Genome Rearrangements In Man and Mouse

Abhinav Tiwari
Department of Bioengineering

Genome Rearrangement

- Scrambling of the order of the genome during evolution
- Operations on chromosomes
  - Reversal
  - Translocation
  - Fusion
  - Fission
- The “genomic distance” between multichromosomal genomes is defined as the number of such rearrangements in the most parsimonious scenario

Examples

Reversal

1 2 3 4 5 6 → 1 2 -5 -4 -3 6

Translocation

Chromosome 1: 1 2 3 4 → 1 2 6 4 5
Chromosome 2: 5 6 → 3

Fusion

Chromosome 1: 1 2 3 4
Chromosome 2: 5 6

→
←

1 2 3 4 5 6

Fission

Why study genome rearrangements?

- Useful in studying evolution
- Less ambiguity in interpreting the mutations
- A larger scale of data which is more appropriate for studying evolution

Adjacencies and Breakpoints

$\pi = \pi_1 \pi_2 \pi_3 \ldots \pi_{n-1} \pi_n$

- A pair of elements π_i and π_{i+1} are adjacent if $\pi_{i+1} = \pi_i \pm 1$
- An adjacency - a pair of adjacent elements that are consecutive
- A breakpoint - a pair of adjacent elements that are not consecutive
- For example:

  $\pi = 1 \mid 9 \mid \underline{3\ \ 4} \mid \underline{7\ \ 8} \mid 2 \mid \underline{6\ \ 5}$

