Branching architecture limits the rate of somatic mutation accumulation in trees

Frank Johannes[1]

[1] Technical University of Munich, Plant Epigenomics, Emil-Ramann-Str. 4, 85354, Freising, Germany; correspondence: frank@johanneslab.org

Trees are long-lived plants that develop complex, highly branched shoot systems as they grow. Their extended lifespan allows somatic mutations to accumulate along these branching structures, ultimately becoming fixed in reproductive and vegetative tissues such as leaves, flowers, and fruits. Mature trees can easily sustain tens of thousands of terminal branches, each potentially carrying mutated gametes. To avoid mutational meltdown and inbreeding depression, long-lived plants appear to have evolved mechanisms that slow mutation accumulation per unit time. How this is achieved remains unclear. Here, we show that branching architecture can limit mutation accumulation to the same extent as reducing the mutation rate itself. Tree structures that maximize branch path sharing during development constrain mutational diversity in the crown. These architectural factors can drive differences in mutation burden by orders of magnitude, even when mutation rates and terminal branch numbers are identical. Building on these insights, we show that current estimates of somatic mutation rates in trees are actually upward biased by a factor that scales with the topological characteristics of the tree. These results raise the deeper question whether somatic mutation rate differences, recently detected among tree species, reflect variation in branching strategies rather than variation in the rates themselves. It is possible that specific branching architectures have evolved not only to optimize resource allocation and structural stability but also to limit mutational load.

Keywords: somatic mutations, somatic epimutations, mutation rate, epimutation rate, life-history traits, genomics, branching, development.

Trees are long-lived organisms that develop complex, highly branched shoot systems over time (Fig.1A). In many species, mature individuals can sustain tens of thousands of terminal branches, each derived from a small cluster of precursor cells in the shoot apical meristem (SAM)—a self-renewing population of stem cells at the shoot apex (Schoen and Schultz (2019)). Repeated branching during growth creates cellular bottlenecks (Fig.2A), leading to somatic drift and the fixation of *de*

novo mutations in large clonal sectors of the crown (Chen et al. (2024), Tomimoto and Satake (2023), Iwasa et al. (2023)). Over time, these mutations accumulate along the branching structure, potentially affecting both vegetative functions, such as leaf/fruit physiology and stem integrity, and reproductive tissues that generate gametes for the next generation (Schoen and Schultz (2019), Quiroz et al. (2023)).

Long-lived plants appear to have evolved several strategies to mitigate the detrimental effects of mutation load and reduce the risk of inbreeding depression. These include low selfing rates and large effective population sizes, both of which help maintain genetic diversity and minimize inbreeding coefficients (Lesaffre (2021), Scofield and Schultz (2006), Schoen and Schultz (2019)). Beyond these population-level adaptations, emerging evidence points to intrinsic mechanisms that actively slow the accumulation of somatic mutations in long-lived plants (Lanfear et al. (2013)). This is apparent in the inverse relationship between mutation rates and generation time that has been detected in genomic analyses of diverse tree species (Johannes (2024), De La Torre et al. (2017)). Indeed, somatic mutation rates in long-lived trees, such as *O. fragrans*, can be about two orders of magnitute lower per year than in shorter-lived trees (e.g. *S. suchowensis*) (Johannes (2024)). These patterns extend beyond genetic mutations. Similar scaling has been observed at the epigenetic level, where epimutation rates—measured as stochastic gains and losses of DNA methylation—are significantly reduced in long-lived species. This parallel reduction suggests a shared underlying biological mechanism that affects both genetic and epigenetic mutation processes (Johannes (2024)). However, the exact nature of this mechanism remains unknown. One prominent hypothesis proposes that long-lived trees simply slow meristematic cell divisions per unit time. This would reduce errors arising from DNA synthesis and DNA methylation maintenance during replication. Although direct *in vivo* validation of this hypothesis remains challenging, it is strongly supported by the observation that somatic (epi)mutation rates correlate positively with growth rates and cell division frequency in vegetative tissues (Lanfear et al. (2013), Zhou et al. (2024)).

The link between growth dynamics and somatic mutation accumulation suggests a broader connection between molecular evolution and life-history traits. Since SAM-derived mutations are primarily incorporated into lateral organs during branch initiation, the spatial and temporal dynamics of branching may present a major rate limiting step for the accumulation of mutations in the tree crown - a possibility that has received little attention. Here, we present a theoretical analysis to investigate this relationship. We demonstrate that branching architectures and branch age distributions can substantially modulate somatic mutation accumulation, with effects comparable to reducing the mutation rate itself. Tree structures that maximize developmental path sharing—where many terminal branches trace back to a common meristematic lineage—tend to suppress mutational diversity. In some cases, these architectural features can reduce mutation burdens by orders of magnitude, even when baseline mutation rates and terminal branch numbers remain unchanged.

