Between but not within species variation in the Distribution of Fitness Effects

Jennifer James[1,2*], Chedly Kastally[3,4], Katharina B. Budde[5,6], Santiago C. González-Martínez[7], Pascal Milesi[1,8], Tanja Pyhäjärvi[3,4], GenTree Consortium[✦], Martin Lascoux[1]

[1]Department of Ecology and Genetics, Uppsala University; Uppsala, Sweden
[2]Swedish Collegium of Advanced Study, Uppsala University; Uppsala, Sweden
[3]Department of Forest Sciences, University of Helsinki; Helsinki, Finland
[4]Viikki Plant Science Centre, University of Helsinki; Helsinki, Finland
[5]Department of Forest Genetics and Forest Tree Breeding, Georg-August-University Goettingen; Goettingen, Germany
[6]Center of Biodiversity and Sustainable Land Use (CBL), University of Goettingen, Goettingen, Germany
[7]National Research Institute for Agriculture, Food and the Environment (INRAE); University of Bordeaux; BIOGECO; Cestas, France
[8]Science for Life Laboratory (SciLifeLab); Uppsala, Sweden

[*]Corresponding author. Email: Jennifer.James@ebc.uu.se

[✦]GenTree Consortium:
Paraskevi Alizoti[1], Ricardo Alía[2], Olivier Ambrosio[3], Filippos A Aravanopoulos[1], Georg von Arx[4], Albet Audrey[5], Francisco Auñón[2], Camilla Avanzi[6], Evangelia Avramidou[1], Francesca Bagnoli[7], Marko Bajc[8], Eduardo Ballesteros[2], Evangelos Barbas[1], José M García del Barrio[2], Cristina C Bastias[9], Catherine Bastien[10], Giorgia Beffa[11], Raquel Benavides[12], Vanina Benoit[13], Frédéric Bernier[5], Henri Bignalet[5], Guillaume Bodineau[14], Damien Bouic[5], Sabine Brodbeck[11], William Brunetto[15], Jurata Buchovska[16], Corinne Buret[13], Melanie Buy[17], Ana M Cabanillas-Saldaña[18], Bárbara Carvalho[12], Stephen Cavers[19], Fernando Del Caño[2], Sandra Cervantes[20,21], Nicolas Cheval[5], José M Climent[2], Marianne Correard[22], Eva Cremer[23], Darius Danusevičius[16], Benjamin Dauphin[24], Jean-Luc Denou[5], Bernard Dokhelar[5], Alexis Ducousso[25], Bruno Fady[26], Patricia Faivre-Rampant[27], Anna-Maria Farsakoglou[1], Patrick Fonti[4], Ioannis Ganopoulos[28], Olivier Gilg[22], Nicolas De Girardi[29], René Graf[11], Alan Gray[30], Delphine Grivet[31], Felix Gugerli[24], Christoph Hartleitner[32], Katrin Heer[33], Enja Hollenbach[34], Agathe Hurel[25], Bernard Issenhuth[5], Florence Jean[15], Véronique Jorge[35], Arnaud Jouineau[36], Jan-Philipp Kappner[34], Robert Kesälahti[37], Florian Knutzen[23], Sonja T Kujala[38], Timo A Kumpula[37], Katri Kärkkäinen[38], Mariaceleste Labriola[39], Celine Lalanne[25], Johannes Lambertz[34], Gregoire Le-Provost[25], Vincent Lejeune[14], Isabelle Lesur-Kupin[40,41],

Joseph Levillain[42], Mirko Liesebach[43], David López-Quiroga[12], Ermioni Malliarou[1], Jérémy Marchon[11], Nicolas Mariotte[36], Antonio Mas[12], Silvia Matesanz[44], Benjamin Meier[11], Helge Meischner[34], Célia Michotey[17], Sandro Morganti[11], Tor Myking[45], Daniel Nievergelt[4], Anne Eskild Nilsen[45], Eduardo Notivol[46], Dario I. Ojeda[47], Sanna Olsson[31], Lars Opgenoorth[24,48], Geir Ostreng[45], Birte Pakull[43], Annika Perry[30], Sara Pinosio[7,49], Andrea Piotti[6], Christophe Plomion[40], Nicolas Poinot[5], Mehdi Pringarbe[22], Luc Puzos[5], Annie Raffin[5], José A Ramírez-Valiente[2], Christian Rellstab[24], Dourthe Remi[5], Oliver Reutimann[11], Sebastian Richter[34], Juan J Robledo-Arnuncio[2], Odile Rogier[35], Elisabet Martínez Sancho[4], Outi Savolainen[37], Simone Scalabrin[50], Volker Schneck[51], Silvio Schueler[52], Ivan Scotti[26], Sergio San Segundo[2], Vladimir Semerikov[53], Lenka Slámová[4], Ilaria Spanu[54], Jørn Henrik Sønstebø[45], Jean Thevenet[22], Mari Mette Tollefsrud[45], Norbert Turion[22], Fernando Valladares[12], Giovanni G. Vendramin[7], Marc Villar[55], Marjana Westergren[56], Johan Westin[57]

[1]Aristotle University of Thessaloniki, School of Forestry and Natural Environment, Laboratory of Forest Genetics and Tree Improvement, 541
[2]Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria - Centro de Investigación Forestal (INIA-CIFOR), Ctra. de la Coruña km 7.5, 28040, Madrid, Spain
[3]INRAE, URFM F-84914, Avignon, France
[4]Forest Dynamics, Swiss Federal Research Institute WSL, 8903 Birmensdorf, Switzerland
[5]INRAE, UEFP, F-33610, Cestas, France
[6]Institute of Biosciences and Bioresources, National Reaseach Council of Italy
[7]Institute of Biosciences and Bioresources, National Reasearch Council of Italy (IBBR-CNR), 50019 Sesto Fiorentino, Italy
[8]Slovenian Forestry Institute, Vecna pot 2, 1000 Ljubljana, Slovenia
[9]Centre d'Ecologie Fonctionnelle et Evolutive (CEFE), CNRS, UMR 51
[10]INRAE, Dept ECODIV, F-45075, Orléans, France
[11]Biodiversity & Conservation Biology, Swiss Federal Research Institute WSL, 8
[12]LINCGlobal, Department of Biogeography and Global Change, Museo Nacional de Ciencias Naturales, CSIC, Serrano
[13]INRAE, ONF, BioForA, F-45075, Orléans, France
[14]INRAE, GBFOR, F-45075, Orléans, France
[15]INRAE, URFM, F-849
[16]Vytautas Magnus University, Studentu Street 11, 53361, Akademija, Lithuania
[17]INRAE, URGI, F-78026, Versailles, France
[18]Departamento de Agricultura, Ganadería y Medio Ambiente, Gobierno de Aragón, P. Mª Agustín 36, 50071, Zaragoza, Spain
[19]UK Centre for Ecology & Hydrology (UKCEH), EH26 0QB Bush Estate, United Kingdom
[20]Department of Ecology and Genetics, University of Oulu, 90014 Oulu, Finland
[21]Biocenter Oulu, University of Oulu, 90014 Oulu, Finland
[22]INRAE, UEFM, F-84914, Avignon, France
[23]Bavarian Institute for Forest Genetics, Forstamtsplatz 1, 83317, Teisendorf, Germany
[24]Biodiversity and Conservation Biology, Swiss Federal Research Institute WSL, 8903 Birmensdorf, Switzerland
[25]INRAE, Université de Bordeaux, BIOGECO, F-33770, Cestas, France

[26]National Research Institute for Agriculture, Food and the Environment (INRAE), 84914 Avignon, France
[27]University of Paris-Saclay, INRAE, Study of Plant Genome Polymorphism, 91000 Evry-Courcouronnes, France
[28]Institute of Plant Breeding and Genetic Resources, Hellenic Agricultural Organization DEMETER (ex NAGREF), 57001, Thermi, Greece
[29]Swiss Federal Research Institute WSL, 8903 Birmensdorf, Switzerland
[30]UK Centre for Ecology and Hydrology, Bush Estate Penicuik, EH26 0QB, Edinburgh, UK
[31]Institute of Forest Sciences (ICIFOR-INIA), CSIC, 28040 Madrid, Spain
[32]LIECO GmbH & Co KG
[33]Forest Genetics, Albert-Ludwigs Universität Freiburg, Bertoldstraße 17, 79098 Freiburg, Germany
[34]Philipps University Marburg, Faculty of Biology, Plant Ecology and Geobotany, Karl-von-Frisch Strasse 8, 35043, Marburg, Germany
[35]INRAE, ONF, BioForA, 45075 Orléans, France
[36]INRAE, URFM, F-84914, Avignon, France
[37]University of Oulu, Pentti Kaiteran katu 1, 90014, University of Oulu, Finland
[38]Natural Resources Institute Finland, Paavo Havaksentie 3, 90014, University of Oulu, Finland
[39]Institute of Biosciences and BioResources, National Research Council (CNR), via Madonna del Piano 10, 50019, Sesto, Fiorentino, Italy
[40]University of Bordeaux, INRAE, BIOGECO, 33610 Cestas, France
[41]Helix Venture, 33700 Mérignac, France
[42]Université de Lorraine, AgroParisTech, INRAE, SILVA, 54000, Nancy, France
[43]Thünen Institute of Forest Genetics, Sieker Landstr. 2, 22927, Grosshansdorf, Germany
[44]Área de Biodiversidad y Conservación, Universidad Rey Juan Carlos, Calle Tulipán s/n, 28933, Móstoles, Spain
[45]Division of Forestry and Forest Resources, Norwegian Institute of Bioeconomy Research (NIBIO), P.O. Box 115, 1431, Ås, Norway
[46]Centro de Investigación y Tecnología Agroalimentaria de Aragón -Dpto. de Sistemas Agrarios, Forestales y Medio Ambiente (CITA), Avda. Montañana 930, 50059, Zaragoza, Spain
[47]Norwegian Institute of Bioeconomy Research (NIBIO), 8027 Bodø, Norway
[48]Plant Ecology and Geobotany, Philipps-Universität Marburg, 35043 Marburg, Germany
[49]Institute of Applied Genomics (IGA), 33100 Udine, Italy
[50]IGA Technology Services S.r.l., 33100 Udine, Italy
[51]Thünen Institute of Forest Genetics, Eberswalder Chaussee 3a, 15377, Waldsieversdorf, Germany
[52]Austrian Research Centre for Forests (BFW), Seckendorff-Gudent-Weg 8, 1131, Wien, Austria
[53]Institute of Plant and Animal Ecology, Ural branch of RAS, 8 Marta St. 202, 620144, Ekaterinburg, Russia
[54]Institute of Biosciences and BioResources, National Research Council (CNR), via Madonna del Piano 10, 50019, Sesto Fiorentino, Italy
[55]INRAE, ONF, BioForA, F-45075 Orléans, France
[56]Slovenian Forestry Institute, 1000 Ljubljana, Slovenia
[57]Skogforsk, Tomterna 1, 91821, Sävar, Sweden

Abstract

New mutations provide the raw material for evolution and adaptation. The distribution of fitness effects (DFE) describes the spectrum of effects of new mutations that can occur along a genome, and is therefore of vital interest in evolutionary biology. Recent work has uncovered striking similarities in the DFE

between closely related species, prompting us to ask whether there is variation in the DFE among populations of the same species, or among species with different degrees of divergence, *i.e.*, whether there is variation in the DFE at different levels of evolution. Using exome capture data from six tree species sampled across Europe we characterised the DFE for multiple species, and for each species, multiple populations, and investigated the factors potentially influencing the DFE, such as demography, population divergence and genetic background. We find statistical support for there being variation in the DFE at the species level, even among relatively closely related species. However, we find very little difference at the population level, suggesting that differences in the DFE are primarily driven by deep features of species biology, and that evolutionarily recent events, such as demographic changes and local adaptation, have little impact.

