

Evolution of gene order conservation in prokaryotes

ABSTRACT

Gene order conservation is a genomic measure that can be useful for studying relationships between prokaryotes and the evolutionary forces shaping their genomes. Gene organization is extensively conserved in some genomic regions, and further studies are needed to elucidate the reason for this conservation.

INTRODUCTION

Background Completely sequenced genomes enable the study of relations between organisms in terms of the complete set of genes they possess. Genomic properties have been proposed as the most convenient tool for studying these relationships, as they are global properties that may circumvent many of the difficulties of classical molecular phylogenies. Common gene content or conservation of families of proteins are examples of this kind of genomic information. From this genomic perspective, conservation of gene order is a very informative measure that may provide information both about the function and interactions of the proteins these genes encode, and about the evolution of the genomes and the organisms themselves. Gene order is generally well preserved at close phylogenetic distances. When the species are not closely related, the degree of gene order conservation is usually low, and consequently it was proposed that conservation of gene order is easily lost during evolution. This loss also extends to the disruption of operons, in some cases wiping them out completely. Nevertheless, some instances of especially well-preserved clusters of genes are known, even in divergent species. The best examples are the genes for ribosomal proteins and the *dcw* cluster. Lathe and co-workers recently identified genomic regions in which gene order is especially highly conserved. Even if some rearrangement does occur in these regions, the general trend is to keep the genes closer together than in other regions. This shows that selection for gene location and ordering could exist in some cases. The operon structure and common regulation cannot easily account for the conservation, as these conserved regions extend for more than a single operon; hence the proposed nomenclature of uber-operons. Conservation of gene order can be due to any one of the following three reasons. First, the species have diverged only recently and gene order has not yet been destroyed; second, there has been lateral gene transfer of a block of genes; and third, the integrity of the cluster is important to the fitness of the cell. Only in this latter case is gene order conservation selectable. Proposed explanations for selection for gene ordering include helping the interaction of proteins encoded by the genes of the cluster, favoring lateral gene transfer, or co-localization of the mRNAs in the same region of the cell. These explanations are not mutually exclusive. Recent studies of the structure of the *dcw* cluster suggest that, in this particular case, conservation of gene order within the cluster may be linked to cellular morphology, thus connecting gene order with a selectable phenotype. The importance of gene order in the study of evolution is starting to be recognized. Even if the loss of gene order conservation is faster than the loss of sequence similarity, a large amount of conservation remains at medium phylogenetic distances, such as that between *Escherichia coli* and *Bacillus subtilis*. Conservation is a valuable clue to the relationships between organisms and the influence of events such as lateral gene transfer on the evolution of genomes. I present here an analysis of the extent and characteristics of gene order conservation in prokaryotes and attempt to answer two questions. Does conservation of

gene order occur similarly throughout the prokaryotes? Are the conserved regions distributed uniformly within the genomes?

CONCLUSION

Conclusions Gene order is a labile genomic characteristic. The level of conservation is high when organisms are phylogenetically closely related, but conservation is lost rapidly, probably to a higher degree than other genetic or genomic features. Thus, the instances in which gene order is conserved between phylogenetically distant organisms may indicate that strong selection pressures are keeping them together, in the cases in which lateral gene transfer is unlikely to be the origin of the conservation. Selection could be because the operon controls the assembly of a multifunctional enzymatic complex or the performance of an important stage in a metabolic pathway. But in some cases, other explanations should be considered, in which the gene order could influence the phenotype. The existence of conserved units bigger than operons seems to argue in favor of other explanations. Gene order conservation can be valuable for establishing the relationships between organisms as it is not influenced by parameters that affect other genomic measures, such as the content of unique genes, that are ultimately dependent on the lifestyle of the species. Genomic properties have been proposed as alternatives to classical molecular phylogenies as they measure global features of the genomes. So far, no genomic property by itself can represent that alternative, and integration of information on different properties is desirable. In this perspective, the information offered by gene order conservation is crucial.