



Announcement of Population Data

Y chromosome STR allelic and haplotype diversity in five ethnic Tamil populations from Tamil Nadu, India

Kuppareddi Balamurugan^{a,*}, G. Suhasini^b, M. Vijaya^b, S. Kanthimathi^b, Nicole Mullins^a, Martin Tracey^c, George Duncan^d^aSchool of Criminal Justice, University of Southern Mississippi, 118 College Drive, P.O. Box 5127, Hattiesburg, MS 39406, USA^bDepartment of Genetics, Postgraduate Institute of Basic Medical Sciences, University of Madras, Chennai 600 113, India^cDepartment of Biology, Florida International University, University Park Campus, Miami, FL 33199, USA^dDNA Laboratory, Broward County Sheriff's Office, 201 SE 6th Street, Fort Lauderdale, FL 33301, USA

ARTICLE INFO

Article history:

Received 12 February 2010

Received in revised form 10 May 2010

Accepted 26 May 2010

Available online 10 July 2010

Keywords:

Y STR

Indian population

Y STR haplotypes

Gene diversity

Genetic distance

Y STR duplication

ABSTRACT

We have analyzed 17 Y chromosomal STR loci in a population sample of 154 unrelated male individuals of the Tamil ethnic group residing in the state of Tamil Nadu, Southern India using AmpFISTR[®] Yfiler[™] PCR amplification kit. The population samples consist of the following castes: Kongu Gounder (KOG), Nadar Hindu (NAH), Agamudayar (AGA), Parayar (PAR) and other Tamil individuals (MCT) of mixed castes. A total of 152 unique haplotypes were identified among the 154 individuals studied. The haplotype diversity was found to be 0.9935 or higher for all the five groups. The results of population pairwise *F_{st}* *p* values indicate no statistically significant differentiation between the five populations in this study, but the results were highly significant when compared with 12 other global populations (*p* < 0.05). Comparison of populations in this study with other national and global populations using Principal co-ordinate analysis (PCA) using Rst distance matrix indicates a delineation of all the Indian populations from other unrelated populations.

© 2010 Elsevier Ireland Ltd. All rights reserved.

Populations: A total of 154 samples from the state of Tamil Nadu, India, were analyzed for 16 Y chromosomal STR markers. The populations analyzed in this study include 100 Tamils from four endogamous populations namely Agamudayar (AGA, *n* = 18), Kongu Gounder (KOG, *n* = 20), Nadar Hindu (NAH, *n* = 32), Parayar (PAR, *n* = 30) and Tamils of unknown caste origin (MCT, *n* = 54). The four Tamil populations belong to different social hierarchical groups within the Hindu caste system. The hierarchical system is generally designated as forward, backward, most backward, scheduled castes and tribes. The sampling locations of the four Tamil caste populations were described elsewhere [1]. **Agamudayar:** The estimated population size is about 0.82 million [2]. **Kongu Gounder:** The estimated population size is about five million. **Nadar Hindu:** They constitute approximately 11 million of the Indian population with 90% practicing Hinduism. **Parayar:** This group is one of the scheduled caste populations of Tamil Nadu with an estimated population of 1.8 million (census report of India 2001).

Extraction: Blood samples were collected with informed consent from 100 unrelated individuals comprised of four endogamous populations and 54 unrelated individuals with no known caste affiliation. Samples were collected in accordance with the

ethical guidelines stipulated by the institutions involved in this study. DNA was isolated following the method as described in Miller et al. [3]. Total human DNA was quantitated using Quantifiler[®] Human DNA quantitation kit as per manufacturer's recommendations (Applied Biosystems, Foster City, CA).

PCR: Approximately 1 ng of DNA template was amplified at 17 Y chromosomal STR loci in a multiplex reaction using the AmpFISTR[®] Yfiler[™] PCR amplification kit (Applied Biosystems, Foster City, CA). The 17 Y STR loci amplified are DYS19, DYS385a, DYS385b, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS448, DYS456, DYS458, DYS635 and YGATAH4.

Typing: Amplicons were separated by capillary electrophoresis on an automated ABI 310 Genetic Analyzer (Applied Biosystems, Foster City, CA) and sized with Genescan-500 LIZ size standard. Allele calls and genotyping were carried out using allelic ladder provided in the kit and GeneMapper[®] ID v3.2 software (Applied Biosystems, Foster City, CA).

