LD at Functional Sites Across Global Populations

The population genomics of an immune gene family

Roxanne Leiter and Christopher Tillquist, University of Louisville

UNIVERSITY OF LOUISVILLE. COLLEGE OF ARTS & SCIENCES

Background

Linkage disequilibrium, or LD, is defined as a non–random association between two or more genetic loci (Lewontin and Kojima, 1960; Slatkin, 2008). The ability to establish correlations between disease phenotypes and particular polymorphisms within the genome gave the concept of linkage disequilibrium practical significance. LD patterns can also reflect the influence of past selection events, population histories, population structure, gene flow and mutation events in the context of local recombination rates, sex differences and population characteristics (Ardlie et al., 2002; Slatkin, 2008). Several studies have established more extensive LD in populations outside of Africa compared to those within Africa, which are assumed to be older and more historically established (Flint-Garcia et al., 2003). A point of contention within LD research—and a difficulty that will be addressed here—is to what extent do populations evince similar patterns of LD in functional parts of the genome. The answer to this question has practical implications for endeavors such as the International HapMap Project, which seeks to advance personalized medicine through the determination of disease phenotypes with an underlying genetic basis (Liu et al., 2004).

Materials and Methods

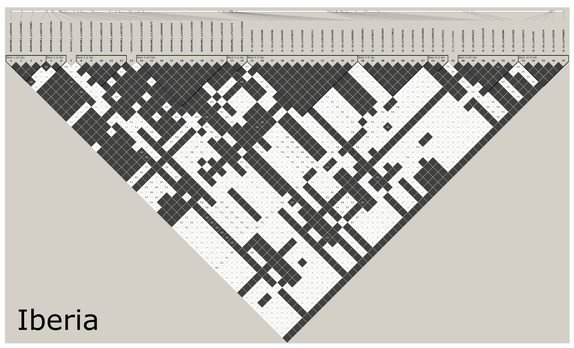
We targeted five genes on the q–arm of chromosome 1, four of which are part of the IL10 family cluster. Interleukins are a particular class of cytokines, which participate in immune function by acting as a messenger between immune cells (Brocker et al., 2010). Interleukins are primarily involved in immune cell development, activation and differentiation, and can inhibit or promote inflammatory responses. Cis to this gene family is MAPKAP-K2, a gene encoding a kinase that is part of the serine/threonine-protein kinase family. Much like the interleukins, these kinases exhibit a diverse array of functions in biological systems leading to the activation of downstream metabolic pathways through further phosphorylation (Uniprot, 2014).

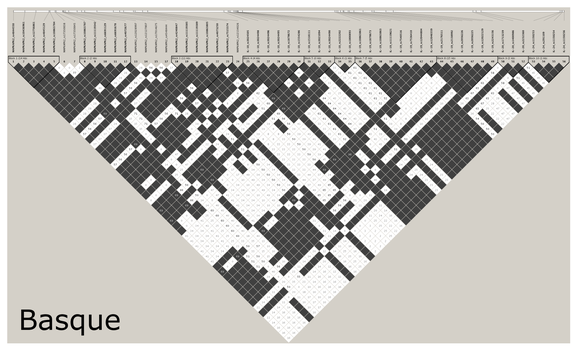
The SNP data at 57 loci for these analyses were generated on the Sequenom genotyping platform at the Genomic Analysis and Technology Core at the University of Arizona using custom primers for SNPs within genes. A total of 82 samples represent Basque, Chinese, Iberian, Indo-Pakistan, Middle East, North African, Pacific, Russian and South African populations (Coriell). Linkage disequilibrium and haplotype analyses were performed using the program Haploview, which allows data input, LD analysis, triangular heat map generation, haplotype block estimation and visualization (Barrett et al., 2005).

