Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Host-tree association and evolutionary history shape components of genomic diversity across a saproxylic beetle assemblage

Rama Krovi

sarvaniramz@gmail.com

Institute of Systematics and Evolution of Animals, Polish Academy of Sciences
https://orcid.org/0000-0003-4798-3705

Nermeen Amer
Institute of Systematics and Evolution of Animals, Polish Academy of Sciences

Alicja Wierzbicka
National Research Institute of Animal Production

Tomasz Szmatoła
University of Agriculture in Kraków

Maria Oczkowicz
National Research Institute of Animal Production

Radosław Plewa
Forest Research Institute

Tomasz Jaworski
Forest Research Institute

Marcin Kadej
University of Wrocław

Adrian Smolis
University of Wrocław

Łukasz Kajtoch
Institute of Systematics and Evolution of Animals, Polish Academy of Sciences

Article

Keywords: comparative population genomics, ddRAD-seq, genetic diversity, host-tree association, macrogenetics, phylogenetic signal, saproxylic beetles

Posted Date: October 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11127962/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is no duality of interest

Abstract

Intraspecific genetic diversity is a crucial yet incompletely mapped component of biodiversity, and comparative studies frequently treat heterozygosity and nucleotide diversity as interchangeable despite their differing genomic denominators and demographic timescales. Using a standardized de novo ddRAD-seq workflow, we generated genome-wide data for 759 individuals representing 149 populations, 26 saproxylic beetle species and 14 families across eight forest complexes, and tested whether species occurrence, host-tree association and phylogenetic history predict distinct components of genomic diversity. Population-level linear mixed models ranked the intercept-only null model highest for all-position nucleotide diversity (π), variant-site observed and expected heterozygosity (H_o and H_e) and the inbreeding coefficient (F_{IS}), indicating that species identity and site structure govern most variance. Species-level Phylogenetic Generalized Least-Squares models, however, resolved a clear division among metrics: H_e was significantly elevated in deciduous-associated species relative to conifer specialists, whereas π carried a consistent phylogenetic signal that persisted after excluding the influential lineage *Rhysodes sulcatus*. Occurrence tier predicted no metric after correction, and forest-management effects were minor and confounded with forest-complex identity. Ecological and evolutionary predictors thus shape distinct components of genomic diversity: H_e aligns with contemporary ecological association, while π reflects shared evolutionary history.

Introduction

Intraspecific genetic diversity contributes to population persistence and adaptive potential, yet it remains the least understood level of biodiversity. Macrogenetics seeks general explanations for variation in genetic diversity by comparing many populations and species across space (Blanchet et al. 2017; Leigh et al. 2021). Candidate explanations include abundance, geographic range, ecological specialization, dispersal, evolutionary history and anthropogenic disturbance. Their relative importance remains uncertain because comparative datasets often combine different markers, sampling designs and analytical pipelines.

Genome-wide data generated under a common workflow can reduce some of this heterogeneity (Yahara et al. 2010, Coates et al. 2018). They do not, however, eliminate differences among diversity metrics. Nucleotide diversity calculated across all callable positions incorporates invariant sites and represents per-site sequence diversity, whereas heterozygosity reported only among retained variant sites is conditional on discovery and filtering (Schmidt et al. 2021). Other metrics are mostly derived from one of these two basic variables (i.e. inbreeding (F_{IS}) is measured based on observed and expected heterozygosity), and are not independent measures of diversity. Treating these responses as interchangeable can obscure whether a predictor is associated with a general genomic property or with a particular metric and denominator (Brault et al. 2026).

Saproxylic beetles provide a useful assemblage for testing ecological and evolutionary hypotheses. They depend on deadwood and tree-related microhabitats, contribute substantially to decomposition, and

include both abundant taxa (some treated as pests) and rare specialists of conservation concern (Stokland et al. 2012; Seibold et al. 2015). Host-tree association, microclimatic niche, and species distribution (range) may influence effective population size and population connectivity, which can be estimated using genomic data. At the same time, mutation rate, life-history traits and demographic history may be conserved among related lineages, generating similar genetic patterns among congeners or confamilial species.

Forest management may also be one of the factors that correlate with beetle genetic diversity because it alters the quantity, quality and continuity of deadwood, although human-made alterations are relatively recent on an evolutionary timescale and therefore their effect may not yet be detectable.

Previous studies have already demonstrated that both the geographic location and management regime of forests influence the genetic diversity of saproxylic beetles (Amer et al. 2026), and that deadwood diversity, together with the duration of forest protection, are key determinants of genetic diversity in this ecological group (Krovi et al., 2026). However, these studies did not examine how phylogenetic affinity, biogeographic history, microclimatic preferences, trophic specialization, and other ecological traits shape patterns of genetic diversity across multiple saproxylic beetle species. The novel contribution of the present study is therefore a cross-taxonomic, assemblage-scale assessment of the effects of species traits, phylogenetic affiliation, and forest management on genome-wide within-population genetic diversity across 26 saproxylic beetle species. This addresses a macrogenetic question that has not been explored in previous studies. Specifically, we tested the following three hypotheses:

H1. Species with high population abundance are expected to exhibit greater genomic diversity than habitat specialists, particularly taxa associated with old-growth forest remnants, owing to greater habitat connectivity and larger, more persistent populations.

H2. Species associated with deciduous deadwood are expected to exhibit greater genomic diversity than conifer specialists, reflecting the historical predominance of broadleaved forests and contrasting outbreak dynamics.

H3. Shared evolutionary history is predicted to structure genomic diversity, most clearly for all-position nucleotide diversity, whereas variant-site heterozygosity may be more responsive to recent ecological and demographic processes.

Materials and Methods

Study system and sampling

We analysed 759 beetle specimens representing 26 species and 14 families Table S1. Samples formed 149 species-by-site population records across 37 sites in eight Polish forest complexes: Augustów (AF), Białowieża (BF), Carpathian Mountains (CM), Dolina Baryczy (DB), Holy Cross Mountains (HM), Knyszyń (KF), Oder Forests (OF) and Silesian Forests (SF). Three complexes were classified as primeval (BF, CM,

HM) and the remaining five as managed, shown in Figure 1. Because every site within a complex had the same management category, management was completely nested within the forest complex. We consequently treat management as an exploratory between-forest comparison. Population records contained five individuals on average; species were represented by 3–10 population records.

Collection of protected beetle species and sampling in protected areas was conducted under permissions granted by the Polish Ministry of Environment (DOP-1.61.61.2021.TP 1516071.4994962.3977511 and DOP-1.61.61.2021.TP 1516071.4994586.3977515), the General Directorate of Environmental Protection (DZP-WF.6401.68.2021.AS), and the Regional Directorates of Environmental Protection in Białystok (WPN.6205.74.2021.MM), Rzeszów (WPN.6205.116.2021.MP.2 and WPN.6205.114.2021.ŁL.2), Kielce (WPN.I.6205.1.49.2021.EJP.2) and Wrocław (WPN.6205.143.2021.AR).