Quality control: Positive and negative controls as specified in the Yfiler[™] kit user's manual. The data were analyzed and verified by two independent analysts.

Analyses of data: STR allele frequencies and gene diversity estimates were calculated using Genepop version 4.0 [4,5]. The haplotype diversity was estimated using Nei's formula: $h = [n(1 - \sum x_i^2)] / (n - 1)$, where *n* is the number of samples analyzed

* Corresponding author. Tel.: +1 601 266 6048; fax: +1 601 266 4391.

E-mail address: Kuppareddi.Balamurugan@usm.edu (K. Balamurugan).

Table 1
The unique Y STR haplotypes and their frequency in the Tamil population.

Haplotypes	Sample	DYS19	DYS385	DYS389I	DYS389II	DYS390	DYS391	DYS392	DYS393	DYS437	DYS438	DYS439	DYS448	DYS456	DYS458	DYS635	YGATA H4	Frequency
1	KOG-3	16	16, 18	12	31	21	10	11	13	14	11	12	20	16	17	24	11	1
2	KOG-4	15	14, 16	12	28	22	10	14	11	15	10	14	19	15	14	21	12	1
3	KOG-5	15	15, 15	13	31	22	9	11	15	16	10	11	19	16	18	21	12	1
4	KOG-6	14	12, 19	14	31	23	10	10	14	16	11	10	19	15	20	24	12	1
5	KOG-8	16	13, 14	13	30	25	11	11	13	14	11	10	20	16	15	23	13	1
6	KOG-10	14	13, 16	12	28	22	10	14	11	15	10	12	20	15	15	22	12	1
7	KOG-11	17	16, 16	14	31	21	11	11	13	15	10	13	21	15	15	22	11	1
8	KOG-15	14	13, 17	12	29	22	10	14	11	15	10	12	20	15	15	24	11	1
9	KOG-16	15	13, 17	12	28	22	10	14	11	15	10	12	20	15	15	23	12	1
10	KOG-17	14	13, 16	12	28	22	10	14	11	15	10	13	19	15	14	24	12	1
11	KOG-18	16	11, 14	13	31	24	11	11	13	14	11	11	20	15	16	23	11	1
12	KOG-19	14	13, 17	12	28	22	11	14	11	15	10	12	19	15	14	23	12	1
13	KOG-20	14	13, 17	12	28	22	10	14	11	15	10	11	20	15	15	23	12	1
14	KOG-21	14	14, 17	13	30	23	10	11	13	15	9	12	20	15	16	21	12	1
15	KOG-22	15	15, 18	14	30	21	10	11	14	15	10	11	18	15	17	21	11	1
16	KOG-23	13	11, 17	13	31	23	9	12	13	14	10	12	19	15	18	21	12	1
17	KOG-24	15	11, 14	13	31	25	11	11	13	14	11	11	20	15	16	24	12	1
18	KOG-32	15	16, 17	14	30	22	10	11	12	14	9	11	19	15	15	20	12	1
19	KOG-33	15	15, 18	13	29	22	10	11	12	14	9	11	19	17	19	20	12	1
20	KOG-34	15	14, 17	13	30	22	10	11	12	14	9	11	19	16	17	20	12	1
21	NAH-1	15	16, 17	12	29	24	9	11	13	16	10	10	19	15	16	23	10	1
22	NAH-2	14	13, 19	13	28	23	10	10	15	16	11	10	19	15	17	26	12	1
23	NAH-3	15	13, 15	13	30	24	10	11	13	14	9	12	20	15	19	23	12	1
24	NAH-4	14	14, 17	14	30	22	11	15	13	14	11	12	19	15	20	22	10	1
25	NAH-5	14	15, 18	14	31	23	10	11	12	14	9	12	20	13	19	22	11	1
26	NAH-6	14	13, 17	11	27	22	10	14	11	15	10	12	18	15	16	24	12	1
27	NAH-7, 10	14	13, 17	12	28	22	10	14	11	15	10	12	18	15	16	24	12	2
28	NAH-8	14	13, 16	12	28	22	10	14	11	15	10	13	19	16	14	25	11	1
29	NAH-9	14	13, 18	12	28	22	10	14	11	15	10	12	19	16	15	24	12	1
30	NAH-11	14	13, 17	12	28	22	10	14	11	15	10	12	19	16	15	24	12	1
31	NAH-12	15	13, 19	13	29	23	10	10	14	16	11	10	18	16	19	24	12	1
32	NAH-13	15	13, 15	12	28	25	10	12	15	14	10	11	22	15	14	22	11	1
