Raju *et al.*
Egyptian Journal of Medical Human Genetics (2022) 23:162
https://doi.org/10.1186/s43042-022-00378-7

Egyptian Journal of Medical Human Genetics

RESEARCH

Open Access

The HLA profile and genetic affinities of three primitive Tamil-speaking endogamous groups: Kallars of Thanjavur, Piramalai Kallar and Vanniyar

Kamaraj Raju[1], Balakrishnan Karuppiah[2*] and Rathika Chinniah[2]

Abstract

Background: The present study was aimed to study the frequencies of HLA-DRB1/-DQB1 alleles and haplotypes of three endogamous groups of Tamil Nadu state, South India. PCR-SSP typing of HLA-DRB1 and -DQB1 alleles were performed on 111 Kallars of Thanjavur, 80 Piramalai Kallar of Madurai and 119 Vanniyar. Genetic distances, neighbor-joining phylogenetic dendrograms and correspondence analysis have been performed.

Results: The HLA class II alleles, DRB1*07 (25.2%), DRB1*15 (15.7%), DRB1*14 (11.7%) and DRB1*12 (9.90%) among Kallars of Thanjavur; DRB1*15 (28.7%), DRB1*04 (15.6%), DRB1*10 (14.3%), DRB1*13 (11.2%) and DRB1*03 (9.37%) among Piamalai Kallar and DRB1*15 (24.7%), DRB1*04 (15.9%), DRB1*07 (11.7%), DRB1*12 (11.3%) and DRB1*10 (10.0%) among Vanniyar were more frequent. Similarly, alleles DQB1*06 (31.0%), DQB1*02 (26.5%) and DQB1*05 (24.7%) among Kallars of Thanjavur; DQB1*05 (32.5%), DQB1*06 (31.8%), DQB1*02 (16.2%) and DQB1*03:02 (12.5%) among Piramalai Kallar and DQB1*05 (52.9%), DQB1*06 (22.6%) and DQB1*02 (11.3%) among Vanniyar were more frequent. We genotyped the two most frequent two-locus haplotypes, such as DRB1*15-DQB1*06 and DRB1*07-DQB1*02 for HLA-A/-B/–C alleles to identify the 5-locus extended haplotypes to extrapolate global affinities. We identified a number of five locus extended haplotypes among south Indian population with stronger global affinities. Further, we identified the presence of a highly unique extended haplotypes such as A*11-B*35-C*12-DRB1*07-DQB1*02 (HF:0.1458) in Kallars of Thanjavur, A*03-B*35-C*04-DRB1*15-DQB1*06 (HF:0.1833) in Piramalai Kallar and A*03-B*07-C*07-DRB1*15-DQB1*06 (HF: 0.1800) in Kallars of Thanjavur and (HF: 0.1081) in Vanniyar population.

Conclusions: Allele distribution and haplotype analysis have demonstrated that the Kallars of Thanjavur, Piramalai Kallar and Vanniyar populations shared HLA alleles with other ethnic and other Indian populations, while showing population specific haplotypes. Analysis of population-specific distribution of HLA alleles is proved to be important in finding out the relatedness of the ethnic groups across continents. The extensive polymorphism of the HLA system also has useful application in the study of the origin, evolution and migration patterns of human populations.

Keywords: HLA allele frequencies, Haplotype frequencies, Kallars of Thanjavur, Piramalai Kallar, Vanniyar, South Indian population

Background

Tamil Nadu is a land of enormous cultural, geographical and linguistic diversity and harbors appreciably greater genetic diversity than any other comparable global region after Africa. From time immemorial, Central Asia is

*Correspondence: immunobala@mkuniversity.org

[2] Department of Immunology, School of Biological Sciences, Madurai Kamaraj University (University with Potential for Excellence), Tamil Nadu 625021 Madurai, India
Full list of author information is available at the end of the article

Springer Open

© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

thought to be an important reservoir of genetic diversity and the source of at least three major waves of migration leading into Europe, the Americas, and India [1]. We recently established that two Tamil-speaking populations of the South India, the Mukkuvar, an endogamous group from coastal regions of South Tamil Nadu state, more closely related to ancient Hispanic and Guanche populations and the Valayar, a group inhabiting predominantly inland hilly regions with forest cover, are related to Austronesian and Micronesian populations [2].

The human leukocyte antigen (HLA) polymorphism in human populations has been studied for long to investigate immunogenetics, human genetic relationships and in reconstructing past migrational histories of populations. The extensive allelic variation among the HLA class I and class II genes distinguishes these as the most polymorphic coding sequence loci in the human genome [3]. The analysis of these genes has been a valuable tool in unraveling the historical relationships between ethnic groups and has greatly increased the knowledge of the ancestry and migration patterns of many populations including various endogamous groups in South India. South Indian populations are of interest for the study of genetic polymorphisms, because of the relatively long isolation of these populations from other ethnic groups. As Dobzhansky put it: 'South Indian caste systems are the biggest experiment ever done on Homo sapiens.'

However, in a recent South Indian population base study, an increased frequency of class II alleles DRB1*07, DRB1*15, DQB1*02, DQB1*05 and DQB1*06 have been reported. On the other hand, DRB1*01, DRB1*09 and DRB1*16 have been identified with low frequency, particularly among North Indian populations [4].

The purpose of this study was to analyze the alleles of the HLA class II (DRB1, DQB1) loci in three primitive Tamil-speaking Dravidian endogamous population groups such as Kallars of Thanjavur, Piramalai Kallar (of Madurai) and Vanniyar. Furthermore, the DNAs of most common two-locus haplotypes, such as DRB1*15-DQB1*06 and DRB1*07-DQB1*02 were typed for HLA-A, -B and -C alleles to identify the possible 5-locus extended haplotypes (EH) in these population groups. The frequencies of HLA-DRB1 and DQB1 alleles and haplotypes and phylogenetic relationships of these populations with nineteen other global populations are also constructed and discussed.

Methods

Study population

The Kallars of Thanjavur samples (*n*, 111) were collected from Thanjavur, Piramalai Kallar samples (*n*, 80) were collected from Madurai and Theni and Vanniyar samples (*n*, 119) were collected from Salem and Tiruvannamalai in the state of Tamil Nadu. Peripheral blood was collected via venipuncture from randomly chosen healthy individuals. Institutional ethical clearance was obtained from Madurai Kamaraj University (Ethical and Review Board Committee). Ethnicity was determined based upon current residence, place of birth, and family language history. Demographic data obtained included sex and age. Genomic DNA was extracted using salting out method as described previously [5].

