Denisova specimen genome (Reich *et al.*, 2010), and reference chimpanzee (*Pan troglodytes*) genome

(CGSC 2.1/panTro2) were obtained and analyzed using tools present in the USCS Genome Browser (Fujita *et al*., 2011). Three different approaches were used:

First, a total of 77 missense and one nonsense mutations were selected, as identified by Green *et al.* (2010), where the derived allele (taking in consideration that a mutation event produces a new or derived allele that is different from the "original" or ancestral allele) is fixed in humans, whereas the ancestral allele is present in chimpanzee and Neanderthals. We then verified whether they were present in the Denisova genome. The first comparison between this set of data was performed using the Blat tool of the USCS genome browser with query type protein and default conditions. The UCSC track configurations used were: (a) Comparative Genomics-Vertebrate Multiz Alignment & Conservation (44 Species); (b) Neanderthal Assembly and Analysis; (c) Denisova Assembly and Analysis; and (d) Variation and Repeats - Simple Nucleotide Polymorphisms (dbSNP build 130). We used the Convert tool to verify the 78 positions in the chimp genome. Finally, we also confirmed the condition of the allele (ancestral or derived) in Denisova with an alignment with the human, chimp and all archaic contigs, comprising a region of 200 bp surrounding the allele of interest. The alignments were performed using the Codon Code Aligner (Trial Version), with a local alignment using the human reference sequence.

Second, using the USCS Table Browser for the human genome assembly NCBI Build 36/hg18, with the group "Neanderthal Assembly and Analysis", track "H-C Coding Diffs" and the table "ntHumChimpCodingDiff", we retrieved all non-synonymous substitutions in which the derived allele is fixed or nearly fixed (presents higher frequency) in humans when compared to the chimpanzee ancestral allele, and also the ones available for the draft genome of Neanderthals and/or Denisova. All non-synonymous substitutions which are also known as human polymorphisms (dbSNP build 130) were visually checked on the Human UCSC Genome Browser. The Grantham score was then used to categorize all amino acid changes into classes of chemical similarity. The sites were classified as conservative (Grantham score 0-50), moderately conservative (51-100), moderately radical (101-150) and radical (> 151; Grantham, 1974; Li *et al.*, 1985). All amino acid changes classified as radical were visually checked on the Human UCSC Genome Browser. For those classified as radical, an additional analysis using the GeneDecks V3 software was performed to check whether they belonged to a specific class of genes and/or a functional cluster.

Third, in a previous selection of genes for which earlier studies (Carroll, 2003; Varki, 2004, Hill and Walsh, 2005; Varki and Altheide, 2005; Varki *et al*., 2008; Fu *et al*., 2011) indicated positive selection in the human lineages, the degree and nature of the variation present were compared with those in *Homo neanderthalensis* and the Denisova hominin. The bioinformatic tools used in the other approaches were also used here. We also evaluated, for certain substitutions of specific interest, the impact of the amino acid substitutions on the structure and function of the protein using PolyPhen-2 (Polymorphism Phenotyping v2; Adzhubei *et al*., 2010).

Results and Discussion

Human-specific non-synonymous substitutions

Green *et al.* (2010) identified mutations related to 78 amino acid changes that are specific to present-day humans, their ancestral allelic states being present in both Neanderthal and chimpanzees. These changes occurred in 73 different genes with diverse functional activities. For instance, four of them (*OR2AT4, OR4D9, OR52W1* and *OR1K1*) encode olfactory receptors andare located in chromosomes 11 and 9, respectively. We observed that 26 (33%) of them are also present in the Denisova hominin, including those of the above-mentioned chromosome 11 olfactory genes (Supplementary Material Table S1).

Initially, three explanations for this sharing of derived alleles among modern humans and Denisovans were advanced. First, contamination with modern human DNA. Nonetheless, it is unlikely that contamination could explain this degree of sharing, since the level of contamination was calculated to be less than 1% (Reich *et al.*, 2010). Second, C $\rightarrow$ T and G $\rightarrow$ A artifactual mutations due to 5-methylcytosine post-mortem deamination (Briggs *et al.*, 2009), but the frequencies of these artefactual transitions were estimated to be only 4.5% and 5.9%, respectively, in the Neanderthal genome (Green *et al.*, 2010), and in the Denisova study this kind of deamination was chemically reversed, allowing proper sequencing. In addition, the drier and cooler climate at the Denisova Cave resulted in DNA samples that were about ten times less damaged than those of other sites (Reich *et al*., 2010; Liang and Nielsen, 2011). Even if we ignore these transitions, considering that the post-mortem deamination interfered little in the Denisova sequencing due to the reasons presented above, a significant number of shared derived alleles still remains (11; 14%). Third, intermarriages could have occurred between modern human and Denisovan individuals. But this type of admixture was estimated to have a frequency of 4%-6% only (Reich *et al*., 2010).

With these three hypotheses being inconsistent with the results, a more probable one would be that the ancestral species of *Homo sapiens, Homo neanderthalensis* and the Denisova hominin could have been polymorphic at at least 33% of these loci. Following the split of these lineages, different evolutionary trajectories could have then occurred, resulting in the fixation (or near fixation) of derived alleles in some, but not in other lineages. The conclusion then is

that alleles could reach fixation in specific branch(es) of a determined phylogeny from a polymorphic ancestral condition.