Traits data collection

Species traits were compiled from the project trait database (organised based on literature, e.g. Seibold et al. 2015, Wende et al. 2017, Hagge et al. 2021, including expert knowledge of co-authors of this study), and checked for consistency within species. Candidate traits included: “family” (taxonomic assignment to the beetle family), “abundance” (having populations with numerous individuals vs rare), “distribution” (boreal vs temperate vs boreal & temperate), “conservation protection status” (protected vs not), “shade tolerance” (moist and cold vs dry and warm), “economic value” (neutral vs pest, in respect to timber production), “trophic guild” (invertivorous vs xylovorous vs fungivorous), “microhabitat” (living under bark vs wood vs tree cavities vs fungi), “settlement” (deadwood decay: fresh vs old), “tree type” (coniferous vs deciduous vs various =both), “host tree” (host tree species), “body size” (imago: small <10 mm, medium 10-20 mm, large >20 mm), “dispersal ability” (low vs moderate vs high), “voltinism” (uni- vs multivoltine= number of generations/year). Categorical predictors were retained for comparative testing only when every level contained at least five species. This rule retained the “abundance”, “host-tree” category and “microclimatic” niche among the hypothesis-level predictors. The complete trait list and level counts are provided in Supporting Tables S2–S3a.

DNA isolation, ddRAD sequencing and variant processing

High molecular weight DNA was extracted using the Sherlock AX kit (A&A Biotechnology). ddRAD-seq libraries were prepared following Bayona-Vásquez et al., (2019) with modifications by Amer et al., (2026). Double-digest RAD libraries were prepared using NheI and EcoRI, size-selected to 400–600 bp and sequenced as 150-bp paired-end reads on an Illumina NovaSeq X Plus. Reads were demultiplexed and quality filtered before species-specific de novo assembly in Stacks v2.68 (Rochette et al. 2019). The documented workflow used a minimum stack depth of three and species-specific assembly optimization. An initial Stacks populations run required loci in at least two populations and at least 80% of individuals per population. PLINK filtering (Purcell et al. 2007) then retained markers with minor-allele frequency ≥0.03 and no more than 30% missing genotypes; individuals with >50% missing data were screened before the final whitelist-based populations run.

Genetic diversity metrics

Population statistics were extracted from the Stacks populations.sumstats_summary.tsv files. The primary response, nucleotide diversity (π), was taken from the “All positions” section and therefore included invariant callable sites in the denominator. Observed heterozygosity (H□), expected heterozygosity (H□) and F_{IS} were taken from the “Variant positions” section and describe retained variable sites. We retained both all-position and variant-position values in the audit dataset, but did not treat the denominators as equivalent. Allelic richness was excluded because the available values exceeded the theoretical maximum for correctly rarefied biallelic SNPs. Per-species mean callable-site counts and sample coverage, together with sampling design and species-mean diversity, are reported in Supporting Table S2&S4.

Phylogeny

The 26-species phylogeny was derived from the Open Tree of Life using rotl (Michonneau et al. 2016), then processed using ape (Paradis & Schliep 2019) and phytools (Revell 2012) after taxonomic reconciliation. Polytomies were resolved deterministically and Grafen branch lengths were applied where the induced tree was not ultrametric (Grafen 1989). Because this synthetic phylogeny is not time calibrated, phylogenetic estimates are interpreted as comparative structure rather than evolutionary rates. Pagel’s λ and Blomberg’s K were calculated for raw and log-transformed metrics where transformation was defined, using 1,000 permutations for K. Sensitivity analyses were repeated after excluding Rhysodes sulcatus, a relict species associated with European primeval forests (Kostanjsek et al. 2018), the sole representative of the suborder Adephaga (all other taxa belonging to Polyphaga; McKenna et al. 2019), which exhibited substantial phylogenetic leverage.

Statistical analysis

Analyses were conducted at both population and species levels in R version 4.3.0 (R Core Team 2023)

1. Population-level Linear Mixed Models (LMMs):

Population-level variations were evaluated using Linear Mixed Models (LMMs) implemented in the lme4 package (Bates et al. 2015) in R. To account for hierarchical experimental structure and spatial non-independence, nested random effects were included for taxonomic group (Family/Species) and sampling design (Forest/Site) (Bolker et al. 2009). The candidate model set was restricted a priori to: (i) Intercept-only null, (ii) Occurrence tier, (iii) Host-tree category, (iv) Microclimatic niche, (v) Standardized body size, and (vi) Occurrence + Host-tree category. To facilitate convergence for small numerical values (e.g., π), responses were scaled by a constant prior to modeling, and parameter estimates were back-transformed to original units for reporting. Models were fitted via Maximum Likelihood (ML) and ranked using corrected Akaike Information Criterion (AICc) with MuMIn package (Bartoń 2010; Burnham &

Anderson 2004). To avoid over-fitting species-level traits evaluated on sub-sampled populations, AICc small-sample corrections were calculated using the number of species (N = 26) rather than total population records (N = 149). The top-performing models were re-fitted using Restricted Maximum Likelihood (REML). Fixed-effect degrees of freedom and P-values were obtained using Satterthwaite's approximation as implemented in lmerTest (Kuznetsova et al. 2017). Marginal and conditional (R^2) values were calculated using performance (Lüdecke et al. 2021).

2. Species-level Phylogenetic Generalized Least Squares (PGLS):

Species-level metric means were evaluated using PGLS models in caper (Orme et al. 2011), estimating Pagel's λ by Maximum Likelihood. Response variables π, H□, and H□ were log-transformed; F_{IS} was analyzed on its original scale due to zero/negative values. Single-predictor models evaluated occurrence tier, host-tree category, microclimatic niche, and body size. Benjamini–Hochberg (BH) false discovery rate corrections were applied to p-values across single-predictor models within each response. Sensitivity runs evaluated the full 26-species dataset against the 25-species subset excluding *R. sulcatus*.

3. Forest Management Evaluation:

Forest management category was excluded from candidate ecological model selection to prevent confounding effects. Instead, management was evaluated as an exploratory fixed factor (primeval vs. managed) within an LMM retaining the Forest/Site random hierarchy. An additional exploratory LMM evaluated protected vs. unprotected population records (74 vs. 75 populations). Model diagnostics included residual normality checks, variance inflation factors (VIF), spatial autocorrelation tests (Moran's I), and leave-one-species-out cross-validation.

