

Characterizing MHC Ancestral Haplotypes from genotyping MHC loci in parent-child trio's of the Thousand Genome Project

Ting Wang, David Sayer
Thermo Fisher Scientific, West Hills, CA, USA

ThermoFisher
SCIENTIFIC

Background and Aims

- Characterization of the MHC block structure and understanding the extent of diversity of each block is critical for the understanding of non-HLA MHC contributions to outcomes in transplantation and contributions to disease.
- “Strings” of MHC blocks, referred to as polymorphic frozen blocks (PFB), conserved through generations are known as Ancestral / Conserved Extended Haplotypes (AH/CEH), have remained unchanged since the early 2000's [PMID: 1464552, 11029610].
- Our aim is to create a catalogue of AH/CEH and PFB characterized at loci spanning the MHC. By typing whole genome sequence data from geographically distinct samples from the Thousand Genome Project (1KGP) [PMID: 18327223] parent-child trio's, we aim to:
 - Detect MHC haplotype recombination hotspot and infer MHC genomic block structure.
 - Characterize existing AH/CEH at classical HLA loci and non-classical MHC loci at high-resolution.
 - Identify new AH/CEH.

Materials and Methods

- Whole genome sequencing (WGS) data of 599 parent-child family trios (totally 1797 samples) of 5 population groups (**AFR** - African, **AMR** - American, **EAS** - East Asian, **EUR** - European, and **SAS** - South Asian) in the 1KGP.
- Genotypes of 25 MHC loci (**Fig 1**) were determined using (TSV) TypeStream Visual NGS Analysis Software from One Lambda.
- Haplotypes were estimated by *haplo.stats* R package and were further refined/verified by allele segregation in parent-child trios.
- MHC block structure was determined by Asymmetric LD calculation [PMID: 25023400] and an inhouse method - BLOB - for calculating recombination frequency between loci.
- AH/CEHs were identified in at least 3 unrelated individuals according to the previous definitions and were compared with the previously reported AH/CEHs [PMID: 1464552, 11029610].
- Due to high rates of recombination between HLA loci F~V, A~C and DQB1~DPB1, and small population sizes, the minimum requirements for defining an AH/CEH was HLA-C~HLA-DQB1.

Results and Conclusion

- 1564 unique haplotypes of the 25 MHC loci are identified, and 7 frequent (>0.5%) recombination hotspots were detected (**Fig 1 and 2**), indicating 8 MHC genomic blocks (**Fig 3**).
 - 108, 126 and 84 polymorphic frozen blocks (PFB) were identified in Alpha, Beta, and Delta blocks, and the distributions of unique PFBs by population are shown in **Fig 4**.
 - 20, 36, and 84 unique AH/CEHs, spanning F~HDPB1, V~DQB1, and C~DQB1, were identified respectively (**Fig 5A**). Analysis for the presence of AH/CEH or their PFB for each population revealed that <20% of individuals in each population had MHC that didn't include PFB from any of the identified AH/CE (**Fig 5B**).
- Larger studies of more MHC loci are likely to identify more AH/CEH and further characterize MHC block structure.
 - Our study represents the early stages of a larger effort to create a comprehensive and dynamic catalog of global AH/CEH.
 - A comprehensive catalogue of PFB will guide the development of future MHC typing and matching strategies.

Fig 1 - Recombination frequency between MHC loci using BLOB identifies MHC block boundaries; additional blocks may be identified if loci within the recombination hotspots are included, particularly between J-L, L-E, and E-C