Genome-wide search for changes in protein coding sequences

We found 10,447 non-synonymous substitutions in which the derived allele is fixed (9,555) or nearly fixed (892) in humans, including Africans, when compared to chimpanzees. Table 1 lists the number of these sites for each human chromosome. The Grantham score indicated that 43% of the changes in the 10,447 sites could be classified as conservative, 43% as moderately conservative, 10% as moderately radical and 4% as radical. A higher frequency of non-synonymous changes per gene is found in chromosome 21, which is one of the smallest of our genome and a historical landmark in biomedical research (Antonarakis *et al.*, 2002).

Overall, the radical amino acid changes are present in all three hominin genomes, while ancestral alleles were found in nine genes only: three in Neanderthal (*DMRT3, FAM111A* and *RASAL1*) and six in Denisova (*C7orf46, DNAJC12, DUOX1, FREM2, RDM1* and *NKAIN4*). In 14 sites, both ancestral and derived alleles were found (Neanderthal genes: *FASTKD3, CMYA5, VCAN, SLC22A1, AMZ1, STK31, KRT75, OR6T1, MESP2, C17orf66, C17orf78, CCDC57*; Denisova genes: *GGH* and *TMEM99*). See details concerning these genes in the Supplementary Material Table S2.

Table 1 - Non-synonymous substitutions in modern humans when compared with chimpanzees for each human chromosome.

Chromosome	Length (bp)	Known protein-coding genes	Fixed	Nearly fixed	Non-synonymous substitutions per gene
1	247249719	2107	1085	96	0.5605
2	242951149	1333	675	63	0.5536
3	199501827	1095	568	54	0.5680
4	191273063	774	402	31	0.5594
5	180857866	893	523	50	0.6417
6	170899992	1082	509	49	0.5157
7	158821424	983	374	40	0.4212
8	146274826	731	356	30	0.5280
9	140273252	846	317	38	0.4196
10	135374737	812	422	42	0.5714
11	134452384	1358	597	55	0.4801
12	132349534	1055	528	44	0.5422
13	114142980	353	165	10	0.4958
14	106368585	638	329	40	0.5784
15	100338915	652	300	33	0.5107
16	88827254	900	424	36	0.5111
17	78774742	1228	536	43	0.4715
18	76117153	285	136	11	0.5158
19	63811651	1449	499	37	0.3699
20	62435964	576	275	31	0.5313
21	46944323	253	141	26	0.6601
22	49691432	504	221	23	0.4841
X	154913754	968	158	9	0.1725
Y	57772954	89	15	1	0.1798
Total		20964	9555	892	

There are 381 genes that determine radical amino acid changes. Of these, 243 do not have a significant association or form any functional cluster. Forty genes participate in a single independent functional cluster, while 21 are olfactory receptor genes previously reported as being human-specific and the target of positive selection (Gilad *et al.*, 2003; Green *et al.*, 2010; Reich *et al.*, 2010). Among these, the *OR1K1* olfactory gene shows the ancestral allele in both archaic humans, whereas *OR6T1* presents ancestral and derived alleles in Neanderthals only. The other 77 genes are listed as participating in several functional groups and possibly influence multiple phenotypic traits. One example is the *BRCA1* gene, which, besides playing a role on cell cycle regulation, a also affects the immune and nervous systems, as well as metabolic conditions. Table S3 lists all the 381 genes.

Interestingly, eight new alleles could be identified in the Neanderthal and/or in Desinova genomes (Table 2). These variations were found in eight different genes located in seven chromosomes. For example, at position 88,634,660 of chromosome 4 (*SPARCL1* gene) both the ancestral (G) and derived (C) alleles are present in humans (Grantham score 81). In the other hominins, G was detected in Denisova, while in Neanderthal the derived state is represented by a new mutant allele (A), leading to the presence of an aspartic acid (Asn) at position 106 in the amino acid sequence (Grantham score 23). Another illustrative example occurs at position 73,452,685 of chromosome 7 (*CLIP2*). In this case, the new allele (T) was detected in the Denisova genome (Grantham score: 98) only.

Our analyses furthermore indicate that seven of these eight mutations apparently have no functional consequences (Table 2). This apparent neutrality, however, should be considered with caution, since an allele, when considered individually, can have a modest effect, undetectable by the analyses made in the present study.

The change in *ZNF772* (Zinc finger protein 772 - Trp $\rightarrow$ Ter), on the other hand, leads to a premature stop codon. The exact function of this gene is unknown, but the UniProt database (The UniProt Consortium, 2011) indicates, by similarity, that the translated protein may be a transcription factor. Frankel *et al.* (2011) showed for the first time that multiple single-nucleotide substitutions in transcriptional elements could explain the evolution of complex morphologies. It is therefore possible that the 772 - Trp $\rightarrow$ Ter mutation could represent one of a series as those mentioned above.

Table 2 - New alleles identified in the Neanderthal and/or Denisova genomes[1].