Results

Population-Level Baseline and Multi-Level Variance (Null Models)

In population-level linear mixed-effects models (LMMs), the intercept-only null model ranked first across all four genomic diversity metrics. For all-position nucleotide diversity (π), the second-ranked candidate model was standardized body size (ΔAICc = 2.21), whereas the occurrence tier and host-tree category yielded ΔAICc values of 3.78 and 4.62, respectively. For variant-site observed (H□) and expected (H□) heterozygosity, host-tree association ranked second but remained behind the null (ΔAICc =1.30 for H_o; ΔAICc = 1.91 for H_e). The null model was also preferred for the inbreeding coefficient (F_{IS}). Conditional R^2 values for null models were high (0.69-0.92), showing that species identity and multi-level site clustering account for most variance across all metrics rather than ecological predictor effects. (Supporting Table S3b)

Occurrence Tier and Macroecological Prevalence (H1)

Occurrence pattern did not consistently predict any genomic metric after false discovery rate (BH) correction (Table 1). For all-position π, common species showed an initial positive unadjusted contrast relative to rare species (β= 0.267, raw P = 0.034), but this effect was not retained following Benjamini–Hochberg correction (adjusted P = 0.172). For variant-site H and H, common and pest species tended toward lower point estimates than rare species, though no contrasts reached statistical significance. F_{IS} contrasts across occurrence tiers were consistently weak and non-significant.

Interspecific Trait Specialization and Host-Tree Association (H2)

In species-level PGLS models, host-tree association significantly predicted variant-site H (Table 1; Figure 2). Deciduous-associated species exhibited significantly higher H than conifer-associated species (log-scale β = 0.212, SE = 0.066, P = 0.0041, adjusted P = 0.036). In contrast, host-tree association showed no detectable effect on all-position π (β = −0.056, SE = 0.116, P = 0.635). Although host-tree category and microclimatic niche were partially correlated, generalized variance inflation factors (GVIF 1.21 and 1.59) confirmed predictor separability. Microclimatic niche alone did not predict any metric (H β = 0.067, P = 0.444; H β = 0.111, P = 0.438; π β = 0.080, P = 0.496), and species utilizing "various" host trees exhibited expected heterozygosity levels comparable to conifer specialists, confirming that elevated H specifically tracks deciduous host-tree association independently of shade or moisture preferences.

Evolutionary History and Metric-Specific Phylogenetic Signal (H3)

All-position π exhibited strong, consistent phylogenetic signal across both log and raw scales (Table 2; Figure 3). On the log scale, Pagel's λ was 0.583 (P = 0.0099) for the full 26-species assemblage and remained stable at λ = 0.580 (P = 0.0126) in sensitivity tests excluding *Rhysodes sulcatus*. Raw-scale estimates were similarly stable (λ = 0.472, P = 0.032 for full data; λ = 0.468, P = 0.037 without *R. sulcatus*). This strong signal in all-position π contrasted sharply with variant-site metrics, which showed weak or lineage-sensitive signals. Variant-site H displayed a lineage-sensitive signal in the full dataset (λ = 0.657, P = 0.0296), but this signal was not retained after removal of *R. sulcatus* (λ = 0.257, P = 0.232). Variant-site H_e did not exhibit significant Pagel's λ in either dataset, and its Blomberg's K signal was non-significant in the sensitivity run.

Forest Management and Landscape Nesting

In exploratory LMMs, managed forest complexes showed slightly higher point estimates across all four metrics, but all 95% confidence intervals overlapped zero (0.00011 for π, 95% CI: −0.00008 to 0.00030; 0.005 for H, 95% CI: (−0.014 to 0.025; 0.010 for H, 95% CI: −0.007 to 0.027; 0.009 for F_{IS}, 95% CI: −0.003 to 0.020). In exploratory models comparing protected versus unprotected population records (74 vs. 75 records), protected status showed no detectable association with any metric (H β = 0.005, 95% CI −0.009 to 0.018, P = 0.497; H β = −0.001, 95% CI −0.011 to 0.009, P = 0.803; F_{IS} β = −0.011, 95% CI

−0.023 to 0.000, P = 0.058). Because the management category was represented by eight forest complexes and was completely nested within forest complex identity, these LMM estimates might reflect broad spatial variation among sampled landscapes rather than direct management or protection effects.

Discussion

Previous research on saproxylic beetles examined associations between ecological and morphological traits and extinction risk, community composition and environmental responses (Seibold et al. 2015; Hagge et al. 2021), without testing how these traits relate to within-species genomic diversity. Our study links these approaches by comparing multiple measures of genomic diversity across 26 species spanning several beetle families and ecological strategies. Although limited species coverage constrains broad generalisation across Coleoptera, it provides a comparative test of ecological and phylogenetic associations with population genetic diversity in saproxylic beetles.

Our comparative analysis reveals that genomic diversity is shaped by distinct, scale-dependent processes. Variance in genomic diversity is primarily structured by baseline species identity, with ecological traits explaining specific, metric-dependent fractions of diversity. Crucially, our findings show that occurrence patterns do not consistently dictate genomic diversity (H_1), whereas host-tree specialization specifically drives variant-site expected heterozygosity (H_e) in broadleaf-associated taxa (H_2). Furthermore, deep evolutionary history strongly constrains all-position nucleotide diversity (π), whereas variant-site heterozygosity reflects recent ecological specialization (H_3).

Occurrence patterns do not dictate genomic diversity (H_1)

Contrary to the hypothesis that species with populations comprising numerous individuals maintain elevated genomic diversity relative to rare or localized specialists, abundance tier showed no significant association with any metric after false discovery rate correction. While abundant species exhibited a non-significant positive trend for all-position π, variant-site heterozygosity (H_o and H_e) tended to be slightly lower in abundant (including pest) species than in rare taxa. This disconnect highlights that macroecological prevalence (geographic range and local abundance) does not directly reflect evolutionary effective population size (Ne). Specialized, habitat-restricted saproxylic beetles may maintain historically stable, highly connected populations in primeval forest refugia, buffering them against rapid genetic drift despite localized distribution ranges (Seibold et al. 2015). Conversely, abundant (including pest) species frequently undergo local population outbreaks followed by bottlenecks- (Lantschner & Corley 2023), inducing recurring founder effects that temporarily inflate variant-site homozygosity, and may have implications for the genomics of irruptive beetles (Mykhailenko et al. 2024). Finally, it should be emphasized that not all rare species have restricted distributions and, conversely, some range-restricted species may have abundant populations. Therefore, assigning taxa to two simple categories can be ambiguous, and abundance should not be treated as equivalent to

commonness or widespread distribution. More quantitative measures of population abundance and range occupancy would provide a stronger test of these relationships.

Host-tree association drives variant-site expected heterozygosity (H2)

Species associated with deciduous host trees exhibited significantly higher expected heterozygosity than conifer specialists. This relationship was independent of microclimatic niche preferences (shade and moisture), as confirmed by low variance inflation factors. Temperate forest habitats in Central Europe represent ancestral, structurally complex broadleaf ecosystems that have maintained continuous long-term microhabitat availability throughout post-glacial recolonization (Roberts et al. 2018). Conifers were also present in these forests, although the dominance of pines and spruces is relatively recent and was caused by human forest management (Jansen et al. 2017). The broad dietary and structural niches provided by decaying broadleaf wood support larger, long-term effective population sizes in deciduous specialists (Seibold et al. 2015; Ulyshen 2018), which is why most primaeval relics are among taxa associated with deciduous trees. In contrast, conifer-associated saproxylic beetles historically occupied restricted boreo- montane refugia prior to modern commercial forestry expansions (Kajtoch et al. 2022, Krovi et al. 2025). The lower H_e observed in conifer specialists likely reflects historical demographic bottlenecks rather than modern forestry practices, as recent conifer plantation expansions have occurred too recently to rebuild standing variant-site heterozygosity (Lachat et al. 2012).