Keywords: DFE, deleterious mutations, population structure, forest trees

Introduction

The distribution of fitness effects of new mutations (DFE), *i.e.*, the proportion of new mutations that are expected to be adaptive, neutral, slightly deleterious or strongly deleterious, is at the heart of any evolutionary model, yet, in spite of recent progress (for a review, see Johri et al. 2022) it is still hard to estimate and is poorly understood. While there is variation in the DFE across distantly related species with dissimilar biological features (Huber et al. 2017), on shorter evolutionary timescales it is not clear how the DFE might come to differ among species or populations, although we can make some predictions from the Nearly Neutral Theory (Ohta 1973). In particular, the strength of selection acting on new mutations is expected to scale with effective population size, N_e, and therefore to be affected by demographic processes. We also expect that the fraction of mutations inferred to be nearly neutral, *i.e.*, slightly deleterious, will be related to proxies of N_e. In particular, the ratio of slightly deleterious to neutral diversity will be smaller in high N_e populations (Welch, Eyre-Walker, and Waxman 2008).

Despite these predictions, empirical evidence has been mixed. Major evolutionary transitions do affect the DFE. For instance, a shift in mating systems from outcrossing to selfing leads to a lower N_e and a significant increase in the fraction of slightly deleterious mutations (e.g., Douglas et al. 2015), as predicted under Nearly Neutral Theory. However, a number of studies have found that across closely related species the DFE and related summary statistics, such as the ratio of nonsynonymous to synonymous nucleotide diversity, π_N/π_S, are remarkably stable (Castellano et al. 2019; Grivet et al. 2017; Liu et al. 2022), even when comparing domesticated species and their wild relatives (Chen, Glémin, and

Lascoux 2017). In the latter, domestication has a very strong effect on synonymous nucleotide diversity but the ratio of nonsynonymous to synonymous nucleotide variation, a good proxy of the slightly deleterious class of mutations for populations at demographic equilibrium (Ohta 1973), was barely affected. Additionally, while some studies have found associations between parameters associated with the DFE and demographic processes such as range expansion (González-Martínez, Ridout, and Pannell 2017; Willi et al. 2020), others have not (Takou et al. 2021).

These contrasting results may reflect real biological and demographic differences across species and populations. Species may also experience different environmental conditions across their ranges, which could result in changes in the parameters of the DFE. For example, Martin and Lenormand (2006) found evidence to support a scenario in which mutations have more variable fitness effects when an organism exists in an environment to which it is less well adapted. They interpreted this result in terms of a simple fitness landscape model. A recent study in *A. thaliana (*Weng et al. 2021) also found that mutational variance was greater in populations growing in stressful environments in which their fitness was low. However, not all of the results in Weng et al. agree with the predictions of a simple fitness landscape model. For example, the authors found that beneficial mutations were more common in populations in less stressful environments. Additionally, a review of the impact of environment on the effects of new mutations found that environmental stress can both decrease and increase the mean strength of selection acting on new mutations, as well as its variance (Agrawal and Whitlock 2010). Population differentiation may also be important, with more differentiated populations appearing to have less similar strengths of selection acting on shared mutations than less differentiated populations (Huang et al. 2021). Whether this could lead to differences in the DFE between populations given enough evolutionary time has not yet been systematically investigated.

However, contrasting results across species and populations might also be due to differences in metrics used to characterize patterns of deleterious and neutral diversity. It has been argued that while summary statistics such as the ratio of nonsynonymous to synonymous nucleotide diversity provide a good measure of the efficiency of selection, they are poor measures of the deleterious genetic load experienced by a population due to the effects of demography and non-equilibrium dynamics. For instance, after a demographic event, slightly deleterious nonsynonymous mutations will reach their equilibrium frequency spectra more rapidly than synonymous mutations, simply because the equilibrium frequencies of slightly deleterious mutations are lower (Simons et al. 2014; Simons and Sella 2016). Counts of nonsynonymous derived alleles are more robust to non-equilibrium dynamics, and give a good measure of load if mutations are deleterious, and their effects are additive. Therefore, metrics such as R_{xy}, which were

specifically developed for the purpose of estimating asymmetries in counts of derived mutations between populations, provide a better proxy of genetic load (Do et al. 2015). A combination of such metrics, in addition to those based on the site frequency spectrum, may allow for a greater understanding of how new mutations affect the molecular evolution of populations and species differ.

In the present study we investigated variation in the DFE at both the species and population levels by leveraging exome capture data collected from range-wide populations of six forest tree species, comprising four angiosperms and three conifers, at different degrees of phylogenetic distance. These trees are keystones of European forests with a range of life history traits. All species are widely distributed, but there are marked differences in levels of population differentiation within species (Supplementary Table 1). By using orthologous genomic regions, we were able to compare the DFE among species while controlling for gene content. Additionally, all species have been sampled broadly across their natural ranges, following the same sampling scheme, providing us with an ideal dataset to assess the constancy of the DFE at the within-species level. Finally, we also explored variation in patterns of genetic load between populations.

2. Methods

2.1 Samples

The data consists of six wind-pollinated forest tree species (6), two conifers (*Picea abies*, and *Pinus pinaster*), and four angiosperms (*Betula pendula*, *Fagus sylvatica*, *Populus nigra*, and *Quercus petraea*), distributed across Eurasia from the boreal to the Mediterranean region, and with either animal-, wind-, or water-dispersed seeds. The species vary in both life history and population structure (Milesi et al. 2023; Supplementary Table 1).

2.2 Sequencing and SNP calling

Sequencing and SNP calling was as described in Milesi et al. (2023). Briefly, the data is the result of targeted nuclear DNA sequencing (~10,000 species-specific probes that covered ~3 Mb of sequence) on a total of 3,407 adult trees collected from 19 to 26 locations per species (~25 samples each) across their distribution range. The targeted regions primarily consisted of orthologous regions among species, in addition to regions that had previously been identified as targets of selection. Site based annotation (4-

fold degenerate, 0-fold degenerate sites) of detected SNPs was generated using the python script NewAnnotateRef.py available at https://github.com/fabbyrob/science/blob/master/pileup_analyzers/NewAnnotateRef.py (Williamson et al. 2014). Detected SNPs were functionally annotated in order to predict their effects on protein sequences using the tool ANNOVAR (Wang, Li, and Hakonarson 2010). SNPs were classified as 'non-coding'; 'coding 4-fold degenerate synonymous'; 'coding 0-fold degenerate nonsynonymous'; and 'nonsense' (determining a premature STOP codon or a STOP loss). Filtering steps were applied in order to remove incorrectly assigned or clear hybrid samples. Full documentation of bioinformatics pipelines used to generate these VCF files are available at https://github.com/GenTree-h2020-eu/GenTree. The VCF files used in the present study correspond to version 5.3.2, available at https://entrepot.recherche.data.gouv.fr/dataset.xhtml?persistentId=doi:10.57745/DV2X0M. In order to be included in our analyses, both polymorphic and monomorphic sites had to have a call depth >8 or genotype quality >20. Loci with >50% missing calls were also removed. SNPs and monomorphic sites were further restricted to those that are either 4-fold or 0-fold degenerate. An additional subdivision of our SNP dataset was created, which included only those SNPs that occur in orthologous genomic regions found in all six tree species.

2.3 SNP polarisation

To increase the power of our DFE estimation methods, we inferred the ancestral state at each SNP. This was achieved by considering the state of the site in either a single outgroup species (two species in our dataset, Supplementary Table 2) or two outgroup species (four species, Suppl. Table 2). For each species, we therefore mapped the genome of one or more outgroup species to the same reference genome used for SNP calling for that species using the bwa software package (Li and Durbin 2009); for further details and commands used, see Suppl. Table 2. We also retained SNP sites that could not be matched to a site in an outgroup species (see below), due, for example, to being missing in the outgroup species genome. We used the maximum likelihood method implemented in Est-SFS (Keightley and Jackson 2018) for assigning the ancestral allele at polymorphic sites, assuming the Kimura 2-parameter substitution model. To conduct this step, we first down-sampled to a maximum number of 100 haplotypes per species by sampling randomly from a hypergeometric distribution to account for missing data and to not exceed the maximum permissible number of haplotypes for Est-SFS. We then used the probability associated with the state of each SNP to assign likely ancestral states, removing SNPs for which the probability of the major allele being the ancestral state was between 0.4 and 0.6, and which we are therefore not able to polarise with confidence. So, SNPs for which there was no outgroup information available could therefore

still be assigned an ancestral state based on their minor allele frequency; however, we note that this is a small fraction of SNPs, and that all downstream analyses account for errors in ancestral state identification. We used a model averaging procedure to assess the effect of accounting for error in ancestral state identification on DFE inference (see section 2.6); additionally, we assessed how restricting our dataset to GC conservative mutations, which are less likely to be affected by polarisation error due to the exclusion of CpG hypermutable sites, affects our results.

2.4 Grouping samples

For downstream analyses, we were interested in investigating variation in the DFE across a species range. DFE inference power depends on the number of sequenced individuals, and number of available SNPs; we therefore pooled individuals into groups based on sampling location (see Supplementary Fig. 1 for a map of sampling locations). This was first achieved by taking all individuals per country; subsequently, if multiple distinct admixture groups were present in this 'country' pool of individuals, as identified in Milesi et al. (2023), this pool was subdivided further based on the admixture groups. If any pool contained fewer than 20 individuals it was not included in our analysis in the interest of maintaining sufficient power to achieve accurate results. We will refer to these pools as 'populations', full details of which can be found in Supplementary Table 3. We also calculated the mean latitude and longitude of each sampling location per population.

2.5 Summary statistics

We inferred a number of standard population genetic summary statistics including Wright's fixation index, F_{ST} (as calculated over 4-fold sites), and 0- and 4-fold pairwise nucleotide site diversity, π_0 and π_4, respectively, after first projecting our data down to an SFS of 40 haplotypes, *i.e.*, 20 individuals per species or population, to account for any sites with missing data. Projecting takes the average across every possible resampling of the data, as implemented in python using functions in the δaδi package (Gutenkunst et al. 2010). Those pairwise nucleotide diversity estimates were then used to calculate the ratio π_0/π_4. For each population, pairwise F_{ST} was calculated with python scripts, as implemented in δaδi (Gutenkunst et al. 2010). For each species, we identified the median longitude and latitude of sampling locations, and chose as a reference the pooled population sampled closest to this location, which represents a 'central' population to the species range. These 'central' populations were DE for *B. pendula*, CH for *F. sylvatica*, LT for *P. abies*, FR-North for *P. pinaster*, IT-North for *P. nigra*, and CH for *Q. petraea* (for location details, see Supplementary Table 3).

