Refinement of association signals and residual heritability in host control of HIV viral load

Paul J McLaren[1,2], Istvan Bartha[1,2], Jacques Fellay[1,2] and the International Collaboration for the Genomics of HIV
[1]School of Life Sciences, École Polytechnique Fédérale de Lausanne, [2]Swiss Institute of Bioinformatics

Summary

Rate of HIV disease progression is related to host viral load. Host genetics are known to play a role in establishing patient viral load.

Through the International Collaboration for the Genomics of HIV, we have obtained genome-wide SNP data on 6,315 HIV+ individuals of European ancestry.

After imputation we tested >8M SNPs for association with viral load and observed peaks of association on chromosome 6 and chromosome 3 (Figure 1).

Fine mapping of both signals demonstrated multiple, independent variants at associated loci (Figures 2 and 3).

Based on genotyped SNP data, we estimate the narrow-sense heritability of viral load to be 14% (3% SD). Removing associated regions we observed residual heritability of 4% (3% SD, Figure 4).

By combining multiple data sets of HIV viral load we identified previously unknown, independent signals at associated loci and evidence for limited residual heritability not explained by known factors.

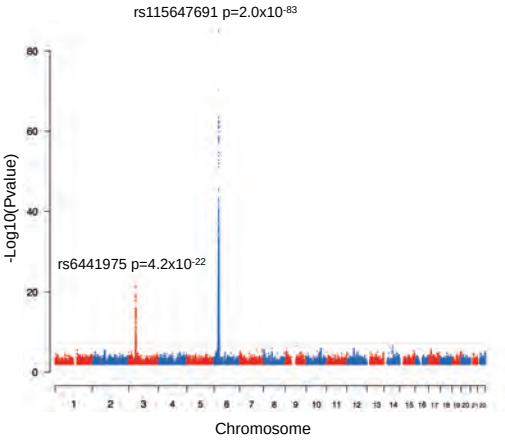
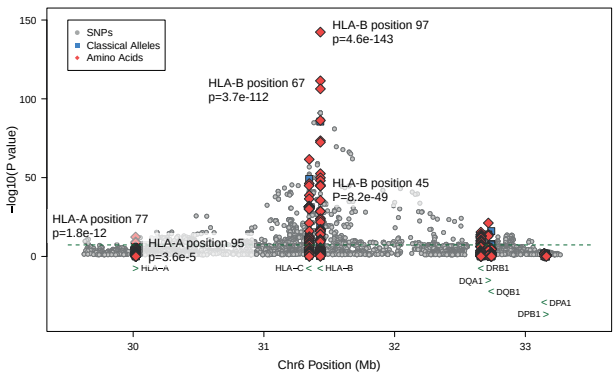


Figure 1: SNP association results. We tested >8M SNPs for association with viral load in 6,315 HIV+ patients. We observed genome-wide significant associations on chromosomes 3 and 6.

Results



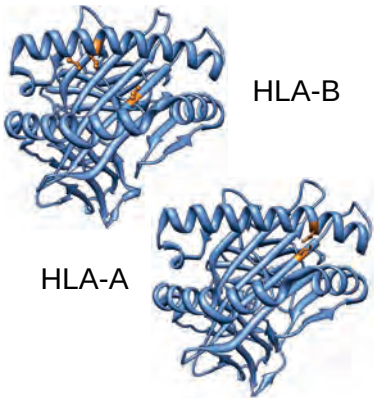
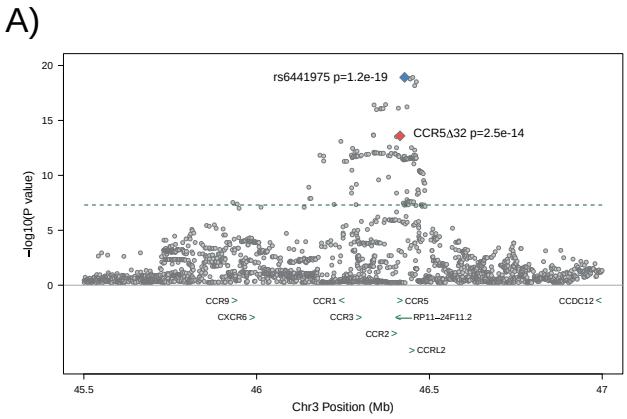


Figure 2: Chromosome 6 association results. A) We imputed classical HLA alleles (blue) and amino acids (red) in HLA genes and tested these for association along with SNPs (grey). We identified 5 independent amino acids (3 in HLA-B; 2 in HLA-A) using forward conditional regression. B) Independently associated amino acid positions (orange) line the peptide binding groove, implicating this biological process as the key determinate of HIV viral load.



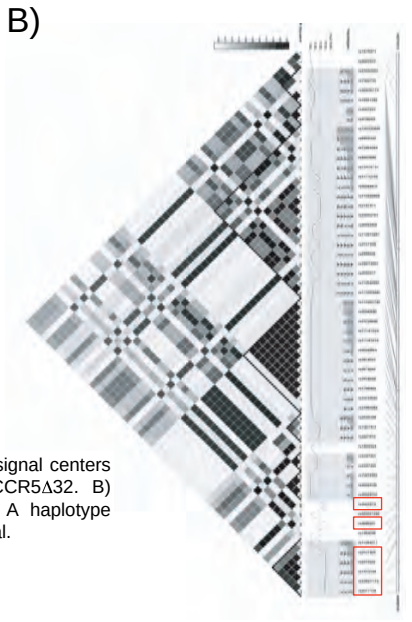


Figure 3 Chromosome 3 association results. A) The chr3 association signal centers on *CCR5*. The top SNP partially tags the known causal variant CCR5Δ32. B) Conditioned on CCR5Δ32 57 SNPs remain genome-wide significant. A haplotype containing 7 regulatory SNPs (red boxes) is the strongest association signal.

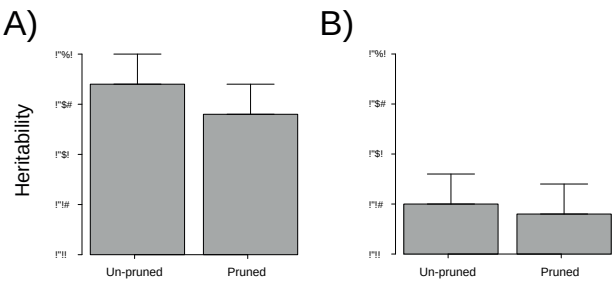


Figure 4: Estimates of narrow sense heritability of viral load. A) Heritability was calculated on all genotyped SNPs or pruned by LD (R^2 < 0.001) resulting in estimates of 17% and 14% (3% SD). B) After removing the associated regions, residual heritability is detected (5% and 4% with 3% SD).

Methods

Genome-wide genotype data were collected from 25 cohort studies and clinical centers participating in the International collaboration for the genomics of HIV.

Genotype imputation was performed using the 1000 Genomes project reference panel. Association was tested per cohort using linear regression followed by meta-analysis.

HLA allele and amino acid types were imputed using the SNP2HLA pipeline (Jia et al *PLoS One* 2013). Independent associations were determined by forward conditional regression.

Narrow-sense heritability was calculated on genotyped SNPs using GCTA (Yang et al *AJHG* 2011).