Int J Legal Med (2011) 125:367–375
DOI 10.1007/s00414-010-0485-x

ORIGINAL ARTICLE

Y-STR diversity in the Himalayas

Tenzin Gayden · Shilpa Chennakrishnaiah · Joel La Salvia · Sacha Jimenez · Maria Regueiro · Trisha Maloney · Patrice J. Persad · Areej Bukhari · Annabel Perez · Oliver Stojkovic · Rene J. Herrera

Received: 16 April 2010 / Accepted: 28 June 2010 / Published online: 21 July 2010

Abstract Linguistic and ethnic diversity throughout the Himalayas suggests that this mountain range played an important role in shaping the genetic landscapes of the region. Previous Y-chromosome work revealed that the Himalayas acted as a biased bidirectional barrier to gene flow across the cordillera. In the present study, 17 Y-chromosomal short tandem repeat (Y-STR) loci included in the AmpF*l*STR® Yfiler kit were analyzed in 344 unrelated males from three Nepalese populations (Tamang, Newar, and Kathmandu) and a general collection from Tibet. The latter displays the highest haplotype diversity (0.9990) followed by Kathmandu (0.9977), Newar (0.9570), and Tamang (0.9545). The overall haplotype diversity for the Himalayan populations at 17 Y-STR loci was 0.9973, and the corresponding values for the extended (11 loci) and minimal (nine loci) haplotypes were 0.9955 and 0.9942, respectively. No Y-STR profiles are shared across the four Himalayan collections at the 17-, 11-, and nine-locus resolutions considered, indicating a lack of recent gene flow among them. Phylogenetic analyses support our previous findings that Kathmandu, and to some extent Newar, received significant genetic influence from India while Tamang and Tibet exhibit limited or no gene flow from the subcontinent. A median-joining network of haplogroup O3a3c-M134 based on 15 Y-STR loci from our four Himalayan populations suggests either a male founder effect in Tamang, possibly from Tibet, or a recent bottleneck following their arrival south of the Himalayas from Tibet leading to their highly reduced Y single-nucleotide polymorphism and Y-STR diversity. The genetic uniqueness of the four Himalayan populations examined in this study merits the creation of separate databases for individual identification, parentage analysis, and population genetic studies.

Keywords Himalayas · Nepal · Tibet · Y-STR

Electronic supplementary material The online version of this article (doi:10.1007/s00414-010-0485-x) contains supplementary material, which is available to authorized users.

T. Gayden · T. Maloney · P. J. Persad · A. Bukhari · A. Perez · R. J. Herrera
Department of Biological Sciences,
Florida International University,
Miami, FL, USA 33199

T. Gayden · S. Chennakrishnaiah · J. La Salvia · S. Jimenez · M. Regueiro · T. Maloney · P. J. Persad · A. Bukhari · A. Perez · R. J. Herrera (✉)
Department of Molecular and Human Genetics,
College of Medicine, Florida International University,
University Park, OE 304,
Miami, FL 33199, USA
e-mail: herrerar@fiu.edu

S. Chennakrishnaiah · J. La Salvia · S. Jimenez
Department of Chemistry and Biochemistry,
Florida International University,
Miami, FL, USA 33199

O. Stojkovic
Institute of Forensic Medicine, Faculty of Medicine,
University of Belgrade,
Deligradska 31,
11000 Belgrade, Serbia

Introduction

Binary polymorphisms, including single-nucleotide polymorphisms (SNPs) and insertions/deletions, located within the non-recombining region of the Y-chromosome have been deemed useful for tracing ancient paternal lineages [1–3]. Y-chromosome short tandem repeats (Y-STRs), on the other hand, mutate rapidly and are more suitable for reconstructing the phylogeny of relatively recent demographic events as well as for forensic casework and paternity testing [2, 4–6].

Recently, Y-STR data have been increasingly used to investigate the evolution, migration, and genetic diversity of modern human populations [2, 6–11]. This interest in Y-STRs as markers for ancestry and population studies have prompted the creation of a comprehensive worldwide Y-STR database called Y-Chromosome Haplotype Reference Database (YHRD, www.yhrd.org) for calculating haplotype frequencies, matching probabilities, and performing comparative population genetic analyses [12–15]. This global repository currently supports most frequently used haplotype formats, namely, nine-locus minimal (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, and DYS385a/b), SWGDAM-recommended 11-locus extended (minimal + DYS438 and DYS439), 12-locus Powerplex (SWGDAM + DYS437), and 17-locus Yfiler haplotypes [1].

Our previous studies of Y-chromosomal biallelic [16] and autosomal STR polymorphisms [17] of Tibet and Nepal have revealed that these Himalayan groups arrived in the area during the Neolithic time from Northeast Asia with subsequent gene flow from the Indian subcontinent into the Kathmandu valley and Newar population. The latter conclusion is congruent with recent mtDNA studies [18, 19], which reported shared maternal lineage between Indian and Nepalese populations. In contrast, Tibet and the Tamang population display limited influence from the Indian subcontinent, suggesting that the Himalayan massif has acted as a barrier for gene flow from the south into the Tibetan plateau [16, 17]. In addition to the Northeast Asian influence, the high frequency of the Asian-specific *Alu* insertion at the Y *Alu* polymorphism locus in the Tibetan Y-chromosomes had previously led some researchers to argue for Central Asian contribution in the Tibetan gene pool [20–22].

