

## ARTICLE

# Africans in Yorkshire? The deepest-rooting clade of the Y phylogeny within an English genealogy

Turi E King<sup>1</sup>, Emma J Parkin<sup>1</sup>, Geoff Swinfield<sup>2</sup>, Fulvio Cruciani<sup>3</sup>, Rosaria Scozzari<sup>3</sup>, Alexandra Rosa<sup>4</sup>, Si-Keun Lim<sup>5</sup>, Yali Xue<sup>5</sup>, Chris Tyler-Smith<sup>5</sup> and Mark A Jobling<sup>\*,1</sup>

<sup>1</sup>Department of Genetics, University of Leicester, Leicester, UK; <sup>2</sup>GSGS, 14 Beaconsfield Road, Mottingham, London, UK; <sup>3</sup>Department of Genetics and Molecular Biology, Università degli Studi di Roma 'La Sapienza', Rome, Italy; <sup>4</sup>Human Genetics Laboratory, University of Madeira, Funchal, Portugal; <sup>5</sup>Wellcome Trust Sanger Institute, Hinxton, UK

The presence of Africans in Britain has been recorded since Roman times, but has left no apparent genetic trace among modern inhabitants. Y chromosomes belonging to the deepest-rooting clade of the Y phylogeny, haplogroup (hg) A, are regarded as African-specific, and no examples have been reported from Britain or elsewhere in Western Europe. We describe the presence of an hgA1 chromosome in an indigenous British male; comparison with African examples suggests a Western African origin. Seven out of 18 men carrying the same rare east-Yorkshire surname as the original male also carry hgA1 chromosomes, and documentary research resolves them into two genealogies with most-recent-common-ancestors living in Yorkshire in the late 18th century. Analysis using 77 Y-short tandem repeats (STRs) is consistent with coalescence a few generations earlier. Our findings represent the first genetic evidence of Africans among 'indigenous' British, and emphasize the complexity of human migration history as well as the pitfalls of assigning geographical origin from Y-chromosomal haplotypes.

*European Journal of Human Genetics* (2007) 15, 288–293. doi:10.1038/sj.ejhg.5201771; published online 24 January 2007

**Keywords:** Y chromosome; haplogroup; African; surnames; genealogy; Y-STRs

## Introduction

The population of the UK today is culturally diverse, with 8% of its 54 million inhabitants belonging to ethnic minorities, and over one million classifying themselves as 'Black or Black British' in the 2001 census. These people owe their origins to immigration from the Caribbean and Africa beginning in the mid-20th century; before this time, the population has been seen as typically Western European, and its history has been interpreted in terms of more local immigration, including that of the Saxons, Vikings and Normans.<sup>1</sup> However, in reality, Britain has a long history of contact with Africa (reviewed by Fryer<sup>2</sup>).

Africans were first recorded in the north 1800 years ago, as Roman soldiers defending Hadrian's wall – 'a division of Moors'. Some historians suggest that Vikings brought captured North Africans to Britain in the 9th century. After a hiatus of several hundred years, the influence of the Atlantic slave trade began to be felt, with the first group of West Africans being brought to Britain in 1555. African domestic servants, musicians, entertainers and slaves then became common in the Tudor period, prompting an unsuccessful attempt by Elizabeth I to expel them in 1601. By the last third of the 18th century, there were an estimated 10 000 black people in Britain,<sup>3</sup> mostly concentrated in cities such as London.

Has this presence left a genetic trace among people regarded as 'indigenous' British? In principle, Y-chromosomal haplotyping offers a means to detect long-established African lineages. Haplotypes of the non-recombining region of the Y, defined by slowly mutating binary markers

\*Correspondence: Professor MA Jobling, Department of Genetics, University of Leicester, University Road, Leicester LE1 7RH, UK.