Ethnographic notes

Kallars of Thanjavur

Kallars are a wide-spread, ancient population living in southern parts of Tamil Nadu. Traditionally, they were described as semi agriculturists and semi warriors. Kallars are known to be the oldest immigrants of Neolithic period with Mediterranean racial elements [6]. The Chola country or Tanjore was their original abode of the Kallars, and from there they migrated to the Pandya Kingdom following its subjugation by the Cholas around eleventh century A.D [7]. The Cullaries are said to be in general a brave people, expert in the use of the lance and in throwing the curved stick called 'vullaree taddee' (Kallar Taddee or Valari) [8]. Portions of the Madurai and Thanjavur districts are divided into areas known as nadus, a name which is specially applicable to Kallar tracts [9]. In each "nadu" a certain caste, called the Nattan, is the predominant factor in the settlement of social questions which arise among the various castes living within the nadu (Pinnaiyur Nadu, Kasavala Nadu, Keel Vengai Nadu, Konoor Nadu, etc.). The Kallars of Thanjavur comprise an endogamous subgroup of the Kallar community exhibiting exogamy at the clan level while maintaining strict endogamy at the subgroup level (Kandiyar,Vandayar, etc.).

Piramalai Kallar

An endogamous subset of Kallars, called 'Piramalai Kallars', live west of Madurai city (78.1E, 9.6N). In 1975, 95% of the population of Usilampatti and Thirumangalam taluks and 60% of Periyakulam taluks of Madurai district and 60% of Uthamapalayam taluk of Theni district were Piramalai Kallars, with a population size of approximately 350,000 [10]. In some old records they are referred to as Anaiyur Kallars which place was the stronghold of the Piramalai Kallars during the early years of their settlement [7].

Vanniyar

The Vanniyans, Pallis or Padaiyachis, are found in all the Tamil districts. The bulk of them are still labourers, but many now farm their own lands, while others are engaged in trade [11]. The Vanniyans are at the present

time a small and obscure agricultural caste, but there is reason to believe that they are descendants of ancestors who, in former times, held a good position among the tribes of South India [9].

Vanniyas are mentioned in Ceylon archives. Wanni (Wavuniya, in the recent times) is the name of a district in Ceylon. It is situated towards Trincomalee in the northeast quarter [12]. The name "Vanniyan", seems to have been introduced by the Brahmans, possibly to gratify the desire of the Pallis for genealogical distinction. Padaiyachi means a soldier, and is also of late origin. After the fall of the Pallava dynasty, the Pallis became agricultural servants, and it is only since the advent of British rule that they have begun to assert their claims to a higher position [13].

HLA class I and II DNA typing

HLA-A,-B,-C,-DRB1,-DQB1 alleles were genotyped by PCR with sequence specific primers (SSP) [14]. HLA-A/-B/-C genotyping was performed for selected DNA samples which possessed highly frequent two locus haplotypes DRB1*15-DQB1*06 and DRB1*07-DQB1*02 to detect the five locus ancestral extended haplotype (AEH).

Statistical analysis

Frequency of haplotypes of two and five loci was obtained by direct counting and expressed as indicated in the corresponding tables. The frequency data were converted into genetic distances in Arlequin (v3. 1). Dendrogram was constructed using Molecular and Evolutionary Genetics Analysis (MEGAv3.1) [15]. The genetic distances were also used for Principal Component Analysis (PCA) using GenAlEx 6 [16].

Results

Allele frequencies

The frequency distribution of HLA-DRB1 and -DQB1 alleles among three populations are presented (Table 1). The study revealed the presence of thirteen HLA-DRB1* and seven HLA-DQB1* alleles. The most frequent class II alleles, based on their frequency were: DRB1*07 (25.2%), DRB1*15 (15.7%), DRB1*14 (11.7%) and DRB1*12 (9.90%) among Kallars of Thanjavur; DRB1*15 (28.7%), DRB1*04 (15.6%), DRB1*10 (14.3%), DRB1*13 (11.2%) and DRB1*03 (9.37%) among Piramalai Kallar and DRB1*15 (24.7%), DRB1*04 (15.9%), DRB1*07 (11.7%), DRB1*12 (11.3%) and DRB1*10 (10.0%) among Vanniyar populations. Similarly, the alleles DQB1*06 (31.0%), DQB1*02 (26.5%) and DQB1*05 (24.7%) revealed higher frequencies among Kallars of Thanjavur; alleles DQB1*05 (32.5%), DQB1*06 (31.8%), DQB1*02 (16.2%) and DQB1*03:02 (12.5%) revealed higher frequencies among Piramalai Kallar and alleles DQB1*05 (52.9%), DQB1*06 (22.6%) and DQB1*02 (11.3%) revealed higher frequencies among Vanniyar (Table 1).

Table 1 Frequencies of HLA-DRB1 and -DQB1 alleles in selected populations of South India

DRB1/DQB1 alleles	Kallars of Thanjavur (*n*, 111)	Piramalai Kallar (*n*, 80)	Vanniyar (*n*, 119)
DRB1*01	1.35 (03)	3.12 (05)	1.26 (03)
DRB1*03	3.15 (07)	9.37 (15)	5.88 (14)
DRB1*04	8.10 (18)	15.6 (25)	15.9 (38)
DRB1*07	25.2 (56)	4.37 (07)	11.7 (28)
DRB1*08	6.30 (14)	4.37(07)	2.52 (06)
DRB1*09	0.45 (01)	ND	0.42 (01)
DRB1*10	11.7 (26)	14.3 (23)	10.0 (24)
DRB1*11	1.80 (04)	0.62 (01)	3.36 (08)
DRB1*12	9.90 (22)	8.12 (13)	11.3 (27)
DRB1*13	3.60 (08)	11.2 (18)	4.62 (11)
DRB1*14	11.7 (26)	ND	7.98 (19)
DRB1*15	15.7 (35)	28.7 (46)	24.7 (59)
DRB1*16	0.90 (02)	ND	ND
DQB1*02	26.5 (59)	16.2 (26)	11.3 (27)
DQB1*04	0.90 (02)	1.87 (03)	0.42 (01)
DQB1*05	24.7 (55)	32.5 (52)	52.9 (126)
DQB1*06 (01/02/03)	31.0 (69)	31.8 (51)	22.6 (54)
DQB1*03:01, 03:04 (DQ7)	9.00 (20)	3.75 (06)	5.04 (12)
DQB1*03:02 (DQ 8)	5.85 (13)	12.5 (20)	5.88 (14)
DQB1*03:03 (DQ 9)	1.80 (04)	1.25 (02)	1.68 (04)

ND Not detected

Hardy–Weinberg equilibrium examination

Hardy–Weinberg exact tests were performed on two HLA loci (DRB1 and DQB1). The observed, expected homozygosities, Chi-square (X^2) and the statistical *P* value are given in Table 2. Significant *P*- values were categorized by their levels ($P<0.0001$, $P<0.0003$) for convenience of discussion and whether the heterozygotes were in deficit or in excess. In order to make biologically meaningful comparisons, only heterozygotes which have a difference larger than two between the 'observed' and 'expected' counts were presented [17].

