February 21st 2012

Evolutionary Origins of Compartmentalized Cells

Peter Parham

Departments of Structural Biology and Microbiology & Immunology, Stanford University

International Center for Theoretical Sciences National Center for Biological Sciences, Bangalore, India

Laurent Abi-Rached Pa

Lisbeth Guethlein

Achim Moesta

Paul Norman

Ketty Genzekhaze

Ana Older Aguilar

Hugo Hilton

Richard Green, University of California, Santa Cruz

Ashley Moffett, University of Cambridge

Bone-marrow transplant donor registries worldwide

[1] HLA class I : variable ligands for lymphocyte receptors.

[2] Human migration Out-of-Siberia to colonize the Americas

[3] Human migration Out-of-Africa to colonize Eurasia.

[1] HLA class I : variable ligands for lymphocyte receptors.

[1] ~ 5% of mammalian genes are genes of the immune system.

[1] ~ 5% of mammalian genes are genes of the immune system.

[2] Comparison of the mouse and human genomes identified two classes of genes that differ both in number and nucleotide sequence: genes of the immune system and genes of the reproductive system. *Waterston 2002 Nature*

[1] ~ 5% of mammalian genes are genes of the immune system.

[2] Comparison of the mouse and human genomes identified two classes of genes that differ both in number and nucleotide sequence: genes of the immune system and genes of the reproductive system. *Waterston 2002 Nature*

[3] The combination of immune defence and reproduction is essential for the health and survival of individuals, populations and species.

[1] ~ 5% of mammalian genes are genes of the immune system.

[2] Comparison of the mouse and human genomes identified two classes of genes that differ both in number and nucleotide sequence: genes of the immune system and genes of the reproductive system. *Waterston 2002 Nature*

[3] The combination of immune defence and reproduction is essential for the health and survival of individuals, populations and species.

[4] The class I genes of the major histocompatibility complex (MHC) are the most variable and rapidly evolving genes in mammalian genomes.

[1] ~ 5% of mammalian genes are genes of the immune system.

[2] Comparison of the mouse and human genomes identified two classes of genes that differ both in number and nucleotide sequence: genes of the immune system and genes of the reproductive system. *Waterston 2002 Nature*

[3] The combination of immune defence and reproduction is essential for the health and survival of individuals, populations and species.

[4] The class I genes of the major histocompatibility complex (MHC) are the most variable and rapidly evolving genes in mammalian genomes.

[5] MHC class I molecules function as ligands for the receptors of natural killer cells and killer T cells and contribute to both immune defence and reproduction.

Three Genomic Complexes Encode the Ligands and Receptors that Control and Diversify Human NK-cell Function



[1] The human MHC is called the HLA complex.

[1] The human MHC is called the HLA complex.

[2] The highly polymorphic HLA class I genes are named HLA-A, -B and -C.

[1] The human MHC is called the HLA complex.

- [2] The highly polymorphic HLA class I genes are named HLA-A, -B and -C.
- [3] HLA-A, -B, -C were discovered in the late 1950s as major genetic factors that cause complications (graft rejection and graft-versus-host disease) in clinical transplantation.

[6] The human MHC is called the HLA complex.

- [7] The highly polymorphic HLA class I genes are named HLA-A, -B and -C.
- [8] HLA-A, -B, -C were discovered in the late 1950s as major genetic factors that cause complications (graft rejection and graft-versus-host
- [9] HLA gene polymorphisms also associate with a wide range of human diseases, frequently giving the strongest genetic association.

[1] The human MHC is called the HLA complex.

- [2] The highly polymorphic HLA class I genes are named HLA-A, -B and -C.
- [3] HLA-A, -B, -C were discovered in the late 1950s as major genetic factors that cause complications (graft rejection and graft-versus-host disease) in clinical transplantation.
- [4] HLA gene polymorphisms also associate with a wide range of human diseases, frequently giving the strongest genetic association.
- [5] Because of the clinical importance some 20 million human individuals, representing a wide geographical and ethnical range, have been genotyped for HLA-A, -B, -C.

