

Genomátrendeződés és genom architektúra humán tumor genomokban

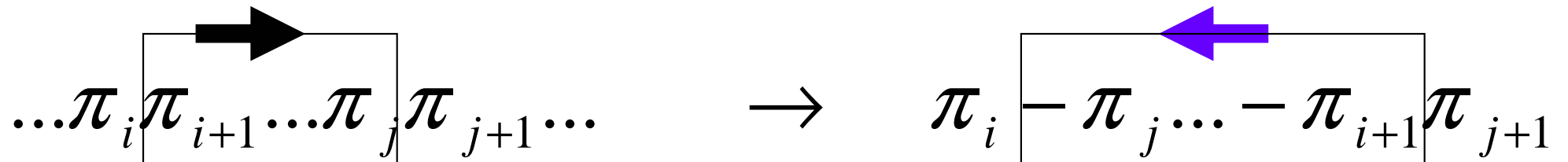
Miklós István

***MTA-ELTE Elméleti Biológiai és Ökológiai Modellező csoport,
Budapest, Hungary***

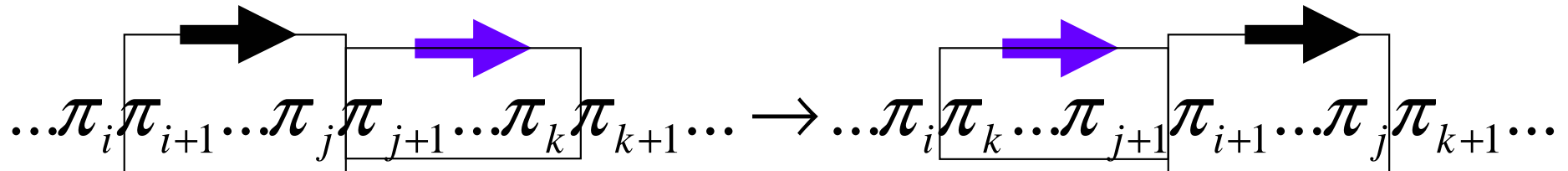
Bioinformatika konferencia, 2005. május 21. Balatonfüred

Genomátrendező mutációk

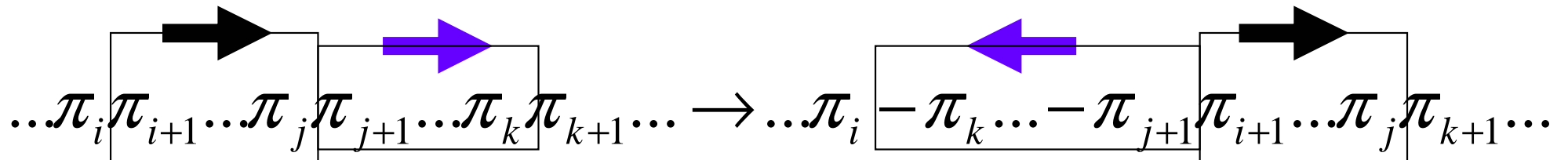
Inversions



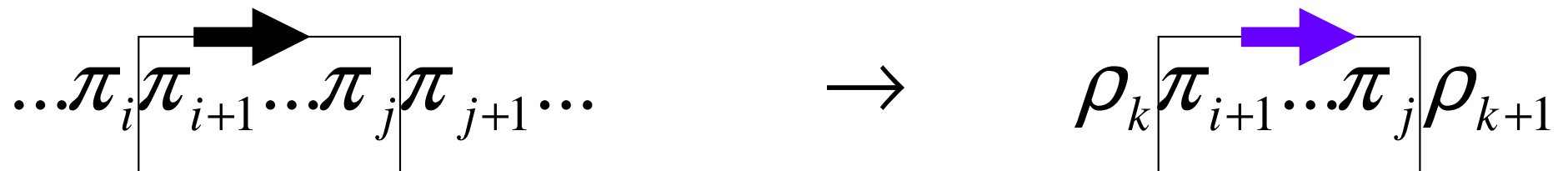
Transpositions



Inverted Transpositions



Translocations



Irodalmi áttekintés

Sturtevant & Novitski (1941) The homologies of chromosome elements and the genus *Drosophyla*. *Genetics* 26:517-541.

Watterson *et al.* (1982) The chromosome inversion problem. *J. Theor Biol.* 99:1-7.

Berman & Hannenhalli (1996) Fast Sorting by Reversal. *CPM1996*: 168-185

Miklós (2003) MCMC Genome Rearrangement. *Bioinformatics*, 19(Suppl.2):ii130-ii137.

Miklós, Ittész, Hein (2005) ParIS Genome Rearrangement server. *Bioinformatics*, 21:817 – 820.

Génátrendeződés tumor genomokban



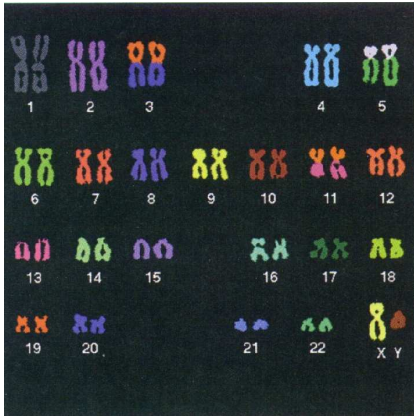
mFISH

- Minden szín egy kromoszóma egészséges genomban
- Többszínű kromoszóma az ábrán: átrendezett
- Alacsony felbontás, pontos térkép nem állapítható meg

End sequence profiling

Volik et al (2003) End-sequence profiling: Sequence-based analysis of aberrant genomes *PNAS* 100: 7696-7701