33	NAH-14	16	14, 15	12	28	24	10	12	15	14	10	11	22	15	14	22	11	1
34	NAH-15	15	15, 16	12	29	24	10	12	14	14	10	11	22	16	17	21	11	1
35	NAH-17	14	13, 17	12	28	22	10	14	11	15	10	12	18	15	15	24	12	1
36	NAH-18	14	12, 18	13	29	24	10	10	14	16	11	11	19	17	17	24	12	1
37	NAH-19	14	13, 15	14	31	23	9	11	13	15	9	11	20	15	16	23	11	1
38	NAH-20	14	15, 18	14	30	23	10	11	12	14	9	12	20	13	18	22	11	1
39	NAH-22	15	13, 16	12	28	22	10	14	11	15	10	12	19	17	14	23	12	1
40	NAH-23	14	14, 16	13	30	24	10	14	13	14	11	11	19	14	19	22	12	1
41	NAH-25	15	14, 15	13	29	22	10	11	13	15	10	11	19	16	19	22	11	1
42	NAH-26	15	15, 17	13	29	22	10	11	12	14	9	11	19	16	17	20	11	1
43	NAH-27	14	13, 18	13	30	22	10	14	11	14	10	12	19	15	14	26	11	1
44	NAH-28	15	13, 16	12	28	22	10	14	11	15	10	12	19	17	14	22	12	1
45	NAH-29	14	13, 15	14	30	23	10	11	13	15	9	11	21	15	17	24	11	1
46	NAH-30	14	13, 18	13	30	22	10	14	11	15	10	12	19	15	14	26	11	1
47	NAH 31	15	13, 17	14,15	30,31	21	10	11	13	14	10	12	18	15	17	24,26	11	1
48	NAH-32	15	13, 16	12	28	24	10	11	12	15	9	11	19	14	18	22	10	1
49	NAH-33	16	12, 15	13	31	22	11	11	12	14	10	12	19	14	17	21	12	1
50	NAH-34	14	14, 18	14	29	24	10	10	14	15	11	10	19	15	16	24	11	1
51	NAH-35	14	14, 17	12	28	22	10	14	11	15	11	13	18	15	15	24	12	1
52	AGA-3	14	18, 19	14	32	22	10	10	14	16	12	10	19	16	18	24	12	1
53	AGA-7	16	11, 14	12	30	25	10	11	13	14	11	10	20	11	17	23	12	1
54	AGA-8, 14	14	13, 17	13	31	22	10	14	11	15	11	13	19	15	15	22	13	2
55	AGA-10	15	14, 17	13	29	22	10	11	12	14	9	11	19	14	17	20	12	1
56	AGA-13	14	14, 14	13	29	23	10	11	13	14	9	11	20	15	15	21	12	1
57	AGA-16	14	13, 20	13	29	23	10	10	15	16	11	12	19	16	17	25	13	1
58	AGA-19	14	13, 17	13	31	22	10	14	11	15	11	13	19	15	16	22	13	1
59	AGA-20	15	16, 17	12	30	21	10	11	13	14	10	13	20	15	20	23	10	1
60	AGA-21	15	15, 17	12	28	21	10	11	12	14	8	12	19	15	14	20	11	1
61	AGA-22	15	12, 15	13	29	25	10	11	13	14	9	11	19	14	16	21	11	1
62	AGA-23	13	13, 18	13	29	23	10	10	13	16	11	10	20	15	18	25	12	1
63	AGA-24	14	15, 18	13	29	23	10	11	12	15	9	12	21	16	16	21	12	1
64	AGA-25	15	15, 17	14	30	22	10	11	12	14	8	12	19	15	17	20	12	1
65	AGA-28	15	15, 15	13	31	22	9	11	15	16	10	11	19	18	17	22	11	1
66	AGA-30	14	13, 17	12	28	24	10	11	12	14	9	12	19	14	16	21	11	1
67	AGA-31	15	14, 15	13	31	22	9	11	15	16	10	11	19	17	18	22	11	1
68	AGA-35	15	15, 15	13	30	22	9	11	15	16	10	11	19	18	17	22	11	1
69	PAR-2	16	11, 13	13	30	25	10	11	13	14	11	10	20	15	16	23	12	1
70	PAR-3	14	14, 16	12	28	22	10	14	11	15	10	11	19	16	15	24	12	1
71	PAR-4	15	12, 13	12	28	23	10	12	13	14	10	12	19	12	14	27	11	1
72	PAR-5	14	13, 18	12	29	22	11	14	11	16	11	13	20	15	15	25	12	1
73	PAR-6	16	15, 15	13	31	22	10	11	15	16	10	11	19	16	17	22	12	1
74	PAR-7	15	12, 19	12	28	23	10	10	14	16	11	10	19	16	19	24	12	1
75	PAR-8	15	9, 16	13	29	22	10	14	12	16	10	12	20	15	17	23	11	1
76	PAR-10	15	7, 16	13	29	22	11	14	12	16	10	11	19	15	19	23	11	1
77	PAR-11	14	14, 17	12	29	22	10	14	11	15	11	12	19	16	15	23	12	1

