© Kamla-Raj 2008 Int J Hum Genet, 8(1-2): 143-153 (2008)

HLA Genomic Diversity of India and its Implications in HIV Pandemic

R M. Pitchappan, V. J. Kavitha and M. Jayalakshmi

Department of Immunology, School of Biological Sciences, Centre for Excellence in Genomic Sciences, Madurai Kamaraj University, Madurai 625021, Tamil Nadu, India

KEYWORDS HLA-B57; HIV non progression; genomic diversity; India

ABSTRACT Both culturally and biologically, India is one of the mega diversity countries of the world. This pertains to Human Genome Diversity as well. Although little explored so far, HLA is no exception to this. HLA profile varies based on regional, linguistic and caste profiles. HLA alleles identified as common in HIV non progressors in other parts of the world are the commonest in the Vellala related communities of southern India. Migration and selection might be responsible for this. In the context of the well established MHC restriction phenomenon and the role of CTL in controlling HIV viral replication, it has become essential to revisit these areas and better understand the underlying phenomena so that better experiments are designed in future research. The article highlights these aspects in the context of HIV pandemic in India.

INTRODUCTION

India is one of the megadiversity countries of the world with unique fauna and flora. Similarly India is also a mega diversity country in terms of Human species as well. This can be attributed to the early settlement of Man in India, since ~ 50,000 years, through coastal (southern route) and land routes. From hunter gatherer he started leading a settled life as a result of 'social' hunting and organized agriculture. Man thus settled in more ideal ecological niches such as Merghargh of present day Afghanistan, Swat Pirak culture in northern India, IndoGangetic doab riverine culture and Kaveri Delta and Athichanallur in Tamil Nadu (Kochhar 2000; Wells et al. 2001). The ecological factors such as climate, rainfall temperature and monsoon (seasonal cycles) were then ideal for human survival in India. Hence having settled in a particular place He expanded in population size, learnt better water management and irrigation techniques which transformed Him into a social creature. These populations in earlier settlements must have been subjected to the prevailing or brought in infectious agents such as virus and bacteria that might have acted as bottle necks and/or selection pressures in terms of immunology. Two factors thus might have effectively contributed to the shaping of the genomic diversity of the Major Histocompatibility (MHC) system more particularly the Human Leucocyte Antigen system A (HLA): i) place of origin and/or migration and ii) the prevailing infectious diseases enroute to and in the place of settlement. This probably resulted in the prevailing scenario of HLA distribution in India and the world.

In this article, we describe the genomic diversity of the most important immunogenetic system of our body, viz. HLA that is responsible for the recognition of self from non-self and most importantly in protecting all of us from particular infections. We present here the basic premise on which the immunogenetics has been built on the past four decades. It is the most extensively studied genetic polymorphisms of our system. Any clinical or population based study as on date invoking a immunological/genetic basis for a disease or disorder needs to essentially consider the host gene pool particularly the HLA/MHC for drawing better conclusions.

I. MHC AND HLA

Major histocompatibility complex is present in all the Chordates studied so far, with the increasing order of complexity of structure, function and polymorphism through vertebrate evolution (HLA is a part of the major histocompatibility system of humans).

Address for Correspondence: Prof. RM. Pitchappan, Professor Emeritus, Department of Immunology, School of Biological Sciecnes, Madurai Kamaraj University, Madurai 625021, Tamil Nadu, India
Telephone: (+91 452) 2458418, *Fax:* 2450181;
E-mail: pitchappanrm@yahoo.co.uk

Discovery of MHC: Genetic basis of immune responses have now been amply confirmed. The pioneering work of Jean Dausset describing the white cell agglutinins during the Second World War has lead to the discovery of the MHC system. The recognition for this was the Nobel Prize in Physiology and Medicine awarded to him in 1980, jointly with George D. Snell, discoverer of the Mouse strains, and Baruj Benaceraaf for his Ir gene complementation. Having received the Nobel Prize although many were in queue to fund his research the earliest research grant of this great scientist was from Lady Tata Trust (personal conversation with RMP in 1981).

Rolf Zinkernagel and Peter Doherty described the MHC restriction phenomenon in 70's the first evidence for a probable role of MHC that turned out to be true till today, for which they received the Nobel Prize in the year 1996. This seminal discovery led to the newer understanding of the immune system and functions, biophysical basis of peptide binding and thymic education, explaining the molecular basis of generation of diversity in immune repertoire (cf Roitt et al. 2006). X-ray crystallography and better understanding of the structure of the MHC molecules and its peptide binding grooves provided the convincing evidence for the bio-physical basis of the trigger that initiates an immune response (Madden et al. 1989).

Of late, the HIV/AIDS pandemic and the search for its HLA association have brought out the direct clinical evidence of this MHC restriction phenomenon in Cytotoxic T Lymphocyte (CTL) generation and its role in disease control.

MHC Genetics: Major Histocompatibility Complex encompasses a region of 3500 kb, in the short arm of chromosome 6 p21.3. There are 220 loci and majority of them are involved in immune response (Bhattacharya et al. 2007). Two of these clusters called Klein class I and Klein class II originated by gene duplications and each one of them consists of a handful of closely (functionally and structurally) related loci (Table 1). Klein performed his studies in 1988 and Klein Class I and II are more popularly known as HLA class I and II loci. Classical examples of Klein class I are HLA-A, B and C loci and that of Class II are HLA-DR, DP and DQ, earlier all defined serologically by employing Human allo-antisera. These alleles are inherited in a Mendelian fashion and are co-dominant. With the advent of DNA technologies, the alleles are defined at sequence level. The allelic polymorphism is the highest in HLA B locus (894 alleles) and HLA DRB1 locus (577 alleles) as on May 2007 (Table 1). Klein class I alleles codes for the heavy chain of the molecules and a dimer formed with beta 2 microglobulin are expressed on all the cells and tissues of our body ('Identity Card', called Tissue or Transplantation antigens). Klein class II expression is specific to the cells of the immune system and further their expression is restricted to B lymphocytes and activated T cells of the immune system.

