Are there Fragile Regions in the Human Genome?

Combinatorial Algorithms

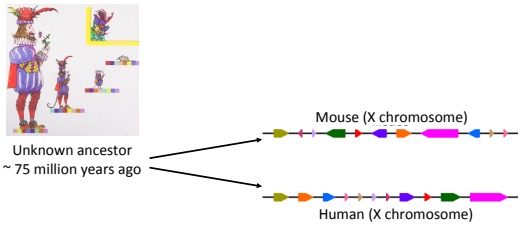
Phillip Compeau and Pavel Pevzner

Bioinformatics Algorithms: an Active Learning Approach

Are There Fragile Regions in the Human Genome?

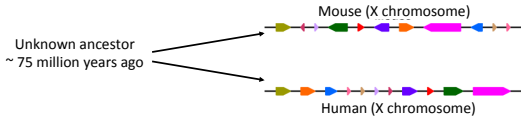
- **Transforming Men into Mice**
- Sorting by Reversals
- Breakpoint Theorem
- Rearrangements in Tumor Genomes
- 2-Break Distance Problem
- Breakpoint Graphs
- 2-Break Distance Theorem
- Rearrangement Hotspots in the Human Genome
- Synteny Block Construction

Genome Rearrangements



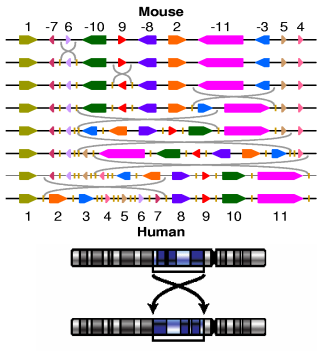
- What are the similarity blocks and how to find them?
- What is the evolutionary scenario for transforming one genome into the other?

Genome rearrangements

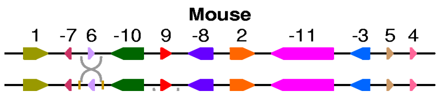


- What are the similarity blocks and how to find them?
- What is the evolutionary scenario for transforming one genome into the other?

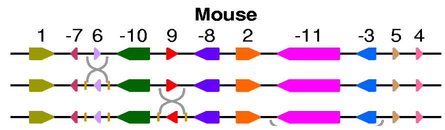
Transforming Mice into Men



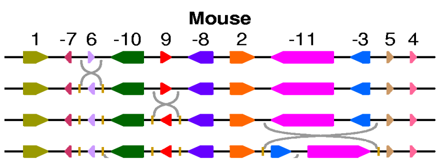
Series of Reversals



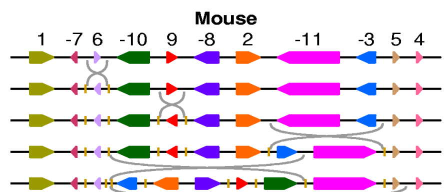
Series of Reversals



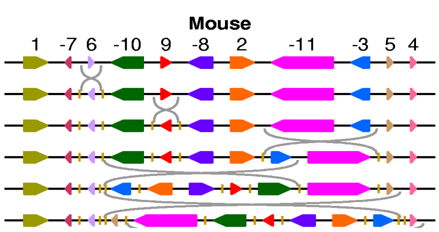
Series of Reversals



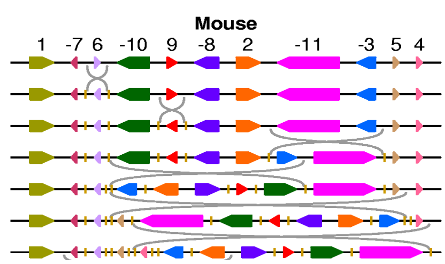
Series of Reversals



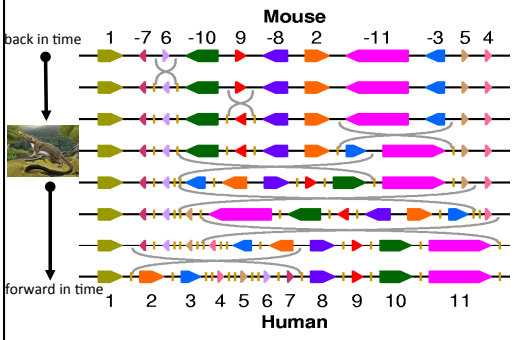
Series of Reversals



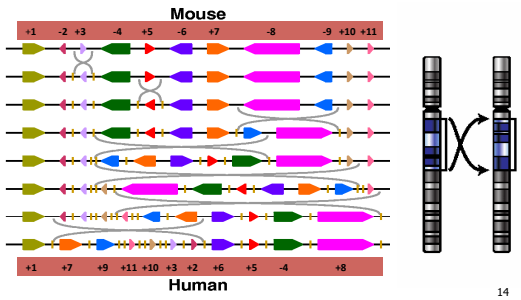
Series of Reversals



Human-Mouse Transformation Accomplished!

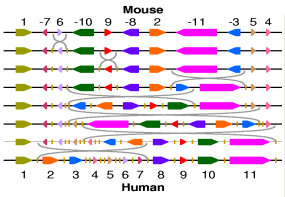


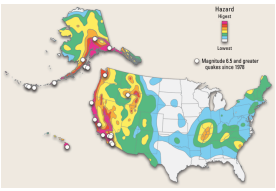
Transforming mice into men (X chromosome)



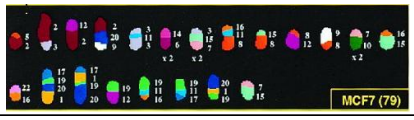
14

Evolutionary Earthquakes

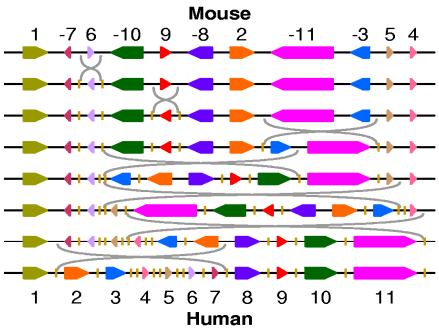




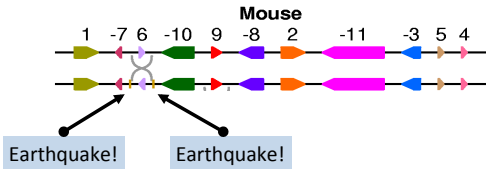
Are there any rearrangement hotspots in mammalian genomes?



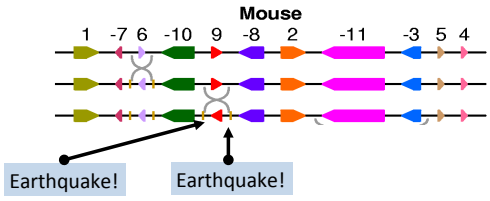
Are There Any Rearrangement Hotspots in This Scenario?



