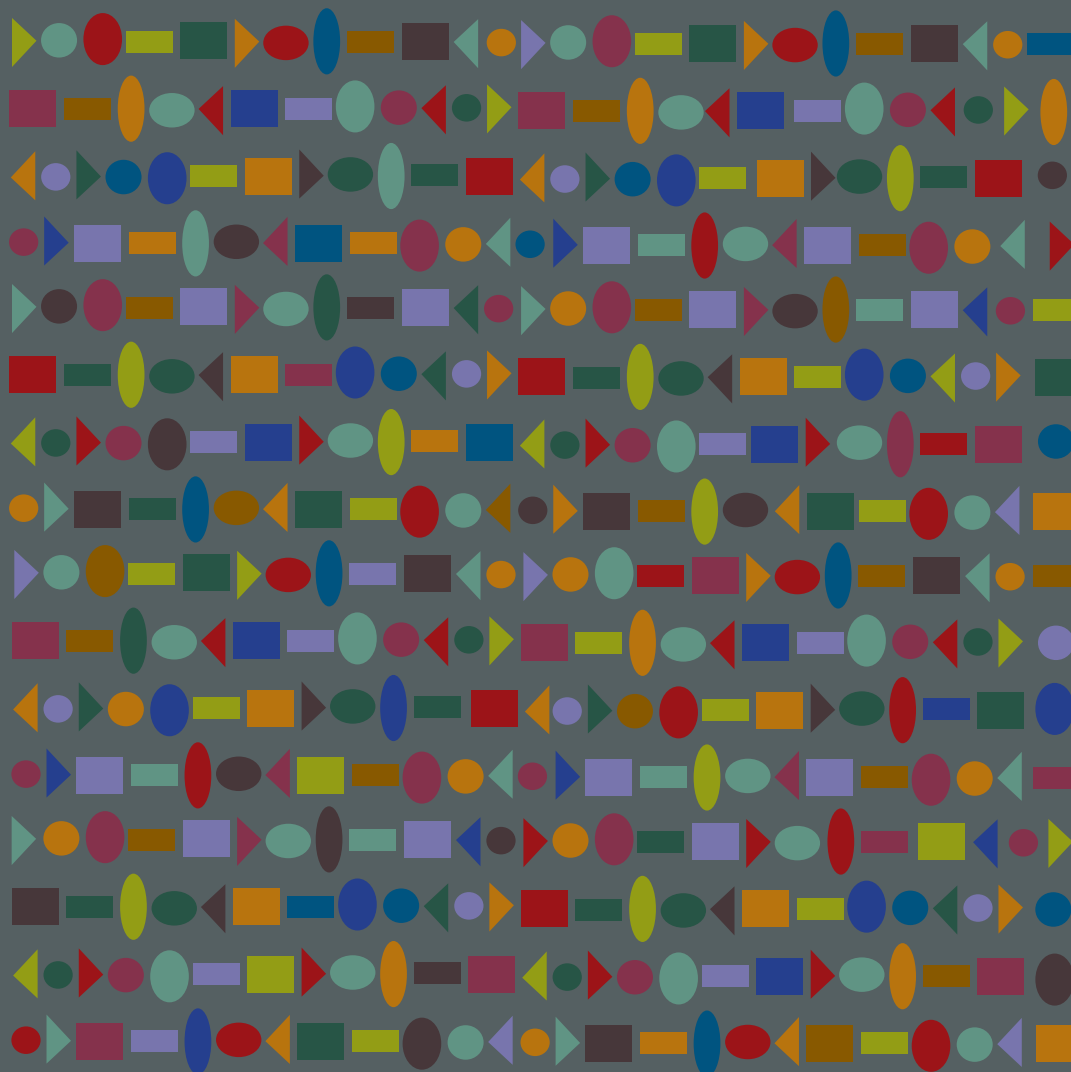


COMBINATORICS OF GENOME REARRANGEMENTS

Guillaume Fertin, Anthony Labarre, Irena Rusu, Éric Tannier, and Stéphane Vialette



Combinatorics of Genome Rearrangements



Computational Molecular Biology

Sorin Istrail, Pavel Pevzner, and Michael Waterman, editors

Computational molecular biology is a new discipline, bringing together computational, statistical, experimental, and technological methods, which is energizing and dramatically accelerating the discovery of new technologies and tools for molecular biology. The MIT Press series on Computational Molecular Biology is intended to provide a unique and effective venue for the rapid publication of monographs, textbooks, edited collections, reference works, and lecture notes of the highest quality.

Computational Molecular Biology: An Algorithmic Approach

Pavel A. Pevzner, 2000

Computational Methods for Modeling Biochemical Networks

James M. Bower and Hamid Bolouri, editors, 2001

Current Topics in Computational Molecular Biology

Tao Jiang, Ying Xu, and Michael Q. Zhang, editors, 2002

Gene Regulation and Metabolism: Postgenomic Computation Approaches

Julio Collado-Vides, editor, 2002

Microarrays for an Integrative Genomics

Isaac S. Kohane, Alvin Kho, and Atul J. Butte, 2002

Kernel Methods in Computational Biology

Bernhard Schölkopf, Koji Tsuda, and Jean-Philippe Vert, editors, 2004

An Introduction to Bioinformatics Algorithms

Neil C. Jones and Pavel A. Pevzner, 2004

Immunological Bioinformatics

Ole Lund, Morten Nielsen, Claus Lundegaard, Can Keşmir, and Søren Brunak, 2005

Ontologies for Bioinformatics

Kenneth Baclawski and Tianhua Niu, 2005

Biological Modeling and Simulation: A Survey of Practical Models, Algorithms, and Numerical Methods

Russell Schwartz, 2008

Combinatorics of Genome Rearrangements

Guillaume Fertin, Anthony Labarre, Irena Rusu, Éric Tannier, and Stéphane Vialette, 2009

Combinatorics of Genome Rearrangements

Guillaume Fertin, Anthony Labarre, Irena Rusu, Éric Tannier and Stéphane Vialette

**The MIT Press
Cambridge, Massachusetts
London, England**

© 2009 Massachusetts Institute of Technology

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

For information about special quantity discounts, please email special_sales@mitpress.mit.edu

This book was set in Times New Roman and Syntax on 3B2 by Asco Typesetters, Hong Kong.
Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Combinatorics of genome rearrangements / Guillaume Fertin . . . [et al.].

p. cm. — (Computational molecular biology)

Includes bibliographical references and index.

ISBN 978-0-262-06282-4 (hardcover : alk. paper) 1. Translocation (Genetics)—Mathematical models.
2. Translocation (Genetics)—Data processing. 3. Combinatorial analysis. 4. Genomics—Mathematics.

I. Fertin, Guillaume, 1972– II. Series.

[DNLM: 1. Gene Rearrangement. 2. Genome. 3. Models, Genetic. QU 470 C731 2009]

QH462.T7C66 2009

572.8'77—dc22

2008042152

10 9 8 7 6 5 4 3 2 1

Contents

Preface	xiii
Acknowledgments	xv

1	Introduction	1
1.1	A Minimalist Introduction to Molecular Evolution	1
1.2	Birth of the Combinatorics of Genome Rearrangements	4
1.3	Statement of the Problem	6
1.4	Scope of This Survey	7
1.5	Overview of the Models	7
1.6	Organization of the Book	8
I	DUPLICATION-FREE MODELS: PERMUTATIONS	11
2	Genomes as Permutations	13
2.1	The Symmetric Group	13
2.2	The Cycles of a Permutation	14
2.3	Signed Permutations	15
2.4	Distances on Permutation Groups	15
2.4.1	Rearrangements as Generators	16
2.4.2	Invariant Distances	17
2.5	Circular Permutations	18
2.5.1	Classical Circular Permutations	19
2.5.2	Genomic Circular Permutations	19
2.6	First Measures of Similarity between Permutations	20
2.6.1	Breakpoints	20
2.6.2	Common Intervals and Semipartitive Families	21
3	Distances between Unsigned Permutations	25
3.1	Transposition Distance	25
3.1.1	Lower Bounds on the Transposition Distance	26
3.1.2	Upper Bounds	29
3.1.3	Improving Bounds Using Toric Permutations	32

3.1.4	Easy Cases	33
3.1.5	Approximation Algorithms	34
3.1.6	Conjectures and Open Problems	35
3.2	Prefix Transposition Distance	36
3.2.1	Lower Bounds	37
3.2.2	Upper Bounds	38
3.2.3	Diameter	38
3.2.4	Easy Cases	39
3.2.5	Approximation Algorithms	39
3.2.6	Variant: Insertion of the Leading Element	40
3.3	Reversal Distance	40
3.3.1	Lower Bounds	40
3.3.2	Upper Bounds	43
3.3.3	Easy Cases	43
3.3.4	Computational Complexity	44
3.3.5	Approximation Algorithms	45
3.3.6	Exact Algorithms	46
3.4	Prefix Reversal Distance (Pancake-Flipping)	47
3.4.1	Lower Bounds	47
3.4.2	History	48
3.4.3	Variants	48
3.5	Variants	49
3.5.1	Block Interchange Distance	49
3.5.2	Element Interchange Distances	50
3.5.3	Weighted Reversals	52
3.5.4	Fixed-Length Reversals	54
3.5.5	Bounded Variants	54
3.5.6	Cut-and-Paste	55
3.5.7	Strip Moves	55
3.5.8	Stack-Sorting	56
3.5.9	Tandem Duplications and Random Losses	58
3.5.10	Combined Operations: Reversals and Transpositions	59
3.6	Relations between Distances on Unsigned Permutations	61
4	Distances between Signed Permutations	63
4.1	Conserved Interval Distance	63
4.2	Signed Reversal Distance	64
4.2.1	Reversals	64
4.2.2	The Distance Formula	65
4.2.3	The Scenario of Reversals	67
4.2.4	The Space of All Optimal Solutions	68
4.2.5	Experimental Results	69
4.3	Variants of Sorting by Reversals	69
4.3.1	Perfect Signed Reversal Distance	69
4.3.2	Prefix Reversals (Burned Pancakes)	70
4.3.3	Reversals That Are Symmetric around a Point	70

4.3.4	Weighted Reversals	71
4.3.5	Fixed-Length Reversals	71
4.4	Combined Operations	72
4.4.1	Reversals and Transpositions	72
4.4.2	Reversals, Transpositions, Transreversals, Revrevs	72
4.5	Double Cut-and-Joins	73
5	Rearrangements of Partial Orders	75
5.1	Genomes as Partially Ordered Sets	75
5.2	Partially Ordered Sets	75
5.2.1	Basic Definitions	75
5.2.2	Representing Posets	77
5.2.3	Topological Sorting	77
5.3	Constructing a Poset	78
5.4	Reversal Distance	79
5.5	Breakpoint Distance	80
5.5.1	Exact Algorithms	80
5.5.2	Heuristics for Computing the Breakpoint Distance	81
6	Graph-Theoretic and Linear Algebra Formulations	83
6.1	Simple Permutations and the Interleaving Graph	83
6.2	The Overlap Graph	84
6.3	The Local Complementation of a Graph	85
6.4	The Matrix Tightness Problem	85
6.5	Extension to Sorting by Transpositions	86
6.6	The Intermediate Case of Directed Local Complementation	87
II	MODELS HANDLING DUPLICATIONS: STRINGS	89
7	Generalities	91
7.1	Biological Motivations	91
7.2	Strings and Rearrangements on Strings	92
7.3	Balanced Strings	94
7.4	How to Deal with Multiple Copies?	95
8	Distances between Arbitrary Strings	97
8.1	The Match-and-Prune Model	98
8.1.1	Breakpoint Distance	100
8.1.2	Signed Reversal Distance	106
8.1.3	Adjacency Similarity	108
8.1.4	Common Intervals Similarity	111
8.1.5	Conserved Intervals Similarity	113
8.1.6	Conserved Intervals Distance	114
8.1.7	MAD and SAD Numbers	118
8.1.8	Heuristics	119

8.2	The Block Edit Model	123
8.2.1	Block Covering Distance	123
8.2.2	Symmetric Block Edit Distance	126
8.2.3	Large Block Edit Distance	129
8.2.4	String Edit Distance with Transpositions	130
8.2.5	Signed Strings	131
9	Distances between Balanced Strings	133
9.1	Minimum Common String Partition Problems	133
9.1.1	Unsigned MCSP	134
9.1.2	Signed MCSP	135
9.1.3	Reversed MCSP	137
9.1.4	Full Breakpoint Distance	138
9.2	Reversal Distance	138
9.2.1	Unsigned Reversals	138
9.2.2	Signed Reversals	141
9.2.3	Sorting by Reversals with Length-Weighted Costs	142
9.2.4	Prefix Reversals on Unsigned Strings (Pancake-Flipping)	144
9.2.5	Reversals of Length at Most 2	147
9.3	Unsigned Transpositions	147
9.3.1	Unit Cost Transpositions	147
9.3.2	Length-Weighted Transpositions	150
9.3.3	Restricted Length-Weighted Transpositions	150
9.3.4	Prefix Transpositions	152
9.3.5	Adjacent Swaps	153
9.4	Unsigned Block Interchanges	153
9.4.1	Unit-Cost Block Interchanges	153
9.4.2	Character Swaps	155
9.5	Relations between Distances	157
III	MULTICHROMOSOMAL MODELS	159
10	Paths and Cycles	161
10.1	Genomes	161
10.2	Breakpoints	162
10.3	Intervals	163
10.4	Translocation Distance	164
10.4.1	Feasibility	166
10.4.2	Unsigned Genomes	166
10.4.3	Signed Genomes	167
10.4.4	Translocations Preserving Centromeres	168
10.4.5	Variants and Special Cases	169
10.5	Double Cut-and-Joins (2-Break Rearrangement)	170
10.6	k -Break Rearrangement	171
10.7	Fusions, Fissions, Translocations, and Reversals	172
10.8	Rearrangements with Partially Ordered Chromosomes	174

11	Cycles of a Permutation	175
11.1	A Model for Multichromosomal Circular Genomes	175
11.2	A Generalization to Signed Genomes	178
11.2.1	A Different Kind of Signed Permutation	178
11.2.2	The Operations	179
11.2.3	Some Results	179
12	Set Systems and the Syntenic Distance	181
12.1	Introduction	181
12.2	Structural Properties	182
12.2.1	Compact Representation	182
12.3	Lower Bounds	184
12.4	Diameter	185
12.5	Algorithmic Results	185
12.5.1	Syntenic Distance	185
12.5.2	Easy Cases	186
12.6	Conjectures and Open Problems	189
IV	MULTIGENOMIC MODELS	191
13	Median and Halving Problems	193
13.1	Breakpoint Median	194
13.1.1	Complexity	194
13.1.2	Algorithms	195
13.2	Reversal and DCJ Median	197
13.2.1	Complexity	197
13.2.2	Algorithms	197
13.2.3	Variants	198
13.3	Duplicated Genomes	199
13.3.1	The Double Distance	199
13.3.2	Genome Halving	201
13.3.3	Solving Tetraploidy	202
13.3.4	Guided Halving	202
13.3.5	Genome Halving with Unordered Chromosomes	203
13.4	Other Variants, Generalizations, and Discussion	205
13.4.1	Other Operations	205
13.4.2	More Permutations in the Input	205
13.4.3	Medians and Centers	205
13.4.4	Discussion	206
14	Rearrangement Phylogenies	207
14.1	The Large Parsimony Problem	207
14.2	The Large Parsimony Problem with Gene Orders	209
14.2.1	Breakpoint and Reversal Phylogenies on Permutations	209
14.2.2	Variants	211

14.3	Heuristics for the Breakpoint/Reversal Phylogeny Problem	211
14.3.1	Tree Steinerization	212
14.3.2	Sequential Addition	216
14.3.3	Character Encodings	217
14.4	Variants	220
V	MISCELLANEOUS	221
15	Software	223
15.1	Pairwise Rearrangements	223
15.1.1	Unichromosomal Models	223
15.1.2	Multichromosomal Models	225
15.2	Phylogeny Reconstruction and Medians	226
15.2.1	BPAanalysis	226
15.2.2	MGR	226
15.2.3	GRIL	226
15.2.4	GRAPPA	227
15.2.5	MedRbyLS	227
15.2.6	rEvoluzer and amGRP	227
15.2.7	GENESIS	228
16	Open Problems	229
16.1	Complexity Issues	229
16.1.1	Hardness	229
16.1.2	Approximability	230
16.1.3	Polynomial Complexity	231
16.2	Diameter	231
16.3	Tightness of Bounds	232
	APPENDICES	233
A	Graph Theory	235
A.1	Undirected Graphs	235
A.1.1	Basic Definitions	235
A.1.2	Paths and Cycles	236
A.1.3	Connectivity	237
A.1.4	Bipartite Graphs	238
A.1.5	Trees and Forests	238
A.1.6	Matching	238
A.1.7	Adjacency Matrix	239
A.2	Directed Graphs	240
A.2.1	Basic Definitions	240
A.2.2	Paths and Cycles	241
A.2.3	Connectivity	241
A.2.4	Directed Acyclic Graphs	241

B Complexity Theory 243

- B.1 The Class **NP** 243
 - B.1.1 **NP**-Optimization Problems: From **PTAS** to **APX** 246
 - B.1.2 **NP**-Optimization Problems: Beyond **APX** 250
 - B.1.3 Parameterized Complexity 250
- B.2 Some **NP**-Complete Problems 252

Glossary 257

Bibliography 263

Index 283

Preface

In 1984, at a congress in Paris, François Jacob, one of the most famous evolutionary scientists, stated that “La molécule de l’hérédité est raboutée, modifiée, coupée, rallongée, raccourcie, retournée” (the molecule of heredity is sewed together, modified, cut, lengthened, shortened, reversed) during evolution. From one cell to another, one individual to another, one species to another, the content of the DNA molecules is often similar, but their organization often differs dramatically. The mutations that affect this organization are called *genome rearrangements*, and the structural differences between molecules in two genomes motivate the study of their combinatorics. Indeed, the inference of the evolutionary events that can explain the multiple combinations of observed genomes can often be formalized as combinatorial optimization problems.