Haplotype frequencies

The two-locus haplotypes such as DRB1*07-DQB1*02 ($n=48$; 0.2162), DRB1*15-DQB1*06 ($n=25$; 0.1126), DRB1*14-DQB1*05 ($n=21$; 0.0945), DRB1*14-DQB1*06 ($n=19$; 0.0855), DRB1*10-DQB1*05 ($n=18$; 0.0810), DRB1*12-DQB1*06 ($n=18$; 0.0810), DRB1*07-DQB1*06 ($n=17$; 0.0765), DRB1*07-DQB1*05 ($n=16$; 0.0720), DRB1*12-DQB1*05 ($n=14$; 0.0630), DRB1*15-DQB1*05 ($n=13$; 0.0585) and DRB1*15-DQB1*02 ($n=12$; 0.0540)

Table 2 The Hardy–Weinberg equilibrium of HLA-DRB1 and -DQB1 loci in selected populations of South India

Kallars of Thanjavur						Piramalai Kallar				Vanniyar			
Locus	**Genotypes**	**Obs**	**Exp**	X^2	***p*-value**	**Obs**	**Exp**	X^2	***p*-value**	**Obs**	**Exp**	X^2	***p*-value**
DRB1	Homozygote Reference	41	50.2	14.735	0.0001	27	34.9	14.821	0.0001	44	53.1		
	Heterozygote	70	51.5			53	37.2			75	56.7	12.975	0.0003
	Homozygote Variant	04	13.2			02	9.9			06	15.1		
DQB1	Homozygote Reference	49	57.1	14.214	0.0001	33	39.9	13.839	0.0001	48	56.7		
	Heterozygote	62	45.7			47	33.2			71	53.6	12.902	0.0003
	Homozygote Variant	01	9.1			ND	ND			04	12.7		

Obs. Observed, *Exp*. expected, *ND* not detected

were the most frequent (> 0.05) among Kallars of Thanjavur. Haplotypes, DRB1*15-DQB1*06 ($n=30$; 0.1875), DRB1*15-DQB1*05 ($n=27$; 0.1687), DRB1*10-DQB1*05 ($n=21$; 0.1312), DRB1*10-DQB1*06 ($n=17$; 0.1062), DRB1*03-DQB1*02 ($n=15$; 0.0937), DRB1*13-DQB1*06 ($n=12$; 0.0750) and DRB1*12-DQB1*05 ($n=10$; 0.0625) were predominant among Piramalai Kallar. Similarly, haplotypes DRB1*15-DQB1*05 ($n=39$; 0.1638), DRB1*15-DQB1*06 ($n=37$; 0.1554), DRB1*12-DQB1*05 ($n=27$; 0.1134), DRB1*04-DQB1*05 ($n=24$; 0.1008), DRB1*10-DQB1*05 ($n=21$; 0.0882) and DRB1*07-DQB1*05 ($n=18$; 0.0756) were the most frequent (> 0.05) among Vanniyar (Table 3).

Based on the frequencies, the dominant two-locus haplotypes in these three populations were selected and typed for HLA-A, -B and -C locus alleles to find out the five-locus ancestral/ extended haplotype (AEH). The AEH represented more than five numbers was only presented for further analysis. The predominant five locus-extended haplotypes are presented in Table 4. The Ancestral Extended Haplotypes (AEH), A*11-B*35-C*12-DRB1*07-DQB1*02 (HF:0.1458; $n=14/48$), A*03-B*35-C*12-DRB1*07-DQB1*02 (HF:0.1041; $n=10/48$), A*03-B*07-C*07-DRB1*15-DQB1*06 (HF:0.1800; $n=09/25$),A*24-B*35-C*04-DRB1*15-DQB1*06(HF:0.1600; $n=08/25$), A*01-B*35-C*04-DRB1*07-DQB1*02 (HF:0.0833; $n=08/48$), A*03-B*40-C*04-DRB1*15-DQB1*06 (HF:0.1400; $n=07/25$), A*11-B*35-C*04-DRB1*07-DQB1*02 (HF:0.0729; $n=07/48$), A*03-B*35-C*04-DRB1*07-DQB1*02 (HF:0.0729; $n=07/48$), A*03-B*07-C*12-DRB1*15-DQB1*06 (HF:0.1200; $n=06/25$), A*11-B*07-C*07-DRB1*15-DQB1*06 (HF:0.1000; $n=05/25$), A*24-B*35-C*04-DRB1*07-DQB1*02 (HF:0.0520; $n=05/48$) and A*11-B*40-C*07-DRB1*15-DQB1*06 (HF:0.1000; $n=05/25$) were the most frequent in Kallars of Thanjavur. The Ancestral Extended Haplotypes, A*03-B*35-C*04-DRB1*15-DQB1*06 (HF:0.1833; $n=11/30$), A*24-B*35-C*04-DRB1*15-DQB1*06 (HF:0.1166; $n=07/30$), A*24-B*35-C*12-DRB1*15-DQB1*06 (HF:0.0833; $n=05/30$) and A*24-B*35-C*07-DRB1*15-DQB1*06 (HF:0.0666; $n=04/30$) were the most frequent in Piramalai Kallar. Similarly, A*03-B*07-C*07-DRB1*15-DQB1*06 (HF:0.1081; $n=08/37$), and A*11-B*07-C*07-DRB1*15-DQB1*06 (HF:0.0675; $n=05/37$) were the frequent ones in Vanniyar population (Table 4).

Phylogenetic affinities

In order to assess the relationship between the populations, genetic distance (DA) was calculated and presented as neighbor-joining (NJ) tree and Principal Component Analysis. The neighbor-joining of the population relationships is presented in Fig. 2. The constructed NJ tree showed population relationships as a series of bifurcations, which are commonly interpreted as population splits.

The present study has been carried out to explore the utility of frequencies of alleles of HLA-DRB1 loci in inferring phylogenetic relationships. The NJ dendrogram revealed a unique genetic background of the study populations (Kallars of Thanjavur, Piramalai Kallar and Vanniyar). Based on the analysis, the N6 group may be genetically related to both N4 (Micronesian, Australian Aborigine and Austronesian) and N5 nodes (Melanesian) or they might reveal an admixture of Polynesian, N4 and N5 nodes. The Mediterranean populations are divided from N1 node and form a different cluster and cluster together with Black population (Fig. 1).

The PCA performed for HLA-DRB1 frequencies revealed the position of each population in two dimensions. The plot shows that the Kallars of Thanjavur belong to the same cluster as Austronesian, Australian Aborigine, indigenous groups and shared genetic similarities with the Mediterranean and Blacks. Similarly, Piramalai Kallar and Vanniyar belong to the same cluster as Austronesian and Australian Aborigine indigenous groups, Micronesian and Oriental indigenous populations. The PCA notably separated the admixed populations of Amerindian (who are Native Indians of the

Table 3 Frequencies of HLA-DRB1-DQB1 two-locus haplotypes in selected populations of South India