Although several studies based on Y-STR data from Tibet have been previously published [23–26], the present study is the first of its kind to perform comprehensive phylogenetic analyses involving a considerable sample size. The knowledge of the phylogenetic relationships among populations is essential in order to assess whether or not populations should be considered as a separate entity as databases for forensic analysis and paternity testing. This work also improves on an earlier Nepalese study [27] by examining populations that are anthropologically well characterized. In addition, this report complements previous Y-chromosomal biallelic [16] and autosomal STR [17] data, thus providing a comprehensive analysis of the genetic diversity in the Himalayas. In the present study, 17 Y-STR loci were typed in the three Nepalese populations of Tamang, Newar, and Kathmandu as well as a general collection from Tibet to investigate their genetic ancestry and phylogenetic relationships to previously published geographically targeted groups from Northeast Asia, Southeast Asia, South Central Asia, and Central Asia using nine-locus minimal haplotypes (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, and DYS385a/b). Furthermore, the haplotype data generated from the 17 Y-STR loci were utilized to ascertain the values of parameters of forensic interests.

Materials and methods

Sample collection and DNA isolation

Blood samples were collected with informed consent from 344 unrelated males from Tibet (156) and three populations from Nepal (188), namely Tamang (45), Newar (66), and Kathmandu (77). These samples have been previously typed for Y-SNP markers in one of our earlier studies [16] (Supplementary Table 1). Genealogical information of the donors was recorded for a minimum of two generations to ascertain their paternal ancestry. Sample collections were performed in strict compliance with the ethical guidelines put forth by the institutions involved in this study. DNA was extracted by the standard phenol–chloroform method and ethanol precipitated as previously described [28] and stored at −80°C.

Table 1 Populations analyzed

Population	Abbreviation	Number	References
Northeast Asia			
Osaka	OSA	131	Hashiyada et al. 2008 [36]
Korea	KOR	252	Kwak et al. 2005 [2]
Shandong Han (NE China)	CSH	131	Yan et al. 2007 [37]
Southeast Asia			
Malaysia	MAL	334	Chang et al. 2007 [38]
Philippines	PHI	76	Kwak et al. 2005 [2]
Taiwan	TAI	200	Huang et al. 2008 [39]
Thailand	THA	41	Kwak et al. 2005 [2]
Vietnam	VIE	43	Kwak et al. 2005 [2]
South Central Asia			
Bangladesh	BAN	72	Dobashi et al. 2005 [40]
Haryana	HAR	84	Nagy et al. 2007 [41]
Punjab	PUN	80	Nagy et al. 2007 [41]
Central Asia			
Mongolia	MON	92	Kwak et al. 2006 [2]
Buryat	BUR	215	Woźniak et al. 2006 [42]
Himalayas			
Bhutan	BHU	856	Parkin et al. 2006 [43]
Nepal (general)	NEP	765	Parkin et al. 2007 [27]
Newar	NEW	66	present study
Kathmandu	KAT	77	present study
Lhasa	LHA	112	Zhang et al. 2006 [23]
Tamang	TAM	45	Present study
Tibet	TIB	156	Present study

DNA amplification and STR genotyping

DNA samples were amplified at 17 Y-STR loci in a multiplex reaction using the AmpF*l*STR® Yfiler kit [29]. PCR amplifications were performed as specified by the manufacturer using the recommended DNA amounts (0.5-1.25 ng) in an Eppendorf Master gradient cycler (Eppendorf AG, Germany). Amplicons were separated by multicapillary electrophoresis in an ABI Prism 3130 Genetic Analyzer, and the ABI GeneScan 500 LIZ internal size standard was used as a basis for comparison. The software GeneMapper® v3.1 [29] was employed to determine fragment sizes, while alleles were designated by comparison to an allelic ladder supplied by the manufacturer. The nomenclature of the Y-STR loci studied is as recommended by the DNA Commission of the International Society of Forensic Genetics for analysis of Y-STR systems [30] while the Y-SNP nomenclature followed is in accordance with the Y-Chromosome Consortium [31, 32].

Statistical and phylogenetic analyses

Allelic frequencies were calculated by the gene counting method [6]. Gene and haplotype diversities were computed using the software package Arlequin v. 3.1 [33]. Chromosomes carrying null alleles or duplicated loci were excluded from the haplotype calculation at the 17 Y-STR loci level. Discrimination capacity (DC) and fraction of unique haplotypes (FUH) were estimated as percentage proportions of different and unique haplotypes, respectively, in a given population. All the statistical parameters were calculated for both minimal and extended haplotypes for comparison purposes. The dw_{min} (minimum diversity within the population), mw_{min} (minimum matching probability within the population), mw_{max} (maximum matching probability within the population), and mb_{min} (minimum matching probability between two populations) were calculated as previously described [34, 35]. The ratio mw_{max}/mb_{min} gives an upper estimate of how much more probable it is to find a match within a population than between two populations, and the ratio mw_{min}/mb_{min} provides a lower estimate for the same parameter.