Gene	Description	Function	Chromosome	Strand	Position	Chimpanzee	Neanderthal	Denisova	Modern humans	AA position	AA change humans	AA change others	SNP	GS humans	GS others	PolyPhen-2 prediction
CD1C	CD1c molecule	Antigen-presenting protein that binds self and non-self lipid and glycolipid antigens and presents them to T-cell receptors on natural killer T-cells	1	+	156527695	C	C/G	C	C/A	70	Asn > Thr	Asn > Ser	rs3138100	65	46	benign
CLIP2	CAP-GLY domain containing linker protein 2	May be link microtubules to dendritic lamellar body (DLB), a membranous organelle predominantly present in bulbous dendritic appendages of neurons linked by dendrodendritic gap junctions. May operates in the control of brain-specific organelle translocations (By similarity)	7	+	73452685	C	C	C/T	C/G	977	Pro > Arg	Pro > Leu	rs2522943	103	98	benign
COL28A1	collagen, type XXVIII, alpha 1	May act as a cell-binding protein	7	-	7512216	G	G/A	G	G/C	327	Ser > Thr	Ser > Asn	rs10486176	58	46	benign
JDP2	Jun dimerization protein 2	Component of the AP-1 transcription factor that represses transactivation mediated by the Jun family of proteins.May control transcription via direct regulation of the modification of histones and the assembly of chromatin	14	+	74974413	G	ND	G/T	G/A	13	Ala > Thr	Ala > Ser	rs3625	58	99	benign
KAZALD1	Kazal-type serine peptidase inhibitor domain 1	Involved in the proliferation of osteoblasts during bone formation and bone regeneration. Promotes matrix assembly (By similarity)	10	+	102814339	C	C/T	C	C/G	255	Ala > Gly	Ala > Val	rs807037	60	64	benign
KCNJ5	potassium inwardly-rectifying channel, subfamily J, member 5	Potassium channels are present in most mammalian cells, where they participate in a wide range of physiologic responses	11	+	128287222	G	G/A	G	G/C	282	Glu > Gln	Glu > Lys	rs7102584	29	56	benign
SPARCL1	SPARC-like 1 (hevin)	May participate in signal transduction. Proteins are expected to have molecular function (calcium ion binding) and to localize in various compartments (extracellular space, extracellular region, proteinaceous extracellular matrix, synapse)	4	-	88634660	G	A	G	G/C	106	Asp > His	Asp > Asn	rs1049544	81	23	benign
ZNF772	zinc finger protein 772	May be involved in transcriptional regulation (By similarity)	19	-	62677378	G	G/A	G	G/C	182	Trp > Cys	Trp > Ter	rs2074060	215	Stop codon	-

[1]AA = amino acid; SNP = single nucleotide polymorphism; GS= Grantham score; Others = Neanderthal and Denisova; Data from Genecards, Uniprot and AceView databases.

It worth stating that a unique variant may change gene expression or protein function only slightly, but may be decisive for modifications in a gene regulatory/functional pathway (Goldstein, 2009) implications of such "exclusive" variants and conditions, thus, deserve additional investigation.

It is important to note that Green *et al.* (2010) and Reich *et al.* (2010) choose not to consider sites with changes leading to more than one derived allele in their analyses, although they did not rule out the hypothesis that these sites may present real mutational events rather than artifacts due to sequencing errors (see supplementary data in Green *et al.* (2010) and Reich *et al.* (2010).

Genes with evidenve for positive selection in the human lineage when compared to those of other primates

In this category, the most cited gene is *FOXP2* (Enard *et al.*, 2002), where two amino acid changes (Thr303Asn and Asn325Ser) were associated with the emergence of modern language and other unique cognitive human abilities. Krause *et al.* (2007) established that these two amino acid modifications were also present in Neanderthals. They also found that the changes occur in the most common modern human haplotype, defined by nine nucleotides located in the upstream intronic region of *FOXP2* exon 7. We verified changes in the same location in the Denisova specimen, in accordance with a scenario suggested by Krause *et al.* (2007), of a selective sweep which started before the divergence of modern and archaic humans.

Other candidate genes involved in neurogenesis and cognition (*ASPM, MCPH1*, *AHI1,* and *KLK8;* Dorus *et al*., 2004; Gilbert *et al*., 2005; Hill and Walsh, 2005; Evans *et al.*, 2006; Vallender, 2008; Montgomery and Mundy, 2010; Montgomery *et al.*, 2011) were also investigated by us (Table 3). Overall it is possible to see that the derived alleles are commonly found in the archaic genomes, but in some cases both derived and ancestral alleles are found.

An intriguing difference between modern humans *vs.* chimpanzee and other primates is the alternative splicing for *KLK8*. The KLK8 protein, a neuropsin, is preferentially expressed in the central nervous system and is involved in learning and memory. The longer spliced form of this mRNA, due to a mutation (c.71-127T → A), was considered until recently to be expressed only in *H. sapiens* (Li *et al.*, 2004, 2007; Varki *et al.*, 2008). Our results, however, show that both the Neanderthal and Denisova hominins presented this alteration, stressing that cognitive traits seen in modern humans could already be present in archaic hominins.

In conclusion, the findings presented in this study provide clues about the possible course of hominid evolution and illustrate the importance of comparative genomic approaches.

Table 3 - Cognition genes and genetic differences between modern and archaic humans[1].

