

FOR THE RECORD

G. Hima Bindu,¹ M.Sc.; R. Trivedi,¹ Ph.D.; and V. K. Kashyap,¹ Ph.D.

Genotypic Polymorphisms at Seventeen Autosomal Short Tandem Repeat Loci in Four Tribal Populations of Andhra Pradesh, India

POPULATIONS: Four tribal populations of Andhra Pradesh, South India (1), Chenchu ($n = 100$), Lambadi ($n = 107$), Naikpod Gond ($n = 104$) and Yerukula ($n = 101$) were analyzed for DNA polymorphisms at 15 tetranucleotide and 2 pentanucleotide short tandem repeat (STR) loci in the present study.

KEYWORDS: forensic sciences, short tandem repeats, human identification, Andhra Pradesh, tribal populations, THO1, D3S1358, FGA, D8S1179, CSF1PO, D13S317, D16S539, D21S11, D7S820, TPOX, vWA, D18S51, D2S1338, D19S433, D5S818, Penta D and Penta E, DNA typing, population genetics

Peripheral blood samples were obtained by venipuncture from consenting healthy, unrelated individuals of the four tribal populations of South India. Genomic DNA was extracted from blood using the standard phenol-chloroform extraction method (2), followed by ethanol precipitation. Extracted DNA was quantitated using Quantiblot™ Human DNA Quantitation Kit according to the manufacturer's instructions (Perkin Elmer, New Jersey).

The fifteen tetranucleotide repeat loci were co-amplified in 1 to 10 ng of individual genomic DNA samples using the AmpF/STR® Identifiler™ PCR Amplification Kit (Applied Biosystems, Foster City), under conditions recommended by the manufacturer (3), in a GeneAmp PCR 9700 thermal cycler (Applied Biosystems, Foster City). The highly polymorphic pentanucleotide repeats Penta D and Penta E were also co-amplified using the primers and conditions as mentioned in Krenke et al. (4). PCR-amplified fragments were then separated and detected on a 5% denaturing polyacrylamide-sequencing gel using the ABI Prism™ 377 DNA Sequencer (Applied Biosystems, Foster City). The genotypic classification was made in comparison with allelic ladders provided with AmpF/STR® Identifiler™ PCR Amplification Kit. The data analysis and allele designations were carried out using the GeneScan™ Analysis Software (Version 3.7; Applied Biosystems, Foster City) and Genotyper™ software (Version 3.7; Applied Biosystems, Foster City).

Allele frequencies were computed using the gene counting method. Possible divergence from Hardy-Weinberg expectations (HWE) was determined by calculating observed heterozygosity, probability of homozygosity (5), exact test (6) and the likelihood ratio test (7), based on 2000 shuffling experiments performed using DNATYPE program (8). Measures of the discriminatory power of genetic markers, the power of discrimination (PD) (9), and the

probability of match (pM) (10) were computed. In addition, the power of exclusion (PE) (11) was also estimated for all studied populations at each locus. The most frequent and rare genotype for every population at each locus was also determined.

Tables 1–4 show the observed allele frequencies and forensically important statistical parameters based on seventeen STR loci in the four tribal populations of Andhra Pradesh. The results indicate high degree of polymorphism in all the analyzed loci of the studied populations.

A high degree of polymorphism for the selected loci was indicated by the observed heterozygosities in the studied groups. Locus Penta E (greater than 92%) and D2S1338 (greater than 85.2%) were most heterozygous and the locus CSF1PO (greater than 56%) was least heterozygous among the analyzed STR loci. The agreement with the Hardy-Weinberg Expectations (HWE) (tested by exact test and likelihood ratio) was confirmed for all the studied loci with the exception of locus D3S1358 for Naikpod Gond, loci D3S1358 and D13S317 for Yerukula, loci D13S317, D16S539 and Penta E for the Lambadi population.

