In the following we will examine restricted models considering linear genomes that are singular. We start by describing in more detail the special operations that occur in restricted scenarios.

6.2.1 Transposition and block-interchange via (ei)-composition

As we have seen, besides the structural rearrangements that correspond to a single DCJ operation, some additional rearrangements correspond to two DCJ operations. A *block-interchange* occurs when two segments exchange their positions. A particular case is a *transposition*, in which one of the two exchanged segments is empty. When a block interchange affects one single chromosome it is said to be *internal*, otherwise *external*. These rearrangements require at least three distinct cuts and cannot be represented by a single DCJ operation. Instead, they can be obtained by a composition of two DCJ operations. While external block interchanges and transpositions can always be mimicked by two consecutive translocations (Figure 6.3 (i)), internal ones can only be mimicked by two DCJs if the first is a circular excision and the second is a circular integration (Figure 6.3 (ii)). We call such a pair of operations an (ei)-*composition*. Note that, without (ei)-compositions, an internal block-interchange or transposition requires three inversions, as illustrated in Figure 6.3 (iii).

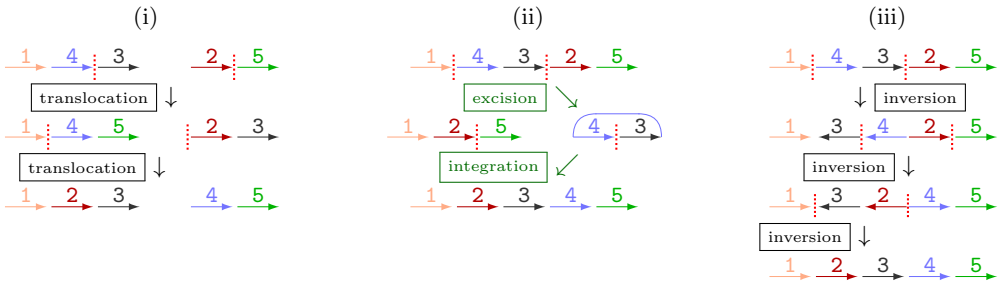

Figure 6.3: (i) External block interchange of markers 4 and 2 mimicked by two translocations. (ii) Internal block interchange of markers 4 and 2 mimicked by an (ei)-composition. (iii) Without a circular excision, the internal block interchange of markers 4 and 2 requires at least three inversions to be mimicked.

6.2.2 Formalizing restricted DCJ models

Let $\mathbb{A}$ and $\mathbb{B}$ be two linear genomes. A scenario sorting $\mathbb{A}$ into $\mathbb{B}$ is said to be *restricted* when each circular excision is immediately followed by a circular integration, forming an (ei)-composition. The *restricted distance* of $\mathbb{A}$ and $\mathbb{B}$, is the minimum cost of a restricted scenario sorting $\mathbb{A}$ into $\mathbb{B}$.

- If linear genomes $\mathbb{A}$ and $\mathbb{B}$ are canonical, we have the *restricted DCJ distance* of $\mathbb{A}$ and $\mathbb{B}$, denoted by $\mathrm{rd}_{\mathrm{DCJ}}(\mathbb{A},\mathbb{B})$. It is clear that $\mathrm{d}_{\mathrm{DCJ}}(\mathbb{A},\mathbb{B}) \leq \mathrm{rd}_{\mathrm{DCJ}}(\mathbb{A},\mathbb{B})$.
- If $\mathbb{A}$ and $\mathbb{B}$ are singular, we have the *restricted DCJ-indel distance* of $\mathbb{A}$ and $\mathbb{B}$, denoted by $\mathrm{rd}^{\mathrm{ID}}_{\mathrm{DCJ}}(\mathbb{A},\mathbb{B})$, and the *restricted DCJ-substitutuion distance* of $\mathbb{A}$ and $\mathbb{B}$, denoted by $\mathrm{rd}^{\mathrm{SB}}_{\mathrm{DCJ}}(\mathbb{A},\mathbb{B})$. Again, $\mathrm{d}^{\mathrm{ID}}_{\mathrm{DCJ}}(\mathbb{A},\mathbb{B}) \leq \mathrm{rd}^{\mathrm{ID}}_{\mathrm{DCJ}}(\mathbb{A},\mathbb{B})$ and $\mathrm{d}^{\mathrm{SB}}_{\mathrm{DCJ}}(\mathbb{A},\mathbb{B}) \leq \mathrm{rd}^{\mathrm{SB}}_{\mathrm{DCJ}}(\mathbb{A},\mathbb{B})$.

6.2.3 Restricted and unrestricted DCJ distances are the same

Let $\mathbb{A}$ and $\mathbb{B}$ be two canonical linear genomes. Consider a scenario sorting $\mathbb{A}$ into $\mathbb{B}$ in which an intermediate step is the excision of circular chromosome $K_\circ$. Observe that $K_\circ$ must have at least one adjacency $\alpha_1 = \gamma_1\gamma_2$ that is not part of any chromosome of $\mathbb{B}$. Moreover, $\mathbb{B}$ must have a cutpoint $\psi_2 = \gamma_1\gamma_3$ such that γ_3 is not part of any adjacency in $K_\circ$. Note that marker extremities $\gamma_1, \gamma_2, \gamma_3$ are distinct. Furthermore, if γ_3 is empty, then ψ_2 is a telomere. By creating the cutpoint ψ_2, circular chromosome $K_\circ$ is integrated into a linear chromosome by a gaining DCJ. This guarantees that $\mathrm{rd}_{\mathrm{DCJ}}(\mathbb{A}, \mathbb{B}) = \mathrm{d}_{\mathrm{DCJ}}(\mathbb{A}, \mathbb{B})$ [99].

Finding an optimal restricted DCJ-indel sorting scenario

Below we will describe the most efficient available algorithm for the restricted canonical DCJ sorting, which mimics the solution as the sorting of a permutation represented as a balanced tree structure and runs in time $O(n_* \log n_*)$ [60].

