PHILOSOPHICAL TRANSACTIONS B

royalsocietypublishing.org/journal/rstb

Research

Cite this article: Dickey EH, Rosenberg NA. 2025 Labelled histories with multifurcation and simultaneity. *Phil. Trans. R. Soc. B* **380**: 20230307.
https://doi.org/10.1098/rstb.2023.0307

Received: 12 June 2024
Accepted: 11 September 2024

One contribution of 18 to a theme issue '"A mathematical theory of evolution": phylogenetic models dating back 100 years'.

Subject Areas:
evolution, genetics

Keywords:
labelled histories, multifurcation, phylogenetic models

Author for correspondence:
Noah A. Rosenberg
e-mail: noahr@stanford.edu

Labelled histories with multifurcation and simultaneity

Emily H. Dickey and Noah A. Rosenberg

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

EHD, 0000-0002-1554-8859; NAR, 0000-0002-1829-8664

In mathematical models of phylogenetic trees evolving in time, a labelled history for a rooted labelled bifurcating tree is a temporal sequence of the branchings that give rise to the tree. That is, given a leaf-labelled tree with n leaves and $n-1$ internal nodes, a labelled history is an identification between the internal nodes and the set $\{1, 2, \ldots, n-1\}$, such that the label assigned to a given node is strictly greater than the labels assigned to its descendants. We generalize the concept of labelled histories to r-furcating trees. Consider a rooted labelled tree in which each internal node has exactly r children, $r \geq 2$. We first generalize the enumeration of labelled histories for a bifurcating tree ($r = 2$) to enumerate labelled histories for an r-furcating tree with arbitrary $r \geq 2$. We formulate a conjecture for the rooted unlabelled r-furcating tree shape on n internal nodes whose labelled topologies have the most labelled histories. Finally, we enumerate labelled histories for r-furcating trees in a setting that allows for simultaneous branchings. These results advance mathematical phylogenetic modelling by extending computations concerning fundamental features of bifurcating phylogenetic trees to a more general class of multifurcating trees.

This article is part of the theme issue '"A mathematical theory of evolution": phylogenetic models dating back 100 years'.

1. Introduction

Stochastic phylogenetic models describe the evolutionary process, beginning from a single lineage that eventually gives rise to many lineages [1–5]. Models tracing back to the foundational 1925 paper of Yule [6] consider stochastic processes for both the discrete structure of trees and the time intervals between branchings [7].

Edwards, in a pioneering article [8] that recognized Yule's contribution to phylogenetic modelling, and Harding, in a detailed treatment that soon followed [9], studied the discrete structure of evolving phylogenetic trees separately from the time intervals separating divergences. Following Yule, they considered branchings as non-simultaneous bifurcations. Extinction is disallowed, and throughout the process, each extant lineage is equally likely to be the next to bifurcate. Conditioning on the appearance of n lineages at the end of the process, each of the possible sequences of bifurcations is equally likely to describe the branching history.

The Yule or Yule–Harding model for the discrete branching structure—or sometimes the Yule–Harding–Kingman model, after its use in Kingman's coalescent process for population genetics [10]—has come to serve as a foundational model for the discrete structure of evolving trees. What is the probability that the evolutionary process produces a given topological relationship among n labelled lineages? With what probability does a tree of n lineages have a specified number of clades of a given size? What are the mean and variance of measures of 'tree balance'? Many such questions have been studied under the Yule–Harding model, providing an understanding of the features expected for evolutionary trees under simple assumptions [3,11,12].

THE ROYAL SOCIETY PUBLISHING

A central concept in the study of the discrete structures produced by evolutionary models is that of a *labelled history*. For a tree that has n leaves with distinct labels, a labelled history gives the topological relationship of the n leaves together with the sequence of branchings that has given rise to it [3, p. 47]. The labelled histories describe the discrete outcomes of an evolutionary model that reaches n labelled leaves. Probabilistic analysis of a phylogenetic model often involves a computation on the discrete probability space that takes the labelled histories as the set of possible outcomes—with the Yule–Harding model assuming that each outcome is equally likely. Combinatorial features of the labelled histories provide basic results for application of such models.

How many labelled histories exist with n leaves? How many share a specific labelled topology? What labelled topologies possess the largest number of labelled histories? The first question was answered by Edwards [8]. The second was answered by Brown [13] but reported earlier without a biological context [14, p. 67]. A solution to the third question was presented by Hammersley & Grimmett [15] based on a conjecture of Harding [9].

Biological scenarios often suggest the importance of models that extend beyond the assumption of non-simultaneous bifurcations [16]. Multifurcation is relevant to species radiations, in which many lineages can diverge simultaneously. For pathogen lineages, which can be transmitted from one to many hosts, multifurcation may describe descent more accurately than a sequence of bifurcations. Multifurcation is also suited to genealogies of marine invertebrates, in which some individuals possess large numbers of surviving offspring while others have none. Simultaneous bifurcations can be important in population-genetic models of large samples in small populations, in which they are too probable to ignore. Mathematical analysis of models with multifurcation and simultaneity can therefore contribute to understanding evolutionary descent in diverse biological scenarios.

Here, we study the combinatorics of labelled histories in settings with multifurcation and simultaneity. First, we allow r-furcations, supposing that each internal node is permitted to branch into precisely r descendants, $r \geq 2$. Next, we consider r-furcations with simultaneity, allowing multiple concurrent r-furcating divergences. We extend enumerative results that have been available only in select scenarios beyond non-simultaneous bifurcations [12,17,18]. We also conjecture that aspects of subtree sizes that give rise to the largest number of labelled histories in the bifurcating case naturally extend to scenarios with r-furcations.

2. Definitions

Definitions largely follow Steel [3] and King & Rosenberg [12]. We consider leaf-labelled, rooted trees T, in which leaves have unique labels. For a tree T, non-leaf nodes, including the *root node*, are *internal nodes*. The *labelled topology* of T is its topological structure together with its leaf labels. We indicate the number of leaves of T by $|T|$ and also by n. We also have occasion to refer to *unlabelled* topologies; the unlabelled topology of T is its topological structure with its leaf labels omitted.

A node u of T is *descended* from internal node v if v lies on the path from the root node to u. Node v is *ancestral* to u; trivially, a node is both ancestral to and descended from itself. We consider *r-furcating trees*, in which r is fixed and each internal node has exactly r immediate descendants, $r \geq 2$. Most mathematical phylogenetic studies assume $r = 2$. For an r-furcating tree T whose root has immediate subtrees $T_1, T_2, \ldots, T_r$, we write $T = T_1 \oplus T_2 \oplus \ldots \oplus T_r$.

Given a labelled topology for a rooted tree T with w internal nodes, a *labelled history* for T is a bijection f from $\{1, 2, \ldots, w\}$ to the internal nodes of T, so that if node u is descended from node v in T, then $f^{-1}(u) \leq f^{-1}(v)$. A labelled history can be viewed as the temporal sequence of internal nodes, with a convention in this study that the numbers assigned to nodes increase backward in time along genealogical lines. Note that w, the number of internal nodes, is $|T| - 1$ for bifurcating trees, and more generally, $w = (|T| - 1)/(r - 1)$ for r-furcating trees. Figure 1*a* shows two labelled histories for the same bifurcating labelled topology.

The classic Yule–Harding model for bifurcating trees assumes that each internal node occurs at a unique point in time. We consider r-furcating trees (figure 1*c*) with non-simultaneous r-furcations; we also allow *simultaneity*, in which multiple internal nodes share the same point in time (figure 1*b*,*d*). We use the term *event* to refer to a set of simultaneous internal nodes. If internal node u is descended from internal node v and $u \neq v$, then u and v are not part of the same event. With simultaneity, the temporal sequence of internal nodes that encodes a labelled history contains simultaneous nodes. For emphasis, we sometimes use the term *tie-permitting* to refer to labelled histories that allow simultaneity, as such labelled histories allow 'ties' in node times.

For a given scenario with a fixed number of leaves, a *maximally probable* labelled topology is a labelled topology whose number of labelled histories is greater than or equal to that of all other labelled topologies [19]. Because each labelling of an unlabelled topology by a distinct set of leaf labels gives rise to the same number of labelled histories, we use *unlabelled* topologies to indicate the maximally probable labelled topologies.

3. Results

We investigate three aspects of labelled histories in four settings. In particular, we examine (i) the number of labelled histories across all labelled topologies with a given number of leaves, (ii) the number of labelled histories for a specific labelled topology, including cases of fully symmetric trees and (iii) the characterization of maximally probable labelled topologies. The four settings are (i) bifurcating trees with non-simultaneous branching, (ii) bifurcating trees allowing simultaneous branching, (iii) r-furcating trees with non-simultaneous branching and (iv) r-furcating trees allowing simultaneous branching. Section 3(a) reviews known results. Section 3(b) reviews the main result of King & Rosenberg [12] on simultaneous bifurcation, adding one new result (proposition 5). Sections 3(c) and 3(d) contain the main new results of the study.

