

# HLA-A allele frequency and haplotype distribution in the Dravidian tribal communities of South India

R. Thomas, M. Banerjee

Human Molecular Genetics Laboratory, Rajiv Gandhi Centre for Biotechnology, Thiruvananthapuram, Kerala 695 014, India

**BACKGROUND:** The tribal communities of South India are considered to be the original inhabitants of the Indian sub-continent, belonging to the most primitive Dravidian speaking communities. These Dravidian speaking forest dwelling tribal populations have remained isolated from any intermingling with other non-tribal communities.

**AIMS AND OBJECTIVES:** We propose to understand the evolutionary processes mediated by molecular, functional and immunological information based on human leukocyte antigen (HLA) genetic system.

**MATERIAL AND METHODS:** The HLA-A diversity was analyzed in seven Dravidian tribal populations namely Malapandaram, Paniya, Kurichiya, Kanikkar, Adiya, Kattunaikka and Kuruma of Kerala, South India using the PCR-SSP method. The tribal communities were compared with a group comprising of random non-tribal Dravidian samples of southern India.

**RESULTS:** In the present study, 11 HLA-A alleles were identified in the South Indian population. The most frequent alleles included HLA-A\*24, A\*02, A\*33 and \*A11. HLA-A\*24 had the highest frequency in all the tribal groups while, A\*02 was the highest frequency allele in the RND group. The haplotype Cw\*14-A\*24 was present in all the populations. The three-locus haplotype B\*52-Cw\*14-A\*24 was observed in all the populations except Kurichiya and Kanikkar.

**CONCLUSION:** The study suggests that the RND population is highly diverse and more likely to have an admixed origin.

**Key words:** Dravidian, HLA-A, haplotype, South India, Tribals

inhabitants of the Indian sub-continent,<sup>[2]</sup> belonging to the most primitive Dravidian speaking communities. The tribal communities of India continue to extensively hunt and gather as well as practice low input shifting cultivation. The highlands of the Western Ghats are home to 34 tribal communities of Kerala.<sup>[3]</sup> These Dravidian speaking forest dwelling tribal populations have remained isolated from any intermingling with other non-tribal communities. These communities are scattered in different regions of the forests of the Western Ghats and comprise 1.8% of the total population. Class and caste distinctions isolate the tribal populations, which remain geographically and culturally isolated in small endogamous groups and have very little contact with the mainstream population. The endogamous tribal communities in Kerala are some of the oldest and most evolutionarily conserved groups in India. Consanguineous marriages are very common among these isolated communities predisposing them to several genetic disorders, leading to drastic population decline.

The human leukocyte antigen (HLA) genetic system has been extensively used to understand the evolutionary processes mediated by molecular, functional and immunological information.<sup>[4-5]</sup> Several investigators have used the HLA system to understand human genetic relatedness and migrations of different world populations.<sup>[4,6-9]</sup> In an earlier study we have reported the HLA -B and -C polymorphism in various tribal communities of South India.<sup>[10]</sup> In the present study, we have analyzed the HLA-A allele frequency and HLA class I haplotypes in seven tribal populations of Kerala and one non-tribal population. The HLA-A data is pertinent in constructing the complete HLA class I haplotypes and

Southern India is one of the oldest geophysical regions of the world.<sup>[1]</sup> This peninsular region is surrounded by the Bay of Bengal in the east, the Indian Ocean on the south and the Arabian Sea on the west. Majority of the people of southern India speak languages belonging to the Dravidian language family. The tribal communities of South India are considered to be the original

in understanding the migration patterns of different populations in India.

## Materials and Methods

Blood samples were collected from individuals who voluntarily consented to participate in the study. The mode of sampling was random and only unrelated individuals were selected for the study. All the seven tribal communities used in the study dwell in isolated colonies in the deep forest and due care was taken to avoid sampling from related individuals and siblings. The seven tribal communities (N=107) include the Malapandaram, Paniya, Kurichiya, Kanikkar, Adiya, Kattunaikka and the Kuruma. Tribal community leaders were approached with the help of Non Governmental Organizations working with the seven tribal populations from different regions of Kerala. The purpose of the study was explained to the tribal leaders who in turn informed their communities.

