# What Can the Y Chromosome Tell Us about the Origin of Modern Humans?

CHRIS TYLER-SMITH

**Summary.** Crow has proposed that a change to an X–Y homologous gene was an important event in human speciation. This chapter reviews how our current understanding of the human Y chromosome can contribute to an evaluation of this hypothesis. Human and ape Y lineages are generally believed to have split about 5–7 million years ago, while extant human Y lineages trace back to a common ancestor that probably lived between 40 and 200 thousand years ago. Between these dates, two substantial segments of DNA on the Y chromosome were duplicated on the Y: the Yq pseudoautosomal region and the Xq/Yp homology region. The former does not contain any good candidate speciation genes but the latter may, and these are discussed by Carole Sargent *et al.* later in these *Proceedings*. The Xq–Yp transposition probably occurred soon after the ape–human split and, at the same time or subsequently, was divided in two by an inversion. An exhaustive evaluation of the genes contained in this region provides the best way to test Crow's hypothesis.

## INTRODUCTION

### Why consider the Y chromosome?

DURING THE LAST DECADE, Crow has developed the hypothesis that a change in the expression of a pair of genes on the Y and Y chromosomes was a critical event in the evolution of language and the speciation of modern humans; these ideas are outlined in Tim Crow's chapter in these *Proceedings*. If the relevant genes had already been identified, there would be no need for the present chapter. The most direct test of the hypothesis is to search for such genes, and this approach is described by Carole Sargent *et al*. in the next chapter of these *Proceedings*. In the absence of clear evidence for the crucial genes, the present chapter reviews relevant aspects of our knowledge of the human Y chromosome and considers what this indirect information can contribute to the debate.

*Proceedings of the British Academy*, **106**, 217–229, © The British Academy 2002.

**Genetic properties of the Y chromosome**

Humans have 46 chromosomes, consisting of one pair of sex chromosomes (designated X and Y) and 22 other pairs (designated autosomes). The Y chromosome carries a gene, *SRY* (*S*ex-determining *R*egion of the *Y* chromosome; Sinclair *et al*., 1990), that provides the primary sex-determining signal and directs development away from the default female pathway to the male pathway; thus in females the karyotype is 46,XX, while in males it is 46,XY. The primary sex-determining role of the Y chromosome has several important consequences for its genetics and evolution, some obvious but others less so. *SRY* must be haploid (present in only one copy per genome) in order for this sex-determining mechanism to work. It therefore has no homologue and so cannot recombine. However, recombination is required for successful meiosis, and so the Y chromosome itself must recombine. This paradox is resolved by dividing the Y into distinct sections: pseudoautosomal, which recombines with the X chromosome, and Y-specific, which carries *SRY* and does not recombine. A priori, it would seem that one pseudoautosomal region and one Y-specific region would be sufficient, and their sizes cannot be predicted. In fact, the human Y has two pseudoautosomal regions of 2.4 Mb and 0.3 Mb, one at each end, but the majority of the chromosome, the size of which is conventionally estimated at about 60 Mb, shows Y-specific inheritance (Figure 1a). In addition to *SRY*, a considerable number of other genes are located in the Y-specific region. For several reasons, it is difficult to give the exact number of these: despite the progress of the human genome sequencing project, our knowledge is incomplete and new genes are still being identified. Some genes are duplicated or are members of multigene families, and it is not always clear whether a gene is active or an inactive pseudogene. Furthermore, polymorphisms are found in the population so that a gene may be present in some individuals and absent from others, or the number of copies in a multigene family may vary between individuals. Nevertheless, about 20 different protein-coding genes with diverse functions have been described and are shown in Figure 1d. This is, however, a small number compared with other chromosomes. Chromosome 22, for example, is smaller but at least 545 genes were predicted from the sequence (Dunham *et al*., 1999).