Table 1: Major HLA Loci, their Alleles and Proteins involved in antigen presentation and immune responses, May 2007

(cf: IMGT/HLA data base, http://www.ebi.ac.uk/imgt/hla/stats.html

Numbers of HLA Alleles			
HLA Class I Alleles			1,839
HLA Class II Alleles			875
HLA Alleles			2,714
Other non-HLA Alleles			102
HLA Class I			
Gene	A	B	C
Alleles	545	894	307
Proteins	436	766	244
Nulls	39	31	7
HLA Class II			
Gene	DRB	DQB1	DPB1
Alleles	577	83	126
Proteins	476	61	113
Nulls	7	1	2
HLA Class II - DRB Alleles			
Gene			DRB1
Alleles			494
Proteins			418
Nulls			2

The serological specificities of yesteryears are designated as HLA-A2, HLA-B57, HLA-DR15 etc. Their exact alleles defined at the sequence level are designated as HLA-A*0201, HLA-B*5701 and HLA-DRB1*150101, respectively. Each serological specificity, identified by an antiserum, is now described at the exact sequence level and is comprised of many alleles. For example, HLA-B35 has 85 alleles as on date, which are characterized by exact sequence differences in the alpha helixes and Beta sheet of peptide binding groove of HLA molecule.

MHC Peptide Binding Groove: The peptide binding grooves of these HLA molecules hold the key for the specificity of recognition in the immune system. Only if there is a good fit of this peptide binding groove with that of the MHC-peptide complex (neoantigen hypothesis), the peptide will be presented and antigen recognition

by the T cell will take place. Class I molecules bind peptides of 9 amino acid length and class II, 9-12 amino acid long ones. Furthermore, not all amino acids are important, only specific pockets of the peptide binding groove and given residues of the peptide are important in this binding and these are called 'Consensus Motifs' (cf. Roitt et al. 2006). There is no guarantee that all the peptides generated by Antigen Presenting Cells (APC) of an individual will bind and be presented by all the HLA molecules and only those with good fit (consensus motifs) will be presented and may be recognized by T cells, thus providing scope for the enormous amount of diversity and immune response that can be created in the process.

II. HLA IN THE POPULATIONS OF INDIA

Allelic Polymorphisms: Table 2-3 presents the phenotype frequencies (at the serological level) of various populations from southern India studied in our laboratory (www.geocities.com/rdbgy; Pitchappan 2002). The Indian populations have not been extensively studied for HLA polymorphisms and there are only a few sporadic literatures on exact allele frequencies as on date.

Table 4-5 presents the HLA DRB1, HLA-A, B allelic and selected haplotype frequencies reported in various Indian studies . A few important observations can be made from this table: i) a gradient of an allele frequency from east to west, ii) higher frequency of a given allele/haplotype in a geographical region/ caste and iii) one daughter allele of a given parent allele occurring in one geographical region while the other in another geographical region. For example, HLA-DRB1*16, one split (daughter allele) of HLA DR2 specificity being prevalent in Eastern Europe while the other split, (HLA DRB1*15) is common in India and its east. Similarly HLA-DRB1*11 and *12, splits of DR5, and DRB1*13 and DRB1*14, splits of DR6 are prevalent in two different geographical areas (Shanmugalakshmi et al 2003). Simplest explanation for this kind of pattern of two different daughter alleles of a parent allele seen in two different geographical locations/ isolates could be the common origin of the daughter alleles in the ancient settlement and their subsequent dispersal in two different directions due to migration. The picture is more clear with DRB1* and Haplotypes (Table 6). The picture with HLA A and B alleles in Indian populations are the tip of the iceberg, though these estimates suffer because of low sample numbers in many studies. Thus one cannot be sure of the contributions by migration and bottle neck effects leading to the present scenario: one may have to examine these issues in the context of other genomic diversities, with better study designs and sample size.

HLA Profile of Indian Population and Its Implications

From our studies of HLA and from those available in literature, we found differences between different regions and populations of India (Pitchappan et al. 1997; cf. Shanmugalakshmi et al. 2003) as many report (Gosh 2007). Each south Indian caste group studied has their own HLA profile and 'more related' caste groups have more similar profiles (Pitchappan et al. 1988). This picture has been correlated to many diseases and immunological phenomena as indicated in the publications from our laboratory (Pitchappan et al. 1988, 1989, 2002). This can be further exemplified by correlating this profile to their migratory backgrounds and genetic epidemiology.

An ancient isolated population Piramalai Kallars numbering about 0.3 million, living in two provinces (taluks) west of Madurai have the highest frequency of high risk allele for TB, HLA-DRB1*1501; almost every alternate individual of this population possesses HLA DRB1 *1501 but TB was not common among them, probably for want of close contacts and the resultant transmission from the affected individuals and due to the lack of conducive environment for manifesting the disease (Pitchappan et al., unpublished results). The villages are situated in highly arid, healthy atmosphere amidst paddy fields with lot of fresh air. Every alternate individual from this caste also possesses M20 NRY chromosome marker (Wells et al. 2001). This marker presumably originated 30,000 years ago in India representing one of the earlier migrations into India (www.nationalgeographic.com/genographic). Thus whether the higher frequency of DRB1*1501 and M20 are due to founder effect, bottle neck or selection need to be further explored.

Sourashtrians of Madurai who are thought to have originated in Eurasian Steppes representing Indo Iranian nomadic culture, and responsible for Indo-European languages possess NRY-M17

Table 2: HLA A frequencies in selected Indian populations (castes or tribes).

