[Downloaded free from http://www.ijhg.com on Monday, July 25, 2011, IP: 110.234.118.27] || Click here to download free Android application fo

Ind J Hum Genet 4: 52-61 (1998)

A Polymorphic Human Y-chromosomal G to A Transition Found in India

Arpita Pandya[1], Turi E. King[2], Fabrício R. Santos[1], Paul G. Taylor[2], Kumarasamy Thangaraj[3], Lalji Singh[3], Mark A. Jobling[2] and Chris Tyler-Smith[1]*

[1] Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK.
[2] Department of Genetics, University of Leicester, University Road, Leicester LE1 7RH, UK.
[3] Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad 500 007, India.

Abstract

We describe a new G to A transition polymorphism, designated Apt, on the human Y chromosome. The A allele chromosomes and the most closely related G allele chromosomes have been found only in India, where they make up about 1.3% of the population. They are found in several different social groups and show considerable microsatellite, minisatellite and major satellite diversity, suggesting that they do not have a recent origin.

Key words: Human Y chromosome, base substitution polymorphism, microsatellite and minisatellite diversity, Indian genetic history

Introduction

DNA polymorphisms on the non-recombining part of the Y chromosome are useful reagents for studying human genetic history (Jobling and Tyler-Smith, 1995; Sautes Tyler Smith, 1996; Mitchell and Hammer, 1997). They allow male lineages to be distinguished and their relationships to be determined, and thus provide insights into male-specific aspects of population structure and movements. The number of copies of the Y chromosome in the population is about one quarter of the number of each autosome and there is a large variance in the number of offspring left by males, so Y-chromosomal lineages are affected more by genetic drift than autosomal lineages. If the mutation rate is low and migration is slower than drift, Y-chromosomal markers will show a high degree of geographical specificity in their localisation.

The markers available have a wide range of mutation rates. Base substitutions (Seilsted *et al*, 1994), small insertions and deletions (Underhill *et al*, 1997), and retroposon insertions (Hammer, 1994), collectively called 'biallelic markers', have a low mutation rate. Consequently, these polymorphisms usually have a single origin and may be present in a restricted part of the world, such as 47z in Japan, Korea and Taiwan (Nakahori *et al*, 1989; Lin *et al*, 1994), and *DYS199* in the Americas (Underhill *et al*, 1996). In contrast, microsatellites (Roewer *et al*, 1992) and the minisatellite MSY1 (Joling *et al*, 1998) have high mutation rates; for a microsatellite, each allele size has usually arisen several times and its presence does not indicate shared ancestry. For MSY1, internal allele structures can be determined by MVR-PCR, which reveals greater diversity; here, too, some kinds of structures have clearly arisen several times, but others have probably arisen very rarely, and thus are good markers for shared ancestry. Both microsatellites and MSY1 are polymorphic in all populations (Joling *et al*, 1998; Kayser *et al*,1997). Combinations of markers are

* Author for correspondence and reprints:

[Downloaded free from http://www.ijhg.com on Monday, July 25, 2011, IP: 110.234.118.27] || Click here to download free Android application fo

particularly informative and have been used to investigate the peopling of Japan (Hammer and Horai, 1995) and the origins of some of the populations in Asia and northern Europe (Zerjal *et al*, 1997).

There are likely to be biallelic Y markers specific to each part of the world, and many more are needed for a comprehensive understanding of Y lineages. Here we present one that is found in India but not in the other regions tested.

Materials and methods

DNA samples

Some of the DNA samples used (Table 1) were collected by the authors, some have previously been described (Zerjal *et al*, 1997; Mathias *et al*, 1994), and others were provided by our colleagues: Bamileke and Biaka Pygmies (Giovanni Destro-Bisol, Gabriella Spedini), Chinese (Aiping Liu, John Mitchell), Cook Islanders (Bryan Sykes), Indian Gujaratis (Ken McElreavey, Reiner Veitia), Indonesians (John Mitchell), Slovak Gypsies (Vladimír Ferák).

TABLE 1: Distribution of the haplogroup 15 chromosomes

Continent	Population	G	A	Total
Africa	San	9		9
	Biaka Pygmy	25		25
	Kenyan	14		14
	Bamileke	40		40
	Algerian	27		27
	Other	12		12
Europe	Icelandic	28		28
	British	27		27
	Basque	26		25
	Slovak Gypsy	74		74
	Other	35		35
Asia	Indian, N	275	2	277
	Indian, S	98	3	101
	Sri Lankan	24		24
	Mongolian	65		65
	Chinese	80		80
	Indonesian	18		18
	Other	19		19
Oceania	Cook Islander	10		10
	Australian	3		3
	Other	3		3
America	Amerindians, N	2		2
	Amerindians, S	5		5
Total		**919**	**5**	**924**
	Chimpanzee	2		2
	Gorilla	1		1

