Agri Articles
(e-Magazine for Agricultural Articles)
Volume: 03, Issue: 06 (NOV-DEC, 2023)
Available online at http://www.agriarticles.com
©Agri Articles, ISSN: 2582-9882

Consensus Mapping: In an Era of Comparative Genetics

(*H. P. Vadodariya, H. H. Patel and D. V. Makwana)
Department of Genetics and Plant Breeding, Navsari Agricultural University, Gujarat
*Corresponding Author's email: himanivadodariya@gmail.com

Abstract
Genetic maps are a tool to study genetic linkage, genome structure and evolution and evolutionary biology. Despite the development in the number of marker types as well as the fast use of high-throughput genotyping techniques, order conflicts frequently arise among these distinct genetic maps, primarily as a result of experimental mistakes. The main goal of integrating several genetic maps is to identify new information regarding marker orders and to resolve order disputes between separate maps. In this paper, a range of map integration strategies have been presented, all of which aim to maximise the goal function of determining a consensus genetic map. The most current method is based on graph theory and use directed acyclic graphs (DAGs) to describe mappings from individual populations.

Introduction
Researchers investigating genetic linkage, genomic structure, and evolution can benefit from genetic maps. They have been employed as scaffolding during genome assembly and in experiment design. In the past, scientists have concentrated on producing a individual reference map derived from a single population; but, in light of the new trend towards comparative genetics, scientists are now collecting data from other populations and lineages within the single species.

Multiple genetic maps for a single species are now accessible because to the quick uptake of high-throughput genotyping techniques like recombination analysis and physical imaging in recent years. We may generate a larger density of markers and, hence, a better genome coverage by combining these maps into a consensus genetic map, as opposed to any single genetic map. Nevertheless, order conflicts frequently arise among these distinct genetic maps, primarily as a result of experimental mistakes. Thus, creating consensus genomic maps is one of computational biology's most difficult tasks. On the basis of only the marker order relationships given by a particular set of individual genetic maps, several computational techniques have been presented.

Genetic linkage map and Consensus map
- The relative locations of genetic markers along a chromosome are depicted in genetic linkage maps. The likelihood of genetic loci getting separated during segregation and recombination is correlated with the genetic distance between markers.
- The challenge of creating genetic linkage maps using genotyping data dates back to the early 1900s, when biologists studying chromosomal recombinational activity and structure while studying the molecule.
- Only a few of phenotypic markers, primarily resulting from mutation, were recorded in early genetic linkage maps by tracking the biochemical and phenotypic alterations of the

organism under investigation. Genetic maps are far more dense with DNA-based markers (such as RFLPs, RAPDs, SSRs, and AFLPs) have been introduced.

- In recent years, the development in the number of marker types as well as the fast use of high-throughput genotyping technologies have occurred simultaneously. Multiple genetic maps are becoming more and more frequent for the same organism, generally for various genotyping methods and sets of genetic markers.
- When there are several genetic maps accessible, they usually include some common markers. on these situations, creating a larger map referred to as a consensus map. Because it offers a larger marker density and hence a more genome coverage level than the individual map construct, a consensus map is preferable.

Error in consensus mapping

It is not always possible to construct a consensus map that is consistent with the individual maps since genotyping errors are likely to result in ordering conflicts between the different maps. There are two kinds of inaccuracies that might be seen because of the manner that individual genetic maps are put together using genotyping data.

1. Local reshuffles: Local reshuffles which can be defined as a inaccuracies shown in a sequential order of nearby markers

2. Global displacement: A situation where a few markers are positioned far from the proper ones is referred to as a global displacement.

Steps for Consensus mapping

1. Constructing individual genetic map: A linear sequence of bins, each of which may contain one or more genetic markers, makes up an individual genetic map. The partial order on chromosomal markers is defined by this mapping investigation, which is the source of its generation. The relative ordering of markers in the same bin are unknown, however markers from separate bins are arranged according to their respective bins. Take the genetic map 2 {8 5} 4 3, for instance, where two markers from the same bin are enclosed in a curly bracket. Markers 8 and 5 are arranged in the same sequence as marker 2, however there is no indication of a relative hierarchy between them.

2. Constructing Consensus genetic map: The main goals of integrating several genetic maps are to identify new information regarding marker orders and to resolve order disputes between separate maps. A consensus genetic map is the term commonly used to describe the final product of map integration exercises. With more coverage and precision than any component individual genetic map, it defines a partial order on markers. Similar to an individual genetic map, a directed acyclic graph may also be used to describe a consensus genetic map, however this model frequently needs a generic graphical framework. These days, a wide range of map integration strategies have been put forth, all of which aim to maximise the goal function of determining a consensus genetic map. Finding a consensus genetic map, or an acyclic subgraph generated from the aggregate graph, is a widely used method that involves deleting the minimal feedback edge set, which is the smallest set of edges, as seen in figure.