HLA A Allele	*West Parsis*	*West Maratha*	*Pawra*	*Bhils*	*Golla*	*Namboothiri*	*Syrian Christian*	*Nair*	*Malabar muslim*	*Ezhavas*	*Pulaya*	*Iyer*	*Sourashtra*	*Nadars*	*Nadar**	*Kallars**	*Irula**	*Naidu**	*Malayali**
References	a	a	a	a	g	a	a	a	a	a	a	e	f	d	c	c	b	c	b
Language	IE	IE	IE	IE	DR	DR	DR	DR	DR	DR	DR	DR	DR	DR	DR	DR	DR	DR	DR
States of India	MAH	MAH	MAH	MAH	AP	KE	KE	KE	KE	KE	KE	TN	TN	TN	TN	TN	TN	TN	TN
Sample size$	50	91	50		111	40	41	41	34	24	16	74	76	61	101	36	191	57	42
A*1	9	12.3	6		6	6.3	4.8	8.5	5.9	10.4		17	11.2	9	15.6	15	6.8	11.1	4.9
A*2		15.4	21	22	15.3	15	12.9	19.5	19.1	25	6.3	7	7	10.7	9.3	8.7	10.8	20.5	18.4
A*3	3	9.3		9	7.8	3.8	11.3	14.6	8.8	8.3	6.3	7	3.9	20.5	12.1	4.3	5.1	4.5	1.2
A*9	11	22.9	22	15	17.8	25	27.4	24.4	17.7	12.5	25	16.3	13.4	15.6	13.8	21.8	22.1	19.4	19.8
A*23								1.2		2.1		0.7							
A*24	11	22.9	22	15	17.3	25	27.4	23.2	17.7	10.4	25	17	13.4	15.6					
A*10	5			4	3.1	2.5	1.6	2.4	5.9	6.3		4.8	7	1.6	4	11.8	13.5	0.9	3.6
A*26	4			4	3.1	2.5	1.6	2.4	5.9	6.3		3.5	3.9	1.6					
A*11	10	12.3	21		8.7	7.5	11.3	8.5	14.7	10.4	53.1	9.2	21.6	9.8	18.6	16.7	18.2	18.4	22.8
A*28	3	7.5	12	7	4.4	11.3	3.2	4.9	2.9	4.2		2.7	7	7.4	6.7	2.8	2.9	8.2	8.7
A*68	3	7.5	12	7	4.4	11.3	3.2	4.9	2.9	4.2		2.7	1	7.4					
A*19	57	17.9	18	20	19.7	28.9	27.4	17	22.1	22.9	9.3	27.1	30.7	25.4	9.9	5.7	17	3.6	
A*29	4							2.4				0.7	1				2.7		
A*30	16		6	1	1.1	2.5	1.6	1.2	1.5			2.7	2.9						
A*31	15	4.3	6		3.9	1.3	4.8	6.1	2.9	8.3	3	5.6	3.9	18.9			9		12.7
A*32	5	0.6	3	1	2.6	6.3	6.5	1.2	1.5			3.5	1				4		
A*33	17	13	3	18	11.7	18.8	14.5	6.1	16.2	14.6	6.3	16.2	17.8	6.6			1.3		2.4
Total	97	92.6	88	74	78.4	89	96.7	94.9	94.2	95.8	100	89	101	92.6	90	86.8	96.4	86.6	79.4

The maximum observed frequency of a given allele and that are >50% of this, are highlighted.
* = many studies did not investigate the split alleles; $ = Sample size were too small in many populations. Alleles represented in one or two populaitons only that too at lower frequencies were omitted in the table. Blank - not studied or zero frequency. Abb: IE = Indo European, DR = Dravidian, ST = Sino Tibetean, MAH = Maharashtra, AP = Andhra Pradesh, KE = Kerala, TN = Tamil Nadu

References Cited a - "<http://www.allelefrequencies.net>"; b - Pitchappan et al. 1997; c - Rajasekar et al 1987; d - Shankarkumar et al. 2004; e - Balakrishnan et al 1996; f - Pitchappan et al. 1986; g - Crawford et al 2001.

Table 3: HLA-B allele frequencies in selected Indian populations (castes / tribes)

HLA B Allele	*Parsis*	*Marathas*	*Bhils*	*Pawar*	*Golla*	*Adiya*	*Paniya*	*Kanikkar*	*Kurumba*	*Kurichiya*	*Ezhavas*	*Pulaya*	*Malapanadram*	*Random non-tribal Dravidian group*	*Nair*	*Syrian Christian*	*Malabar Muslim*	*Namboothiri*	*Kattunaikka*	*Irula*	*Nadar*	*Kallars**	*Nadar**	*Iyer*	*Sourashtra*	*Naidu**	*Malayali**
Reference	b	b	b	b	h	a	a	a	a	a	b	b	a	a	b	b	b	b	a	c	e	d	d	f	g	d	c
Language	IE	IE	IE	IE	DR	DR	DR	DR	DR	DR	DR	DR	DR	DR	DR	DR	DR	DR	DR	DR	DR	DR	DR	DR	DR	DR	DR
State	MAH	MAH	MAH	MAH	AP	KE	KE	KE	KE	KE	KE	KE	KE	KE	KE	KE	KE	KE	KE	TN	TN	TN	TN	TN	TN	TN	TN
Sample size$	50	91	50	50	111	21	10	22	15	10	24	16	10	78	41	41	34	40	17	191	61	36	101	74	52	57	42
B*5	14	12.4	21	1	28.2	33.3	50	18.1	13.3	15	14.6	3.1	10	15.3	17.1	16.1	11.7	8.8	26.5	18.1	15.6	15	20.4	17.8	15.7		11.4
B*51	7	6.8	12	1	14.3	11.9	15	13.6	10	15	14.6	3.1	5	10.2	9.8	14.5	8.8	5		9	15.6			9.2	2.9		
B*52	5	5.6			13.4	21.4	35	4.5	3.3				5	5.1		1.6	2.9	3.8	26.5	5.1				6.3	10.1		
B*53	2		9		0.6										7.3					4				1.4			
B*07	1	13.5			5.1	4.8	5	9.1	16.7	35	1.5	28.1		19.2	24.4	11.3	25	28.8		1.6	1.6	11.8	11	10	6		11.4
B*08	1	1.9	8		6.3	2.4				10				1.3	2.4	1.6	4.4	2.5	2.9	14.1		8.7	1	3.4	2.9	5.2	2.4
B*12	12	14.8	3	4	1.7	2.4								10.9	8.5	4.8		1.3		0.8	7.4	4.3	6.1	9.2	7		3.6
B*13	19	4.3	4	12	3.8	2.4		4.5	3.3		2.1			3.8	6.1		1.5	1.3	2.9	5.1	1.6	4.3	0.5	4.1	4.9		7.4
B*14	26													0.6		1.6				0.3							
B*15	3	5.6	5	15	7.5	2.4		22.7		5	10.5	18.8		4.3	12.2	8	4.5	2.6	5.9	1.1	20.5	2.8	15	9.2	12.3	10.8	11.4
B*16			4		0.9	2.4					6.3	3.1			2.4			1.3						0.7	1		
B*17	3	7.4	9	15	9.7			4.5	36.6	10	8.3		45	11.5	4.8	13	10.3	8.8	5.9	8.8	8.2	10.2	10.4	8.5	9.1		11.4
B*57					2.7				3.3	5			20	6.4	2.4	6.5	8.8	1.3			8.2				4.9		
B 58	3	7.4	9	15	7			4.5	33.3	5	8.3		25	5.1	2.4	6.5	1.5	7.5	5.9						3.9		
B*18		1.9	5	11	1.6	2.4		2.3						2.6		3.2		6.3	5.9								
B*21					1			4.5			2.1			0.6	1.2			6.3				5.7	1	2.1			
B*22	2	2.5			3.4					10		6.2		1.9	4.8	3.2	3	2.5		3.2	5.7	5.7	3	5.6	1		15.5
B*27		3.7	2	5										1.9							4.1		1	0.7	1.9		
B*35	12	10.5	14	20	9.8	4.8		4.5	10		12.5	9.4		8.3	6.1	17.7	23.5	16.3	32.4	9.9	17.3	2.8	4	13	16.8	32.4	1
B*37		2.4	1		7.3	9.5				5	4.2					1.6				1.3	9			3.4	2.9		
B*40	7	17.9	21	17	12.1	33.4	45	29.5	20	10	25	25	45	16	9.8	17.7	10.3	20.1	17.6	18.5	9	8.7	7.7	7	13.4	26.2	15.5
	100	99	100	100	100	100	100	100	100	100	87	94	100	100	100	100	94	107	100	83	100	80	81	96	95	75	91