Humán genom

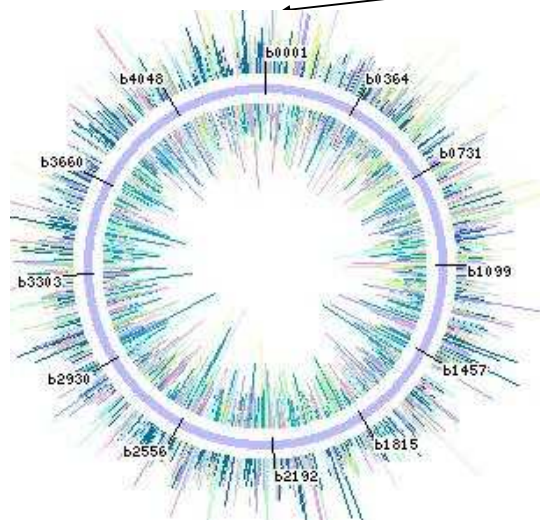


Hind III

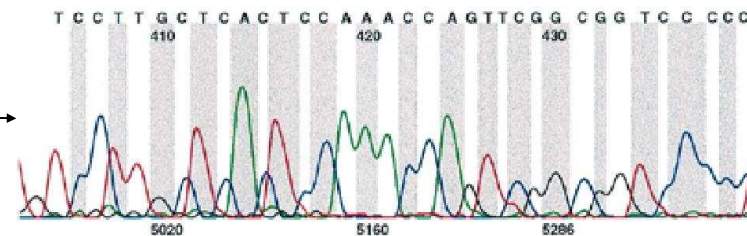


HPLC

Klónozás



Amplifikálás



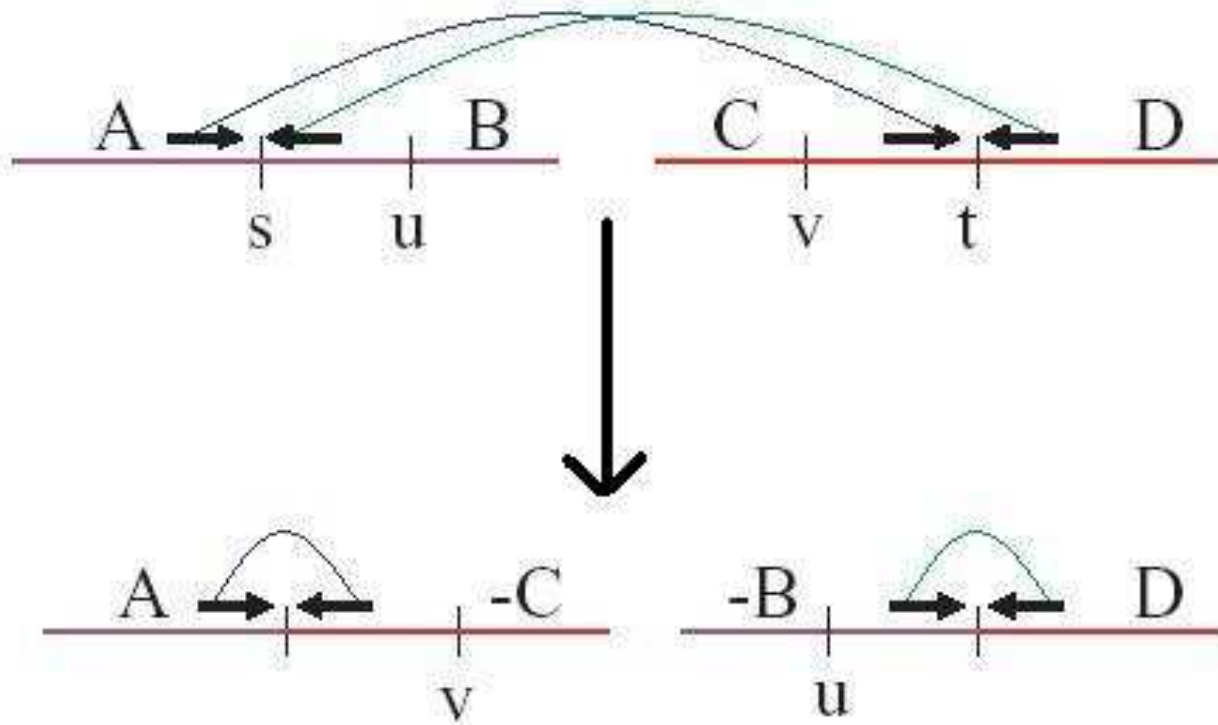
Szekvenálás

BAC

End sequence profiling

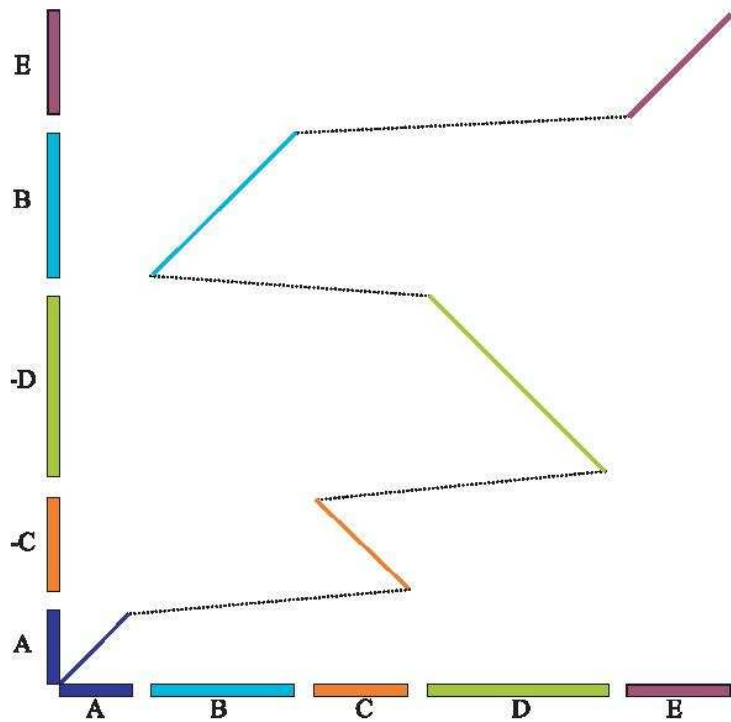
A szekvenálásokat BLASToljuk a humán genomra