A total of 16 reference populations (Table 1) from the published literature were included for comparison across nine-locus minimal haplotypes (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, and DYS385a/b), since the data for the remaining loci typed in this study were not reported for all reference collections. Correspondence analysis (CA) plots were generated using NTSYSpc 2.02i [44], while a neighbor joining (NJ) dendrogram was built utilizing the PHYLIP v3.6 program [45]. The statistical robustness of the phylogenetic relationships within the NJ tree was assessed using bootstrap analysis involving 1,000 replications. A median-joining network based on Y-haplogroup O3a3c-M134 chromosomes in the Himalayan populations was constructed utilizing the 15 Y-STR loci, excluding the bilocal marker DYS385a/b, using the program NETWORK 4.2.0.0 (www.fluxus-engineering.com). For network calculations, microsatellite loci were weighted inversely to their variance such that higher weights were assigned to the least variable loci [22, 46]. The network generated was postprocessed by employing the maximum parsimony parameter [47, 48] to obtain the simplest possible projection.

Results and discussion

Allelic frequencies and gene diversities for the 17 Y-STR loci analyzed in the Tamang, Newar, Kathmandu, and Tibet collections are listed in Supplementary Tables 2 to 5. Tamang exhibits a relatively (to the other three Himalayan populations) high degree of genetic homogeneity, with 11 of the 17 loci displaying gene diversities lower than 0.5, whereas Kathmandu represents the other extreme with all loci registering values above 0.5 for the same parameter. As expected, Kathmandu also possesses the highest average gene diversity (0.6657) followed by Newar (0.6411), Tibet (0.6352), and Tamang (0.4195).

Null alleles were detected in two Newari and four Tibetan samples at locus DYS448 as previously described [6, 27], while five Kathmandu males were null at DYS458, consistent with the location of a deletion in the short arm of the Y-chromosome encompassing the amelogenin locus (Supplementary Table 1) [16, 49]. The frequency of the amelogenin Y deletion in the Kathmandu populace (6.49%) is higher than levels previously reported in Nepal (1.2%) [27] and India (1.8%) [50] but lower than that of Sri Lankan males (8%) [51]. Duplication was observed in one individual from Newar at locus DYS458 (alleles 16, 18). All null alleles and duplications were confirmed by repeating the amplification process.

Table 2 shows the number of different and unique haplotypes, FUH, DC, and haplotype diversity for the three haplotype resolutions considered: Yfiler (17 loci), extended (11 loci), and minimal (nine loci) haplotypes. Overall, a total of 262 different haplotypes were identified, of which 228 (66.28%) were unique to a single individual (Table 2). Chromosomes carrying null alleles ($n=11$) or duplicated loci ($n=1$) were excluded from the haplotype calculation at the 17 Y-STR loci level. When analyzed at the 11- and nine-locus haplotype resolutions, the number of different haplotypes decreased to 236 and 214, respectively, among which 189 (54.94%) and 161 (46.80%) were distinct haplotypes, respectively (Table 2). The significant increase in the proportion of unique haplotypes using the Yfiler system (66.28%) compared to the minimal haplotype (46.80%) reflects the power of discrimination at 17 Y-STR loci.

Table 2 Parameters of forensic interest in Himalayan populations using the nine-locus, 11-locus, and the Yfiler haplotypes

Haplotypes	Tamang	Newar	Kathmandu	Tibet	All populations
9-locus Y-STR haplotype					
Sample size	45	66	77	156	344
Number of different haplotypes	18	30	61	116	214
Number of unique haplotypes	10	18	49	95	161
Fraction of unique haplotypes	0.2222	0.2727	0.6364	0.6090	0.468
Discrimination capacity	0.4000	0.4545	0.7922	0.7436	0.6221
Haplotype diversity±SD	0.9010±0.0253	0.9585±0.0097	0.9932±0.0034	0.9940±0.0018	0.9942±0.0009
11-locus Y-STR haplotype					
Sample size	45	66	77	156	344
Number of haplotypes	19	30	66	129	236
Unique haplotypes	12	18	56	115	189
Fraction of unique haplotypes	0.2667	0.2727	0.7272	0.7372	0.5494
Discrimination capacity	0.4222	0.4545	0.8571	0.8462	0.6860
Haplotype diversity±SD	0.9020±0.0256	0.9585±0.0097	0.9959±0.0029	0.9967±0.0013	0.9955±0.0008
17-locus Y-STR haplotype					
Sample size[a]	45	63	72	152	332
Number of different haplotypes	27	29	67	141	262
Number of unique haplotypes	21	17	63	130	228
Fraction of unique haplotypes	0.4667	0.2698	0.8750	0.8553	0.6867
Discrimination capacity	0.6000	0.4603	0.9306	0.9276	0.7892
Haplotype diversity ± SD	0.9545±0.0167	0.9570±0.0105	0.9977±0.0029	0.9990±0.0009	0.9970±0.0007

[a] Excludes samples carrying null alleles and duplicated loci

The overall haplotype diversity for the Himalayan population at 17 Y-STR loci was 0.9973 while the corresponding values for the extended and minimal haplotypes were 0.9955 and 0.9942, respectively (Table 2). The latter is slightly lower than values previously reported for East Asian (0.9996, n=700) [2] and European (0.9976, n=11,610) populations [14]. This relatively lower diversity may be attributed to the reduced heterogeneity observed in Tamang (0.9010) and Newar (0.9585), which in turn may be the result of bottlenecks and/or founder effects in the case of former and genetic drift in the latter. The genetic homogeneity in Tamang is also reflected in their reduced average gene diversity (0.4195), consistent with previous reports [16, 17]. In addition, the discrimination capacity in Tamang and Newar are considerably lower than that of Kathmandu and Tibet at all the three haplotype resolutions (Table 2).

