



haplogroup M was brought to Asia from East Africa, along the southern route, by the earliest migration wave of anatomically modern humans, ~60,000 years ago (Kivisild et al. 1999a, 1999b, 2000; Quintana-Murci et al. 1999). Another deep late Pleistocene link through haplogroup U was found to connect western Eurasian and Indian populations. Less than 10% of the maternal lineages of the caste populations had an ancestor outside India in the past 12,000 years (Kivisild et al. 1999a, 1999b). mtDNA profiles from a larger set of populations all over the subcontinent have bolstered the view of fundamental genomic unity of Indians (Roychoudhury et al. 2001). In contrast, the Y-chromosome genetic distance estimates showed that the chromosomes of Indian caste populations were more closely related to Europeans than to eastern Asians (Bamshad et al. 2001). The tendency of higher caste status to associate with increasing affinities to European (specifically to eastern European) populations hinted at a recent male-mediated introduction of western Eurasian genes into the Indian castes' gene pool. The similarities with Europeans were specifically expressed in substantial frequencies of clades J and R1a (according to Y Chromosome Consortium [YCC] 2002 nomenclature) in India. The exact location of the origin of these haplogroups is still uncertain, as is the timing of their spread (Zerjal et al. 1999; Bamshad et al. 2001; Passarino et al. 2001; Quintana-Murci et al. 2001; Wells et al. 2001).

To address the question of the origin of Indian maternal and paternal lineages further, we analyzed variation in mtDNA, the Y chromosome, and one autosomal locus (Jin et al. 1999) in two southern Indian tribal groups from Andhra Pradesh and compared them with Indian caste groups and populations from Iran, the Middle East, Europe, and central Asia.

## Material and Methods

Chenchus were first described as shy hunter-gatherers by the Mohammedan army in 1694. They reside in the ranges of Amrabad Plateau, Andhra Pradesh, and have a population size of ~17,000. Their society is patriarchal and patrilineal, with marriage occurring mostly between clans (kulam) of equal status. Chenchus are described as an australoid population, when physical anthropological features are used as criteria (Bhowmick 1992; Singh 1997; Thurin 1999). The Chenchu language belongs to the Dravidian language family.

More than 300,000 Koyas live in the plains and forests on both sides of the Godavari River in Andhra Pradesh. Their language is related to the Gondi, which connects a large group of Dravidian languages in southern India. They are primarily farmers and live in villages. Exoga-

mous patrilocal clans make up their social structure, as they do for the Chenchus (Singh 1997).

After informed consent was obtained, 180 blood samples were collected from healthy and maternally unrelated volunteers belonging to Chenchu and Koya tribes from Andhra Pradesh, 106 West Bengalis of different caste ranks, 58 Konkanastha Brahmins from Bombay, and 53 Gujaratis; in addition, 132 samples were collected from Sri Lanka (including 40 Sinhalese), 112 from Punjabis of different caste rank, and 139 samples from Uttar Pradesh (including those described by Kivisild et al. 1999a). The Lambadi ( $n = 86$ ) sample from Andhra Pradesh, the Boksas ( $n = 18$ ) from Uttar Pradesh, and the Lobanas ( $n = 62$ ) from Punjab are described by Kivisild et al. (1999a) and were further analyzed here for Y-chromosomal and additional mtDNA markers. The Y-chromosome sample size of each population is shown in figure 3. In addition, 388 Turks from central Turkey (Cappadocia), 202 Kuwaitis, 202 Saudi Arabians, and 440 Iranians were used in mtDNA haplogroup frequency comparisons. In addition, Y-chromosome STR data from six loci were used for comparing intra- and interhaplogroup variances in selected haplogroups and populations. These included 88 central Asians (Altai, Kirghiz, Uzbek, and Tajik) belonging to haplogroup R1a, and Y chromosomes from haplogroups I (12 Estonians and 9 Czechs), J (6 Czechs), and R1 (39 Estonians and 30 Czechs). Further details about these populations will be published elsewhere. DNA was extracted using standard phenol-chloroform methods (Sambrook et al. 1989).

D A A —

Hypervariable segments (HVS) I (nucleotide positions [nps] 16024–16400) and II (nps 16520–300) of the control region were sequenced in 96 Chenchu and 81 Koya samples. In addition, three segments of the coding region (nps 1674–1880, 4761–5260, and 8250–8710) were sequenced, and informative RFLP positions (Macaulay et al. 1999; Quintana-Murci et al. 1999) were checked (table 1) in selected individuals from different haplotypes, to define haplogroup affiliations. Published HVS-I sequence data used for haplotype comparisons included 250 Telugus from Andhra Pradesh (Kivisild et al. 1999a; Bamshad et al. 2001), 48 Haviks, 43 Mukris, and 7 Kaders from Kerala and Karnataka (Mountain et al. 1995).

-C — A —

Y-chromosomal haplogroups were determined by RFLP and denaturing high-performance liquid chromatography (dHPLC) methods, using 35 biallelic markers (Rosser et al. 2000; Underhill et al. 2000, 2001b) that are shown in hierarchical relation to one another

in figure 3. Length variation at six STR loci (DYS19, DYS388, DYS390, DYS391, DYS392, and DYS393) was typed using Cy5-labeled primers, and amplification products were subjected to electrophoresis on ALF Express (Pharmacia-Amersham). Scoring of repeat lengths was standardized by use of controls sequenced by P. de Knijff.

1

A 246-bp segment of the MX1 locus of chromosome 21, containing eight polymorphic sites in humans, was sequenced, and a nearby *StuI* recognition site polymorphism was determined according to Jin et al. (1999) in 42 Chenchus, 28 Koyas, 35 West Bengalis, 34 Punjabis, and 35 Turks from Cappadocia.

Preparation of sequencing templates was performed according to methods described by Kaessmann et al. (1999). Purified products were sequenced with the DYE-namic ET terminator cycle sequencing kit (Amersham Pharmacia Biotech) and were analyzed on an ABI 377 DNA Sequencer. Sequences were aligned and analyzed with the Wisconsin Package (GCG).

D A

Median networks (Bandelt et al. 1995, 1999) were constructed using the Network 2.0 program (A. Röhl; Shareware Phylogenetic Network Software Web site) with default settings. Cluster ages were calculated using  $\rho$ , the averaged distance to a specified founder haplotype, according to Forster et al. (1996), as well as standard errors as described by Saillard et al. (2000a). In mtDNA coalescent calculations, using the estimator  $\rho$ , we use a mutation rate of one transition in the segment between nps 16090 and 16365 per 20,180 years, calibrated with the inference that links Eskimo-Aleutian haplogroup A diversity to their post-Younger Dryas population expansion (Forster et al. 1996).