The variety of problems that have been raised in this domain is so interesting from a combinatorial point of view that this field has grown and become partly independent of the application, so that it now belongs as much to mathematics as to biology. The mathematics and algorithmics related to genome rearrangements have witnessed a huge expansion over these last years, and this dynamics seems to be continuing at the present time. Due to this success, the field has swallowed other studies that were developed earlier and without biological motivations. For example, many problems about sorting permutations with constraints are now presented as rearrangement problems, without considering the biological relevance of the constraint.

Although molecular biology gave birth to it, combinatorics of genome rearrangements is now a mathematical and algorithmic field that has found its own coherence. It has its own important results, many peripheral developments, and its famous open problems. A great interest of this domain is the simplicity of the formulation of the problems, compared to the sometimes great complexity or even nonexistence of solutions. Moreover, the fact that the subject has now been studied for nearly two decades and has been discussed only in specialized research literature motivates both a thorough survey of the topic and an introduction to a broader audience. This book intends to fulfill both goals.

Acknowledgments

S verine B rard, Ga lle Giberti, and Julien Moncel participated in the birth of the project that led to the present book. The decision to undertake this work was made during a meeting in Lyon funded by the ACI “Nouvelles Interfaces des Math matiques,” π -vert, led by Marie-France Sagot.

Anthony Labarre was funded by the Fonds pour la Formation   la Recherche dans l’Industrie et dans l’Agriculture (FRIA), by a grant from the Fonds National de la Recherche Scientifique (FNRS), and by Communaut  Franaise de Belgique—Actions de Recherche Concert es. He wishes to thank his family and friends, as well as his adviser Jean-Paul Doignon. Eric Tannier was funded by the French National Research Institute for Computer Science (INRIA) and by the Agence Nationale de la Recherche (ANR), projects JC05_49162 (GENOMICRO) and NT05-3_45205 (REGLIS).

1 Introduction

Although this book is combinatorially inclined and does not devote much discussion to the biological issues, we will start with a short introduction to molecular evolution, for conceptual and historical purposes; indeed, this is where the combinatorial study of genome rearrangements originates, and the invention of most variants of genome rearrangement problems are still driven by biological constraints. This introduction is not necessary to understand the combinatorial problems and their solutions, but it allows us to place them in their context and explain why they are important, independently of their mathematical value.

1.1 A Minimalist Introduction to Molecular Evolution

The “molecules of heredity,” the support of genetic information, are present in every cell of all living organisms (bacteria, plants, animals, etc.). Each molecule is called a *chromosome*, and the set of all chromosomes is what we will call the *genome*. Chromosomes are made of *DNA* (**d**eoxy**r**ibonucleic **a**cid), a double-stranded molecule in which each *strand* is a long succession of *nucleotides* (a *sequence*). Nucleotides can be of four types—*A*, *C*, *G*, and *T*—and the two DNA strands are coupled in such a way that an *A* on one strand is always coupled with a *T* on the other strand, and a *C* on one strand is always coupled with a *G* on the other strand. Those strands are said to be *complementary*: the sequence on one strand determines the sequence on the other one. Figure 1.1 shows a representation of the above concepts.

Because of complementarity, a DNA molecule is usually represented as a single sequence (one arbitrary strand), but the organization in two strands will often be crucial for our purpose. Chromosomes can be either circular (the sequence forms a circle and has no ends), which is often the case in bacteria, or linear (the sequence has two ends, called *telomeres*), which is often the case in animals and plants. A *segment* of DNA is a part of this molecule made of consecutive nucleotides. A *gene* is a segment of DNA that contains the information needed to construct the other molecules in the cell.

Chromosome

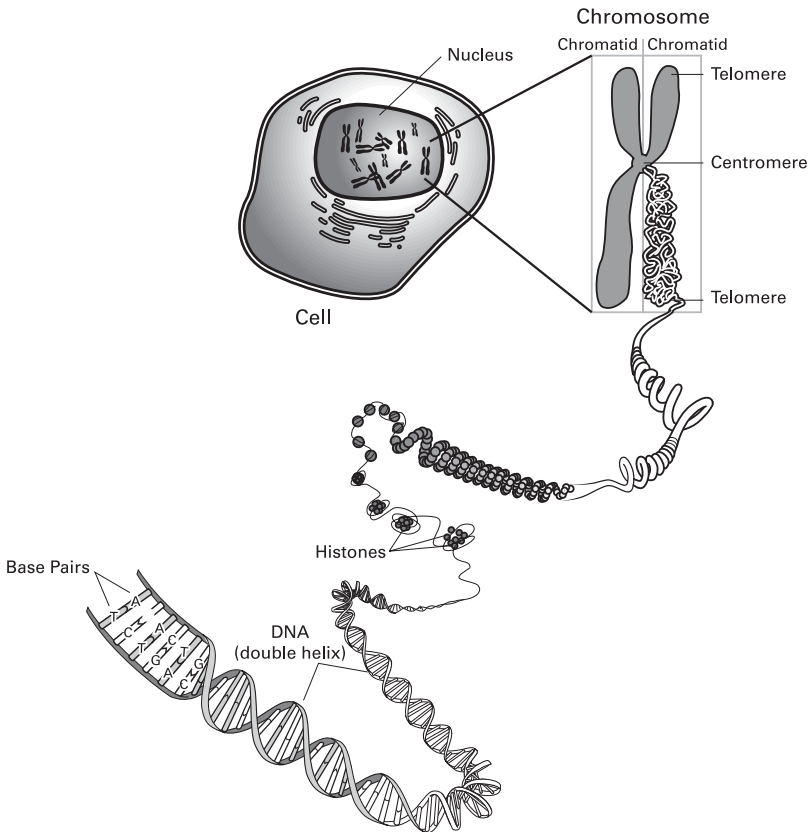


Figure 1.1

A chromosome and a fragment of a DNA molecule

Source: National Institutes of Health, National Human Genome Research Institute, Division of Intramural Research

What accounts for the diversity of living organisms is the possibility for DNA to *replicate* itself with some inaccuracy: one genome is used to produce another, almost identical genome. This inaccuracy is the principle of molecular *evolution*.

A DNA molecule may evolve by *point mutations* (i.e., mutations at the level of nucleotides). There are three different kinds of point mutations: substitutions (one nucleotide is replaced with another), insertions (a nucleotide is added to the sequence), and deletions (a nucleotide is removed from the sequence). Detecting these events is the goal of *sequence alignment* (for a presentation of this topic, see, for example, Setubal and Meidanis [333] or Jones and Pevzner [224]).

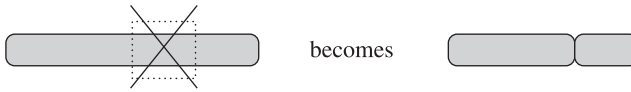


Figure 1.2
Deletion of the dotted region in a chromosome



Figure 1.3
Transposition of the dotted region in a chromosome

CCGTGCGTACACTGC becomes CCGTGTACGCACTGC

Figure 1.4
Reversal of the underlined segment, resulting in the boxed segment

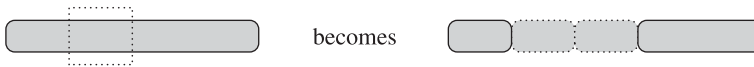
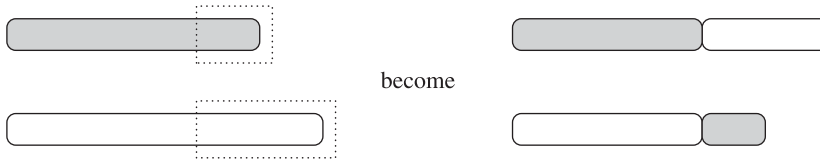


Figure 1.5
Tandem duplication of the dotted region in a chromosome

However, a sequence may also evolve by modifying its organization at a larger scale. These large-scale mutations are called *rearrangements*, or *structural variations*, and detecting them is the goal of *genome rearrangement* problems. The main rearrangements include the following:

- *Deletions*. A segment of the genome is lost (see figure 1.2).
- *Transpositions*. A segment of the genome moves to another location (see figure 1.3). Transpositions are sometimes referred to as translocations or insertions, but transposition is well adopted in the field of combinatorics of genome rearrangements.
- *Inversions* or *reversals*. A segment of the genome is reversed and the strands are exchanged (see figure 1.4).
- *Duplications*. A segment of DNA is copied and inserted in the genome. There are three main standard types of duplications: *tandem duplications*, illustrated by figure 1.5, which insert the copy next to the original; *retrotranspositions*, which insert a copy of a gene at an arbitrary location in the genome; and *whole genome duplications*, which copy either the whole genome or some of its chromosomes.

**Figure 1.6**

Reciprocal translocation of the dotted regions in two chromosomes

**Figure 1.7**

Fusion of two chromosomes

- *Reciprocal translocation*. A segment of a chromosome that contains a telomere is exchanged with a segment of another chromosome that also contains a telomere (see figure 1.6).
- *Fusion*. Two chromosomes are joined into one (see figure 1.7).
- *Fission*. One chromosome splits into two (this is the inverse of a fusion).
- *Horizontal, or lateral, transfer*. A segment of the genome is copied from one genome to another. This is common mainly in unicellular organisms.

All these operations act on a genome at the level of DNA segments rather than on nucleotides. This is why a genome is often represented by a sequence of segments in that setting: they are the segments that are found in an almost identical state in several species, not cut by rearrangements. Two segments are said to be *homologous* if they derive from a common ancestor and are distinguished by a replication event (they end up in two different genomes) or by a duplication event (they both belong to the same genome).

Genes are often taken as those homologous segments because, due to their functional utility, they are less subject to small mutations and are rarely cut by rearrangements, which is not the case for other parts of the genome.

1.2 Birth of the Combinatorics of Genome Rearrangements

In 1936 two renowned biologists, Dobzhansky (the inventor of the synthetic theory of evolution) and Sturtevant (the discoverer of rearrangement processes in genomes at the beginning of the twentieth century) proposed for the first time to use the degree

of disorder between the organization of genes in two different genomes as an indicator of an evolutionary distance between organisms (see Dobzhansky and Sturtevant [145, 146]). They proposed a scenario of inversions to explain chromosomal differences between 17 groups of flies, as well as a reconstruction of putative ancestral gene arrangements and species histories from the observation of the gene order along the chromosomes.

Since rearrangements are relatively rare events, scenarios minimizing their number are more likely to be close to reality. In 1941 Sturtevant and Novitski [343] formulated the problem of minimizing the number of inversions that may explain the differences in arrangements between two species: "... for each such sequence there was determined the minimum number of successive inversions required to reduce it to the ordinal sequence chosen as 'standard.' For numbers of loci above nine the determination of this minimum number proved too laborious, and too uncertain, to be carried out...."

The reconstruction of genome rearrangements from the examination of chromosomes, using techniques such as "chromosome banding" or "in-situ hybridisation" [301] were numerous, all focusing on relatively close species, so that the number of rearrangements was small. All these studies were based on the parsimony criterion, which makes molecular biologists often prefer explanations of differences between genomes that involve as few mutations as possible. This principle makes the connection with combinatorial optimization possible, because the optimization principle meets the parsimony criterion.

As we entered the genome sequencing era, the importance of rearrangements in evolution or illnesses was pointed out by several biologists, such as Palmer and Herbon [289], who examined the differences in the gene order of the mitochondrial genomes of cabbage and turnip, which are very similar in sequence but dramatically different in structure. It was not until 1982 that some researchers working in combinatorial optimization started to formalize and become involved in this problem, in order to overcome the limit of nine genes stated by Sturtevant and Novitski [343]. Watterson et al. [369] proposed to represent the relative positions of genes in different genomes as *permutations*. In order to propose an evolutionary scenario between two species, one had to solve the problem of transforming one circular permutation into another with a minimum number of inversions. The problem was far from being solved after this first article, but it was well stated.

Transforming one permutation into another by means of a minimum number of allowed operations is often equivalent to *sorting* a permutation by means of the same operations (see page 17). Though it took a decisive start in a biological context, the problem of sorting permutations with constraints was not new: a few mathematicians and computer scientists had already tackled that kind of problem in the past. Those problems were not, however, motivated by biology: constraints were related

to data structures as stacks, or were simply introduced as games that later turned out to be particular cases of genome rearrangement problems, or found uses in other fields.

New models were later proposed to handle more operations, duplicated segments, and several chromosomes. Shortly after Watterson et al. [369], the field started its dramatic expansion.

1.3 Statement of the Problem

The *genome rearrangement problem* is formulated in its most general form as follows: given a set of genomes and a set of possible evolutionary events, find a shortest set of events transforming those genomes into one another.

What “genome” means here, and what events are, makes the diversity of the problem. Miscellaneous models have been proposed, depending on various parameters, and we briefly review them in section 1.5. “Shortest” usually refers to the number of events, but it may also mean “of least weight” if events are weighted (e.g., according to their probability of occurrence).

The length (or weight) of an optimal sequence of events transforming one genome into another is called the *distance* between the two genomes. We will often require that this distance be a metric on the set of genomes, in the mathematical sense, and we recall its definition here.

Definition 1.1 A **metric** d on a set S is an application

$$d : S \times S \rightarrow \mathbb{R} : (s, t) \mapsto r$$

satisfying the following three axioms:

1. For all $s, t \in S$, $d(s, t) \geq 0$ and $d(s, t) = 0$ if and only if $s = t$ (positivity).
2. For all $s, t \in S$, $d(s, t) = d(t, s)$ (symmetry).
3. For all $s, t, u \in S$, $d(s, u) \leq d(s, t) + d(t, u)$ (triangular inequality).

A set S equipped with a metric d is called a **metric space** and is denoted (S, d) . Finding an optimal sequence of events between two genomes of course yields the distance between the two genomes, but the converse is not always true. Therefore, most of the time we will examine both aspects of the problem. A related problem we will be interested in is that of determining the *diameter* of a metric space.

Definition 1.2 The **diameter** of a metric space (S, d) is the maximal value the distance can reach, that is,

$$\max_{s, t \in S} d(s, t).$$

1.4 Scope of This Survey

This survey is restricted to the algorithmic and combinatorial aspects of genome rearrangements, but it also encompasses a few problems that are similar in spirit, even if they were not motivated by biology in the first place. The motivation for this is two-fold: first, those problems deserve at least to be mentioned here, since they are closely related to genome rearrangement problems; and second, the study of related problems or variants of our problems may provide insight on the original problems we are interested in.

There has been a lot of work on probabilistic models and statistical studies of genome rearrangement problems, which we will not consider in this work. We refer the reader to the surveys of Eriksson [165] and Durrett [150]. As the reader may have guessed by reading section 1.1, we also will not delve much into the biological aspects, and will focus on the mathematical aspects of genome rearrangements, though we will mention applications and biological contexts where appropriate.

Some partial surveys have been published in earlier articles or book chapters. In 1995, Hannenhalli and Pevzner [197] wrote the first survey on the combinatorics of genome rearrangements, mainly based on their success at sorting signed permutations by reversals (see sections 3.3 and 4.2). The chapters by Pevzner [296] and Setubal and Meidanis [333] dedicated to genome rearrangements mainly focus on sorting permutations by reversals. The books edited by Gascuel [182], Sankoff and Nadeau [322], Böckenhauer and Bongartz [74], Jiang et al. [222], and Tseng and Zerkowitz [360] contain chapters that survey part of the field or try to give a quick overview of it. A survey article by Li et al. [247] reveals the importance and popularity of the field, and this book intends to be a more developed version of it.