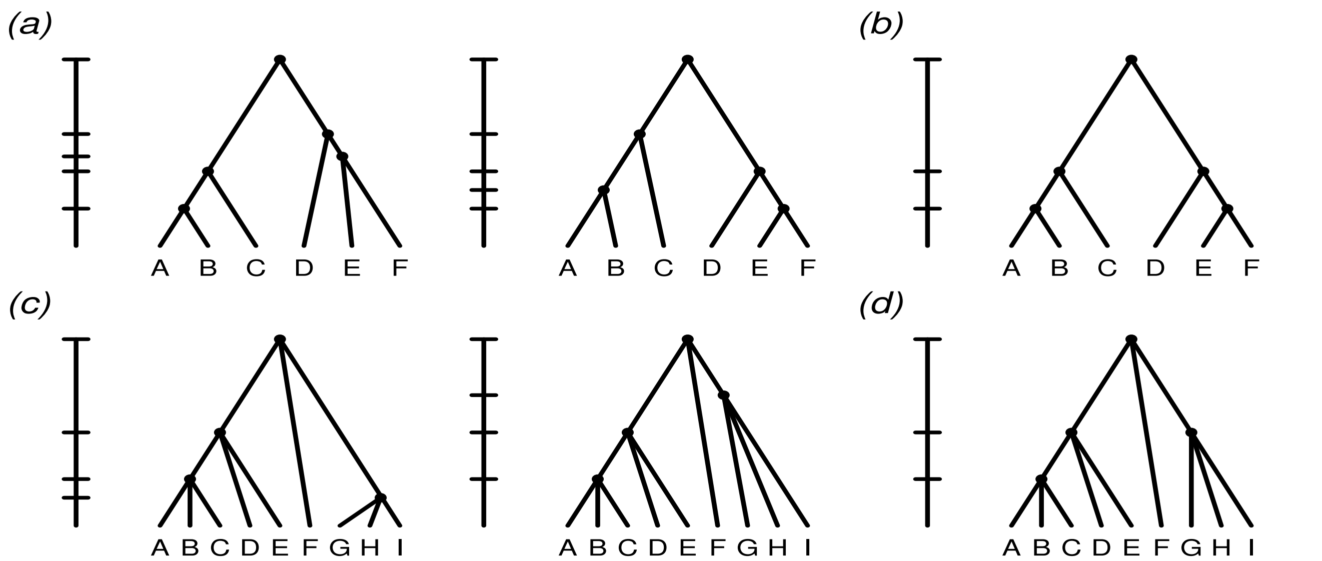

Figure 1. Labelled topologies. (*a*) Two labelled histories for the same bifurcating labelled topology. (*b*) A labelled history for the labelled topology in (*a*), permitting simultaneity. (*c*) Two labelled histories for the same trifurcating labelled topology. (*d*) A labelled history for the labelled topology in (*c*), permitting simultaneity. Tick marks in the bars adjacent to trees indicate the times of events. Using subscripts to label internal nodes from the leaves towards the root, the labelled histories are $(((A, B)_1, C)_2, (D, (E, F)_3)_4)_5$ and $(((A, B)_2, C)_4, (D, (E, F)_1)_3)_5$ in (*a*), $(((A, B)_1, C)_2, (D, (E, F)_1)_2)_3$ in (*b*), $(((A, B, C)_2, D, E)_3, F, (G, H, I)_1)_4$ and $(((A, B, C)_1, D, E)_2, F, (G, H, I)_3)_4$ in (*c*) and $(((A, B, C)_1, D, E)_2, F, (G, H, I)_2)_3$ in (*d*).

(a) Bifurcating trees, non-simultaneous branching

Consider bifurcating trees with non-simultaneous bifurcations and $n \geq 2$ leaves; we also include the trivial tree with $n = 1$ that consists of a single leaf.

(i) Total number of labelled histories

Let $A_2(n)$ denote the total number of labelled histories across all labelled topologies with n leaves. Trivially, $A_2(1) = 1$. To count labelled histories, we proceed backward in time from the n lineages. For $n \geq 2$, each of $\binom{n}{2}$ pairs can be the first pair to coalesce, leaving $n-1$ lineages. Recursively, $A_2(n) = \binom{n}{2} A_2(n-1)$. After the first coalescence, each of $\binom{n-1}{2}$ pairs can be next to coalesce, and so on, until only a single pair of lineages remains.

Proposition 1 ([8]). *Permitting only non-simultaneous bifurcations, the total number of labelled histories on n leaves, $A_2(n)$, satisfies $A_2(1) = 1$, and for $n \geq 2$,*

$$A_2(n) = \binom{n}{2}\binom{n-1}{2}\cdots\binom{3}{2}\binom{2}{2} = \frac{n!\,(n-1)!}{2^{n-1}}.$$

Sequence $A_2(n)$ begins 1, 1, 3, 18, 180, 2700 and is sequence A006472 in OEIS, the On-Line Encyclopedia of Integer Sequences [20].

(ii) Number of labelled histories for a specific labelled topology

Next, consider a labelled topology T for a bifurcating tree with non-simultaneous bifurcations and n leaves. Denote by T_1 and T_2 the two immediate subtrees of the root of T, with $|T_1|$ and $|T_2|$ leaves, respectively. For $n \geq 2$, the number of labelled histories for T, $N(T)$, is obtained recursively:

$$N(T) = \binom{|T|-2}{|T_1|-1} N(T_1)\, N(T_2), \tag{3.1}$$

with $N(T) = 1$ for $n = 1$. To obtain a non-recursive formula, define $V^0(T)$ as the set of internal nodes of T, including the root. Furthermore, define for internal nodes v of T the function $m : V^0(T) \to \mathbb{N}$, with $m(v)$ denoting the number of leaves contained in the subtree rooted at node v. By definition, $2 \leq m(v) \leq n$. We expand the recursive equation (3.1) and multiply by $(n-1)/(n-1)$ so that the product in the denominator includes the root node.

Proposition 2 ([13]). *Permitting only non-simultaneous bifurcations, the number of labelled histories for a labelled topology T with n leaves, $N(T)$, satisfies $N(T) = 1$ for $n = 1$, and for $n \geq 2$,*

$$N(T) = \frac{(n-1)!}{\prod_{v \in V^0(T)} (m(v) - 1)}.$$

Let $T_{2,k}$ denote a fully symmetric bifurcating labelled topology with 2^k leaves, $k \geq 1$: a tree in which each internal node has two subtrees with the same unlabelled topology. Denoting by $S_2(k)$ the number of labelled histories for $T_{2,k}$, equation (3.1) gives

$$S_2(k) = \binom{2^k - 2}{2^{k-1} - 1} S_2(k-1)^2.$$

With $S_2(1) = 1$, this recurrence can be solved; the solution can also be obtained from proposition 2 by noting that $T_{2,k}$ has 2^0 internal nodes with $m(v) = 2^k$, 2^1 internal nodes with $m(v) = 2^{k-1}$, and so on, up to 2^{k-2} internal nodes with $m(v) = 2^2$ and 2^{k-1} internal nodes with $m(v) = 2^1$. The sequence $S_2(k)$ is OEIS entry A056972.

Corollary 3 ([12]). *Permitting only non-simultaneous bifurcations, the number of labelled histories for a fully symmetric labelled topology $T_{2,k}$ with 2^k leaves satisfies*

$$S_2(k) = \prod_{j=2}^{k} \binom{2^j - 2}{2^{j-1} - 1}^{2^{k-j}} = \frac{(2^k - 1)!}{\prod_{j=2}^{k} (2^j - 1)^{2^{k-j}}}$$

for $k \geq 2$, with $S_2(1) = 1$.

(iii) Maximally probable labelled topologies

For labelled topologies with non-simultaneous bifurcations, let U_n^* denote the unlabelled topology on n leaves whose labellings give rise to the largest number of labelled histories. Recall the notation $T = T_1 \oplus T_2$ for a tree T whose root has subtrees T_1 and T_2.

Theorem 4 ([15]). *Permitting only non-simultaneous bifurcations, the unlabelled topology whose labellings have the largest number of labelled histories among unlabelled topologies with n leaves takes the form $U_n^* = U_t^* \oplus U_{n-t}^*$, where for $n \geq 3$,*

$$t = 2^{\lfloor \log_2(\frac{n-1}{3}) \rfloor + 1}.$$

The decomposition at the root contains trees of size t and $n - t$, where t is a specified power of 2. We have $(t, n-t) = (1,1)$ for the trivial $n = 2$, and for $n = 3$ to $n = 16$, we have $(t, n-t) =$ (1,2), (2,2), (2,3), (2,4), (4,3), (4,4), (4,5), (4,6), (4,7), (4,8), (8,5), (8,6), (8,7) and (8,8). For $n = 2$ to $n = 14$, these topologies can be viewed in figure 2.

Letting $G_2(n)$ denote the number of labelled histories for the unlabelled topology whose labellings have the largest number of labelled histories among unlabelled topologies with n leaves, the proof of theorem 4 finds that the value of t in the theorem maximizes the recursion $G_2(n) = [2/(n-1)] \max_{t \in \{1,2,\ldots n-1\}} [G_2(t)\, G_2(n-t)]$.

(b) Bifurcating trees, simultaneous branching

King & Rosenberg [12] introduced the study of labelled histories in the scenario that allows simultaneous bifurcations. The labelled histories with simultaneity allowed subsume the labelled histories without simultaneity.

(i) Total number of labelled histories

We provide a new result for the total number of labelled histories across all bifurcating labelled topologies, allowing simultaneous bifurcations, denoting this quantity $Y_2(n)$. Proceeding recursively backward in time, we choose i pairs of lineages to coalesce simultaneously in the first 'event', with $1 \leq i \leq \lfloor n/2 \rfloor$. The number of ways of choosing i pairs of lineages from among n lineages is $\frac{1}{i!}\prod_{j=0}^{i-1}\binom{n-2j}{2}$, as there are $\binom{n}{2}$ choices for the first pair, $\binom{n-2}{2}$ for the second pair, and so on, with $\binom{n-2(i-1)}{2}$ choices for the ith pair. The same i pairs can be chosen in any of $i!$ orders. Once the i pairs coalesce, we are left with $n - i$ lineages.

