

## LETTERS

# Chimpanzee and human Y chromosomes are remarkably divergent in structure and gene content

Jennifer F. Hughes<sup>1</sup>, Helen Skaletsky<sup>1</sup>, Tatyana Pyntikova<sup>1</sup>, Tina A. Graves<sup>2</sup>, Saskia K. M. van Daalen<sup>3</sup>, Patrick J. Minx<sup>2</sup>, Robert S. Fulton<sup>2</sup>, Sean D. McGrath<sup>2</sup>, Devin P. Locke<sup>2</sup>, Cynthia Friedman<sup>4</sup>, Barbara J. Trask<sup>4</sup>, Elaine R. Mardis<sup>2</sup>, Wesley C. Warren<sup>2</sup>, Sjoerd Repping<sup>3</sup>, Steve Rozen<sup>1</sup>, Richard K. Wilson<sup>2</sup> & David C. Page<sup>1</sup>

The human Y chromosome began to evolve from an autosome hundreds of millions of years ago, acquiring a sex-determining function and undergoing a series of inversions that suppressed crossing over with the X chromosome<sup>1,2</sup>. Little is known about the recent evolution of the Y chromosome because only the human Y chromosome has been fully sequenced. Prevailing theories hold that Y chromosomes evolve by gene loss, the pace of which slows over time, eventually leading to a paucity of genes, and stasis<sup>3,4</sup>. These theories have been buttressed by partial sequence data from newly emergent plant and animal Y chromosomes<sup>5–8</sup>, but they have not been tested in older, highly evolved Y chromosomes such as that of humans. Here we finished sequencing of the male-specific region of the Y chromosome (MSY) in our closest living relative, the chimpanzee, achieving levels of accuracy and completion previously reached for the human MSY. By comparing the MSYs of the two species we show that they differ radically in sequence structure and gene content, indicating rapid evolution during the past 6 million years. The chimpanzee MSY contains twice as many massive palindromes as the human MSY, yet it has lost large fractions of the MSY protein-coding genes and gene families present in the last common ancestor. We suggest that the extraordinary divergence of the chimpanzee and human MSYs was driven by four synergistic factors: the prominent role of the MSY in sperm production, ‘genetic hitchhiking’ effects in the absence of meiotic crossing over, frequent ectopic recombination within the MSY, and species differences in mating behaviour. Although genetic decay may be the principal dynamic in the evolution of newly emergent Y chromosomes, wholesale renovation is the paramount theme in the continuing evolution of chimpanzee, human and perhaps other older MSYs.

This study required that we first complete the sequencing of the chimpanzee MSY<sup>9–11</sup> using large-insert bacterial artificial chromosome (BAC) clones and the iterative mapping and sequencing strategy used to comprehensively sequence the human MSY<sup>12,13</sup> (Supplementary Fig. 1 and Supplementary Note 1). This protracted approach is essential because much of the MSY consists of lengthy, highly similar repeat units, or ‘amplicons’, that cannot be distinguished by conventional mapping methods. BACs deriving from different copies of the same amplicon must be fully sequenced to identify subtle differences that distinguish one amplicon copy from another. These sequence variants are then used to sort other potentially overlapping BACs (Supplementary Fig. 2 and Supplementary File 1). By iterating this process, we assembled and analysed a tiling path of 219 BAC and 12 fosmid clones from the chimpanzee MSY

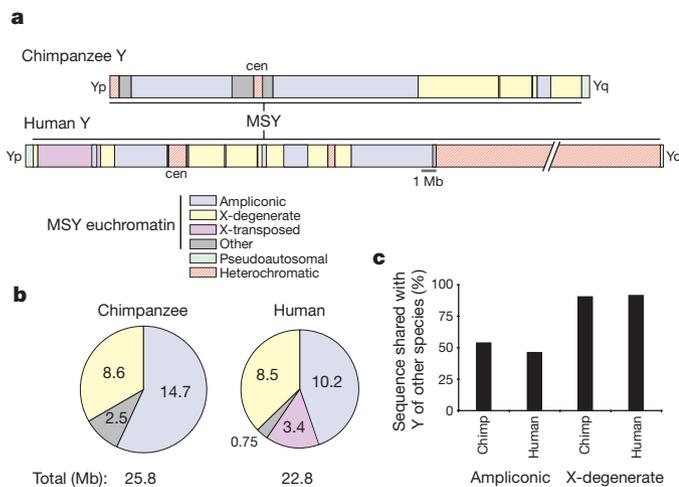
(Supplementary Fig. 1 and Supplementary Table 1). To avoid polymorphic differences between chimpanzee Y chromosomes that might confound identification of subtle sequence differences between amplicon copies<sup>12,13</sup>, we intended to sequence the Y chromosome of one male chimpanzee. In fact, all but 17 of the 230 sequenced BAC or fosmid clones were derived from one male.

