Restricted DCJ-indel model revisited

Marília Braga
Inmetro
Brazil

Jens Stoye
Bielefeld University
Germany

BSB 2013
Overview

1 Motivation and Background

2 Restricted DCJ model with ER composition

3 The restricted DCJ-indel model

4 Conclusions and Future Work
Overview

1. Motivation and Background

2. Restricted DCJ model with ER composition

3. The restricted DCJ-indel model

4. Conclusions and Future Work
Motivation and Background

Genomes are composed of one or more DNA molecules:

**Prokaryotic Cell**
- Chromosome
- Plasmids

**Eukaryotic Cell**
- Nucleus
- Chromosomes
- Nucleotides
- Antiparallel strands
Motivation and Background

Each chromosome contains a set of genes:
Motivation and Background

Each chromosome contains a set of genes:

\[ \text{The DNA strand in which each gene lies gives its orientation} \]
Motivation and Background

Each chromosome contains a set of genes:

\[ \text{a} \quad \text{b} \]

the DNA strand in which each gene lies gives its orientation
Motivation and Background

Each chromosome contains a set of genes:

\[ a \quad b \quad c \quad d \]

the DNA strand in which each gene lies gives its orientation...
Motivation and Background

Each chromosome contains a set of genes:

```
  a  b  c  d
```

Genomes are subject to large scale mutations...
Motivation and Background

Each chromosome contains a set of genes:

An inversion changes the order and the orientation of genes.

Genomes are subject to large scale mutations...
Motivation and Background

Each chromosome contains a set of genes:

- $a$ - $b$ - $c$ - $d$

An inversion changes the order and the orientation of genes.

- $a$ - $d$ - $c$ - $b$

Genomes are subject to large scale mutations...
...inversions, translocations, fusions, fissions...
Motivation and Background

Each chromosome contains a set of genes:

An inversion changes the order and the orientation of genes.

Genomes are subject to large scale mutations...
...inversions, translocations, fusions, fissions...
...insertions, deletions, duplications...
Motivation and Background

Comparing two distinct genomes

Common genes: $G = \{a, b, c, d, e\}$

Unique genes:

$A = \{u, v, w\}$

$B = \{x, z\}$

Insertions and Deletions (Indels), Substitutions and Duplications change the content of the genome

Rearrangements change the organization of the genome and can be modeled by the Double Cut and Join (DCJ) ([Yancopoulos, Attie and Friedberg, 2005])

Genomic distance: minimum number of operations that sort one genome into another

Restricted DCJ-indel
Motivation and Background

Comparing two distinct genomes

Common genes:  \[ G = \{a, b, c, d, e\} \] Unique genes:  \[ \mathcal{A} = \{u, v, w\} \]
\[ \mathcal{B} = \{x, z\} \]

A  \[ b \quad a \quad u \quad d \quad e \quad v \quad w \quad c \]
\[ b \quad a \quad e \quad d \quad u \quad v \quad w \quad c \]
\[ \downarrow \text{inversion} \]

B  \[ a \quad b \quad c \quad x \quad d \quad z \quad e \]
Motivation and Background

Comparing two distinct genomes

Common genes: $\mathcal{G} = \{a, b, c, d, e\}$

Unique genes:
- $\mathcal{A} = \{u, v, w\}$
- $\mathcal{B} = \{x, z\}$

A: $\begin{align*}
\text{inversion} & \quad \downarrow \\
\text{deletion} & \quad \downarrow
\end{align*}$

B: $\begin{align*}
\end{align*}$

Insertions and Deletions, Substitutions and Duplications change the content of the genome. Rearrangements change the organization of the genome and can be modeled by the Double Cut and Join (DCJ) [Yancopoulos, Attie and Friedberg, 2005].