[6] HLA-A, -B and -C genes do not have a wild-type form. Instead they are represented by a variety of allelic forms, each present at significant frequency (~1-25%). Such numbers of alleles are maintained by balancing selection.

[6] HLA-A, -B and -C genes do not have a wild-type form. Instead they are represented by a variety of allelic forms, each present at significant frequency (~1-25%). Such numbers of alleles are maintained by balancing selection.

[7] The amino-acid substitutions that distinguish HLA variant forms alter the specificity and strength of the immune response to infection. HLA class I variability concentrates in the binding site for peptides



Figure 8-30 The Immune System, 2/e (© Garland Science 2005)

[6] HLA-A, -B and -C genes do not have a wild-type form. Instead they are represented by a variety of allelic forms, each present at significant frequency (~1-25%). Such numbers of alleles are maintained by balancing selection.

[7] The amino-acid substitutions that distinguish HLA variant forms alter the specificity and strength of the immune response to infection.

[8] Almost all individuals are heterozygous for HLA-A, -B and -C.

[6] HLA-A, -B and -C genes do not have a wild-type form. Instead they are represented by a variety of allelic forms, each present at significant frequency (~1-25%). Such numbers of alleles are maintained by balancing selection.

[7] The amino-acid substitutions that distinguish HLA variant forms alter the specificity and strength of the immune response to infection.

[8] Almost all individuals are heterozygous for HLA-A, -B and -C.

[9] Most individuals have different HLA types and thus distinctive immune systems with different strengths and weaknesses.

How much HLA class I diversity does a human population need to survive?

[1] HLA class I : variable ligands for lymphocyte receptors.

[1] The first human colonization of the Americas began with the arrival at Alaska ~17,000 years ago of Asian migrants who walked from Siberia.



~17,000 years ago Asian migrants from Siberia were the first human populations to survive long term in America

- [1] The first human colonization of the Americas began with the arrival at Alaska ~17,000 years ago of Asian migrants who walked from Siberia.
- [2] The migrant bands were small and took with them only a fraction of the genetic diversity in the Siberian population.

- [1] The first human colonization of the Americas began with the arrival at Alaska ~17,000 years ago of Asian migrants who walked from Siberia.
- [2] The migrant bands were small and took with them only a fraction of the genetic diversity in the Siberian population.
- [3] Some of this genetic diversity was lost during the long, arduous journey as a consequence of disease, conflict, population splitting and population bottleneck.

- [1] The first human colonization of the Americas began with the arrival at Alaska ~17,000 years ago of Asian migrants who walked from Siberia.
- [2] The migrant bands were small and took with them only a fraction of the genetic diversity in the Siberian population.
- [3] Some of this genetic diversity was lost during the long, arduous journey as a consequence of disease, conflict, population splitting and population bottleneck.
- [4] These trends continued as the emigrants' journey continued to colonize all of North, Central and South America.

- [1] The first human colonization of the Americas began with the arrival at Alaska ~17,000 years ago of Asian migrants who walked from Siberia.
- [2] The migrant bands were small and took with them only a fraction of the genetic diversity in the Siberian population.
- [3] Some of this genetic diversity was lost during the long, arduous journey as a consequence of disease, conflict, population splitting and population bottleneck.
- [4] These trends continued as the emigrants' journey continued to colonize all of North, Central and South America.
- [5] A second colonization of the Americas began in 1492 with the arrival by boat of Europeans, precipitating >500 years of conflict, disease, and bottleneck for Amerindian populations.

CONCLUSIONS

CONCLUSIONS

[1] From their beginning Amerindian populations have always had less genetic diversity than other large continental population groups: Africans, Asians, Europeans.

CONCLUSIONS

[1] From their beginning Amerindian populations have always had less genetic diversity than other large continental population groups: Africans, Asians, Europeans.

[2] Throughout their history Amerindians have have been subject to episodes of strong natural selection from infectious disease and conflict (warfare) leading to population bottleneck.