The resulting euchromatic sequence comprises 25.8 megabases (Mb) in eight contigs, the largest of which spans 10.1 Mb (Supplementary Fig. 1, Supplementary Table 1 and Supplementary File 2). We ordered and oriented the contigs by metaphase fluorescence *in situ* hybridization (FISH) (Supplementary Figs 3 and 4), and confirmed amplicon copy numbers by interphase FISH (Supplementary Fig. 5). To test the completeness of our effort, we also analysed 14 Mb of shotgun sequencing reads from flow-sorted chimpanzee Y chromosomes. This independent sampling of the chromosome confirmed that our sequencing of the chimpanzee MSY was essentially complete (Supplementary Note 2). We estimate that the finished sequence has an error rate of about one nucleotide per Mb.

Our laboratories previously demonstrated that the human MSY euchromatin is largely comprised of two sequence classes: ampliconic and X-degenerate<sup>13</sup>. We find that the same two sequence classes dominate the chimpanzee MSY euchromatin (Fig. 1a, b), and thus the same was probably true in the common ancestor. The ampliconic segments are composed of large, nearly identical repeat units, most often arrayed as palindromes, and they contain multi-copy gene families expressed predominantly or exclusively in the testis<sup>13</sup> (Supplementary Figs 6 and 7). In contrast, the X-degenerate segments are dotted with single-copy homologues of X-linked genes. These single-copy MSY genes, most of which are expressed ubiquitously, are surviving relics of ancient autosomes from which the X and Y chromosomes evolved<sup>2</sup>. Together, the ampliconic and X-degenerate sequences comprise the bulk of the MSY euchromatin in both chimpanzee and human (Fig. 1b). A third sequence class in the human MSY euchromatin—the X-transposed sequences—has no counterpart in the chimpanzee MSY. The presence of these sequences in the human MSY is the result of an X-to-Y transposition that occurred in the human lineage after its divergence from the chimpanzee lineage<sup>14</sup>.

Given that primate sex chromosomes are hundreds of millions of years old<sup>2</sup>, theories of decelerating decay would predict that the chimpanzee and human MSYs should have changed little since the separation of these two lineages just 6 million years ago. To test this prediction, we aligned and compared the nucleotide sequences of the chimpanzee and human MSYs (Supplementary File 3). As expected,

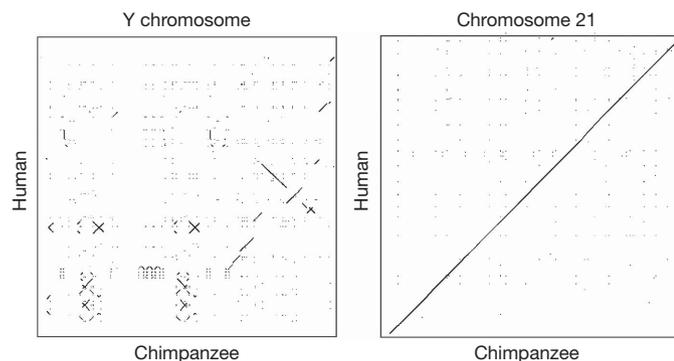
<sup>1</sup>Howard Hughes Medical Institute, Whitehead Institute, and Department of Biology, Massachusetts Institute of Technology, 9 Cambridge Center, Cambridge, Massachusetts 02142, USA. <sup>2</sup>The Genome Center, Washington University School of Medicine, 4444 Forest Park Boulevard, St Louis, Missouri 63108, USA. <sup>3</sup>Center for Reproductive Medicine, Department of Obstetrics and Gynecology, Academic Medical Center, Amsterdam 1105 AZ, The Netherlands. <sup>4</sup>Division of Human Biology, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North C3-168, Seattle, Washington 98109, USA.



