

Molecular Dissection of the Basal Clades in the Human Y Chromosome Phylogenetic Tree

Rosaria Scozzari¹, Andrea Massaia¹, Eugenia D'Atanasio¹, Natalie M. Myres^{2,3}, Ugo A. Perego^{2,4}, Beniamino Trombetta¹, Fulvio Cruciani^{1*}

1 Dipartimento di Biologia e Biotechnologie "Charles Darwin", Sapienza Università di Roma, Rome, Italy, **2** Sorenson Molecular Genealogy Foundation, Salt Lake City, Utah, United States of America, **3** AncestryDNA, Provo, Utah, United States of America, **4** Dipartimento di Biologia e Biotechnologie "Lazzaro Spallanzani", Università di Pavia, Pavia, Italy

Abstract

One hundred and forty-six previously detected mutations were more precisely positioned in the human Y chromosome phylogeny by the analysis of 51 representative Y chromosome haplogroups and the use of 59 mutations from literature. Twenty-two new mutations were also described and incorporated in the revised phylogeny. This analysis made it possible to identify new haplogroups and to resolve a deep trifurcation within haplogroup B2. Our data provide a highly resolved branching in the African-specific portion of the Y tree and support the hypothesis of an origin in the north-western quadrant of the African continent for the human MSY diversity.

Citation: Scozzari R, Massaia A, D'Atanasio E, Myres NM, Perego UA, et al. (2012) Molecular Dissection of the Basal Clades in the Human Y Chromosome Phylogenetic Tree. PLoS ONE 7(11): e49170. doi:10.1371/journal.pone.0049170

Editor: David Caramelli, University of Florence, Italy

Received: July 4, 2012; **Accepted:** October 4, 2012; **Published:** November 7, 2012

Copyright: © 2012 Scozzari et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the Italian Ministry of Education, University and Research (grant "Progetti di Ricerca di Interesse Nazionale", n. 2009P2CNKK). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: NMM is affiliated to AncestryDNA based in Utah. NMM and UAP are affiliated to Sorenson Molecular Genealogy Foundation. AncestryDNA sell kits to determine ancestry (<http://dna.ancestry.com/>). Sorenson Molecular Genealogy Foundation sold their database to Ancestry.com this year: <http://corporate.ancestry.com/press/press-releases/2012/05/ancestry.com-dna-launches/>. There are no other patents, products in development or marketed products to declare. This does not alter the authors' adherence to all the PLoS ONE policies on sharing data and materials.

* E-mail: fulvio.cruciani@uniroma1.it

Introduction

The male specific portion of the Y chromosome (MSY) is the largest haplotypic block in the human genome. Due to the lack of allelic recombination, the history of mutational accumulation in the MSY can be described in a unique evolutionary haplogroup tree. The MSY tree has been largely used for the study of the human evolutionary history [1,2], and is currently becoming an increasingly important resource in forensic [3,4] and sex-chromosome evolution studies [5–13].

The number of phylogenetically characterized Y-specific markers has substantially increased in recent years. In the MSY phylogeny reported by Karafet et al. [14], 80 mutations belong to the deepest portion of the tree (haplogroups A1a, A1b, A2, A3 and B), defining a total of 34 distinct African-specific haplogroups/paragroups. An independent high resolution MSY phylogeny has been recently obtained from 2,870 Y-SNPs discovered (or re-discovered) in the course of a large whole-genome re-sequencing study, but the observed variable sites all belong to the recent "out of Africa" CT clade [15]. Recently, in a re-sequencing study of the Y chromosome, the root of the tree moved to a new position and several changes at the basal nodes of the phylogeny were introduced [16]. Interestingly, the estimated coalescence age and deep branching pattern of the revised MSY tree appear to be more similar to those of the mtDNA phylogeny [17,18] than previously reported [1].

To increase the level of phylogenetic resolution in the African-specific portion of the MSY tree, we mapped 146 previously detected mutations [16] with respect to 59 mutations defining

internal nodes in the Y chromosome phylogeny reported by Karafet et al. [14]. We also mapped eleven previously unreported mutations from our database, and nine new mutations that had been identified through the re-sequencing of two Y chromosomes belonging to the widespread African-specific haplogroup A3b. Finally, two additional mutations identified during the process of phylogenetic mapping were also included in the analysis. The new integrated MSY tree represents an improved resource for the study of human prehistory and genome evolution.