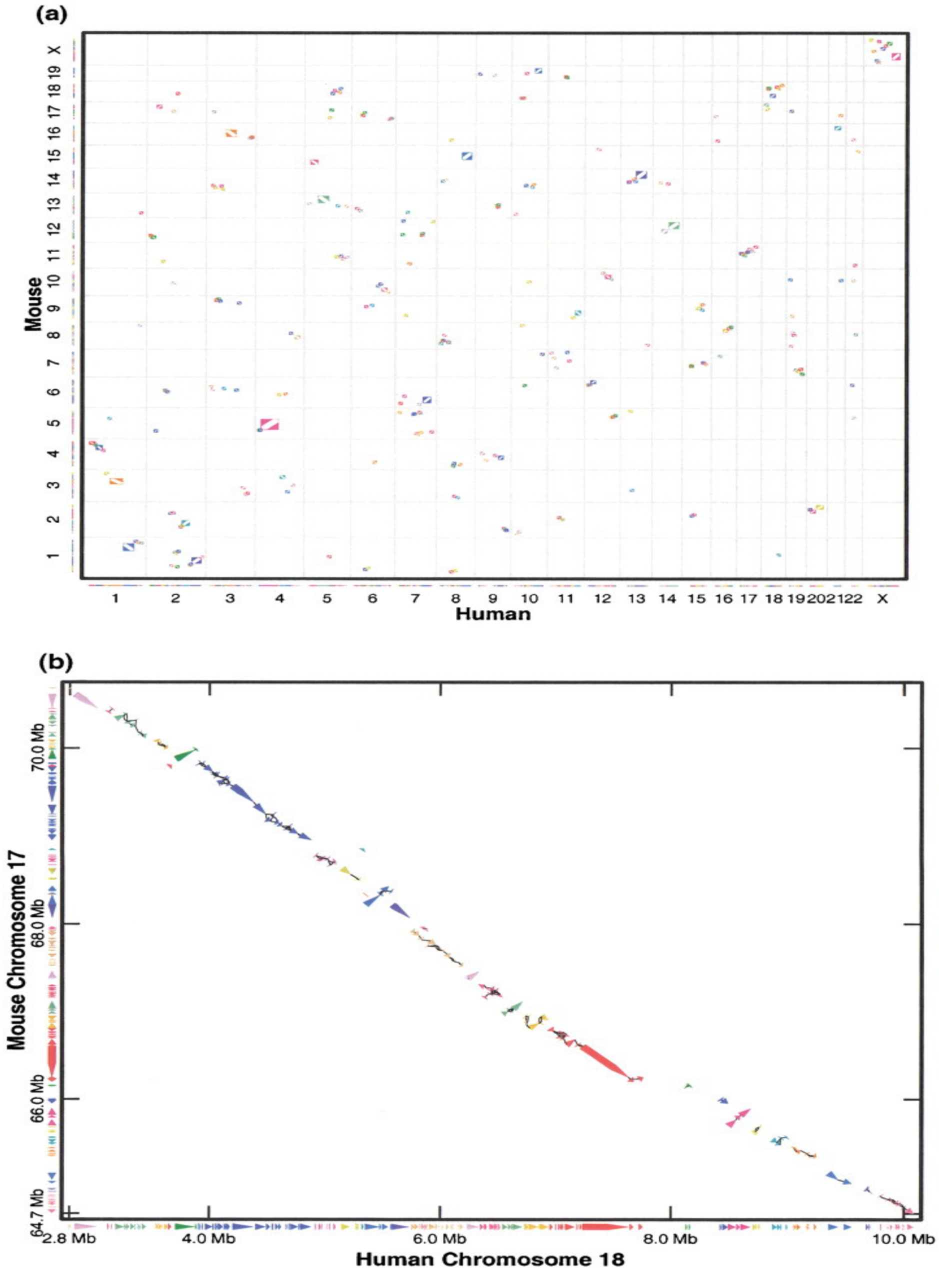
Shortcomings of earlier works

- Do not distinguish between micro- and macro-rearrangements
- Unreliable assignment of orthologs
- Conserved gene order can be disrupted by recent duplications and insertions

Problem

“ To obtain a meaningful estimate of the number of rearrangement events on the evolutionary path from mouse to human”

Human and mouse synteny blocks



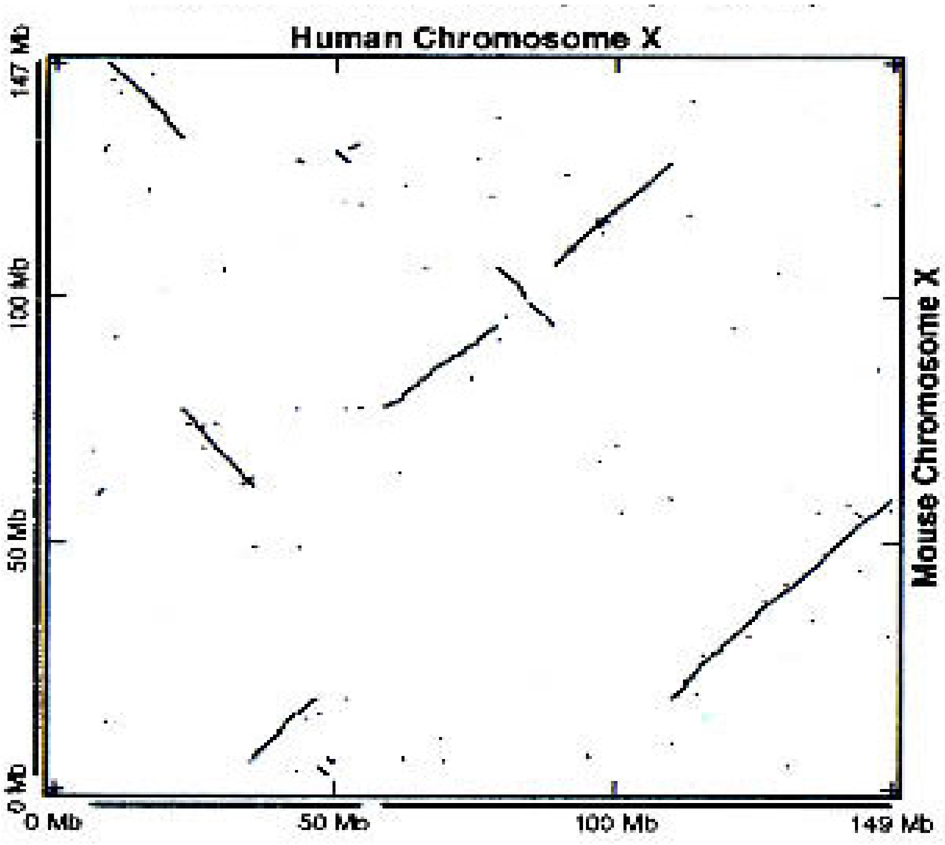
Synteny blocks are segments that can be converted into conserved segments by micro-rearrangements

