A selective difference between human Y-chromosomal DNA haplotypes

M.A. Jobling*, G. Williams†, K. Schiebel‡, A. Pandya†, K. McElreavey§, L. Salas§, G.A. Rappold‡, N.A. Affara¶ and C. Tyler-Smith†

DNA analysis is making a valuable contribution to the understanding of human evolution [1]. Much attention has focused on mitochondrial DNA (mtDNA) [2] and the Y chromosome [3,4], both of which escape recombination and so provide information on maternal and paternal lineages, respectively. It is often assumed that the polymorphisms observed at loci on mtDNA and the Y chromosome are selectively neutral and, therefore, that existing patterns of molecular variation can be used to deduce the histories of populations in terms of drift, population movements, and cultural practices. The coalescence of the molecular phylogenies of mtDNA and the Y chromosome to recent common ancestors in Africa [5,6], for example, has been taken to reflect a recent origin of modern human populations in Africa. An alternative explanation, though, could be the recent selective spread of mtDNA and Y chromosome haplotypes from Africa in a population with a more complex history [7]. It is therefore important to establish whether there are selective differences between classes (haplotypes) of mtDNA and Y chromosomes and, if so, whether these differences could have been sufficient to influence the distributions of haplotypes in existing populations. A precedent for this hypothesis has been established for mtDNA in that one mtDNA background increases susceptibility to Leber hereditary optic neuropathy [8]. Although studies of nucleotide diversity in global samples of Y chromosomes have suggested an absence of recent selective sweeps or bottlenecks [9], selection may, in principle, be very important for the Y chromosome because it carries several loci affecting male fertility [10,11] and as many as 5% of males are infertile [11,12]. Here, we show that one class of infertile males, *PRKX/PRKY* translocation XX males, arises predominantly on a particular Y haplotypic background. Selection is, therefore, acting on Y haplotype distributions in the population.

Addresses: *Department of Genetics, University of Leicester, University Road, Leicester LE1 7RH, UK. †CRC Chromosome Molecular Biology Group, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK. ‡Institute of Human Genetics, Ruprecht-Karls-University, Im Neuenheimer Feld 328, D69120 Heidelberg, Germany. §Institut Pasteur, 25 rue du Dr. Roux, 75724 Paris Cedex, France. ¶Human Molecular Genetics Group, Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QP, UK.

Correspondence: C. Tyler-Smith
E-mail: chris@bioch.ox.ac.uk

Received: **10 September 1998**
Revised: **30 October 1998**
Accepted: **9 November 1998**

Published: **14 December 1998**

Current Biology 1998, 8:1391–1394
http://biomednet.com/elecref/0960982200801391

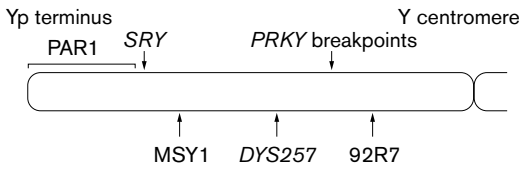
Results and discussion

The human Y chromosome is responsible for the determination of maleness through the action of the sex-determining gene *SRY*, which initiates the differentiation of the testis early in development [13]. Other genes specific to the Y chromosome are subsequently necessary for males to produce adequate numbers of reproductively proficient sperm, and therefore to be fertile. *SRY* lies on the short arm of the Y chromosome (Yp) just proximal to the pseudoautosomal region, in which an obligate homologous recombination event occurs during male meiosis between the X and Y chromosomes [14]. Occasionally this recombination event occurs ectopically, resulting in the translocation of extra Y-chromosomal material, including *SRY*, to the X chromosome. Such translocations give rise to XX males and their reciprocals, XY females: individuals whose karyotypes are discordant with their sex phenotypes. Translocation XX males lack the Y-specific genes necessary for the formation of active sperm and so are completely infertile. The commonest class of XX males, with a frequency of about 1 in 80,000 in the population, results from a recombination between the homologous genes *PRKX* on the short arm of the X chromosome (Xp) and *PRKY* on Yp; the reciprocal translocation accounts for one-third of translocation XY females, although these are rare [15].