Tel: +44 116 252 3427; Fax: +44 116 252 3378;

E-mail: maj4@leicester.ac.uk

Received 19 September 2006; revised 23 November 2006; accepted 28 November 2006; published online 24 January 2007

such as SNPs, can be arranged into a unique phylogeny.<sup>4–6</sup> These binary haplotypes, known as haplogroups (hg), show a high degree of geographical differentiation, reflecting the powerful influence of genetic drift on this chromosome. Some clades of the phylogeny are so specific to particular continents or regions that they have been used to assign population-of-origin to individual DNA samples,<sup>7</sup> and in quantifying the origins of the components of admixed populations using simple allele-counting methods.<sup>8–10</sup>

Studies of British genetic diversity, generally sampling on the criterion of two generations of residence, have found no evidence of African Y-chromosomal lineages,<sup>11–14</sup> suggesting that they either never became assimilated into the general population or have been lost by drift. However, here, we describe a globally rare and archetypically African sublineage in Britain and show that it has been resident there for at least 250 years, representing the first genetic trace of an appreciable African presence that has existed for several centuries.<sup>2</sup>

## Materials and methods

### DNA samples

British and US males were recruited under appropriate informed consent, and DNA extracted from buccal samples as described.<sup>15</sup> R-surnamed British males were sampled randomly using information from electoral rolls, and a specific questionnaire was used to exclude close patrilineal relatives.

### Binary marker typing

A set of 11 binary markers (M9, M69, M89, M145, M170, M172, M173, M201, P25, SRY<sub>10831</sub> and 12f2) was typed using the SNaPshot minisequencing procedure (Applied Biosystems) as described,<sup>15</sup> and the additional hgA-specific markers M91 and M33 typed by DNA sequencing using published primers.<sup>5</sup>

### Y-STR. typing

A set of 17 Y-chromosomal short tandem repeats (Y-STRs: DYS19, DYS388, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS434, DYS435, DYS436, DYS437, DYS438, DYS439, DYS460, DYS461, DYS462) were typed as described.<sup>16</sup> HgA1 chromosomes were typed with an additional 51 Y-STRs (DYF386S1, DYF390S1, DYS406S1, DYS472, DYS476, DYS480, DYS481, DYS485, DYS487, DYS488, DYS490, DYS491, DYS492, DYS494, DYS495, DYS497, DYS505, DYS508, DYS511, DYS525, DYS530, DYS531, DYS533, DYS537, DYS540, DYS549, DYS554, DYS556, DYS565, DYS567, DYS568, DYS569, DYS570, DYS572, DYS573, DYS575, DYS576, DYS578, DYS579, DYS580, DYS583, DYS589, DYS590, DYS594, DYS617, DYS618, DYS636, DYS638, DYS640, DYS641 and DYS643) according to Lim *et al.*<sup>17</sup> and with an additional nine

(DYS385a/b, DYS425, DYS426, DYS447, DYS448, Y-GATA-H4.1 and YCAIIa/b) according to Parkin *et al.*<sup>18</sup>

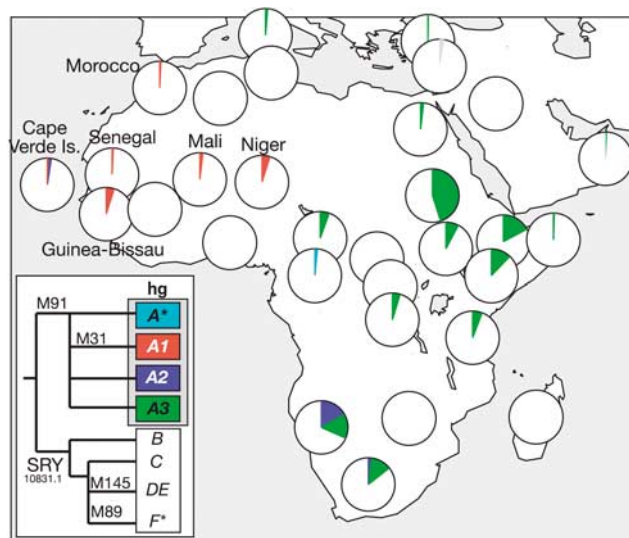
### Network analysis and dating

Median joining networks<sup>19</sup> were constructed within the program Network 4.1.0.9 ([www.fluxus-engineering.com/sharenet.htm](http://www.fluxus-engineering.com/sharenet.htm)) using variance-based weighting as described.<sup>20</sup> Weighting for hgA1 was based on variance in all hgA1 chromosomes, whereas weighting for the R surname network was based on variance observed in 291 British hgR1b chromosomes.