To calculate the 95% credible regions (CR85(W)r/F2-345-ucaoCR85(W,)]T485(trl20(wA30.grages)-u-475(po1 aration)-41 Tc[(,)

**Table 1**

**mtDNA Haplotypes in Chenchu and Koya Populations**

Type	Group	HVS-I (minus 16000)	HVS-II	1674-1880;	4761-5260;	8250-8710	
		1					1 1 1 1 1 1 1 1 1 1 1 1 1 1
		6		6			3 4 5 5 7 7 8 9 0 0 0 0 0 1 2 2 2 2 3 4 5 5
		5		4 6			5 5 5 8 0 5 2 8 2 3 3 3 8 4 3 4 4 7 2 7 6 9
		1		4 3			3 7 8 2 2 9 4 2 3 6 9 9 7 6 0 0 0 0 5 6 0 2
		9		7 e			7 7 4 3 5 8 9 0 7 4 4 7 1 5 8 3 6 4 9 6 6 5
							a q a a a f b g y e c a z u g z o w o u a <del>28609</del> TD0 Tca





Table 2

## Major mtDNA Lineage Clusters in India and Western Asia

POPULATION (n)	<i>D</i> <sup>b</sup>	NO. OF LINEAGES IN CLUSTER (95% CR FOR PROPORTION) <sup>a</sup>										
		L1-L3	M	M2	M3	M5	MΔ9bp	U2i, U7	U1, U3-U6, U*	HV <sup>c</sup> , TJ, N1, X	B, F <sup>d</sup>	R*
Tribal:												
Chenchus (96) <sup>e</sup>	.87	0 (.00-.03)	93 (.91-.99)	17 (.11-.27)	1 (.00-.06)	18 (.12-.28)	3 (.01-.09)	0 (.00-.03)	0 (.00-.03)	0 (.00-.03)	0 (.00-.03)	1 (.00-.06)
Koyas (81) <sup>e</sup>	.94	0 (.00-.04)	56 (.58-.78)	15 (.12-.28)	5 (.03-.14)	0 (.00-.04)	17 (.14-.31)	1 (.00-.07)	0 (.00-.04)	0 (.00-.04)	0 (.00-.04)	25 (.22-.42)
Tamil Nadu (49) <sup>f</sup>	.96	0 (.00-.06)	35 (.58-.82)	1 (.01-.11)	12 (.15-.38)	0 (.00-.06)	0 (.00-.06)	6 (.06-.24)	2 (.01-.14)	0 (.00-.06)	0 (.00-.06)	6 (.06-.24)
Western Bengal (34) <sup>f</sup>	.99	0 (.00-.08)	22 (.48-.79)	2 (.02-.19)	3 (.03-.23)	0 (.00-.08)	0 (.00-.08)	7 (.10-.37)	0 (.00-.08)	0 (.00-.08)	0 (.00-.08)	5 (.07-.30)
Caste:												
Western Bengalis (106) <sup>e</sup>	.97	0 (.00-.03)	76 (.63-.79)	4 (.02-.09)	7 (.03-.13)	6 (.03-.12)	0 (.00-.03)	10 (.05-.17)	1 (.01-.05)	6 (.03-.12)	0 (.00-.03)	12 (.07-.19)
Gujaratis and Konkanastha Br. (111) <sup>e</sup>	.99	0 (.00-.03)	53 (.39-.57)	5 (.02-.10)	7 (.03-.13)	0 (.00-.03)	0 (.00-.03)	20 (.12-.26)	5 (.02-.10)	11 (.06-.17)	5 (.02-.10)	12 (.06-.18)
Kerala/Karnataka (99) <sup>g</sup>	.96	0 (.00-.03)	63 (.54-.72)	15 (.09-.24)	6 (.03-.13)	15 (.09-.24)	0 (.00-.03)	20 (.14-.29)	1 (.01-.05)	...	0 (.00-.03)	9 (.05-.16)
Lambadis (86) <sup>h</sup>	.99	0 (.00-.03)	55 (.53-.73)	9 (.06-.19)	4 (.02-.11)	9 (.06-.19)	0 (.00-.03)	9 (.06-.19)	2 (.01-.08)	7 (.04-.16)	0 (.00-.03)	11 (.07-.22)
Lobanas (62) <sup>h</sup>	.98	0 (.00-.05)	34 (.43-.67)	3 (.02-.13)	3 (.02-.13)	5 (.04-.18)	0 (.00-.05)	2 (.01-.11)	1 (.00-.09)	5 (.04-.18)	0 (.00-.05)	11 (.10-.29)
Punjabis (112) <sup>e</sup>	.99	0 (.00-.03)	46 (.32-.50)	1 (.00-.05)	5 (.02-.10)	1 (.00-.05)	0 (.00-.03)	15 (.08-.21)	8 (.04-.14)	21 (.13-.27)	6 (.03-.11)	11 (.06-.17)
Sri Lanka (132) <sup>e</sup>	.99	0 (.00-.02)	77 (.50-.66)	9 (.04-.13)	6 (.02-.10)	2 (.01-.05)	0 (.00-.02)	19 (.09-.21)	5 (.02-.09)	11 (.05-.14)	2 (.01-.05)	18 (.09-.21)
Telugu, upper (59) <sup>i</sup>	.99	0 (.00-.05)	36 (.48-.72)	3 (.02-.14)	11 (.11-.30)	0 (.00-.05)	0 (.00-.05)	10 (.10-.29)	1 (.00-.09)	2 (.01-.11)	0 (.00-.05)	9 (.08-.27)
Telugu, middle (114) <sup>i</sup>	.99	0 (.00-.03)	73 (.55-.72)	7 (.03-.12)	4 (.01-.09)	5 (.02-.10)	0 (.00-.03)	10 (.05-.15)	1 (.01-.05)	6 (.03-.11)	0 (.00-.03)	24 (.15-.29)
Telugu, lower (70) <sup>i</sup>	.99	0 (.00-.04)	50 (.60-.81)	7 (.05-.19)	1 (.00-.08)	3 (.02-.12)	0 (.00-.04)	5 (.03-.16)	0 (.00-.04)	1 (.00-.08)	0 (.00-.04)	15 (.14-.32)
Uttar Pradesh (139) <sup>e,h</sup>	.99	0 (.00-.02)	79 (.49-.65)	4 (.01-.07)	14 (.06-.16)	0 (.00-.02)	0 (.00-.02)	21 (.10-.22)	3 (.01-.06)	9 (.04-.12)	2 (.00-.05)	20 (.10-.21)
West Asian:												
Iranians (440) <sup>e</sup>	.99	2 (.00-.02)	24 (.04-.08)	0 (.00-.01)	5 (.01-.03)	0 (.00-.01)	0 (.00-.02)	41 <sup>j</sup> (.07-.12)	90 (.17-.25)	245 (.51-.60)	2 (.00-.02)	17 (.02-.06)
Turks, Cappadocia (388) <sup>e</sup>	.99	1 (.00-.01)	16 (.03-.07)	0 (.00-.01)	0 (.00-.01)	0 (.00-.01)	0 (.00-.01)	7 (.01-.04)	93 (.20-.29)	244 (.58-.68)	1 (.00-.01)	7 (.01-.04)
Middle East (406) <sup>e,k</sup>	.99	41 (.08-.13)	30 (.05-.10)	2 (.00-.02)	1 (.00-.01)	0 (.00-.01)	0 (.00-.03)	15 (.02-.06)	40 (.07-.13)	269 (.62-.71)	1 (.00-.01)	3 (.00-.02)