The mapping of deleted Y chromosomes [16,17] had suggested an inversion polymorphism on Yp within the European population, and this has been confirmed in normal Y chromosomes using fluorescence *in situ* hybridisation (FISH) [18]. As the *PRKY* gene lies within this inversion (our unpublished observations), we hypothesised that *PRKX/PRKY* translocations might preferentially take place when the *PRKY* gene lies in one particular orientation and can, therefore, align with *PRKX*. If this were so, it would be expected that haplotypes associated with the susceptible orientation would be found preferentially in *PRKX/PRKY* translocation XX males and XY females. We therefore investigated the Y

Table 1

Frequency of Y alleles in the European population and in *PRKX/PRKY* translocation individuals.



	MSY1 class 1	*DYS257*(A)	92R7(T)
European population	53/81	54/81	54/81
XX males	2/24	2/24	n.p.
XX males' fathers	n.t.	n.t.	0/2
XY females	n.p.	n.p.	0/5
XY female's father	0/1	0/1	0/1
Total independent chromosomes	2/25	2/25	0/7
P value (χ^2 test)	<0.001	<0.001	<0.005

PAR1, pseudoautosomal region 1; n.p., not present; n.t., not tested.

chromosome haplotypes in *PRKX/PRKY* translocation individuals of European origin and compared them with the haplotypes found in randomly chosen individuals from the European population.

A wide variety of polymorphic markers is now available to define Y chromosome haplotypes [3], but for the purposes of this study, our choice of markers was limited by two criteria. Firstly, because our sample of translocation individuals were of European origin, we required markers that were usefully polymorphic in Europe. Secondly, because most of the translocation individuals were XX males, and for most of these their fathers' DNAs were unavailable, we preferred markers that lay in the translocated segment of the Y chromosome — the distal ~7 megabases (Mb) of Y-specific material. Three markers fitted the first criterion: the hypervariable minisatellite MSY1 [19] and the two biallelic base-substitutional markers *DYS257* [6] (substitutions occurring between G and A) and 92R7 [20] (substitutions occurring between C and T). The first two of these markers also fitted the second criterion.

We first examined the haplotypes identified by these markers in normal European Y chromosomes. The minisatellite MSY1 consists of an array of around 50–115 repeat units of 25 base pairs which vary in sequence, and the positions of which can be mapped along the array by PCR to provide 'MSY1 codes' [19]. The different repeat types, the most common of which are designated types 1, 3 and 4, are arranged in blocks within the array, and the order of these blocks is described as the 'modular structure' of an array; more than 50 such structures are known [19,21]. Of the 81 chromosomes examined, 53 had MSY1 codes with the same, simple, modular structure (a 5′ block of type 1 repeats, a central block of type 3 repeats, and a 3′ block of type 4 repeats), here denoted 'MSY1 class 1' (Table 1). There is strong association between MSY1 class 1, the *DYS257* A allele and the 92R7 T allele. Between the two biallelic markers the association is complete: 54 of the 81 chromosomes were *DYS257*(A) 92R7(T) — 'A/T' — and the remaining 27 were *DYS257*(G) 92R7(C) — 'G/C'. Of the 54 A/T chromosomes, 53 belonged to MSY1 class 1, whereas all of the G/C chromosomes belonged to other MSY1 modular structural classes. Thus, the majority (80/81) of our sample of European Y chromosomes fell into two broad categories: 'class 1/A/T' (about two-thirds) and 'non-class 1/G/C' (about one-third).

We next assembled a collection of 32 DNA samples representing 29 independent Y chromosomes that had undergone *PRKX/PRKY* translocations and three paternal Y chromosomes. We then asked to which of the two haplotypic classes they belonged. In this set, only the three parental Y chromosomes (from the fathers of two XX males and one XY female) were complete and could be tested with all three markers. A further 24 Y samples (from XX males) could be examined with *DYS257* and MSY1 but not with 92R7. The remaining five samples (from XY females) lacked the distal ~7 Mb of Y-specific material and so could be analysed with 92R7 but not with *DYS257* or MSY1. The three markers were tested, where possible, on the 32 DNA samples. *PRKX/PRKY* translocation chromosomes were found to have predominantly (27/29) the less common non-class 1/G/C haplotype (Table 1), with highly significant associations for all three markers. (Because there is a strong association between the three markers, these tests are not, of course, independent.)

In the investigation of the Yp inversion polymorphism in normal chromosomes by FISH [18] described above, only three chromosomes were analysed. One chromosome had the class 1/A/T haplotype, and the other two — with the opposite orientation — had the non-class 1/G/C haplotype. Thus, although the sample size is too small to provide statistical significance, the orientation of the *PRKY* gene also appears to correlate with the two broad categories of European chromosomes.