Time-to-most-recent-common-ancestor (TMRCA) was estimated within Network from the  $\rho$ -statistic, using a 35-year generation time,<sup>15</sup> and a mean per-locus, per-generation mutation rate of  $2 \times 10^{-3}$ .<sup>21–23</sup>

## Results

As part of a survey of British Y-chromosome diversity, we recruited a set of 421 males, who described themselves as British, and whose paternal grandfathers were born in Britain. The Y chromosomes of these males were typed using a set of 11 binary markers,<sup>15</sup> including M145 (defining superhaplogroup DE) and M89 (defining superhaplogroup F). All chromosomes carried the derived allele at one or other of these two markers, with a single exception, in male GB1757, which could in principle belong to hgA, B or C (see phylogeny in Figure 1). Further testing, including the markers M91 and M31, gave the



**Figure 1** Distribution of Y chromosomes belonging to hgA. Distribution of hgA among 4516 Y chromosomes in and around Africa, taken from the literature (see text). The colored sector of each pie chart is proportional to the frequency of subgroups of hgA, defined according to the phylogeny lower left. The gray sector in Cyprus indicates hgA chromosomes that were not further subclassified.<sup>32</sup> The phylogeny includes selected markers referred to in the text.

surprising result that it belonged to hgA, within the sublineage A1.

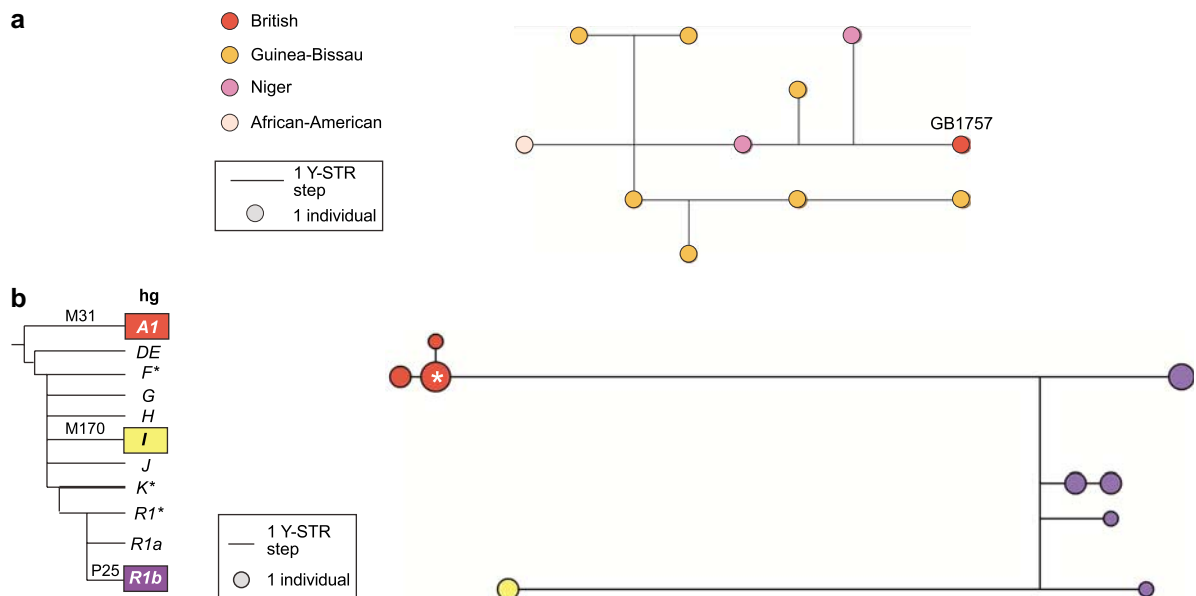
HgA is the deepest-rooting clade of the Y phylogeny, and shows a particularly specific localization to the African continent (Figure 1), which is compatible with an African origin for modern human Y chromosomes. It constitutes 5.4% of a composite sample of 3551 Africans,<sup>4,8,10,24–30</sup> whereas in non-African indigenous populations only seven cases have been described, from Turkey,<sup>31</sup> Cyprus,<sup>32</sup> Sardinia<sup>33,34</sup> and Oman.<sup>35</sup> Extensive surveys of Western European populations have failed to find any examples of these chromosomes.<sup>13,36–39</sup> The subhaplogroup A1 was first reported in a single individual among a sample of 44 males from Mali.<sup>4</sup> Subsequently, this scarce Western African hg has been found in only 25 more males (Figure 1): 2/64 Moroccan Berbers,<sup>24</sup> 3/766 African Americans,<sup>40,41</sup> 2/39 Mandinka from Gambia/Senegal, 1/55 Malian Dogon,<sup>30</sup> 1/201 Cape Verde Islanders, 14/276 males from Guinea-Bissau,<sup>10</sup> and 2/39 males from Niger (F Cruciani and R Scozzari, unpublished data).