A consequence of the lack of recombination over most of the Y chromosome and the low gene density, is that large-scale rearrangements are tolerated more readily than on other chromosomes. These rearrangements can take forms such as duplications, deletions, inversions and translocations. Because of the low gene density, the rearrangements themselves are unlikely to disrupt the expression of crucial genes. In the absence of recombination, changes in the position of the centromere do not lead to the generation of acentric or dicentric chromosomes.

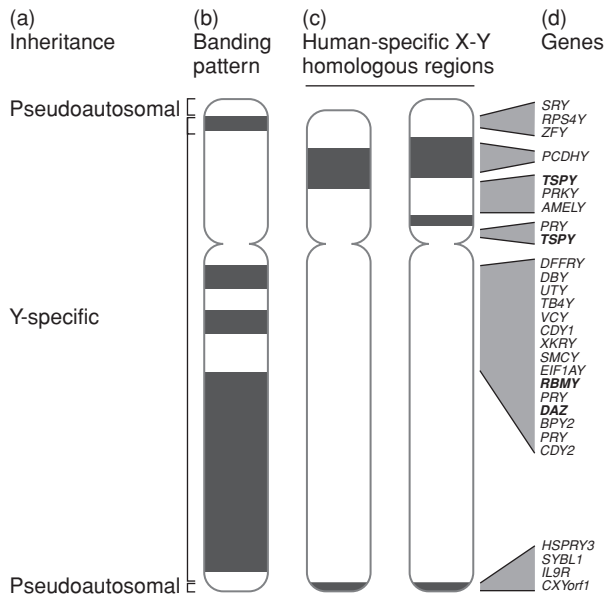


**Figure 1.** (a) Inheritance of different sections of the Y chromosome. (b) Conventional representation of the banding pattern. (c) X–Y homologous regions specific to the human Y lineage, indicated in black. The proximal Yp region is absent from some males (left). (d) Genes from the Y-specific and Yq pseudoautosomal regions (Lahn & Page, 1997; Jobling & Tyler-Smith, 2000). Loci shown in bold are present in multiple copies. Several genes are known from the Yp pseudoautosomal region, but are not listed here.

A consequence of the haploid state of the Y chromosome is that there are fewer copies of the Y present in the population than of the X or any autosome: a couple have, between them, four copies of each autosome and three copies of the X, but only one copy of the Y. This means that changes in gene frequency due to chance (genetic drift) will occur more rapidly on the Y than on other chromosomes, and so variant Y chromosomes will have a better chance of being fixed in the population. All genomes or sections of genomes that do not recombine must be able to trace their ancestry back to a single DNA molecule. The time taken is called the time to the most recent common ancestor (TMRCA), divergence time or coalescence time. If everything else is equal, this time will depend on the effective population size, and so the Y coalescence time will be one-third of that for a region on the X, or one-quarter of that of an autosome.

A final important property is also a consequence of the lack of recombination. If a variant Y chromosome has a selective advantage in the population, it

will tend to increase in frequency and may become fixed. If this occurs, all polymorphisms that happen to be present on the selected Y will also be fixed, and all other variants will be lost from the population, even though these extra variants themselves may be neutral. This is sometimes referred to as a selective sweep or hitch-hiking. If such an event had occurred in the recent past, an unexpectedly low amount of variation would be seen on current Ys.

Thus, overall, the Y chromosome is distinguished from the rest of the genome by its ability to tolerate large-scale rearrangements and the higher probability of such changes being fixed. In effect, it is reinventing itself more rapidly.

## What can Y chromosomal studies contribute?

Studies of the genomes, including the Y chromosomes, of contemporary humans and apes can contribute to our understanding of three relevant areas: (1) the timing of the divergence between apes and humans; (2) the changes that have occurred to the Y chromosome along the human lineage; and (3) the timing of the divergence of modern human Y chromosomes (Figure 2). The next three sections will consider these areas.

