Int J Legal Med (2005) 119: 195–201
DOI 10.1007/s00414-004-0518-4

ORIGINAL ARTICLE

Kyoung Don Kwak · Han Jun Jin · Dong Jik Shin · Jung Min Kim · Lutz Roewer · Michael Krawczak · Chris Tyler-Smith · Wook Kim

Y-chromosomal STR haplotypes and their applications to forensic and population studies in east Asia

Received: 16 March 2004 / Accepted: 8 December 2004 / Published online: 27 April 2005

Abstract We have analyzed 11 Y-STR loci (DYS19, the two DYS385 loci, DYS388, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DXYS156Y) in 700 males from ten ethnic groups in east Asia in order to evaluate their usefulness for forensic and population genetic studies. A total of 644 different haplotypes were identified, among which 603 (86.14%) were individual-specific. The haplotype diversity averaged over all populations was 0.9997; using only the nine Y-STRs comprising the "minimal haplotype" (excluding DYS388 and DXYS156Y) it was 0.9996, a value similar to that found in 1924 samples from other Asian populations (0.9996; Lessig et al. Legal Medicine 5 (2003) 160–163), and slightly higher than in European populations (0.9976; n=11,610; Roewer et al. For Sci International (2001) 118:103–111). All of the individual east Asian populations examined here had high haplotype diversity (≥0.997), except for the Mongolians (0.992) and Manchurians (0.960). The most frequent haplotype identified by the nine markers was present at only 1% (7/700). Population comparisons based on Φ_{ST} or ρ genetic distance measures revealed clustering according to the traditional northeast–southeast distinction, but with exceptions. For example, the Yunnan population from southern China lay among the northern populations, possibly reflecting recent migration, while the Korean population, traditionally considered northern, lay at the boundary between northern and southern populations. An admixture estimate suggested 55 (51–59)% northern, 45(41–49)% southern contribution to the Koreans, illustrating the complexity of the genetic history of this region.

Keywords Y chromosome · STRs · Forensic genetics · Phylogeny · East Asia

Electronic Supplementary Material Supplementary material is available for this article at http://dx.doi.org/10.1007/s00414-0518-4.

K. D. Kwak · H. J. Jin · D. J. Shin · J. M. Kim · W. Kim (✉)
Department of Biological Sciences,
Dankook University,
Cheonan, 330-714, South Korea
e-mail: wookkim@dankook.ac.kr
Tel.: +82-41-5503441
Fax: +82-41-5503441

L. Roewer
Institute of Legal Medicine,
Charité-University Medicine Berlin,
Hannoversche Strasse 6,
10115 Berlin, Germany

M. Krawczak
Institute of Medical Informatics and Statistics,
Christian-Albrechts-University,
24105 Kiel, Germany

C. Tyler-Smith
The Wellcome Trust Sanger Institute,
Wellcome Trust Genome Campus,Hinxton,
Cambridge, CB10 1SA, UK

Introduction

The usefulness of Y-chromosomal markers for studies of human population genetics and forensic analysis has recently been recognized. The human Y chromosome has special features not found in autosomes: a haploid state and father-to-son transmission. The DNA sequence of the non-recombining portion of the Y therefore contains a record only of the mutational events that occurred in the past. As a consequence, haplotypes constructed from Y-chromosomal alleles can be used to study paternal lineages (Hammer 1995; Jobling and Tyler-Smith 1995; Underhill et al. 2001) and to differentiate human population groups (Hammer et al. 1997; Su et al. 1999; Kayser et al. 2001).

Studies of the ancient divergences in human evolution are simplest using polymorphisms with low probabilities of back and parallel mutation, for which ancestral states can be determined (Hammer and Zegura 1996). Binary polymorphisms, single-base substitution and small insertions or deletions (indels), are well-suited for this purpose. In contrast, most microsatellite or short tandem repeat (STR) loci have a high level of variability due to variation in repeat number. Thus, Y chromosome short tandem re-

peat (Y-STR) haplotypes are more useful for investigating and reconstructing the phylogeny of more recently diverged paternal lineages (Roewer et al. 1996; Santos and Tyler-Smith 1996), as well as for forensic/paternity testing (Edwards et al. 1992; Hammond et al. 1994; de Knijff et al. 1997; Kayser et al. 1997). In addition, binary marker ascertainment bias can lead to quite different conclusions about the same populations (Su et al. 1999; Karafet et al. 2001), but this should not occur when unbiased markers are used that are variable in all populations (Zerjal et al. 2002). Since microsatellites are so variable, Y-STR haplotype relationships and frequencies may provide a less biased measure of diversity and Y-chromosomal history, albeit more noisy than that from binary markers (Jobling and Tyler-Smith 2003).

