

Evolution of X-Degenerate Y Chromosome Genes in Greater Apes: Conservation of Gene Content in Human and Gorilla, But Not Chimpanzee

Hiroki Goto · Lei Peng · Kateryna D. Makova

Received: 9 June 2008 / Accepted: 1 December 2008 / Published online: 14 January 2009
© Springer Science+Business Media, LLC 2009

Abstract Compared with the X chromosome, the mammalian Y chromosome is considerably diminished in size and has lost most of its ancestral genes during evolution. Interestingly, for the X-degenerate region on the Y chromosome, human has retained all 16 genes, while chimpanzee has lost 4 of the 16 genes since the divergence of the two species. To uncover the evolutionary forces governing ape Y chromosome degeneration, we determined the complete sequences of the coding exons and splice sites for 16 gorilla Y chromosome genes of the X-degenerate region. We discovered that all studied reading frames and splice sites were intact, and thus, this genomic region experienced no gene loss in the gorilla lineage. Higher nucleotide divergence was observed in the chimpanzee than the human lineage, particularly for genes with disruptive mutations, suggesting a lack of functional constraints for these genes in chimpanzee. Surprisingly, our results indicate that the human and gorilla orthologues of the genes disrupted in chimpanzee evolve under relaxed functional constraints and might not be essential. Taking mating patterns and effective population sizes of ape species into account, we conclude that genetic hitchhiking associated with positive selection due to sperm competition might explain the rapid decline in the Y chromosome gene number in chimpanzee. As we found no evidence of positive selection acting on the X-degenerate genes, such

selection likely targets other genes on the chimpanzee Y chromosome.

Keywords Y chromosome · Genetic degeneration · Sperm competition · Sex chromosomes · Ape evolution · Gene content

Introduction

The human Y chromosome has attracted special attention from geneticists, evolutionary biologists, and the general public because of its distinctive role in sex determination and fertility. Muller (1914) first suggested that mammalian sex chromosome pairs evolved from a pair of autosomes, and Ohno (1967) developed this idea to explain the evolution of the ZW chromosome pair. The mammalian sex chromosomes, X and Y, evolved from a pair of autosomes ~166 million years ago (MYA; Lahn and Page 1999; Graves 2006; Veyrunes et al. 2008; Potrzebowski et al. 2008). Since then, the Y chromosome has undergone a series of inversions, preventing recombination with the X chromosome over most of its length, except for short pseudoautosomal regions. Consequently, the Y chromosome has degenerated substantially in both size and gene content (Charlesworth and Charlesworth 2000). For instance, in human, while the X chromosome spans ~155 megabases (Mb) and contains 1,098 genes (Ross et al. 2005), the Y chromosome is only ~58 Mb long and includes 86 genes, which encode only 23 distinct proteins (Skaletsky et al. 2003).

The male-specific (nonrecombining) portion of the human Y chromosome consists of three regions: ampliconic, X-transposed, and X-degenerate (Skaletsky et al. 2003). The ampliconic region is comprised of genes

Electronic supplementary material The online version of this article (doi:10.1007/s00239-008-9189-y) contains supplementary material, which is available to authorized users.

H. Goto · L. Peng · K. D. Makova (✉)
Department of Biology and Center for Comparative Genomics
and Bioinformatics, 305 Wartik Laboratory, The Pennsylvania
State University, University Park, PA 16802, USA
e-mail: kdm16@psu.edu

located within palindromes in which gene decay might have been forestalled because of intrachromosomal gene conversion (Rozen et al. 2003; Skaletsky et al. 2003). The X-transposed region harbors only two functional protein-coding genes originating from an X-to-Y transposition that occurred in the human lineage after its divergence from the chimpanzee lineage (Skaletsky et al. 2003). The X-degenerate region on the Y chromosome (XDY) encodes 16 protein-coding genes that are the remnants of the ancient homologous genes between the X and the Y chromosome.

Several population genetic models have been proposed to explain Y chromosome degeneration (Charlesworth and Charlesworth 2000). Under relaxed purifying selection models (background selection, Hill-Robertson effect with weak selection, and Muller's ratchet), deleterious mutations accumulate due to random genetic drift (Charlesworth and Charlesworth 2000). Under the model of genetic hitchhiking, the fixation of deleterious mutations is caused by their linkage to positively selected advantageous mutations (Maynard Smith and Haigh 1974; Charlesworth and Charlesworth 2000; Hughes et al. 2005). Thus, the relaxed purifying selection models are most potent under strong genetic drift (i.e., in populations with small effective size), while the hitchhiking model works best under intense positive selection. The relative contribution of each of these models to Y chromosome evolution is presently unknown.

A recent study revealed that 4 of a total of 16 XDY genes are disrupted in the chimpanzee compared to the human Y chromosome by virtue of point mutations (Hughes et al. 2005; Kuroki et al. 2006; Perry et al. 2007). Perry et al. (2007) found that the majority of these mutations originated in the chimpanzee lineage after its split from the human lineage, but before its split from the bonobo lineage. Hughes et al. (2005) speculated that the decay of the chimpanzee XDY genes might be a by-product of genetic hitchhiking due to positive selection related to sperm competition. Indeed, since each chimpanzee female usually has multiple sexual partners during the same periovulatory period (polyandrous species), while each human female typically has a single partner (monoandrous species), sperm competition is more intense in chimpanzee than in human (Goodall 1986; Hasegawa and Hiraiwai-Hasegawa 1990; Dixson 1998). As a result, potential positive selection acting on genes important for sperm competition and the associated hitchhiking effect are expected to be stronger in chimpanzee than in human. This hypothesis has been based on an analysis of just two species (human and chimpanzee) and it is uncertain whether a higher degree of gene preservation in species with less sperm competition represents a general phenomenon. Additionally, it has not been tested whether XDY genes in chimpanzee evolved under selection.