Analysis of existing Y polymorphisms

DNA preparation, restriction enzyme digestion, gel electrophoresis and hybridisations using 50f2 (Guelläen *et al*, 1984), and the major satellite probes (Oakey and Tyler-Smith, 1990) were carried out as described previously (Oakey and Tyler-Smith, 1990 with reference therein), except that a low-stringency wash (2 x SSC, 1% SDS, 65°C) was used for 50f2. Typing of the alphoid heteroduplex (Santos *et al*, 1995), *DYS19*, *DYS389I*, *DYS389II*, *DYS390*, *DYS391*, *DYS392* and *DYS393* microsatellite (Kayser et al, 1997) and MSY1 minisatellite (Joling *et al*, 1998) polymorphisms was according to the published protocols. Median networks (Bandelt *et al*, 1995) linking microsatellite or minisatellite haplotypes were constructed using the program Network 1.1 (Arne Röhl and Peter Forster, personal communication). Dates were estimated using a method based on that described by Bertranpetit and Calafell (1996).

Establishment of a PCR assay for the Apt polymorphism

Cosmid M2H4 was isolated from the library LL0YNC03 by screening with 50f2, and the 9 kb *Taq*I fragment at the 50f2/E locus subcloned into the *Cla*I site of Bluescript (Stratagene). Further subcloning was carried out to isolate the region of this fragment predicted to contain the polymorphic *Taq*I site, and sequencing of this subclone allowed a pair of primers (TEK E, 5'-TGGATTGCATTCAACTTCACTTAC-3' and TEK G, 5'-CTGAGTTCAAATGCTCGGGTCTC-3') to be designed flanking the site. PCR amplification was carried out in an MJR PTC-200 using the buffer described (Jeffreys *et al*, 1990) and the

[Downloaded free from http://www.ijhg.com on Monday, July 25, 2011, IP: 110.234.118.27] || Click here to download free Android application fo

cycle conditions 94°C 30 sec, 65.5°C 30 sec, 72°C 60 sec for 33 cycles.

Results

A new G to A transition

We carried out a survey of known Y polymorphisms in a set of 349 males from around the world. These experiments included 14 probe hybridisations to restriction enzyme digests and so constituted a search for additional polymorphisms in the 51 bands visualised. Some new polymorphisms were discovered, and one is shown in Fig. 1a. In *Taq*I digests, the probe 50f2 commonly detected the pattern of five bands shown in track G, but in some individuals (track A) the 9.8 kb band (50f2/E; assigned according to Vergnaud *et al*, 1986) was absent and a novel 4.0 kb band was present. Other digests, such as *Eco*RI (not shown), did not reveal an altered hybridisation pattern in these males, and so the polymorphism was interpreted as a probable point mutation introducing a new *Taq*I site 4.0 kb from one end of the 9.8 kb fragment (Fig. 1d).

We then set out to determine the nucleotide change(s) causing the polymorphism and establish a more convenient assay for it. The 9.8 kb *Taq*I fragment was cloned and partially sequenced. The sequence information was used to design primers that amplified an 850 bp fragment around the 4.0 kb position (Fig. 1b). As expected, this fragment was not cleaved by *Taq*I in most individuals (track G), but it was cleaved into ~420 bp and ~430 bp subfragments in those individuals with the 4.0 kb hybridisation band, confirming the interpretation of the hybridisation pattern.

Sequencing of the PCR products from both types of male revealed a single difference: the nucleotide at position 419 was a G in individuals with the 9.8 kb *Taq*I fragment/additional site absent and an A in individuals with the 4.0 kb *Taq*I fragment/additional site present. The change converted the sequence TCGG into TCGA, the *Taq*I recognition site, accounting for the changes in *Taq*I digestion pattern. Inspection of the sequence revealed that the next two bases were CC. Therefore the G allele chromosomes were predicted to carry the sequence GGCC, the *Hae*III recognition site, while the A allele chromosomes carried GACC. The loss of this *Hae*III site in A allele chromosomes was confirmed experimentally (Fig. 1b). The allele present in two male chimpanzees and one male gorilla was also determined by sequencing PCR products. All carried the G, indicating that this is the ancestral state. Thus the Apt polymorphism was found to be a G to A transition, and a PCR amplification/restriction enzyme digestion assay that positively identified both alleles was established.

Relationship to other Y-chromosomal markers

A network showing the relationships of 23 haplogroups defined by 22 unique or rare-event markers has been constructed. It is an extension of published networks (Jobling and Tyler-Smith, 1995; Jobling *et al*, 1996) and details will be described elsewhere. Here we note that the Apt A allele defines haplogroup 15, and that haplogroup 15 is derived from haplogroup 2 (Fig. 2a).