Results

Newly Described Mutations

Information regarding 22 previously undescribed mutations is given in Table 1. The first set of eleven mutations came from our database (mutations in the range V2-V249 in Table 1). Nine new mutations (mutations in the range of V262-V317 in Table 1) were discovered by sequencing two A3b subjects, while two additional mutations (V254 and V341) were identified during the process of phylogenetic mapping. Apart from one exception (V306, corresponding to the SNP rs113042298), none of these mutations is present in the current SNP database (dbSNP, build 135).

Phylogenetic Mapping

Most of the mutations here analyzed belong to the African portion of the MSY phylogeny, which is comprised of haplogroups A1b, A1a, A2, A3 and B [16]. Through phylogenetic mapping it was possible to identify 15 new African haplogroups and to resolve

Table 1. Information on the mutations here reported and analyzed for the first time.

SNP	Y-Position ^a	Mutation	Forward Primer	Reverse Primer
V2	6778215	A to C	TTGCTGAGTGTACTGGGATCTT	AAGACACCCAGTTGTCATTTTC
V3	6778229	T to C	TTGCTGAGTGTACTGGGATCTT	AAGACACCCAGTTGTCATTTTC
V11	6892902	A to T	CTGGCCTTAGGAAATAGGTCAA	AGGGTCATTATCGTTGAGGAAG
V33	6894717	G to T	AACACCAGGATTTGTTTTGGAG	TTCAGCAACTGTAGAATCAGG
V34	6894718	C to T	AACACCAGGATTTGTTTTGGAG	TTCAGCAACTGTAGAATCAGG
V37	6818279	G to A	ATTATTTGAATGCAAGTGGGGA	TGAAACTGAATTAAGGAAGGTGG
V73	16691696	A to G	TGAAGGAGATTAATTGGGGTAGA	TTCCATACCATGCCTTCTTT
V87	17947454	A to T	CATCCCTGGTCATCCCTC	CTAGACCCTGATTCATGTAAGCC
V147	6739492	G to A	GTTTTGTGGGTGAGAGAGGAAG	CCATTTCTACATGGAGGAAGTTTT
V248	7589991	C to T	GGCAGCCCTCAGGATATGTA	GATTCCACTAAACCCGACGA
V249 ^b	25207704; 26841450; 27120952	T/T/T to G/G/G	GGCCAAATCAGAGAAATGGA	CGCAGAACCTGAAATTGTGA
V254	6870497	G to A	AAATGCAGTGTCCAGGGAGT	TTAAGTAGTCCCAGAAAGTTAAAG
V262	6659209	C to G	CTACAACGCCAGCTGATTT	CTGTCTTATGGCCACCCAAG
V265	6661164	A to G	CTCTAGCAATTAGGGCTTCAG	TGTTGCCTAGATGACAAGCA
V303	2798066 - 2798068	del TTT	GGCACCCCTGTAGAACCATA	GAAAAAGAGCAGGCATGGTG
V304	2796955 - 2796957	del GAA	GCTTTGGGGGAGTTAGGAAA	TACAGGGTGCCAGATGGACT
V305	2854573	T to C	ACCCCTGGGTTCAACTATCC	ATTACTCAGGGTGGTGACG
V306	7594967	G to C	CCTTAAAGGGCTCGGAGAGA	CACTCGACATCGACCTCTCA
V313	7622390	G to A	GAGACTCAGGCAGGCATCAT	TCCAAAGTTTACAGCTGATT
V314	7642949	A to T	AGGCCTCATCCAGACCTATG	CCACCATGCTGGCTTATTTT
V317	2908553	G to A	AAGATGCCAGCCTCGAGTTA	TTTTTGATCTGAGGCCATC
V341	4840884	A to T	TTCAGCTATTGCCTCTATGG	TTAAGCTCGTGGGAAATGC

^aPosition according to the February 2009 human Y-chromosome reference sequence (GRCh37).

^bPCR primers amplify three paralogous MSY regions.

doi:10.1371/journal.pone.0049170.t001

one basal trifurcation (Figure 1). A new deep branch within the “out of Africa” haplogroup C was also identified (Figure S1).

