

FOR THE RECORD

V. K. Kashyap,¹ Ph.D.; Saurav Guha,¹ M.Sc.; and R. Trivedi,¹ Ph.D.

Concordance Study on 15 STR Loci in Three Major Populations of Himalayan State Sikkim*

POPULATION: "Sikkim," an erstwhile independent nation, now is a northeastern Himalayan hilly state of India. This state is inhabited by three major communities Nepali (collective term for different caste and communities of Nepal Kingdom, population size: 14 lakhs approximately), and two tribal groups (1) Bhutia (population size: 52 thousand approximately), & Lepcha (population size: 30 thousand approximately). All these communities practice very high degree of endogamy.

KEYWORDS: forensic science DNA typing, STR, D3S1358, TH01, D21S11, D18S51, Penta E, D5S818, D13S317, D7S820, D16S539, CSF1PO, Penta D, vWA, D8S1179, TPOX, FGA, Sikkim population

Samples were collected randomly from different individuals of East District of Sikkim. Nepali ($N = 110$), Bhutia ($N = 75$) and Lepcha ($N = 48$) in form of either blood or buccal swab as per cooperation of participant. Genomic DNA was extracted by using standard phenol/chloroform procedure (2). Quantitation of DNA was carried out using the Quantiblot kit (PE Applied Biosystems) and subsequent PCR amplification was performed using the Powerplex™ 16 multiplex System (Promega Corp, Madison, U.S.A.) The products were detected on a 5% denaturing polyacrylamide sequencing gel using the ABI Prism™ 377 DNA Sequencer (PE Applied Biosystems) and genotype classification was made by comparison with allelic ladders provided with the Powerplex™ 16 System.

Statistical analyses included tests for possible divergence from Hardy-Weinberg expectations, unbiased estimate of the expected homozygote frequencies (3), likelihood ratio test (4) and the exact test (5) based on 2000 shuffling experiments performed using DNATYPE program (6).

The distribution of observed allele frequencies at the 15 STR loci in the Sikkim population are shown in Tables 1 to 4. Statistical results for observed homozygosity, in occurrence with probability of homozygosity, G-square test & Exact test are presented in Tables 5–6. Frequencies obtained were compared with some other previously published data on Indian populations (7–10) and unpublished data on Bihar and Tamilnadu. Data were also compared with US Caucasians, African Americans, and Hispanics with no significant difference (11) (data not shown). The data generated from this study will contribute to the expansion of Indian DNA database suitable for population studies and forensic applications.

DNA Typing Unit, Central Forensic Science Laboratory, 30 Gorachand Road, Kolkata—700 014, West Bengal, India.

* Supported by a grant under IX Plan of Ministry of Home Affairs, Government of India.

The complete data set is available to any interested researcher upon request.

References

1. Singh KS. India's Communities, National Series. People of India. Oxford University Press. 1998.
2. Sambrook J, Fritsch EF, Maniatis T. Molecular cloning. A laboratory manual. 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. 1989.
3. Nei M. Estimate of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 1978;89:583–90.
4. Weir BS. Independence of VNTR alleles defined as fixed bins. *Genetics* 1992;130:873–87.
5. Guo SW, Thompson EA. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* 1992;48:361–72.
6. DNATYPE, (Windows 95/NT ver). Chakraborty R, Stivers D, Zhong Y, CHG. University of Texas, Houston, Texas. USA.
7. Kashyap VK, Sarkar N, Trivedi R. Allele frequencies for STR loci of the Powerplex16 multiplex system in five endogamous populations of India. *Forensic Sci Int* 2002. In press.
8. Sahoo S, Kashyap VK. Allele frequency of data for Powerplex 16 loci in four major populations of Orissa, India. *J Forensic Sci* 2002. In press.
9. Rajkumar R, Kashyap VK. Distribution of alleles of fifteen STR loci of the PowerPlex 16 multiplex system in four predominant population groups of South India. *Forensic Sci International USA*. 2002. In press.
10. Gaikwad S, Kashyap VK. Polymorphism at fifteen hypervariable microsatellite loci in four population of Maharashtra, India. *Forensic Sci International USA*. 2002. In press.
11. Levedakou EN, Freeman DA, Budzynski MJ, Early BE, McElfresh KC, Schumm JW, et al. Allele frequencies for fourteen STR Loci of the Powerplex™ 1.1 and 2.1 Multiplex Systems and Penta D locus in Caucasians, African-Americans, Hispanics, and other populations of the United States of America and Brazil. *J Forensic Sci* 46(3);2001:423–767.

Additional information and reprint requests:
V.K. Kashyap, Ph.D.

