Quantifying phylogenetic and nonphylogenetic patterns in the richness, frequency, and identity of links in a herbivore-parasitoid interaction network

Frazer H. Sinclair[1,2], Chang-Ti Tang[1], Richard I. Bailey[3], György L. Csóka[4], George Melika[5], James A. Nicholls[6], José-Luis Nieves-Aldrey[7], Alex Reiss[1], Y. Miles Zhang[1,8], Albert B. Phillimore[1], Karsten Schönrogge,[9] & Graham N Stone[1]

1. Institute of Ecology and Evolution, University of Edinburgh, Charlotte Auerbach Road, Edinburgh, EH9 3FL, Scotland, United Kingdom;
2. Current address: Tropical Biology Association, David Attenborough Building, Pembroke Street, Cambridge, CB2 3QZ, United Kingdom.
3. Department of Ecology and Vertebrate Zoology, University of Lodz, Faculty of Biology and Environmental Protection, 12/16 Banacha St., 90-237 Łódź, Poland;
4. University of Sopron, Forest Research Institute, 3232 Mátrafüred, Hungary;
5. Plant Health Diagnostic National Reference Laboratory, National Food Chain Safety Office, Budaörsi u. 141-145, 1118 Budapest, Hungary;
6. CSIRO, Australian National Insect Collection, Clunies Ross Street, Acton, ACT 2601, Australia. Current address: Royal Botanic Garden Edinburgh, Inverleith Row, Edinburgh, United Kingdom;
7. Museo Nacional de Ciencias Naturales (CSIC), Departamento de Biodiversidad y Biología Evolutiva, C/ José Gutiérrez Abascal 2, ES-28006 Madrid, Spain;
8. Current address: Daniel K. Inouye U.S. Pacific Basin Agricultural Research Center, Hilo, Hawaii, USA.
9. UK Centre for Ecology & Hydrology, Maclean Building, Benson Lane, Wallingford, OX10 8BB, UK

Orcids and email:
Frazer Sinclair 0000-0001-5017-3215 (frazer_sinclair@yahoo.co.uk)

Chang-Ti Tang 0000-0002-5693-9645 (cynipidsman@gmail.com)

Richard I Bailey 0000-0001-9870-410X (richard.bailey@biol.uni.lodz.pl)

György Csóka 0000-0001-9132-4825 (csoka.gyorgy@uni-sopron.hu)

George Melika 0000–0002–5204–6890 (melikageorge@gmail.com)

James A Nicholls 0000-0002-9325-563X (james.nicholls@csiro.au)

José-Luis Nieves-Aldrey 0000-0002-4711-7455 (aldrey@mncn.csic.es)

Alexander Reiss 0000-0002-1230-8032 (alreiss993@gmail.com)

Yuanmeng M Zhang 0000-0003-4801-8624 (yuanmeng.zhang@gmail.com)

Albert B Phillimore 0000-0002-6553-1553 (albert.phillimore@ed.ac.uk)

Karsten Schönrogge 0000-0003-0122-6493 (ksc@ceh.ac.uk)

Graham N Stone. 0000-0002-2737-696X (graham.stone@ed.ac.uk)

Corresponding author: graham.stone@ed.ac.uk. Tel: 44 (0)131 6507194

Article type: E-article

Key words: Trophic interaction network, cynipid, parasitoid, phylogenetic structure, community assembly.

Short title/running head: Herbivore-parasitoid network structure

Elements: Main text with 5 figures and 1 table; Appendix with 1 table; Supplement.

Abstract

Uncovering the patterns and structure in species interactions is central to understanding community assembly and dynamics. Species interact via their phenotypes, but identifying and quantifying the traits that structure species-specific interactions (links) can be challenging. Where these traits show phylogenetic signal, link properties (such as which species interact, and how often) may be predictable using models that incorporate phylogenies in place of trait data. However, quantification of phylogenetic patterns in link properties is conceptually and methodologically challenging because it requires co-estimation of multiple phylogenetic and non-phylogenetic pattern types in interaction data for multiple sites, while controlling for confounding effects and making biologically plausible assumptions about which species can interact. Here we show how this can be done in a Bayesian mixed modelling framework, using data for trophic interactions between oak cynipid galls and parasitoid natural enemies. We find strong signatures of cophylogeny (i.e. related parasitoids attack related host galls) in both link incidence (presence/absence) and link frequency data, alongside patterns in link incidence/richness and identity across sites that are independent of either parasitoid or gall wasp phylogeny. Our results are robust to substantially reduced sample completeness, and are consistent with structuring of trophic interactions by a combination of phylogenetically conserved and phylogenetically labile traits in both trophic levels. We show that incorporation of phylogenetic relationships into analyses of species interactions has substantial explanatory power even in the absence of trait data, with potential applied use in prediction of natural enemies of invading pests and non-target hosts of biocontrol agents.