The orientation of the *PRKY* gene could, therefore, be responsible for the strong bias towards non-class 1/G/C chromosomes in the *PRKX/PRKY* translocation individuals. The two exceptional translocation individuals with the class 1/A/T haplotype, then, could have arisen in a number of ways. It is possible, for example, that a different *PRKX/PRKY* recombination event could have taken place at a low frequency in the second orientation. Alternatively,

the inversion may have taken place after the mutations defining the class 1/A/T haplotype, so that a few class 1/A/T chromosomes still had the same orientation of *PRKY* as the non-class 1/G/C chromosomes; or the inversion may have reverted. Determining the orientation of the inversion in a greater number of chromosomes would allow us to test these hypotheses and the proposed mechanism for the bias.

The bias towards non-class 1/G/C chromosomes among the translocation individuals is clear whatever its mechanism and this bias will lead to a selective advantage of class 1/A/T chromosomes relative to the susceptible non-class 1/G/C chromosomes in the population. Because the translocation individuals are entirely sterile, we can estimate the selective advantage of the class 1/A/T chromosomes from the frequency of *PRKX*/*PRKY* translocation XX males in the population (1 in 80,000) and the proportion of non-class 1/G/C translocations (27/29): this selective advantage is approximately 1 in 90,000. If the product of selective advantage (s) and effective population size (N_e) is less than or equal to one, alleles are considered neutral [22]. The value of N_e for the Y chromosome has been estimated at approximately 5,000 [23] (with considerable uncertainty — see [24]), so $N_e s$ is approximately 0.06 and selection, although acting, is unlikely to have been responsible for the rise in frequency of class 1/A/T chromosomes in the population because it would have been overwhelmed by stochastic or other selective events. This estimate of N_e, though, is assumed to reflect the small size of the population over evolutionary history. The population is currently much larger and N_e, estimated from the census size and the proportion of reproductively active individuals [25], may be as high as 10^8–10^9 for males. If so, $N_e s$ would be 10^3–10^4. Consequently, the small selective difference may no longer be overwhelmed by random factors and selection may therefore be effective. Nevertheless, the loss of the susceptible haplotype in response to this selective force will be slow. In conclusion, this work shows that selective differences between Y haplotypes can be detected and also provides a model for studies that may show larger selective advantages.

Materials and methods

Patients

The XX males studied were CHM347, WHT1856, WHT2127, GA/WHT1123, RH/WHT1209 and LGL105 [15]; RA and ED (K.S. and G.A.R., unpublished observations); JM, AG, TA, WB and DR [16,26]; KM1, KM6, KM7, KM8, KM10, KM13 and KM15 (K.M., unpublished observations); and Kru, Mcb, Cha, Wla and Cri (M.A.J., unpublished observations). The fathers were those of JM and AG. The XY females were WHT745, WHT1615, WHT1327 and #651 [15] and AM [26]. The father was that of AM. These patients, and the normal males used as controls, were of western European origin.

Scoring of polymorphisms

DYS257 and MSY1 were scored as described [6,19]. 92R7 [20] was scored in most samples by PCR followed by *Hind*III digestion (M.E. Hurles, F.R. Santos and C.T-S., unpublished observations).

Acknowledgements

We thank the DNA donors for making this study possible, and Rosalind Harding and Peter Donnelly for comments on the manuscript and discussions. M.A.J. is supported by a Wellcome Trust Career Development Fellowship (grant number 044910), G.W. by the MRC, K.S. and G.A.R. by the DFG, A.P. by the BBSRC and C.T-S. by the CRC.