The British male carrying the hgA1 chromosome knew of no familial African connection, and he displays a typical European appearance. To investigate the relationship of his Y chromosome with African examples, we compared its 10-locus Y-STR haplotype (see Supplementary Information) with those of 10 other available hgA1 chromosomes,<sup>10,24,40</sup> (F Cruciani and R Scozzari, unpublished data). Figure 2a shows a median-joining network of these haplotypes: they

are diverse, and all are unique. Although the British haplotype is peripheral, it lies equidistant (four mutational steps) from Niger and Guinea-Bissau haplotypes, and similar distances (2–4 steps) exist between other haplotypes in the network. This is compatible with a Western African origin for the British chromosome, but does not point to a particular population. Using the British haplotype (11 loci) to search the Y Chromosome Haplotype Reference Database (<http://www.yhrd.org>) finds no matches among 15 815 chromosomes worldwide, emphasizing its rarity. Also, when the haplotypes of the other hgA1 chromosomes are used in similar searches, they find only self-matches in the populations from which they derive, underlining the scarcity and African-specificity of hgA1.

How long has this archetypically African Y chromosome been in England? To address this question, our strategy was to seek patrilineally related individuals who would share the hg, but whose Y-STR haplotype diversity could be used to estimate a TMRCA. To do this, we exploited the relationship between surnames and Y chromosome haplotypes,<sup>15,42–44</sup> noting that the upper bound of any estimated age would be limited by the fact that hereditary English surnames did not exist before the 11th century.<sup>45</sup>

The hgA1-bearing male bears a locative surname, which we refer to here as R, deriving from an East Yorkshire village.<sup>46</sup> Only 121 people carried this name in 1998 (<http://www.spatial-literacy.org/uclnames>), and it still has



**Figure 2** Diversity of Y-STR haplotypes of chromosomes belonging to hgA1 and within the R surname. (a) Relationships of Y-STR haplotypes within hgA1. Weighted median joining network containing the 10-locus Y-STR haplotypes (DYS19, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS389I, DYS389II-I) of eleven hgA1 chromosomes. Circles represent haplotypes, with area proportional to frequency and colored according to population. (b) Relationships of Y-STR haplotypes carried by men with the R surname. Weighted median joining network containing the 17-locus Y-STR haplotypes of 18 men named R. Circles represent haplotypes, with area proportional to frequency and colored according to hg, as shown in the phylogeny to the left. The node used as a root for dating hgA1 chromosomes is indicated with an asterisk.

a strong east Yorkshire focus. We recruited 18 apparently unrelated men carrying this name (or a close variant spelling, carried by 50 individuals) and typed a set of 11 binary markers and 17 Y-STRs,<sup>15</sup> supplemented with the binary marker M31, allowing us to identify hgA1.

Figure 2b shows a median-joining network of 17-locus Y-STR haplotypes (see Supplementary Information) of Y chromosomes carried by 18 R-surnamed males. The chromosomes belong to three hgs, and include four clusters, indicating either multiple foundation or historical non-paternity within the name. However, a total of seven of the males carry hgA1 chromosomes, belonging to three closely related Y-STR haplotypes, and based on the  $\rho$ -statistic within Network, having a TMRCA of  $440 \pm 330$  years.