<sup>a</sup> U2i excludes variants of U2 with 16129C; U\* = other derivatives of haplogroup U; R\* = derivatives of haplogroup R that do not belong to HV, TJ, U, B, and F.

<sup>b</sup> HVS-I haplotype diversity.

<sup>c</sup> Including pre-HV, as defined by Saillard et al. 2000b.

<sup>d</sup> None of the Indian samples belonged to haplogroup B.

<sup>e</sup> From the present study.

<sup>f</sup> From Roychoudhury et al. 2001.

<sup>g</sup> From Moutain et al. 1995.

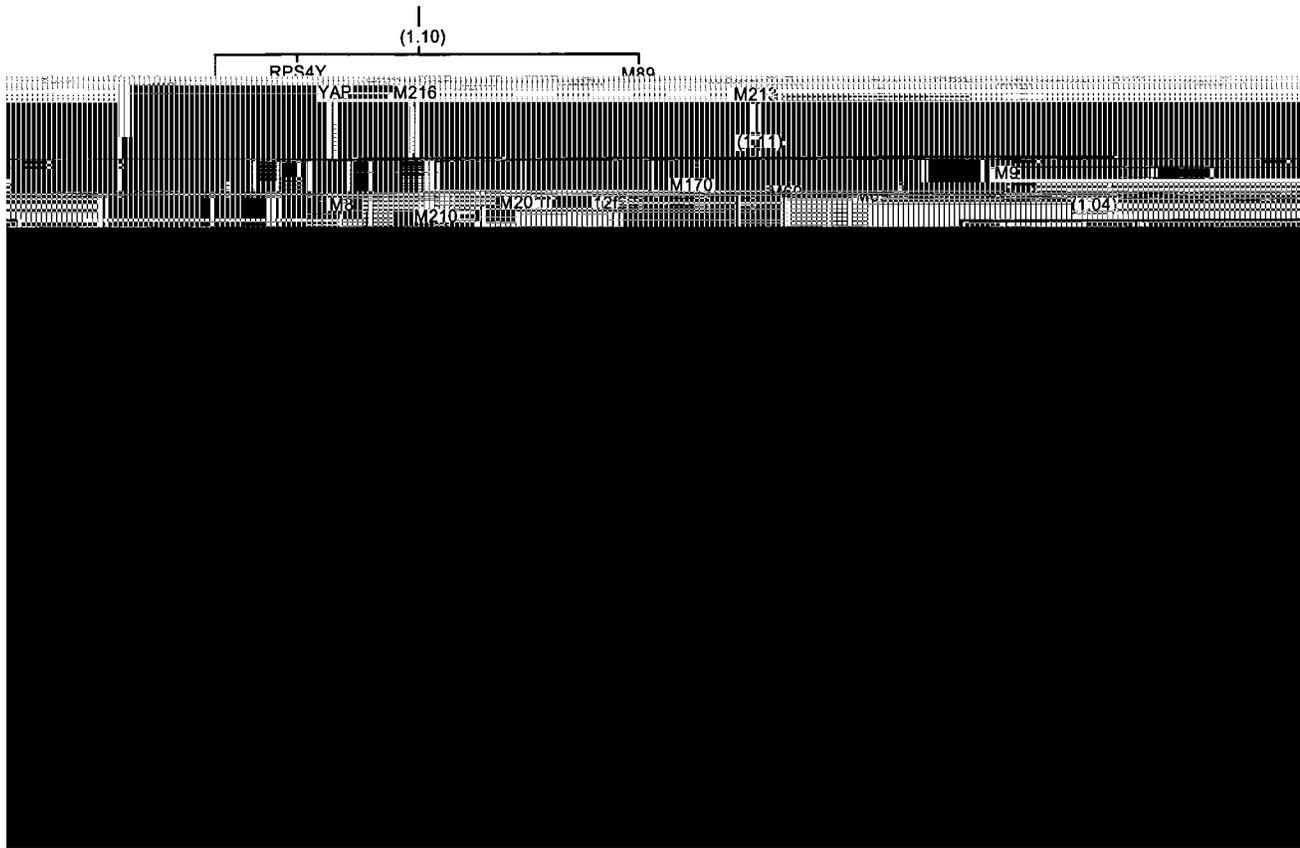
<sup>h</sup> From Kivisild et al. 1999.

<sup>i</sup> From Bamshad et al. 2001.

<sup>j</sup> Forty Iranians belonged to U7 and one belonged to U2i.

<sup>k</sup> The Middle East sample includes 202 Kuwaitis and 204 Saudi Arabians.





**Figure 3** Y-chromosomal SNP tree and haplogroup frequencies in 8 Indian populations. Haplogroup defining markers (and their background average variances of 6 STR loci) are shown along the branches of the tree.

