Distribution of Lyme disease spirochete cp32 prophages and natural diversity among their lipoprotein-encoding erp loci

Abstract:

Lyme disease spirochetes possess complex genomes, consisting of a main chromosome and 20 or more smaller replicons. Among those small DNAs are the cp32 elements, a family of prophages that replicate as circular episomes. All complete cp32s contain an erp locus, which encodes surface-exposed proteins. Sequences were compared for all 193 erp alleles carried by 22 different strains of Lyme disease-causing spirochete to investigate their natural diversity and evolutionary histories. These included multiple iso- lates from a focus where Lyme disease is endemic in the northeastern United States and isolates from across North America and Europe. Bacteria were derived from diseased humans and from vector ticks and included members of 5 different Borrelia geno- species. All erp operon 5=-noncoding regions were found to be highly conserved, as were the initial 70 to 80 bp of all erp open reading frames, traits indicative of a common evolutionary origin. However, the majority of the protein-coding regions are highly diverse, due to numerous intra- and intergenic recombination events. Most erp alleles are chimeras derived from se- quences of closely related and distantly related erp sequences and from unknown origins. Since known functions of Erp surface proteins involve interactions with various host tissue components, this diversity may reflect both their multiple functions and the abilities of Lyme disease-causing spirochetes to successfully infect a wide variety of vertebrate host species.