To shed light on the evolutionary forces leading to genetic degeneration of ape Y chromosomes, we determined the complete coding sequences of 16 XDY genes in gorilla (*Gorilla gorilla*), a monoandrous species (Dixson 1998). The effective population size of gorilla is similar to or higher than that of chimpanzee (Stone et al. 2002; Yu et al. 2003, 2004; Fischer et al. 2004, 2006; Thalmann et al. 2007). In contrast, human effective population size is approximately two or three times smaller (Kaessmann et al. 1999; Yu et al. 2004). Specifically, we asked two questions. First, have XDY genes decayed in the gorilla lineage? While Perry et al. (2007) sequenced gorilla XDY exons mutated in chimpanzee, the complete coding sequences of gorilla XDY genes have not been determined prior to the present study. Second, using gorilla as an outgroup, we analyzed the lineage-specific divergence and tested for positive selection in the human and chimpanzee XDY genes. If selection due to sperm competition (accompanied by genetic hitchhiking) is important for Y chromosome evolution, we expect elevated divergence at chimpanzee XDY genes. If mechanisms determined by genetic drift, most potent with small effective population size, are important, we anticipate elevated divergence in the human lineage. Due to a short divergence time between gorilla and the human-chimpanzee common ancestor (~1 million years [Glazko and Nei 2003]), the former might not represent an ideal outgroup for polarizing human-chimpanzee substitutions that occurred on autosomes or chromosome X (Chen and Li 2001; Perry et al. 2007). This problem is less prominent for the gorilla Y chromosome than other chromosomes, because the Y chromosome has a shorter coalescent time due to the smaller effective population size compared to X and autosomes. Thus gorilla Y is appropriate for separating human- vs. chimpanzee-specific substitutions. Moreover, gorilla XDY gene sequences allow us to polarize human vs. chimpanzee substitutions and to compute lineage-specific divergence values from a larger number of sites than when using human X homologues as an outgroup.

Materials and Methods

Sequence Data for Human and Chimpanzee

Using the mRNA annotations listed by Skaletsky et al. (2003) as a reference (Supplementary Table S1), we retrieved the coding sequences for 16 genes in the X-degenerate region of human and chimpanzee Y chromosomes (human and chimpanzee XDY) from the UCSC Genome Browser (hg18 and panTro2 for human and chimpanzee, respectively; <http://genome.ucsc.edu> [Hinrichs et al. 2006]). The genomic sequences for 14

homologous genes on the human X chromosome were obtained similarly. The 16 XDY genes correspond to only 14 homologous X chromosome genes because two XDY gene pairs are each homologous to a single gene on the X. Indeed, *RPS4Y1* and *RPS4Y2* originated from a Y-specific gene duplication preceding the human-chimpanzee divergence and thus correspond to a single X chromosome homologue, *RPS4X* (Skaletsky et al. 2003; Hughes et al. 2005; Kuroki et al. 2006). Similarly, *CYorf15A* and *CYorf15B* are homologous to the 5' and 3' regions of *CXorf15*, respectively (Skaletsky et al. 2003).

DNA Samples

DNA samples were obtained from the Integrated Biomaterials and Information Resource (IPBIR; <http://www.ipbir.org>), Aging Cell Repository (NIA; <http://www.nia.nih.gov>), and Natural Science Research Laboratory at Texas Tech University. We determined the complete coding sequences of the XDY genes in a male gorilla (*Gorilla gorilla*; PR00573). DNA of a female gorilla (*Gorilla gorilla*; NG05251) was used as a negative control to ensure Y-specific amplification (see below). To clarify the gene structure differences among species for some exons, we analyzed DNA samples from three human males, one chimpanzee male (*Pan troglodytes verus*; NA03450), and one additional gorilla male (*Gorilla gorilla*; TK26847).

Amplification, Cloning, and Sequencing

The Y-chromosome-specific primers were designed based on alignments of X and Y chromosome genes and using either Oligo Lite Version 6.71 (Molecular Biology Insights, Inc., Cascade, CO) or Primer3 (Rozen and Skaletsky 2000). The first amplification was performed in 10 μ l with 50 ng genomic DNA, 1 \times buffer, 0.2 mM dNTPs (PCR grade; Roche Applied Science, Indianapolis, IN), 0.7 units Expand High Fidelity PCR Enzyme mix (Roche Applied Science), 1.5 mM MgCl₂, and 0.28 μ M forward and reverse Y-specific primers (Integrated DNA Technologies, Inc., Skokie, IL). Using PCR products of the first reaction as a template, a second amplification was carried out in 25 μ l under the same conditions. Primer sequences and PCR conditions are described in Supplementary Table S2.

For each pair of primers, DNA from male and female gorilla was amplified in parallel. A PCR product was considered to be Y-chromosome-specific when it was present only in reactions with male template DNA (and not female). Following the amplification, the Y-chromosome-specific PCR products were extracted from gel (Qiagen, Inc., Valencia, CA). For exon 5 of *PRKY*, it was challenging to design Y-chromosome-specific primers. Therefore, universal primers amplifying this exon from

both *PRKX* and *PRKY* were designed and amplification products were subcloned into the pCR2.1-TOPO vector with the TOPO TA Cloning kit (Invitrogen Life Technologies, Carlsbad, CA). Seven clones were screened by sequencing and the Y-specific exon was identified by a comparison to the X-chromosome-specific exon amplified from females.

All sequencing reactions were carried out in 10 μ l with 35 fmol purified PCR product, 1 μ M sequencing primer, and 2 μ l DTCS Quick Start Master Mix (Beckman Coulter, Inc., Fullerton, CA) or BigDye Terminator v.3.1 (Applied Biosystems, Foster City, CA). Thermal cycling conditions were 96°C for 20 s, 50°C for 20 s, and 60°C for 4 min for 40 cycles. Sequencing products were purified by ethanol precipitation, then separated and analyzed on a CEQ8000 (Beckman Coulter, Inc.) or an ABI Hitachi 3730XL (Applied Biosystems). Sequencing contigs for each of the PCR fragments were assembled in the SeqManII module of the Lasergene sequence analysis software (DNASTAR, Inc., Madison, WI). All sequences generated here were deposited in GenBank (accession numbers FJ532255–FJ532278).

Data Analysis

Pairwise nucleotide divergences were estimated with the TN93 model (Tamura and Nei 1993) implemented in MEGA4.0 (Tamura et al. 2007). The use of different models did not alter the results qualitatively. Standard errors were computed by the bootstrap method (1,000 replicas). Maximum likelihood (ML) methods were employed for the lineage-specific analyses using the modules of PAML (version 3.15; Yang 1997). The lineage-specific divergences were calculated according to the TN93 model (Tamura and Nei 1993) as implemented in the baseml module of PAML (Yang 1997). The nonsynonymous and synonymous lineage-specific rates (K_A and K_S , respectively) were estimated by the modified NG (Nei and Gojobori 1986) and ML (Goldman and Yang 1994) methods. Fisher's exact tests were carried out to evaluate whether the differences in divergence or in the K_A/K_S ratio were significant.

Monte Carlo simulations were performed to determine whether the ratio of the numbers of nonsynonymous-to-synonymous substitutions (N/S) was significantly different from expected for the disrupted gene group in a human-gorilla comparison. First, we calculated the number of synonymous (S) and nonsynonymous (N) substitutions for each gene. Four human-gorilla homologous gene pairs were randomly picked up from 16 gene pairs, and the N/S ratio for them was calculated. This step was repeated 10,000 times. As a result, a simulated frequency distribution of the N/S ratio was generated. An empirical p value

was calculated by comparing this distribution to the observed N/S ratio of the disrupted orthogroup.