Table 3 Y-STR haplotype matching probabilities within and between the Himalayan populations

Parameters	Tamang (Tam)	Newar (New)	Kathmandu (Kat)	Tibet (Tib)
N[a]	45	63	72	152
dw_{min}	0.9333	0.9418	0.9838	0.9926
mw_{max}	0.0667	0.0582	0.0162	0.0074
mw_{min}	45/990	56/1,953	6/2,556	9/11,476
mb_{min}	Tam/New 0.0000	New/Kat 0.0002	Kat/Tib 0.0000	Tib/Tam 0.0000
	Tam/Kat 0.0018	New/Tib 0.0000	Kat/Tam 0.0018	Tib/New 0.0000
	Tam/Tib 0.0000	New/Tam 0.0000	Kat/New 0.0002	Tib/Kat 0.0000
mw_{max}/mb_{min}	Tam/New 0.00	New/Kat 291	Kat/Tib 0.00	Tib/Tam 0.00
	Tam/Kat 37.06	New/Tib 0.00	Kat/Tam 9.00	Tib/New 0.00
	Tam/New 0.00	New/Tam 0.00	Kat/New 81.00	Tib/Kat 0.00
mw_{min}/mb_{min}	Tam/New 0.00	New/Kat 143.50	Kat/Tib 0.00	Tib/Tam 0.00
	Tam/Kat 25.25	New/Tib 0.00	Kat/Tam 1.28	Tib/New 0.00
	Tam/New 0.00	New/Tam 0.00	Kat/New 11.50	Tib/Kat 0.00

[a] Excludes samples carrying null alleles and duplicated loci

Table 3 shows the Y-STR haplotype matching probabilities within and between the Himalayan populations. Similar to the limited haplotype and gene diversities values, Tamang (0.0667) shows the highest maximum match probability within the population followed by Newar (0.0582), Kathmandu (0.0162), and Tibet (0.0072) (Table 3). When compared among the four Himalayan collections, the maximum probabilities ($db_{max,}$ which is $1-mb_{min}$) of obtaining two different Y-haplotypes when sampling a pair of individuals from Tamang and Kathmandu, Tamang and Newar,