Finally, we also inferred R_{xy}, an estimator of the differences of genetic load between populations, as defined in Do et al. (2015). Briefly, R_{xy} measures the average difference in the accumulation of mutations between two genomes sampled in different populations at all sites for which the ancestral state is known. One counts the number of derived mutations in genome X that are not present in genome Y and vice versa, and R_{xy} is defined as the ratio of these two counts. If selection has been equally effective and mutation rates have been the same since the populations diverged, R_{xy} is expected to equal 1. This statistic was shown to be monotonically related to the difference in mutation load between the two populations. We followed Do et al. (2015) in calculating confidence intervals on this estimate using a weighted block jack-knife procedure whereby SNP data was divided into 100 'consecutive' blocks and R_{xy} was recalculated, removing 1 block per run. Each VCF was first split into chunks of length 2Mb, based on SNP position in the assembled genome, and then these chunks were combined into 100 groups of similar length. This grouping was done such that consecutive parts of the genome were kept together, although small scaffolds meant that occasionally different scaffolds were combined into a single group. As before, we used our estimated 'central' population per species as the reference when presenting results, but results are very similar when different populations are used as reference. We also calculate R'_{xy} for 0-fold degenerate, nonsynonymous sites, a measure which is normalised using putatively neutral 4-fold degenerate synonymous sites, by dividing R_{xy} for 0-fold sites by R_{xy} for 4-fold sites. The custom scripts used to calculate all summary statistics are available at: https://github.com/j-e-james/TreeDFEScripts.

2.6 DFE inference

The DFE was primarily inferred using polyDFE (Tataru et al. 2017; Tataru and Bataillon 2019). PolyDFE implements a likelihood-based approach, and simultaneously infers the DFE while also accounting for the effects of other distorters of the SFS such as demography and errors in SNP polarization through the incorporation of nuisance parameters (Eyre-Walker, Woolfit, and Phelps 2006), which are inferred for each category of the SFS. This method requires the specification of a class of neutral and a class of non-neutral sites, for which we used 0- and 4-fold degenerate sites, thereby avoiding site-counting issues that arise with 2- and 3-fold degenerate sites. As no change in 4-fold degenerate sites results in a change in amino acid, they are the best proxy for neutrally evolving sites in coding DNA, although it is possible that there is selection on synonymous codon usage in the species (Duret 2002). We then inferred the neutral and non-neutral site frequency spectra across species and populations, after first projecting our data down to the same number of individuals to account for any sites with missing data. Our analyses were run on

data projected down to 40 haplotypes, *i.e.*, 20 individuals. All scripts required for the processing of data, and the polyDFE input files used in this analysis, are available at (https://github.com/j-e-james/TreeDFEScripts). PolyDFE allows for the fitting of both deleterious-only DFEs and mixed DFEs, which account for the possible effects of beneficial mutations on DFE inference. From the DFE fitted for beneficial mutations, polyDFE is also able to estimate α, henceforth referred to as α_{DFE}, the rate of adaptive molecular evolution. In polyDFE, this is calculated from the full DFE for beneficial mutations, however, this may inflate estimates of α_{DFE} due to the inclusion of beneficial mutations with very small selective effects. We therefore followed Galtier et al. (2016) by incorporating a lower bound of 5 for the population selection coefficients of positive mutations to be used in the calculation of α_{DFE}, as implemented in polyDFE v. 2.0 (Tataru et al. 2017). This lower bound is arbitrary, and changing it will have an impact on the estimated value of α_{DFE}. Finally, we note that we only use polymorphism data when running PolyDFE, to avoid having to make the assumption that the DFE is invariant between the ingroup and outgroup species. PolyDFE is able to estimate the deleterious DFE accurately without divergence information, and the inclusion of divergence data provides little or no improvement to estimates of the beneficial DFE (Booker 2020, Tataru et al. 2017).

To ensure that, as far as possible, our polyDFE runs explored the full range of parameter space when estimating the DFE, we ran polyDFE a minimum of five times per species, using different starting parameters for each run (Supplementary Table 4). Runs in which parameters were close to the edges of their permitted ranges were removed; we then assessed whether our runs reliably returned similar estimated DFE parameters and had a small gradient of the likelihood. As DFE estimation entails considerable uncertainty, we then ran polyDFE a further four times per species, initialising runs using both analytically estimated parameters and those parameters previously found to return the smallest gradient of the likelihood, fitting a different model per run: in model 1 we fit a full (deleterious and advantageous mutations) DFE, including an estimation of the rate of misidentification of the ancestral allele, ε_{anc}; in model 2 we fit a full DFE, without including the estimation of ε_{anc}; in model 3 we fit a deleterious mutation-only DFE; in model 4 we again fit a deleterious mutation-only DFE, but without including an estimation of ε_{anc}. All models include an estimation of nuisance parameters, which account for the effects of demography. We then performed model averaging over the four models, as described in Muyle et al. (2021), such that models are weighted by their AICs to account for uncertainty in parameter estimation. We calculated AIC weights and generated bootstrap datasets using the R functions provided in polyDFE (Tataru et al. 2017), available at https://github.com/paula-tataru/polyDFE. All other scripts used to conduct these analyses are available at https://github.com/j-e-james/TreeDFEScripts.

PolyDFE v2.0 (Tataru and Bataillon 2019) enables the simultaneous fitting of DFEs to multiple datasets, allowing for model comparisons to assess whether models in which DFE parameters differ between datasets provide a significantly better fit than models in which the DFE parameters are shared between datasets. In situations in which we were interested in comparing models (e.g. comparing populations), we inferred a full DFE, including nuisance parameters and ε_{anc}, allowing these parameters to vary between datasets, as recommended by Tataru et al. (2019), to account for differences in ancestral identification error and demographic processes between comparisons.

2.7 Statistical analyses

We considered correlations between our inferred DFE parameters, life history traits, and population genetics summary statistics. All statistical analyses and plotting was conducted in R, using scripts available at https://github.com/j-e-james/TreeDFEScripts.

3. Results

3.1 Summary statistics

Our dataset comprises polarized SNPs from approximately 3 Mb of targeted genome sequencing for six European tree species which were sampled broadly across their range, with approximately 25 individuals sequenced per location. For all populations across species (see 'population' definition above), we calculated population genetic summary statistics to investigate the efficiency of selection across species and among populations within species.

Species vary broadly in π_0/π_4, the efficiency of purifying selection (Fig. 1A), with selection appearing to be comparatively inefficient in *P. pinaster* and *P. abies* relative to broad-leaved species such as *B. pendula*. This may reflect differences among species in effective population size. However, it does not clearly relate to levels of panmixia, despite the species exhibiting different degrees of genetic differentiation across their ranges (Fig. 1B). *B. pendula* exhibits very little population differentiation and has the most efficient selection of the six species. For *F. sylvatica*, *Q. petraea*, *B. pendula* and *P. abies*, despite their broad geographic ranges, the efficiency of purifying selection was similar among populations within a species. By contrast, *P. nigra* and *P. pinaster* have the highest levels of population genetic structure, with strongly differentiated and isolated Moroccan populations, and there is a relationship with

latitude and π_0/π_4 for both species (*P. nigra* $R^2 = 0.91$, $p = 0.0019$ and *P. pinaster* $R^2 = 0.55$, $p = 0.04$), with π_0/π_4 being lowest in populations at lower latitudes for both *P. nigra* and *P. pinaster*. However, these species have intermediate values of π_0/π_4 when comparing among species. *P. abies*, *F. sylvatica* and *Q. petraea* show moderate levels of structure, with population F_{ST} increasing with latitude, and while *F. sylvatica* and *Q. petraea* have intermediate values of π_0/π_4, *P. abies* has the least efficient selection of any of the species studied.

It has been argued that metrics measuring the ratio of nonsynonymous to synonymous (or 0- to 4-fold degenerate) diversity are poor measures of genetic load (Do et al. 2015). We therefore also estimated the statistic R_{xy}, which compares the frequency of derived alleles between a focal (X) and reference (Y) population. The neutral expectation is that the number of derived alleles is the same in the focal population as in the reference, while values of R_{xy} above 1 indicate that the focal population has an excess of derived alleles.

Comparing focal populations to a single reference population (for which we used a population that was approximately central for the sampling locations per species, Fig. 1C), the most striking results are for *P. nigra* populations MA and GB, which show a deficit of derived alleles relative to the reference population. We also note a slight tendency for low latitude populations of *P. abies* to show a relative deficit of derived alleles, which agrees with our 0- to 4-fold diversity results. However, in the vast majority of populations, we find no deviation from the neutral expectation that the number of derived alleles at 0-fold degenerate sites is the same in the focal as in the reference population. If we use 4-fold degenerate synonymous sites to normalise R_{xy} (R'_{xy}, Fig. 1D), which has been suggested to account for the effects of population structure (Do et al. 2015; Grossen et al. 2020), we find that no population has a significant deficit relative to the focal population. Therefore, although population structure has resulted in a deficit or excess of mutations in some populations, there is little evidence that populations differ in their genetic load.

3.2 Species DFE

We inferred the full DFE for all species, incorporating a gamma-distributed deleterious DFE and an exponential-distributed beneficial DFE, using only polymorphism data (Tataru et al. 2017). The gamma distribution is a flexible and commonly used distribution, and is parameterised by two values: the shape parameter, b, which is inversely related to the coefficient of variation of the strengths of selection acting on new mutations, and the scale parameter, S_d, which is the mean scaled strength of selection ($N_e s$) acting

on new mutations. We also inferred the purely deleterious DFE for all six species to assess whether incorporating beneficial mutations improves our DFE model inference (Fig. 2). We fitted the full DFE and the deleterious-only DFE models both with and without incorporating an estimation of the error rate for the inference of the ancestral state of alleles, ε_{anc}, and conducted a model averaging procedure (see methods and Supplementary Fig. 2 to see the fit of each model, and Supplementary Table 5 for all model averaged inferred parameters for the deleterious and beneficial DFEs), such that the estimated DFE parameters presented here incorporate the degree of model uncertainty (Fig. 3, Table 1). In three of the species in our dataset, incorporating the rate of ancestral misidentification did not significantly improve the fit of the DFE model; while in *F. sylvatica*, *Q. petraea* and *P. abies* we see a small model improvement (p-values of likelihood ratio tests comparing models are 0.018, 0.016, and 0.0039 respectively); these species have high S_d values and a high proportion of adaptive substitutions, which is the selective regime in which we expect the rate of ancestral error to be inflated in polyDFE analyses (Tataru et al. 2017). Generally, the species in our dataset have similar values of b, but vary considerably in S_d. However, S_d values should be interpreted with caution, because S_d is not related to the distribution of selection coefficients of segregating mutations in a straightforward way (see Supplementary Fig. 3 and supplementary text).