Table 1 (continued)

Haplotypes	Sample	DYS19	DYS385	DYS389I	DYS389II	DYS390	DYS391	DYS392	DYS393	DYS437	DYS438	DYS439	DYS448	DYS456	DYS458	DYS635	YGATA H4	Frequency
78	PAR-12	15	11, 14	13	29	25	10	11	13	14	11	10	20	14	16	23	13	1
79	PAR-13	15	15, 18	13	30	23	10	11	12	14	9	11	19	16	19	20	14	1
80	PAR-14	14	13, 15	12	28	22	10	14	11	15	10	12	19	17	14	24	12	1
81	PAR-15	15	15, 16	14	30	22	10	11	12	14	9	11	20	15	16	22	12	1
82	PAR-16	14	12, 17	12	28	22	10	14	11	15	10	12	19	16	16	25	11	1
83	PAR-18	15	16, 17	14	29	22	10	12	12	14	9	12	19	15	19	20	12	1
84	PAR-19	14	13, 17	12	28	22	10	14	11	16	10	13	19	15	15	23	12	1
85	PAR-20	16	13, 18	13	29	21	10	11	13	15	10	12	19	15	17	21	11	1
86	PAR-21	14	13, 16	12	28	22	11	14	9	15	10	12	19	15	15	24	12	1
87	PAR-22	13	13, 17	13	29	25	10	13	13	15	11	13	19	16	17	23	12	1
88	PAR-23	14	12, 18	14	30	23	10	10	14	16	11	10	19	15	16	24	11	1
89	PAR-24	15	7, 16	13	29	22	10	14	12	16	10	12	19	15	18	21	11	1
90	PAR-25	15	15, 16	13	29	22	10	11	12	14	9	11	19	17	16	20	12	1
91	PAR-26	17	15, 14	14	31	21	11	11	12	14	10	11	18	15	16	21	11	1
92	PAR-27	15	15, 17	12	30	24	10	11	14	14	10	11	19	15	16	18	12	1
93	PAR-28	15	15, 17	13	31	22	10	11	13	15	9	11	19	16	17	20	12	1
94	PAR-30	15	14, 17	12	28	23	10	14	11	15	11	13	19	15	16	23	12	1
95	PAR-31	15	13, 17	13	29	24	10	11	12	15	9	12	19	13	15	21	11	1
96	PAR-32	15	14, 17	12	29	24	10	11	13	16	10	10	19	16	16	24	11	1
97	PAR-33	15	16, 16	13	30	21	10	11	12	14	8	12	18	16	16	21	11	1
98	PAR-35	16	15, 16	13	29	22	10	11	12	14	9	12	19	16	17	20	10	1
99	MCT-2	15	16, 17	13	29	22	10	11	12	14	9	12	19	15	17	20	12	1
100	MCT-3	15	14, 16	12	28	22	10	11	13	15	10	12	20	15	18	24	11	1
101	MCT-5	16	12, 18	14	30	21	10	11	15	15	10	12	19	15	17	21	12	1
102	MCT-8	15	17, 17	14	30	22	11	11	12	14	9	11	19	14	17	20	11	1
103	MCT-9	15	11, 14	14	31	25	10	11	13	14	11	12	20	15	17	23	11	1
104	MCT-14	15	11, 15	13	31	23	10	11	14	14	11	10	20	15	15	23	13	1
105	MCT-16	15	13, 16	12	28	22	10	11	15	16	10	12	21	15	16	21	12	1
106	MCT-17	14	13, 15	13	29	22	10	14	11	15	10	10	19	16	14	24	12	1
107	MCT-18	15	12, 20	13	29	23	10	10	14	15	11	10	19	15	19	24	12	1
108	MCT-20	14	15, 17	13	31	24	10	11	14	14	10	12	19	15	16	18	12	1
109	MCT-21	15	11, 11	14	32	25	10	11	13	14	11	10	18	15	15	23	12	1
110	MCT-22	15	14, 19	14	30	23	10	10	13	16	11	10	19	15	18	25	12	1
111	MCT-32	16	11, 14	13	30	25	11	11	13	14	11	11	20	14	16	23	12	1
112	MCT-34	16	13, 20	13	30	23	10	10	13	16	11	12	19	15	17	24	12	1
113	MCT-38	15	14, 17	14	30	21	11	12	13	15	10	12	18	15	16	21	12	1