Introduction

Biological communities comprise sets of species that interact via processes along a continuum from antagonism (e.g. predation, parasitism, competition) to mutualism (e.g. pollination, seed dispersal, parasite removal) (Bascompte et al. 2006; Cagnolo et al. 2011; Peralta 2016; Caves 2021). Revealing why some species interact but not others is central to understanding community structure, assembly and dynamics, and remains a fundamental goal of ecology (Cattin et al. 2004; Singer and Stireman 2005; Cavender-Bares et al. 2009; Yeakel et al. 2012; Bramon Mora et al. 2020). This is particularly true for tritrophic communities of plants, insect herbivores and insect natural enemies that, between them, comprise more than 50% of terrestrial biodiversity (Smith et al. 2008; Novotny et al. 2010). At a community level, interactions can be summarised as networks of pairwise (bipartite) species links (Bascompte et al. 2006; Cagnolo et al. 2011). Species vary in three attributes of their link distribution: richness (how many species each is linked to), identity (which species they are linked to), and frequency (how often a link is realised, a measure of interaction strength) (Memmott et al. 1994; Yeakel et al. 2012; Maia et al. 2019; Braga et al. 2020; Heimpel et al. 2021). Species interact via phenotypes, and links are mediated by phenotypic traits (Ehrlich and Raven 1964; Thompson 2005; Blasco-Costa et al. 2021). Well-studied examples include pollinator tongue length and floral structure in pollination mutualisms (Darwin 1862; Whittall and Hodges 2007; Anderson and Johnson 2008), and plant chemical defences and herbivore countermeasures in herbivory (Ehrlich and Raven 1964; Janz 2011). Identifying such traits is fundamental to studying the impacts of natural selection and coevolution in community assembly (Thompson 2005; Rezende et al. 2007a; Pearse and Hipp 2009; Janz 2011; Fontaine and Thébault 2015; Endara et al. 2018), and allows modelling of network structures based on traits rather than species (Blasco-Costa et al. 2021). This enhances mechanistic understanding of

structuring processes, and should improve predictive power (McGill et al. 2006; Truitt et al. 2019; Marjakangas et al. 2022; Pinilla-Gallego et al. 2022).

In all but a few systems, however, the key traits that structure species interactions are either unknown or quantified only for a subset of interacting species (Belshaw et al. 2003; Pakeman 2014; Penone et al. 2014; Gripenberg et al. 2019). If phylogeny is a valid proxy for functional trait variation then species' link properties may be predictable from phylogenetic relationships in each trophic level (Ives and Godfray 2006; Rezende et al. 2007b; Ives and Helmus 2011; Peralta 2016; Poisot and Stouffer 2018; Gripenberg et al. 2019; Gallinat and Pearse 2021; Perez-Lamarque et al. 2022; Benadi et al. 2022). Where this is true, more closely related predators, for example, will be more similar in traits that structure their links with herbivores, and *vice versa* (Ives and Godfray 2006; Losos 2008; Poulin et al. 2011; Stouffer et al. 2012; Naisbit et al. 2012; Ives 2022). While our study is framed in terms of interactions between natural enemies and host defensive phenotypes, the same rationale applies to other bipartite interactions, such as pollen transfer links in flower visitation networks (Ballantyne et al. 2015) and competition between species (Carvalheiro et al. 2014; Lemos-Costa et al. 2023).

Where phylogenetic effects are strong, statistical frameworks incorporating them may be valuable in predicting the identity and strength of links for unsampled species (Ives and Godfray 2006; Pearse and Altermatt 2013, 2015). Potential applications include predicting native natural enemies of invading plants or animal pests, predicting potential non-target victims of imported biocontrol agents (Pearse and Altermatt 2013, 2015; Davies 2021; Heimpel et al. 2021), and predicting (co-)extinction risk (Rezende et al. 2007b). A major attraction of phylogeny-based prediction is that addition of a missing species to an existing model only requires adding it to a

phylogeny. For a molecular phylogeny, this only requires a single DNA sample from a single example of any life stage. This is often possible from archival material, and is substantially less labour intensive than measurement of trait values, which requires identification of structuring traits followed by measurements from multiple living examples of a specific life stage (Albert et al. 2011; Wong and Carmona 2021).