Genes	Description	Function	Chromosome	Modern Humans/ Chimpanzee non-synonymous diferences	Polymorphic sites in Modern Humans[2]	Derived allele present in Neanderthal	Ancestral allele present in Neanderthal	Exclusive allele present in Neanderthal	Derived allele present in Denisova	Ancestral allele present in Denisova	Exclusive allele present in Denisova
ASPM	Abnormal spindle protein homolog	Probable role in mitotic spindle regulation and coordination of mitotic processes (By similarity). May have a preferential role in regulating neurogenesis	1	21	4	14	3	0	21	0	0
MCHP1	Microcephalin	Implicated in chromosome condensation and DNA damage induced cellular responses. May play a role in neurogenesis and regulation of the size of the cerebral cortex	8	11	4	5	2	1	3	3	0
AHI1	Abelson helper integration site 1 protein homolog	This gene is apparently required for both cerebellar and cortical development in humans	6	12	0	7	0	0	10	0	0
KLK8[3]	Kallikrein-related peptidase 8, Neuropsin	Plays a role in the formation and maturation of orphan and small synaptic boutons in the Schaffer-collateral pathway, regulates Schaffer-collateral long-term potentiation in the hippocampus and is required for memory acquisition and synaptic plasticity	19	1	0	1	0	0	1	0	0

[1]Considering the 45 sites analyzed for the Neanderthal and Denisova genomes, respectively 28% and 5% of the data are missing. [2]Ancestral and derived alleles present. [3]The diference occurs in the non-coding region and changes a splice site. Data from USC Genome, Genecards, Uniprot and NEXTBIO databases.

Acknowledgments

The financial support of Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS, Programa de Apoio a Núcleos de Excelência, PRONEX) is gratefully acknowledged.

References

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS and Sunyaev SR (2010) A method and server for predicting damaging missense mutations. Nat Meth 7:248-249.

Antonarakis SE, Lyle R, Deutsch S and Reymond A (2002) Chromosome 21: A small land of fascinating disorders with unknown pathophysiology. Int J Dev Biol 46:89-96.

Armitage SJ, Jasim SA, Marks AE, Parker AG, Usik VI and Uerpmann HP (2011) The southern route "Out of Africa": Evidence for an early expansion of modern humans into Arabia. Science 331:453-456.

Briggs AW, Stenzel U, Meyer M, Krause J, Kircher M and Pääbo S (2009) Removal of deaminated cytosines and detection of *in vivo* methylation in ancient DNA. Nucleic Acids Res 22:386-397.

Brown P, Sutikna T, Morwood MJ, Soejono RP, Jatmiko, E. Wayhu Saptomo and Rokus AD (2004) A new small-bodied hominin from the Late Pleistocene of Flores. Nature 431:1055-1061.

Brunet M, Guy F, Pilbeam D, Mackaye HT, Likius A, Ahounta D, Beauvilain A, Blondel A, Bocherens CH, Boisserie J, *et al*. (2002) A new hominid from the Upper Miocene of Chad, Central Africa. Nature 418:145-151.

Carroll SB (2003) Genetics and the making of *Homo sapiens*. Nature 422:849-857.

Derevianko AP, Shunkov MV, Agadjaniav AK, Baryshnikov GF, Malaeva EM, Ulianov VA, Kulik NA, Postnov AV and Anoikin AA (2003) Paleoenvironment and Paleolithic Human Occupation of Gorny Altai (Subsistence and Adaptation in the Vicinity of Denisova Cave). Russian Academy of Sciences, Novosibirsk (in Russian, with abstract in Enghish and French).

Dodge DR (2012) A molecular approach to Neanderthal extinction. Quatern Int 259:22-32.

Dorus S, Vallender EJ, Evans PD, Anderson JR, Gilbert SL, Mahowald M, Wyckoff GJ, Malcom CM and Lahn, BT (2004) Accelerated evolution of nervous system genes in the origin of *Homo sapiens*. Cell 119:1027-1040.

Enard W, Przeworski M, Fisher SE, Lai CSL, Wiebe V, Kitano T, Monaco AP and Pääbo S (2002) Molecular evolution of *FOXP2*, a gene involved in speech and language. Nature 418:869-872.

Endicott P, Ho SYW and Stringer C (2010) Using genetic evidence to evaluate four palaeoanthropological hypotheses for the timing of neanderthal and modern human origins. J Hum Evol 59:87-95.

Evans PD, Mekel-Bobrov N, Vallender EJ, Hudson RR and Lahn BT(2006) Evidence that the adaptive allele of the brain size gene *microcephalin* introgressed into *Homo sapiens* from an archaic *Homo* lineage. Proc Natl Acad Sci USA 103:18178-18183.

Frankel N, Erezyilmaz DF, McGregor AP, Wang S, Payre F and Stern DL (2011) Morphological evolution caused by many subtle-effect substitutions in regulatory DNA. Nature 474:598-603.

Fu X, Giavalisco P, Liu X, Catchpole G, Fu N, Ning Z-B, Guo S, Yan Z, Somel M, Pääbo S, *et al.* (2011) Rapid metabolic evolution in human prefrontal cortex. Proc Natl Acad Sci USA 108:6181-6186.

Fujita PA, Rhead B, Zweig AS, Hinrichs AS, Karolchik D, Cline MS, Goldman M, Barber GP, Clawson H, Coelho A, *et al*. (2011) The UCSC Genome Browser database: Update 2011. Nucleic Acids Res 876-882.