1.5 Overview of the Models

Depending on the assumptions that are made on the data, or the events we want to study, different models can be used. The basic objects will be *homologous markers* (i.e., segments of genomes that can be found in several species, leading to the belief that they belonged to the common ancestor of these species). Genes are a good example of such markers, though they are not the only ones; but since genes were historically first used as markers for genome rearrangement studies, we often say “genes” for “homologous markers,” as a simplification.

We will start with the simplest possible model, and progressively extend it by dropping restrictions. In the case where

1. the order of genes in each genome is known,
2. all genomes share the same set of genes,

3. all genomes contain a single copy of each gene, and
4. all genomes consist of a single chromosome,

genomes are modeled by *permutations* (see page 13): each gene can be assigned a unique number and is found exactly once in each genome.

As explained in section 1.1, a DNA molecule has two strands, and some rearrangements may change the strand that a segment belongs to. Therefore, each segment may be assigned a + or a - sign (+ is omitted most of the time) to indicate the strand it resides on, leading to the model of *signed permutations* (see page 15). We have also seen in section 1.1 that chromosomes can be linear or circular, and the latter case can be modeled using circular permutations rather than linear (classical) ones.

In spite of all the technical progress that has been made over the last decades and the large number of genomes that have been completely sequenced, many genomes have been only partially sequenced, which means that we cannot model them using permutations because genes are not totally ordered. In that case, genomes can nevertheless be modeled by *partially ordered sets*, and some studies can still be conducted using that model, as we will see in chapter 5.

In general, however, genes do not appear exactly once in each genome: due to duplications and deletions, there can be several copies of a gene in a given genome, or no copy at all. In that case, genomes are modeled by *strings* (on the alphabet of genes, see page 91) rather than permutations. Of course, it is possible to sign the elements of the string or to deal with circular strings.

A great part of living organisms have a genome that consists of several chromosomes (in a variable number, which can lie between 1 and 100), as is the case for all animals, and permutations as we have presented them are no longer a realistic model in that case. One can use the *disjoint cycle decomposition* of permutations (see page 14) to represent each chromosome using a cycle, in the case where chromosomes are circular, but this concept does not extend to linear chromosomes or strings since it cannot model duplicated genes. We may therefore want to extend our model to disjoint sets of paths and cycles (page 159), where each path or cycle models a chromosome.

Finally, one may not care about or simply not know the order of genes in each chromosome, and care only about whether two genes are in *synteny* (i.e., whether they belong to the same chromosome). In that case, genomes are modeled by *collections* of *sets* of genes (see chapter 12).

1.6 Organization of the Book

The first three parts of this book are organized according to the models presented in the previous section, each part being devoted to a mathematical object that has been

used to construct genome rearrangement problems. Part IV is devoted to an important generalization of the basic genome rearrangement problem, known as the *median problem*, which aims at considering more than two genomes at the same time and inferring their common ancestors. It surveys the attempts to reconstruct the kin relationships between genomes by drawing *phylogenetic trees* in which nodes are ancestral configurations and branches (edges) account for evolutionary events.

Part V is a collection of summaries that provide useful additional information on the field, such as a list of available software based on the algorithms that we describe in the book and a list of open problems. This book also includes two appendices: appendix A is devoted to basic concepts of graph theory, and appendix B recalls the basics of the algorithmic theory of complexity, as well as a few **NP**-complete problems.

Bibliography

- [1] Z. ADAM AND D. SANKOFF, *The ABCs of MGR with DCJ*, *Evolutionary Bioinformatics*, 4 (2008), pp. 69–74.
- [2] S. AHN AND S. TANKSLEY, *Comparative linkage maps of the rice and maize genomes*, *Proceedings of the National Academy of Sciences (USA)*, 90 (1993), pp. 7980–7984.
- [3] M. AIGNER AND D. B. WEST, *Sorting by insertion of leading element*, *Journal of Combinatorial Theory*, ser. A, 45 (1987), pp. 306–309.
- [4] Y. AJANA, J.-F. LEFEBVRE, E. R. M. TILLIER, AND N. EL-MABROUK, *Exploring the set of all minimal sequences of reversals—an application to test the replication-directed reversal hypothesis*, in *Proceedings of the Second International Workshop on Algorithms in Bioinformatics (WABI)*, R. Guigo and D. Gusfield, eds., vol. 2452 of *Lecture Notes in Computer Science*, Springer-Verlag, 2002, pp. 300–315.
- [5] S. B. AKERS AND B. KRISHNAMURTHY, *A group-theoretic model for symmetric interconnection networks*, *IEEE Transactions on Computers*, 38 (Apr. 1989), pp. 555–566.
- [6] S. B. AKERS, D. HAREL, AND B. KRISHNAMURTHY, *The star graph: An attractive alternative to the n -cube*, in *Proceedings of the Fourth International Conference on Parallel Processing (ICPP)*, Pennsylvania State University Press, 1987, pp. 393–400.
- [7] M. A. ALEKSEYEV, *Multi-break rearrangements: From circular to linear genomes*, in *Proceedings of the Fifth RECOMB Comparative Genomics Satellite Workshop (RECOMB-CG)*, G. Tesler and D. Durand, eds., vol. 4751 of *Lecture Notes in Computer Science*, Springer-Verlag, 2007, pp. 1–15.
- [8] M. A. ALEKSEYEV AND P. A. PEVZNER, *Genome halving problem revisited*, in *Proceedings of the Twenty-fourth Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*, K. Lodaya and M. Mahajan, eds., vol. 3328 of *Lecture Notes in Computer Science*, Springer-Verlag, 2004, pp. 1–15.
- [9] M. A. ALEKSEYEV AND P. A. PEVZNER, *Colored de Bruijn graphs and the genome halving problem*, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 4, no. 1 (2007), pp. 98–107.
- [10] M. A. ALEKSEYEV AND P. A. PEVZNER, *Whole genome duplications, multi-break rearrangements, and genome halving problem*, in *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, N. Bansal, K. Pruhs, and C. Stein, eds., *Society for Industrial and Applied Mathematics*, 2007, pp. 665–679.
- [11] M. A. ALEKSEYEV AND P. A. PEVZNER, *Whole genome duplications and contracted breakpoint graphs*, *SIAM Journal of Computing*, 36 (2007), pp. 1748–1763.
- [12] A. AMIR, Y. AUMANN, G. BENSON, A. LEVY, O. LIPSKY, E. PORAT, S. SKIENA, AND U. VISHNE, *Pattern matching with address errors: Rearrangement distances*, in *Proceedings of the Seventeenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, ACM, 2006, pp. 1221–1229.
- [13] A. AMIR, Y. AUMANN, P. INDYK, A. LEVY, AND E. PORAT, *Efficient computations of l_1 and l_∞ rearrangement distances*, in *Proceedings of the Fourteenth International Symposium on String Processing and Information Retrieval (SPIRE)*, N. Ziviani and R. Baeza-Yates, eds., vol. 4726 of *Lecture Notes in Computer Science*, Springer-Verlag, 2007, pp. 39–49.

- [14] A. AMIR, T. HARTMAN, O. KAPAH, A. LEVY, AND E. PORAT, *On the cost of interchange rearrangement in strings*, in Proceedings of the Fifteenth Annual European Symposium on Algorithms (ESA), L. Arge, M. Hoffmann, and E. Welzl, eds., vol. 4698 of Lecture Notes in Computer Science, Springer-Verlag, 2007, pp. 99–110.
- [15] S. ANGIBAUD, G. FERTIN, I. RUSU, A. THÉVENIN, AND S. VIALETTE, *Efficient tools for computing the number of breakpoints and the number of adjacencies between two genomes with duplicate genes*, Journal of Computational Biology, 15, no. 8 (2008), pp. 1093–1115.
- [16] S. ANGIBAUD, G. FERTIN, I. RUSU, AND S. VIALETTE, *How pseudo-Boolean programming can help genome rearrangement distance computation*, in Proceedings of the Fourth RECOMB Comparative Genomics Satellite Workshop (RECOMB-CG), G. Bourque and N. El-Mabrouk, eds., vol. 4205 of Lecture Notes in Computer Science, Springer-Verlag, 2006, pp. 75–86.
- [17] S. ANGIBAUD, G. FERTIN, I. RUSU, A. THÉVENIN, AND S. VIALETTE, *A pseudo-Boolean programming approach for computing the breakpoint distance between two genomes with duplicate genes*, in Proceedings of the Fifth RECOMB Comparative Genomics Satellite Workshop (RECOMB-CG), G. Tesler and D. Durand, eds., vol. 4751 of Lecture Notes in Computer Science, Springer-Verlag, 2007, pp. 16–29.
- [18] S. ANGIBAUD, G. FERTIN, I. RUSU, AND S. VIALETTE, *A pseudo-Boolean framework for computing rearrangement distances between genomes with duplicates*, Journal of Computational Biology, 14 (2007), pp. 379–393.
- [19] S. ANGIBAUD, G. FERTIN, I. RUSU, A. THÉVENIN, AND S. VIALETTE, *On the approximability of comparing genomes with duplicates*, Journal of Graph Algorithms and Applications (2008). In press.
- [20] D. L. APPLGATE, R. E. BIXBY, V. CHVÁTAL, AND W. J. COOK, *The Traveling Salesman Problem: A Computational Study*, Princeton Series in Applied Mathematics, Princeton University Press, 2007.
- [21] W. ARNDT AND J. TANG, *Improving inversion median computation using commuting reversals and cycle information*, in Proceedings of the Fifth RECOMB Comparative Genomics Satellite Workshop (RECOMB-CG), G. Tesler and D. Durand, eds., vol. 4751 of Lecture Notes in Computer Science, Springer-Verlag, 2007, pp. 30–44.
- [22] G. AUSIELLO, P. CRESCENZI, G. GAMBOSI, V. KANN, A. MARCHETTI-SPACCAMELA, AND M. PROTASI, *Complexity and Approximation: Combinatorial Optimization Problems and Their Approximability Properties*, Springer-Verlag, 1999.
- [23] A. AUYEUNG AND A. ABRAHAM, *Estimating genome reversal distance by genetic algorithm*, in IEEE Congress on Evolutionary Computation, vol. 2, IEEE Computer Society Press, 2003, pp. 1157–1161.
- [24] D. AVIS AND M. NEWBORN, *On pop-stacks in series*, Utilitas Mathematica, 19 (1981), pp. 129–140.
- [25] A. BACHRACH, K. CHEN, C. HARRELSON, R. MIHAESCU, S. RAO, AND A. SHAH, *Lower bounds for maximum parsimony with gene order data*, in Proceedings of the Third RECOMB Comparative Genomics Satellite Workshop (RECOMB-CG), A. McLysaght and D. H. Huson, eds., vol. 3678 of Lecture Notes in Computer Science, Springer-Verlag, 2005, pp. 1–10.
- [26] D. A. BADER, B. M. E. MORET, AND M. YAN, *A linear-time algorithm for computing inversion distance between signed permutations with an experimental study*, Journal of Computational Biology, 8, no. 5 (2001), pp. 483–491.
- [27] M. BADER AND E. OHLEBUSCH, *Sorting by weighted reversals, transpositions, and inverted transpositions*, in Proceedings of the Tenth Annual International Conference on Research in Computational Molecular Biology (RECOMB), A. Apostolico, C. Guerra, S. Istrail, P. A. Pevzner, and M. S. Waterman, eds., vol. 3909 of Lecture Notes in Computer Science, Springer-Verlag, 2006, pp. 563–577.
- [28] M. BADER AND E. OHLEBUSCH, *Sorting by weighted reversals, transpositions, and inverted transpositions*, Journal of Computational Biology, 14, no. 5 (2007), pp. 615–636.
- [29] V. BAFNA AND P. A. PEVZNER, *Genome rearrangements and sorting by reversals*, in Proceedings of the Thirty-fourth Annual Symposium on Foundations of Computer Science (FOCS), IEEE Computer Society Press, 1993, pp. 148–157.
- [30] V. BAFNA AND P. A. PEVZNER, *Sorting by transpositions*, SIAM Journal on Discrete Mathematics, 11 (May 1998), pp. 224–240 (electronic).

- [31] V. BAFNA, D. BEAVER, M. FÜRER, AND P. A. PEVZNER, *Circular permutations and genome shuffling*, in *Comparative Genomics: Empirical and Analytical Approaches to Gene Order Dynamics*, Map Alignment and the Evolution of Gene Families, vol. 1 of Sankoff and Nadeau [322], 2000, pp. 199–206.
- [32] B. BAKER, *Approximation algorithms for NP-complete problems on planar graphs*, *Journal of the ACM*, 41 (Jan. 1994), pp. 153–180.
- [33] D. W. BASS AND I. H. SUDBOROUGH, *On the shuffle-exchange permutation network*, in *Proceedings of the Third International Symposium on Parallel Architectures, Algorithms, and Networks (ISPAN)*, IEEE Computer Society Press, 1997, pp. 165–171.
- [34] D. W. BASS AND I. H. SUDBOROUGH, *Pancake problems with restricted prefix reversals and some corresponding Cayley networks*, *Journal of Parallel and Distributed Computing*, 63, no. 3 (2003), pp. 327–336.
- [35] T. BATU AND S. C. SAHINALP, *Locally consistent parsing with applications to approximate string comparisons*, in *Developments in Language Theory*, C. de Felice and A. Restivo, eds., vol. 3572 of *Lecture Notes in Computer Science*, Springer-Verlag, 2005, pp. 22–35.
- [36] W. W. BEIN, L. L. LARMORE, S. LATIFI, AND I. H. SUDBOROUGH, *Block sorting is hard*, *International Journal of Foundations of Computer Science*, 14, no. 3 (2003), pp. 425–437.
- [37] W. W. BEIN, L. L. LARMORE, L. MORALES, AND I. H. SUDBOROUGH, *A faster and simpler 2-approximation algorithm for block sorting*, in *Proceedings of the Fifteenth International Symposium on Fundamentals of Computation Theory (FCT)*, M. Liskiewicz and R. Reischuk, eds., vol. 3623 of *Lecture Notes in Computer Science*, Springer-Verlag, 2005, pp. 115–124.
- [38] M. A. BENDER, D. GE, S. HE, H. HU, R. PINTER, S. SKIENA, AND F. SWIDAN, *Improved bounds on sorting with length-weighted reversals*, in *Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, J. I. Munro, ed., *Society for Industrial and Applied Mathematics*, 2004, pp. 919–928.
- [39] M. A. BENDER, D. GE, S. HE, H. HU, R. PINTER, S. SKIENA, AND F. SWIDAN, *Improved bounds on sorting by length-weighted reversals*, *Journal of Computer and System Sciences*, 74, no. 5 (2008), pp. 744–774.
- [40] M. BENOÎT-GAGNÉ AND S. HAMEL, *A new and faster method of sorting by transpositions*, in *Proceedings of the Eighteenth Annual Symposium on Combinatorial Pattern Matching (CPM)*, B. Ma and K. Zhang, eds., vol. 4580 of *Lecture Notes in Computer Science*, Springer-Verlag, 2007, pp. 131–141.
- [41] S. BÉRARD, A. BERGERON, C. CHAUVE, AND C. PAUL, *Perfect sorting by reversals is not always difficult*, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 4 (Jan.–Mar. 2007), pp. 4–16.
- [42] S. BÉRARD, C. CHAUVE, AND C. PAUL, *A more efficient algorithm for perfect sorting by reversals*, *Information Processing Letters*, 106 (2008), pp. 90–95.
- [43] C. BERGE, *Graphes et hypergraphes*, *Monographies Universitaires de Mathématiques*, no. 37, Dunod, 1970.
- [44] A. BERGERON, *A very elementary presentation of the Hannenhalli-Pevzner theory*, *Discrete Applied Mathematics*, 146, no. 2 (2005), pp. 134–145.
- [45] A. BERGERON AND J. STOYE, *On the similarity of sets of permutations and its applications to genome comparison*, in *Proceedings of the Ninth International Computing and Combinatorics Conference (COCOON)*, T. Warnow and B. Zhu, eds., vol. 2697 of *Lecture Notes in Computer Science*, Springer-Verlag, 2003, pp. 68–79.
- [46] A. BERGERON AND F. STRASBOURG, *Experiments in computing sequences of reversals*, in *Proceedings of the First Workshop on Algorithms in Bioinformatics (WABI)*, O. Gascuel and B. M. E. Moret, eds., vol. 2149 of *Lecture Notes in Computer Science*, Springer-Verlag, 2001, pp. 164–174.
- [47] A. BERGERON, J. MIXTACKI, AND J. STOYE, *Reversal distance without hurdles and fortresses*, in *Proceedings of the Fifth Annual Symposium on Combinatorial Pattern Matching (CPM)*, M. Crochemore and D. Gusfield, eds., vol. 807 of *Lecture Notes in Computer Science*, Springer-Verlag, 1994, pp. 388–399.
- [48] A. BERGERON, C. CHAUVE, T. HARTMAN, AND K. ST-ONGE, *On the properties of sequences of reversals that sort a signed permutation*, *JOBIM*, 2002, pp. 99–108.
- [49] A. BERGERON, S. HEBER, AND J. STOYE, *Common intervals and sorting by reversals: A marriage of necessity*, *Bioinformatics*, 18, supp. 2 (2002), pp. S54–S63.
- [50] A. BERGERON, J. MIXTACKI, AND J. STOYE, *The reversal distance problem*, in Gascuel [182].