Blood samples were collected with informed consent from a Dravidian population group (n=77) consisting of randomly selected non-tribal Malayalam speaking individuals from the staff and students of Rajiv Gandhi Centre for Biotechnology, Thiruvananthapuram, Kerala. We refer to them as the Random Non-tribal Dravidian (RND) population group as they belong to different castes and communities and form a heterogeneous Dravidian population of South India. This group was used to understand the genetic structuring of tribal and non-tribal communities of the region. This research work was approved by the Institutional Human ethical committee of the Rajiv Gandhi Centre for Biotechnology, Thiruvananthapuram, Kerala. Around 5-10 mL of peripheral blood samples were collected in EDTA vials from the subjects and stored at 4°C.

Molecular HLA typing was carried out using the polymerase chain reaction with sequence specific primer (PCR-SSP) method for the HLA-A locus. Amplification primers for HLA-A were synthesized based on published literature.<sup>[11]</sup> The basic principle of the PCR-SSP typing utilizes the same PCR protocol and parameters for typing all the HLA-A alleles. As an additional quality control measure, 10 reference DNA samples from the 12<sup>th</sup>

International Histocompatibility Workshop (1995) were also typed simultaneously. Most of the alleles are detected in group-specific amplifications corresponding to serological specificities although some allele specific detection is possible. HLA antigens are defined here as two digit -"alleles" based on allelic-group molecular resolution.

## Results

All the 107 samples used in the present study did not show any ambiguity on amplification. Hardy-Weinberg equilibrium (HWE), maximum-likelihood estimate of allele frequencies, haplotype frequencies, linkage disequilibrium, relative linkage disequilibrium and their significance as determined by chi-square test were calculated using the Arlequin v.2.0. (<http://anthropologie.unige.ch/arlequin>).<sup>[12-13]</sup>

The HLA-A allele diversity was characterized in seven tribal communities of Kerala, South India using the PCR-SSP method. In the present study, 11 HLA-A alleles were detected in the tribal and RND populations. There was no significant deviation ( $P>0.05$ ) from HWE in any of the populations. The HLA-A allele distribution is summarized in Table 1. The RND had the highest polymorphism of HLA-A alleles with all 11 alleles detected. The number of HLA-A alleles in each community ranged between 2 and 8 alleles. The Kurichiya community had the highest number of alleles while the Paniya had the least number of alleles. HLA-A\*24 was present in all the communities and had the highest allele frequency in all the tribal groups. HLA A\*02 and A\*33 alleles were present in all the communities except Paniya. HLA\*A11 was absent only in the Malapandaram population. HLA A\*02 was the highest frequency allele in the RND group. The HLA bi-locus haplotype frequencies occurring at a frequency above 10 percent are summarized in Table 2. The haplotype Cw\*14-A\*24 was present in all the populations. The other major shared haplotypes include B\*52-A\*24, B\*58-A\*33, Cw\*07-A\*02, Cw\*10-A\*33, Cw\*15-A\*24. The three loci haplotypes at a frequency above 10 percent are shown in Table 3. The three-locus haplotype B\*52-Cw\*14-A\*24 was present in all the populations except Kurichiya and Kanikkar.

**Table 1: HLA-A allele frequencies in the tribal communities of Kerala**

HLA-A	PN n = 10	MP n = 10	AD n = 23	KR n = 10	KA n = 22	KN n = 17	KU n = 15	RND n = 77
A*01		0.1		0.05			0.1	0.091
A*02		0.1	0.239	0.25	0.114	0.118	0.033	0.195
A*03			0.044	0.05				0.058
A*11	0.15		0.044	0.05	0.318	0.118	0.033	0.156
A*24	0.85	0.35	0.544	0.25	0.409	0.706	0.433	0.169
A*26		0.05	0.022					0.026
A*30						0.029		0.006
A*31		0.05		0.05				0.065
A*32					0.046		0.067	0.026
A*33		0.35	0.044	0.1	0.091	0.029	0.333	0.162
A*68			0.022	0.2	0.023			0.045