Evolutionary history and the denominator distinction (H3)

Previous beetle phylogenomic studies have examined deep evolutionary relationships and diversification across Coleoptera (Zhang et al. 2018; McKenna et al. 2019), while studies within particular herbivorous groups have linked diversification to host-plant use (McKenna et al. 2009). Until now, however, no research has linked phylogeny with intraspecific genomic diversity across multiple taxa for deadwood-dwelling organisms.

Our results strongly support H_3 for sequence-level divergence while demonstrating that phylogenetic signal is metric-dependent. All-position nucleotide diversity retained a strong, stable phylogenetic signal, remaining invariant to the removal of the basal lineage *Rhysodes sulcatus*. In contrast, variant-site heterozygosity (H_e) showed no significant phylogenetic signal. This distinction points at a methodological denominator distinction in restriction site-associated DNA sequencing (RAD-seq) (Harvey et al. 2016). All-position π incorporates invariant callable sites into the denominator, causing it to reflect deep, genome-wide mutation accumulation across evolutionary time (Nei & Li 1979; Lynch & Conery 2003). Because mutation accumulation is constrained by lineage age and historical metabolic rates, π exhibits strong phylogenetic conservatism (Pagel 1999; Blomberg et al. 2003). Conversely, variant-site metrics (H_o and H_e) restrict the denominator to retained polymorphic SNPs, isolating active allele frequency distributions that respond rapidly to recent ecological selection, host specialization, and population dynamics (Luikart et al. 2003; Allendorf et al. 2012). Inbreeding (F_{IS}) also retained evidence of signal after lineage removal, but it is a composite of H□ and H□ and should not be interpreted as an

independent measure of genetic diversity. Evaluating both nucleotide diversity and heterozygosity simultaneously prevents erroneous conclusions regarding species-wide genetic vulnerability.

Spatial nesting masks direct forest management effects

Exploratory comparisons between managed and primeval forest complexes revealed no consistent management effects across any diversity metric. Although managed forests showed slightly higher point estimates, the 95% confidence intervals overlapped zero in all models. This absence of a direct management signal stems from spatial nesting: because management status was uniform within entire forest complexes, management effects were completely confounded with broad-scale geographic identity and regional forest history (Dormann et al. 2007). Population-level linear mixed models confirmed that random site and species identity accounted for 69–92% of total variance.

First, although our sampling encompassed both primeval and managed forests, populations of rare species and most common forest specialists in commercially managed forests were almost exclusively found in relatively well-preserved old-growth stands, including not only nature reserves but also forest patches that had remained free from logging for extended periods (Amer et al. 2026). In contrast, only abundant species, including pests, were consistently sampled in intensively managed stands. Second, forests in Central Europe, including those in Poland, comprise a heterogeneous mosaic of stand types, and even commercially managed forests often contain remnants of old-growth habitat (only some being primaeval) (Banaś et al. 2014, Bujoczek et al. 2021). The absence of a detectable management effect may partly reflect this complex forest structure, as well as the necessarily simplified classification of beetle populations into either primeval or managed forest categories. Consequently, evaluating forest management impacts requires finer-scale microhabitat measurements such as local deadwood volume and diversity, or the availability of old trees, rather than categorical landscape designations (Müller et al. 2010).

Conclusions

The main methodological outcome of this study is that some hypotheses can be evaluated with particular genomic diversity metrics, whereas others remain inconclusive. This is especially important in cross-species de novo SNP datasets, where catalogue size, locus recovery and the pool of retained variants differ among lineages.

Our findings caution against using macroecological rarity or categorical protection status as sole proxies for genomic vulnerability in the conservation of saproxylic beetles (Calix et al. 2018; Eckelt et al. 2018). Rare saproxylic specialists inhabiting broadleaf forest remnants often harbor substantial variant-site heterozygosity, underscoring the conservation value of primeval broadleaf forest fragments (Krovi et al. 2025; Amer et al. 2026). For conservation (rare taxa) or management (pests), the data support lineage-aware monitoring and careful reporting of the genetic component being measured. Maintaining diverse and continuous deadwood resources remains ecologically well motivated (Krovi et al. 2026), but a causal genetic benefit should be tested with direct resource and demographic measurements.

Across the studied saproxylic beetle assemblage, host-tree association and phylogenetic history were linked to distinct patterns of genomic variation. Deciduous-tree association predicted expected heterozygosity among retained variant sites, but not all-position nucleotide diversity. In contrast, π exhibited the strongest and most consistent phylogenetic signal. Abundance patterns did not consistently predict diversity, and the clustered management comparison did not permit a general causal conclusion. These findings emphasize that genomic diversity is not a single property but a hierarchy-dependent characteristic whose interpretation depends on the metric used.

References

1. Allendorf, F. W., Luikart, G. H., & Aitken, S. N. (2012). *Conservation and the Genetics of Populations* (2nd ed.). Wiley-Blackwell.
2. Amer, N. R., Krovi, R. S., Wierzbicka, A., Plewa, R., Kadej, M., Jaworski, T., Smolis, A., Oleksa, A., Szmatoła, T., Oczkowicz, M., & Kajtoch, Ł. (2026). Drivers of Genetic Diversity in Keystone Forest Species: Contrasting Responses of Rare and Common Saproxylic Beetles to Habitat Location and Naturalness. *Molecular Ecology*, *35*(14), e70454. https://doi.org/10.1111/mec.70454
3. Banaś, J., Bujoczek, L., Zięba, S., & Drozd, M. (2014). The effects of different types of management, functions, and characteristics of stands in Polish forests on the amount of coarse woody debris. *European Journal of Forest Research*, *133*(6), 1095–1107. https://doi.org/10.1007/s10342-014-0825-3
4. Bartoń, K. (2010). *MuMIn: Multi-Model Inference* (p. 1.48.19) [Dataset]. https://doi.org/10.32614/CRAN.package.MuMIn
5. Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01
6. Bayona-Vásquez, N. J., Glenn, T. C., Kieran, T. J., Pierson, T. W., Hoffberg, S. L., Scott, P. A., Bentley, K. E., Finger, J. W., Louha, S., Troendle, N., Diaz-Jaimes, P., Mauricio, R., & Faircloth, B. C. (2019). Adapterama III: Quadruple-indexed, double/triple-enzyme RADseq libraries (2RAD/3RAD). *PeerJ*, *7*, e7724. https://doi.org/10.7717/peerj.7724
7. Blanchet, S., Prunier, J. G., & De Kort, H. (2017). Time to Go Bigger: Emerging Patterns in Macrogenetics. *Trends in Genetics*, *33*(9), 579–580. https://doi.org/10.1016/j.tig.2017.06.007
8. Blomberg, S. P., Garland, T., & Ives, A. R. (2003). TESTING FOR PHYLOGENETIC SIGNAL IN COMPARATIVE DATA: BEHAVIORAL TRAITS ARE MORE LABILE. *Evolution*, *57*(4), 717–745. https://doi.org/10.1111/j.0014-3820.2003.tb00285.x
9. Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J.-S. S. (2009). Generalized linear mixed models: A practical guide for ecology and evolution. *Trends in Ecology & Evolution*, *24*(3), 127–135. https://doi.org/10.1016/j.tree.2008.10.008
10. Brault, M., Brazier, T., Mackintosh, A., Paupe, A., Lascoux, M., & Glémin, S. (2026). A Recipe for a Good π. How to Properly Estimate Population Genetics Summary Statistics and Why we Should