Data structure for handling permutations. Our sorting algorithm uses a data structure for handling permutations by Kaplan and Verbin [57]. It can be traced back to Chrobak *et al.* [34], where it was used to improve heuristics for the traveling salesman problem. It supports the following three operations in logarithmic time: find the i^{th} marker in a linear chromosome, return the position of marker `X`, and perform a reversal operation. Linear chromosomes can be represented by a balanced tree supporting operations split and merge (e.g., red-black tree or splay tree). The order is the same as the left-to-right order of markers on the chromosome. In each node of the tree, we store one marker, its orientation, number of descendants, and a reverse flag. A reverse flag being "on" signifies that the whole subtree is reversed. The reverse flag of node v can be cleared ("pushed down") by changing v's orientation, swapping its children and flipping their reverse flags. Reversing a segment from i to j can be implemented as follows:

1. Find the i^{th} and j^{th} markers (using the information about sizes of subtrees and reverse flags).
2. Split the tree into three parts: T_1 with markers before i, T_3 with markers after j, and T_2 with the segment from i to j.
3. Flip the reverse flag in the root of T_2, and
4. Merge T_1, T_2 and T_3.

We store a lookup table with a pointer to the corresponding node of a tree for every marker. In this way, we can find the position of any marker in logarithmic time. To support multi-linear genomes, we simply concatenate the chromosomes with a delimiter between each pair, and in each node we store the number of delimiters in its subtree. This way, given a marker `X`, we can tell on which chromosome it is by counting the number of delimiters before `X`. To support different rearrangement operations, we can express them as a sequence of reversals. For example, block interchange can be mimicked by four reversals; if we add sufficiently many delimiters at the end of the sequence (representing empty chromosomes), we can also mimic fusions and fissions.

Algorithm description. As already mentioned and described in Chapter 3, Bergeron *et al.* [10] gave a linear-time algorithm for DCJ sorting disregarding the constraint of reincorporating circular chromosomes immediately. The solution can be easily adapted to a quadratic-time algorithm for the restricted version: after each step, check whether a circular chromosome was created and if so, find the appropriate DCJ operation acting on adjacencies in the circular and the original linear chromosome that reintegrates the circular chromosome. It is not obvious how to do this efficiently (say in polylogarithmic time).

Yancopoulos *et al.* [99] had proposed to transform $\mathbb{A}$ into $\mathbb{B}$ by restricted sorting in four stages: (0) Add caps to the ends of linear chromosomes. (1) By translocations, fusions and fissions transform $\mathbb{A}$ into $\mathbb{A}'$ such that chromosomes in $\mathbb{A}'$ and $\mathbb{B}$ have the same marker contents. (2) Perform oriented reversals to get $\mathbb{A}''$ with all markers in the same direction as in $\mathbb{B}$. (3) Finally, use block interchanges to transform $\mathbb{A}''$ into $\mathbb{B}$. Stages 2 and 3 can be implemented in $O(n_* \log n_*)$ time using the data structure described above (Swenson *et al.* [92]; Feng and Zhu [47]). Thus, a unichromosomal restricted DCJ sorting can be solved in $O(n_* \log n_*)$ time. However, it is not obvious how to implement stage 1 efficiently. Our algorithm is based on the following observation:

Observation 1 *Let* $\mathtt{X}$, $\mathtt{Y}$ *be two markers that are adjacent in* $\mathbb{B}$*, but not in* $\mathbb{A}$*. If* $\mathtt{X}$ *and* $\mathtt{Y}$ *are on different chromosomes in* $\mathbb{A}$*, there is a translocation that puts them together. If* $\mathtt{X}$ *and* $\mathtt{Y}$ *are on the same chromosome and have a different orientation, there is a reversal that puts them together. These operations are optimal in the DCJ model. Transposition and block interchange take two DCJ operations. These operations are optimal if they create two new common adjacencies and destroy none.*

This is simply because, even more generally, k operations, that create k new adjacencies and destroy none, create k new cycles in the relational graph, and thus decrease the distance by k.

Theorem 12 *A restricted optimal DCJ scenario transforming multilinear genome* $\mathbb{A}$ *into multilinear genome* $\mathbb{B}$ *can be found in* $O(n_* \log n_*)$ *time.*

Proof: The ends of linear chromosomes, telomeres, produce some difficulties and nasty special cases. Here, again, we can bypass these special cases with the capping technique, but this time transforming $\mathbb{A}$ and $\mathbb{B}$ into multilinear co-tailed and not circular genomes: we adjoin new markers (caps) to the ends of the lienar chromosomes so that we do not change the distance and we do not have to worry about telomeres any more. We find all the paths in the relational graph $\mathsf{G}_\mathrm{R}(\mathbb{A}, \mathbb{B})$. Paths of odd length have one end in $\mathbb{A}$ and one in $\mathbb{B}$ – simply adjoin a new marker (properly oriented) to the two telomeres. This increases the number of markers by one, but instead of an odd path, we have a cycle and a 1-path, so the distance does not change. For paths starting and ending in $\mathbb{A}$, add two new markers to the ends of $\mathbb{A}$ and add a new chromosome consisting of just these two markers (properly oriented) to $\mathbb{B}$. The case with a path starting and ending in $\mathbb{B}$ is symmetric. The number of markers increases by 2, but instead of an even path, we have a cycle and two odd paths, so the distance does not change. Capping of all chromosomes can be done in linear time.

Let $\mathbb{A}'$ and $\mathbb{B}'$ be the capped co-tailed multilinear genomes and let p be the number of caps added to them. Note that the number of chromosomes in $\mathbb{A}'$ and in $\mathbb{B}'$ is $m = \frac{p}{2}$. Without loss of generality, we may assume that the $p + n_*$ markers in the m capped linear chromosomes

of $\mathbb{A}'$ and $\mathbb{B}'$ are the numbers from 1 to $p+n_*$; and that the target genome $\mathbb{B}'$ is the identity permutation split into the successive m chromosomes:

$$\mathbb{B}' = \{\, [1\ 2 \ldots \mathtt{K}_1{-}1\ \mathtt{K}_1]\ [\mathtt{K}_1{+}1\ \mathtt{K}_1{+}2 \ldots \mathtt{K}_2{-}1\ \mathtt{K}_2] \ldots [\mathtt{K}_{m-1}{+}1\ \mathtt{K}_{m-1}{+}2 \ldots \mathtt{K}_m{-}1\ \mathtt{K}_m] \,\}.$$

The caps are increasing from left to right: $1 < \mathtt{K}_1 < \mathtt{K}_2 < \mathtt{K}_3 < \ldots < \mathtt{K}_{m-1} < \mathtt{K}_m = p + n_*$. The representation of the initial genome is

$$\mathbb{A}' = \{\, [1 \ldots \mathtt{K}_1]\ [\mathtt{K}_1{+}1 \ldots \mathtt{K}_2] \ldots [\mathtt{K}_{m-1}{+}1 \ldots \mathtt{K}_m] \,\}.$$

We will be transforming $\mathbb{A}'$ into $\mathbb{B}'$ gradually "from left to right": once we have transformed the beginning of a chromosome in $\mathbb{A}'$ to $\mathtt{X}\ \mathtt{X}{+}1\ \mathtt{X}{+}2 \ldots \mathtt{Y}$ we extend it by moving $\mathtt{Y}{+}1$ next to $\mathtt{Y}$. We adopt the notation $\ddot{\mathtt{Z}}$ to represent a marker $\mathtt{Z}$ with unkown orientation.