Table 3

## Major Y-Chromosomal Haplogroups in India Compared with Western Eurasia

POPULATION (n)	FREQUENCY (95% CR FOR PROPORTION)			REFERENCE
	H1 (M52)	L (M11)	R2 (M124)	
India:				
Punjab (66)	.03 (.01–.10)	.12 (.06–.22)	.05 (.02–.13)	Present study
West Bengal (31)	.10 (.04–.25)	.00 (.00–.09)	.23 (.12–.40)	Present study
AP tribes (82)	.49 (.43–.64)	.07 (.04–.15)	.04 (.01–.10)	Present study
AP tribes (67)	.10 (.05–.20)	.04 (.02–.12)	...	Ramana et al. 2001
AP castes (125)	.10 (.06–.16)	.08 (.04–.14)	...	Ramana et al. 2001
Tamil Nadu (259)	.17 (.13–.22)	.29 (.24–.35)	.07 (.05–.11)	Wells et al. 2001
Tadjikistan (168)	.01 (.00–.04)	.11 (.07–.16)	.06 (.03–.11)	Wells et al. 2001
Uzbeks (366)	.02 (.01–.04)	.03 (.02–.05)	.02 (.01–.04)	Wells et al. 2001
Kyrgyzstan (92)	.00 (.00–.03)	.00 (.00–.03)	.02 (.01–.08)	Wells et al. 2001
Kazakistan (95)	.01 (.00–.06)	.01 (.00–.06)	.01 (.00–.06)	Wells et al. 2001
Iran (52)	.00 (.00–.06)	.04 (.01–.13)	.02 (.01–.10)	Wells et al. 2001
Near East (101)	.01 (.00–.05)	.02 (.01–.07)	.00 (.00–.03)	Semino et al. 2000
Caucasus (147)	.00 (.00–.02)	.01 (.00–.05)	.00 (.00–.02)	Wells et al. 2001
Europe (839)	.00 (.00–.01)	.01 (.00–.01)	.00 (.00–.01)	Semino et al. 2000

(table 2). Furthermore, 32% of the Koya M\* HVS-I sequences shared an A at hypervariable np 16129, which is characteristic of a likely polyphyletic HVS-I clade M5 (Bamshad et al. 2001). The loss of 12403 *MnII*, one of the four defining markers of African M1 cluster (Maca-Meyer et al. 2001), was not found in either tribal sample (table 1).

A 9-bp deletion between COII/trRNA<sup>Lys</sup> occurs in high frequency in eastern Asian and some African populations, because of its independent origins at different phylogenetic backgrounds (Soodyall et al. 1996). It was shown recently that some Indian populations also harbor the 9-bp deletion while clustering separately from Asian and African deleted lineages (Watkins et al. 1999). We found that 21% of Koyas and 3% of Chenchus harbored the deletion at the haplogroup M background. The Chenchu type (16184-16223-16256G-16362) has been previously observed at notable frequencies (44%) among Irulas, another tribe from Andhra Pradesh with australoid anthropological features (Watkins et al. 1999). The presence of the 9-bp deletion at the haplogroup M background was also observed among Kadars of Tamil Nadu and Kerala (Edwin et al. 2002). The HVS-I motif associated with the 9-bp deletion in Koyas has not been observed in previously published studies. Whether the Koya and Chenchu 9-bp deletion types stem from the same deletion event is difficult to judge. They differ by seven HVS-I mutations, suggesting either an ancient common root or independent origins of the deletion.

The haplogroup R lineages of the Koyas (31%) and Chenchus (1%) did not further subdivide into western Eurasian-specific (HV, U, TJ, and R1; Macaulay et al. 1999) or eastern Eurasian-specific branches (B and R9; Kivisild et al. 2002) and showed a coalescence time of 73,000 ± 20,900 years, which overlaps with the age es-

timate 0(o)0(uEint)]TJojig70(Chenlj/F7 1 Tfi41(and)-365(showe65(tatio

**Table 4**

**Compound Y-Chromosomal Haplotypes in Chenchus and Koyas**

HAPLOTYPE	CLADE <sup>a</sup>	NO. OF REPEATS AT LOCUS						CHENCHU	KOYA	MATCH <sup>b</sup>
		DYS019	DYS388	DYS390	DYS391	DYS392	DYS393			
1	C	15	13	23	10	11	12	1		
2	C	16	13	24	10	11	12	1		
3	F	15	13	21	11	11	14		1	
4	F	15	14	21	11	11	14		1	
5	F	16	13	21	10	12	14		1	
6	F	16	13	21	11	10	14		1	
7	F	16	13	21	11	11	13		1	
8	F	16	13	21	11	11	14		2	
9	F	16	14	21	11	11	14		1	
10	F	17	13	21	11	10	14		2	
11	F	17	13	21	11	11	13		1	
12	H1	13	12	23	11	11	12	1		
13	H1	14	12	21	9	11	12	1		
14	H1	14	12	22	10	11	12		2	
15	H1	14	13	22	10	11	13		1	
16	H1	14	13	22	11	11	12		1	
17	H1	15	12	21	9	11	12	7		
18	H1	15	12	21	10	11	12	2		
19	H1	15	12	21	11	11	12	1		
20	H1	15	12	22	10	11	12	3	11	*
21	H1	15	12	22	10	11	13		3	*
22	H1	15	13	21	11	11	12		1	
23	H1	15	13	22	10	11	13		2	
24	H1	15	13	22	11	11	12		1	
25	H1	16	12	22	10	11	13		1	
26	H1	16	13	22	10	11	12		2	
27	H2	16	12	21	10	11	13		4	*
28	J2*	14	15	23	10	11	13	1		
29	J2e	15	14	24	11	11	12	1		*
30	J2e	15	15	24	10	11	12	1		*
31	L	14	12	22	10	13	11	2		
32	L	14	12	22	10	14	11	4		*
33	R1a	15	12	24	11	11	12		1	
34	R1a	15	12	25	10	11	13	1		*
35	R1a	15	12	26	10	11	13	1		*
36	R1a	16	12	24	10	11	13	1		*
37	R1a	17	12	24	11	11	13	1		*
38	R1a	16	12	24	11	11	13	7		*
39	R1b	14	13	24	11	13	13	1		*
40	R2	14	12	22	10	10	14	1		
41	R2	14	12	23	10	10	14	1		
42	R2	14	13	22	10	10	14	1		
Total								41	41	

<sup>a</sup> According to YCC nomenclature.