In addition to the forensic benefit, the Y-STR haplotype database has useful information about the phylogeography of the corresponding populations. Thus, surveys of Y-STR variation can provide valuable information for studies of the peopling of east Asia, which has been the subject of some controversy. There have been two major models for the explanation of the early migration routes into east Asia. The first model postulates that a southeast Asian origin is most likely, followed by a northward migration (Turner 1990). Recent genetic surveys using autosomal microsatellite markers (Chu et al. 1998) and Y-chromosomal binary markers (Su et al. 1999) have been interpreted as supporting this model. In contrast, the second model suggests a bi- and/or multidirectional route: one migration through central Asia and one through southeast Asia (Nei and Roychoudhury 1993; Cavalli-Sforza et al. 1994; Karafet et al. 2001). Ding et al. (2000) also noted that southeast Asia is not the homeland for northeast Asian populations, but the potential importance of more recent gene flow from central Asia needs to be stressed for the peopling of northeast Asia. In this regard, since Korea and Japan lie between the southeast and northeast Asian gene pools, their population genetic data can also give us valuable information about the prehistoric migration route(s) and population expansions in east Asia. Although several databases of Y-STRs have been published and are highly contributed to the field of population and forensic genetics in east Asia (Shin et al. 2001; Tsai et al. 2002; Tang et al. 2003; Uchihi et al. 2003), the amount of data for a combining Y-haplotypes from diverse regions of Asia is still limited.

The present study, therefore, was designed to evaluate the usefulness of 11 Y-STR markers for forensic and population genetic studies in east Asia. The Y-STR loci (DYS19, two DYS385 loci, DYS388, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DXYS156Y) were analyzed in 700 males from ten east Asian populations using a set of two multiplex PCR genotyping systems. We report that forensic parameters obtained here appeared to be highly informative, and that the east Asian settlement must be seen as a complex process involving a bi- and/or multidirectional range expansions. The data presented herein are included in the worldwide Y-STR haplotype reference database (YHRD) available at http://www.yhrd.org.

Materials and methods

DNA samples

We studied 700 male DNA samples, collected from ten east Asian populations, as shown in Fig. 1 in ESM.

Geographical origin and linguistic affiliation of the populations analyzed here are given in Table 1. The DNA samples included subsets of the samples examined by Kim et al. (2000) and Jin et al. (2003), although the exact number of subjects for each population occasionally varies between these studies. DNAs were prepared from whole blood by the standard method (Sambrook et al. 1989) or extracted from buccal cells according to the procedure of Richards et al. (1993).

PCR amplification and genotyping

Two multiplex PCR amplifications for typing 11 Y-STR markers were constructed: multiplex GK1 (DYS19 and DYS389I/II labeled with 6FAM, DYS390 and DYS393 labeled with NED, and DYS392 labeled with HEX); multiplex GK2 (DYS385 loci and DYS388 labeled with 6 FAM, DYS391 labeled with NED, and DXYS156Y labeled with HEX). Only the forward primer was labeled in each case with the fluorescent dye. Multiplex PCR amplifications were performed in a volume of 15 μl containing 5–10 ng of genomic DNA, 50 mmol/l KCl, 10 mmol/l Tris–HCl (pH 8.3), 0.01% gelatin, 160 μg/ml BSA, 2.0 mmol/l $MgCl_2$, 200 μmol/l dNTPs, 2.5–5 pmol/l each primer and 1.5 U AmpliTaq Gold DNA polymerase (PE

Table 1 Geographical origin and linguistic affiliation of population samples studied

Population	Number	Geographical origin	Linguistic affiliation[a]
Korean	252	East Asia	Korean-Japanese
Chinese			
Beijing-Han	49	East Asia	Sino-Tibetan
Manchurian	32	East Asia	Mongolic (Manchu)
Yunnan	29	Southeast Asia	Sino-Tibetan
Japanese	104	East Asia	Korean-Japanese
Mongolian			
Buryat	42	East Asia	Mongolic (Buryat)
Vietnamese	43	Southeast Asia	Austroasiatic (Mon-Khmer)
Thai	41	Southeast Asia	Austroasiatic (Daic)
Indonesian	32	Southeast Asia	Austronesian (Western Malayo-Polynesian)
Philippine	76	Southeast Asia	Austronesian (Western Malayo-Polynesian)

[a]According to Ruhlen (1991)

Biosystems). Primer sequences for DYS19, DYS388, DYS389I/II, DYS390, DYS391, DYS392 and DYS393 are described in Kayser et al. (1997), while DXYS156Y and DYS385a/b were taken from Chen et al. (1994) and Schneider et al. (1998), respectively. In case of DXYS156Y, we also confirmed that males tested here appeared to have an inserted adenine in the fourth repeat motif from the 5′ end but never in females (Cali et al. 2002). Thus, the Y alleles can be distinguished from the X alleles by typing the adenine insertion. The two DYS385 repeats were identified simultaneously using a conventional non-locus-discriminating analysis, but not from a locus-discrimination protocol (Niederstätter et al. 2004).