There are several cases we need to consider:

1. If $\mathtt{Y}{+}1$ is already next to $\mathtt{Y}$ we are done!
2. If $\mathtt{Y}{+}1$ is on a different chromosome than $\mathtt{Y}$, we can always use a translocation. In the rest of the proof, we assume that $\mathtt{Y}{+}1$ is on the same chromosome, to the right of $\mathtt{Y}$.
3. If $\mathtt{Y}$ and $\mathtt{Y}{+}1$ have different orientation, we can use a reversal.

   Otherwise, $\mathtt{Y}$ and $\mathtt{Y}{+}1$ have the same orientation. Following Christie [33], find marker $\mathtt{Z}$ with the highest number between $\mathtt{Y}$ and $\mathtt{Y}{+}1$ and find marker $\mathtt{Z}{+}1$.
4. If $\mathtt{Z}{+}1$ is on a different chromosome, we can use a translocation to move it next to $\mathtt{Z}$; this operation also moves $\mathtt{Y}{+}1$ to another chromosome, so we can use another translocation to move it next to $\mathtt{Y}$.

   Otherwise the situation is $\mathtt{Y} \ldots \ddot{\mathtt{Z}} \ldots \mathtt{Y}{+}1 \ldots \ddot{\mathtt{Z}{+}1}$ (since $\mathtt{Z}$ is the highest number between $\mathtt{Y}$ and $\mathtt{Y}{+}1$ and the part of the chromosome to the left of $\mathtt{Y}$ is already sorted, $\mathtt{Z}{+}1$ must be to the right of $\mathtt{Y}{+}1$).
5. If $\mathtt{Z}$ and $\mathtt{Z}{+}1$ have different orientations, we can use a reversal to move $\mathtt{Z}{+}1$ next to $\mathtt{Z}$; this will also change the orientation of $\mathtt{Y}{+}1$, so in the next step, we can use another reversal to move $\mathtt{Y}{+}1$ next to $\mathtt{Y}$.
6. Finally, if $\mathtt{Z}$ and $\mathtt{Z}{+}1$ have the same orientation, we interchange blocks

   (i) $\mathtt{Y}\,\langle \ldots \mathtt{Z}\rangle \ldots \langle \mathtt{Y}{+}1 \ldots \rangle\, \mathtt{Z}{+}1 \;\rightsquigarrow\; \mathtt{Y}\,\langle \mathtt{Y}{+}1 \ldots \rangle \ldots \langle \ldots \mathtt{Z}\rangle\, \mathtt{Z}{+}1$

   if both $\mathtt{Z}$ and $\mathtt{Z}{+}1$ have direct orientation; or

   (ii) $\mathtt{Y}\,\langle \ldots \rangle\, \overline{\mathtt{Z}} \ldots \langle \mathtt{Y}{+}1 \ldots \overline{\mathtt{Z}{+}1}\rangle \;\rightsquigarrow\; \mathtt{Y}\,\langle \mathtt{Y}{+}1 \ldots \overline{\mathtt{Z}{+}1}\rangle\, \overline{\mathtt{Z}} \ldots \langle \ldots \rangle$

   if both $\mathtt{Z}$ and $\mathtt{Z}{+}1$ have reverse orientation.

   In both cases (i) and (ii), with two DCJs we move $\mathtt{Y}{+}1$ next to $\mathtt{Y}$ and $\mathtt{Z}{+}1$ next to $\mathtt{Z}$.

Note that $p \leq 2n_*$. Every step can be implemented in $O(\log n_*)$ time using an extended version of the data structure described above. We need the data structure to support the following operations: (1) Given a marker, find the chromosome that contains it. (2) Perform a DCJ operation. (3) Given interval $\mathtt{Y} \ldots \mathtt{Z}$, find the marker with the highest number on the chromosome between $\mathtt{Y}$ and $\mathtt{Z}$. To support this query, we store the highest number in the subtree in each node. □

Bibliography

[1] Mark D. Adams, Susan E. Celniker, Robert A. Holt, Cheryl A. Evans, Jeannine D. Gocayne, Peter G. Amanatides, Steven E. Scherer, Peter W. Li, Roger A. Hoskins, Richard F. Galle, Reed A. George, Suzanna E. Lewis, Stephen Richards, Michael Ashburner, Scott N. Henderson, Granger G. Sutton, Jennifer R. Wortman, Mark D. Yandell, Qing Zhang, Lin X. Chen, Rhonda C. Brandon, Yu-Hui C. Rogers, Robert G. Blazej, Mark Champe, Barret D. Pfeiffer, Kenneth H. Wan, Clare Doyle, Evan G. Baxter, Gregg Helt, Catherine R. Nelson, George L. Gabor, Miklos, Josep F. Abril, Anna Agbayani, Hui-Jin An, Cynthia Andrews-Pfannkoch, Danita Baldwin, Richard M. Ballew, Anand Basu, James Baxendale, Leyla Bayraktaroglu, Ellen M. Beasley, Karen Y. Beeson, P. V. Benos, Benjamin P. Berman, Deepali Bhandari, Slava Bolshakov, Dana Borkova, Michael R. Botchan, John Bouck, Peter Brokstein, Phillipe Brottier, Kenneth C. Burtis, Dana A. Busam, Heather Butler, Edouard Cadieu, Angela Center, Ishwar Chandra, J. Michael Cherry, Simon Cawley, Carl Dahlke, Lionel B. Davenport, Peter Davies, Beatriz de Pablos, Arthur Delcher, Zuoming Deng, Anne Deslattes Mays, Ian Dew, Suzanne M. Dietz, Kristina Dodson, Lisa E. Doup, Michael Downes, Shannon Dugan-Rocha, Boris C. Dunkov, Patrick Dunn, Kenneth J. Durbin, Carlos C. Evangelista, Concepcion Ferraz, Steven Ferriera, Wolfgang Fleischmann, Carl Fosler, Andrei E. Gabrielian, Neha S. Garg, William M. Gelbart, Ken Glasser, Anna Glodek, Fangcheng Gong, J. Harley Gorrell, Zhiping Gu, Ping Guan, Michael Harris, Nomi L. Harris, Damon Harvey, Thomas J. Heiman, Judith R. Hernandez, Jarrett Houck, Damon Hostin, Kathryn A. Houston, Timothy J. Howland, Ming-Hui Wei, Chinyere Ibegwam, Mena Jalali, Francis Kalush, Gary H. Karpen, Zhaoxi Ke, James A. Kennison, Karen A. Ketchum, Bruce E. Kimmel, Chinnappa D. Kodira, Cheryl Kraft, Saul Kravitz, David Kulp, Zhongwu Lai, Paul Lasko, Yiding Lei, Alexander A. Levitsky, Jiayin Li, Zhenya Li, Yong Liang, Xiaoying Lin, Xiangjun Liu, Bettina Mattei, Tina C. McIntosh, Michael P. McLeod, Duncan McPherson, Gennady Merkulov, Natalia V. Milshina, Clark Mobarry, Joe Morris, Ali Moshrefi, Stephen M. Mount, Mee Moy, Brian Murphy, Lee Murphy, Donna M. Muzny, David L. Nelson, David R. Nelson, Keith A. Nelson, Katherine Nixon, Deborah R. Nusskern, Joanne M. Pacleb, Michael Palazzolo, Gjange S. Pittman, Sue Pan, John Pollard, Vinita Puri, Martin G. Reese, Knut Reinert, Karin Remington, Robert D. C. Saunders, Frederick Scheeler, Hua Shen, Bixiang Christopher Shue, Inga Sidén-Kiamos, Michael Simpson,