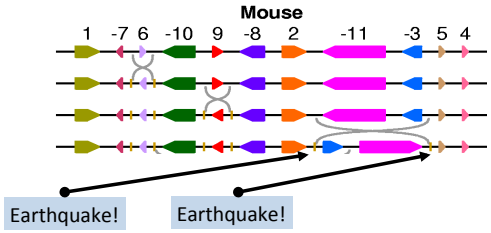
Are There Any Rearrangement Hotspots in This Scenario?



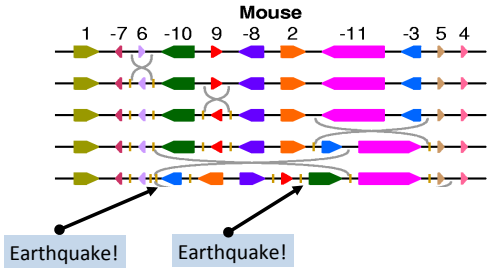
Are There Any Rearrangement Hotspots in This Scenario?



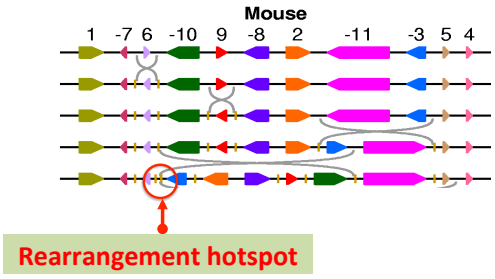
Are There Any Rearrangement Hotspots in This Scenario?



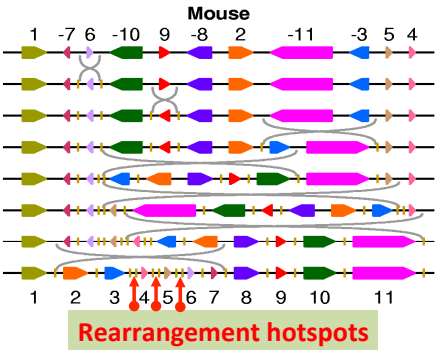
Are There Any Rearrangement Hotspots in This Scenario?



A Rearrangement Hotspots Found!



More Rearrangement Hotspots



Random Breakage Model of Chromosome Evolution

- *Susumu Ohno, 1973*: Rearrangements occur at random positions (no fragile regions).
- *Nadeau & Taylor, 1984:* the first statistical arguments in favor of the Random Breakage Model (RBM).

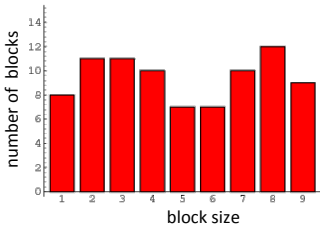
Does RBM Have Predictive Power?

Apply *N* **random reversals** to a chromosome consisting of *M* "genes."
Can we predict how many blocks of length *k* will be generated?

If we can predict it, will these predictions fit what we observe in real genomes?

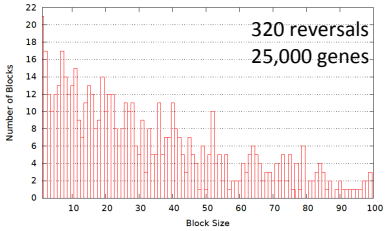
Does RBM Have Predictive Power?

Apply *N* **random reversals** to a chromosome consisting of *M* "genes."
Can we predict how many blocks of length *k* will be generated?



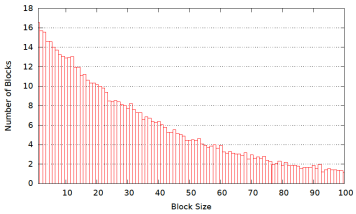
Does RBM Have Predictive Power?

Apply *N* **random reversals** to a chromosome consisting of *M* "genes."
Can we predict how many blocks of length *k* will be generated?



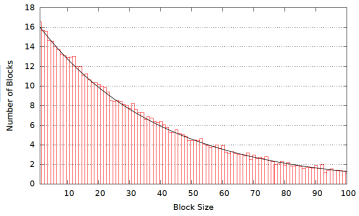
Despite the fact that reversals occur at random positions, we can predict (roughly) how many blocks of each length will be generated!

Histogram of Synteny Block Lengths (averaged over 100 simulations)



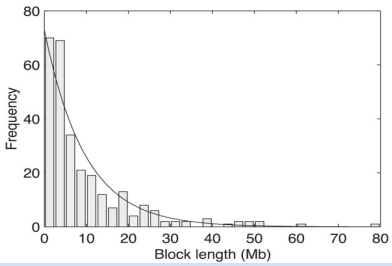
Exponential Distribution (simulated data)

$y=\lambda e^{-\lambda \cdot x}$



Lengths of Real Human-Mouse Synteny Blocks

$y=\lambda e^{-\lambda \cdot x}$

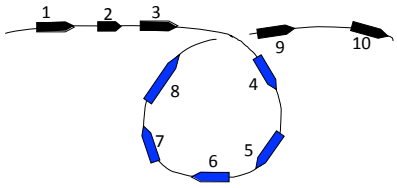


1990s: RBM was embraced by biologists and has become *de facto* theory of chromosome evolution.

Are There Fragile Regions in the Human Genome?

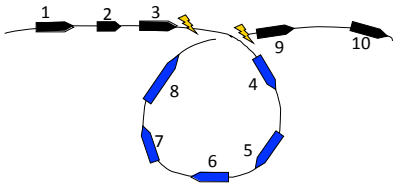
- Transforming Men into Mice
- **Sorting by Reversals**
- Breakpoint Theorem
- Rearrangements in Tumor Genomes
- 2-Breaks
- Breakpoint Graphs
- 2-Break Distance Theorem
- Rearrangement Hotspots in the Human Genome
- Synteny Block Construction

Reversals



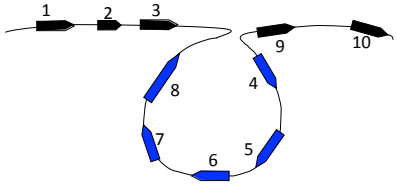
+1 +2 +3 +4 +5 +6 +7 +8 +9 +10

Reversals



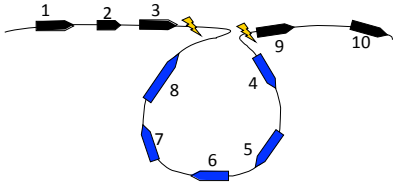
+1 +2 +3 +4 +5 +6 +7 +8 +9 +10

Reversals



+1 +2 +3 -8 -7 -6 -5 -4 +9 +10

Reversals



+1 +2 +3 -8 -7 -6 -5 -4 +9 +10

The reversal introduced two *breakpoints* (disruptions in gene order).