Systematically Report Them. *Genome Biology and Evolution*, *18*(6), evag103. https://doi.org/10.1093/gbe/evag103

11. Bujoczek, L., Bujoczek, M., & Zięba, S. (2021). How much, why and where? Deadwood in forest ecosystems: The case of Poland. *Ecological Indicators*, *121*, 107027. https://doi.org/10.1016/j.ecolind.2020.107027
12. Burnham, K. P., & Anderson, D. R. (Eds.). (2004). *Model Selection and Multimodel Inference*. Springer New York. https://doi.org/10.1007/b97636
13. Cálix, M., Alexander, K. N. A., Nieto, A., Dodelin, B., Soldati, F., Telnov, D., Vazquez-Albalate, X., Aleksandrowicz, O., Audisio, P., Istrate, P., Jansson, N., Legakis, A., Liberto, A., Makris, C., Merkl, O., Mugerwa Pettersson, R., Schlaghamersky, J., Bologna, M. A., Brustel, H., … Purchart, L. (2018). *European Red List of Saproxylic Beetles*. https://portals.iucn.org/library/sites/library/files/documents/RL-4-023.pdf
14. Coates, D. J., Byrne, M., & Moritz, C. (2018). Genetic Diversity and Conservation Units: Dealing With the Species-Population Continuum in the Age of Genomics. *Frontiers in Ecology and Evolution*, *6*, 165. https://doi.org/10.3389/fevo.2018.00165
15. Dormann, C. F., McPherson, J. M., Araújo, M. B., Bivand, R., Bolliger, J., Carl, G., Davies, R. G., Hirzel, A., Jetz, W., Kissling, W. D., Kühn, I., Ohlemüller, R., Peres-Neto, P. R., Reineking, B., Schröder, B., Schurr, F. M., & Wilson, R. (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data: A review. *Ecography*, *30*(5), 609–628. https://doi.org/10.1111/j.2007.0906-7590.05171.x
16. Eckelt, A., Müller, J., Bense, U., Brustel, H., Bußler, H., Chittaro, Y., Cizek, L., Frei, A., Holzer, E., Kadej, M., Kahlen, M., Köhler, F., Möller, G., Mühle, H., Sanchez, A., Schaffrath, U., Schmidl, J., Smolis, A., Szallies, A., … Seibold, S. (2018). "Primeval forest relict beetles" of Central Europe: A set of 168 umbrella species for the protection of primeval forest remnants. *Journal of Insect Conservation*, *22*(1), 15–28. https://doi.org/10.1007/s10841-017-0028-6
17. Grafen, A. (1989). The phylogenetic regression. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *326*(1233), 119–157. https://doi.org/10.1098/rstb.1989.0106
18. Hagge, J., Müller, J., Birkemoe, T., Buse, J., Christensen, R. H. B., Gossner, M. M., Gruppe, A., Heibl, C., Jarzabek-Müller, A., Seibold, S., Siitonen, J., Soutinho, J. G., Sverdrup-Thygeson, A., Thorn, S., & Drag, L. (2021). What does a threatened saproxylic beetle look like? Modelling extinction risk using a new morphological trait database. *Journal of Animal Ecology*, *90*(8), 1934–1947. https://doi.org/10.1111/1365-2656.13512
19. Harvey, M. G., Smith, B. T., Glenn, T. C., Faircloth, B. C., & Brumfield, R. T. (2016). Sequence Capture versus Restriction Site Associated DNA Sequencing for Shallow Systematics. *Systematic Biology*, *65*(5), 910–924. https://doi.org/10.1093/sysbio/syw036
20. Jansen, S., Konrad, H., & Geburek, T. (2017). The extent of historic translocation of Norway spruce forest reproductive material in Europe. *Annals of Forest Science*, *74*(3), 56. https://doi.org/10.1007/s13595-017-0644-z

21. Kajtoch, Ł., Gronowska, M., Plewa, R., Kadej, M., Smolis, A., Jaworski, T., & Gutowski, J. M. (2022). A review of saproxylic beetle intra- and interspecific genetics: Current state of the knowledge and perspectives. *The European Zoological Journal*, *89*(1), 481–501. https://doi.org/10.1080/24750263.2022.2048717
22. Kostanjsek, F., Sebek, P., Baranova, B., Seric Jelaska, L., Riedl, V., & Cizek, L. (2018). Size matters! Habitat preferences of the wrinkled bark beetle, *Rhysodes sulcatus* , the relict species of European primeval forests. *Insect Conservation and Diversity*, *11*(6), 545–553. https://doi.org/10.1111/icad.12295
23. Krovi, R. S., Amer, N. R., Oczkowicz, M., & Kajtoch, Ł. (2025). Meta-analysis of spatial genetic patterns among European saproxylic beetles. *Biodiversity and Conservation*, *34*(1), 1–27. https://doi.org/10.1007/s10531-024-02940-8
24. Krovi, R. S., Amer, N. R., Wierzbicka, A., Plewa, R., Kadej, M., Jaworski, T., Smolis, A., Szmatoła, T., Oczkowicz, M., & Kajtoch, Ł. (2026). *Beyond deadwood volume: Structural heterogeneity and species-specific dispersal shape saproxylic beetle genetics*. Evolutionary Biology. https://doi.org/10.64898/2026.07.17.738652
25. Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13). https://doi.org/10.18637/jss.v082.i13
26. Lachat, T., Wermelinger, B., Gossner, M. M., Bussler, H., Isacsson, G., & Müller, J. (2012). Saproxylic beetles as indicator species for dead-wood amount and temperature in European beech forests. *Ecological Indicators*, *23*, 323–331. https://doi.org/10.1016/j.ecolind.2012.04.013
27. Lantschner, M. V., & Corley, J. C. (2023). Spatiotemporal outbreak dynamics of bark and wood-boring insects. *Current Opinion in Insect Science*, *55*, 101003. https://doi.org/10.1016/j.cois.2022.101003
28. Leigh, D. M., Van Rees, C. B., Millette, K. L., Breed, M. F., Schmidt, C., Bertola, L. D., Hand, B. K., Hunter, M. E., Jensen, E. L., Kershaw, F., Liggins, L., Luikart, G., Manel, S., Mergeay, J., Miller, J. M., Segelbacher, G., Hoban, S., & Paz-Vinas, I. (2021). Opportunities and challenges of macrogenetic studies. *Nature Reviews Genetics*, *22*(12), 791–807. https://doi.org/10.1038/s41576-021-00394-0
29. Lüdecke, D., Ben-Shachar, M., Patil, I., Waggoner, P., & Makowski, D. (2021). performance: An R Package for Assessment, Comparison and Testing of Statistical Models. *Journal of Open Source Software*, *6*(60), 3139. https://doi.org/10.21105/joss.03139
30. Luikart, G., England, P. R., Tallmon, D., Jordan, S., & Taberlet, P. (2003). The power and promise of population genomics: From genotyping to genome typing. *Nature Reviews Genetics*, *4*(12), 981–994. https://doi.org/10.1038/nrg1226
31. Lynch, M., & Conery, J. S. (2003). The origins of genome complexity. *Science*, *302*(5649), 1401–1404. https://doi.org/10.1126/science.1089370
32. McKenna, D. D., Sequeira, A. S., Marvaldi, A. E., & Farrell, B. D. (2009). Temporal lags and overlap in the diversification of weevils and flowering plants. *Proceedings of the National Academy of Sciences*, *106*(17), 7083–7088. https://doi.org/10.1073/pnas.0810618106