**Figure 1 | Comparison of chimpanzee and human Y chromosomes.**

**a**, Schematic representations of chromosomes. cen, centromere; Yp, short arm; Yq, long arm. For both chromosomes, the MSY is indicated. Six sequence classes are shown, four of which are MSY euchromatin. ('Other' denotes MSY single-copy sequences that are not X-degenerate or X-transposed.) Chromosomes are drawn to scale, with the exception of the large heterochromatic block on human Yq. **b**, Sizes (in Mb) of four MSY euchromatin sequence classes in chimpanzee and human. **c**, Percentages of ampliconic and X-degenerate sequences present on chimpanzee Y chromosome that are also present on human Y chromosome, and vice versa.

we found that the degree of similarity between orthologous chimpanzee and human MSY sequences (98.3% nucleotide identity) differs only modestly from that reported when comparing the rest of the chimpanzee and human genomes (98.8%)<sup>15</sup>. Surprisingly, however, >30% of chimpanzee MSY sequence has no homologous, alignable counterpart in the human MSY, and vice versa (Supplementary Fig. 8 and Supplementary Note 3). In this respect, the MSY differs radically from the remainder of the genome, where <2% of chimpanzee euchromatic sequence lacks a homologous, alignable counterpart in humans, and vice versa<sup>15</sup>. We conclude that, since the separation of the chimpanzee and human lineages, sequence gain and loss have been far more concentrated in the MSY than in the balance of the genome. Moreover, the MSY sequences retained in both lineages have been extraordinarily subject to rearrangement: whole-chromosome dot-plot comparison of chimpanzee and human MSYs shows marked differences in gross structure (Fig. 2 and Supplementary Fig. 9), which contrasts starkly with chromosome 21, the only other chromosome comprehensively mapped and



**Figure 2 | Dot plots of DNA sequence identity between chimpanzee and human Y chromosomes and chromosome 21.** Each dot represents 100% chimpanzee–human identity within a 200-base-pair (bp) window. In the Y-chromosome plot, the human chromosome is oriented with short arm to top and long arm to bottom, and the chimpanzee chromosome is oriented with short arm to left and long arm to right. For chromosome 21, which is acrocentric, the plot represents only the long arm.

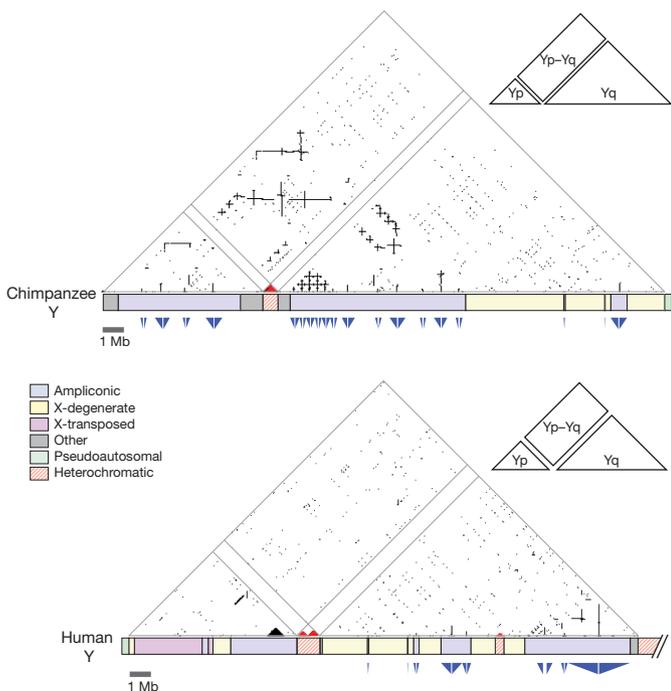
sequenced in both species<sup>16</sup>. Contrary to the decelerating decay theory, the chimpanzee and human MSYs differ markedly in sequence structure.

