PNAS RESEARCH ARTICLE | GENETICS

Low-frequency somatic mutations are heritable in tropical trees *Dicorynia guianensis* and *Sextonia rubra*

Sylvain Schmitt[a,b,c,1], Patrick Heuret[d], Valérie Troispoux[e], Mélanie Beraud[f], Jocelyn Cazal[e], Émilie Chancerel[g], Charlotte Cravero[h], Erwan Guichoux[g], Olivier Lepais[g], João Loureiro[i], William Marande[h], Olivier Martin-Ducup[j], Gregoire Vincent[d], Jérôme Chave[k], Christophe Plomion[g], Thibault Leroy[l,m,2], Myriam Heuertz[g,2], and Niklas Tysklind[e,2]

Edited by Detlef Weigel, Max-Planck-Institut fur Biologie Tubingen, Tübingen, Germany; received August 3, 2023; accepted January 22, 2024

Somatic mutations potentially play a role in plant evolution, but common expectations pertaining to plant somatic mutations remain insufficiently tested. Unlike in most animals, the plant germline is assumed to be set aside late in development, leading to the expectation that plants accumulate somatic mutations along growth. Therefore, several predictions were made on the fate of somatic mutations: mutations have generally low frequency in plant tissues; mutations at high frequency have a higher chance of intergenerational transmission; branching topology of the tree dictates mutation distribution; and exposure to UV (ultraviolet) radiation increases mutagenesis. To provide insights into mutation accumulation and transmission in plants, we produced two high-quality reference genomes and a unique dataset of 60 high-coverage whole-genome sequences of two tropical tree species, *Dicorynia guianensis* (Fabaceae) and *Sextonia rubra* (Lauraceae). We identified 15,066 de novo somatic mutations in *D. guianensis* and 3,208 in *S. rubra*, surprisingly almost all found at low frequency. We demonstrate that 1) low-frequency mutations can be transmitted to the next generation; 2) mutation phylogenies deviate from the branching topology of the tree; and 3) mutation rates and mutation spectra are not demonstrably affected by differences in UV exposure. Altogether, our results suggest far more complex links between plant growth, aging, UV exposure, and mutation rates than commonly thought.

somatic mutations | tree crown | mutation transmission

Significance

The origin and fate of new mutations have received less attention in plants than in animals. Similarly to animals, plant mutations are expected to accumulate with growth and time and under exposure to ultraviolet light. However, contrary to animals, plant reproductive organs form late in an individual's development, allowing the transmission to the progeny of mutations accumulated along growth. Here, we resequenced DNA from different branches differentially exposed to sunlight of two tropical tree species. We showed that new mutations are generally rare in plant tissues and do not mimic branching patterns but can nevertheless be transmitted to the progeny. Our findings provide a perspective on heritable plant mutation and its pivotal role as the engine of evolution.

The Weismann theory (1) states that hereditary traits are transmitted exclusively from the germline. The theory is valid in most animals (2) where germline cells are set aside early in development (1). In plants, germline segregation is generally assumed to occur late in development (3, 4 but see ref. 5), which leads to several predictions on the fate of somatic mutations occurring in plant tissues: mutations have generally low frequency in plant tissues (6); mutations at high frequency have a higher chance of intergenerational transmission; branching topology of the tree dictates mutation distribution (7); and exposure to UV (ultraviolet) radiation increases mutagenesis (8). At present, all these hypotheses, albeit crucial for plant science, have been poorly tested empirically.

To identify a large set of de novo plant somatic mutations, we resequenced 60 samples in total for two tropical tree species, *Dicorynia guianensis* (Amshoff) and *Sextonia rubra* (Mez) van der Werff (*SI Appendix*, *Note S1*), corresponding to 3 leaves per branch for a total of up to 10 branches per tree, in addition to cambium tissues from the base of the trunk for comparison (*SI Appendix*, *Note S2*). The branches were selected as growing in either low or high light exposure, getting the benefits of the maximum contrast of forests located near the equator (5°N). UV light exposure was assessed directly at the sampling points and additionally estimated with a canopy transmittance model inferred using terrestrial and drone lidar scans for the *D. guianensis* tree (*SI Appendix*, *Note S3*). Given that the quality of the reference genomes is known to be a key aspect to ensure accurate mutation detection, we used a combination of high-fidelity reads and optical maps to generate near chromosome-level assemblies for two wild tropical tree species, *D. guianensis* and *S. rubra*. The two genome assemblies differ in size (550 and 991 Mb) and in their genomic content for guanine–cytosine (GC), TE, and genes, with highly heterogeneous patterns along chromosomes in *D. guianensis* vs. relatively homogeneous ones in *S. rubra* (Fig. 1 and *SI Appendix*, *Note S4*). These two high-quality annotated genomes were used as a reference to detect somatic mutations.

Using a mutation detection methodology initially developed for human cancer mutations (9) and later adapted to plants (6), we identified 15,066 unique somatic mutations in *D. guianensis* and 3,208 in *S. rubra*. Only a few were restricted to a single branch (5 to 9%, Fig. 2 *A* and *D* and *SI Appendix*, *Note S5*), whereas most mutations were shared by at least two branches whose nearest shared branching point was the base of the crown

Author contributions: J. Chave, C.P., T.L., M.H., and N.T. designed research; S.S., T.L., M.H., and N.T. performed research; P.H., V.T., M.B., J. Cazal, É.C., C.C., E.G., O.L., J.L., W.M., O.M.-D., and G.V. contributed new reagents/analytic tools; S.S., T.L., M.H., and N.T. analyzed data; and S.S., T.L., M.H., and N.T. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: sylvain.schmitt@cirad.fr.

[2]T.L., M.H., and N.T. contributed equally to this work.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2313312121/-/DCSupplemental.

Published February 27, 2024.