In five of the six species, a model incorporating beneficial mutations into estimates of the DFE was the most highly weighted, although only in four species was this model a significantly better fit to the data. Ignoring the contribution of beneficial mutations to the DFE in these species leads to a reduction in the inferred value of b and to an increase in the inferred value of S_d (Figs 2 and 3). We used polyDFE to estimate the rate of adaptive molecular evolution (*i.e.*, the proportion of nonsynonymous substitutions that are beneficial), α_{DFE}, incorporating a lower bound for the minimum strength of selection acting on new mutations. We demonstrate the effects of different bounds on the estimate of α_{DFE} in Supplementary Fig. 4. We find that the rate of adaptive evolution is fairly high in some of the tree species, particularly in *B. pendula* and *P. nigra*, suggesting that adaptive substitutions are common in these forest tree species (Fig. 3).

P. abies and *P. pinaster*, the two conifer species included in this study, are an interesting pair to compare. *P. abies* has remarkably low values of α_{DFE}, which is particularly surprising given the fairly high inferred α_{DFE} in the other conifer in the dataset, *P. pinaster*. These differences may arise due to *P. pinaster* having relatively differentiated populations, which could facilitate local adaptation due to the limited influx of alleles from other populations, whereas *P. abies* has less population structure, *i.e.*, greater levels of admixture, and uniformly high levels of purifying selection across its range (Figs. 1 and 3). It is also

interesting to note that *P. pinaster* has quite a distinct discretised DFE compared to the other species in the dataset, with a high inferred b, a low inferred absolute S_d, and a relatively small estimated fraction of mutations falling into the most strongly deleterious category ($N_e s < -100$). *P. pinaster* has fewer SNPs compared to the other species in the dataset, and so we have less confidence in these results, however, these differences could represent the greater phylogenetic distance between *P. pinaster* with the other forest trees in the dataset. The most closely phylogenetically related species in this dataset are *F. sylvatica* and *Q. petraea*, which do have similar DFEs (shown in Figs. 2 and 3). However, we find that for these two species, DFE models that are fitted independently per species have significantly better log likelihoods than models in which either both b and S_d are shared between species ($p = 0.05$), or models in which only b is shared between species ($p = 0.03$), as might be the case if the two species shared a DFE but had different effective population sizes.

3.3 Drivers of differences in the DFE at the species level

GC-biased gene conversion

It has been demonstrated that GC-biased gene conversion can result in misinference of the DFE (Bolívar et al. 2018). CpG sites are highly mutable, and prone to polarization error. We therefore repeated our analyses restricting our dataset to GC-conservative mutations (Supplementary Table 6, Supplementary Fig. 5). We found that fitting the DFE parameters independently for GC-conservative mutations does not provide a better model fit than allowing DFE parameters to be shared between GC-conservative mutations and the full SNP dataset. Our inferred DFEs are similar across datasets (Supplementary Fig. 6 and 7), it is therefore unlikely that differences in GC-biased gene conversion, due, for example, to differences in recombination rate among species, explain differences in the DFE among species.

Life history traits and N_e

There are no significant correlations between any of the estimated parameters of the DFE and the two life history traits that we tested, maximum longevity and age at first reproduction (*i.e.*, minimum age at flowering), which were previously shown to predict genetic diversity and the efficiency of selection in plants (Chen, Glemin, and Lascoux 2017). However, the mean scaled strength of selection acting on deleterious variants, S_d, varies across species, increasing with a proxy of N_e, the level of neutral nucleotide site diversity π_4, which reflects the stronger effect of drift in smaller populations, as expected under Nearly-Neutral Theory (Spearman's $rho = -0.79$, $p = 0.048$, Pearson's $R = 0.72$, $p = 0.065$).

As expected under nearly neutral theory, the fraction of mutations that we infer to be nearly neutral from the DFE is correlated to our estimate of π_0/π_4 (Fig. 2E). However, π_0/π_4 is always greater than the nearly neutral fraction of mutations as estimated from the DFE. This is likely to be due to the contribution of segregating beneficial and slightly beneficial mutations to diversity in these species. Indeed, if we consider results from models in which we fit the deleterious DFE only (Fig. 2E), this systematic difference between π_0/π_4 and the nearly neutral fraction is reduced. *B. pendula* and *P. nigra* are particular outliers, highlighting the effect that beneficial variants have on patterns of molecular evolution in these species.

Gene content

The differences in the DFE that we observe between species are unlikely to be due to differences in gene content, or differences in genes sequenced, between species. Indeed, the parameters of the DFE were very similar when calculated across all genes in the dataset, and when calculated only for those common orthologs that were sequenced in all six species (for details of the relative proportions of all species orthologs, see Supplementary Table 7). Only in *P. pinaster* do likelihood ratio tests suggest that an independent DFE for orthologs found in all species is a better fit to the data than a shared DFE for all genes. We found that a slightly higher fraction of new mutations is inferred to be strongly deleterious in orthologs (Fig. 3), which we might expect as such genes are likely to be older, involved in many important biological functions, and under strong purifying selection. This suggests that in *P. pinaster*, genes in our dataset that are not part of the all-species ortholog set might experience differences in selective effects; they may be under less strong purifying selection. We also note a lower fraction of adaptive substitutions in all-species orthologs.

3.4 Differences among populations within species

The DFE is similar across populations of the same species, with species explaining a considerable proportion of the variation in the parameters of the deleterious and beneficial DFE as calculated across populations (Fig. 5; for deleterious-only DFE inferences see Supplementary Fig. 8). For the majority of populations, we could not reject the null model that the DFE of the population is the same as the DFE inferred for the species as a whole. This was true even under a very conservative scenario in which we fit models assuming that both b and S_d are shared across the populations and the species on average. In other words, the mean scaled strength of selection and the variance in fitness due to new mutations is consistent across populations, despite any differences in demographic history and local adaptation to environmental conditions between populations.

There are some exceptions to these general trends. We note that while results for *P. pinaster* populations indicate a considerably greater spread of inferred *b* among populations (Fig. 5A), no differences between populations are statistically supported. However, model comparison results indicate that two *B. pendula* populations might have different DFEs from the species on average (ES and IT, Supplementary Fig. 9). We also find that one *Q. petraea* population (LT), four *F. sylvatica* populations (AT, GB, NO, and SI Supplementary Fig. 9) and three *P. nigra* populations (GB, IT-S and MA, Supplementary Fig. 9), have significantly different DFEs from the DFE as calculated over all populations of each species (Fig. 4). Our results suggest both S_d and *b* differ between these populations and the dataset as a whole. We note that of these outlier populations only two, *F. sylvatica* NO and *P. nigra* GB, are significant after performing a strict Bonferroni correction for multiple testing.

In these analyses, we compare a population-specific DFE to a species-level DFE inferred over all populations, which might reduce differences between populations and the species-level 'pooled' DFE. We therefore repeated our analyses, comparing each focal population to a central reference population. We again found little difference in the DFE between populations within a species. The only populations that had significantly different DFEs to the central population were *F. sylvatica* NO and *P. nigra* MA and GB, and thus we can conclude that our results are consistent.

3.5 Drivers of differences in the DFE at the population level

At the species level, evidence of a relationship between population differentiation and variation in the effectiveness of selection, or in the shape of the DFE, is not clear. *P. pinaster*, *P. abies* and *P. nigra*, have higher mean F_{ST} values by approximately an order of magnitude, however, among-population variation in parameters of the DFE and the estimated effectiveness of selection are similar for both these and other species in the dataset with lower levels of population differentiation (see Fig. 4).

To investigate more systematically the possibility that differentiation leads to differences in the DFE, we correlated population pairwise F_{ST} with population differences in the DFE parameters S_d and *b*, and population differences in π_0/π_4. Although in some species, greater population differentiation appears to correlate with larger differences in parameters of the DFE, this relationship is not consistent (Supplementary Fig. 10).

Discussion

Here we have shown that both the efficiency of selection and the DFE differ among species, but that there is relatively little variation among populations within species. Our results suggest striking differences between different tree species, with conifers generally having a smaller fraction of highly deleterious mutations. This variation is not driven by differences in gene content between species. The nature of our exome capture dataset has resulted in a dataset that contains a high proportion of genes which are orthologs, with only 140 out of a total of 1042 genes sequenced in only a single species. Even when we restrict our analysis to orthologous genes sequenced across every species, which constituted an average of 32% of genes sequenced per species (ranging from 26% in *P. nigra* to 38% in *F. sylvatica*), differences between species in both the mean scaled strength of selection (S_d) and the coefficient of variation in the strength of selection (b) remain the same.

An important caveat of the present study is that in order to estimate the DFE, we have assumed that the DFE can be reasonably well approximated as a continuous gamma distribution. This allowed us to conduct straightforward comparative analyses across species and populations. However, it is important to acknowledge that although the DFE is commonly modelled as a gamma distribution (Bataillon and Bailey, 2014, Martin and Lenormand, 2006), other distributions can be theoretically justified (Loewe and Charlesworth, 2006), and in some studies better support has been found for alternative distributions such as the lognormal or multimodal (Kousathanas and Keightley, 2013, Loewe and Charlesworth, 2006, Sawyer et al. 2003). Alternative distributions may be better able to model high concentrations of strongly deleterious or lethal mutations more accurately, a feature that has been observed in some mutation accumulation experiments (Eyre-Walker and Keightley, 2007). Such mutations have little chance of being observed in most datasets due to their rarity, and as such the shape of the most deleterious class of mutations is based on projecting from the DFE. However, previous analyses have found that the inferred parameters of the gamma DFE are not greatly affected by including or excluding an additional parameter that takes these most deleterious mutations into account (Eyre-Walker et al. 2006). At the other end of the selective scale, alternative models may also be better able to cope with the fact that it is difficult to examine the DFE for mutations that are either neutral or have very small selection coefficients (Welch et al. 2008). Some studies have considered models that consist of a distribution of selected effects plus a point mass of neutral mutations, which have sometimes been found to fit data well (Loewe and Charlesworth, 2006, Kim et al. 2017).

Our analysis was conducted on targeted resequencing data. This approach allowed for the sampling of a high number of individuals per population and per species, increasing the amount of power we had to

infer DFE's parameters, which are notoriously difficult to estimate. While it is possible that the genomic regions used in this analysis do not reflect processes across the genome, a dataset restricted to all-species orthologs has a similar DFE to the dataset as a whole, making it likely that our DFEs are representative of the whole coding genomes of the species included in this study. Interestingly, recent work (Simons et al. 2022) has argued that in a scenario in which most traits are highly polygenic and experiencing stabilizing selection, the distribution of selection coefficients will be similar across loci that underlie all such traits. The orthologous genes which make up the majority of the coding sequences included in this analysis are perhaps likely to experience both stabilising selection, and to underlie traits that are highly polygenic, and hence be well described by the model developed by Simons et al.