An interesting question is whether these evolutionary changes have involved the ampliconic and X-degenerate regions in equal measure. Previous models of Y-chromosome evolution treated the chromosome as a uniform, homogeneous substrate for evolutionary change<sup>1,3,4,17</sup>. In fact, the evolution of ampliconic sequences has outpaced that of X-degenerate sequences, and to such a degree that the ampliconic architecture of the common ancestor's MSY may be difficult to reconstruct even after an outgroup MSY has been sequenced. Comparing the chimpanzee and human MSY sequences, we find that the average length of uninterrupted, alignable segments in the ampliconic regions is only one-third of that in X-degenerate regions: 0.5 versus 1.5 Mb (Supplementary Fig. 10 and Supplementary Note 3). This reflects extensive rearrangement (Supplementary Figs 8 and 9) and rampant sequence gain and loss in the ampliconic regions. About half of the chimpanzee ampliconic sequence has no homologous, alignable counterpart in the human MSY, and vice versa, compared to <10% of the X-degenerate sequence (Fig. 1c).

The molecular mechanisms that enabled this wholesale remodelling of ampliconic regions merit consideration. Although the chimpanzee and human MSYs do not normally participate in meiotic exchange with a partner chromosome, the mirroring of sequences in the ampliconic regions provides ample opportunity for ectopic homologous recombination within the MSY. This recombinational proclivity is well documented in the human MSY, where it has repeatedly given rise to large-scale structural polymorphisms during the past 100,000 years of human history<sup>18</sup> as well as to Y-chromosomal anomalies that cause spermatogenic failure and sex reversal in current generations<sup>12,19–21</sup>. We suggest that ectopic homologous recombination between MSY amplicons has similarly accelerated structural remodelling of the MSY in the chimpanzee and human lineages during the past 6 million years.

The chimpanzee ampliconic regions are particularly massive (44% larger than in human; Fig. 1b) and architecturally ornate, with 19 palindromes (compared to eight in human) and elaborate mirroring of nucleotide sequences between the short and long arms of the chromosome, a feature not found in the human MSY (Fig. 3 and Supplementary Fig. 11). Of the 19 chimpanzee palindromes, only 7 are also found in the human MSY; the other 12 are chimpanzee-specific. Unlike the human MSY, nearly all of the chimpanzee MSY palindromes exist in multiple copies (Supplementary Fig. 11), so that each palindrome arm has potential partners for both intra- and inter-palindrome gene conversion (non-reciprocal transfer)<sup>22</sup>. This may help explain why arm-to-arm nucleotide sequence divergence in some chimpanzee MSY palindromes (as much as 0.5%; Supplementary Figs 12 and 13) is more pronounced than in human MSY palindromes (<0.06%)<sup>13</sup>.

Gene conversion may also account for the relatively low density of retrotransposable elements in ampliconic regions. In the chimpanzee and human MSYs, retrotransposon content is markedly lower in ampliconic than in X-degenerate regions—41% versus 63% in both species ( $P < 0.000001$ , Z-test; Supplementary Fig. 1i and Supplementary Table 2). Although it is possible that retrotransposons preferentially integrate in X-degenerate sequences, this seems unlikely given the similarity in C+G content and gene density in ampliconic and X-degenerate regions (Supplementary Table 2 and Supplementary Fig. 1h). An alternative explanation is that gene conversion between amplicon copies removes retrotransposons, especially recently integrated ones. Tellingly, an endogenous retrovirus that colonized the chimpanzee genome after the chimpanzee–human split<sup>23</sup> is present in 23 copies in the chimpanzee MSY, but only two of these copies are located in ampliconic regions (14.7 Mb), whereas 21 copies are located in X-degenerate regions (8.6 Mb;  $P < 0.000001$ , chi-square test; Supplementary Fig. 14). These findings offer



**Figure 3 | Triangular dot plots of DNA sequence identities within euchromatic MSY of chimpanzee and human.** Each dot represents 100% intra-chromosomal identity within a 200-bp window. Red dots represent matches between heterochromatic sequences. Direct repeats appear as horizontal lines, inverted repeats as vertical lines, and palindromes as vertical lines that nearly intersect the baseline. Insets indicate that each large triangular plot contains two smaller triangles (one revealing sequence identities within Yp and one revealing identities within Yq) and a rectangle (revealing sequence identities between Yp and Yq). Immediately below the plots are schematic representations of chromosomes. Triangles below chromosome schematics denote sizes and locations of palindromes. Gaps between opposed triangles represent the non-duplicated spacers between palindrome arms.

counterpoint to models of unchecked retrotransposon integration as a driving force in Y-chromosome evolution<sup>17,24</sup>.