- Nem csak a pozíció, az irány is informatív
- Felbontás kb. 140 kb

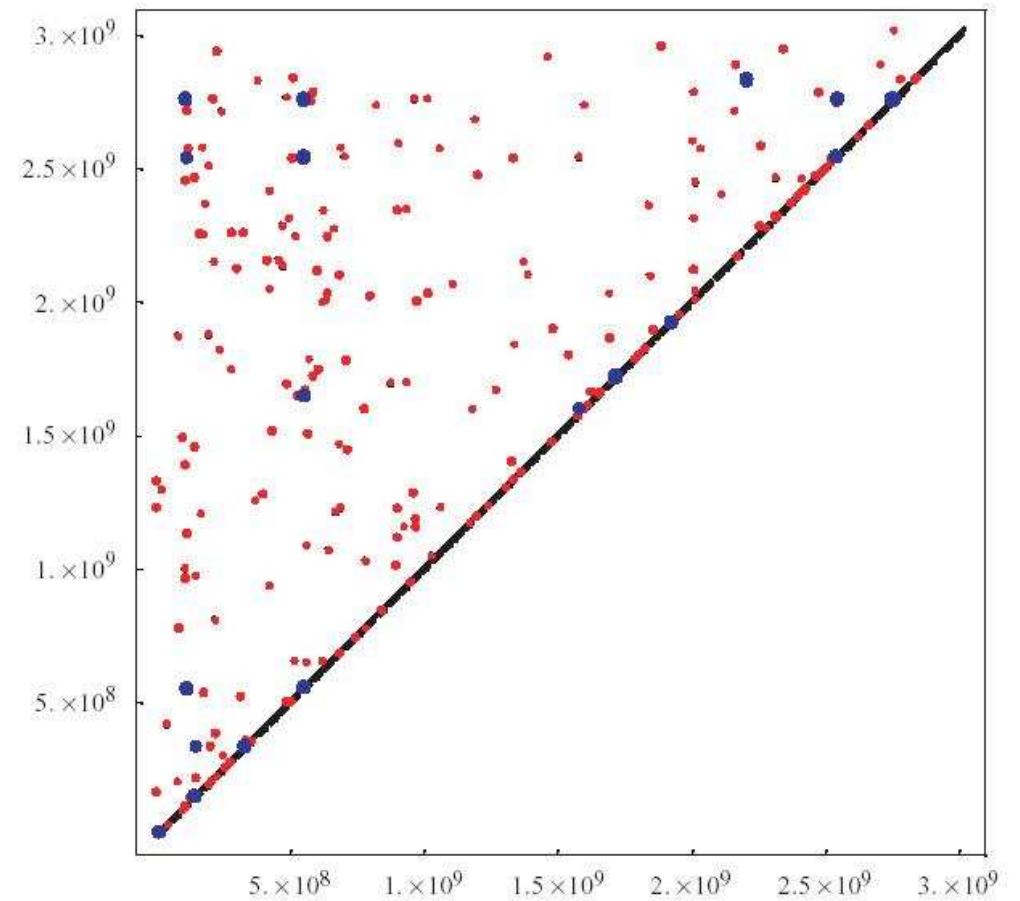


End sequence profiling

Ideális eset:



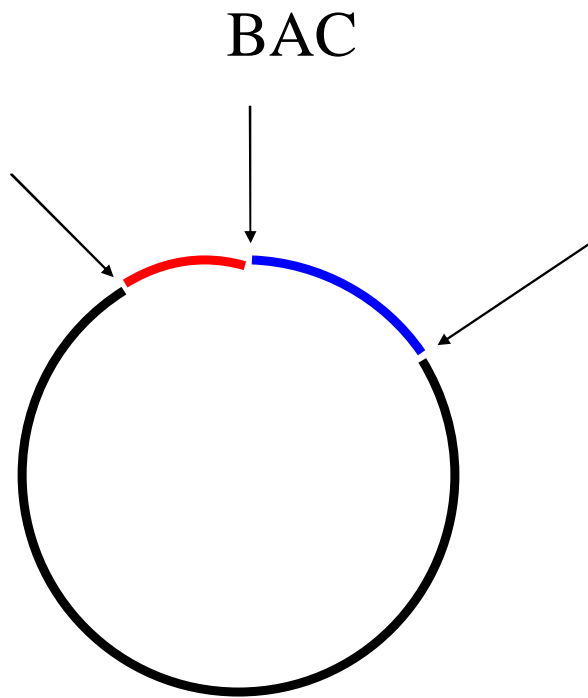
Gyakorlatban:



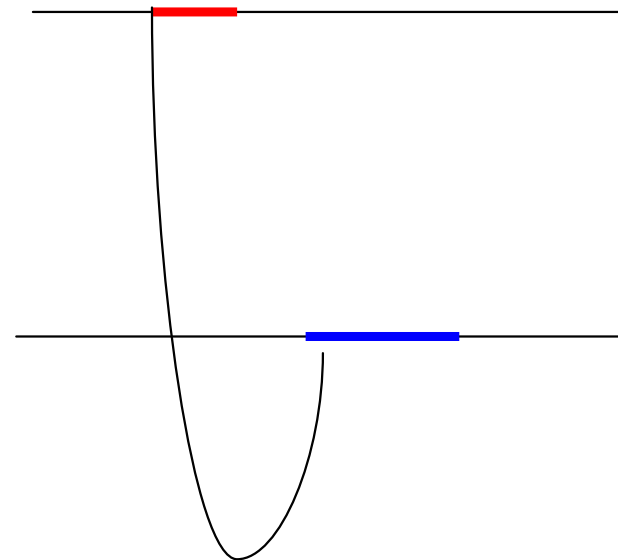
Problémák I.