Genomic distance: minimum number of operations that sort one genome into another.
Motivation and Background

Comparing two distinct genomes

Common genes: \( S = \{a, b, c, d, e\} \)

Unique genes:

\( \mathcal{A} = \{u, v, w\} \)

\( \mathcal{B} = \{x, z\} \)

\[ \begin{align*}
A & \quad b \quad a \quad u \quad d \quad e \quad v \quad w \quad c \\
& \quad \downarrow \text{inversion} \\
& \quad b \quad a \quad e \quad d \quad U \quad V \quad W \quad c \\
& \quad \downarrow \text{deletion} \\
& \quad b \quad a \quad e \quad d \quad c \\
& \quad \downarrow \text{insertion} \\
& \quad b \quad a \quad e \quad d \quad z \quad x \quad c
\end{align*} \]

\[ B \quad a \quad b \quad c \quad x \quad d \quad z \quad e \]
Motivation and Background

Comparing two distinct genomes

Common genes: $\mathcal{G} = \{a, b, c, d, e\}$
Unique genes: $\mathcal{A} = \{u, v, w\}$ $\mathcal{B} = \{x, z\}$

Diagram:

A: $b \rightarrow a \rightarrow u \rightarrow d \rightarrow e \rightarrow v \rightarrow w \rightarrow c$
  - Inversion
  - Deletion
  - Insertion
  - Fission

B: $a \rightarrow b \rightarrow c \rightarrow x \rightarrow d \rightarrow z \rightarrow e$
Motivation and Background

Comparing two distinct genomes

Common genes: \( \mathcal{G} = \{a, b, c, d, e\} \)

Unique genes:
\( \mathcal{A} = \{u, v, w\} \)
\( \mathcal{B} = \{x, z\} \)

A
\[ b \quad a \quad u \quad d \quad e \quad v \quad w \quad c \]
\[ \downarrow \text{inversion} \]
\[ b \quad a \quad e \quad d \quad U \quad v \quad w \quad c \]
\[ \downarrow \text{deletion} \]
\[ b \quad a \quad e \quad d \quad c \]
\[ \downarrow \text{insertion} \]
\[ b \quad a \quad e \quad d \quad z \quad x \quad c \]
\[ \downarrow \text{fission} \]
\[ b \]
\[ \downarrow \text{translocation} \]
\[ b \quad a \quad e \quad d \quad z \quad x \quad c \]
\[ a \quad b \quad c \quad x \quad z \quad d \quad e \]

B
\[ a \quad b \quad c \quad x \quad d \quad z \quad e \]
Motivation and Background

Comparing two distinct genomes

Common genes: \( G = \{a, b, c, d, e\} \)
Unique genes:

\[ A = \{u, v, w\} \]
\[ B = \{x, z\} \]

\[ \Delta \text{ - inversion} \]
\[ \nabla \text{ - deletion} \]
\[ \nabla \text{ - insertion} \]
\[ \nabla \text{ - fission} \]
\[ \Delta \text{ - translocation} \]

Genomic distance: minimum number of operations that sort one genome into another


Restricted DCJ-indel

Insertions and Deletions (Indels), Substitutions and Duplications change the content of the genome.
Rearrangements change the organization of the genome and can be modeled by the Double Cut and Join (DCJ).

Marilia Braga
Motivation and Background

Comparing two distinct genomes

Common genes: \( S = \{ a, b, c, d, e \} \)
Unique genes: \( A = \{ u, v, w \} \)
\( B = \{ x, z \} \)

A

\[ \begin{array}{l}
\text{b} \quad \text{a} \quad \text{u} \quad \text{d} \quad \text{e} \quad \text{v} \quad \text{w} \quad \text{c} \\
\downarrow \text{inversion} \\
\text{b} \quad \text{a} \quad \text{e} \quad \text{d} \quad \text{u} \quad \text{v} \quad \text{w} \quad \text{c} \\
\downarrow \text{deletion} \\
\text{b} \quad \text{a} \quad \text{e} \quad \text{d} \quad \text{c} \\
\downarrow \text{insertion} \\
\text{b} \quad \text{a} \quad \text{e} \quad \text{d} \quad \text{z} \quad \text{x} \quad \text{c} \\
\downarrow \text{fission} \\
\text{b} \quad \text{a} \quad \text{e} \quad \text{d} \quad \text{z} \quad \text{x} \quad \text{c} \quad \text{translocation} \\
\downarrow \text{translocation} \\
\text{a} \quad \text{b} \quad \text{c} \quad \text{z} \quad \text{x} \quad \text{d} \quad \text{e} \\
\downarrow \text{inversion} \\
\text{a} \quad \text{b} \quad \text{c} \quad \text{x} \quad \text{z} \quad \text{d} \quad \text{e} \\
\end{array} \]

B

\[ \begin{array}{l}
\text{a} \quad \text{b} \quad \text{c} \quad \text{x} \quad \text{d} \quad \text{z} \quad \text{e} \\
\end{array} \]

Insertions and Deletions (Indels), Substitutions and Duplications change the content of the genome. Rearrangements change the organization of the genome and can be modeled by the Double Cut and Join (DCJ) [Yancopoulos, Attie and Friedberg, 2005].