Using PAML software (version 3.15; Yang 1997), the likelihood ratio test (LRT) was applied to examine the following three hypotheses: (1) that the K_A/K_S ratio was significantly higher in the branch of interest than the background branch (tests B and C of Yang [1998]); (2) that the K_A/K_S ratio in the branch of interest was significantly different from 1 (test B' or C' of Yang [1998]); and (3) that positive selection acted on the branch of interest (improved branch-site model; test 2 of Zhang et al. [2005]). The tests were performed by comparing the log-likelihood values between the null and the alternate hypotheses. Bonferroni correction was applied to correct for multiple tests. Additionally, we compared M1a-M2a, M7-M8, and M8-M8a models but did not detect any significant results (data not shown).

Results

Gorilla X-Degenerate Y Chromosome Genes

We determined the complete sequences of coding exons and their splice sites for 14 gorilla XDY genes, including *AMELY*, *CYorf15A*, *CYorf15B*, *DDX3Y*, *EIF1AY*, *JARID1D*, *NLGN4Y*, *PRKY*, *RPS4Y1*, *RPS4Y2*, *TBLIY*, *TMSB4Y*, *USP9Y*, and *UTY*. The sequences of two additional gorilla XDY genes, *SRY* and *ZFY*, were obtained from GenBank (accession numbers X86382 and AY913764, respectively). After aligning the sequences of coding regions and splice sites of human, chimpanzee, and gorilla XDY genes, we noticed that seven gene-disruptive mutations found in chimpanzee *CYorf15B*, *TBLIY*, *TMSB4Y*, and *USP9Y* genes were absent from orthologous genes in gorilla and human. This agrees with observations of Perry et al. (2007).

Our results indicated that neither loss nor substantial alterations in function of the XDY genes occurred in the gorilla lineage after its divergence from the human-chimpanzee-gorilla common ancestor. Indeed, no frameshift or splice site mutations were detected in gorilla compared with human XDY genes. However, in two instances, gene lengths differed between the two species (Supplementary

Fig. S1). First, in the last exon of *TBLIY*, the stop codon was present two codons upstream in human compared to gorilla. Sequencing of this exon in three human males, one chimpanzee male, and an additional gorilla male indicated that this mutation occurred in the human lineage. This human-specific premature stop codon is located outside of the WD40 repeats important for interactions between *TBLIY* and other proteins (Smith et al. 1999; Ono 2003) and thus is unlikely to affect the protein's function. Second, we discovered a mutation introducing a premature stop codon in the gorilla *UTY* gene (occurring 16 codons upstream compared to the human or chimpanzee *UTY* gene). Sequencing of this exon in three human males, one chimpanzee male, and an additional gorilla male showed that this mutation was specific to the gorilla lineage. Again, since the premature stop codon in gorilla *UTY* is located outside of the tetratricopeptide repeats (TPRs) shown to be involved in protein-protein interactions (Greenfield et al. 1998), this mutation is not expected to have an impact on the protein's function.

Pairwise Nucleotide Divergence at XDY Genes

We present the pairwise nucleotide divergences at total and synonymous sites for concatenated XDY genes between human, chimpanzee, and gorilla in Table 1. Several observations were evident. First, the human-chimpanzee divergence was lower than the human-gorilla and gorilla-chimpanzee divergences, consistent with previous studies (Chen and Li 2001). Indeed, at total sites, the human-chimpanzee divergence was 1.02%, while the corresponding values between human and gorilla and between chimpanzee and gorilla were 1.05% and 1.25%, respectively. Similarly, at synonymous sites, the divergence between human and chimpanzee was 1.30%, lower than the human-gorilla divergence (1.77%) and the chimpanzee-gorilla divergence (2.00%).

Second, the human-gorilla divergence at synonymous sites was higher for the Y chromosome genes than for the X chromosome and autosomal genes, consistent with rapid evolution of the Y chromosome due to male mutation bias (Makova and Li 2002). Whereas the human-gorilla synonymous divergence for XDY genes was 1.77%, the corresponding values for X chromosome genes and

Table 1 The pairwise nucleotide divergence per 100 sites at synonymous and total sites for 16 X-degenerate Y chromosome genes in greater apes

Sites	Human vs. chimpanzee	Human vs. gorilla	Chimpanzee vs. gorilla
Total	1.015 ± 0.041	1.049 ± 0.069	1.245 ± 0.059
Synonymous	1.296 ± 0.110	1.767 ± 0.129	2.003 ± 0.146

Note: A total of 9,928 synonymous and 29,784 total sites were analyzed

autosomal genes were 0.61% (calculated here from the available 3,399 bp of human and gorilla X chromosome orthologues) and 1.11% (Shi et al. 2003), respectively. The male-to-female mutation rate ratio (α) was not estimated here due to the small number of sites analyzed.

Phylogenetic Analysis

An initial phylogenetic analysis of human, chimpanzee, and gorilla XDY genes, using their human X homologues as an outgroup and utilizing the neighbor-joining method (Saitou and Nei 1987), led to the following results. The concatenated sequences of genes supported the human-chimpanzee clade with 100% bootstrap support value (data not shown). Next, we divided human, chimpanzee, and gorilla XDY genes into two groups of orthologues (“orthogroups”) based on gene impairment in chimpanzee—disrupted orthogroup (*CYorf15B*, *TBL1Y*, *TMSB4Y*, and *USP9Y*) and nondisrupted orthogroup (the other 12 XDY genes). The phylogenetic analysis of concatenated gene sequences within each orthogroup, using human X homologues as an outgroup, again supported the human-chimpanzee clustering with high bootstrap values (93% and 96% for the disrupted and nondisrupted orthogroups, respectively; data not shown). When we built the phylogenies for each XDY gene individually (again using the human X homologue as an outgroup), 10 of 16 genes (*CYorf15B*, *EIF1AY*, *PRKY*, *RPS4Y1*, *RPS4Y2*, *SRY*, *TBL1Y*, *USP9Y*, *UTY*, and *ZFY*) favored the human-chimpanzee clade, while 4 genes (*AMELY*, *CYorf15A*, *NLGN4Y*, and *TMSB4Y*) and two genes (*DDX3Y* and *JARID1D*) attested to the human-gorilla and gorilla-chimpanzee clades, respectively (data not shown). The non-chimpanzee-human clades were usually supported by low bootstrap values (<70% for four of the genes) and thus are likely due to a limited number of sites examined.