Chimeric BACs

Két fragment kerül be a baci kromoszómába

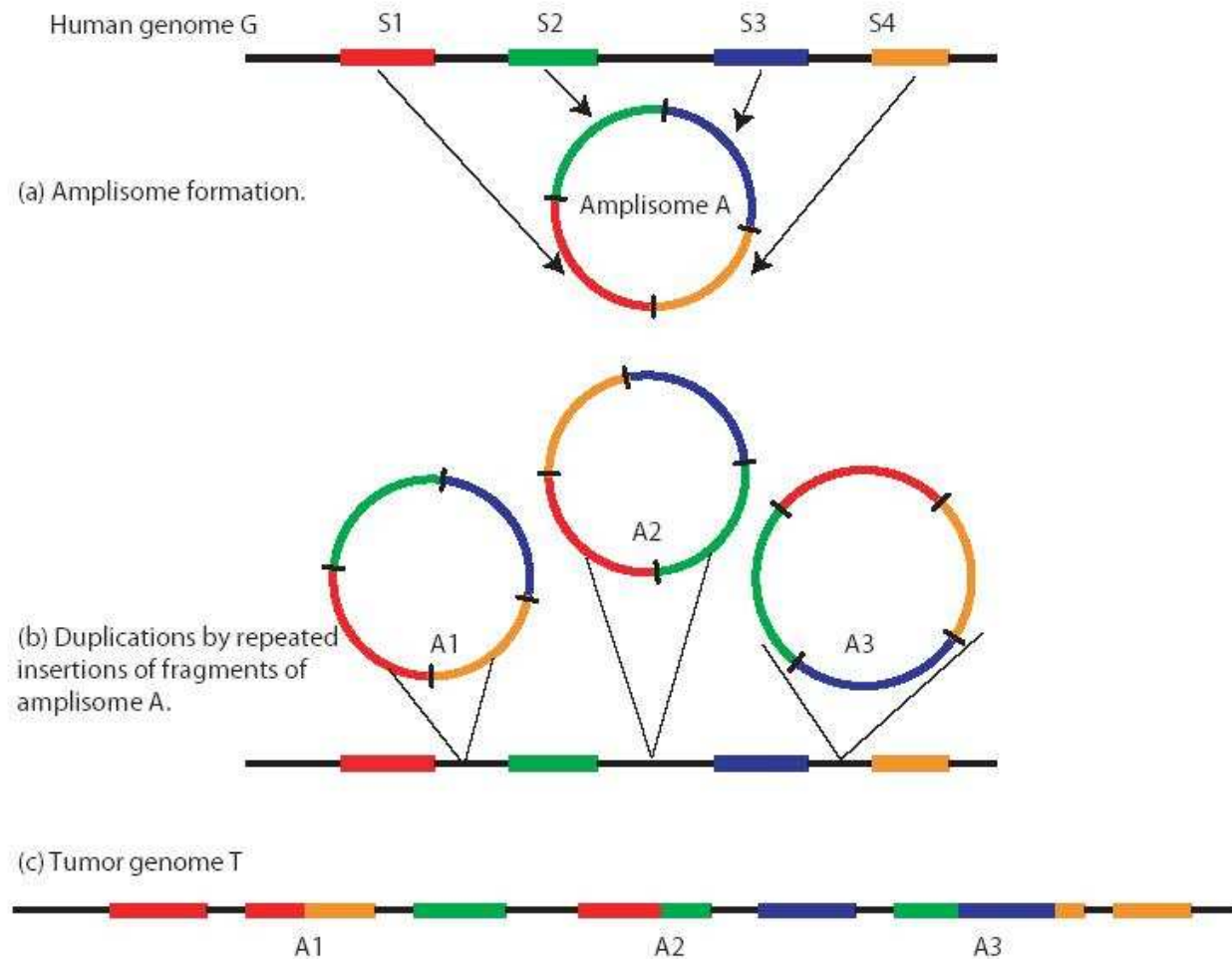


Hibás értelmezés



Problémák II.

Ampliszóma duplikációt okozhat



(Részleges) megoldások I.-II.

BIOINFORMATICS

Vol. 19 Suppl. 2 2003, pages ii162–ii171
DOI: 10.1093/bioinformatics/btg1074



Reconstructing tumor genome architectures

Benjamin J. Raphael^{1,}, Stanislav Volik², Colin Collins² and Pavel A. Pevzner¹*

¹*Department of Computer Science and Engineering, University of California San Diego, La Jolla, CA 92093-0114, USA and* ²*Cancer Research Institute, University of California Comprehensive Cancer Center, San Francisco, CA 94115, USA*

