Restriction-site associated DNA markers provide new insights into the evolutionary history of the bark beetle genus Dendroctonus

Martin Godefroid, Andrea S. Meseguer, Laure Saune, Guénaëlle Genson, Jean-Claude Streito, Jean-Pierre Rossi, Alejandro Zaldivar Riveron, François Mayer, Astrid Cruaud, Jean Yves Rasplus

Restriction-site associated DNA markers provide new insights into the evolutionary history of the bark beetle genus *Dendroctonus*.

Martin Godefroid[1,2,Ŧ], Andrea S. Meseguer[3], Laure Sauné[1], Guenaëlle Genson[1], Jean-Claude Streito[1], Jean-Pierre Rossi[1], Alejandro Zaldivar Riveron[4], François Mayer[5], Astrid Cruaud[1]* & Jean-Yves Rasplus[1]*

*Co-last authorship

[1]*CBGP, Univ Montpellier, CIRAD, INRA, IRD, Montpellier SupAgro, Montpellier, France*

[2]*Consejo Superior de Investigaciones Cientificas – Instituto de Ciencias Agrarias (ICA – CSIC), Madrid, Spain*

[3]*CNRS, UMR 5554 Institut des Sciences de l'Evolution de Montpellier (Université de Montpellier | CNRS | IRD | EPHE), Montpellier, France.*

[4]*Departamento de Zoología, Instituto de Biología, Universidad Nacional Autónoma de Mexico, Mexico, D.F., Mexico*

[5]*Biological Control and Spatial Ecology Laboratory* (*LUBIES*), *University of Brussels*

[Ŧ] *corresponding author: martin.godefroid@gmail.com*

Keywords Tomicini, phylogeny, biogeography, RAD sequencing, inter-continental disjunctions

ABSTRACT

The bark beetle genus *Dendroctonus* contains some of the most economically important pests of conifers worldwide. Despite many attempts, there is no agreement today on the phylogenetic relationships within the genus, which limits our understanding of its evolutionary history. Here, using restriction-site associated DNA (RAD) markers from 70 specimens representing 17 species (85% of the known diversity) we inferred the phylogeny of the genus, its time of origin and biogeographic history, as well as the evolution of key ecological traits (host plants, larval behavior and adults' attack strategies). For all combinations of tested parameters (from 6,444 to 23,570 RAD tags analysed), the same, fully resolved topology was inferred. Our analyses suggest that the most recent common ancestor (mrca) of all extant *Dendroctonus* species was widely distributed across eastern Palearctic and western Nearctic during the early Miocene, from where species dispersed to other Holarctic regions. A first main inter-continental vicariance event occurred during early Miocene isolating the ancestors of *D. armandi* in the Palearctic, which was followed by the radiation of the main *Dendroctonus* lineages in North America. During the Late Miocene, the ancestor of the *'rufipennis'* species group colonized north-east Palearctic regions from western North America, which was followed by a second main inter-continental vicariance event isolating Pleistocene populations in Asia (*D. micans*) and western North America (*D. murrayanae* and *D. punctatus*). The present study supports previous hypotheses explaining intercontinental range disjunctions across the Northern Hemisphere by the fragmentation of a continuous distribution due to climatic cooling, host range fragmentation and geological changes during the late Cenozoic. The reconstruction of ancestral ecological traits indicates that the mrca bored individual galleries and mass attacked the boles of pines. The gregarious feeding behavior of the larvae as well as the individual attack of the base of trees have apparently independently evolved twice in North America (in the '*rufipennis*' and the '*valens*' species groups), which suggests a higher adaptive potential than previously thought and may be of interest for plant protection and biodiversity conservation in a rapidly changing world.

Introduction

Understanding the environmental processes and eco-evolutionary forces that shape spatial patterns of biodiversity is a main goal in ecological and evolutionary research. In addition, phylogenetic and macro-evolutionary information has been proven crucial for plant protection, invasive species control and biodiversity management. Indeed, it is increasingly recognized that phylogenetic and macro-evolutionary information are valuable to depict species' niche dimensions and anticipate their adaptive potential to novel environmental conditions (Larson and Olden, 2012; Morales- Castilla et al., 2017; Williams et al., 2008).

This perspective is crucial for pest management in a rapidly changing world, where managers need accurate predictions of the potential impact of pests under current and future climate conditions (Bentz et al., 2010; Godefroid et al., 2016). In this study, we focus on *Dendroctonus*, one of the most economically important conifer-feeding genera of bark beetles and a long-standing major concern for forest managers (Six and Bracewell, 2015). The genus *Dendroctonus* includes 20 species feeding under the bark of conifer species belonging to the genera *Picea*, *Pinus*, *Larix* and *Pseudotsuga* (Armendáriz-Toledano et al., 2015; Wood, 1982). Most *Dendroctonus* species are native to North America, while two species naturally occur in Eurasia (*D. micans* (Kugelann) and *D. armandi* Tsai & Li). Several species have shifted their distribution during the last decades and cause unprecedented damages to forests in newly colonized areas. For example, *D. valens* was introduced in China from western America via international trade (Yan et al., 2005), while global warming allowed *D. frontalis* and *D. ponderosae* to expand their distribution northwards (Cullingham et al., 2011; Weed et al., 2013).

Many authors have attempted to resolve the phylogenetic relations within *Dendroctonus* (Bentz and Stock, 1986; Havill et al., 2019; Hopkins, 1909; Kelley and Farrell, 1998; Lanier, 1981; Reeve et al., 2012; Sequeira and Farrell, 2001; Víctor and Zúñiga, 2016; Wood, 1982, 1963; Zúñiga et al., 2002). However, topologies differed among studies and resolution and support values at the deepest nodes were generally low, which highlights the need of further phylogenetic investigations based on additional DNA markers (Víctor and Zúñiga, 2016). Notably, two phylogenetic relationships are still debated. First, the position of the Eurasian species *D. armandi* is still unresolved. *D. armandi* appears sister to all other *Dendroctonus* in the phylogenies of Kelley and Farrell (1998) and Víctor and Zúñiga (2016), while it is closely related to *D. simplex* and *D. pseudotsugae* in the topology inferred by Reeve et al. (2012), though

with low support. Second, the relationships between *D. rufipennis*, *D. murrayanae*, *D. punctatus* and *D. micans* are controversial. While Kelley and Farrell (1998) and Víctor and Zúñiga (2016) inferred *D. punctatus* as sister to *D. micans*, other studies suggested that *D. punctatus* and *D. murrayanae* were sister species, though with low support (Furniss, 1996; Reeve et al., 2012).

The *Dendroctonus* genus provides a promising opportunity to uncover the role of ecological innovation, plant-insect interactions and geographic events in shaping biodiversity of phytophagous insects. For instance, the existing phylogenies of *Dendroctonus* support hypotheses of host genus conservatism on macro-evolutionary time scale (Kelley and Farrell, 1998), which is similar to patterns observed in other phytophagous insects taxa such as aphids and seed-beetles (Kergoat et al., 2007; Meseguer et al., 2015). In addition, *Dendroctonus* species display a high diversity of life-history traits such as different degree of host specificity, a wide range of phenology-related traits (e.g., some species are strictly univoltine while others have a more flexible phenology) and larval feeding behaviors (gregarious *versus* solitary) (Kelley and Farrell, 1998; Reeve et al., 2012). Particularly, two main attack strategies characterize *Dendroctonus* species i.e., (i) one strategy consists in adults mass-attacking the bole of living trees, which may induces severe outbreaks and high tree mortality rate; (ii) the alternative strategy consists in few adults attacking the base or roots of trees, which is usually associated to low rates of tree mortality. The 'mass-attack' strategy, what is supposed to make the beetles able to overcome host defenses, was inferred to be an ancestral character in *Dendroctonus* (Reeve et al., 2012). The pattern and the timing of the evolution of the non-massive attack strategy remains, however, uncertain according to existing ancestral state reconstructions (Reeve et al., 2012).

The geographical pattern and the timing of cladogenetic events in *Dendroctonus* are still debated. Based on a mummified gallery found on a larch wood in Canadian high artic, *Dendroctonus* was speculated to have occurred in Nearctic during the Eocene (Labandeira et al., 2001). Wood (1985) suggested that *Dendroctonus* originated in Mexico from species that fed on *Araucaria*, which is supported by the high diversity of species occurring in Mexican mountain ranges (13 spp). Zúñiga et al. (2002) conversely claimed that *Dendroctonus* originated in the northern areas of North America from karyological and distribution data. A biogeographic analysis performed on Scolytinae by Pistone et al. (2017), inferred the most recent common ancestor (mrca) of *D. terebrans* and *D. ponderosae* to have lived in the Nearctic during the

Middle Miocene. More recently, Havill et al. (2019) inferred a Middle Miocene origin for the *Dendroctonus frontalis* species group using mitochondrial and nuclear DNA sequences together with microsatellite loci.