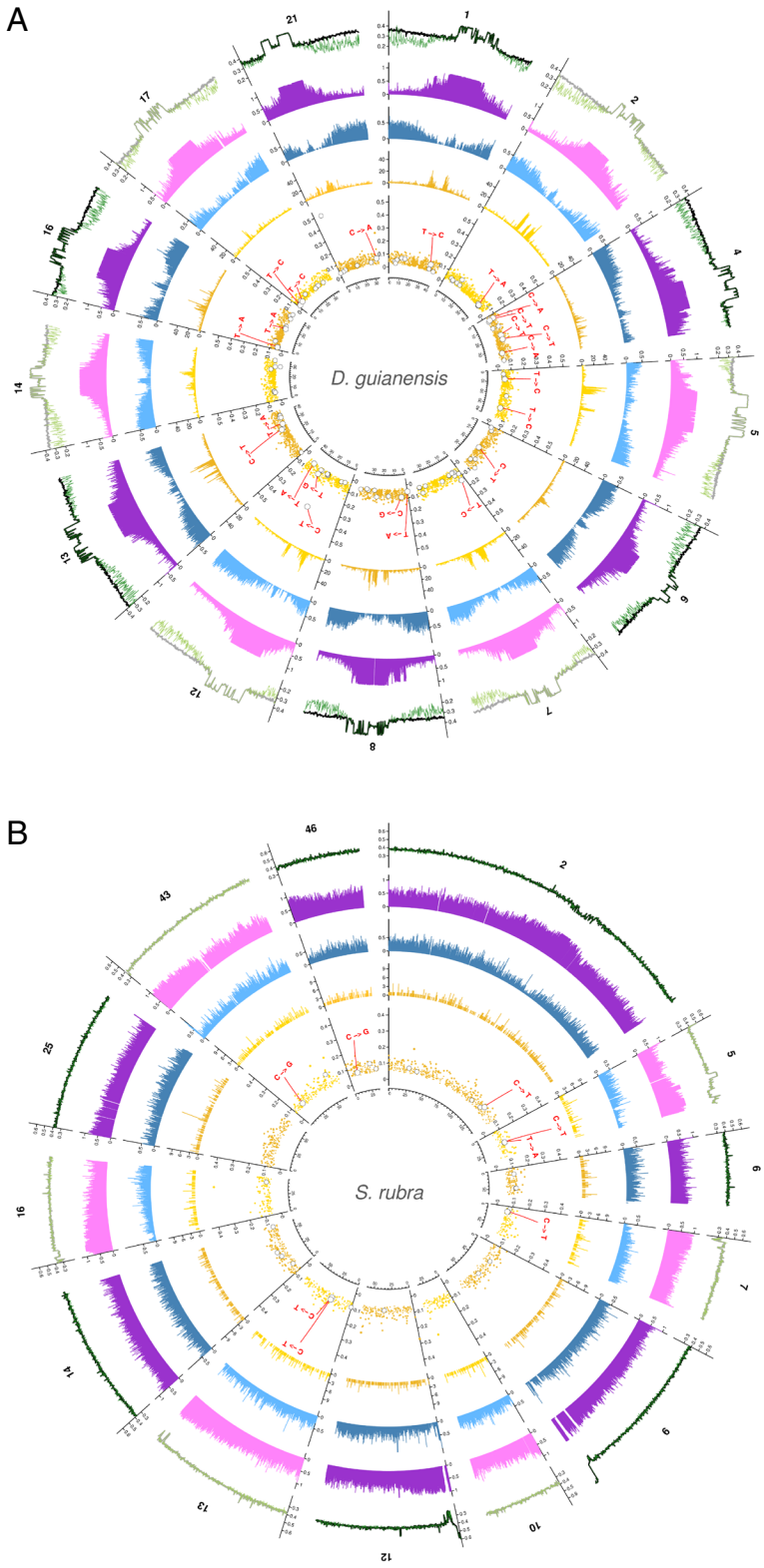

Fig. 1. Crown mutations and transmitted mutations in the genomic landscape of the *D. guianensis* and the *S. rubra* trees' assembled pseudochromosomes. The genomic landscape is similarly portrayed for the two tropical trees: the *D. guianensis* tree (*A*) and the *S. rubra* tree (*B*). The first (most external) track represents the percentage of GC in the whole genome with the black line and in the TE with the green line. The second (least external) track represents the percentage of TE with purple bars. The third track (*Middle*) represents the percentage of genes with blue bars. The fourth (least internal) track represents the number of somatic mutations detected in the tree crown with yellow bars. The number of somatic mutations correlates with genomic landscapes in *D. guianensis*, the species exhibiting a higher genomic heterogeneity in terms of percentage of genes and TEs (Poisson regression, percentage of TEs b = −0.37(0.04), $P < 1.10\text{-}16$, percentage of genes b = −2.31(0.15), $P < 1.10\text{-}16$), whereas this is not always significant in *S. rubra* (Poisson regression, percentage of TEs b = −0.62(0.10), $P < 1.10\text{-}9$, percentage of genes b = −0.31(0.18), $P = 0.746$). The fifth (innermost) track represents the allelic fraction of the somatic mutations detected in the crown in yellow, the mutations tested for transmission in gray, and the mutations found transmitted to the embryos in red. The inner labels indicate the type of mutations for somatic mutations transmitted to embryos. All measurements are calculated in nonoverlapping windows of 100 kb. A ruler is drawn on each pseudochromosome, with tick marks every 2 Mb. The genome heterozygosities estimated with K-mer distributions were high for both species, at 0.9% for *D. guianensis* and at 0.7% for *S. rubra*.