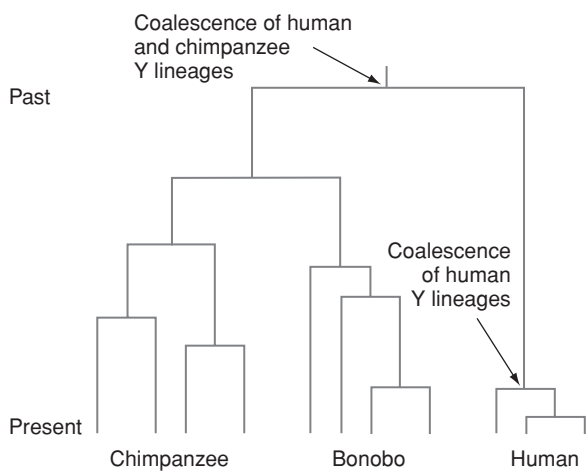


**Figure 2.** Schematic representation of ape and human Y lineages. Representation is schematic because no suitable data are available for ape Ys.

## THE APE–HUMAN Y CHROMOSOMAL COALESCENCE

### Comparison of ape and human Y chromosomes

So far, our knowledge of the overall organisation of ape and human Y chromosomes is derived mainly from cytogenetic studies. Comparisons of high-resolution (late prophase) G-banded chromosomes were described by Yunis & Prakash (1982). They observed large differences in the heterochromatic part of the Y, but judged the euchromatic part to be homologous between the different species. However, a more recent comparison, carried out using *in situ* hybridisation (Archidiacono *et al*., 1998), which allows the positions of specific sequences on the chromosome to be visualised, showed that the situation was more complex. Sequences from the short arm of the human Y were homologous to the long arm of the chimpanzee Y. In addition, several differences in the copy number and order of the sequences were also detected, but Archidiacono *et al.* (1998) did not attempt to propose an evolutionary scenario for the chromosome.

### Timing of the ape–human coalescence

The date of the split between apes and humans can be estimated by measuring the DNA sequence difference (or an indirect measure of this, such as protein structure difference) and translating this into years using an event in the fossil record that is considered well-calibrated. The initial molecular dating of the ape–human split used antisera to serum albumins (Sarich & Wilson, 1967) and an assumed divergence between hominoids and Old World monkeys of 30 million years ago. It produced a date of about 5 million years ago, which initially caused some surprise, but has subsequently been supported by a large number of other studies, and a date of 5–7 million years ago is now generally accepted. One exception is the work of Arnason *et al.* (1996), who estimated a divergence time of about 13.5 million years ago using mitochondrial DNA sequence information and the split between whales and cows at 60 million years ago as their standard.

## CHANGES ON THE HUMAN Y LINEAGE

Although their structures are incompletely understood, human and chimpanzee Y chromosomes have clearly accumulated many gross differences (Archidiacono *et al*., 1998). Half of these have probably occurred on the human lineage, and the human-specific changes can be identified by comparison with additional apes such as gorilla or orang-utan. Among them, the

differences resulting in new X–Y homologous regions are of particular interest here.

**The Xq–Yp translocation and Yp inversion**

The focus of the Crow hypothesis is on a region of Xq that has transposed onto Yp since the divergence between apes and humans (Page *et al.*, 1984). This region is being studied intensively by Affara and colleagues, and their work is described in the next chapter. Here, I will comment only on three aspects.

First, when did the transposition occur? It is present in all human Y chromosomes that have been examined, and no chimpanzee Ys, so must have taken place after the divergence between the human and chimpanzee Y lineages and before the TMRCA of modern human Ys (Figure 2). This would place it after 5–7 million years ago and before 40–200 thousand years ago (see below). These, however, are wide limits. An alternative, and to some extent independent, date can be obtained from the extent of sequence divergence between the X and Y copies. This date is only partially independent because it may be calibrated against the same events in the fossil record as the ape–human divergence. Schwartz *et al.* (1998) measured the X–Y similarity in about 5 kb of sequence at 99.3 ± 0.2% and translated this into a date of *c.* 3–4 million years ago using an estimate of 0.2% divergence per million years (Shimmin *et al.*, 1993). A comparison of 100 kb of sequence from two bacterial artificial chromosome (BAC) clones (AC004388 from the X and AC010722 from the Y) reveals a similarity of 99.0%, which would correspond to a slightly older date of *c.* 5 million years ago using the same divergence rate. This suggests that the translocation occurred soon after the split.