Originally designed for population genetics (Baird et al., 2008), Restriction-site Associated DNA sequencing (RAD-seq) has proven to be a powerful and efficient tool to decipher phylogenetic relationships within genera. Indeed, by providing access to thousands of markers, RAD-seq approaches successfully resolved ambiguous phylogenies in various taxa: plants (Hipp et al., 2018), insects (Cruaud et al., 2014), reptiles (Leaché et al., 2014) and fishes (Jones et al., 2013) among others. In the present study, we used RAD-seq to investigate the evolutionary history of the genus *Dendroctonus*. Based on a fully resolved tree, we reconstructed the ancestral states of key ecological traits, performed divergence time estimates and inferred the biogeographic history of the genus.

Material & Methods

Field sampling and species identification

The data set contained 70 specimens representing 17 *Dendroctonus* species (Appendix S1). All described species of *Dendroctonus* were sampled with the exception of *D. mesoamericanus*, *D. vitei* and *D. parallelocollis*. To account for intraspecific diversity, two to seven specimens per species were included. Individuals belonging to the genera *Hylastes* and *Tomicus* were also sampled for use as outgroups in further phylogenetic inferences. All species of *Dendroctonus* were identified using the taxonomic key of Wood (1982).

To confirm our identifications and to conduct divergence time estimates (see below), a *ca.* 800 bp fragment of the *COI* gene was amplified with PCR using the C1-J-2183 (Jerry) and TL2-N-3014 (Pat) primers (Simon et al., 1994) following Víctor and Zúñiga (2016). Unpurified PCR products were sent to Eurofins Genomics for Sanger sequencing. Forward and reverse strands were assembled using Geneious 10.2.3 (https://www.geneious.com). Consensuses were translated to amino acids using MEGA 7 (Kumar et al., 2016) to detect frame-shift mutations and premature stop codons, which may indicate the presence of pseudogenes. *COI* sequences were compared with reference sequences published by Kelley and Farrell (1998) and Víctor and Zúñiga (2016) using Geneious 10.2.3. As morphological identification in the *frontalis* group is challenging, a maximum likelihood tree was inferred for this group with raxmlHPC-HYBRID

(version 8.2.10) (Stamatakis, 2014) using the computational facilities provided by the CIPRES Science Gateway (Miller et al., 2010). A rapid bootstrap search (100 replicates) followed by a thorough ML search (-m GTRGAMMA) was performed (Appendix S2).

Construction of the RAD library

DNA extraction was performed with the Qiagen DNeasy 96 Blood & Tissue Kit according to manufacturer instructions with the following modifications: two successive elutions (50 µL each) with heated buffer AE (55°C) and an incubation step of 15 minutes followed by plate centrifugation (6000 rpm for 2 minutes) were performed. An incision was made between the thorax and the abdomen of each specimen, to facilitate action of the proteinase K. DNA was quantified with a Qubit® 2.0 Fluorometer (Invitrogen). Library construction followed Baird et al. (2008) and Etter et al. (2011) with modifications detailed in Cruaud et al. (2014). The *PstI* enzyme was chosen as cutter. The number of expected cut sites was estimated with an *in silico* digestion of the genome of *D. ponderosae* (assembly DendPond_female_1.0, 261Mb) using a custom script (Number of expected RAD tags = 86,256). About 250ng of input DNA was used for each sample. The quantity of P1 adapters (100nM) to be added to saturate restriction sites (result=2.5uL) as well as the optimal time for DNA sonication on a Covaris S220 ultrasonicator to obtain fragments of 300 – 600 bp (results = duty cycle 10%, intensity 5, cycles/burst 200, duration 55s) were evaluated in a preliminary experiment on a bulk DNA from all species. After tagging with barcoded P1 adapters and prior to sonication, samples were pooled in groups of 16. After size selection on gel (300-600bp), end-repair and 3' A-tailing, each pool was tagged with a different barcoded P2 adapter. Final enrichment PCR was carried out in 5 independent 25µL reactions. The number of PCR cycles was set to 13 to limit PCR duplicates. The 5 reactions were pooled and quantified by qPCR using the « Library Quantification Kit - Illumina/Universal » from KAPA (KK4824). Libraries were pooled at equimolar ratio and sent to MGX-Montpellier GenomiX for 2*125nt paired-end sequencing on an Illumina HiSeq 2500 flow cell.

Bioinformatic analyses and phylogenetic inferences

The Perl pipeline *RADIS* (Cruaud et al., 2016) that relies on *Stacks* (Catchen et al., 2013, 2011) for demultiplexing of data, removing PCR duplicates and building individual and catalog loci was used to cluster raw reads into loci.

Individual loci were built using *ustacks* [m=3; M=2; N=4]. Different combinations of two key parameters were explored: the number of mismatches allowed when merging individual loci (parameter *n* in cstacks) and the minimum number of samples that should have sequences for a locus to be included in the analysis (parameter radis_nsample_min in RADIS). Six combinations were tested with *n* = 10 or 12, and radis_nsample_min = 35, 50 or 75 % of the samples. High values of the *n* parameter were used as we focused on relatively old phylogenetic divergences. Only samples for which at least 10,000 loci were included in the catalog (radis_nloci_min = 10,000) were kept in the analysis. To avoid paralogy issues, loci for which at least one sample had three or more sequences (radis_npbloci_cutoff = 3) were removed from the data set.

Phylogenetic inferences were performed on the concatenation of all loci (without partitioning) with raxmlHPC-PTHREADS-AVX (version 8.0.20; -f a -x 12345 -p 12345 -# 100 -m GTRGAMMA) (Stamatakis, 2014). Analyses were performed on the Genotoul bioinformatics platform Toulouse Midi-Pyrenees (Bioinfo Genotoul) and the CIPRES Science Gateway (Miller et al., 2010).

Divergence time estimation

Unambiguous fossil calibration points for divergence time estimates in the genus *Dendroctonus* are lacking. A middle Eocene gallery found in *Larix altoborealis* was supposed to have been made by an early member of *Dendroctonus* by Labandeira et al. (2001) but this could serve only as a minimum age for the stem *Dendroctonus*. In addition, we deliberately did not use this fossil to calibrate the crown node of *Dendroctonus* because a generic assignation based on mummified galleries remains uncertain. As a consequence, we decided to use more informative calibration priors based on the mitochondrial DNA and the results of previous molecular dating studies.

First, we estimated minimum and maximum ages for the mrcas of i) *D. punctatus* and *D. murrayanae*; ii) *D. ponderosae* and *D. jeffreyi*; iii) *D. valens* and *D. rhizophagus*; iv) *D. brevicomis* and *D. approximatus* by calculating the minimum and maximum sequence divergences between the species pairs from the *COI* sequences generated in the present study and published by Kelley and Farrell (1998). Pairwise genetic distances were then divided by the insect *COI* clock rate of 3.54 ± 0.38 % divergence My^{-1} inferred by Papadopoulou et al. (2010). The minimum and maximum ages obtained from the mtDNA ranged from 1.7 to 3.6 Ma (Appendix S3). To avoid under- or overestimation of those ages, we took a conservative

approach and assigned a minimum of 1 Ma and a maximum of 5 Ma for those four nodes in subsequent dating analyses. Second, the node (*Hylastes versus Tomicus* + *Dendroctonus*) was assigned a minimum of 37 Ma and a maximum of 54 Ma derived from the 95% Highest Posterior Density (HPD) of the *BEAST* analyses of Jordal et al. (2011) and Pistone et al. (2017).

Divergence time estimates were performed with *BEAST* version 1.8.4 (Drummond et al., 2012). As sequence-based relaxed-clock dating methods are computationally unfeasible when many missing data occur in DNA alignments, we ran the *BEAST* analysis on a subset of RAD loci, which was obtained after removal of large percentage of missing data. This data subset was obtained by constructing a consensus sequence for each species from the original RAD alignment and then by discarding all loci positions with missing data for the outgroup *Hylastes*, which was the species with the largest amount of missing data in the original RAD dataset. This data subset was composed of 137, 802 bases, which corresponds to 1,351 RAD loci.

BEAST was run with a starting tree for which we forced the topology to remain unchanged. *BEAST* starting tree was generated with the penalized maximum likelihood approach (Sanderson, 2002, 1997) implemented in the function *chronos* of the R package ape (Paradis et al., 2004). Before running the *chronos* function, the best resolved RAxML tree with *Hylastes* specified as the outgroup was pruned in ape to include only one specimen per *Dendroctonus* species. This pruned tree was then used as input in *chronos* function. To account for potential stochasticity of the results, the analysis was repeated 100 times and the mean and standard errors of nodes ages were calculated. The analyses were run under a relaxed DNA substitution model and using 10 different values of the rate-smoothing parameters λ, ranging from 0 (i.e. a high variation in substitution rates) to 1 (i.e. complete smoothing, which is equivalent to a strict molecular clock) with incremental steps of 0.1. The model yielding the higher penalized likelihood value was selected.

The best substitution model was selected with JModelTest version 2 (Darriba et al., 2012). A Yule speciation model with an uncorrelated lognormal molecular clock, a GTR+Gamma substitution model (i.e. the best substitution model identified by JModelTest) and uniform distributions as priors for the five calibration points were implemented. We ran 200 million generations of Markov chains Monte Carlo (MCMC) and checked for convergence of all parameters in Tracer v.1.6.0 (Rambaut and Drummond, 2009). 25% of the samples were removed as a burn-in to generate a maximum clade credibility tree using the software

TreeAnnotator. The *BEAST* analysis was repeated twice to ensure repeatability and convergence of all parameters.