Amplification reactions were carried out in a Perkin Elmer GeneAmp PCR System 9700 thermal cycler (Applied Biosystems), and cycling conditions were as follows: multiplex GK1, initial denaturation at 95°C for 10 min, followed by 35 cycles of 94°C for 1 min, 54°C for 2 min, 72°C for 2 min, and a final extension at 60°C for 40 min; multiplex GK2: initial denaturation at 95°C for 10 min, followed by 35 cycles of 94°C for 1 min, 57°C for 1 min, 72°C for 1.5 min, and final extension at 60°C for 30 min.

Detection of the amplified products was accomplished with the ABI 310 genetic analyzer (Applied Biosystems), and Y-STR alleles were named according to the number of repeat units they contain. Samples were prepared with 11.5 μl Hi-Di formamide (Applied Biosystems, P/N 4311320), 0.5 μl GS500 ROX (P/N 401734) size standard as an internal-sizing ladder, and with 1 μl PCR product. Samples were injected onto the capillary array for 5 s at 15,000 V. Separation were performed at 15,000 V for 34 min using the POP-4 polymer (Applied Biosystems, P/N 402838), 1× Genetic Analyzer Buffer with EDTA (P/N 402824), and a 47-cm array (P/N 402839) with a run temperature of 60°C. Following data collection, samples were analyzed with GeneScan 3.1.2 (For Macintosh, Applied Biosystems). Allele typing was performed in comparison to control DNA samples provided by B. Brinkmann (Institut für Rechtsmedizin, Westfälische Wihelms-Universität, Münster) and home-made allelic ladders by using Genotyper 2.5 (For Macintosh, Applied Biosystems). The nomenclature of the Y-STR loci studied here have also followed the recommendations of the International Society for Forensic Genetics (ISFG) guidelines for the Y-STR analyses (Gill et al. 2001).

Data analyses

Allele frequencies were calculated by counting from the observed phenotypes. Gene or haplotype diversity was calculated according to the formula $h=n(1-\sum x_i^2)/(n-1)$, where n represents the number of chromosomes sampled and x_i is the frequency of the ith haplotype or allele (Nei 1987). The discrimination capacity was estimated as the percentage proportion of individual-specific haplotypes (Gené et al. 1999). In Y-linked systems, haplotype diversity value is identical to the discrimination index (DI) (Sensabaugh 1982) and to the power of exclusion (POE) (Chakravarti and Li 1983).

The genetic relationship between different populations was assessed by means of pair-wise Φ_{ST}, an analogue of F_{ST} that takes the evolutionary distance between individual haplotypes into account (Excoffier et al. 1992; Excoffier and Smouse 1994). Estimates of Φ_{ST} were obtained using the Arlequin software (Schneider et al. 2000) and tested for statistical significance by means of randomisation (1,000 replicates per comparison). We initially build a neighbour-joining tree with PHYLIP v. 3.57c (Felsenstein 1995) tree using $D=-\ln(1-\Phi_{ST})$, a transformation of the pair-wise Φ_{ST} values that took the evolutionary process of Y-STR haplotype divergence into account. Following the theory of Wright's F_{ST}, D should be proportional to separation time assuming equal population sizes and a constant mutation rate in an infinite sites model (Nei 1987). However, the transformation still led to an intolerable number of negative branch lengths, indicative of non-additive distances. Owing to the high backward rate and stepwise nature of the Y-STR mutation process, Φ_{ST} may indeed increase at a slower rate than expected, and distance D was therefore modified into $D_x=-\ln(1-\Phi_{ST}^x)$. Delay parameter x was chosen such that the relative proportion of total negative branch lengths within the neighbour-joining tree attained its first local minimum at $x<1$. The transformed D_x values were then displayed as a multidimensional scaling plot, using SPSS 12.0, for ease of comparison to other distances. The second measure of the genetic relationship between populations used was ρ, the distance from an STR haplotype in one population to the closest haplotype in the second population, averaged over all chromosomes. ρ distances were also displayed in a multidimensional scaling plot.

Admixture proportions cannot be calculated from Y-STR haplotype data using standard measures because haplotype diversity is very high and most haplotypes in the admixed population are not found at all in the source populations. We therefore used the estimator m_ρ (Helgason et al., 2000), which takes into account the number of mutational differences between STR haplotypes and allows the most closely related haplotypes in the source populations to be used.

