

We aim to identify recombinants in complicated multiple-source partial alignments, what we did is to employ the division strategy into algorithm, specifically, we identify triples of sequences which consist of the target sequence and two consecutive source sequences. In order to illustrate its reasonability for sufficing to detect recombinants, the following plot is one example of phylogenetic network, where two recombinants are involved in these four sequences a, b, c, d , which leads to three-source partial alignment. In fact, the identified recombinants from every two adjacent segments are the same with true recombinants in each possible phylogenetic network.

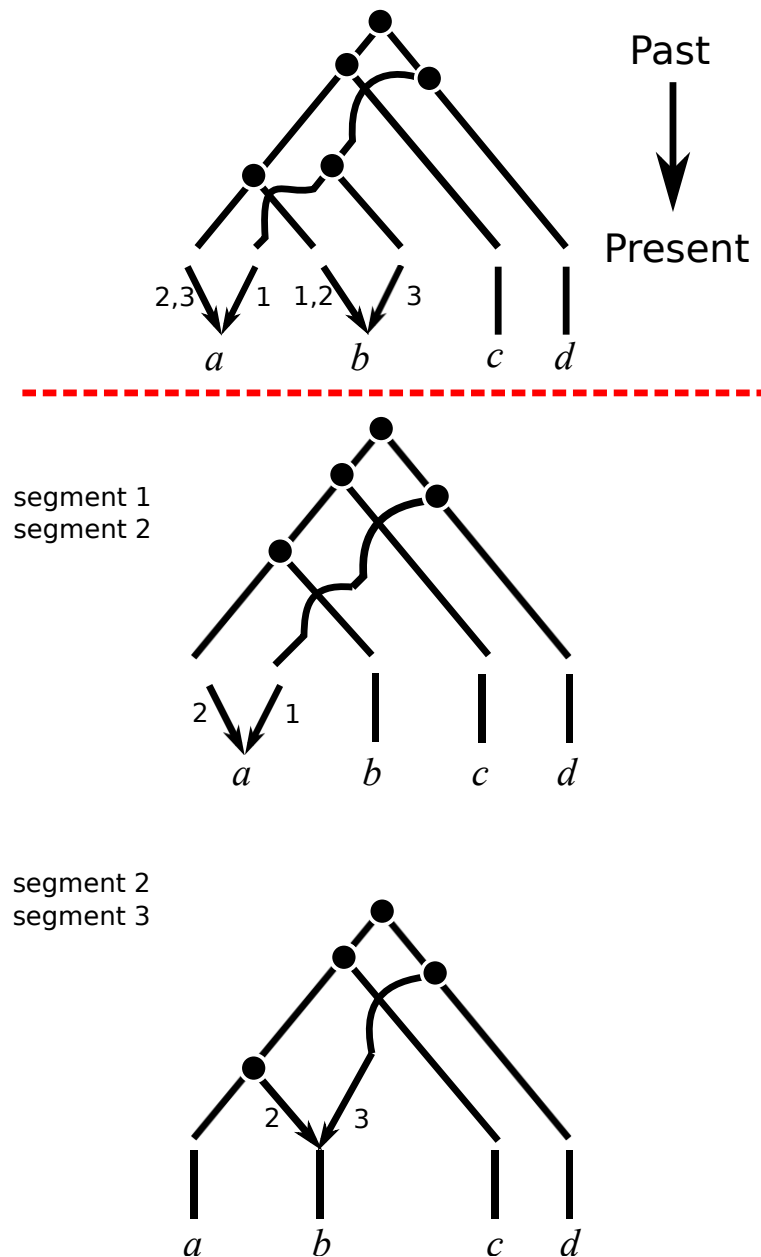


Figure 1: Numbers beside the line represent the alignment segment. Below two sub-networks are part of the top network, but each is related with two consecutive segments (segment 1,2 and segment 2,3). Based on recombinant identification algorithm in simpler two-source alignment, we get all identified recombinants.

segment 1 and segment 2	
triple	identified recombinant
abc	a
acd	a
abd	a
bcd	b, c or d

segment 2 and segment 3	
triple	identified recombinant
abc	b
bcd	b
abd	b
acd	a, c or d

Above tables illustrate the identification results from network in Figure 1. It suffices to detect recombinant sequences. We are able to detect recombinants accurately from aligned triples as long as that particular triple contains one true recombinant.

There is also a certain possibility to detect recombinants incorrectly, for instance, the triple bcd in segment 1 and 2, then b, c, d has the equal probability ($1/3$) to be identified as recombinant separately. Therefore, we tend to get more false positive detection when each network involves more and more triples. This actually explains the lower specificity for multiple-source recombinants than simpler two-source or three-source simulation scenarios.