Ancestral trait reconstruction

Ancestral states of four relevant ecological traits to explain invasion biology in *Dendroctonus* were inferred using the feature matrix constructed by Reeve et al. (2012), which relies on ecological information available in the scientific literature (Baker, 1972; Cibrián Tovar et al., 1995; Furniss and Johnson, 2002; Henigman et al., 1999; Six and Klepzig, 2004; Wood, 1982). The studied traits were: larval behavior (gregarious or solitary), attack behavior (solitary or massive), parts of the tree preferentially attacked (bole, roots or base), host plant genus (*Pinus*, *Picea*, *Larix* or *Pseudotsuga*) (Appendix S4). We estimated ancestral states for these traits over the *BEAST* chronogram using the function 'ace' in the R package ape (Paradis et al., 2004).

Biogeographic analyses

The biogeographic history of *Dendroctonus* was reconstructed using the model 'dispersal–extinction–cladogenesis' (DEC) implemented in Lagrange (Ree et al., 2005; Ree and Smith, 2008) and using the fast C++ version (Smith, 2009). Analyses were run over the *BEAST* chronogram. We divided the Holarctic into four areas by considering the paleogeographic history of the continents: western Nearctic "WN", eastern Nearctic "EN", Western Palearctic 'WP", and eastern Palearctic "EP". We collected information on the native distribution of the species from the scientific literature (Wood, 1982) (Appendix S5) and considered that all ranges comprised of four areas could be an ancestral state (maxareas = 4).

DEC allows time-dependent constraints on dispersal and area connectivity (Buerki et al., 2011). We set up three distinct DEC models of increasing complexity: M0 used only distributions of extant species with no additional constraint (equal area connectivity and dispersal probabilities through time); M1 incorporated an area connectivity (adjacency) matrix where dispersal was only allowed (with equal probability) between adjacent areas, preventing unrealistic disjoint distributions at ancestral nodes in eastern Palearctic and eastern Nearctic and in western Palearctic and western Nearctic regions; M2 incorporated, in addition to an adjacency matrix, a transition matrix where dispersal rates were scaled to reflect changing dispersal opportunities between areas and through time, based on the major tectonic events hypothesized

to have affected rates of migration (Sanmartin et al., 2001). Dispersal rates in the biogeographical model (Q transition matrix) were scaled to 1 when the areas were in direct contact (adjacent regions or regions connected by a land bridge, e.g., Beringia, connecting eastern Palearctic and western Nearctic in the Eocene). Dispersal rates were reduced to 0.01 for areas separated by large barriers, such as wide oceans (e.g., eastern Palearctic and eastern Nearctic in the present). Finally, a scalar value of 0.5 reflected an intermediate probability of movement (e.g., for dispersal between neighboring areas not connected by land bridge, such as western Palearctic and eastern Nearctic in the Eocene). Based on these scalars, we allowed the transition matrix to vary across 6 time slices "TS" (3.5, 10, 20, 30, 40 and 50 Ma) according to the changing palaeogeographic configuration of the continents through time (Appendix S5).

Results

RAD library and phylogenetic inferences

Information on RAD library and alignments are available in Table 1 and in Appendix S6. Depending on the value of the parameters n (number of mismatches allowed when merging individual loci) and radis_nsample_min (minimum number of samples that should have sequences for a locus to be included in the analysis) matrices were composed of 6,444 to 23,570 RAD tags and contained only 15.97% to 39.66% missing data.

All RAxML analyses yielded the same fully resolved tree topology (Fig. 1; Appendix S7). Particularly, when *n*=12 and radis_nsample_min =35% (23,570 RAD tags, 39.08% missing data, 18.3 % parsimony informative sites), bootstrap values of every node connecting two different species was 100 (Fig. 1; Appendix S7).

The genus *Dendroctonus* was recovered monophyletic and the Chinese *D. armandi* was sister to all other Nearctic *Dendroctonus* (Fig. 1; Appendix S7). American species of *Dendroctonus* were subdivided into four clades:

(i) the *pseudotsugae* species group including *D. simplex* LeConte, 1868, and *D. pseudotsugae* Hopkins, 1905, which feed on species of *Larix* Mill and *Pseudotsuga* Carrière respectively (Fig. 1; Appendix S7);

(ii) the *rufipennis* species group (Wood, 1963, 1982) encompassing the three species of *Dendroctonus* developing under the bark of *Picea* species (*D. rufipennis*, Kirby 1837,

D. micans Kug. and *D. punctatus*) and *D. murrayanae* that develops on *Pinus banksiana* Lamb and *Pinus contorta* Douglas ex Loudon (Fig. 1; Appendix S7);

(iii) the *frontalis* species group encompassing the pine-feeding species *D. ponderosae*, *D. jeffreyi, D. frontalis, D. mexicanus, D. approximatus, D. adjunctus* Blandford, 1897, and *D. brevicomis*. In this clade *D. jeffreyi* and *D. ponderosae* cluster in a well-supported subgroup.

(iv) the *valens* species group comprising the pine-feeding *D. rhizophagus, D. valens* and *D. terebrans* (Fig. 1; Appendix S7).

In all our analyses the *pseudotsugae* species group was recovered sister to all other species groups and the *rufipennis* species group was sister to the *frontalis* + *valens* species groups.

Dating analyses

Maximal values of penalized-likelihood were reached when the smoothing parameter was turned to zero; i.e. when high variation in substitution rate was allowed. *BEAST* dating analyses using the *chronos* chronogram as input inferred that *Hylastes* diverged from the *Dendroctonus* + *Tomicus* stem lineage during the Eocene *ca.* 43 Ma (95% HPD 37-52 Ma; Fig. 2) and that the stem lineage of *Dendroctonus* diverged from *Tomicus* during the Eocene-Oligocene transition *ca.* 36 Ma (95% HPD 25-48 Ma; Fig. 2). The split between *D. armandi* and all other *Dendroctonus* lineages is inferred to have occurred *ca.* 20 Ma during the early Miocene (95% HPD 13-27 Ma; Fig. 2). The diversification of main lineages of *Dendroctonus* in North America started during the middle Miocene (around 16 Ma; Fig. 2).

Biogeographic analyses

Among the biogeographic models investigated here (M0-M2, see methods), the model including an area connectivity (adjacency) and a dispersal matrix (M2) received the best likelihood scores (*-lnlik* = 32.12; Appendix S8a-c). Biogeographic reconstructions were however similar between models. The most likely reconstruction (Fig. 3) suggested the mrca of current *Dendroctonus* species was widespread in eastern Palearctic and western Nearctic regions (p= 0.71) during the early Miocene, from where an early vicariant event isolated Asian (*D. armandi*) and Nearctic (i.e. the other species of the genus) populations (Fig. 3). North American species diversified in the western parts of the region and from this area two independent colonizations occurred to the

eastern Nearctic before the Pliocene, in the ancestors of *D. rhizophagus* - *D. terebrans* and the ancestors of *D. rufipennis* - *D. punctatus* (Fig. 3). The ancestor of *D. rufipennis* and *D. punctatus* also returned to the eastern Palearctic region during this period, but a second vicariance event divided the Holarctic populations during the Pliocene, *ca*. 4 Ma (95% HPD 2-5 Ma; Fig. 3), splitting *D. micans* in the Palearctic from *D. murrayanae* and *D. punctatus* in the Nearctic. The uncertainty associated to this reconstruction is presented on Appendix S8c.

Reconstruction of ancestral ecological traits

Maximum likelihood inferences suggested that the mrca of all *Dendroctonus* species displayed a solitary larval feeding behavior and a mass attack strategy, fed on *Pinus* and preferentially attacked the bole of trees. According to ML inferences, adaptation to other hosts (*Picea*, *Larix* and *Pseudotsuga*) occurred later in the evolutionary history of *Dendroctonus* (Fig. 4), after the colonization of the Nearctic. Interestingly, a reverse host shift from *Picea* to *Pinus* occurred recently in the '*rufipennis*' clade (Fig. 4). Similarly, there were two independent shifts from ancestral ecological traits to gregarious larval feeding behavior, solitary attack strategy of the roots and/or the base of trees during the diversification of both '*rufipennis*' *and 'valens'* species groups (Fig. 4).

Discussion

A fully resolved Dendroctonus *phylogeny*

All analyses consistently converged on a single and fully resolved topology (Fig. 1). The Chinese species *D. armandi* was recovered sister to all other species of *Dendroctonus.* This hypothesis is supported by previous mitochondrial analyses (Kelley and Farrell, 1998; Víctor and Zúñiga, 2016) and by morphology (Wood, 1982). Indeed, Wood (1982) highlighted that *D. armandi* was morphologically different from all other *Dendroctonus* species, suggesting it differentiated from other lineages in the early stages of the genus radiation.