In addition to the predictive value of statistical patterns (Pearse and Altermatt 2013, 2015, Ives 2022), there is considerable interest in the extent to which phylogenetic and other patterns in link properties can be used to infer underlying structuring processes (Forister and Feldman 2010; Althoff et al. 2014; Russo et al. 2018; Harmon et al. 2019; Hembry and Weber 2020; Blasco-Costa et al. 2021; Dismukes et al. 2022; Perez-Lamarque and Morlon 2024). For example, in antagonistic bipartite networks (plant-herbivore, herbivore-enemy) a strong co-phylogenetic pattern – in which related species in one trophic level are often linked to related species in another trophic level (Fig 1.D), can indicate phylogenetic conservatism in structuring traits. This pattern is (to varying extents) associated with escape-and-radiate coevolution, and with herbivore tracking of pre-existing plant traits through preferential host switching (Janz 2011; Perez-Lamarque and Morlon 2024). Additional information on the relative timing and spatial distribution of events in each phylogeny is required to discriminate between these and alternative scenarios, such as shared vicariance events (Althoff et al. 2014; Perez-Lamarque and Morlon 2024). In contrast, demonstration that links between species are highly non-random but not associated with phylogenetic relationships in one or both trophic levels is compatible with community assembly by processes that reduce phylogenetic signal in structuring trait distributions, such as convergent or divergent evolution (Janz 2011; Endara et al. 2017, 2018; Ward et al. 2024), bounded evolution (Boucher and Démerey 2016), or evolution towards an attractor (e.g., an Ornstein–Uhlenbeck

process) (Boucher et al. 2018). Associating specific phylogenetic and non-phylogenetic patterns in link properties with specific underlying processes (or combinations of processes) nevertheless remains inherently difficult (Losos 2011; Hembry and Weber 2020).

Whatever their application, quantification of alternative patterns in link properties (fig. 1) is a conceptual and methodological challenge (Pearse and Altermatt 2013, Pearse and Altermatt 2015; Russo et al. 2018; Harmon et al. 2019; Dismukes et al. 2022; Ives 2022). Most methods and empirical analyses have focussed on estimating co-phylogenetic effects (fig. 1D) without incorporating potential contributions of multiple alternative phylogenetic and non-phylogenetic patterns (fig. 1) (Legendre et al. 2002; Hommola et al. 2009; Leppänen et al. 2013; Eklöf and Stouffer 2016; Poisot and Stouffer 2018; Russo et al. 2018; Braga et al. 2020; Cruz-Laufer et al. 2022). However, accurate estimation of the contribution of any single pattern (and hence any prediction or pattern-based inference of associated processes) requires a modelling framework that simultaneously estimates terms capturing alternative link richness and identity patterns, while controlling for potential confounding effects (Rafferty and Ives 2013; Hadfield et al. 2014; Gallinat and Pearse 2021).

Several linear mixed modelling (LMM) approaches and their generalised extensions (GLMM) – here collectively termed two phylogeny mixed models (2PhyGLMMs) – have been developed that allow such co-estimation, and have been applied to links recorded in terms of incidence (a binary response, whether two species interact or not; Hadfield et al. 2014; Endara et al. 2018) and counts (how frequently two species interact; Rafferty and Ives 2013; Hadfield et al. 2014). Advantages of generalised linear model-based approaches include ability to incorporate alternative error structures associated with different data types, and additional variables

representing species traits or spatial structure in the data (Rafferty and Ives 2013; Hadfield et al. 2014). The individual bipartite interactions that contribute signal to each model term (and hence to each type of link pattern) can also be identified (Hadfield et al. 2014). To date, however, very few interaction networks have been analysed using 2PhyGLMMs (for examples, see Rafferty and Ives 2013; Hadfield et al. 2014; Endara et al. 2018; Galen et al. 2019; Lajoie and Kembel 2021). It also remains unclear how sample sizes (i.e. number of trophic link records for each species at each trophic level) and data completeness affect statistical power (Bersier et al. 2002; Olesen et al. 2011; Rivera-Hutinel et al. 2012; Hadfield et al. 2014; De Aguiar et al. 2019).