As an empirical adjunct to TMRCA calculations, we undertook extensive genealogical research to ask if the seven R-surnamed males carrying hgA1 chromosomes could be connected into a single genealogy with a historically verifiable MRCA. This research resolved the males into two well-supported genealogies (Figure 3a), with MRCAs born in 1788 and 1789, respectively. However, although both of these ancestors were resident in Yorkshire, evidence could not be found for a familial relationship between them. Patterns of forename usage in the two genealogies are quite distinct, which argues against a very recent connection. We recruited 12 unrelated R-surnamed men from the USA, hoping that the presence of hgA1 would indicate that the chromosome had been associated with the surname before emigration from Britain. However, none of these men carried a chromosome from this hg (data not shown), so the approach was uninformative.

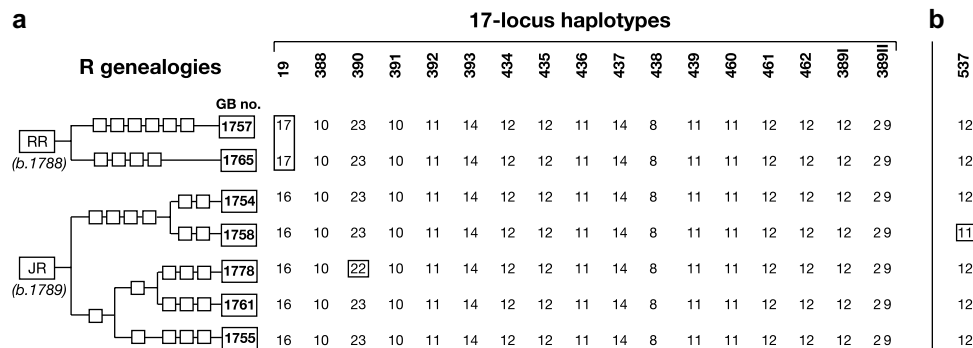
Finally, we exploited a new resource of multiple novel Y-STRs<sup>47</sup> in an attempt to refine TMRCA for the two genealogies. We typed the hgA1 chromosomes with an additional 60 Y-STRs (see Materials and methods), bringing the total to 77. Surprisingly, this analysis revealed only one new mutation – a single repeat decrease at Y-STR DYS537 in

male GB1758 (Figure 3b). Applying the  $\rho$ -statistic within Network, as described above, to the 73-locus Y-STR haplotypes (excluding the bilocal DYS385a/b and YCAIIa/b; see Supplementary Information) in the two genealogical clusters yields a TMRCA of  $140 \pm 80$  years, which, adding the average age of the living individuals (52 years), equates to a likely oldest date of 1734 for the coalescence of the two genealogies. The TMRCA range overlaps that obtained using 17 markers (see above), and suggests that only a small number of generations separates the two genealogies from their common ancestor.

### Discussion

Our study shows that a globally rare Y-chromosome type, belonging to the deepest-rooting African branch of the Y-phylogeny, has been present in Northern England since at least the mid-18th century. HgE3a is by far the most frequent Y-chromosomal lineage in Africa, existing at 48% in a continent-wide sample of 1122 chromosomes,<sup>30</sup> so we would expect any substantial past immigration from Africa to Britain to have left examples of chromosomes belonging to this common hg. However, a survey of 1772 Y chromosomes from the British Isles found none,<sup>13</sup> and they are also absent from our control sample of 421 chromosomes. The general rarity of African lineages may reflect a low level of initial introgression, later loss through drift, or sampling bias – for example, the large British survey<sup>13</sup> sampled from small towns, in which the descendants of early British Africans, who were concentrated in cities, may be depleted.