The alphoid heteroduplex polymorphic system (Santos *et al*, 1995) has also been used to characterise the 349 males and additional individuals. A allele (haplogroup 15) chromosomes have a characteristic heteroduplex pattern designated αhXXVIII (Fig. 1c, track A). This pattern was found in all haplogroup 15 chromosomes but in only one additional individual out of approximately 2 000 tested. The additional male was a member of haplogroup 2. We can therefore propose the phylogeny shown in Fig. 2b, in which haplogroup 2 chromosomes with the common αhIII pattern (Fig. 1c, track G) give rise to haplogroup 2 variants with αhXXVIII, and the G to A transition subsequently occurs on this background. Haplogroup 2αhXXVIII chromosomes are the closest relatives of the haplogroup 15 chromosomes.

Geographical distribution of haplogroup 15 chromosomes

We have used the hybridisation and PCR assays to test a total of 924 individuals from around the world for the G or A allele (Table 1). There was

[Downloaded free from http://www.ijhg.com on Monday, July 25, 2011, IP: 110.234.118.27] || Click here to download free Android application fo

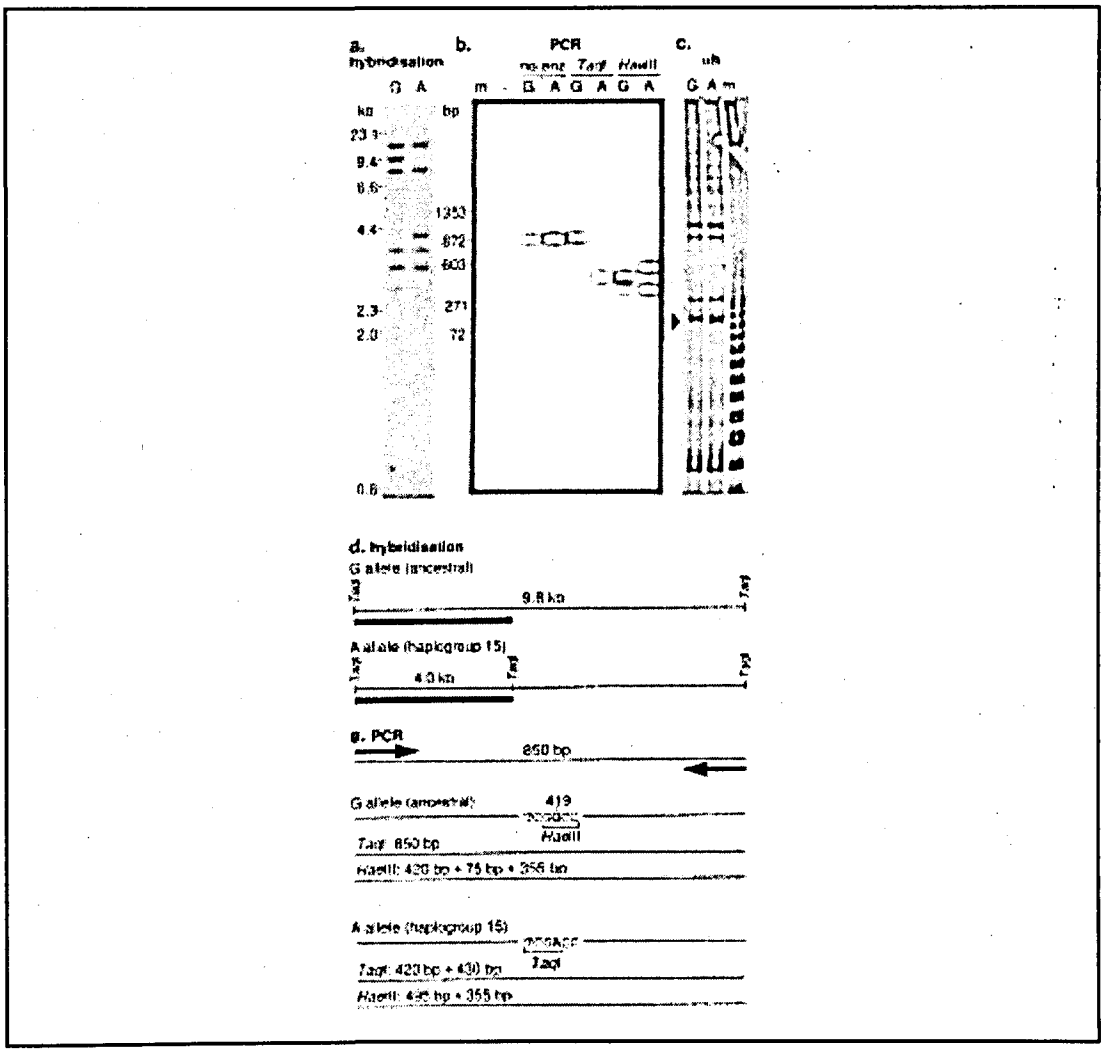

FIGURE 1: Detection and characterisation of the G to A transition polymorphism

a. Genomic DNA from a G allele male (G) or an A allele male (A) was digested with TaqI*, fractionated by agarose gel electrophoresis and probed with 50f2.* lHin*dIII size markers are shown.*

b. PCR products from G allele and A allele males are shown without digestion or after digestion with TaqI *or* Hae*III. m = fx174/*Hae*III size markers, - = no DNA.*

c. Alphoid heteroduplex patterns of typical G allele and A allele males. The arrowhead marks the additional heteroduplex band that distinguishes ahIII from ahXXVIII. m = 50 bp ladder size marker.