Haplogroup A1b. The P114 mutation, which defines haplogroup A1b according to Karafet et al. [14], had been detected in central-western Africa at very low frequencies (in total, three chromosomes from Cameroon) [16,19]. Thirty-two additional biallelic mutations have been recently discovered for haplogroup A1b, which now appears to be one of the two deepest-rooting branches of the MSY tree [16]. Here, these mutations were analyzed in three Y chromosomes that were ancestral at both P114 and the markers that, according to [16], define macro-haplogroup A1a-T (samples 1–3 in Table S1) and in one Y chromosome derived at P114 (sample 4 in Table S1). The four chromosomes turned out to carry the derived allele at 19 of the 32 markers analyzed, resulting in one terminal haplogroup, defined by P114 and three other mutations, and three new paragroups (Figure 1). We identified a third allele (A) for the V161 polymorphism, which had been previously reported as a biallelic G to C transversion on the A1b branch [16]. The presence of nucleotide A at the orthologous MSY position in the chimp reference sequence (October 2010 chimp assembly, UCSC Genome Browser), along with the structure of the human MSY tree as shown in Figure 1, suggest that this triallelic polymorphism may have originated from two independent mutations, from A to C (V161.1) within the A1b branch, and from A to G (V161.2) at the root of macro-haplogroup A1a-T.

Haplogroup A1a. Haplogroup A1a is present at relatively low frequencies in western and northern Africa [20–25]. In the

phylogeny by Karafet et al. [14], A1a is defined by mutations M31 and P82. Twenty additional mutations have been reported for this branch [16]. These mutations, along with M31, P82 and a mutation identified within the P82 amplicon (V147), were analyzed in two A-M31 chromosomes (samples 5 and 6 in Table S1). This analysis led to the splitting of A1a into two branches that shared all but four of the 23 mutations analyzed (Figure 1).

Haplogroup A2. Haplogroup A2 was first reported as a clade limited to Khoisan-speaking populations in southern Africa [26]. Recently, a new deep A2 paragon was found among western Pygmies from central Africa by Batini et al. [19]. In the tree by Karafet et al. [14], haplogroup A2 is defined by P3 and other 17 phylogenetically equivalent markers, with three mutations that define three terminal branches (A-M114, A-P28 and A-P262). Here, 19 mutations that had been identified for haplogroup A2 [16], as well as the markers P3, M114, P28, P262, were genotyped in three A2 Y chromosomes (samples 7–9 in Table S1). This analysis made it possible to identify a new branch (A-V218) within haplogroup A-P262. The newly reported V73 mutation was found to be phylogenetically equivalent to M114 (Figure 1).

Haplogroup A3. In the tree by Karafet et al. [14], haplogroup A3 is defined by the M32 mutation and contains two African-specific clades, A3a and A3b. A3a is a rare branch that has been observed exclusively in eastern Africa [19–21]. The sister clade, A3b, is further subdivided into a southern African (A3b1 or A-M51) and a mainly eastern African (A3b2 or A-M13) haplogroup [19–21,23,27–30]. Since both A3b1 and A3b2 are quite common, but are yet poorly resolved haplogroups, we