Our analysis unambiguously supports the five species groups defined on morphological characters (Víctor and Zúñiga, 2016) and proposes a fully resolved evolutionary hypothesis between species groups of *Dendroctonus,* that differs from all previous studies. The close relationship between *D. simplex* and *D. pseudotsugae* as well as relationships within the *frontalis* species group are well-supported and corroborate previous phylogenetic hypotheses based on

morphology and/or molecules (Havill et al., 2019; Kelley and Farrell, 1998; Víctor and Zúñiga, 2016). The relationships within the *rufipennis* species group, that were still debated (Víctor and Zúñiga, 2016) are fully resolved in our tree. Our RAD-based phylogeny suggests that *D. punctatus* and *D. murrayanae* are sister species, a relationship first hypothesized by Reeve et al. (2012). *Dendroctonus micans* is inferred sister to the clade encompassing *D. punctatus* and *D. murrayanae*. These relationships are corroborated by the analysis of ecological characteristics since *D. micans*, *D. punctatus* and *D. murrayanae* share common life-history traits (i.e., gregarious larval behavior and non-massive attack strategy) presumably inherited from a common ancestor (Fig. 4).

Noticeably, our topology differs from the phylogeny obtained by Víctor and Zúñiga (2016) in the position of the *rufipennis* species group. In their study, Víctor and Zúñiga (2016) inferred a sister-group relationships between the *valens* and *rufipennis* species groups but this poorly-supported node was mostly based on three larval characters: *(i)* spiracle openings surrounded by conspicuous sclerotized tubercles; *(ii)* dorsopleural lobes beneath spiracles, heavily sclerotized and pigmented, bearing a pair of evident setae (with the noticeable exception of *D. rufipennis*); *(iii)* presence of sclerotized dorsal plates on abdominal segments VIII and IX. In our topology, the *rufipennis* species group is sister to the *valens* + *frontalis* groups, which suggests that the above-mentioned morphological character states evolved two times independently and may represent convergent adaptation to similar conditions rather than represent an ancestral state lost by the *frontalis* group. To the exception of *D. rufipennis*, species belonging to the *rufipennis* and *valens* groups show similar larval gregarious behavior and these characters may therefore be correlated to this behavior.

Our results also reveal that all specimens of *D. valens* included in this study do not form a monophyletic group (Fig. 1, Appendix S7). Indeed, the Mexican populations of *D. valens* appear more closely related to *D. rhizophagus* than to other North American populations of *D. valens* (Appendix S4). This result strongly suggests that specimens of *D. valens* collected in Mexico may belong to another species. This hypothesis was already considered by several authors, who claimed that specimens of *D. valens* occurring in Mexico might in fact belong to *D. beckeri* (Thatcher, 1954), a species occurring in Central America and considered to be a synonym of *D. valens* (Wood, 1963). The present study provides additional evidence for the validity of *D.*

beckeri, warranting the need of an in-depth analysis of the *D. valens* species complex (Cai et al., 2008).

Host conservatism and convergent evolution of gregarious larval behavior and non-massive attack strategy

Species of *Dendroctonus* exhibit a high diversity of ecological traits (Reeve et al., 2012) including larval behavior, feeding preferences and the way they develop in conifer trees. Following the colonization of North America, *Dendroctonus* diversified and shifted to several new host-plants during the Miocene (*Pseudotsugae*/*Larix* lineage, *Picea*). High levels of host-plant conservatism followed this period, with only *D. murrayanae* shifting to a new genus of plants during the last 15 Myr suggesting that insect-plant associations tend to be stable on macro-evolutionary time (Chen et al., 2016; Jousselin et al., 2013; Kergoat et al., 2007; Meseguer et al., 2015). Our reconstruction of ancestral ecological traits confirms the statement of Reeve et al. (2012), who suggested that the mrca of all extant *Dendroctonus* species was probably a species that dug individual galleries and mass attacked the boles of pine trees. This set of ecological traits is indeed shared by numerous *Dendroctonus* species except a few species from the *rufipennis* and the *valens* species groups. We inferred that the gregarious feeding behavior of the *Dendroctonus* larvae as well as the non-massive attack strategy have independently evolved in the *rufipennis* and the *valens* species group (Fig. 4).

Early and Middle Miocene evolution of the genus Dendroctonus*: inter-continental vicariance and host-driven radiation in the Nearctic*

Our analyses suggest that the mrca of all extant *Dendroctonus* species originated during the early Miocene and was widely distributed across eastern Palearctic and western Nearctic regions (Fig. 3). The inferred origin for *Dendroctonus* postdates the speculated age of the only fossil assigned to the genus, a mummified gallery found on a larch wood in Canadian high artic from the middle Eocene (Labandeira et al., 2001). Thus, the discrepancy found between the age estimates of these mummified galleries and the crown node of *Dendroctonus* suggests this fossil could belong to the stem lineage of *Dendroctonus* or represent an extinct genus of Tomicini.

During the early Miocene and until the Pliocene, western North America and eastern Asia were connected through the Beringia Land Bridge (BLB), which was covered by temperate

woodlands with coniferous forests (Wolfe, 1994). Several insect groups associated with conifers, such as *Dendroctonus*, may have been able to dispersed or establish widely distributed populations across these continents at that time (Sanmartin et al., 2001). For instance, multiple dispersals across the BLB corridor are described in several insect taxa during Miocene (23.3-16.3 Ma); e.g., Apidae: *Bombus* (Hines, 2008) & *Eucera* (Dorchin et al., 2018); Aphididae: *Cinara* (Meseguer et al., 2015) ; Formicidae: *Myrmica* (Jansen et al., 2010) & *Premnothorax* (Prebus, 2017); Nymphalidae: *Polyommatus* (Vila et al., 2011); Psocidae: *Trichadenotecnum* (Yoshizawa et al., 2017).

A first inter-continental vicariance event presumably occurred during early Miocene (Fig. 3) isolating the ancestors of *D. armandi* in the Palearctic. This is a solitary pine-feeder species presently restricted to the Qinling Mountain Range in China and considered a relic of the Cenozoic fauna of the region associated with conifers. Subsequently, ancestral lineages of *Dendroctonus* diversified in North America, which was partly induced by specialization on different coniferous hosts (Figs. 2–3). Several *Dendroctonus* species are inferred to have specialized on *Larix* and *Pseudotsuga* during the Miocene (Fig. 3). Another *Dendroctonus* lineage (i.e. the '*rufipennis*' group) also diversified in North America by specializing on *Picea* during the same period (Fig. 3). Noticeably, the radiation of the *frontalis* species group is inferred to have started during Middle Miocene (*ca.* 11.2 Ma- 95% HPD 7.1-15.8*)*, which is highly consistent with the recent study of Havill et al. (2019) (*ca.* 12 Ma – 95% HPD 8.2-15.7).

The spatial pattern and the timing of this first vicariance event and subsequent specialization events in the Nearctic show notable resemblances with biogeographic patterns described in another conifer-feeding groups e.g., the aphid genus *Cinara* (Hemiptera: Aphididae: Lachninae) (Meseguer et al., 2015) or the leafroller moths genus *Choristoneura* (Lepidoptera: Tortricidae) (Fagua et al., 2019); indeed, the crown diversification of the main clades of *Dendroctonus* was somewhat contemporaneous with the diversification of *Choristoneura* [21 Ma, 95 HPD 17-26] (Fagua et al., 2019) and several main lineages of *Cinara* (Meseguer et al., 2015). Fagua et al. (2019) also inferred a main inter-continental vicariance event isolating Nearctic from Palearctic populations followed by a boost of speciation during the early Miocene in *Choristoneura* (*ca.* 16 Ma – 95 HPD 13-21), which mimics the early biogeographic history of *Dendroctonus* (Fig. 3). In *Cinara*, several vicariance events (at least four) separating western Nearctic and eastern Palearctic lineages were also inferred to occur during early Miocene

(Meseguer et al., 2015). Beside this, the host switch of both *Dendroctonus* and *Cinara* to *Larix and Pseudotsuga* from *Pinus*-associated ancestors are inferred to have occurred simultaneously in western Nearctic regions after the late Oligocene/early Miocene (Fig. 3) (Meseguer et al., 2015). Similarly, specialization to *Picea* were inferred to occur during the early Miocene for both insect taxa (Meseguer et al., 2015). Although further research is needed to refine the age estimates of these events, independent studies conducted on different insect groups associated with conifers, converged towards relatively similar biogeographic and host-evolution scenarios, which reinforces their credibility and suggests that the results obtained for *Dendroctonus* extend beyond the particular history of this group and might reflect a common macroevolutionary pattern for conifer-feeding insects in the Holarctic.

Late Miocene Dendroctonus *evolution: climate cooling and range expansion across the boreal forest belt*

From the Late Miocene and until the Quaternary, three dispersal events from western to eastern Nearctic regions occurred in the ancestors of the '*valens*', '*rufipennis*' and '*pseudotsugae*' species groups (Fig. 3). Several cold-adapted species belong to these lineages and currently occur in northern parts of North America (e.g., *D. simplex, D. valens, D. rufipennis, D. punctatus, D. murrayanae*), suggesting that northern corridors to dispersal might have connected Nearctic regions during this period. In addition, the species of the '*valens*' and '*rufipennis*' groups exhibit large distribution ranges within North America and Eurasia (e.g., *D. micans*) and are able to colonize large areas of coniferous forests. They have relatively broad ecological amplitudes (Six and Bracewell, 2015) and are relatively polyphagous (Kelley and Farrell, 1998), which probably facilitated dispersal of those lineages across the Holarctic.