Gibbons A (2002) In search of the first hominids. Science 295:1214-1219.

Gilad Y, Bustamante C, Lancet D and Pääbo S (2003) Natural selection on the olfactory receptor gene family in humans and chimpanzees. Am J Hum Genet 73:489-501.

Gilbert SL, Dobyns WB and Lahn BT (2005) Genetic links between brain development and brain evolution. Nat Rev Genet 6:581-590.

Goldstein DB (2009) Common genetic variation and human traits. N *e.g.* J Med 360:1696-1698.

González-José R, Escapa I, Neves WA, Cúneo R and Pucciarelli HM (2008) Cladistic analysis of continuous modularized traits provides phylogenetic signals in *Homo* evolution. Nature 453:775-778.

Grantham R (1974) Amino acid difference formula to help explain protein evolution Science 185:862.

Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai W, Fritz MH, *et al*. (2010) A draft sequence of the Neandertal genome. Science 328:710-722.

Haile-Selassie Y (2001) Late Miocene hominids from the Middle Awash, Ethiopia. Nature 412:178-181.

Haile-Selassie Y, Asfaw B and White TD (2004) Hominid cranial remains from Upper Pleistocene deposits at Aduma, Middle Awash, Ethiopia. Am J Phys Anthropol 123:1-10.

Hill RS and Walsh CA (2005) Molecular insights into human brain evolution. Nature 437:64-67.

Hünemeier T, Ruiz-Linares A, Silveira A, Paixão-Côrtes VR, Salzano FM and Bortolini MC (2010) Population data support the adaptive nature of HACNS1 *sapiens*/neandertal-chimpanzee differences in a limb expression domain. Am J Phys Anthropol 143:478-481.

Johanson DC and Edgar B (1996) From Lucy to Language. Simon & Schuster, New York, 272 pp.

Kimbel WH and Delezene LK (2009) "Lucy" redux: A review of research on *Australopithecus afarensis*. Am J Phys Anthropol 140:42-48.

Krause J, Lalueza-Fox C, OrlandoL, Enard W, Green REE, Burbano HAA, Hublin J-JJ, Hänni C, Fortea J, de la Rasilla M, *et al.* (2007) The derived FOXP2 variant of modern humans was shared with neandertals. Curr Biol 17:1908-1912.

Leakey MG, Feibel CS, McDougall I, Ward C and Walker A (1998) New specimens and confirmation of an early age for *Australopithecus anamensis*. Nature, 393:62-66.

Li WH, Wu CI and Luo CC (1985) A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes Mol Biol Evol 2:150-174.

Li Y, Qian Y-P, Yu X-J, Wang Y-Q, Dong D-G, Sun W, Ma R-M and Su B (2004) Recent origin of a hominoid-specific splice

form of neuropsin, a gene involved in learning and memory. Mol Biol Evol 21:2111-2115.

Liang M and Nielsen R (2011) Who is *H. sapiens* really, and how do we know? BMC Biology 9:e20.

Lu Z-X, Peng J and Su B (2007) A human-specific mutation leads to the origin of a novel splice form of neuropsin (*KLK8*), a gene involved in learning and memory. Hum Mut 28:978-984.

Montgomery SH and Mundy NI (2010) Brain evolution: Microcephaly genes weigh in. Curr Biol 20:244-246.

Montgomery SH, Capellini I, Venditti C, Barton RA and Mundy NI (2011) Adaptive evolution of four microcephaly genes and the evolution of brain size in anthropoid primates. Mol Biol Evol 28:625-638.

Rasmussen M, Li Y, Lindgreen S, Pedersen JS, Albrechtsen A, Moltke I, Metspalu M, Metspalu E, Kivisild T, Gupta R, *et al.* (2010) Ancient human genome sequence of an extinct Palaeo-Eskimo. Nature 463:757-762.

Reich D, Green RE, Kircher M, Krause J, Patterson N, Durand EY, Viola B, Briggs AW, Stenzel U, Johnson PLF, *et al.* (2010) Genetic history of an archaic hominin group from Denisova Cave in Siberia. Nature 468:1053-1060.

Schuster SC, Miller W, Ratan A, Tomsho LP, Giardine B, Kasson LR, Harris RS, Petersen DC, Zhao F, Qi J, *et al.* (2010) Complete Khoisan and Bantu genomes from southern Africa. Nature 463:943-947.

Schwartz JH and Tattersall I (2010) Fossil evidence for the origin of *Homo sapiens*. Am J Phys Anthropol 143:94-121.

Suwa G, Asfaw B, Kono RT, Kubo D, Lovejoy OC and White TD (2009) The *Ardipithecus ramidus* skull and its implications for hominid origins. Science 326:681-687.

Tattersall I (1997) Out of Africa again... and again? Sci Am 276:60-67.

The 1000 Genomes Project Consortium (2010) A map of human genome variation from population-scale sequencing. Nature 467:1061-1073.

The UniProt Consortium (2011) Ongoing and future developments at the Universal Protein Resource. Nucleic Acids Res 39:D214-D219.

Vallender EJ (2008) Exploring the origins of the human brain through molecular evolution. Brain, Behav Evol 72:168-177.

Varki A (2004) How to make an ape brain. Nat Genet 10:1061-1063.