References

1. Jorde LB, Bamshad M, Rogers AR: **Using mitochondrial and nuclear DNA markers to reconstruct human evolution.** *Bioessays* 1998, **20:**126-136.
2. Stoneking M, Soodyall H: **Human evolution and the mitochondrial genome.** *Curr Opin Genet Dev* 1996, **6:**731-736.
3. Jobling MA, Tyler-Smith C: **Fathers and sons: the Y chromosome and human evolution.** *Trends Genet* 1995, **11:**449-456.
4. Mitchell RJ, Hammer MF: **Human evolution and the Y chromosome.** *Curr Opin Genet Dev* 1996, **6:**737-742.
5. Cann RL, Stoneking M, Wilson AC: **Mitochondrial DNA and human evolution.** *Nature* 1987, **325:**31-36.
6. Hammer MF, Karafet T, Rasanayagam A, Wood ET, Altheide TK, Jenkins T, *et al.*: **Out of Africa and back again: nested cladistic analysis of human Y chromosome variation.** *Mol Biol Evol* 1998, **15:**427-441.
7. Wise CA, Sraml M, Rubinsztein DC, Easteal S: **Comparative nuclear and mitochondrial genome diversity in humans and chimpanzees.** *Mol Biol Evol* 1997, **14:**707-716.
8. Torroni A, Petrozzi M, D'Urbano L, Sellitto D, Zeviani M, Carrara F, *et al.*: **Haplotype and phylogenetic analyses suggest that one European-specific mtDNA background plays a role in the expression of Leber hereditary optic neuropathy by increasing the penetrance of the primary mutations 11778 and 14484.** *Am J Hum Genet* 1997, **60:**1107-1121.
9. Underhill PA, Jin L, Lin AA, Mehdi SQ, Jenkins T, Vollrath D, *et al.*: **Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography.** *Genome Res* 1997, **7:**996-1005.
10. Vogt PH, Edelmann A, Kirsch S, Henegariu O, Hirschmann P, Kiesewetter F, *et al.*: **Human Y chromosome azoospermia factors (AZF) mapped to different subregions in Yq11.** *Hum Mol Genet* 1996, **5:**933-943.
11. Elliot DJ, Cooke HJ: **The molecular genetics of male infertility.** *Bioessays* 1997, **19:**801-809.
12. Okabe M, Ikawa M, Ashkenas J: **Male infertility and the genetics of spermatogenesis.** *Am J Hum Genet* 1998, **62:**1274-1281.
13. Sinclair AH, Berta P, Palmer MS, Hawkins JR, Griffiths BL, Smith MJ, *et al.*: **A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif.** *Nature* 1990, **346:**240-244.
14. Rappold GA: **The pseudoautosomal regions of the human sex chromosomes.** *Hum Genet* 1993, **92:**315-324.
15. Schiebel K, Winkelmann M, Mertz A, Xu X, Page DC, Weil D, *et al.*: **Abnormal XY interchange between a novel isolated protein kinase gene, *PRKY*, and its homologue, *PRKX*, accounts for one third of all (Y+)XX males and (Y−)XY females.** *Hum Mol Genet* 1997, **6:**1985-1989.
16. Affara NA, Ferguson-Smith MA, Tolmie J, Kwok K, Mitchell M, Jamieson D, *et al.*: **Variable transfer of Y-specific sequences in XX males.** *Nucleic Acids Res* 1986, **14:**5375-5387.
17. Vergnaud G, Page DC, Simmler M-C, Brown L, Rouyer F, Noel B, *et al.*: **A deletion map of the human Y chromosome based on DNA hybridization.** *Am J Hum Genet* 1986, **38:**109-124.
18. Vogt PH, Affara N, Davey P, Hammer M, Jobling MA, Lau Y-FC, *et al.*: **Report of the third international workshop on human Y chromosome mapping 1997.** *Cytogenet Cell Genet* 1997, **79:**1-20.
19. Jobling MA, Bouzekri N, Taylor PG: **Hypervariable digital DNA codes for human paternal lineages: MVR-PCR at the Y-specific minisatellite, MSY1 (DYF155S1).** *Hum Mol Genet* 1998, **7:**643-653.
20. Mathias N, Bayés M, Tyler-Smith C: **Highly informative compound haplotypes for the human Y chromosome.** *Hum Mol Genet* 1994, **3:**115-123.
21. Bouzekri N, Taylor PG, Hammer MF, Jobling MA: **Novel mutational processes in the evolution of a haploid minisatellite, MSY1: array homogenization without homogenization.** *Hum Mol Genet* 1998, **7:**655-659.
22. Kimura M: **Evolutionary rate at the molecular level.** *Nature* 1968, **217:**624-626.

23. Hammer MF: **A recent common ancestry for human Y chromosomes.** *Nature* 1995, **378:**376-378.
24. Tavaré S, Balding DJ, Griffiths RC, Donnelly P: **Inferring coalescence times from DNA sequence data.** *Genetics* 1997, **145:**505-518.
25. Li W-H, Graur D: *Fundamentals of Molecular Evolution.* Massachusetts: Sinauer Associates; 1991.
26. Affara NA, Ferguson-Smith MA, Magenis RE, Tolmie JL, Boyd E, Cooke A, *et al.*: **Mapping the testis determinants by an analysis of Y-specific sequences in males with apparent XX and XO karyotypes and females with XY karyotypes.** *Nucleic Acids Res* 1987, **15:**7325-7342.