Admixture between populations of African and European origin is often sex-biased, with a greater proportion of the African component of the hybrid population being contributed by females.<sup>48</sup> Assuming an equal number of males and females of African origin migrating to Britain, we might therefore expect mitochondrial DNA to reveal a



**Figure 3** Genealogical relationships and Y-STR haplotypes of hgA1 R-surnamed men. (a) Genealogies of R-surnamed men bearing hgA1 Y chromosomes (left). To the right are their 17-locus Y-STR haplotypes, with differences highlighted by boxes. (b) Additional difference revealed at DYS537; haplotypes based on the other 59 Y-STRs (see text and Supplementary Information) are identical in all males.

stronger signal of African admixture than the Y chromosome. There is little published evidence, but a study of mitochondrial DNA sequence diversity among 100 'white Caucasian' British<sup>49</sup> does contain one haplotype, which represents an hg L1c sequence (defined according to Salas *et al.*<sup>50</sup>), with a probable origin in West Central Africa.<sup>51</sup> This could represent a possible maternal counterpart to the Y-lineage we describe here.

The presence of a very rare type of Y chromosome in two distinct branches of a genealogy that coalesce before the late 18th century demonstrates clearly that the chromosome must have been introduced before this time. However, the upper limit on the time of linkage with the R surname is less certain. In principle, the association could have been formed many generations earlier, with the descendants of any earlier-branching lineages having drifted to extinction, or not sampled in our study. As the pattern of Y-diversity in the surname suggests either multiple origins for the name, or non-paternities introducing diverse Y chromosomes, the hgA1 lineage need not have been a founding type at the time of surname establishment (about 700 years ago<sup>45</sup>). The contributor of the chromosome to the surname may himself have been a first-generation immigrant African, or an admixed and phenotypically European man carrying an African Y chromosome introduced into Britain some time earlier.

The remarkable Y chromosome present in the R surname provides the first genetic evidence of a long-lived African presence within Britain. It represents a cautionary tale for those who would predict population-of-origin from a Y hg,<sup>7</sup> and emphasizes the complex nature of the history of human migration.

### Acknowledgements

We thank all DNA donors, John Butler for access to samples, and anonymous reviewers for helpful comments. TEK was supported by a Wellcome Prize Studentship (061129), MAJ by a Wellcome Trust Senior Fellowship in Basic Biomedical Science (057559), and YX and CTS by the Wellcome Trust. EJP was supported by the Arts and Humanities Research Board and the EC Sixth Framework Program under Contract No. ERAS-CT-2003-980409, as part of the European Science Foundation EUROCORES Program OMLL.