One of the most striking plant-herbivore interactions involves galling, which is a parasitic or mutualistic life history strategy in which non-plant organisms manipulate plant development to create novel gall structures (Giron et al. 2016; Harris and Pizschke 2020). An estimated 21,000-211,000 insect species induce galls, and this life history has evolved multiple times in multiple insect orders (Espírito-Santo & Fernandes, 2007). Insect galls commonly harbour other (non-inducer) insects, which exploit gall tissues (gallivores, herbivorous inquilines), other insects (predators, parasitoids), or both (Raman et al. 2005). Because these non-inducer assemblages are often specialists in particular gall types, galls represent convenient natural microcosms for the study of traits and processes structuring species interactions. Gall phenotypes are under inducer control, and a question attracting attention across many galling systems is the extent to which patterns in galler-controlled traits structure associated animal communities (e.g. Farache et al. 2018; Oliveira et al. 2020; De Araújo and Maia 2021). Candidate structuring traits in many galling systems are poorly known, and an informative step in their identification is to quantify the relative strengths of (co-)phylogenetic and non-phylogenetic patterns in links between gall-inhabiting guilds.

Here we use 2PhyGLMMs to model link properties in an antagonistic bipartite network comprising galls induced by oak gall wasps (Hymenoptera: Cynipidae; Cynipini) and chalcid parasitoids (Hymenoptera: Chalcidoidea) (fig. 2). Oak gall wasps are herbivores that develop inside galls that they induce on oaks and related Fagaceae. Most have a life cycle with strict alternation between a single spring (sexual) and a single summer/autumn (asexual) generation, each of which develops in a gall with a specific structural phenotype on a specific organ (acorn, bud, catkin, leaf, root) on specific host plant taxa (fig.2). The two generations support species-rich assemblages of chalcid parasitoids, which in Europe belong to the families Eulophidae, Eupelmidae, Eurytomidae, Megastigmidae, Ormyridae, Pteromalidae and Torymidae (Bailey et al. 2009; Askew et al. 2013). The parasitoids attacking the two gall generations of a single host species are typically very different, while those attacking the same generation of related species are often similar (Askew et al. 2013). This suggests that although the two generations of the same gall wasp necessarily share the same evolutionary history, their associated parasitoid communities have assembled largely independently (Askew 1961a; Bailey et al. 2009). This feature of gall wasp biology means that differences in patterns between gall wasp generations cannot be caused by shared phylogenetic and biogeographical history, and hence may be attributable to contrasting patterns in the evolution of structuring traits between generations (Moran 1994; Hood and Ott 2017).

The gall wasp system has several attributes that make it well suited for analysis of patterns in network structure. First, the parasitoids that attack oak cynipid galls are almost all specialists in this system (Askew et al. 2013), allowing it to be considered in ecological isolation (Bailey et al. 2009). Second, both trophic levels show wide variation in link properties, including richness (also termed the degree, specialisation or generality of a parasitoid species and the vulnerability of a

host; Schoener 1989; Bersier et al. 2002), identity (also termed host range or host repertoire; Braga et al. 2020; Heimpel et al. 2021) and frequency (Askew et al. 2013). Both attributes are shared with parasitoid communities centred on other herbivore guilds (Askew 1980; Stireman and Singer 2003; Leppänen et al. 2013; Santos et al. 2022). Third, there are good reasons to predict both phylogenetic and non-phylogenetic patterns in link properties in this system. Previous research on oak gall communities has identified traits that structure trophic links by influencing the ability of parasitoids to exploit specific host gall types (Bailey et al. 2009). These include parasitoid ovipositor toughness and length (fig. 2I-K), which determine the depth to which a female parasitoid can exploit hosts in target galls (Askew 1965; Quicke et al. 1998; Egan and Ott 2007), and a suite of gall wasp traits that, for parasitoids, influence their availability (phenology, host plant association, gall location on the host plant), accessibility (gall structural defences), and resource quality (host size) (Bailey et al. 2009). For both trophic levels, some link-structuring traits are phylogenetically conserved (e.g. gall wasp-oak associations, some gall defences, ovipositor length in some parasitoid lineages), while others appear multiple times independently across the phylogeny (e.g. gall location on the oak, some gall defences, ovipositor length across parasitoid lineages) (Quicke and Belshaw 1999; Cook et al. 2002; Stone and Schönrogge 2003; Stone et al. 2009; Nicholls et al. 2018; Elias et al. 2018).