The maximum observed frequency of a given allele and that are >50% of this, are highlighted.
References cited a - Thomas et al 2004; b - "<http://www.allelefrequencies.net>" ; c - Pitchappan et al. 1997; d - Rajasekar et al 1987; e - Shankar kumar et al. 2004; f - Balakrishnan et al 1996; g - Pitchappan et al. 1986; h - Crawford et al 2001.
Other legends as per Table 2

Table 4: HLA DRB1* Allele frequencies (x100) of selected southern Indian populations (castes & tribes)

HLA DRB1 Allele*	*Asian Indians*	*South asia*	*Golla*	*Malayali**	*Iyer*	*Yadhavas*	*Vanniyar*	*Piramalai kallars*	*Irula**
References	a	b	c	d	e	g	f	g	d
Language			DR	DR	DR	DR	DR	DR	DR
States of India			AP	TN	TN	TN	TN	TN	TN
Sample size$	*59*	*196*	*111*	*42*	*74*	*233*	*132*	*202*	*191*
01		2.6		7.4	8.8	3.3	2.9	3.8	9.3
10		4.6	14.7	3.6	5.4	8.2	7.0	11.9	14.4
03	22.9	4.1		4.9	6.1	14.3	13.8	11.6	2.2
DR5	13.6	9.7	6.4	16.4	10.2	6.1	3.2	4.0	4.9
11 (5)		8.2	4.1		7.5	4.3	2.9	0.9	
12 (5)		1.5	2.3		2.7	1.8	0.3	3.0	
08	1.7	1.5	2.3	4.9	8.1	5.7	5.1	4.7	22.6
DR6	9.8	19.4	23.0	10.0	4.1			8.3	7.2
13 (6)		9.2	15.2		2.7	4.1	2.2	5.7	
14 (6)		10.2	7.8		1.4	0.8	5.4	2.6	
04	7.0	4.6	11.5	12.7	16.9	18.5	7.9	10.8	6.6
07	17.6	17.3	6.9	12.7	15.6	13.7	12.4	8.2	9.8
DR2	25.1	31.6	15.8	21.3	13.5	12.9	13.6	26.0	15.3
1501 (2)		29.6	7.0		10.8	6.5	9.5	22.5	
1502 (2)			8.8			6.4	4.1	3.5	
16 (2)		2.0			2.7				
09		0.5	1.8	6.1	4.8	1.6	1.9	0.5	6.1
Total	96.0	94.4	80.1	95.1	85.2	78.6	62.9	84.9	75.8

The maximum observed frequency of a given allele and that are >50% of this, are highlighted. DR5, DR6, DR2 are the serological specificites, obtained adding the split alleles (with brackets)
References cited: a - Suciu-Foca et al. 1981; b - Carrington et al. 2002; c - "<http://www.allelefrequencies.net>"; d - Pitchappan et al. 1997; e - Balakrishnan et al. 1996; f - Ravikumar et al. 1999; g - Shanmugalakshmi et al. 2003. Other legends as per table 2

and this allele is the commonest in Central Asia and common in northern parts of India as well. Many of the Indian populations possess HLA-B17 and HLA A1-B17 haplotypes (Table 5): this allele and haplotype were the commonest in India and called as "Telugu haplotype" by Hammond in Indians settled at Durban, South Africa (Hammond et al. 1979). This allele is distributed through out India at a lower but consistant frequency of 0.05 to 0.1. The difference in the HLA frequencies and NRY of the two populations mentioned above suggest the possibility of different origin of these populations, their expansion in varied ancient settlements and probable differential susceptibility to infection in these settlements. Molecular epidemiology may throw further light on these aspects.