Variation in the efficiency of selection at the among-population level in the species in our dataset is low. It is perhaps not surprising that many populations have highly similar DFEs to that inferred for the species overall, given the remarkably similar levels of π_0/π_4 across populations in most species (Fig. 1), and their often similar demographic histories. Although many of the species are important economically, their use by humans is unlikely to have affected the DFE, especially given that the domestication of other crop plants has had little effect on their DFEs compared to their wild relatives (Chen, Glémin, and Lascoux 2017). Comparatively, forest tree domestication and breeding is in its infancy, and the increasing effects of human activity have not yet had sufficient time to have a large impact on the tree species in this study. Recent work on the demographic histories of these species found that populations were remarkably stable in recent time, with little detectible effective population size reductions even in the face of periods of glaciation (Milesi et al. 2023). The two populations for which we have the strongest evidence for differences in the DFE, on the other hand, are somewhat unusual in terms of their demographic histories. The *P. nigra* GB population experienced a sharp population decrease in the past, and subsequently recovered. *F. sylvatica* NO also experienced a fairly extreme decrease in N_e, from which it has since recovered. Both populations differ from the species as a whole in that they have a comparatively high fraction of strongly deleterious mutations.

It has been suggested that differences in genetic load among populations might drive differences in the DFE and that populations at the edge of a species' range will have a temporarily increased mutation load relative to central populations, due to the increased importance of drift in these populations (Peischl et al. 2013; Willi et al. 2018). While for most populations we find no evidence that mutation load differs between populations, in two *P. nigra* populations, GB and MA, there is a reduction in the proportion of derived alleles relative to other *P. nigra* populations. *P. nigra* is also one of the two species that show a relationship between the efficiency of selection and latitude. For the GB population, greater purging of

deleterious derived alleles is in line with our finding that a high fraction of new mutations in this population are strongly deleterious, and that the mean strength of selection acting on new deleterious mutations is greater.

However, for the Moroccan (MA) *P. nigra* population, a comparatively low fraction of new mutations is inferred to be strongly deleterious (Supplementary Fig. 9). This population is differentiated, and in addition, there is little correlation in the frequency of alleles between this and other *P. nigra* populations. There are also a number of fixed differences between MA and other *P. nigra* populations. The DFE might differ due to these fixed differences; for example, new mutations may be less strongly deleterious when they occur on a genetic background in which many deleterious mutations are already present. However, the differences we observe are not due to inbreeding; we do not see any evidence for a shift in mating system in the MA population. There are no clonal individuals in the MA population, nor any increase in the degree of relatedness between individuals (as estimated via the KING algorithm, implemented in PLINK (Manichaikul et al. 2010; Chang et al. 2015)).

The fact that we generally do not find evidence for variation in the DFE at the population level does not mean that there is no local adaptation occurring in response to different environmental conditions across populations. Tree species generally show high levels of local adaptation, e.g., for phenological traits (Savolainen, Pyhäjärvi, and Knürr 2007), and the species in this study were generally inferred to have a high proportion of beneficial substitutions, with the exception of *P. abies* (α_{DFE}, Fig. 3C). Infrequent, strong selective sweeps are expected to leave little signature on the SFS (Booker 2020), and thus have a relatively small effect on statistics calculated from it, including the DFE. Therefore, it is possible that the tree populations do experience local adaptation through selective sweeps, the effects of which we will not detect with the summary statistics considered here. However, the DFE is informative about the strength of purifying selection and the variance of mutational effects, which do not differ among populations in the tree species in this study.

It has been hypothesised that higher population differentiation might lead to greater differences in the parameters of the DFE between populations. Our general finding is that there is some relationship between population differentiation and differences in the DFE, particularly in the strength of deleterious selection (Supplementary Fig. 10), but it is not consistent. It is interesting to consider this finding in light of the scattering and collecting phase of the coalescent (Wakeley 1999). During the collecting phase, the more ancient part of a species' history, the rate of coalescence is independent of the current geographic distribution of individuals. However, demographic history and geography will determine coalescence

during the more recent part of a species history, the scattering phase. From this study, it seems that the DFE is more strongly affected by ancient events, *i.e.*, the collecting phase of the coalescent, and the long-term N_e, leading to similar strengths of purifying selection across most populations of the same species. Whether this finding is generally true remains to be seen; the tree species in this study have moderate to high dispersal rates, however, stronger patterns of isolation by distance will lead to a stronger signal during the scattering phase (Wilkins 2004), which may result in the scattering phase having a greater impact on the DFE.

Why do differences in the DFE exist at the species level? Neither of the life history traits that we examined, maximum longevity and average age at first flowering, showed a relationship with any parameters of the DFE. We focussed on these two traits as they have been previously shown to be predictive of genetic diversity in plants (Chen, Glemin, and Lascoux 2017), although it is possible that other life history traits might affect the DFE. Previous work suggests that there might be a relationship between the DFE and large life history changes, such as transitioning from selfing to outcrossing. For example, in the herb *Arabis alpina*, selfing was associated with a reduction in the fraction of mutations inferred to experience strong negative selection, and a general reduction in the efficiency of purifying selection, while populations with mixed mating systems had very similar DFEs to outcrossing populations, with no signal of increased genetic load (Laenen et al. 2018). Relatedness also clearly plays a part- previous studies on closely related species have found that they share the same DFE (Chen, Glemin, and Lascoux 2017; Castellano et al. 2019; Liu et al. 2022). The most closely related species in our dataset, *F. sylvatica* and *Q. petraea*, also appear to have more similar DFEs (see, for example, Fig. 3), although model comparison tests indicate that fitting DFEs independently to these species provides a better fit to the data, albeit only slightly (log-likelihoods of independent and shared models: -523.6194, -526.7272, $p = 0.045$). It may be that some slow evolving aspect of genome biology, for example, gene interaction networks, methods of gene expression regulation, or genome organisation or size, eventually lead to differences in DFEs between species. The possibility that genome organisation could affect the DFE was previously investigated by Hämälä and Tiffin (2020), who showed that a number of genome features could influence selective constraint, including expression level, expression variability, and gene network connectivity, while Castellano et al. (2019), found that gene density was negatively correlated to nonsynonymous diversity, possibly due to greater constraint acting on gene dense regions. This is of particular relevance to the species included in this study, because conifer genomes are considerably larger than the genomes of other tree species (De La Torre et al. 2014).

In summary, genome and species biology are important determinants of the DFE, whose long-term effects dominate short-term processes. Our findings indicate that despite differences among populations in environmental challenges faced, the mean strength of selection experienced by new mutations and their variation in selective effects remain similar across populations. The DFEs of the tree species in this study are stable, reflecting deep processes. A large change, such as a shift in breeding system e.g., from outcrossing to inbreeding, or genome structure, may be required before the DFE differs between populations or species.

Acknowledgments: JJ was supported by a grant from the Wenner-Gren Foundation. The present project was supported by the European Union's Horizon 2020 Research and Innovation Programme grant agreement no. 676876 (Gentree). SCG-M was supported by the French government in the framework of the IdEX Bordeaux University 'Investments for the Future' program / GPR Bordeaux Plant Sciences. The computations were enabled by resources in project SNIC 2022/22-910 provided by the Swedish National Infrastructure for Computing (SNIC) at UPPMAX, partially funded by the Swedish Research Council (VR) through grant agreement no. 2018-05973. We are grateful to Sylvain Glémin for his input on an earlier draft of this manuscript, and to two anonymous reviewers for their valuable insights into this work.

Data Availability

The genetic data underlying this article are available as VCF files at: https://entrepot.recherche.data.gouv.fr/dataset.xhtml?persistentId=doi:10.57745/DV2X0M.

Full documentation of bioinformatics pipelines used to generate the VCF files, including SNP filtering steps, are available at https://github.com/GenTree-h2020-eu/GenTree.

Code for all other analysis and bioinformatic steps is available at https://github.com/j-e-james/TreeDFEScripts.

References

Agrawal, Aneil F., and Michael C. Whitlock. 2010. "Environmental Duress and Epistasis: How Does Stress Affect the Strength of Selection on New Mutations?" *Trends in Ecology and Evolution* 25 (8): 450–58. https://doi.org/10.1016/j.tree.2010.05.003.

Bolívar, Paulina, Carina F. Mugal, Matteo Rossi, Alexander Nater, Mi Wang, Ludovic Dutoit, and Hans Ellegren. 2018. "Biased Inference of Selection Due to GC-Biased Gene Conversion and the Rate of

Protein Evolution in Flycatchers When Accounting for It." *Molecular Biology and Evolution* 35 (10): 2475–86. https://doi.org/10.1093/molbev/msy149.

Booker, Tom R. 2020. "Inferring Parameters of the Distribution of Fitness Effects of New Mutations When Beneficial Mutations Are Strongly Advantageous and Rare." *G3* 10 (7): 2317–26. https://doi.org/10.1534/g3.120.401052.

Castellano, David, Adam Eyre-Walker, and Kasper Munch. 2019. "Impact of Mutation Rate and Selection at Linked Sites on DNA Variation across the Genomes of Humans and Other Homininae." *Genome Biology and Evolution* 12 (1): 3550–61. https://doi.org/10.1093/gbe/evz215.

Castellano, David, Moisès Coll MacIà, Paula Tataru, Thomas Bataillon, and Kasper Munch. 2019. "Comparison of the Full Distribution of Fitness Effects of New Amino Acid Mutations across Great Apes." *Genetics* 213 (3): 953–66. https://doi.org/10.1534/genetics.119.302494.

Chang, Christopher C., Carson C. Chow, Laurent C.A.M. Tellier, Shashaank Vattikuti, Shaun M. Purcell, and James J. Lee. 2015. "Second-Generation PLINK: Rising to the Challenge of Larger and Richer Datasets." *GigaScience* 4 (1): 1–16. https://doi.org/10.1186/s13742-015-0047-8.

Chen, Jun, Sylvain Glemin, and Martin Lascoux. 2017. "Genetic Diversity and the Efficacy of Purifying Selection across Plant and Animal Species." *Mol. Biol. Evol.* 34 (6): 1417–28. https://doi.org/10.1093/molbev/mst012.

De La Torre AR, Birol I, Bousquet J, Ingvarsson PK, Jansson S, Jones SJ, Keeling CI, MacKay J, Nilsson O, Ritland K, Street N, Yanchuk A, Zerbe P, Bohlmann J. Insights into conifer giga-genomes. Plant Physiol. 2014 Dec;166(4):1724-32. doi: 10.1104/pp.114.248708. Epub 2014 Oct 27. PMID: 25349325; PMCID: PMC4256843.

Do, Ron, Daniel Balick, Heng Li, Ivan Adzhubei, Shamil Sunyaev, and David Reich. 2015. "No Evidence That Selection Has Been Less Effective at Removing Deleterious Mutations in Europeans than in Africans." *Nature Genetics* 47 (2): 126–31. https://doi.org/10.1038/ng.3186.

Douglas, Gavin M., Gesseca Gos, Kim A. Steige, Adriana Salcedo, Karl Holm, Emily B. Josephs, Ramesh Arunkumar, et al. 2015. "Hybrid Origins and the Earliest Stages of Diploidization in the Highly Successful Recent Polyploid Capsella Bursa-Pastoris." *PNAS* 112 (9): 2806–11. https://doi.org/10.1073/pnas.1412277112.

Duret, Laurent. 2002. "Evolution of Synonymous Codon Usage in Metazoans." *Current Opinion in Genetics & Development* 12 (6): 640–49. https://doi.org/10.1016/S0959-437X(02)00353-2.

