

Paper Reference

Guryev et al. Haplotype Block Structure Is Conserved across Mammals. *PLoS Genetics*, July 2006.

Abstract

Haplotypes are blocks composed of correlated single-nucleotide polymorphisms (SNPs). In an attempt to understand the forces that drive the formation and preservation of haplotype structure, the authors examine similarities in haplotype structures when comparing mouse, rat, and human genomes. Due to the lack of genome-wide high-resolution genotyping data for all 3 species, a limited region of 5 Mb on mouse chromosome 1 was selected to ensure that a dense haplotype map can be constructed. The authors used pair-wise comparisons to compare haplotype block partitioning for the 3 different species. They found that haplotype organization is significantly conserved across these species. The 3 potential forces driving haplotype structure conservation are recombination rate conservation, reduced recombination in certain regions, and selection on large regions of the genome. Evidence show that contributions of the first 2 forces are unlikely while the effect of selection on haplotype conservation is significant. Because selective pressure is sparsely observed in mammalian genomes, the authors reason that only selective sweeps, background selection, and haplotype-driven selection are possible mechanisms responsible for the mild selection observed. Analysis show that LD values peak in promoter and polymorphic positions within genic regions. This supports the view that advantageous combinations of specific alleles (haplotype-driven selection) shape haplotype conservation across mammalian species.

Discussion

Previously, studying the effect of selection on haplotype structure has not been very successful due to the lack of loci under strong selective pressure. To accomplish this goal, comparative genomics is effectively used in this paper to compare different mammalian genomes. Since human, mouse, and rat have significantly different haplotype block sizes and have distinct polymorphic positions, the significant level of haplotype structure conservation observed across the species indicate selection as a powerful mechanism. The authors argue against the conservation of recombination rate as a major aspect of selection since recombination hotspots evolve at a fast rate and not every hotspot results in a haplotype boundary. The authors also refute the contribution of lower recombination in certain regions because their analysis shows that genic regions in fact have higher recombination rates than intergenic. The remaining route through which selection can manifest, haplotype-driven selection, is then supported by evidence illustrating tighter linkages between promoter regions and evenly spaced genic regions than between intergenic regions. Previous studies have suggested the possibility of higher LD near and within genes, but did not provide quantitative analysis. Here, the authors went further and applied a gene-centered approach to human HapMap data to study LD decay profiles in genes and their flanking regions. They found that elevated LD values within and near genic regions are due to peaks of very high LD's as opposed to uniformly higher LD's across the entire region. Also, the LD is much higher in the 5' than 3' flanking regions of genic regions. This leads to their conclusion that although individual allelic variations are not consequential, the combination of alleles can provide synergy via interacting residues in proteins or via gene regulation. Finally, there are advantages and disadvantages associated with conserved haplotype structures. A disadvantage is that we can no longer reliably pinpoint disease-causing phenotypes to one SNP. The major advantage and future implication is that we can use conserved blocks to identify highly-selected, functionally important genomic regions such as promoters and enhancers.