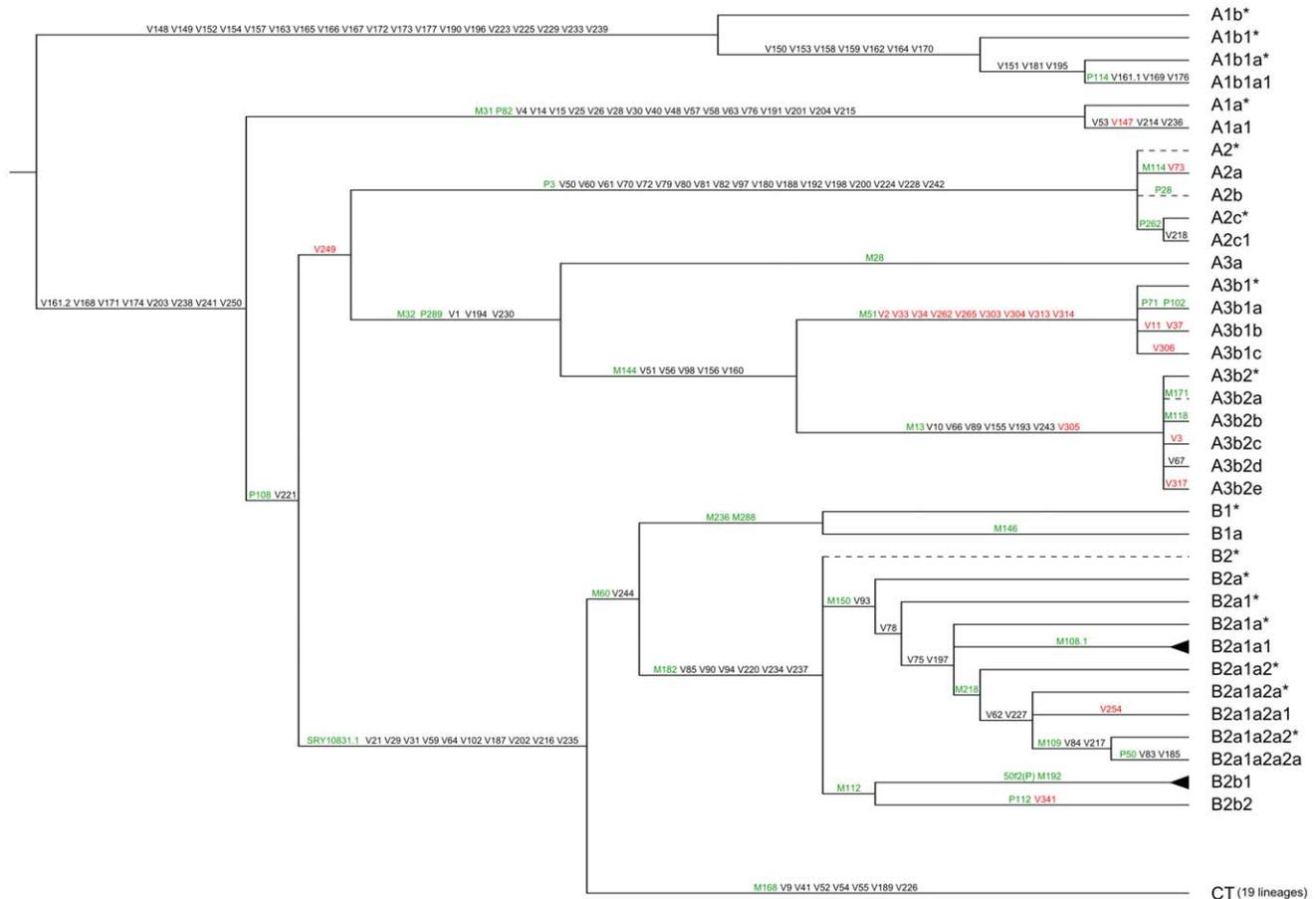


Figure 1. Revised topology of the deepest portion of the human MSY tree. The names of the mutations genotyped are indicated on the branches (green, mutations from the paper by Karafet et al. [14]; black, mutations from the paper by Cruciani et al. [16]; red, previously undescribed mutations, see text). For the sake of clarity, the internal structure of haplogroups B-M108.1 (2 branches) and B-50f2(P) (8 branches) is not shown (black triangles). The phylogenetic position of mutations mapping within haplogroup CT is shown in Figure S1. Dashed lines indicate putative branchings (no positive control available). The microsatellite intermediate allele *DYS449.2*, that was found to delineate new phylogenetic structure in human Y chromosome haplogroup tree [42], was not observed in 19 Y*(xBT) and 4 B chromosomes analyzed. doi:10.1371/journal.pone.0049170.g001

performed a MSY re-sequencing analysis of about 90 kb for each of two A3b chromosomes (one A-M51* and one A-M13*) to find additional informative markers. We detected a total of 9 new mutations (V262–V317 in Table 1). A total of 41 markers (9 new mutations, 7 mutations from our database, 15 mutations identified in a previous study [16], and 10 mutations defining A3 branches in Karafet et al. [14]) were analyzed in ten subjects (samples 10–19 in Table S1). Mutation V249 joined haplogroups A2 and A3 and should be considered to be phylogenetically equivalent to the PK1 marker, which was originally thought to be only associated with the A2 lineage, but which has recently been found to cluster haplogroups A2 and A3 [19]. The phylogenetic mapping of the other mutations led to the identification of five new haplogroups, doubling the number of both A3b1 and A3b2 terminal branches. Finally, the P289 marker [14] was positioned upstream of both A3a and A3b (Figure 1).