114	MCT-39	16	11, 14	13	30	25	11	12	13	14	11	10	20	16	16	23	12	1
115	MCT-41	14	13, 17	12	28	22	10	14	11	15	10	12	19	15	15	25	12	1
116	MCT-42	15	15, 16	12	29	24	10	12	13	14	10	12	22	16	18	22	11	1
117	MCT-43	14	16, 19	14	30	23	10	12	12	14	9	12	20	15	18	23	11	1
118	MCT-44	14	14, 17	12	28	22	10	14	11	14	10	13	19	15	13	22	12	1
119	MCT-45	15	14, 17	13	29	22	10	11	14	14	10	11	19	14	18	21	11	1
120	MCT-48	14	13, 17	12	28	22	10	14	11	15	10	11	19	15	15	24	12	1
121	MCT-49	15	17, 18	14	32	22	10	11	12	14	9	11	19	16	16	20	12	1
122	MCT-54	15	16, 17	13	27	22	10	11	12	14	9	11	19	16	16	20	13	1
123	MCT-55	14	13, 16	12	28	22	10	14	11	15	10	13	19	17	14	24	12	1
124	MCT-57	14	14, 17	14	30	22	10	11	13	15	10	12	18	16	15	21	13	1
125	MCT-59	14	13, 17	12	28	22	10	13	11	15	10	12	19	16	15	24	12	1
126	MCT-60	15	16, 16	13	30	21	10	11	12	14	8	11	17	16	16	23	11	1
127	MCT-61	15	13, 18	15	32	22	11	11	12	14	9	12	19	16	18	20	12	1
128	MCT-63	14	12, 18	13	29	24	11	10	14	16	11	11	19	15	18	24	12	1
129	MCT-66	14	13, 18	14	31	24	10	10	13	16	11	10	18	15	17	24	12	1
130	MCT-67	15	11, 14	14	32	25	10	11	13	14	11	11	20	15	18	23	13	1
131	MCT-68	16	11, 14	14	32	25	10	11	13	14	11	10	20	16	15	23	12	1
132	MCT-69	13	14, 16	13	30	23	11	11	14	14	10	11	19	17	18	18	12	1
133	MCT-71	15	15, 17	13	29	22	10	11	12	14	9	11	18	15	17	20	11	1
134	MCT-72	15	14, 16	12	28	22	10	14	11	15	11	11	18	16	15	24	12	1
135	MCT-74	15	15, 16	12	29	24	10	12	14	14	10	12	22	15	17	21	11	1
136	MCT-75	16	15, 18	13	29	22	10	11	12	14	9	12	18	15	17	20	11	1
137	MCT-76	16	9, 15	13	29	25	10	11	13	14	11	10	20	16	17	23	14	1
138	MCT-77	13	15, 15	12	28	22	10	14	11	15	10	13	19	15	14	24	12	1
139	MCT-78	15	15, 18	13	29	21	10	11	12	14	8	11	19	15	15	20	11	1
140	MCT-83	13	10, 13	12	29	23	10	11	13	14	10	12	19	15	17	19	11	1
141	MCT-84	14	12, 19	14	29	22	10	10	15	16	11	10	20	15	18	25	12	1
142	MCT-87	14	13, 17	13	29	22	9	14	11	15	10	12	20	16	15	23	12	1
143	MCT-88	16	11, 14	14	32	25	10	11	13	14	11	10	20	15	16	23	13	1
144	MCT-91	15	14, 18	13	29	22	11	11	12	14	9	11	19	16	18	20	12	1
145	MCT-94	14	13, 15	12	28	21	10	14	11	15	10	12	19	15	15	24	12	1
146	MCT-95	15	11, 11	14	32	25	10	11	13	14	11	10	19	16	16	23	13	1
147	MCT-96	16	11, 15	13	30	25	11	11	13	14	12	10	19	15	16	23	13	1
148	MCT-97	15	11, 14	14	31	25	11	11	13	14	12	11	19	15	17	23	13	1
149	MCT-98	15	12, 14	14	31	25	10	11	13	14	11	10	20	15	18	23	12	1
150	MCT-99	16	13, 20	13	29	23	11	10	14	16	11	10	19	15	15	25	12	1
151	MCT-101	15	13, 16	13	29	25	10	11	13	14	9	11	19	15	17	22	12	1
152	MCT-103	15	11, 14	14	30	25	10	11	13	14	11	10	20	15	16	23	13	1