Rearrangement Scenario with 5 Reversals

Step 0:	2	-4	-3	5	-8	-7	-6	1
Step 1:	2	3	4	5	-8	-7	-6	1
Step 2:	2	3	4	5	6	7	8	1
Step 3:	2	3	4	5	6	7	8	-1
Step 4:	-8	-7	-6	-5	-4	-3	-2	-1
Step 5:	1	2	3	4	5	6	7	8

Rearrangement Scenario with **4** Reversals

Step 0:	2	-4	-3	5	-8	-7	-6	1
Step 1:	2	3	4	5	-8	-7	-6	1
Step 2:	-5	-4	-3	-2	-8	-7	-6	1
Step 3:	-5	-4	-3	-2	-1	6	7	8
Step 4:	1	2	3	4	5	6	7	8

Reversal distance: the minimum number of reversals to transform one permutation into another.

Sorting by 4 Reversals

Step 0:	2	-4	-3	5	-8	-7	-6	1
Step 1:	2	3	4	5	-8	-7	-6	1
Step 2:	-5	-4	-3	-2	-8	-7	-6	1
Step 3:	-5	-4	-3	-2	-1	6	7	8
Step 4:	1	2	3	4	5	6	7	8

Sorting by Reversals Problem: Calculate the reversal distance between a permutation and the identity permutation (+1 +2 … +*n*).

- **Input**: A permutation *P*.
- **Output**: The reversal distance between *P* and the identity permutation.

Greedy Sorting by Reversals

(**+1** −7 +6 −10 +9 −8 **+2** −11 −3 +5 +4)

Greedy Sorting by Reversals

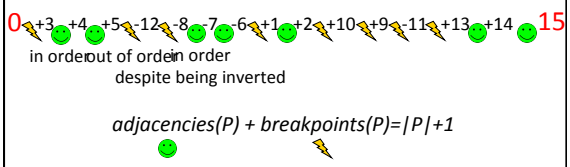
(+1 **−7 +6 −10 +9 −8 +2** −11 −3 +5 +4)
(+1 **−2** +8 −9 +10 −6 +7 −11 −3 +5 +4)
(+1 +2 **+8 −9 +10 −6 +7 −11 −3** +5 +4)
(+1 +2 +3 **+11 −7 +6 −10 +9 −8 +5 +4**)
(+1 +2 +3 **−4** −5 +8 −9 +10 −6 +7 −11)
(+1 +2 +3 +4 **−5** +8 −9 +10 −6 +7 −11)
(+1 +2 +3 +4 +5 **+8 −9 +10 −6** +7 −11)
(+1 +2 +3 +4 +5 +6 **−10 +9 −8 +7** −11)
(+1 +2 +3 +4 +5 +6 **−7** +8 −9 +10 −11)
(+1 +2 +3 +4 +5 +6 +7 +8 **−9** +10 −11)
(+1 +2 +3 +4 +5 +6 +7 +8 +9 +10 **−11**)

(+1 +2 +3 +4 +5 +6 +7 +8 +9 +10 +11)

Are there Fragile Regions in the Human Genome?

- Transforming Men into Mice
- Sorting by Reversals
- **Breakpoint Theorem**
- Rearrangements in Tumor Genomes
- 2-Breaks
- Breakpoint Graphs
- 2-Break Distance Theorem
- Rearrangement Hotspots in the Human Genome
- Synteny Block Construction

Adjacencies and Breakpoints



adjacencies(P) + *breakpoints(P)*=|*P*|+1

What is the number of breakpoints in the identity permutation (+1 +2 … +*n*)?

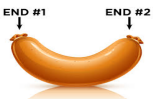
Sorting by Reversals as Breakpoint Elimination

									breakpoints(P)
Step 0:	2	-4	-3	5	-8	-7	-6	1	**6**
Step 1:	2	3	4	5	-8	-7	-6	1	**4**
Step 2:	-5	-4	-3	-2	-8	-7	-6	1	**4**
Step 3:	-5	-4	-3	-2	-1	6	7	8	**2**
Step 4:	1	2	3	4	5	6	7	8	**0**

How many breakpoints can be eliminated by a single reversal?

Breakpoint Theorem ≥ *breakpoints(P)*/**2**

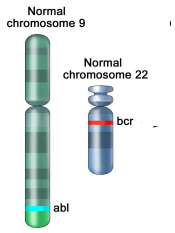
Hint:



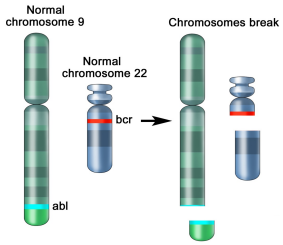
Are There Fragile Regions in the Human Genome?

- Transforming Men into Mice
- Sorting by Reversals
- Breakpoint Theorem
- **Rearrangements in Tumor Genomes**
- 2-Breaks
- Breakpoint Graphs
- 2-Break Distance Theorem
- Rearrangement Hotspots in the Human Genome
- Synteny Block Construction

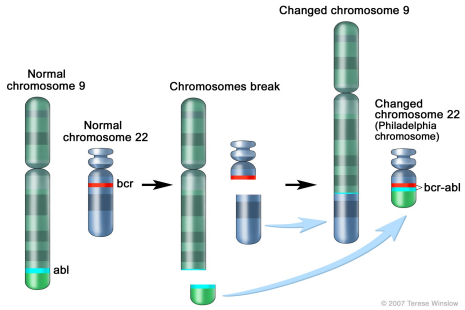
Philadelphia Chromosome



Philadelphia Chromosome



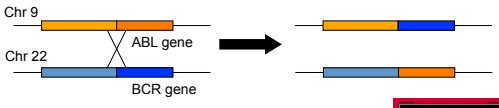
Philadelphia Chromosome



Oncogenic **bcr-abl** fusion gene:
biomarker for chronic myeloid leukemia (CML)

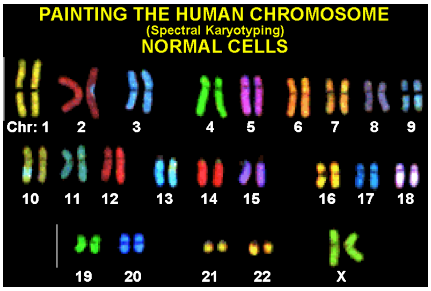
Translocations

Translocation in CML yields Philadelphia chromosome:



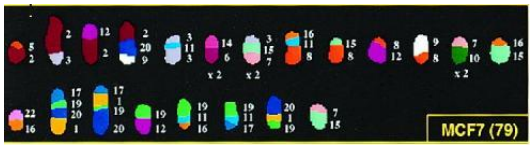
Thousands of rearrangements known for different tumors.