Based on these results and to rescue the maximum number of informative sites (that sometimes correspond to gaps in alignments with X homologues), we rebuilt the phylogenetic trees in each case, utilizing gorilla XDY genes (and not the human X homologous sequences) as an

outgroup for the comparison between human and chimpanzee genes. This phylogenetic tree was particularly valid for the subsequent analyses, since the majority of them utilized the concatenated gene sequences that have a sufficient number of informative sites to support the human-chimpanzee clade. When individual genes were analyzed, we still conducted the three-way lineage-specific analyses even for the six XDY genes contradicting the human-chimpanzee grouping (although with low bootstrap values), but we exercised caution during interpretation of the results.

Lineage-Specific Nucleotide Divergence at XDY Genes

Since 4 of 16 XDY genes are disrupted in chimpanzee but not in human, we initially contrasted the chimpanzee and human lineage-specific divergences and, in several comparisons, found the former to be higher than the latter (Table 2). When the sequences of all 16 XDY genes were concatenated and gorilla was used as an outgroup, the chimpanzee divergence was elevated compared with the human divergence at either total (0.605% vs. 0.410%) or synonymous (0.784% vs. 0.536%) sites. Tajima’s (1993) relative rate test was highly significant for the difference between human and chimpanzee lineage-specific divergences at total sites ($p < 0.001$) and marginally significant for the corresponding difference at synonymous sites ($p = 0.056$). Similar results were obtained when only the eight genes supporting human-chimpanzee grouping were considered (data not shown).

What can account for the higher divergence of XDY genes in the chimpanzee compared with the human lineage? We hypothesized that, due to relaxed selection, disrupted chimpanzee XDY genes (concatenated) might have higher divergence from their human orthologues than undisrupted chimpanzee XDY genes. In agreement with this expectation, the accelerated accumulation of substitutions in chimpanzee compared with human XDY genes was more pronounced in the disrupted than the nondisrupted orthogroup (Table 2). Indeed, Tajima’s test comparing human- and chimpanzee-specific rates was significant or marginally significant for the disrupted orthogroup

Table 2 Human and chimpanzee lineage-specific divergences per 100 sites for 16 X-degenerate Y chromosome genes

Genes	Total sites			Synonymous sites		
	No. sites	Human	Chimpanzee	No. sites	Human	Chimpanzee
16 genes	29,784	0.410	0.605	8513.4	0.536	0.784
Nondisrupted orthogroup	19,899	0.427	0.526	5693.1	0.582	0.752
Disrupted orthogroup	9,885	0.376	0.766	2688.5	0.463	0.896
Upstream region	6,321	0.381	0.863	4636.8	0.495	0.932
Downstream region	3,564	0.366	0.596	2551.1	0.405	0.832

(0.376% vs. 0.766%, $p < 0.001$, for total sites; 0.463% vs. 0.896%, $p = 0.064$, for synonymous sites), while it was not significant for the nondisrupted orthogroup (0.427% vs. 0.526%, $p = 0.169$, for total sites; 0.582% vs. 0.752%, $p = 0.300$, for synonymous sites). Moreover, compared within the same species, the chimpanzee-specific divergence was higher for the disrupted than the nondisrupted genes at total sites (0.766% vs. 0.526%; $p = 0.012$, Fisher's exact test) as well as at synonymous sites (0.896% vs. 0.752%; $p = 0.513$), again consistent with relaxed selection at disrupted genes. In contrast, the human-specific divergence was lower in the disrupted than the nondisrupted orthogroup at both total sites (0.376% vs. 0.427%; $p = 0.568$) and synonymous sites (0.463% vs. 0.582%; $p = 0.523$), although the difference was not significant.

To examine whether functional constraints were different between the disrupted and the nondisrupted orthogroups in human and gorilla lineages, we counted the number of synonymous (S) and nonsynonymous (N) substitutions between human and gorilla for each of these two

orthogroups (Table 3). The proportion of nonsynonymous substitutions to synonymous substitutions was significantly higher for the disrupted than the nondisrupted orthogroup (2.000 vs. 0.835; $p = 0.0006$, Fisher's exact test). By enumerating all possible pairs of 4 of 16 genes (i.e., a Monte Carlo simulation), we obtained a frequency distribution of the nonsynonymous-to-synonymous substitution ratio (N/S ratio; Fig. 1). The N/S ratio for the disruptive orthogroup (2.000) was higher than the mean of the distribution (1.150) and significantly unlikely to occur by chance alone ($p = 0.023$, one-tailed test; Fig. 1). Thus, the four genes disrupted in chimpanzee evolve under relaxed functional constraints not only in this species, where they are disrupted, but also in human and gorilla.

Selection Tests

Intrigued by the high divergence of the XDY genes in the chimpanzee lineage, we investigated whether this observation could be explained by positive selection. The chimpanzee and human lineage-specific nonsynonymous and synonymous substitution rates (K_A and K_S , respectively) and their ratio (K_A/K_S) were estimated for the XDY genes (Table 4). Notably, the K_A/K_S ratios for the concatenated XDY genes in both human (0.686) and chimpanzee (0.697) lineages were at least twice as high as the genome-wide mean K_A/K_S ratios for genes located outside of chromosome Y (0.259 and 0.245 for human and chimpanzee, respectively [Bakewell et al. 2007]). High K_A/K_S ratios for the XDY genes imply an accumulation of deleterious mutations on the Y chromosome due to its small effective population size and lack of recombination

Table 3 Numbers of synonymous and nonsynonymous substitutions in the human-gorilla comparison for disrupted and nondisrupted orthogroups

Orthogroup	Nonsynonymous sites	Synonymous sites
Disrupted	66	33
Nondisrupted	96	115

Note: $p = 0.0006$; Fisher's exact test. The total numbers of nonsynonymous and synonymous sites are XXXX and XXXX bp for the disrupted orthogroup and XXXX and XXXX bp for the nondisrupted orthogroup, respectively. The total number of sites is 29,820 bp

Fig. 1 The simulated frequency distribution for the ratio of the numbers of nonsynonymous-to-synonymous substitutions (N/S ratio) for four human-gorilla orthologous gene pairs randomly picked up from 16 such gene pairs. This distribution was generated by a Monte Carlo simulation with 10,000 replicates. The observed N/S ratio for the disrupted orthogroup is indicated by the arrow