Proposition 5. *Permitting simultaneous bifurcations, the total number of labelled histories on n leaves, $Y_2(n)$, satisfies $Y_2(1) = 1$, and for $n \geq 2$,*

$$Y_2(n) = \sum_{i=1}^{\lfloor n/2 \rfloor} \left[\frac{1}{i!} \prod_{j=0}^{i-1} \binom{n-2j}{2} \right] Y_2(n-i) \tag{3.2}$$

$$= \sum_{i=1}^{\lfloor n/2 \rfloor} \frac{1}{i!} \left[\frac{n!}{2^i (n-2i)!} \right] Y_2(n-i). \tag{3.3}$$

If only the $i = 1$ term in the sum in equation (3.3) is tabulated—so that no simultaneous bifurcations are allowed—then the recursion reduces to the form in proposition 1, or $A_2(n) = \binom{n}{2} A_2(n-1)$.

The first terms in proposition 5 are $Y_2(1) = 1$, $Y_2(2) = 1$, $Y_2(3) = 3$, $Y_2(4) = 21$, $Y_2(5) = 255$ and $Y_2(6) = 4815$, compared with $A_2(1) = 1$, $A_2(2) = 1$, $A_2(3) = 3$, $A_2(4) = 18$, $A_2(5) = 180$ and $A_2(6) = 2700$. $Y_2(n)$ was previously obtained as the solution to a related problem in graph theory, appearing as OEIS sequence A317059 (see proposition 4.14 of [21]).

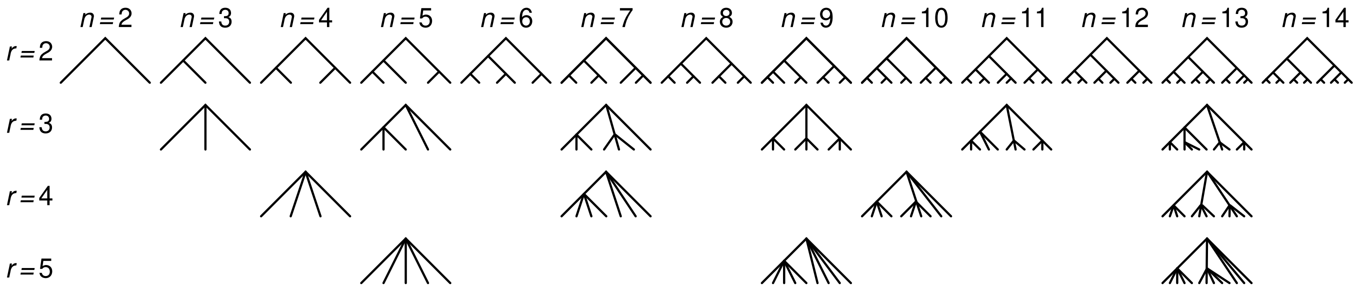

Figure 2. The maximally probable topology of an r-furcating tree on n leaves with non-simultaneous branching for small r and n. For bifurcating topologies, the maximally probable unlabelled shape is identified by theorem 4. More generally, for $r \geq 2$, the maximally probable unlabelled shapes shown are obtained by exhaustive computation, and they follow conjecture 13.

(ii) Number of labelled histories for a specific labelled topology

For a labelled topology T, King & Rosenberg [12] derived a recursion for the number of labelled histories, allowing simultaneous bifurcations. Let $N(T)$ denote the number of tie-permitting labelled histories for T, with $E(T, z)$ denoting the number that have exactly z events, or z separate times at which bifurcations occur. If simultaneity is disallowed, then $z = |T| - 1$.

Let $\delta(T)$ denote the 'height' of T, the maximum over leaves of T of the number of edges separating the root of T from the leaf. $\delta(T)$ is the minimal number of events in which the coalescences of topology T can occur. We then have $N(T) = \sum_{z=\delta(T)}^{|T|-1} E(T, z)$.

For a labelled topology T with at least two leaves, denote by T_1 and T_2 the labelled topologies for the two immediate subtrees of the root of T. Theorem 3 of [12] gives the number of tie-permitting labelled histories for T. The theorem sums tie-permitting labelled histories across all possible numbers of events z; for each z, enumeration proceeds recursively from the root, tabulating ways that the left and right subtrees can coalesce with a total of z events.

Theorem 6 ([12]). *Permitting simultaneous bifurcations, the number of labelled histories for a labelled topology T with n leaves, N(T), satisfies*

$$N(T) = \sum_{z=\delta(T)}^{|T|-1} E(T, z).$$

The number of tie-permitting labelled histories $E(T, z)$ satisfies

(i) *If T is a labelled topology with 1 leaf, then $E(T, 0) = 1$ and $E(T, z) = 0$ for $z \neq 0$.*
(ii) *If $|T_1| = 1$ and $|T_2| = 1$, then $E(T, 1) = 1$ and $E(T, z) = 0$ for $z \neq 1$.*
(iii) *If at least one of $|T_1|, |T_2|$ exceeds 1, then*

$$E(T, z) = \sum_{a = \max(\delta(T_1), z - |T_2|)}^{\min(|T_1| - 1, z-1)} \; \sum_{b = \max(\delta(T_2), z - a - 1)}^{\min(|T_2| - 1, z-1)} F(z, a, b)\, E(T_1, a)\, E(T_2, b),$$

where

$$F(z, a, b) = \binom{z-1}{(z-1)-b,\ (z-1)-a,\ a+b-(z-1)}.$$

(c) r-furcating trees, non-simultaneous branching

We now generalize to r-furcating trees, where $r \geq 2$ is a fixed value. First, notice that a tree in which each internal node has r immediate descendants must have $n = 1 + w(r-1)$ leaves, for some non-negative integer w, which again counts the number of internal nodes ($w = 0$ is the case of a single leaf). This result can be understood by noticing that beginning with a single r-furcating root, each sequential replacement of a leaf with an internal node adds $r - 1$ additional leaves.

(i) Total number of labelled histories

Let $A_r(n)$ denote the total number of labelled histories across all r-furcating trees with non-simultaneous r-furcations. As in the bifurcating case, $A_r(1) = 1$, and for larger values of n, $A_r(n) = \binom{n}{r} A_r(n - r + 1)$, as each of $\binom{n}{r}$ sets of r can be the first to coalesce.

Proposition 7. *For $w \geq 0$, let $n = 1 + w(r-1)$. Permitting only non-simultaneous r-furcations, the total number of labelled histories on n leaves, $A_r(n)$, satisfies $A_r(1) = 1$, and for $n \geq r$,*

$$A_r(n) = \binom{n}{r}\binom{n-r+1}{r}\cdots\binom{3r-2}{r}\binom{2r-1}{r}\binom{r}{r} = \frac{n!}{(r!)^w}\prod_{i=1}^{w}\left[n-(r-1)i\right].$$

royalsocietypublishing.org/journal/rstb *Phil. Trans. R. Soc. B* **380**: 20230307

Table 1 provides the total numbers of labelled histories for small r and small n. For $r = 2$, the proposition recovers the value in proposition 1. For $r = 3$, $A_r(n)$ simplifies to

$$A_3(n) = \frac{n!}{6^{(n-1)/2}} \prod_{i=1}^{(n-1)/2} (n - 2i) = \frac{2\,n!\,(n-2)!}{12^{(n-1)/2}\,(\frac{n-3}{2})!}, \tag{3.4}$$

The expression corresponds to OEIS sequence A339411.

More generally, note that $\prod_{i=1}^{w} [n - (r-1)i] = (r-1)^w \prod_{i=1}^{w} \left(\frac{n}{r-1} - i\right) = (r-1)^w \Gamma\left(\frac{n}{r-1}\right) / \Gamma\left(\frac{1}{r-1}\right)$, so that

$$A_r(n) = \frac{n!\,(r-1)^w\,\Gamma\left(\frac{n}{r-1}\right)}{(r!)^w \Gamma\left(\frac{1}{r-1}\right)}.$$

(ii) Number of labelled histories for a specific labelled topology

We next count labelled histories for a labelled topology T of an r-furcating tree with non-simultaneous r-furcations. For a tree with $|T| = n$ leaves, starting with a single leaf, each replacement of a leaf by an internal node with r descendants adds $r - 1$ leaves, so that $|T| = 1 + w(r-1)$, or

$$w = \frac{|T| - 1}{r - 1}. \tag{3.5}$$

Again, we count bijections between internal nodes and node ranks, with the constraint that the rank for an internal node must exceed those of its descendants. First, $N(T) = 1$ for $|T| = 1$. For $|T| > 1$, the root is assigned rank $(|T| - 1)/(r-1)$. A multinomial coefficient gives us the number of ways of allocating ranks $1, 2, \ldots, (|T| - 1)/(r-1) - 1$ to the r subtrees $T_1, T_2, \ldots, T_r$—which possess $(|T_1| - 1)/(r-1), (|T_2| - 1)/(r-1), \ldots, (|T_r| - 1)/(r-1)$ internal nodes, respectively. Multiplying by the numbers of labelled histories for the r subtrees, we obtain

$$N(T) = \binom{\frac{|T|-1}{r-1} - 1}{\frac{|T_1|-1}{r-1}, \frac{|T_2|-1}{r-1}, \ldots, \frac{|T_r|-1}{r-1}} N(T_1)\,N(T_2) \cdots N(T_r). \tag{3.6}$$

Next, $N(T_1), N(T_2), \ldots, N(T_r)$ can each be defined by the same recursive decomposition of their r immediate subtrees. We continue the decomposition until we reach subtrees with 1 leaf, obtaining a product of multinomial coefficients, one for each internal node of T. With $V^0(T)$ as the set of all internal nodes of T and $m(v)$ as the number of leaves for the subtree with root v, we multiply by an additional term $\frac{|T|-1}{r-1} / \frac{|T|-1}{r-1} = \frac{n-1}{r-1} / \frac{n-1}{r-1}$.