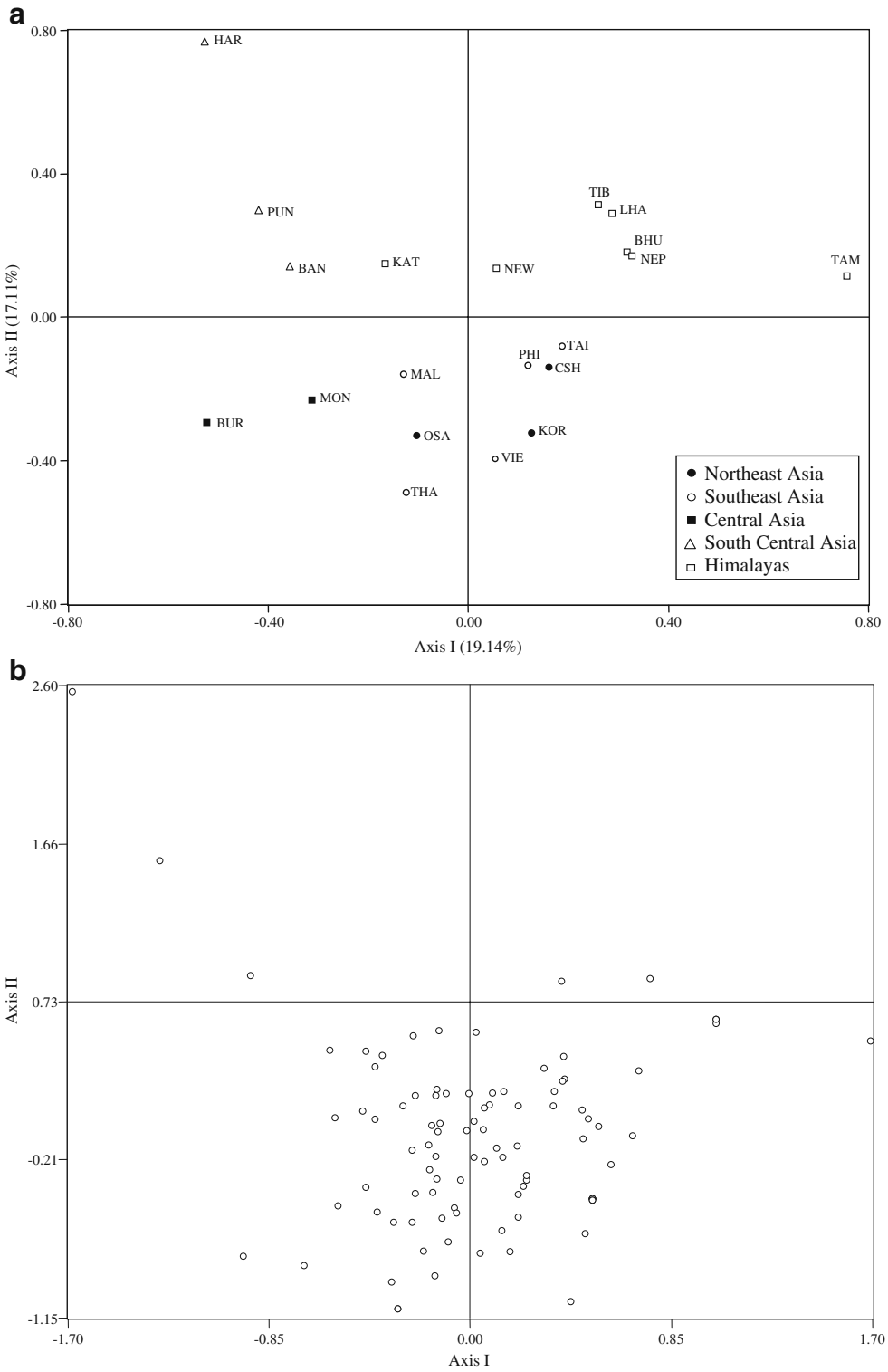

Fig. 1 **a** Correspondence analysis based on allelic frequencies of nine Y-STR loci from 20 populations; **b** contribution of each of the 94 alleles of the nine Y-STR loci to the partitions of populations in **a**

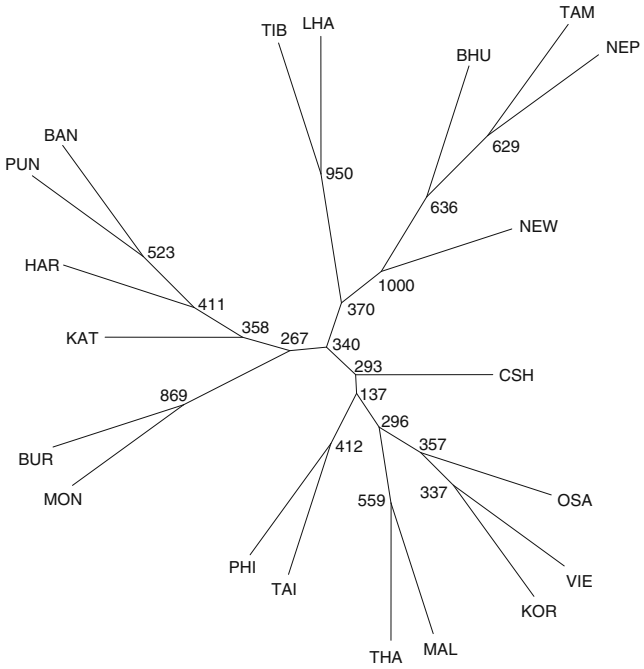

Fig. 2 NJ tree based on Nei's genetic distance. The numbers at the nodes represent bootstrap values estimated from 1,000 replications

Tamang and Tibet, Kathmandu and Newar, Kathmandu and Tibet, and Newar and Tibet are 99.82%, 100%, 100%, 99.98%, 100%, and 100%, respectively. This high power of discrimination underscores the genetic uniqueness of the four Himalayan populations as well as the need for a well-defined ethnic and geographic sampling of populations for forensic applications.

There are no profiles shared across all four collections in any of the three data sets. Of the 34 shared haplotypes at the 17 Y-STR loci level (Table 2), 32 are common within a single population while two haplotypes are shared between two Nepalese groups—one between Newar and Kathmandu and another between Tamang and Kathmandu. However, when examined at the minimal nine-locus resolution, two haplotypes are shared across three Himalayan populations, with the exception of Newar. These two haplotypes differ from each other by a one-step mutation at the bilocal marker DYS385a/b (13, 18 vs. 13, 19 alleles) and one of them (DYS385a/b = 13, 19 alleles) is the most frequent minimal haplotype observed in our Himalayan populations at a frequency of 3.2% (14-12-28-23-10-14-12-13-19). This most common haplotype is represented by six Tibetans, two Tamangs, and two individuals from Kathmandu. Comparison of the latter haplotype with the YHRD database (Release 32) returned 75 exact matches, the majority of which are from Bhutan (26), China (19; 13 of the 19 are from Eastern China), and Nepal (17), while the remaining matches are from East Asia, including South Korea (four), Japan (three), and Malaysia (two), with the exceptions of two Indian males, an admixed Hispanic American, and a Tibetan from Lhasa. It is not