Our findings have major implications for understanding the rate and topological

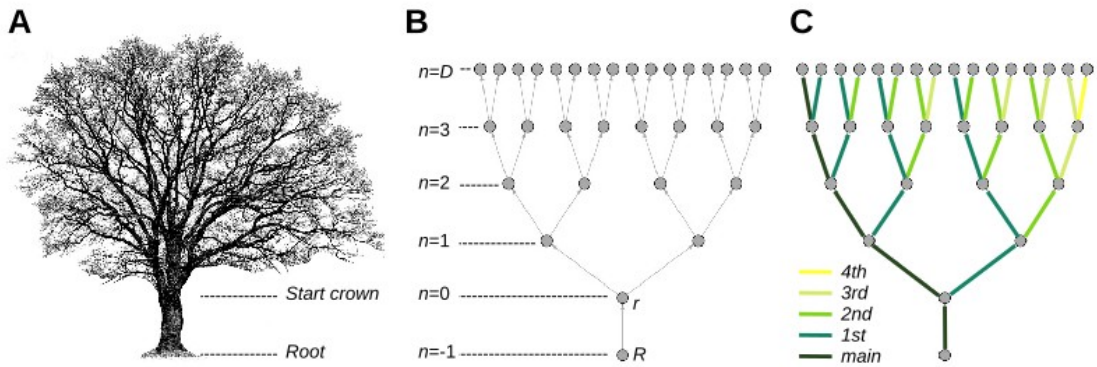

Figure 1: A graph presentation of the tree crown A. Trees develop complex, highly branched shoot systems as they grow. An example of an oak tree is shown. Mature specimen can sustain hundreds of thousands of terminal branches. **B.** The branching structure of a tree can be presented as a connected, acyclic graph with edges representing branches and nodes branch points. Node R represents the actual root of the tree and node r marks the start of the tree crown. We assume that the crown itself can be modeled as a *b-ary* tree, where each internal node has b children. For simplicity only a binary tree ($b = 2$) is shown, although the example picture in (A.) is approximately ternary ($b = 3$). The hierarchical structure of the tree can be quantified by its depth, D. It represents the maximum distance, measured in terms of branching levels, from the root to the terminal nodes. The branching levels are indexed as $n \in \{-1, 0, \ldots, D\}$, where $n = -1$ corresponds to R, $n = 0$ corresponds to r, and $n = D$ corresponds to the terminal nodes. **C.** Branch "order" is superimposed on the binary tree. It refers to the position of a branch in the hierarchy relative to the main stem. First order branches are those that split from the main stem; second order branches split from first order branches, third from second, and so on. Branch order indexing reveals that that there are unique branches starting at each level n that end in distinct terminal nodes at $n = D$, without switching branch order along the way.

distribution of somatic mutations in trees, and highlight life-history traits as key determinants of intra-organism somatic heterogeneity.

Results and discussion

A graph representation of the tree crown

Let us represent the hierarchically branching structure of a tree as a connected, acyclic graph $A = (V, E)$, where V is the set of nodes (i.e. branch points), and E is the set of directed edges (i.e. branches). The graph contains a root $R \in V$, which connects to node r by a single directed edge (e_{Rr})(Fig. 1B). Node r represents the lowest branch point of the tree and marks the start of the tree crown. We assume that the crown itself can be modeled as a *b-ary* tree, where each internal node has exactly b children. Hence, every other node $v \in V \setminus \{R\}$ has exactly one parent, ensuring a unique path from R to any other node. The hierarchical structure of the tree can be quantified by its depth, D. It represents the maximum distance, measured in terms of branching levels, from the root to the terminal nodes. The branching levels are indexed as $n \in \{-1, 0, \ldots, D\}$, where $n = -1$ corresponds to R, $n = 0$ corresponds to r (the lowest branch point in the crown), and $n = D$ corresponds to the terminal nodes (Fig. 1B). Hence, at the maximum level $n = D$,

the number of terminal nodes is b^D.

Branch order indexing and path counts

In a real tree, it is important to distinguish branch "order", which refers to the position of a branch in the hierarchy relative to the main stem. Consider an internal node along the main stem. One of its outgoing edges continues as the main stem, while the other $b-1$ outgoing edges are first-order lateral branches (Fig. 1C). Likewise, an internal node along a first-order lateral branch produces one outgoing edge that continues as a first-order lateral branch, while the others are second-order lateral branches, and so on. To keep track of this, each branch can be given a label p, determined recursively. For a node at level n, let k index the parent node within its level, and let i index the outgoing edge of the parent node, where $i = 0, 1, \ldots, b-1$. The branch order of the edge connecting a parent node k at level n to its i-th child node at level $n+1$ is defined as:

$$p(n+1, bk+i) = \begin{cases} p(n,k), & \text{if } i = 0, \\ p(n,k)+1, & \text{if } i > 0. \end{cases}$$

This recursive definition applies to all edges in the tree. At $n = -1$, the e_{Rr} edge is assigned $p(-1, 0) = 0$, marking it as the main stem. At $n = 0$, the first outgoing edge from r ($i = 0$) continues the main stem with $p = 0$, while the other outgoing edges ($i > 0$) initiate first-order lateral branches with $p = 1, 2, \ldots$. At all subsequent levels $n \geq 0$, the first edge of each internal node ($i = 0$) inherits the branch order of its parent (p), while the lateral branches ($i > 0$) increment p by 1. To ensure uniqueness, each edge is indexed by its parent and child node indices, n, k, and i, with the child node uniquely determined as $bk + i$.