Haplogroup B. Haplogroup B (B-M60 [14,20]) is essentially a sub-Saharan clade. The majority of haplogroup B lineages show well defined and restricted geographic and ethnic distributions. An exception is represented by haplogroup B-M109, which is present in central, eastern, and southern Africa [19–21,23,27,29,30]. Extensive sequencing of one B-M109 chromosome recently led to the identification of 17 mutations for this haplogroup [16]. Here, a

total of 33 mutations (17 previously described mutations [16], 2 mutations - V254 and V341 - identified during the mapping process, and 14 mutations defining haplogroup B branches in Karafet et al. [14]) were analyzed in 13 haplogroup B chromosomes (Table S1, samples 20–32). We substantially increased the resolution of the B2a clade (B-M150), with five new branches detected (Figure 1). The trifurcation B2a/B2b/B2c within the major clade B2 reported by Karafet et al. [14] was resolved by repositioning the M112 mutation (Figure 1).

“Out of Africa” haplogroups. All Y-clades that are not exclusively African belong to the macro-haplogroup CT, which is defined by mutations M168, M294 and P9.1 [14,31] and is subdivided into two major clades, DE and CF [1,14]. In a recent study [16], sequencing of two chromosomes belonging to haplogroups C and R, led to the identification of 25 new mutations, eleven of which were in the C-chromosome and seven in the R-chromosome. Here, the seven mutations which were found to be shared by chromosomes of haplogroups C and R [16], were also found to be present in one DE sample (sample 33 in Table S1), and positioned at the root of macro-haplogroup CT (Figure 1 and Figure S1). Six haplogroup C chromosomes (samples 34–39 in Table S1) were analyzed for the eleven haplogroup C-specific mutations [16] and for SNPs defining branches C1 to C6

in the tree by Karafet et al. [14] (Figure S1). Through this analysis we identified a chromosome from southern Europe as a new deep branch within haplogroup C (C-V20 or C7, Figure S1). Previously, only a few examples of C chromosomes (only defined by the marker RPS4Y₇₁₁) had been found in southern Europe [32,33]. To improve our knowledge regarding the distribution of haplogroup C in Europe, we surveyed 1965 European subjects for the mutation RPS4Y₇₁₁ and identified one additional haplogroup C chromosome from southern Europe, which has also been classified as C7 (data not shown). Further studies are needed to establish whether C7 chromosomes are the relics of an ancient European gene pool or the signal of a recent geographical spread from Asia. Two mutations, V248 and V87, which had never been previously described, were found to be specific to haplogroups C2 and C3, respectively (Figure S1). Three of the seven R-specific mutations (V45, V69 and V88) were previously mapped within haplogroup R [34], whereas the remaining four mutations have been here positioned at the root of haplogroups F (V186 and V205), K (V104) and P (V231) (Figure S1) through the analysis of 12 haplogroup F samples (samples 40–51, in Table S1).

Discussion

Here we report the mapping of 227 Y chromosome mutations mainly belonging to the African specific clades, greatly increasing the level of resolution for the deepest portion of the tree.

The structure of haplogroup A3b has been greatly refined. Based on the relative number of mutations accumulated, both the southern African A3b1 (A-M51) and the eastern African A3b2 (A-M13) clades seem to have experienced relatively recent expansions resulting in multiple terminal short sister clades. As for haplogroup B, a recent analysis of the microsatellite diversity of the B2a (B-M150) clade had shown a relatively deep coalescence for this haplogroup, suggesting a scenario that predates the diffusion of the Bantu languages [19]. Here, the fine dissection of the B2a clade, with five new branches identified, opens up the possibility to better understand the phylogeographic pattern associated with this widespread and still poorly studied sub-Saharan clade. Finally, the basal A1b haplogroup, an extremely rare haplogroup [16,19], appears to be subdivided into a number of deep branches (Figure 1), a finding that highlights the importance of targeted studies of rare haplogroups in phylogenetic analyses. This haplogroup represents one of the two deepest-rooting branches of the most recently reported MSY tree [16]. New whole-chromosome African sequences are likely to provide a finer resolution of the basal portion of the Y chromosome phylogeny and, possibly, more deep-rooting Y chromosome branches.

Two A1b chromosomes from a previous work (one from Algeria and one from Cameroon) [16] were included in this study together with two newly identified A1b chromosomes, whose geographic origin can be traced back to west-central Africa (Ghana) on the basis of the microsatellite profile (data not shown). It is worth noting that three additional A1b chromosomes have been recently found in Caribbean populations, which exhibit substantial Y-STR haplotype sharing with Y chromosomes from Gabon [35,36]. Taken together, all these data reinforce the hypothesis of an origin in the north-western quadrant of the African continent for the A1b haplogroup [16], and, together with recent findings of ancient Y-lineages in central-western Africa [19], provide new evidence regarding the geographical origin of human MSY diversity.