CONCLUSIONS

[1] From their beginning Amerindian populations have always had less genetic diversity than other large continental population groups: Africans, Asians, Europeans.

[2] Throughout their history Amerindians have have been subject to episodes of strong natural selection from infectious disease and conflict (warfare) leading to population bottleneck.

[3] Amerindians provide a test case to address the question:
Human Migration Out-of-Siberia to Colonize the American Continents

CONCLUSIONS

[1] From their beginning Amerindian populations have always had less genetic diversity than other large continental population groups: Africans, Asians, Europeans.

[2] Throughout their history Amerindians have have been subject to episodes of strong natural selection from infectious disease and conflict (warfare) leading to population bottleneck.

[3] Amerindians provide a test case to address the question:

CONCLUSIONS

- [1] From their beginning Amerindian populations have always had less genetic diversity than other large continental population groups: Africans, Asians, Europeans.
- [2] Throughout their history Amerindians have have been subject to episodes of strong natural selection from infectious disease and conflict (warfare) leading to population bottleneck.
- [3] Amerindians provide a test case to address the question: How much HLA class I diversity do human populations need to survive long term?

The Yucpa population of Venezuela have a typical Amerindian history



Allele	HLA-A KIR Ligand	Frequency %	Allele	HLA- KIR Ligand	B Frequency %	Allele	HLA- KIR Ligand	C Frequency %
*6801		46.7	*3909		41.8	*0702	C1	76.2
*0204		27.8	*3905		34.4	*1503	C2	9.8
*0212		0.8	*5201	Bw4	9.8	*0401	C2	7.4
*0213		0.8	*3512		7.4	*0302	C1	4.1
*2402	Bw4	17.2	*3543		1.6	*0304	C1	0.8
*3101		6.5	*4002		4.1	*0102	C1	1.6
			*4004		0.8			
Total	Bw4	17.2	Total	Bw4	9.8	Total	C1	82.8 (97)
	Pheno		C2	17.2 (33)				

1. A balance of six HLA-A, seven HLA-B and six HLA-C alleles

Allele	HLA-A KIR Ligand	Frequency %	Allele	HLA- KIR Ligand	B Frequency %	Allele	HLA- KIR Ligand	C Frequency %
*6801		46.7	*3909		41.8	*0702	C1	76.2
*0204		27.8	*3905		34.4	*1503	C2	9.8
*0212		0.8	*5201	Bw4	9.8	*0401	C2	7.4
*0213		0.8	*3512		7.4	*0302	C1	4.1
*2402	Bw4	17.2	*3543		1.6	*0304	C1	0.8
*3101		6.5	*4002		4.1	*0102	C1	1.6
			*4004		0.8			
Total	Bw4	17.2	Total	Bw4	9.8	Total	C1	82.8 (97)
	Pheno		C2	17.2 (33)				

1. A balance of six HLA-A, seven HLA-B and six HLA-C alleles

2. Amerindian-specific alleles: one HLA-A, three HLA-B, no HLA-C

Allele	HLA-A KIR Ligand	Frequency %	Allele	HLA- KIR Ligand	B Frequency %	Allele	HLA- KIR Ligand	C Frequency %
*6801		46.7	*3909		41.8	*0702	C1	76.2
*0204		27.8	*3905		34.4	*1503	C2	9.8
*0212		0.8	*5201	Bw4	9.8	*0401	C2	7.4
*0213		0.8	*3512		7.4	*0302	C1	4.1
*2402	Bw4	17.2	*3543		1.6	*0304	C1	0.8
*3101		6.5	*4002		4.1	*0102	C1	1.6
			*4004		0.8			
Total	Bw4	17.2	Total	Bw4	9.8	Total	C1	82.8 (97)
	Pheno		C2	17.2 (33)				