Proposition 8. *Permitting only non-simultaneous r-furcations, the number of labelled histories for a labelled topology T with n leaves, N(T), satisfies* $N(T) = 1$ *for* $n = 1$, *and for* $n = 1 + w(r-1)$ *with* $w \geq 1$,

$$N(T) = \frac{\left(\frac{n-1}{r-1}\right)!}{\prod_{v \in V_0(T)} \left(\frac{m(v)-1}{r-1}\right)}.$$

With $r = 2$, we recover proposition 2. That proposition 2 generalizes to multifurcation had been noted by Semple & Steel [17, p. 23], referencing a result of Stanley [22, p. 312] in the context of partially ordered sets.

Let $T_{r,k}$ denote a fully symmetric r-furcating labelled topology with r^k leaves, $k \geq 1$: a tree in which each internal node has r subtrees with the same unlabelled topology. Denoting by $S_r(k)$ the number of labelled histories for $T_{r,k}$, we can use the recursive equation 3.6 or the closed-form proposition 8 to obtain $S_r(k)$. In applying proposition 8, we have r^0 internal nodes with $m(v) = r^k$, r^1 internal nodes with $m(v) = r^{k-1}$, and so on, until r^{k-2} internal nodes with $m(v) = r^2$ and r^{k-1} internal nodes with $m(v) = r^1$.

Corollary 9. *Permitting only non-simultaneous r-furcations, the number of labelled histories for a fully symmetric labelled topology* $T_{r,k}$ *with* r^k *leaves satisfies*

$$S_r(k) = \frac{\left(\frac{r^k-1}{r-1}\right)!}{\prod_{j=2}^{k} \left(\frac{r^j-1}{r-1}\right)^{r^{k-j}}}$$

for $k \geq 2$, *with* $S_r(1) = 1$.

Table 2 shows the values of $S_r(k)$ for small r and k. With $r = 2$, the corollary recovers corollary 3. Inserting $r = 3$ and $r = 4$, we obtain the sequences OEIS A273723 and OEIS A273725.

(iii) Maximally probable labelled topologies

With our general formula counting labelled histories for any r-furcating tree, we investigate the unlabelled topology on n leaves whose labelled topologies have the most labelled histories.

Table 1. The total numbers of labelled histories $A_r(n)$ for small r-furcating trees with non-simultaneous r-furcations. Values of n are obtained from $n = 1 + w(r - 1)$ and $A_r(n)$ is computed from proposition 7. $A_2(n)$ follows OEIS A006472, and $A_3(n)$ follows OEIS 339411.

	$r = 2$		$r = 3$		$r = 4$		$r = 5$	
w	n	$A_2(n)$	n	$A_3(n)$	n	$A_4(n)$	n	$A_5(n)$
1	2	1	3	1	4	1	5	1
2	3	3	5	10	7	35	9	126
3	4	18	7	350	10	7350	13	162 162
4	5	180	9	29 400	13	5 255 250	17	1 003 458 456
5	6	2700	11	4 851 000	16	9 564 555 000	21	20 419 376 121 144
6	7	56 700	13	1 387 386 000	19	37 072 215 180 000	25	1 084 881 453 316 380 720

Table 2. The number of labelled histories $S_r(k)$ for a fully symmetric r-furcating labelled topology $T_{r,k}$ with $n = r^k$ leaves and non-simultaneous r-furcations, with $k \geq 1$. Values are computed from corollary 9. Numbers in scientific notation are approximate. $S_2(k)$ follows OEIS A056972, $S_3(k)$ follows OEIS A273723, and $S_4(k)$ follows OEIS A273725.

	$r = 2$		$r = 3$		$r = 4$	
k	n	$S_2(k)$	n	$S_3(k)$	n	$S_4(k)$
1	2	1	3	1	4	1
2	4	2	9	6	16	24
3	8	80	27	7 484 400	64	3 892 643 213 082 624
4	16	21 964 800	64	3.542×10^{37}	256	1.117×10^{110}

We first discuss trifurcation ($r = 3$). Consider a trifurcating tree T with n leaves. For $n = 3$, only one unlabelled topology is possible for T. For any labelling of its leaves, only one labelled history is possible. Similarly, for $n = 5$, only one unlabelled topology is possible, and only one labelled history for a labelled topology with that unlabelled topology. For each $n = |T|$, the maximum is obtained by recursively maximizing equation (3.6):

$$N(T) = \binom{\frac{n-1}{2} - 1}{\frac{|T_1|-1}{2}, \frac{|T_2|-1}{2}, \frac{|T_3|-1}{2}} N(T_1)\, N(T_2)\, N(T_3). \tag{3.7}$$

For fixed $|T_1|$, $|T_2|$, $|T_3|$, this equation is maximal if T_1, T_2, T_3 are the maximally probable subtrees of sizes $|T_1|$, $|T_2|$, $|T_3|$. At what allocation of $|T_1|$, $|T_2|$, $|T_3|$ with $|T_1| + |T_2| + |T_3| = n$ is the product in equation (3.7) maximized?

Following Harding [9], we increment n by 2, starting with $n = 7$, considering each decomposition $(|T_1|, |T_2|, |T_3|)$ with $n = |T_1| + |T_2| + |T_3|$. The decompositions that numerically produce the maximally probable labelled topologies appear in table 3; the associated numbers of labelled histories follow OEIS sequence A178008. A leaf configuration with each subtree at the same power of 3 increments subtree by subtree to the next power of 3. We formalize a conjecture.

Consider trifurcating tree $T = T_1 \oplus T_2 \oplus T_3$. By equation (3.7), its number of labelled histories satisfies

$$\frac{N(T)}{\left(\frac{|T|-1}{2}\right)!} = \frac{2}{|T|-1} \frac{N(T_1)}{\left(\frac{|T_1|-1}{2}\right)!} \frac{N(T_2)}{\left(\frac{|T_2|-1}{2}\right)!} \frac{N(T_3)}{\left(\frac{|T_3|-1}{2}\right)!}. \tag{3.8}$$

Let $M(T) = N(T)/(\frac{|T|-1}{2})!$, so that equation (3.8) becomes $M(T) = \frac{2}{|T|-1} M(T_1)\, M(T_2)\, M(T_3)$. To find the tree T^* that maximizes $N(T)$ across all trifurcating trees with n leaves, we must solve the maximization

$$\max_{\{T_1, T_2, T_3 \,|\, |T_1| + |T_2| + |T_3| = n\}} \frac{2}{n-1} M(T_1)\, M(T_2)\, M(T_3).$$

At this point, it is convenient to recast the maximization over sets of subtrees as a maximization over positive integer vectors; the tree is recovered from the vectors by recursively noting that the subtrees of a maximally probable tree are each maximally probable for their size. The maximal number of labelled histories among trees of size $n = 1 + w(r - 1)$ leaves, $r = 3$, is obtained as

$$G_3(n) = \max_{\{t_1, t_2, t_3 \,|\, t_1, t_2, t_3 > 0,\ t_1 + t_2 + t_3 = n\}} \frac{2}{n-1} G_3(t_1)\, G_3(t_2)\, G_3(t_3),$$

where $G_3(1) = 1$ and $(t_1, t_2, t_3) = 1 + (r - 1)(w_1, w_2, w_3)$ for non-negative integers w_1, w_2, w_3.

Taking the logarithm of both sides, this maximization is equivalent to a minimization of $\phi(n) = -\log G_3(n)$:

$$\phi(n) = \min_{\{t_1, t_2, t_3 \,|\, t_1, t_2, t_3 > 0,\ t_1 + t_2 + t_3 = n\}} \left[\phi(t_1) + \phi(t_2) + \phi(t_3) + \log\left(\frac{n-1}{2}\right)\right].$$

Table 3. The division of a tree of n leaves into subtrees with sizes $|T_1|$, $|T_2|$ and $|T_3|$ leaves that, by numerical computation and by conjecture 12, produces the maximally probable labelled topology with n leaves. Numbers in scientific notation are approximate. The numbers of labelled histories follow OEIS A178008.

number of leaves (n)	number of labelled histories	leaf configuration ($\|T_1\|, \|T_2\|, \|T_3\|$)	number of leaves (n)	number of labelled histories	leaf configuration ($\|T_1\|, \|T_2\|, \|T_3\|$)
3	1	(1,1,1)	35	20 432 412 000	(17, 9, 9)
5	1	(3,1,1)	37	205 837 632 000	(19, 9, 9)
7	2	(3,3,1)	39	2 500 927 228 800	(21, 9, 9)
9	6	(3,3,3)	41	21 598 916 976 000	(23, 9, 9)
11	12	(5,3,3)	43	263 986 763 040 000	(25, 9, 9)
13	40	(7,3,3)	45	3 837 961 401 120 000	(27, 9, 9)
15	180	(9,3,3)	47	3.377×10^{16}	(27, 11, 9)
17	630	(9,5,3)	49	4.316×10^{17}	(27, 13, 9)
19	3360	(9,7,3)	51	6.658×10^{18}	(27, 15, 9)
21	22 680	(9,9,3)	53	7.283×10^{19}	(27, 17, 9)
23	113 400	(9,9,5)	55	1.122×10^{21}	(27, 19, 9)
25	831 600	(9,9,7)	57	2.045×10^{22}	(27, 21, 9)
27	7 484 400	(9,9,9)	59	2.603×10^{23}	(27, 23, 9)
29	38 918 880	(11,9,9)	61	4.612×10^{24}	(27, 25, 9)
31	302 702 400	(13,9,9)	63	9.580×10^{25}	(27, 27, 9)
33	2 918 916 000	(15,9,9)	65	1.188×10^{27}	(27, 27, 11)

This minimization takes the form of a recursive minimization related to a problem solved by Batty *et al*. [23]. We restate their corollary 5.3 as theorem 11. We expand the assertion in their equation 5.1 as lemma 10, indicating that each positive integer has a unique partition of a certain type; the partition has r terms that are scalar multiples of one of two consecutive powers of r, and possibly one more term bounded by the difference between the two multiples.