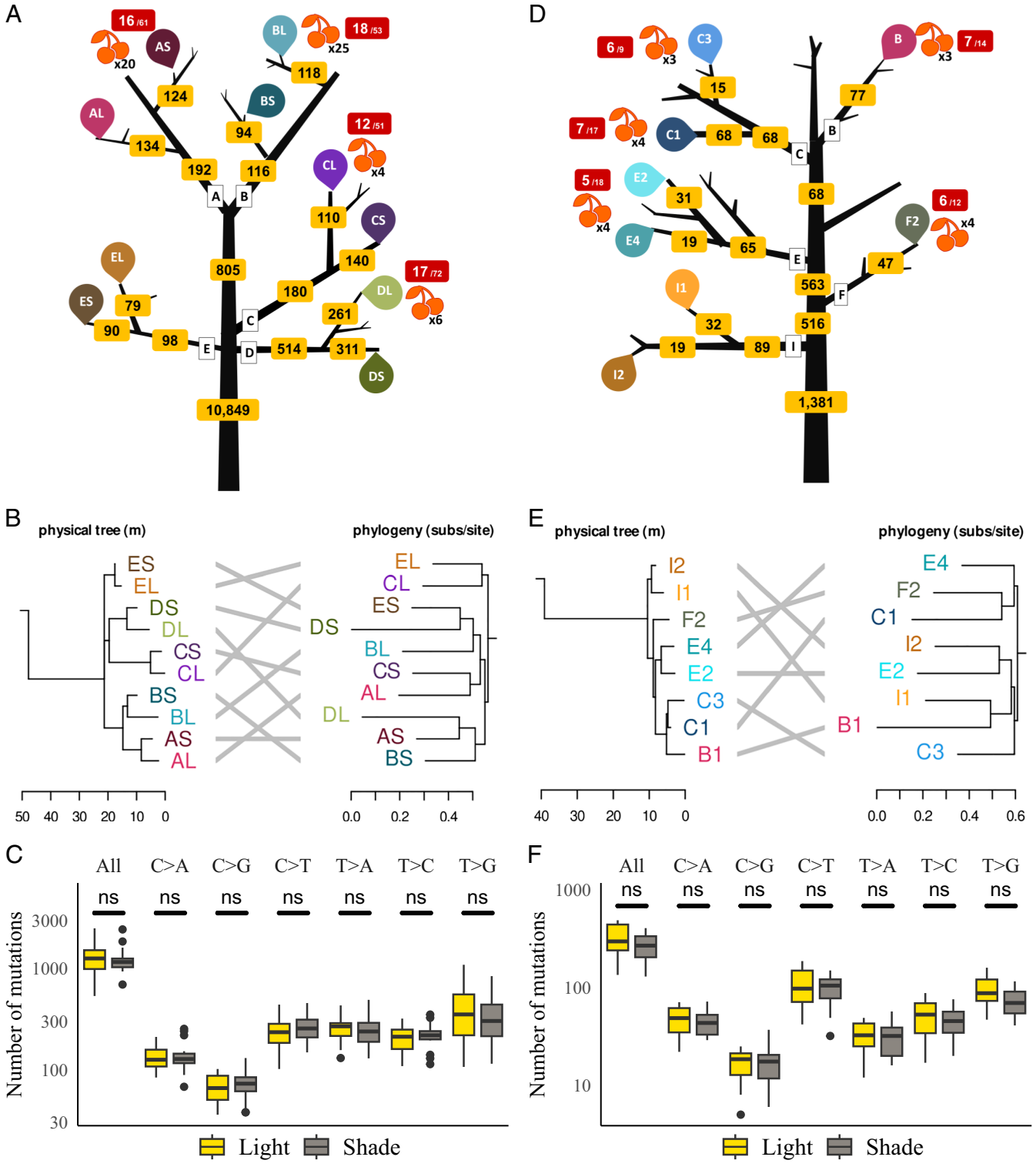

Fig. 2. Distributions of somatic mutations through branching topology of the tree, phylogenies, and with light. The distributions of somatic mutations through physical trees, phylogenies, and with light are similarly shown for the two tropical trees: the *D. guianensis* tree (*A*, *B*, and *C*) and the *S. rubra* tree (*D*, *E*, and *F*). (*A* and *D*) The branching topology of the tree is shown in black with the branch names in white boxes. The number of somatic mutations through the crown is indicated in the yellow boxes before the original branching event. The sampling points of three leaves in the light-exposed branches ("L" in letter codes, light colors) and in the shaded branches ("S", dark colors) are indicated with unique letter codes and colored drop symbols. Fruit sampling points are represented by red fruits, with the number of fruits sampled indicated in black. The red boxes with white labels indicate the transmission of mutations to fruit embryos out of the total number of mutations tested. (*B* and *E*) A side-by-side comparison of the physical tree (*Left*, branch length in meters) and the maximum likelihood phylogeny of mutations (*Right*, branch length in substitutions per site). The letters on the ends of the branches indicate the sampling points shown in (*A* and *D*). (*C* and *F*) Different mutagens may cause specific mutation types, i.e., changing from base X to base Y (X > Y). The effect of light exposure on the accumulation of somatic mutations as a function of mutation type (X > Y) is represented in yellow and gray boxes. The yellow boxes represent the number of mutations accumulated in all leaves of light-exposed branches and the gray boxes in all leaves of shaded branches. Boxplots show the median (center line), upper and lower quartiles (box limits), 1.5x interquartile range (whiskers), and outliers (points). The "ns" labels indicate nonsignificant differences in Student's *t* tests (two-sided). Mutation types include all mutations and all types of transitions and transversions. The *y*-axis was logarithmically scaled to facilitate reading of low values.

(43 to 72%), thus originating below the base of the crown. We further tested the correspondence between the topology of the physical tree and the phylogenies obtained from the somatic mutations and found no correspondence (Fig. 2 *B* and *E* and *SI Appendix*, *Note S6*). These results challenge the expectation in plants that the distribution of mutations corresponds to the branching topology of the tree following the growth of the shoot apical meristems (7). We also found no difference in the number of mutations, the type of mutations (nucleotide changes), or the mutation spectra (mutation context with 5′ and 3′ amino acids) between the branches exposed to high vs. low light conditions (Fig. 2 *C* and *F* and *SI Appendix*, *Note S7*), which suggests a shielding from UVs in the bud layers (10).

As compared to previous reports about somatic mutations in plants (4, 10, 11), we have detected far more mutations (10 to 100 times more). This discrepancy is likely associated with the methodology (6) since the vast majority of identified somatic mutations had a low allelic fraction, i.e., the fraction of genomic reads with the mutation, which indicates the frequency of mutated cells in the analyzed sample (Fig. 3 *A* and *B* and *SI Appendix*, *Note S8*). The higher total number of mutations detected in *D. guianensis* can be explained by an enrichment in low fraction mutations in the *D. guianensis* tree detected through deeper sequencing (*SI Appendix*, Fig. S15) because increasing the number of reads of a genomic region increases the chances of finding a mutation present in only a few cells of a sample. We generalized the result of the predominance of low fraction mutations in two pedunculate oaks (4, 10) and a dataset from one tortuous beech *Fagus sylvatica* L. using the same methodology (Fig. 3*C*). We then considered mutations at a high allelic fraction (>0.25), a category of mutations for which methodological differences are expected to have a limited impact. The two tropical trees had 3 and 6 somatic mutations with allelic fraction>0.25, as compared to 56 to 421 somatic mutations for the reanalyzed oaks and beech trees from temperate regions (Fig. 3*C*). Future large-scale investigations are needed to properly test whether there is a difference in mutation rates