Marian P. Skupski, Tom Smith, Eugene Spier, Allan C. Spradling, Mark Stapleton, Renee Strong, Eric Sun, Robert Svirskas, Cyndee Tector, Russell Turner, Eli Venter, Aihui H. Wang, Xin Wang, Zhen-Yuan Wang, David A. Wassarman, George M. Weinstock, Jean Weissenbach, Sherita M. Williams, Trevor Woodage, Kim C. Worley, David Wu, Song Yang, Q. Alison Yao, Jane Ye, Ru-Fang Yeh, Jayshree S. Zaveri, Ming Zhan, Guangren Zhang, Qi Zhao, Liansheng Zheng, Xiangqun H. Zheng, Fei N. Zhong, Wenyan Zhong, Xiaojun Zhou, Shiaoping Zhu, Xiaohong Zhu, Hamilton O. Smith, Richard A. Gibbs, Eugene W. Myers, Gerald M. Rubin, and J. Craig Venter. The genome sequence of *Drosophila melanogaster*. *Science*, 287:2185–2195, 2000.

[2] Max Alekseyev and Pavel A. Pevzner. Colored de Bruijn graphs and the genome halving problem. *IEEE/ACM Trans. Comput. Biol. Bioinform.*, 4(1):98–107, 2008.

[3] Adrian M. Altenhoff and Christophe Dessimoz. Phylogenetic and functional assessment of orthologs inference projects and methods. *PLOS Computational Biology*, 5(1):1–11, 01 2009.

[4] Adrian M. Altenhoff, Jeremy Levy, Magdalena Zarowiecki, Bartłomiej Tomiczek, Alex Warwick Vesztrocy, Daniel A. Dalquen, Steven Müller, Maximilian J. Telford, Natasha M. Glover, David Dylus, and Christophe Dessimoz. OMA standalone: orthology inference among public and custom genomes and transcriptomes. *Genome Research*, 29:1152–1163, 2019.

[5] Stephen F. Altschul, Warren Gish, Webb Miller, Eugene W. Myers, and David J. Lipman. Basic local alignment search tool. *J Mol Biol*, 215(3):403–410, 1990.

[6] Sébastien Angibaud, Guillaume Fertin, Irena Rusu, Annelyse Thévenin, and Stéphane Vialette. On the approximability of comparing genomes with duplicates. *J Graph Algo App*, 13(1):19–53, 2009.

[7] Vineet Bafna and Pavel A. Pevzner. Genome rearrangements and sorting by reversals. In *Proceedings of FOCS 1993*, pages 148–157, 1993.

[8] Anne Bergeron. A very elementary presentation of the hannenhalli-pevzner theory. In *Proc. of CPM*, volume 2089 of *LNCS*, pages 106–117, 2001.

[9] Anne Bergeron, Steffen Heber, and Jens Stoye. Common intervals and sorting by reversals: a marriage of necessity. *Bioinformatics*, 18(Suppl. 2):S54–G63, 2002.

[10] Anne Bergeron, Julia Mixtacki, and Jens Stoye. A unifying view of genome rearrangements. In *Proc. of WABI*, volume 4175 of *Lecture Notes in Bioinformatics*, pages 163–173, 2006.

[11] Anne Bergeron, Julia Mixtacki, and Jens Stoye. A new linear time algorithm to compute the genomic distance via the double cut and join distance. *Theoretical Computer Science*, 410(51):5300–5316, 2009.

[12] Matthias Bernt, Daniel Merkle, and Martin Middendorf. Genome rearrangement based on reversals that preserve conserved intervals. *IEEE/ACM Trans Comput Biol Bioinform*, 3(3):275–288, 2006.

[13] Priscila Biller, Laurent Guéguen, Carole Knibbe, and Eric Tannier. Breaking Good: Accounting for Fragility of Genomic Regions in Rearrangement Distance Estimation. *Genome Biology and Evolution*, 5(8):1427—-1439, 2016.

[14] Leonard Bohnenkämper. The Floor Is Lava: Halving Natural Genomes with Viaducts, Piers, and Pontoons. *J Comput Biol*, 31(4):294–311, 2024. A preliminary version appeared in the Proc. of Recomb-CG 2023.

[15] Leonard Bohnenkämper, Marília D. V. Braga, Daniel Doerr, and Jens Stoye. Computing the rearrangement distance of natural genomes. *J Comput Biol*, 28(4):410–431, 2021. A preliminary version appeared in Proc. of RECOMB 2020, LNCS 12074, 3–18.

[16] Leonard Bohnenkämper. Recombinations, chains and caps: resolving problems with the DCJ-indel model. *Algorithms Mol Biol*, 19(8), 2024. A preliminary version appeared in the Proc. of WABI 2023.

[17] Jeffrey L. Boore. The duplication/random loss model for gene rearrangement exemplified by mitochondrial genomes of deuterostome animals. In David Sankoff and Joseph H. Nadeau, editors, *Comparative Genomics*, pages 133–148. Springer, 2000.