PN-Paniya, MP-Malapandaram, AD-Adiya, KR-Kurichiya, KA-Kanikkar, KN-Kattunaikka, KU-Kuruma, RND- Random Non-tribal Dravidian

**Table 2: HLA bi-locus haplotypes occurring with a frequency of more than ten percent**

Haplotype	PN	MP	AD	KR	KA	KN	KU	RND
B*07-A*24	0.05		0.043	0.2			0.099	0.015
B*07-A*68				0.15	0.023			
B*35-A*24			0.044			0.26	0.094*	
B*51-A*24	0.1		0.087*				0.033	
B*52-A*24	0.25	0.05	0.196		0.045	0.265*	0.033	0.026
B*57-A*01		0.1*		0.05*				0.026
B*57-A*02		0.1*						0.026
B*58-A*33		0.135		0.05	0.046*	0.029*	0.265	0.045*
B*61-A*02			0.109	0.1				0.023
B*61-A*24	0.4*		0.152		0.182	0.147	0.2*	0.028
B*62-A*24				0.05*	0.159			
Cw*01-A*11					0.11			0.007
Cw*04-A*24			0.044	0.05		0.147		
Cw*06-A*01		0.1*						0.013
Cw*06-A*02		0.1*			0.068*	0.029*		0.017
Cw*07-A*02			0.022	0.1	0.023	0.029	0.033	0.059
Cw*07-A*68				0.15				
Cw*10-A*24	0.1	0.05			0.027			
Cw*10-A*33		0.25		0.05*	0.063	0.029	0.2*	0.052*
Cw*12-A*11	0.15							
Cw*12-A*24	0.15		0.119			0.029		
Cw*14-A*24	0.5	0.2	0.314	0.2	0.035	0.382	0.333	0.135*
Cw*15-A*02			0.107	0.05				0.053
Cw*15-A*24	0.05	0.05			0.05	0.118	0.333	

\* Haplotypes in linkage disequilibrium,  $P < 0.001$

PN-Paniya, MP-Malapandaram, AD-Adiya, KR-Kurichiya, KA-Kanikkar, KN-Kattunaikka, KU-Kuruma, RND- Random Non-tribal Dravidian

## Discussion

There is considerable debate about the evolutionary histories of the Indian tribals. The proto-Australoid tribals, who speak dialects belonging to the Austric linguistic group, are believed to be the basic element in the Indian population.<sup>[14]</sup> Other anthropologists, historians and linguists have supported the view that the Austro-Asi-

atic (a subfamily of the Austric language family) speaking tribals are the original inhabitants of India.<sup>[15-17]</sup> Elsewhere it has been proposed that the Dravidians are the original inhabitants, the Austro-Asiatics being later immigrants.<sup>[2]</sup> Recent findings based on genomic and mitochondrial markers<sup>[18]</sup> support the view that Dravidians were spread throughout India<sup>[19]</sup> and were later restricted to Southern India to avoid linguistic dominance.

**Table 3: HLA-B, -C, -A haplotypes occurring with a frequency of more than 10%**

Haplotype	PN	MP	AD	KR	KA	KN	KU	RND
B*07-Cw*07-A*24	0.05		0.022	0.2				0.013 <sup>†</sup>
B*07-Cw*07-A*68				0.1				
B*07-Cw*14-A*24			0.022 <sup>†</sup>				0.1	
B*52-Cw*12-A*11	0.1							
B*52-Cw*12-A*24	0.1 <sup>†</sup>		0.022 <sup>†</sup>					
B*52-Cw*14-A*24	0.15 <sup>†</sup>	0.05	0.152			0.265	0.033	0.026
B*57-Cw*06-A*01		0.1		0.05				0.019
B*57-Cw*06-A*02		0.1						0.019
B*58-Cw*10-A*33		0.1			0.046	0.029	0.2	0.039
B*61-Cw*10-A*24	0.1	0.2						
B*61-Cw*14-A*24	0.3				0.137	0.029 <sup>†</sup>		0.006 <sup>†</sup>
B*61-Cw*14-A*33		0.15						
B*61-Cw*15-A*24			0.022 <sup>†</sup>		0.068	0.118	0.033	