33. McKenna, D. D., Shin, S., Ahrens, D., Balke, M., Beza-Beza, C., Clarke, D. J., Donath, A., Escalona, H. E., Friedrich, F., Letsch, H., Liu, S., Maddison, D., Mayer, C., Misof, B., Murin, P. J., Niehuis, O., Peters, R. S., Podsiadlowski, L., Pohl, H., … Beutel, R. G. (2019). The evolution and genomic basis of beetle diversity. *Proceedings of the National Academy of Sciences*, *116*(49), 24729–24737. https://doi.org/10.1073/pnas.1909655116
34. Michonneau, F., Brown, J. W., & Winter, D. J. (2016). rotl: An R package to interact with the Open Tree of Life data. *Methods in Ecology and Evolution*, *7*(12), 1476–1481. https://doi.org/10.1111/2041-210X.12593
35. Müller, J., & Bütler, R. (2010). A review of habitat thresholds for dead wood: A baseline for management recommendations in European forests. *European Journal of Forest Research*, *129*(6), 981–992. https://doi.org/10.1007/s10342-010-0400-5
36. Mykhailenko, A., Zieliński, P., Bednarz, A., Schlyter, F., Andersson, M. N., Antunes, B., Borowski, Z., Krokene, P., Melin, M., Morales-García, J., Müller, J., Nowak, Z., Schebeck, M., Stauffer, C., Viiri, H., Zaborowska, J., Babik, W., & Nadachowska-Brzyska, K. (2023). *Complex genomic landscape of inversion polymorphism in Europe's most destructive forest pest*. https://doi.org/10.1101/2023.10.10.561670
37. Nei, M., & Li, W.-H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences of the United States of America*, *76*(10), 5269–5273. https://doi.org/10.1073/pnas.76.10.5269
38. Orme, D., Freckleton, R., Thomas, G., Petzoldt, T., Fritz, S., Isaac, N., & Pearse, W. (2011). *caper: Comparative Analyses of Phylogenetics and Evolution in R* (p. 1.0.4) [Dataset]. https://doi.org/10.32614/CRAN.package.caper
39. Pagel, M. (1999). Inferring the historical patterns of biological evolution. *Nature*, *401*(6756), 877–884. https://doi.org/10.1038/44766
40. Paradis, E., & Schliep, K. (2019). ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*(3), 526–528. https://doi.org/10.1093/bioinformatics/bty633
41. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., De Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795
42. R Core Team. (2023). *R: A Language and Environment for Statistical Computing* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/
43. Revell, L. J. (2012). phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, *3*(2), 217–223. https://doi.org/10.1111/j.2041-210X.2011.00169.x
44. Roberts, N., Fyfe, R. M., Woodbridge, J., Gaillard, M.-J., Davis, B. A. S., Kaplan, J. O., Marquer, L., Mazier, F., Nielsen, A. B., Sugita, S., Trondman, A.-K., & Leydet, M. (2018). Europe's lost forests: A pollen-based synthesis for the last 11,000 years. *Scientific Reports*, *8*(1), 716. https://doi.org/10.1038/s41598-017-18646-7

45. Rochette, N. C., Rivera-Colón, A. G., & Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*, *28*(21), 4737–4754. https://doi.org/10.1111/mec.15253
46. Schmidt, T. L., Jasper, M., Weeks, A. R., & Hoffmann, A. A. (2021). Unbiased population heterozygosity estimates from genome-wide sequence data. *Methods in Ecology and Evolution*, *12*(10), 1888–1898. https://doi.org/10.1111/2041-210X.13659
47. Seibold, S., Brandl, R., Buse, J., Hothorn, T., Schmidl, J., Thorn, S., & Müller, J. (2015). Association of extinction risk of saproxylic beetles with ecological degradation of forests in Europe. *Conservation Biology*, *29*(2), 382–390. https://doi.org/10.1111/cobi.12427
48. Stokland, J. N., Siitonen, J., & Jonsson, B. G. (2012). *Biodiversity in Dead Wood* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139025843
49. Ulyshen, M. D., & Šobotník, J. (2018). An Introduction to the Diversity, Ecology, and Conservation of Saproxylic Insects. In M. D. Ulyshen (Ed.), *Saproxylic Insects* (Vol. 1, pp. 1–47). Springer International Publishing. https://doi.org/10.1007/978-3-319-75937-1_1
50. Wende, B., Gossner, M. M., Grass, I., Arnstadt, T., Hofrichter, M., Floren, A., Linsenmair, K. E., Weisser, W. W., & Steffan-Dewenter, I. (2017). Trophic level, successional age and trait matching determine specialization of deadwood-based interaction networks of saproxylic beetles. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1854), 20170198. https://doi.org/10.1098/rspb.2017.0198
51. Yahara, T., Donoghue, M., Zardoya, R., Faith, D. P., & Cracraft, J. (2010). Genetic diversity assessments in the century of genome science. *Current Opinion in Environmental Sustainability*, *2*(1–2), 43–49. https://doi.org/10.1016/j.cosust.2010.03.008
52. Zhang, S.-Q., Che, L.-H., Li, Y., Dan Liang, Pang, H., Ślipiński, A., & Zhang, P. (2018). Evolutionary history of Coleoptera revealed by extensive sampling of genes and species. *Nature Communications*, *9*(1), 205. https://doi.org/10.1038/s41467-017-02644-4

Tables

Table 1. Principal species-level ecological-trait results.