### References

- Davies N: *The Isles – a history*. Basingstoke: Papermac, 1999.
- Fryer P: *Staying power: the history of black people in Britain*. London: Pluto Press, 1984.
- Ramdin R: *Reimagining Britain: 500 years of Black and Asian History*. London: Pluto Press, 1999.
- Underhill PA, Shen P, Lin AA *et al*: Y chromosome sequence variation and the history of human populations. *Nat Genet* 2000; **26**: 358–361.
- Y Chromosome Consortium: A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Res* 2002; **12**: 339–348.
- Jobling MA, Tyler-Smith C: The human Y chromosome: an evolutionary marker comes of age. *Nat Rev Genet* 2003; **4**: 598–612.
- Brion M, Sanchez JJ, Balogh K *et al*: Introduction of a single nucleotide polymorphism-based 'major Y-chromosome haplogroup typing kit' suitable for predicting the geographical origin of male lineages. *Electrophoresis* 2005; **26**: 4411–4420.
- Bosch E, Calafell F, Comas D, Oefner PJ, Underhill PA, Bertranpetit J: High-resolution analysis of human Y-chromosome variation shows a sharp discontinuity and limited gene flow between Northwestern Africa and the Iberian Peninsula. *Am J Hum Genet* 2001; **68**: 1019–1029.
- Bosch E, Calafell F, Rosser ZH *et al*: High level of male-biased Scandinavian admixture in Greenlandic Inuit shown by Y-chromosomal analysis. *Hum Genet* 2003; **112**: 353–363.
- Gonçalves R, Rosa A, Freitas A *et al*: Y-chromosome lineages in Cabo Verde Islands witness the diverse geographic origin of its first male settlers. *Hum Genet* 2003; **113**: 467–472.
- Wilson JF, Weiss DA, Richards M, Thomas MG, Bradman N, Goldstein DB: Genetic evidence for different male and female roles during cultural transitions in the British Isles. *Proc Natl Acad Sci USA* 2001; **98**: 5078–5083.
- Weale ME, Weiss DA, Jager RF, Bradman N, Thomas MG: Y chromosome evidence for Anglo-Saxon mass migration. *Mol Biol Evol* 2002; **19**: 1008–1021.
- Capelli C, Redhead N, Abernethy JK *et al*: A Y chromosome census of the British Isles. *Curr Biol* 2003; **13**: 979–984.
- Goodacre S, Helgason A, Nicholson J *et al*: Genetic evidence for a family-based Scandinavian settlement of Shetland and Orkney during the Viking periods. *Heredity* 2005; **95**: 129–135.
- King TE, Ballereau SJ, Schürer K, Jobling MA: Genetic signatures of coancestry within surnames. *Curr Biol* 2006; **16**: 384–388.
- Bosch E, Lee AC, Calafell F *et al*: High resolution Y chromosome typing: 19 STRs amplified in three multiplex reactions. *Forensic Sci Int* 2002; **125**: 42–51.
- Lim S, Xue X, Parkin EJ, Tyler-Smith C: Variation of 52 new Y-STR loci in the Y Chromosome Consortium worldwide panel of 76 diverse individuals. *Int J Legal Med* 2006, in press, doi 10.1007/s00414-006-0124-8.
- Parkin EJ, Kraayenbrink T, van Driem GL, Tshering K, de Knijff P, Jobling MA: 26-locus Y-STR typing in a Bhutanese population sample. *Forensic Sci Int* 2006; **161**: 1–7.
- Bandelt H-J, Forster P, Röhl A: Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 1999; **16**: 37–48.
- Qamar R, Ayub Q, Mohyuddin A *et al*: Y-chromosomal DNA variation in Pakistan. *Am J Hum Genet* 2002; **70**: 1107–1124.
- Heyer E, Puymirat J, Dieltjes P, Bakker E, de Knijff P: Estimating Y chromosome specific microsatellite mutation frequencies using deep rooting pedigrees. *Hum Mol Genet* 1997; **6**: 799–803.
- Jobling MA, Heyer E, Dieltjes P, de Knijff P: Y-chromosome-specific microsatellite mutation rates re-examined using a minisatellite, MSY1. *Hum Mol Genet* 1999; **8**: 2117–2120.
- Gusmão L, Sanchez-Diz P, Calafell F *et al*: Mutation rates at Y chromosome specific microsatellites. *Hum Mutat* 2005; **26**: 520–528.
- Cruciani F, Santolamazza P, Shen PD *et al*: A back migration from Asia to sub-Saharan Africa is supported by high-resolution analysis of human Y-chromosome haplotypes. *Am J Hum Genet* 2002; **70**: 1197–1214.
- Semino O, Santachiara-Benerecetti AS, Falaschi F, Cavalli-Sforza LL, Underhill PA: Ethiopians and Khoisan share the deepest clades of the human Y-chromosome phylogeny. *Am J Hum Genet* 2002; **70**: 265–268.
- Al-Zahery N, Semino O, Benuzzi G *et al*: Y-chromosome and mtDNA polymorphisms in Iraq, a crossroad of the early human dispersal and of post-Neolithic migrations. *Mol Phylogenet Evol* 2003; **28**: 458–472.
- Arredi B, Poloni ES, Paracchini S *et al*: A predominantly Neolithic origin for Y-chromosomal DNA variation in North Africa. *Am J Hum Genet* 2004; **75**: 338–345.