DNA Typing Unit, Central Forensic Science Laboratory, 30 Gorachand Rd, Kolkata—700 014, West Bengal, India.

TABLE 1—Genotype frequencies for 13 tetranucleotide repeat loci in Nepali population (N = 110) of Sikkim.

Allele	D3S1358	TH01	D18S51	D5S818	D13S317	D7S820	D16S539	CSF1PO	vWA	D8S1179	TPOX	Allele	D21S11	FGA
5												17		0.016
6		0.182										18		0.016
7		0.158		0.007	0.007	0.008						19		0.050
8		0.095		0.055	0.174	0.185	0.040				0.375	19.2		0.008
9		0.492			0.134	0.120	0.193	0.064		0.007	0.108	20		0.025
9.3		0.071										20.2		0.008
10			0.015	0.206	0.095	0.209	0.153	0.193	0.023	0.111	0.066	21		0.033
11			0.023	0.349	0.277	0.193	0.266	0.233	0.007	0.047	0.358	22		0.191
12			0.023	0.253	0.166	0.233	0.241	0.411	0.015	0.071	0.033	22.2		0.025
13			0.174	0.126	0.142	0.032	0.080	0.088	0.031	0.190	0.058	23		0.191
14			0.230			0.016	0.016	0.008	0.111	0.166		23.2		0.025
15	0.293		0.129				0.008		0.031	0.253		24		0.158
15.2	0.015											24.2		0.108
16	0.341		0.119						0.261	0.095		25		0.008
17	0.222		0.063						0.277	0.047		26		0.108
18	0.079		0.039						0.142	0.007		27	0.007	0.008
19	0.015		0.079						0.095			28	0.095	
19.2												28.2	0.007	
20	0.007		0.023									29	0.269	
20.2												29.2	0.015	
21			0.039									30	0.214	0.008
22			0.007									30.2	0.039	
22.2												31	0.071	
23			0.023									31.2	0.079	0.008
23.2												32	0.015	
24			0.007									32.2	0.119	
												33	0.007	
												33.2	0.039	
												34.2	0.007	
												35.2	0.007	

TABLE 2—Genotype frequencies for 13 tetranucleotide repeat loci in Bhutia population (N = 75) of Sikkim.

Allele	D3S1358	TH01	D18S51	D5S818	D13S317	D7S820	D16S539	CSFIPO	vWA	D8S1179	TPOX	Allele	D21S11	FGA
5												18.2		0.015
6		0.156										19		0.062
7		0.296										20		0.140
8		0.093									0.359	20.2		
9		0.343		0.109	0.203	0.296	0.218				0.078	21		
9.3		0.078										22		0.515
10		0.015		0.125	0.140	0.078	0.140	0.109		0.031	0.093	22.2		0.078
11		0.015		0.140	0.343	0.250	0.281	0.265		0.140	0.265	23		0.062
12				0.234	0.187	0.312	0.203	0.546		0.171	0.093	23.2		0.078
13	0.015		0.093	0.250	0.046	0.015	0.125	0.062		0.281	0.109	24		
14	0.062		0.218	0.140		0.046	0.031	0.015	0.281	0.218		24.2		
15	0.265		0.156						0.187	0.062		25		0.031
15.2												26		
16	0.343		0.203						0.171	0.015		27		
17	0.234		0.125						0.281	0.046		27.2		
18	0.078		0.015						0.078	0.031		28	0.062	
18.2												28.2		
19			0.031									29	0.312	0.015
20												29.2		
20.2												30	0.171	
21			0.015									30.2	0.046	
22												31	0.093	
22.2												31.2	0.015	
23			0.015									32	0.046	
23.2												32.2	0.218	
24												33	0.015	
24.2												33.2		
												34.2		
												35.2		
												36	0.015	

TABLE 3—Genotype frequencies for 13 tetranucleotide repeat loci in Lepcha population (N = 48) of Sikkim.

Allele	D3S1358	TH01	D18S51	D5S818	D13S317	D7S820	D16S539	CSFIPO	vWA	D8S1179	TPOX	Allele	D21S11	FGA
5												22		0.102
6		0.125										22.2		0.034
7		0.227						0.056				23		0.193
8		0.147		0.068		0.045	0.034			0.113	0.522	23.2		0.079
9		0.500	0.013	0.045	0.136	0.102	0.170	0.102		0.034	0.045	24		0.147
9.3												24.2		0.102
10				0.272	0.136	0.136	0.181	0.170	0.045	0.170	0.431	25		0.340
11			0.045	0.212	0.284	0.284	0.204	0.193	0.045	0.125		26	0.022	
12			0.034	0.215	0.170	0.170	0.261	0.295	0.011	0.102		27	0.022	
13			0.125	0.125	0.068	0.125	0.147	0.079	0.022	0.261		27.2	0.011	
14	0.013		0.136					0.102	0.193			28	0.102	
15	0.204		0.170						0.022	0.056		28.2	0.045	
15.2												29	0.193	
16	0.352		0.147						0.147	0.125		29.2	0.034	
17	0.284		0.045						0.420	0.011		30	0.136	
18	0.125		0.102						0.079			30.2	0.091	
19	0.022		0.079						0.011			31	0.045	
20			0.102						0.011			31.2	0.056	
21												32	0.022	
												32.2	0.159	
												33	0.011	
												33.2	0.022	
												34	0.022	
												34.2		
												35.2		