III. IMPLICATIONS

Recognition of MHC Peptide Complex: The Zinkernagel and Doherty phenomenon mentioned earlier revolves around the MHC (HLA)-peptide binding: a CTL generated in response to a given MHC-peptide complex knows to kill the homologous MHC peptide complex bearing cells only i.e., the cells of the same host genetic H2 (similar to HLA in mouse) make up, infected by the same virus. The CTL generated under one MHC background does not know to identify any other antigens or MHC background (Doherty 1997)

This is what exactly happens in every viral infection including HIV infection. Nature has invented CTL to kill only the homologous virus infected cells in our body and not other uninfected cells and tissues in the vicinity or any other virus/ epitope (the 9 or 12 amino acid antigen determinant generated by the APC, bound and presented by the MHC and recognized by the T cell/ B cell of the immune system).

HLA, CTL and Long Term Non Progressors: The importance of CTL in keeping at bay the HIV infection has been shown in many HLA association studies. Table 6 provides a list of recent studies wherein the HLA is reported to be

Table 5: HLA two locus haplotypes identified in various studies on Indian populations and nearby regions (per 1000)

	West India												*North India*			*South India*										
	Africa	*Europe*	*Jewish*	*Iranian*	*Pakistan*	*W. India*	*India*	*Hindu*	*Muslim*	*W. India*	*Marathas*	*Punjab*	*N India*	*N. India*	*N India*	*S.India*	*Telugu*	*Golla*	*Sourasthra*	*Iyers*	*Nadars*	*Kallars*	*Naidus*	*Gudiyatham*	*Tamil Nadu*	*Tamils*
Reference	a	b	c	d	e	i	f	o	o	b	k	g	l	j	d	d	m	a	n	p	p	p	p	r	o	m
Language	IE	IE	IE	IE	IE	IE	IE	IE	IE	IE	IE	IE	IE	IE	IE	IE	IE	IE	IE	IE	DR	DR	DR	DR	DR	DR
State											Mah						AP	AP	TN	TN	TN	TN	TN	TN	TN	TN
Sample size$			130	100	60		138	133	60		91	150	400	153	156	344	122	111	52	54	101	36	57	62	385	288
A1-B8			76	31																						
A10(26)-B8		**17**			**87**		**22**			**18**		**57**	**21**	**13**	**19**						**5**	**71**			**12**	
A2-B40		**16**	**51**	**40**		**118**		**5**	**45**	**55**					**61**	**31**	**41**			**8**	**27**	**0**	**25**	**32**	**28**	**37**
A1-B17		**16**	**28**		**41**		**81**	**27**	**28**	**27**	**39**		**12**	**29**		**17**	**96**		**28**	**86**	**25**	**34**	**32**	**28**	**40**	62
A3-B7			34										18					59		**14**	**28**	**27**	**7**	**18**		
A33-B44							**39**	**26**	7						**27**	**39**	4		48	19			**9**	**15**	**4**	**13**
A1-B37							**29**	7			16		12			**27**	16							13		**20**
B37-DR10																**29**			12							
B57-DR7													19			17			**50**							
A2-B46																										
A33-B17																13								16		
A33-B58			8										20			13										
B58-DR3																										
A2-B44						19.6												23								
A24-B7									9							**36**	**21**								**29**	**30**

Refrence sited a - AF; b - Gosh et al. 2003; c - Aizawa et al. 1986; d - Aizawa 1986; e - Solheim et al. 1972; f - Mittal et al. 1982; g - Harris et al. 1972; h - Suciu-Forca et al. 1981; i - Shankarkumar 2003; j - Singal 1972; k - Jaini et al. 2002; I - Mehra et al. 1986, m - Hammond et al. 1979; n - Pitchappan 1986; o - Rajeshekar et al. 1985; p - Rajeshekar et al. 1987; q - 3rd Asia Oceania Workshop 1986; r - Wolf et al. 1980; s - Albert et al. 1972

Table 5: Contd.....

	East												
	Sherpa Nepal	*Nepalese*	*Thai*	*Malay*	*Vietnam*	*Thaiwanese*	*Sichuan*	*Chinese*	*S. Chinese*	*N. Chinese*	*Korean*	*Japanese*	*Japanese*
Reference	t	q	q	q	q	q	q	q	q	q	q	q	k
Language	ST	ST	ST	ST	ST	ST	ST	ST	ST	ST	ST	ST	ST
State													
Sample size$	112	68	138	160	126	85	266	86	407	403	315	472	
A1-B8													
A10(26)-B8		39		54									
A2-B40	6										12	12	
A1-B17	55		19	15	19			24					
A3-B7				10									
A33-B44			97	55				36	7	9	68	41	341
A1-B37			68	5						15		5	
B37-DR10													
B57-DR7		16											
A2-B46			68				133	64	106	43			
A33-B17			46	39	35								
A33-B58		47	46	59		115		133					
B58-DR3		38		58		55							
A2-B44			18								38	31	
A24-B7											27	57	
A26-B61											18	31	

Table 6: Selected Literatures on HLA B association in HIV pathogenesis

Reference	*Population*	*Risk Alleles*	*Protective allele*	*Effect*
Lazaryan *et al* 2006	Zambians, Rwandans	HLA-B58 Supertype	B*5703, B*5703-Cw*18	Viral load lower in B*5703 but not in B-58
Bailey *et al* 2006	USA		HLA-B*57	Low plasma viral load (<50 copies / ml)
Stewart-Jones *et al* 2005	UK		HLA-B*5701 or HLA-B*5703.	Long Term Non Progression; crystal structure
Frahm *et al* 2005	USA		HLA-B*63	Presents B-57 binding epitopes, low viral load (3,280 RNA copies/ml)
Munkanta et al 2005	Japan	HLA-B*5401	HLA-B*1507	
Gaudieri *et al* 2005	Australia		HLA-B*5701, HLA-B*2705	Protection against viraemia
Papasteriades *et al 2005*	Greece	HLA-A28, B21 and DR3,		Fast progressors
Brown *et al* 2005	USA	HLA-A2		27-kD Nef protein down-modulate
Lopez-Vazquez *et al 2005*	Zambia		B57	Protection
Stephens 2005	Caucasoid, African		B57, B58	low viraemia, delayed onset of AIDS
Tomiyama *et al 2005*	Japan		HLA-B*5101	slow progression to AIDS
den Uyl *et al* 2004			HLA-B27	strong CTL response against the p24 epitope
Kiepiela *et al* 2004	South Africa		HLA-B57, B58	Low viral load,
Altfeld *et al* 2003			HLA-B57	Low viral load