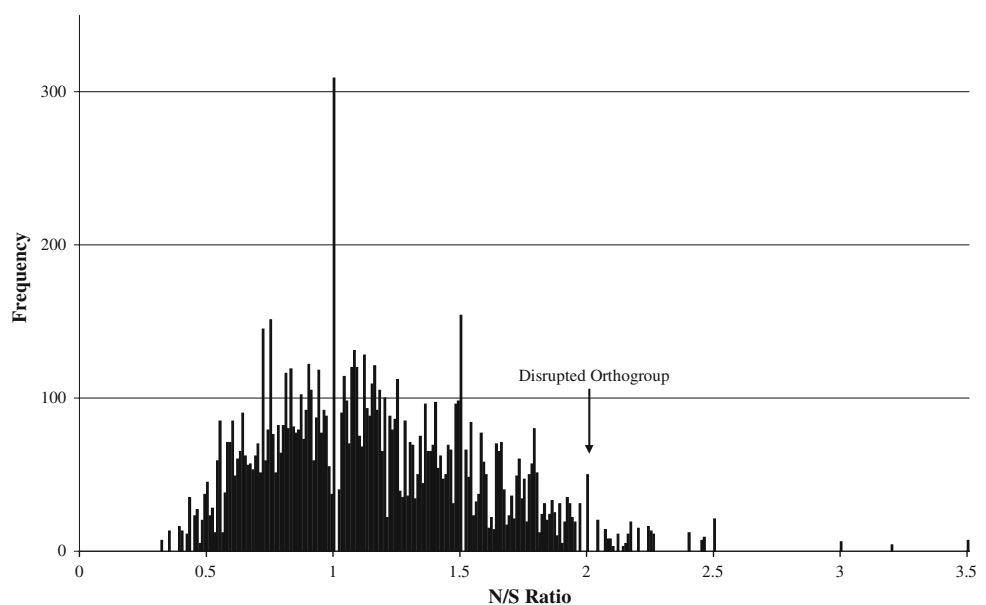


Table 4 Nonsynonymous (K_A) and synonymous (K_S) rates and their ratio (K_A/K_S) for 16 X-degenerate Y chromosome genes in the human and chimpanzee lineages

Gene(s)	No. of sites		Human			Chimpanzee		
	Nonsyn.	Syn.	K_A	K_S	K_A/K_S	K_A	K_S	K_A/K_S
<i>AMELY</i>	398.2	177.8	0.508	0.573	0.884	0.507	0	NA
<i>CYorf15A</i>	268.4	124.2	1.144	0	NA	0.755	1.604	0.471
<i>CYorf15B</i> ^a	386.3	153.7	0.521	0	NA	0.256	1.328	0.193
<i>DDX3Y</i>	1,505.1	474.9	0.200	0.217	0.923	0.603	0.438	1.377
<i>EIF1AY</i>	303.5	128.5	0	0.776	0	0.328	0	NA
<i>JARID1D</i>	3,224.7	1,380.3	0.127	0.811	0.156	0.253	0.284	0.891
<i>NLGN4Y</i>	1,820.9	627.1	0.279	0.817	0.341	0.565	2.825	0.200
<i>PRKY</i>	663.8	167.2	1.228	2.952	0.416	1.522	1.302	1.169
<i>RPS4Y1</i>	565.0	224.0	0	0.446	0	0.181	1.830	0.099
<i>RPS4Y2</i>	564.4	224.6	1.105	0	NA	0.176	0.446	0.394
<i>SRY</i>	471.8	140.2	0.422	0	NA	1.062	0.772	1.375
<i>TBL1Y</i> ^a	1,105.5	451.5	0.461	1.158	0.398	0.741	2.106	0.352
<i>TMSB4Y</i> ^a	100.1	31.9	1.002	0	NA	1.018	3.206	0.318
<i>USP9Y</i> ^a	5,703.0	1,953.0	0.301	0.371	0.811	0.746	0.589	1.266
<i>UTY</i>	3,010.1	1,030.9	0.541	0.601	0.899	0.337	0.603	0.559
<i>ZFY</i>	1,758.3	644.7	0.172	0.323	0.532	0.172	0.479	0.359
16 genes (concatenated)	21,270.6	8,513.4	0.367	0.536	0.686	0.547	0.784	0.697
Nondisrupted orthogroup	14,205.9	5,693.1	0.372	0.582	0.638	0.445	0.752	0.592
Disrupted orthogroup	7,196.5	2,688.5	0.350	0.463	0.757	0.733	0.896	0.818
Upstream region	1,684.2	4,636.8	0.348	0.495	0.702	0.854	0.932	0.916
Downstream region	1,012.9	2,551.1	0.356	0.405	0.881	0.517	0.832	0.621

Note: NA, not available (division by zero). K_A and K_S per 100 sites

^a Genes belonging to the disrupted orthogroup

(Charlesworth and Charlesworth 2000). Additionally, positive selection might contribute to the elevated K_A/K_S ratio on the Y chromosome (Gerrard and Filatov 2005).

Several ML ratio tests (implemented in PAML; Yang 1997) were employed to examine the K_A/K_S ratio for the XDY genes in the chimpanzee vs. the other (human and gorilla) lineages and to evaluate whether selection affected the evolution of these genes. First, we compared the K_A/K_S ratios in the chimpanzee vs. background (human and gorilla) lineages (Supplementary Table S3). After applying the Bonferroni correction for multiple tests, we found that the chimpanzee-specific K_A/K_S ratio was not significantly different from the background ratio for each of the 16 XDY genes examined. Second, we applied the improved branch-site model (Zhang et al. 2005), a test optimized for detection of positive selection, to assess whether positive selection acted on the XDY genes in chimpanzee and human. After applying the Bonferroni correction, we could not detect any significant indication of positive selection in the either chimpanzee or the human lineage (Supplementary Table S4). Even for four XDY genes (*DDX3Y*, *PRKY*, *SRY*, and *USP9Y*) for which the chimpanzee

lineage-specific K_A/K_S ratio was >1 (Table 4), potentially indicating positive selection, none of them was significantly different from 1 after the Bonferroni correction for multiple tests (Supplementary Table S5). Third, we tested whether, for the three disrupted genes (*CYorf15B*, *TMSB4Y*, and *TBL1Y*) in chimpanzee with lineage-specific K_A/K_S ratios <1 (Table 4), these ratios were significantly <1 , suggestive of purifying selection. However, after applying the Bonferroni correction for multiple tests, the K_A/K_S ratio for each of these genes was not significantly different from 1 (Supplementary Table S5). The selection tests performed above might lack statistical power due to a small number of sites mutated between closely related species (Shi et al. 2003). Note that in the K_A/K_S ratio analysis of individual genes, for genes with the human-gorilla clade (*AMELY*, *CYorf15A*, *NLGN4Y*, and *TMSB4Y*), the chimpanzee-specific values include both chimpanzee lineage-specific substitutions and substitutions supporting the human-gorilla grouping. The analysis of *DDX3Y* and *JARID1D*, supporting the gorilla-chimpanzee grouping, has a similar limitation. This does not change the conclusions of these analyses, though.