The last inter-continental dispersal event detected in the genus occurred when the ancestor of the *'rufipennis'* group colonized north-east Palearctic regions from western North America during the Late Miocene. At this time, direct dispersal between Eurasia and western North America was still possible through Beringia, although a global cooling trend promoted the expansion of a boreal forest belt over this corridor replacing more mesic vegetation (Sanmartin et al., 2001; Schneck et al., 2012; Tiffney, 1985). It is thought that the cooling of the climate during Miocene may have acted as a barrier to dispersal across Beringia for several insect taxa with limited cold tolerances such as *Cinara* (Meseguer et al., 2015) or *Polyommatus,* (Vila et al.,

2011), however, the boreal forest belt presumably acted as a suitable dispersal corridor for the '*rufipennis*' lineage, which currently occurs in boreal zones and tolerates extremely cold climate conditions. Thus, late Miocene temperatures on the BLB might have not exceeded the thermal tolerances of the ancestor of the '*rufipennis'* species group. Direct land connection between North America and Asia through BLB might have existed until 4.8-5.5 Ma, when a marine transgression opened the Bering Strait (Gladenkov et al., 2002). This is congruent in our study, with a vicariant event dividing Pliocene populations in Asia (*D. micans*) and western North America (*D. murrayanae* and *D. punctatus*).

Our results support previous hypotheses explaining Holarctic intercontinental range disjunctions in multiple plant and insect taxa (Meseguer et al., 2015; Ian Milne, 2006; Sanmartin et al., 2001; Wen et al., 2016) – i.e. the existence of closely related lineages in Eastern Asia and Western America – by the fragmentation of a continuous distribution across the Northern Hemisphere due to climatic cooling, plant-host range fragmentation and/or geological changes during the late Cenozoic. Further studies would certainly help to better understand the role of Holarctic boreal forests, host specialization, and past climate change on the dispersal and diversification of insect groups associated with conifers. This may help to better understand mechanisms of adaptation to novel environmental conditions and then assess invasion risk and forecast responses to global change of major forest pests.

Figure and table captions

Table 1. Characteristics of the RAD data sets obtained with different parameters combinations. n = number of mismatches allowed when merging individual loci in cstacks, radis_nsample_min = minimum number of samples that should have sequences for a RAD locus to be included in the analysis. In all data sets, loci for which at least one sample has three or more sequences (radis_npbloci_cutoff = 3) were removed.

Figure 1. Phylogenetic reconstruction of the genus *Dendroctonus*. The tree was inferred with RAxML from 23,570 RAD tags. The analyzed data set was obtained with the number of mismatches allowed when merging individual loci in cstacks set to 12 and the minimum number of samples that should have sequences for a RAD locus to be included in the analysis set to 35% of the samples. Loci for which at least one sample has three or more sequences were removed.

Bootstrap values (100 replicates) are indicated at nodes. Missing data for each sample is indicated between brackets. Species names are colored according to host genus tree (*Larix*, *Picea*, *Pinus*, *Pseudotsuga*). Habitus are printed to scale (Photos G. Fleck, INRA ©).

Figure 2 Maximum clade credibility dated tree obtained with BEAST (see text for details). Median node ages and 95% highest posterior density intervals are represented at nodes (Purple rectangle). The time scale unit is in millions of years.

Figure 3. Reconstruction of the historical biogeography for the genus *Dendroctonus* inferred from a Lagrange (see text for details). Colored pies besides species names show current distributions. Colored pies at nodes represent the inferred ancestral ranges reconstructed with the highest relative probability. Squares on branches reflect how the ancestral range was divided at the speciation event, while arrows represent estimated dispersal events. Colors represent the four biogeographic areas under study; eastern Nearctic (NE), western Nearctic (WN), western Palearctic (WP) and eastern Palearctic (EP). The unit of the time scale is in millions of years.

Figure 4 Reconstruction of relevant ancestral ecological traits including larval behavior (solitary or gregarious), host plant (*Picea*, *Pinus*, *Pseudotsuga*, *Larix*), attack strategy of adults (mass- or solitary attacks) and part of the tree preferentially attacked (bole, base, roots). Pies at nodes represent probabilities of character states based on maximum-likelihood phylogenetic comparative methods.

Acknowledgement

This work was supported by the INRA (SMaCH metaprogram and the recurring funding to AC & JYR). We thank the Genotoul bioinformatics platform Toulouse Midi-Pyrenees (Bioinfo Genotoul) for providing computing and storage resources and Montpellier GenomiX (Montpellier, France) for sequencing of the library. We are grateful to Loic Dohet, Richard Hofstetter, Ronald F. Billings, Marie-Anne Auger-Rosenberg, Matthew P. Ayres and Thomas Atkinson for providing specimens. We thank Flavia Termignoni and Miguel Vera, who provided support to sample insects in Mexico. We are grateful to Bjarte Jordal, Anthony Cognato, Jiri Hulcr, Dario Pistone, Duane McKenna, James Basinger, Thomas Marcussen and Conrad

Labandeira for discussion on dating methods and node calibration. We thank Eric Pierre for help with species identifications and Günter Fleck for imaging of the specimens. We thank the two anonymous reviewers who provided constructive comments on a previous version of the manuscript.

Author contributions

M.G., A.C. & J.Y.R designed the study; M.G., A.S.M., J.P.R., A.C & J.Y.R. performed research; M.G., J.C.S., A.Z.R., F.M., A.S.M, A.C & J.Y.R sampled specimens; M.G., L.S., G.G. performed molecular experiments, under the supervision of A.C.; M.G., A.S.M. & A.C. analyzed the data; M.G., A.S.M., A.C. & J.Y.R wrote the paper; all authors discussed the results and commented on the manuscript.

Data accessibility

Sanger sequences were deposited on Genbank (accessions MK734343: 734375). Cleaned RAD reads are available as a NCBI Sequence Read Archive (accession PRJNA530572).

The authors declare no conflict of interest.

Supplementary Information

Appendix S1 Sampling information on specimens included in the RAD library.

Appendix S2 RAxML tree inferred from *COI* sequences for *D. mexicanus, D. vitei, D. frontalis, D. mesoamericanus* and *D. brevicomis*. The *COI* fragments sequenced in the herein study were aligned with the reference sequences published by Kelley and Farrell (1998) and Victor & Zuniga (2015). Genbank accessions numbers for references sequences samples are included in the tips labels. Nodes labels represent bootstrap support.

Appendix S3 Estimates of the four node ages used as calibration points in *BEAST* and penalized-likelihood dating analyses.

Appendix S4 Ecological characteristics and geographic distributions of extant species of *Dendroctonus*.

Appendix S5 Transition matrix in Lagrange reconstructions.

Appendix S6 Sequencing data obtained for each sample. Number of raw reads, cleaned reads, ustacks loci and percentage of missing data for each analysed data set.

Appendix S7 RAxML trees obtained from all data sets. Each data set was obtained with a different combination of parameters. n = number of mismatches allowed when merging individual loci in cstacks, radis_nsample_min = minimum number of samples that should have sequences for a locus to be included in the analysis. In all data sets, loci for which at least one sample has three or more sequences (radis_npbloci_cutoff = 3) were removed. Bootstrap values (100 replicates) are indicated at nodes.

Appendix S8 Full results of the M0 unconstrained biogeographical model (see methods).

References

Armendáriz-Toledano, F., Niño, A., Sullivan, B.T., Kirkendall, L.R., Zúñiga, G., 2015. A New Species of Bark Beetle, *Dendroctonus mesoamericanus* sp. nov.(Curculionidae: Scolytinae), in Southern Mexico and Central America. Ann. Entomol. Soc. Am. sav020.

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A., Johnson, E.A., 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS One 3, e3376. https://doi.org/10.1371/journal.pone.0003376

Baker, W.L., 1972. Eastern forest insects. US Forest Service.

Bentz, B.J., Régnière, J., Fettig, C.J., Hansen, E.M., Hayes, J.L., Hicke, J.A., Kelsey, R.G., Negron, J.F., Seybold, S.J., 2010. Climate change and bark beetles of the western United States and Canada: direct and indirect effects. Bioscience 60, 602–613.

Bentz, B.J., Stock, M.W., 1986. Phenetic and phylogenetic relationships among ten species of *Dendroctonus* bark beetles (Coleoptera: Scolytidae). Ann. Entomol. Soc. Am. 79, 527–534.

Buerki, S., Forest, F., Alvarez, N., Nylander, J.A.A., Arrigo, N., Sanmartin, I., 2011. An evaluation of new parsimony-based versus parametric inference methods in biogeography: a case study using the globally distributed plant family Sapindaceae. J. Biogeogr. 38, 531–550.

Cai, Y.-W., Cheng, X.-Y., Xu, R.-M., Duan, D.-H., Kirkendall, L.R., 2008. Genetic diversity and biogeography of red turpentine beetle *Dendroctonus valens* in its native and invasive regions. Insect Sci. 15, 291–301. https://doi.org/10.1111/j.1744-7917.2008.00213.x

Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A., Cresko, W.A., 2013. *Stacks*: an analysis tool set for population genomics. Mol. Ecol. 22, 3124–3140.