- [51] A. BERGERON, J. MIXTACKI, AND J. STOYE, *On sorting by translocations*, in Proceedings of the Ninth Annual International Conference on Research in Computational Molecular Biology (RECOMB), S. Miyano, J. Mesirov, S. Kasif, S. Istrail, P. A. Pevzner, and M. Waterman, eds., vol. 3500 of Lecture Notes in Bioinformatics, Springer-Verlag, 2005, pp. 615–629.
- [52] A. BERGERON, J. MIXTACKI, AND J. STOYE, *A unifying view of genome rearrangements*, in Proceedings of the Sixth International Workshop on Algorithms in Bioinformatics (WABI), R. L. Malmberg, L. Cai, P. Berman, S. Rahmann, L. Zhang, and R. Sharan, eds., vol. 4175 of Lecture Notes in Computer Science, Springer-Verlag, 2006, pp. 163–173.
- [53] A. BERGERON, J. MIXTACKI, AND J. STOYE, *HP distance via double cut and join distance*, in Proceedings of the Nineteenth Annual Symposium on Combinatorial Pattern Matching (CPM), P. Ferragina and G. M. Landau, eds., vol. 5029 of Lecture Notes in Computer Science, Springer-Verlag, 2008, pp. 56–68.
- [54] P. BERMAN AND T. FUJITO, *On approximation properties of the independent set problem for degree 3 graphs*, in Proceedings of the Fourth International Workshop on Algorithms and Data Structures (WADS), S. G. Akl, F. K. H. A. Dehne, J.-R. Sack, and N. Santoro, eds., vol. 955 of Lecture Notes in Computer Science, Springer-Verlag, 1995, pp. 449–460.
- [55] P. BERMAN AND M. FÜRER, *Approximating maximum independent set in bounded degree graphs*, in Proceedings of the Fifth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), Society for Industrial and Applied Mathematics, 1994, pp. 365–371.
- [56] P. BERMAN AND S. HANNENHALLI, *Fast sorting by reversal*, in Proceedings of the Seventh Annual Symposium on Combinatorial Pattern Matching (CPM), D. S. Hirschberg and E. W. Myers, eds., vol. 1075 of Lecture Notes in Computer Science, Springer-Verlag, 1996, pp. 168–185.
- [57] P. BERMAN AND M. KARPINSKI, *On some tighter inapproximability results (extended abstract)*, in Proceedings of the Twenty-sixth International Colloquium on Automata, Languages and Programming (ICALP), J. Wiedermann, P. van Emde Boas, and M. Nielsen, eds., vol. 1644 of Lecture Notes in Computer Science, Springer-Verlag, 1999, pp. 200–209.
- [58] P. BERMAN, S. HANNENHALLI, AND M. KARPINSKI, *1.375-approximation algorithm for sorting by reversals*, in Proceedings of the Tenth Annual European Symposium on Algorithms (ESA), R. H. Möhring and R. Raman, eds., vol. 2461 of Lecture Notes in Computer Science, Springer-Verlag, 2002, pp. 200–210.
- [59] M. BERN AND P. PLASSMANN, *The Steiner problem with edge lengths 1 and 2*, Information Processing Letters, 32 (1989), pp. 171–176.
- [60] M. BERNT, D. MERKLE, AND M. MIDDENDORF, *Genome rearrangement based on reversals that preserve conserved intervals*, IEEE/ACM Transactions on Computational Biology and Bioinformatics, 3, no. 3 (2006), pp. 275–288.
- [61] M. BERNT, D. MERKLE, AND M. MIDDENDORF, *A parallel algorithm for solving the reversal median problem*, in Proceedings of the Sixth International Conference on Parallel Processing and Applied Mathematics (PPAM), R. Wyrzykowski, J. Dongarra, N. Meyer, and J. Wasniewski, eds., vol. 3911 of Lecture Notes in Computer Science, Springer-Verlag, 2006, pp. 1089–1096.
- [62] M. BERNT, D. MERKLE, AND M. MIDDENDORF, *The reversal median problem, common intervals, and mitochondrial gene orders*, in Proceedings of the Second International Symposium on Computational Life Sciences (COMPLIFE), M. Berthold, R. Glen, and I. Fischer, eds., vol. 4216 of Lecture Notes in Bioinformatics, 2006, pp. 52–63.
- [63] M. BERNT, D. MERKLE, AND M. MIDDENDORF, *A fast and exact algorithm for the perfect reversal median problem*, in Proceedings of the Third International Symposium on Bioinformatics Research and Applications (ISBRA), I. Măndoiu and A. Zelikovsky, eds., vol. 4463 of Lecture Notes in Bioinformatics, Springer-Verlag, 2007, pp. 305–316.
- [64] M. BERNT, D. MERKLE, AND M. MIDDENDORF, *Using median sets for inferring phylogenetic trees*, Bioinformatics, 23 (2007), pp. e129–e135. Special issue of ECCB 2006.
- [65] M. BERNT, D. MERKLE, K. RAMSCH, G. FRITZSCH, M. PERSEKE, D. BERNHARD, M. SCHLEGEL, P. F. STADLER, AND M. MIDDENDORF, *CREX: Inferring genomic rearrangements based on common intervals*, Bioinformatics, 23, no. 21 (2007), pp. 2957–2958.
- [66] M. BLANCHETTE, T. KUNISAWA, AND D. SANKOFF, *Parametric genome rearrangement*, Gene, 172 (1996), pp. GC11–GC17.

- [67] M. BLANCHETTE, G. BOURQUE, AND D. SANKOFF, *Breakpoint phylogenies*, in vol. 8 of S. Miyano and Tagaki, eds., *Genome Informatics*, Universal Academy Press, 1997, pp. 25–34.
- [68] G. BLIN AND R. RIZZI, *Conserved interval distance computation between non-trivial genomes*, in Proceedings of the Eleventh International Computing and Combinatorics Conference (COCOON), L. Wang, ed., vol. 3595 of *Lecture Notes in Computer Science*, Springer-Verlag, 2005, pp. 22–31.
- [69] G. BLIN, C. CHAUVE, AND G. FERTIN, *The breakpoint distance for signed sequences*, in Proceedings of the First International Conference on Algorithms and Computational Methods for Biochemical and Evolutionary Networks, vol. 3 of *Texts in Algorithms*, King's College (London) and KCL Publications, 2004, pp. 3–16.
- [70] G. BLIN, C. CHAUVE, AND G. FERTIN, *Genes order and phylogenetic reconstruction: Application to gamma-proteobacteria*, in Proceedings of the Third RECOMB Comparative Genomics Satellite Workshop (RECOMB-CG), A. McLysaght and D. H. Huson, eds., vol. 3678 of *Lecture Notes in Computer Science*, Springer-Verlag, 2005, pp. 11–20.
- [71] G. BLIN, A. CHATEAU, C. CHAUVE, AND Y. GINGRAS, *Inferring positional homologs with common intervals of sequences*, in Proceedings of the Fourth RECOMB Comparative Genomics Satellite Workshop (RECOMB-CG), G. Bourque and N. El-Mabrouk, eds., vol. 4205 of *Lecture Notes in Computer Science*, Springer-Verlag, 2006, pp. 24–38.
- [72] G. BLIN, E. BLAIS, D. HERMELIN, P. GUILLON, M. BLANCHETTE, AND N. EL-MABROUK, *Gene maps linearization using genomic rearrangement distances*, *Journal of Computational Biology*, 14, no. 4 (2007), pp. 394–407.
- [73] G. BLIN, C. CHAUVE, G. FERTIN, R. RIZZI, AND S. VIALETTE, *Comparing genomes with duplications: A computational complexity point of view*, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 4, no. 4 (2007), pp. 523–534.
- [74] H.-J. BÖCKENHAUER AND D. BONGARTZ, eds., *Algorithmic Aspects of Bioinformatics*, Natural Computing Series, Springer-Verlag, 2007. Original German edition published by Teubner.
- [75] M. BÓNA, *A survey of stack-sorting disciplines*, *Electronic Journal of Combinatorics*, 9 (2003).
- [76] M. BÓNA, *Combinatorics of permutations*, in *Discrete Mathematics and Its Applications*, Chapman & Hall/CRC, 2004.
- [77] D. BONGARTZ, *Algorithmic Aspects of Some Combinatorial Problems in Bioinformatics*, Ph.D. thesis, University of Viersen, Germany, 2006.
- [78] J. BONIN AND A. DE MIER, *Lattice path matroids: Structural properties*, *European Journal of Combinatorics*, 27 (2006), pp. 701–738.
- [79] R. BOPANA AND M. HALLDÖRSSON, *Approximating maximum independent sets by excluding subgraphs*, *BIT*, 32, no. 2 (1992), pp. 180–196.
- [80] P. BOSE, J. F. BUSS, AND A. LUBIW, *Pattern matching for permutations*, *Information Processing Letters*, 65 (1998), pp. 277–283.
- [81] G. BOURQUE AND P. A. PEVZNER, *Genome-scale evolution: Reconstructing gene orders in the ancestral species*, *Genome Research*, 12 (2002), pp. 26–36.
- [82] G. BOURQUE, P. A. PEVZNER, AND G. TESLER, *Reconstructing the genomic architecture of ancestral mammals: Lessons from human, mouse, and rat genomes*, *Genome Research*, 14 (2004), pp. 507–516.
- [83] G. BOURQUE, Y. YACEF, AND N. EL-MABROUK, *Maximizing synteny blocks to identify ancestral homologs*, in Proceedings of the Third RECOMB Comparative Genomics Satellite Workshop (RECOMB-CG), A. McLysaght and D. H. Huson, eds., vol. 3678 of *Lecture Notes in Computer Science*, Springer-Verlag, 2005, pp. 21–34.
- [84] G. BOURQUE, E. M. ZDOBNOV, P. BORK, P. A. PEVZNER, AND G. TESLER, *Comparative architectures of mammalian and chicken genomes reveal highly variable rates of genomic rearrangements across different lineages*, *Genome Research*, 15 (Jan. 2005), pp. 98–110.
- [85] G. BOURQUE, G. TESLER, AND P. A. PEVZNER, *The convergence of cytogenetics and rearrangement-based models for ancestral genome reconstruction*, *Genome Research*, 16 (2006), pp. 311–313.
- [86] M. BOUVEL AND D. ROSSIN, *A variant of the tandem duplication-random loss model of genome rearrangement*. Available at <http://arxiv.org/abs/0801.2524>, 2008.

- [87] M. BRAGA, M.-F. SAGOT, C. SCORNAVACCA, AND E. TANNIER, *The solution space of sorting by reversals*, in Proceedings of the Third International Symposium on Bioinformatics Research and Applications (ISBRA), I. Măndoiu and A. Zelikovsky, eds., vol. 4463 of Lecture Notes in Computer Science, Springer-Verlag, 2007, pp. 293–304.
- [88] D. BRYANT, *The Complexity of the Breakpoint Median Problem*, tech. rep. CRM-2579, Centre de Recherches Mathématiques, Université de Montréal, 1998.
- [89] D. BRYANT, *The complexity of calculating exemplar distances*, in vol. 1 of Sankoff and Nadeau [322], pp. 207–212.
- [90] D. BRYANT, *A lower bound for the breakpoint phylogeny problem*, Journal of Discrete Algorithms, 2 (2004), pp. 229–255.
- [91] B.-M. BUI-XUAN, M. HABIB, AND C. PAUL, *Revisiting T. Uno and M. Yagiura's algorithm*, in Proceedings of the Sixteenth Annual International Symposium on Algorithms and Computation (ISAAC), X. Deng and D. Du, eds., Springer-Verlag, 2005, pp. 146–155.
- [92] A. CAPRARA, *Formulations and hardness of multiple sorting by reversals*, in Proceedings of the Third Annual International Conference on Computational Molecular Biology (RECOMB), S. Istrail, P. Pevzner, and M. Waterman, eds., ACM, 1999, pp. 84–93.
- [93] A. CAPRARA, *Sorting permutations by reversals and Eulerian cycle decompositions*, SIAM Journal on Discrete Mathematics, 12 (Feb. 1999), pp. 91–110 (electronic).
- [94] A. CAPRARA, *On the tightness of the alternating-cycle lower bound for sorting by reversals*, Journal of Combinatorial Optimization, 3 (1999), pp. 149–182.
- [95] A. CAPRARA, *On the practical solution of the reversal median problem*, in Proceedings of the First Workshop on Algorithms in Bioinformatics (WABI), O. Gascuel and B. M. E. Moret, eds., vol. 2149 of Lecture Notes in Computer Science, Springer-Verlag, 2001, pp. 238–251.
- [96] A. CAPRARA, *Additive bounding, worst-case analysis and the breakpoint median problem*, SIAM Journal on Optimization, 13 (2002), pp. 508–519.
- [97] A. CAPRARA, *The reversal median problem*, INFORMS Journal on Computing, 15, no. 1 (2003), pp. 93–113.
- [98] A. CAPRARA AND R. RIZZI, *Improved approximation for breakpoint graph decomposition and sorting by reversals*, Journal of Combinatorial Optimization, 6 (2002), pp. 157–182.
- [99] A. CAPRARA, G. LANCIA, AND S.-K. NG, *Sorting permutations by reversals through branch-and-price*, INFORMS Journal on Computing, 13 (2001), pp. 224–244.
- [100] A. CAYLEY, *Note on the theory of permutations*, Philosophical Magazine, 34 (1849), pp. 527–529.
- [101] M. CHARIKAR, K. MAKARYCHEV, AND Y. MAKARYCHEV, *Near-optimal algorithms for maximum constraint satisfaction problems*, in Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), N. Bansal, K. Pruhs, and C. Stein, eds., Society for Industrial and Applied Mathematics, 2007, pp. 62–68.
- [102] K. CHAUDHURI, K. CHEN, R. MIHAESCU, AND S. RAO, *On the tandem duplication-random loss model of genome rearrangement*, in Proceedings of the Seventeenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), ACM, 2006, pp. 564–570.
- [103] C. CHAUVE AND G. FERTIN, *On maximal instances for the original syntenic distance*, Theoretical Computer Science, 326, no. 1–3 (2004), pp. 29–43.
- [104] C. CHAUVE, G. FERTIN, R. RIZZI, AND S. VIALETTE, *Genomes containing duplicates are hard to compare*, in Proceedings of the Second International Workshop on Bioinformatics Research and Applications (IWBRA), V. N. Alexandrov, G. D. van Albada, P. M. A. Sloot, and J. Dongarra, eds., vol. 3992 of Lecture Notes in Computer Science, Springer-Verlag, 2006, pp. 783–790.
- [105] T. CHEN AND S. SKIENA, *Sorting with fixed-length reversals*, Discrete Applied Mathematics, 71 (1996), pp. 269–295.
- [106] X. CHEN, J. ZHENG, Z. FU, P. NAN, Y. ZHONG, S. LONARDI, AND T. JIANG, *Assignment of orthologous genes via genome rearrangement*, IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2, no. 4 (2005), pp. 302–315.