[18] Marília D. V. Braga. An overview of genomic distances modeled with indels. In *Proc. of CiE*, volume 7921 of *LNCS*, pages 22–31. Springer, 2013.

[19] Marília D. V. Braga, Leonie R. Brockmann, Katharina Klerx, and Jens Stoye. Investigating the complexity of the double distance problems. *Algorithms Mol Biol*, 19(1), 2024. Preliminary versions appeared in the Proc. of WABI 2022 and RECOMB-CG 2023.

[20] Marília D. V. Braga, Cedric Chauve, Daniel Doerr, Katharina Jahn, Jens Stoye, Annelyse Thévenin, and Roland Wittler. The potential of family-free genome comparison. In C. Chauve, N. El-Mabrouk, and E. Tannier, editors, *Models and Algorithms for Genome Evolution*, volume 19 of *Computational Biology Series*, chapter 13, pages 287–307. Springer, Berlin, 2013.

[21] Marília D. V. Braga and Jens Stoye. The solution space of sorting by DCJ. *J Comput Biol*, 17(9):1145–1165, 2010. A preliminary version appeared in the Proc. of RECOMB-CG 2009.

[22] Marília D. V. Braga, Eyla Willing, and Jens Stoye. Double cut and join with insertions and deletions. *J Comput Biol*, 18(9):1167–1184, 2011. A preliminary version appeared in the Proc. of WABI 2010.

[23] Marília D. V. Braga, Daniel Doerr, Diego P. Rubert, and Jens Stoye. Family-free genome comparison. In João Carlos Setubal, Peter F. Stadler, and Jens Stoye, editors, *Comparative Genomics: Methods and Protocols*, volume 2802 of *Methods in Molecular Biology*, pages 57–72. Springer, New York, 2024. Second edition; for the first edition see [42, from 2018].

[24] Marília D. V. Braga, Raphael Machado, Leonardo C. Ribeiro, and Jens Stoye. Genomic distance under gene substitutions. *BMC Bioinform*, 12(Suppl 9):S8, 2011.

[25] Marília D. V. Braga, Raphael Machado, Leonardo C. Ribeiro, and Jens Stoye. On the weight of indels in genomic distances. *BMC Bioinformatics*, 12(Suppl. 9):S13, 2011.

[26] Marília D. V. Braga and Jens Stoye. Sorting linear genomes with rearrangements and indels. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 12(3):500–506, 2015.

[27] David Bryant. The complexity of calculating exemplar distances. In David Sankoff and Joseph H. Nadeau, editors, *Comparative Genomics*, volume 1 of *Computational Biology Series*, pages 207–211. Kluver Academic Publishers, London, 2000.

[28] Benjamin Buchfink, Chao Xie, and Daniel H. Huson. Fast and sensitive protein alignment using DIAMOND. *Nat Methods*, 12:59–60, 2015.

[29] Sèverine Bérard, Annie Chateau, Cedric Chauve, Christophe Paul, and Eric Tannier. Perfect DCJ rearrangement. In *Proceedings of RECOMB-CG*, volume 5267 of *LNBI*, page 158–169, 2008.

[30] Cedric Chauve. Personal communication in Dagstuhl Seminar no. 18451 - Genomics, Pattern Avoidance, and Statistical Mechanics, November 2018.

[31] Xin Chen. On sorting unsigned permutations by double-cut-and-joins. *J Comb Optim*, 25(1):339–351, 2013.

[32] Xin Chen, Jie Zheng, Zheng Fu, Peng Nan, Yang Zhong, S. Lonardi, and Tao Jiang. Assignment of orthologous genes via genome rearrangement. *IEEE/ACM Trans Comput Biol Bioinform*, 2(4):302–315, 2005.

[33] David A. Christie. Sorting permutations by block-interchanges. *Inf Process Lett*, 60(4):165–169, 1996.

[34] M. Chrobak, T. Szymacha, and A. Krawczyk. A data structure useful for finding hamiltonian cycles. *Theor Comput Sci*, 71(3):419–424, 1990.

[35] Andrew G. Clark, Michael B. Eisen, Douglas R. Smith, Casey M. Bergman, Brian Oliver, Therese A. Markow, Thomas C. Kaufman, Manolis Kellis, and Drosophila12GenomesConsortium. Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature*, 450:203–218, 2007.

[36] Philip E. C. Compeau. DCJ-indel sorting revisited. *Algorithms Mol Biol*, 8(1), 2013. A preliminary version appeared in the Proc. of WABI 2012.

[37] Poly H. da Silva, Raphael Machado, Simone Dantas, and Marília D. V. Braga. Restricted DCJ-indel model: sorting linear genomes with DCJ and indels. *BMC Bioinformatics*, 13(Suppl. 19):S14.

[38] Poly H. da Silva, Raphael Machado, Simone Dantas, and Marília D. V. Braga. DCJ-indel and DCJ-substitution distances with distinct operation costs. *Algorithms Mol Biol*, 8(21), 2013. A preliminary version appeared in Proc. of WABI 2012.

[39] Poly H. da Silva, Raphael Machado, Simone Dantas, and Marília D.V. Braga. Genomic distance with high indel costs. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 14(3):728–732, 2017.

[40] C. Dessimoz, G. Cannarozzi, M. Gil, D. Margadant, A. C. J. Roth, A. Schneider, and G. H. Gonnet. OMA, a comprehensive, automated project for the identification of orthologs from complete genome data: introduction and first achievements. In *Proc. of RECOMB-CG*, volume 3678 of *Lecture Notes in Bioinformatics*, pages 61–72, 2005.

[41] Daniel Doerr and Cedric Chauve. Small parsimony for natural genomes in the DCJ-indel model. *J Bioinform Comput Biol*, 19(6):2140009, 2021.

[42] Daniel Doerr, Pedro Feijão, and Jens Stoye. Family-free genome comparison. In João C. Setubal, Jens Stoye, and Peter F. Stadler, editors, *Comparative Genomics: Methods and Protocols*, volume 1704 of *Methods in Molecular Biology*, pages 331–342. Springer Nature, New York, 2018. First edition; for the second edition see [23, from 2024].

[43] Daniel Doerr, Annelyse Thévenin, and Jens Stoye. Gene family assignment-free comparative genomics. *BMC Bioinform*, 13(Suppl 19):S3, 2012.