This indexing reveals that there are unique branches starting at each level n that end in distinct terminal nodes at $n = D$, without switching branch order along the way (Fig. 1C). The count, C_n, of such unique branches is given by:

$$C_n = \begin{cases} 1, & \text{if } n = -1, \\ b^n - \sum_{k=-1}^{n-1} C_k, & \text{if } 0 \leq n < D, \\ 0, & \text{if } n = D. \end{cases}$$

Their corresponding lengths, measured in terms of the number of nodes (i.e. branch points) that are traversed from n to D, are:

$$B_n = \begin{cases} D - n, & \text{if } -1 \leq n < D, \\ 0, & \text{if } n = D. \end{cases}$$

These lengths are important, as each of the nodes along the branch represents a cellular bottleneck, where stem cells are "sampled" for the initiation of lateral branches.

The *realized* effective stem cell population size

Stem cells in the SAM have the capacity for self-renewal and for producing progenitor cells that differentiate (Heidstra and Sabatini (2014)). Live imaging and cell lineage tracing have identified a small number of resident stem cells at the very tip of the SAM surface that distinguish themselves from surrounding cells by their extremely slow division rates (Lyndon (1998), Kwiatkowska (2008)). These so-called apical stem cells (ASCs) are the origin of all SAM cell lineages, including those forming lateral organs and the germline (Burian et al. (2016)). The number of ASC cells, m, must be viewed as the *actual* effective stem cell population size of the plant, as all other cells are just clonal descendants. Estimates in Arabidopsis and tomato suggest that m is small, consisting of only 3-4 stem cells (Burian et al. (2016)), but this number may vary across species. Given the cell biological details of shoot formation (Burian et al. (2016), Shi et al. (2016)) as well as recent theoretical arguments (Johannes (2025)), it is reasonable to assume that only a single ASC is sampled via its clonal descendants to initiate the formation of lateral branches (Fig. 2A). The number of sampling events along a branch, defined by B_n, therefore affects the degree to which mutations that have arisen in ASCs over time will be "released" into terminal branches and leaves. To clarify this point, consider a single branch, starting at branch level n and ending in one of the terminal nodes at D. It can be shown that the expected number of *unique* ASC that are sampled along this branch is given by:

$$\dot{m}_n = m - m \cdot \left(1 - \frac{1}{m}\right)^{B_n}, \tag{1}$$

where $\dot{m}_n$ is the defined as the *realized* effective stem cell population size of branches of length B_n. The equation shows that $\dot{m}_n$ is smaller than m by a factor that scales with B_n (Fig. 2B). That is, as the number of sampling events increases, $\dot{m}_n$ converges to the *actual* effective stem cell population size (m):

$$\lim_{B_n \to \infty} \dot{m}_n = m.$$

Although this limiting case technically never occurs during ontogeny, values of B_n tend to be sufficiently large in adult trees, so that the limiting case is approximately met. However, this may not be true for young trees that have not yet undergone extensive branching; a limitation that should be kept in mind.

Expected number of fixed somatic mutations

To obtain an expression for the expected number of fixed somatic mutations in the tree crown, let us first consider the total age of the tree T, measured in units of time (e.g. years or number of cell divisions). T denotes the time from the root node (R) at $n = -1$ to the terminal nodes at $n = D$. All unique branches, starting at branch level n, can be assigned a corresponding age, t_n, which is a fraction of T ($t_n \leq T$). Let μ be the mutation rate per haploid genome, per base pair per unit time, and G be the haploid genome size. In a diploid species, the

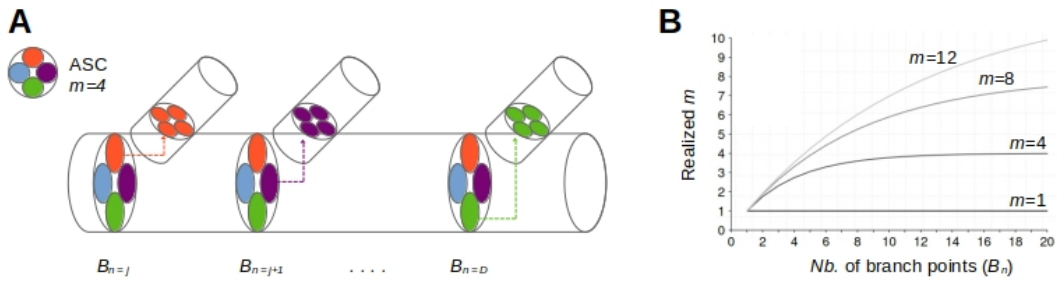

Figure 2: *Realized* effective stem cell population size. A. Three branching events are shown $B_{n=j}, B_{n=j+1}, \ldots, B_{n=D}$. At each branch point only one cell is selected indirectly from the ASC population to act as a precursor for lateral branch initiation. A new ASC population is established in the lateral branch, with all cells sharing the cell-lineage identity of its founder ASC (see color code). Although the effective stem cell population size equals four ($m = 4$) in this case, the *realized* effective stem cell population size is three ($\dot{m} = 3$), because only three distinct cells were sampled along the stem. **B.** The *realized* effective stem cell population size converges to the actual population size when the number of branch points becomes large. The rate of convergence depends on m and follows Eq.1.