DRB1-DQB1	HF (%)		
	Kallars of Thanjavur (*n*, 111)	Piramalai Kallar (*n*, 80)	Vanniyar (*n*, 119)
DRB1*03-DQB1*02	**0.0270 (06)**	**0.0937 (15)**	**0.0462 (11)**
DRB1*03-DQB1*05	0.0045 (01)	0.0312 (05)	0.0420 (10)
DRB1*03-DQB1*06	0.0045 (01)	0.0062 (01)	0.0210 (05)
DRB1*04-DQB1*02	0.0270 (06)	0.0187 (03)	0.0210 (05)
DRB1*04- DQB1*05	0.0270 (06)	0.0437 (07)	**0.1008 (24)**
DRB1*04- DQB1*06	0.0495 (11)	0.0437 (07)	0.0378 (09)
DRB1*07- DQB1*02	**0.2162 (48)**	0.0312 (05)	**0.0462 (11)**
DRB1*07- DQB1*04	ND	ND	0.0168 (04)
DRB1*07- DQB1*05	**0.0720 (16)**	0.0125 (02)	**0.0756 (18)**
DRB1*07- DQB1*06	**0.0765 (17)**	ND	0.0336 (08)
DRB1*07- DQB1*07	0.0270 (06)	ND	ND
DRB1*07- DQB1*08	0.0180 (04)	ND	0.0294 (07)
DRB1*07- DQB1*09	0.0270 (06)	0.0062 (01)	0.0084 (02)
DRB1*08- DQB1*05	0.0090 (02)	0.0148 (03)	0.0210 (05)
DRB1*08- DQB1*06	0.0225 (05)	0.0250 (04)	ND
DRB1*10- DQB1*02	0.0315 (07)	ND	0.0168 (04)
DRB1*10- DQB1*05	**0.0810 (18)**	**0.1312 (21)**	**0.0882 (21)**
DRB1*10- DQB1*06	0.0405 (09)	**0.1062 (17)**	0.0210 (05)
DRB1*11- DQB1*05	ND	ND	0.0336 (08)
DRB1*12- DQB1*05	**0.0630 (14)**	**0.0625 (10)**	**0.1134 (27)**
DRB1*12- DQB1*06	**0.0810 (18)**	0.0312 (05)	0.0294 (07)
DRB1*12- DQB1*07	0.0180 (04)	0.0312 (05)	0.0210 (05)
DRB1*13- DQB1*05	0.0045 (01)	0.0437 (07)	0.0420 (10)
DRB1*13- DQB1*06	0.0225 (05)	0.0750 (12)	0.0126 (03)
DRB1*14- DQB1*05	**0.0945 (21)**	ND	0.0462 (11)
DRB1*14- DQB1*06	0.0855 (19)	ND	0.0168 (04)
DRB1*15- DQB1*02	0.0540 (12)	0.0312 (05)	0.0378 (09)
DRB1*15- DQB1*05	**0.0585 (13)**	**0.1687 (27)**	**0.1638 (39)**
DRB1*15- DQB1*06	**0.1126 (25)**	**0.1875 (30)**	**0.1554 (37)**
DRB1*15- DQB1*07	0.0315 (07)	ND	0.0042 (01)

ND-Not Detected

Americas), Mestizo (Spain and Latin America which originally meant to denote a person of combined European and Amerindian descent) and West Asia populations (Persian, Mulatto, Jew and Hispanic) from our study populations (Fig. 2).

Discussion

Southern India is one of the oldest geophysical regions of the world [18]. The majority of the people of southern India speak languages belonging to the Dravidian language family. The original name for the Dravidian family was 'Tamulic', but the term 'Dravidian' was substituted by Bishop Caldwell, in order that the designation 'Tamil' might be reserved for the language of that name [19]. Studies on HLA polymorphism allow the delineation of their origin, geographic distribution patterns of these gene variants (alleles) and phylogenetic relatedness with world populations. In this study, the immunogenetic polymorphism of HLA-A/-B/-C/-DRB1*/-DQB1* alleles were genotyped by PCR using sequence specific primers (SSP). The present study on Kallars of Thanjavur, Piramalai Kallar and Vanniyar populations, three ancient and sympatrically isolated endogamous caste groups of Tamil Nadu, have revealed their unique HLA-DRB1/DQB1 allele and haplotype profiles.

The HLA class II alleles DRB1*04, DRB1*07, DRB1*10, DRB1*12, DRB1*14 and DRB1*15, were found with higher frequencies in study populations of south India. However, these alleles were completely absent in some ethnic populations: such as the allele DRB1*04, was completely absent in Berber and Guanche; allele DRB1*07, was completely absent in Australian Aborigine,

Table 4 Frequencies of extended haplotypes (EH) associated with the most frequent two-locus haplotypes DRB1*15-DQB1*06 and DRB1*07-DQB1*02 in selected populations and their possible origin/ affinities

A-B-C-DR-DQ	Haplotype frequency (%)			Total nos	Possible origin/affinities*
	KLT (2N)	PKL (2N)	VAN (2N)		
A*24-B*35-C*04-DRB1*15-DQB1*06	0.1600 (08)	0.1166 (07)	0.0405 (03)	18	Asian, Hispanic, Caucasoid
A*03-B*07-C*07-DRB1*15-DQB1*06	0.1800 (09)	ND	0.1081 (08)	17	Caucasoid, Asian, Hispanic
A*11-B*35-C*12-DRB1*07-DQB1*02	0.1458 (14)	0.1000 (01)	0.0454 (01)	16	Caucasoid
A*03-B*40-C*04-DRB1*15-DQB1*06	0.1400 (07)	0.01 (06)	0.0405 (03)	16	Caucasoid
A*03-B*35-C*04-DRB1*15-DQB1*06	0.0400 (02)	0.1833 (11)	0.0270 (02)	15	Hispanic, Black
A*24-B*35-C*07-DRB1*15-DQB1*06	0.1800 (09)	0.0666 (04)	0.0135 (01)	14	Asian, Caucasoid
A*03-B*35-C*12-DRB1*07-DQB1*02	0.1041 (10)	0.1000 (01)	ND	11	Not available
A*24-B*35-C*12-DRB1*15-DQB1*06	0.0200 (01)	0.0833 (05)	0.0540 (04)	10	Asian, Caucasoid
A*11-B*07-C*07-DRB1*15-DQB1*06	0.1000 (05)	ND	0.0675 (05)	10	Asian, Hispanic, Caucasoid
A*11-B*35-C*04-DRB1*07-DQB1*02	0.0729 (07)	0.2000 (02)	ND	9	Asian, Hispanic, Caucasoid
A*01-B*35-C*04-DRB1*07-DQB1*02	0.0833 (08)	ND	ND	8	Caucasoid
A*03-B*35-C*04-DRB1*07-DQB1*02	0.0729 (07)	0.1000 (01)	ND	8	Asian, Hispanic, Caucasoid
A*03-B*07-C*12-DRB1*15-DQB1*06	0.1200 (06)	ND	0.0270 (02)	8	Asian, Caucasoid
A*11-B*35-C*04-DRB1*15-DQB1*06	0.0400(02)	0.0166 (01)	0.0540 (04)	7	Asian, Hispanic
A*11-B*40-C*07-DRB1*15-DQB1*06	0.1000 (05)	ND	0.0135 (01)	6	Asian
A*24-B*40-C*07-DRB1*15-DQB1*06	0.0400 (02)	0.0333 (02)	0.0270 (02)	6	Asian
A*24-B*07-C*07-DRB1*15-DQB1*06	0.0400 (02)	ND	0.0405 (03)	5	Asian, Hispanic, Caucasoid
A*24-B*35-C*04-DRB1*07-DQB1*02	0.0520 (05)	ND	ND	5	Asian, Hispanic, Caucasoid
A*11-B*40-C*04-DRB1*15-DQB1*06	0.0600 (03)	ND	0.0270 (02)	5	Asian
A*24-B*40-C*04-DRB1*15-DQB1*06	0.0200 (01)	0.0333 (02)	0.0270 (02)	5	Asian