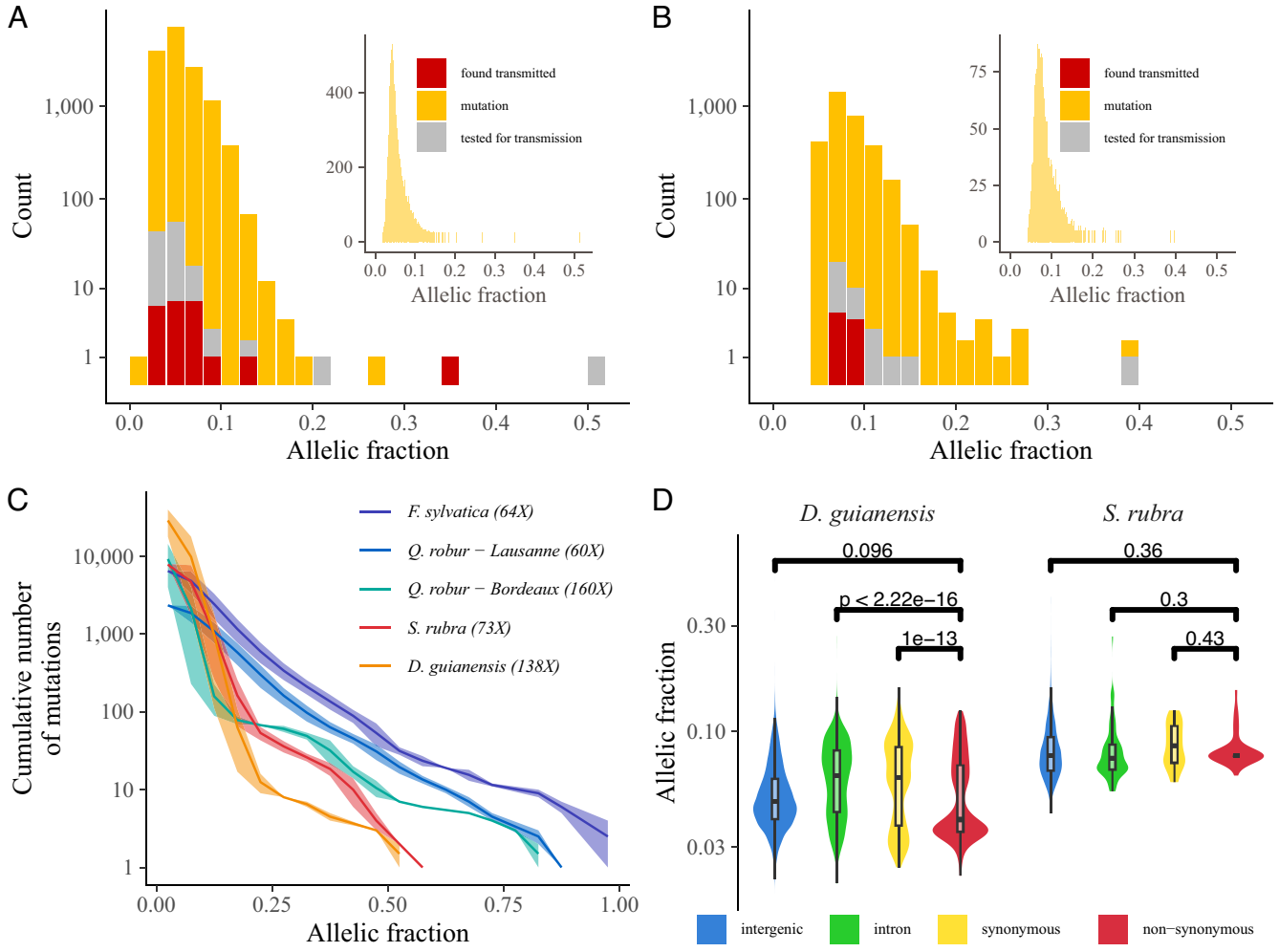

Fig. 3. Allelic fractions of somatic mutations among trees and among genomic elements. Histogram of allelic fractions of mutations detected in the crown of the two tropical trees: the *D. guianensis* tree (*A*) and the *S. rubra* tree (*B*). The main histograms show the allelic fractions of the somatic mutations using a bin of 0.02 and a log-transformed count with the mutations detected in the crown in yellow, the mutations tested for transmission in gray, and the mutations found transmitted to the embryos in red. The histograms in *insets* show the allelic fractions of the somatic mutations using a bin of 0.001 and a natural count. (*C*) Cumulative number of somatic mutations per branch with decreasing allelic fraction for five trees reanalyzed with the same pipeline. The five trees include the two tropical trees studied, i.e., the *D. guianensis* tree in orange and the *S. rubra* tree in red, and three temperate trees, two pedunculate oaks *Q. robur* L. from Bordeaux in green and Lausanne in blue and a tortuous phenotype of common beech *F. sylvatica* L. in purple. All trees were analyzed with the same pipeline (*Materials and Methods*) but were sequenced with a different depth indicated in brackets. The line represents the median value while the area represents the minimum and maximum values on the 2 to 10 branches per tree. (*D*) Comparisons of allelic fractions for nonsynonymous mutations in red with synonymous mutations in yellow, intronic mutations in green and intergenic mutations in blue for the two tropical trees: the *D. guianensis* tree (*Left*) and the *S. rubra* tree (*Right*). Boxplots show the median (center line), upper and lower quartiles (box limits), 1.5x interquartile range (whiskers), and outliers (points). The *P*-value above the bars indicates the significance of the Student's *t* test (two-sided) for the pairs of groups.

between temperate and tropical trees, after accounting for the phylogenetic signal, differences in tree age, among other null hypotheses. Overall, our results suggest that low-frequency mutations account for the vast majority of within-individual somatic diversity in plants (for all species, >90% with f < 0.25).

The origin of the somatic mutations' spatial distribution in the physical tree lies in the functioning of the shoot apical meristems. Shoot apical meristems divide either symmetrically into two stem cells or asymmetrically into one stem cell and one differentiated cell (12), resulting in the three-dimensional spatial distribution of stem cells and the somatic mutations they carry during tree growth. In eudicots, the layered structure of shoot apical meristems limits cell movement through the prevalence of anticlinal cell divisions, which favors the retention of mutated cell clones, e.g., in the form of stable periclinal chimeras (13). This mechanism could lead to sectoral chimerism through somatic mutations, which may explain both the discrepancy between the physical tree and phylogeny (Fig. 2 *B* and *E*, 12) and the prevalence of numerous low-frequency somatic mutations (Fig. 3 *A* and *B*).

Somatic mutations are often viewed as a source of within-tree adaptive variation (14). To test this hypothesis, we investigated whether nonsynonymous somatic mutations exhibit differences in allelic fraction as compared to synonymous ones or to noncoding regions. Higher, or lower fractions would be evidence for positive, or negative selection, respectively. For both species, we detected that the average allelic fraction at nonsynonymous sites was lower than those at synonymous sites (Fig. 3*D*). This difference is highly significant in *D. guianensis* (Student's *t* test, *P*-value < 10^{-13}) but not significant in *S. rubra* ($P = 0.43$), likely because a limited number of mutations was detected (31 synonymous and 9 nonsynonymous mutations). All together, these results are consistent with the intraorganismal purifying selection of nonsynonymous mutations, as also observed in seagrass (14), supporting that far more de novo mutations are detrimental than beneficial.