Despite the elaborate structure of the chimpanzee MSY, its gene repertoire is considerably smaller and simpler than that of the human MSY (Table 1) as a result of gene loss in the chimpanzee lineage and gene acquisition in the human lineage. For example, we previously discovered that the chimpanzee X-degenerate regions had lost 4 out of 16 genes through inactivating mutations, whereas the human X-degenerate regions had not lost any genes since the time of the last common ancestor<sup>9</sup>. We also reported that two X-transposed genes in the human MSY had been acquired since the time of the last common ancestor<sup>13</sup>.

To investigate whether the gene content of the ampliconic regions differs in chimpanzee and human, we searched the chimpanzee MSY sequence for homologues of all known human ampliconic genes, and we assessed their open reading frames, splice sites, and transcriptional activity electronically and experimentally (Supplementary Tables 3–5 and Supplementary Fig. 7). In addition, we searched for new chimpanzee ampliconic genes using a combination of electronic prediction and shotgun sequencing (>38 Mb total) of chimpanzee testis complementary DNA. We found no new chimpanzee ampliconic genes. We did discover that, within the ampliconic regions, three out of nine multi-copy, testis-expressed gene families present in human have been mutationally disabled or are simply absent in chimpanzee (Table 1). For example, the chimpanzee MSY contains five loci homologous to the human *XKRY* gene family, but all five copies share a frameshift mutation that severely truncates the open reading frame and predicted protein (Supplementary Table 3). We

**Table 1 | Genes and gene families in chimpanzee and human Y chromosomes**

| Sequence class              | Gene            | Numbers of intact copies |            | Difference (human – chimp) |
|-----------------------------|-----------------|--------------------------|------------|----------------------------|
|                             |                 | Human                    | Chimpanzee |                            |
| X-degenerate                | <i>AMELY</i>    | 1                        | 1          | -                          |
|                             | <i>CYorf15A</i> | 1                        | 1          | -                          |
|                             | <i>CYorf15B</i> | 1                        | -          | 1                          |
|                             | <i>DDX3Y</i>    | 1                        | 1          | -                          |
|                             | <i>EIF1AY</i>   | 1                        | 1          | -                          |
|                             | <i>KDM5D</i>    | 1                        | 1          | -                          |
|                             | <i>NLGN4Y</i>   | 1                        | 1          | -                          |
|                             | <i>PRKY</i>     | 1                        | 1          | -                          |
|                             | <i>RPS4Y1</i>   | 1                        | 1          | -                          |
|                             | <i>RPS4Y2</i>   | 1                        | 1          | -                          |
|                             | <i>SRY</i>      | 1                        | 1          | -                          |
|                             | <i>TBL1Y</i>    | 1                        | -          | 1                          |
|                             | <i>TMSB4Y</i>   | 1                        | -          | 1                          |
|                             | <i>USP9Y</i>    | 1                        | -          | 1                          |
|                             | <i>UTY</i>      | 1                        | 1          | -                          |
|                             | <i>ZFY</i>      | 1                        | 1          | -                          |
| Total                       | 16              | 12                       | 4          |                            |
| Ampliconic                  | <i>BPY2</i>     | 3                        | 2          | 1                          |
|                             | <i>CDY</i>      | 4                        | 5          | -1                         |
|                             | <i>DAZ</i>      | 4                        | 4          | -                          |
|                             | <i>HSFY</i>     | 2                        | -          | 2                          |
|                             | <i>PRY</i>      | 2                        | -          | 2                          |
|                             | <i>RBMV</i>     | 6                        | 6          | -                          |
|                             | <i>TSPY</i>     | ~35                      | 6          | 29                         |
|                             | <i>VCY</i>      | 2                        | 2          | -                          |
|                             | <i>XKRY</i>     | 2                        | -          | 2                          |
|                             | Total           | 60                       | 25         | 35                         |
|                             | X-transposed    | <i>TGIF2LY</i>           | 1          | -                          |
| <i>PCDH11Y</i>              |                 | 1                        | -          | 1                          |
| Grand total (genes)         | 78              | 37                       | 41 (53%)   |                            |
| Grand total (gene families) | 27              | 18                       | 7 (33%)    |                            |