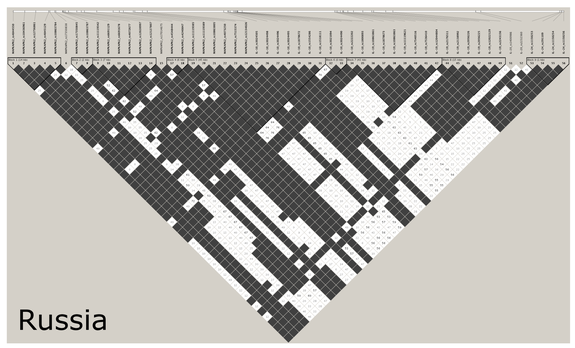
Table 1 (below): This table shows haplotype block boundaries, which are indicated by bolded lines between markers. SNPs that are contained within a block but are not tagSNPs are marked with an X while tagSNPs are marked with a T in blue. Blank cells were not contained within haplotype blocks or did not have data included in a particular population due to data trimming.

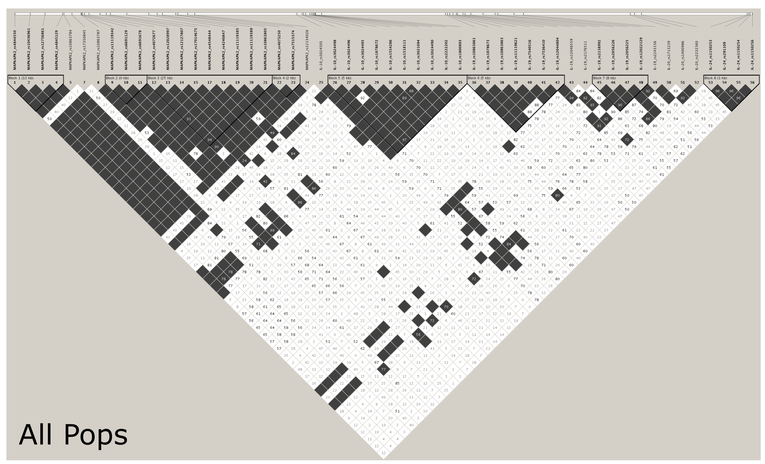
	MAPKAPK2																							IL10												IL19																			IL20			IL24					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57
All Populations	T	X	X	X	T	.	.	T	T	X	X	T	T	T	X	.	X	T	T	X	X	X	T	T	T	T	T	T	T	T	T	X	T	T	T	.	T	T	T	X	T	T	T	T	T	T	X	X	.
Basque (10)	T	X	X	X	T	.	.	T	T	X	X	T	.	.	.	.	.	X	T	X	X	T	T	.	T	T	T	X	X	T	T	T	X	T	X	T	X	X	X	X	X	T	T	X	X	X	X	X	.
China (10)	X	T	X	X	X	.	T	.	X	X	X	T	T	X	X	.	X	X	X	X	X	X	X	X	.	.	.	.	T	X	T	T	.	X	T	X	X	X	X	X	T	T	T	X	T	X	X	T	.	T	T	X	X	X	X	T
Iberia (10)	T	X	X	X	T	T	.	T	T	T	X	T	.	T	T	X	X	T	T	X	T	X	T	T	T	T	T	X	X	T	T	T	X	X	X	T	X	X	X	X	X	T	T	T	.	T	T	X	.	T	T	X	T	X	X	X	X
Indo-Pakistan (9)	T	X	X	X	T	T	.	X	T	X	X	T	T	X	T	X	X	.	T	X	T	X	T	T	.	T	T	T	X	T	X	T	T	X	X	X	X	X	X	X	T	.	T	.	X	T	X	T	T	T	T	X	X	X	.
Middle East (10)	X	T	X	X	T	.	.	.	.	T	.	T	T	X	T	X	X	T	X	T	X	T	T	T	T	T	T	X	T	T	X	T	T	X	X	X	T	X	X	X	X	T	T	X	T	X	T	X	X	T	X	X	X	X	.
N. Africa (7)	T	X	X	X	T	.	.	T	T	X	X	T	.	X	.	.	X	T	T	X	X	.	T	T	X	.	T	X	X	T	X	T	X	T	X	T	X	X	X	X	X	X	T	T	T	T	X	.	T	T	T	T	X	X	X	.
Pacific (7)	X	X	X	T	T	.	.	T	X	T	X	T	.	X	.	.	X	.	X	T	X	X	X	X	.	.	T	.	.	T	X	X	T	.	X	X	T	X	T	X	X	X	.	X	T	X	T	X	X	X	.	T	T	T	X	X	X	.
Russia (10)	T	X	X	X	T	.	T	T	T	X	X	T	X	X	.	.	T	T	T	X	X	T	T	X	T	T	X	X	T	T	T	T	T	T	X	T	X	X	X	X	X	T	.	T	T	X	X	X	X	.	.	.	T	T	X	X	.
S. Africa (9)	T	X	X	X	T	.	.	T	T	X	X	T	T	T	X	.	X	T	T	X	X	X	T	T	T	T	T	T	T	T	T	T	X	T	T	T	X	T	X	X	X	.	T	X	T	T	T	X	T	T	T	T	T	X	X	.