The highly polymorphic pentanucleotide repeat loci, Penta D and Penta E have added significantly to the discriminatory power of the system. The most informative locus among the seventeen STR systems was Penta E (PD = 0.919 in Lambadi and 0.890 in Yerukula) and FGA (PD = 0.917 in Chenchu and 0.920 in Naikpod Gond), while the least informative was CSF1PO (PD = 0.570 in Chenchu and 0.665 in Yerukula) and TPOX (PD \geq 0.673 in Lambadi and Naikpod Gond). The validation and efficacy of these STR loci in human identification are thus facilitated by a high degree of observed discriminatory values. The power of exclusion calculated for these populations varied at different loci, but the cumulative power of exclusion of the seventeen loci, exhibited the expected value of 0.999. The present analysis is consistent with the database of allele frequencies of polymorphic loci of other studied Indian populations (12–15).

¹ National DNA Analysis Center, Central Forensic Science Laboratory, Kolkata, India.

TABLE 1—Allele frequencies at seventeen STR Loci in Chenchu Tribe (N = 100) of Andhra Pradesh, India.

Alleles	D5S818	FGA	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	THO1	D13S317	D16S539	D2S1338	D19S433	vWA	TPOX	D18S51	Penta D	Penta E
4.0																0.010	
5.0																	0.110
6.0								0.205								0.030	
7.0					0.015	0.005		0.130	0.010					0.005	0.005		0.080
8.0					0.275			0.160	0.230	0.140				0.375		0.010	
9.0	0.015		0.010		0.030	0.030		0.440	0.070	0.100				0.125		0.300	0.010
9.3								0.065									
10.0	0.100		0.170		0.185	0.095			0.115	0.120				0.070		0.180	0.010
11.0	0.450		0.045		0.300	0.245			0.315	0.355		0.015		0.410	0.010	0.270	0.200
11.2												0.010					
12.0	0.260		0.105		0.190	0.600			0.180	0.235				0.015	0.056	0.080	0.150
13.0	0.155		0.120		0.005	0.015			0.065	0.050				0.230	0.005	0.136	0.060
14.0	0.010		0.145			0.010	0.065		0.015					0.210	0.205	0.242	0.050
14.2														0.070			0.010
15.0			0.260				0.195							0.170	0.045	0.187	0.100
15.2														0.035			
16.0			0.120				0.480							0.045	0.160	0.162	0.110
16.2														0.060			
17.0			0.025				0.220				0.045			0.005	0.320	0.141	0.070
17.2	0.010													0.015			
18.0		0.010					0.030				0.160			0.200	0.025		0.090
19.0		0.061					0.010				0.265			0.055	0.015		
20.0		0.121									0.040			0.010	0.010		
20.2		0.005															
21.0		0.086									0.080				0.010		0.010
21.2		0.035															
22.0		0.121									0.045						
22.2		0.040															
23.0		0.096									0.135						
23.2		0.025															
24.0		0.066									0.180						
24.2		0.045															
25.0		0.106									0.050						
25.2		0.045															
26.0		0.091															
26.2		0.035															
27.0		0.010		0.005													
28.0				0.160													
29.0				0.175													
30.0				0.160													
30.2				0.020													
31.0				0.040													
31.2				0.095													
32.0				0.010													
32.2				0.230													
33.0				0.005													
33.2				0.070													
34.2				0.030													
H _{obs}	0.710	0.838	0.780	0.830	0.760	0.560	0.750	0.680	0.720	0.680	0.910	0.750	0.720	0.660	0.828	0.760	0.740
H(p)	0.752	0.004	0.102	0.610	0.946	0.845	0.123	0.403	0.072	0.028	0.053	0.010	0.115	0.823	0.808	0.590	0.002
ET	0.943	0.552	0.066	0.197	0.140	0.276	0.076	0.547	0.112	0.511	0.769	0.395	0.757	0.797	0.339	0.543	0.488
LR	0.893	0.550	0.051	0.239	0.201	0.245	0.036	0.643	0.235	0.498	0.573	0.257	0.808	0.612	0.368	0.631	0.698
pM	0.305	0.083	0.160	0.152	0.237	0.430	0.322	0.282	0.207	0.228	0.161	0.156	0.215	0.330	0.163	0.209	0.107
PD	0.695	0.917	0.840	0.848	0.763	0.570	0.678	0.718	0.793	0.772	0.839	0.844	0.785	0.671	0.837	0.791	0.893
PE	0.444	0.672	0.562	0.656	0.527	0.246	0.510	0.398	0.460	0.398	0.816	0.510	0.460	0.369	0.653	0.527	0.493

N: number of individuals; H_{obs}: observed heterozygosity; H(p): probability of homozygosity; ET: exact test; LR: likelihood ratio; pM: match of probability; PD: power of discrimination; PE: power of exclusion.