Eyre-Walker, Adam, Megan Woolfit, and Ted Phelps. 2006. "The Distribution of Fitness Effects of New Deleterious Amino Acid Mutations in Humans." *Genetics* 173 (2): 891–900. https://doi.org/10.1534/genetics.106.057570.

González-Martínez, Santiago C., Kate Ridout, and John R. Pannell. 2017. "Range Expansion

Compromises Adaptive Evolution in an Outcrossing Plant." *Current Biology* 27 (16): 2544-2551.e4. https://doi.org/10.1016/j.cub.2017.07.007.

Grivet, Delphine, Komlan Avia, Aleksia Vaattovaara, Andrew J. Eckert, David B. Neale, Outi Savolainen, and Santiago C. González-Martínez. 2017. "High Rate of Adaptive Evolution in Two Widespread European Pines." *Molecular Ecology* 26 (24): 6857–70. https://doi.org/10.1111/mec.14402.

Grossen, Christine, Frédéric Guillaume, Lukas F. Keller, and Daniel Croll. 2020. "Purging of Highly Deleterious Mutations through Severe Bottlenecks in Alpine Ibex." *Nature Communications* 11 (1001). https://doi.org/10.1038/s41467-020-14803-1.

Gutenkunst, Ryan, Ryan Hernandez, Scott Williamson, and Carlos Bustamante. 2010. "Diffusion Approximations for Demographic Inference: DaDi." *Nature Precedings*. https://doi.org/10.1038/npre.2010.4594.1.

Haller, B.C., & Messer, P.W. 2023. "SLiM 4: Multispecies eco-evolutionary modeling." *The American Naturalist 201 (5). https://doi/10.1086/723601*

Hämälä, Tuomas, and Peter Tiffin. 2020. "Biased Gene Conversion Constrains Adaptation in Arabidopsis Thaliana." *Genetics* 215 (3): 831–46. https://doi.org/10.1534/genetics.120.303335.

Huang, Xin, Alyssa Lyn Fortier, Alec J Coffman, Travis J Struck, Megan N Irby, Jennifer E James, José E León-Burguete, Aaron P Ragsdale, and Ryan N Gutenkunst. 2021. "Inferring Genome-Wide Correlations of Mutation Fitness Effects between Populations." *Molecular Biology and Evolution* 38 (10): 4588–4602. https://doi.org/10.1093/molbev/msab162.

Huber, Christian D., Bernard Y. Kim, Clare D. Marsden, and Kirk E. Lohmueller. 2017. "Determining the Factors Driving Selective Effects of New Nonsynonymous Mutations." *Proceedings of the National Academy of Sciences of the United States of America* 114 (17): 4465–70. https://doi.org/10.1073/pnas.1619508114.

Johri, Parul, Adam Eyre-walker, Jeffrey D Jensen, and Ryan N Gutenkunst. 2022. "On the Prospect of Achieving Accurate Joint Estimation of Selection with Population History." *GBE* 14 (7): evac088.

Keightley, Peter D., and Benjamin C. Jackson. 2018. "Inferring the Probability of the Derived vs. The Ancestral Allelic State at a Polymorphic Site." *Genetics* 209 (3): 897–906. https://doi.org/10.1534/genetics.118.301120.

Laenen, Benjamin, Andrew Tedder, Michael D. Nowak, Per Toräng, Jörg Wunder, Stefan Wötzel, Kim A. Steige, et al. 2018. "Demography and Mating System Shape the Genome-Wide Impact of Purifying Selection in Arabis Alpina." *Proceedings of the National Academy of Sciences of the United States of America* 115 (4): 816–21. https://doi.org/10.1073/pnas.1707492115.

Li, Heng, and Richard Durbin. 2009. "Fast and Accurate Short Read Alignment with Burrows-Wheeler

Transform." *Bioinformatics* 25 (14): 1754–60. https://doi.org/10.1093/bioinformatics/btp324.
Liu, Shuyu, Lei Zhang, Yupeng Sang, Qiang Lai, Xinxin Zhang, Changfu Jia, Zhiqin Long, et al. 2022. "Demographic History and Natural Selection Shape Patterns of Deleterious Mutation Load and Barriers to Introgression across Populus Genome." *Molecular Biology and Evolution* 39 (2): msac008. https://doi.org/10.1093/molbev/msac008.
Manichaikul, Ani, Josyf C. Mychaleckyj, Stephen S. Rich, Kathy Daly, Michèle Sale, and Wei Min Chen. 2010. "Robust Relationship Inference in Genome-Wide Association Studies." *Bioinformatics* 26 (22): 2867–73. https://doi.org/10.1093/bioinformatics/btq559.
Martin, Guillaume, and Thomas Lenormand. 2006. "The Fitness Effect of Mutations across Environments: A Survey in Light of Fitness Landscape Models." *Evolution* 60 (12): 2413–27. https://doi.org/https://doi.org/10.1111/j.0014-3820.2006.tb01878.x.
Milesi, Pascal, Chedly Kastally, Benjamin Dauphin, Sandra Cervantes, Francesca Bagnoli, Katharina B Budde, Stephen Cavers, et al. 2023. "Synchronous Effective Population Size Changes and Genetic Stability of Forest Trees through Glacial Cycles." *BioRxiv* 17: 2023.01.05.522822. https://doi.org/https://doi.org/10.1101/2023.01.05.522822.
Muyle, Aline, Hélène Martin, Niklaus Zemp, Maéva Mollion, Sophie Gallina, Raquel Tavares, Alexandre Silva, et al. 2021. "Dioecy Is Associated with High Genetic Diversity and Adaptation Rates in the Plant Genus Silene." *Molecular Biology and Evolution* 38 (3): 805–18. https://doi.org/10.1093/molbev/msaa229.
Ohta, Tomoko. 1973. "Slightly Deleterious Mutant Substitutions in Evolution." *Nature* 246: 96–98.
Peischl, S, I Dupanloup, M Kirkpatrick, and L Excoffier. 2013. "On the Accumulation of Deleterious Mutations during Range Expansions." *Molecular Ecology* 22 (24): 5972–82. https://doi.org/10.1111/mec.12524.
Rousset, François. 1997. "Genetic Differentiation and Estimation of Gene Flow from F-Statistics Under Isolation by Distance." *Genetics* 145 (4): 1219–28. https://doi.org/10.1002/ajmg.c.30221.
Savolainen, Outi, Tanja Pyhäjärvi, and Timo Knürr. 2007. "Gene Flow and Local Adaptation in Trees." *Annual Review of Ecology, Evolution, and Systematics* 38: 595–619. https://doi.org/10.1146/annurev.ecolsys.38.091206.095646.
Simons, Yuval B., and Guy Sella. 2016. "The Impact of Recent Population History on the Deleterious Mutation Load in Humans and Close Evolutionary Relatives." *Current Opinion in Genetics and Development* 41: 150–58. https://doi.org/10.1016/j.gde.2016.09.006.
Simons, Yuval B., Michael C. Turchin, Jonathan K. Pritchard, and Guy Sella. 2014. "The Deleterious Mutation Load Is Insensitive to Recent Population History." *Nature Genetics* 46 (3): 220–24. https://doi.org/10.1038/ng.2896.

Takou, Margarita, Tuomas Hämälä, Evan M Koch, Kim A Steige, Hannes Dittberner, Levi Yant, Mathieu Genete, et al. 2021. "Maintenance of Adaptive Dynamics and No Detectable Load in a Range-Edge Outcrossing Plant Population." *Molecular Biology and Evolution* 38 (5): 1820–36. https://doi.org/10.1093/molbev/msaa322.

Tataru, Paula, and Thomas Bataillon. 2019. "PolyDFEv2.0: Testing for Invariance of the Distribution of Fitness Effects within and across Species." *Bioinformatics* 35 (16): 2868–69. https://doi.org/10.1093/bioinformatics/bty1060.

Tataru, Paula, Maéva Mollion, Sylvain Glémin, and Thomas Bataillon. 2017. "Inference of Distribution of Fitness Effects and Proportion of Adaptive Substitutions from Polymorphism Data." *Genetics* 207 (3): 1103–19. https://doi.org/https://doi.org/10.1534/genetics.117.300323.

Wakeley, John. 1999. "Nonequilibrium Migration in Human History." *Genetics* 153 (4): 1863–71. https://doi.org/10.1093/genetics/153.4.1863.

Wang, Kai, Mingyao Li, and Hakon Hakonarson. 2010. "ANNOVAR: Functional Annotation of Genetic Variants from High-Throughput Sequencing Data." *Nucleic Acids Research* 38 (16): e164. https://doi.org/10.1093/nar/gkq603.

Welch, John J., Adam Eyre-Walker, and David Waxman. 2008. "Divergence and Polymorphism under the Nearly Neutral Theory of Molecular Evolution." *Journal of Molecular Evolution* 67 (4): 418–26. https://doi.org/10.1007/s00239-008-9146-9.

Weng, Mao Lun, Jon Ågren, Eric Imbert, Henning Nottebrock, Matthew T. Rutter, and Charles B. Fenster. 2021. "Fitness Effects of Mutation in Natural Populations of Arabidopsis Thaliana Reveal a Complex Influence of Local Adaptation." *Evolution* 75 (2): 330–48. https://doi.org/10.1111/evo.14152.

Wilkins, Jon F. 2004. "A Separation-of-Timescales Approach to the Coalescent in a Continuous Population." *Genetics* 168 (4): 2227–44. https://doi.org/10.1534/genetics.103.022830.

Willi, Yvonne, Marco Fracassetti, Olivier Bachmann, and Josh Van Buskirk. 2020. "Demographic Processes Linked to Genetic Diversity and Positive Selection across a Species' Range." *Plant Communications* 1 (6): 100111. https://doi.org/10.1016/j.xplc.2020.100111.

Willi, Yvonne, Marco Fracassetti, Stefan Zoller, and Josh Van Buskirk. 2018. "Accumulation of Mutational Load at the Edges of a Species Range." *Molecular Biology and Evolution* 35 (4): 781–91. https://doi.org/10.1093/molbev/msy003.

Williamson, Robert J., Emily B. Josephs, Adrian E. Platts, Khaled M. Hazzouri, Annabelle Haudry, Mathieu Blanchette, and Stephen I. Wright. 2014. "Evidence for Widespread Positive and Negative Selection in Coding and Conserved Noncoding Regions of Capsella Grandiflora." *PLoS Genetics* 10 (9). https://doi.org/10.1371/journal.pgen.1004622.

Wright, S. 1946. “Isolation by Distance under Diverse Systems of Mating.” *Genetics* 31 (1): 39–59. https://doi.org/10.1093/genetics/31.1.39.