Painting Chromosomes (Normal Cells)



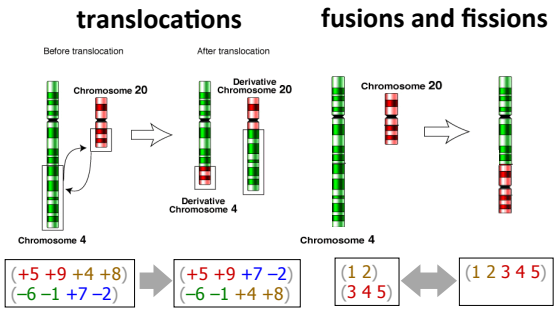
Painting Chromosomes (Tumor Cells)

- MCF7 is human breast cancer cell line:



What sequence of rearrangements has produced MCF7?

Rearrangements in Multichromosomal Genomes



Are There Fragile Regions in the Human Genome?

- Transforming Men into Mice
- Sorting by Reversals
- Breakpoint Theorem
- Rearrangements in Tumor Genomes
- **2-Break Distance Problem**
- Breakpoint Graphs
- 2-Break Distance Theorem
- Rearrangement Hotspots in the Human Genome
- Synteny Block Construction

From Linear to Circular Chromosomes

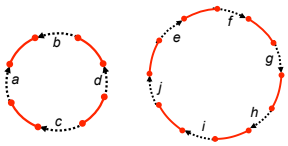
$(+a\ -b\ -d\ +c)$ $(+e\ +f\ +g\ +h\ +i\ +j)$

From Linear to Circular Chromosomes

$(+a\ -b\ -d\ +c)$ $(+e\ +f\ +g\ +h\ +i\ +j)$

Black directed edges represent synteny blocks.

From Linear to Circular Chromosomes



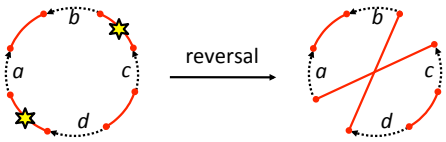
$(+a\ -b\ -d\ +c)$ $(+e\ +f\ +g\ +h\ +i\ +j)$

Black directed edges represent synteny blocks.
Red undirected edges connect adjacent synteny blocks.

Equivalent Representations of a Circular Genome

Q=(+a-b-d+c) *Q=(+a-b-d+c)*

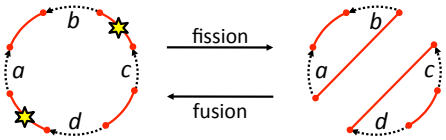
Reversals on Circular Chromosomes



P=(+a-b-c+d) *Q=(+a-b-d+c)*

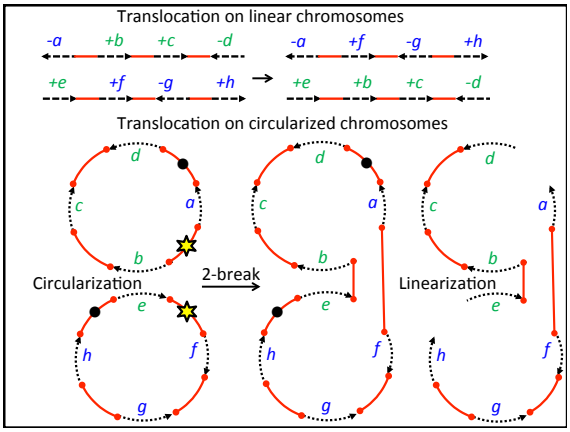
A reversal deletes two red edges and replaces them by two other red edges (on the same 4 nodes)

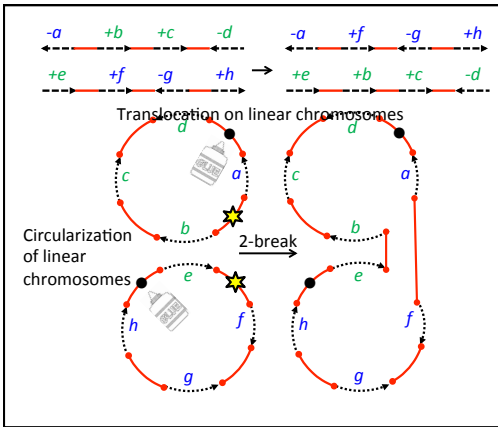
Fusion/Fission as Replacement of Two Red Edges

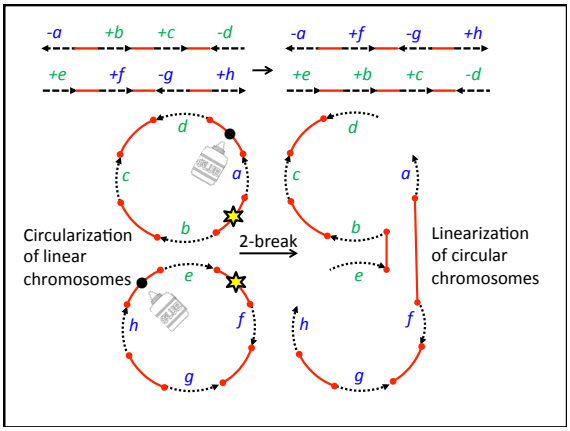


P=(+a-b-c+d) *Q= (+a-b) (-c+d)*

2-break: a rearrangement that replaces two red edges with two new red edges on the same 4 nodes.







2-Break Distance

2-Break distance *d(P,Q)*:
minimum number of 2-breaks transforming genome *P* into genome *Q*

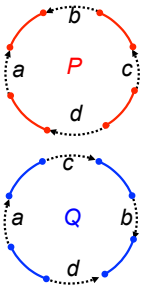
2-Break Distance Problem. Find the 2-break distance between two genomes.

- **Input.** Two genomes on the same set of synteny blocks.
- **Output.** The 2-break distance between these genomes.

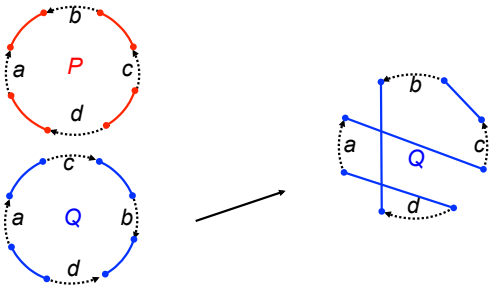
Are There Fragile Regions in the Human Genome?

