Supplementary Material:
Genome Rearrangement with ILP

Tom Hartmann\textsuperscript{1}, Nicolas Wieseke\textsuperscript{1}, Roded Sharan\textsuperscript{2} Martin Middendorf\textsuperscript{1}, and Matthias Bernt\textsuperscript{1}

\textsuperscript{1} Parallel Computing and Complex Systems Group, Institute of Computer Science, University Leipzig, Germany
\{thartmann,wieseke,middendorf,bernt\}@informatik.uni-leipzig.de
\textsuperscript{2} Blavatnik School of Computer Science, Tel Aviv University, Tel Aviv, Israel
roded@post.tau.ac.il

1 Supplementary Proofs

**Theorem 1.** Let $x_1, x_2, N \in \mathbb{R}$ such that $N > |x_1 - x_2| \geq 0$ and $A_1, \ldots, A_n, B_1, \ldots, B_m$ be binary variables with $n + m > 0$ such that $n, m \in \mathbb{N}_0$. Then the implication $A_1 = 1 \land \ldots \land A_n = 1 \land B_1 = 0 \land \ldots \land B_m = 0 \Rightarrow x_1 = x_2$ is satisfied by using the following constraints:

\[
    N(1 - A_1) + \ldots + N(1 - A_n) + NB_1 + \ldots + NB_m + x_1 \geq x_2
\]

\[
    N(1 - A_1) + \ldots + N(1 - A_n) + NB_1 + \ldots + NB_m + x_2 \geq x_1
\]

**Proof.** Assume $A_i = 1$ and $B_j = 0$ for all $i \in \{1, \ldots, n\}, j \in \{1, \ldots, m\}$. Then, the constraints imply $x_1 \geq x_2$ and $x_2 \geq x_1$. Therefore, $x_1 = x_2$ holds. Note that both constraints are always satisfied if there exist at least one $A_i$ with $A_i = 0$ or one $B_j$ with $B_j = 1$. This is true since $N > |x_1 - x_2| \geq 0$ holds by the assumption.

The following corollaries follow directly from Theorem 1.

**Corollary 1.** Implication (5) is satisfied by

\[
    (1 - B_{cl}^k) + (1 - B_{fl}^M) + O_{efl} \geq 1.
\]

**Corollary 2.** Implication (6) is satisfied by

\[
    (1 - T_l) + (1 - B_{cl}^M) + (1 - B_{fl}^R) + O_{efl} \geq 1.
\]

**Corollary 3.** Implication (7) is satisfied by

\[
    (1 - U_l) + (1 - B_{cl}^M) + (1 - B_{fl}^R) + O_{efl} \geq 1.
\]

**Corollary 4.** Implication (8) is satisfied by

\[
    (1 - I_l) + (1 - B_{cl}^M) + (1 - B_{fl}^R) + e \geq f,
\]

\[
    (1 - I_l) + (1 - B_{cl}^M) + (1 - B_{fl}^R) + f \geq e.
\]
Corollary 5. Implication (10) is satisfied by
\[ I_l \leq C^L_l, \]
\[ (1 - I_l) \geq C^R_l. \]

Corollary 6. Implication (11) is satisfied by
\[ (1 - T_l) \geq C^L_l, \]
\[ (1 - T_l) \geq C^R_l. \]

Corollary 7. Implication (12) is satisfied by
\[ (1 - U_l) + 1 \geq C^L_l + C^R_l \]
\[ (1 - U_l) + C^L_l + C^R_l \geq 1. \]

Corollary 8. Implication (13) is satisfied by
\[ (1 - B^X_{el}) \geq L^X_{el}. \]

Corollary 9. Implication (14) is satisfied by
\[ (1 - O_{efl}) + (1 - B^X_{fl}) + L^X_{fl} \geq 1. \]

Corollary 10. Implication (15) is satisfied by
\[ (1 - O_{efl}) + (1 - B^X_{fl}) \geq L^X_{el}. \]

Corollary 11. Implication (16) is satisfied by
\[ B^X_{el} \leq R^X_{el}. \]