- [107] Z. CHEN, R. H. FOWLER, B. FU, AND B. ZHU, *Lower bounds on the approximation of the exemplar conserved distance problem of genomes*, in Proceedings of the Twelfth International Computing and Combinatorics Conference (COCOON), D. Z. Chen and D. T. Lee, eds., vol. 4112 of Lecture Notes in Computer Science, Springer-Verlag, 2006, pp. 245–254.
- [108] Z. CHEN, B. FU, AND B. ZHU, *The approximability of the exemplar breakpoint distance problem*, in Proceedings of the Second International Conference on Algorithmic Aspects in Information and Management (AAIM), S.-W. Cheng and C. K. Poon, eds., vol. 4041 of Lecture Notes in Computer Science, Springer-Verlag, 2006, pp. 291–302.
- [109] Z. CHEN, B. FU, J. XU, B. YANG, Z. ZHAO, AND B. ZHU, *Non-breaking similarity of genomes with gene repetitions*, in Proceedings of the Eighteenth Annual Symposium on Combinatorial Pattern Matching (CPM), B. Ma and K. Zhang, eds., vol. 4580 of Lecture Notes in Computer Science, Springer-Verlag, 2007, pp. 119–130.
- [110] B. CHITTURI AND I. H. SUDBOROUGH, *Bounding prefix transposition distance for strings and permutations*, in Proceedings of the Forty-first Annual Hawaii International Conference on System Sciences (HICSS), IEEE Computer Society Press, 2008, p. 468.
- [111] B. CHITTURI, W. FAHLE, Z. MENG, L. MORALES, C. SHIELDS, I. SUDBOROUGH, AND W. VOIT, *An $(18/11)n$ upper bound for sorting by prefix reversals*, Theoretical Computer Science, (2008).
- [112] B. CHITTURI, H. SUDBOROUGH, W. VOIT, AND X. FENG, *Adjacent swaps on strings*, in Proceedings of the Fourteenth International Computing and Combinatorics Conference (COCOON), X. Hu and J. Wang, eds., vol. 5092 of Lecture Notes in Computer Science, Springer-Verlag, 2008, pp. 299–308.
- [113] D. A. CHRISTIE, *Sorting permutations by block-interchanges*, Information Processing Letters, 60 (1996), pp. 165–169.
- [114] D. A. CHRISTIE, *A $3/2$ -approximation algorithm for sorting by reversals*, in Proceedings of the Ninth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), Society for Industrial and Applied Mathematics, 1998, pp. 244–252.
- [115] D. A. CHRISTIE, *Genome Rearrangement Problems*, Ph.D. thesis, University of Glasgow, Scotland, 1998.
- [116] D. A. CHRISTIE AND R. W. IRVING, *Sorting strings by reversals and by transpositions*, SIAM Journal on Discrete Mathematics, 14, no. 2 (2001), pp. 193–206.
- [117] M. CHROBAK, P. KOLMAN, AND J. SGALL, *The greedy algorithm for the minimum common string partition problem*, in Proceedings of the Seventh International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX), vol. 3122 of Lecture Notes in Computer Science, Springer-Verlag, 2004, pp. 84–95.
- [118] D. S. COHEN AND M. BLUM, *On the problem of sorting burnt pancakes*, Discrete Applied Mathematics, 61 (1995), pp. 105–120.
- [119] E. COISSAC, E. MAILLIER, AND P. NETTER, *A comparative study of duplication in bacteria and eukaryotes: The importance of telomeres*, Molecular Biology and Evolution, 14 (1997), pp. 1062–1074.
- [120] S. COOK, *The complexity of theorem proving procedures*, in Proceedings of the Third Annual ACM Symposium on Theory of Computing (STOC), ACM, 1971, pp. 151–158.
- [121] T. H. CORMEN, C. E. LEISERSON, R. L. RIVEST, AND C. STEIN, *Introduction to Algorithms*, 2nd ed., MIT Press, 2001.
- [122] G. CORMODE AND S. MUTHUKRISHNAN, *The string edit distance matching problem with moves*, in Proceedings of the Thirteenth Annual ACM-SIAM Symposium on Discrete Mathematics (SODA), Society for Industrial and Applied Mathematics, 2002, pp. 667–676.
- [123] G. CORMODE, S. MUTHUKRISHNAN, M. PATERSON, S. C. SAHINALP, AND U. VISHKIN, *Techniques and applications for approximating string distances—rough draft*, 2000. <http://citeseer.ist.psu.edu/320221.html>.
- [124] M. E. COSNER, *Phylogenetic and Molecular Evolutionary Studies of Chloroplast DNA Variations in the Campanulaceae*, Ph.D. thesis, Ohio State University, 1993.
- [125] M. E. COSNER, R. K. JANSEN, B. M. E. MORET, L. A. RAUBESON, L.-S. WANG, T. WARNOW, AND S. WYMAN, *An empirical comparison of phylogenetic methods on chloroplast gene order data in Campanulaceae*, in vol. 1 of Sankoff and Nadeau [322], pp. 99–121.

- [126] D. CRANSTON, I. H. SUDBOROUGH, AND D. B. WEST, *Short proofs for cut-and-paste sorting of permutations*, Discrete Mathematics, 307, no. 22 (2007), pp. 2866–2870.
- [127] M. CROCHMORE, D. HERMELIN, G. LANDAU, AND S. VIALETTE, *Approximating the 2-interval pattern problem*, in Proceedings of the Thirteenth Annual European Symposium on Algorithms (ESA), G. S. Brodal and S. Leonardi, eds., vol. 3669 of Lecture Notes in Computer Science, Springer-Verlag, 2005, pp. 426–437.
- [128] Y. CUI, L. WANG, AND D. ZHU, *A 1.75-approximation algorithm for unsigned translocation distance*, Journal of Computer and System Sciences, 73 (2007), pp. 1045–1059.
- [129] Y. CUI, L. WANG, D. ZHU, AND X. LIU, *A $(1.5 + \varepsilon)$ -approximation algorithm for unsigned translocation distance*, IEEE/ACM Transactions on Computational Biology and Bioinformatics, 5, no. 1 (2008), pp. 56–66.
- [130] A. C. E. DARLING, B. MAU, F. R. BLATTNER, AND N. T. PERNA, *GRIL: Genome rearrangement and inversion locator*, Bioinformatics, 20 (Jan. 2004), pp. 122–124.
- [131] B. DASGUPTA, T. JIANG, S. KANNAN, M. LI, AND E. SWEDYK, *On the complexity and approximation of syntenic distance*, Discrete Applied Mathematics, 88 (1998), pp. 59–82.
- [132] B. DAVEY AND H. PRIESTLEY, Introduction to Lattices and Order, 2nd ed., Cambridge University Press, 2002.
- [133] W. H. E. DAY, *Computationally difficult parsimony problems in phylogenetic systematics*, Journal of Theoretical Biology, 103 (1983), pp. 429–438.
- [134] C. DEMETRESCU AND I. FINOCCHI, *Combinatorial algorithms for feedback problems in directed graphs*, Information Processing Letters, 86 (2003), pp. 129–136.
- [135] P. DIACONIS, Group Representations in Probability and Statistics, vol. 11 of Lecture Notes—Monograph, Institute of Mathematical Sciences, 1988.
- [136] Z. DIAS AND C. CARVALHO DE SOUZA, *Polynomial-sized ILP models for rearrangement distance problems*, in Proceedings of the Second Brazilian Symposium on Bioinformatics (BSB), M.-F. Sagot and M. E. M. T. Walter, eds., vol. 4643 in Lecture Notes in Computer Science, Springer-Verlag, 2007, pp. 74–85.
- [137] Z. DIAS AND J. MEIDANIS, *Genome rearrangements distance by fusion, fission, and transposition is easy*, in Proceedings of the Eighth International Symposium on String Processing and Information Retrieval (SPIRE), IEEE Computer Society Press, 2001, pp. 250–253.
- [138] Z. DIAS AND J. MEIDANIS, *Sorting by prefix transpositions*, in Proceedings of the Ninth International Symposium on String Processing and Information Retrieval (SPIRE), A. H. F. Laender and A. L. Oliveira, eds., vol. 2476 of Lecture Notes in Computer Science, Springer-Verlag, 2002, pp. 65–76.
- [139] Z. DIAS AND J. MEIDANIS, *The syntenic distance problem using only fusions and fissions*, in Proceedings of the Second Brazilian Workshop on Bioinformatics (WOB), S. Lifschitz, N. F. Almeida Jr., G. J. Pappas Jr., and R. Linden, eds., 2003, pp. 72–79.
- [140] Z. DIAS, J. MEIDANIS, AND M. E. M. T. WALTER, *A new approach for approximating the transposition distance*, in Proceedings of the Seventh International Symposium on String Processing and Information Retrieval (SPIRE), IEEE Computer Society Press, 2000, pp. 199–208.
- [141] G. DIDIER, *Common intervals of two sequences*, in Proceedings of the Third Workshop on Algorithms in Bioinformatics (WABI), G. Benson and R. Page, eds., vol. 2812 of Lecture Notes in Computer Science, Springer-Verlag, 2003, pp. 17–24.
- [142] Y. DIEKMANN, M.-F. SAGOT, AND E. TANNIER, *Evolution under reversals: Parsimony and conservation of common intervals*, IEEE/ACM Transactions on Computational Biology and Bioinformatics, 4, no. 2 (2007), pp. 301–309.
- [143] R. DIESTEL, Graph Theory, vol. 173 of Graduate Texts in Mathematics, 3rd ed., Springer-Verlag, 2005.
- [144] R. DILWORTH, *A decomposition theorem for partially ordered sets*, Annals of Mathematics, 51 (1950), pp. 161–166.
- [145] T. DOBZHANSKY AND A. H. STURTEVANT, *Inversions in the third chromosome of wild races of Drosophila pseudoobscura, and their use in the study of the history of the species*, Proceedings of the National Academy of Sciences (USA), 22, no. 7 (1936), 448–450.

- [146] T. DOBZHANSKY AND A. H. STURTEVANT, *Inversions in the chromosomes of Drosophila pseudoobscura*, Genetics, 23 (1938), pp. 28–64.
- [147] J.-P. DOIGNON AND A. LABARRE, *On Hultman numbers*, Journal of Integer Sequences, 10 (2007), art. 07.6.2.
- [148] R. DOWNEY AND M. R. FELLOWS, *Parameterized Complexity*, Springer-Verlag, 1999.
- [149] T. DUH AND M. FÜRER, *Approximation of k -set cover by semi-local optimization*, in Proceedings of the Twenty-ninth Annual ACM Symposium on the Theory of Computing (STOC), ACM, 1997, pp. 256–265.
- [150] R. DURRETT, *Genome rearrangement: Recent progress and open problems*, 2003. <http://www.math.cornell.edu/~durrett/FGR>.
- [151] H. DWEIGHTER, *Elementary problems and solutions, problem E2569*, American Mathematical Monthly, 82 (1975), 1010.
- [152] J. V. EARNEST-DEYOUNG, E. LERAT, AND B. M. E. MORET, *Reversing gene erosion: Reconstructing ancestral bacterial genomes from gene-content and order data*, in Proceedings of the Fourth Workshop on Algorithms in Bioinformatics (WABI), I. Jonassen and J. Kim, eds., vol. 3240 of Lecture Notes in Computer Science, Springer-Verlag, 2004, pp. 1–13.
- [153] J. EDMONDS, *Paths, trees and flowers*, Canadian Journal of Mathematics, 17 (1965), 449–467.
- [154] N. EL-MABROUK, *Sorting signed permutations by reversals and insertions/deletions of contiguous segments*, Journal of Discrete Algorithms, 1 (2000), pp. 105–121.
- [155] N. EL-MABROUK, *Reconstructing an ancestral genome using minimum segments duplications and reversals*, Journal of Computer and System Sciences, 65, no. 3 (2002), pp. 442–464.
- [156] N. EL-MABROUK, *Genome rearrangements with gene families*, in Gascuel [182], pp. 291–320.
- [157] N. EL-MABROUK AND D. SANKOFF, *Hybridization and genome rearrangement*, in Proceedings of the Tenth Annual Symposium on Combinatorial Pattern Matching (CPM), M. Crochemore and M. Paterson, eds., vol. 1645 of Lecture Notes in Computer Science, Springer-Verlag, 1999, pp. 78–87.
- [158] N. EL-MABROUK AND D. SANKOFF, *The reconstruction of doubled genomes*, SIAM Journal of Computing, 32, no. 3 (2003), pp. 754–792.
- [159] N. EL-MABROUK, J. NADEAU, AND D. SANKOFF, *Genome halving*, in Proceedings of the Ninth Annual Symposium on Combinatorial Pattern Matching (CPM), G. Goos, J. Hartmanis, and J. van Leeuwen, eds., vol. 1448 of Lecture Notes in Computer Science, Springer-Verlag, 1998, pp. 235–250.
- [160] I. ELIAS AND T. HARTMAN, *A 1.375-approximation algorithm for sorting by transpositions*, IEEE/ACM Transactions on Computational Biology and Bioinformatics, 3, no. 4 (2006), pp. 369–379.
- [161] F. ERGUN, S. MUTHUKRISHNAN, AND S. C. SAHINALP, *Comparing sequences with segment rearrangements*, in Proceedings of the Twenty-third Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS), P. K. Pandya and J. Radhakrishnan, eds., vol. 2914 of Lecture Notes in Computer Science, Springer-Verlag, 2003, pp. 183–194.
- [162] N. ERIKSEN, *$(1 + \varepsilon)$ -approximation of sorting by reversals and transpositions*, Theoretical Computer Science, 289, no. 1 (2002), pp. 517–529.
- [163] N. ERIKSEN, *Reversal and transposition medians*, Theoretical Computer Science, 374, no. 1–3 (2007), pp. 111–126.
- [164] H. ERIKSSON, K. ERIKSSON, J. KARLANDER, L. SVENSSON, AND J. WÄSTLUND, *Sorting a bridge hand*, Discrete Mathematics, 241 (2001), pp. 289–300. Selected papers in honor of Helge Tverberg.
- [165] K. ERIKSSON, *Statistical and combinatorial aspects of comparative genomics*, Scandinavian Journal of Statistics, 31, no. 2 (2004), pp. 203–216.
- [166] V. ESTIVILL-CASTRO AND D. WOOD, *A survey of adaptive sorting algorithms*, ACM Computing Surveys, 24, no. 4 (1992), pp. 441–476.
- [167] L. EULER, *Solutio problematis ad geometriam situs pertinentis*, Commentarii Academiae Scientiarum Imperialis Petropolitanae, 8 (1736), pp. 128–140.
- [168] S. EVEN AND O. GOLDREICH, *The minimum-length generator sequence problem is NP-hard*, Journal of Algorithms, 2 (1981), pp. 311–313.

- [169] U. FEIGE, *A threshold of $\log n$ for approximating set cover*, Journal of the ACM, 45, no. 4 (1998), pp. 634–652.
- [170] J. FELSENSTEIN, *Inferring Phylogenies*, 2nd ed., Sinauer Associates, 2003.
- [171] J. FENG AND D. ZHU, *Faster algorithms for sorting by transpositions and sorting by block interchanges*, ACM Transactions on Algorithms, 3, no. 3 (2007), pp. 1–14.
- [172] W. FENG, L. WANG, AND D. ZHU, *CTRD: A fast applet for computing signed translocation distance between genomes*, Bioinformatics, 20, no. 17 (2004), pp. 3256–3257.
- [173] X. FENG, I. H. SUDBOROUGH, AND E. LU, *A fast algorithm for sorting by short swap*, in Proceedings of the IASTED International Conference on Computational and Systems Biology (CASB), D.-Z. Du, ed., ACTA Press, 2006, pp. 62–67.
- [174] V. FERRETTI, J. H. NADEAU, AND D. SANKOFF, *Original synteny*, in Proceedings of the Seventh Annual Symposium on Combinatorial Pattern Matching (CPM), D. S. Hirschberg and E. W. Myers, eds., vol. 1075 of Lecture Notes in Computer Science, Springer-Verlag, 1996, pp. 159–167.
- [175] M. FIGEAC AND J.-S. VARRÉ, *Sorting by reversals with common intervals*, in Proceedings of the Fourth Workshop on Algorithms in Bioinformatics (WABI), I. Jonassen and J. Kim, eds., vol. 3240 of Lecture Notes in Computer Science, Springer-Verlag, 2004, pp. 26–37.
- [176] J. FISCHER AND S. W. GINZINGER, *A 2-approximation algorithm for sorting by prefix reversals*, in Proceedings of the Thirteenth Annual European Symposium on Algorithms (ESA), G. S. Brodal and S. Leonardi, eds., vol. 3669 of Lecture Notes in Computer Science, Springer-Verlag, 2005, pp. 415–425.
- [177] V. J. FORTUNA, *Distâncias de Transposição entre Genomas*, master's thesis, Universidade Estadual de Campinas, Brazil, 2005.
- [178] L. R. FOULDS AND R. L. GRAHAM, *The Steiner tree problem in phylogeny is NP-complete*, Advances in Applied Mathematics, 3 (1982), pp. 43–49.
- [179] L. FROENICKE, M. G. CALDÉS, A. GRAPHODATSKY, S. MÜLLER, L. A. LYONS, T. J. ROBINSON, M. VOLLETH, F. YANG, AND J. WIENBERG, *Are molecular cytogenetics and bioinformatics suggesting diverging models of ancestral mammalian genomes?*, Genome Research, 16, no. 3 (2006), pp. 306–310.
- [180] Z. FU AND T. JIANG, *Computing the breakpoint distance between partially ordered genomes*, Journal of Bioinformatics and Computational Biology, 5, no. 5 (2007), pp. 1087–1101.
- [181] M. R. GAREY AND D. S. JOHNSON, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W. H. Freeman, 1979.
- [182] O. GASCUEL, ed., *Mathematics of Evolution and Phylogeny*, Oxford University Press, 2005.
- [183] W. H. GATES AND C. H. PAPADIMITRIOU, *Bounds for sorting by prefix reversal*, Discrete Mathematics, 27 (1979), pp. 47–57.
- [184] B. GAUT AND J. DOEBLEY, *DNA sequence evidence for the segmental allotetraploid origin of maize*, Proceedings of the National Academy of Sciences (USA), 94 (1997), pp. 6809–6814.
- [185] T. GIBSON AND J. SPRING, *Evidence in favor of ancient octaploidy in the vertebrate genome*, Biochemical Society Transactions, 28 (2000), pp. 259–264.
- [186] S. GOG AND M. BADER, *How to achieve an equivalent simple permutation in linear time*, in Proceedings of the Fifth RECOMB Comparative Genomics Satellite Workshop (RECOMB-CG), G. Tesler and D. Durand, eds., vol. 4751 of Lecture Notes in Computer Science, Springer-Verlag, 2007, pp. 58–68.
- [187] S. GOG, M. BADER, AND E. OHLEBUSCH, *GENESIS: Genome evolution scenarios*, Bioinformatics, 24, no. 5 (2008), pp. 711–712.
- [188] A. GOLDSTEIN, P. KOLMAN, AND J. ZHENG, *Minimum common string partition problem: Hardness and approximations*, Electronic Journal of Combinatorics, 12 (2005), no. R50.
- [189] J. GRAMM AND R. NIEDERMEIER, *Breakpoint medians and breakpoint phylogenies: A fixed-parameter approach*, Bioinformatics, 18, supp. 1 (2002), pp. S128–S139.
- [190] Q.-P. GU, S. PENG, AND Q. M. CHEN, *Sorting permutations and its applications in genome analysis*, Lectures on Mathematics in the Life Sciences, 26 (1999), pp. 191–201.
- [191] Q.-P. GU, S. PENG, AND I. H. SUDBOROUGH, *A 2-approximation algorithm for genome rearrangements by reversals and transpositions*, Theoretical Computer Science, 210 (1999), pp. 327–339.