Until now, low-frequency somatic mutations have been neglected because they were assumed not to be transmitted and therefore to have no evolutionary future. We explored the transmission of somatic mutations to the next generation through their redetection in the embryos of developing fruits. We used amplicon resequencing for 160 candidate mutations highly shared between sampled leaves and branches, including low-frequency mutations. Using stringent quality filters (*SI Appendix*, *Note S9*), we demonstrated the transmission of 23 out of 160 tested mutations to embryos in *D. guianensis* and 9 out of 36 in *S. rubra* (Fig. 1). The transmitted mutations were found in several branches of the *D. guianensis* tree but in only one branch of the *S. rubra* tree (Fig. 2). Surprisingly, almost all the mutations for which we found empirical support for their transmission were at low frequency within the plant. Consistently, we observed that the distributions of the allelic fraction of the transmitted mutations (red bars in Fig. 3 *A* and *B*) were similar to the distributions of the allelic fraction of all mutations in the crown of the trees (yellow bars in Fig. 3 *A* and *B*, two-sided Student's *t* test t = 1.41 [–0.40, 0.07], df = 22, $P = 0.17$ for *D. guianensis* and t = –0.34 [–0.12, 0.09], df = 8, $P = 0.07$ for *S. rubra*), resulting in all transmitted mutations having low allelic fractions. By using only mutations with high empirical variant scores, we validated the robustness of our results, namely the abundance of low-frequency mutations, their transmission, the lack of correspondence between mutation phylogenies and the topology of the tree, and the absence of a spectrum associated with UV (*SI Appendix*, *Note S10*). Hence, we found that low-frequency somatic mutations are heritable and thus contribute to increased within-species diversity, which challenges current tacit assumptions that only high-frequency mutations would matter for evolution. Despite their low frequency and scarcity across the genome, low-frequency somatic mutations could substantially contribute to standing genetic variation, which is the engine of evolution (*SI Appendix*, *Note S11*). We therefore call for a new view on somatic mutations in plants with renewed assumptions: i) the distribution of somatic mutations does not necessarily correspond to the branching topology of the tree, ii) most somatic mutations are low-frequency mutations, and iii) low-frequency mutations can be transmitted to embryos in trees. Our results are consistent with far more complex links between growth, aging, and mutation rates than commonly thought in plants, along the lines of recent empirical evidence in animals (2, 15).

Materials and Methods

Choice of Species and Individuals. The study was conducted in the Amazon forest, in the coastal forests of French Guiana. A database of 710 tree species containing available information on the presence of tree rings, maximum diameter at breast height, architectural type, reproductive phenology, and ecological and economic importance was constructed. A set of 15 candidate species was selected, and their genome size was estimated by flow cytometric analyses. On this basis, we chose to work on *D. guianensis* (Amshoff) and *S. rubra* (Mez) van der Werff, which are common in French Guiana, and are ecologically and economically important species. We selected large-stature trees above 40 m without signs of dieback or senescence to maximize the potential for mutations with an increased number of cell divisions. The architecture of the trees was studied with binoculars and by climbing to selected trees where in each bough we could sample pairs of branches with contrasting light exposure. We finally selected a *D. guianensis* tree in the Saint George area (4°01′N, 51°59′W), which has an annual rainfall of 3,665 mm and a mean air temperature of 27 °C, and a *S. rubra* tree near the Paracou research station (5°18′N, 52°53′W), which has an average annual rainfall of 3,041 mm and a mean air temperature of 25.7 °C (*SI Appendix*, *Note S1*).

Sampling and Tree Structure. On 13 October 2020, we sampled the *S. rubra* tree: three cambium tissues at the base of the trunk, about 1.3 m above the ground and equidistant around the perimeter of the trunk, and three leaves from the same twig per branch for a total of eight branches were sampled. The branches were selected in three pairs, each from a different bough, plus two independent branches from two other boughs. Fruits were sampled from 5 different branches where leaves had been collected. On 22 April 2021, we sampled the *D. guianensis* tree: three cambium tissues at the base of the trunk, about 1.3 m above the ground and equidistant from the perimeter of the trunk, and three leaves from a single twig per branch for a total of 10 branches were sampled. The branches were selected in five pairs, each from a different bough. Fruits were sampled from 5 different branches where leaves had been collected. On 13 October 2020 and 15, 16, and 22 November 2021, we described the structure of both trees: branching patterns were recorded, and all branch lengths, as well as basal and terminal diameters, were measured for branches with a basal diameter greater than 10 cm, in addition to the trunk. On 21, 22, and 23 March 2022, 30 wood cores were collected with a drill in branches from throughout the crown and in the trunk of *D. guianensis*. In both species, leaf, cambium, and fruit samples were frozen in liquid nitrogen and stored at −80 °C until DNA and RNA extraction (*SI Appendix*, *Note S2*).

Characterization of Light Conditions. A linear photosynthetically active radiation (PAR) ceptometer (AccuPar, Decagon Devices, Pullman, WA) was used at each sampling position on both trees during sampling to measure direct incident light in the 400- to 700-nm wavelength range around noon in comparison to open incident light measured on the nearest road. A ground (TLS) and drone (DLS) lidar (light detection and ranging) campaign (TLS, Faro Focus3D 120; DLS, Yellowscan Vx20-100) was conducted on 3 May 2021 to map the transmittance of the *D. guianensis* tree canopy. TLS scans were performed horizontally from 0 to 360° and vertically from −60 to 90°, resulting in 174.8 million points per scan for 10 scans in a forest gap near the tree and 4 scans from the nearby road. DLS scans were taken at 35 m above the focal tree in two perpendicular flights with flight lines spaced 10 m apart in a circular area 150 m in diameter above the focal tree, resulting in 46.6 million points. Prior to the lidar acquisition, reflective strips were

placed on the sample points by tree climbers to detect the sample points in the lidar cloud. AMAPvox software was used to calculate an annual illuminance index from the aerial laser scanning. The plant area density (PAD, m2/m3) was calculated for the focal tree in context (with a diameter of 30 m around the tree) using 1 m3 voxels. An estimate of the annual proportion of solar radiation above the canopy received at the sample point was then simulated considering a brightness index of 0.5 and a latitude of 5 degrees. The uncertainty in transmittance due to uncertainties in the location of the sampling point was further assessed by randomly sampling 10 positions around the sampling points to 0.5 m and revealed small variations in transmittance. The estimates were in agreement with the light/shade classification of branches identified by the tree climbers (*SI Appendix, Note S3*).