Results and discussion

The first aim of our genetic survey was to study the forensic DNA parameters and features of east Asian male lineages. The assessment of probabilities for matches between haplotyped male persons or traces persons requires the DNA profiling of a large number of haplotypes in the appropriate reference populations (Lessig et al. 2003). We therefore used the nine 'core' Y-STR markers ("minimal haplotype"), but also included two additional loci to make a total of 11, which were typed in samples from 700 male individuals from ten different populations. Allele frequencies calculated from each population and all haplotypes for the 11 Y-STR loci in 700 males are listed in Appendices A and B, respectively. A total of 644 different haplotypes were identified, among which 603 (86.14%) were indi-

Table 2 Number of different haplotypes and unique haplotypes found in east Asian populations using nine Y-STR markers

Population (*n*)	Number of different haplotype	Number of unique haplotype	Haplotype diversity	Discrimination capacity (%)
Northeast Asians				
Korean (252)	233	215	0.9994	85.32
N. Chinese (81)				
Beijing-Han (49)	48	47	0.9991	95.92
Manchurian (32)	22	16	0.9597	32.65
Japanese (104)	102	100	0.9996	96.15
Mongolian (42)	37	34	0.9919	80.95
Southeast Asians				
S. Chinese (29)				
Yunnan (29)	28	27	0.9975	93.10
Vietnamese (43)	41	39	0.9978	90.70
Thai (41)	41	41	1.0000	100.00
Indonesian (32)	32	32	1.0000	100.00
Philippine (76)	72	69	0.9982	90.79
Total (700)	(630)[a]	(580)[a]	(0.9996)[a]	(82.86)[a]

[a]Number of haplotypes and values calculated from a total of 700 males

vidual-specific (Appendix B, Table 2). Among 41 shared haplotypes, 25 were shared within the same population, 15 shared between two populations, and one found in common among four populations (Table 3). The haplotype diversity averaged over all populations was 0.9997, corresponding to 0.9996 using only the nine 'forensic' markers (i.e. excluding DYS388 and DXYS156) and is thus similar to the value for the other Asian populations (0.9996, *n*=1,924, Lessig et al. 2003), and slightly higher than the value for European populations (0.9976, *n*= 11,610, Roewer et al. 2001). The core set can be chosen as a backbone for the Y-STR database (Fig. 2 in ESM, the core set *h*=0.9996), the respective profiles were termed "minimal haplotypes" in order to emphasize the authors' view that they represent the minimum requirement for sufficiently informative haplotyping in forensic casework (Roewer et al. 2001). Thus, our results confirm that the minimal haplotype using only the nine 'core' Y-STR markers would also be useful for personal identification in forensic fields as well as population genetic studies of Y chromosome lineages. With all 11 markers, there were no population-specific shared haplotypes, but an immense variety of different haplotypes for each population; discrimination capacity was 86% (83% using nine markers, compared to 78% in the East Asian database with 3562 haplotypes (included in the YHRD, see http://www.yhrd.org, Statistics, release 15) and only 45% in the European database with 18711 haplotypes (included in the YHRD, see http://www.yhrd.org, Statistics, release 15). All the east Asian populations examined here have a highly informative haplotype diversity (≥0.995), except for the Mongolians (0.992) and Manchurians (0.960) (Table 2). The observed low Y-STR diversity of the Manchurians and Mongolians compared with other east Asian populations could be explained by genetic drift resulting from small effective population sizes (Pakendorf et al. 2002; Zerjal et al. 2002).

Table 3 Number of common Y-STR haplotypes shared within, between, and among populations

Population (*n*)	Within population	Two populations	Four populations
Koreans (252)	11	11 (CB1, J3, CY1, V5, I1,)	1
Beijing-Han (49)	1	2 (K1, M1)	
Manchurians (32)	3	3 (J1, M2)	1
Japanese (104)	2	4 (K3, CM1)	1
Mongolians (42)	3	3 (CB1, CM2)	
Yunnan (29)	1	1 (K1)	
Vietnamese (43)	1	5 (K5)	
Thais (41)	0		1
Indonesians (32)	0	1 (K1)	
Philippines (76)	3		
Total haplotypes shared	25	15	1

Capital letters refer to abbreviations of population as follows: *K* Koreans, *CH* Beijing-Han, *CM* Manchurians, *CY* Yunnan, *J* Japanese, *M* Mongolians, *V* Vietnamese, *T* Thais, *I* Indonesians, *P* Philippines. Number of common types are indicated by suffixes to abbreviations of the population. A common haplotype is shared by four populations: K/CM/J/T

Butler et al. (2002) have recently developed a Y-STR 20plex that offers a potential increased power of discrimination compared to most previous Y-STR multiplex typing. The 20plex includes a simultaneous amplification of all the markers within the European "minimal" and "extended" haplotypes. Our extension of the number of Y-STRs is required to evaluate and a selection of them typed in the next studies of the east Asian populations.