Our aim in this study is to assess the extent to which gall-parasitoid links show phylogenetic and non-phylogenetic patterns in the gall wasps and parasitoids, and hence are structured by phylogenetically conserved versus non phylogenetically-conserved species traits in these guilds. We address the following questions: (1). Do we see a strong signature of cophylogeny, consistent with simultaneous codiversification or sequential radiation (Nyman et al. 2007). (2). Does the strength of phylogenetic patterns differ for gall wasps and parasitoids? Peralta

(2016) suggests that we might expect stronger phylogenetic signal in hosts than in parasitoids because traits determining enemy ability to exploit hosts/prey evolve faster than traits determining host/prey vulnerability, which should thus have a stronger phylogenetic signal in network structure (Rossberg et al. 2006; Fontaine and Thébault 2015). We also might expect stronger phylogenetic patterns for galls than for parasitoids because, while the inducing cynipids are a single radiating lineage, the parasitoids attacking each gall comprise representatives of several chalcidoid families that diverged before the origin of Cynipini (see Discussion). (3). Are patterns in link identity stronger than patterns in link richness or frequency? We might also expect weaker patterns in link richness or frequency because these are strongly influenced by site-specific distribution and population dynamic effects on a range of temporal and spatial scales in gall wasp communities (Washburn and Cornell 1981; Bunnefeld et al. 2018) and other systems (Berlow et al. 2004). (4) Do we see similar patterns in assemblages associated with the two gall wasp generations? While similar patterns could in principle be explained by processes (such as vicariance) that are necessarily shared by both generations, substantial differences suggest shaping by contrasting evolutionary processes (Moran 2004; Hood and Ott 2017).

We identify the links contributing to significant patterns, and use the flexibility inherent in a 2PhyGLMM modelling framework to explore the robustness of our inference and predictive power to key aspects of network sampling: use of incidence versus frequency data; variation in sampling completeness (Goldwasser and Roughgarden 1997; Jordano 2016), sampling effort (Gotelli and Colwell 2011; Chao et al. 2020) and spatial scale (Thompson and Townsend 2005; Brimacombe et al. 2023). We also explore alternative approaches to identifying sets of species potentially able to form links (Hadfield et al. 2014).

Materials and Methods

Gall Wasp–Parasitoid Community Networks

Our analysis uses 27,445 records of 54 parasitoid species reared from 38,638 galls of 78 generations of 60 cynipid species, sampled from four oak species (*Quercus cerris* L., *Q. pubescens* Willd., *Q. robur* L., *Q. petraea* (Matt.) . Liebl.) at six sites in Hungary (for locations see Supplement figure S1) between 2000–2003. The datasets include 26,679 parasitoid records for 53 parasitoid species reared from 26,847 galls of 45 generations of 36 gallwasp species published in Bailey et al. (2009) for five sites, with additional data for further gall wasp species and generations from these sites and for a sixth site (Köszeg) sampled in 2000. At each site, we collected galls from >100 individual trees of available oak species at fortnightly intervals between April and October, with each gall type harvested as far as possible across the full site area. Galls were collected haphazardly with respect to height and aspect from branches sampled to a height of 8m with a long-handled pruner. Galls were always harvested in their first year of development, across a range of dates to allow full development of gall inhabitants but prior to their emergence. Resulting sample sizes varied across sites and gall types. All galls were identified to generation and species, reared individually, and monitored for emerging insects for 2.5 years. Images of gall phenotypes are available at https://doi.org/10.1371/journal.pbio.1000179.s002 and in Roskam (2019). The specific host(s) of the parasitoids within each galls were not determined – they could have attacked the gall inducer, herbivorous cynipid inquilines or (much less frequently) each other (Askew 1961a; Askew et al., 2013); our data therefore represent association networks (links between parasitoids and host galls) rather than a true trophic network (links between parasitoids and specific insect hosts). Because the associations in our data are based primarily on trophic

interactions between parasitoids and herbivores, we refer to trophic levels as a convenient shorthand, and consider implications of this aspect of our data in the Discussion.