Genome Assemblies and Annotations. High-molecular-weight (HMW) DNA was extracted from 0.7 g of three leaves of both individuals using CTAB and isopropanol precipitation before RNAase treatment and bead purification. High-fidelity (HiFi) genomic reads were produced with two sequencing runs on the PacBio Sequel II system on 2 (*D. guianensis*) to 4 (*S. rubra*) SMRTCells for each run. We obtained 1,898,004 corrected reads for *D. guianensis* (N50 = 21,233; DP = 58.7X), which we assembled into 562 contigs (N50 = 37.76Mb; L50 = 8 contigs; GC = 37.25%) using the HiFiasm assembler (v0.15.5). Similarly, we obtained 6,624,997 corrected reads for *S. rubra* (N50 = 17,577; DP = 114X), which we assembled into 747 contigs (N50 = 16.513Mb; L50 = 17 contigs; GC = 38.50%). HMW DNA was also used to produce optical maps to construct hybrid scaffolds with optical reads produced by two passages of Bionano saphyr. For *D. guianensis*, we obtained 54 hybrid scaffolds (N50 = 38,450Mb; N = 0.76%; 571 gaps), while 515 contigs remained unanchored with a total length of 28,784 Mb representing 4.97% of the genome. For *S. rubra*, we obtained 35 hybrid scaffolds (N50 = 60.458 Mb; N = 1%; 1.923 gaps), while 609 contigs remained unanchored with a total length of 53.067 Mb representing 5.08% of the genome. Genome quality was evaluated using BUSCO and Merqury. We constructed an automated genome annotation workflow that performs: i) de novo and known TE detection, ii) de novo and known gene models detection, and iii) functional gene annotation. De novo TE detection uses RepeatModeler2 (v2.0.3) followed by classification using RepeatClassifier (v2.0.3) and TEclass (v2.1.3). The de novo TEs obtained were merged with the known TE accessions for Viridiplantae from RepBase (v27.07). This consolidated database is used for TE detection in each genome prior to soft masking using RepeatMasker (v2.0.3). Detection of de novo and known gene models is based on BRAKER2 and its dependencies. Finally, functional annotation of candidate genes is based on the Trinotate (v3.2.1) pipeline using TransDecoder (v5.5.0), TMHMM, HMMER, BLAST (v2.13.0), RNAmmer (v1.2), and SignalP (v4.1) with UniProt and Pfam databases (*SI Appendix, Note S4*).

Leaf and Cambium Mutation Detection. Genomic DNA was extracted from 30 mg of frozen leaf or cambium tissue per sample point for both trees using a CTAB protocol with chloroform–isoamyl alcohol extraction (24:1), isopropanol precipitation and resuspension of the pellet in 1x Low TE (10 mM Tris-HCl + 0.1 mM EDTA). DNA was quantified using a Qubit HS assay (Thermo Fisher Scientific, Waltham, MA) and purified with AMPure XP beads (Beckman Coulter Genomics, Danvers, MA) where necessary to allow library preparation. An Illumina sequencing library was produced for each leaf using an optimized NEBNext Ultra II DNA library protocol (New England Biolabs, Ipswich, MA). Libraries were pooled into multiplexes after independently labeling each library prior to whole genome sequencing on an S4 flow cell and NovaSeq 6000 instrument with v1.5 chemistry (2 × 150 PE mode). We obtained 33 cambium and leaf libraries for *D. guianensis* with a sequencing depth of about 160× and 27 libraries with a depth of about 80× for *S. rubra*. We took advantage of a workflow previously developed by some of us to detect somatic mutations from sequencing reads mapped to a genomic ref. 6. Paired sequencing reads from each library were quality controlled using FastQC (v0.11.9) before being trimmed using Trimmomatic (v0.39), which retains only paired-end reads without adapters and with a phred score greater than 15 in a 4-base sliding window. The reads are aligned against the reference genome using BWA mem with the option to mark shorter splits (v0.7.17). The alignments are then compressed using Samtools view in CRAM format, sorted by coordinates using Samtools sort, and indexed using Samtools index (v1.10). Duplicate reads in the alignments are marked using GATK MarkDuplicates (v4.2.6.1). Sequencing depth was estimated along the genome using Mosdepth (v0.2.4) globally and over a sliding window of 1 kb. We used Jellyfish (v1.1.12) and GenomeScope to estimate heterozygosity up to 21-mer. We used GATK (HaplotypeCaller, GatherGVCFs, GenomicsDBImport, GenotypeGVCFs) to call heterozygous sites from the previously obtained alignments. We filtered single-nucleotide polymorphisms (SNPs) using bcftools (v1.10.2), GATK VariantFiltration (v4.2.6.1), and plink (v1.90), retaining only biallelic SNPs and discarding those with quality less than 30, quality per depth less than 2, Fisher strand ratio greater than 60, and strand odds ratio greater than three. To eliminate all truly heterozygous sites, we further filtered out SNPs present in all sampled genotypes and tissues (no missing data) and shared by at least all but one tissue. Finally, the workflow uses Strelka2 (v2.9.10) to detect mutations by comparing two samples, a mutated sample and a normal (directional) sample. To detect cambium mutations present at the base of the tree, we compared all potential pairs (six in total) among the three cambium libraries. To detect leaf mutations, we compared each leaf library to the first cambium library as a reference sample. We filtered candidate leaf mutations discarding previously identified heterozygous sites and all candidate mutations from all cambium comparisons using BEDTools subtract (v2.29.2). We also filtered candidate leaf mutations using the following criteria: i) no copies of the mutated allele in the reference sample, in this case the cambium sample; ii) a read depth for both samples between the 5th quantile and the 95th quantile of the corresponding library coverage; and iii) the presence of the mutation in at least two biological replicates (at least 2 leaves of the crown) We used the same pipeline and compared mutations detected in two pedunculate oaks *Quercus robur* L. (4, 10), and in a dataset from a tortuous phenotype of common beech *F. sylvatica* (16, 17) (*SI Appendix, Note S5*).

Somatic Mutation Distributions through Physical Trees, Phylogenies, and with Light Exposure. We explored mutation distribution along tree architecture by assuming the origin of the mutation in the tree architecture was at the latest the most recent common branching event among all branches harboring the mutation (11). We further built mutation phylogenies using iqtree rooting the tree with the nonmutated library from the cambium mean genotype without mutations. We compared phylogenies to the physical architecture of both trees with the dendextend R package (*SI Appendix, Note S6*). We explored the effects of light on the occurrence of mutations in the trees using Student's *t* tests and Kolmogorov–Smirnov tests. We compared the number of mutations detected in branches exposed to high vs. low light conditions using the leaves as an observation. We further compared mutation types (base change) and mutation spectra (mutation context with 5′ and 3′ bases) between high and low light conditions among branches of each tree (*SI Appendix, Note S7*).