We then investigated the Y-chromosomal relationships between the populations. In all of these analyses nine Y-STRs were used, excluding DYS385 because the two alleles were not assigned to individual loci in this study. Pair-wise Φ_{ST} values between the ten east Asian populations were calculated (Table 4). The values were significantly greater than zero for most population pairs, except for the eight distances indicated by underlining. A multidimensional scaling plot of the transformed distances is

Table 4 Population pairwise Φ_{ST} values between ten east-Asian populations using nine Y-STRs

	Beijing-Han	Indonesians	Japanese	Koreans	Manchurians	Mongolians	Filipinos	Thais	Vietnamese	Yunnan
Beijing-Han	–	0.001	<0.001	0.066	<0.001	<0.001	0.005	<0.001	0.074	0.085
Indonesians	0.0423	–	0.002	0.057	0.001	<0.001	0.003	<0.001	0.126	0.034
Japanese	0.0733	0.0424	–	<0.001	<0.001	0.027	<0.001	<0.001	0.003	<0.001
Koreans	0.0095	0.0132	0.0522	–	<0.001	<0.001	<0.001	<0.001	0.230	0.151
Manchurians	0.0695	0.1271	0.1219	0.0812	–	<0.001	<0.001	<0.001	0.001	0.041
Mongolians	0.1066	0.0765	0.0199	0.0898	0.1442	–	<0.001	<0.001	<0.001	0.003
Filipinos	0.0205	0.0300	0.0943	0.0307	0.0977	0.1189	–	<0.001	<0.001	0.018
Thais	0.1087	0.0643	0.0930	0.0986	0.1614	0.1127	0.1085	–	<0.001	<0.001
Vietnamese	0.0159	0.0131	0.0334	0.0028	0.0915	0.0598	0.0440	0.0738	–	0.118
Yunnan	0.0158	0.0237	0.0616	0.0082	0.0411	0.0770	0.0221	0.0802	0.0154	–

Lower left half Pair-wise Φ_{ST} values between populations (non-significant values are underlined), *upper right halfp* value for Φ_{ST}=0 determined by randomisation (1,000 replicates)

shown in Fig. 1A. Most of the populations form a loose cluster towards the centre, with the Thais, Mongolians and Manchurians as the furthest outliers. The positions of the Mongolians and Manchurians may be accounted for by drift, as discussed above, but this explanation does not apply to the Thais, who have a haplotype diversity of 1.000 (and a reasonably large sample size, n=41). They do, however, differ from all the other populations in following matrilocal, rather than patrilocal customs: married couples live with the wife's rather than the husband's family, and this may be responsible for their distinct pattern of Y-chromosomal variation. In addition, further analyses of the Y-STR variation using a larger sample size should be required to better understand genetic profiles of the several populations with <50 individuals surveyed here.

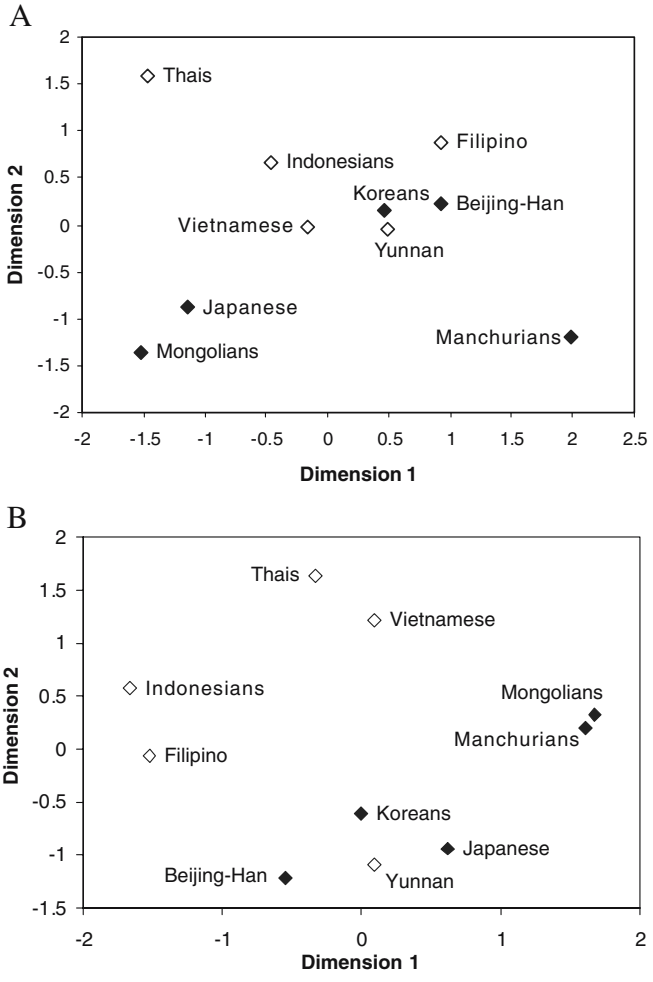

Fig. 1 Population comparisons. **A** Multidimensional scaling plot of transformed pair-wise Φ_{ST} distances; stress=0.12. **B** Multidimensional scaling plot of ρ genetic distances; stress=0.28. In both sections, southeast Asian populations are represented by open diamonds, and northeast populations by closed diamonds