1. Six HLA-A, seven HLA-B and six HLA-C alleles

2. Amerindian-specific alleles: one HLA-A, three HLA-B, no HLA-C

3. Have C1, C2 and Bw4 epitopes but no A3/11 epitope.

Variable NK cell receptors Recognize Four Epitopes of HLA-A,-B,and -C



Allele	HLA-A KIR Ligand	Frequency %	Allele	HLA- KIR Ligand	B Frequency %	Allele	HLA- KIR Ligand	C Frequency %
*6801		46.7	*3909		41.8	*0702	C1	76.2
*0204		27.8	*3905		34.4	*1503	C2	9.8
*0212		0.8	*5201	Bw4	9.8	*0401	C2	7.4
*0213		0.8	*3512		7.4	*0302	C1	4.1
*2402	Bw4	17.2	*3543		1.6	*0304	C1	0.8
*3101		6.5	*4002		4.1	*0102	C1	1.6
			*4004		0.8			
Total	Bw4	17.2	Total	Bw4	9.8	Total	C1	82.8 (97)
	Pheno		C2	17.2 (33)				

- 1. Six HLA-A, seven HLA-B and six HLA-C alleles
- 2. Amerindian-specific alleles: one HLA-A, three HLA-B, no HLA-C
- 3. Have C1, C2 and Bw4 epitopes but no A3/11 epitope.
- 4. High frequency of *HLA-C*0702*.

[1] Compared to other genetic markers, Amerindian populations preserve a diversity of HLA-A, -B, -C variants that ensures a majority of individuals are heterozygous for HLA-A and HLA-B.

- [1] Compared to other genetic markers, Amerindian populations preserve a diversity of HLA-A, -B, -C variants that ensures a majority of individuals are heterozygous for HLA-A and HLA-B.
- [2] Lost HLA class I diversity was replaced by selection for new HLA-A and HLA-B variants. Some new variants arose by point mutation of an allele brought from Siberia, but the majority arose by recombination between two alleles brought from Siberia, which introduces two or more substitutions to the peptide-binding site.

- [1] Compared to other genetic markers, Amerindian populations preserve a diversity of HLA-A, -B, -C variants that ensures a majority of individuals are heterozygous for HLA-A and HLA-B.
- [2] Lost HLA class I diversity was replaced by selection for new HLA-A and HLA-B variants. Some new variants arose by point mutation of an allele brought from Siberia, but the majority arose by recombination between two alleles brought from Siberia, which introduces two or more substitutions to the peptide-binding site.
- [3] Amerindians retain three of the four HLA class I epitopes that interact with variable NK cell receptors.

- [1] Compared to other genetic markers, Amerindian populations preserve a diversity of HLA-A, -B, -C variants that ensures a majority of individuals are heterozygous for HLA-A and HLA-B.
- [2] Lost HLA class I diversity was replaced by selection for new HLA-A and HLA-B variants. Some new variants arose by point mutation of an allele brought from Siberia, but the majority arose by recombination between two alleles brought from Siberia, which introduces two or more substitutions to the peptide-binding site.
- [3] Amerindians retain three of the four HLA class I epitopes that interact with variable NK cell receptors.
- [4] Human populations fpr which HLA class I diversity falls below that observed in Amerindians either died out or were assimilated by another population.

Are there genetic benefits from the second human colonization of the Americas ?

- [1] Overall the European conquest of the Americas did not greatly decrease the overall genetic diversity of Amerindians as a whole.
- [2] During the early and formative period of colonisation the European populations were small and likely carried only a fraction of the genetic diversity from their population of origin.
- [3] During this formative period the well-documented parenting between Europeans and Amerindians could have been mutually beneficial due to the introduction into both populations of new HLA class I variants.
- [4] The degree of assimilation has varied greatly between countries and is still an ongoing evolution. For example, in many urban Mexican populations there is now an approximately 50:50 ratio of European and Amerindian HLA variants.

Boosting Eurasian Immunity by Alliance and Dalliance of Humans, Ancient and Modern

[3] Human migration Out-of-Africa to colonize Eurasia.

~350,000 years ago Neandertal ancestors went Out-of-Africa



~70,000 years ago modern humans went Out-of-Africa and co-existed with Neandertals in Europe and Asia for ~30,000 years before Neandertals disappeared from the fossil record Two human migrations Out-of-Africa to colonise Europe and Asia: did modern humans Neandertals meet, mate and conflict?