d. Map of TaqI *sites in genomic DNA. The black box indicates the region that hybridises to the 50f2 probe.*

e. Map of the PCR products showing the Hae*III site found in the G allele chromosomes and the* TaqI *site found in A allele chromosomes.*

[Downloaded free from http://www.ijhg.com on Monday, July 25, 2011, IP: 110.234.118.27] || Click here to download free Android application fo

complete agreement between the two assays. The A allele was rare in this sample and was found in just five individuals (~0.5%). All five came from India (Table 2), where the A allele was present in about 1.3% of those tested. The related haplogroup 2/αhXXVIII individual was also from India, so, in our sample, these lineages are specific to India.

Variation within the haplogroup 15 chromosomes

The haplogroup 15 chromosomes were analysed with a set of seven Y-specific microsatellites (Fig. 4). All had different haplotypes; the average difference was 5.6 steps per haplotype or 0.8 steps per locus. These chromosomes and the haplogroup 2αhXXVIII chromosome were also typed at the

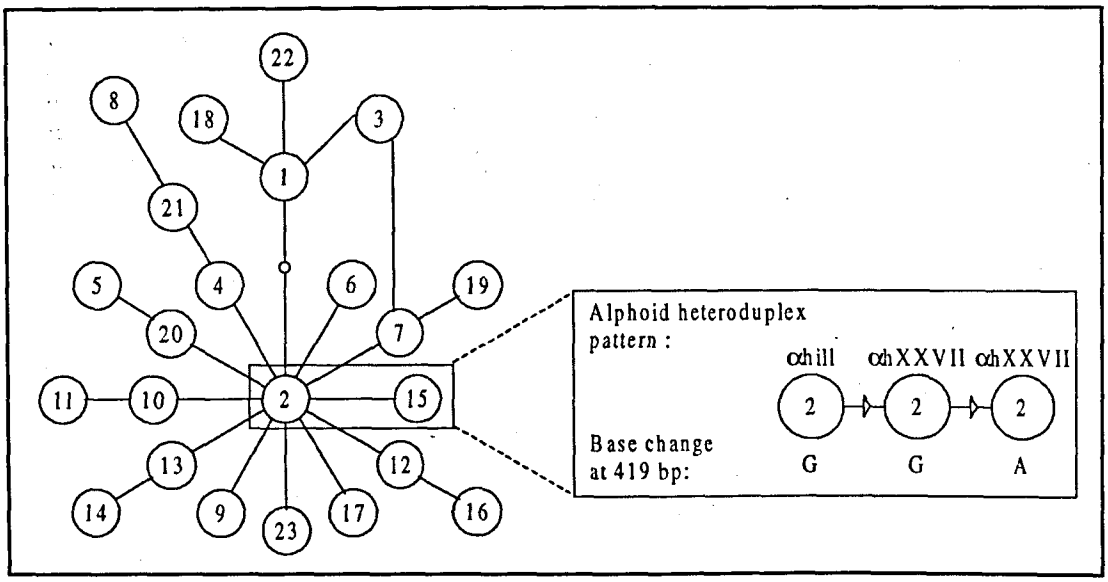

FIGURE 2: Phylogenetic position of the A allele (haplogroup 15) chromosomes

a. Unrooted network illustrating the relationships between 23 haplogroups (numbers in circles) defined by 22 polymorphisms (lines).

b. Detail showing the sequence of changes leading to the haplogroup 15 chromosomes.

TABLE 2: Details of the haplogroup 15 and related individuals

Code	*Apt*	*Haplo-group*	*αh type*	*Comm[1].*	*Geographical origin*	*Language*	*MSY[1] haplotype[2]*	*DYZ[3] (kb)*	*DYZ[5] (kb)*
m554	G	2	XXVIII	ST[3]	Andhra Pradesh	Telugu	$(3)_8(1)_{11}(3)_{28}(4)_{15}$	1350	740
m431	A	15	XXVIII	OBC[4]	Andhra Pradesh	Telugu	$(3)_5(1)_{12}(3)_{32}(4)_{15}$	>1600	640
m478	A	15	XXVIII		Andhra Pradesh		$(3)_6(1)_{11}(3)_{31}(4)_8$	>1600	840
m513	A	15	XXVIII	Dalit	Karnataka	Kannada	$(3)_6(1)_{12}(3)_{30}(4)_{15}$	>1600	750
m593	A	15	XXVIII	ST	Bihar	Hindi	$(3)_4(1)_{10}(3)_{36}(4)_{15}$	1380	770
m622	A	15	XXVIII	ST	Bihar	Hindi	$(3)_6(1)_{12}(3)_{27}(4)_{20}$	1400	510