<sup>†</sup> Haplotypes not in linkage disequilibrium,  $P > 0.001$

PN-Paniya, MP-Malapandaram, AD-Adiya, KR-Kurichiya, KA-Kanikkar, KN-Kattunaikka, KU-Kuruma, RND- Random Non-tribal Dravidian

In the present study we observe that the major HLA-A alleles in the tribals (HLA A\*24, A\*02, A\*11 and A\*33) are shared with the South Indian RND group and North Indian populations.<sup>[8]</sup> However A\*01 is present at a very high frequency in the North Indian population while it is sparsely distributed in the Dravidian communities. A\*33 and A\*11 alleles were restricted to the Asian populations of Pakistan, North India and Orientals.<sup>[7-9]</sup> It is notable that A\*24 is present with the highest frequency in all the tribal communities and North Indian population, but not in the other world populations.<sup>[7-9]</sup> A\*31 has been reported to be the most frequent allele in the Dravidian caste groups of South India (TamilNadu),<sup>[20]</sup> however in the present study it was present only at a lower frequency in the RND group and mostly absent among the tribals. This shows that there might be considerable difference between the genetic make-up of Dravidian communities in TamilNadu and Kerala.

The presence of the bi-locus haplotype Cw\*14-A\*24 in all the South Indian populations may be characteristic of the Dravidian population. The presence of the haplotype B\*52-Cw\*14- A\*24 in five of the seven tribal groups is very interesting. This haplotype may be an ancestral haplogroup identifying the earliest inhabitants of Southern India. This haplotype is absent in the Kurichiya and Kanikkar populations. It is present at a lesser frequency among the RND group. The Kurichiya have a high social status and are thought to represent undifferentiated stock of proto-historic Dravidian speakers of Kerala comprising the Nair and Ezhava.<sup>[21]</sup> Previ-

ous studies on HLA-B and -C allele diversity<sup>[10]</sup> also show that the Kurichiya cluster with the mainstream population compared to the other tribal communities. The Kanikkar who are considered typical of Malids belong to the pre-Dravidian stock, which is believed to be closest to the non-tribal communities.<sup>[22]</sup> The Kanikkar tribal group has been progressively shifting from the tribal traditions and has more contact with the communities lying near the forest fringes. This has led to admixture in their gene pool, leading to loss of ancestral haplotypes shared by more primitive isolated communities. Paniya, Adiya and Kattunaikka may be the most primitive communities in comparison to the others due to the high frequency of the B\*52-Cw\*14- A\*24 haplotype. The RND group though a non-tribal population is more admixed and may be a reason for the presence of this haplotype.

The present study on HLA-A corroborates the earlier observation based on HLA-B and -C that the RND population is more diverse and likely to have an admixed origin. The complete HLA class I data will be more valuable to bone marrow transplantation registries and useful in the study of molecular anthropology and genetic predisposition to autoimmune and infectious diseases, vaccine development and vaccinogenomics.