combined number of ASC mutations that arise across all unique branches (C_n) with corresponding age (t_n) is:

$$s_n(\Psi) = 2G\mu \cdot \dot{m}_n \cdot C_n t_n.$$

The parameter vector $\Psi = \{D, b, t\}$ emphasizes that the expected number of mutations is determined by the specific topological characteristics of the tree. Using Eq.1 from above, we have

$$s_n(\Psi) = MC_n t_n - MC_n t_n \left(1 - \frac{1}{m}\right)^{B_n},$$

where $M = 2G\mu \cdot m$ denotes the total mutation rate of the ASC population in a single meristem per unit time. Summing over all branching levels $n = -1, \ldots, D$, the total number of fixed somatic mutations, S, in the tree as a whole is:

$$\begin{aligned} S(\Psi) &= M\mathcal{A}(\Psi) - M \sum_{n=-1}^{D} C_n t_n \left(1 - \frac{1}{m}\right)^{B_n} \\ &\approx M\mathcal{A}(\Psi) \cdot \left[1 - \left(1 - \frac{1}{m}\right)^{\overline{B}}\right]. \end{aligned}$$

where $\mathcal{A}(\Psi) = \sum_{n=-1}^{D} C_n t_n$ is the total age of all unique branches in units of time. Since M is just a constant, we can write:

$$S(\Psi) \propto \mathcal{A}(\Psi).$$

This expression should hold for an adult tree where the average branch length, $\overline{B}$, is sufficiently large. It highlights that the expected number of fixed somatic mutations in the tree crown is proportional to the spatio-temporal features of the branching architecture. These features affect mutation accumulation independently of the mutation rate and the effective stem cell population size. Our goal is to explore the impact of these architectural aspects in more detail.

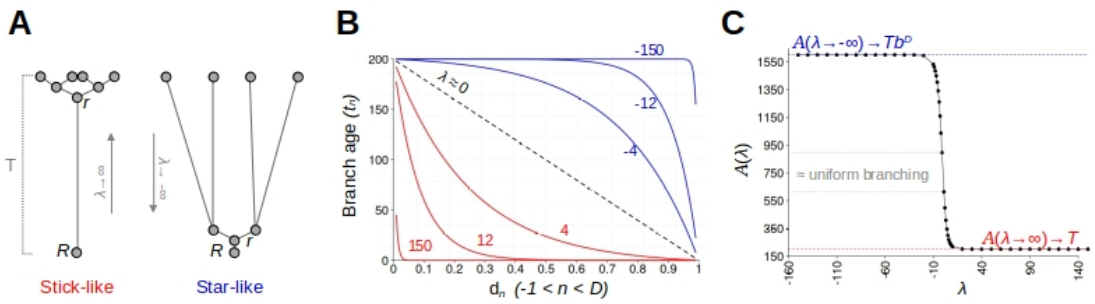

Figure 3: Branch age scaling and its impact on $\mathcal{A}(\Psi, \lambda)$. A. A b-ary tree can converge to a stick-like or a star-like tree, as parameter λ goes to ∞ or $-\infty$, respectively. **B.** Transformation of the relative distance d_n to branch ages t_n for different positive (red) or negative (blue) values of λ. For visualization we assumed $T = 200$. **C.** Values of $\mathcal{A}(\Psi, \lambda)$ as a function of λ. In the limit, a stick-like and star-like tree can differ by a factor of b^D, independently of the mutation rate. Trees with uniform branching over time display intermediate values (gray dotted lines). For visualization we assumed $b = 2$, $D = 3$, $T = 200$.

Impact of branching architecture and age distribution

We had introduced t_n as the age of branches that start at branch level n and end in one of the terminal nodes without changing branch order along the way. Let

$$t_n = T \cdot f(n), \tag{2}$$

where T is the total age of the tree and $f(n)$ is a scaling function that dictates how branch ages are distributed across the time axis, giving more or less emphasis to early or late branching events. The function can be linear or non-linear, but must satisfy the normalization condition: $\sum_{n=-1}^{D} f(n) = 1$. To motivate a flexible form for the function $f(n)$ let us first express the distance of each branch level to the terminal node as a fraction of the total branching depth of the tree:

$$d_n = \frac{B_n}{D+1} = \frac{D-n}{D+1}.$$

Here $d_D = 0$ and $d_{-1} = 1$ are boundary cases. We define $f(n)$ as follows

$$f(n) = \begin{cases} 1, & \text{if } n = -1, \\ \frac{e^{-\lambda d_n} - e^{-\lambda}}{1 - e^{-\lambda}}, & \text{if } -1 < n < D, \\ 0, & \text{if } n = D. \end{cases} \tag{3}$$

where we have chosen an exponential decay function to act on d_n $(-1 < n < D)$. For small values of λ this function expands all d_n toward 1, meaning that it moves branching events close to the root node, so that all branch ages become equivalent to the total age of the tree (Fig. 3A and B). Conversely, for large values of λ, d_n is shrunk toward 0, meaning that branching events are compressed toward terminal nodes (Fig. 3A and B). Hence, $\mathcal{A}(\Psi, \lambda)$ can be expressed as:

$$\mathcal{A}(\Psi, \lambda) = T \left(1 + \sum_{n=0}^{D-1} C_n \cdot \frac{e^{-\lambda d_n} - e^{-\lambda}}{1 - e^{-\lambda}} \right) \tag{4}$$