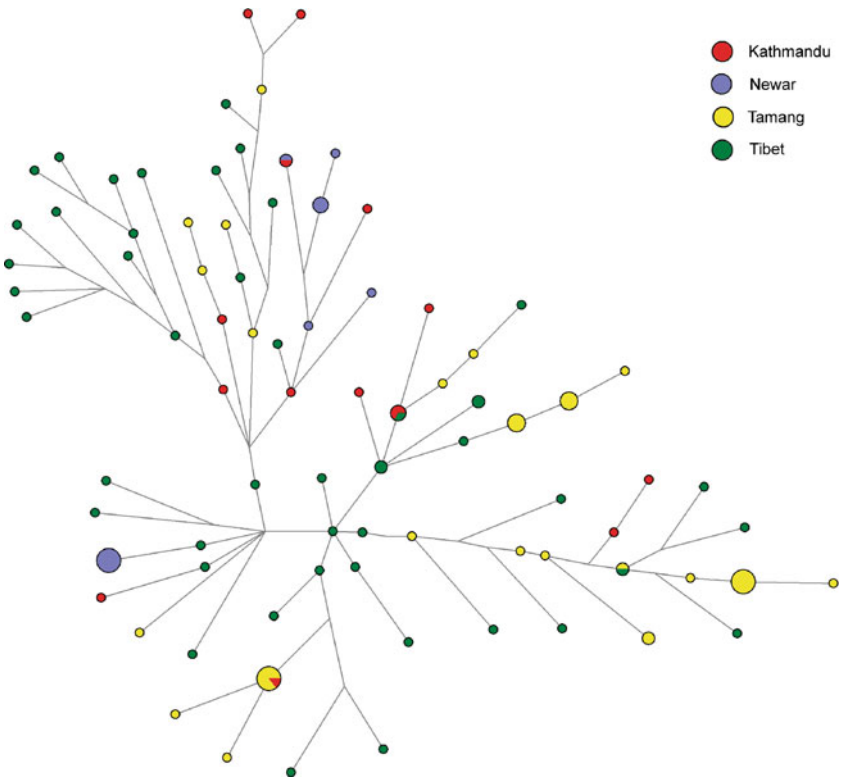

Fig. 3 A median-joining network based on 15 Y-STR loci within the Y-haplogroup O3a3c-M134 of the four Himalayan populations. The area of the circles is proportional to the haplotype frequency and the smallest circle corresponds to one Y-chromosome

surprising to find matching profiles from Nepalese samples in the YHRD database since three of the populations under study are from the same region. Matches with Bhutanese and Chinese samples, however, suggest recent gene flow and/or shared common patrimonies among these groups concordant with previous findings [16, 17, 20, 21, 43].

Phylogenetic relationships between the four Himalayan collections and other geographically targeted populations were assessed using CA (Fig. 1) and NJ (Fig. 2) analyses. Figure 1b represents the contribution of each of 94 alleles of nine Y-STR loci to the partition of the populations. The Himalayans cluster loosely in the upper right quadrant of the plot (Fig. 1a) and share the same clade in the tree (Fig. 2), with the exception of Kathmandu which maps closer to the South Central Asian group in both the analyses [16, 17]. Newar also seems to display some affinity to the South Central Asian assemblage along the *X*-axis, whereas Tamang and Haryana are outliers from their respective groups (Fig. 1a). The genetic similarities observed within the Himalayas based on their Y-STR loci (Figs. 1a and 2) are reflected in the high frequencies of Y-haplogroup O3a3c-M134, common among Tibeto-Burman speakers [16, 21, 52]. This inference is supported by the fact that the most frequent minimal haplotype (14-12-28-23-10-14-12-13-19), as well as its one-step mutation neighbor at DYS385 (DYS385a/b = 13, 18 allele), both belong to haplogroup O3a3c-M134 [16]. On the other hand, Kathmandu and Newar's affinity to the South Central Asian cluster in the CA and NJ tree (Figs. 1a and 2, respectively) may be due to the presence of Indian Y-lineages (haplogroups R, H, and C5) in their gene pools [16].

The East Asian collections map toward the middle of the lower half of the graph while the Mongolians and Buryats segregate to the left of the chart (Fig. 1a). There is no clear genetic partitioning between the northern and southern East Asian populations in both the CA and NJ tree (Figs. 1a and 2) [2]. Overall, the NJ dendrogram mirrors the distributions of populations in the CA with the exceptions of Mongolia and Buryat which form a sister clade with the South Central Asian branch and the general population of Nepal showing more affinities with Tamang than with Bhutan (Fig. 2).

The lack of phylogenetic affinities exhibited by Tamang in relation to Tibet (Figs. 1a and 2) is of interest given its proposed close genetic association with the latter in previous studies [16, 17]. Although both Tamang and Tibet share high frequencies of haplogroup O3a3c-M134 [16], their Y-STR profiles differ considerably. In order to gain further insight into the recent demographic history of these two groups, a median-joining network based solely on Y-haplogroup O3a3c-M134 was constructed at the level of the 15 Y-STR loci utilizing our four Himalayan populations (Fig. 3). It is notable that Tamang and Newar form distinct clusters due to their shared or closely related haplotypes, while Tibet and Kathmandu are highly divergent (Fig. 3). This finding suggests either a male founder effect in Tamang, possibly from Tibet, or a recent bottleneck event as they settled south of the Himalayas from Tibet, leading to their highly reduced Y-SNP [18] and Y-STR diversity (Table 2). On the other hand, Newar's unique genetic profile may be due to isolation and/or drift [17].

The 17 Y-STR loci data obtained in this study from the four Himalayan populations were submitted to the YHRD (www.yhrd.org) and they were assigned to the East Asian—Sino-Tibetan—Tibeto-Burman metapopulation [1].

In summary, our results confirmed previous Y-chromosomal and autosomal STR reports that Newar and Kathmandu experienced substantial gene flow from India whereas Tamang and Tibet display no genetic influences from the subcontinent. A median-joining network of haplogroup O3a3c-M134 based on 15 Y-STR loci suggests recent bottleneck and/or founder effect in Tamang. A high value of combined haplotype diversity (0.9970) from the four Himalayan populations is indicative of genetic heterogeneity within the region. In addition, very high percentages (99.82–100%) of the maximum probability (db_{max}) of finding two different Y-haplotypes when sampling a pair of individuals between two different Himalayan populations underscores the genetic singularity of the four Himalayan collections reported. The uniqueness of our four Himalayan populations argues for independent databases for forensic analysis and paternity testing.

Acknowledgments The authors gratefully acknowledge Sheyla Mirabal and Kristian J. Herrera for their constructive criticism of the manuscript.