Varki A and Altheide TK (2005) Comparing the human and chimpanzee genomes: Searching for needles in a haystack. Genome Res 15:1746-1758.

Varki A, Geschwind DH and Eichler EE (2008) Human uniqueness: Genome interactions with environment, behaviour and culture. Nat Rev Genet 9:749-763.

White TD, Asfaw B, Beyene Y, Haile-Selassie Y, Lovejoy OC, Suwa G and Wolde Gabriel G (2009) *Ardipithecus ramidus* and the paleobiology of early hominids. Science 326:75-86.

White TD, Asfaw B, Degusta D, Gilbert H, Richards GD, Suwa G and Clark H (2003) Pleistocene *Homo sapiens* from Middle Awash, Ethiopia. Nature 423:742-747.

Internet Resources

Homo neanderthalensis draft genome, http://genome.ucsc.edu/Neandertal/ (May 30, 2011).

USCS Genome Browser, http://genome.ucsc.edu/ (May 30, 2011).

USCS Table Browser, http://genome.ucsc.edu/cgi-bin/hgTables?org=human (May 30, 2011).

USCS Blat tool, http://genome.ucsc.edu/cgi-bin/hgBlat?org=human, (May 30, 2011).

GeneDecks V3 software, http://www.genecards.org/ (June 30, 2011).

PolyPhen-2 (Polymorphism Phenotyping v2), http://genetics.bwh.harvard.edu/pph2/index.shtml (May 30, 2011).

UniProt database, http://www.uniprot.org/ (May 30, 2011).

NESTBIO database, http://www.nextbio.com/b/nextbio.nb (Junho 30, 2011)

AceView database, http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/index.html (Junho 30, 2012).

Supplementary Material

The following online material is available for this article:

Table S1 - Amino acid changes determined by derived alleles present in extant humans and Denisova.

Table S2 - Genes with radical amino acid changes in Neanderthal or Denisova genomes.

Table S3 - 381 genes with amino acid radical changes between modern and archaic humans.

This material is available as part of the online article from http://www.scielo.br/gmb.

Table S2- Twenty-three genes with radical amino acid changes with ancestral or both (ancestral and derived) alleles found in Neanderthal or Denisova genomes

Symbol	Gene	Chr	Strand	Start	End	Size	Function	Neanderthal allele	Denisovan allele
FASTKD3	FAST kinase domain-containing protein 3	5	-	7859272	7869150	9878	May be mitochondrial protein essential for cellular respiration.	A/D	D
CMYA5	Cardiomyopathy-associated protein 5	5	+	78985659	79096063	110404	May be involved in protein kinase A signalling and vesicular trafficking. Was associated to a cardiac disease.	A/D	D
VCAN	Chondroitin sulfate proteoglycan core protein 2	5	+	82767284	82878122	110838	May play a role in intercellular signaling and in connecting cells with the extracellular matrix. May take part in the regulation of cell motility, growth and differentiation. Binds hyaluronic acid.	A/D	D
SLC22A1	Solute carrier family 22 member 1	6	+	160542805	160579750	36945	Translocates a broad array of organic cations with various structures and molecular weights.	A/D	D
C7orf46	Chromosome 7 open reading frame 46	7	+	23719749	23742868	23119	Unknown	D	A
AMZ1	Archeobacterial metalloproteinase-like protein 1	7	+	2719156	2804759	85603	Zinc metalloprotease. Exhibits aminopeptidase activity against neurogranin in vitro.	A/D	A
STK31	Serine/threonine kinase 31	7	+	23749786	23872132	122346	May have a role in reorganization of sperm chromatin during spermiogenesis.	A/D	D
GGH	Gamma-Glu-X carboxypeptidase	8	-	63927638	63951730	24092	May play an important role in the bioavailability of dietary pteroylpolyglutamates and in the metabolism of pteroylpolyglutamates and antifolates.	D	A/D
DMRT3	Doublesex- and mab-3-related transcription factor 3	9	+	976964	991732	14768	May regulate transcription during sexual development (By similarity).	A	D
DNAJC12	DnaJ homolog subfamily C member 12	10	-	69556427	69597937	41510	Members of this family of proteins are associated with complex assembly, protein folding, and export.	D	A
FAM111A	Family with sequence similarity 111, member A	11	+	58910221	58922512	12291	May be involved with methylation.	A	D
OR6T1	Olfactory receptor, family 6, subfamily T, member 1	11	-	123813492	123814580	1088	Odorant receptor (Potential).	A/D	D
RASAL1	RasGAP-activating-like protein 1	12	-	113536624	113574044	37420	Probable inhibitory regulator of the Ras-cyclic AMP pathway.	D	D
KRT75	Keratin type II cytoskeletal 75	12	-	52817854	52828309	10455	Plays a central role in hair and nail formation. Essential component of keratin intermediate filaments in the companion layer of the hair follicle.	A/D	D

Cont.

Table S2- Cont.