confirmed the presence of this disabling mutation in five other chimpanzees and two bonobos, close relatives of common chimpanzees (data not shown). Similarly, the *HSFY* and *PRY* gene families are well represented in the human MSY but absent from the chimpanzee MSY. Although it is unclear whether the *PRY* family was gained in the human lineage or lost in the chimpanzee lineage, the presence of *HSFY* in the cat<sup>25</sup>, rhesus macaque and bull MSYs (H.S., personal communication) leads us to conclude that this gene family was deleted outright in the chimpanzee lineage.

In aggregate, the consequence of gene loss and gain in the chimpanzee and human lineages, respectively, is that the chimpanzee MSY contains only two-thirds as many distinct genes or gene families as the human MSY, and only half as many protein-coding transcription units (Table 1). In contrast, in the remainder of the genome, comparison of chimpanzee draft sequence with human reference sequence suggests that the gene content of the two species differs by <1% (ref. 15). Indeed, at 6 million years of separation, the difference in MSY gene content in chimpanzee and human is more comparable to the difference in autosomal gene content in chicken and human, at 310 million years of separation<sup>26</sup>.

We have conducted the first, to our knowledge, comprehensive comparison of Y chromosomes from two species, providing empirical insight into Y-chromosome evolution and a test of decelerating-decay theories. These theories elegantly account for the degeneration observed in neo-Y chromosomes recently evolved from autosomes<sup>3–8</sup>. However, they did not predict and cannot account for the rapid divergence of the older, highly evolved chimpanzee and human MSYs described here. Instead, remodelling and regeneration have dominated chimpanzee and human MSY evolution during the past 6 million years. We suggest that this renovation, involving both architecture and genetic repertoire, was propelled by a combination of factors acting in synergy. Three of these factors distinguished the evolving hominid MSY from the bulk of the genome: (1) the highly disproportionate role of MSY genes—especially ampliconic gene families—in sperm production<sup>13</sup>, (2) the brisk kinetics of ectopic

recombination and resultant structural change in ampliconic regions<sup>18</sup>, and (3) the absence of crossing over with a homologue, which creates the opportunity for a single advantageous mutation to dictate the evolutionary fate of the MSY ('genetic hitchhiking')<sup>1,3</sup>. The evolutionary effect of these three MSY features was probably multiplied by sperm competition, especially in the lineage of the modern chimpanzee, in which several males mate with the same female at each oestrus<sup>27</sup>. This heightened sperm competition in the chimpanzee lineage, along with positive selection and hitchhiking effects, may account for greater MSY sequence amplification than in the human MSY, and extensive gene loss compared with little or none in the human MSY. In the future, complete Y-chromosome sequences from other species will shed further light on these hypotheses.

## METHODS SUMMARY

**BAC selection and sequencing.** The iterative mapping and sequencing strategy<sup>12</sup> was used to assemble a path of sequenced clones selected from the CHORI-251 and RPCI-43 BAC libraries and the CHORI-1251 fosmid library (<http://bacpac.chori.org>). The rate of error in the finished sequence was estimated by counting mismatches in overlapping clones.

**454 sequencing of flow-sorted Y chromosomes and testis cDNA.** Chromosomes were collected from a lymphoblastoid cell line (Coriell repository number S00600, derived from the same chimpanzee used to construct the CHORI-251 BAC and CHORI-1251 fosmid libraries), prepared as described<sup>28</sup>, and sorted to enrich for Y chromosomes using an Influx cell sorter. The resulting Y-enriched DNA sample was amplified using the GenomiPhi amplification kit (GE Healthcare) to obtain enough template (>1 µg) for 454 sequencing on a GS20 machine. Chimpanzee testis cDNA was generated from total RNA isolated using the RNeasy kit (Qiagen). The cDNA was normalized using the Trimmer kit (Evrogen) and sequenced on a GS20 (454) machine.