Human and mouse genomes share 281 synteny blocks

GRIMM-Synteny Algorithm

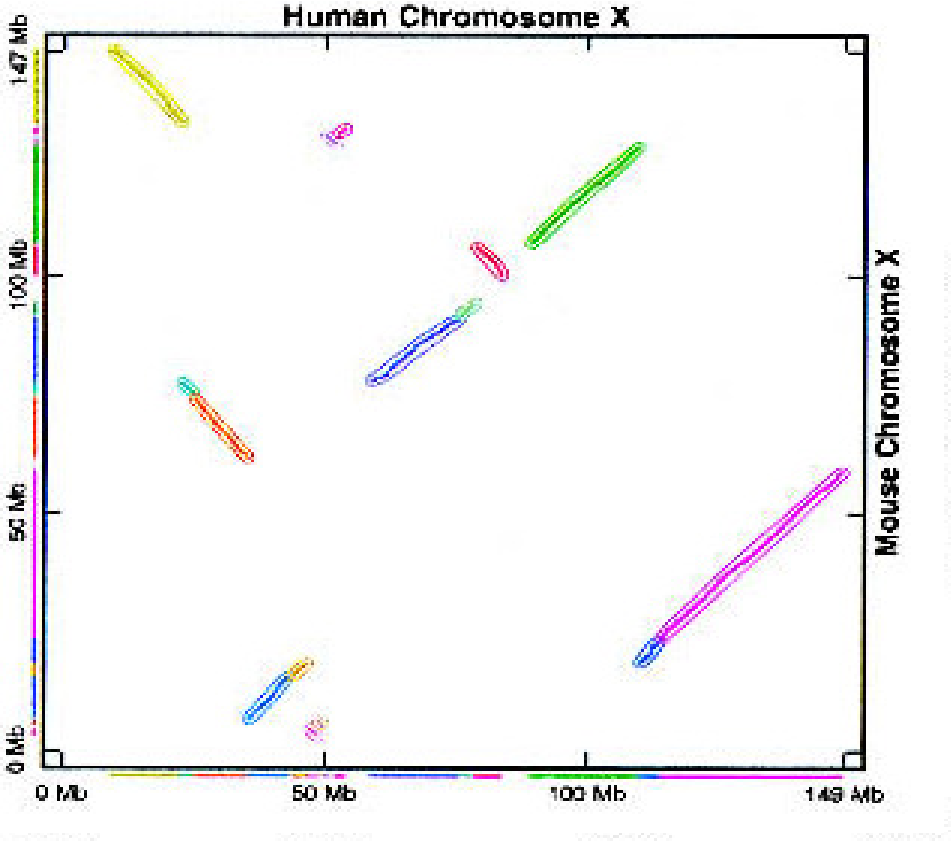
- Form an anchor graph whose vertex set is the set of anchors (bi-directional best local similarities called anchors).
- Connect vertices in the anchor graph by an edge if the distance between them is smaller than the gap size G.
- Determine the connected components of the anchor graph. Each connected component is called a cluster.
- Delete “small” clusters (shorter than the minimum cluster size C in length).
- Determine the cluster order and signs for each genome.
- Output the strips in the resulting cluster order as synteny blocks