**Cf* AFND, *ND* not detected

Melanesian, Micronesian and Polynesian populations; allele DRB1*10, was completely absent in Amerindian, Australian Aborigine, Berber, Caucasoid, Guanche, Hispanic, Jew, Kurd, Melanesian, Mestizo, Micronesian, Mulatto, Oriental, Persian and Siberian; allele DRB1*12, was completely absent in Amerindian, Arab, Berber, Guanche, Hispanic, Jew, Kurd, Mestizo, Mulatto and Persian and allele DRB1*15 was completely absent in Guanche and Mulatto populations [20].

The alleles DQB1*05, DQB1*06, DQB1*02 and DQB1*03:02 were identified with higher frequencies in all the three study populations. Further, the two-locus haplotypes such as DRB1*15-DQB1*06, DRB1*07-DQB1*02, DRB1*15-DQB1*05, DRB1*10-DQB1*05, DRB1*12-DQB1*05, DRB1*10-DQB1*06 and DRB1*04-DQB1*05 were present in significant frequencies. The predominant extended haplotypes (more than 10 nos in all the 3 populations combined together) were: A*24-B*35-C*04-DRB1*15-DQB1*06 (*n*, 18), A*03-B*07-C*07-DRB1*15-DQB1*06 (*n*, 17), A*11-B*35-C*12-DRB1*07-DQB1*02 (*n*, 16), A*03-B*40-C*04-DRB1*15-DQB1*06 (*n*, 16), A*03-B*35-C*04-DRB1*15-DQB1*06 (*n*, 15), A*24-B*35-C*07-DRB1*15-DQB1*06 (*n*, 14), A*03-B*35-C*12-DRB1*07-DQB1*02 (*n*, 11), A*24-B*35-C*12-DRB1*15-DQB1*06 (*n*, 10) and A*11-B*07-C*07-DRB1*15-DQB1*06 (*n*, 10). Thus, the presence of HLA-DRB1*/DQB1* alleles, two-locus and 5-locus extended haplotypes, are evolutionarily conserved for long periods of time in South India. These observations throw interesting insights in the understanding of mosaic immunogenetic architecture in the context of health and disease in this ancient geo-physical region.

However, AEH represented less than 10 includes 12 numbers in Kallars of Thanjavur, 4 numbers in Piramalai Kallar and 2 in Vanniyar (Table 4). Thus, the number of AEH is more in Kallars of Thanjavur than Piramalai Kallar and Vanniyar: this could be attributed to a number of socio-political and biological attributes. The Kallars of Thajavur were strategically located in a region of rich agricultural practice in the fertile plains of river Cauvery river water-fed area bordered on its eastern side by the Bay of Bengal coastal region. This region is more prone for biological / genetic amalgamation from outsiders, new immigrants into their region through coastal routes and also from plains of all direction (due to its fertile agriculture) that resulted in population admixture and genetic amalgamation. Also, Kallars of Thanjavur are less rigid in terms of their mate selection by having more number of marital clans (1300 clans), whereas the other two populations (Piramalai Kallar and Vanniyar) were relatively more rigid for genetic amalgamation (including miscegenation by socio-biological amalgamation) by way

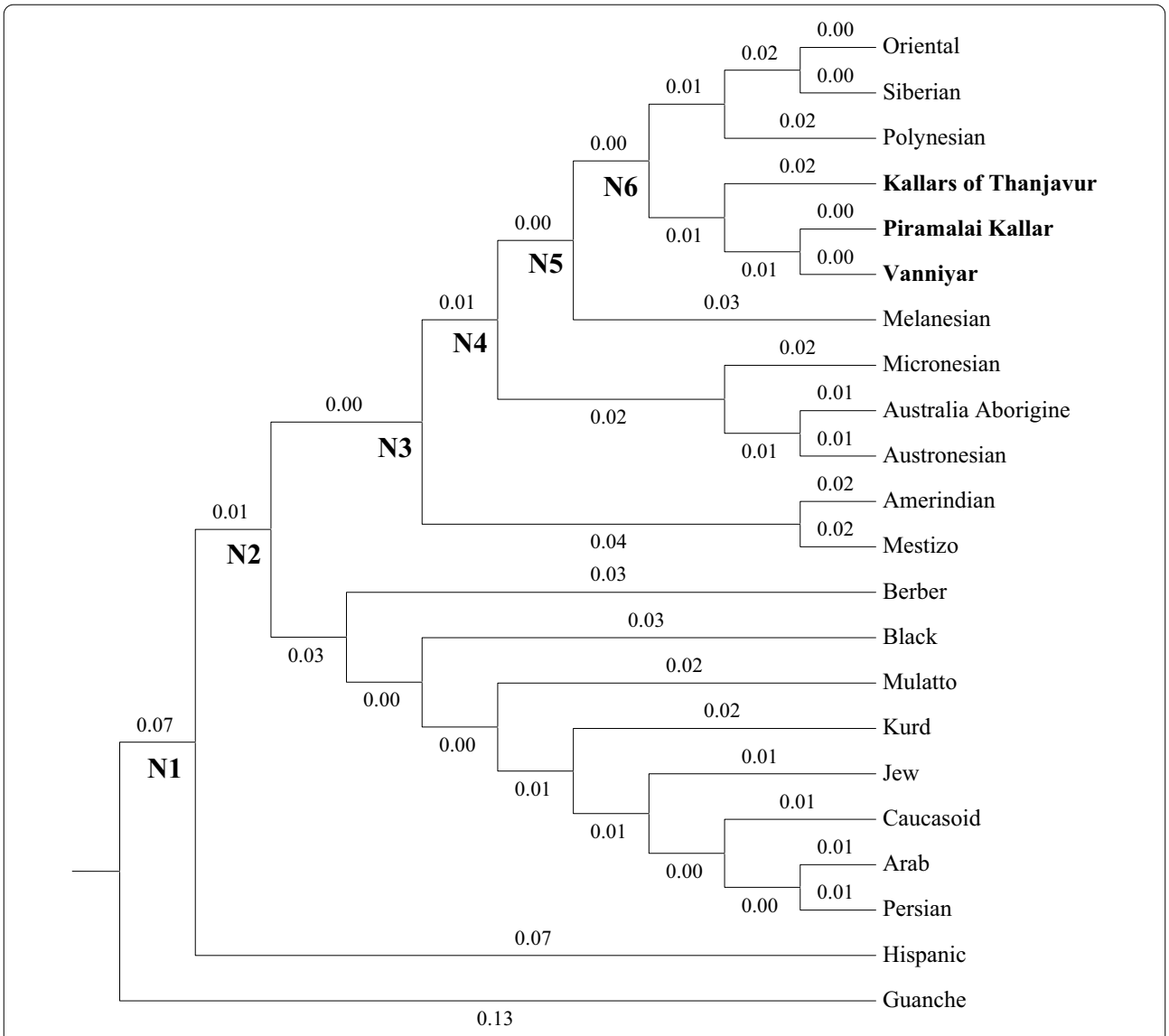

Fig. 1 Neighbor-joining dendrogram showing relatedness between South Indian (Kallars of Thanjavur, Piramalai Kallar and Vanniyar) and other selected global populations. Genetic distances between populations (DA) were calculated using frequencies of HLA-DRB1 alleles. Frequency data for other ethnic populations were from AFND

of mate selection outside their group. Also, the marital clans in these later groups are comparatively less than the former.