and x_i is the frequency of the i th haplotype [6]. Population pairwise analysis of Rst and Fst genetic distances for comparison with other

populations was carried out using Arlequin version 3.1 [7]. Graphical representation of genetic distances (Rst) of the populations in

this study along with other national and global populations was performed based on principal co-ordinate analysis (PCA) plot using GenAlEx software package version 6.3 [8]. Analysis of Molecular variance was also performed by the same software package.

Results: A total of 152 unique haplotypes were identified among the 154 individuals studied with one haplotype shared by two individuals of the NAH population and another shared by two individuals of the AGA population. The unique haplotypes observed in this study along with their frequency are presented in Table 1. Allele frequencies, gene diversity and haplotype diversity are presented in Table 2 (Supplementary material). The lowest gene diversity (0.2391) was observed at locus DYS391 in the PAR population while the highest gene diversity (0.8891) was observed at locus DYS385 in the MCT population. The haplotype diversity was found to be 0.9935 or higher for all the five groups. The overall haplotype diversity for all the samples studied was 0.9998. One of the samples (NAH 31) showed a duplication of loci at DYS389I, DYS389II and DYS635 with biallelic results (Fig. 2, Supplementary material). These results were reconfirmed by a second amplification of the sample with similar results. Also, analysis of this sample with 15 autosomal STR loci shows a clean, single source DNA profile with no evidence of contamination and these results validate the duplication of the three loci in this sample.

Several Y STR allele duplications or triplications have either been published or referenced in the databases. In a recent study, Butler et al. [9] have compiled 42 such duplications or triplications in Y chromosomal STRs. The available information on duplication events in Y chromosomal regions is probably an underestimate, because several duplications may not be identifiable because duplicated alleles of the same size are not discernible or they are not reported. The observation of biallelic results in one out of 154 samples studied suggests that these events are rare, but have been reported in other populations as well [9–11]. The relative frequency of these duplications is known to be little over one percent [9]. Since the DYS635 and DYS389I/II are separated only by 233 kbp [9], it is possible that an entire segment of the Y chromosome, Yq11.21 may be duplicated in the NAH31 sample rather than individual locus duplications. Four other Y STR loci, DYS434, DYS435, DYS437 and DYS439 are located between DYS635 and DYS389 loci. The peak height of locus DYS437 in NAH31 sample is twice the height of the peaks neighboring this locus in the electropherogram, and highly suggestive of duplication of this locus also (Fig. 2, Supplementary material, bottom panel). It is possible that the DYS439 locus is also duplicated considering the location of this locus on the Y chromosome (between DYS437 and DYS389), but this cannot be substantiated.

Haplotype database comparisons: All the 152 haplotypes with 16 markers were searched against the YHRD database (www.yhrd.org)

with no matching results indicating the uniqueness of the haplotypes in these groups. A haplotype search against the Applied Biosystems Y STR database (www.appliedbiosystems.com) generated eight matches, with two from a Thailand population and six from an Asian Indian population. Only the two non-Indian haplotype matches suggest these haplotypes are unique to the Indian subcontinent.