Catchen, J.M., Amores, A., Hohenlohe, P., Cresko, W., Postlethwait, J.H., 2011. *Stacks*: building and genotyping loci de novo from short-read sequences. G3 Genes, Genomes, Genet. 1, 171–182.

Chen, R., Favret, C., Jiang, L., Wang, Z., Qiao, G., 2016. An aphid lineage maintains a bark-feeding niche while switching to and diversifying on conifers. Cladistics 32, 555–572.

Cibrián Tovar, D., Lagunes Tejeda, Á., Bravo Mojica, H., Vera Graziano, J., Carrillo Sánchez, J.L., Sosa Moss, C., 1995. Insectos forestales de México. Colegio de Postgraduados, Instituto de Fitosanidad, Especialidad en Entomología.

Cruaud, A., Gautier, M., Galan, M., Foucaud, J., Sauné, L., Genson, G., Dubois, E., Nidelet, S., Deuve, T., Rasplus, J.-Y., 2014. Empirical assessment of RAD sequencing for interspecific phylogeny. Mol. Biol. Evol. 31, 1272–1274.

Cruaud, A., Gautier, M., Rossi, J.-P., Rasplus, J.-Y., Gouzy, J., 2016. RADIS: analysis of RAD-seq data for interspecific phylogeny. Bioinformatics 32, 3027–3028.

Cullingham, C.I., Cooke, J.E.K., Dang, S., Davis, C.S., Cooke, B.J., Coltman, D.W., 2011. Mountain pine beetle host-range expansion threatens the boreal forest. Mol. Ecol. 20, 2157–2171. https://doi.org/10.1111/j.1365-294X.2011.05086.x

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9, 772.

Dorchin, A., López-Uribe, M.M., Praz, C.J., Griswold, T., Danforth, B.N., 2018. Phylogeny, new generic-level classification, and historical biogeography of the *Eucera* complex (Hymenoptera: Apidae). Mol. Phylogenet. Evol. 119, 81–92.

Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with

BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969–1973.

Etter, P.D., Preston, J.L., Bassham, S., Cresko, W.A., Johnson, E.A., 2011. Local de novo assembly of RAD paired-end contigs using short sequencing reads. PLoS One 6, e18561.

Fagua, G., Condamine, F.L., Dombroskie, J.J., Byun, B.-K., De Prins, J., Simonsen, T.J., Baez, M., Brunet, B.M.T., Sperling, F.A.H., 2019. Genus delimitation, biogeography and diversification of *Choristoneura* Lederer (Lepidoptera: Tortricidae) based on molecular evidence. Syst. Entomol. 44, 19–38.

Furniss, M.M., 1996. Taxonomic status of *Dendroctonus punctatus* and *D. micans* (Coleoptera: Scolytidae). Ann. Entomol. Soc. Am. 89, 328–333.

Furniss, M.M., Johnson, J.B., 2002. Field guide to the bark beetles of Idaho and adjacent regions. Idaho Forest, Wildlife, and Range Experiment Station, University of Idaho.

Gladenkov, A.Y., Oleinik, A.E., Marincovich Jr, L., Barinov, K.B., 2002. A refined age for the earliest opening of Bering Strait. Palaeogeogr. Palaeoclimatol. Palaeoecol. 183, 321–328.

Godefroid, M., Rasplus, J., Rossi, J., 2016. Is phylogeography helpful for invasive species risk assessment? The case study of the bark beetle genus *Dendroctonus*? Ecography (Cop.). 39, 1–13.

Havill, N.P., Cognato, A.I., Del-Val, E., Rabaglia, R.J., Garrick, R.C., 2019. New Molecular Tools for *Dendroctonus frontalis* (Coleoptera: Curculionidae: Scolytinae) Reveal an East--West Genetic Subdivision of Early Pleistocene Originker. Insect Syst. Divers. 3, 2.

Henigman, J., Ebata, T., Allen, E., Holt, J., Pollard, A., 1999. Field guide to forest damage in British Columbia. British Columbia Ministry of Forests.

Hines, H.M., 2008. Historical biogeography, divergence times, and diversification patterns of bumble bees (Hymenoptera: Apidae: *Bombus*). Syst. Biol. 57, 58–75.

Hipp, A.L., Manos, P.S., González-Rodríguez, A., Hahn, M., Kaproth, M., McVay, J.D., Avalos, S.V., Cavender-Bares, J., 2018. Sympatric parallel diversification of major oak clades in the Americas and the origins of Mexican species diversity. New Phytol. 217, 439–452.

Hopkins, A.D., 1909. Contributions Toward a Monograph of the Scolytid Beetles. The genus *Dendroctonus*. US Government Printing Office, Washington.

Ian Milne, R., 2006. Northern hemisphere plant disjunctions: a window on Tertiary land bridges and climate change? Ann. Bot. 98, 465–472.

Jansen, G., Savolainen, R., Vepsäläinen, K., 2010. Phylogeny, divergence-time estimation,

biogeography and social parasite--host relationships of the Holarctic ant genus *Myrmica* (Hymenoptera: Formicidae). Mol. Phylogenet. Evol. 56, 294–304.

Jones, J.C., Fan, S., Franchini, P., Schartl, M., Meyer, A., 2013. The evolutionary history of *Xiphophorus* fish and their sexually selected sword: a genome-wide approach using restriction site-associated DNA sequencing. Mol. Ecol. 22, 2986–3001.

Jordal, B.H., Sequeira, A.S., Cognato, A.I., 2011. The age and phylogeny of wood boring weevils and the origin of subsociality. Mol. Phylogenet. Evol. 59, 708–724.

Jousselin, E., Cruaud, A., Genson, G., Chevenet, F., Foottit, R.G., Cœur d'acier, A., 2013. Is ecological speciation a major trend in aphids? Insights from a molecular phylogeny of the conifer-feeding genus *Cinara*. Front. Zool. 10, 56.

Kelley, S.T., Farrell, B.D., 1998. Is specialization a dead end? The phylogeny of host use in *Dendroctonus* bark beetles (Scolytidae). Evolution (N. Y). 1731–1743.

Kergoat, G.J., Silvain, J.-F., Delobel, A., Tuda, M., Anton, K.-W., 2007. Defining the limits of taxonomic conservatism in host--plant use for phytophagous insects: Molecular systematics and evolution of host--plant associations in the seed-beetle *genus Bruchus* Linnaeus (Coleoptera: Chrysomelidae: Bruchinae). Mol. Phylogenet. Evol. 43, 251–269.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874.

Labandeira, C.C., LePage, B.A., Johnson, A.H., 2001. A *Dendroctonus* bark engraving (Coleoptera: Scolytidae) from a middle Eocene *Larix* (Coniferales: Pinaceae): early or delayed colonization? Am. J. Bot. 88, 2026–2039.

Lanier, G.N., 1981. Cytotaxonomy of *Dendroctonus*, in: Application of Genetics and Cytology in Insect Systematics and Evolution, Proceedings of the 1980 Annual Meeting of the Entomological Society of America. University of Moscow, Idaho, pp. 33–66.

Larson, E.R., Olden, J.D., 2012. Using avatar species to model the potential distribution of emerging invaders. Glob. Ecol. Biogeogr. 21, 1114–1125. https://doi.org/10.1111/j.1466-8238.2012.00758.x

Leaché, A.D., Fujita, M.K., Minin, V.N., Bouckaert, R.R., 2014. Species delimitation using genome-wide SNP data. Syst. Biol. 63, 534–542.

Meseguer, A.S., Coeur d'acier, A., Genson, G., Jousselin, E., 2015. Unravelling the historical biogeography and diversification dynamics of a highly diverse conifer-feeding aphid genus.

J. Biogeogr. 42, 1482–1492.

Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees, in: 2010 Gateway Computing Environments Workshop (GCE). pp. 1–8.

Morales-Castilla, I., Davies, T.J., Pearse, W.D., Peres-Neto, P., 2017. Combining phylogeny and co-occurrence to improve single species distribution models. Glob. Ecol. Biogeogr.

Papadopoulou, A., Anastasiou, I., Vogler, A.P., 2010. Revisiting the insect mitochondrial molecular clock: the mid-Aegean trench calibration. Mol. Biol. Evol. 27, 1659–1672.

Paradis, E., Claude, J., Strimmer, K., 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20, 289–290.

Pistone, D., Gohli, J., Jordal, B.H., 2017. Molecular phylogeny of bark and ambrosia beetles (Curculionidae: Scolytinae) based on 18 molecular markers. Syst. Entomol.

Prebus, M., 2017. Insights into the evolution, biogeography and natural history of the acorn ants, genus *Temnothorax* Mayr (hymenoptera: Formicidae). BMC Evol. Biol. 17, 250.

Rambaut, A., Drummond, A.J., 2009. Tracer: MCMC trace analysis tool, version 1.5. Available at website http://tree.bio.ed.ac.uk/software/tracer.