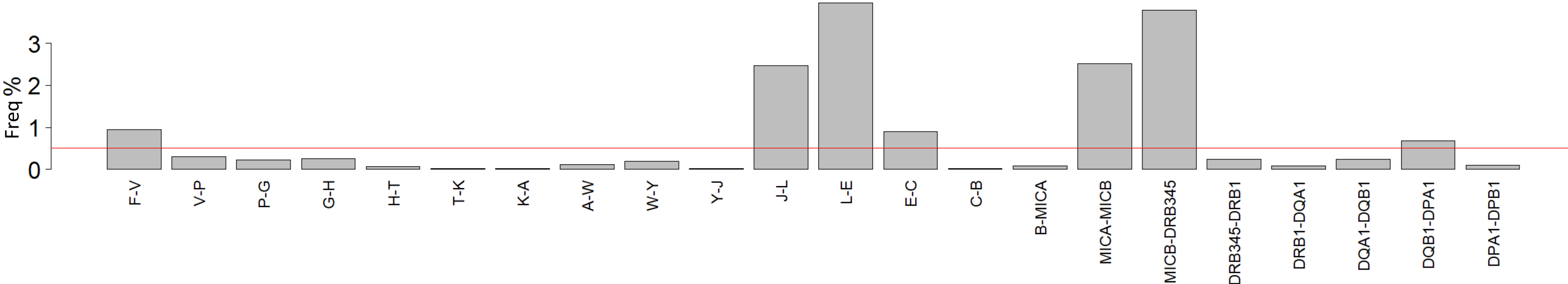


Fig 2 - Asymmetric LD Matrix Identifies PFB (LD>.8)