All sampled parasitoids are in the superfamily Chalcidoidea and were identified to species by expert taxonomists using morphological keys (Bailey et al. 2009). Full lists of sampled gall wasp and parasitoid taxa and an additional comment on taxonomic resolution are provided in Supplement section S1 and tables S1 and S2. As in Bailey et al. (2009), we analysed data for sexual and asexual generations separately. Gall-parasitoid interaction matrices are shown for each generation in fig. 3. Summaries of sample sizes, link richness and link frequency for all gall generations and parasitoid taxa are provided in Supplement tables S1 and S2, and the link datasets are available online (Sinclair et al. 2025). The sexual generation dataset comprises 255 distinct link types identified from rearing 18210 adult parasitoids of 46 species from 11791 galls of 26 cynipid species, while the asexual generation dataset comprises 439 link types identified from rearing 9235 adult parasitoids of 43 species from 26847 galls of 52 cynipid species.

For both generation datasets, we analysed link properties using incidence and frequency data. Incidence data incorporate no information on interaction strength (Berlow et al. 2004) and thus cannot discriminate between alternative scenarios in which a parasitoid species attacks two alternative hosts at ratios of 99:1 and 50:50. In contrast, frequency data allow detection of cross-species variation in interaction strength. When designing sampling experiments, a valid question is whether additional inferences are possible from count data that justify the additional sampling effort required. The same phylogenetic and non-phylogenetic effects can be estimated in incidence and frequency models, though the interpretation of terms differs between them. For example, for incidence data the parasitoid phylogeny effect (fig. 1B) increases as related parasitoids attack an increasingly similar richness of host species, while for count data this occurs as related parasitoids

inflict increasingly similar average attack rates on hosts. Although in this case the incidence-based interpretation is more straightforward, by not discriminating between weak and strong links (Berlow et al. 2004) the presence/absence approach is less informative about patterns of enemy richness.

Gall Wasp and Parasitoid Phylogenies

We generated gall wasp and parasitoid phylogenies using partial sequences of one mitochondrial coding locus (cytochrome b for gall wasps and cytochrome c subunit 1 for parasitoids) and one nuclear non-coding locus (28S D2 for both trophic groups), totalling 999 and 1304 base pairs for gall wasps and parasitoids respectively. Full molecular and phylogenetic methods are provided in Supplement section S2, tables S2-S5, and Genbank accession numbers of all sequences are provided in Supplement table S7. To incorporate phylogenetic information into a GLMM, all root-tip distances are scaled to 1 and the inverse of the phylogenetic covariance matrix is generated for the gall wasp and parasitoid ultrametric trees (Hadfield 2010; Hadfield and Nakagawa 2010). To allow us to place observed phylogenetic patterns in a temporal context, we inferred node ages for both trophic levels (see Supplement section S2). The phylogenies for each generation dataset used in MCMCglmm analyses are shown in simplified form in fig. 3, and are provided for all taxa in each trophic level and with full node age and branch support information in Supplement figures S2 and S3.

Two Phylogeny Mixed Models

In 2014 Hadfield *et al.* developed and applied a GLMM approach to analysis of incidence data for links between fleas and their mammalian hosts. Independently, Rafferty and Ives (2013) developed and applied an equivalent LMM to link frequency (count) data for insect visitors to flowering plants. Both approaches combine bipartite community interaction data with a phylogeny for each trophic level, and allow simultaneous estimation of variance components for link richness and link identity model terms, all of which are fitted as random effects. Frequentist forms (e.g. penalised quasi-likelihood) of 2PhyGLMMs can be fitted in software such as ASReml (Butler et al. 2017) or the phyr R package (Li et al. 2020). However, while these are several thousand times faster than equivalent Bayesian approaches such as MCMCglmm, they behave poorly for Bernoulli data (as for incidence networks) where the grouping factors include phylogenetic or pedigree data, with likely downward bias in estimation of variance components (Gilmour et al. 1985; Hadfield et al. 2014). We therefore used the Bayesian MCMC framework incorporated in the package MCMCglmm (v2.32) (Hadfield 2010) in R (v4.1.0) (R Core Team 2021). R code for our analyses (including intraclass correlation calculations and figure plotting) is available online (Sinclair et al. 2025). We summarise the terms fitted in our MCMCglmm models below, largely following Hadfield et al. (2014), and give their biological interpretation in the context of host-parasitoid systems. Selected terms are illustrated diagrammatically for an incidence model in fig. 1 and all terms are described in full for both incidence and frequency models in Appendix A. The patterns expected in a host-parasitoid link matrix when each term is substantial are illustrated in Hadfield et al. (2014).