Other remarks: Comparison of populations in this study with other national and global populations using principal co-ordinate analysis (PCA) of Rst genetic distance matrix indicates grouping of all the Indian populations in the center of the plot with other unrelated populations segregated away from the center (Fig. 1). The two African populations (SAF, NAM) occupied one quadrant of the plot, the Caucasian, Brazil and Croatian populations occupied a second quadrant while the Chinese and Japanese populations occupied a third quadrant indicating similarities of the related population while maintaining distance from dissimilar populations. The results of population pairwise F_{st} p values indicate no statistically significant differentiation between the five populations in this study, but the results were highly significant when compared with 12 other global populations ($p < 0.05$). Fifty eight of the 60 comparisons showed significant differentiation at or below the 5% level. Additionally, analysis of molecular variance (AMOVA) indicated a 99% variance within populations and 1% variance among the five populations in this study, which adds more evidence to the level of homogeneity observed in these populations. This is not surprising due to the fact that all the populations in this study reside in the state of Tamil Nadu, and speak the common language Tamil which is one of the Dravidian languages of Southern India. The limited Y chromosomal differentiation among the five populations suggests a common founding Y gene pool for the Southern Indian male populations in this study. Other Y STR data presented by Frank et al. [12] reported 118 unique haplotypes in a population of 123 male individuals with a haplotype diversity of 0.9989. These high numbers of unique haplotypes and haplotype diversity are consistent with the results of this study. The haplotype data from Frank et al. [12] which is part of the YHRD database did not yield any matching haplotypes with the results of the present study further suggesting the uniqueness of these haplotypes in the Southern region of India.

Recently, Watkins et al. [13] studied several caste populations spanning different social hierarchy from Tamil Nadu using autosomal STRs, Y chromosomal SNPs, and mitochondrial haplotypes and concluded that the Tamil south Indian castes are only modestly differentiated from one another with 0.96% of STR variance. Estimates for heterozygosity and repeat variance in these populations indicated no substantial between-caste differences or excess Homozygosity in these caste groups. Also, there is extensive sharing of Y

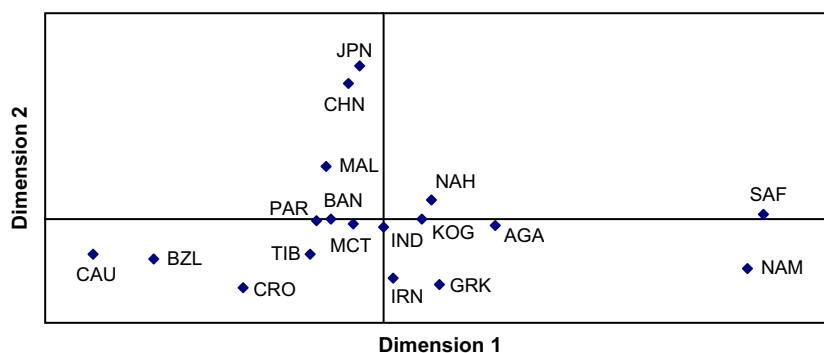


Fig. 1. Principal co-ordinate analysis (PCA) plot derived from the Rst genetic distance matrix of the five populations from this study compared to other national and global populations. Population key: AGA – Agamudayar, present study; BAN – Bangladesh [14]; BZL – Brazil [15]; CAU – Caucasian [16]; CHN – Chinese [17]; CRO – Croatia [18]; GRK – Greek [19]; IND – Indian [17]; IRN – Iran [20]; JPN – Japanese [21]; KOG – Tamil, present study; MAL – Malaysian [17]; MCT – Tamil, present study; NAH – Tamil present study; NAM – Namibia, Africa [22]; PAR – Tamil, present study; SAF – South Africa [23]; TIB – Tibet [24].

and mitochondrial haplogroups among castes, and the overall level of inter-caste differentiation is low. Additionally the high diversity of certain specific haplotypes in the South Indian and Indus valley populations has led to the suggestion that there is little genetic influence from other Eurasians on the castes of south India.

The data presented here can be used for regional, national, and global haplotype comparisons, human identification, population differentiation and for purposes of human migration, dispersal and settlement.

Conflict of interest: The authors declare no conflict of financial or other interests.

Acknowledgements

The authors thank Dr. A. Ramesh, Chairman, Department of Genetics, University of Madras, Chennai, India, for sharing the samples for this study. This study was partly supported by a startup grant for one of the authors (K.B.) from the University of Southern Mississippi, Hattiesburg, MS. Partial financial support was provided to the Department of Genetics, University of Madras by the Department of Biotechnology, Government of India. The authors also thank the volunteers for providing the samples for this study.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.legalmed.2010.05.008.