The Effect of Disruptive Mutations on the K_A/K_S Ratios

To investigate whether the disruptive mutations led to two distinct selective regimes at chimpanzee XDY genes, each gene with one or more such mutations (*CYorf15B*, *TBL1Y*, *TMSB4Y*, and *USP9Y*) was broken into two regions—located upstream and downstream of the first disruptive mutation, respectively. By utilizing a sequence of the third species (gorilla), we compared the human and chimpanzee lineage-specific K_A/K_S ratios between the concatenated upstream vs. downstream regions of the four disrupted genes. If the upstream regions still encode functional proteins in chimpanzee, such regions are expected to evolve under stronger purifying selection compared with the downstream regions, and thus the K_A/K_S ratio should be lower in the former than the latter regions. In contrast with this expectation, the chimpanzee-specific K_A/K_S ratio for the concatenated upstream regions was close to 1 (0.916; Table 4) and higher than that for downstream regions (0.621), although the difference was not significant ($p = 0.423$). However, in agreement with the chimpanzee-specific gene decay, the K_A/K_S ratio of the upstream regions was higher in chimpanzee than in human (0.916 vs. 0.702; $p = 0.791$).

We next examined each of the four disrupted genes separately (Table 5). Two of them (*CYorf15B* and *TMSB4Y*) were too short for any meaningful statistical comparisons to be made. For *TBL1Y*, the K_A/K_S ratio in the chimpanzee lineage was similar between upstream and downstream regions (0.327 vs. 0.387). For *USP9Y*, the K_A/K_S ratio in the chimpanzee lineage was higher in the downstream than the upstream region (1.490 vs. 1.226), although the difference was not significant and the K_A/K_S ratios were not significantly >1 according to a ML test (data not shown). Interestingly, for this gene, the K_A/K_S ratio was higher in the chimpanzee than the human lineage

for both upstream (1.226 vs. 0.805) and downstream (0.842 vs. 1.490) regions, although this was not significant (data not shown).

Discussion

Conservation of the XDY Gene Content in Human and Gorilla, But Not in Chimpanzee

Here we demonstrated that in gorilla, just as in human, the coding exons and splice sites for 16 XDY genes remain intact. In contrast, in chimpanzee, 4 of 16 XDY genes have been disrupted by inactivating mutations (Hughes et al. 2005; Kuroki et al. 2006; Perry et al. 2007). In addition to these 16 genes, the XDY region of both human and chimpanzee harbors 11 pseudogenes (Skaletsky et al. 2003; Hughes et al. 2005; Kuroki et al. 2006), with most of the pseudogenizing mutations shared between the two species. Therefore, it is plausible that these 11 genes became pseudogenes before the human-chimpanzee split (~6 MYA [Glazko and Nei 2003; Hughes et al. 2005; Kuroki et al. 2006]) and presumably even before the divergence of the gorilla from the human-chimpanzee lineage (~7 MYA; Glazko and Nei 2003]). Thus, the XDY gene content is likely to be conserved between human and gorilla (each species has 16 XDY genes), but is drastically different in chimpanzee, which possesses only 12 XDY genes. This conclusion will have to be re-evaluated by sequencing of the homologues of 11 human pseudogenes in gorilla and, ideally, by sequencing of the complete gorilla Y chromosome. The translocation of genes from autosomes to the Y chromosome as well as lineage-specific preservation of X-degenerate genes has been noted in some species (Murphy et al. 2006), and this cannot be excluded for gorilla. However, here we aimed to trace the

Table 5 Human and chimpanzee lineage-specific substitutions in the upstream and downstream regions of the disrupted genes

Gene	Lineage	Upstream of disruptive mutation					Downstream of disruptive mutation				
		Nonsynonymous		Synonymous		K_A/K_S	Nonsynonymous		Synonymous		K_A/K_S
		Sites	Diffs ^a	Sites	Diffs		Sites	Diffs	Sites	Diffs	
<i>CYorf15B</i>	Human	127.1	0	13.9	0	N.A.	282.4	2	116.6	0	N.A.
	Chimpanzee	127.1	0	13.9	0	N.A.	282.4	1	116.6	2.1	0.200
<i>TBL1Y</i>	Human	632.5	4.1	255.5	3.1	0.524	475.4	1	193.6	2	0.204
	Chimpanzee	632.5	5.1	255.5	6.3	0.327	475.4	3.1	193.6	3.3	0.387
<i>TMSB4Y</i>	Human	75.4	0	23.6	0	N.A.	24.2	1.1	8.8	0	N.A.
	Chimpanzee	75.4	1	23.6	0	N.A.	24.2	0	8.8	1.1	0
<i>USP9Y</i>	Human	3862.6	12.1	1330.4	5.2	0.805	1833.2	5	629.8	2.1	0.842
	Chimpanzee	3862.6	33.5	1330.4	9.4	1.226	1833.2	9.1	629.8	2.1	1.490

^a Number of differences

evolutionary fates of known primate XDY genes (homologous to human) in gorilla and not to assemble the complete list of gorilla Y chromosome genes.

We observed that the chimpanzee lineage-specific K_A/K_S ratio was <1 for three of four disrupted genes, although none of these ratios was significantly different from 1. Thus, our results do not indicate that these genes evolved under purifying selection, in contrast to what was shown for some mouse retrogenes (Gayral et al. 2007). Moreover, according to the K_A/K_S tests, we could not detect significant differences in selective pressures between upstream and downstream regions of the four disrupted chimpanzee genes. This calls into question the possibility that the upstream regions of these genes encode functional proteins in chimpanzee. Hughes et al. (2005) suggested that *TMSB4Y* and *USP9Y* lost their protein functions completely (the former gene is not expressed and the latter gene possesses a defect in its catalytic domain in chimpanzee). Our results provide no support for the functionality of any of the four disrupted genes in chimpanzee. However, despite the high divergence in the disrupted orthogroup in the chimpanzee lineage shown here, the mRNA of some of these genes might still have a regulatory role (Perry et al. 2007).

The Mechanisms of Primate Y Chromosome Evolution

Our results allow us to speculate about the evolutionary mechanisms behind ape Y chromosome degeneration by contrasting two groups of models: the relaxed purifying selection models (see Introduction) vs. the genetic hitchhiking model (Charlesworth and Charlesworth 2000). First, the conserved gene content between human and gorilla and the different gene content in chimpanzee contradict the relaxed purifying selection models. These models are expected to be more efficient in species with a small effective population size (Charlesworth and Charlesworth 2000). According to the models, the human lineage, due to the smaller effective population size in human (Kaessmann et al. 1999; Yu et al. 2004), is expected to accumulate more mutations (including disruptive mutations) compared with the chimpanzee and gorilla lineages. However, this goes against our finding of high divergence in the chimpanzee lineage and an earlier report of disruptive mutations in this lineage (Hughes et al. 2005).