Patterns in Link Richness

Patterns in link richness (species degree, for incidence data) and average link frequency (interaction strength, for count data) are captured by four main (i.e., non-interaction) random effects. These terms capture variance partitions for each of gall hosts and parasitoids with and without phylogenetic patterning. Explained variance in richness/frequency across parasitoid and host species that lacks any phylogenetic pattern is allocated to a Parasitoid species effect (fig. 1A) and a Host species effect, respectively. The equivalent phylogenetically patterned effects are the Parasitoid phylogeny effect (fig. 1B) and the Host phylogeny effect, which in incidence models capture, respectively, the extent to which related sets of parasitoids attack a similar richness of hosts, and sets of related hosts are attacked by a similar richness of parasitoids. In frequency models, these effects capture among-species variation in average link frequency (average interaction strength, i.e., the extent to which parasitoids are rare to abundant across host species, and the extent to which hosts are rarely to heavily attacked across parasitoid species). Incorporation of datasets for multiple discrete samples (sites in our analysis) allows estimation of additional model terms. A Site main effect captures among-site variation in the proportion of realised links (equivalent to unweighted connectance) in an incidence model, and in average link frequency in a frequency model. The Site:parasitoid and Site:host interaction effects capture among-site variation in the richness of parasitoid and host species in incidence models, and in average link frequency in frequency models.

Patterns in Link Identity

Patterns in link identity (i.e. the identity of the species forming links) are captured by four random effect interaction terms between hosts and parasitoids. The Parasitoid:host species interaction

captures the extent to which specific sets of parasitoids attack (incidence model) – or have similar link frequency with (frequency model) – specific sets of hosts across samples, regardless of phylogenetic relationships in either trophic level (fig. 1E). This is equivalent to the species interaction of Hadfield et al. (2014), renamed here to underline the involvement of both trophic levels, and can only be directly estimated in models that utilise information from patterns in species interactions across samples (here, across sites). Three additional terms capture the extent to which variation in link identity is predicted by one or both phylogenies. For the incidence model the Parasitoid phylogenetic interaction captures the extent to which related parasitoids attack similar sets of unrelated hosts (fig. 1C), the Host phylogenetic interaction captures the extent to which related hosts are attacked by similar sets of unrelated enemies, and the Cophylogenetic interaction captures the extent to which related taxa in one trophic level link with sets of related taxa in the other (fig. 1D). For the frequency model the same phylogenetic interactions are instead framed in terms of average link frequency (interaction strength). These three types of phylogenetic interactions are equivalent to the parasite evolutionary interaction, host evolutionary interaction and coevolutionary interaction of Hadfield et al. (2014). We prefer our terminology because we are referring only to patterns in data, while use of the term coevolutionary implies additional demonstration of reciprocal adaptation (Althoff et al. 2014; Poisot and Stouffer 2018), which is not what we are testing and would require experimental validation. The Cophylogenetic interaction captures cophylogenetic signal, as defined by Perez-Lamarque and Morlon (2024).

Model Fitting

Our study focuses on four core analyses: of incidence and frequency data, in both sexual and asexual generation bipartite networks. Our modelling of the above terms followed Hadfield et al.

(2014), as detailed in their equation 3 for phylogenetic terms (i.e. Parasitoid phylogeny effect, Host phylogeny effect, Parasitoid phylogenetic interaction, Host phylogenetic interaction and Cophylogenetic interaction), and their equation 5 for non-phylogenetic equivalents (i.e. Parasitoid species effect, Host species effect, and Parasitoid:host species interaction). Modelling of sample sites followed treatment of regions by Hadfield et al. (2014), with Site included as a random effect (as exemplified in their equation 7). We included two additional random effect terms not mentioned in Hadfield et al. (2014) – the Site:parasitoid interaction and Site:host interaction – to account for variation in the richness (incidence models) and mean link frequency (frequency models) of hosts and parasitoids among sites. For ease of reference, interpretations of all fitted terms for incidence and count-based models are summarised in Appendix A.