TABLE 2—Allele frequencies of seventeen STR Loci in Lambadi Tribe (N = 107) of Andhra Pradesh, India.

Alleles	D5S818	FGA	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	TH01	D13S317	D16S539	D2S1338	D19S433	vWA	TPOX	D18S51	Penta D	Penta E
3.2																0.028	
5.0								0.005									0.078
6.0								0.313									
7.0					0.029			0.150	0.042							0.009	0.088
8.0	0.005		0.005		0.252	0.009		0.122	0.215	0.123				0.416		0.019	
9.0	0.019				0.148	0.014		0.318	0.098	0.137				0.126		0.176	0.020
9.3								0.089									
10.0	0.136		0.224		0.195	0.168		0.005	0.103	0.118		0.019		0.070	0.005	0.213	0.039
11.0	0.369		0.070		0.210	0.262			0.210	0.255		0.005		0.364	0.010	0.194	0.127
11.2												0.009					
12.0	0.308		0.079		0.143	0.435			0.215	0.278		0.098		0.023	0.052	0.185	0.059
12.2												0.009					
13.0	0.145		0.126		0.024	0.070	0.009		0.075	0.085		0.327	0.005		0.148	0.130	0.020
13.2												0.009					
14.0	0.005		0.178			0.037	0.033		0.042	0.005		0.192	0.140		0.248	0.009	0.069
14.2												0.047					
15.0	0.014		0.224			0.005	0.318					0.117	0.070		0.210	0.028	0.088
15.2												0.084					
16.0			0.070				0.262				0.019	0.042	0.164		0.138	0.009	0.088
16.2												0.023					
17.0			0.023				0.304				0.066	0.019	0.294		0.081		0.098
18.0	0.019						0.070				0.142		0.210		0.057		0.049
19.0	0.043						0.005				0.151		0.098		0.038		0.098
20.0	0.176										0.170		0.019		0.010		0.049
20.2	0.010																
21.0	0.129										0.038						0.020
22.0	0.190										0.066				0.005		
22.2	0.019																
23.0	0.157										0.179						0.010
23.2	0.019																
24.0	0.114										0.127						
25.0	0.076										0.024						
26.0	0.038										0.014						
27.0	0.010			0.005													
28.0				0.179							0.005						
28.2				0.005													
29.0				0.170													
29.2				0.005													
30.0				0.170													
30.2				0.038													
31.0				0.085													
31.2				0.104													
32.0				0.009													
32.2				0.151													
33.2				0.038													
34.0				0.005													
34.2				0.038													
H _{obs}	0.720	0.867	0.925	0.887	0.800	0.710	0.804	0.720	0.813	0.745	0.821	0.851	0.879	0.720	0.848	0.815	0.961
H(p)	0.833	0.967	0.012	0.511	0.780	0.954	0.095	0.382	0.560	0.137	0.144	0.419	0.063	0.302	0.825	0.726	0.276
ET	0.978	0.250	0.126	0.251	0.316	0.185	0.298	0.601	0.019	0.017	0.663	0.070	0.965	0.411	0.996	0.286	0.016
LR	0.910	0.071	0.113	0.170	0.294	0.263	0.226	0.664	0.032	0.053	0.832	0.023	0.819	0.378	0.998	0.239	0.011
pM	0.271	0.132	0.165	0.135	0.189	0.292	0.268	0.244	0.166	0.197	0.131	0.180	0.192	0.327	0.160	0.167	0.081
PD	0.729	0.868	0.835	0.865	0.811	0.708	0.732	0.756	0.834	0.803	0.869	0.820	0.808	0.673	0.840	0.833	0.919
PE	0.459	0.728	0.847	0.769	0.599	0.444	0.606	0.459	0.624	0.502	0.638	0.696	0.752	0.459	0.690	0.627	0.920

N: number of individuals; H_{obs}: observed heterozygosity; H(p): probability of homozygosity; ET: exact test; LR: likelihood ratio; pM: match of probability; PD: power of discrimination; PE: power of exclusion.

