

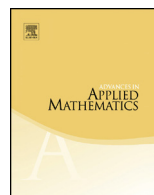


ELSEVIER

Contents lists available at ScienceDirect

Advances in Applied Mathematics

www.elsevier.com/locate/yaama



# Enumeration of lonely pairs of gene trees and species trees by means of antipodal cherries



Noah A. Rosenberg

Department of Biology, Stanford University, Stanford, CA 94305, USA

## ARTICLE INFO

*Article history:*

Received 28 February 2018

Received in revised form 1

September 2018

Accepted 4 September 2018

Available online xxxxx

*MSC:*

05A15

05C05

92D15

*Keywords:*

Cherries

Coalescent histories

Phylogenetics

## ABSTRACT

In mathematical phylogenetics, given a rooted binary leaf-labeled gene tree topology  $G$  and a rooted binary leaf-labeled species tree topology  $S$  with the same leaf labels, a coalescent history represents a possible mapping of the list of gene tree coalescences to the associated branches of the species tree on which those coalescences take place. For certain families of ordered pairs  $(G, S)$ , the number of coalescent histories increases exponentially or even faster than exponentially with the number of leaves  $n$ . Other pairs have only a single coalescent history. We term a pair  $(G, S)$  *lonely* if it has only one coalescent history. Here, we characterize the set of all lonely pairs  $(G, S)$ . Further, we characterize the set of pairs of rooted binary unlabeled tree shapes at least one of the labelings of which is lonely. We provide formulas for counting lonely pairs and pairs of unlabeled tree shapes with at least one lonely labeling. The lonely pairs provide a set of examples of pairs  $(G, S)$  for which the number of compact coalescent histories—which condense coalescent histories into a set of equivalence classes—is equal to the number of coalescent histories. Application of the condition that characterizes lonely pairs can also be used to reduce computation time for the enumeration of coalescent histories.

© 2018 The Author. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

*E-mail address:* [noahr@stanford.edu](mailto:noahr@stanford.edu).

<https://doi.org/10.1016/j.aam.2018.09.001>

0196-8858/© 2018 The Author. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

The study of evolutionary trees makes a distinction between *species trees*, trees that describe relationships among sets of species, and *gene trees*, trees that describe relationships among genetic lineages of members of those species [3,11,12]. In considering the evolution of gene trees in relation to species trees, a variety of new types of combinatorial structures have emerged, each specifying some feature of the relationship between the branching patterns of gene trees and those of species trees [2,4,5,11,17,19,21,22].

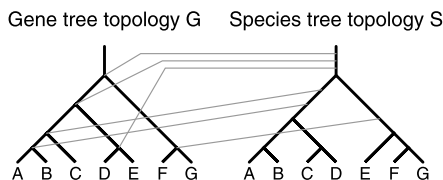
*Coalescent histories* are prominent among the structures useful in the study of gene trees and species trees. For a given gene tree and species tree, a coalescent history describes an evolutionary scenario for gene lineage evolution on the branches of the species tree. More formally, viewing a rooted binary tree “backward in time” from the leaves to the root, each internal node of the tree, including the root node, represents a *coalescence*: an instance at which lineages represented by a set of leaves find common ancestors. We term a node or edge  $v$  of a tree an *ancestor* of a node or edge  $u$  if  $u$  lies on a path from  $v$  to a leaf;  $u$  is said to be a *descendant* of  $v$ . Trivially,  $v$  is an ancestor or descendant of itself. We then have the following definition.

**Definition 1.** Consider a rooted binary leaf-labeled tree  $G$  (the “gene tree topology”) and a rooted binary leaf-labeled tree  $S$  (the “species tree topology”), labeled by the same set of mutually distinct leaf labels. A *coalescent history*  $f$  associates with each coalescence  $v$  in  $G$  an edge  $f(v)$  in  $S$ , such that two properties are satisfied.

(i) For each gene tree coalescence  $v$  in  $G$ , the species tree edge  $f(v)$  in  $S$  is ancestral to each leaf node of  $S$  that shares a label with a leaf that descends from  $v$ .

(ii) For each pair of gene tree coalescences  $u, v$  with the property that  $v$  is ancestral to  $u$  in  $G$ ,  $f(v)$  is ancestral to  $f(u)$  in  $S$ .

Treating a gene tree as evolving on the branches of a species tree, coalescent histories describe permissible lists of edges of the species tree—including as a possibility an edge ancestral to its root—where the coalescences of the gene tree can take place [5,14]. From a biological perspective, the pair of constraints in the definition encodes the rules that (i) a set of gene lineages can coalesce only in a branch of the species tree that is possible for them all to reach, and (ii) ancestors can coalesce no more recently than their descendants (Fig. 1).



**Fig. 1.** A coalescent history for a labeled gene tree topology  $G$  and labeled species tree topology  $S$ . Gray lines represent the mapping of coalescences of the gene tree to edges of the species tree.

Download English Version:

<https://daneshyari.com/en/article/10224175>

Download Persian Version:

<https://daneshyari.com/article/10224175>

[Daneshyari.com](https://daneshyari.com)