Genomic distance: minimum number of operations that sort one genome into another.
Motivation and Background

Comparing two distinct genomes

Common genes: \( \mathcal{G} = \{a, b, c, d, e\} \)

Unique genes: \( \mathcal{A} = \{u, v, w\} \)
\( \mathcal{B} = \{x, z\} \)

Insertions and Deletions - (Indels), Substitutions and Duplications change the content of the genome
Motivation and Background

Comparing two distinct genomes

Common genes: \( \mathcal{G} = \{a, b, c, d, e\} \)

Unique genes: \( \mathcal{A} = \{u, v, w\} \)
\( \mathcal{B} = \{x, z\} \)

Rearrangements change the organization of the genome and can be modeled by the Double Cut and Join - (DCJ)

[Yancopoulos, Attie and Friedberg, 2005]

Insertions and Deletions - (Indels), Substitutions and Duplications change the content of the genome
Motivation and Background

Comparing two distinct genomes

Common genes: \( G = \{a, b, c, d, e\} \)

Unique genes:
- \( A = \{u, v, w\} \)
- \( B = \{x, z\} \)

Rearrangements change the organization of the genome and can be modeled by the Double Cut and Join - (DCJ)

[Yancopoulos, Attie and Friedberg, 2005]

Insertions and Deletions - (Indels), Substitutions and Duplications change the content of the genome

Genomic distance: minimum number of operations that sort one genome into another
Motivation and Background

Some polynomial models for the genomic distance

**Inversion model**  [Hannenhalli & Pevzner 1995, Meidanis et al. 2000]
- unichromosomal genomes with the **same content** and no duplications
- only inversions
Motivation and Background

Some polynomial models for the genomic distance

**Inversion model** [Hannenhalli & Pevzner 1995, Meidanis *et al.* 2000]
- unichromosomal genomes with the **same content** and no duplications
- only inversions

**HP model** [Hannenhalli & Pevzner 1995]
- multichromosomal linear genomes with the **same content** and no duplications
- inversions, translocations, fusions and fissions
Motivation and Background

Some polynomial models for the genomic distance

**Inversion model**  [Hannenhalli & Pevzner 1995, Meidanis *et al.* 2000]
- unichromosomal genomes with the **same content** and no duplications
- only inversions

**HP model**  [Hannenhalli & Pevzner 1995]
- multichromosomal linear genomes with the **same content** and no duplications
- inversions, translocations, fusions and fissions

**The double cut and join (DCJ) model**  [Yancopoulos *et al.* 2005, Bergeron *et al.* 2006]
- multichromosomal genomes with the **same content** and no duplications
- inversions, translocations, fusions, fissions, **circular excisions** and **reincorporations**
- **distance** can be easily computed in **linear time**
Motivation and Background

For genomes with unequal contents but no duplications

**Inversions with indels** [El-Mabrouk 2000, RECOMB-CG 2013]

- unichromosomal, allows **inversions** and **indels** with the **same cost**
- no exact solution yet, but for an important set of instances, the inversion-indel distance can be computed in linear time
Motivation and Background

For genomes with unequal contents but no duplications

**Inversions with indels**  [El-Mabrouk 2000, RECOMB-CG 2013]
- unichromosomal, allows *inversions* and *indels* with the *same cost*
- no exact solution yet, but for an important set of instances, the inversion-indel distance can be computed in linear time

**DCJ with indels**  [Yancopoulou & Friedberg 2008, WABI 2010/2012, Compeau 2012, submit. to LATIN 2014]
- multichromosomal, allows *DCJ operations* and *indels* with *distinct costs*
- *distance* can be computed in *linear time*
Motivation and Background