An example: X-chromosome



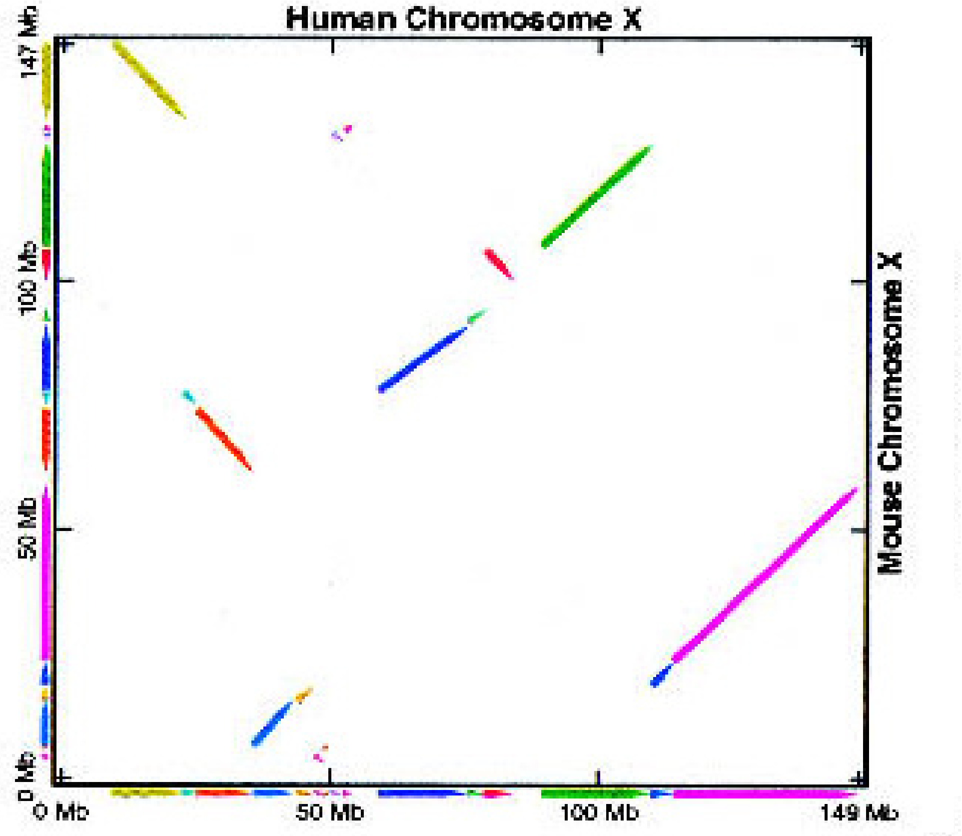
Dot plot of anchors

An example: X-chromosome



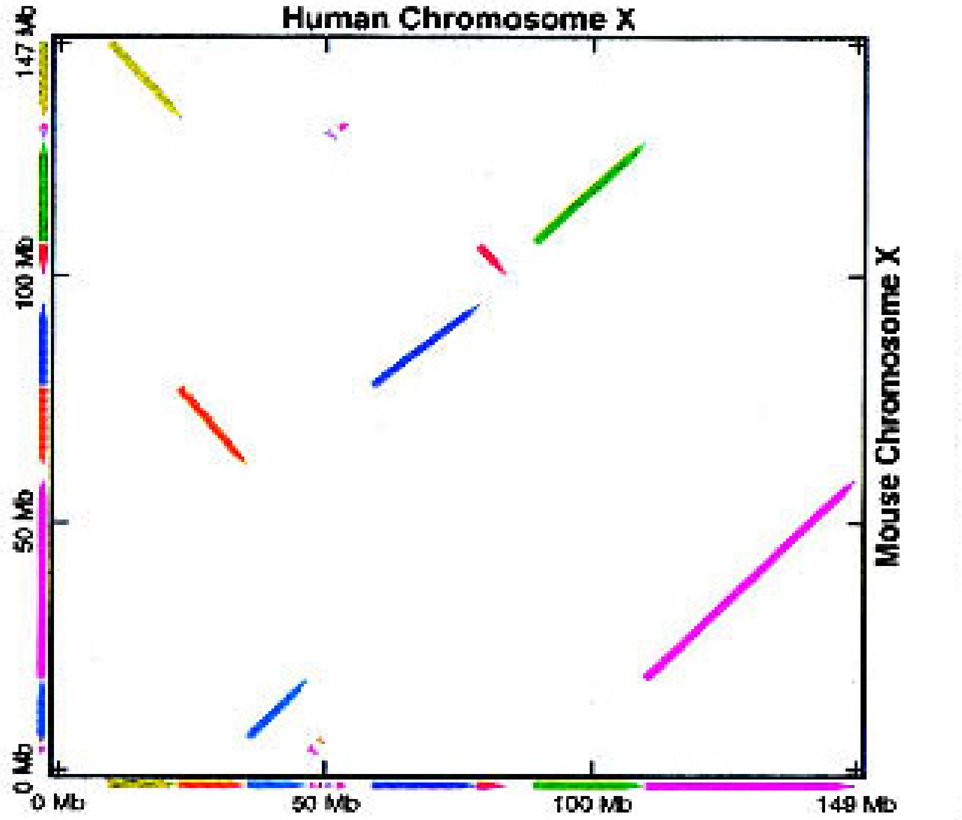
Cluster of anchors

An example: X-chromosome