There is some separation between the southern and northern populations, but it is incomplete and populations would not readily be assigned to two distinct groups in the absence of other information. In general, linguistic and geographic classifications of the populations examined here did not match with classification by Y-STR variation: the Y-STR data in these populations must reflect male-specific aspects of genetic structuring and population history as well as drift. This result is consistent with a recent report on mtDNA variation in Chinese ethnic populations (Yao et al. 2002b). The plot of ρ distances (Fig. 1B) shows a clearer north–south division, except for the Yunnan population which comes from southern China but lies within the northern cluster. The clearer division may be seen because ρ distances are less influenced by very ancient phylogenetic divergences between lineages, which may predate the entry of modern humans into east Asia, and so may reflect the more recent genetic history of these populations. The explanation for the anomalous genetic position of the Yunnan population may lie in their origin further north: many Han only moved to Yunnan during the Ming dynasty (1368–1644; Du and Yip 1993; Yao et al. 2002a).

The classification of the Korean population has been of particular interest because it lies geographically at the boundary of the northern/southern geographical areas and its affinities have previously been uncertain. In the present analyses it lies close to both northern and southern populations, and an admixture estimate using the combined Beijing Han, Japanese, Manchurian and Mongolian data to represent northern populations, and Indonesian, Filipino, Thai, Vietnamese and Yunnan data to represent southern

populations, produced a figure of 55(51–59)% northern, 45 (41–49)% southern contributions. If the Yunnan data were omitted, a slightly higher northern contribution was estimated: 61(58–65)%, and a correspondingly lower southern contribution. Modern Koreans are of course derived from ancient sources, not contemporary populations, but this result illustrates the likely complexity of their origins. Recent studies of mtDNA (Horai et al. 1996; Kivisild et al. 2002) and the Y chromosome (Hammer and Horai, 1995; Karafet et al. 2001; Jin et al. 2003) have also noted that the Koreans and Japanese possess lineages from both the southern and the northern haplogroup complexes, although they are generally considered a northeast Asian group. Our Y-STR data are consistent with the previous findings since both contributions are present, but the estimated northern one is slightly greater than the southern. It means that the peopling of east Asia must be seen as a complex process with genetic contributions involving multidirectional range expansions and genetic admixture.

The Y-STR information derived from the multiplexes developed here can, thus, be useful for forensic analysis and paternity testing where it can be of great benefit by providing information not available from autosomal DNA systems. In addition, the Y-STR population genetic data presented here provide the basic information required by the forensic community, as well as for studies of genetic history. It is now possible to evaluate evidence for population movements and recent range expansions in east Asian populations from a Y-chromosomal perspective, and see how the traditional northeast/southeast classification of east Asian populations does not fit well with classification by Y-STR variation.

Acknowledgements We would like to thank M.F. Hammer and B. Brinkmann for providing reference DNA samples. Special thanks go to the volunteers for proving DNA samples. This work was supported by a grant from the Korean Science and Engineering Foundation (KOSEF F05-2001-000-00572-0), and C.T.S. is supported by The Wellcome Trust.

References

Butler JM, Schoske R, Vallone PM, Kline MC, Redd AJ, Hammer MF (2002) A novel multiplex for simultaneous amplification of 20 Y chromosome STR markers. Forensic Sci Int 129:10–24

Cali F, Forster P, Kersting C, Mirisola MG, D'Anna R, De Leo G, Romano V (2002) DXYS156: a multi-purpose short tandem repeat locus for determination of sex, paternal and maternal geographic origins and DNA fingerprinting. Int J Leg Med 116:133–138

Cavalli-Sforza LL, Menozzi P, Piazza A (1994) The history and geography of human genes. Princeton University Press, Princeton

Chakravarti A, Li CC (1983) The probability of exclusion based on the HLA locus. Am J Hum Genet 35:1048–1052

Chen H, Lowther W, Avramopoulos D, Antonarakis SE (1994) Homologous loci DXYS156X and DXYS156Y contain a polymorphic pentanucleotide repeat (TAAAA)n and map to human X and Y chromosomes. Hum Mutat 4:208–211

Chu JY, Huang W, Kuang SQ, Wang JM, Xu JJ, Chu ZT, Yang ZQ, Lin KQ, Li P, Wu M, Geng ZC, Tan CC, Du RF, Jin L (1998) Genetic relationship of populations in China. Proc Natl Acad Sci U S A 95:11763–11768

Ding YC, Wooding S, Harpending HC, Chi HC, Li HP, Fu YX, Pang JF, Yao YG, Yu JG, Moyzis R, Zhang Y (2000) Population structure and history in East Asia. Proc Natl Acad Sci U S A 97:14003–14006

Du R, Yip VF (1993) Ethnic groups in China. Science Press, Beijing

Edwards A, Hammond HA, Jin L, Caskey CT, Chakraborty R (1992) Genetic variation at five trimeric and tetrameric tandem loci in four human population groups. Genomics 12:241–253

Excoffier L, Smouse P (1994) Using allele frequencies and geographic subdivision to reconstruct gene trees within a species: molecular variance parsimony. Genetics 136:343–359