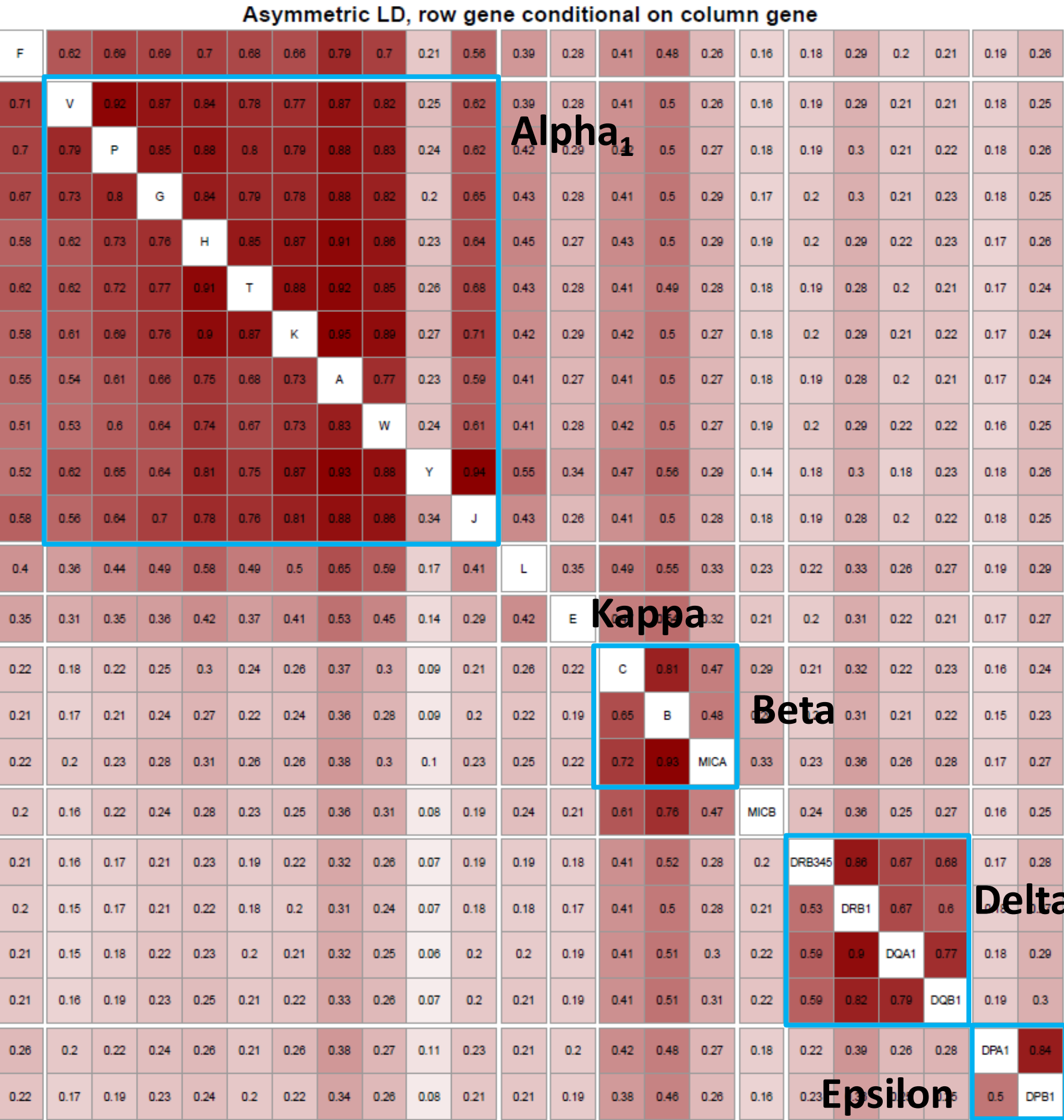


Fig 3 - MHC Genomic Block Structure

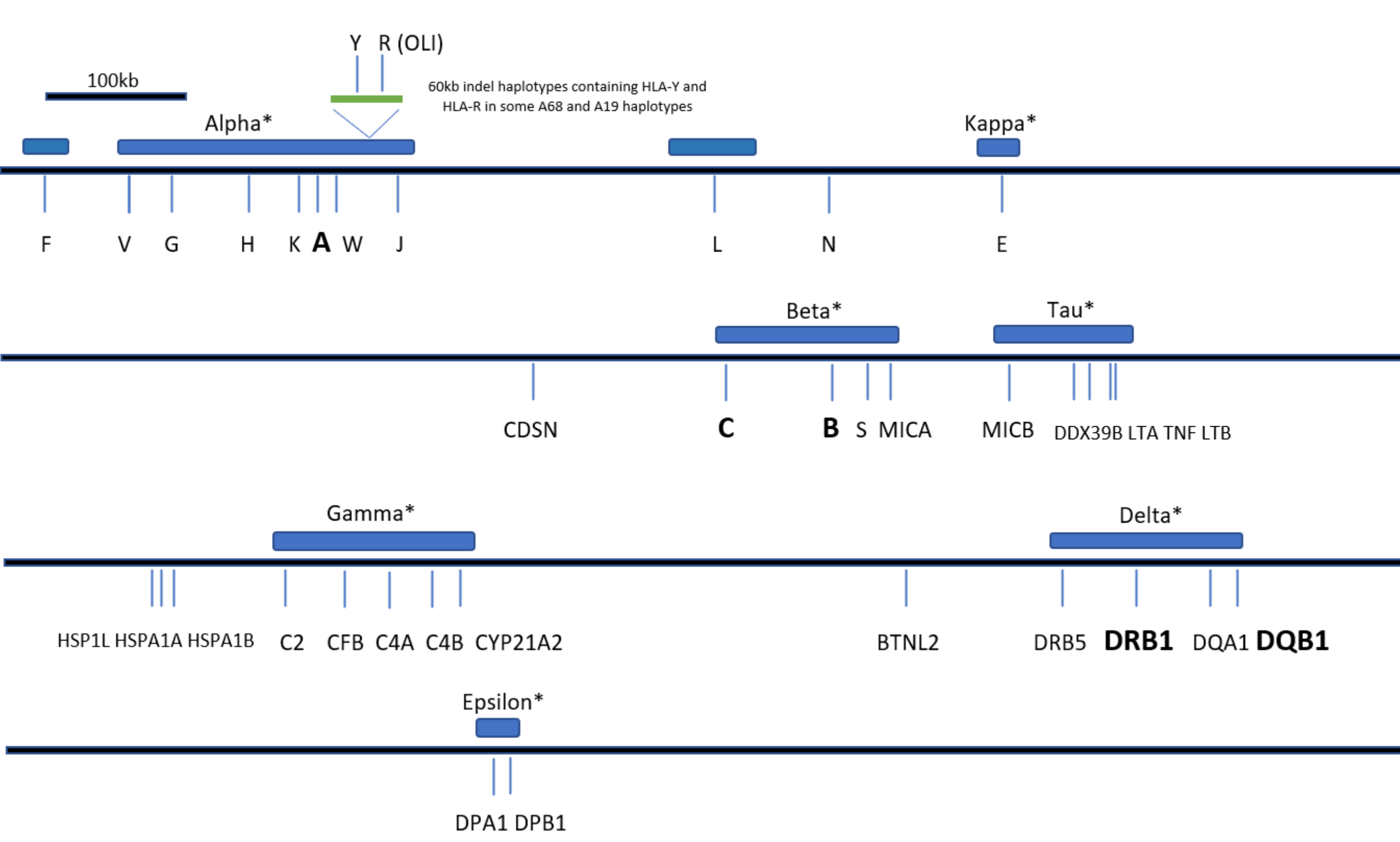


Fig 4 - Distribution of Unique PFB by 1KGP Population

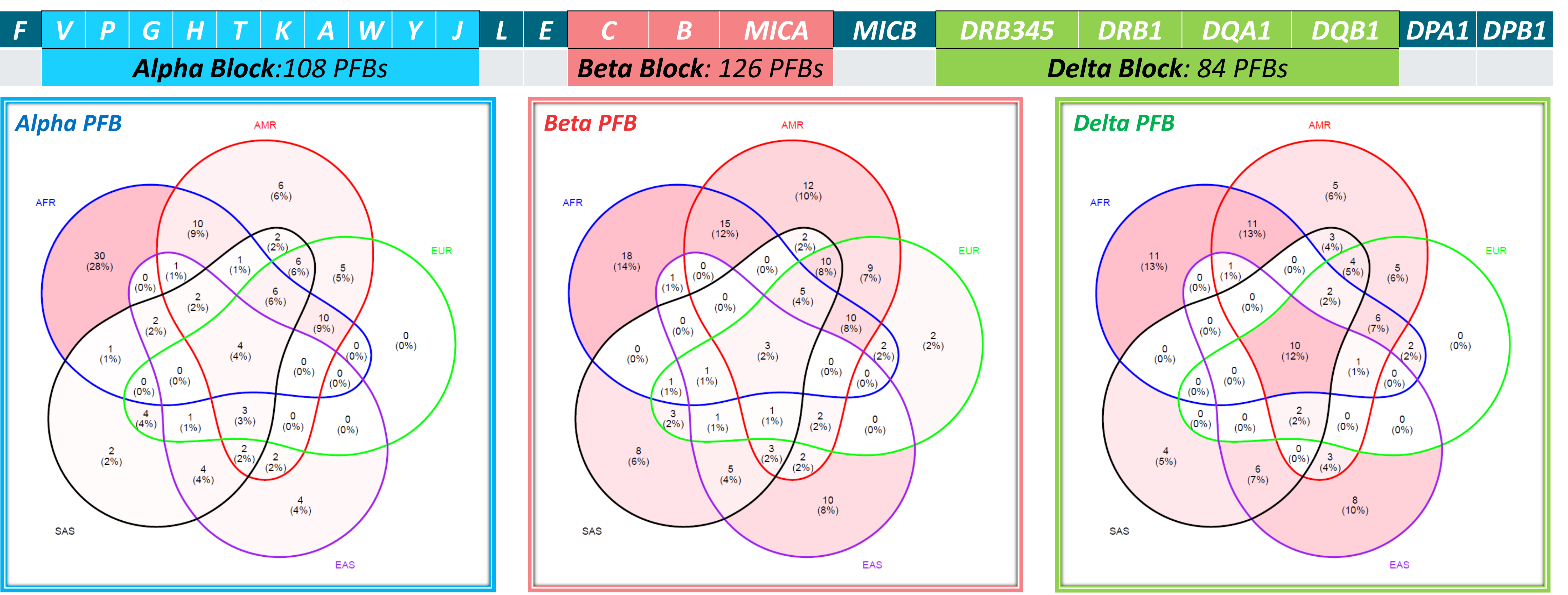


Fig 5 - AH and Recombinant PFB in the 1KGP Populations

A	F	V	P	G	H	T	K	A	W	Y	J	L	E	C	B	MICA	MICB	DRB345	DRB1	DQA1	DQB1	DPA1	DPB1				
Complete AH	20 unique F~DPB1 AH (13 known + 7 new)																										
		36 unique V~DQB1 AH (15 known + 21 new)																									
															84 unique C~DQB1 AH (30 known + 54 new)												
Recombinant with ≥ 1 recognizable PFB from the complete AH		Alpha										×	Beta				×	Delta									
														Beta				×	Delta								
		Alpha										×	Beta														
	Alpha										×										Delta						
	Alpha																										
															Beta												
																		Delta									
Non-PFB Hap.																											