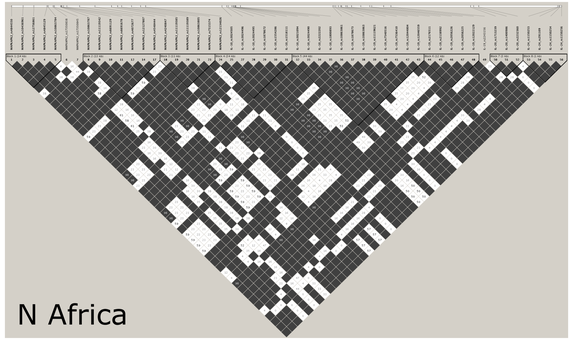
Figure 1: These figures show triangular heat maps generated in Haploview depicting LD patterns across the loci included in this study. Dark shaded boxes indicate pairs of loci in high LD while white boxes indicate that the associated pair of SNPs were not in LD.

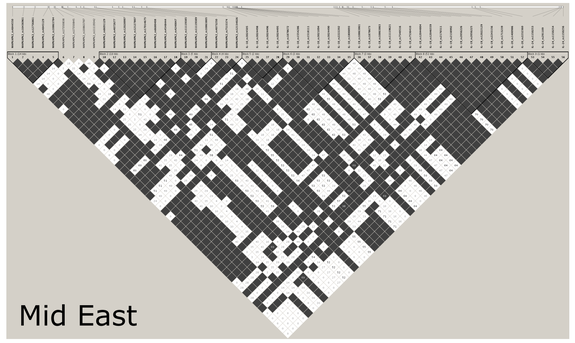


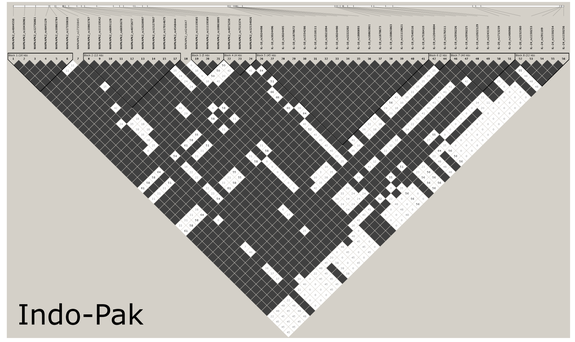


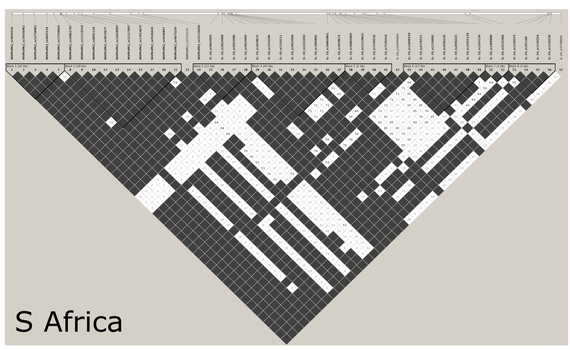


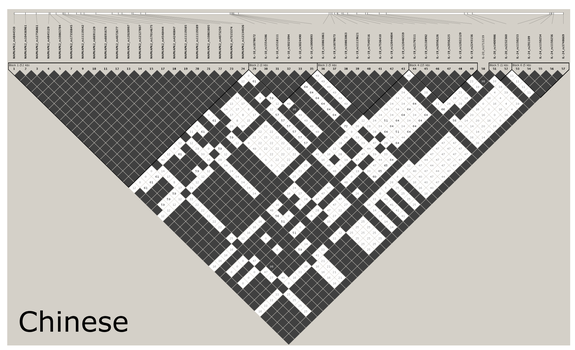


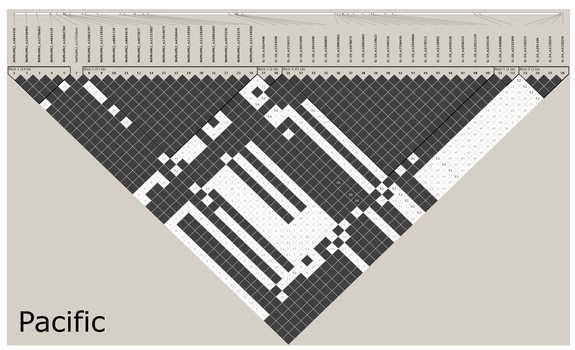












Discussion

Generally, gene boundaries mark haplotype boundaries such that with few exceptions, haplotype block structure is determined by genes which are likely under purifying selection. The LD map including all populations confirms that recombination has occurred in this region on chromosome 1 and that LD is generally preserved within genes. Several populations, namely Pacific, Indo-Pakistan, and, to some extent China, exhibit more cohesive LD patterns. This has been suggested to be the result of serial founder effect (Jakobsson et al., 2006). This is not the case with the Iberian and Basque samples.