Different geographical regions in the African continent have been proposed for the origin of different portions of the human genome [18,37,38]. Current whole-genome sequencing projects [e.g. 15,39] and new methods to analyze ancient DNA samples

[40] are likely to provide a more coherent picture of the demographic history and origin of our species.

Materials and Methods

Ethics Statement

This study is part of a research project approved by the “Policlinico Umberto I, Sapienza Università di Roma” Ethical Committee (protocol number 1016/10, according to the DM 15/7/1997 and following). The data were analyzed anonymously.

DNA Samples

Samples were chosen on the basis of their SNP/microsatellite genotype as determined in the present or previous studies [21,41]. A total of 51 DNA samples from the collections of the Authors were analyzed. Haplogroup information and country of origin of each sample is reported in Table S1. Haplogroup nomenclature follows the same criterion as reported in [16].

Identification of New Mutations by Sequencing

Overall, about 90 Kb were sequenced for each of two unrelated Y chromosomes belonging to haplogroups A3b1 (89.8 kb) and A3b2 (89.3 kb).

We designed polymerase chain reaction (PCR) and sequence primers on the basis of the Y-chromosome sequence reported in the February 2009 assembly of the UCSC Genome Browser (<http://genome.ucsc.edu/>) using Primer3 software (<http://frodo.wi.mit.edu/primer3/>). Sequencing templates were obtained through PCR in a 50- μ l reaction containing 50 ng of genomic DNA, 200 μ M each deoxyribonucleotide (dNTP), 2.5 mM MgCl₂, 1 unit of Taq polymerase, and 10 pmoles of each primer. A touch-down PCR program was used with an annealing temperature that decreased from 62°C to 55°C over 14 cycles, followed by 30 cycles with an annealing temperature of 55°C.

Following DNA amplification, PCR products were purified using the QIAquick PCR purification kit (Qiagen, Hilden, Germany). Cycle sequencing was performed using the BigDye Terminator Cycle Sequencing Kit with Amplitaq DNA polymerase (Applied Biosystems, Foster City, CA) and an internal or PCR primer. Cycle sequencing products were purified by ethanol precipitation and run on an ABI Prism 3730XL DNA sequencer (Applied Biosystems). Chromatograms were aligned and analyzed for mutations using Sequencher 4.8 (Gene Codes Corporation, Ann Arbor, MI).

Genotyping and Phylogenetic Mapping

To obtain a refined MSY tree, we determined the allelic state at 168 markers (22 newly described mutations and 146 previously reported mutations [16]) in 51 Y chromosomes which were representative of different Y haplogroups (Table S1). Fifty-nine mutations reported by Karafet et al. [14] were also analyzed (Figure 1 and Figure S1). In order to detect the presence, if any, of the intermediate variant allele DYS449.2 [42], nineteen Y*(xBT) and 4 B chromosomes were typed for the DYS449 microsatellite by sequencing.

Supporting Information

Figure S1 Structure of the macro-haplogroup CT. For details on mutations see legend to Figure 1. Dashed lines indicate putative branchings (no positive control available). The position of V248 (haplogroup C2) and V87 (haplogroup C3) compared to mutations that define internal branches was not determined. Note

that mutations V45, V69 and V88 have been previously mapped (Cruciani et al. 2010; Eur J Hum Genet 18:800–807). (TIF)

Table S1 Haplogroup affiliation for 51 Y chromosomes analyzed in this study. (XLS)

Acknowledgments

We are grateful to all the donors for providing DNA samples.