- 28 Hurles ME, Sykes BC, Jobling MA, Forster P: The dual origin of the Malagasy in island southeast Asia and East Africa: evidence from maternal and paternal lineages. *Am J Hum Genet* 2005; **76**: 894–901.
- 29 Sanchez JJ, Hallenberg C, Borsting C, Hernandez A, Morling N: High frequencies of Y chromosome lineages characterized by E3b1, DYS19-11, DYS392-12 in Somali males. *Eur J Hum Genet* 2005; **13**: 856–866.
- 30 Wood ET, Stover DA, Ehret C *et al*: Contrasting patterns of Y chromosome and mtDNA variation in Africa: evidence for sex-biased demographic processes. *Eur J Hum Genet* 2005; **13**: 867–876.
- 31 Cinnioglu C, King R, Kivisild T *et al*: Excavating Y-chromosome haplotype strata in Anatolia. *Hum Genet* 2004; **114**: 127–148.
- 32 Capelli C, Redhead N, Romano V *et al*: Population structure in the Mediterranean basin: a Y chromosome perspective. *Ann Hum Genet* 2006; **70**: 207–225.
- 33 Scozzari R, Cruciani F, Pangrazio A *et al*: Human Y-chromosome variation in the western Mediterranean area: implications for the peopling of the region. *Hum Immunol* 2001; **62**: 871–884.
- 34 Francalacci P, Morelli L, Underhill PA *et al*: Peopling of three Mediterranean islands (Corsica, Sardinia, and Sicily) inferred by Y-chromosome biallelic variability. *Am J Phys Anthropol* 2003; **121**: 270–279.
- 35 Luis JR, Rowold DJ, Regueiro M *et al*: The Levant versus the Horn of Africa: evidence for bidirectional corridors of human migrations. *Am J Hum Genet* 2004; **74**: 532–544.
- 36 Semino O, Passarino G, Oefner PJ *et al*: The genetic legacy of Paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective. *Science* 2000; **290**: 1155–1159.
- 37 Rosser ZH, Zerjal T, Hurles ME *et al*: Y-chromosomal diversity within Europe is clinal and influenced primarily by geography, rather than by language. *Am J Hum Genet* 2000; **67**: 1526–1543.
- 38 Beza S, Gusmão L, Lopes A *et al*: Micro-phylogeographic and demographic history of Portuguese male lineages. *Ann Hum Genet* 2006; **70**: 181–194.
- 39 Flores C, Maca-Meyer N, Gonzalez AM *et al*: Reduced genetic structure of the Iberian peninsula revealed by Y-chromosome analysis: implications for population demography. *Eur J Hum Genet* 2004; **12**: 855–863.
- 40 Vallone PM, Butler JM: Y-SNP typing of US African American and Caucasian samples using allele-specific hybridization and primer extension. *J Forensic Sci* 2004; **49**: 723–732.
- 41 Hammer MF, Chamberlain VF, Kearney VF *et al*: Population structure of Y chromosome SNP haplogroups in the United States and forensic implications for constructing Y chromosome STR databases. *Forens Sci Int* 2005; **164**: 45–55.
- 42 Sykes B, Irven C: Surnames and the Y chromosome. *Am J Hum Genet* 2000; **66**: 1417–1419.
- 43 Jobling MA: In the name of the father: surnames and genetics. *Trends Genet* 2001; **17**: 353–357.
- 44 McEvoy B, Bradley DG: Y-chromosomes and the extent of patrilineal ancestry in Irish surnames. *Hum Genet* 2006; **119**: 212–219.
- 45 McKinley RA: *A history of British surnames*. London: Longman, 1990.
- 46 Reaney PH, Wilson RM: *A dictionary of English surnames*. Oxford: Oxford University Press, 1997.
- 47 Kayser M, Kittler R, Erler A *et al*: A comprehensive survey of human Y-chromosomal microsatellites. *Am J Hum Genet* 2004; **74**: 1183–1197.
- 48 Jobling MA, Hurles ME, Tyler-Smith C: *Human evolutionary genetics: origins, peoples and disease*. New York/Abingdon: Garland Science, 2003, pp 386–388.
- 49 Piercy R, Sullivan KM, Benson N, Gill P: The application of mitochondrial DNA typing to the study of white Caucasian genetic identification. *Int J Legal Med* 1993; **106**: 85–90.
- 50 Salas A, Richards M, Lareu MV *et al*: The African diaspora: mitochondrial DNA and the Atlantic slave trade. *Am J Hum Genet* 2004; **74**: 454–465.
- 51 Coia V, Destro-Bisol G, Verginelli F *et al*: Brief communication: mtDNA variation in North Cameroon: lack of Asian lineages and implications for back migration from Asia to sub-Saharan Africa. *Am J Phys Anthropol* 2005; **128**: 678–681.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (<http://www.nature.com/ejhg>)