TABLE 3—Allele frequencies at seventeen STR Loci in Naikpod Gond Tribe (N = 104) of Andhra Pradesh, India.

Alleles	D5S818	FGA	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	THO1	D13S317	D16S539	D2S1338	D19S433	vWA	TPOX	D18S51	Penta D	Penta E
5.0																	0.080
6.0								0.130									
7.0					0.029			0.159								0.003	0.010
8.0			0.005		0.264			0.154	0.255	0.034				0.192		0.001	
9.0	0.005				0.048	0.014		0.500	0.173	0.101				0.240	0.005	0.200	0.020
9.3								0.058									
10.0	0.154		0.216		0.197	0.260			0.139	0.159				0.091		0.130	
11.0	0.442		0.034		0.293	0.279			0.163	0.346		0.010		0.471	0.034	0.336	0.190
12.0	0.250		0.034		0.163	0.332	0.005		0.139	0.231				0.072	0.005	0.190	0.150
12.2														0.005			
13.0	0.144		0.082		0.005	0.115	0.005		0.087	0.101					0.112	0.090	0.020
13.2														0.010			
14.0	0.005		0.096				0.043		0.043	0.029				0.293	0.183	0.180	0.020
14.2														0.024		0.020	0.060
15.0			0.341				0.341							0.091	0.120	0.189	0.020
15.2														0.067		0.020	0.050
16.0			0.168				0.288				0.005			0.043	0.173	0.184	0.010
16.2														0.019		0.130	
17.0			0.014				0.183				0.029			0.005	0.183	0.092	0.130
17.2														0.010			
18.0	0.015	0.010					0.120				0.125			0.255	0.010	0.050	
19.0	0.059						0.014				0.226			0.063	0.063	0.040	
19.2	0.005																
20.0	0.054										0.149			0.024	0.005	0.010	
20.2	0.034																
21.0	0.074										0.072				0.024	0.030	
21.2	0.034																
22.0	0.108										0.043				0.024	0.030	
22.2	0.029																
23.0	0.137										0.226						
23.2	0.034																
24.0	0.123										0.067						
24.2	0.059																
25.0	0.098										0.029				0.010		
25.2	0.049																
26.0	0.049										0.029						
26.2	0.020																
27.0	0.010																
28.0	0.010		0.087														
29.0			0.178														
29.2			0.005														
30.0			0.216														
30.2			0.024														
31.0			0.058														
31.2			0.159														
32.0			0.034														
32.2			0.139														
33.0			0.014														
33.2			0.077														
34.2			0.010														
H _{obs}	0.702	0.833	0.760	0.769	0.692	0.750	0.702	0.654	0.837	0.692	0.769	0.779	0.856	0.683	0.913	0.800	0.920
H(p)	0.919	0.001	0.449	0.009	0.042	0.667	0.254	0.553	0.862	0.032	0.030	0.931	0.335	0.865	0.163	0.900	0.475
ET	0.641	0.060	0.398	0.583	0.410	0.186	0.020	0.432	0.279	0.589	0.744	0.538	0.541	0.664	0.694	0.695	0.675
LR	0.640	0.023	0.475	0.663	0.255	0.225	0.014	0.540	0.324	0.871	0.988	0.793	0.469	0.735	0.513	0.843	0.664
pM	0.303	0.080	0.210	0.142	0.225	0.269	0.250	0.319	0.170	0.221	0.154	0.218	0.181	0.325	0.134	0.215	0.112
PD	0.697	0.920	0.790	0.858	0.775	0.731	0.750	0.681	0.830	0.779	0.846	0.782	0.819	0.675	0.866	0.785	0.888
PE	0.431	0.662	0.526	0.543	0.416	0.510	0.431	0.630	0.668	0.416	0.543	0.560	0.706	0.402	0.821	0.543	0.843

N: number of individuals; H_{obs}: observed heterozygosity; H(p): probability of homozygosity; ET: exact test; LR: likelihood ratio; pM: match of probability; PD: power of discrimination; PE: power of exclusion.