Corollary 12. Implication (17) is satisfied by
\[ (1 - O_{efl}) + (1 - B^X_{fl}) + R^X_{el} \geq 1. \]

Corollary 13. Implication (18) is satisfied by
\[ (1 - O_{efl}) + (1 - B^X_{fl}) \geq R^X_{el}. \]

Corollary 14. Implication (19) is satisfied by
\[ (1 - R^L_{el}) + (1 - L^M_{el}) + W^L_{el} \geq 1. \]

Corollary 15. Implication (20) is satisfied by
\[ (1 - R^L_{el}) + L^M_{el} \geq W^L_{el}. \]

Corollary 16. Implication (21) is satisfied by
\[ (1 - L^M_{el}) + R^L_{el} \geq W^L_{el}. \]
Corollary 17. Implication (22) is satisfied by
\((1 - R_{el}^M) + (1 - L_{el}^R) + W_{el}^R \geq 1\).

Corollary 18. Implication (23) is satisfied by
\((1 - R_{el}^M) + L_{el}^R \geq W_{el}^R\).

Corollary 19. Implication (24) is satisfied by
\((1 - L_{el}^R) + R_{el}^M \geq W_{el}^R\).

Corollary 20. Implication (25) is satisfied by
\((1 - C_l^X) + (1 - W_{el}^X) + (1 - W_{fl}^X) + O_{efl+1} \geq (1 - O_{efl})\),
\((1 - C_l^X) + (1 - W_{el}^X) + (1 - W_{fl}^X) + (1 - O_{efl}) \geq O_{efl+1}\).

Corollary 21. Implication (26) is satisfied by
\((1 - W_{el}^L) + (1 - W_{fl}^R) + O_{efl+1} \geq (1 - O_{efl})\),
\((1 - W_{el}^L) + (1 - W_{fl}^R) + (1 - O_{efl}) \geq O_{efl+1}\).

Corollary 22. Implication (27) is satisfied by
\(W_{el}^L + W_{el}^R + O_{efl+1} \geq O_{efl}\),
\(W_{el}^L + W_{el}^R + O_{efl} \geq O_{efl+1}\).

Corollary 23. Implication (28) is satisfied by
\(W_{fl}^L + W_{fl}^R + O_{efl+1} \geq O_{efl}\),
\(W_{fl}^L + W_{fl}^R + O_{efl} \geq O_{efl+1}\).

Corollary 24. Implication (29) is satisfied by
\(C_l^X + (1 - W_{el}^X) + (1 - W_{fl}^X) + O_{efl+1} \geq O_{efl}\),
\(C_l^X + (1 - W_{el}^X) + (1 - W_{fl}^X) + O_{efl} \geq O_{efl+1}\).

Corollary 25. Implication (30) is satisfied by
\((1 - C_l^X) + (1 - W_{el}^X) + O_{el+1} \geq (1 - S_{el})\),
\((1 - C_l^X) + (1 - W_{el}^X) + (1 - S_{el}) \geq S_{el+1}\).

Corollary 26. Implication (31) is satisfied by
\(C_l^X + (1 - W_{el}^X) + S_{el+1} \geq S_{el}\),
\(C_l^X + (1 - W_{el}^X) + S_{el} \geq S_{el+1}\).

Corollary 27. Implication (32) is satisfied by
\(W_{el}^L + W_{el}^R + S_{el+1} \geq S_{el}\),
\(W_{el}^L + W_{el}^R + S_{el} \geq S_{el+1}\).
Fig. S1:
Fraction of inversions of all scenarios which are obtained by solving (a) $\Pi_1$, (b) $\Pi_2$, and (c) $\Pi_3$. Only instances that have been solved in less than 90000s are included. The dotted polygon shows the all-type area and dashed lines show weighting schemes $(w_I, w_T, w_U)$ such that $w_T = w_U$ (respectively $w_I = w_T$ and $w_I = w_U$).
Fig. S2:
Fraction of transpositions of all scenarios which are obtained by solving a) $\Pi_1$, b) $\Pi_2$, and c) $\Pi_3$. Only instances that have been solved in less than 90000s are included. The dotted polygon shows the all-type area and dashed lines show weighting schemes $(w_I, w_T, w_U)$ such that $w_T = w_U$ (respectively $w_I = w_T$ and $w_I = w_U$).
Fig. S3:
Fraction of inverse transpositions of all scenarios which are obtained by solving (a) $\Pi_1$, (b) $\Pi_2$, and (c) $\Pi_3$. Only instances that have been solved in less than 90000s are included. The dotted polygon shows the all-type area and dashed lines show weighting schemes ($w_I, w_T, w_U$) such that $w_T = w_U$ (respectively $w_I = w_T$ and $w_I = w_U$).