References

- [1] Kanthimathi S, Vijaya M, Ramesh A. Genetic study of Dravidian castes of Tamil Nadu. *J Genet* 2008;87:175–9.
- [2] Second Backward class commission of Tamil Nadu, 1989.
- [3] Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988;6:1215.
- [4] Raymond M, Rousset F. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J Hered* 1995;86:248–9.
- [5] Rousset F. Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Mol Ecol Resources* 2008;8:103–6.
- [6] Nei M. *Molecular evolutionary genetics*. New York: Columbia University Press; 1987.
- [7] Excoffier L, Laval G, Schneider S. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evol Bioinform Online* 2005; 1:47–50.
- [8] Peakall R, Smouse PE. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* 2006;6:288–95.
- [9] Butler JM, Decker AE, Kline MC, Vallone PM. Chromosomal duplications along the Y-chromosome and their potential impact on Y-STR interpretation. *J Forensic Sci* 2005;50:853–9.
- [10] Sanchez JJ, Brion M, Parson W, Blanco-Verea AJ, Børsting C, Lareu M, et al. Duplications of the Y-chromosome specific loci P25 and 92R7 and forensic implications. *Forensic Sci Int* 2004;140:241–50.
- [11] Kirsch S, Münch C, Jiang Z, Cheng Z, Chen L, Batz C, et al. Evolutionary dynamics of segmental duplications from human Y-chromosomal euchromatin/heterochromatin transition regions. *Genome Res* 2008;18:1030–42.
- [12] Frank WE, Ralph HC, Tahir MA. Y chromosome STR haplotypes and allele frequencies in a Southern Indian male population. *J Forensic Sci* 2008;53:248–51.
- [13] Watkins WS, Thara R, Mowry BJ, Zhang Y, Witherspoon DJ, Tolpinrud W, et al. Genetic variation in South Indian castes: evidence from Y-chromosome, mitochondrial, and autosomal polymorphisms. *BMC Genet* 2008;9:86.
- [14] Alam S, Ali ME, Ferdous A, Hossain T, Hasan M, Akhteruzzaman S. Haplotype diversity of 17 Y-chromosomal STR loci in the Bangladeshi population. *Forensic Sci Int Genet* 2010;4:e59–60.
- [15] Schwengber SP, Kommers T, Matte CHF, Raimann PE, Carvalho BA, Leite FPN, et al. Population data of 17 Y-STR loci from Rio Grande do Sul state (South Brazil). *Forensic Sci Int Genet* 2009;4:e31–3.
- [16] Available from: <http://www.cstl.nist.gov/strbase/NISTpop.htm>.
- [17] Chang YM, Perumal R, Keat PY, Kuehn DLC. Haplotype diversity of 16 Y-chromosomal STRs in three main ethnic populations (Malays, Chinese and Indians) in Malaysia. *Forensic Sci Int* 2007;167:70–6.
- [18] Haliti N, Carapina M, Masic M, Strinovic D, Klaric IM, Kubat M. Evaluation of population variation at 17 autosomal STR and 16 Y-STR haplotype loci in Croatians. *Forensic Sci Int Genet* 2009;3:e137–8.
- [19] Kovatsi L, Saunier J, Irwin JA. Population genetics of Y-chromosome STRs in a population of Northern Greeks. *Forensic Sci Int Genet* 2009;4:e21–2.
- [20] Roewer L, Willuweit S, Stoneking M, Nasidze I. A Y-STR database of Iranian and Azerbaijanian minority populations. *Forensic Sci Int Genet* 2009;4:e53–5.
- [21] Mizuno N, Kitayama T, Fujii K, Nakahara H, Yoshida K, Sekiguchi K, et al. A forensic method for the simultaneous analysis of biallelic markers identifying Y chromosome haplogroups inferred as having originated in Asia and the Japanese archipelago. *Forensic Sci Int Genet* 2010;4:73–9.
- [22] Fujihara J, Yuasa I, Muro T, Iida R, Tsubota E, Nakamura H, et al. Allele frequencies and haplotypes for 28 Y-STRs in Ovambo population. *Leg Med* 2009;11:205–8.
- [23] Kido A, Fujitani N, Hara M, Kimura H. Genetic data of 16 Y-chromosomal short tandem repeat loci in Africans from South Africa. *J Forensic Sci* 2006;51:1414–6.
- [24] Li KL, Kai L, Yuming M. Y chromosome STR haplotypes of Tibetan Living Tibet lassa. *Forensic Sci Int* 2007;172:79–83.