Lemma 10. *Choose positive integers $q, r \geq 2$. Let $k_0 = 0$ and $k_p = (q-1)r^{p-1}$ for integers $p \geq 1$, and let $j_p = k_{p+1} - k_p$ for $p \geq 0$. Each integer $n \geq q$ has a unique decomposition specified by integers (p, s, b) as $n = sk_p + (r-s)k_{p+1} + b$, where $1 \leq s \leq r$, $p \geq 0$ and $0 \leq b < j_p$.*

Proof. Existence is verified by finding a suitable decomposition. First, considering all s in $[1, r]$ and b in $[0, j_p)$, $sk_p + (r-s)k_{p+1} + b$ ranges in interval $I_0 = [0, (q-1)r)$ for $p = 0$ and in $I_p = \left[(q-1)r^p, (q-1)r^{p+1}\right)$ for $p \geq 1$. The intervals do not overlap, so that in seeking to decompose n, p is the unique solution of the inequality $(q-1)r^p \leq n < (q-1)r^{p+1}$, or $p = \left\lfloor \log_r\left(\frac{n}{q-1}\right)\right\rfloor$.

Next, noting $0 \leq b < (q-1)r^p - (q-1)r^{p-1}$, we find s by solving the inequality

$$s(q-1)r^{p-1} + (r-s)(q-1)r^p \leq n < s(q-1)r^{p-1} + (r-s)(q-1)r^p + [(q-1)r^p - (q-1)r^{p-1}].$$

The solution is also unique,

$$s = \left\lceil \frac{(q-1)r^{p+1} - n}{(q-1)(r^p - r^{p-1})} \right\rceil.$$

We confirm that $s \leq r$. Using $(q-1)r^p \leq n$, we see that

$$s \leq \left\lceil \frac{(q-1)r^{p+1} - (q-1)r^p}{(q-1)(r^p - r^{p-1})} \right\rceil = \lceil r \rceil = r,$$

as is needed. Finally, b is uniquely specified as $b = n - [s(q-1)r^{p-1} + (r-s)(q-1)r^p]$. ∎

Theorem 11 ([23]). *Choose positive integers $q, r \geq 2$. Consider a recursion of the form*

$$f(n) = \begin{cases} g(n), & \text{if } n < q, \\ \min\left\{\sum_{i=1}^r f(a_i)\right\} + g(n), & \text{if } n \geq q. \end{cases}$$

where the minimum is over r-tuples $a = (a_1, a_2, \ldots, a_r)$ of integers a_i with $0 \leq a_i < n$ and $\sum_{i=1}^r a_i = n$.

Consider the special r-tuple $\sigma(n)$ formed by the unique decomposition (p, s, b) of $n \geq q$, with

$$\sigma_i(n) = \begin{cases} k_p, & 1 \leq i < s, \\ k_p + b, & i = s, \\ k_{p+1}, & s < i \leq r, \end{cases}$$

where $0 \leq b < k_{p+1} - k_p$. *If* g *is strictly increasing and strictly concave on each interval* $k_p \leq n \leq k_{p+1}$ *for* $p \geq 0$ *and* $g(0) \geq 0$, *then the minimum of* $f(n)$ *is obtained by* $a = \sigma(n)$.

Theorem 11 does not precisely apply to $\phi(n)$ in the role of $f(n)$, as our scenario has a number of differences; in particular, the theorem requires that $f(n)$ be defined for non-negative integers, whereas $\phi(n)$ is defined only for $n = 1 + w(r-1)$; in addition, the theorem allows the a_i to range over all non-negative integers less than n, whereas only values $1 + w(r-1)$ are permissible for the t_i. Nevertheless, we find that the minimum of $\phi(n)$ obtained numerically in table 3 accords with the location of $\sigma(n)$.

Following theorem 11, with $r = 3$, we construct $\sigma(n)$: $k_0 = 0$, $k_p = 3^{p-1}$ for $p \geq 1$, and

$$j_p = \begin{cases} 1, & p = 0, \\ 3^p - 3^{p-1}, & p \geq 1. \end{cases}$$

By lemma 10, each integer $n \geq 2$ has a unique decomposition as $n = sk_p + (3-s)k_{p+1} + b$, where $1 \leq s \leq 3$, $p \geq 0$ and $0 \leq b < j_p$. Vector σ calculated from theorem 11 gives

$$\sigma_i(n) = \begin{cases} k_p, & 1 \leq i < s, \\ k_p + b, & i = s, \\ k_{p+1}, & s < i \leq 3. \end{cases} \tag{3.9}$$

The vector σ accords with table 3. For example, $n = 57$ has decomposition $57 = 2 \cdot 9 + 1 \cdot 27 + 12$, with $p = 3$, $s = 2$, $b = 12$, $k_p = 9$ and $k_{p+1} = 27$. We obtain $\sigma_1(57) = 9$, $\sigma_2(57) = 21$ and $\sigma_3(57) = 27$, which, when rearranged, gives the entry for $n = 57$ in table 3.

Recall the notation $\oplus$ to indicate subtrees of a tree T; for an unlabelled topology U that consists of k subtrees with identical unlabelled topology $U_1 = U_2 = \ldots = U_k$, we simplify to $U = kU_1 = U_1 \oplus U_1 \oplus \ldots \oplus U_1$. Note that empty subtrees are ignored in the notation, so that, for example, $T_1 \oplus T_2 \oplus \varnothing = T_1 \oplus T_2$. Also recall that U_n^* denotes the unlabelled topology on n leaves whose labellings produce the largest number of labelled histories. We use the decomposition σ to provide a conjecture for the form of U_n^*.

Conjecture 12. *Permitting only non-simultaneous trifurcations, the unlabelled topology whose labellings have the largest number of labelled histories among unlabelled topologies with* $n \geq 3$ *leaves takes the form* $U_n^* = (s-1)U_{3^{p-1}}^* \oplus U_{3^{p-1}+b}^* \oplus (3-s)U_{3^p}^*$, *where*

$$\begin{aligned} p &= \lfloor \log_3 n \rfloor, \\ s &= \left\lceil \frac{3^{p+1} - n}{3^p - 3^{p-1}} \right\rceil, \\ b &= n - \left(s \cdot 3^{p-1} + (3-s)3^p\right). \end{aligned}$$

Generalizing to r-furcating trees, we conjecture that the maximally probable tree is decomposed at the root into subtrees of sizes $(r^{p-1}, \ldots, r^{p-1}, r^{p-1} + b, r^p, \ldots, r^p)$.

Conjecture 13. *Permitting only non-simultaneous r-furcations, the unlabelled topology whose labellings have the largest number of labelled histories among unlabelled topologies with* $n \geq r$ *leaves takes the form* $U_n^* = (s-1)U_{r^{p-1}}^* \oplus U_{r^{p-1}+b}^* \oplus (r-s)U_{r^p}^*$, *where*

$$\begin{aligned} p &= \lfloor \log_r n \rfloor, \\ s &= \left\lceil \frac{r^{p+1} - n}{r^p - r^{p-1}} \right\rceil, \\ b &= n - \left(sr^{p-1} + (r-s)r^p\right). \end{aligned}$$

If $r = 2$, then we can see that this conjecture recovers theorem 4. For $n \geq 2$, the conjecture gives $p = \lfloor \log_2 n \rfloor$, $s = \lceil \frac{2^{p+1} - n}{2^p - 2^{p-1}} \rceil = 4 - \lfloor \frac{n}{2^{p-1}} \rfloor$ and $b = n - \left(s \cdot 2^{p-1} + (2-s)2^p\right) = n - (4-s)2^{p-1}$. First, suppose $k \geq 1$ and $2^k \leq n \leq 3 \times 2^{k-1} - 1$. We obtain $(p, s, b) = (k, 2, n - 2^k)$, from which conjecture 13 gives $U_n^* = U_{2^{k-1}}^* \oplus U_{n-2^{k-1}}^*$. Next, suppose $k \geq 1$ and $3 \times 2^{k-1} \leq n \leq 2^{k+1} - 1$. We obtain $(p, s, b) = (k, 1, n - 3 \times 2^{k-1})$, from which the conjecture gives $U_n^* = U_{n-2^k}^* \oplus U_{2^k}^*$.

In theorem 4, we begin with the trivial $U_2^* = U_1^* \oplus U_1^*$ and $U_3^* = U_1^* \oplus U_2^*$; for $k \geq 2$, suppose $2^k \leq n \leq 3 \times 2^{k-1}$. We obtain $\lfloor \log_2(\frac{n-1}{3}) \rfloor = k - 2$, from which the quantity t in the theorem is $t = 2^{k-1}$ and $U_n^* = U_{2^{k-1}}^* \oplus U_{n-2^{k-1}}^*$. If $3 \times 2^{k-1} + 1 \leq n \leq 2^{k+1} - 1$, then we have $\lfloor \log_2(\frac{n-1}{3}) \rfloor = k - 1$, producing $t = 2^k$ and $U_n^* = U_{2^k}^* \oplus U_{n-2^k}^*$. Although the intervals of fixed $\lfloor \log_2(\frac{n-1}{3}) \rfloor$ defined by the construction in theorem 4 differ from the intervals defined by the floor and ceiling functions in conjecture 13, they lead to the same decomposition of a tree into subtrees.