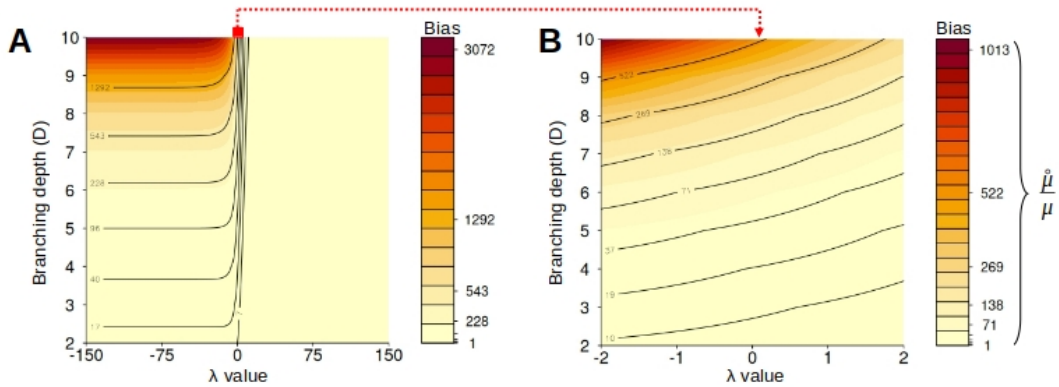

Figure 4: Biases in somatic mutation rate estimates. A. Shown is the bias of the *naive* estimator of the somatic mutation rate when branching structure and branch age distributions are ignored. The biases increases when D becomes large and λ small. But even in situations where branching is uniform $\lambda \approx 0$ and branching depth is modest (e.g. $D = 5$), the biases can be at least one order of magnitude. Panel **B.** shows a zoom-in for λ range $-2 to 2$.

where $t_{-1} = T \cdot 1$ and $t_D = T \cdot 0$ are the ages of branches starting at the root node (R) and the terminal nodes at D, respectively. It can be shown that:

$$\lim_{\lambda \to +\infty} \mathcal{A}(\Psi, \lambda) = T \quad \text{and} \quad \lim_{\lambda \to -\infty} \mathcal{A}(\Psi, \lambda) = T b^D. \tag{5}$$

(Fig. 3C). These limiting cases mean that for $\lambda \to +\infty$, mutation accumulation in the time interval T is restricted to the main stem, with essentially no contribution from younger branches. The resulting tree topology resembles a "stick-tree" (Fig. 3A), which maximizes branch path sharing from the root to the terminal nodes. Conversely, for $\lambda \to -\infty$, the tree architecture converges to a "star-tree", where branch path sharing is essentially absent (Fig. 3A). Instead, b^D branches start at the root node and independently accumulate mutations in the time interval T. An intermediate situation is reached when branch levels are uniformly distributed across time (i.e. $\lambda \approx 0$). Taken together, these results show that the number of somatic mutations in the crown of a mature tree can differ by a factor of b^D, depending on the specific branching strategy. For example, in a binary tree (i.e. $b = 2$) with 5 branching levels (i.e. $D = 5$), star-like branching configurations would accumulate 35 times more somatic mutations in terminal nodes than stick-like configurations.

Implications for somatic mutation rate estimates

The fact that the number of somatic mutations is not onlwanrchitey a function of the mutation rate and time, but also influenced by the branching architecture and branch age distribution has implications for somatic mutation rate estimation. Several genomic studies have estimated rates by relating the number of somatic mutations (S) detected in the crown to the age of the tree (T) (Duan et al. (2022), Hanlon et al. (2019), Hofmeister et al. (2020), Wang et al. (2019), Orr et al. (2020), Perez-Roman et al. (2022), Satake et al. (2023), Schmid-Siegert

et al. (2017), Shahryary et al. (2020a), Xie et al. (2016),). A common *naive* estimator used in many of these studies is:

$$\tilde{\mu} = \frac{S}{T \cdot 2G}$$

Substituting $S = 2G\mu \cdot m \cdot \mathcal{A}(\Psi, \lambda)$ from above, we can express the bias associated with the *naive* estimator as follows:

$$\text{Bias} := \frac{\tilde{\mu}}{\mu} = \frac{m\mathcal{A}(\Psi, \lambda)}{T}.$$

When we apply the limits from Eq.5, we have:

$$\lim_{\lambda \to +\infty} \text{Bias} = m \quad \text{and} \quad \lim_{\lambda \to -\infty} \text{Bias} = b^D.$$

Hence, in the extreme case where the tree converges to a stick-tree (i.e. $\lambda \to \infty$), $\tilde{\mu}$ is overestimted by a factor that scales with the effective stem cell population size. In the other extreme where the tree converges to a star-tree ($i.e. \lambda \to -\infty$), it is overestimated by a factor that scales with the number of terminal branches. This result shows that, in all cases (expect when the tree is a stick-tree with $m = 1$), the *naive* estimator $\tilde{\mu}$ is upward biased, even if the mutation rate remains unchanged and the number of terminal branches is the same. As illustrated in (Fig. 4A and B), the scale of this bias can be substantial. Even in situations where branch points are uniformally distributed across time (i.e. $\lambda \approx 0$), $\tilde{\mu}$ would overestimate the somatic mutation rate by at least one order of magnitude. This observation reinforces the need to use intra-organism phylogenetic methods for somatic mutation rate estimation (Shahryary et al. (2020b), Orr et al. (2020)), as these methods take the topological structure of the tree into account.