HLA-B*73 is an unusual HLA-B that has the C1 Epitope, a common feature of chimpanzee HLA-B



HLA-B*73 unusually combines ancient sequence divergence with modern sequence homogeneity



Abi-Rached et al 2011 Science

HLA-B*73 localizes to western Asia and is in strong linkage disequilibrium with HLA-C*1505



HLA-B*73 localizes to western Asia and is in strong linkage disequilibrium with HLA-C*1505



HLA-B*73 is particularly frequent in Parsees

Its unusual properties and the results from demographic simulations point to HLA-B*73 having been passed from archaic humans to modern humans by horizontal transfer



Abi-Rached et al 2011 Science 334: 89-94. A unique functional feature of the HLA-B*7301 haplotype is that carries both the C1 and C2 epitopes



Whole Genome Sequence Comparison of Neutral Markers Gave Evidence For up to 6% Archaic DNA in Modern Eurasian Genomes





Ancient DNA from preserved finger bone



Denisova, Neanderthal and Modern Humans all co-existed



Estimated ~4-6% Denisovan contribution to modern Melanesian genomes and 1-4% Neandertal contribution to modern Eurasian genomes.

HLA typing of the Denisovan woman.

Denisovan HLA class I									
Allele				Closest modern ty	Next best type				
Locus	#	Coverage	Reads (#)	Name	Differences	Name	Differences		
HLA-A	1	15%	15	A*02:01/03/07/48	0	A*68	7		
	2	21%	17	A*11:01/53	0	A*03/*30	4		
HI A_B	1	34%	35	B*15:58	3	B*46	5		
IILA-D	2	39%	43	B*35:63 ¹	0	B*53	9		
HLA-C	1	33%	30	C*12:02:02	0	C*06	7		
	2	19%	16	C*15:02/05/17 [§]	1	C*02	5		

She shares five of six HLA-A, -B, -C alleles with modern humans

Denisovan HLA class I									
Allele				Closest modern ty	Next best type				
Locus	#	Coverage	Reads (#)	Name	Name Differences				
	1	15%	15	A*02:01/03/07/48	0	A*68	7		
TILA-A	2	21%	17	A*11:01/53	0	A*03/*30	4		
ШЛ В	1	34%	35	B*15:58	3	B*46	5		
TILA-D	2	39%	43	B*35:63 ¹	0	B*53	9		
HLA-C	1	33%	30	C*12:02:02	0	C*06	7		
	2	19%	16	C*15:02/05/17 [§]	1	C*02	5		

She carried all four epitopes that are ligands for variable NK cell receptors: A3/11, Bw4, C1, and C2.

Denisovan HLA class I										
Allele				Closest modern ty	Next best type					
Locus	#	Coverage	Reads (#)	Name	Differences	Name	Differences			
	1	15%	15	A*02:01/03/07/48	0	A*68	7			
TILA-A	2	21%	17	A*11:01/53	0	A*03/*30	4			
HI A_B	1	34%	35	B*15:58	3	B*46	5			
TILA-D	2	39%	43	B*35:63 [¶]	0	B*53	9			
HLA-C	1	33%	30	C*12:02:02	0	C*06	7			
	2	19%	16	C*15:02/05/17 [§]	1	C*02	5			

The HLA-A to HLA-C haplotype



All possible Denisovan *HLA-A/C* haplotypes are absent from Africa and present in Asia



All possible Denisovan *HLA-A/C* haplotypes are absent from Africa and present in Asia



Modern humans acquired these haplotypes from archaic humans





These three alleles were passed from archaic to modern humans.



These three alleles were passed from archaic to modern humans.

Their frequencies in modern humans exceed the neutral estimates.



These three alleles were passed from archaic to modern humans.

Their frequencies in modern humans exceed the neutral estimates.

On acquisition by modern humans they were subject to selection.