[44] Nadia El-Mabrouk. Sorting signed permutations by reversals and insertions/deletions of contiguous segments. *Journal of Discrete Algorithms*, 1(1):105–122, 2001.

[45] Nadia El-Mabrouk and David Sankoff. The reconstruction of doubled genomes. *SIAM Journal on Computing*, 32(3):754–792, 2003.

[46] Péter L. Erdős, Lajos Soukup, and Jens Stoye. Balanced vertices in trees and a simpler algorithm to compute the genomic distance. *Applied Mathematics Letters*, 24(1):82–86, 2011.

[47] Jianxing Feng and Daming Zhu. Faster algorithms for sorting by transpositions and sorting by block interchanges. *ACM Trans Algorithms*, 3(3), 2007.

[48] Paul Feyerabend. *Against Method*. New Left Books, 1975.

[49] Paul Feyerabend. *Farewell to Reason*. New Left Books, 1987.

[50] Walter M. Fitch. Distinguishing homologous from analogous proteins. *Syst Zool*, 19:99–113, 1970.

[51] Shmuel Friedland. An upper bound for the number of perfect matchings in graphs, 2008.

[52] Marija Gimbutas. *The Goddesses and Gods of Old Europe - Myths and Cult Images 6500 - 3500 BC*. University of California Press, 1982.

[53] David Graeber and David Wengrow. *The Dawn of Everything, a new history of humanity*. Allen Lane Imprint, 2021.

[54] P. Hall. On representatives of subsets. *Journal of the London Mathematical Society*, s1-10(1):26–30, 1935.

[55] Sridhar Hannenhalli and Pavel A. Pevzner. Transforming men into mice (polynomial algorithm for genomic distance problem). In *Proc. of FOCS*, pages 581–592, 1995.

[56] Sridhar Hannenhalli and Pavel A. Pevzner. Transforming cabbage into turnip: polynomial algorithm for sorting signed permutations by reversals. *Journal of the ACM*, 46(1):1–27, 1999. A preliminary version appeared in the Proc. of ACM Symposium on the Theory of Computing 1995.

[57] Haim Kaplan and Elad Verbin. Sorting signed permutations by reversals, revisited. *J Comput Syst Sci*, 70(3):321–341, 2005.

[58] Eugene V. Koonin. Orthologs, paralogs, and evolutionary genomics. *Annual Review of Genetics*, 39(1):309–338, 2005.

[59] Davi Kopenawa and Bruce Albert. *The Falling Sky: Words of a Yanomami Shaman*. Harvard University Press, 2013. Originally published in French in 2010.

[60] Jakub Kováč, Robert Warren, Marília D. V. Braga, and Jens Stoye. Restricted dcj model: Rearrangement problems with chromosome reincorporation. *J. Comput. Biol.*, 18:1231–1241, 2011. A preliminary version appeared in Proc. of RECOMB-CG 2010.

[61] Ailton Krenak. *Ideas to Postpone the End of the World.* House Of Anansi Press Ltd, 2019.

[62] Sudhir Kumar, Glen Stecher, Michael Li, Christina Knyaz, and Koichiro Tamura. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol*, 35(6):1547–1549, 2018.

[63] Sudhir Kumar, Glen Stecher, Michael Suleski, and S. Blair Hedges. Timetree: a resource for timelines, timetrees, and divergence times. *Mol Biol Evol*, 34(7):1812–1819, 2017.

[64] Marcus Lechner, Sven Findeiß, Lydia Steiner, Manja Marz, Peter F. Stadler, and Sonja J. Prohaska. Proteinortho: Detection of (co-)orthologs in large-scale analysis. *BMC Bioinform*, 12(124), 2011.

[65] Marcus Lechner, Maribel Hernandez-Rosales, Daniel Doerr, Nicolas Wieseke, Annelyse Thévenin, Jens Stoye, Roland K. Hartmann, Sonja J. Prohaska, and Peter F. Stadler. Orthology detection combining clustering and synteny for very large datasets. *PLoS One*, 9(8:e105015), 2014.

[66] Fábio V. Martinez, Pedro Feijão, Marília D. V. Braga, and Jens Stoye. On the family-free DCJ distance and similarity. *Algorithms Mol Biol*, 13(10), 2015. A preliminary version appeared in the Proc. of WABI 2014.

[67] Humberto R. Maturana and Francisco J. Varela. *The Tree of Knowledge: The Biological Roots of Human Understanding.* Shambhala Publications, 1998.

[68] Compiled by T. C. McLuhan. *Touch the Earth: a self-portrait of Indian existence.* Abacus, 1989.

[69] J. Meidanis, M. E. M. T. Walter, and Z. Dias. Reversal distance of signed circular chromosomes. Relatório Técnico IC-00-23, Institute of Computing, University of Campinas, Brazil, 2000.

[70] Carolyn Merchant. *The Death of Nature: Women, Ecology and the Scientific Revolution.* New York: Harper and Row, 1980.

[71] Julia Mixtacki. Genome halving under DCJ revisited. In *Proceedings of COCOON 2008*, volume 5092 of *LNCS*, pages 276–286, 2008.

[72] C. Moritz, T. E. Dowling, and W. M. Brown. Evolution of animal mitochondrial dna: relevance for population biology and systematics. In *Annu. Rev. Ecol. Syst*, volume 18, pages 269–292, 1987.

[73] Mihai Nadin. *Disrupt Science: The Future Matters.* Springer, 2023.

[74] Aïda Ouangraoua and Anne Bergeron. Combinatorial structure of genome rearrangements scenarios. *J Comput Biol*, 17(9):1129–1144, 2010. A preliminary version appeared in the Proc. of RECOMB-CG 2009.

[75] Catia Pesquita, Daniel Faria, Hugo Bastos, António EN Ferreira, André O Falcão, and Francisco M Couto. Metrics for GO based protein semantic similarity: a systematic evaluation. *BMC Bioinformatics*, 9(Suppl 5):S4, 2008.