Ree, R.H., Moore, B.R., Webb, C.O., Donoghue, M.J., 2005. A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. Evolution (N. Y). 59, 2299–2311.

Ree, R.H., Smith, S.A., 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. Syst. Biol. 57, 4–14.

Reeve, J.D., Anderson, F.E., Kelley, S.T., 2012. Ancestral state reconstruction for *Dendroctonus* bark beetles: evolution of a tree killer. Environ. Entomol. 41, 723–730.

Sanderson, M.J., 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Mol. Biol. Evol. 19, 101–109.

Sanderson, M.J., 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy.

Sanmartin, I., Enghoff, H., Ronquist, F., 2001. Patterns of animal dispersal, vicariance and diversification in the Holarctic. Biol. J. Linn. Soc. 73, 345–390.

Schneck, R., Micheels, A., Mosbrugger, V., 2012. Climate impact of high northern vegetation: Late Miocene and present. Int. J. Earth Sci. 101, 323–338.

Sequeira, A.S., Farrell, B.D., 2001. Evolutionary origins of Gondwanan interactions: How old are *Araucaria* beetle herbivores? Biol. J. Linn. Soc. 74, 459–474.

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., Flook, P., 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 87, 651–701.

Six, D.L., Bracewell, R., 2015. *Dendroctonus*, in: Bark Beetles: Biology and Ecology of Native and Invasive Species. Academic Press, Elsevier Publishing, US, pp. 305–350.

Six, D.L., Klepzig, K.D., 2004. *Dendroctonus* bark beetles as model systems for studies on symbiosis. Symbiosis 37, 207–232.

Smith, S.A., 2009. Taking into account phylogenetic and divergence-time uncertainty in a parametric biogeographical analysis of the Northern Hemisphere plant clade Caprifolieae. J. Biogeogr. 36, 2324–2337.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.

Team, R.C., others, 2014. R: A language and environment for statistical computing.

Thatcher, T.O., 1954. A New Species of *Dendroctonus* from Guatemala (Scolytidae) Publication No. 6. Coleopt. Bull. 3–6.

Tiffney, B.H., 1985. The Eocene North Atlantic land bridge: its importance in Tertiary and modern phytogeography of the Northern Hemisphere. J. Arnold Arbor. 66, 243–273.

Víctor, J., Zúñiga, G., 2016. Phylogeny of *Dendroctonus* bark beetles (Coleoptera: Curculionidae: Scolytinae) inferred from morphological and molecular data. Syst. Entomol. 41, 162–177.

Vila, R., Bell, C.D., Macniven, R., Goldman-Huertas, B., Ree, R.H., Marshall, C.R., Bálint, Z., Johnson, K., Benyamini, D., Pierce, N.E., 2011. Phylogeny and palaeoecology of *Polyommatus* blue butterflies show Beringia was a climate-regulated gateway to the New World. Proc. R. Soc. London B Biol. Sci. 278, 2737–2744.

Weed, A.S., Ayres, M.P., Hicke, J.A., 2013. Consequences of climate change for biotic disturbances in North American forests. Ecol. Monogr. 83, 441–470.

Wen, J., Nie, Z.-L., Ickert-Bond, S.M., 2016. Intercontinental disjunctions between eastern Asia and western North America in vascular plants highlight the biogeographic importance of the Bering land bridge from late Cretaceous to Neogene. J. Syst. Evol. 54, 469–490.

Williams, S.E., Shoo, L.P., Isaac, J.L., Hoffmann, A.A., Langham, G., 2008. Towards an integrated framework for assessing the vulnerability of species to climate change. PLoS Biol. 6, e325.

Wolfe, J.A., 1994. An analysis of Neogene climates in Beringia. Palaeogeogr. Palaeoclimatol. Palaeoecol. 108, 207–216.

Wood, S.L., 1985. Aspectos taxonómicos de los Scolytidae. Proc. 2nd Natl. 170–174.

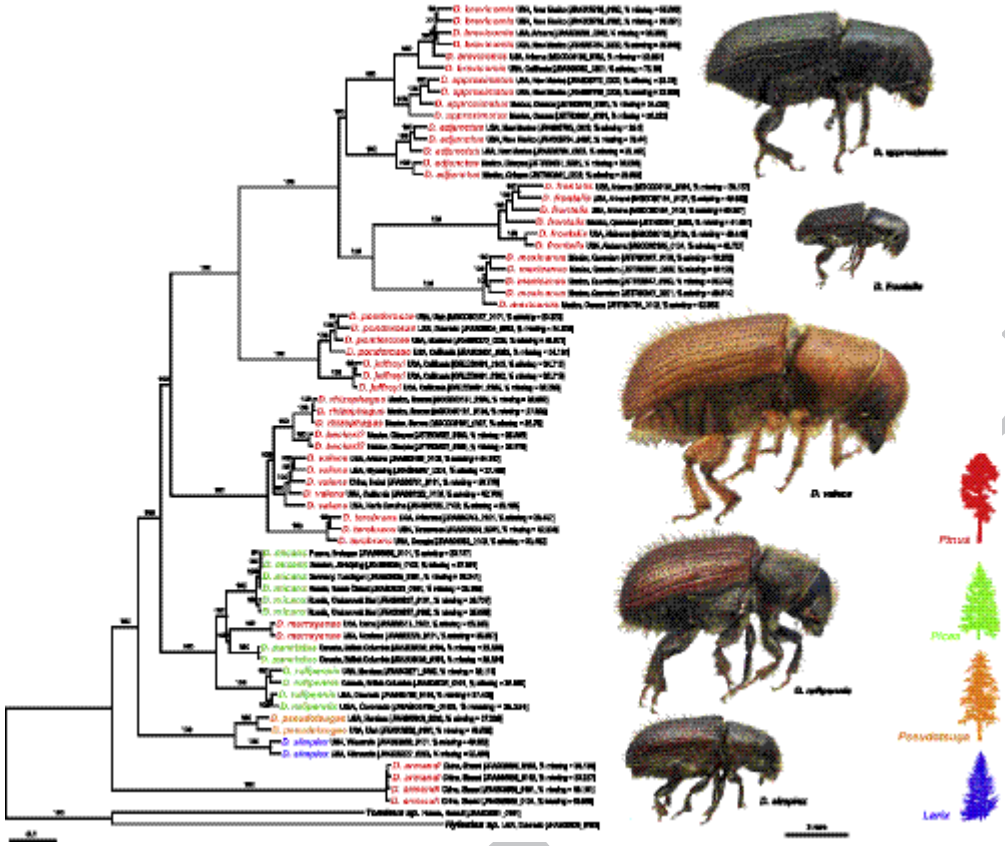
Wood, S.L., 1982. The bark and ambrosia beetles of North and Central America (Coleoptera: Scolytidae), a taxonomic monograph. Gt. Basin Nat. Mem. 6, 1–625.

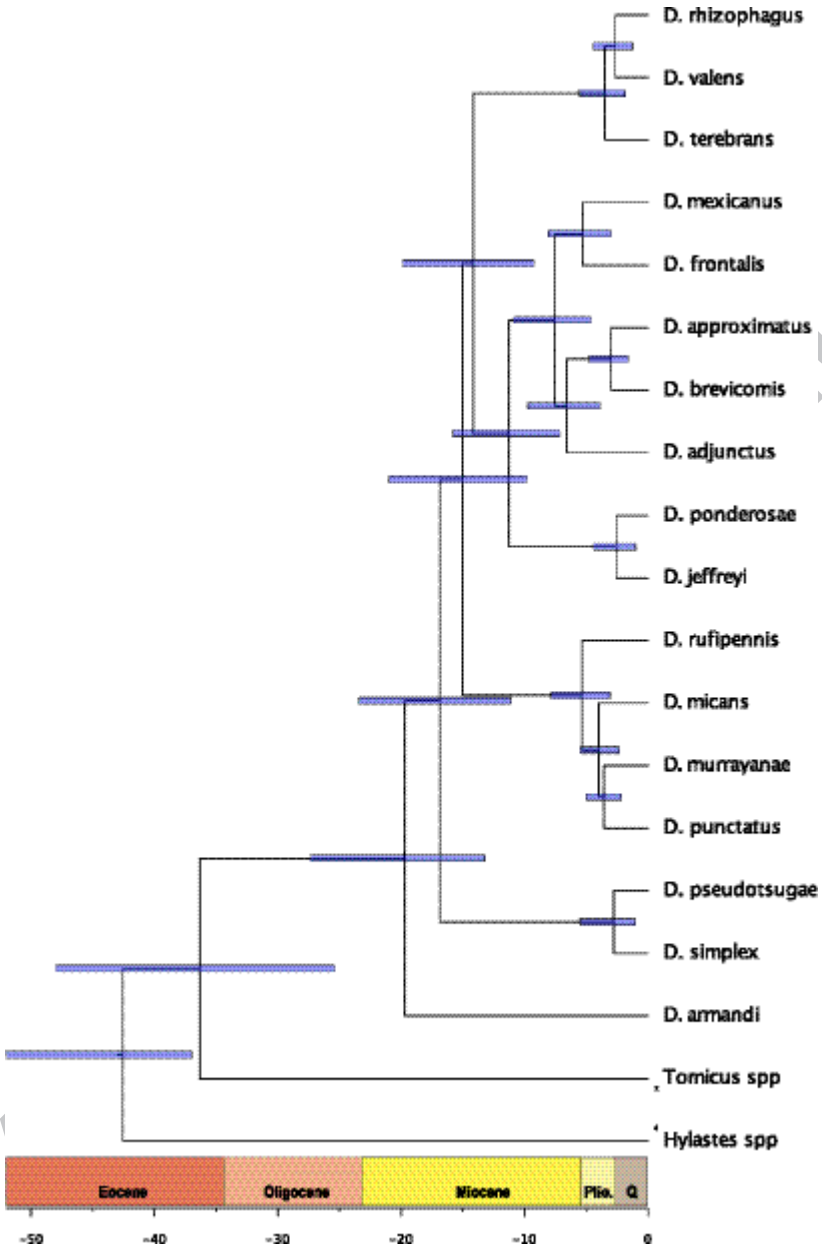
Wood, S.L., 1963. A revision of the bark beetle genus *Dendroctonus* Erichson (Coleoptera: Scolytidae). West. North Am. Nat. 23, 1–116.