Secondly, the Xq/Yp homology region forms one contiguous block on the X, but two distinct blocks on the Y, one distal and one proximal. It is thought that the Y copy has undergone an inversion with one endpoint within the homology region and the other elsewhere on Yp (Schwartz *et al.*, 1998). When did this inversion occur? It cannot have happened before the transposition, but could have been contemporaneous with it, or have occurred at any time between then and the modern human Y TMRCA. It is difficult to see how any analysis of modern Ys can narrow this interval.

Thirdly, the proximal Yp homology region is entirely absent from some males as a result of a deletion polymorphism (Figure 1c) (Santos *et al.*, 1998). The individuals carrying this deletion are quite rare (two out of 350 in the initial study, 0.6%) and their phenotypes have not been studied in detail, but they were detected in a survey of normal males and there is no evidence of any abnormalities. Thus it is unlikely that a crucial gene lies within this region.

### The Xq–Yq pseudoautosomal region

A second relevant difference between ape and human Y chromosomes is that the human Y has a long arm pseudoautosomal region (Yq PAR; Figure 1) but ape Ys do not. Like the Xq/Yp homology region, the Yq PAR was formed between the times of the ape–human Y divergence and the human Y coalescence. However, unlike the Xq/Yp region, it continues to undergo exchange with the X (Freije *et al*., 1992) and thus has not diverged in sequence, so that the timing of its origin cannot be refined using sequence information. The region has now been completely sequenced (Ciccodicola *et al*., 2000) and is 330 kb long. Four genes have been predicted or detected: *HSPRY3*, *SYBL1*, *IL9R* and *CXYorf1*. *HSPRY* and *SYBL1* appear to be inactive on the Y, and are thus poor candidates for genes leading to speciation through their presence on the Y. This conclusion must, however, be considered provisional because gene activity has only been measured in a small number of cell types, and it remains possible that one or both of these Y genes are active in cells that have not yet been tested. *IL9R* and *CXYorf1* are probably both active on the Y, but additional factors make them poor candidates for speciation genes. *IL9R* is a growth factor for some of the cells in the blood, and thus seems excluded by its function. The function of *CXYorf1* is unknown, and its sequence provides no clues, but highly homologous genes are present on at least six other chromosomes, and it seems unlikely that such a small increase in gene number (and gene product level) could have a large phenotypic effect.

Thus the Yq PAR provides a human-specific region of X–Y homology, but no good candidate genes for speciation events.

## THE HUMAN Y CHROMOSOMAL COALESCENCE

As outlined in the introduction, all copies of the non-recombining section of the Y chromosome must have descended from a single individual. Any change to the Y that contributed to speciation must have been present in this individual. It is therefore important to establish when this individual lived: the coalescence time of surviving Y chromosomes.

### Variation on the human Y chromosome

Y DNA variation takes many forms, but the variants that have been most useful have been binary polymorphisms and microsatellites (also called short tandem repeats; STRs). Binary polymorphisms derive their name from the finding that they have just two alleles; the most abundant are single nucleotide polymorphisms (SNPs), where one nucleotide is replaced by another (for example

T by C), but they also include insertions or deletions of a few nucleotides and insertions of retroposon sequences. Microsatellites consist of small units (for example GATA) that are repeated in tandem. The number of copies varies between individuals: for example 11 on one Y chromosome and 12 on another. Binary polymorphisms have low mutation rates, which have been estimated from comparisons of human and chimpanzee DNA at $1.2 \times 10^{-9}$ per nucleotide per year, assuming no selection and a split 5 million years ago (Thomson *et al.*, 2000). If the generation time is 25 years, this would correspond to $3 \times 10^{-8}$ per nucleotide per generation. In contrast, microsatellites have much higher mutation rates, measured in modern families at $2.1 \times 10^{-3}$ per locus per generation (Heyer *et al.*, 1997).