TABLE 4—Allele frequencies at seventeen STR Loci in Yerukula Tribe (N = 101) of Andhra Pradesh, India.

Alleles	D5S818	FGA	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	THO1	D13S317	D16S539	D2S1338	D19S433	vWA	TPOX	D18S51	Penta D	Penta E
5.0								0.010									0.147
6.0								0.332									0.010
7.0	0.005				0.035			0.010	0.005								0.059
8.0	0.010		0.015		0.243			0.134	0.134	0.045				0.317		0.010	0.010
9.0	0.020				0.198			0.257	0.243	0.144			0.025	0.183		0.225	0.020
9.3								0.257									
10.0	0.104		0.173		0.233	0.178			0.050	0.030				0.079	0.005	0.196	
11.0	0.292		0.040		0.183	0.351			0.149	0.292				0.416	0.015	0.284	0.078
12.0	0.218		0.144		0.109	0.421			0.322	0.267			0.020	0.005	0.030	0.127	0.176
12.2													0.005				
13.0	0.267		0.015			0.050			0.030	0.168					0.167	0.088	0.069
14.0	0.040		0.243				0.025		0.020	0.054			0.282	0.302	0.364	0.039	0.020
14.2													0.015				
15.0	0.045		0.252				0.238		0.050				0.074	0.188	0.217	0.029	0.088
15.2													0.050				
16.0			0.119				0.134				0.010		0.109	0.099	0.071		0.127
16.2													0.015				
17.0							0.342				0.010			0.188	0.035		0.108
18.0							0.173				0.139			0.144	0.071		0.078
19.0		0.064					0.089				0.079			0.079	0.005		
20.0		0.134									0.134						0.010
20.2		0.015															
21.0		0.128									0.020				0.015		
21.2		0.010															
22.0		0.168									0.019				0.005		
22.2		0.005															
23.0		0.188									0.104						
23.2		0.030															
24.0		0.199									0.218						
25.0		0.025									0.213						
26.0		0.025									0.054						
27.0		0.010		0.005													
28.0				0.139													
29.0				0.149													
30.0				0.322													
30.2				0.114													
31.0				0.045													
31.2				0.104													
32.2				0.104													
33.2				0.020													
H _{obs}	0.792	0.812	0.762	0.812	0.802	0.703	0.713	0.733	0.772	0.713	0.852	0.673	0.792	0.762	0.859	0.765	0.824
H(p)	0.786	0.207	0.216	0.869	0.984	0.421	0.169	0.878	0.635	0.063	0.947	0.165	0.815	0.102	0.059	0.488	0.129
ET	0.675	0.336	0.129	0.144	0.228	0.262	0.013	0.174	0.023	0.510	0.268	0.248	0.326	0.506	0.561	0.565	0.150
LR	0.559	0.377	0.132	0.094	0.290	0.248	0.013	0.294	0.042	0.479	0.236	0.271	0.262	0.578	0.476	0.531	0.267
pM	0.219	0.144	0.189	0.182	0.199	0.335	0.230	0.261	0.209	0.212	0.151	0.266	0.199	0.313	0.220	0.197	0.110
PD	0.781	0.856	0.811	0.818	0.801	0.665	0.770	0.739	0.791	0.788	0.849	0.734	0.801	0.687	0.780	0.803	0.890
PE	0.584	0.621	0.531	0.621	0.603	0.433	0.448	0.481	0.549	0.448	0.698	0.388	0.584	0.531	0.712	0.535	0.643

N: number of individuals; H_{obs}: observed heterozygosity; H(p): probability of homozygosity; ET: exact test; LR: likelihood ratio; pM: match of probability; PD: power of discrimination; PE: power of exclusion.