Predictor	Response	Estimate	Raw P	BH P	Sensitivity	Conclusion
Host: deciduous vs coniferous	H (log)	β = 0.212; SE = 0.066	0.0041	0.036	β = 0.216; P = 0.0027 without *R. sulcatus*	Supported
Host: deciduous vs coniferous	H (log)	β = 0.286; SE = 0.109	0.015	0.103	Positive but not retained after correction	Suggestive
Host: deciduous vs coniferous	π (log)	β = −0.056; SE = 0.116	0.635	0.635	No relationship	Unsupported
Occurrence: common vs rare	π (log)	β = 0.267	0.034	0.172	Not retained after correction	Suggestive
Occurrence tier	H, H and F_{IS}	Directions inconsistent	> 0.05	> 0.05	No corrected association	Unsupported

Note: π was calculated across all callable positions. H and H were calculated among retained variant positions. BH correction was applied within response.

Table 2 Summary of phylogenetic-signal tests and lineage-removal sensitivity.

Response	Scale	Full λ	Full P	Without *R. sulcatus* λ	Sensitivity P	Interpretation
π	log	0.583	0.0099	0.580	0.0126	Stable signal
π	raw	0.472	0.032	0.468	0.037	Stable signal
F_{IS}	raw	0.938	0.001	0.483	0.047	Signal, less stable
H	log	0.657	0.0296	0.257	0.232	Lineage-sensitive
H	log	0.306	0.446	0.0001	1.000	Unsupported

Note: The synthetic Open Tree of Life phylogeny used Grafen branch lengths and is not time calibrated. Full Pagel's λ and Blomberg's K results are in Supporting Table S5.

Figures

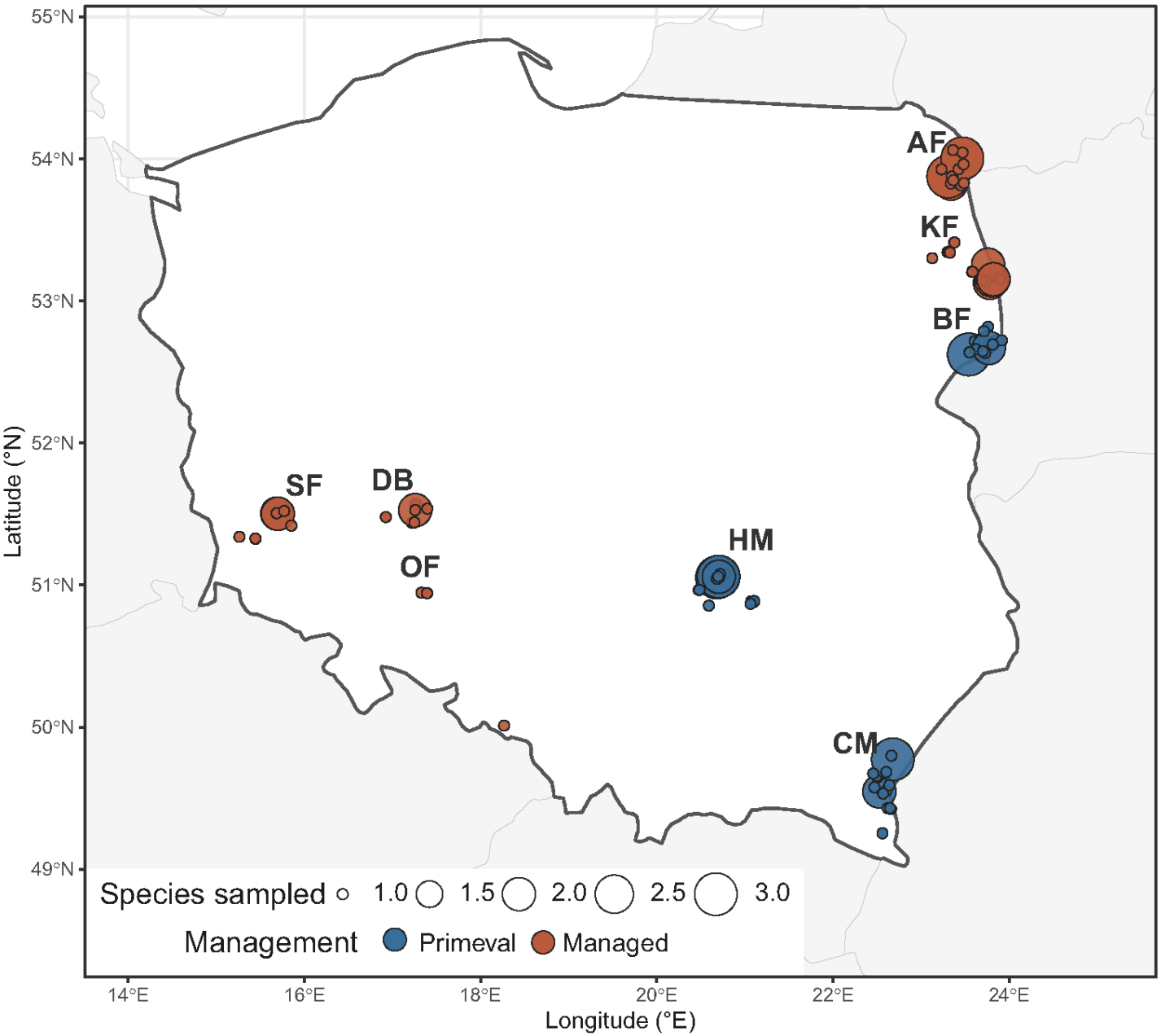

Figure 1

Distribution of sampled beetle populations across Poland. Points represent 105 distinct recorded coordinates among 149 species-population records. Point size indicates the number of population records sharing the same recorded coordinates, and colour indicates the management category of the forest complex.