**Timing of the human Y coalescence**

There have been several estimates of the Y chromosomal coalescence time and these are summarised in Figure 3, where they are compared with autosomal and X chromosomal coalescence times. Six of the first seven estimates were around 150 thousand years ago, with wide confidence limits (Hammer, 1995; Whitfield *et al.*, 1995; Tavare *et al.*, 1997; Underhill *et al.*, 1997; Hammer *et al.*, 1998; Karafet *et al.*, 1999). The exception was estimate 2 (Whitfield *et al.*, 1995). However, reconsideration of estimate 2 revealed that it was an estimate of the coalescence time of five chromosomes, and that the coalescence time of the entire population of Y chromosomes would be significantly older (estimate 3) (Tavare *et al.*, 1997); thus estimates 1–7 are all consistent. They contrast with estimates 8 and 9 (Pritchard *et al.*, 1999; Thomson *et al.*, 2000), which suggest a more recent date of about half of the earlier estimates, but again with wide confidence limits. Estimate 8 is based on microsatellite variation, but estimate 9, like 1–7, is based on binary variation.

Which estimate is correct? This apparently esoteric question is important, because it raises the possibility that selection may have acted on the Y. According to simple population genetics models, in the absence of selection coalescence time is proportional to effective population size. As the ratio of Y chromosomes : X chromosomes : autosomes is 1 : 3 : 4, coalescence times should also show this ratio. Y coalescence times 1–7 are in reasonable agreement with this expectation; 8 and 9 less so. Several explanations are possible. The confidence limits are very wide, so there may be no significant departure from neutral expectation to explain. The effective population size for the Y may be smaller than assumed, or mutation rates may differ between loci in ways that are not understood. Alternatively, the Y may have undergone a selective sweep. According to this scenario, an advantageous Y variant arose about 60,000 years ago and spread through the population because of selection. Could this hypothetical variant have increased language ability?

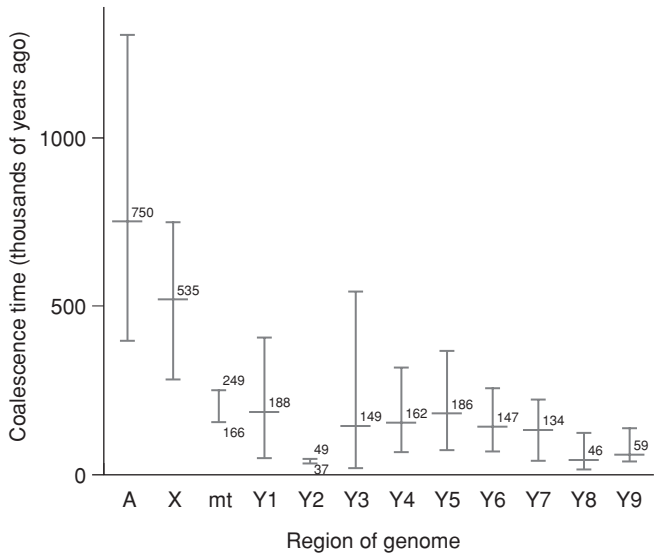


**Figure 3.** Published coalescence times for autosomal (A), X chromosomal (X), mitochondrial DNA (mt) and Y chromosomal (Y1–Y9) lineages. In most cases the estimate is shown, together with the 95% confidence limits given by the authors, or two times the standard error as an approximation to these. For mt and Y2, only the limits are given. Sources of data are: A, Harding *et al.* (1997); X, Kaessmann *et al.* (1999); mt, Vigilant *et al.* (1991); Y1, Hammer (1995); Y2, Whitfield *et al.* (1995); Y3, Tavare *et al.* (1997); Y4 and Y5, Underhill *et al.* (1997); Y6, Hammer *et al.* (1998); Y7 Karafet *et al.* (1999); Y8, Pritchard *et al.* (1999); Y9, Thomson *et al.* (2000).