Received on March 17, 2003; accepted on June 9, 2003

BIOINFORMATICS

Vol. 20 Suppl. 1 2004, pages i265–i273
DOI: 10.1093/bioinformatics/bth931



Reconstructing tumor amplisomes

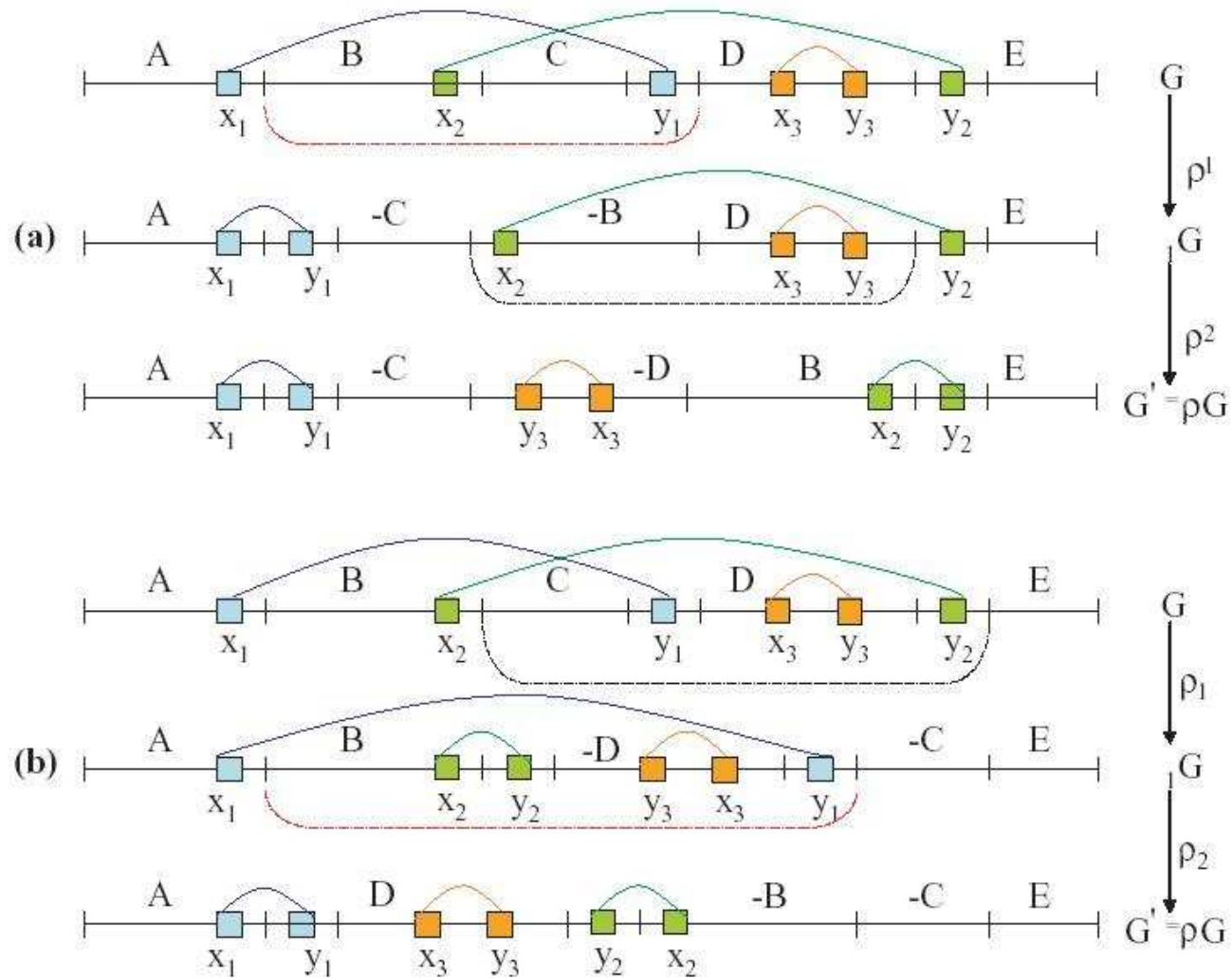
Benjamin J. Raphael and Pavel A. Pevzner*

Department of Computer Science and Engineering, University of California, San Diego, La Jolla, CA 92093-0114, USA

Received on January 15, 2004; accepted on March 1, 2004

Problémák III.

Heurisztikus, egyetlen optimális megoldást kereső eljárások



Problémák IV.

Csak inverziók és reciprok transzlokációk megengedettek

Bár van közvetett bizonyíték reciprok transzlokációra, semmi nem zárja ki azt, hogy történnek transzpozíciók, transzlokációk

Az indok nagyon prózai

Csak inverziók és reciprok transzlokációk vizsgálatára léteznek polinomiális idejű optimalizáló algoritmusok

Biológiailag relevánsabb modellek szükségesek

Megoldások

Bayes-statisztikai keretmunka, Markov lánc Monte Carlo mintavételezés

BIOINFORMATICS APPLICATIONS NOTE Vol. 21 no. 6 2005, pages 817–820
doi:10.1093/bioinformatics/bti060

Genome analysis

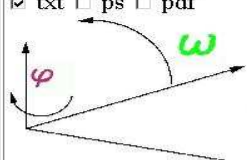
ParIS Genome Rearrangement server

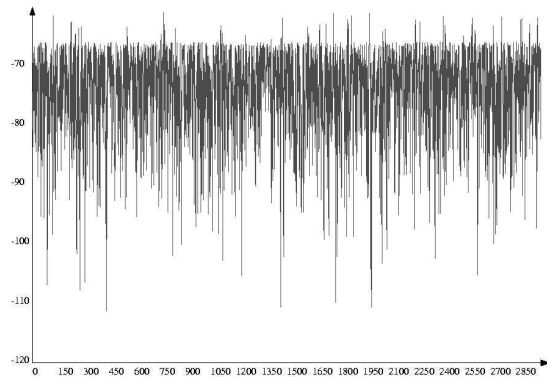
I. Miklós^{1,*}, P. Ittész² and J. Hein³