TABLE 4—Genotype frequencies for two pentanucleotide repeat loci in Nepali, Bhutia and Lepcha population of Sikkim.

Allele	Nepali		Bhutia		Lepcha	
	Penta E	Penta D	Penta E	Penta D	Penta E	Penta D
5	0.049		0.015			
6						
7	0.090	0.008	0.062	0.015		
8	0.016	0.081	0.031	0.031		
9	0.008	0.172	0.015	0.140	0.136	0.284
9.3						
10	0.057	0.147		0.171		
11	0.106	0.196		0.218	0.079	0.170
12	0.131	0.139		0.031	0.045	0.113
13	0.016	0.180		0.281	0.147	0.147
14	0.040	0.040		0.078	0.045	0.090
15	0.139	0.032		0.031	0.079	0.102
15.2					0.113	0.022
16	0.131					
17	0.090				0.090	0.011
18	0.065				0.136	0.056
19	0.049				0.045	
19.2					0.034	
20	0.008				0.045	
21						
24						
25						

TABLE 5—Statistical results of observed and probability of homozygosity, G-square test, exact test for 13 tetranucleotide repeat loci in population of Sikkim.

Parameters	D3S1358	TH01	D21S11	D18S51	D5S818	D13S317	D7S820	D16S539	CSF1PO	vWA	D8S1179	TPOX	FGA
Nepali													
Obs. homozygosity	0.333	0.285	0.079	0.157	0.269	0.222	0.145	0.258	0.258	0.190	0.142	0.366	0.200
Homozygosity (<i>p</i>)	0.893	0.241	0.986	0.981	0.163	0.309	0.350	0.777	0.554	0.993	0.878	0.088	0.999
Exact test (<i>p</i>)	0.365	0.559	0.043	0.225	0.109	0.220	0.128	0.463	0.558	0.977	0.792	0.570	0.587
G-square test (<i>p</i>)	0.560	0.333	0.037	0.305	0.096	0.275	0.139	0.624	0.472	0.973	0.716	0.423	0.656
Bhutia													
Obs. homozygosity	0.250	0.343	0.348	0.156	0.283	0.312	0.281	0.125	0.250	0.250	0.187	0.343	0.437
Homozygosity (<i>p</i>)	0.369	0.060	0.408	0.748	0.033	0.204	0.960	0.083	0.496	0.248	0.424	0.096	0.847
Exact test (<i>p</i>)	0.349	0.004	0.318	0.125	0.039	0.719	0.967	0.189	0.646	0.376	0.120	0.019	0.015
G-square test (<i>p</i>)	0.267	0.003	0.235	0.082	0.063	0.821	0.961	0.092	0.363	0.370	0.091	0.039	0.015
Lepcha													
Obs. homozygosity	0.272	0.568	0.181	0.188	0.227	0.250	0.363	0.386	0.340	0.204	0.318	0.454	0.477
Homozygosity (<i>p</i>)	0.477	0.035	0.955	0.538	0.318	0.442	0.002	0.090	0.007	0.988	0.005	0.972	0.099
Exact test (<i>p</i>)	0.394	0.010	0.011	0.242	0.428	0.144	0.372	0.284	0.317	0.426	0.293	1.000	0.019
G-square test (<i>p</i>)	0.298	0.056	0.008	0.388	0.411	0.228	0.342	0.137	0.124	0.428	0.125	1.000	0.008

TABLE 6—Statistical results of observed and probability of homozygosity, G-square test, exact test for two pentanucleotide repeat loci in population of Sikkim.

Parameters	Nepali		Bhutia		Lepcha	
	Penta E	Penta D	Penta E	Penta D	Penta E	Penta D
Obs. homozygosity	0.098	0.196	0.098	0.125	0.159	0.340
Homozygosity (<i>p</i>)	0.999	0.346	0.999	0.833	0.236	0.019
Exact test (<i>p</i>)	0.236	0.187	0.236	0.660	0.053	0.001
G-square test (<i>p</i>)	0.176	0.247	0.176	0.530	0.087	0.003