A major reason for the different coalescent times is that estimates 1–7 assumed a constant population size, while 8 and 9 assumed population growth. As the autosomal and X chromosomal estimates also assumed constant population size, and all the different loci have been present in the same world-wide population, it may be necessary to compare estimates 8 and 9 with autosomal and X chromosomal coalescence times calculated assuming population growth. Thus the data and analysis are still too preliminary to show whether or not there has been a selective sweep.

## CONCLUSIONS

There have been substantial changes to the human Y after the split from apes, including the addition of genes from the X, and some of these are thus

expressed from both the X and the Y in humans but not in apes. These changes occurred after the ape–human split at 5–7 million years ago and before the common ancestor of modern human Ys at *c.* 40–200 thousand years ago. The complete sequences of the X–Y homologous regions should be available shortly and will provide an excellent starting point for the identification of all the genes they contain.

The origin of the unique characteristics of modern humans must involve changes to a small number of genes, but the nature of the changes and the number and identity of the genes are unknown. Few specific testable hypotheses have been suggested. The Crow hypothesis proposes that one of the human-specific X–Y homologous genes influences linguistic ability and contributed to speciation. We have the tools to test this hypothesis and should soon know whether there is such a gene.

## DISCUSSION

**Questioner:** Does the fact you can find certain sets of haplotypes localised suggest a certain period of time for the change to have taken place?

**Tyler-Smith:** We can estimate the times at which the various haplogroups appeared from the amounts of variation we find within them when we use other markers like microsatellites, and in some cases these estimates are of just a few hundred to a few thousand years. In other cases they can be 10,000 or a few 10,000 years. The one that I showed with the most extreme localisation in China is a quite diverse set of chromosomes where we estimated a time of perhaps 10,000 years. So I think the distribution we see will depend on the population history, and if a chromosome arose in a sedentary population then it may remain quite localised. But some that are widespread, like the one that was spread across the north of Asia and Europe, have a younger estimated time.

**Questioner:** Does the distribution of Y variants tell us anything about the peopling of the Americas?

**Tyler-Smith:** Quite a lot of work has been done in that area that shows a high proportion of the native American populations have a single Y type that is not found outside America except in a few rare instances that are explained by back migration. So since that type is found through most of the populations including the different linguistic groups then that supports the idea that there was only one migration to the Americas and only one peopling, and goes against some hypotheses of multiple waves of migration. If we try and trace back the origins of that lineage it traces back to central Siberia: northern parts of Asia rather than southern China/Asia. What we would really like to know is the tim-

ing: did it occur perhaps just before 12–15 thousand years ago or did it occur 30–40,000 years ago? In that case the evidence is not just good enough to give a definitive answer.

**Questioner:** I am puzzled about the whole of your theory: it seems that it hinges on the fact that there's a later change in a part of the Y chromosome and it's transferred and then it splits up. This is what Tim Crow realised and related to lateralisation and the rest. Then you go at great length to show that there is considerable variation in the Y chromosome, not only between individuals but in areas itself. Why is it that one of them has a very significant change and another not at all?

**Tyler-Smith:** That depends on whether gene expression is affected. Some of the gross chromosomal rearrangements that you see, like in the movements of the centromere or an inversion of a large part of the chromosomes, if they do not affect a gene or interrupt a gene then they may have no phenotypic consequences whatever. Whereas a single base change in a crucial gene that inactivates it may be lethal, so I think that the magnitude of the genetic change is not a measure of how phenotypically significant it is. So far as Tim's hypothesis is concerned, it is essential for that hypothesis that there are genes or a gene within the region that transposed from the X to the Y.

**Questioner:** So you think there is a gene there?

**Tyler-Smith:** It is a hypothesis that there is a gene in that region and this is a testable hypothesis. And since these entire regions are now being sequenced and there are reasonably efficient ways of detecting genes, if there is such a gene it will become apparent. If there is not such a gene, then the hypothesis is wrong. I think Nabeel Affara will tell us that.