**FISH analysis.** All assays were performed on the chimpanzee lymphoblastoid cell line S00600. Interphase FISH analysis was performed as previously described<sup>29</sup>. For each probe set, 200 nuclei were scored. Extended metaphase FISH was performed as previously described<sup>30</sup>.

**Sequence analysis, dot plots and alignments.** Chimpanzee and human gene sequences were aligned using CLUSTAL W with default parameters (<http://www.clustal.org>). The search for new chimpanzee Y-chromosome genes was performed using GenomeScan (<http://genes.mit.edu/genomescan.html>). Square dot plot and triangular dot plot analyses were performed using custom Perl codes that are available at [http://jura.wi.mit.edu/page/papers/Hughes\\_et\\_al\\_2005/tables/dot\\_plot.pl](http://jura.wi.mit.edu/page/papers/Hughes_et_al_2005/tables/dot_plot.pl) and [http://jura.wi.mit.edu/page/Y/azfc/self\\_dot\\_plot.pl](http://jura.wi.mit.edu/page/Y/azfc/self_dot_plot.pl), respectively.

**RT-PCR.** Total RNAs were isolated from male chimpanzee tissues (testis, liver, lung and spleen; Yerkes National Primate Research Center) using the RNeasy kit (Qiagen). PCR with reverse transcription (RT-PCR) primer sequences and product sizes are listed in Supplementary Table 5.

Received 3 August; accepted 24 November 2009.

Published online 13 January 2010.

- Rice, W. R. Evolution of the Y sex chromosome in animals. *Bioscience* **46**, 331–343 (1996).
- Lahn, B. T. & Page, D. C. Four evolutionary strata on the human X chromosome. *Science* **286**, 964–967 (1999).
- Charlesworth, B. & Charlesworth, D. The degeneration of Y chromosomes. *Phil. Trans. R. Soc. Lond. B* **355**, 1563–1572 (2000).
- Bachtrog, D. The temporal dynamics of processes underlying Y chromosome degeneration. *Genetics* **179**, 1513–1525 (2008).
- Filatov, D. A., Moneger, F., Negrutiu, I. & Charlesworth, D. Low variability in a Y-linked plant gene and its implications for Y-chromosome evolution. *Nature* **404**, 388–390 (2000).
- Bachtrog, D. & Charlesworth, B. Reduced adaptation of a non-recombining neo-Y chromosome. *Nature* **416**, 323–326 (2002).
- Liu, Z. *et al.* A primitive Y chromosome in papaya marks incipient sex chromosome evolution. *Nature* **427**, 348–352 (2004).
- Peichel, C. L. *et al.* The master sex-determination locus in threespine sticklebacks is on a nascent Y chromosome. *Curr. Biol.* **14**, 1416–1424 (2004).
- Hughes, J. F. *et al.* Conservation of Y-linked genes during human evolution revealed by comparative sequencing in chimpanzee. *Nature* **437**, 100–103 (2005).