Excoffier L, Smouse P, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491

Felsenstein J (1995) PHYLIP (phylogeny inference package) version 3.57c. Department of Genetics, University of Washington, Seattle

Gené M, Borrego N, Xifró A, Piqué E, Moreno P (1999) Haplotype frequencies of eight Y-chromosome STR loci in Barcelona (north-east Spain). Int J Leg Med 112:403–405

Gill P, Brenner C, Brinkmann B et al. (2001) DNA Commission of the International Society of Forensic Genetics: recommendations on forensic analysis using Y-chromosome STRs. Int J Leg Med 114:305–309

Hammer MF (1995) A recent common ancestry for human Y chromosomes. Nature 378:376–378

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

Hammer MF, Zegura SL (1996) The role of the Y chromosome in human evolutionary studies. Evol Anthropol 5:116–134

Hammer MF, Spurdle AB, Karafet T, Bonner MR, Wood ET, Novelletto A, Malaspina P, Mitchell RJ, Horai S, Jenkins T, Zegura SL (1997) The geographic distribution of human Y chromosome variation. Genetics 145:787–805

Hammond HA, Jin L, Zhong Y, Caskey CT, Chakraborty R (1994) Evaluation of 13 short tandem repeat loci for use in personal identification applications. Am J Hum Genet 55:175–189

Helgason A, Sigurðardóttir S, Nicholson J, Sykes B, Hill EW, Bradley DG, Bosnes V, Gulcher JR, Ward R, Stefánsson K (2000) Estimating Scandinavian and Galic ancestry in the male settlers of Iceland. Am J Hum Genet 67:697–717

Horai S, Murayama K, Hayasaka K, Matsubayashi S, Hattori Y, Fucharoen G, Harihara S, Park KS, Omoto K, Pan IH (1996) mtDNA polymorphism in East Asian populations, with special reference to the peopling of Japan. Am J Hum Genet 59:579–590

Jin HJ, Kwak KD, Hammer MF, Nakahori Y, Shinka T, Lee JW, Jin F, Jia X, Tyler-Smith C, Kim W (2003) Y-chromosomal DNA haplogroups and their implications for the dual origins of the Koreans. Hum Genet 114:27–35

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

Jobling MA, Tyler-Smith C (2003) The human Y chromosome: an evolutionary marker comes of age. Nat Rev Genet 4:598–612

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

Kayser M, Caglia A, Corach D, Fertwell N, Gehrig C, Graziosi G, Heidorn F, Herrmann S, Herzog B, Hidding M, Honda K, Jobling M, Krawczak M, Leim K, Meuser S, Meyer E, Oesterreich W, Pandya A, Parson W, Penacino G, Perez-Lezaun A, Piccinini A, Prinz M, Schmitt C, Schneider PM, Szibor R, Teifel-Greding J, Weichhold G, Knijff PD, Roewer L (1997) Evaluation of Y-chromosomal STRs: a multicenter study. Int J Leg Med 110:125–133

Kayser M, Krawczak M, Excoffier L, Dieltjes P, Corach D, Pascali V, Gehrig C, Bernini LF, Jespersen J, Bakker E, Roewer L, de Knijff P (2001) An extensive analysis of Y-chromosomal microsatellite haplotypes in globally dispersed human populations. Am J Hum Genet 68:990–1018

Kim W, Shin DJ, Harihara S, Kim YJ (2000) Y chromosomal DNA variation in East Asian populations and its potential for inferring the peopling of Korea. J Hum Genet 45:76–83

Kivisild T, Tolk HV, Parik J, Wang Y, Papiha SS, Bandelt HJ, Villems R (2002) The emerging limbs and twigs of the east Asian mtDNA tree. Mol Biol Evol 19:1737–1751

de Knijff P, Kayser M, Caglia A, Corach D, Fretwell N, Gehrig C, Graziosi G, Heidorn F, Herrmann S, Herzog B, Hidding M, Honda K, Jobling M, Krawczak M, Leim K, Meuser S, Meyer E, Oesterreich W, Pandya A, Parson W, Penacino G, Perez-Lezaun A, Piccinini A, Prinz M, Roewer L (1997) Chromosome Y microsatellites: population genetic and evolutionary aspects. Int J Leg Med 110:134–149

Lessig R, Willuweit S, Krawczak M, Wu F-C, Pu C-E, Kim W, Henke L, Henke J, Miranda J, Hidding M, Benecke M, Schmitt C, Magno M., Calacal G, Delfin FC, De Ungria MCA, Elias S, Augustin C, Tun Z, Honda K, Kayser M, Gusmao L, Amorim A, Alves C, Hou Y, Keyser C, Ludes B, Roewer L (2003) Asian online Y-STR haplotype reference database. Legal Med 5:160–163

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Nei M, Roychoudhury AK (1993) Evolutionary relationships of human populations on a global scale. Mol Biol Evol 10:927–943