References

1. Willuweit S, Roewer L, International Forensic Y Chromosome User Group (2007) Y chromosome haplotype reference database (YHRD): update. Forensic Sci Int Genet 1:83–87
2. Jobling MA, Tyler-Smith C (1993) The human Y chromosome: an evolutionary marker comes of age. Nat Rev Genet 8:598–612
3. Kwak KD, Jin HJ, Shin DJ, Kim JM, Roewer L, Krawczak M, Tyler-Smith C, Kim W (2005) Y-chromosomal STR haplotypes and their applications to forensic and population studies in east Asia. Int J Leg Med 119:195–201
4. Gayden T, Regueiro M, Martinez L, Cadenas AM, Herrera RJ (2008) Human Y-chromosome haplotyping by allele-specific polymerase chain reaction. Electrophoresis 29:2419–2423
5. Iida R, Kishi K (2005) Identification, characterization and forensic application of novel Y-STRs. Leg Med 7:255–258
6. Hanson EK, Ballantyne J (2007) An ultra-high discrimination Y chromosome short tandem repeat multiplex DNA typing system. PLoS ONE 8:e688
7. Budowle B, Ge J, Low J, Lai C, Yee WH, Law G, Tan WF, Chang YM, Perumal R, Keat PY, Mizuno N, Kasai K, Sekiguchi K,