The two Denisovan HLA-C alleles are those associated with HLA-B*73

HLA typing of the Neandertal woman.

Neandertal HLA class I										
Allele				Closest modern ty	Next best type					
Locus	#	Coverage	Reads (#)	Name	Differences	Name	Differences			
	1	30%	40	A*02[not :05]	0	A*68	14			
ΠLΑ-Α	2	16%	16	A*26/*66	0	A*34	2			
	1	28%	34	B*07:02/03/06 [§]	0	B*48	2			
TILA-D	2	32%	43	B*51:01/08	0	B*52/*78	2			
HLA-C	1	35%	52	C*07:02 [§]	0	C*08/*18	46			
	2	25%	31	C*16:02 [§]	0	C*05	9			
She shares all six *HLA-A*, *-B*, *-C* alleles with modern humans

Neandertal HLA class I							
Allele				Closest modern type		Next best type	
Locus	#	Coverage	Reads (#)	Name	Differences	Name	Differences
HLA-A	1	30%	40	A*02[not :05]	0	A*68	14
	2	16%	16	A*26/*66	0	A*34	2
HLA-B	1	28%	34	B*07:02/03/06 [§]	0	B*48	2
	2	32%	43	B*51:01/08	0	B*52/*78	2
HLA-C	1	35%	52	C*07:02 [§]	0	C*08/*18	46
	2	25%	31	C*16:02 ^s	0	C*05	9

She has three of the four epitopes for NK cell receptors: C1, C2, and Bw4.

Neandertal HLA class I							
Allele				Closest modern type		Next best type	
Locus	#	Coverage	Reads (#)	Name	Differences	Name	Differences
HLA-A	1	30%	40	A*02[not :05]	0	A*68	14
	2	16%	16	A*26/*66	0	A*34	2
HLA-B	1	28%	34	B*07:02/03/06 [§]	0	B*48	2
	2	32%	43	B*51:01/08	0	B*52/*78	2
HLA-C	1	35%	52	C*07:02 [§]	0	C*08/*18	46
	2	25%	31	C*16:02 [§]	0	C*05	9

All possible Neandertal *HLA-A/C* haplotypes are absent from Africa and present in Eurasia



All possible Neandertal *HLA-A/C* haplotypes are absent from Africa and present in Eurasia



Modern humans acquired these haplotypes from archaic humans

Modern distribution of Neandertal HLA-C alleles



The frequencies of the Neandertal allele far exceeds those of neutral markers

On entering modern humans the Neandertal alleles were positively selected

In Today's Modern Human Population Some 50% of MHC Class I alleles Were Acquired by Social Interactions with Archaic Humans

Population	Putative archaic ancestry atHLA-A
African	6.7%
European	51.7%
Chinese	72.2%
Japanese	80.7%
Papua New Guinea	82.3% [65.9% - 95.3%]

Archaic contribution to HLA-A in modern Eurasians and Melanesians is >50%



Summary: I

Improving HLA diversity through population admixture

- 1. The contribution of archaic HLA class I to modern Eurasian and Melanesian HLA class I is much greater than the total genome estimates, pointing to selective advantage.
- 2. The advantage of archaic HLA class I to modern humans was to replenish lost HLA diversity and provide ready-made alleles already adapted to the local environment.
- 3. Adaptive introgression is unlikely to be restricted to HLA : all polymorphic gene families of the immune system and elsewhere are potential candidates for examination.
- 4. Population admixture is essentially another form of recombination that can introduce particularly divergent new variants with a single mating event.

Summary: II

Improving HLA diversity through population admixture

5. Acquisition of archaic HLA class I alleles may have been essential for the survival of modern humans in Eurasia.

6. The Neandertals may not have gone extinct but eventually became assimilated by a much larger population of Moderns.

Gene conversion between HLA-C and HLA-B formed C1-bearing HLA-B*46 which was selected and has spread to reach high frequencies in South East Asian populations





Abi-Rached et al 2010 PLoS Genetics

*KIR3D51*013* is also a candidate for adaptive introgression