.

associated with HIV infection and progression. While there was no unanimity of the identified associations in the yesteryears, many studies now have shown a correlation between HLA-B*57, 27, 15, and 51 with protection mediated by strong CTL responses, low viraemia, and delayed onset and slow progression towards AIDS. A few other alleles such as A2, A28, B21, DR3 and DR5 are high risk alleles for HIV susceptibility. It is essential that an active CTL repertoire is maintained to check the disease under control, with low viral load and non progression towards

AIDS. Sometimes escape mutants also arise under certain MHC background (Leslie et al. 2004). Another study employing HIV viral overlap peptides and CTL responses has shown a dominant effect of HLA-B restricted CTL responses (Kiepiela et al. 2004). The mechanism of HIV susceptibility and progression thus includes (i) HLA B locus alleles associated with peptide specific responses, (ii) plasma viraemia and CD4 counts varying according to the HLA-B allelic polymorphism, (iii) HLA B alleles associated in general with non-progression compared to other HLA loci and (iv) HLA B allele homozygosity disadvantage reflecting in the severity of infection.

All these results need to be viewed in the context of the host genomic diversity of the involved genes, such as alleles of various HLA loci and also the diversity at the level of viral clades and its dynamics resulting from the strong immune attack mounted by the host. This molecular-genetic epidemiology of the host genetics, parasite diversity, host immune response and occurrence of viral variants at the population level looks too intricate (Bhattacharya et al. 2007), but not impossible to study. Recent studies on KIR polymorphisms operating at the level of innate immunity have thrown more light on the possible effect of such genomic diversity in the selection processes (Norman et al 2007).

HLA, HIV and Viral Diversity: As mentioned earlier, many of the HLA alleles implicated in the HIV pathogenesis and CTL responses are the commonest in India. For example, the protective allele HLA B57 is one of the most common in India and is present in higher frequencies in Indian populations (Kiepiela et al. 2004). Similarly B35, A11, DR5 are also quite common. Further the frequencies of various HLA alleles are also quite different in various parts of India and in various linguistic and caste groups (Pitchapan 2002). In southern India, we have observed that HLA-B*17 HIV seropositives possessed significantly higher CD4 count compared to patients carrying other HLA B locus alleles and HLA A* or DRB1* loci were not involved in this association (Pitchappan et al. unpublished results). The infecting organism in majority of the patients in India belongs to Clade C and the virus infecting HLA B57 host is known to develop 'escape mutants' by changing the amino acid in TW10 epitope of Gag protein, thus blunting the onslaught of CTL mediated killing of infected cells (Leslie et al. 2004). However, the virus mutates back to original epitope moiety when transmitted to a new non-B57 host. If an escape variant reaches fixation in the population, the epitope will be lost as a potential target to the immune system, thus explaining how HIV is evolving at a population level. Understanding the direction of HIV evolution thus has important implications for vaccine development (Bhattacharya et al. 2007; Leslie et al. 2005).

Considering the enormous number of known HLA B alleles (N=751) and those new variants hitherto unknown in India, it may be an intelligent guess to imagine the diversity created and added to the population, as a function of time. The mutations fixed in the escape mutants described above thus add up to viral diversity in circulation. In a couple of years, the HLA B alleles that are protective against HIV may increase in the population. Alternatively, many "ineffective" circulating viruses/clades themselves may also lead to a status of 'herd immunity'. The infected Commercial/Casual Sex Workers (CSW) with HLA B57 in India may be a good reservoir spreading the disease, though with low viral load and without developing a clinical disease or even becoming seropositive by themselves. From public health point of view, case finding and targeted anti-retroviral therapy (ART) or microbicide treatment of HLA-B57 and related high risk allele carrying, and non-progressing CSW might be one way of reducing the load and reservoir of HIV virus in circulation. Any new generation vaccine contemplated should induce a sterilizing immunity in this group of people; else the virus infecting non-B57 will revert back with sufficient virulence and the pandemic may continue for long.

The statistics on HIV prevalence in India makes the policy makers to worry of their responsibilities. According to 2004 sentinel survey, estimated prevalence of infection is 5.134 million, but not all the states are equally bearing the brunt. The highest prevalence is in Mumbai and Tamil Nadu: 44.7% among the female CSW and 9.6% in MSM (Men who have Sex with Men) of Mumbai, 39.9% among IVDU (Intra Venous Drug Users) in TamilNadu and 15.65% STD clinics of Mumbai (NACO 2006). It is essential to consider the disease with the prevailing genetic epidemiological background, rather than considering it in isolation.

IV. FUTURE DIRECTION

In India, a multifaceted, multicentric, large scale study to understand its Immuno-genetic-epidemiological profile and molecular-epidemiology of HIV infections, various diseases and disorders is the need of the hour. A partnership between various laboratories with requisite expertise may be highly advantageous, though less happen than done; though the world goes the same way! The expertise and knowledge available in the country and in one country may thus need to be respected, acknowledged and harnessed to work in tandem with the meager resources available globally. The great and long "Indian Heritage" that has resulted in such a fascinating sympatricaly isolated gene pools, may need to be better employed to understand the differential disease susceptibility. We might have answers to most of the diseases in India!

REFERENCES

Aizawa M (Ed.) 1986. *Pre Data Analysis Book* II. *III Asia Oceania Histocompatibility Workshop and Conference*. Sapporo; Hokkai do University Press

Aizawa M, Natori T, Wakisaka A and Konoeda Y (Eds.) 1986 *HLA in Asia – Oceanaia, Proceedings of III Asia Oceania Histocompatibility Workshop and Conference* (Sapporo; Hokkai do University Press)

Albert E D, McClelland J D, Hammer C, Zinkr and Brendel W 1972. Study of the HLA System in the Nepalese Sherpa Population. In: J Dausset, J Colombani (Eds.) *Histocompatibility Testing* (Copenhagen: Munksgaard) pp. 227-232.