Literature Cited: (1) Ardlie KG, Kruglyak L, Seielstad M. 2002. Patterns of Linkage Disequilibrium in the Human Genome. Nat Rev Genet 3:299–309. (2) Barrett JC, Fry B, Maller J, Daly MJ. 2005. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21:263–265. (3) Brocker C, Thompson D, Matsumoto A, Nebert DW, Vasiliou V. 2010. Evolutionary divergence and functions of the human interleukin (IL) gene family. Hum Genomics 5:30. (4) Conrad DF, Jakobsson M, Coop G, Wen X, Wall JD, Rosenberg NA, Pritchard, JK. 2005. A worldwide survey of haplotype variation and linkage disequilibrium in the human genome. Nat Genet 38:1251-1260. (5) Flint-Garcia SA, Thornsberry JM, S E, IV B. 2003. Structure of Linkage Disequilibrium in Plants. Annu Rev Plant Biol 54:357–374. (6) Lewontin R., Kojima K. 1960. Evolutionary Dynamics of Complex Polymorphisms. Evolution 14:458–472. (7) Liu N, Sawyer SL, Mukherjee N, Pakstis AJ, Kidd JR, Kidd KK, Brookes AJ, Zhao H. 2004. Haplotype Block Structures Show Significant Variation among Populations. Genet Epidemiol 27:385–400. (8) Slatkin M. 2008. Linkage disequilibrium — understanding the evolutionary past and mapping the medical future. Nat Rev Genet 9:477–485. (9) UniProt. 2014. MAPKAP-K2 Gene. Univers Protein Resource [Internet]. Available from: http://www.uniprot.org/uniprot/P49137.