References

- Underhill PA, Kivisild T (2007) Use of Y chromosome and mitochondrial DNA population structure in tracing human migrations. *Annu Rev Genet* 41: 539–564.
- Chiaroni J, Underhill PA, Cavalli-Sforza LL (2009) Y chromosome diversity, human expansion, drift, and cultural evolution. *Proc Natl Acad Sci U S A* 106: 20174–20179.
- Butler JM (2005) *Forensic DNA typing*. London: Elsevier Academic Press.
- Kayser M (2007) Uni-parental markers in human identity testing including forensic DNA analysis. *BioTechniques* 43: S16–S21.
- Rozen S, Skaletsky H, Marszalek JD, Minx PJ, Cordum HS, et al. (2003) Abundant gene conversion between arms of palindromes in human and ape Y chromosomes. *Nature* 423: 873–876.
- Rozen S, Marszalek JD, Alagappan RK, Skaletsky H, Page DC (2009) Remarkably little variation in proteins encoded by the Y chromosome's single-copy genes, implying effective purifying selection. *Am J Hum Genet* 85: 923–928.
- Bosch E, Hurles ME, Navarro A, Jobling MA (2004) Dynamics of a human interparalog gene conversion hotspot. *Genome Res* 14: 835–844.
- Repping S, Korver CM, Oates RD, Silber S, van der Veen F, et al. (2004) Are sequence family variants useful for identifying deletions in the human Y chromosome? *Am J Hum Genet* 75: 514–517.
- Repping S, van Daalen SKM, Brown LG, Korver CM, Lange J, et al. (2006) High mutation rates have driven extensive structural polymorphism among human Y chromosomes. *Nat Genet* 38: 463–467.
- Jobling MA (2008) Copy number variation on the human Y chromosome. *Cytogenet Genome Res* 123: 253–262.
- Rosser ZH, Balaesque P, Jobling MA (2009) Gene conversion between the X chromosome and the male-specific region of the Y chromosome at a translocation hotspot. *Am J Hum Genet* 85: 130–134.
- Cruciani F, Trombetta B, Macaulay V, Scozzari R (2010) About the X-to-Y gene conversion rate. *Am J Hum Genet* 86: 495–497.
- Trombetta B, Cruciani F, Underhill PA, Sellitto D, Scozzari R (2010) Footprints of X-to-Y gene conversion in recent human evolution. *Mol Biol Evol* 27: 714–725.
- Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, et al. (2008) New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Res* 18: 830–838.
- The 1000 Genomes Project Consortium (2010) A map of human genome variation from population-scale sequencing. *Nature* 467: 1061–1073.
- Cruciani F, Trombetta B, Massaia A, Destro-Bisol G, Sellitto D, et al. (2011) A revised root for the human Y chromosomal phylogenetic tree: the origin of patrilineal diversity in Africa. *Am J Hum Genet* 88: 814–818.
- Mishmar D, Ruiz-Pesini E, Golik P, Macaulay V, Clark AG, et al. (2003) Natural selection shaped regional mtDNA variation in humans. *Proc Natl Acad Sci U S A* 100: 171–176.
- Behar DM, Villemers R, Soodyall H, Blue-Smith J, Pereira L, et al. (2008) The dawn of human matrilineal diversity. *Am J Hum Genet* 82: 1130–1140.
- Batini C, Ferri G, Destro-Bisol G, Brisighelli F, Luiselli D, et al. (2011) Signatures of the preagricultural peopling processes in sub-Saharan Africa as revealed by the phylogeography of early Y chromosome lineages. *Mol Biol Evol* 28: 2603–2613.
- Underhill PA, Shen P, Lin AA, Jin L, Passarino G, et al. (2000) Y chromosome sequence variation and the history of human populations. *Nat Genet* 26: 358–361.
- Cruciani F, Santolamazza P, Shen P, Macaulay V, Moral P, et al. (2002) A back migration from Asia to sub-Saharan Africa is supported by high-resolution analysis of human Y-chromosome haplotypes. *Am J Hum Genet* 70: 1197–1214.
- Gonçalves R, Rosa A, Freitas A, Fernandes A, Kivisild T, et al. (2003) Y-chromosome lineages in Cabo Verde Islands witness the diverse geographic origin of its first male settlers. *Hum Genet* 113: 467–472.
- Wood ET, Stover DA, Ehret C, Destro-Bisol G, Spedini G, et al. (2005) Contrasting patterns of Y chromosome and mtDNA variation in Africa: evidence for sex-biased demographic processes. *Eur J Hum Genet* 13: 867–876.
