

P033 Significance of the distribution of chromosomal CRISPR matches to viruses and plasmids in the crenarchaeal order Sulfolobales

Shiraz Ali Shah and Roger A. Garrett

University of Copenhagen

Clusters of regularly inter-spaced palindromic repeats (CRISPR) are found within large intergenic genomic regions in almost all archaea and in many of the bacteria sequenced to date. CRISPRs consist of unique spacer sequences separating the repeats, and sequence analyses have suggested that at least some spacers derive from virus or plasmid genomes. CRISPRs, together with the physically linked genes of Cas proteins, are considered provide the cell with an adaptive immune-response against invading viruses and plasmids via an unknown RNA-based mechanism. Here we compare spacer sequences from all the sequenced genomes of acidothermophiles, of the order Sulfolobales, with the many sequenced viruses and plasmids of this order, at both a nucleotide and an amino acid sequence level. More than 1000 matching spacers were identified from ~3000 spacers analysed, and they were mapped on the viral and plasmid genomes. Here, we present and interpret the biases of the matches, to selected viruses and plasmids, with respect to: (a) DNA strand, (b) coding and non coding genome regions; (c) coding and non coding DNA strands, and (d) conserved versus less conserved genes. Finally, the CRISPRs are divided into families based on sequence analyses of their flanking sequences and we use this information to investigate whether any biases are inherent to specific CRISPR families.