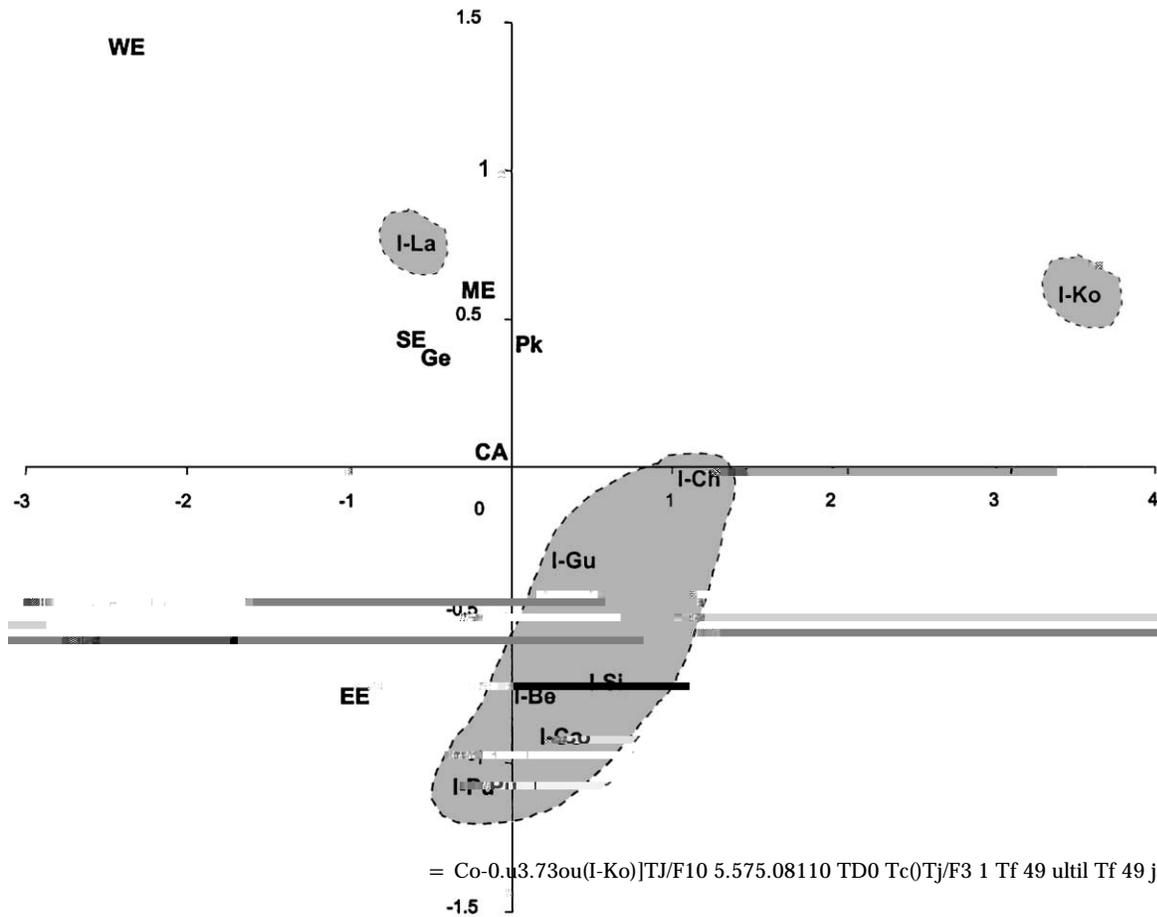
<sup>b</sup> Compared to 239 Indian caste samples.

in Middle Eastern populations to the west (table 3). Unclassified derivatives of the general Eurasian clade F were observed most frequently (27%) in the Koyas.

In comparison with caste groups (see fig. 3 and table 3), both tribal populations showed significantly ( $P < .01$ ) higher frequencies of haplogroup H1. The characteristic M52 A→C transversion has also been described at relatively high frequencies in populations of Tamil Nadu, in southern India (Wells et al. 2001). Among the caste groups, its frequency is the lowest among Punjabis

in the northwest. Interestingly, more than one-third of Andhra Pradesh middle and lower caste Y chromosomes were defined as clade 1R in a previous study (Bamshad

**Table 5**



**Figure 4** Multidimensional scaling plot of eight Indian and seven western Eurasian populations, using  $F_{st}$  distances calculated for 16 Y-chromosomal SNP haplogroups. From India: I-Ch = Chenchus, I-Ko

**Table 6**

**MX1 Haplotypes of Chromosome 21 in Indian Populations, Compared with Continental Groups of the World**

POPULATION	NO. OF LINEAGES (95% CR FOR PROPORTION)									
	Ht1	Ht2	Ht3	Ht4	Ht5	Ht6	Ht7	Ht8	Ht9	Ht10
Chenchus (84) <sup>a</sup>	0 (.00-.04)	11 (.08-.22)	0 (.00-.04)	0 (.00-.04)	16 (.12-.29)	0 (.00-.04)	40 (.37-.58)	17 (.13-.30)	0 (.00-.04)	0 (.00-.04)
Koyas (56) <sup>a</sup>	0 (.00-.05)	4 (.03-.17)	0 (.00-.05)	0 (.00-.05)	15 (.17-.40)	0 (.00-.05)	16 (.18-.42)	21 (.26-.51)	0 (.00-.05)	0 (.00-.05)
Punjab (68) <sup>a</sup>	0 (.00-.04)	10 (.08-.25)	0 (.00-.04)	0 (.00-.04)	7 (.05-.20)	0 (.00-.04)	31 (.34-.57)	20 (.20-.41)	0 (.00-.04)	0 (.00-.04)
West Bengal (70) <sup>a</sup>	1 (.00-.08)	2 (.01-.10)	0 (.00-.04)	0 (.00-.04)	16 (.15-.34)	0 (.00-.04)	34 (.37-.60)	17 (.16-.36)	0 (.00-.04)	0 (.00-.04)
Pakistan (72) <sup>b</sup>	0 (.00-.04)	5 (.03-.15)	0 (.00-.04)	0 (.00-.04)	6 (.04-.17)	0 (.00-.04)	27 (.27-.49)	34 (.36-.59)	0 (.00-.04)	0 (.00-.04)
Anatolia (70) <sup>a</sup>	2 (.01-.10)	7 (.05-.19)	0 (.00-.04)	0 (.00-.04)	3 (.02-.12)	1 (.00-.08)	29 (.31-.53)	28 (.29-.52)	0 (.00-.04)	0 (.00-.04)
Europe (192) <sup>b</sup>	4 (.01-.05)	17 (.06-.14)	0 (.00-.02)	0 (.00-.02)	3 (.01-.05)	0 (.00-.02)	88 (.39-.53)	79 (.34-.48)	1 (.00-.03)	0 (.00-.02)
East Asia (118) <sup>b</sup>	0 (.00-.03)	3 (.01-.07)	0 (.00-.03)	0 (.00-.03)	59 (.41-.59)	11 (.05-.16)	45 (.30-.47)	0 (.00-.03)	0 (.00-.03)	0 (.00-.03)
Sub-Saharan Africa (102) <sup>b</sup>	9 (.05-.16)	3 (.01-.08)	0 (.00-.03)	0 (.00-.03)	22 (.15-.31)	23 (.16-.32)	33 (.24-.42)	9 (.05-.16)	0 (.00-.03)	3 (.01-.08)
Amerinds (120) <sup>b</sup>	0 (.00-.02)	14 (.07-.19)	0 (.00-.02)	0 (.00-.02)	78 (.56-.73)	0 (.00-.02)	2 (.01-.06)	26 (.15-.30)	0 (.00-.02)	0 (.00-.02)
Australia and PNG (76) <sup>b</sup>	2 (.01-.09)	10 (.07-.23)	16 (.13-.32)	2 (.01-.09)	27 (.26-.47)	0 (.00-.04)	15 (.12-.30)	4 (.02-.13)	0 (.00-.04)	0 (.00-.04)

<sup>a</sup> From the present study.