- King TE, Parkin EJ, Swinfield G, Cruciani F, Scozzari R, et al. (2007) Africans in Yorkshire? The deepest-rooting clade of the Y phylogeny within an English genealogy. *Eur J Hum Genet* 15: 288–293.
- Rosa A, Ornelas C, Jobling MA, Brehm A, Villemers R. (2007) Y-chromosomal diversity in the population of Guinea-Bissau: a multiethnic perspective. *BMC Evol Biol* 7: 124.
- Hammer MF, Karafet T, Rasanayagam A, Wood ET, Altheide TK, et al. (1998) Out of Africa and back again: nested cladistic analysis of human Y chromosome variation. *Mol Biol Evol* 15: 427–441.
- Semino O, Santachiara-Benerecetti AS, Falaschi F, Cavalli-Sforza LL, Underhill PA (2002) Ethiopians and Khoisans share the deepest clades of the human Y-chromosome phylogeny. *Am J Hum Genet* 70: 265–268.
- Hassan HY, Underhill PA, Cavalli-Sforza LL, Ibrahim ME (2008) Y-chromosome variation among Sudanese: restricted gene flow, concordance with language, geography, and history. *Am J Phys Anthropol* 137: 316–323.
- Gomes V, Sánchez-Diz P, Amorim A, Carracedo A, Gusmão L (2010) Digging deeper into East African human Y chromosome lineages. *Hum Genet* 127: 603–613.
- Naidoo T, Schlebusch CM, Makkani H, Patel P, Mahabeer R, et al. (2010) Development of a single base extension method to resolve Y chromosome haplogroups in sub-Saharan African populations. *Investig Genet* 1: 6.
- Underhill PA, Passarino G, Lin AA, Shen P, Mirazón Lahr M, et al. (2001) The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Ann Hum Genet* 65: 43–62.
- Semino O, Passarino G, Oefner PJ, Lin AA, Arbuzova S, et al. (2000) The genetic legacy of Paleolithic *Homo sapiens sapiens* in extant Europeans: a Y chromosome perspective. *Science* 290: 1155–1159.
- Battaglia V, Fornarino S, Al-Zahery N, Olivieri A, Pala M, et al. (2009) Y-chromosomal evidence of the cultural diffusion of agriculture in Southeast Europe. *Eur J Hum Genet* 17: 820–830.
- Cruciani F, Trombetta B, Sellitto D, Massaia A, Destro-Bisol G, et al. (2010) Human Y chromosome haplogroup R-V88: a paternal genetic record of early mid Holocene trans-Saharan connections and the spread of Chadic languages. *Eur J Hum Genet* 18: 800–807.
- Simms TM, Martinez E, Herrera KJ, Wright MR, Perez OA, et al. (2011) Paternal lineages signal distinct genetic contributions from British Loyalists and continental Africans among different Bahamian islands. *Am J Phys Anthropol* 146: 594–608.
- Simms TM, Wright MR, Hernandez M, Perez OA, Ramirez EC, et al. (2012) Y-chromosomal diversity in Haiti and Jamaica: contrasting levels of sex-biased gene flow. *Am J Phys Anthropol* 148: 618–631.
- Tishkoff SA, Reed FA, Friedlaender FR, Ehret C, Ranciaro A, et al. (2009) The genetic structure and history of Africans and African Americans. *Science* 324: 1035–1044.
- Henn BM, Gignoux CR, Jobin M, Granka JM, Macpherson JM, et al. (2011) Hunter-gatherer genomic diversity suggests a southern African origin for modern humans. *Proc Natl Acad Sci U S A* 108: 5154–5162.
- Wei W, Ayub Q, Chen Y, McCharty S, Hou Y, et al. (2012) A calibrated human Y-chromosomal phylogeny based on resequencing. *Genome Res*, epub ahead of print, DOI: 10.1101/gr.143198.112.
- Meyer M, Kircher M, Gansauge MT, Li H, Racimo F, et al. (2012) A high-coverage genome sequence from an archaic Denisovan individual. *Science*, epub ahead of print, DOI: 10.1126/science.1224344.
- Scozzari R, Cruciani F, Pangrazio A, Santolamazza P, Vona G, et al. (2001) Human Y-chromosome variation in the western Mediterranean area: implications for the peopling of the region. *Hum Immunol* 62: 871–884.
- Myres NM, Ritchie KH, Lin AA, Hughes RH, Woodward SR, et al. (2009) Y-chromosome short tandem repeat intermediate variant alleles DYS392.2, DYS449.2, and DYS385.2 delineate new phylogenetic substructure in human Y-chromosome haplogroup tree. *Croat Med J* 50: 239–249.

Author Contributions

Conceived and designed the experiments: RS FC. Performed the experiments: AM EDA. Analyzed the data: AM BT. Contributed reagents/materials/analysis tools: RS NMM UAP. Wrote the paper: RS FC. Reviewed the manuscript prior to submission: RS NMM UAP FC.