Tables

Species	S_d	b	Fraction of mutations $-1 < N_e s < 0$	π_0/π_4	Model
Fagus sylvatica	-25,000	0.36	0.20	0.23	- ε_{anc}
Quercus petraea	-9500	0.41	0.23	0.27	+- ε_{anc}
Betula pendula	-190	1.59	0.17	0.20	+-
Populus nigra	-571	0.39	0.24	0.26	+-
Picea abies	-47,000	0.097	0.30	0.35	- ε_{anc}
Pinus pinaster	-64	0.73	0.22	0.28	+-

Table 1: Model averaged estimates of the DFE parameters, for all species.
S_d: the mean scaled strength of deleterious selection acting on new mutations rounded to two S. F., *i.e.*, the scale parameter of the gamma-shaped deleterious DFE; b: the shape parameter of the gamma-shaped deleterious DFE, which is inversely related to the coefficient of variation in the fitness effects of new deleterious mutations; the inferred fraction of mutations with fitness effects between -1 and 0, *i.e.*, the nearly neutral fraction of slightly deleterious mutations; and π_0/π_4. DFE parameters shown are model-averaged, such that estimates are weighted by model AIC. The best model, as ascertained using likelihood ratio tests, is indicated in the last column; whether fitting a deleterious only DFE (-) or a full DFE including beneficial mutations (+-), and whether including the rate of error in the inference of the ancestral state improves the model fit (ε_{anc}).

Figure Legends

Figure 1. Levels of non-neutral diversity and mutation load are similar across populations within a species, despite different levels of population differentiation.

The average π_0/π_4 per population for all species (A), and ‘focal’ population pairwise F_{ST} (Wright’s F_{ST}) for all species (B), using a ‘central’ (see methods) population per species as a reference, plotted against average sampling latitude for the focal population. Lines shown are linear regression slopes, along with their 95% confidence intervals. In C), we plot R_{XY} for 0-fold degenerate, nonsynonymous sites, calculated per population, comparing focal populations (X) to a ‘central’ (see methods) reference population (Y), while in D) we plot R'_{XY} for 0-fold degenerate, nonsynonymous sites, a measure which is normalised using putatively neutral 4-fold degenerate synonymous sites. In C) and D) black diamonds indicate the calculated values, while coloured error bars are 95% confidence intervals on the estimate, calculated through jack-knifing. X axis labels are population codes, which

begin with the two-letter country code of the sampling locations for each population, ordered by increasing latitude. The third letter provides additional location information for populations: C = Corsica, S = South, N = North, E = East, W = West (for exact sampling locations, see Supplementary Table 3). Colour codes for species are consistent across all figure panels, with species ordered such that more closely related species are closer together.

Figure 2. Species differences in the deleterious-only DFE. A) Shows the model- averaged discretised DFE, *i.e.*, the fraction of new mutations in each scaled fitness effect ($N_e s$) category. Black bars indicate 95% confidence intervals on the estimated fraction, as estimated from model-averaged bootstrap replicates. (B) Violin plots of the shape parameter, b, and (C) Violin plots of the scale parameter S_d, for the gamma distribution of deleterious fitness effects per species. Black diamonds are the inferred model-averaged parameters, while violins show the 95% confidence intervals, as estimated from model-averaged bootstrap replicates.

Figure 3. Species differences in the full DFE. A) Shows the model- averaged discretised DFE, *i.e.*, the fraction of new mutations in each scaled fitness effect ($N_e s$) category. Black bars indicate 95% confidence intervals on the estimated fraction, as estimated from model averaged bootstrap replicates. (B) Violin plots show the shape parameter, b, the scale parameter, (C) S_d, for the gamma distribution of deleterious fitness effects per species, D) α_{DFE}, the estimated fraction of substitutions inferred to be adaptive. Black diamonds are the inferred model-averaged parameters, while violins show the 95% confidence intervals, as estimated from model-averaged bootstrap replicates. In E), we show the fraction of slightly deleterious ($-1 < N_e s < 0$) mutations plotted against the ratio of 0- to 4-fold degenerate nucleotide diversity. Circles represent the fraction as inferred from the deleterious-only DFE model, diamonds represent the fraction as inferred from the full (advantageous and deleterious) DFE model. The dashed line indicates x = y. For *P. abies*, the diamond and circle overlap.

Figure 4. **Discretised DFEs for each species, showing model comparisons for different categories of genes.** Darkest bars show the independent fit for all genes, lightest bars show the independent fit for orthologs found in all species, intermediate bar shows the fit if the parameters are inferred to be shared across the all-species orthologs and the full dataset. We show model fits for the full DFE, including an estimate of the rate of ancestral allele misidentification, ε_{anc}, for all species.

Figure 5. DFE parameters are consistent across populations within a species for the full DFE. Shown are the model averaged inferred parameters. We plot the shape (A) and scale (B) parameter of the gamma deleterious distribution of fitness effects, C) α_{DFE}, the proportion of substitutions that are expected to be adaptive, (D) the proportion of mutations inferred to be effectively neutral, *i.e.*, the fraction of mutations for which $-1 < N_e s < 1$. Boxplots show the distribution of values per species, with outlier points indicated as black dots, and labelled by their population codes. Population codes always start with two letter country codes, S = South. For exact sampling locations, see Supplementary Table 3.

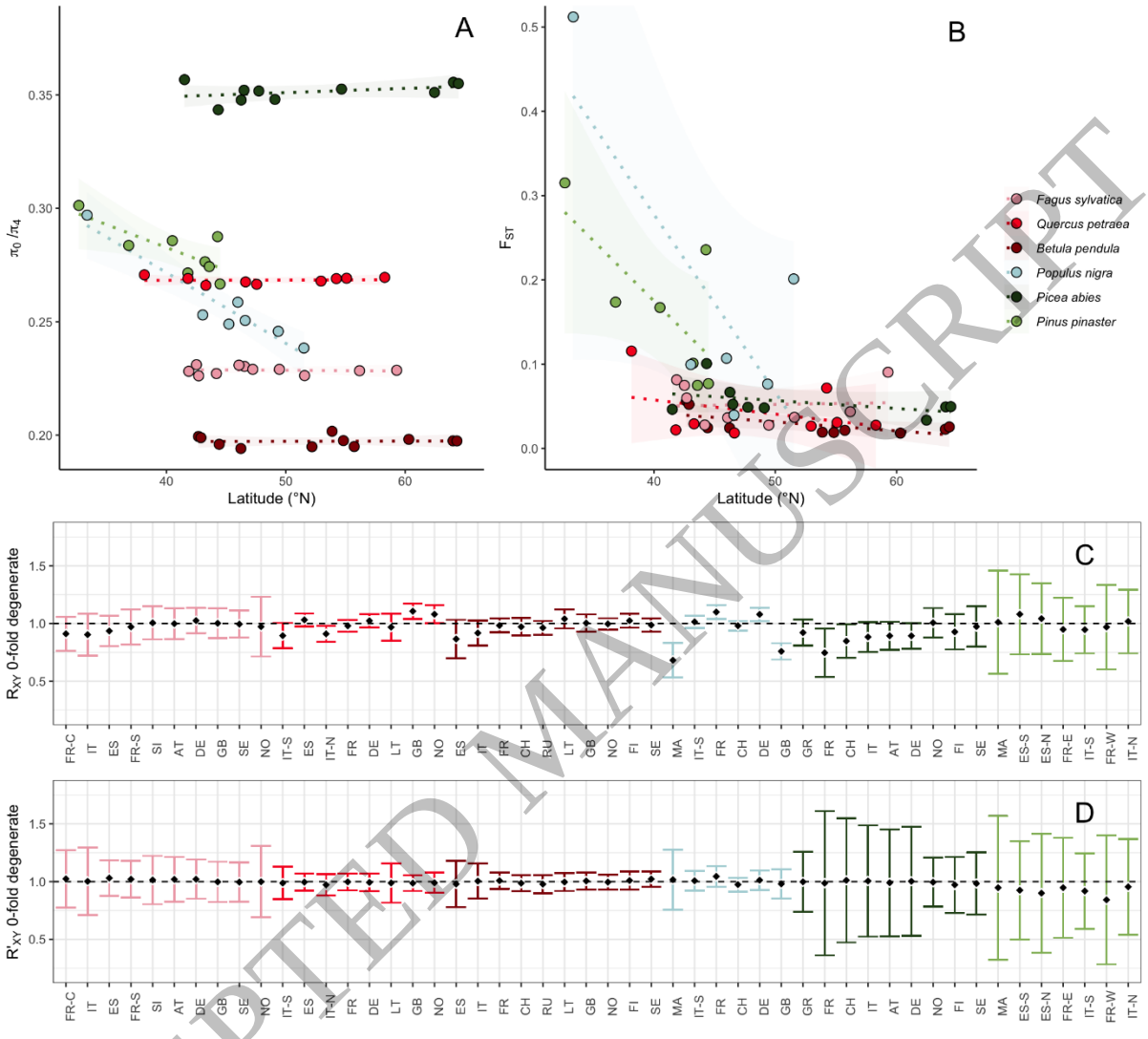

Figure 1
159x143 mm (x DPI)

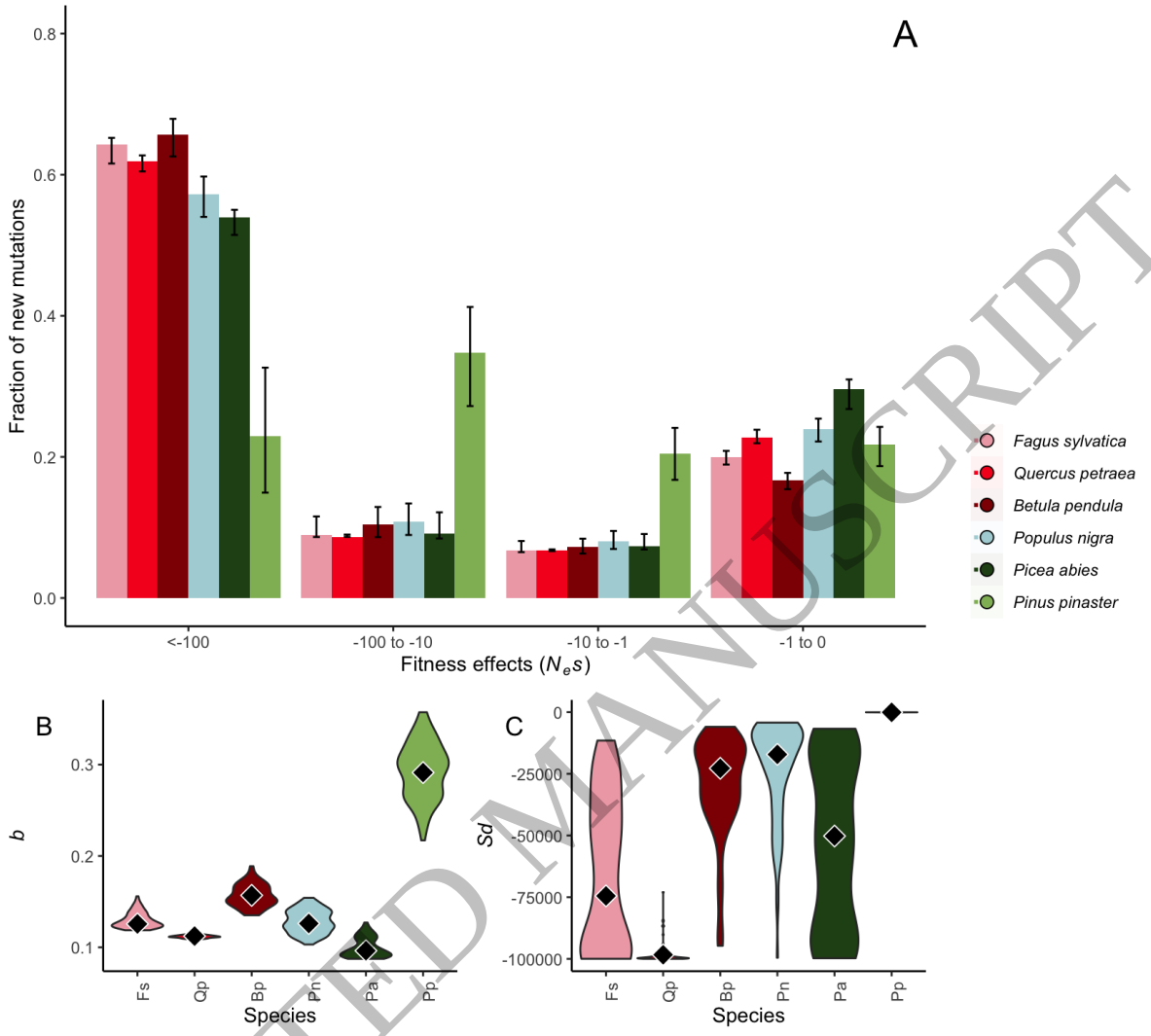

Figure 2
159x141 mm (x DPI)