Niederstätter H, Berger B, Oberacher H, Brandstätter A, Huber CG, Parson W (2004) Separate analysis of DYS385a and b versus conventional DYS385 typing: is there forensic relevance? Int J Leg Med (Epub ahead of print)

Pakendorf B, Morar B, Tarskaia LA, Kayser M, Soodyall H, Rodewald A, Stoneking M (2002) Y-chromosomal evidence for a strong reduction in male population size of Yakuts. Hum Genet 110:198–200

Richards B, Skoletsky J, Shuber AP, Balfour R, Stern RC, Dorkin HL, Parad RB, Witt D, Klinger KW (1993) Multiplex PCR amplification from the CFTR gene using DNA prepared from buccal brushes/swabs. Hum Mol Genet 2:159–163

Roewer L, Kayser M, Dieltjes P, Nagy M, Bakker E, Krawczak M, de Knijff P (1996) Analysis of molecular variance (AMOVA) of Y-chromosome-specific microsatellites in two closely related human populations. Hum Mol Genet 5:1029–1033

Roewer L, Krawczak M, Willuweit S, Nagy M, Alves C, Amorim A, Anslinger K, Augustin C, Betz A, Bosch E, Caglia A, Carracedo A, Corach D, Dekairelle AF, Dobosz T, Dupuy BM, Furedi S, Gehrig C, Gusmao L, Henke J, Henke L, Hidding M, Hohoff C, Hoste B, Jobling MA, Kargel HJ, de Knijff P, Lessig R, Liebeherr E, Lorente M, Martinez-Jarreta B, Nievas P, Nowak M, Parson W, Pascali VL, Penacino G, Ploski R, Rolf B, Sala A, Schmidt U, Schmitt C, Schneider PM, Szibor R, Teifel-Greding J, Kayser M (2001) Online reference database of European Y-chromosomal short tandem repeat (STR) haplotypes. Forensic Sci Int 118:106–113

Ruhlen M (1991) Postscript. In: Guide to the world's languages (ed. M. Ruhlen), pp. 379–407. Stanford University Press, Stanford, CA

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: A laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, New York

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

Schneider PM, Meuser S, Waiyawuth W, Seo Y, Rittner C (1998) Tandem repeat structure of the duplicated Y-chromosomal STR locus DYS385A/B and frequency studies in the German and three Asian populations. Forensic Sci Int 97:61–70

Schneider S, Roessli D, Excoffier L (2000) Arlequin v. 2.000: a software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Geneva, Switzerland

Sensabaugh GF (1982) Isozymes in forensic science. Isozymes Curr Top Biol Med Res 6:247–282

Shin DJ, Jin HJ, Kwak KD, Choi JW, Han MS, Kang PW, Choi SK, Kim W (2001) Y-chromosome multiplexes and their potential for the DNA profiling of Koreans. Int J Leg Med 115:109–117

Su B, Xiao J, Underhill P, Deka R, Zhang W, Akey J, Huang W, Shen D, Lu D, Chu J, Tan J, Shen P, Davis R, Cavalli-Sforza LL, 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

Tang JP, Hou YP, Li YB, Wu J, Zhang J, Zhang HJ (2003) Characterization of eight Y-STR loci and haplotypes in a Chinese Han population. Int J Leg Med 117:263–270

Tsai L-C, Yuen T-Y, Hsieh H-M, Lin M, Tzeng C-H, Huang N-W, Linacre A, Lee JG-I (2002) Haplotype frequencies of nine Y-chromosome STR loci in the Taiwanese Han population. Int J Leg Med 116:179–183

Turner CG (1990) Major features of sundadonty and sinodonty, including suggestions about east Asian microevolution, population history and late Pleistocene relationships with Australian aboriginals. Am J Phys Anthropol 82:295–317

Uchihi R, Yamamoto T, Usuda K, Yoshimoto T, Tanaka M, Tokunaga S, Kurihara R, Tokunaga K, Katsumata Y (2003) Haplotype analysis with 14 Y-STR loci using 2 multiplex amplification and typing systems in 2 regional populations in Japan. Int J Leg Med 117:34–38

Underhill PA, Passarino G, Lin AA, Shen P, Lahr MM, Foley RA, Oeffner PJ, Cavalli-Sforza LL (2001) The phylogeography of Y-chromosome binary haplotypes and the origins of modern human populations. Ann Hum Genet 65:43–62

Yao Y-G, Kong Q-P, Bandelt H-J, Kivisild T, Zhang Y-P (2002a) Phylogenetic differentiation of mitochondrial DNA in Han Chinese. Am J Hum Genet 70:635–651

Yao Y-G, Nie L, Harpending H, Fu Y-X, Yuan Z-G, Zhang Y-P (2002b) Genetic relationship of Chinese ethnic populations revealed by mtDNA sequence diversity. Am J Phys Anthropol 118:63–76

Zerjal T, Wells RS, 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