Second, higher divergence per million years (Table 6) and more disrupted genes in the chimpanzee lineage than in the other analyzed lineages are consistent with genetic hitchhiking driven by sperm competition. Indeed, genetic hitchhiking anticipates more rapid accumulation of mutations in chimpanzee, a polyandrous species, than in human or gorilla, monoandrous species, assuming that positive selection due to sperm competition is a major driving force. However, surprisingly, we found no evidence of lineage-

Table 6 A summary of parameters important for the evolutionary models discussed

	Human	Chimpanzee	Gorilla
Effective population size	~10,000 ^a	~25,000 ^a	~25,000 ^a
Sperm competition	Weak	Strong	Weak
Nucleotide divergence in a lineage	0.415 ^b	0.586 ^b	0.485 ^b
Divergence time (MY)	6 ^c	6 ^c	7 ^c
Nucleotide divergence per MY	0.069 ^d	0.098 ^d	0.069 ^d

^a From Kaessmann et al. (1999); Yu et al. (2004), and Thalmann et al. (2007)

^b The nucleotide divergence per 100 sites was calculated (according to the TN93 model) for 16 XDY genes (concatenated) using the sequences of the homologous human X chromosome genes as an outgroup. Only human, chimpanzee, and gorilla lineage-specific values are presented here

^c From Glazko and Nei (2003)

^d Nucleotide divergence per 100 sites

specific positive selection acting on any of the 16 XDY genes in chimpanzee and human. This might be explained by the low power of statistical tests due to a recent divergence among the studied species or by positive selection directed at genes outside of the XDY region, e.g., in the ampliconic region, as suggested by Hughes et al. (2005). In support of the latter hypothesis, all human ampliconic genes are expressed only in testis (Skaletsky et al. 2003) and thus are plausible targets for sperm competition. It will be of great interest to further evaluate the models of relaxed purifying selection due to genetic drift vs. that of genetic hitchhiking due to sperm competition by analyzing other primate species.

Essential Genes on Primate and Mammalian Y Chromosomes

The data on the XDY genes in yet another species, gorilla, provide an opportunity to narrow down the list of essential Y chromosome genes for primates and, more generally, for eutherian mammals. Four XDY genes (*CYorf15B*, *TBL1Y*, *TMSB4Y*, and *USP9Y*) are disrupted in chimpanzee and, according to our results, evolve under relatively low functional constraints in human and gorilla. *USP9Y* is also disrupted in spider monkey (Gerrard and Filatov 2005). Thus, the other 12 genes (*AMELY*, *CYorf15A*, *DDX3Y*, *EIF1AY*, *JARID1D*, *NLGN4Y*, *PRKY*, *RPS4Y1*, *RPS4Y2*, *SRY*, *UTY*, and *ZFY*) are the candidates for being indispensable on primate Y chromosomes. Of these 12 genes, 7 (*AMELY*, *DDX3Y*, *EIF1AY*, *JARID1D*, *SRY*, *UTY*, and *ZFY*) are also found on cat Y chromosome (Murphy et al. 2006). *AMELY* is known to be absent in mouse (Lahn et al. 2001). Therefore, the remaining six genes might be essential for eutherian Y chromosomes. Future research

including sequencing of multiple complete mammalian Y chromosomes will be required to assess whether any genes are indispensable on the Y in all eutherian mammals. Note that even the sex-determining gene, *SRY*, is known to be lost from two rodent groups, *Ellobius* and *Tokudaia*, and thus might not be essential for eutherian males but might be required for maintaining the Y chromosome, as the two rodent groups mentioned above lack not only *SRY*, but also the Y chromosome altogether (Graves 2006).

Interestingly, some of the genes disrupted in chimpanzee are known to be important for spermatogenesis in human. For instance, mutations in *USP9Y* result in lack of sperm in semen or in azoospermia (Sun et al. 1999; Blagosklonova et al. 2000). This indicates that particular genes might be essential for some primate species but not others. It is likely that the Y chromosome encodes species-specific reproductive and potentially adaptive traits.

Acknowledgments We are grateful to Chungoo Park, Saby Das, Allison N. Lau, Masafumi Nozawa, and Melissa Wilson for valuable suggestions. We also thank M. Paula Goetting-Minesky for experimental support at the early stages of this project. Meg Bakewell kindly provided the control files for PAML. One of the male gorilla DNA samples was provided by Robert Baker from the Natural Sciences Research Laboratory at Texas Tech University. This research was supported by startup funds from the Eberly College of Sciences at Pennsylvania State University to K.D.M.