Fig. S4:
Runtime for all weighting schemes which were used to solve $\Pi_i$ with $i \in [1 : 3]$. 
3 Supplementary Tables

<table>
<thead>
<tr>
<th>Constants</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>number of genes of $\pi$</td>
</tr>
<tr>
<td>$d$</td>
<td>distance $d(\pi, \iota)$ between $\pi$ and $\iota$</td>
</tr>
<tr>
<td>$l$</td>
<td>$l \in [0 : d]$</td>
</tr>
<tr>
<td>$w_I$</td>
<td>cost of an inversion</td>
</tr>
<tr>
<td>$w_T$</td>
<td>cost of a transposition</td>
</tr>
<tr>
<td>$w_U$</td>
<td>cost of an inverse transposition</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Binary Variables</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$O_{\mathrm{eft}}$</td>
<td>$O_{\mathrm{eft}} = 1 \Leftrightarrow$ gene $e$ is to the left of gene $f$ in permutation $\pi_l$</td>
</tr>
<tr>
<td>$S_{\mathrm{el}}$</td>
<td>$S_{\mathrm{el}} = 1 \Leftrightarrow$ $e$ has a negative sign in $\pi_l$</td>
</tr>
<tr>
<td>$I_l$</td>
<td>$I_l = 1 \Leftrightarrow$ $\rho_l$ is an inversion with $\rho_l(\pi_l) = \pi_{l+1}$</td>
</tr>
<tr>
<td>$T_l$</td>
<td>$T_l = 1 \Leftrightarrow$ $\rho_l$ is a transposition with $\rho_l(\pi_l) = \pi_{l+1}$</td>
</tr>
<tr>
<td>$U_l$</td>
<td>$U_l = 1 \Leftrightarrow$ $\rho_l$ is an inverse transposition with $\rho_l(\pi_l) = \pi_{l+1}$</td>
</tr>
<tr>
<td>$C^X_l$</td>
<td>$C^X_l = 1 \Leftrightarrow$ $X$-th interval gets inverted from $\pi_l$ to $\pi_{l+1}, X \in {L, R}$</td>
</tr>
<tr>
<td>$B^X_{\mathrm{el}}$</td>
<td>$B^X_{\mathrm{el}} = 1 \Leftrightarrow \pi_l(e)$ is the $X$-th bounding element, $X \in {L, M, R}$</td>
</tr>
<tr>
<td>$L^X_{\mathrm{el}}$</td>
<td>$L^X_{\mathrm{el}} = 1 \Leftrightarrow \pi_l(e)$ is left of $X$-th bounding element, $X \in {M, R}$</td>
</tr>
<tr>
<td>$R^X_{\mathrm{el}}$</td>
<td>$R^X_{\mathrm{el}} = 1 \Leftrightarrow \pi_l(e)$ is right of $X$-th bounding element, $X \in {L, M}$</td>
</tr>
<tr>
<td>$W^X_{\mathrm{el}}$</td>
<td>$W^X_{\mathrm{el}} = 1 \Leftrightarrow \pi_l(e)$ is between $B^L_l$ and $B^M_l$ or $B^M_l$ and $B^R_l$ with $X \in {L, R}$, respectively</td>
</tr>
</tbody>
</table>

**Table S1:**
Variables that are used in the ILP formulation.