(d) r-furcating trees, simultaneous branching

We next proceed to investigate r-furcating trees with simultaneous branching.

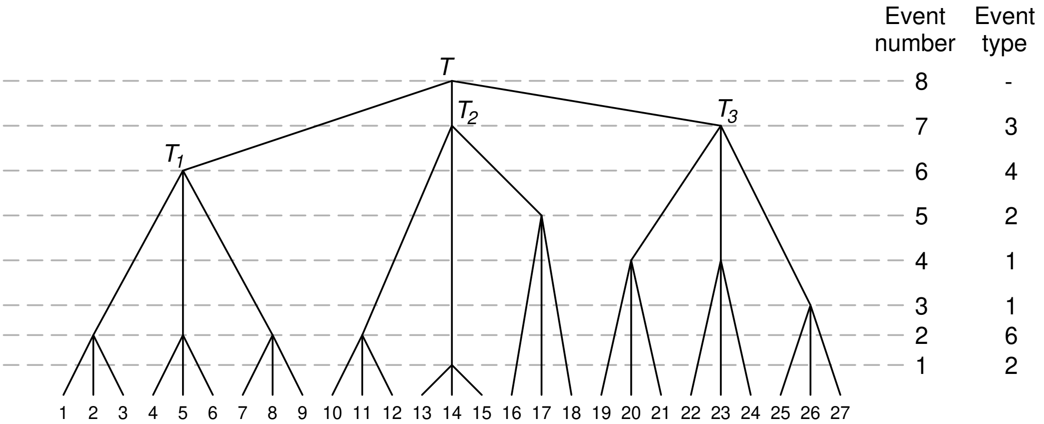

Figure 3. Counting labelled histories with simultaneity. A 27-leaf trifurcating labelled topology T is depicted; the root of T is indicated, as are the roots of subtrees T_1, T_2 and T_3. Events are numbered backward in time; dashed lines indicate time points and simultaneity. Each event is assigned a type based on simultaneity across T_1, T_2, T_3. For a 27-leaf trifurcating tree, the number of possible events z ranges from $\delta(T) = 3$ to $(|T| - 1)/(r - 1) = 13$; $z = 8$ is shown. Given z, $E(T, z)$ is calculated by summing labelled histories across counts (a_1, a_2, a_3) of the numbers of events in subtrees (T_1, T_2, T_3); $(a_1, a_2, a_3) = (2,4,3)$ is depicted. With z and (a_1, a_2, a_3) fixed, the sum proceeds over compositions $C(z + 2^r - 2, 2^r - 1)$; composition $c = (3,3,2,2,1,2,1)$ of 14 into 7 parts is shown, or $c^* = (2,2,1,1,0,1,0)$. Thus, for example, $c_6^* = 1$ indicates a simultaneous event between subtrees T_1 and T_2, as the binary representation 110_2 for 6 has a 1 in in the leftmost digit (T_1), a 1 in the middle digit (T_2) and a 0 in the rightmost digit (T_3). The labelled history depicted of $a_1 = 2$ events in subtree T_1 is the only arrangement of events in T_1 with $a_1 = 2$; the labelled history in subtree T_2 is one of six arrangements with $a_2 = 4$, and that in subtree T_3 is one of six arrangements with $a_3 = 3$. Once the simultaneity configuration c^* is chosen, the number of ways of assigning the $z - 1 = 7$ non-root events of T to its event types is $\binom{7}{2,2,1,1,0,1,0} = 1260$. The total number of labelled histories for the given z, (a_1, a_2, a_3), and c^* equals (1)(6)(6)(1260) = 45360.

Table 4. The total numbers of labelled histories $Y_r(n)$ for small r-furcating trees with simultaneous r-furcations. Values of n are obtained from $n = 1 + w(r - 1)$ and $Y_r(n)$ is computed from proposition 14. $Y_2(n)$ follows A317059.

	$r = 2$		$r = 3$		$r = 4$		$r = 5$	
w	n	$Y_2(n)$	n	$Y_3(n)$	n	$Y_4(n)$	n	$Y_5(n)$
1	2	1	3	1	4	1	5	1
2	3	3	5	10	7	35	9	126
3	4	21	7	420	10	8925	13	198 198
4	5	255	9	43 960	13	8 033 025	17	1 552 358 808
5	6	4815	11	9 347 800	16	19 010 866 875	21	41 269 930 621 920
6	7	130 095	13	3 513 910 400	19	97 622 651 251 125	25	2 917 021 792 126 858 416

(i) Total number of labelled histories

We extend our results for labelled histories permitting simultaneous bifurcations (proposition 5) and non-simultaneous r-furcations (proposition 7) to count labelled histories permitting simultaneous r-furcations.

Proposition 14. *For $w \geq 0$, let $n = 1 + w(r - 1)$. Permitting simultaneous r-furcations, the total number of labelled histories on n leaves, $Y_r(n)$, satisfies $Y_r(1) = 1$, and for $n \geq r$,*

$$Y_r(n) = \sum_{i=1}^{\lfloor n/r \rfloor} \left[\frac{1}{i!} \prod_{j=0}^{i-1} \binom{n - rj}{r} \right] Y_r(n - (r-1)i)$$
$$= \sum_{i=1}^{\lfloor n/r \rfloor} \frac{1}{i!} \left[\frac{n!}{(r!)^i (n - ri)!} \right] Y_r(n - (r-1)i).$$

Proof. We choose i groups of r lineages to coalesce simultaneously in the first event, with $1 \leq i \leq \lfloor n/r \rfloor$. The number of ways of choosing i groups of r lineages is $\frac{1}{i!}\prod_{j=0}^{i-1}\binom{n-rj}{r}$. The remaining $n - (r - 1)i$ lineages have $Y_r(n - (r - 1)i)$ labelled histories. ∎

Table 4 gives the total numbers of labelled histories for small r and the smallest permissible n, starting at $Y_r(1) = Y_r(r) = 1$, and

$$Y_r(2r - 1) = \frac{(2r-1)!}{r!\,(r-1)!}$$
$$Y_r(3r - 2) = \frac{(2r-1)(3r-2)!}{r!^2\,(r-1)!} + \frac{(3r-2)!}{2\,r!^2\,(r-2)!}.$$

royalsocietypublishing.org/journal/rstb Phil. Trans. R. Soc. B 380: 20230307

For $r = 2$, proposition 14 recovers proposition 5. For non-simultaneous multifurcation, $i = 1$, and proposition 14 recovers the recursion $A_r(n) = \binom{n}{r} A_r(n - r + 1)$ that underlies proposition 7.

(ii) Number of labelled histories for a specific labelled topology

We next count labelled histories for labelled topology T of an r-furcating tree with simultaneous r-furcations, generalizing the result for simultaneous bifurcations from theorem 6.

Recall that $E(T, z)$ counts tie-permitting labelled histories of labelled topology T with z events. Across all r-furcating trees T, the maximal number of events is $(|T| - 1)/(r - 1)$, if each internal node occurs at a distinct time point. $E(T, (|T| - 1)/(r - 1))$ produces the result of proposition 8, the number of labelled histories for T if only non-simultaneous r-furcations are permitted.

We will also need the minimum number of events permissible for an r-furcating tree T. This quantity is the height $\delta(T)$, as the longest path from a leaf to the root contains $\delta(T)$ internal nodes; internal nodes that do not lie on the longest path can all be made simultaneous with internal nodes that do lie on that path. Hence, for an r-furcating tree T, the number of events, z, must satisfy

$$\delta(T) \le z \le \frac{|T| - 1}{r - 1}.$$

As in theorem 6, the number of labelled histories $N(T)$ satisfies $N(T) = \sum_{z=\delta(T)}^{(|T|-1)/(r-1)} E(T, z)$.

The 1-leaf tree has $\delta(T) = 0$, so that $E(T, 0) = 1$ and $E(T, z) = 0$ for $z \ne 0$. If T has r leaves, then $E(T, 1) = 1$ and $E(T, z) = 0$ for $z \ne 1$. For the non-trivial case, in which at least one of $|T_1|, |T_2|, \ldots, |T_r|$ exceeds 1, suppose that for each j, $1 \le j \le r$, a_j distinct events occur in subtree T_j. Each a_j is bounded below by $\delta(T_j)$ and above by $(|T_j| - 1)/(r - 1)$. The events in different subtrees are not necessarily distinct. Indeed, considering the r-furcating tree T, there exist $2^r - 1$ possible sets that could be the set of subtrees in which a single point in time is associated with a collection of simultaneous nodes. To encode them, write the number k, $1 \le k \le 2^r - 1$, in binary, with r digits. Reading left to right, the jth digit, $1 \le j \le r$, indicates presence or absence of an event in subtree j. Event type k is an event with simultaneous r-furcations in all the subtrees with entries of 1 in its binary representation.

Recall that a composition of positive integer n into k parts, $1 \le k \le n$, is an ordered list of k positive integers whose sum is n. In a labelled history with simultaneity and z events, the $z - 1$ non-root events must each have one of the types $1, 2, \ldots, 2^r - 1$. The simultaneity configuration of a labelled history, counting the numbers of events of different types, can be encoded $c^* = (c_1 - 1, c_2 - 1, \ldots, c_{2^r-1} - 1)$, where $c = (c_1, c_2, \ldots, c_{2^r-1})$ is a composition of $(z - 1) + (2^r - 1)$ into $2^r - 1$ positive integer parts. Subtracting 1 element-wise to obtain c^* from c, with $c_k^* = c_k - 1$ for each k, we decompose $z - 1$ into $2^r - 1$ non-negative integers.