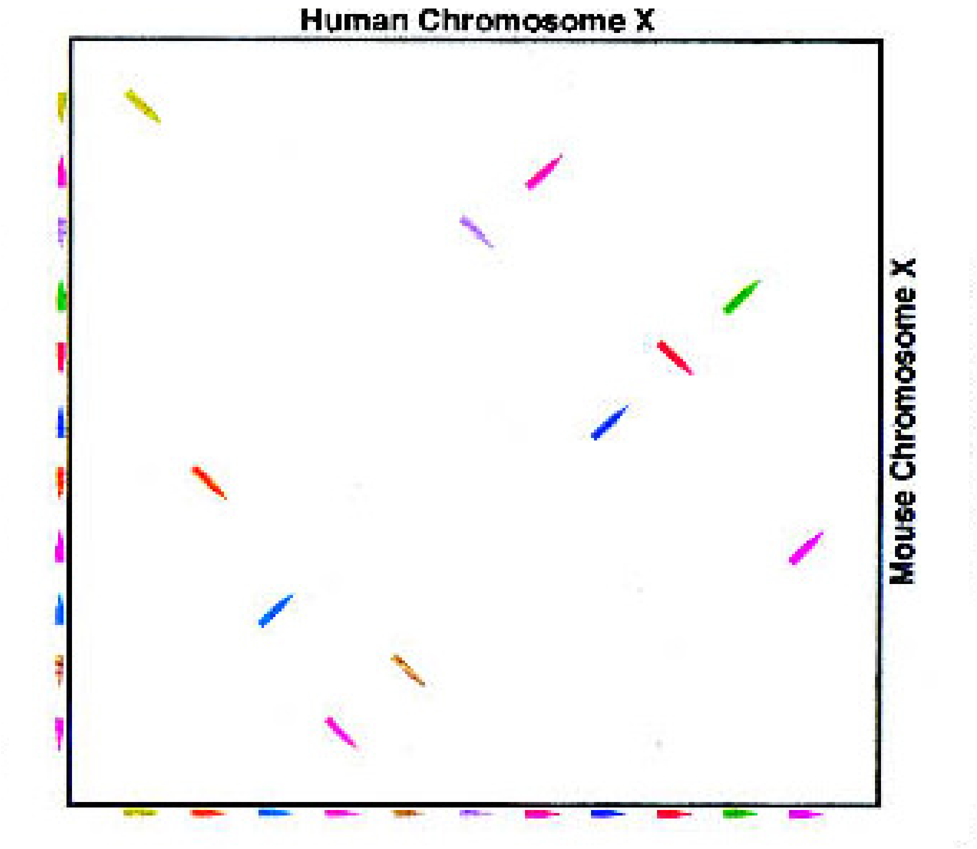
Rectified anchors

An example: X-chromosome



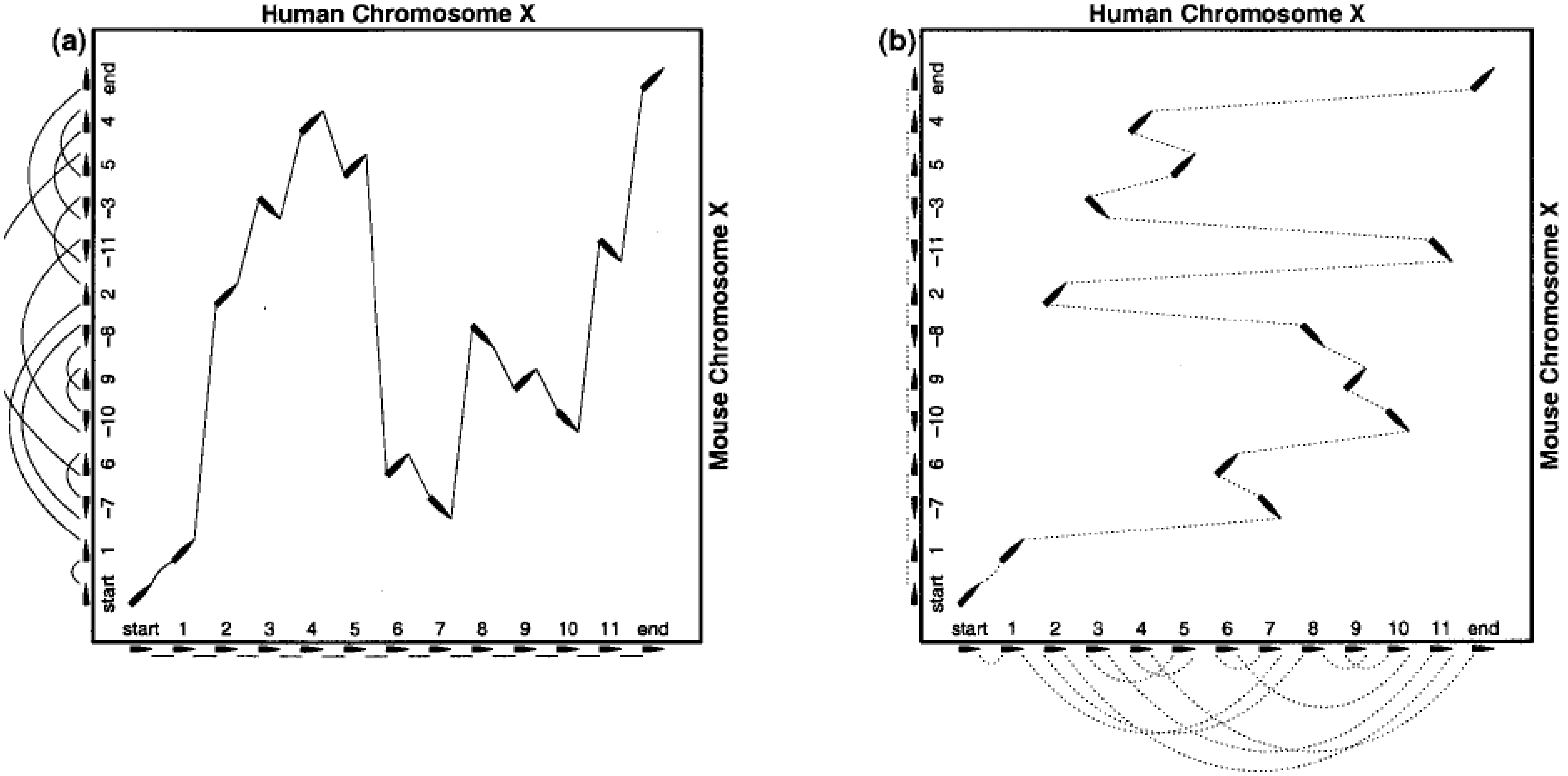
Synteny blocks

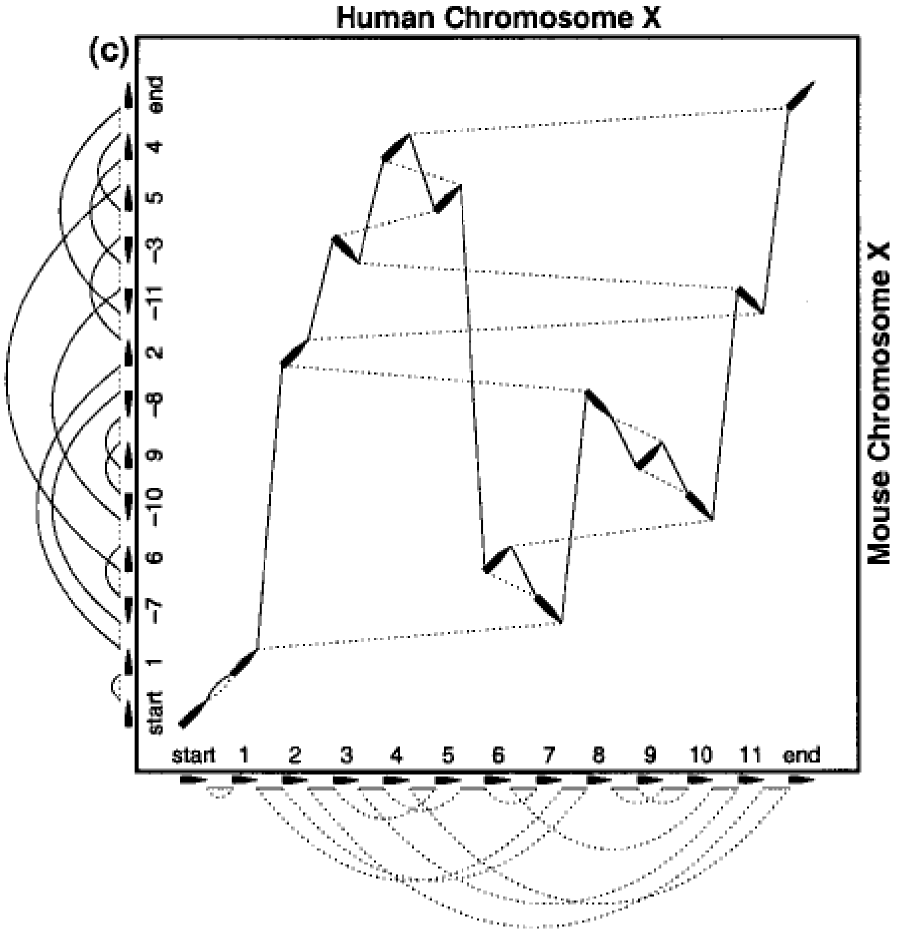
An example: X-chromosome