Symbol	Gene	Chro	Strand	Start	End	Size	Function	Neanderthal allele	Denisovans allele
FREM2	FRAS1-related extracellular matrix protein 2	13	+	39261173	39461268	200095	Extracellular matrix protein required for maintenance of the integrity of the skin epithelium and for maintenance of renal epithelia. May be required for epidermal adhesion.	ND	D
DUOX1	Nicotinamide adenine dinucleotide phosphate oxidase	15	+	45422192	45457774	35582	Plays a role in thyroid hormones synthesis and lactoperoxidase-mediated antimicrobial defense at the surface of mucosa. May have its own peroxidase activity through its N-terminal peroxidase-like domain.	D	A
MESP2	Mesoderm posterior protein 2 BHLHC6	15	+	90319589	90321985	2396	Transcription factor with important role in somitogenesis. Together with MESP1 is involved in the epithelialization of somitic mesoderm and in the development of cardiac mesoderm.	A/D	D
RDM1	RAD52 homolog B	17	-	34245070	34257780	12710	May confer resistance to the antitumor agent cisplatin. Binds to DNA and RNA.	A/D	D
C17orf66	Chromosome 17 open reading frame 66	17	-	34181955	34195895	13940	Unknown	A/D	D
C17orf78	Chromosome 17 open reading frame 78	17	+	35732985	35749662	16677	Unknown	A/D	D
CCDC57	Coiled-coil domain-containing protein 57	17	-	80059346	80170689	111343	Unknown	A/D	D
TMEM99	transmembrane protein 99	17	+	38975358	38992526	17168	Unknown	ND	A/D
NKAIN4	Na+/K+ transporting ATPase interacting 4	20	-	61869286	61904046	34760	May interact with the beta subunit of Na,K-ATPase (By similarity)	D	A

*Chr = Chromosome; A=ancestral allele; D= derived allele; ND= no data; data from Genecards, Uniprot and NESTBIO databases.

Table S1- Amino acid changes determined by derived alleles present in extant humans and Denisova

Genes	Description	Human (derived)					Chimp	Neand.	Denisova
		Base	AA Change	Chrom.	Strand	Position	Base	Base	Base
ACCN4	Amiloride-sensitive cation channel 4, pituitary	T	V/A	2	+	220087788	C	C	T
ARSF	Arylsulfatase F	A	I/V	X	+	3012475	G	G	A
CDCA2	Cell division cycle associated 2	A	I/V	8	+	25416950	G	G	A
CF170	Chromosome 6 open reading frame 170	T	S/C	6	+	121644484	A	A	T
C12orf66	Chromomosome 12 open reading frame 66	C	V/L	12	+	62873951	G	G	C
EMIL2	Elastin microfibril interfacer 2	G	R/K	18	+	2880589	A	A	G
EYA2	Eyes absent homolog 2 (Drosophila)	T	S/P	20	+	45078304	C	C	T
FSTL4	Follistatin-like 4	T	T/A	5	+	132562844	C	C	T
KRT16	Keratin 16	T	T/A	17	+	37020864	C	T	T
LIMS2	LIM and senescent cell antigen-like domains 2	T	T/A	2	+	128113346	C	C	T
NSUN3	NOP2/Sun domain family, member 3	C	S/F	3	+	95285751	T	T	C
OR2AT4	Olfactory receptor, family 2, subfamily AT, member 4	A	V/A	11	+	74477736	G	G	A
OR52W1	Olfactory receptor, family 52, subfamily W, member 1	C	P/L	11	+	6177181	T	T	C
OR4D9	Olfactory receptor, family 4, subfamily D, member 9	G	R/K	11	+	59039869	A	A	G
DCHS1	Dachsous 1 (Drosophila)	C	E/Q	11	+	6611387	G	G	C
PEG3	Paternally expressed 3	T	S/G	19	+	62017061	C	C	T
PHLP	Phosducin-like	T	K/R	9	+	124622444	C	C	T
RGS16	Regulator of G-protein signaling 16	T	D/A	1	+	180836069	G	G	T
RNAS7	Ribonuclease, RNase A family, 7	A	M/V	14	+	20581121	G	G	A
RPTN	Repetin	A	*/R	1	+	150393846	G	G	A
RPTN	Repetin	T	K/E	1	+	150393996	C	C	T
SLC36A4	Solute carrier family 36	T	H/R	11	+	92535568	C	C	T
SNTG1	Syntrophin, gamma 1	C	T/S	8	+	51628204	G	G	C
STEA1	Six transmembrane epithelial antigen of the prostate 1	G	C/S	7	+	89631971	C	C	G
TRPM5	Transient receptor potential cation channel, subfamily M, member 5	T	I/V	11	+	2383887	C	T	T
ZFY26	Zinc finger, FYVE domain containing 26	T	H/R	14	+	67344044	C	C	T

Table S3 - 381 genes that exhibit amino acid radical changes between modern and archaic humans