The alleles DRB1*01, DRB1*08, DRB1*09, DRB1*11 and DRB1*16, were found in low frequency in all the three study populations. Interestingly, these alleles have been reported with higher frequencies in different ethnic populations. The allele DRB1*01, was highly frequent in Berber (25.0%), Mulatto (25.0%), Caucasoid (23.45%), Persian (10.0%), Jew (10.0%) and Mestizo (9.37). The allele DRB1*08 showed higher frequencies in Polynesian (33.33%), Amerindian (32.81%), Melanesian (30.0%), Mestizo (25.0%), Micronesian (25.0%), Hispanic (18.75%), Siberian (11.36%) and Oriental (11.35) populations. Allele DRB1*09 showed higher frequencies in Oriental (30.85%), Siberian (27.27%) and Polynesian (26.66%) populations. The allele DRB1*11 showed higher frequencies in Kurd (50%), Persian (40.0%), Jew (35.0%), Polynesian (26.66%), Arab (25.61%), Berber (25.0%), Caucasoid (25.0%), Mulatto (25.0%), Black (23.33%), Melanesian (12.0%) and Siberian (11.36%) populations. However, the allele DRB1*16 was highly frequent in Melanesian (28.0%) and Amerindian (19.53%) populations [20].

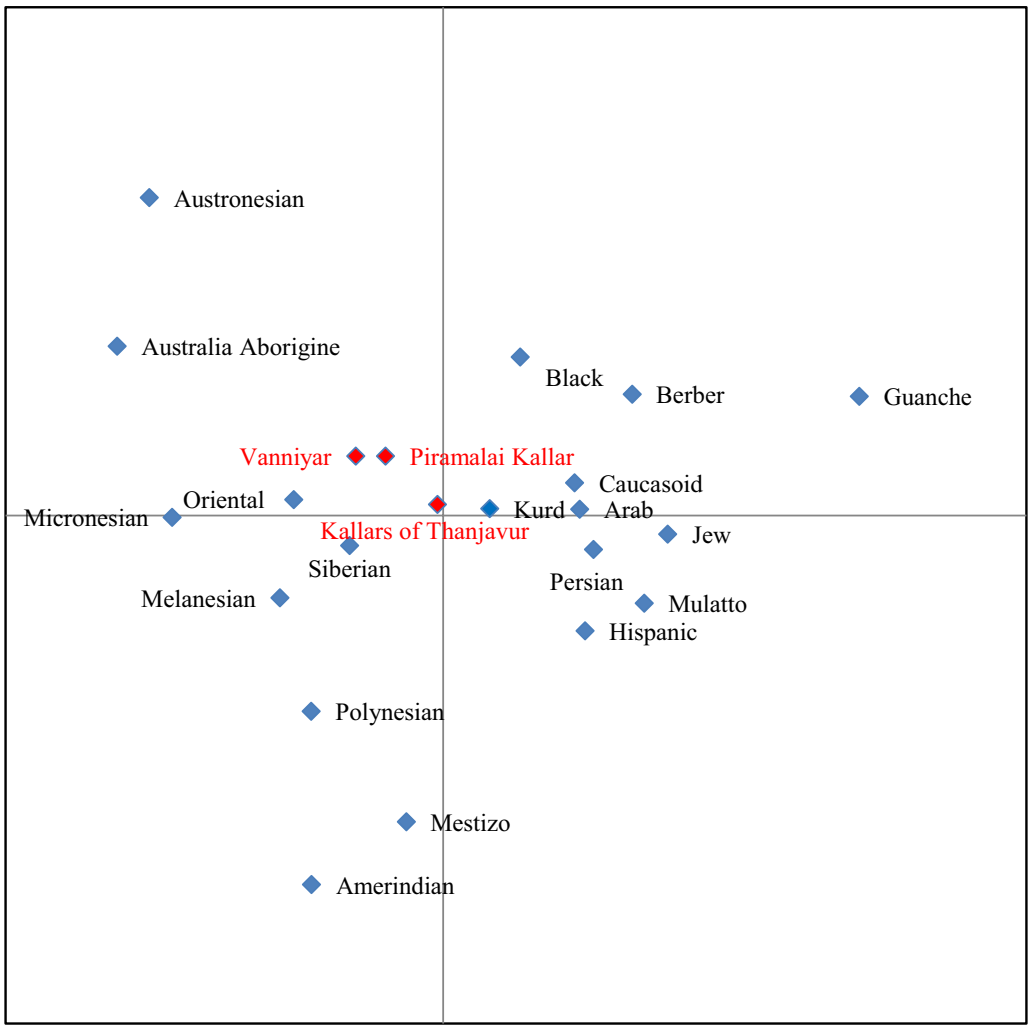

Fig. 2 Principal component analysis (PCA) showing of the relationship between the South Indian and other selected global populations based on frequencies of HLA-DRB1 alleles. Frequency data for other populations were from AFND

We have analyzed and compared the DRB1* allele frequency results of 3 endogamous groups with 19 ethnic populations using neighbor-joining (NJ) method and Principal Component Analysis (PCA) (Figs. 1 and 2). The dendrogram analysis based on HLA-DRB1 profile, showed a unique genetic architecture of the south Indian study populations. These three populations groups have been connected to both N4 (Micronesian, Australian Aborigine and Austronesian) and N5 (Melanesian) nodes in one side, and joined with Polynesian cluster on the other side. The Australian continent holds some of the earliest archaeological evidence for the expansion of modern humans out of Africa, with the initial occupation ~ 40,000 years ago [21]. Austronesian are populations living mainly on islands in the north and east of New Guinea and coastal patches, and also in places as far away as Taiwan, Easter Islands, New Zealand and Madagascar, and in many densely populated areas such as Malaysia, Indonesia and the Philippines [22]. The origins of the Polynesians remain an enigma. Linguistic reconstructions of proto-Austronesian languages suggest a shared origin for Polynesians. Micronesians are from northern Borneo and Sulawesi [23] and Melanesians are from the Papua New Guinean coast (Madang), islands (Rabaul) and highlands (Goroka), and from New Caledonia and Fiji [24]. We recently reported that the phylogenetic affinities of South Indian populations with the South East

Asia, most particularly with Austronesians possibly by forward migration of Indian population into South East Asia [2]. Perhaps, it was established that, the ancestral population from Africa left a genetic trail in Indian sub-continent *en route* Australia, in the distant past.