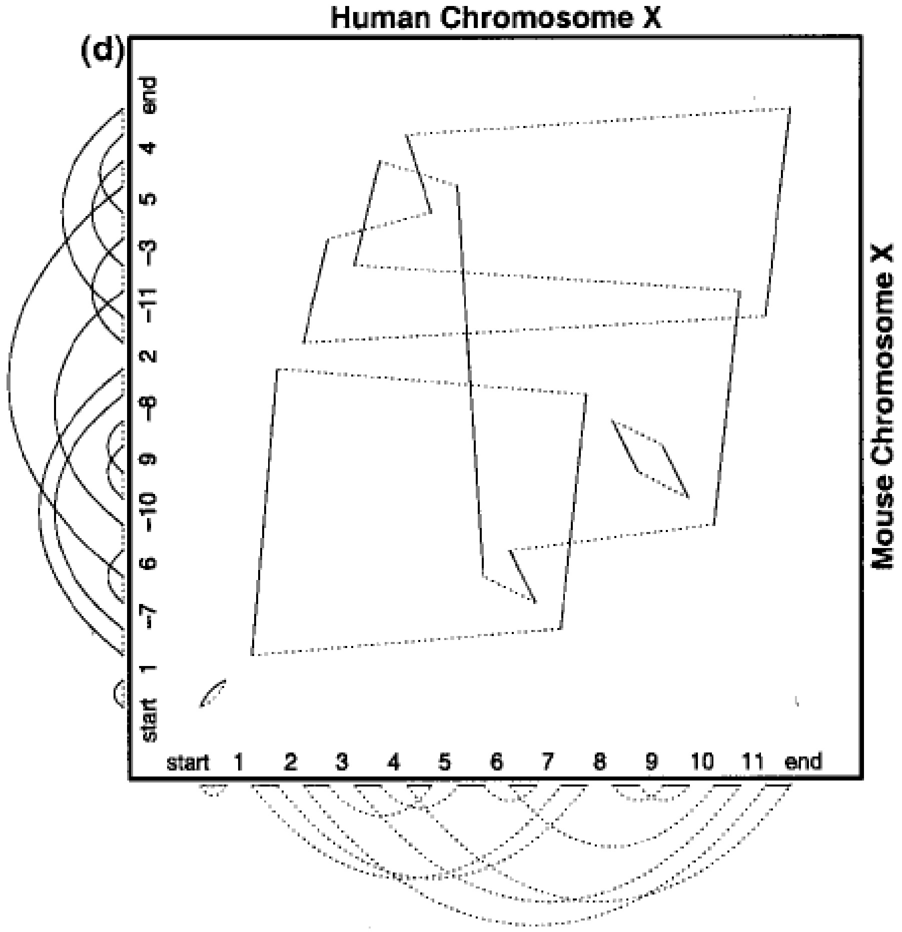


Synteny blocks as units of same size

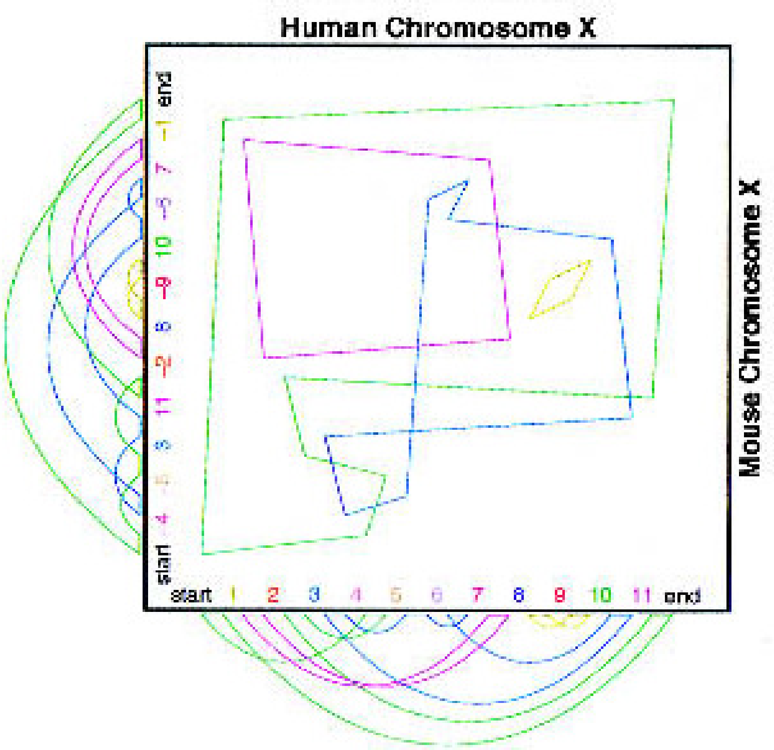
A new way to construct Breakpoint Graph

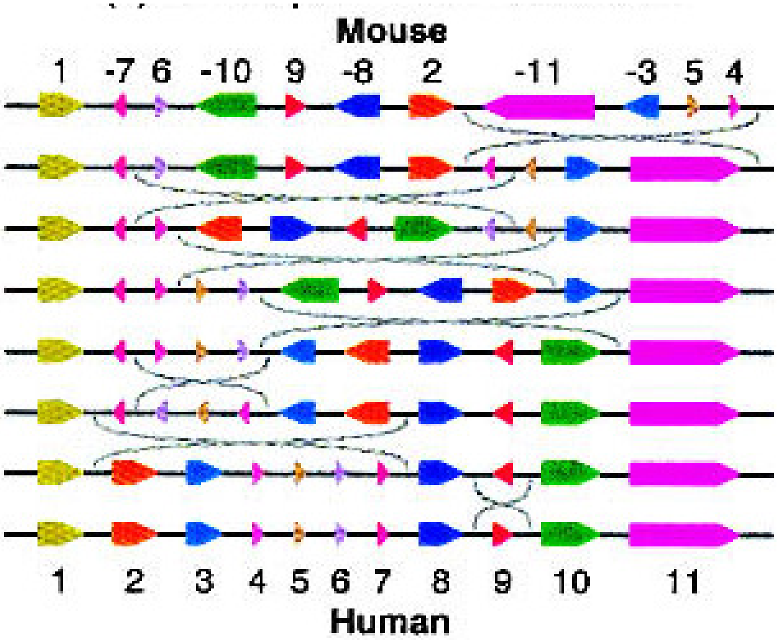



(c)
Human Chromosome X
Mouse Chromosome X
start
1
2
3
4
5
6
7
8
9
10
11
end
start
1
−7
6
−10
9
−8
2
−11
−3
5
4
end


(d)
Human Chromosome X
Mouse Chromosome X
start
1
2
3
4
5
6
7
8
9
10
11
end
start
1
−7
6
−10
9
−8
2
−11
−3
5
4
end