In the present study, evaluation of 15 tetranucleotides (D5S818, FGA, TPOX, CSF1PO, D3S1358, THO1, D13S317, D16S539, vWA, D8S1179, D21S11, D7S820, D2S1338, D19S433 and D18S51) and 2 pentanucleotides loci (Penta D and Penta E) among the four populations revealed that these genetic markers are highly variable and hence ideal for human identification and analysis of genetic diversity in populations. For quality control, we followed laboratory internal control standards and kit controls.

The complete dataset is available upon request through electronic mail from the corresponding author at cfslkolkata@indiatimes.com, dtubprd@giascl01.vsnl.net.in and the webpage of Central Forensic Science Laboratory <http://humstr.ndacdb.org/>

Acknowledgment

The authors are thankful to the participants for providing their blood samples and Tribal Welfare Department, Govt. of Andhra Pradesh and the various NGO's for facilitating in this study. The work was carried out from the grant under Plan Project of CFSL, Kolkata. G. Hima Bindu was assisted by a national fellowship from Council of Scientific and Industrial Research (CSIR), New Delhi, India.

References

1. Singh KS. People of India, national series vol. IV, India's communities (A-Z), ASI. New Delhi: Oxford University Press, 1998.

2. Sambrook J, Fritsch EF, Maniatis T. Molecular cloning. A laboratory manual. 2nd ed. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press, 1989.
3. Applied Biosystems AmpF/STR® Identifiler™ PCR Amplification Kit. User's manual, Instructions for use of products.
4. Krenke BE, Treba A, Anderson SJ, et al. Validation of a 16-locus Fluorescent multiplex system. *J Forensic Sci.* 2002;47:773–85.
5. Nei M. Estimate of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 1978;89:583–90.
6. Guo SW, Thompson EA. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* 1992;48:361–72. [\[PubMed\]](#)
7. Weir BS. Independence of VNTR alleles defined as fixed bins. *Genetics* 1992;130:873–87. [\[PubMed\]](#)
8. Chakraborty R, Strivers D, Zhong Y, CHG. DNATYPE [computer program; Windows 95/NT ver]. Houston, TX: University of Texas, 1998.
9. Fisher RA. Standard calculations for evaluating a blood-group system. *Heredity* 1951;5:95–102. [\[PubMed\]](#)
10. Jones DA. Blood samples: Probability of discrimination. *J Forensic Sci Soc* 1972;12:355–59. [\[PubMed\]](#)
11. Garber RA, Morris JW. General equations for the average power of exclusion for genetic systems of n co-dominant alleles in one-parent cases of disputed parentage. In: Walker RH editors. Inclusion probabilities in

parentage testing. Arlington, VA: American Association of Blood Banks, U.S.A.; 1983;277–80.

12. Sahoo S, Kashyap VK. Allele frequency data for Powerplex 16 loci in four major populations of Orissa, India. *J Forensic Sci* 2002;47:912–15. [\[PubMed\]](#)
13. Ashma R, Kashyap VK. Genetic study of 15 important STR loci among four major ethnic groups of Bihar, India. *J Forensic Sci* 2002;47:1139–42. [\[PubMed\]](#)
14. Gaikwad S, Kashyap VK. Polymorphism at fifteen hypervariable microsatellite loci in four populations of Maharashtra, India. *Forensic Sci Int* 2002;126:267–71. [\[PubMed\]](#)
15. Sitalaximi T, Trivedi R, Kashyap VK. Genotype profile for thirteen tetranucleotide repeat loci and two pentanucleotide repeat loci in four endogamous Tamil population groups of India. *J Forensic Sci* 2002;47:1168–73. [\[PubMed\]](#)

Additional information and reprints requests:

V. K. Kashyap, Ph.D.
National DNA Analysis Center
Central Forensic Science Laboratory
30, Gorachand Road
Kolkata-700014, West Bengal
India

E-mail: dtubprd@giascl01.vsnl.net.in, cfslnkolkata@indiatimes.com