Chakraborty R (2009) The effects of Asian population substructure on Y STR forensic analyses. Leg Med 11:64–69
8. Alves C, Gusmão L, Barbosa J, Amorim A (2003) Evaluating the informative power of Y-STRs: a comparative study using European and new African haplotype data. Forensic Sci Int 34:126–133
9. Roewer L, Croucher PJ, Willuweit S, Lu TT, Kayser M, Lessig R, de Knijff P, Jobling MA, Tyler-Smith C, Krawczak M (2005) Signature of recent historical events in the European Y-chromosomal STR haplotype distribution. Hum Genet 116:279–291
10. Immel UD, Krawczak M, Udolph J, Richter A, Rodig H, Kleiber M, Klintschar M (2006) Y-chromosomal STR haplotype analysis reveals surname-associated strata in the East-German population. Eur J Hum Genet 14:577–582
11. Rebała K, Mikulich AI, Tsybovsky IS, Siváková D, Dzupinková Z, Szczerkowska-Dobosz A, Szczerkowska Z (2007) Y-STR variation among Slavs: evidence for the Slavic homeland in the middle Dnieper basin. J Hum Genet 52:406–414
12. Rodríguez V, Tomàs C, Sánchez JJ, Castro JA, Ramon MM, Barbaro A, Morling N, Picornell A (2009) Genetic sub-structure in western Mediterranean populations revealed by 12 Y chromosome STR loci. Int J Leg Med 123:137–141
13. Roewer L, Krawczak M, Willuweit S et al (2001) Online reference database of European Y-chromosomal short tandem repeat (STR) haplotypes. Forensic Sci Int 118:106–113
14. Kayser M, Brauer S, Willuweit S et al (2002) Online Y-chromosomal short tandem repeat haplotype reference database (YHRD) for U.S. populations. J Forensic Sci 47:513–519
15. Lessig R, Willuweit S, Krawczak M et al (2003) Asian online Y-STR haplotype reference database. Leg Med 5:S160–S163
16. Gayden T, Cadenas AM, Regueiro M, Singh NB, Zhivotovsky LA, Underhill PA, Cavalli-Sforza LL, Herrera RJ (2007) The Himalayas as a directional barrier to gene flow. Am J Hum Genet 80:884–894
17. Gayden T, Mirabal S, Cadenas AM, Lacau H, Simms TM, Morlote D, Chennakrishnaiah S, Herrera RJ (2009) Genetic insights into the origins of Tibeto-Burman populations in the Himalayas. J Hum Genet 54:216–223
18. Thangaraj K, Chaubey G, Kivisild T et al (2008) Maternal footprints of Southeast Asians in North India. Hum Hered 66:1–6
19. Fornarino S, Pala M, Battaglia V, Maranta R, Achilli A, Modiano G, Torroni A, Semino O, Santachiara-Benerecetti SA (2009) Mitochondrial and Y-chromosome diversity of the Tharus (Nepal): a reservoir of genetic variation. BMC Evol Biol 9:154
20. Qian Y, Qian B, Su B et al (2000) Multiple origins of Tibetan Y chromosomes. Hum Genet 106:453–454
21. Su B, Xiao C, Deka R et al (2000) Y chromosome haplotypes reveal prehistorical migrations to the Himalayas. Hum Genet 107:582–590
22. Hammer MF, Karafet TM, Park H, Omoto K, Harihara S, Stoneking M, Horai S (2006) Dual origins of the Japanese: common ground for hunter–gatherer and farmer Y chromosomes. J Hum Genet 51:47–58
23. Zhang Q, Yan J, Tang H, Jiao Z, Liu Y (2006) Genetic polymorphisms of 17 Y-STRs haplotypes in Tibetan ethnic minority group of China. Leg Med 8:300–305
24. Zhu B, Liu S, Ci D et al (2006) Population genetics for Y-chromosomal STRs haplotypes of Chinese Tibetan ethnic minority group in Tibet. Forensic Sci Int 161:78–83
25. Kang LL, Liu K, Ma Y (2007) Y chromosome STR haplotypes of Tibetan living Tibet Lassa. Forensic Sci Int 172:79–83
26. Zhu B, Wu Y, Shen C et al (2008) Genetic analysis of 17 Y-chromosomal STRs haplotypes of Chinese Tibetan ethnic group residing in Qinghai province of China. Forensic Sci Int 175:238–243
27. Parkin EJ, Kraaijenbrink T, Opgenort JRML, van Driem GL, Tuladhar NM, de Knijff P, Jobling MA (2007) Diversity of 26-locus Y-STR haplotypes in a Nepalese population sample: isolation and drift in the Himalayas. Forensic Sci Int 166:176–181
28. Antunez de Mayolo G, Antúñez de Mayolo A, Antúñez de Mayolo P (2002) Phylogenetics of worldwide human populations as determined by polymorphic Alu insertions. Electrophoresis 23:3346–3356
29. Applied Biosystems (2006) AmpF/STR Yfiler PCR amplification kit user's manual. Applied Biosystems, Foster City
30. Gusmão L, Butler JM, Carracedo A, Gill P, Kayser M, Mayr WR, Morling N, Prinz M, Roewer L, Tyler-Smith C, Schneider PM, DNA Commission of the International Society of Forensic Genetics (2006) DNA Commission of the International Society of Forensic Genetics (ISFG): an update of the recommendations on the use of Y-STRs in forensic analysis. Forensic Sci Int 157:187–197
31. Y Chromosome Consortium (2002) A nomenclature system for the tree of human Y-chromosomal binary haplogroups. Genome Res 12:339–348
32. Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, Hammer MF (2008) New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. Genome Res 18:830–838
33. Excoffier L, Laval LG, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. Evol Bioinform Online 1:47–50
34. Brinkmann C, Forster P, Schürenkamp M, Horst J, Rolf B, Brinkmann B (1999) Human Y-chromosomal STR haplotypes in a Kurdish population sample. Int J Leg Med 112:181–183
35. Pfeiffer H, Brinkmann B, Hühne J, Rolf B, Morris AA, Steighner R, Holland MM, Forster P (1999) Expanding the forensic German mitochondrial DNA control region database: genetic diversity as a function of sample size and microgeography. Int J Leg Med 112:291–298
36. Hashiyada M, Umetsu K, Yuasa I et al (2008) Population genetics of 17 Y-chromosomal STR loci in Japanese. Forensic Sci Int Genet 2: e69–e70
37. Yan J, Tang H, Liu Y et al (2007) Genetic polymorphisms of 17 Y-STRs haplotypes in Chinese Han population residing in Shandong province of China. Leg Med 94:196–202
38. Chang YM, Perumal R, Keat PY, Kuehn DL (2007) Haplotype diversity of 16 Y-chromosomal STRs in three main ethnic populations (Malays, Chinese and Indians) in Malaysia. Forensic Sci Int 167:70–76
39. Huang TY, Hsu YT, Li JM, Chung JH, Shun CT (2008) Polymorphism of 17 Y-STR loci in Taiwan population. Forensic Sci Int 174:249–254
40. Dobashi Y, Kido A, Fujitani N, Susukida R, Hara M, Oya M (2005) Y-chromosome STR haplotypes in a Bangladeshi population. Leg Med 7:122–126
41. Nagy M, Henke L, Henke J et al (2007) Searching for the origin of Romanies: Slovakian Romani, Jats of Haryana and Jat Sikhs Y-STR data in comparison with different Romani populations. Forensic Sci Int 169:19–26
42. Woźniak M, Derenko M, Malyarchuk B et al (2006) Allelic and haplotypic frequencies at 11 Y-STR loci in Buryats from South-East Siberia. Forensic Sci Int 164:271–275
43. Parkin EJ, Kraayenbrink T, van Driem GL, Tshering Of Gaselô K, de Knijff P, Jobling MA (2006) 26-Locus Y-STR typing in a Bhutanese population sample. Forensic Sci Int 161:1–7
44. Rohlf F (2002) NTSYSpc. Exter, Setauket
45. Felsentein J (2002) Phylogeny inference package (PHYLIP), version 3.6a3. Department of Genetics, University of Washington, Seattle, distributed by author
46. Kayser M, Brauer S, Weiss G et al (2003) Reduced Y-chromosome, but not mitochondrial DNA, diversity in human populations from West New Guinea. Am J Hum Genet 72:281–302
47. Polzin T, Daneschmand SV (2002) On Steiner trees and minimum spanning trees in hypergraphs. Oper Res Lett 31:12–20

48. Ferri G, Robino C, Alù M et al. (2008) Molecular characterisation and population genetics of the DYS458.2 allelic variant. Forensic Sci Int Genet: Genetics Supplement Series, Progress in Forensic Genetics 12—Proceedings of the 22nd International ISFG Congress, vol. 1, pp 203–205
49. Cadenas AM, Regueiro M, Gayden T, Singh N, Zhivotovsky LA, Underhill PA, Herrera RJ (2007) Male amelogenin dropouts: phylogenetic context, origins and implications. Forensic Sci Int 166:155–163
50. Thangaraj K, Reddy AG, Singh L (2002) Is the amelogenin gene reliable for gender identification in forensic casework and prenatal diagnosis? Int J Leg Med 116:121–123
51. Santos FR, Pandya A, Tyler-Smith C (1998) Reliability of DNA-based sex tests. Nat Genet 18:103
52. Wen B, Xie X, Gao S (2004) Analyses of genetic structure of Tibeto-Burman populations reveals sex-biased admixture in southern Tibeto-Burmans. Am J Hum Genet 74:856–865