- Transforming Men into Mice
- Sorting by Reversals
- Breakpoint Theorem
- Rearrangements in Tumor Genomes
- 2-Break Distance Problem
- **Breakpoint Graphs**
- 2-Break Distance Theorem
- Rearrangement Hotspots in the Human Genome
- Synteny Block Construction

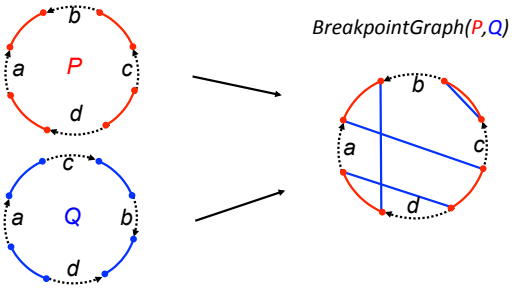
Comparing Genomes *P* and *Q*



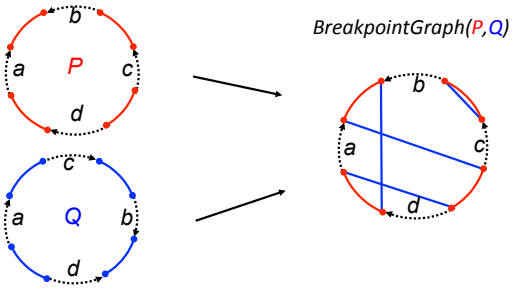
Different Drawing of *Q*



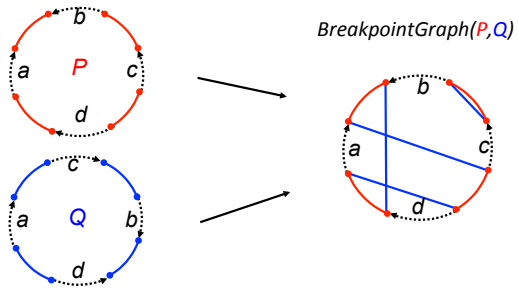
Superimposing *P* and *Q*



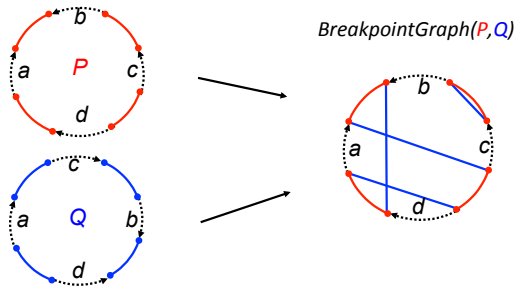
Red and Black Edges in Breakpoint Graph Form..



Blue and Black Edges in Breakpoint Graph Form…

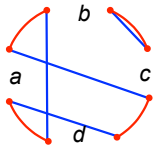


What About Red and Blue Edges???



Alternating Red-Blue Cycles

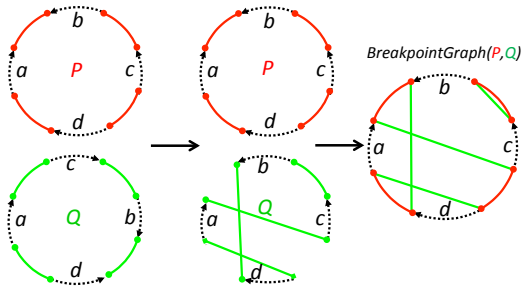
BreakpointGraph(P,Q)

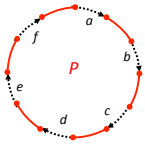


Red and blue edges form **alternating red-blue cycles**.

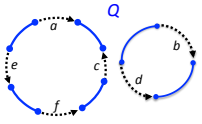
cycle(P,Q): number of red-blue alternating cycles

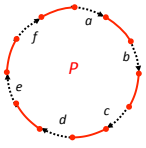
Breakpoint Graph



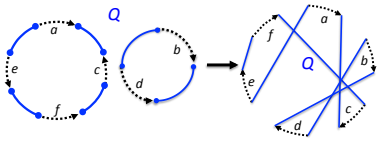


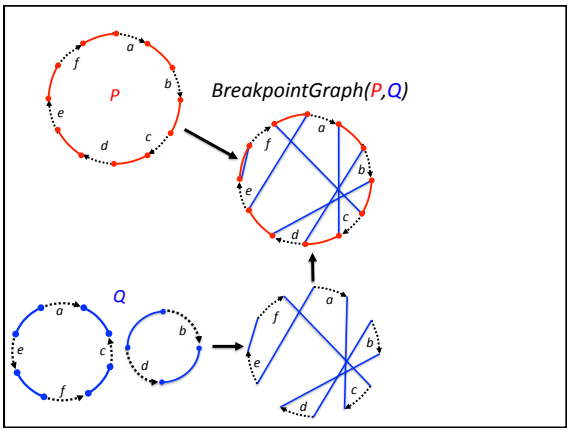
What is *cycle(P,Q)*?

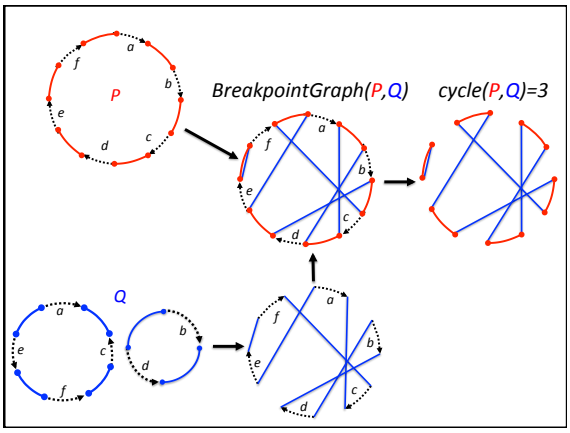




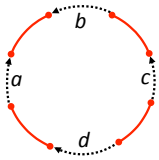
Arranging blacks edges of *Q*
in the same order as black edges in *P*







Given *P*, What Genome *Q* Maximizes *cycle(P,Q)*?

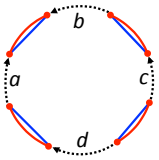


P=(+a-b-c+d)

Given *P*, What Genome *Q* Maximizes *cycle(P,Q)*?

BreakpointGraph(P,P)

Identity Breakpoint Graph



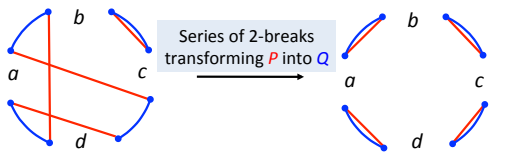
cycle(P,P)=#blocks=4

P=(+a-b-c+d)

Genome Rearrangements Affect Red-Blue Cycles

Each transformation *P* → *Q* corresponds to a transformation:

