

Identifiability in the network multispecies coalescent model using quintet concordance factors

Joseph Cummings

We study identifiability questions for the network multispecies coalescent model using concordance factors on small subsets of taxa. Building on foundational work on trees using quintet concordance factors by Allman, Degnan, and Rhodes, as well as advances for networks using quartet concordance factors by Allman, Baños, Garrote-Lopez, and Rhodes, we focus on the additional information provided by quintets.

Our approach is algebraic. These models admit polynomial parameterizations and thus define algebraic varieties. By studying the vanishing ideals of these varieties and identifying (partial) generating sets, we translate identifiability questions into algebraic ones. For level-1 phylogenetic networks, we show that quintet concordance factors often determine key structural features of the network, including the location of hybridization events and, in many cases, the position of the root.