- [192] S. A. GUYER, L. S. HEATH, AND J. P. VERGARA, *Subsequence and run heuristics for sorting by transpositions*, Technical Report, Virginia State University, 1997.
- [193] E. GYÖRI AND G. TURÁN, *Stack of pancakes*, *Studia Scientiarum Mathematicarum Hungarica*, 13 (1978), pp. 133–137.
- [194] Y. HAN, *Improving the efficiency of sorting by reversals*, in Proceedings of the 2006 International Conference on Bioinformatics and Computational Biology, CSREA Press, 2006, pp. 406–409.
- [195] S. HANNENHALLI, *Polynomial-time algorithm for computing translocation distance between genomes*, *Discrete Applied Mathematics*, 71 (1996), pp. 137–151.
- [196] S. HANNENHALLI AND P. A. PEVZNER, *Transforming men into mice: Polynomial algorithm for genomic distance problem*, in Proceedings of the Thirty-sixth Annual Symposium on Foundations of Computer Science (FOCS), IEEE Computer Society Press, 1995, pp. 581–592.
- [197] S. HANNENHALLI AND P. A. PEVZNER, *Towards a computational theory of genome rearrangements*, in *Computer Science Today: Recent Trends and Developments*, Jan van Leeuwen, ed., vol. 1000 of Lecture Notes in Computer Science, Springer-Verlag, 1995, pp. 184–202.
- [198] S. HANNENHALLI AND P. A. PEVZNER, *To cut . . . or not to cut (applications of comparative physical maps in molecular evolution)*, in Proceedings of the Seventh Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), Society for Industrial and Applied Mathematics, 1996, pp. 304–313.
- [199] S. HANNENHALLI AND P. A. PEVZNER, *Transforming cabbage into turnip: Polynomial algorithm for sorting signed permutations by reversals*, *Journal of the ACM*, 46, no. 1 (1999), pp. 1–27.
- [200] S. HANNENHALLI, C. CHAPPEY, E. V. KOONIN, AND P. A. PEVZNER, *Genome sequence comparison and scenarios for gene rearrangements: A test case*, *Genomics*, 30, no. 2 (1995), pp. 299–311.
- [201] T. HARTMAN, *A simpler 1.5-approximation algorithm for sorting by transpositions*, in Proceedings of the Fourteenth Annual Symposium on Combinatorial Pattern Matching (CPM), R. A. Baeza-Yates, E. Chávez, and M. Crochemore, eds., vol. 2676 of Lecture Notes in Computer Science, Springer-Verlag, 2003, pp. 156–169.
- [202] T. HARTMAN AND R. SHAMIR, *A simpler and faster 1.5-approximation algorithm for sorting by transpositions*, *Information and Computation*, 204, no. 2 (2006), pp. 275–290.
- [203] T. HARTMAN AND R. SHARAN, *A 1.5-approximation algorithm for sorting by transpositions and transreversals*, *Journal of Computer and System Sciences*, 70 (2005), pp. 300–320.
- [204] T. HARTMAN AND E. VERBIN, *Matrix tightness: A linear-algebraic framework for sorting by transpositions*, in Proceedings of the Thirteenth International Conference on String Processing and Information Retrieval (SPIRE), F. Crestani, P. Ferragina, and M. Sanderson, eds., vol. 4209 of Lecture Notes in Computer Science, Springer-Verlag, 2006, pp. 279–290.
- [205] J. HÅSTAD, *Clique is hard to approximate within $n^{1-\epsilon}$* , *Acta Mathematica*, 182 (1999), pp. 105–142.
- [206] D. HE, *A novel greedy algorithm for the minimum common string partition problem*, in Proceedings of the Third International Workshop on Bioinformatics Research and Applications (IWBRA), I. Mândoiu and A. Zelikovsky, eds., vol. 4463 of Lecture Notes in Computer Science, Springer-Verlag, 2007, pp. 441–452.
- [207] L. S. HEATH AND J. P. C. VERGARA, *Sorting by bounded block-moves*, *Discrete Applied Mathematics*, 88 (1998), pp. 181–206.
- [208] L. S. HEATH AND J. P. C. VERGARA, *Sorting by short block-moves*, *Algorithmica*, 28 (2000), pp. 323–352.
- [209] L. S. HEATH AND J. P. C. VERGARA, *Sorting by short swaps*, *Journal of Computational Biology*, 10, no. 2 (2003), pp. 775–789.
- [210] S. HEBER AND J. STOYE, *Finding all common intervals of k permutations*, in Proceedings of the Twelfth Annual Symposium on Combinatorial Pattern Matching (CPM), A. Amir and G. M. Landau, eds., vol. 2089 of Lecture Notes in Computer Science, Springer-Verlag, 2001, pp. 207–218.
- [211] S. HEDETNIEMI, S. HEDETNIEMI, AND A. LIESTMAN, *A survey of gossiping and broadcasting in communication networks*, *Networks*, 18 (1988), pp. 319–349.
- [212] M. H. HEYDARI AND I. H. SUDBOROUGH, *On the diameter of the pancake network*, *Journal of Algorithms*, 25, no. 1 (1997), pp. 67–94.

- [213] I. HOLYER, *The NP-completeness of some edge-partition problems*, SIAM Journal of Computing, 10 (1981), pp. 713–717.
- [214] A. HULTMAN, *Toric Permutations*, master's thesis, Department of Mathematics, Kungliga Tekniska Högskolan, Stockholm, 1999.
- [215] C. HURKENS, L. VAN IERSEL, J. KEIJSPER, S. KELK, L. STOUGIE, AND J. TROMP, *Prefix reversals on binary and ternary strings*, SIAM Journal on Discrete Mathematics, 21, no. 3 (2007), pp. 592–611.
- [216] Y. INTERIAN AND R. DURRETT, *Genomic midpoints: Computation and evolutionary implications*. Submitted.
- [217] Y. INTERIAN AND R. DURRETT, *Computing genomic midpoints*. <http://www.math.cornell.edu/~durrett/mdpt/mdpt.html>.
- [218] B. N. JACKSON, P. S. SCHNABLE, AND S. ALURU, *Consensus genetic maps as median orders from inconsistent sources*, IEEE/ACM Transactions on Computational Biology and Bioinformatics, 5, no. 2 (2008), pp. 161–171.
- [219] O. JAILLON, J.-M. AURY, F. BRUNET, J.-L. PETIT, N. STANGE-THOMANN, E. MAUCELLI, L. BOUNEAU, C. FISCHER, C. OZOUF-COSTAZ, A. BERNOT, S. NICAUD, D. JAFFE, S. FISHER, G. LUTFALLA, C. DOSSAT, B. SEGURENS, C. DASILVA, M. SALANOUBAT, M. LEVY, N. BOUDET, S. CASTELLANO, V. ANTHOUARD, C. JUBIN, V. CASTELLI, M. KATINKA, B. VACHERIE, C. BIÉMONT, Z. SKALLI, L. CATTOLICO, J. POULAIN, V. DE BERARDINIS, C. CRUAUD, S. DUPRAT, P. BROTTIER, J.-P. COUTANCEAU, J. GOUZY, G. PARRA, G. LARDIER, C. CHAPPELLE, K. J. MCKERNAN, P. MCEWAN, S. BOSAK, M. KELLIS, J.-N. VOLFF, R. GUIGÓ, M. C. ZODY, J. MESIROV, K. LINDBLAD-TOH, B. BIRREN, C. NUSBAUM, D. KAHN, M. ROBINSON-RECHAVI, V. LAUDET, V. SCHACHTER, F. QUÉTIER, W. SAURIN, C. SCARPELLI, P. WINCKER, E. S. LANDER, J. WEISSENBACH, AND H. R. CROLLIUS, *Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype*, Nature, 431, no. 7011 (2004), pp. 946–957.
- [220] G. JEAN AND M. NIKOLSKI, *Genome rearrangements: A correct algorithm for optimal capping*, Information Processing Letters, 104, no. 1 (2007), pp. 14–20.
- [221] M. R. JERRUM, *The complexity of finding minimum-length generator sequences*, Theoretical Computer Science, 36 (1985), pp. 265–289.
- [222] T. JIANG, Y. XU, AND M. Q. ZHANG, eds., *Current Topics in Computational Molecular Biology*, MIT Press, 2002.
- [223] D. JOHNSON, *Approximation algorithms for combinatorial problems*, Journal of Computer and System Sciences, 9 (1974), pp. 256–278.
- [224] N. C. JONES AND P. A. PEVZNER, *An Introduction to Bioinformatics Algorithms*, MIT Press, 2004.
- [225] V. KANN, *On the Approximability of NP-Complete Optimization Problems*, Ph.D. thesis, Department of Numerical Analysis and Computing Science, Royal Institute of Technology, Stockholm, 1992.
- [226] H. KAPLAN AND N. SHAFRIR, *The greedy algorithm for edit distance with moves*, Information Processing Letters, 97, no. 1 (2006), pp. 23–27.
- [227] H. KAPLAN AND E. VERBIN, *Sorting signed permutations by reversals, revisited*, Journal of Computer and System Sciences, 70, no. 3 (2005), pp. 321–341.
- [228] H. KAPLAN, R. SHAMIR, AND R. E. TARJAN, *A faster and simpler algorithm for sorting signed permutations by reversals*, SIAM Journal of Computing, 29, no. 3 (2000), pp. 880–892 (electronic).
- [229] D. KARGER, R. MOTWANI, AND M. SUDAN, *Approximate graph coloring by semidefinite programming*, Journal of the ACM, 45, no. 2 (1998), pp. 246–265.
- [230] R. M. KARP, *Reducibility among combinatorial problems*, in *Complexity of Computer Computations*, R. E. Miller and J. W. Thatcher, eds., Plenum, 1972, pp. 85–103.
- [231] M. KARPINSKI AND A. ZELIKOVSKY, *Approximating dense cases of covering problems*, Electronic Colloquium on Computational Complexity, tech. rep. TR97-004 (1997), p. 4.
- [232] J. KECECIOGLU AND D. SANKOFF, *Exact and approximation algorithms for sorting by reversals, with application to genome rearrangement*, Algorithmica, 13 (1995), pp. 180–210.
- [233] J. D. KECECIOGLU AND R. RAVI, *Of mice and men: Algorithms for evolutionary distances between genomes with translocation*, in *Proceedings of the Sixth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, Society for Industrial and Applied Mathematics, 1995, pp. 604–613.

- [234] M. KELLIS, B. BIRREN, AND E. LANDER, *Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae**, *Nature*, 428 (2004), pp. 617–624.
- [235] J. D. KLEIN, *Problem Kernels for Genome Rearrangement Problems*, diplomarbeit, Friedrich-Schiller-Universität Jena, 2007.
- [236] J. KLEINBERG AND D. LIBEN-NOWELL, *The syntenic diameter of the space of N -chromosome genomes*, in vol. 1 of Sankoff and Nadeau [322], pp. 185–197.
- [237] D. E. KNUTH, *Sorting and Searching*, vol. 3 of his *The Art of Computer Programming*, 2nd ed., Addison-Wesley, 1995.
- [238] P. KOLMAN, *Approximating reversal distance for strings with bounded number of duplicates*, in Proceedings of the Thirtieth International Symposium on Mathematical Foundations of Computer Science (MFCS), J. Jedrzejowicz and A. Szepietowski, eds., vol. 3618 of Lecture Notes in Computer Science, Springer-Verlag, 2005, pp. 580–590.
- [239] P. KOLMAN AND T. WALEŃ, *Reversal distance for strings with duplicates: Linear time approximation using hitting set*, in Proceedings of the Fourth International Workshop on Approximation and Online Algorithms (WAOA), T. Erlebach and C. Kaklamani, eds., vol. 4368 of Lecture Notes in Computer Science, Springer-Verlag, 2006, pp. 279–289.
- [240] A. LABARRE, *New bounds and tractable instances for the transposition distance*, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 3, no. 4 (2006), pp. 380–394.
- [241] A. LABARRE, *Edit distances and factorisations of even permutations*, in Proceedings of the Sixteenth Annual European Symposium on Algorithms (ESA), Lecture Notes in Computer Science, Springer-Verlag, 5193 (2008), pp. 635–646.
- [242] S. LAKSHMIVARAHAN, J.-S. JWO, AND S. K. DHALL, *Symmetry in interconnection networks based on Cayley graphs of permutation groups: A survey*, *Parallel Computing*, 19 (1993), pp. 361–407.
- [243] R. LANGKJÆR, P. CLIFTON, M. JOHNSTON, AND J. PISKUR, *Yeast genome duplication was followed by asynchronous differentiation of duplicated genes*, *Nature*, 421 (2003), pp. 848–852.
- [244] A. LEMPEL, *Matrix factorization over $GF(2)$ and trace-orthogonal bases of $GF(2^n)$* , *SIAM Journal of Computing*, 4, no. 2 (1975), pp. 175–186.
- [245] R. LENNE, C. SOLNON, T. STÜTZLE, E. TANNIER, AND M. BIRATTARI, *Reactive stochastic local search algorithms for the genomic median problem*, in Proceedings of the Eighth European Conference on Evolutionary Computation in Combinatorial Optimization (EVO-COP), J. I. van Hemert and C. Cotta, eds., vol. 4972 of Lecture Notes in Computer Science, Springer-Verlag, 2008, pp. 266–276.
- [246] G. LI, X. QI, X. WANG, AND B. ZHU, *A linear-time algorithm for computing translocation distance between signed genomes*, in Proceedings of the Fifteenth Annual Symposium on Combinatorial Pattern Matching (CPM), S. C. Sahinalp, S. Muthukrishnan, and U. Dogrusoz, eds., vol. 3109 of Lecture Notes in Computer Science, Springer-Verlag, 2004, pp. 323–332.
- [247] Z. LI, L. WANG, AND K. ZHANG, *Algorithmic approaches for genome rearrangement: A review*, *IEEE Transactions on Systems, Man and Cybernetics, part C*, 36, no. 5 (2006), pp. 636–648.
- [248] D. LIBEN-NOWELL, *On the structure of syntenic distance*, *Journal of Computational Biology*, 8, no. 1 (2001), pp. 53–67.
- [249] D. LIBEN-NOWELL, *Gossip is syntenic: Incomplete gossip and the syntenic distance between genomes*, *Journal of Algorithms*, 43, no. 2 (2002), pp. 264–283.
- [250] D. LIBEN-NOWELL AND J. KLEINBERG, *Structural properties and tractability results for linear syntenic*, *Journal of Discrete Algorithms*, 2, no. 2 (2004), pp. 207–228.
- [251] G.-H. LIN AND T. JIANG, *A further improved approximation algorithm for breakpoint graph decomposition*, *Journal of Combinatorial Optimization*, 8, no. 4 (2004), pp. 183–194.
- [252] G.-H. LIN AND G. XUE, *Signed genome rearrangement by reversals and transpositions: Models and approximations*, *Theoretical Computer Science*, 259, no. 1–2 (2001), pp. 513–531.
- [253] G.-H. LIN AND G. XUE, *On the terminal Steiner tree problem*, *Information Processing Letters*, 84, no. 2 (2002), pp. 103–107.
- [254] Y. C. LIN, C. L. LU, H.-Y. CHANG, AND C. Y. TANG, *An efficient algorithm for sorting by block-interchanges and its application to the evolution of *Vibrio* species*, *Journal of Computational Biology*, 12, no. 1 (2005), pp. 102–112.