[76] Stephen Richards, Yue Liu, Brian R. Bettencourt, Pavel Hradecky, Stan Letovsky, Rasmus Nielsen, Kevin Thornton, Melissa J. Hubisz, Rui Chen, Richard P. Meisel, Olivier Couronne, Sujun Hua, Mark A. Smith, Peili Zhang, Jing Liu, Harmen J. Bussemaker, Marinus F. van Batenburg, Sally L. Howells, Steven E. Scherer, Erica Sodergren, Beverly B. Matthews, Madeline A. Crosby, Andrew J. Schroeder, Daniel Ortiz-Barrientos, Catharine M. Rives, Michael L. Metzker, Donna M. Muzny, Graham Scott, David Steffen, David A. Wheeler, Kim C. Worley, Paul Havlak, K. James Durbin, Amy Egan, Rachel Gill, Jennifer Hume, Margaret B. Morgan, George Miner, Cerissa Hamilton, Yanmei Huang, Lenée Waldron, Daniel Verduzco, Kerstin P. Clerc-Blankenburg, Inna Dubchak, Mohamed A.F. Noor, Wyatt Anderson, Kevin P. White, Andrew G. Clark, Stephen W. Schaeffer, William Gelbart, George M. Weinstock, and Richard A. Gibbs. Comparative genome sequencing of *Drosophila pseudoobscura*: Chromosomal, gene, and cis-element evolution. *Genome Res*, 15:1–18, 2005.

[77] Robert Rosen. *Life Itself: A Comprehensive Inquiry into the Nature, Origin, and Fabrication of Life*. Columbia University Press,, 1991.

[78] Robert Rosen. *Essays on Life Itself*. Columbia University Press, 2000.

[79] Alexander C. J. Roth, Gaston H. Gonnet, and Christophe Dessimoz. Algorithm of OMA for large-scale orthology inference. *BMC Bioinform*, 9(518), 2008.

[80] Diego P. Rubert and Marília D. V. Braga. Efficient gene orthology inference via large-scale rearrangements. *Algorithms Mol Biol*, 18(14), 2023. A preliminary version appeared in the Proc. of WABI 2022.

[81] Diego P. Rubert, Daniel Doerr, and Marília D. V. Braga. The potential of family-free rearrangements towards gene orthology inference. *J Bioinform Comput Biol*, 19(6):2140014, 2021.

[82] Diego P. Rubert, Fábio V. Martinez, and Marília D. V. Braga. Natural Family-Free Genomic Distance. *Algorithms Mol Biol*, 16(4), 2021. A preliminary version appeared in the Proc. of WABI 2020.

[83] Marie-France Sagot and Eric Tannier. Perfect sorting by reversals. In *Proc. of COCOON*, volume 3595 of *LNCS*, pages 42–51, 2005.

[84] Naruya Saitou and Masatoshi Nei. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol*, 4(4):406–425, 1987.

[85] David Sankoff. Edit distance for genome comparison based on non-local operations. In *Proc. of CPM*, volume 644 of *Lecture Notes in Computer Science*, pages 121–135, 1992.

[86] David Sankoff. Genome rearrangement with gene families. *Bioinformatics*, 15(11):909–917, 1999.

[87] João C. Setubal and Peter F. Stadler. Gene phylogenies and orthologous groups. In João C. Setubal, Jens Stoye, and Peter F. Stadler, editors, *Comparative Genomics: Methods and Protocols*, pages 1–28. Springer, New York, 2018.

[88] Mingfu Shao, Yu Lin, and Bernard Moret. An exact algorithm to compute the double-cut-and-join distance for genomes with duplicate genes. *J Comput Biol*, 22(5):425–435, 2015.

[89] Guanqun Shi, Meng-Chih Peng, and Tao Jiang. MultiMSOAR 2.0: an accurate tool to identify ortholog groups among multiple genomes. *PLoS One*, 6(6:e20892), 2011.

[90] Guanqun Shi, Liqing Zhang, and Tao Jiang. MSOAR 2.0: Incorporating tandem duplications into ortholog assignment based on genome rearrangement. *BMC Bioinform*, 11(10), 2010.

[91] Alfred H. Sturtevant. A case of rearrangement of genes in drosophila. In *Proc. of Natl Acad Sci USA*, volume 7, pages 235–237, 1921.

[92] Krister M. Swenson, Vaibhav Rajan, Yu Lin, and Bernard M.E. Moret. Sorting signed permutations by inversions in $O(n \log n)$ time. *J Comput Biol*, 17(3):489–501, 2010.

[93] Eric Tannier, Chunfang Zheng, and David Sankoff. Multichromosomal median and halving problems under different genomic distances. *BMC Bioinformatics*, 10:120, 2009.

[94] Tamir Tassa. Finding all maximally-matchable edges in a bipartite graph. *Theoretical Computer Science*, 423:50–58, 2012.

[95] Roman L. Tatusov, Eugene V. Koonin, and David J. Lipman. A genomic perspective on protein families. *Science*, 278(5338):631–637, 1997.

[96] Fredj Tekaia. Inferring orthologs: Open questions and perspectives. *Genomics Insights*, 9:GEI.S37925, 2016.

[97] Eyla Willing, Jens Stoye, and Marília D. V. Braga. Computing the inversion-indel distance. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 18(6):2314–2326, 2021.

[98] Eyla Willing, Simone Zaccaria, Marília D. V. Braga, and Jens Stoye. On the Inversion-Indel Distance. *BMC Bioinformatics*, 14(Suppl. 15):S3, 2013.

[99] Sophia Yancopoulos, Oliver Attie, and Richard Friedberg. Efficient sorting of genomic permutations by translocation, inversion and block interchange. *Bioinformatics*, 21(16):3340–3346, 2005.

[100] Sophia Yancopoulos and Richard Friedberg. DCJ path formulation for genome transformations which include insertions, deletions, and duplications. *J Comput Biol*, 16(10):1311–1338, 2009.

[101] Zhaoming Yin, Jijun Tang, Stephen W. Schaeffer, and David A. Bader. Exemplar or matching: modeling DCJ problems with unequal content genome data. *Journal of Combinatorial Optimization*, 32(4):1165–1181, 2016.

[102] Doron Zeilberger. Yet another proof of cayley's formula for the number of labelled trees: `http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimPDF/cayley.pdf`, included in Appendix B.

[103] Qi Zhou and Doris Bachtrog. Ancestral chromatin configuration constrains chromatin evolution on differentiating sex chromosomes in *Drosophila*. *PLoS Genet*, 11(6), 2015.