- Kuroki, Y. *et al.* Comparative analysis of chimpanzee and human Y chromosomes unveils complex evolutionary pathway. *Nature Genet.* **38**, 158–167 (2006).
- Hughes, J. F. *et al.* Has the chimpanzee Y chromosome been sequenced? *Nature Genet.* **38**, 853–854 (2006).
- Kuroda-Kawaguchi, T. *et al.* The AZFc region of the Y chromosome features massive palindromes and uniform recurrent deletions in infertile men. *Nature Genet.* **29**, 279–286 (2001).
- Skaletsky, H. *et al.* The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* **423**, 825–837 (2003).
- Page, D. C., Harper, M. E., Love, J. & Botstein, D. Occurrence of a transposition from the X-chromosome long arm to the Y-chromosome short arm during human evolution. *Nature* **311**, 119–123 (1984).
- The Chimpanzee Sequencing and Analysis Consortium. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* **437**, 69–87 (2005).
- Watanabe, H. *et al.* DNA sequence and comparative analysis of chimpanzee chromosome 22. *Nature* **429**, 382–388 (2004).
- Charlesworth, B. The evolution of sex chromosomes. *Science* **251**, 1030–1033 (1991).
- Repping, S. *et al.* High mutation rates have driven extensive structural polymorphism among human Y chromosomes. *Nature Genet.* **38**, 463–467 (2006).
- Repping, S. *et al.* Recombination between palindromes P5 and P1 on the human Y chromosome causes massive deletions and spermatogenic failure. *Am. J. Hum. Genet.* **71**, 906–922 (2002).
- Repping, S. *et al.* Polymorphism for a 1.6-Mb deletion of the human Y chromosome persists through balance between recurrent mutation and haploid selection. *Nature Genet.* **35**, 247–251 (2003).
- Lange, J. *et al.* Isodicentric Y chromosomes and sex disorders as byproducts of homologous recombination that maintains palindromes. *Cell* **138**, 855–869 (2009).
- Rozen, S. *et al.* Abundant gene conversion between arms of palindromes in human and ape Y chromosomes. *Nature* **423**, 873–876 (2003).
- Yohn, C. T. *et al.* Lineage-specific expansions of retroviral insertions within the genomes of African great apes but not humans and orangutans. *PLoS Biol.* **3**, e110 (2005).
- Steinemann, S. & Steinemann, M. Retroelements: tools for sex chromosome evolution. *Cytogenet. Genome Res.* **110**, 134–143 (2005).
- Murphy, W. J. *et al.* Novel gene acquisition on carnivore Y chromosomes. *PLoS Genet.* **2**, e43 (2006).
- Hillier, L. W. *et al.* Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* **432**, 695–716 (2004).
- Dixson, A. F. *Primate Sexuality: Comparative Studies of the Prosimians, Monkeys, Apes and Human Beings* 37–44 (Univ. Chicago Press, 1998).
- Mefford, H. C. *et al.* Comparative sequencing of a multicopy subtelomeric region containing olfactory receptor genes reveals multiple interactions between non-homologous chromosomes. *Hum. Mol. Genet.* **10**, 2363–2372 (2001).
- Saxena, R. *et al.* Four DAZ genes in two clusters found in the AZFc region of the human Y chromosome. *Genomics* **67**, 256–267 (2000).
- Yeh, R. F., Lim, L. P. & Burge, C. B. Computational inference of homologous gene structures in the human genome. *Genome Res.* **11**, 803–816 (2001).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank L. Brown, V. Frazzoni, G. Rogers and A. Berger for technical assistance, and L. Brown, M. Carmell, G. Dokshin, M. Gill, J. Mueller, K. Romer and Y. Soh for comments on the manuscript. This work was supported by the National Institutes of Health and the Howard Hughes Medical Institute.

**Author Contributions** J.F.H., H.S., W.C.W., S. Rozen, R.K.W. and D.C.P. planned the project. J.F.H., H.S. and T.P. performed BAC mapping and RT-PCR analysis. T.A.G., P.J.M., R.S.F. and S.D.M. were responsible for finished BAC sequencing. J.F.H. and H.S. performed comparative sequence analyses. S.K.M.v.D. and S. Repping performed FISH analyses. C.F. and B.J.T. flow-sorted chromosomes. D.P.L. and E.R.M. carried out 454 sequencing. J.F.H. and D.C.P. wrote the paper.

**Author Information** cDNA sequences of chimpanzee ampliconic genes have been deposited in GenBank (<http://www.ncbi.nlm.nih.gov>) under accession numbers AY958084 (BPY2), AY958081 (CDY), AY958083 (DAZ), AY995201/2 (RBMV), AY958082 (TSPY) and AY995203 (XKRY). Chimpanzee testis cDNA sequences have been deposited in GenBank under accession numbers EX654441–EX654483. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests: details accompany the full-text HTML version of the paper at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to D.C.P. ([dcpage@wi.mit.edu](mailto:dcpage@wi.mit.edu)).