Conclusion

The factors influencing somatic mutation accumulation in long-lived plants are of fundamental interest to ecology and evolution. Recent genomic studies have reported unexpectedly low somatic mutation rates in trees compared to annuals (Johannes (2024))—an observation that has sparked considerable interest. In a recent review, Schoen and Schultz (2019) calculated that adult trees should accumulate far more mutations than reported in sequencing studies. Using oak as an example, and considering approximately 15 to 20 branching events, a genome size of ∼500 million base pairs, ∼7 cell divisions per branching event, and a per-cell division mutation rate of $\sim 5 \cdot 10^{-8}$ (borrowed from yeast), they estimated that 2,600 to 3,500 mutations should accumulate in the crown—two orders of magnitude more than observed. They attributed this discrepancy to incomplete mutation detection, likely due to technical challenges in identifying low-frequency variants in bulk sequencing data.

This explanation is plausible because somatic mutations originate in distinct layers of the SAM, giving rise to lineage-restricted subpopulations in lateral organs

like leaves and fruits. In bulk sequencing, these layer-specific variants often appear at low frequencies and may be misclassified as noise, leading to false negatives. However, recent layer-enriched sequencing studies have reported mutation rates consistent with bulk estimates (Goel et al. (2024), Johannes (2025)), suggesting that data heterogeneity and low variant allele frequencies are unlikely to explain the discrepancy. Similarly, applying the same calculations to germline mutation rates in Arabidopsis mutation accumulation lines, where there are only about 34 seed-to-seed cell divisions (Watson et al. (2016)), would predict mutation rates far exceeding empirical estimates. Since these germline rates are more reliably measured due to the single-layer and single-cell bottleneck at the start of each generation (Quiroz et al. (2023)), technical artifacts from variant calling are unlikely to be the primary cause. Instead, these patterns point to biological mechanisms that actively limit somatic mutation accumulation in long-lived species (Lanfear et al. (2013), Hofmeister et al. (2020), Watson et al. (2016), Burian (2021))).

In this study, we showed that developmental factors associated with branching strategies significantly influence somatic mutation accumulation. Perhaps surprisingly, our results suggest that current studies may have overestimated somatic mutation rates by up to an order of magnitude, as tree branching architecture has not been properly accounted for. This means that the already low somatic mutation rates in trees are actually even lower. While phylogenetically-informed methods have been developed to address these biases (Shahryary et al. (2020b), Orr et al. (2020)), they require accurate information about branch ages and branching topology, which remains challenging to obtain.

Although we have developed our ideas in the context of somatic DNA sequence mutations, the theory is general enough to be also applicable to somatic epimutations, which have been studied in the form of stochastic DNA methylation changes in model and non-model systems (van der Graaf et al. (2015), Denkena et al. (2021), Hazarika et al. (2022), Hofmeister et al. (2020), Zhou et al. (2024)). Spontaneous epimutations typically occur at rates that are several orders of magnitude higher than DNA sequence mutations (Johannes and Schmitz (2019), and therefore have considerably high reversion probabilities. As a consequence, they are often modeled using relaxed infinite-site assumptions (van der Graaf et al. (2015), Shahryary et al. (2020b), Charlesworth and Jain (2014), Vidalis et al. (2016), Sellinger et al. (2024), Yao et al. (2023)). While this may be necessary over evolutionary timescales, where the rate of reversions is non-negligible, over the typical lifetime of a trees these reversion probabilities can be ignored, so that our theory holds without modifications.

A key insight from our work is that, similar to mechanisms that slow meristematic cell division, branching architecture can reduce somatic (epi)mutation accumulation without altering the intrinsic (epi)mutation rate (i.e., the error rate per cell division). The next step is to investigate how branching strategies vary across species and whether they mediate the observed correlation between generation time and somatic (epi)mutation rates. A related question is to which extent genetic and environmental factors that influence branching rate and patterning may

indirectly affect the rate of somatic (epi)mutation accumulation in trees. These questions open up an interesting research program at the intersection of life-history traits and molecular evolution in long-lived plants.

Acknowledgments

I thank Lutz Johannes and Tomàs Johannes-Colomé for their help in generating Fig. 1A. Thanks also to Benjamin Werner for his quick feedback.

References

Burian, A. (2021). Does shoot apical meristem function as the germline in safeguarding against excess of mutations? *Front. Plant Sci.*, 12:707740.