Write $I(c, j) = \sum_{k=1}^{2^r-1} c_k^* f(k, j)$, where $f(k, j) = 1$ if the r-digit binary representation of k has a 1 in position j. $I(c, j)$ counts internal nodes of subtree j for a simultaneity configuration c^* encoded by composition c. Given that subtree j has a_j distinct events, for a composition c in the set $C(z + 2^r - 2, 2^r - 1)$ of compositions of $(z - 1) + (2^r - 1)$ into $2^r - 1$ parts, labelled histories with the simultaneity configuration c^* are possible if and only if for each j from 1 to r, $I(c, j) = a_j$.

For a simultaneity configuration c^* that has the specified numbers of subtree events $a_1, a_2, \ldots, a_r$, the number of labelled histories for subtree j is $E(T_j, a_j)$. Given the labelled histories for the subtrees, the number of ways of assigning the $z - 1$ non-root events of T across the $2^r - 1$ entries in the configuration is $\binom{z-1}{c_1^*, c_2^*, \ldots, c_{2^r-1}^*}$. We assign events in the $2^r - 1$ categories to the $z - 1$ sequential positions in the list of events; once events are assigned, the $a_1, a_2, \ldots, a_r$ subsequences follow based on the orders of events in the subtrees. The derivation is depicted in figure 3. We have obtained the following theorem.

Theorem 15. *Permitting simultaneous r-furcations, the number of labelled histories for a labelled topology T with n leaves, N(T), satisfies*

$$N(T) = \sum_{z=\delta(T)}^{\frac{|T|-1}{r-1}} E(T, z).$$

The number of tie-permitting labelled histories $E(T, z)$ satisfies

(i) *If T is a labelled topology with 1 leaf, then $E(T, 0) = 1$ and $E(T, z) = 0$ for $z \ne 0$.*
(ii) *If $|T_j| = 1$ for all j, $1 \le j \le r$, then $E(T, 1) = 1$ and $E(T, z) = 0$ for $z \ne 1$.*
(iii) *If $|T_j|$ exceeds 1 for at least one index j, $1 \le j \le r$, then*

$$E(T, z) = \sum_{a_1=\delta(T_1)}^{\min\left(\frac{|T_1|-1}{r-1}, z-1\right)} \sum_{a_2=\delta(T_2)}^{\min\left(\frac{|T_2|-1}{r-1}, z-1\right)} \cdots \sum_{a_r=\delta(T_r)}^{\min\left(\frac{|T_r|-1}{r-1}, z-1\right)} \sum_{c \in C(z+2^r-2, 2^r-1)} \prod_{j=1}^{r} [\![I(c, j) = a_j]\!] \times \left(\prod_{j=1}^{r} E(T_j, a_j) \right) \binom{z-1}{c_1^*, c_2^*, \ldots, c_{2^r-1}^*}.$$

$[\![\cdot]\!]$ denotes the Iverson bracket, equalling 1 if the statement in the brackets holds and 0 otherwise.

royalsocietypublishing.org/journal/rstb *Phil. Trans. R. Soc. B* **380:** 20230307

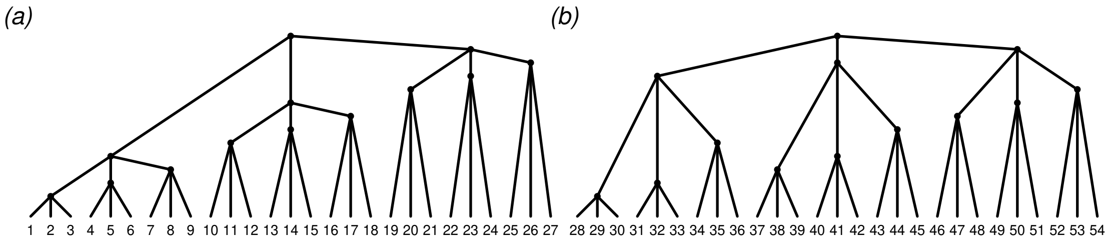

Figure 4. Two labelled histories for a 27-leaf labelled topology. (*a*) Labelled history for the 2022–2023 tournament of *Celebrity Jeopardy*. (*b*) Labelled history for the 2023–2024 tournament of *Celebrity Jeopardy*. Labels 1–54 represent distinct contestants.

Table 5. The number of labelled histories $L_r(k)$ for a fully symmetric r-furcating labelled topology $T_{r,k}$ with $n = r^k$ leaves and simultaneous r-furcations, with $k \geq 1$. Values are computed from theorem 15. Numbers in scientific notation are approximate.

	$r = 2$		$r = 3$		$r = 4$	
k	n	$L_2(k)$	n	$L_3(k)$	n	$L_4(k)$
1	2	1	3	1	4	1
2	4	3	9	13	16	75
3	8	365	27	308 682 013	64	3.017×10^{18}
4	16	1 323 338 487	81	2.044×10^{43}	256	9.402×10^{122}

The theorem collapses in two different ways to results we have already obtained. First, if only non-simultaneous r-furcations are allowed, then $z = (|T| - 1)/(r - 1)$ and $a_j = (|T_j| - 1)/(r - 1)$. All terms in the modified compositions c^* equal 0 except those corresponding to non-simultaneous coalescences, so that

$$N(T) = E\big(T, (|T| - 1)/(r - 1)\big) = \left(\prod_{j=1}^{r} E(T_j, a_j)\right)\binom{z-1}{a_1, a_2, \ldots, a_r}.$$

Noting that $N(T_j) = E(T_j, a_j)$, this equation accords with the recursive equation (3.6) and hence with the non-recursive proposition 8.

Second, if $r = 2$ and simultaneous bifurcations are allowed, then

$$E(T, z) = \sum_{a_1 = \delta(T_1)}^{\min\left(\frac{|T_1|-1}{r-1}, z-1\right)} \sum_{a_2 = \delta(T_2)}^{\min\left(\frac{|T_2|-1}{r-1}, z-1\right)} \sum_{c \in C(z+2,3)} [\![I(c,1) = a_1]\!][\![I(c,2) = a_2]\!] \; E(T_1, a_2)\, E(T_2, a_2) \binom{z-1}{c_1^*, c_2^*, c_3^*}.$$

Because $c_1^* + c_2^* + c_3^* = z - 1$ by construction, non-zero terms in the sum must have $I(c,1) = c_2^* + c_3^* = a_1$ and $I(c,2) = c_1^* + c_3^* = a_2$. Considering vectors (c_1^*, c_2^*, c_3^*) with $c_1^* + c_2^* + c_3^* = z - 1$, $c_2^* + c_3^* = a_1$ and $c_1^* + c_3^* = a_2$, one can obtain the same lower limits of summation as in theorem 6.

In the earlier theorem, we require that $a_1 + a_2 \geq z - 1$, as the sum of the numbers of events in the subsequences a_1 and a_2 for the two subtrees T_1 and T_2 is at least the number of non-root events $z - 1$ for the full tree T. Hence, $a_1 \geq (z-1) - a_2 \geq (z-1) - (|T_2| - 1) = z - |T_2|$; also, $a_2 \geq (z-1) - a_1$. Here, terms with $a_1 + a_2 < z - 1$ are not summed; to verify this fact, suppose such a term is summed as a nonzero quantity. A requirement for the Iverson brackets in the summation to both equal 1 is for $a_1 + a_2 = (c_2^* + c_3^*) + (c_1^* + c_3^*)$, so that $z - 1 + c_3^* < z - 1$ and $c_3^* < 0$, contradicting the fact that c_3^* is a non-negative count. Finally, the trinomial coefficient is $\binom{z-1}{(z-1)-a_1,\, (z-1)-a_2,\, a_1+a_2-(z-1)}$, verifying the equivalence to the earlier theorem.

We compute the numbers of labelled histories for fully symmetric trees with simultaneous r-furcations in table 5. The values quickly become substantially larger than the corresponding quantities in table 2 with only non-simultaneous r-furcations.

4. Discussion

We have enumerated labelled histories in a variety of settings, generalizing classic results on the enumeration of labelled histories with non-simultaneous bifurcations across all bifurcating labelled topologies (proposition 1) and for a specific bifurcating labelled topology (proposition 2). Our generalizations proceed in two directions. In particular, we allow simultaneity. In proposition 5, we count labelled histories, allowing for simultaneous bifurcations, across all bifurcating labelled topologies. In theorem 6, we report the recently derived number of labelled histories for a specific bifurcating labelled topology, allowing simultaneous bifurcations.

The other direction of generalization is to allow r-furcations. In proposition 7, we count labelled histories across all r-furcating labelled topologies, without simultaneity. In proposition 8, we report an earlier result counting labelled topologies for

Table 6. Summary of the main mathematical results.

	number of labelled histories across all possible labelled topologies	number of labelled histories for specific labelled topologies	labelled topology with the most labelled histories
bifurcating, non-simultaneous	proposition 1 [8], table 1, OEIS A006472	proposition 2 [13], table 2	theorem 4 [15], figure 2
bifurcating with simultaneity	proposition 5, table 4, OEIS A317059	theorem 6 [12], table 5	open problem
r-furcating, non-simultaneous	proposition 7, table 1	proposition 8, table 2	open problem, conjecture 13, figure 2
r-furcating with simultaneity	proposition 14, table 4	theorem 15, table 5	open problem

a specific r-furcating labelled topology, without simultaneity. In conjecture 13, we suggest a candidate for the r-furcating unlabelled topology whose labelled topologies have the largest number of labelled histories.