<sup>b</sup> From Jin et al. 1999.

dian populations considered here. Similarly, haplotype 8, which is common in Europe but absent in eastern Asia, was found in India at low frequencies. As is the case in other Eurasian and African populations, haplotypes 3 and 4, which are specific to Australian and Papuan populations, were not found in India. In contrast to the significant differences of haplotype frequencies that were observed between Indian and other world populations, none of the differences in haplotype frequencies was significant within India between caste and tribal groups.

## Discussion

Phylogeographic patterns of the Y chromosome and mtDNA support the concept that the Indian subcontinent played a pivotal role in the late Pleistocene genetic differentiation of the western and eastern Eurasian gene pools. All non-Africans, including Indian populations, have inherited a subset of African mtDNA haplogroup L3 lineages, differentiated into groups M and N. Al-





- Female gene flow stratifies Hindu castes. *Nature* 395:651–652
- Bandelt H-J, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16:37–48
- Bandelt H-J, Forster P, Sykes BC, Richards MB (1995) Mitochondrial portraits of human populations using median networks. *Genetics* 141:743–753
- Bergen AW, Wang CY, Tsai J, Jefferson K, Dey C, Smith KD, Park SC, Tsai S-J, Goldman D (1999) An Asian–Native American paternal lineage identified by RPS4Y resequencing and by microsatellite haplotyping. *Ann Hum Genet* 63:63–80
- Bertorelle G, Excoffier L (1998) Inferring admixture proportions from molecular data. *Mol Biol Evol* 15:1298–1311
- Bhattacharyya NP, Basu P, Das M, Pramanik S, Banerjee R, Roy B, Roychoudhury S, Majumder PP (1999) Negligible male gene flow across ethnic boundaries in India, revealed by analysis of Y-chromosomal DNA polymorphisms. *Genome Res* 9:711–719
- Bhowmick PK (1992) *The Chenchus of the forests and plateaux: Calcutta*. Institute of Anthropology, Calcutta
- Cavalli-Sforza LL, Menozzi P, Piazza A (1994) *The history and geography of human genes*. Princeton University Press, Princeton
- Chikhi L, Bruford MW, Beaumont MA (2001) Estimation of admixture proportions: a likelihood-based approach using Markov chain Monte Carlo. *Genetics* 158:1347–1362
- Chikhi L, Nichols RA, Barbujani G, Beaumont MA (2002) Y genetic data support the Neolithic demic diffusion model. *Proc Natl Acad Sci USA* 99:11008–11013
- Cruciani F, Santolamazza P, Shen P, Macaulay V, Moral P, Olckers A, Modiano D, Holmes S, Destro-Bisol G, Coia V, Wallace DC, Oefner PJ, Torroni A, Cavalli-Sforza LL, Scozzari R, Underhill PA (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
- Edwin D, Vishwanathan H, Roy S, Usha Rani M, Majumder P (2002) Mitochondrial DNA diversity among five tribal populations of southern India. *Curr Sci* 83:158–163
- Forster P, Harding R, Torroni A, Bandelt H-J (1996) Origin and evolution of Native American mtDNA variation: a reappraisal. *Am J Hum Genet* 59:935–945
- Forster P, Torroni A, Renfrew C, Röhl A (2001) Phylogenetic star contraction applied to Asian and Papuan mtDNA evolution. *Mol Biol Evol* 18:1864–1881
- Hammer MF, Redd AJ, Wood ET, Bonner MR, Jarjanazi H, Karafet T, Santachiara-Benerecetti S, Oppenheim A, Jobling MA, Jenkins T, Ostrer H, Bonne-Tamir B (2000) Jewish and Middle Eastern non-Jewish populations share a common pool of Y-chromosome biallelic haplotypes. *Proc Natl Acad Sci USA* 97:6769–6774
- Helgason A, Sigurðadóttir S, Gulcher J, Ward R, Stefánson K (2000a) mtDNA and the origins of the Icelanders: deciphering signals of recent population history. *Am J Hum Genet* 66:999–1016
- Helgason A, Sigurðadóttir S, Nicholson J, Sykes B, Hill EW, Bradley DG, Bosnes V, Gulcher JR, Ward R, Stefánsson K (2000b) Estimating Scandinavian and Gaelic ancestry in the male settlers of Iceland. *Am J Hum Genet* 67:697–717
- Herrnstadt C, Elson JL, Fahy E, Preston G, Turnbull DM, Anderson C, Ghosh SS, Olefsky JM, Beal MF, Davis RE, Howell N (2002) Reduced-median-network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups. *Am J Hum Genet* 70:1152–1171
- Ingman M, Kaessmann H, Pääbo S, Gyllensten U (2000) Mitochondrial genome variation and the origin of modern humans. *Nature* 408:708–713
- Jin L, Underhill PA, Doornik V, Davis RW, Shen P, Cavalli-Sforza LL, Oefner PJ (1999) Distribution of haplotypes from a chromosome 21 region distinguishes multiple prehistoric human migrations. *Proc Natl Acad Sci USA* 96:3796–3800
- Kaessmann H, Heissig F, von Haeseler A, Pääbo S (1999) DNA sequence variation in a non-coding region of low recombination on the human X chromosome. *Nat Genet* 22:78–81
- Kalaydjieva L, Calafell F, Jobling MA, Angelicheva D, de Knijff P, Rosser ZH, Hurles ME, Underhill P, Tournev I, Marushiakova E, Popov V (2001) Patterns of inter- and intra-group genetic diversity in the Vlach Roma as revealed by Y chromosome and mitochondrial DNA lineages. *Eur J Hum Genet* 9:97–104
- Karafet T, Xu L, Du R, Wang W, Feng S, Wells RS, Redd AJ, Zegura SL, Hammer MF (2001) Paternal population history of East Asia: sources, patterns, and microevolutionary processes. *Am J Hum Genet* 69:615–628
- Karafet TM, Zegura SL, Posukh O, Osipova L, Bergen A, Long J, Goldman D, Klitz W, Harihara S, deKnijff P, Wiebe V, Griffiths RC, Templeton AR, Hammer MF (1999) Ancestral Asian source(s) of new world Y-chromosome founder haplotypes. *Am J Hum Genet* 64:817–831
- Kayser M, Brauer S, Weiss G, Schiefenovel W, Underhill PA, Stoneking M (2001) Independent histories of human Y chromosomes from Melanesia and Australia. *Am J Hum Genet* 68:173–190
- Kayser M, Brauer S, Weiss G, Underhill PA, Roewer L, Schiefenovel W, Stoneking M (2000) Melanesian origin of Polynesian Y chromosomes. *Curr Biol* 10:1237–1246
- Ke Y, Su B, Song X, Lu D, Chen L, Li H, Qi C, Marzuki S, Deka R, Underhill P, Xiao C, Shriver M, Lell J, Wallace D, Wells RS, Seielstad M, Oefner P, Zhu D, Jin J, Huang W, Chakraborty R, Chen Z, Jin L (2001) African origin of modern humans in East Asia: a tale of 12,000 Y chromosomes. *Science* 292:1151–1153
- Kivisild T, Bamshad MJ, Kaldma K, Metspalu M, Metspalu E, Reidla M, Laos S, Parik J, Watkins WS, Dixon ME, Papiha SS, Mastana SS, Mir MR, Ferak V, VILLEMS R (1999a) Deep common ancestry of Indian and western-Eurasian mitochondrial DNA lineages. *Curr Biol* 9:1331–1334
- Kivisild T, Kaldma K, Metspalu M, Parik J, Papiha SS, VILLEMS R (1999b) The place of the Indian mitochondrial DNA variants in the global network of maternal lineages and the peopling of the Old World. In: Deka R, Papiha SS (eds) *Genomic diversity*. Kluwer/Academic/Plenum Publishers, New York, pp 135–152
- Kivisild T, Papiha SS, Rootsi S, Parik J, Kaldma K, Reidla M, Laos S, Metspalu M, Pielberg G, Adojaan M, Metspalu E,

