

SORTING GENOMES BY TRANSLOCATIONS AND DELETIONS

Xingqin Qi

Department of Applied Mathematics, Shandong University at Weihai,
Weihai, 264213, China

School of Mathematics and System Sciences, Shandong University,
Jinan, 250100, China

Guojun Li

School of Mathematics and System Sciences, Shandong University,
Jinan, 250100, China

Department of Biochemistry and Molecular Biology, University of Georgia,
Athens, Georgia 30602, USA

Shuguang Li

School of Mathematics and System Sciences, Shandong University,
Jinan, 250100, China

Department of Mathematics and Information Science, Yantai University,
Yantai, 264005, China

Ying Xu

Department of Biochemistry and Molecular Biology, University of Georgia,
Athens, Georgia 30602, USA

Given two signed multi-chromosomal genomes Π and Γ with the same gene set, the problem of sorting by translocations (SBT) is to find a shortest sequence of translocations transforming Π to Γ , where the length of the sequence is called the translocation distance between Π and Γ . In 1996, Hannenhalli gave the formula of the translocation distance for the first time, based on which an $O(n^3)$ algorithm for SBT was given. In 2005, Anne Bergeron et al. revisited this problem and gave an elementary proof for the formula of the translocation distance which leads to a new $O(n^3)$ algorithm for SBT. In this paper, we show how to extend Anne Bergeron's algorithm for SBT to include deletions, which allows us to compare genomes containing different genes. We present an asymptotically optimal algorithm for transforming Π to Γ by translocations and deletions, providing a feasible sequence with length at most $OPT + 2$, where OPT is the minimum number of translocations and deletions transforming Π to Γ . Furthermore, this analysis can be used to approximate the minimum number of translocations and insertions transforming one genome to another.