**Questioner:** I think my question is a follow-up on that one. You particularly did not tell us about the homologous regions and your haplotypes are clearly all very different. Are all your haplotypes contained within the homologous region, or do some not have this region?

**Tyler-Smith:** One of the deletions that we characterised removes the proximal part of that homology region. But as far as we know, all of the other haplotypes contain the homology region and particularly all of them contain the large block of homology.

**Questioner:** Can you explain how the translocation and inversion are related? Specifically what is the calibration that you used to obtain its age?

**Tyler-Smith:** The translocation was dated by comparing the sequence and the X and Y chromosomes. Also we know it occurred after the divergence of human and chimp Y chromosomes. So that provides broad limits. The more

precise calibration that Tim mentioned of 3 million years or so is obtained from comparing sequences. That will have broad confidence limits. As for the timing of the inversion, I don't know how it can be dated except to say that it must have been later than the translocation. I wouldn't like to say what the 95% confidence limits are, but I would suspect that they are pretty broad and I would like to say that I don't know how to date the inversion.

## References

Archidiacono, N., Storlazzi, C.T., Spalluto, C., Ricco, A.S., Marzella, R. & Rocchi, M. (1998) Evolution of chromosome Y in primates. *Chromosoma*, **107**, 241–6.

Arnason, U., Gullberg, A., Janke, A. & Xu, X. (1996) Pattern and timing of evolutionary divergences among hominoids based on analyses of complete mtDNAs. *Journal of Molecular Evolution*, **43**, 650–61.

Ciccodicola, A., D'Esposito, M., Esposito, T., Gianfrancesco, F., Migliaccio, C., Miano, M.G., Matarazzo, M.R., Vacca, M., Franze, A., Cuccurese, M., Cocchia, M., Curci, A., Terracciano, A., Torino, A., Cocchia, S., Mercadante, G., Pannone, E., Archidiacono, N., Rocchi, M., Schlessinger, D. & D'Urso, M. (2000) Differentially regulated and evolved genes in the fully sequenced Xq/Yq pseudoautosomal region. *Human Molecular Genetics*, **9**, 395–401.

Dunham, I., Shimizu, N., Roe, B.A., Chissoe, S., Hunt, A.R., Collins, J.E., Bruskiewich, R., Beare, D.M., Clamp, M., Smink, L.J., Ainscough, R., Almeida, J.P., Babbage, A., Bagguley, C., Bailey, J., Barlow, K., Bates, K.N., Beasley, O., Bird, C.P., Blakey, S., Bridgeman, A.M., Buck, D., Burgess, J., Burrill, W.D., O'Brien, K.P., *et al.* (1999) The DNA sequence of human chromosome 22. *Nature*, **402**, 489–95.

Freije, D., Helms, C., Watson, M.S. & Donis-Keller, H. (1992) Identification of a second pseudoautosomal region near the Xq and Yq telomeres. *Science*, **258**, 1784–7.

Hammer, M.F. (1995) A recent common ancestry for human Y chromosomes. *Nature*, **378**, 376–8.

Hammer, M.F., Karafet, T., Rasanayagam, A., Wood, E.T., Altheide, T.K., Jenkins, T., Griffiths, R.C., Templeton, A.R. & Zegura, S.L. (1998) Out of Africa and back again: nested cladistic analysis of human Y chromosome variation. *Molecular Biology and Evolution*, **15**, 427–41.

Harding, R.M., Fullerton, S.M., Griffiths, R.C., Bond, J., Cox, M.J., Schneider, J.A., Moulin, D.S. & Clegg, J.B. (1997) Archaic African and Asian lineages in the genetic ancestry of modern humans. *American Journal of Human Genetics*, **60**, 772–89.