Altfeld M, Addo MM, Shankarappa R, Lee PK, Allen TM, et al. 2003. Enhanced Detection of Human Immunodeficiency Virus Type 1-Specific T-Cell Responses to Highly Variable Regions by Using Peptides Based on Autologous Virus Sequences. *J Virol*, **77:** 330-340.

Balakrishnan K, Pitchappan RM, Suzuki K, Shankarkumar U, Santhakumari R, et al. 1996. HLA affinities of Iyers, a Brahmin population of Tamil Nadu, South India. *Human Biology,* **68(4)**: 523-537

Bailey JR, Williams TM, Siliciano RF, Blankson JN 2006. Maintenance of viral suppression in HIV-1-infected HLA-B*57+ elite suppressors despite CTL escape mutations. *J Exp Med*, **203**: 1357-1369.

Bhattacharya T, Daniels M, Heckerman D, Foley B, Frahm N, et al. 2007. Founder Effects in the assessment of HIV polymorphisms and HLA allele associations. *Science*, **315**:1583-1586.

Brown A, Gartner S, Kawano T, Benoit N, Cheng-Mayer C 2005. HLA-A2 down-regulation on primary human macrophages infected with an M-tropic EGFP-tagged HIV-1 reporter virus. *J Leukoc Bio*, **78**: 675-685.

Carrington C.V.F, Kondeatis E, Ramdath D.D, Norman P. J, Vaughan R.W, et al. 2002. A Comparison of HLA-DR and -DQ Allele and Haplotype Frequencies in Trinidadian Populations of African, South Asian, and Mixed Ancestry. *Human Immunology*, **63:** 1045–1054.

Crawford M.H, Reddy B.M, Martinez-Laso J, Steven J. Mack, et al. 2001. Caste Subdivisions of Andhra Pradesh, India, According to the HLA System. *Human Immunology*, **62**: 1031–1041.

den Uyl D, van der Horst-Bruinsma IE and van Agtmael M 2004. Progression of HIV to AIDS: a protective role for HLA-B27? *AIDS Rev*, **6**: 89-96.

Doherty PC 1997. The Nobel Lectures in Immunology. The Nobel Prize for Physiology or Medicine, 1996. Cell mediated immunity in virus infections. *Scand J Immunol*, **46(6):** 527-540

Frahm N, Adams S, Kiepiela P, Linde CH, Hewitt HS et al. 2005. HLA-B63 Presents HLA-B57/B58-Restricted Cytotoxic T-Lymphocyte Epitopes and Is Associated with Low Human Immunodeficiency Virus Load. *J Virol*, **79**: 10218-10225.

Gaudieri S, DeSantis D, McKinnon E, Moore C, Nolan D et al. 2005. Killer immunoglobulin-like receptors and HLA act both independently and synergistically to modify HIV disease progression. *Genes Immun*, **6**: 683-690.

Gosh S 2007. Interpreting a genetic case-control finding: What can be said, what cannot be said and implications in Indian populations. *Ind J Hum Gen*, **17**: 1-4.

Gosh K, Shankarkumar U, and Mohanty D, 2002. Common Human Leucocyte Antigen haplotypes in Indians – its implications in finding unrelated compatible bone marrow donors. *Transfus. Med*, **12**: 43-48.

Hammond MG, Appadoo B, Brain P 1979.HLA and cancer in South African Indians. *Tissue Antigens*, **14(4):** 296-302.

Harris R, Wentzel J, Carroll C A and Jennison R F 1972 HLA Frequencies in West Pakistanis (Punjabi) in the United Kingdom. In: J Dausset and J Colombani (Eds.). *Histocompatibility Testing*. Copenhagen: Munksgaard pp. 163-170.

Jaini R, Naruse T, Kanga U, Kikkawa E, Kaur G, et al, 2002. Molecular diversity of the HLA-A*19 group of alleles in North Indians: Possible oriental influence. *Tissue Antigen*, **59**: 487 : 491.

Kiepiela P, Leslie AJ, Honeyborne I, Ramduth D, Thobakgale C et al. 2004. Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HLA. *Nature*, **432**: 769-775.

Kochhar R 2000. *The Vedic People - Their History and Geography* Orient Longman, New Delhi.

Lazaryan A, Lobashevsky E, Mulenga J, Karita E, Allen S et al. 2006. Human Leukocyte Antigen B58 Supertype and Human Immunodeficiency Virus Type 1 Infection in Native Africans. *J Virol*, **80**: 6056–6060.

Leslie AJ, Pfafferott KJ, Chetty P, Draenert R, Addo MM et al. 2004. HIV evolution: CTL escape mutation and reversion after transmission. *Nat Med*, **10**: 282-289.

Leslie A, Kavanagh D, Honeyborne I, Pfafferott K, Edwards C et al. 2005. Transmission and accumulation of CTL escape variants drive negative associations between HIV polymorphisms and HLA. *J Exp Med*, **201**: 891-902.

Lopez-Vazquez A, Mina-Blanco A, Martinez-Borra J, Njobvu PD, Suarez-Alvarez B et al. 2005. Interaction between KIR3DL1 and HLA-B*57 supertype alleles influences the progression of HIV-1 infection in a Zambian population. *Hum Immunol*, **66**:285-289.

Madden DR, Saper MA, Garrett TP, Bjorkman PJ, Strominger JL et al.1989. Comparison of orthorhombic and monoclinic crystal structures of HLA-A2.*Cold Spring Harb Symp Quant Biol*, 54 **Pt 1:** 353-359.

Mehra NK, Jaini R, Rajalingam R, Kaur G, and Balamurugan A, 2001. Molecular diversity of HLA-A*02 in Asian Indians: predominance of A*0211. *Tissue Antigens*, **57**: 502–507.

Mittal KK, Naik S, Sansonetti N, Cowherd R, and Kumar R 1982. The HLA antigens in Indian Hindus. *Tissue Antigen*, **20**: 223-226.