Notes: [1]Comm. = community [2]The number in parentheses indicates repeat type (type 1, 3 or 4), and the following number in subscript indicates the number of repeats of that type. [3]ST = Scheduled Tribe [4]OBC = Other Backward Caste

[Downloaded free from http://www.ijhg.com on Monday, July 25, 2011, IP: 110.234.118.27] || Click here to download free Android application fo

minisatellite MSY1 locus (Fig. 3). Again, all chromosomes had different codes; the mean difference among the haplogroup 15 chromosomes was 10.8 single repeat unit steps and there were no differences in modular structure. The networks relating the microsatellite or MSY1 haplotypes were not strikingly similar (Fig. 3), although both separate m432 and m593 from the other chromosomes. The different topologies of the networks may reflect the limited number of mutations and their stochastic occurrence. All six chromosomes were also typed at the major satellite loci *DYZ3* and *DYZ5* (Table 2). *DYZ3* arrays were among the largest seen in any country and had the largest average size of any haplogroup; *DYZ5* arrays were very variable.

Discussion

The Apt polymorphism provides a new Y-chromosomal marker that appears to be specific to India. Haplotype analysis with 21 other biallelic markers is consistent with a single origin for all A allele chromosomes.

Origin of the haplogroup 15 chromosomes

Where and when did the G to A transition occur? Since it has been found only in India, and the ancestral haplotype, haplogroup 2αhXXVIII, has also been found only in India, the most likely location for the mutation could be in India. However, this conclusion has to be considered with caution because few of the nearby populations have been surveyed (Fig. 4) and the mutation may have occurred a long time ago.

In principle, the date of the G to A transition can be estimated from the diversity of the haplogroup 15 chromosomes (Bertranpetit and Calafell, 1996). The transition must have occurred on a single chromosome and the observed diversity has arisen subsequently by mutation. It is difficult to estimate accurately the time required because the number of chromosomes is small, mutation rates are not known accurately, and the ancestral microsatellite or MSY1 haplotype in not known. If all positions in Figure 4a are considered potential roots and a microsatellite mutation rate of 2.1×10^{-3} is used (Heyer *et al*, 1997), between ~200 and ~350 generations would be required, or 4000 to 7000 years at 20 years per generation. If the 95% confidence interval limits of the mutation rate (0.6×10^{-3} to 4.9×10^{-3}) are also considered, the range of dates extends from about 2000 years to 25000 years. If a similar calculation is done with the MSY1 data (Fig. 4c), using the position of the branch to m554 as the root and a mutation rate of 0.2×10^{-1} to 1.1×10^{-1} (Joling *et al*, 1998), a slightly more recent age of 1400 to 7600 years is obtained. Whatever the absolute date, relative dating is also possible: for example, the higher microsatellite diversity suggests that the G to A transition, despite its limited geographical distribution, occurred earlier than the Tat T to C transition (Zerjal *et al*, 1997) that has spread over much of Asia and northern Europe.

Distribution of haplogroup 15 and related individuals within India

The social structure of modern India is extremely intricate. The 'People of India' project (Singh, 1993) reports 4635 communities which are grouped into three broad categories: *Dalits* (previously known as 'Untouchables'; ~16% of the population in the 1981 census), Scheduled Tribes (~8%), and Other Communities (~76%). The Other Communities practice several different religions, the most common being Hindu. The caste system, usually defined by the Hindu religion, consists mainly of *Brahmans*, *Kshatriyas*, *Vaishyas* and Other Backward Castes. Some of the *Dalits* are also Hindus. There has been mobility between the castes, especially the middle castes, but the mobility of the Y chromosome may have been low because children usually inherit their father's caste. The Scheduled Tribes, once isolated populations of unknown origin with distinct cultural practices and languages, have undergone some modernisation in the last two centuries due to economic development and have now often adopted the language of the surrounding populations. Indo-Aryan family languages are commonly spoken in the north, Dravidian family languages in the south.

[Downloaded free from http://www.ijhg.com on Monday, July 25, 2011, IP: 110.234.118.27] || Click here to download free Android application fo