APPENDIX A

Symbols used in the adopted formalism

$\mathbb{A}, \mathbb{B}, \mathbb{C}$	genomes
$\eta(.)$	family-annotation of marker, marker extremity, adjacency, telomere or chromosome
$\mathsf{adj}(\mathbb{A})$	set of unannotated adjacencies in genome $\mathbb{A}$
$\mathsf{ADJ}(\mathbb{A})$	$= \eta(\mathsf{adj}(\mathbb{A}))$: set of annotated adjacencies in genome $\mathbb{A}$
$\mathsf{ext}(\mathbb{A})$	set of unannotated marker extremities in genome $\mathbb{A}$
$\mathsf{EXT}(\mathbb{A})$	$= \eta(\mathsf{ext}(\mathbb{A}))$: set of annotated marker extremities in genome $\mathbb{A}$
$\mathsf{tel}(\mathbb{A})$	set of unannotated telomeres in genome $\mathbb{A}$
$\mathsf{TEL}(\mathbb{A})$	$= \eta(\mathsf{tel}(\mathbb{A}))$: set of annotated telomeres in genome $\mathbb{A}$
$\psi(\mathbb{A})$	set of cutpoints in genome $\mathbb{A}$
$\mathfrak{C}(\mathbb{A})$	set of unannotated chromosomes in genome $\mathbb{A}$
$\mathcal{F}(\mathbb{A})$	set of families occurring in annotated genome $\mathbb{A}$
$\mathcal{G}(\mathbb{A})$	$= \eta(\mathcal{M}(\mathbb{A}))$: set of annotated markers of genome $\mathbb{A}$
$\mathcal{M}(\mathbb{A})$	set of unannotated markers of genome $\mathbb{A}$
$\Phi(\mathtt{F}, \mathbb{A})$	number of occurrences (markers) of family `F` in genome $\mathbb{A}$

In comparison of two annotated genomes $\mathbb{A}$ and $\mathbb{B}$:

$\mathcal{F}_*$	set of common families of genomes $\mathbb{A}$ and $\mathbb{B}$
$\mathcal{A}$	set of families exclusive to genome $\mathbb{A}$ (see also table below)
$\mathcal{B}$	set of families exclusive to genome $\mathbb{B}$ (see also table below)
$\mathcal{G}_*$	(multi)set of common annotated markers of genomes $\mathbb{A}$ and $\mathbb{B}$

Graphs and respective derived properties:

G_{R}	relational graph of canonical or singular genomes
G_{MR}	multi-relational graph of balanced or natural genomes
$\mathsf{G}^{\mathsf{w}}_{\mathrm{FFR}}$	family-free multi-relational graph of unannotated genomes
$\mathrm{J}(.)$	set of junctions in the whole relational graph or in one of its connected components
$\mathsf{G}_{\mathcal{F}_*}$	family graph of natural genomes
$\mathsf{G}^{\mathsf{w}}_{\sigma}$	similarity graph
$\mathsf{G}^{\mathsf{w}}_{\mathfrak{C}}$	shared-content graph
$\widehat{\mathsf{G}}^{\mathsf{w}}_{\mathfrak{C}}$	perfect shared-content graph
Γ	component (path or cycle) of a relational graph
$\Gamma^{\langle i \rangle}$	component whose length (given by the number of extremity edges) is $i \geq 0$
$\mathbb{O}$	cycle of a relational graph
$\mathbb{O}^{\langle i \rangle}$	cycle of (even) length $i \geq 0$
$\mathbb{AA}$	path of a relational graph starting and ending in a telomere of genome $\mathbb{A}$
$\mathbb{AA}^{\langle i \rangle}$	$\mathbb{AA}$-path of (even) length $i \geq 0$
$\mathbb{AB}$	path of a relational graph starting in a telomere of $\mathbb{A}$ and ending in a telomere of $\mathbb{B}$
$\mathbb{AB}^{\langle i \rangle}$	$\mathbb{AB}$-path of (odd) length $i \geq 1$
$\mathbb{BB}$	path of a relational graph starting and ending in a telomere of genome $\mathbb{B}$
$\mathbb{BB}^{\langle i \rangle}$	$\mathbb{BB}$-path of (even) length $i \geq 0$
$\Upsilon(\Gamma)$	type of component Γ
$\mathcal{S}$	$= \{\Gamma \mid \Gamma \in \mathsf{G}_{\mathrm{R}}(\mathbb{A},\mathbb{B})$ and $\Upsilon(\Gamma) = \mathbb{O}^{\langle 0 \rangle}\}$: set of (circular) singletons
$\mathcal{C}$	$= \{\Gamma \mid \Gamma \in \mathsf{G}_{\mathrm{R}}(\mathbb{A},\mathbb{B})$ and $\Upsilon(\Gamma) = \mathbb{O}^{\langle 2,4,\ldots \rangle}\}$: set of cycles of length at least 2
$\mathcal{P}_{\mathbb{AA}}$	$= \{\Gamma \mid \Gamma \in \mathsf{G}_{\mathrm{R}}(\mathbb{A},\mathbb{B})$ and $\Upsilon(\Gamma) = \mathbb{AA}\}$: set of $\mathbb{AA}$-paths
$\mathcal{P}_{\mathbb{AB}}$	$= \{\Gamma \mid \Gamma \in \mathsf{G}_{\mathrm{R}}(\mathbb{A},\mathbb{B})$ and $\Upsilon(\Gamma) = \mathbb{AB}\}$: set of $\mathbb{AB}$-paths
$\mathcal{P}_{\mathbb{BB}}$	$= \{\Gamma \mid \Gamma \in \mathsf{G}_{\mathrm{R}}(\mathbb{A},\mathbb{B})$ and $\Upsilon(\Gamma) = \mathbb{BB}\}$: set of $\mathbb{BB}$-paths
$\mathcal{A}$	as a path-subscript: odd sequence of runs starting and ending in an $\mathcal{A}$-run
$\mathcal{B}$	as a path-subscript: odd sequence of runs starting and ending in an $\mathcal{B}$-run
$\mathcal{AB}$	path-subscript: even sequence of runs starting in an $\mathcal{A}$-run and ending in a $\mathcal{B}$-run
$\Lambda(\Gamma)$	number of runs in component Γ
$\aleph(\Gamma)$	number of transitions in component Γ
$\lambda(\Gamma)$	indel-potential of component Γ
$\sigma(\Gamma)$	substitution-potential of component Γ
$\mathcal{O}$	ortholog-set
$\widetilde{\mathcal{O}}$	complement of ortholog-set $\mathcal{O}$
$\mathcal{H}$	ortholog-maxset
$\widetilde{\mathcal{H}}$	complement of ortholog-maxset $\mathcal{H}$
$\mathcal{L}$	sibling-set
$\widetilde{\mathcal{L}}$	complement of sibling-set $\mathcal{L}$
$\mathcal{J}$	sibling-maxset
$\widetilde{\mathcal{J}}$	complement of sibling-maxset $\mathcal{J}$
$\mathcal{Q}$	capping-maxset