Number	Classes	Genes
243	Genes not having a significant association or forming a functional cluster	*A2ML1, A4GNT, ABCA8, ACP6, ACTL7B, ADAMTS14, ADC, AGBL2, AGXT2, AKAP11, AMZ1, ANKFN1, ARHGEF26, ARHGEF37, ARSJ, B3GNT4, BIN2, C10orf47, C11orf16, C13orf31, C14orf28, C14orf49, C16orf46, C17orf55, C17orf66, C17orf78, C19orf57, C1orf101, C1orf110, C1orf129, C2orf71, C3orf62, C4orf37, C5orf39, C6orf146, C7orf46, C7orf62, C9orf173, CAMSAP1L1, CASC5, CCDC116, CCDC146, CCDC157, CCDC28A, CCDC51, CCDC57, CCDC8, CCDC83, CCL16, CDCP1, CEP63, CEP70, CMAS, CMYA5, COQ3, CPSF3, CUEDC1, CUL9, CYP4X1, DCXR, DDX42, DEFB132, DEGS2, DIS3L, DMGDH, DNAH3, DNAH7, DNAJC12, DNAJC5G, DTX3L, DUOX1, EIF5B, EME1, ESCO2, EXOSC3, FAM111A, FAM187B, FAM24A, FAM53A, FAM71A, FAM83A, FASTKD3, FAT2, FBLIM1, FCRL6, FDXR, FGFBP1, FMO1, FMR1NB, GALNT5, GCAT, GGH, GJD4, GSTM3, GTPBP5, GTPBP8, GTSE1, HEATR7B2, HERC4, HUS1B, IBTK, IFI44, IFIT2, IL22RA2, IPPK, IQCA1, KCNA10, KCNK17, KCNV2, KIAA0391, KIAA1199, KIAA1328, KIAA1377, KIF15, KIF6, KNTC1, KRT76, KRTAP11-1, KRTAP20-1, LELP1, LRIT2, LRRK1, LY75-CD302, MARCH10, MARS2, METTL7A, MMAB, MPST, MRPL50, MRPS28, MRPS7, MS4A12, MTIF2, MTMR8, MYO9A, MYOM1, NDUFB2, NKAIN4, NLRP14, NTN5, NUP107, NUP188, NUP85, OPLAH, P2RY6, PADI3, PCDHB1, PCM1, PDCL, PDE9A, PGCP, PHACTR2, PHLDB2, PIK3C2A, PLA2G2D, PLA2G4F, PLD2, PLD4, PLEKHH2, PNPLA5, PREPL, PSD2, PTER, PUS10, RAET1E, RANBP17, RASAL1, RBM23, RDM1, RFPL4B, RGL4, RHOBTB1, RNASE9, RNASET2, RNF103-VPS24, RNF125, RNF213, RNF32, RPAP2, RPL10L, RRP12, RUSC2, SAGE1, SARDH, SCML1, SCRN1, SCYL3, SIGLEC15, SLA2, SLC15A3, SLC22A18, SLC36A1, SLC9A11, SLCO1A2, SLCO1B1, SMC4, SNRNP48, SON, SPATA18, SPATA5L1, SPHKAP, SPPL2A, SRBD1, SRFBP1, STH, STK31, SV2C, SYTL3, TAS2R60, TBL1Y, TEP1, TEX9, TMC7, TMCO5A, TMED3, TMEM71, TMEM99, TRAPPC9, TRIM36, TRIM6, TRIM6-TRIM34, TUFT1, TXNDC3, UBA5, UBAP2, UGT3A1, USP2, USP29, USP44, WDR96, ZBBX, ZCWPW2, ZFYVE9, ZNF212, ZNF214, ZNF254, ZNF280A, ZNF408, ZNF438, ZNF474, ZNF606, ZNF648, ZNF778.*
77	Genes forming several functional groups and that would be influencing multiple phenotypic traits	*ABCC4, ANKRD11, BLM, BRCA1, C5, CCL24, CD274, CDK5RAP2, CIITA, EGF, ELOVL3, FAS, FLT4, G6PC2, GPR44, HRG, IL15RA, IL2RA, ITPKC, LGALS3BP, LY75, LY9, NLRP12, PKHD1, POLM, POLQ, PZP, RECQL4, SALL4, SLAMF6, SLC12A1, SQSTM1, SST, SYNE1, TAX1BP1, TNFRSF9, TNS1, TRAFD1, TRPM2, ABCA12, ADRB3, APOB, CER1, COL19A1, COQ7, CRYM, DNAJA3, FREM2, GPRC6A, ITPR3, KIAA1409, LAMA1, MAPT, MESP2, MYH4, NEK1, OFD1, STIL, TRH, TRIP11, UTRN, C2CD3, CDH23, LMTK2, MC5R, NPHS1, P2RY1, CNGA3, GPR98, GRIN2C, MREG, PER3, RNF103, DMRT3, RBBP8, FRAS1, TBX22.*
40	Genes participating of a single independent functional cluster	*CARD6, CLEC7A, ERN2, MAPKAPK3, MS4A2, AKAP6, CDC25C, TRAIP, ABCC2, EXO1, GCKR, NAT2, PAFAH2, SLC22A1, TEX15, C3orf54, DLC1, PDE6C, RBP3, RDH8, RP1, RPGRIP1, USH2A, VCAN, FANCC, HPSE, INSL6, MLH3, RXFP2, SYCP1, ZPBP2, EVPL, IL1RL2, KRT75, MAP3K11, ANK2, PTPRH, SLC4A9, TRPV6, CENPE.*
21	Olfactory receptor genes	*OR13C8, OR14A16, OR14I1, OR2G3, OR2L13, OR2Z1, OR4D6, OR4K1, OR52H1, OR5C1, OR5D13, OR5D14, OR5M11, OR6K2, OR6K6, OR6N1, OR6Q1, OR6T1, OR8J1, OR8K1.*

Data obtained from GeneDecks V3 software (http://www.genecards.org/).