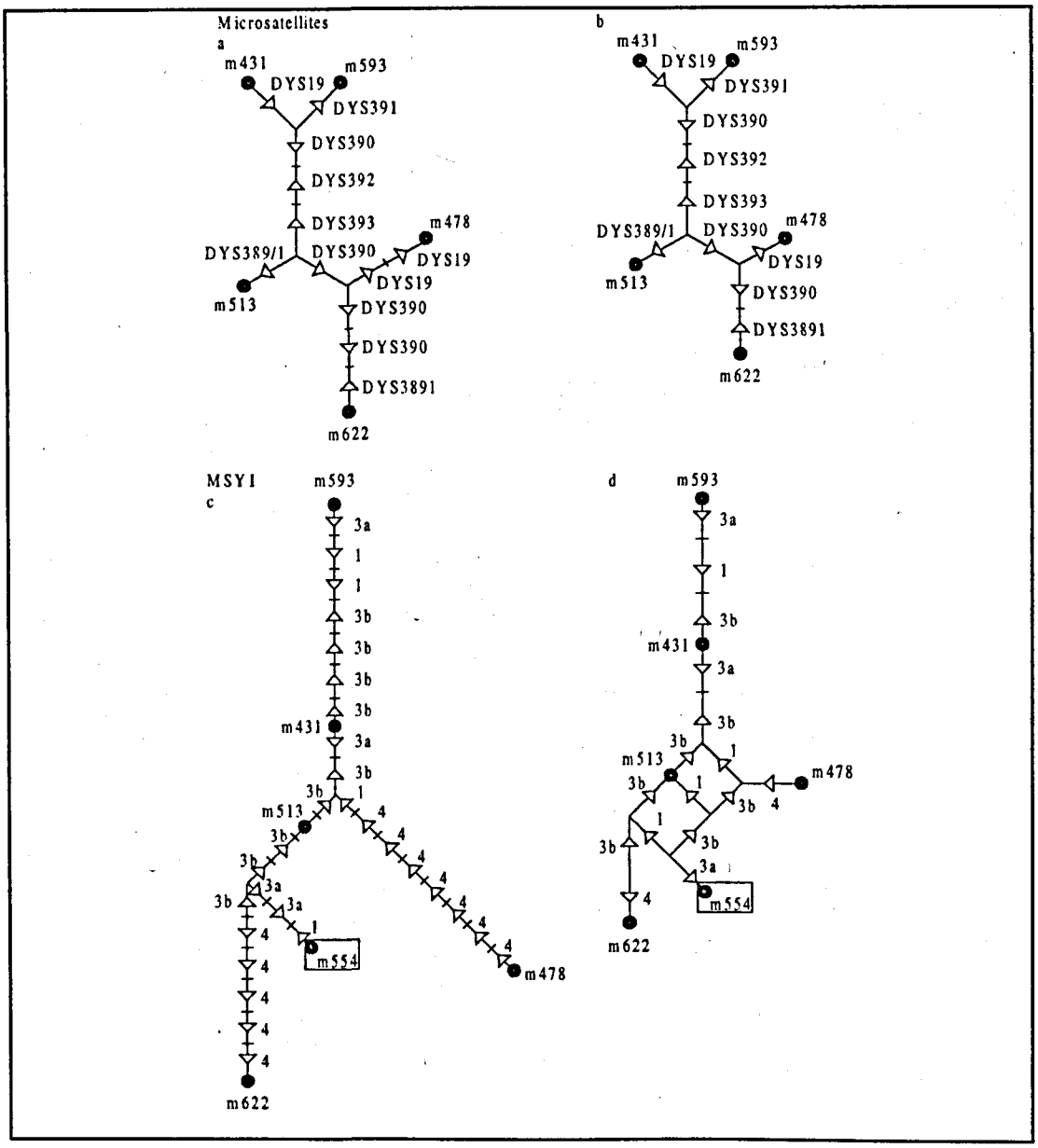

FIGURE 3: *Diversity within the haplogroup 15 chromosomes. Individual haplotypes are represented by circles, mutational changes by arrowheads on lines*

- *a. Median network linking microsatellite haplotypes by single step mutations.*
- *b. Median network linking microsatellite haplotypes, allowing multistep mutations.*
- *c. Median network linking MSY1 haplotypes, single step mutations.*
- *d. Median network linking MSY1 haplotypes, multistep mutations.*

[Downloaded free from http://www.ijhg.com on Monday, July 25, 2011, IP: 110.234.118.27] || Click here to download free Android application fo