BreakpointGraph(P,Q) *BreakpointGraph(Q,Q)*



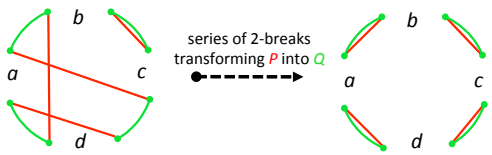
cycle(P,Q)=2 → *cycle(Q,Q)= 4*

Genome Rearrangements Affect Red-Blue Cycles

Each transformation *P* → *Q*

corresponds to a transformation

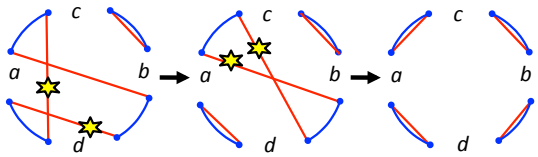
BreakpointGraph(P,Q) → *BreakpointGraph(Q,Q)*



cycle(P,Q)=2 *cycle(Q,Q)= 4*

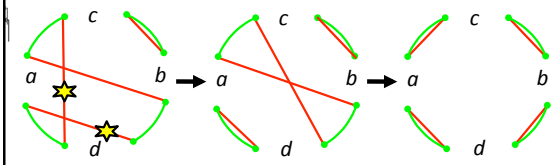
Rearrangements Change *cycle(P,Q)*

P=(+a –b –c **+d**) → *P'*=(+a **–b –c** -d) → *P''*=*Q*=(+a +c +b -d)
BreakpointGraph(P,Q) → *BreakpointGraph(P',Q)* → *BreakpointGraph(Q,Q)*
cycle(P,Q)=2 → *cycle(P',Q)*=3 → *cycle(Q,Q)*=4



Rearrangements Change *cycle(P,Q)*

P=(+a –b –c **+d**) → *P'*=(+a **–b –c** -d) → *P''*=*Q*=(+a +c +b -d)
BreakpointGraph(P,Q) → *BreakpointGraph(P',Q)* → *BreakpointGraph(Q,Q)*
cycle(P,Q)=2 → *cycle(P',Q)*=3 → *cycle(Q,Q)*=4



Are There Fragile Regions in the Human Genome?

- Transforming Men into Mice
- Sorting by Reversals
- Breakpoint Theorem
- Rearrangements in Tumor Genomes
- 2-Break Distance Problem
- Breakpoint Graphs
- **2-Break Distance Theorem**
- Rearrangement Hotspots in the Human Genome
- Synteny Block Construction

Sorting by 2-Breaks

2-breaks
P → ... → *Q*

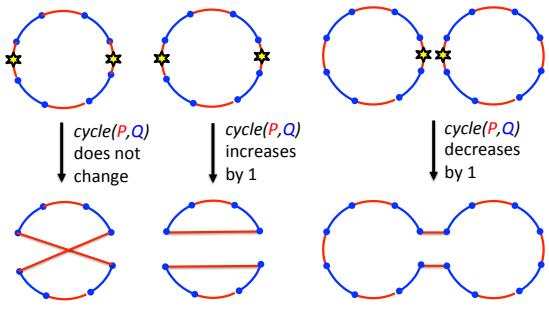
BreakpointGraph(P,Q) → ... → *BreakpointGraph(Q,Q)*

cycle(P,Q) → ... → *cycle(Q,Q)*=*blocks(Q,Q)*

# of red-blue cycles increases by ***blocks(P,Q) - cycle(P,Q)***

How much each 2-break can contribute to this increase?

A 2-Break May Change *cycle(P,Q)* by …



Each 2-Break Increases #Cycles by at Most 1

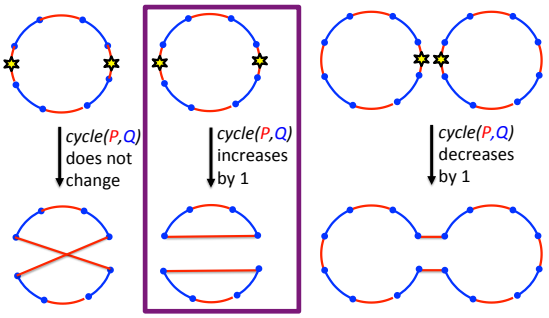
A 2-break:

adds 2 new red edges and thus creates at most **2** new cycles (containing two new red edges)

removes 2 red edges and thus destroys at least **1** old cycle (containing two old edges)

change in the number of cycles ≤ **2-1=1**.

There Exists a 2-Break Increasing *cycle(P,Q)* by 1



2-Break Distance Theorem

- A 2-break increases #cycles by at most 1.
- There exists a 2-break increasing #cycles by 1.
- Every sorting by 2-breaks must increase #cycles by *blocks(P,Q) - cycle(P,Q)*
- 2-break distance between genomes *P* and *Q*:

d(P,Q) = blocks(P,Q) - cycle(P,Q)

2-Break Distance between Human and Mouse Genomes

- *Human* and *Mouse* genomes can be decomposed into **280** synteny blocks (at least 0.5 million nucleotides in length)
- The breakpoint graph on these blocks has **35** cycles
- The 2-break distance between *Human* and *Mouse:*

d(H,M) = blocks(H,M)-cycle(H,M) = **280-35=245**

- There are numerous **245**-step scenarios.
- The true scenario may have more than **245** steps.

Are There Fragile Regions in the Human Genome?

- Transforming Men into Mice
- Sorting by Reversals
- Breakpoint Theorem
- Rearrangements in Tumor Genomes
- 2-Break Distance Problem
- Breakpoint Graphs
- 2-Break Distance Theorem
- **Rearrangement Hotspots in the Human Genome**
- Synteny Block Construction

Are There Fragile Regions in the Human Genome?

Are There Fragile Regions in the Human Genome?

Rearrangement Hotspots Theorem: Yes!

Are There Fragile Regions in the Human Genome?

Rearrangement Hotspots Theorem: Yes!

Proof: *If the Random Breakage Model is correct*, then *N* rearrangements applied to circular chromosomes will produce approximately *2N* synteny blocks.

- Since there are 280 human-mouse synteny blocks, there must have been approximately 280/2 = **140** 2-breaks on the human-mouse evolutionary path.
- However, the 2-Break Distance Theorem implies that there are at least **245** 2-breaks on this path.

STOP and Think. Is **245 ≈ 140**?

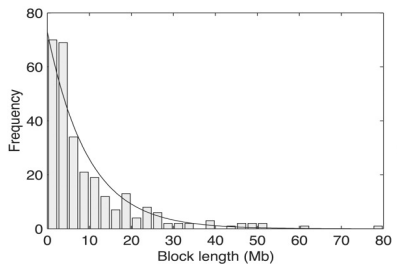
A Contradiction!

Since 245 is much larger than 140, we arrived at a contradiction implying that **one of our assumptions is incorrect! Which one?**

But the only assumption we made in this proof was:

"If the Random Breakage Model is correct…"

If RBM Is Wrong, How Would You Explain the Exponential Distribution?



Computational Tests vs. Biological Models

- Why have biologists embraced the Random Breakage Model?

Computational Tests vs. Biological Models

- Why have biologists embraced the Random Breakage Model?
  - **A logical fallacy**: RBM is not the only model that complies with the "exponential distribution" test.