3. Rearrangement distances between partially ordered genomes: The

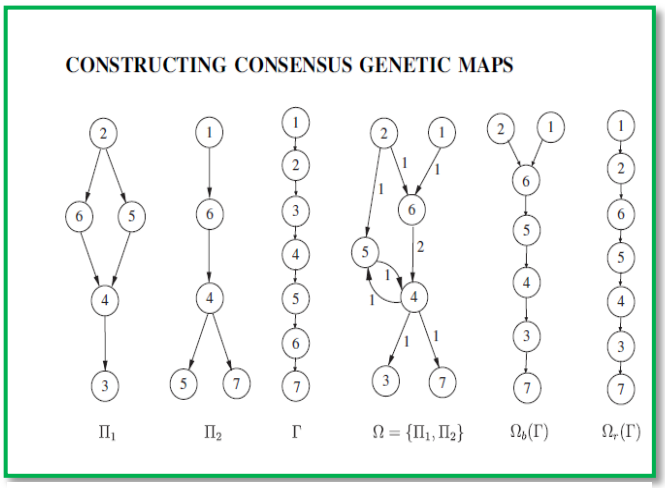

Figure 1: An illustrating example of consensus map in comparative analysis

evolutionary distance depicted between any of the two genomes in comparative genomics investigations is frequently measured using the breakpoint distance between them and the reversal distance. They were first described on two genomes in the entire order, and more recently, by adding the idea of linearization, they were extended to two partly ordered genomes. A topological sort called a linearization of a partly ordered genome P represents a potential overall order of all the markers.

4. Common adjacencies of two genomes: A breakpoint is defined as two markers that are close in one genome yet differ in another, given both genomes in the full order. We shall refer to two markers as forming a common adjacency if they are contiguous in both genomes. One noteworthy finding is that, for every given genome, the sum of the common adjacencies and breakpoints (i.e., breakpoint distance) is one fewer than the genome's size. When calculating the breakpoint distance between two partially-ordered genomes, finding two linearizations that minimize the number of breakpoints is similar to finding two linearizations that reduce the number of common adjacencies.

Different approaches to construct consensus map

The most popular conventional mapping strategy is to combine genotyping data from individual mapping populations and then use common mapping methods to construct the consensus map. Although this pooling technique is frequently employed, it has a number of drawbacks. First of all, not all circumstances might benefit from it. Two data from a double haploid (DH) population and a F2 recombinant inbred lines (RILs) population, for instance, are two examples of distinct populations from which the data cannot be combined and handled consistently downstream. Second, there are a lot of missing observations from the pooling process, and the fraction of missing data rises as more data sets need to be joined.

Another method is to utilise JOINMAP to generate consensus estimates of pairwise genetic distances while accounting for population structure and size. Then, one looks for a map that minimises an objective function that quantifies the map's fit to the distance estimations as well as its overall quality. The disadvantages of this technique are twofold. First, distance estimations based on a limited sample of recombination events are not particularly reliable. The construction of genetic maps using these estimations will result in errors in the ordering of adjacent markers. Second, the computational difficulty of finding the best map with regard to the objective function in use is quite time demanding. Despite these disadvantages, JOINMAP remains the most used software tool for creating consensus maps.

Other commercial tools that are less well-known include MULTIPOINT and CARTEBLANCHE. MULTIPOINT's consensus map-building technique involves reprocessing the initial genotyping data rather than integrating the different maps. The difficulty that MULTIPOINT must solve is computationally highly difficult, limiting the number of markers in the maps significantly.

The most current method is based on graph theory and use directed acyclic graphs (DAG) to describe mappings from individual populations. On the basis of their common vertices, the set of DAGs is subsequently combined into a consensus graph.

Conclusion

In conclusion, genetic maps are important tools for studying genetic linkage and genome structure, but integrating multiple genetic maps into a consensus map can provide a better genome coverage. However, order conflicts often arise among these maps due to experimental mistakes and various strategies, including the use of graph theory, are being used to create a consensus genetic map. The most current method is based on graph theory and use directed acyclic graphs (DAG) to describe mappings from individual populations.

References

1. Chen, X. and Yang, J. Y. (2010). Constructing consensus genetic maps in comparative analysis. *Journal of Computational Biology*, **17**(11): 1561-1573.
2. Jackson, B. N., Aluru, S. and Schnable, P. S. (2005). Consensus genetic maps: a graph theoretic approach. In *2005 IEEE Computational Systems Bioinformatics Conference* (pp. 35-43). IEEE.
3. Wu, Y., Close, T. J. and Lonardi, S. (2010). Accurate construction of consensus genetic maps via integer linear programming. *IEEE/ACM Transactions on computational biology and bioinformatics*, **8**(2): 381-394.