Low-Frequency Mutations Annotation. We explored the allelic fractions of somatic mutations in relation to tree sequencing depth, a known determinant of the sensitivity of somatic mutation detection (8), for the *D. guianensis* tree, the *S. rubra* tree, two pedunculate oaks *Q. robur* (4, 10), and data from one tortuous phenotype of common beech *F. sylvatica* (16, 17). We further compared mutation annotations in terms of their presence in transposable elements (TE) and genes among trees. We assessed mutation functional impact using SNPeff and related nonsynonymous mutations to their functional annotations, gene ontology, and allelic fraction. We finally explored the allelic fraction of mutations depending on genomic contexts using Student's *t* tests (*SI Appendix, Note S8*).

Detection of Fruit Mutations. We explored mutation transmission to fruits using amplicon resequencing. We kept as candidate mutations for redetection only mutations present in at least three leaves from the branches that had fruits during sampling for resequencing, which resulted in 160 candidate mutations (124 for *D. guianensis* and 36 for *S. rubra*). Frozen fruits were dissected in 4 tissues: i) embryo sac, ii) nucellus, iii) pericarp, and iv) fruit base. Genomic DNA was extracted from 10 to 50 mg of frozen fruit tissue for both trees and additional leaf tissue for positive control with a CTAB protocol with chloroform–isoamyl alcohol (24:1) extraction, isopropanol precipitation and pellet resuspension in 1x Low TE (10 mM Tris-HCl + 0.1 mM EDTA). DNA was quantified using a Qubit HS assay. Primer3plus was used to design primer pairs targeting candidate mutations (amplicon size between 100 and 200 pb). Only one *D. guianensis* candidate mutation failed to yield a primer pair. Illumina universal tags were added to the 5′ end of the forward and reverse primer sequences respectively. Oligonucleotides were ordered in a plate format from Integrated DNA Technologies (Coralville, IA) with standard desalt purification at 25 nmoles synthesis scale. Each primer pair was tested using simplex PCR amplification of

one DNA sample per species in a volume of 10 μL containing 2 μL of 5× Hot Firepol Blend master mix (Solis Biodyne, Tartu, Estonia), 1 μL of 2 μM primer pairs, 1 μL of DNA (10 ng/μL), and 6 μL of PCR-grade water. We ran the PCR on a Veriti 96-Well thermal cycler (Applied Biosystems, Waltham, MA) performing an initial denaturation at 95 °C for 15 min, followed by 35 cycles of denaturation at 95 °C for 20 s, annealing at 59 °C for 60 s, extension at 72 °C for 30 s, and a final extension step at 72 °C for 10 min. We checked the amplification on a 3% agarose gel. A total of six *D. guianensis* primer pairs that failed to amplify were discarded at this stage. The remaining 101 *D. guianensis* and 33 *S. rubra* primer pairs, targeting respectively 117 and 36 mutations, were grouped accounting for potential primer dimer formation using PrimerPooler for subsequent multiplex PCR amplification. Four multiplexed PCRs were done for each species in a volume of 10 μL using 2 μL of 5× Hot Firepol Multiplex master mix (Solis Biodyne), 1 μL of multiplex primer mix (0.5 μM of each primer), 2 μL of DNA (10 ng/μL), and 5 μL of PCR-grade water. Amplifications were performed on a Veriti 96-Well thermal cycler (Applied Biosystems) using an initial denaturation at 95 °C for 12 min followed by 35 cycles of denaturation at 95 °C for 30 s, annealing at 59 °C for 180 s, extension at 72 °C for 30 s, and a final extension step at 72 °C for 10 min. The amplicons from the four multiplexed PCRs of each sample were pooled. Illumina (San Diego, CA) adapters and sample-specific Nextera XT index pairs were added to the amplicons by a PCR targeting the Illumina universal tags attached to the locus-specific primers. This indexing PCR was done in a volume of 20 μL using 5× Hot Firepol Multiplex master mix, 5 μL of amplicon, and 0.5 μM of each of the forward and reverse adapters, using an initial denaturation at 95 °C for 12 min followed by 15 cycles of denaturation at 95 °C for 30 s, annealing at 59 °C for 90 s, extension at 72 °C for 30 s, and a final extension step at 72 °C for 10 min. We then pooled the libraries and purified them with 0.9× AMPure XP beads (Beckman Coulter, Brea, CA). We checked the library quality on a Tapestation 4200 (Agilent Technologies, Santa Clara, CA) and quantified it using QIAseq Library Quant Assay kit (Qiagen, Hilden, Germany) on a LightCycler 480 quantitative PCR (Roche, Basel, Switzerland). The sequencing was done on an Illumina MiSeq sequencer using a V2 flow cell with a 2 × 150 bp paired-end sequencing kit. Paired-end sequencing reads of each library were trimmed using Trimmomatic (v0.39) keeping only paired-end reads without adaptors and a phred score above 15 in a sliding window of 4 bases. Reads were aligned against the reference genome using BWA mem with the option to mark shorter splits (v0.7.17). Alignments were then compressed using Samtools view in CRAM format, sorted by coordinates using Samtools sort, and indexed using Samtools index (v1.10). We used GATK (HaplotypeCaller, GatherGVCFs, GenomicsDBImport, GenotypeGVCFs) to call candidate mutations transmitted to fruits from previously obtained alignments. We filtered variants that corresponded to the 160 candidate mutations (124 for the *D. guianensis* tree and 36 for the *S. rubra* tree). For stringency, we highlighted and discussed only candidate mutations that were considered a heterozygous site by the GATK GenotypeGVCFs call (*SI Appendix, Note S9*).

Data, Materials, and Software Availability. Genomic and transcriptomic reads from leaf, cambium, and fruits and corresponding genomes are available in GenBank (18). genomeAnnotation and detectMutations pipelines as well as downstream analyses are available on GitHub (19). Results and intermediary files are available on Zenodo (20).