- Mastana SS, Wang Y, Gölge M, Demirtas H, Schnakenberg E, De Stefano GF, Geberhiwot T, Claustres M, Villems R (2000) An Indian ancestry: a key for understanding human diversity in Europe and beyond. In: Renfrew C, Boyle K (eds) *Archaeogenetics: DNA and the population prehistory of Europe*. McDonald Institute for Archaeological Research, University of Cambridge, Cambridge, pp 267–279
- Kivisild T, Tolk H-V, Parik J, Wang Y, Papiha SS, Bandelt H-J, Villems R (2002) The emerging limbs and twigs of the east Asian mtDNA tree. *Mol Biol Evol* 19:1737–1751
- Maca-Meyer N, Gonzalez AM, Larruga JM, Flores C, Cabrera VM (2001) Major genomic mitochondrial lineages delineate early human expansions. *BMC Genet* 2:13
- Macaulay VA, Richards MB, Hickey E, Vega E, Cruciani F, Guida V, Scozzari R, Bonn -Tamir B, Sykes B, Torroni A (1999) The emerging tree of west Eurasian mtDNAs: a synthesis of control-region sequences and RFLPs. *Am J Hum Genet* 64:232–249
- Majumder PP (2001) Indian caste origins: genomic insights and future outlook. *Genome Res* 11:931–932
- Majumder P, Dey B (2001) Absence of the HIV-1 protective Delta ccr5 allele in most ethnic populations of India. *Eur J Hum Genet* 9:794–796
- Mountain JL, Hebert JM, Bhattacharyya S, Underhill PA, Ottolenghi C, Gadgil M, Cavalli-Sforza LL (1995) Demographic history of India and mtDNA-sequence diversity. *Am J Hum Genet* 56:979–992
- Nebel A, Filon D, Brinkmann B, Majumder PP, Faerman M, Oppenheim A (2001) The Y chromosome pool of Jews as part of the genetic landscape of the Middle East. *Am J Hum Genet* 69:1095–1112
- Nebel A, Landau-Tasserion E, Filon D, Oppenheim A, Faerman M (2002) Genetic evidence for the expansion of Arabian tribes into the Southern Levant and North Africa. *Am J Hum Genet* 70:1594–1596
- Nei M, Roychoudhury A (1993) Evolutionary relationships of human populations on a global scale. *Mol Biol Evol* 10:927–943
- Papiha SS (1996) Genetic variation in India. *Hum Biol* 68:607–628
- Passarino G, Semino O, Magri C, Al-Zahery N, Benuzzi G, Quintana-Murci L, Andellnovic S, Bulc-Jakus F, Liu A, Arslan A, Santachiara-Benerecetti AS (2001) The 49a,f haplotype 11 is a new marker of the EU19 lineage that traces migrations from northern regions of the Black Sea. *Hum Immunol* 62:922–932
- Passarino G, Semino O, Modiano G, Bernini LF, Santachiara Benerecetti AS (1996) mtDNA provides the first known marker distinguishing proto-Indians from the other Caucasoids; it probably predates the diversification between Indians and Orientals. *Ann Hum Biol* 23:121–126
- Qamar R, Ayub Q, Mohyuddin A, Helgason A, Mazhar K, Mansoor A, Zerjal T, Tyler-Smith C, Mehdi SQ (2002) Y-chromosomal DNA variation in Pakistan. *Am J Hum Genet* 70:1107–1124
- Quintana-Murci L, Krausz C, Zerjal T, Sayar SH, Hammer MF, Mehdi SQ, Ayub Q, Qamar R, Mohyuddin A, Radhakrishna U, Jobling MA, Tyler-Smith C, McElreavey K (2001) Y-chromosome lineages trace diffusion of people and languages in southwestern Asia. *Am J Hum Genet* 68:537–542
- Quintana-Murci L, Semino O, Bandelt H-J, Passarino G, McElreavey K, Santachiara-Benerecetti AS (1999) Genetic evidence of an early exit of *Homo sapiens sapiens* from Africa through eastern Africa. *Nat Genet* 23:437–441
- Raitio M, Lindroos K, Laukkanen M, Pastinen T, Sistonen P, Sajantila A, Syvanen A (2001) Y-chromosomal SNPs in Finno-Ugric-speaking populations analyzed by minisequencing on microarrays. *Genome Res* 11:471–482
- Ramana GV, Su B, Jin L, Singh L, Wang N, Underhill P, Chakraborty R (2001) Y-chromosome SNP haplotypes suggest evidence of gene flow among caste, tribe, and the migrant Siddi populations of Andhra Pradesh, South India. *Eur J Hum Genet* 9:695–700
- Redd A, Roberts-Thomson J, Karafet T, Bamshad M, Jorde L, Naidu J, Walsh B, Hammer MF (2002) Gene flow from the Indian subcontinent to Australia: evidence from the Y chromosome. *Curr Biol* 12:673–677
- Richards M, Macaulay V, Hickey E, Vega E, Sykes B, Guida V, Rengo C, et al (2000) Tracing European founder lineages in the near eastern mtDNA pool. *Am J Hum Genet* 67:1251–1276
- Rosser ZH, Zerjal T, Hurles ME, Adojaan M, Alavantic D, Amorim A, Amos W, et al (2000) Y-chromosomal diversity