Model \ Test	Exponential distribution
RBM	**YES**

Computational Tests vs. Biological Models

- Why have biologists embraced the Random Breakage Model?
  - A logical fallacy: RBM is not the only model that complies with the "exponential distribution" test.
- Why was RBM refuted?
  - It does not comply with the observed "breakpoint reuse."

Model \ Test	Exponential distribution	Breakpoint reuse
RBM	**YES**	**NO**

A New Model Needed

- Why have biologists embraced the Random Breakage Model?
  - A logical fallacy: RBM is not the only model that complies with the "exponential distribution" test.
- Why was RBM refuted?
  - It does not comply with the "breakpoint reuse" observed in genomes.
- **Is there a model that complies with both the "exponential distribution" and the "breakpoint reuse" tests?**

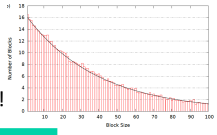
Model \ Test	Exponential distribution	Breakpoint reuse
RBM	YES	NO
???	YES	YES

Fragile Breakage Model

- Genome is a mosaic of:
  - *fragile regions* with high propensity for rearrangements and
  - *solid regions* with low propensity for rearrangements.
- Fragile regions (regions between consecutive synteny blocks) are small, accounting for less than ~ 5% of genome.

Does FBM Explain BOTH Exponential Distribution and Rearrangement Hotspots?

- A small number of short fragile regions explain rearrangement hotspots.
- If the fragile regions are somewhat randomly distributed throughout the genome, the synteny blocks follow the exponential distribution!

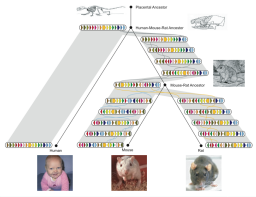


Model \ Test	Exponential distribution	Breakpoint reuse
RBM	YES	NO
FBM	YES	YES

But Is There a Test that **BOTH** RBM and FBM Fail?

Model \ Test	Exponential distribution	Breakpoint reuse	A New Test?
RBM	YES	NO	NO
FBM	YES	YES	NO

Information About Multiple Genomes Enables a New Test



Multiple Breakpoint Reuse Test: analysis of breakpoints across multiple genomes

Model \ Test	Exponential distribution	Breakpoint reuse	A New Test?
RBM	YES	NO	NO
FBM	YES	YES	NO

Birth and Death of Fragile Regions

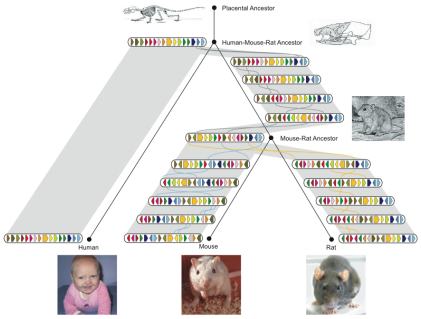
- Recent studies revealed evidence for the "***birth and death***" of the fragile regions, implying that they move to different locations in different lineages.
- This discovery resulted in the ***Turnover Fragile Breakage Model (TFBM)*** that complies with a new Multiple Breakpoint Reuse (**MBR**) Test.
- TFBM points to locations of the ***currently*** fragile regions.

Birth and Death of Fragile Regions

- Recent studies revealed evidence for the "***birth and death***" of the fragile regions, implying that they move to different locations in different lineages.
- This discovery resulted in the ***Turnover Fragile Breakage Model (TFBM)*** that complies with a new Multiple Breakpoint Reuse (**MBR**) Test.
- TFBM points to locations of the ***currently*** fragile regions.

Model \ Test	Exponential distribution	Breakpoint reuse	MBR
RBM	YES	NO	NO
FBM	YES	YES	NO
TFBM	YES	YES	YES

Where Are the Fragile Regions Located? What Causes Fragility?

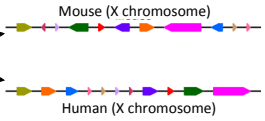


Are There Fragile Regions in the Human Genome?

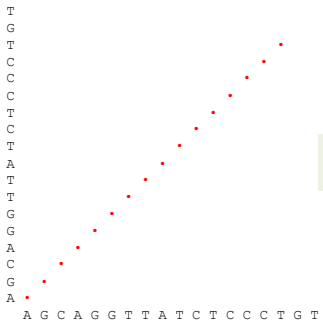
- Transforming Men into Mice
- Sorting by Reversals
- Breakpoint Theorem
- Rearrangements in Tumor Genomes
- 2-Break Distance Problem
- Breakpoint Graphs
- 2-Break Distance Theorem
- Rearrangement Hotspots in the Human Genome
- **Synteny Block Construction**

How Do We Construct the Synteny Blocks?

Unknown ancestor
~ 75 million years ago

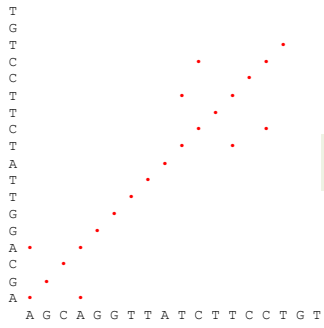


Comparing Genome with Itself (in 2-D)



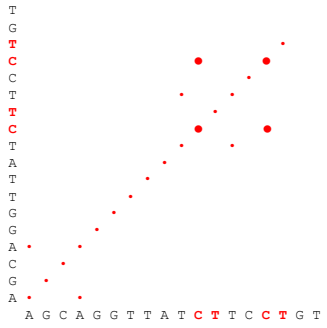
Identical *k*-mers (*k*=3)

From Identical 3-mers to Identical 2-mers

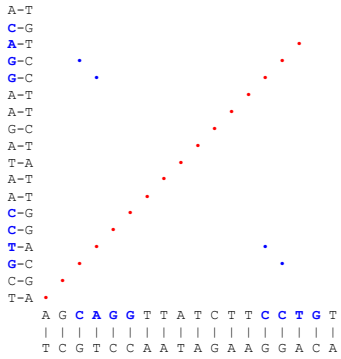


Identical *k*-mers (*k*=**2**)

From Identical 3-mers to Identical 2-mers

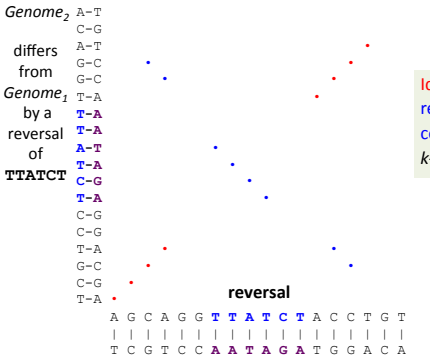


We Forgot about the Complementary Strand



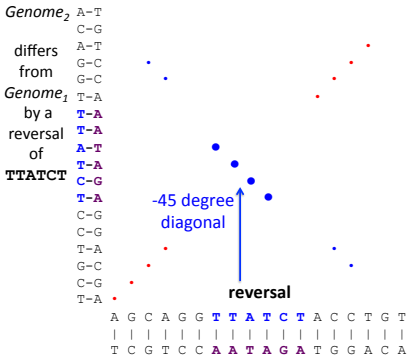
Identical and reverse complementary *k*-mers (*k*=3).