Finally, we consider both simultaneity and r-furcations, enumerating in proposition 14 the total number of labelled histories across all r-furcating labelled topologies, allowing simultaneity, and in theorem 15 the number of labelled histories for a specific r-furcating labelled topology, allowing simultaneity. The results are examined numerically in tables 1–5 and summarized in table 6.

Enumerative phylogenetic results, which appear in such domains as phylogenetic encodings, probabilistic computations on tree spaces and computational complexity calculations, have generally focused on non-simultaneous bifurcating trees. Simultaneity results are suited to settings in which time intervals are discretized, such as in discrete-generation population-genetic models with large samples in relation to the population size [24–26]. Recent mathematical–phylogenetic interest in multifurcation [18,27,28] has potential for application in settings such as pathogen transmission.

The results connect to other settings that use structures that correspond to labelled histories. In computer science, a binary heap is a complete binary tree whose nodes store keys according to a total order; in a *max-heap*, the key of a parent node must exceed the keys of its children. Binary heaps extend to r-ary heaps, where each parent has r children, so that labelled history enumerations have an analogous meaning in counting r-ary heaps. $S_2(k)$ in corollary 3, which counts labelled histories for fully symmetric bifurcating trees, counts binary heaps on k levels (OEIS A056972). Similarly, $S_3(k)$ and $S_4(k)$ (table 2) count labelled histories for fully symmetric trifurcating and quadfurcating trees—and associated ternary and quaternary heaps (OEIS A273723 and A273725). The connection expands the links between phylogenetic enumeration and associated structures in computer science [5].

As an example of labelled histories beyond phylogenetics, King & Rosenberg [12] counted game sequences for single-elimination sports tournaments. A bifurcating tree structure is specified, with teams at the leaves. Teams play games pairwise, the winner advancing in the tree until only one team remains. The analogy extends to multifurcation, in which games involve r players, with r not necessarily equal to 2. For $r = 3$, the television game show *Celebrity Jeopardy* provides an example. A tournament has 27 players who compete in groups of 3 (figure 4). In the 2022−2023 season, the tournament used a sequence in which the finalist of each subtree of nine players was determined before the next subtree competed (figure 4*a*, [29]). In 2023–2024, each player played one game before any player played two (figure 4*b*, [30]). If they had used the same players, then the two seasons would be possible to regard as two of 7 484 400 possibilities (table 2), or 308 682 013 if simultaneous matches are allowed (table 5).

As it is unlikely for biological lineages to always diverge in groups of exactly r lineages, a natural extension is to *at-most-r-furcating trees*, where a parent can have at least 2 and at most r children [28]. Analogous problems can be considered for the total number of labelled histories across all at-most-r-furcating trees and for specific at-most-r-furcating trees on n leaves, both for non-simultaneous and for simultaneous branching. A related setting examines at-most-n-furcating trees on n leaves [18] (see also [31, p. 30]). We have conjectured a characterization of maximally probable labelled topologies for r-furcating trees; this problem, and the characterization of the maximally probable topologies with simultaneity, for r-furcating trees and even for $r = 2$, remain open.

The Yule–Harding tree shape model has been important to the field of mathematical phylogenetics for decades. This work relaxes two of its assumptions: that divergences involve bifurcations and that they are non-simultaneous. Extended phylogenetic models can make use of our relaxed assumptions in order to accommodate r-furcations and simultaneity.

Ethics. This work did not require ethical approval from a human subject or animal welfare committee.

Data accessibility. This article has no additional data.

Declaration of AI use. We have not used AI-assisted technologies in creating this article.

Authors' contributions. E.H.D.: investigation and writing—original draft, writing—review and editing; N.A.R.: conceptualization, investigation, supervision and writing—original draft, writing—review and editing.

Both authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. Grant support was provided by NIH grant R01 HG005855.

Acknowledgements. We are grateful to Hsien-Kuei Hwang for pointing us to [24], and to Michael Fuchs for helpful conversations.

References

1. Aldous DJ. 2001 Stochastic models and descriptive statistics for phylogenetic trees, from Yule to today. *Stat. Sci.* **16**, 23–34. (doi:10.1214/ss/998929474)
2. Lambert A, Stadler T. 2013 Birth–death models and coalescent point processes: The shape and probability of reconstructed phylogenies. *Theor. Popul. Biol*. **90**, 113–128. (doi:10.1016/j.tpb.2013.10.002)
3. Steel M. 2016 *Phylogeny: discrete and random processes in evolution*. Philadelphia, PA: Society for Industrial and Applied Mathematics.
4. Wiehe T. 2021 Counting, grafting and evolving binary trees. In *Probabilistic structures in evolution* (eds E Bääke, A Wakolbinger), pp. 427–450. Berlin, Germany: EMS Press. (doi:10.4171/ecr/17-1/20)
5. Fuchs M. 2025 Shape parameters of evolutionary trees in theoretical computer science. *Phil. Trans. R. Soc. B* **380**, 20230304. (doi:10.1098/rstb.2023.0304)
6. Yule GU. 1925 II.—A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F. R. S. *Phil. Trans. R. Soc. Lond. B* **213**, 21–87. (doi:10.1098/rstb.1925.0002)
7. Pennell M, MacPherson A. 2025 Reading Yule in light of the history and present of macroevolution. *Phil. Trans. R. Soc. B* **380**, 20230299. (doi:10.1098/rstb.202.0299)
8. Edwards AWF. 1970 Estimation of the branch points of a branching diffusion process. *J. R. Stat. Soc. Ser. B* **32**, 155–174. (doi:10.1111/j.2517-6161.1970.tb00828.x)
9. Harding EF. 1971 The probabilities of rooted tree-shapes generated by random bifurcation. *Adv. Appl. Probab*. **3**, 44–77. (doi:10.1017/s0001867800037587)
10. Kingman JFC. 1982 The coalescent. *Stoch. Proc. Appl*. **13**, 235–248. (doi:10.1016/0304-4149(82)90011-4)
11. Fischer M, Herbst L, Kersting S, Kühn AL, Wicke K. 2023 *Tree balance indices: a comprehensive survey*. Berlin, Germany: Springer.
12. King MC, Rosenberg NA. 2023 A Mathematical Connection Between Single-Elimination Sports Tournaments and Evolutionary Trees. *Math. Mag.* **96**, 484–497. (doi:10.1080/0025570X.2023.2266389)
13. Brown JKM. 1994 Probabilities of evolutionary trees. *Syst. Biol*. **43**, 78–91. (doi:10.1093/sysbio/43.1.78)
14. Knuth DE. 1998 *The art of computer programming, volume 3: sorting and searching*, 2nd edn. Upper Saddle River, NJ: Addison-Wesley.
15. Hammersley J, Grimmett G. 1974 Maximal solutions of the generalized subadditive inequality. In *Stochastic geometry* (ed. EF Harding), pp. 270–285. London, UK: Wiley.
16. Eldon B. 2020 Evolutionary genomics of high fecundity. *Annu. Rev. Genet*. **54**, 213–236. (doi:10.1146/annurev-genet-021920-095932)
17. Semple C, Steel M. 2003 *Phylogenetics*. Oxford, UK: Oxford University Press.
18. Wirtz J. 2022 On the enumeration of leaf-labelled increasing trees with arbitary node-degree. *arXiv* 2211.03632. (doi:10.48550/arXiv.2211.03632)
19. Degnan JH, Rosenberg NA. 2006 Discordance of species trees with their most likely gene trees. *PLoS Genet*. **2**, 762–768. (doi:10.1371/journal.pgen.0020068)
20. OEIS Foundation Inc. 2024 The On-Line Encyclopedia of Integer Sequences. See https://oeis.org.
21. Dougherty AL, Mayers N, Short R. 2020 How to build a graph in *n* days: Some variations of graph assembly. *Discr. Appl. Math*. **284**, 20–28. (doi:10.1016/j.dam.2020.03.010)
22. Stanley RP. 1997 *Enumerative combinatorics volume 1*. Cambridge, UK: Cambridge University Press.
23. Batty CJK, Pelling MJ, Rogers DG. 1982 Some recurrence relations of recursive minimization. *SIAM J. Algebr. Discr. Methods* **3**, 13–29. (doi:10.1137/0603002)
24. Fu YX. 2006 Exact coalescent for the Wright–Fisher model. *Theor. Popul. Biol*. **69**, 385–394. (doi:10.1016/j.tpb.2005.11.005)
25. Bhaskar A, Clark AG, Song YS. 2014 Distortion of genealogical properties when the sample is very large. *Proc. Natl Acad. Sci. USA*. **111**, 2385–2390. (doi:10.1073/pnas.1322709111)
26. Melfi A, Viswanath D. 2018 Single and simultaneous binary mergers in Wright-Fisher genealogies. *Theor. Popul. Biol*. **121**, 60–71. (doi:10.1016/j.tpb.2018.04.001)
27. Mir A, Rotger L, Rosselló F. 2018 Sound Colless-like balance indices for multifurcating trees. *PLoS One* **13**, e0203401. (doi:10.1371/journal.pone.0203401)
28. Maranca ARP, Rosenberg NA. 2024 Bijections between the multifurcating unlabeled rooted trees and the positive integers. *Adv. Appl. Math*. **153**, 102612. (doi:10.1016/j.aam.2023.102612)
29. Sony Pictures. Celebrity Jeopardy! Season 1. See https://www.jeopardy.com/track/celebrity-jeopardy/season-1 (accessed 1 June 2024).
30. Sony Pictures. 2023 Celebrity Jeopardy! Season 2. See https://www.jeopardy.com/track/celebrity-jeopardy/season-2 (accessed 1 June 2024).
31. Felsenstein J. 2004 *Inferring phylogenies*. Sunderland, MA: Sinauer.