Burian, A., Barbier de Reuille, P., and Kuhlemeier, C. (2016). Patterns of Stem Cell Divisions Contribute to Plant Longevity. *Curr. Biol.*, 26(11):1385–1394.

Charlesworth, B. and Jain, K. (2014). Purifying selection, drift, and reversible mutation with arbitrarily high mutation rates. *Genetics*, 198(4):1587–1602.

Chen, Y., Burian, A., and Johannes, F. (2024). Somatic epigenetic drift during shoot branching: a cell lineage-based model. *Genetics*, page iyae091.

De La Torre, A. R., Li, Z., Van de Peer, Y., and Ingvarsson, P. K. (2017). Contrasting rates of molecular evolution and patterns of selection among gymnosperms and flowering plants. *Mol. Biol. Evol.*, 34(6):1363–1377.

Denkena, J., Johannes, F., and Colomé-Tatché, M. (2021). Region-level epimutation rates in arabidopsis thaliana. *Heredity*, pages 1–13.

Duan, Y., Yan, J., Zhu, Y., Zhang, C., Tao, X., Ji, H., Zhang, M., Wang, X., and Wang, L. (2022). Limited accumulation of high-frequency somatic mutations in a 1700-year-old osmanthus fragrans tree. *Tree Physiol.*, 42(10):2040–2049.

Goel, M., Campoy, J. A., Krause, K., Baus, L. C., Sahu, A., Sun, H., Walkemeier, B., Marek, M., Beaudry, R., Ruiz, D., Huettel, B., and Schneeberger, K. (2024). The vast majority of somatic mutations in plants are layer-specific. *Genome Biol.*, 25(1):194.

Hanlon, V. C. T., Otto, S. P., and Aitken, S. N. (2019). Somatic mutations substantially increase the per-generation mutation rate in the conifer picea sitchensis. *Evol Lett*, 3(4):348–358.

Hazarika, R. R., Serra, M., Zhang, Z., Zhang, Y., Schmitz, R. J., and Johannes, F. (2022). Molecular properties of epimutation hotspots. *Nat Plants*, 8(2):146–156.

Heidstra, R. and Sabatini, S. (2014). Plant and animal stem cells: similar yet different. *Nat. Rev. Mol. Cell Biol.*, 15(5):301–312.

Hofmeister, B. T., Denkena, J., Colomé-Tatché, M., Shahryary, Y., Hazarika, R., Grimwood, J., Mamidi, S., Jenkins, J., Grabowski, P. P., Sreedasyam, A., Shu, S., Barry, K., Lail, K., Adam, C., Lipzen, A., Sorek, R., Kudrna, D., Talag, J., Wing, R., Hall, D. W., Jacobsen, D., Tuskan, G. A., Schmutz, J., Johannes, F., and Schmitz, R. J. (2020). A genome assembly and the somatic genetic and epigenetic mutation rate in a wild long-lived perennial Populus trichocarpa. *Genome Biol.*, 21(1):259.

Iwasa, Y., Tomimoto, S., and Satake, A. (2023). The genetic structure within a single tree is determined by the behavior of the stem cells in the meristem. *Genetics*, 223(4).

Johannes, F. (2024). Allometric scaling of somatic mutation and epimutation rates in trees. *Evolution*, 79(1):1–5.

Johannes, F. (2025). Somatic evolution of stem cell mutations in long-lived plants. *bioRxiv*.

Johannes, F. and Schmitz, R. J. (2019). Spontaneous epimutations in plants. *New Phytol.*, 221(3):1253–1259.

Kwiatkowska, D. (2008). Flowering and apical meristem growth dynamics. *J. Exp. Bot.*, 59(2):187–201.

Lanfear, R., Ho, S. Y. W., Jonathan Davies, T., Moles, A. T., Aarssen, L., Swenson, N. G., Warman, L., Zanne, A. E., and Allen, A. P. (2013). Taller plants have lower rates of molecular evolution. *Nat. Commun.*, 4(1):1879.

Lesaffre, T. (2021). Population-level consequences of inheritable somatic mutations and the evolution of mutation rates in plants. *Proc. Biol. Sci.*, 288(1958):20211127.

Lyndon, R. F. (1998). *Developmental and cell biology series: The shoot apical meristem: Its growth and development series number 34*. Developmental and cell biology series. Cambridge University Press, Cambridge, England.

Orr, A. J., Padovan, A., Kainer, D., Külheim, C., Bromham, L., Bustos-Segura, C., Foley, W., Haff, T., Hsieh, J.-F., Morales-Suarez, A., Cartwright, R. A., and Lanfear, R. (2020). A phylogenomic approach reveals a low somatic mutation rate in a long-lived plant. *Proceedings of the Royal Society B: Biological Sciences*, 287(1922).

Perez-Roman, E., Borredá, C., López-García Usach, A., and Talon, M. (2022). Single-nucleotide mosaicism in citrus: Estimations of somatic mutation rates and total number of variants. *Plant Genome*, 15(1):e20162.

Quiroz, D., Lensink, M., Kliebenstein, D. J., and Monroe, J. G. (2023). Causes of mutation rate variability in plant genomes. *Annu. Rev. Plant Biol.*, 74:751–775.