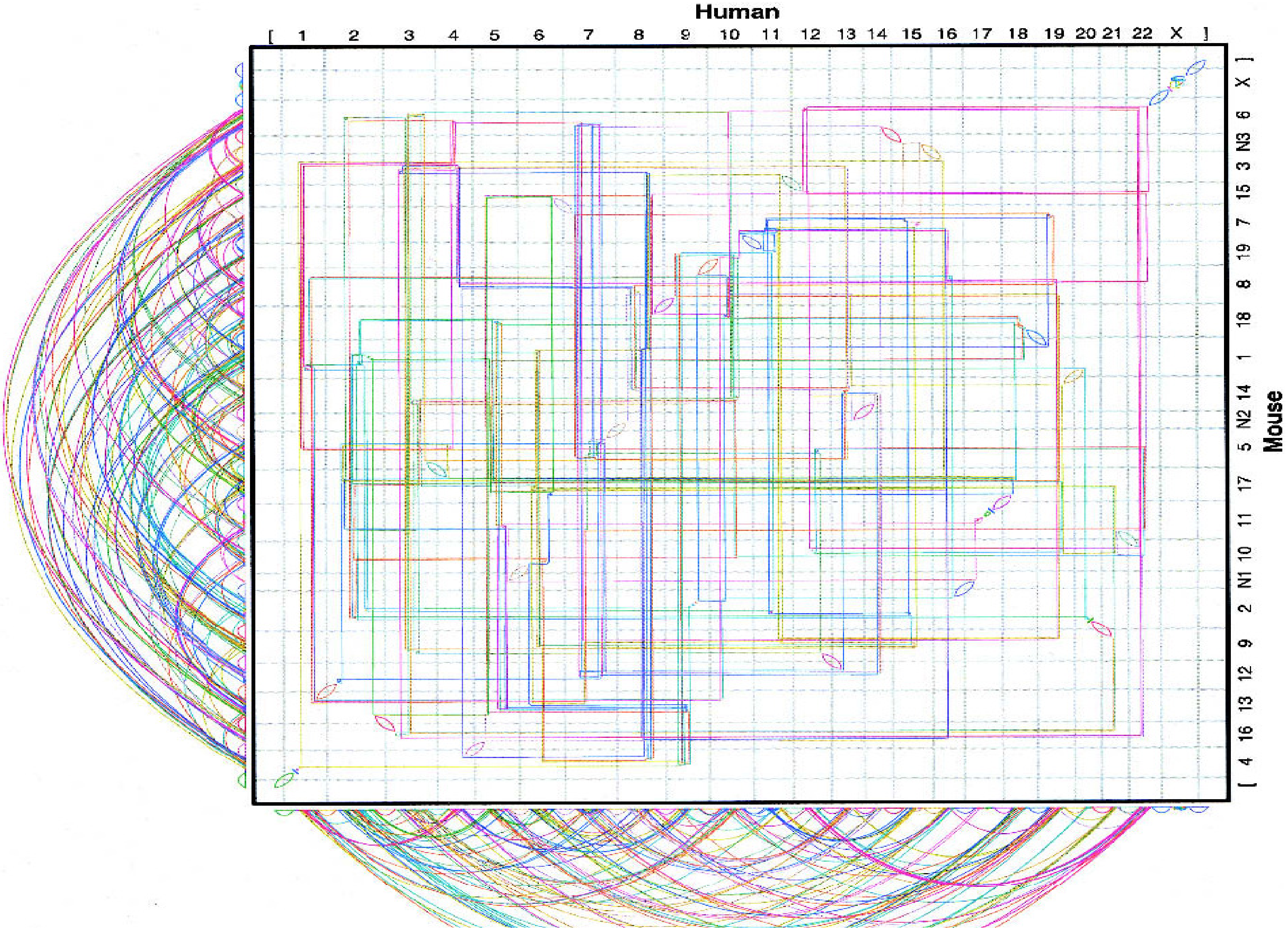
Parsimonious rearrangement scenario





Hannenhalli-Pevzner algorithm uses breakpoint graph to construct the most parsimonious evolutionary scenario

Multichromosomal breakpoint graph of the whole human and mouse genomes



Reversal distance

- Rd is at most ½ the number of breakpoints in the genome
- Inaccurate as breakpoints might be reused in the evolution
- Hannenhalli and Pevzner theorem estimates

$$Rd = n+1-c+h$$

- A similar theorem holds for multichromosomal genomes
- Fast implementation of the Hannenhalli Pezner algorithm available via GRIMM web server
- 245 rearrangements (149 inversions, 93 translocations, 3 fissions)
- 41 out of 281 synteny blocks do not show any rearrangements, 10 are extremely rearranged

Summary

- New algorithm for constructing synteny blocks
- Study arrangement of snyteny blocks in human and mouse
- Derive a most parsimonious human-mouse rearrangement scenario
- Provide evidence that intrachromosomal rearrangements are more frequent than interchromosomal rearrangements