For genomes with unequal contents but no duplications

Inversions with indels [El-Mabrouk 2000, RECOMB-CG 2013]
- unichromosomal, allows inversions and indels with the same cost
- no exact solution yet, but for an important set of instances, the inversion-indel distance can be computed in linear time

- multichromosomal, allows DCJ operations and indels with distinct costs
- distance can be computed in linear time

DCJ with substitutions [RECOMB-CG 2011, BSB 2012, AMB 2013]
- multichromosomal, allows DCJ operations and substitutions with distinct costs
- more parsimonious than DCJ-indel (substitutions include indels)
- distance can be computed in linear time
Overview

1 Motivation and Background

2 Restricted DCJ model with ER composition

3 The restricted DCJ-indel model

4 Conclusions and Future Work
Restricted DCJ model with ER composition

Sorting linear genomes...

General scenario

\[ a ightarrow e \rightarrow d \rightarrow c \rightarrow b \rightarrow f \]

\[ a \rightarrow b \rightarrow c \rightarrow d \rightarrow e \rightarrow f \]
Restricted DCJ model with ER composition

Sorting linear genomes...

**General scenario**

```
<table>
<thead>
<tr>
<th>a</th>
<th>e</th>
<th>d</th>
<th>c</th>
<th>b</th>
<th>f</th>
</tr>
</thead>
</table>
```

- **Excision**

```
<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>f</th>
</tr>
</thead>
</table>
```

```
| e | d | c |
```

- **Reincorporation**

```
<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>f</th>
</tr>
</thead>
</table>
```

```
| e | c |
```

```
| d |
```

```
| a | b | c | e | f |
```

Many circular chromosomes can coexist. A circular chromosome is immediately reincorporated after its excision (ER composition).
Restricted DCJ model with ER composition

Sorting linear genomes...

General scenario

\[ \begin{array}{cccccc}
    a & e & d & c & b & f \\
  \end{array} \]

excision \[ \downarrow \]

\[ \begin{array}{cccccc}
    a & b & f \\
  \end{array} \]

\[ \begin{array}{cccccc}
    e & d & c \\
  \end{array} \]

excision \[ \downarrow \]

\[ \begin{array}{cccccc}
    a & b & f \\
  \end{array} \]

reincorporation \[ \downarrow \]

\[ \begin{array}{cccccc}
    e & c \\
  \end{array} \]

\[ \begin{array}{cccccc}
    d \\
  \end{array} \]

reincorporation \[ \downarrow \]

\[ \begin{array}{cccccc}
    a & b & c & e & f \\
  \end{array} \]

\[ \begin{array}{cccccc}
    d \\
  \end{array} \]

reincorporation \[ \downarrow \]

\[ \begin{array}{cccccc}
    a & b & c & d & e & f \\
  \end{array} \]

Many circular chromosomes can coexist in the intermediate species.
Restricted DCJ model with ER composition

Sorting linear genomes...

**General scenario**

```
  a  e  d  c  b  f
  ↓ excision
  a  b  f  e  d  c
  ↓ excision
  a  b  f  e  c  d
  ↓ reincorporation
  a  b  c  e  f  d
  ↓ reincorporation
  a  b  c  d  e  f
```

**Restricted scenario**

```
  a  e  d  c  b  f
  a  e  d  c  b  f
  a  b  f  e  d  c
  a  b  f  e  d  c
  a  b  c  e  f  d
  a  b  c  e  f  d
  a  b  c  d  e  f
  a  b  c  d  e  f
```

Many circular chromosomes can coexist in the intermediate species.
Restricted DCJ model with ER composition

Sorting linear genomes...