¹Theoretical Biology and Ecology Modelling Group, Hungarian Academy of Sciences and Eötvös Loránd University, Pázmány Péter Sétány 1/c. H-1117 Budapest, Hungary, ²Collegium Budapest, Institute for Advanced Study, Szentháromság u. 2. H-1014 Budapest, Hungary and ³Genome Analysis and Bioinformatics Group, Oxford Centre for Gene Function, Department of Statistics, Oxford University, 1 South Parks Road, Oxford OX1 3TG, UK

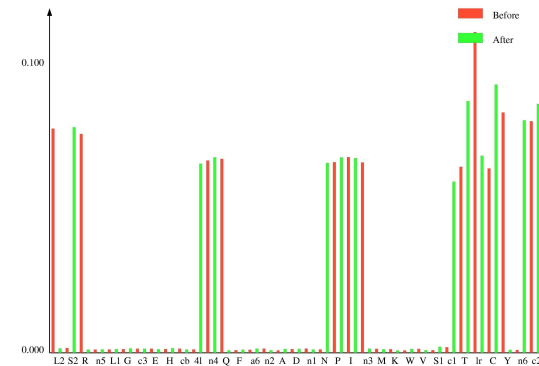
Received on May 25, 2004; revised on June 29, 2004; accepted on July 7, 2004

Advance Access publication October 12, 2004

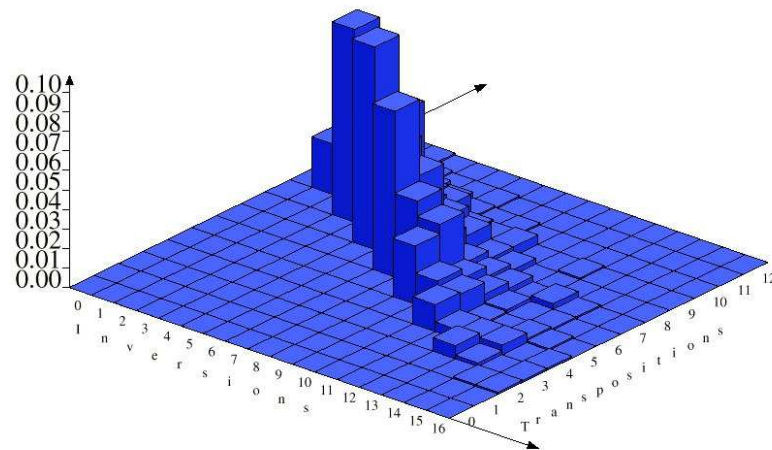
Process name	<input type="text" value="1"/>
Log file	<input checked="" type="checkbox"/> txt <input type="checkbox"/> ps <input type="checkbox"/> pdf
Trajectory file	<input type="checkbox"/> txt
Posterior	<input checked="" type="checkbox"/> txt <input type="checkbox"/> ps <input type="checkbox"/> pdf  <p>phi (between 10° and 80°, default: 40°) <input type="text"/></p> <p>omega (between 20° and 80°, default: 45°) <input type="text"/></p>
Length distribution of mutations	<input type="checkbox"/> txt <input type="checkbox"/> ps <input type="checkbox"/> pdf
Breakpoint	<input type="checkbox"/> txt <input type="checkbox"/> ps <input type="checkbox"/> pdf
Burn-in (Default value: 10000)	<input type="text" value="10000"/>
Number of samples (Default value: 1000)	<input type="text" value="1000"/>
Number of iterations (Default value: 1000)	<input type="text" value="1000"/>
Random seed (Default value: 0)	<input type="text" value="0"/>
Circular genome?	yes <input type="button" value="v"/>
Genome 1	<input type="text"/> <input type="button" value="Browse..."/>
Genome 2	<input type="text"/> <input type="button" value="Browse..."/>
E-mail	<input type="text"/>