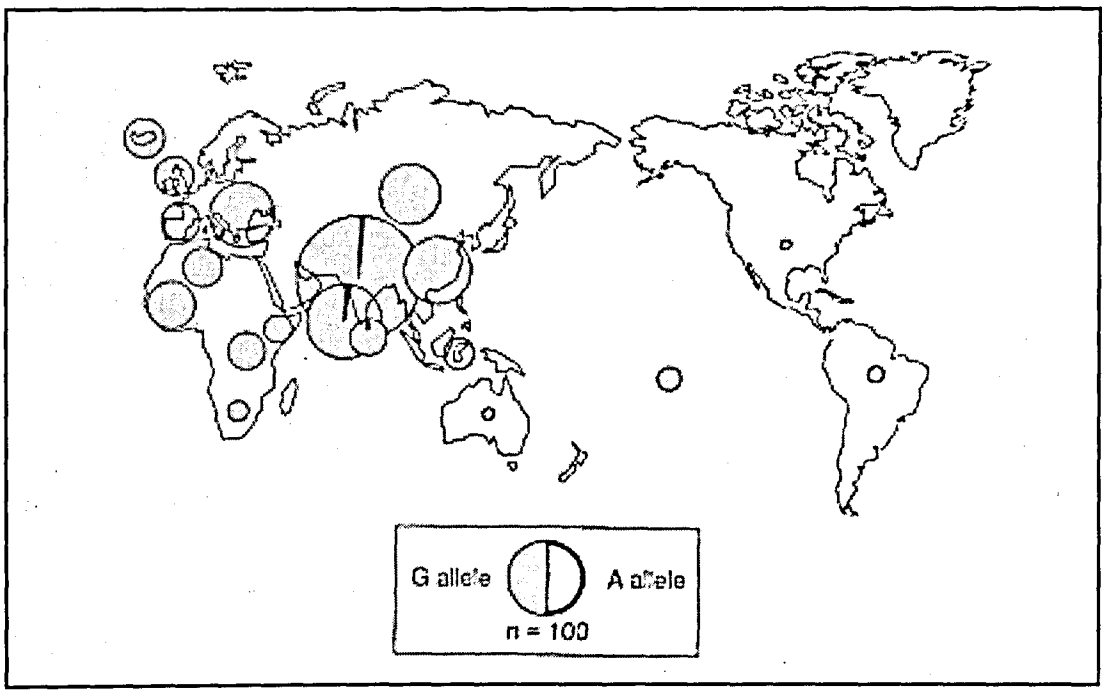

***FIGURE 4:** Worldwide distribution of the G allele (gray) and A allele (black) chromosomes. The area of each circle is proportional to the size of the population sample*

The six individuals sharing hXXVIII have very diverse backgrounds (Table 2). Two come from Bihar in the north of India and speak Hindi, an Indo-Aryan family language; the other four come from Andhra Pradesh or Karnataka in the south and three of them speak Dravidian languages. Their communities include the *Dalits*, the Other Backward Castes, and the Scheduled Tribes.

How might this distribution arise? Several hypotheses can be considered. (1) It could be due to independent recurrences of admixture. However, no source population for such admixture has been identified and family histories rule out recent admixture. (2) The chromosome could have arisen in India sufficiently long ago to spread throughout the country and accumulate diversity, and have drifted to low frequency. It could, for example, predate the entry of the Indo-Aryan language speakers into NW India at around 1500 BC (Thapar, 1996) since it was strikingly absent from the 163 members of the higher castes (*Brahmins*, *Kshatriyas* and *Vaishyas*) tested, whose ancestors formed the major part of the early Indo-Aryan social organisation. (3) In a variant of this hypothesis, the polymorphism could have arisen in one of the ancient source populations who contributed to the Indus Culture of 2500 BC - 1600 BC (Wolpert,1997) and are possible ancestors of modern Dravidian language speakers.

In conclusion, the new G to A transition provides a Y-chromosomal marker that is easy to score and should contribute to future studies of male genetic history in India and nearby countries.

Acknowledgments

We thank everyone who donated their DNA for this study, and Kamal Bagai, Giovanni Destro-Bisol, John Edwards, Nathan Ellis, Vladimír Ferák, Tudevdagva Gerelsaikhan, Michael Hammer, Raoul Heller, Doudja Nafa, Aiping Liu,

[Downloaded free from http://www.ijhg.com on Monday, July 25, 2011, IP: 110.234.118.27] || Click here to download free Android application fo

Adolfo López de Munain, Ken McElreavey, John Mitchell, Elizabeth Robinson, Gabriella Spedini, Bryan Sykes, Reiner Veitia and Upen de Zylva for providing samples. We thank Hans-Jurgen Bandelt, Peter Forster and Arne Röhl for advice on median networks and providing the Network 1.1 program, and Romila Thapar for discussions on Indian history. We also thank Pieter de Jong for the cosmid library LL0YNC03, which was constructed at the Biomedical Sciences Division, Lawrence Livermore National Laboratory, CA 94550, USA, under the auspices of the National Laboratory Gene Library Project sponsored by the U.S. Department of Energy. A.P. was supported by the BBSRC, F.R.S. by The Leverhulme Trust, P.G.T. by The Wellcome Trust, M.A.J. by a Wellcome Trust Career Development Fellowship (grant no. 044918) and C.T-S. by the CRC.

References

Bandelt HJ, Forster P, Sykes BC and Richards MB (1995). Mitochondrial portraits of human populations using median networks. *Genetics* 141:743-753.

Bertranpetit J and Calafell F (1996). Genetic and geographic variation in cystic fibrosis: evolutionary considerations, in: Chadwick D and Cardew G (Eds.), Variation in the human genome, John Wiley & Sons, Chichester, pp. 97-118.

Guelläen G, Casanova M, Bishop C, Geldwerth D, Andre G, Fellous M and Weissenbach J (1984). Human XX males with Y single-copy DNA fragments. *Nature* 307:172-173.

Hammer MF (1994). A recent insertion of an Alu element on the Y chromosome is a useful marker for human population studies. *Mol. Biol, Evol.* 11:749-761.

Hammer MF, and Horai S (1995). Y chromosomal DNA variation and the peopling of Japan. *Am. J. Hum. Genet.* 56:951-962.

