The reconstruction of ancestral genomes for groups of species and the precise identification of associated chromosomal rearrangements are basic questions in evolution.

By comparing chromosomes in different species, blocks of consecutive genes, called syntenic blocks, inherited from their last common ancestor, can be identified. Breakpoints, that are regions between syntenic blocks, result from rearrangements such as: inversions of a DNA segment within a chromosome, reciprocal translocations between two chromosomal arms, duplications, insertions and deletions of DNA segments.

Many models have been already proposed [1,2]. All models are based on the principle of parsimony: the differences between the two genomes have to be explained by a minimum number of rearrangements. The main challenge is to precisely model the information contained in genomes to avoid losing crucial evidences, allowing to trace back as many as possible rearrangements that occurred during evolution.

Our algorithm is designed in order to:
1. Identify all the rearrangements that occurred in the different lineages and their resulting breakpoint regions
2. Reconstruct ancestral genomes with the maximum number of genes and the minimum number of chromosomes.

**Algorithm**

**Introduction**

1. Identification of homologies by Bi-Directional Best Hits (BDBH)
2. Construction of blocks with a BDBH: two consecutive pairs of homologous genes cannot be separated by more than 5 genes failing the BDBH condition
3. Rearrange blocks of genomes: we add homologous genes having at least 30% of similarity over at least 50% of their length
4. Fusion of neighboring blocks in both genomes
5. Definition of the sign for blocks

Tolerating small rearrangements inside blocks in order to keep all the syntenic relationships

**Construction of Synteny Blocks**

- Protein sequences from n species
- Phylogenetic tree of the n species

**Comparison between chromosomes of species G1 and G2:**

We observe several combinatorial arrangements between genes within synteny blocks: perfect order conservation (block 1), inclusion (3 in 2 in G1), microarrangement (H), overlapping (5 and 6 in G1), unassigned blocks (9 in G2, duplication (10 and 11 in G2).

**Algorithm**

**Output**

Ancestral Genome

List of Rearrangements associated to branches

Associated Breakpoints and re-use estimations

**Results**

Complete reconstruction of two ancestral yeast genomes (sharing 90% of their genes with extant species) and identification and localization of the corresponding rearrangements on the branches of the tree

**Conclusion**

1. Precise definition of syntenic blocks, using a parameterized proximity of genes within the blocks and two different measures of homology between genes
2. An algorithm combining the notion of linked breakpoints [1] (adapted to deal with the new notion of synteny pack) and the local adjacency comparison [2] (adapted to exploit the strength of pairwise comparison of several genomes)

we are able to:
1. Reconstruct several complete ancestral genomes with a realistic number of chromosomes
2. Reconstruct the history of chromosomal rearrangements and precisely define the breakpoint regions

These reconstructions are the first step to study genome structures, re-use breakpoints, mechanisms of rearrangement and more generally genome evolution.

---