**General scenario**

- Initial genome: $a\ e\ d\ c\ b\ f$
- Excision: $a\ b\ f$
- Reincorporation: $a\ b\ f$
- Excision: $e\ d\ c$
- Reincorporation: $a\ b\ c\ e\ f$
- Reincorporation: $a\ b\ c\ d\ e\ f$

- Initial genome: $a\ b\ f$
- Excision: $e\ d\ c$
- Reincorporation: $a\ b\ d\ c\ e\ f$
- Excision: $a\ b\ f$
- Reincorporation: $a\ b\ c\ d\ e\ f$

**Restricted scenario**

- Initial genome: $a\ e\ d\ c\ b\ f$
- Excision: $a\ b\ f$
- Reincorporation: $a\ e\ d\ c$
- Excision: $a\ b\ f$
- Reincorporation: $a\ b\ c\ e\ f$
- Reincorporation: $a\ b\ c\ d\ e\ f$

Many circular chromosomes can coexist in the intermediate species.
Restricted DCJ model with ER composition

Sorting linear genomes...

**General scenario**

- $a\rightarrow e\rightarrow d\rightarrow c\rightarrow b\rightarrow f$
- $a\rightarrow b\rightarrow f$
- $a\rightarrow b\rightarrow f$
- $a\rightarrow b\rightarrow f$
- $a\rightarrow b\rightarrow c\rightarrow e\rightarrow f$
- $a\rightarrow b\rightarrow c\rightarrow d\rightarrow e\rightarrow f$

Many circular chromosomes can coexist in the intermediate species.

**Restricted scenario**

- $a\rightarrow e\rightarrow d\rightarrow c\rightarrow b\rightarrow f$
- $a\rightarrow b\rightarrow f$
- $a\rightarrow b\rightarrow f$
- $a\rightarrow b\rightarrow c\rightarrow e\rightarrow f$
- $a\rightarrow b\rightarrow c\rightarrow d\rightarrow e\rightarrow f$

A circular chromosome is immediately reincorporated after its excision (ER composition).
Restricted DCJ model with ER composition

Sorting linear genomes...

**General scenario**

- Initial sequence: \(a\ e\ d\ c\ b\ f\)
- Excision: \(a\ e\ d\ c\ b\ f\) \(\rightarrow\) \(a\ b\ f\) \(\rightarrow\) \(a\ b\ f\) \(\rightarrow\) \(a\ b\ c\ e\ f\) \(\rightarrow\) \(a\ b\ c\ d\ e\ f\)

**Restricted scenario**

- Initial sequence: \(a\ e\ d\ c\ b\ f\)
- Excision: \(a\ e\ d\ c\ b\ f\) \(\rightarrow\) \(a\ b\ f\) \(\rightarrow\) \(a\ b\ f\) \(\rightarrow\) \(a\ b\ c\ d\ e\ f\)

Many circular chromosomes can coexist in the intermediate species.

A circular chromosome is immediately reincorporated after its excision (ER composition).
Restricted DCJ model with ER composition

Sorting linear genomes...

**General scenario**

\[
\begin{align*}
&\text{a} \rightarrow \text{e} \rightarrow \text{d} \rightarrow \text{c} \rightarrow \text{b} \rightarrow \text{f} \\
&\text{a} \rightarrow \text{b} \rightarrow \text{f} \\
&\text{a} \rightarrow \text{b} \rightarrow \text{f} \\
&\text{a} \rightarrow \text{b} \rightarrow \text{f} \\
&\text{a} \rightarrow \text{b} \rightarrow \text{c} \rightarrow \text{e} \rightarrow \text{f} \rightarrow \text{d}
\end{align*}
\]

Many circular chromosomes can coexist in the intermediate species.

**Restricted scenario**

\[
\begin{align*}
&\text{a} \rightarrow \text{e} \rightarrow \text{d} \rightarrow \text{c} \rightarrow \text{b} \rightarrow \text{f} \\
&\text{a} \rightarrow \text{b} \rightarrow \text{f} \\
&\text{a} \rightarrow \text{b} \rightarrow \text{d} \rightarrow \text{c} \rightarrow \text{e} \rightarrow \text{f} \\
&\text{a} \rightarrow \text{b} \rightarrow \text{d} \rightarrow \text{c} \rightarrow \text{e} \rightarrow \text{f} \\
&\text{a} \rightarrow \text{b} \rightarrow \text{c} \rightarrow \text{d} \rightarrow \text{e} \rightarrow \text{f}
\end{align*}
\]

A circular chromosome is immediately reincorporated after its excision (ER composition).
Restricted DCJ model with ER composition

External and Internal Block-Interchange

External (2 translocations)