- [255] Y. C. LIN, C. L. LU, Y.-C. LIU, AND C. Y. TANG, *SPRING: A tool for the analysis of genome rearrangement using reversals and block-interchanges*, *Nucleic Acids Research*, 34 (July 1, 2006), pp. W696–699.
- [256] Y. C. LIN, C. L. LU, AND C. Y. TANG, *Sorting permutation by reversals with fewest block-interchanges*, manuscript, 2006.
- [257] D. LOPRESTI AND A. TOMKINS, *Block edit models for approximate string matching*, *Theoretical Computer Science*, 181 (1997), pp. 159–179.
- [258] C. L. LU, T. C. WANG, Y. C. LIN, AND C. Y. TANG, *ROBIN: A tool for genome rearrangement of block-interchanges*, *Bioinformatics*, 21 (2005), pp. 2780–2782.
- [259] C. L. LU, Y. L. HUANG, T. C. WANG, AND H.-T. CHIU, *Analysis of circular genome rearrangement by fusions, fissions and block-interchanges*, *BMC Bioinformatics*, 7, no. 1 (2006).
- [260] L. LUNDIN, *Evolution of vertebrate genome reflected in paralogous chromosomal regions in man and in the house mouse*, *Genomics*, 16 (1993), pp. 1–19.
- [261] M. MAHAJAN, R. RAMA, V. RAMAN, AND S. VIJAYAKUMAR, *Approximate block sorting*, *International Journal of Foundations of Computer Science*, 17, no. 2 (2006), pp. 337–355.
- [262] M. MAHAJAN, R. RAMA, AND S. VIJAYAKUMAR, *On sorting by 3-bounded transpositions*, *Discrete Mathematics*, 306, no. 14 (2006), pp. 1569–1585.
- [263] M. MARRON, K. M. SWENSON, AND B. M. E. MORET, *Genomic distances under deletions and insertions*, *Theoretical Computer Science*, 325, no. 3 (2004), pp. 347–360.
- [264] J. MEIDANIS AND Z. DIAS, *An alternative algebraic formalism for genome rearrangements*, in vol. 1 of Sankoff and Nadeau [322], pp. 213–223.
- [265] J. MEIDANIS, M. WALTER, AND Z. DIAS, *Reversal distance of signed circular chromosomes*, tech. rep. IC-00-23, University of Campinas, Brazil, 2000.
- [266] J. MEIDANIS, M. E. M. T. WALTER, AND Z. DIAS, *A lower bound on the reversal and transposition diameter*, *Journal of Computational Biology*, 9, no. 5 (2002), pp. 743–745.
- [267] H. MEWES, K. ALBERMANN, M. BAHR, D. FRISHMAN, A. GLEISSNER, J. HANI, K. HEUMANN, K. KLEINE, A. MAIERL, S. OLIVER, F. PFEIFFER, AND A. ZOLLNER, *Overview of the yeast genome*, *Nature*, 387, 6632, supp. (1997), pp. S7–S65.
- [268] S. MICALI AND V. V. VAZIRANI, *An $O(\sqrt{|V|}|E|)$ algorithm for finding maximum matching in general graphs*, in *Proceedings of the Twenty-first Annual Symposium on Foundations of Computer Science (FOCS)*, IEEE Computer Society Press, 1980, pp. 17–27.
- [269] C. MIRA AND J. MEIDANIS, *Sorting by block-interchanges and signed reversals*, in *Proceedings of the Fourth International Conference on Information Technology: New Generations (ITNG)*, IEEE Computer Society Press, 2007, pp. 670–676.
- [270] C. MIRA, Q. PENG, J. MEIDANIS, AND P. A. PEVZNER, *A shortest-cycle based heuristics for the multiple genome rearrangement problem*, tech. rep. IC-06-23, Instituto de Computação, Universidade Estadual de Campinas, Brazil, 2006.
- [271] J. MIXTACKI, *Genome halving under DCJ revisited*, in *Proceedings of the Fourteenth International Computing and Combinatorics Conference (COCOON)*, X. Hu and J. Wang, eds., vol. 5092 of *Lecture Notes in Computer Science*, Springer-Verlag, 2008, pp. 276–286.
- [272] Z. MO AND T. ZENG, *An improved genetic algorithm for problem of genome rearrangement*, *Wuhan University Journal of Natural Sciences*, 11, no. 3 (2006), pp. 498–502.
- [273] B. M. E. MORET, L.-S. WANG, T. WARNOW, AND S. K. WYMAN, *New approaches for reconstructing phylogenies from gene order data*, *Bioinformatics*, 17 (2001), pp. S165–S173.
- [274] B. M. E. MORET, S. K. WYMAN, D. A. BADER, T. WARNOW, AND M. YAN, *A new implementation and detailed study of breakpoint analysis*, in *Proceedings of the Sixth Pacific Symposium on Biocomputing (PSB)*, World Scientific Press, 2001, pp. 583–594.
- [275] B. M. E. MORET, A. C. SIEPEL, J. TANG, AND T. LIU, *Inversion medians outperform breakpoint medians in phylogeny reconstruction from gene-order data*, in *Proceedings of the Second International Workshop on Algorithms in Bioinformatics (WABI)*, R. Guigo and D. Gusfield, eds., vol. 2452 of *Lecture Notes in Computer Science*, Springer-Verlag, 2002, pp. 521–536.

- [276] B. M. E. MORET, J. TANG, L.-S. WANG, AND T. WARNOW, *Steps toward accurate reconstructions of phylogenies from gene-order data*, Journal of Computer and System Sciences, 65, no. 3 (2002), pp. 508–525.
- [277] W. J. MURPHY, D. M. LARKIN, A. EVERTS-VAN DER WIND, G. BOURQUE, G. TESLER, L. AUVIL, J. E. BEEVER, B. P. CHOWDHARY, F. GALIBERT, L. GATZKE, C. HITTE, S. N. MEYERS, D. MILAN, E. A. OSTRANDER, G. PAPE, H. G. PARKER, T. RAUDSEPP, M. B. ROGATCHEVA, L. B. SHOOK, L. C. SKOW, M. WELGE, J. E. WOMACK, S. J. O'BRIEN, P. A. PEVZNER, AND H. A. LEWIN, *Dynamics of mammalian chromosome evolution inferred from multi-species comparative maps*, Science, 309 (2005), pp. 613–617.
- [278] S. MUTHUKRISHNAN AND S. C. SAHINALP, *An efficient algorithm for sequence comparison with block reversals*, Theoretical Computer Science, 321, no. 1 (2004), pp. 95–101.
- [279] C. T. NGUYEN, *Algorithms for Calculating Exemplar Distances*, honours year project report, National University of Singapore, 2005.
- [280] C. T. NGUYEN, Y. C. TAY, AND L. ZHANG, *Divide-and-conquer approach for the exemplar breakpoint distance*, Bioinformatics, 21, no. 10 (2005), pp. 2171–2176.
- [281] T. C. NGUYEN, H. T. NGO, AND N. B. NGUYEN, *Sorting by restricted-length-weighted reversals*, Genomics Proteomics Bioinformatics, 3, no. 2 (2005), pp. 120–127.
- [282] R. NIEDERMEIER, *Invitation to Fixed-Parameter Algorithms*, Oxford Lecture Series in Mathematics and Its Applications, Oxford University Press, 2006.
- [283] E. OHLEBUSCH, M. I. ABOUELHODA, K. HOCKEL, AND J. STALLKAMP, *The median problem for the reversal distance in circular bacterial genomes*, in Proceedings of the Sixteenth Annual Symposium on Combinatorial Pattern Matching (CPM), A. Apostolico, M. Crochemore and K. Park, eds., vol. 3537 of Lecture Notes in Computer Science, Springer-Verlag, 2005, pp. 116–127.
- [284] S. OHNO, *Evolution by Gene Duplication*, Springer-Verlag, 1970.
- [285] M. OZERY-FLATO AND R. SHAMIR, *Two notes on genome rearrangement*, Journal of Bioinformatics and Computational Biology, 1, no. 1 (2003), pp. 71–94.
- [286] M. OZERY-FLATO AND R. SHAMIR, *An $O(n^{3/2} \sqrt{\log(n)})$ algorithm for sorting by reciprocal translocations*, in Proceedings of the Seventeenth Annual Symposium on Combinatorial Pattern Matching (CPM), M. Lewenstein and G. Valiente, eds., vol. 4009 of Lecture Notes in Computer Science, Springer-Verlag, 2006, pp. 259–269.
- [287] M. OZERY-FLATO AND R. SHAMIR, *Rearrangements in genomes with centromeres part I: Translocations*, in Proceedings of the Eleventh Annual International RECOMB Conference (RECOMB), T. Speed and H. Huang, eds., vol. 4453 of Lecture Notes in Computer Science, Springer-Verlag, 2007, pp. 339–353.
- [288] M. OZERY-FLATO AND R. SHAMIR, *Sorting by reciprocal translocations via reversals theory*, Journal of Computational Biology, 14, no. 4 (2007), pp. 408–422.
- [289] J. D. PALMER AND L. A. HERBON, *Plant mitochondrial DNA evolves rapidly in structure, but slowly in sequence*, Journal of Molecular Evolution, 28, no. 1–2 (1988), pp. 87–97.
- [290] C. H. PAPANITRIOU, *Computational Complexity*, Addison-Wesley, 1994.
- [291] C. H. PAPANITRIOU AND M. YANNAKAKIS, *Optimization, approximation and complexity classes*, Journal of Computer and System Sciences, 43 (1991), pp. 425–440.
- [292] I. PE'ER AND R. SHAMIR, *The median problems for breakpoints are NP-complete*, Electronic Colloquium on Computational Complexity, tech. rep. 71, 1998.
- [293] I. PE'ER AND R. SHAMIR, *Approximation algorithms for the permutations median problem in the Breakpoint model*, in Sankoff and Nadeau [322], pp. 225–241.
- [294] Q. PENG, P. A. PEVZNER, AND G. TESLER, *The fragile breakage versus random breakage models of chromosome evolution*, PLoS Computational Biology, 2, no. 2 (2006), p. e14.
- [295] P. PEVZNER AND G. TESLER, *Transforming men into mice: The Nadeau-Taylor chromosomal breakage model revisited*, in Proceedings of the Seventh Annual International Conference on Research in Computational Molecular Biology (RECOMB), ACM Press, 2003, pp. 247–256.
- [296] P. A. PEVZNER, *Computational Molecular Biology: An Algorithmic Approach*, MIT Press, 2000.
- [297] P. A. PEVZNER AND G. TESLER, *Human and mouse genomic sequences reveal extensive breakpoint reuse in mammalian evolution*, Proceedings of the National Academy of Sciences (USA), 100 (2003), pp. 7672–7677.

- [298] P. A. PEVZNER AND M. S. WATERMAN, *Open combinatorial problems in computational molecular biology*, in Third Israel Symposium on the Theory of Computing and Systems, IEEE Computer Society Press, 1995, pp. 158–173.
- [299] R. Y. PINTER AND S. SKIENA, *Genomic sorting with length-weighted reversals*, in Proceedings of the Thirteenth International Conference on Genome Informatics (GIW), R. Lathrop, K. Nakai, S. Miyano, T. Takagi, and M. Kanehisa, eds., vol. 13, Universal Academy Press, 2002, pp. 103–111.
- [300] N. PISANTI AND M.-F. SAGOT, *Further thoughts on the syntenic distance between genomes*, *Algorithmica*, 34, no. 2 (2002), pp. 157–180.
- [301] P. POPESCU, H. HAYES, AND B. DUTRILLAUX, *Techniques in Animal Cytogenetics*, Springer-Verlag, 2000.
- [302] V. Y. POPOV, *Multiple genome rearrangement by swaps and by element duplications*, *Theoretical Computer Science*, 385, no. 1–3 (2007), pp. 115–126.
- [303] F. J. PORTIER AND T. P. VAUGHAN, *Whitney numbers of the second kind for the star poset*, *European Journal of Combinatorics*, 11 (1990), pp. 277–288.
- [304] X. QI, G. LI, J. WU, AND B. LIU, *Sorting signed permutations by fixed-length reversals*, *International Journal of Foundations of Computer Science*, 17, no. 4 (2006), pp. 933–948.
- [305] X. QI, G. LI, S. LI, AND Y. XU, *Sorting genomes by translocations and deletions*, in Proceedings of the Fifth International IEEE Computer Society Computational Systems Bioinformatics Conference (CSB), P. Markstein and Y. Xu, eds., IEEE Computer Society Press, 2006, pp. 157–166.
- [306] X.-Q. QI, *Combinatorial Algorithms of Genome Rearrangements in Bioinformatics*, Ph.D. thesis, University of Shandong, China, 2006.
- [307] X.-Q. QI, J. CAO, AND C. ZHANG, *Sorting circular binary strings with length weighted transpositions*, *Journal of Shandong Normal University (Natural Science)*, 41 (2006), pp. 82–85. In Chinese.
- [308] K. QIU, H. MELJER, AND S. AKL, *Parallel routing and sorting of the pancake network*, in Proceedings of the Third International Conference on Computing and Information (ICCI), F. K. H. A. Dehne, F. Fiala, and W. W. Koczkodaj, eds., vol. 497 of Lecture Notes in Computer Science, Springer-Verlag, 1991, pp. 360–371.
- [309] A. J. RADCLIFFE, A. D. SCOTT, AND E. L. WILMER, *Reversals and transpositions over finite alphabets*, *SIAM Journal on Discrete Mathematics*, 19, no. 1 (2005), pp. 224–244 (electronic).
- [310] A. RAHMAN, S. SHATABDA, AND M. HASAN, *Approximation algorithm for sorting by reversals and transpositions*, in Proceedings of the First Workshop on Algorithms and Computation (WALCOM), M. Kaykobad and M. S. Rahman, eds., Bangladesh Academy of Sciences, 2007, pp. 97–108.
- [311] R. RAZ AND S. SAFRA, *A sub-constant error-probability low-degree test, and sub-constant error-probability PCP characterization of NP*, in Proceedings of the Twenty-ninth Annual ACM Symposium on the Theory of Computing (STOC), ACM, 1997, pp. 475–484.
- [312] K. RICE AND T. WARNOW, *Parsimony is hard to beat*, in Proceedings of the Third International Computing and Combinatorics Conference (COCOON), T. Jiang and D. T. Lee, eds., vol. 1276 of Lecture Notes in Computer Science, Springer-Verlag, 1997, pp. 124–133.
- [313] G. ROBINS AND A. ZELIKOVSKY, *Improved Steiner tree approximation in graphs*, in Proceedings of the Eleventh Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), San Francisco, PA, USA, Society for Industrial and Applied Mathematics, 2000, pp. 770–779.
- [314] M.-F. SAGOT AND E. TANNIER, *Perfect sorting by reversals*, in Proceedings of the Eleventh International Computing and Combinatorics Conference (COCOON), L. Wang, ed., vol. 3595 of Lecture Notes in Computer Science, Springer-Verlag, 2005, pp. 42–51.
- [315] D. SANKOFF, *Edit distance for genome comparison based on non-local operations*, in Proceedings of the Third Annual Symposium on Combinatorial Pattern Matching (CPM), A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, eds., vol. 644 of Lecture Notes in Computer Science, Springer-Verlag, 1992, pp. 121–135.
- [316] D. SANKOFF, *Genome rearrangement with gene families*, *Bioinformatics*, 15, no. 11 (1999), pp. 909–917.