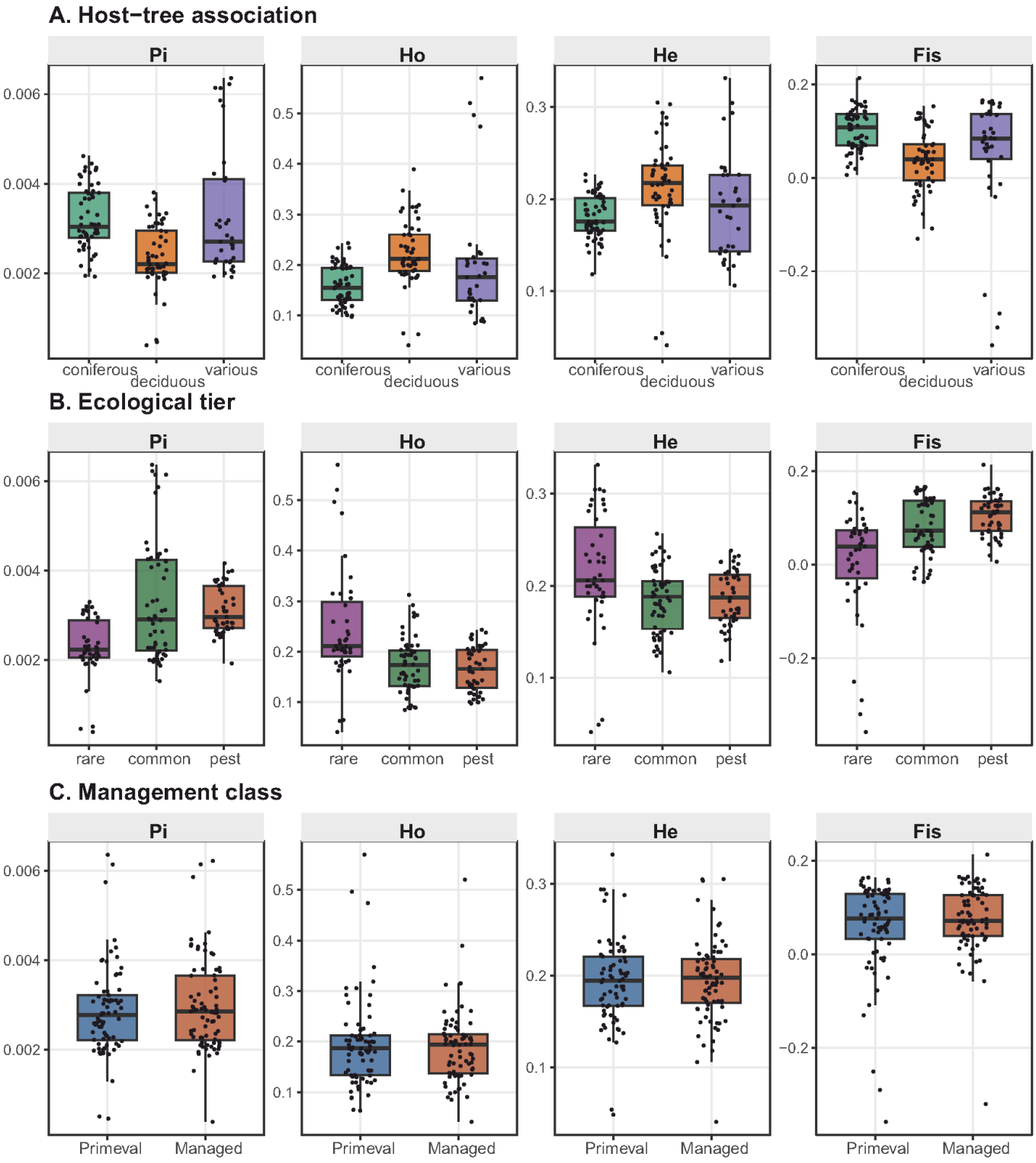

Figure 2

Genetic diversity parameters (π, H_o, H_e, F_{IS}) grouped by (A) host-tree preference, (B) occurrence tier, and (C) forest management regime. Box plots display medians and interquartile ranges, with individual species observations superimposed.

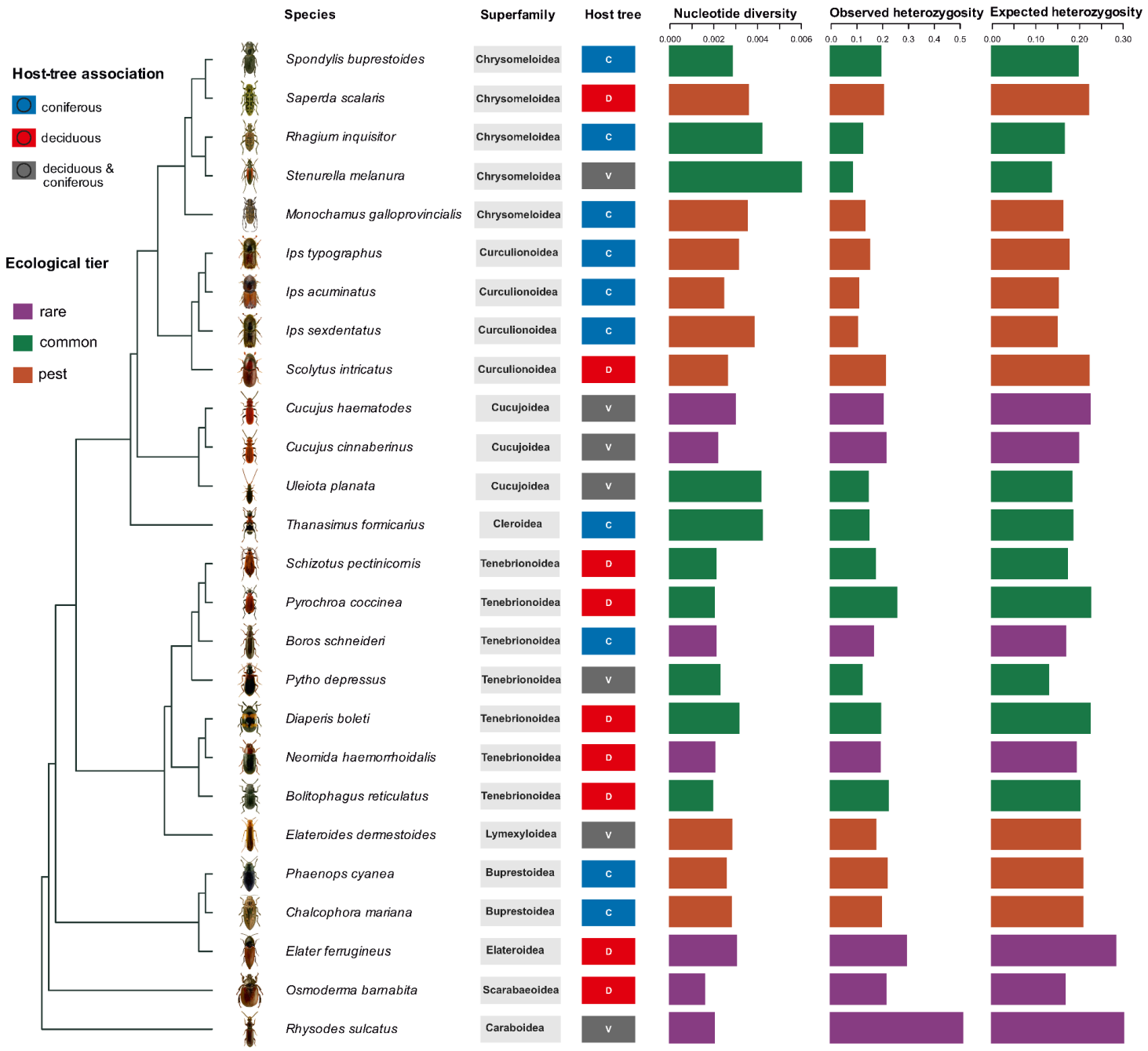

Figure 3

Phylogeny and species-mean genomic diversity for 26 saproxylic beetles. Bars show all-position nucleotide diversity (π) and variant-site observed and expected heterozygosity (H and H). Colours indicate occurrence tier.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupportingInformationHerTrait.docx