Heyer E, Puymirat J, Dieltjes P, Bakker E and de Knijff P (1997). Estimating Y chromosome specific microsatellite mutation frequencies using deep rooting pedigrees. *Hum. Mol. Genet.* 6:799-803.

Jeffreys AJ, Neumann R and Wilson V (1990). Repeat unit sequence variation in minisatellites: a novel source of polymorphism for studying variation and mutation by single molecule analysis. *Cell* 60: 473-485.

Jobling MA and Tyler-Smith C (1995). Fathers and sons: the Y chromosome and human evolution. *Trends Genet.* 11:449-456.

Jobling MA, Samara V, Pandya A, Fretwell N, Bernasconi B, Mitchell RJ, Gerelsaikhan T, Dashnyam B, Sajantila A, Salo PJ, Nakahori Y, Disteche CM, Thangaraj K, Singh L, Crawford MH and Tyler-Smith C (1996). Recurrent duplication and deletion polymorphisms on the long arm of the Y chromosome in normal males. *Hum. Mol. Genet.* 5: 1767-1775.

Jobling, MA, Bouzekri N and Taylor PG (1998). Hypervariable digital DNA codes for human paternal lineages: MVR-PCR at the Y-specific minisatellite, MSY1 (*DYF155S1*). *Hum. Mol. Genet. in press.*

Kayser M, and others (1997). Evaluation of Y-chromosomal STRs: a multicenter study. *Int. J. Legal. Med.* 110:125-133.

Lin SJ, Tanaka K, Leonard W, Gerelsaikhan T, Dashnyam B, Nyamkhishig S, Hida A, Nakahori Y, Omoto K, Crawford MH and Nakagome Y (1994). A Y-associated allele is shared among a few ethnic groups of Asia, Jpn. *J. Hum. Genet.* 39:299-304.

Mathias N, Bayés M and Tyler-Smith C (1994). Highly informative compound haplotypes for the human Y chromosome *Hum. Mol. Genet.* 3:115-123.

Mitchell RJ and Hammer MF (1997). Human evolution and the Y chromosome. *Curr. Op. Genet. Dev.* 6:737-742.

Nakahori Y, Tamura T, Yamada M and Nakagome Y (1989). Two 47z [DXYS5] RFLPs on the X and the Y chromosome. *Nucl. Acids Res.* 17: 2152.

[Downloaded free from http://www.ijhg.com on Monday, July 25, 2011, IP: 110.234.118.27] || Click here to download free Android application fo

Oakey R and Tyler-Smith C (1990). Y chromosome DNA haplotyping suggests that most European and Asian men are descended from one of two males. *Genomics* 7:325-330.

Roewer L, Arnemann J, Spurr NK, Grzeschik KH and Epplen JY (1992). Simple repeat sequences on the Y chromosome are equally polymorphic as their autosomal counterparts. *Hum. Genet.* 89:389-394.

Santos FR, Pena SDJ and Tyler-Smith C (1995). PCR haplotypes for the human Y chromosome based on alphoid satellite variants and heteroduplex analysis. *Gene.* 165:191-198.

Santos FR and Tyler-Smith C (1996). Reading the human Y chromosome: the emerging DNA markers and human genetic history, Brazilian. *J. Genet.* 19: 665-670.

Seielstad MT, Hebert JM, Lin AA, Underhill PA, Ibrahim M, Vollrath D and Cavalli-Sforza LL (1994). Construction of human Y-chromosomal haplotypes using a new polymorphic A to G transition. *Hum. Mol. Genet.* 3:2159-2161.

Singh KS (1993). People of India, volume XI, an anthropological atlas. *Oxford University Press,* Delhi.

Thapar R (1966). A history of India 1, *Penguin Books,* London.

Underhill PA, Jin L, Zemans R, Oefner PJ and Cavalli-Sforza LL (1996). A pre-Colombian Y chromosome-specific transition and its implications for human evolutionary history. *Proc. Natl. Acad. Sci.* USA 93:196-200.

Underhill PA, Jin L, Lin AA, Mehdi SQ, Jenkins T, Vollrath D, Davis RW, Cavalli-Sforza LL and Oefner PJ (1997). Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Res.* 7:996-1005.

Vergnaud G, Page DC, Simmler MC, Brown L, Rouyer F, Noel B, Botstein D, de la Chapelle A and Weissenbach J (1986). A deletion map of the human Y chromosome based on DNA hybridization *Am. J. Hum. Genet.* 38:109-124.

Wolpert S. (1977). A new history of India. *Oxford University Press,* Oxford.

Zerjal T, Dashnyam B, Pandya A, Kayser M, Roewer L, Santos FR, Schiefenhövel W, Fretwell N, Jobling MA, Harihara S, Shimizu K, Semjidmaa D, Sajantila A, Salo P, Crawford MH, Ginter EK, Evgrafov OV and Tyler-Smith C (1997). Genetic relationships of Asians and northern Europeans, revealed by Y-chromosomal DNA analysis. *Am. J. Hum. Genet.* 60:1174-1183.