The Mediterranean populations are divided from N1 node and form a different branch and clustered with Black. We have reported earlier that the Mukkuvar of Tamil Nadu (predominantly the eastern coastal inhabitants of South India) and Hispanic and Guanche populations share a very similar HLA genetic pool not withstanding with the fact of their distant geographic inhabitation [2]. The Guanches confirmed a North African origin and that they were genetically most similar to modern North African Berber people of the nearby North African mainland [25]. Some Guanches were known to be agriculturists and fish hunting communities in Spain. The Mukkuvar of south India were primarily fishers of east coast region of the country. These published evidences corroborate the HLA genetic affinities of Guanche, North African and South Indian Population.

The extended haplotype (EH), A*03-B*35-C*12-DRB1*07-DQB1*02 showed a higher frequency in the Kallars of Thanjavur (HF:0.1041) and Piramalai Kallar (HF:0.1000). Interestingly, this extended haplotype was not reported in any of the world population so far. However, A*24-B*35-C*04- DRB1*15-DQB1*06 haplotype observed in Kallars of Thanjavur (HF:0.1600) and Piramalai Kallar (HF:0.1166), already reported in Asian, Hispanic and Caucasoid populations (Table 4). Similarly, a number of other extended haplotypes observed in the present study with different HLA-A/-B/-C alleles were reported in various ethnic populations: A*11-B*35-C*04-DRB1*07-DQB1*02 in Asian, Hispanic and Caucasoid, A*11-B*35-C*12- DRB1*07-DQB1*02 in Caucasoid, A*24-B*35-C*12-DRB1*15-DQB1*06 in Asian and Hispanic, A*03-B*07-C*07-DRB1*15-DQB1*06 in Asian, Hispanic and Caucasoid and A*03-B*35-C*04-DRB1*15-DQB1*06 Hispanic and Balck. Similarly, A*24-B*35-C*07-DRB1*15-DQB1*06 haplotype was observed in Kallars of Thanjavur (HF-0.1800) (Table 4). Such a complicated patterns of presence of different extended haplotypes in South India, compared to extended haplotypes from global populations throws interesting insights on the population dynamics and turn-over ostensibly influenced by frequent migrations and invasions during prehistoric and historic times.

A substantial amount of research has been conducted on the association of HLA polymorphisms with TB in different populations. HLA-DR2 is most consistently associated with TB in a diverse ethnic populations, including south Indian population [26, 27], Polish [28], Thai [29], Indonesian and Russian [30]. Inter-population variations in HLA-TB associations have been established. DR14 was found at a significantly higher frequency among Iranian TB patients than controls [31]. HLA-DQB1*05:03 was found to influence TB progression in the Cambodian population [32], and DQB1*06:01 was associated with TB susceptibility in the South Indian, Thai and Uganda populations [26, 29, 33]. These alleles are very common in South Indian Population. Such an analysis for other highly frequent HLA alleles and haplotypes of south Indian population is worth exploring further.

The current level of diversity and the variation observed in allelic distributions for different populations could probably result from evolutionary forces that have changed as human populations have encountered new environments in their spread around the globe [34]. Maintenance of high levels of MHC polymorphism is crucial to counteract novel pathogenic challenges and to ensure long-term survival of organisms. Further, HLA alleles found in tropical countries tend to vary a lot from those in temperate parts of the globe, because the pathogens found there are different and highly divergent.

Conclusions

In conclusion, our results show a close relatedness among the study populations and other ethnic group of Oceania such as Polynesian, Melanesian, Micronesian, Australian Aborigine and Austronesian. The Pacific Islands are separate, isolated environments which may have hosted different pathogens from time immemorial and hence mounted differential evolutionary pressure on HLA genes. During the ancient past and at different times, when travel between the islands was not as frequent as it is at present, physical separation and limited inter population contacts may have contributed for the observed differences due to founder effects. But the Mediterranean populations when comparing with other ethnic populations are centrally located and in close vicinity with the Caucasian on one side of Southeast Europe and North African populations on the other side. Although, the Hispanic and Guanche populations were formed a unique ancestral node from the root of the dendrogram, represented that they did not settled in one place, often migrating along the coastal areas of different continents and most probably admixed with other local populations, amalgamated genes from neighboring populations and hence share a common HLA gene pool.

Abbreviations

HLA: Human leukocyte antigen; DNA: Deoxyribonucleic acid; PCR: Polymerase chain reaction; SSP: Sequence specific primers; AFND: Allele frequency net database; MEGA: Molecular and evolutionary genetics analysis; PCA: Principal component analysis; NJ: Neighbor-joining; EH: Extended haplotype; AEH:

Ancestral extended haplotype; NS: Nephritic syndrome; SIPs: South Indian populations; TB: Tuberculosis.

Acknowledgements
We thank all blood donors for their voluntary support. We also acknowledge Mrs. Sheeba Nancy, Research Scholar, Mother Teresa Women's University for statistical support.

Author contributions
Author contributions are as follows. Study design: KR; Sampling: KR, BK, Data Collection: KR, RC; Data Analysis: RK; Manuscript Preparation: KR, BK. All authors read and approved the final manuscript.

Funding
This work was supported by the UGC, Government of India through 'Research Fellowship in Science' for the Meritorious Students Scheme (R. KAMARAJ/F.No.25-1/2014-15 (BSR)/5-67/2007) and Tissue Typing Services, Madurai Kamaraj University (Prof. K. Balakrishnan).

Availability of data and materials
The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
All procedures performed in the study are consistent with the ethical standards of the institutional ethics committee. Institutional ethical clearance was obtained from Madurai Kamaraj University (Ethical and Review Board Committee).

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Department of Biotechnology, PRIST Deemed to be University, Thanjavur, Tamil Nadu 613403, India. [2]Department of Immunology, School of Biological Sciences, Madurai Kamaraj University (University with Potential for Excellence), Tamil Nadu 625021 Madurai, India.

