Contribution of mobile genetic elements on variation in gene content among bacteria

Haruo Suzuki\(^1\)  
hs568@cornell.edu  
Michael J Stanhope\(^1\)  
mjs297@cornell.edu

\(^1\) Department of Population Medicine and Diagnostic Sciences, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA

Keywords: bacteria, gene content, mobile genetic element, plasmid, virus

1   Introduction

Bacterial genomes vary widely in gene content [1]. This variation can be the result of different evolutionary events including gene loss, gene duplication, and horizontal gene transfer [2]. Mobile genetic elements such as plasmids and viruses are major vehicles to transfer genes between different cells, and thus play a key role in the evolution and diversification of bacteria [3]. To understand the importance of the mobile genetic elements in bacterial evolution, we investigated contributions of plasmids and viruses on variation in gene content among bacteria.

2   Method

All analyses were conducted using the G-language Genome Analysis Environment version 1.8.9 [4] and R version 2.9.2 [5]. Complete sequences of bacterial chromosomes, plasmids, and viruses were downloaded in GenBank format from the National Center for Biotechnology Information site [6] in September 2009. We performed homology searches using BLASTN [7] with an E-value threshold of 1e-20 for all nucleotide sequences from a bacterial chromosome as query data, against nucleotide databases of plasmids and viruses. Contributions of plasmids and viruses on variation in gene content among bacteria were measured by Pearson product-moment correlation coefficients (r) of the total number of genes in the bacterial chromosomes with the number of the chromosomal genes found on plasmids and viruses.

3   Results and Discussions

We searched plasmid- and virus-encoded genes in the chromosomes of the 205 bacterial strains from ten genera: *Bacillus* (23), *Clostridium* (21), *Escherichia* (24), *Mycobacterium* (20), *Mycoplasma* (14), *Pseudomonas* (16), *Salmonella* (16), *Staphylococcus* (19), *Streptococcus* (40), and *Yersinia* (12). For the 205 bacterial strains, the total number of genes in the chromosomes ranged from 475 to 6716 (Figure 1-A). The numbers of the chromosomal genes found on plasmids (Figure 1-B) and viruses (Figure 1-C) varied widely among different genera, and often also among strains within the same genus. Genes in the chromosomes of *Mycobacterium*, *Pseudomonas*, and *Yersinia* were found more often on plasmids, while those in the chromosomes of *Staphylococcus* and *Streptococcus* were found more often on viruses. Few genes in the chromosomes of *Mycoplasma* were found on both plasmids and viruses. For the 205 bacterial strains, the total number of chromosomal genes was correlated with the number of genes found on plasmids (Figure 2-A; \(r = 0.78\)) in contrast to viruses (Figure 2-B; \(r = 0.15\)). These results indicate that plasmid-chromosome gene transfers have occurred more often in bacteria with large chromosomes than in bacteria with small chromosomes. The contributions of viruses on chromosomal gene contents were observed in *Escherichia* (\(r = 0.75\)) and *Staphylococcus* (\(r = 0.81\)). Our results suggest that the relative contribution of plasmids and viruses on chromosomal gene content varies among different bacteria.
Figure 1: Box-and-whisker plots showing the distributions of the total number of genes in the bacterial chromosomes (A) and the number of the chromosomal genes found on plasmids (B) and viruses (C).

Figure 2: Scatter plot showing the total number of genes in the bacterial chromosomes, plotted against the number of the chromosomal genes found on plasmids (A) and viruses (B).

References