Heyer, E., Puymirat, J., Dieltjes, P., Bakker, E. & de Knijff, P. (1997) Estimating Y chromosome specific microsatellite mutation frequencies using deep rooting pedigrees. *Human Molecular Genetics*, **6**, 799–803.

Jobling, M.A. & Tyler-Smith, C. (2000) New uses for new haplotypes: the human Y chromosome, disease and selection. *Trends in Genetics*, **16**, 356–62.

Kaessmann, H., Heissig, F., von Haeseler, A. & Paabo, S. (1999) DNA sequence variation in a non-coding region of low recombination on the human X chromosome. *Nature Genetics*, **22**, 78–81.

Karafet, T.M., Zegura, S.L., Posukh, O., Osipova, L., Bergen, A., Long, J., Goldman, D., Klitz, W., Harihara, S., de Knijff, P., Wiebe, V., Griffiths, R.C., Templeton, A.R. & Hammer, M.F. (1999) Ancestral Asian source (s) of new world Y-chromosome founder haplotypes. *American Journal of Human Genetics*, **64**, 817–31.

Lahn, B.T. & Page, D.C. (1997) Functional coherence of the human Y chromosome. *Science*, **278**, 675–80.

Page, D.C., Harper, M.E., Love, J. & Botstein, D. (1984) Occurrence of a transposition from the X-chromosome long arm to the Y-chromosome short arm during human evolution. *Nature*, **311**, 119–23.

Pritchard, J.K., Seielstad, M.T., Perez-Lezaun, A. & Feldman, M.W. (1999) Population growth of human Y chromosomes: a study of Y chromosome microsatellites. *Molecular Biology and Evolution*, **16**, 1791–8.

Santos, F.R., Pandya, A. & Tyler-Smith, C. (1998) Reliability of DNA-based sex tests. *Nature Genetics*, **18**, 103.

Sarich, V.M. & Wilson, A.C. (1967) Immunological time scale for hominid evolution. *Science*, **158**, 1200–3.

Schwartz, A., Chan, D.C., Brown, L.G., Alagappan, R., Pettay, D., Disteche, C., McGillivray, B., de la Chapelle, A. & Page, D.C. (1998) Reconstructing hominid Y evolution: X-homologous block, created by X–Y transposition, was disrupted by Yp inversion through LINE–LINE recombination. *Human Molecular Genetics*, **7**, 1–11.

Shimmin, L.C., Chang, B.H. & Li, W.H. (1993) Male-driven evolution of DNA sequences. *Nature*, **362**, 745–7.

Sinclair, A.H., Berta, P., Palmer, M.S., Hawkins, J.R., Griffiths, B.L., Smith, M.J., Foster, J.W., Frischauf, A.M., Lovell-Badge, R. & Goodfellow, P.N. (1990) A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature*, **346**, 240–4.

Tavare, S., Balding, D.J., Griffiths, R.C. & Donnelly, P. (1997) Inferring coalescence times from DNA sequence data. *Genetics*, **145**, 505–18.

Thomson, R., Pritchard, J.K., Shen, P., Oefner, P.J. & Feldman, M.W. (2000) Recent common ancestry of human Y chromosomes: evidence from DNA sequence data. *Proceedings of the National Academy of Sciences of the USA*, **97**, 7360–5.

Underhill, P.A., Jin, L., Lin, A.A., Mehdi, S.Q., Jenkins, T., Vollrath, D., Davis, R.W., Cavalli-Sforza, L.L. & Oefner, P.J. (1997) Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Research*, **7**, 996–1005.

Vigilant, L., Stoneking, M., Harpending, H., Hawkes, K. & Wilson, A.C. (1991) African populations and the evolution of human mitochondrial DNA. *Science*, **253**, 1503–7.

Whitfield, L.S., Sulston, J.E. & Goodfellow, P.N. (1995) Sequence variation of the human Y chromosome. *Nature*, **378**, 379–80.

Yunis, J.J. & Prakash, O. (1982) The origin of man: a chromosomal pictorial legacy. *Science*, **215**, 1525–30.