1

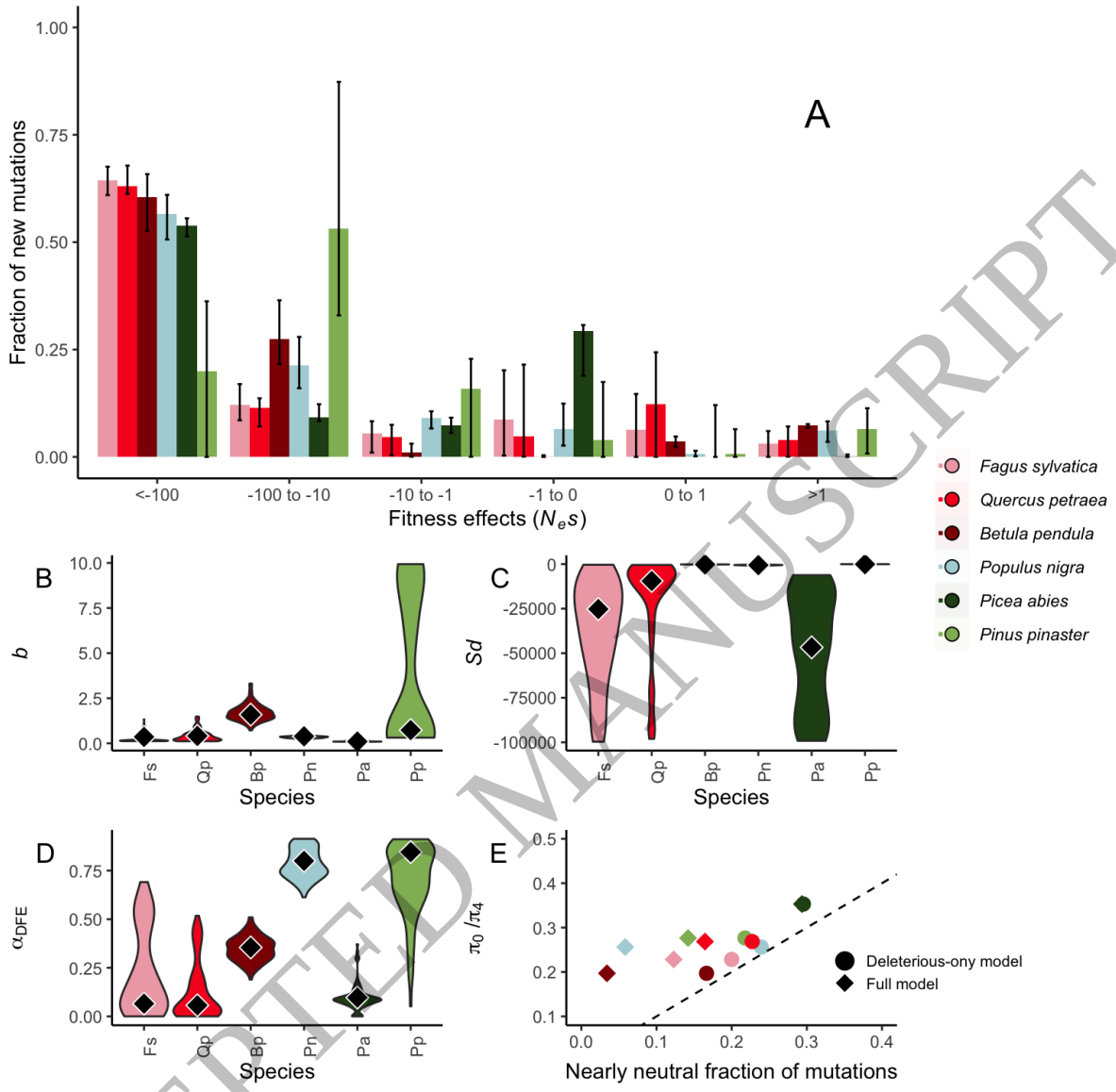

Figure 3
159x154 mm (x DPI)

1

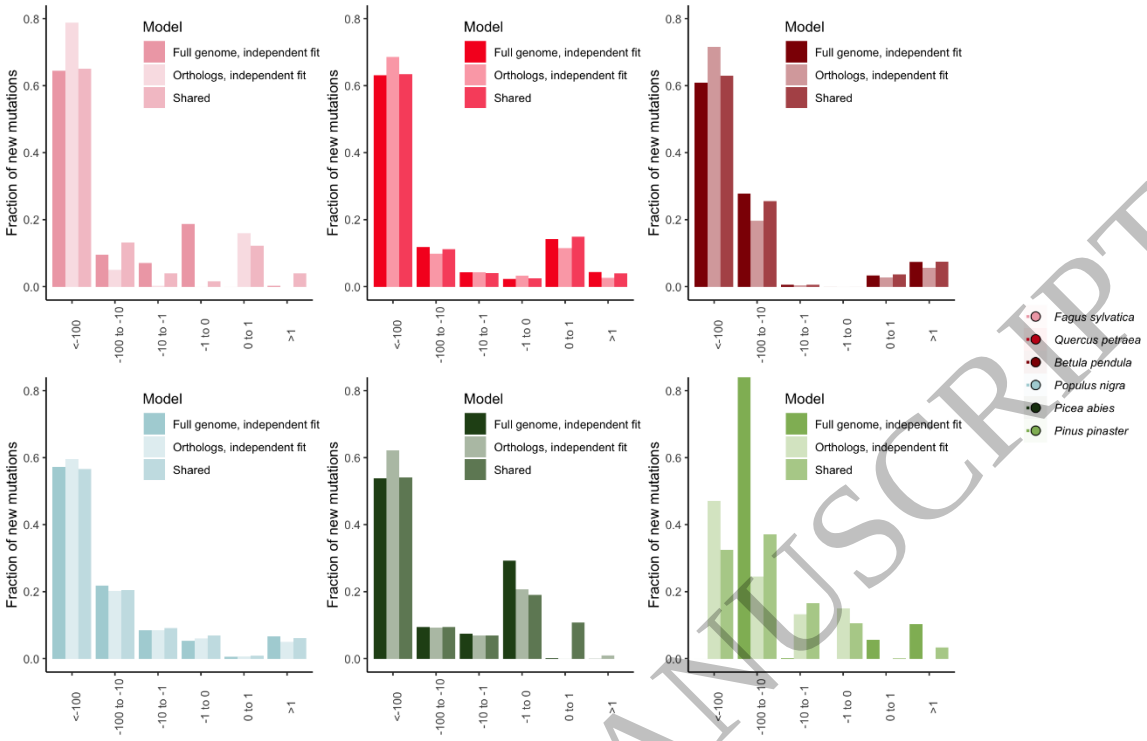

Figure 4
159x98 mm (x DPI)

1

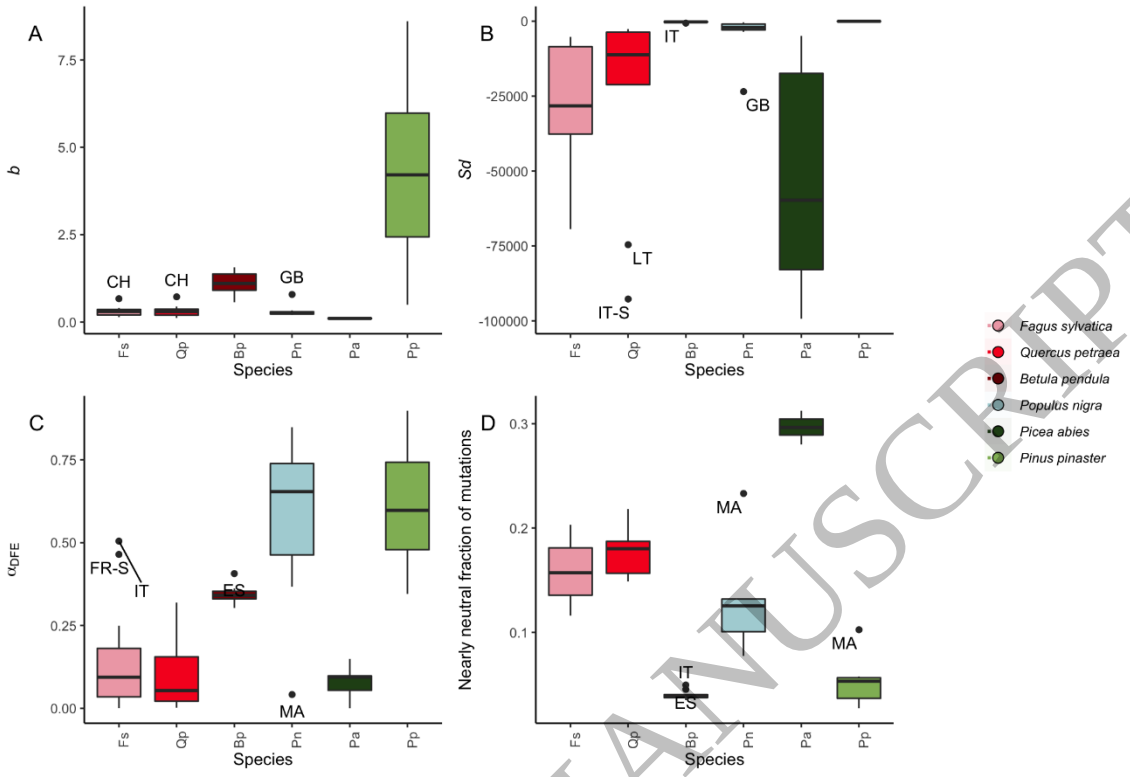

Figure 5
159x103 mm (x DPI)