Received: 17 August 2022 Accepted: 24 November 2022
Published online: 21 December 2022

References

1. Wells RS, Yuldasheva N, Ruzibakiev R, Underhill PA, Evseeva I, Blue-Smith J, Jin L, Su B, Pitchappan R, Shanmugalakshmi S, Balakrishnan K, Read M, Pearson NM, Zerjal T, Webster MT, Zholoshvili I, Jamarjashvili E, Gambarov S, Nikbin B, Dostiev A, Aknazarov O, Zalloua P, Tsoy I, Kitaev M, Mirrakhimov M, Chariev A, Bodmer WF (2001) The Eurasian heartland: a continental perspective on Y-chromosome diversity. Proc Natl Acad Sci USA 98(18):10244–10249. https://doi.org/10.1073/pnas.171305098
2. Kamaraj R, Balakrishnan K, Dhivakar M, Rathika C, Albert C, Justinkumar P, Navaneethakannan K (2020) Distribution of HLA alleles and haplotypes in Tamil-speaking South Indian Populations: affinities with Spanish and Austronesian. Russ J Genet 56(9):1139–1150. https://doi.org/10.1134/S1022795420090100
3. Marsh SG (1997) Nomenclature for factors of the HLA system, update 1997. Eur J Immunogenet 24:239–240
4. Mehra NK (1998) Genetic diversity of HLA polymorphism and new genes. Ind J Hum Genet 4:1–12
5. Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleate cells. Nucleic Acids Res 16:1215. https://doi.org/10.1093/nar/16.3.1215
6. Malhotra KC, Balakrishnan V, Karve I (1981) Anthropometric variation in Tamil Nadu. The Indian Society of human genetics and Calcutta. The Indian Anthropological Society 50–74
7. Francis W (1902) Census of India. Part I. Govt Press, Madras, p 158
8. Turnbull T (1895) "Account of Cullaries," Prefaced to B.S. Ward, Memoir of Madura and Dindigul. Madurai Collectorate 3(6)
9. Nelson JH (1868) The Madura country a manual. The Asylum Press, Madras
10. Shanmugalakshmi S, Balakrishnan K, Manoharan K, Pitchappan RM (2003) HLA DRB1*, DQB1* in Piramalai Kallars and Yadhavas, two Dravidian speaking castes of Tamil Nadu, South India. Tissue Antigens 61:451–464. https://doi.org/10.1034/j.1399-0039.2003.00061.x
11. Thurston E (1896) Anthropology of the Todas and Kotas of the Nilgiri Hills: and of the Brahmans, Kammalans, Pallis, and Pariahs of Madras City
12. Hamilton W (1820) A geographical, statistical, and historical description of Hindostan and the adjacent countries
13. Stuart HA (1893) Census of India 1891 Volume Xiii Madras. The report of the census
14. Olerup O, Zetterquist H (1992) HLA-DR typing by PCR amplification with sequence-specific primers (PCR-SSP) in 2 hours: an alternative to serological DR typing in clinical practice including donor matching in cadaveric transplantation. Tissue Antigens 39:225–235. https://doi.org/10.1111/j.1399-0039.1992.tb01940.x
15. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729. https://doi.org/10.1093/molbev/mst197
16. Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28:2537–2539. https://doi.org/10.1093/bioinformatics/bts460
17. Court MH, Michael H (2012) Court's (2005–2008) online calculator. Tuft University Website.
18. Majumdar DN (1961) Races and cultures of India. Asia Publishing House, Bombay
19. Thurston EB (1909) Castes and Tribes of Southern India. 1:1–98
20. Allele Frequency Net Database (AFND). http://www.allelefrequencies.net/
21. Pugach I, Delfin F, Gunnarsdóttir E, Kayser M, Stoneking M (2013) Genome-wide data substantiate Holocene gene flow from India to Australia. Proc Natl Acad Sci USA 110(5):1803–1808. https://doi.org/10.1073/pnas.1211927110
22. Tryon D (1995) Proto-Austronesian and the major Austronesian subgroups. In: Bellwood P, Fox JJ, Tryon D (eds) The Austronesians: historical and comparative perspectives. National Library of Australia, Canberra, p 17. https://doi.org/10.22459/A.09.2006.02
23. Gao X, Zimmet P, Serjeantson SW (1992) HLA-DR, DQ sequence polymorphisms in Polynesians, Micronesians, and Javanese. Hum Immunol 34(3):153–161. https://doi.org/10.1016/0198-8859(92)90107-X
24. Gao X, Bhatia K, Trent RJ, Serjeantson SW (1992) HLA-DR, DQ nucleotide sequence polymorphisms in five Melanesian populations. Tissue Antigens 40(1):31–37. https://doi.org/10.1111/j.1399-0039.1992.tb01954.x
25. Rodriguez-Varela R, Günther T, Krzewińska M, Stora J, Gillingwater TH, MacCallum M, Arsuaga JL, Dobney K, Valdiosera C, Jakobsson M, Gotherstrom A, Girdland-Flink L (2018) Genomic analyses of pre-European conquest human remains from the Canary Islands reveal close affinity to modern North Africans. Curr Biol 28(10):1677–1679. https://doi.org/10.1016/j.cub.2018.04.083
26. Ravikumar M, Dheenadhayalan V, Rajaram K et al (1999) Association of HLA DRB1, DQB1 and DPB1 alleles with pulmonary tuberculosis in South India. Tuber Lung Dis 79(5):309–317. https://doi.org/10.1054/tuld.1999.0213
27. Shanmugalakshmi S, Pitchappan RM (2002) Genetic basis of tuberculosis susceptibility in India. Indian J Pediatr 69:25–28
28. Dubaniewicz A (2000) HLA-DR antigens in patients with pulmonary tuberculosis in northern Poland. Preliminary report. Arch Immunol Ther Exp (Warsz) 48(1):47–50
29. Vejbaesya S, Chierakul N, Luangtrakool K, Srinak D, Stephens HAF (2002) Associations of HLA class II alleles with pulmonary tuberculosis in Thais. Eur J Immunogenet 29(5):431–434. https://doi.org/10.1046/j.1365-2370.2002.00352.x

30. Khomenko AG, Litvinov VI, Chukanovo VP, Pospelov LE (1990) Tuberculosis in patients with various HLA phenotypes. Tubercle 71:187–192. https://doi.org/10.1016/0041-3879(90)90074-i
31. Mahmoudzadeh-Niknam H, Khalili G, Fadavi P (2003) Allelic distribution of human leukocyte antigen in Iranian patients with pulmonary tuberculosis. Hum Immunol 64(1):124–129. https://doi.org/10.1016/s0198-8859(02)00703-6
32. Goldfeld AE, Delgado JC, Thim S, Bozon MV, Uglialoro AM, Turbay D, Cohen C, Yunis EJ (1998) Association of an HLA-DQ allele with clinical tuberculosis. JAMA 279(3):226–228. https://doi.org/10.1001/jama.279.3.226
33. Wamala D, Buteme HK, Kirimunda S, Kallenius G, Joloba M (2016) Association between human leukocyte antigen class II and pulmonary tuberculosis due to mycobacterium tuberculosis in Uganda. BMC Infect Dis 16:23. https://doi.org/10.1186/s12879-016-1346-0
34. Fernandez Vina MA, Hollenbach JA, Lyke KE, Sztein MB, Maiers M, Klitz W, Cano P, Mack S, Single R, Brautbar C, Israel S, Raimondi E, Khoriaty E, Inati A, Andreani M, Testi M, Moraes ME, Thomson G, Stastny P, Cao K (2012) Tracking human migrations by the analysis of the distribution of HLA alleles, lineages and haplotypes in closed and open populations. Philos Trans R Soc Lond B Biol Sci 367(1590):820–829. https://doi.org/10.1098/rstb.2011.0320

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen journal and benefit from:

- ► Convenient online submission
- ► Rigorous peer review
- ► Open access: articles freely available online
- ► High visibility within the field
- ► Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com