## References

1. Majumdar DN. Races and Cultures of India. Bombay (India): Asia Publishing House; 1961.

2. Sarkar SS. Race and race movements in India. *In*: Chatterjee SK, editor. The cultural heritage of India, Calcutta: The Ramakrishna Mission Institute of Culture; 1958. pp. 17–32.
3. Singh KS. People of India: An introduction Calcutta: Anthropological Survey of India; 1992.
4. Clayton J, Lonjou C. Allele and haplotype frequencies for HLA loci in various ethnic groups. *In*: Charron D, editor. Genetic Diversity of HLA Functional and medical Implications, Paris: EDK Medical and Scientific International Publishers; 1997. pp 665-820.
5. Meyer D, Thomson G. How selection shapes variation of the human major histocompatibility complex: a review. *Ann Hum Genet* 2001;65:1-26.
6. Cox ST, Marsh SG, Scott I, Clayton J, Arguello JR, McWhinnie AJ, *et al.* HLA-A, -B, -C polymorphism in a UK Ashkenazi Jewish potential bone marrow donor population. *Tissue Antigens* 1999;53:41-50.
7. Mohyuddin A, Ayub Q, Khaliq S, Mansoor A, Mazhar K, Rehman S, *et al.* HLA polymorphism in six ethnic groups from Pakistan. *Tissue Antigens* 2002;59:492-501.
8. Rajalingam R, Krausa P, Shilling HG, Stein JB, Balamurugan A, McGinnis MD, *et al.* Distinctive KIR and HLA diversity in a panel of north Indian Hindus. *Immunogenetics* 2002;53:1009-19.
9. Machulla HK, Batnasan D, Steinborn F, Uyar FA, Saruhan-Direskeneli G, Oguz FS, *et al.* Genetic affinities among Mongol ethnic groups and relationship to Turks. *Tissue Antigens* 2003;61:292-9.
10. Thomas R, Nair SB, Banerjee M. HLA-B and HLA-C alleles and haplotypes in the Dravidian tribal populations of southern India *Tissue Antigens* 2004;64:58-65.
11. Tonks S, Marsh SGE, Bunce M, Bodmer, JG. Molecular typing for HLA class I using ARMS-PCR: Further developments following the 12<sup>th</sup> International Histocompatibility Workshop. *Tissue Antigens* 1999;53:175-83.
12. Excoffier L, Slatkin M. Maximum-likelihood estimation of molecular haplotype frequency in a diploid population. *Mol Biol Evol* 1995;12:921-7.
13. Schneider S, Roessli D, Excoffier L. Arlequin Ver 2.0: A Software for Population Genetics Data Analysis. Switzerland: Genetics and Biometry Laboratory, University of Geneva; 2000.
14. Thapar R. A history of India Vol. 1 Middlesex: Penguin; 1966.
15. Risley HH. The people of India Calcutta: Thacker Spink; 1915.
16. Rapson EJ. Peoples and languages. *In*: Rapson EJ, editor. Cambridge history of India, Vol. 1. Delhi: S Chand and Co; 1955.
17. Pattanayak DP. The language heritage of India. *In*: Balasubramanian D, Rao NA, editors. The Indian human heritage. Hyderabad: Universities Press; 1998. pp. 95–9.
18. Basu A, Mukherjee N, Roy S, Sengupta S, Banerjee S, Chakraborty M, *et al.* Ethnic India: a genomic view, with special reference to peopling and structure. *Genome Res* 2003;13:2277-90.
19. Thapar R. Early India: From the origins to AD 1300. California: University of California Press; 2003.
20. Shankarkumar U. HLA A19 subtypes and B loci related haplotype in selected caste groups from the Indian population. *Indian J Hum Genet* 2003; 9:13-16.
21. *In*: Menon TM, editor. The Encyclopaedia of Dravidian Tribes. Vol. I-III. Trivandrum. The DLA publications; 1997.
22. Brahui AR. Brahui. *In*: Menon MT, editor. The encyclopaedia of Dravidian tribes. Vol. III. Trivandrum: The DLA publications; 1997.

## Subscription form

Kindly enter my subscription to “Indian Journal of Human Genetics”

Name of the subscriber .....

Current institutional attachment .....

Designation .....

Delivery address .....

City ..... Pin code .....

State ..... Country .....

Phone no. (with STD/ISD code) .....

E-mail address .....

### Subscription details

Subscription period One year

Subscription type India / Foreign

Subscription starts from January ..... (year)

### Payment details

Cheque No. .... Dated .....

Drawn on ..... Amount .....

Signature ..... Date: .....