Internal (ER composition)

Internal (3 inversions)
Restricted DCJ model with ER composition

- Inversions, translocations, fusions and fissions: mimicked by a single DCJ

- Internal transpositions and block-interchanges: mimicked by two consecutive DCJs (ER composition)

- The distance is the same for both the general and the restricted DCJ models
  [Yancopoulos et al. 2005]

- A restricted DCJ scenario can be obtained in $\Theta(n \log n)$ time
  [Kováč et al. 2010]
Overview

1. Motivation and Background

2. Restricted DCJ model with ER composition

3. The restricted DCJ-indel model

4. Conclusions and Future Work
The restricted DCJ-indel model

- each DCJ costs 1

- each indel costs positive $w$

- inversions, translocations, fusions and fissions: mimicked by a single DCJ

- internal transpositions and block-interchanges: mimicked by two consecutive DCJs (ER composition)
The restricted DCJ-indel model

General scenario

\[ a \rightarrow u \rightarrow e \rightarrow d \rightarrow c \rightarrow v \rightarrow b \rightarrow f \]

\[ a \rightarrow b \rightarrow c \rightarrow d \rightarrow e \rightarrow f \]
The restricted DCJ-indel model

General scenario

- excision (1) ↓
- deletion (w) ↓
- reincorporation (1) ↓
- reincorporation (1) ↓
The restricted DCJ-indel model

**General scenario**

- excision (1)
- deletion (w)
- reincorporation (1)

Many circular chromosomes can coexist in the intermediate species.
The restricted DCJ-indel model

General scenario

Restricted scenario

Many circular chromosomes can coexist in the intermediate species.
Many circular chromosomes can coexist in the intermediate species.
The restricted DCJ-indel model

**General scenario**

```
  a  u  e  d  c  v  b  f
  ↓
  a  b  f
  ↓
  a  b  f
  ↓
  a  b  f
  ↓
  a  b  c  e  f
  ↓
  a  b  c  d  e  f
```

```
  a  u  e  d  c  v  b  f
  ↓
  a  b  f
  ↓
  e  d  c  v  u
  ↓
  a  b  c  d  e  f
```

Many circular chromosomes can coexist in the intermediate species.

**Restricted scenario**

```
  a  u  e  d  c  v  b  f
  ↓
  a  b  f
  ↓
  e  d  c  v  u
  ↓
  a  b  c  d  e  f
  ↓
  a  b  c  d  e  f
```

```
  a  u  e  d  c  v  b  f
  ↓
  a  b  f
  ↓
  e  d  c  v  u
  ↓
  a  b  c  d  e  f
  ↓
  a  b  c  d  e  f
```

A circular chromosome is immediately reincorporated after its excision. (ER composition)
The restricted DCJ-indel model

General scenario

Restricted scenario

Many circular chromosomes can coexist in the intermediate species.

A circular chromosome is immediately reincorporated after its excision. (ER composition)
The restricted DCJ-indel model

**General scenario**

- excision (1)

**Restricted scenario**

- deletion (w)

A circular chromosome is immediately reincorporated after its excision. (ER composition)
The restricted DCJ-indel model

General scenario

A circular chromosome is immediately reincorporated after its excision. (ER composition)

Restricted scenario

Many circular chromosomes can coexist in the intermediate species.

Are both the general and the restricted DCJ-indel distances the same? [RECOMB-CG 2012]
The restricted DCJ-indel model

Moving insertions up and deletions down

A deletion can always be moved down [RECOMB-CG 2012]

Analogously, an insertion can always be moved up [RECOMB-CG 2012]
The restricted DCJ-indel model

Moving insertions up and deletions down

A deletion can always be moved down [RECOMB-CG 2012]
The restricted DCJ-indel model

Moving insertions up and deletions down

A deletion can always be moved down

[RECOMB-CG 2012]
The restricted DCJ-indel model

Moving insertions up and deletions down

A deletion can always be moved down

[RECOMB-CG 2012]
The restricted DCJ-indel model

Moving insertions up and deletions down

A deletion can always be moved down.[RECOMB-CG 2012]

Analogously, an insertion can always be moved up.[RECOMB-CG 2012]
The restricted DCJ-indel model