- [317] D. SANKOFF AND M. BLANCHETTE, *The median problem for breakpoints in comparative genomics*, in Proceedings of the Third International Computing and Combinatorics Conference (COCOON), T. Jiang and D. T. Lee, eds., vol. 1276 of Lecture Notes in Computer Science, Springer-Verlag, 1997, pp. 251–263.
- [318] D. SANKOFF AND M. BLANCHETTE, *Multiple genome rearrangement and breakpoint phylogeny*, Journal of Computational Biology, 5, no. 3 (1998), pp. 555–570.
- [319] D. SANKOFF AND M. BLANCHETTE, *Probability models for genome rearrangement and linear invariants for phylogenetic inference*, in Proceedings of the Third Annual International Conference on Computational Molecular Biology (RECOMB), S. Istrail, P. Pevzner, and M. Waterman, eds., ACM, 1999, pp. 302–309.
- [320] D. SANKOFF AND N. EL-MABROUK, *Genome rearrangement*, in Jiang et al. [222].
- [321] D. SANKOFF AND L. HAQUE, *Power boosts for cluster tests*, in Proceedings of the Third RECOMB Comparative Genomics Satellite Workshop (RECOMB-CG), A. McLysaght and D. H. Huson, eds., vol. 3678 of Lecture Notes in Bioinformatics, Springer-Verlag, 2005, pp. 121–130.
- [322] D. SANKOFF AND J. H. NADEAU, eds., *Comparative Genomics: Empirical and Analytical Approaches to Gene Order Dynamics, Map Alignment and the Evolution of Gene Families*, vol. 1 of Computational Biology, Kluwer Academic Press, 2000.
- [323] D. SANKOFF AND P. TRINH, *Chromosomal breakpoint re-use in genome sequence rearrangement*, Journal of Computational Biology, 12 (2005), pp. 812–821.
- [324] D. SANKOFF, C. MOREL, AND R. J. CEDERGREN, *Evolution of 5S RNA and the non-randomness of base replacement*, Nature New Biology, 245 (October 1973), pp. 232–234.
- [325] D. SANKOFF, R. J. CEDERGREN, AND G. LAPALME, *Frequency of insertion-deletion, transversion, and transition in the evolution of 5s ribosomal RNA*, Journal of Molecular Evolution, 7 (1976), pp. 133–149.
- [326] D. SANKOFF, G. LEDUC, N. ANTOINE, B. PAQUIN, F. LANG, AND R. CEDERGREN, *Gene order comparisons for phylogenetic inference: Evolution of the mitochondrial genome*, Proceedings of the National Academy of Sciences (USA), 89 (1992), pp. 6575–6579.
- [327] D. SANKOFF, G. SUNDARAM, AND J. D. KECECIOGLU, *Steiner points in the space of genome rearrangements*, International Journal of the Foundations of Computer Science, 7 (1996), pp. 1–9.
- [328] D. SANKOFF, D. BRYANT, M. DENEULT, B. F. LANG, AND G. BURGER, *Early eukaryote evolution based on mitochondrial gene order breakpoints*, in Proceedings of the Fourth Annual International Conference on Computational Molecular Biology (RECOMB), R. Shamir, S. Miyano, S. Istrail, P. Pevzner, and M. Waterman, eds., ACM, 2000, pp. 254–262.
- [329] D. SANKOFF, J.-F. LEFEBVRE, E. TILLIER, A. MALER, AND N. EL-MABROUK, *The distribution of inversion lengths in bacteria*, in Proceedings of the Second RECOMB Comparative Genomics Satellite Workshop (RECOMB-CG), J. Lagergren, ed., vol. 3388 of Lecture Notes in Computer Science, Springer-Verlag, 2004, pp. 97–108.
- [330] T. SCHMIDT AND J. STOYE, *Quadratic time algorithms for finding common intervals in two and more sequences*, in Proceedings of the Fifteenth Annual Symposium on Combinatorial Pattern Matching (CPM), S. C. Sahinalp, S. Muthukrishnan, and U. Dogrusoz, eds., vol. 3109 of Lecture Notes in Computer Science, Springer-Verlag, 2004, pp. 347–358.
- [331] C. SEOIGHE AND K. WOLFE, *Extent of genomic rearrangement after genome duplication in yeast*, Proceedings of the National Academy of Sciences (USA), 95 (1998), pp. 4447–4452.
- [332] G. SEROUSSI AND A. LEMPEL, *Factorization of symmetric matrices and trace-orthogonal bases in finite fields*, SIAM Journal of Computing, 9 (1980), pp. 758–767.
- [333] J. MEIDANIS AND J. SETUBAL, *Introduction to Computational Molecular Biology*, PWS Publishing, 1997.
- [334] P. SEYMOUR, *Packing directed circuits fractionally*, Combinatorica, 15, no. 2 (1995), pp. 281–288.
- [335] D. SHAPIRA AND J. A. STORER, *Edit distance with move operations*, in Proceedings of the Thirteenth Annual Symposium on Combinatorial Pattern Matching (CPM), vol. 2373 of Lecture Notes in Computer Science, Springer-Verlag, 2002, pp. 85–98.
- [336] D. SHAPIRA AND J. A. STORER, *Large edit distance with multiple block operations*, in Proceedings of the Tenth International Symposium on String Processing and Information Retrieval (SPIRE), M. A.

- Nascimento, E. S. de Moura, and A. L. Oliveira, eds., vol. 2857 of Lecture Notes in Computer Science, Springer-Verlag, 2003, pp. 369–377.
- [337] D. SHAPIRA AND J. A. STORER, *Edit distance with move operations*, Journal of Discrete Algorithms, 5, no. 2 (2007), pp. 380–392.
- [338] Y.-F. SHE AND G.-L. CHEN, *Parallel algorithm for computing reversal distance*, Proceedings of the Sixth International Conference on Parallel and Distributed Computing Applications and Technologies, 2005, pp. 950–953.
- [339] A. C. SIEPEL, *Exact Algorithms for the Reversal Median Problem*, master's thesis, University of New Mexico, 2001.
- [340] A. C. SIEPEL, *An algorithm to enumerate sorting reversals for signed permutations*, Journal of Computational Biology, 10 (2003), pp. 575–597.
- [341] A. C. SIEPEL AND B. M. E. MORET, *Finding an optimal inversion median: Experimental results*, in Proceedings of the First International Workshop on Algorithms in Bioinformatics (WABI), O. Gascuel and B. M. E. Moret, eds., vol. 2149 of Lecture Notes in Computer Science, Springer-Verlag, 2001, pp. 189–203.
- [342] A. SOLOMON, P. SUTCLIFFE, AND R. LISTER, *Sorting circular permutations by reversal*, in Proceedings of the Eighth International Workshop on Algorithms and Data Structures (WADS), F. K. H. A. Dehne, J.-R. Sack, and M. H. M. Smid, eds., vol. 2748 of Lecture Notes in Computer Science, Springer-Verlag, 2003, pp. 319–328.
- [343] A. H. STURTEVANT AND E. NOVITSKI, *The homologies of the chromosome elements in the genus Drosophila*, Genetics, 26, no. 5 (1941), pp. 517–541.
- [344] J. SUKSAWATCHON, C. LURSINSAP, AND M. BODÉN, *Computing the reversal distance between genomes in the presence of multi-gene families via binary integer programming*, Journal of Bioinformatics and Computational Biology, 5, no. 1 (2007), pp. 117–133.
- [345] K. SWENSON, M. MARRON, J. EARNEST-DEYOUNG, AND B. M. E. MORET, *Approximating the true evolutionary distance between two genomes*, in Proceedings of the Seventh Workshop on Algorithm Engineering and Experiments and the Second Workshop on Analytic Algorithmics and Combinatorics (ALENEX/ANALCO), C. Demetrescu, R. Sedgewick, and R. Tamassia, eds., SIAM, 2005, pp. 121–129.
- [346] F. SWIDAN, M. A. BENDER, D. GE, S. HE, H. HU, AND R. PINTER, *Sorting by length-weighted reversals: Dealing with signs and circularity*, in Proceedings of the Fifteenth Annual Symposium on Combinatorial Pattern Matching (CPM), S. C. Sahinalp, S. Muthukrishnan, and U. Dogrusoz, eds., vol. 3109 of Lecture Notes in Computer Science, Springer-Verlag, 2004, pp. 32–46.
- [347] J. TANG AND B. M. E. MORET, *Phylogenetic reconstruction from gene-rearrangement data with unequal gene content*, in Proceedings of the Eighth International Workshop on Algorithms and Data Structures (WADS), F. K. H. A. Dehne, J.-R. Sack, and M. H. M. Smid, eds., vol. 2748 of Lecture Notes in Computer Science, Springer-Verlag, 2003, pp. 37–46.
- [348] J. TANG AND B. M. E. MORET, *Linear programming for phylogenetic reconstruction based on gene rearrangements*, in Proceedings of the Sixteenth Annual Symposium on Combinatorial Pattern Matching (CPM), A. Apostolico, M. Crochemore and K. Park, eds., vol. 3537 of Lecture Notes in Computer Science, Springer-Verlag, 2005, pp. 406–416.
- [349] J. TANG AND L.-S. WANG, *Improving genome rearrangement phylogeny using sequence-style parsimony*, in Proceedings of the Fifth IEEE International Symposium on Bioinformatics and Bioengineering (BIBE), IEEE Computer Society Press, 2005, pp. 137–144.
- [350] J. TANG, B. M. E. MORET, L. CUI, AND C. W. DEPAMPHILIS, *Phylogenetic reconstruction from arbitrary gene-order data*, in Proceedings of the Fourth IEEE International Symposium on Bioinformatics and Bioengineering (BIBE), IEEE Computer Society Press, 2004, pp. 592–599.
- [351] E. TANNIER AND M.-F. SAGOT, *Sorting by reversals in subquadratic time*, in Proceedings of the Fifteenth Annual Symposium on Combinatorial Pattern Matching (CPM), S. C. Sahinalp, S. Muthukrishnan, and U. Dogrusoz, eds., vol. 3109 of Lecture Notes in Computer Science, Springer-Verlag, 2004, pp. 1–13.
- [352] E. TANNIER, A. BERGERON, AND M.-F. SAGOT, *Advances on sorting by reversals*, Discrete Applied Mathematics, 155, no. 6–7 (2007), pp. 881–888.

- [353] E. TANNIER, C. ZHENG, AND D. SANKOFF, *Multichromosomal median and halving problems*, in Proceedings of the Eighth Workshop on Algorithms in Bioinformatics (WABI), vol. 5251 of Lecture Notes in Bioinformatics, Springer-Verlag, 2008.
- [354] B. TENNER, *Database of permutation pattern avoidance*, published electronically at <http://math.depaul.edu/~bridget/patterns.html>.
- [355] G. TESLER, *Efficient algorithms for multichromosomal genome rearrangements*, Journal of Computer and System Sciences, 65, no. 3 (2002), pp. 587–609.
- [356] G. TESLER, *GRIMM: Genome rearrangements web server*, Bioinformatics, 18, no. 3 (2002), pp. 492–493.
- [357] W. F. TICHY, *The string-to-string correction problem with block moves*, ACM Transactions on Computer Systems, 2, no. 4 (1984), pp. 309–321.
- [358] C. TING AND H. E. YONG, *Optimal algorithms for uncovering synteny problem*, Journal of Combinatorial Optimization, 12 (2006), pp. 421–432.
- [359] N. TRAN, *An easy case of sorting by reversals*, in Proceedings of the Eighth Annual Symposium on Combinatorial Pattern Matching (CPM), A. Apostolico and J. Hein, eds., vol. 1264 of Lecture Notes in Computer Science, Springer-Verlag, 1997, pp. 83–89.
- [360] C.-W. TSENG AND M. V. ZELKOWITZ, eds., *Computational Biology and Bioinformatics*, vol. 68 of Advances in Computers, Elsevier Academic Press, 2006.
- [361] T. UNO AND M. YAGIURA, *Fast algorithms to enumerate all common intervals of two permutations*, Algorithmica, 26, no. 2 (2000), pp. 290–309.
- [362] F. VIDUANI MARTINEZ, J. COELHO DE PINA, AND J. SOARES, *Algorithms for terminal Steiner trees*, in Proceedings of the Eleventh International Computing and Combinatorics Conference (COCOON), L. Wang, ed., vol. 3595 of Lecture Notes in Computer Science, Springer-Verlag, 2005, pp. 369–379.
- [363] M. E. M. T. WALTER, Z. DIAS, AND J. MEIDANIS, *Reversal and transposition distance of linear chromosomes*, in Proceedings of the Fifth International Symposium on String Processing and Information Retrieval (SPIRE), IEEE Computer Society Press, 1998, pp. 96–102.
- [364] L. WANG, *Translocation distance: Algorithms and complexity*, in Tseng and Zelikowitz [360], pp. 106–127.
- [365] L. WANG, D. ZHU, X. LIU, AND S. MA, *An $O(n^2)$ algorithm for signed translocation*, Journal of Computer and System Sciences, 70, no. 3 (2005), pp. 284–299.
- [366] L.-S. WANG, R. K. JANSEN, B. M. E. MORET, L. A. RAUBESON, AND T. WARNOW, *Fast phylogenetic methods for the analysis of genome rearrangement data: An empirical study*, in Proceedings of the Seventh Pacific Symposium on Biocomputing (PSB), R. B. Altman, A. K. Dunker, L. Hunter, K. Lauderdale, and T. E. Klein, eds., World Scientific Press, 2002, pp. 524–535.
- [367] L.-S. WANG, T. WARNOW, B. M. E. MORET, R. K. JANSEN, AND L. A. RAUBESON, *Distance-based genome rearrangement phylogeny*, Journal of Molecular Evolution, 63, no. 4 (2006), pp. 473–483.
- [368] R. WARREN AND D. SANKOFF, *Genome halving with general operations*, in Proceedings of the Sixth Asia-Pacific Bioinformatics Conference (APBC), A. Brazma, S. Miyano, and T. Akutsu, eds., vol. 6 of Advances in Bioinformatics and Computational Biology, Imperial College Press, 2008, pp. 231–240.
- [369] G. A. WATTERSON, W. J. EWENS, T. E. HALL, AND A. MORGAN, *The chromosome inversion problem*, Journal of Theoretical Biology, 99 (1982), pp. 1–7.
- [370] H. WIELANDT, *Finite Permutation Groups*, translated from German by R. Bercov, Academic Press, 1964.
- [371] J. WIENBERG, *The evolution of eutherian chromosomes*, Current Opinion in Genetics & Development, 14 (2004), pp. 657–666.
- [372] K. WOLFE AND D. SHIELDS, *Molecular evidence for an ancient duplication of the entire yeast genome*, Nature, 387 (1997), pp. 708–713.
- [373] S. WU AND X. GU, *Multiple genome rearrangement by reversals*, in Proceedings of the Seventh Pacific Symposium on Biocomputing (PSB), R. B. Altman, A. K. Dunker, L. Hunter, and T. E. Klein, eds., World Scientific Press, 2002, pp. 259–270.

- [374] S. WU AND X. GU, *Algorithms for multiple genome rearrangement by signed reversals*, in Proceedings of the Eighth Pacific Symposium on Biocomputing (PSB), World Scientific Press, 2003, pp. 363–374.
- [375] S. YANCOPOULOS, O. ATTIE, AND R. FRIEDBERG, *Efficient sorting of genomic permutations by translocation, inversion and block interchange*, *Bioinformatics*, 21, no. 16 (2005), pp. 3340–3346.
- [376] I. V. YAP, D. SCHNEIDER, J. KLEINBERG, D. MATTHEWS, S. CARTINHOOR, AND S. R. MCCOUCH, *A graph-theoretic approach to comparing and integrating genetic, physical and sequence-based maps*, *Genetics*, 165 (2003), pp. 2235–2247.
- [377] P. YIN AND A. J. HARTEMINK, *Theoretical and practical advances in genome halving*, *Bioinformatics*, 21 (2005), pp. 869–879.
- [378] F. YUE, M. ZHANG, AND J. TANG, *A heuristic for phylogenetic reconstruction using transposition*, in Proceedings of the Seventh IEEE International Conference on Bioinformatics and Bioengineering (BIBE), J. Y. Yang, M. Qu Yang, M. M. Zhu, Y. Zhang, H. R. Arabnia, Y. Deng, and N. G. Bourbakis, eds., IEEE Computer Society Press, 2007, pp. 802–808.
- [379] C. ZHENG AND D. SANKOFF, *Genome rearrangements with partially ordered chromosomes*, *Journal of Combinatorial Optimization*, 11, no. 2 (2006), pp. 133–144.
- [380] C. ZHENG, A. LENERT, AND D. SANKOFF, *Reversal distance for partially ordered genomes*, *ISMB (supplement of Bioinformatics)*, 21 (2005), pp. i502–i508.
- [381] C. ZHENG, Q. ZHU, AND D. SANKOFF, *Genome halving with an outgroup*, *Evolutionary Bioinformatics*, 2 (2006), pp. 295–302.
- [382] C. ZHENG, Q. ZHU, Z. ADAM, AND D. SANKOFF, *Guided genome halving: Hardness, heuristics and the history of the hemiascomycetes*, *ISMB (supplement to Bioinformatics)*, 24, no. 13 (2008), pp. i96–i104.
- [383] D. ZHU AND S. MA, *Improved polynomial-time algorithm for computing translocation distance between genomes*, *J. Comput.*, 25, no. 2 (2002), pp. 189–196. In Chinese.
- [384] D. ZHU AND L. WANG, *On the complexity of unsigned translocation distance*, *Theoretical Computer Science*, 352, no. 1–3 (2006), pp. 322–328.