Theoretical advances on multichromosomal median computation

Ahmad Mahmoody Ghaidary[1], Cedric Chauve[1], and Ladislav Stacho[1]

Department of Mathematics, Simon Fraser University

In the present work, we consider the classical problem of computing an optimal median for three multichromosomal genomes G_1, G_2, G_3. We denote the breakpoint graph of a pair (G_i, G_j) of genomes by $B(G_i, G_j)$ and the number of alternating cycles of this graph by $c(B(G_i, G_j))$. Formally, the problem we consider is the following:

- G_1, G_2, G_3 are mixed genomes (they can have both linear and circular chromosomes);
- the distance between two genomes G_i and G_j is defined by $d(G, G') = n - c(BG(G, G'))$, which corresponds to the double-cut-and-joint (DCJ) distance for circular genomes,
- median genomes are circular.

Moreover, we define a *3-edge* in a median genome as an edge which connects two vertices of degree three in the breakpoint graph of the three genomes $B(G_1, G_2, G_3)$.

Theorem 1. *Let G_1, G_2, G_3 be three genomes and $B = B(G_1, G_2, G_3)$ their breakpoint graph. If there exists an optimal circular median M for G_1, G_2, G_3 with at most ℓ 3-edges, then computing an optimal circular median can be done in time $O(poly(n) \times 3^m m^\ell)$, where m is the number of vertices of degree 3 in B. In particular*

1. *if m is bounded, there is a polytime algorithm for the median problem.,*
2. *if ℓ is bounded, there is an FPT algorithm for the median problem.*

Our proof relies on an intermediate, weaker result:
If the breakpoint graph of three mixed genomes G_1, G_2 and G_3 has maximum degree 2, then an optimal circular median can be computed in polynomial time.
The proof of this intermediate result is based on the fact that every breakpoint graph with maximum degree two consists of paths and cycles. We show that there exist an optimal median such that: (1) vertices of cycles and paths with even number of edges, are matched between themselves, and (2) the set of all odd paths and odd cycles can be paired, such that no edge in the median connect two vertices in two different parts of the pairing. From there, we prove that is tractable to find optimal adjacencies within even paths and cycles, and that dealing with odd paths and cyles reduces to solving a maximum weight matching problem. Our main result is obtained from Theorem 1 by considering all possibilities for vertices of degree 3: a vertex of degree three can be matched to another vertex of degree three and then by shrinking we remove them from the breakpoint graph and make it simpler, otherwise one of its edges is not in any bicolored cycle so we can remove that edge and make the vertex having degree 2.