Moving insertions up and deletions down

A deletion can always be moved down

[RECOMB-CG 2012]
The restricted DCJ-indel model

Moving insertions up and deletions down

A deletion can always be moved down

Analogously, an insertion can always be moved up
The restricted DCJ-indel model

An algorithm to find a restricted DCJ-indel sorting scenario
The restricted DCJ-indel model

An algorithm to find a restricted DCJ-indel sorting scenario

\[ d' = d = n_1 + w n_2 \]

\[ O(n) \text{ time} \]
The restricted DCJ-indel model

An algorithm to find a restricted DCJ-indel sorting scenario

A \rightarrow B

\begin{align*}
&\text{general DCJ-indel} \\
&\text{insertions up} \\
&\text{deletions down} \\
&n_1 \text{ DCJs} \\
&n_2 \text{ indels} \\
&d = n_1 + wn_2 \\
&\mathcal{O}(n) \text{ time}
\end{align*}

\begin{align*}
&A' \rightarrow B' \\
&\text{general DCJ} \\
&\text{insertions} \\
&\text{deletions}
\end{align*}
The restricted DCJ-indel model

An algorithm to find a restricted DCJ-indel sorting scenario

\[ d = n_1 + w n_2 \]
\[ \mathcal{O}(n) \text{ time} \]

\[ d' = d = n_1 + w n_2 \]
\[ \mathcal{O}(n) \text{ time} \]
The restricted DCJ-indel model

An algorithm to find a restricted DCJ-indel sorting scenario

- $n_1$ DCJs
- $n_2$ indels
- $d = n_1 + wn_2$
- $\mathcal{O}(n)$ time

$A$

B

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>DCJ</td>
<td>DCJ</td>
</tr>
<tr>
<td>$n_1$</td>
<td>$n_2$</td>
</tr>
<tr>
<td>$d = n_1 + wn_2$</td>
<td>$d' = d = n_1 + wn_2$</td>
</tr>
<tr>
<td>$\mathcal{O}(n)$ time</td>
<td>$\mathcal{O}(n)$ time</td>
</tr>
</tbody>
</table>

$A'$

B'

A

B

$A'$

B'

A

B

insertions

deletions

<table>
<thead>
<tr>
<th>A</th>
<th>A'</th>
</tr>
</thead>
<tbody>
<tr>
<td>general DCJ</td>
<td>restricted DCJ</td>
</tr>
</tbody>
</table>

$B$
The restricted DCJ-indel model

An algorithm to find a restricted DCJ-indel sorting scenario

\[ d' = d = n_1 + wn_2 \]
\[ O(n) \text{ time} \]

\[ d'' = d' = d = n_1 + wn_2 \]
\[ O(n \log n) \text{ time} \]
Conclusions and Future Work

Overview

1. Motivation and Background
2. Restricted DCJ model with ER composition
3. The restricted DCJ-indel model
4. Conclusions and Future Work
Conclusions and Future Work

Summary of results

- The distance is the same for both the general and the restricted DCJ-indel models...
Conclusions and Future Work

Summary of results

- The distance is the same for both the general and the restricted DCJ-indel models...

  ... therefore the distance can be computed in $\Theta(n)$ time.
Conclusions and Future Work

Summary of results

▶ The distance is the same for both the general and the restricted DCJ-indel models...

... therefore the distance can be computed in $\Theta(n)$ time.

▶ A restricted scenario can be obtained in $\Theta(n \log n)$ time.
Conclusions and Future Work

Future work

**DCJ-substitution model**: the general and the restricted distances are not the same

**General scenario**

**Restricted scenario**

The restricted version of the DCJ-substitution distance is an open problem.
Conclusions and Future Work

Future work

**DCJ-substitution model**: the general and the restricted distances are not the same

The restricted version of the DCJ-substitution distance is an **open problem**.
Acknowledgements

This research was supported by:

▶ Inmetro
▶ CNPq (grant PROMETRO 563087/2010-2)
▶ Bielefeld University (Germany)
This research was supported by:

- Inmetro
- CNPq (grant PROMETRO 563087/2010-2)
- Bielefeld University (Germany)

Thank you for your attention!