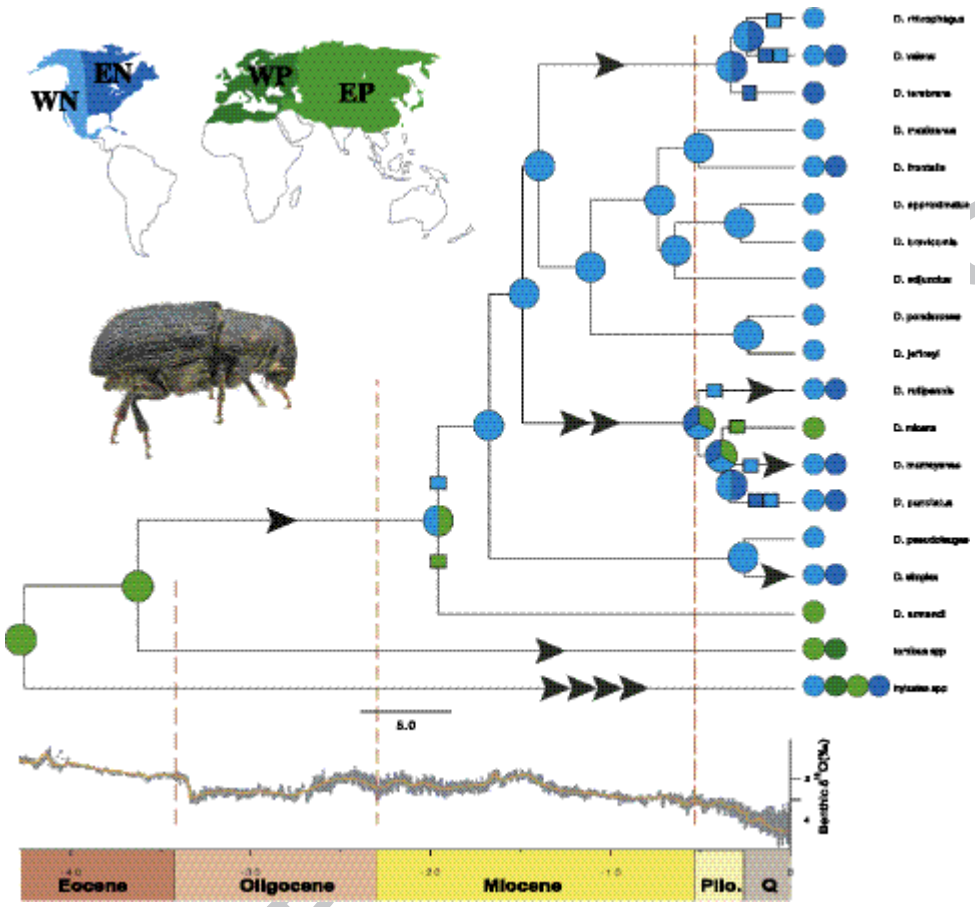
Yan, Z., Sun, J., Don, O., Zhang, Z., 2005. The red turpentine beetle, *Dendroctonus valens* LeConte (Scolytidae): an exotic invasive pest of pine in China. Biodivers. Conserv. 14, 1735–1760. https://doi.org/10.1007/s10531-004-0697-9

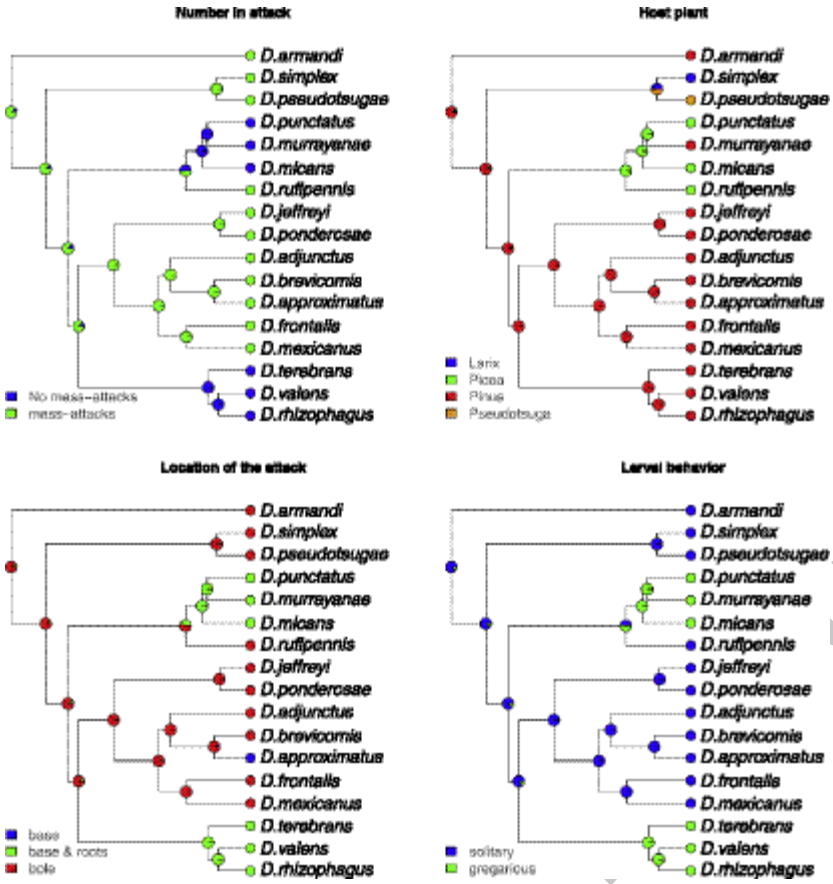
Yoshizawa, K., Johnson, K.P., Yao, I., Casasola González, J.A., Bess, E., Garcia Aldrete, A.N., 2017. Multiple trans-Beringia dispersals of the barklouse genus *Trichadenotecnum* (Insecta: Psocodea: Psocidae). Biol. J. Linn. Soc. 121, 501–513.

Zúñiga, G., Cisneros, R., Hayes, J.L., Macias-Samano, J., 2002. Karyology, geographic distribution, and origin of the genus *Dendroctonus* Erichson (Coleoptera: Scolytidae). Ann. Entomol. Soc. Am. 95, 267–275.



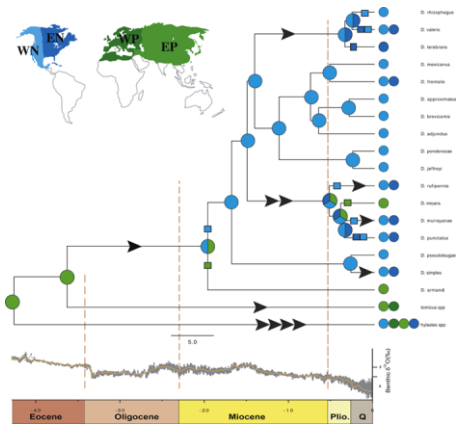






Combination of parameters	# RAD tags	# bases in alignment	% of missing data	% variable sites	% parsimony informative sites	% GC
n = 10, radis_nsample_min = 35% of the samples	22005	2244510	39.66	20.6	16.8	44.1
n = 10, radis_nsample_min = 50% of the samples	13811	1408722	28.22	22.0	18.0	44.4
n = 10, radis_nsample_min = 70% of the samples	6444	657288	15.97	21.8	17.8	44.8
n = 12, radis_nsample_min = 35% of the samples	23570	2404140	39.08	22.3	18.3	44.1
n = 12, radis_nsample_min = 50% of the samples	15173	1547646	28.13	23.6	19.4	44.5
n = 12, radis_nsample_min = 70% of the samples	7100	724200	15.97	23.4	19.1	44.9

Graphical abstract



Highlights

- Using RADseq we inferred the first fully-resolved phylogeny of *Dendroctonus* genus
- *Dendroctonus* presumably originated during the early Miocene
- The *mrca* of all extant *Dendroctonus* was distributed across Palearctic and Nearctic
- The *mrca* presumably bored individual galleries and mass attacked pines
- The gregarious larval behavior has independently evolved twice