ACKNOWLEDGMENTS. We are grateful to Valentine Alt, Emeline Houël, Laetitia Brechet, Saint-Omer Cazal, and Hadrien Lalagüe for their help with tree climbing and sampling. We thank Olivier Brunaux and Caroline Bedeau for their help in accessing the Regina site and the Office National des Forests data. We are grateful to Nicolas Barbier, Ilona Clocher, and Jean-Louis Smock for their help with lidar acquisition. PacBio HiFi reads were produced at Gentyane Genomic Facility (Institut National de Recherche pour l'Agriculture, l'alimentation, et l'Environnement, Clermont-Ferrand, France). Bionano Saphir reads and genome de novo assembly were produced at Centre National de Ressources Génomiques Végétales (Toulouse, France). Whole genome resequencing was performed at Genoscope National Sequencing Centre (Evry, France) with the help of Eric Mahieu, Corinne Cruaud, and Pedro H. Oliveira. Sequence-based genotyping was performed at the Bordeaux Genome Transcriptome Facility (DOI: 10.15454/1.5572396583599417E12) with the help of Zoé Delporte. We are grateful to the GenoToul bioinformatics facility (Castanet-Tolosan, Toulouse, Occitanie, France, DOI: 10.15454/1.5572369328961167E12) for providing help, computing, and data storage resources The climate data were provided by Météo-France to the Unité Mixte de Recherche Écologie des Forêts de Guyane for research purposes within the framework of a MétéoFrance-INRAE AgroClim agreement. This study was funded by an Agence Nationale de la Recherche Investissement d'Avenir grant: Center for the study of Biodiversity in Amazonia (ANR-10-LABEX-0025).

Author affiliations: [a]CNRS, UMR EcoFoG (Agroparistech, Cirad, INRAE, Université des Antilles, Université de la Guyane), Kourou 97310, French Guiana; [b]CIRAD, UPR Forêts et Sociétés, Montpellier 34398, France; [c]Forêts et Sociétés, Université de Montpellier, CIRAD, Montpellier 34398, France; [d]AMAP, Université de Montpellier, CIRAD, CNRS, INRAE, IRD, Montpellier 34980, France; [e]INRAE, UMR EcoFoG (Agroparistech, CNRS, Cirad, Université des Antilles, Université de la Guyane), Kourou 97310, French Guiana; [f]Genoscope, Institut François Jacob, Commissariat à l'Energie Atomique, Université Paris-Saclay, Evry 91057, France; [g]University of Bordeaux, INRAE, BIOGECO, Cestas 33612, France; [h]INRAE, CNRGV, French Plant Genomic Resource Center, Castanet Tolosan 31326, France; [i]Department of Life Sciences, Centre for Functional Ecology, Associate Laboratory TERRA, University of Coimbra, Coimbra 3000-456, Portugal; [j]URFM, INRAE, Avignon 84000, France; [k]Laboratoire Evolution et Diversité Biologique, UMR5174, CNRS, Université Paul Sabatier, IRD, Toulouse, 31077, France; [l]Department of Botany and Biodiversity Research, University of Vienna, Vienna A-1030, Austria; and [m]GenPhySE, Université de Toulouse, INRAE, ENVT, Castanet Tolosan 31326, France

1. A. Weismann, *The Germ-Plasm: A Theory of Heredity* (Scribner's, 1893).
2. L. A. Bergeron *et al.*, Evolution of the germline mutation rate across vertebrates. *Nature* **615**, 285–291 (2023), 10.1038/s41586-023-05752-y.
3. R. Lanfear, Do plants have a segregated germline? *PLoS Biol.* **16**, 1–13 (2018), 10.1371/journal.pbio.2005439.
4. C. Plomion *et al.*, Oak genome reveals facets of long lifespan. *Nat. Plants* **4**, 440–452 (2018), 10.1038/s41477-018-0172-3.
5. L. Wang *et al.*, The architecture of intra-organism mutation rate variation in plants. *PLoS Biol.* **17**, 1–29 (2019), 10.1371/journal.pbio.3000191.
6. S. Schmitt, T. Leroy, M. Heuertz, N. Tysklind, Somatic mutation detection: A critical evaluation through simulations and reanalyses in oaks. *Peer Community J.* **2**, e68 (2022), 10.24072/pcjournal.187.
7. A. J. Orr *et al.*, A phylogenomic approach reveals a low somatic mutation rate in a long-lived plant. *Proc. R. Soc. B.* **287**, 20192364 (2020), 10.1098/rspb.2019.2364.
8. M. Holá, R. Vágnerová, K. J. Angelis, Mutagenesis during plant responses to UVB radiation. *Plant Physiol. Biochem.* **93**, 29–33 (2015), 10.1016/j.plaphy.2014.12.013.
9. S. Kim *et al.*, Strelka2: Fast and accurate calling of germline and somatic variants. *Nat. Methods* **15**, 591–594 (2018), 10.1038/s41592-018-0051-x.
10. E. Schmid-Siegert *et al.*, Low number of fixed somatic mutations in a long-lived oak tree. *Nat. Plants* **3**, 926–929 (2017), 10.1038/s41477-017-0066-9.
11. Y. Duan *et al.*, Limited accumulation of high-frequency somatic mutations in a 1700-year-old Osmanthus fragrans tree. *Tree Physiol.* **42**, 1–10 (2022), 10.1093/treephys/tpac058.
12. J. L. Bowman, Y. Eshed, Formation and maintenance of the shoot apical meristem. *Trends Plant Sci.* **5**, 110–115 (2000), 10.1016/S1360-1385(00)01569-7.
13. A. Burian, Does shoot apical meristem function as the germline in safeguarding against excess of mutations? *Front. Plant Sci.* **12**, 1–9 (2021), 10.3389/fpls.2021.707740.
14. L. Yu *et al.*, Somatic genetic drift and multilevel selection in a clonal seagrass. *Nat. Ecol. Evol.* **4**, 952–962 (2020), 10.1038/s41559-020-1196-4.
15. A. Cagan *et al.*, Somatic mutation rates scale with lifespan across mammals. *Nature* **604**, 517–524 (2022), 10.1038/s41586-022-04618-z.
16. J. M. Aury, C. Plomion, *Fagus sylvatica* genome version 3 (2023). http://www.genoscope.cns.fr/plants.
17. J. M. Aury, C. Plomion, Illumina reads for two branches of the tortuous beach accession 354. ENA BioProject PRJEB70295 (2023). https://www.ebi.ac.uk/ena/browser/view/PRJEB70295.
18. S. Schmitt *et al.*, Mutation in the tropical tree canopy: Genomic and transcriptomic reads from leaf, cambium, and fruits and corresponding genomes. GenBank BioProject PRJNA823677 (2023). https://dataview.ncbi.nlm.nih.gov/object/PRJNA823677.
19. S. Schmitt, sylvainschmitt/treemutation_codes: v0.1.0 (v0.1.0). Zenodo (2023). https://doi.org/10.5281/zenodo.10090430.
20. S. Schmitt, T. Leroy, M. Heuertz, N. Tysklind, Low-frequency somatic mutations are heritable in tropical trees *Dicorynia guianensis* and *Sextonia rubra* (Version v2) [Data set]. Zenodo (2023). https://doi.org/10.5281/zenodo.10089692.