Satake, A., Imai, R., Fujino, T., Tomimoto, S., Ohta, K., Na'iem, M., Indrioko, S., Widiyatno, Purnomo, S., Mollá-Morales, A., Nizhynska, V., Tani, N., Suyama,

Y., Sasaki, E., and Kasahara, M. (2023). Somatic mutation rates scale with time not growth rate in long-lived tropical trees. *eLife*.

Schmid-Siegert, E., Sarkar, N., Iseli, C., Calderon, S., Gouhier-Darimont, C., Chrast, J., Cattaneo, P., Schütz, F., Farinelli, L., Pagni, M., Schneider, M., Voumard, J., Jaboyedoff, M., Fankhauser, C., Hardtke, C. S., Keller, L., Pannell, J. R., Reymond, A., Robinson-Rechavi, M., Xenarios, I., and Reymond, P. (2017). Low number of fixed somatic mutations in a long-lived oak tree. *Nature Plants*, 3(12):926.

Schoen, D. J. and Schultz, S. T. (2019). Somatic mutation and evolution in plants. *Annu. Rev. Ecol. Evol. Syst.*, 50(Volume 50, 2019):49–73.

Scofield, D. G. and Schultz, S. T. (2006). Mitosis, stature and evolution of plant mating systems: low-phi and high-phi plants. *Proc. Biol. Sci.*, 273(1584):275–282.

Sellinger, T., Johannes, F., and Tellier, A. (2024). Improved inference of population histories by integrating genomic and epigenomic data. *eLife*, 12:RP89470.

Shahryary, Y., Hazarika, R. R., and Johannes, F. (2020a). MethylStar: A fast and robust pre-processing pipeline for bulk or single-cell whole-genome bisulfite sequencing data. *BMC Genomics*, 21(1):1–8.

Shahryary, Y., Symeonidi, A., Hazarika, R. R., Denkena, J., Mubeen, T., Hofmeister, B., van Gurp, T., Colomé-Tatché, M., Verhoeven, K. J. F., Tuskan, G., Schmitz, R. J., and Johannes, F. (2020b). AlphaBeta: computational inference of epimutation rates and spectra from high-throughput DNA methylation data in plants. *Genome Biol.*, 21(1):1–22.

Shi, B., Zhang, C., Tian, C., Wang, J., Wang, Q., Xu, T., Xu, Y., Ohno, C., Sablowski, R., Heisler, M. G., Theres, K., Wang, Y., and Jiao, Y. (2016). Two-step regulation of a meristematic cell population acting in shoot branching in arabidopsis. *PLoS Genet.*, 12(7):e1006168.

Tomimoto, S. and Satake, A. (2023). Modelling somatic mutation accumulation and expansion in a long-lived tree with hierarchical modular architecture. *J. Theor. Biol.*, 565:111465.

van der Graaf, A., Wardenaar, R., Neumann, D. A., Taudt, A., Shaw, R. G., Jansen, R. C., Schmitz, R. J., Colomé-Tatché, M., and Johannes, F. (2015). Rate, spectrum, and evolutionary dynamics of spontaneous epimutations. *Proceedings of the National Academy of Sciences*, 112(21):6676–6681.

Vidalis, A., Živković, D., Wardenaar, R., Roquis, D., Tellier, A., and Johannes, F. (2016). Methylome evolution in plants. *Genome Biol.*, 17(1):264.

Wang, L., Ji, Y., Hu, Y., Hu, H., Jia, X., Jiang, M., Zhang, X., Zhao, L., Zhang, Y., Jia, Y., Qin, C., Yu, L., Huang, J., Yang, S., Hurst, L. D., and Tian, D. (2019). The architecture of intra-organism mutation rate variation in plants. *PLoS Biol.*, 17(4):e3000191.

Watson, J. M., Platzer, A., Kazda, A., Akimcheva, S., Valuchova, S., Nizhynska, V., Nordborg, M., and Riha, K. (2016). Germline replications and somatic mutation accumulation are independent of vegetative life span in arabidopsis. *Proc. Natl. Acad. Sci. U. S. A.*, 113(43):12226–12231.

Xie, Z., Wang, L., Wang, L., Wang, Z., Lu, Z., Tian, D., Yang, S., and Hurst, L. D. (2016). Mutation rate analysis via parent-progeny sequencing of the perennial peach. I. a low rate in woody perennials and a higher mutagenicity in hybrids. *Proc. Biol. Sci.*, 283(1841).

Yao, N., Zhang, Z., Yu, L., Hazarika, R., Yu, C., Jang, H., Smith, L. M., Ton, J., Liu, L., Stachowicz, J. J., Reusch, T. B. H., Schmitz, R. J., and Johannes, F. (2023). An evolutionary epigenetic clock in plants. *Science*, 381(6665):1440–1445.

Zhou, M., Schmied, G., Bradatsch, M., Resente, G., Hazarika, R., Kakoulidou, I., Costa, M., Serra, M., Uhl, E., Schmitz, R., Hilmers, T., Toraño Caicoya, A., Crivellaro, A., Pretzsch, H., and Johannes, F. (2024). Accelerated growth increases the somatic epimutation rate in trees. *bioRxiv*.