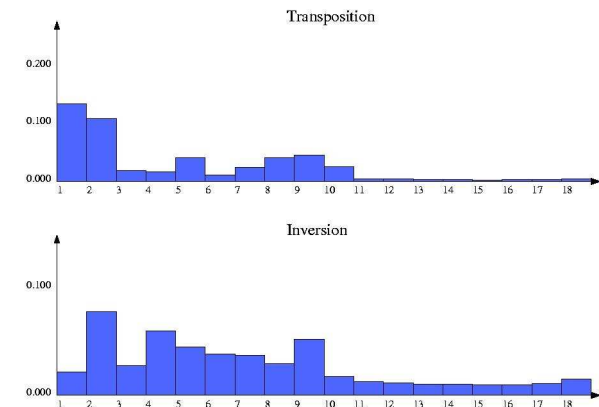
Log likelihood trace



Expected number of breakpoints



Posterior distribution of mutations



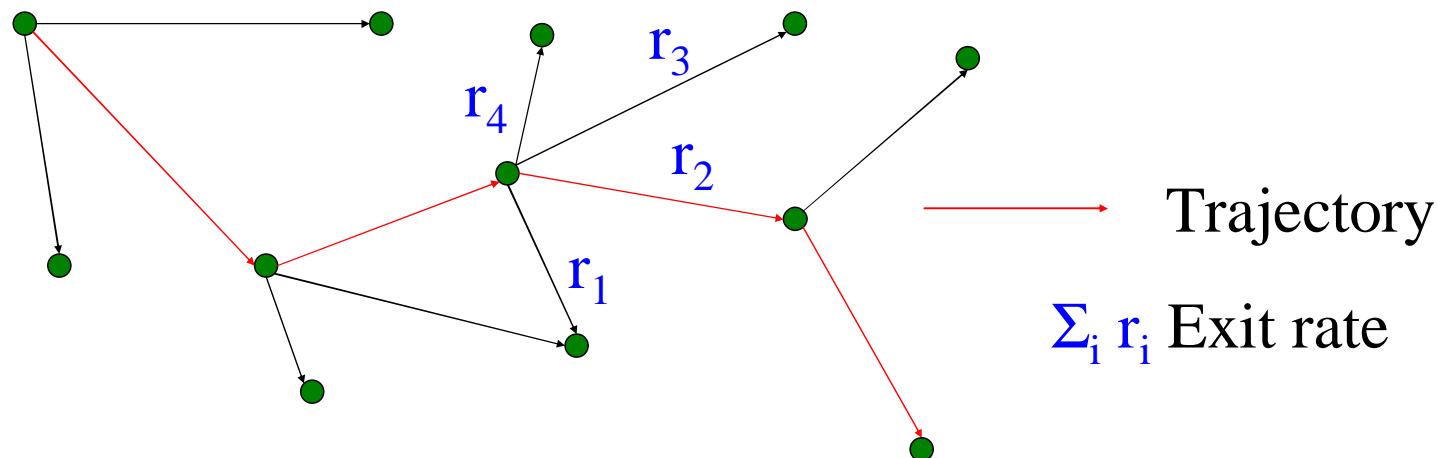
Length distribution

Miklós, Ittész, Hein (2005) Paris Genome Rearrangement server. *Bioinformatics*, 21:817 – 820.

<http://www.stats.ox.ac.uk/~miklos/>

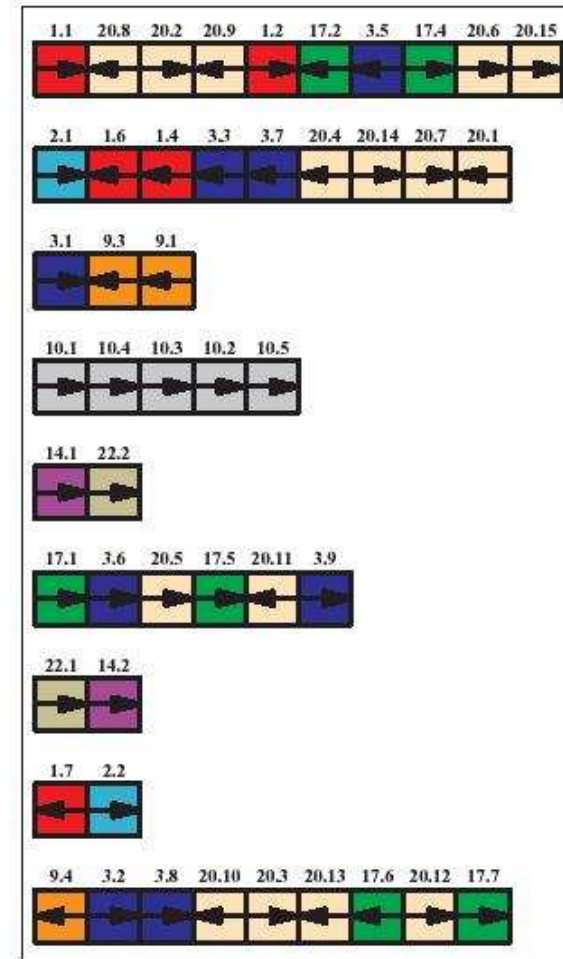
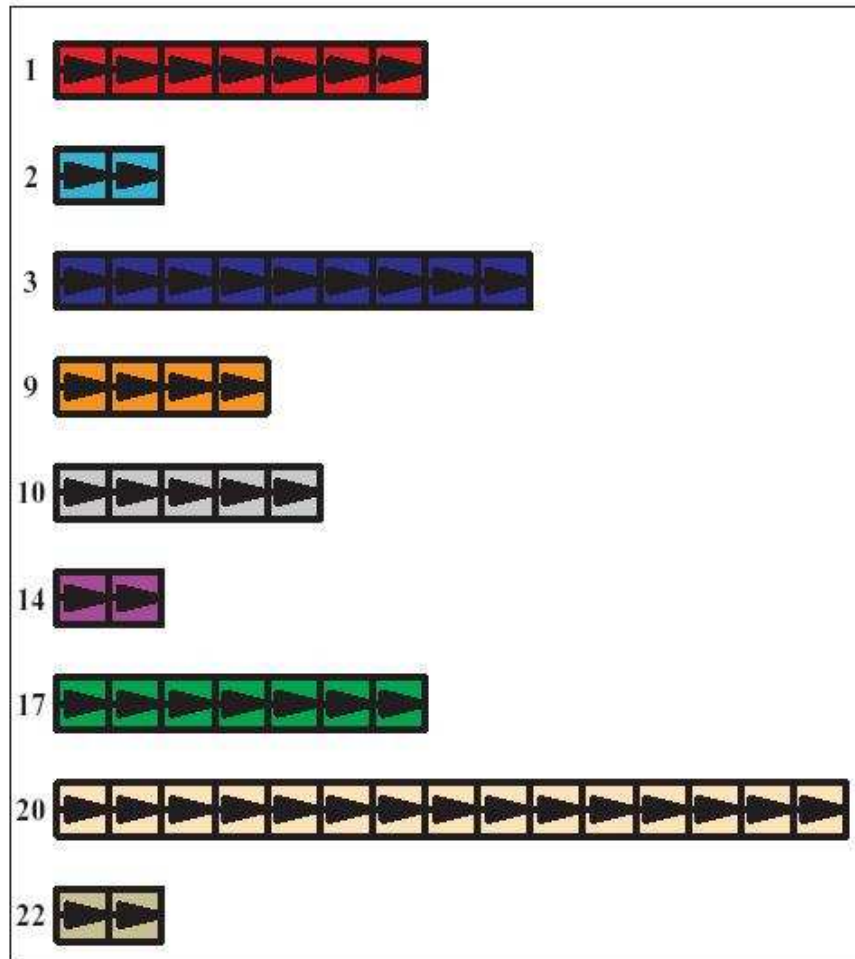
Továbbfejlesztések

- Evolúciós vizsgálat több genomra
- Multikromoszómális genomok vizsgálata
 - Durret, Nielsen, York (2004) *Genetics* 160:621-629.
- Duplikációk, törlések
 - Cselebb adatábrázolással, in prep.
- Mutációs ráták hosszfüggése, egyéb kényszerfeltételek
 - „Trajektória likelihood” számolása
 - Miklós, Lunter, Holmes (2004) *Mol. Biol. Evol.* 21(3):529-540.