Comparing Different Genomes



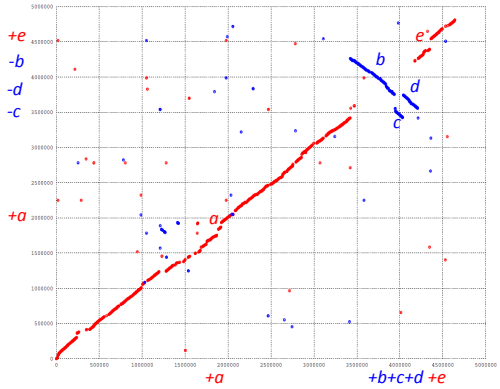
Identical and reverse complementary *k*-mers (*k*=3).

Comparing Different Genomes



Identical and reverse complementary *k*-mers (*k*=3).

E. Coli versus *S. enterica*



Synteny Blocks as Diagonals in Genomic Dot Plot

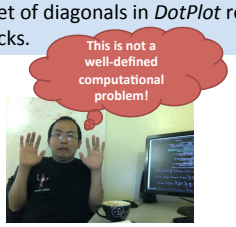
Finding Synteny Blocks Problem. Find diagonals in the genomic dot-plot.

- **Input.** A set of points *DotPlot* in 2-D.
- **Output.** A set of diagonals in *DotPlot* representing synteny blocks.

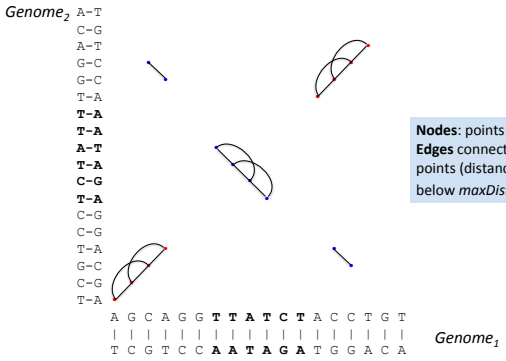
Synteny Blocks as Diagonals in Genomic Dot Plot

Finding Synteny Blocks Problem. Find diagonals in the genomic dot-plot.

- **Input.** A set of points *DotPlot* in 2-D.
- **Output.** A set of diagonals in *DotPlot* representing synteny blocks.



Connecting Closely Located Points in Genomic Dot-Plot



Nodes: points in 2-D
Edges connect close points (distance below *maxDistance*)

Synteny Block Generation Algorithm

Synteny(*DotPlot,maxDistance,minSize*)
maxDistance: gap size
minSize: minimum synteny block size

- Form a graph whose node set is the set of points in *DotPlot*
- Connect two nodes by an edge if the 2-D distance between them is < *maxDistance.* The connected components in the resulting graph define synteny blocks
- Delete small synteny blocks (length < *minSize*)

Another Synteny Block Generation Algorithm

Amalgamate(*DotPlot,maxDistance,minSize*)
maxDistance: gap size
minSize: minimum synteny block size

- Define each point in *DotPlot* as a separate block and iteratively amalgamate the resulting blocks
- Amalgamate two blocks if they contain two points that are separated by < *maxDistance* in another genome.
- Delete small synteny blocks (length < *minSize*)

Two Synteny Block Generation Algorithms: Which One is Better?

Synteny(*DotPlot,maxDistance,minSize*)

- Form a graph whose node set is the set of points in *DotPlot*
- Connect two nodes by an edge if the 2-D distance between them is < *maxDistance.* The connected components in the resulting graph define synteny blocks
- Delete small synteny blocks (length < *minSize*)

Amalgamate(*DotPlot,maxDistance,minSize*)

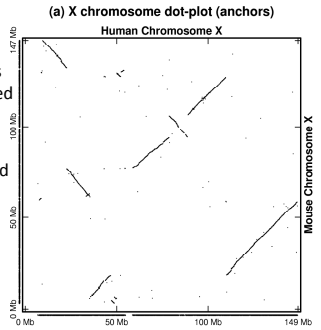
- Define each point in *DotPlot* as a separate block and iteratively amalgamate the resulting blocks
- Amalgamate two blocks if they contain two points that are separated by < *maxDistance* in another genome.
- Delete small synteny blocks (length < *minSize*)

The algorithms look very similar but do they produce similar results?

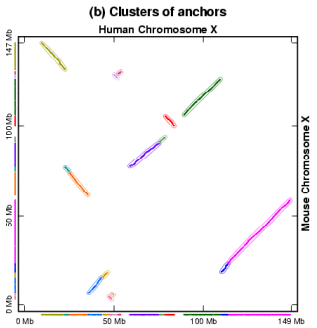
Genomic Dot Plot (Human vs. Mouse X Chromosome)

≈25,000 anchors (regions of shared similarity).

Anchors enlarged for visibility.

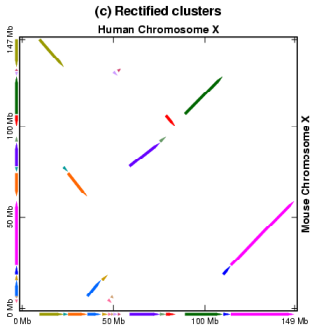


Connected Components in the Synteny Graph



GRIMM-Synteny on X chromosome (c)

From anchors to synteny blocks



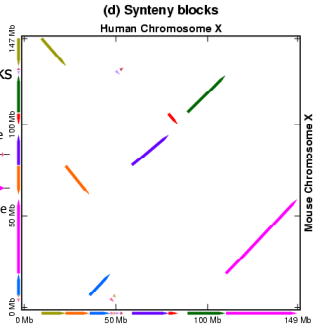
122

Synteny Blocks in 2-D

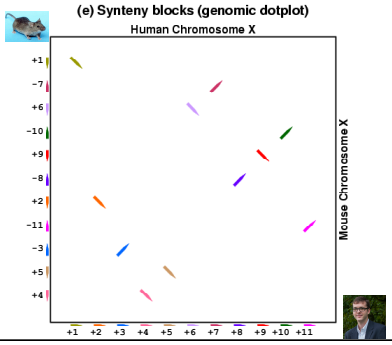
11 synteny blocks

Mouse X chromosome

Human X chromosome



Ignoring the Length of the Synteny Blocks



Constructing Synteny Blocks and Reconstructing Rearrangement History of All Mammals

Science, February, 2013