- Shouse B (2001) Archaeology: spreading the word, scattering the seeds. *Science* 294:988–989
- Shriver MD, Smith MW, Jin L, Marcini A, Akey JM, Deka R, Ferrell RE (1997) Ethnic-affiliation estimation by use of population-specific DNA markers. *Am J Hum Genet* 60:957–964
- Singh KS (1997) The scheduled tribes. In: Singh KS (ed) *People of India*. Vol III. Oxford University Press, Oxford
- Soodyall H, Vigilant L, Hill AV, Stoneking M, Jenkins T (1996) mtDNA control-region sequence variation suggests multiple independent origins of an “Asian-specific” 9-bp deletion in sub-Saharan Africans. *Am J Hum Genet* 58:595–608
- Stringer C (2000) Coasting out of Africa. *Nature* 405:24–25
- Su B, Xiao C, Deka R, Seielstad MT, Kangwanpong D, Xiao J, Lu D, Underhill P, Cavalli-Sforza LL, Chakraborty R, Jin L (2000) Y chromosome haplotypes reveal prehistorical migrations to the Himalayas. *Hum Genet* 107:582–590
- Su B, Xiao J, Underhill P, Deka R, Zhang W, Akey J, Huang W, Shen D, Lu D, Luo J, Chu J, Tan J, Shen P, Davis R, Cavalli-Sforza L, Chakraborty R, Xiong M, Du R, Oefner P, Chen Z, Jin L (1999) Y-chromosome evidence for a northward migration of modern humans into Eastern Asia during the last ice age. *Am J Hum Genet* 65:1718–1724
- Tarn W (1951) *The Greeks in Bactria & India*. Cambridge University Press, Cambridge
- Thangaraj K, Ramana GV, Singh L (1999) Y-chromosome and mitochondrial DNA polymorphisms in Indian populations. *Electrophoresis* 20:1743–1747
- Thurin M (1999) The Chenchu of the Indian Deccan. In: Lee RB, Daily R (eds) *The Cambridge encyclopedia of hunters and gatherers*. Cambridge University Press, Cambridge, pp 252–256
- Underhill PA, Jin L, Lin AA, Mehdi SQ, Jenkins T, Vollrath D, Davis RW, Cavalli-Sforza LL, Oefner PJ (1997) Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Res* 7:996–1005
- Underhill PA, Passarino G, Lin AA, Marzuki S, Oefner PJ, Cavalli-Sforza LL, Chambers GK (2001a) Maori origins, Y-chromosome haplotypes and implications for human history in the Pacific. *Hum Mutat* 17:271–280
- Underhill PA, Passarino G, Lin AA, Shen P, Mirazon Lahr M, Foley R, Oefner PJ, Cavalli-Sforza LL (2001a) The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Ann Hum Genet* 65:43–62
- Underhill PA, Shen P, Lin AA, Jin L, Passarino G, Yang WH, Kauffman E, Bonn -Tamir B, Bertranpetit J, Francalacci P, Ibrahim M, Jenkins T, Kidd JR, Mehdi SQ, Seielstad MT, Wells RS, Piazza A, Davis RW, Feldman MW, Cavalli-Sforza LL, Oefner PJ (2000) Y chromosome sequence variation and the history of human populations. *Nat Genet* 26:358–361
- Walter H, Danker-Hopfe H, Bhasin MK (1991) *Anthropologie Indiens*. Gustav Fischer Verlag, Stuttgart
- Watkins WS, Bamshad M, Dixon ME, Bhaskara Rao B, Naidu JM, Reddy PG, Prasad BV, Das PK, Reddy PC, Gai PB, Bhanu A, Kusuma YS, Lum JK, Fischer P, Jorde LB (1999) Multiple origins of the mtDNA 9-bp deletion in populations of South India. *Am J Phys Anthropol* 109:147–158
- Weale ME, Yepiskoposyan L, Jager RF, Hovhannisyan N, Khudoyan A, Burbage-Hall O, Bradman N, Thomas MG (2001) Armenian Y chromosome haplotypes reveal strong regional structure within a single ethno-national group. *Hum Genet* 109:659–674
- Wells RS, Yuldasheva N, Ruzibakiev R, Underhill PA, Evseeva I, Blue-Smith J, Jin L, et al (2001) The Eurasian heartland: a continental perspective on Y-chromosome diversity. *Proc Natl Acad Sci USA* 98:10244–10249
- Wilson JF, Weiss DA, Richards M, Thomas MG, Bradman N, Goldstein DB (2001) Genetic evidence for different male and female roles during cultural transitions in the British Isles. *Proc Natl Acad Sci USA* 98:5078–5083
- Y Chromosome Consortium, The (2002) A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Res* 12:339–348
- Zerjal T, Pandya A, Santos FR, Adhikari R, Tarazona E, Kayser M, Evgafov O, Singh L, Thangaraj K, Destro-Bisol G, Thomas MG, Qamar R, Mehdi SQ, Rosser ZH, Hurles ME, Jobling MA, Tyler-Smith C (1999) The use of Y-chromosomal DNA variation to investigate population history: recent male spread in Asia and Europe. In: Papiha S, Deka R, Chakraborty R (eds) *Genomic diversity: applications in human population genetics*. Kluwer Academic/Plenum Publishers, New York, pp 91–101
- Zerjal T, Wells R, Yuldasheva N, Ruzibakiev R, Tyler-Smith C (2002) A genetic landscape reshaped by recent events: Y-chromosomal insights into central Asia. *Am J Hum Genet* 71:466–482