References

- Bakewell MA, Shi P, Zhang J (2007) More genes underwent positive selection in chimpanzee evolution than in human evolution. *Proc Natl Acad Sci USA* 104:7489–7494
- Blagosklonova O, Fellmann F, Clavequin MC, Roux C, Bresson JL (2000) AZFa deletions in Sertoli cell-only syndrome: a retrospective study. *Mol Hum Reprod* 6:795–799
- Charlesworth B, Charlesworth D (2000) The degeneration of Y chromosomes. *Philos Trans R Soc Lond B Biol Sci* 355:1563–1572
- Chen FC, Li WH (2001) Genomic divergences between humans and other hominoids and the effective population size of the common ancestor of humans and chimpanzees. *Am J Hum Genet* 68:444–456
- Dixson AF (1998) Primate sexuality: comparative studies of the prosimians, monkeys, apes, and human beings. Oxford University Press, New York
- Fischer A, Wiebe V, Pääbo S, Przeworski M (2004) Evidence for a complex demographic history of chimpanzees. *Mol Biol Evol* 21:799–808
- Fischer A, Pollack J, Thalmann O, Nickel B, Pääbo S (2006) Demographic history and genetic differentiation in apes. *Curr Biol* 16:1133–1138
- Gayral P, Caminade P, Boursot P, Galtier N (2007) The evolutionary fate of recently duplicated retrogenes in mice. *J Evol Biol* 20:617–626
- Gerrard DT, Filatov DA (2005) Positive and negative selection on mammalian Y chromosomes. *Mol Biol Evol* 22:1423–1432
- Glazko GV, Nei M (2003) Estimation of divergence times for major lineages of primate species. *Mol Biol Evol* 20:424–434
- Goldman N, Yang Z (1994) A codon-based model of nucleotide substitution for protein-coding DNA sequences. *Mol Biol Evol* 11:725–736
- Goodall J (1986) The chimpanzees of Gombe: patterns of behavior. Harvard University Press, Cambridge, MA
- Graves JA (2006) Sex chromosome specialization and degeneration in mammals. *Cell* 124:901–914
- Greenfield A, Carrel L, Pennisi D, Philippe C, Quaderi N, Siggers P, Steiner K, Tam PP, Monaco AP, Willard HF, Koopman P (1998) The UTX gene escapes X inactivation in mice and humans. *Hum Mol Genet* 7:737–742
- Hasegawa T, Hiraiwai-Hasegawa M (1990) Sperm competition and mating behavior. In: Nishida T (ed) The chimpanzees of the Mahale mountains: sexual and life history strategies. University of Tokyo Press, Tokyo, pp 115–132
- Hinrichs AS, Karolchik D, Baertsch R, Barber GP, Bejerano G, Clawson H, Diekhans M, Furey TS, Harte RA, Hsu F, Hillman-Jackson J, Kuhn RM, Pedersen JS, Pohl A, Raney BJ, Rosenbloom KR, Siepel A, Smith KE, Sugnet CW, Sultan-Qurraie A, Thomas DJ, Trumbower H, Weber RJ, Weirauch M, Zweig AS, Haussler D, Kent WJ (2006) The UCSC Genome Browser Database: update 2006. *Nucleic Acids Res* 34:D590–D598
- Hughes JF, Skaletsky H, Pyntikova T, Minx PJ, Graves T, Rozen S, Wilson RK, Page DC (2005) Conservation of Y-linked genes during human evolution revealed by comparative sequencing in chimpanzee. *Nature* 437:100–103
- Kaessmann H, Wiebe V, Pääbo S (1999) Extensive nuclear DNA sequence diversity among chimpanzees. *Science* 286:1159–1162
- Kuroki Y, Toyoda A, Noguchi H, Taylor TD, Itoh T, Kim DS, Kim DW, Choi SH, Kim IC, Choi HH, Kim YS, Satta Y, Saitou N, Yamada T, Morishita S, Hattori M, Sakaki Y, Park HS, Fujiyama A (2006) Comparative analysis of chimpanzee and human Y chromosomes unveils complex evolutionary pathway. *Nat Genet* 38:158–167
- Lahn BT, Page DC (1999) Four evolutionary strata on the human X chromosome. *Science* 286:964–967
- Lahn BT, Pearson NM, Jegalian K (2001) The human Y chromosome, in the light of evolution. *Nat Rev Genet* 2:207–216
- Makova KD, Li WH (2002) Strong male-driven evolution of DNA sequences in humans and apes. *Nature* 416:624–626
- Maynard Smith J, Haigh J (1974) The hitch-hiking effect of a favourable gene. *Genet Res* 23:23–35
- Muller HJ (1914) A gene for the fourth chromosome of *Drosophila*. *J Exp Zool* 17:325–336
- Murphy WJ, Pearks Wilkerson AJ, Raudsepp T, Agarwala R, Schäffer AA, Stanyon R, Chowdhary BP (2006) Novel gene acquisition on carnivore Y chromosomes. *PLoS Genet* 2:e43
- Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol* 3:418–426
- Ohno S (1967) Monographs on endocrinology. Sex chromosomes and sex-linked genes, 1st edn. Springer, Heidelberg
- Ono S (2003) Regulation of actin filament dynamics by actin depolymerizing factor/cofilin and actin-interacting protein 1: new blades for twisted filaments. *Biochemistry* 42:13363–13370
- Perry GH, Tito RY, Verrelli BC (2007) The evolutionary history of human and chimpanzee Y-chromosome gene loss. *Mol Biol Evol* 24:853–859
- Potrzebowski L, Vinckenbosch N, Marques AC, Chalmel F, Jégou B, Kaessmann H (2008) Chromosomal gene movements reflect the recent origin and biology of therian sex chromosomes. *PLoS Biol* 6:e80
- Ross MT et al (2005) The DNA sequence of the human X chromosome. *Nature* 434:325–337
- Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* 132:365–386

- Rozen S, Skaletsky H, Marszalek JD, Minx PJ, Cordum HS, Waterston RH, Wilson RK, Page DC (2003) Abundant gene conversion between arms of palindromes in human and ape Y chromosomes. *Nature* 423:873–876
- Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425
- Shi J, Xi H, Wang Y, Zhang C, Jiang Z, Zhang K, Shen Y, Jin L, Zhang K, Yuan W, Wang Y, Lin J, Hua Q, Wang F, Xu S, Ren S, Xu S, Zhao G, Chen Z, Jin L, Huang W (2003) Divergence of the genes on human chromosome 21 between human and other hominoids and variation of substitution rates among transcription units. *Proc Natl Acad Sci USA* 100:8331–8336
- Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, Repping S, Pyntikova T, Ali J, Bieri T, Chinwalla A, Delehaunty A, Delehaunty K, Du H, Fewell G, Fulton L, Fulton R, Graves T, Hou SF, Latrielle P, Leonard S, Mardis E, Maupin R, McPherson J, Miner T, Nash W, Nguyen C, Ozersky P, Pepin K, Rock S, Rohlfling T, Scott K, Schultz B, Strong C, Tin-Wollam A, Yang SP, Waterston RH, Wilson RK, Rozen S, Page DC (2003) The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* 423:825–837
- Smith TF, Gaitatzes C, Saxena K, Neer EJ (1999) The WD repeat: a common architecture for diverse functions. *Trends Biochem Sci* 24:181–185
- Stone AC, Griffiths RC, Zegura SL, Hammer MF (2002) High levels of Y-chromosome nucleotide diversity in the genus *Pan*. *Proc Natl Acad Sci USA* 99:43–48
- Sun C, Skaletsky H, Birren B, Devon K, Tang Z, Silber S, Oates R, Page DC (1999) An azoospermic man with a de novo point mutation in the Y-chromosomal gene *USP9Y*. *Nat Genet* 23:429–432
- Tajima F (1993) Simple methods for testing the molecular evolutionary clock hypothesis. *Genetics* 135:599–607
- Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* 10:512–526
- Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596–1599
- Thalmann O, Fischer A, Lankester F, Pääbo S, Vigilant L (2007) The complex evolutionary history of gorillas: insights from genomic data. *Mol Biol Evol* 24:146–158
- Veyrunes F, Waters PD, Miethke P, Rens W, McMillan D, Alsop AE, Grützner F, Deakin JE, Whittington CM, Schatzkammer K, Kremitzki CL, Graves T, Ferguson-Smith MA, Warren W, Marshall Graves JA (2008) Bird-like sex chromosomes of platypus imply recent origin of mammal sex chromosomes. *Genome Res* 18:965–973
- Yang Z (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci* 13:555–556
- Yang Z (1998) Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. *Mol Biol Evol* 15:568–573
- Yu N, Jensen-Seaman MI, Chemnick L, Kidd JR, Deinard AS, Ryder O, Kidd KK, Li WH (2003) Low nucleotide diversity in chimpanzees and bonobos. *Genetics* 164:1511–1518
- Yu N, Jensen-Seaman MI, Chemnick L, Ryder O, Li WH (2004) Nucleotide diversity in gorillas. *Genetics* 166:1375–1383
- Zhang J, Nielsen R, Yang Z (2005) Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol Biol Evol* 22:2472–2479