Munkanta M, Terunuma H, Takahashi M, Hanabusa H, Miura T et al. 2005. HLA-B polymorphism in Japanese HIV-1-infected long-term surviving hemophiliacs. *Viral Immunol*, **18**: 500-505.

NACO, HSS Govt. India HIV/AIDS epidemiological surveillance & estimation report for the year 2005. *NACO, Govt. India HSS – 2006*.

Norman1 PJ, Abi-Rached L, GendzekhadzeK, Korbel D, Gleimer M, Rowley D, Bruno D, Carrington CVN, Chandanayingyong D, Chang Y, Crespí C, Saruhan-Direskeneli G, A Fraser Hameed1 PK, Kamkamidze G, Koram KA, Layrisse Z, Matamoros N, Milà J, Park1 Pitchappan RM, Ramdath D, Ming-Yuh, Shiau16, Stephens HA,Struik S, David H, , Vaughan R W, Tyan D, Davis R , Riley EM, Ronaghi M Parham1 P 2007. Unusual selection on the KIR3DL1/S1 natural killer cell receptor in Africans. *Nature Genetics* (In Press)

Papasteriades CH, Economidou J, Pappas H, Kapsimali V, Psarra K et al. 2005. HLA antigens as predictors of disease progression in HIV-infected haemophilia patients (a 22 years' follow up). *Haemophilia*, **11**: 371-375.

Pitchappan R M, Manickasundari M, Kakkanaiah N N, Mahendran V, Brahamajothi Bai V, et al. 1986. HLA Antigens in Sourashtrans of Madurai. *HLA in Asia – Oceanaia, Proceedings of III Asia Oceania Histocompatibility Workshop and Conference*

Pitchappan RM 1988. Founder effects explain the distribution of the HLA A1-B17 but not the absence of the A1-B8 haplotypes in India. *J Genetics*, **67**: 101-111.

Pitchappan RM, Koteeswaran A, Kakkaiah VN, Manickasundari M, Rajaram V et al. 1989. HLA Bw57 and DR7 with Psoriasis vulgaris in South India. *Tissue Antigens*, **34**: 133-139.

Pitchappan RM, K.Balakrishnan, V. Sudarsen, V. Brahmajothi, V.Mahendran et al. 1997. Sociobiology and HLA Genetic Polymorphism in Hill Tribes, the Irula of the Nilgiris Hills and the Malayali of the Shevroy hills, South India. *Hum Biol*, **69(1)**: 59-74

Pitchappan RM 2002. Castes, Migration, Immunogenetics, Infectious Diseases & south India. *Community Genetics*, **5**:157-161.

Rajasekar R, 1985. *Studies on the HLA System of Different Population Groups in Madurai*. Ph.D Thesis. Madurai Kamaraj University, Madurai.

Rajasekar R, Kakkanaiah V.N, Pitchappan RM 1987. HLA antigens in South India: II. Selected Caste groups of Tamil Nadu. *Tissue Antigens*, **30**: 113-118.

Ravikumar M, Dheenadhayalan V, Rajaram K, Shanmugalakshmi S, Paul Kumaran P, et al 1999. Associations of HLA-DRB1, DQB1 and DPB1 alleles with pulmonary tuberculosis in South India. *Tubercle Lung Dis*, **79** : 309–317.

Roitt I, Brostoff J, Male TK 2006. *Immunology*. 6th Edition

Shankarkumar U, Prasanavar D, Ghosh K, and Mohanty D, 2003. HLA A*02 Allele Frequencies and B Haplotype Associations in Western Indians. *Human Immunology*, **64**: 562–566.

Shankarkumar U, Sridharan B, Pitchappan RM 2004. HLA diversity among Nadars, a primitive Dravidian caste of South India. *Human Immunology*, **65:** 847–854.

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.

Singal D P 1972 The Distribution of HL-A Leucocyte Antigens in Indians. In: J Dausset, J Colombani (Eds.) *Histocompatibility Testing*. Copenhagen: Munksgaard, pp. 179 – 181.

Solheim B G, Bratile A and Thorsby E 1972 Study of the HLA System in a West Pakistan Population. In: J Dausset, J Colombani (Eds.). *Histocompatibility Testing* (Copenhagen: Munksgaard) pp. 171 – 174

Stephens HA 2005. HIV-1 diversity versus HLA class I polymorphism. *Trends Immunol*, **26**: 41-47.

Stewart-Jones GB, Gillespie G, Overton IM, Kaul R, Roche P et al. 2005. Structures of three HIV-1 HLA-B*5703-peptide complexes and identification of related HLAs potentially associated with long-term non progression. *J Immunol*, **175**:2459-2468.

Suciu-Foca N, Reed E, Khan R, Coburn C, Lewison A, et al. 1981. HLA Antigens in Asian Indians: HLA-D-DR Relations. *Human lmmunology*, **3**: 261-270.

Thomas R, Nair S.B, Banerjee M, 2004. HLA-B and HLA-C alleles and haplotypes in the Dravidian tribal populations of southern India. *Tissue Antigens*, **64**: 58–65

Tomiyama H, Fujiwara M, Oka S and Takiguchi M 2005. Cutting Edge: Epitope-dependent effect of Nef-mediated HLA class I down-regulation on ability of HIV-1-specific CTLs to suppress HIV-1 replication. *J Immunol*, **174**: 36-40.

Wells R S, Yudlasheva N, Ruzibakiev R, Underhill P, Evseeva I et al. 2001. The Eurasian Heartland: A continental perspective of Y-chromosome diversity. *Proc Natl Acad Sci USA*, **98**: 10244-10249.

Wolf E, Fine P E N, Pritchard J, Watson B, Bradley D J, Festenstein H, Chacko C J G nd Stevens A 1980 HLA – A, B and C Antigens in South Indian Families with Leprosy. *Tissue Antigens*, **15**: 436-446.

Zinkernagel RM, Doherty PC 1997. The discovery of MHC restriction. *Immunol Today*, **18(1)**: 14-17.