Seeking Genomic Duplication in Prokaryotic Genomes by Equivalent Tandem Repeats Distributions

Satoshi Mizuta
slmizu@cc.hirosaki-u.ac.jp

Michimasa Koshino
s402069@si.hirosaki-u.ac.jp

Toshio Shimizu
slsimi@si.hirosaki-u.ac.jp

1 Faculty of Science and Technology, Hirosaki University, Hirosaki 036-8561, Japan

Keywords: tandem repeat, whole-genome duplication, prokaryote, bacteria, Neisseria meningitidis

1 Introduction

Gene duplication and large-scale segmental duplication of genome including whole-genome duplication (WGD) are thought to be driving forces of evolution. So far, it has been revealed by recent analysis that WGD had occurred on many eukaryotic genomes, for instance, maize, Arabidopsis, yeast, and vertebrate, even more than once[4]. Although it is not clear whether or not WGD had occurred on prokaryotic genomes[1, 5], there seems not to exist an apparent reason that WGD never occurs on prokaryotic genomes; it is rather natural to think that some prokaryotes had experienced WGD.

The standard manner to investigate genome evolution that is used in recent researches is analyzing homologous genes on the genome. In this study, however, we focused attention on tandem repeats (TRs) instead of genes and analyses equivalent TR-pairs (or, say, repeats of TRs), because TRs are rather free from selection pressure if they are located on inter-genic regions, and thus they might preserve signs of WGD more measurably than genes.

2 Materials and Method

Genome sequences of 44 prokaryotic species were downloaded from GenBank[7] and tandem repeats (TRs) were searched by color coding (CC) method[6]. All the detected TRs were aligned pairwisely by SSEARCH[8] with the default parameter settings, match (= 5), mismatch (= −4), gap creation penalty (= −16), and gap extension penalty (= −4), and the threshold E-value = 10−3. A pair of TRs was identified as an equivalent TR-pair, if the proportion of the overlap to the longer sequence was larger than 50%. The results of detailed analysis of the detected tandem repeats for all the 44 genomes are described elsewhere[2, 3].

In the previous study[3], we derived indications of the occurrences of WGD in several prokaryotic genomes from the location patterns of the equivalent TR-pairs on circles which schematically describe the prokaryotic genomes. In this study, we concentrate on Neisseria meningitidis genome, which is one of the genomes that shows most salient features among the 44 species, in order to investigate the indications of WGD more precisely.

3 Results and Discussions

Figure 1 shows the dot-matrix plots of the equivalent TR-pairs detected on N. meningitidis genome. Each dot represents an equivalent TR-pair, and the positions of each members of an equivalent TR-pair on the genome are presented by x- and y-axis, respectively. The left diagram plots the actual
Figure 1: Dot-matrix plots of equivalent TR-pairs: left for actual data and right for randomly relocated TR-pairs.

positions and the right one plots the positions of randomly relocated TRs preserving the equivalence relations among them.

Comparing the two diagrams, we can definitely recognize some periodic patterns on the left diagram, which possibly suggest the existence of WGD on the genome. Statistical analyses of the patterns will give more reliable conclusions.

References