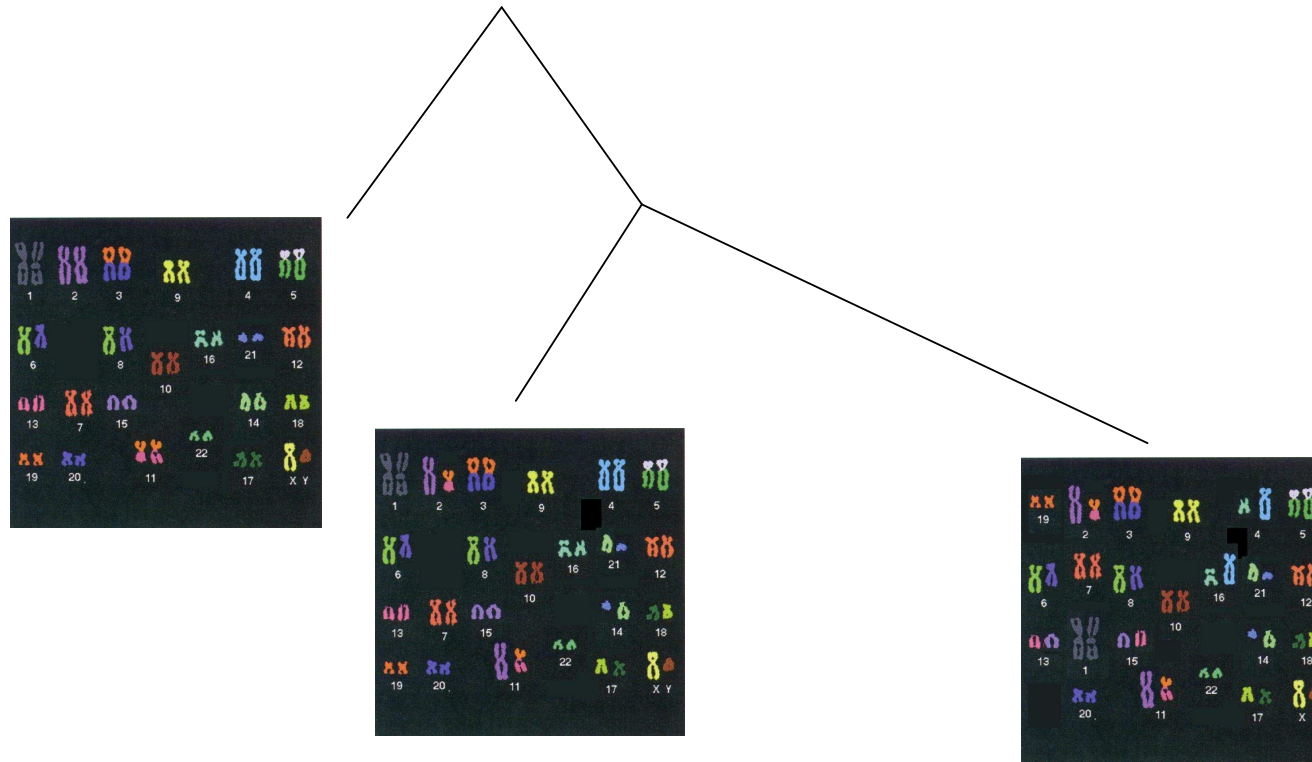
Kérdések

P érték?



Geneológia

ESP a tumor számos fázisában és áttételében



Mi az ősi genom, amely a rákosodási folyamatot kiváltotta?

- Gyógyszer
- Prevenció

Adatbázis

Jelenleg Hind III-as emésztéssel számos ESP-t csinálnak

- Annotáció adatbázisokban?
 - Az ESP-vel nagyon effektíven lehet információt megadni
 - Azonban nincs szabvány ábrázolás
- Rekonstrukciós hibák?
 - A duplikációk és a chimeric BAC-ok miatti hibák
- Más enzimekkel való parciális emésztés?
 - Hogyan hasonlítsunk össze két ESP-t különböző enzimekkel történt emésztésekkel?

Konklúziók

- Genomika: Teljes genom vizsgálat.
 - Paradigmaváltás
 - Új módszerek
- Számos megoldatlan bioinformatikai/informatikai/matematikai probléma
 - Melyek az „érdekes” mintázatok?
 - Geneológia
- Biztató eredmények
 - 2 *Bioinformatics*, 1 *Lecture Notes in Bioinformatics* cikk
 - Webszerver
 - Folyamatban levő munkák
- Hiányzik:
 - Munkaerő (programozó szakdolgozók)
 - Pénz
 - Laboratóriumi kapcsolat

Magamról

- Biológia, kémia és matematika szakos végzettség
- Másfél éves posztdoktori munka az oxfordi egyetemen
- 17 nemzetközi publikáció és két könyvfejezet
 - Genomátrendezés
 - Szekvenciaillesztés és filogenetikai vizsgálat
 - RNS térszerkezetek
 - Modern statisztikai módszerek és MCMC ökológiai adatfeldolgozásban
- Posztdoktori ösztöndíj 2007 januárjáig
- <http://ramet.elte.hu/~miklosi/>