We modelled the sets of species available to form links at each site in each of two ways. In option 1, we assumed that the species recorded at each site in each gall generation represented all species available to interact locally and that the absence of a species is uninformative about species traits and biological processes governing community assembly (for instance absence may be due to biogeographic processes). Host and parasitoid links that did not occur because the two species were not present in the same site and generation were identified as structural zeros and removed from the dataset (Hadfield et al. 2014). In option 2, we assumed that the full set of parasitoid species recorded in a given generation across all six sites was available to interact at each site, such that the absence of an interaction is informative about the focal processes underpinning community assembly. In this version of the dataset all links absent from a single site and generation were given a value of zero (i.e. structural zeros were included). Our rationale is that the ecological reality is somewhere between these two models, such that by fitting them both we can assess the sensitivity of our inference to which one is true. In practice, the two options produced very similar results for

three of the four data sets (Supplement tables S8, S9), though option 2 models took much longer to run. We therefore report results using option 1 and highlight differences between options where these arise. For each dataset, we also fitted a model in which the six site datasets were pooled into a single regional dataset. This requires option 2, and necessarily prevents fitting of model terms dependent on between-site variation. Comparison of these models with their full option 2 equivalents allows assessment of the sensitivity of other model terms to whether site level variation in link properties is accounted for.

We used a binomial model with logit link function for incidence data, and a Poisson model with log-link function for frequency (count) data. For the incidence models we followed Hadfield et al. (2014) and fitted the logarithms of site-specific sample sizes for both hosts and parasitoids as fixed effects to account for the expectation that larger sample sizes for a species would lead to more of its links being observed. These terms were not included for the count-based models, which focus on the relative frequencies of different links rather than their presence-absence (Hadfield et al. 2014).

Parameter expanded priors were used for random effects in all models, with numerator and denominator degrees of freedom set to 1 and a scale parameter of 10^3. This technique is known to improve MCMC chain mixing in situations with complex random effect structures by placing a non-central scaled F-distribution prior on the random effect variance parameters (Hadfield 2010). For incidence based models a residual parameter (see also description below) was fixed at 1, and for frequency-based Poisson models it was estimated from the data with a prior following an inverse Wishart distribution with nu = 0.002. Chains were run for 5 million iterations with a burn-in of 1 million and a thinning interval of 2000, resulting in 2000 sets of parameter estimates.

We assessed the relative importance and statistical support for model terms using intraclass correlation coefficients (ICCs) on the latent scale (Nakagawa and Schielzeth 2010). ICCs are statistical measures that quantify the proportion of total variance in a dataset attributable to differences among groups or classes. For a given model term (or set of terms), the ICC represents the proportion of variance that it explains, and is calculated as its variance parameter (or the sum of a set of variance parameters) divided by a denominator comprising the total variance—i.e., the sum of all terms variance parameters and the residual variance. We report the median and mode for ICC posterior distributions of model terms, based on 2000 posterior samples for each model, and interpret a term as significant where the lower bound of the 95% credible interval is removed from zero (defined here as > 0.01, as in Hadfield et al. (2014)).

Following Nakagawa & Schielzeth (2010), the residual variance in both types of model is the sum of a residual parameter and a distribution specific variance. For binomial models, the residual parameter was fixed at 1, and the distribution specific variance was $\pi^2/3$. This fixing approach is recommended by Hadfield (2017) to anchor the latent scale and aid the identifiability of other model parameters. For Poisson models the residual parameter was estimated from the data as an observation level random effect with non-parameter expanded priors (see previous but one paragraph), and the distribution specific variance was ln(1/exp(intercept))+1 to reflect the relationship between overall mean (i.e. intercept) and variance on the latent scale. Although binomial models included fixed effects to control for sample sizes, the ICCs were 'adjusted' in that the fixed effect variances were not included in the ICC denominator (Nakagawa et al. 2017).

To explore how different gall and parasitoid species contributed to each term's ICC we calculated the posterior modes of the predicted MCMCglmm solutions either for individual species (as appropriate for the Host species, Parasitoid species, Host phylogeny and Parasitoid phylogeny

terms), or pairwise links (as appropriate for the Parasitoid:host species interaction, Host phylogenetic interaction, Parasitoid phylogenetic interaction, and Cophylogenetic interaction terms). These solutions, also sometimes called true effect sizes, are similar to the best linear unbiased predictors (BLUPs) from frequentist mixed models (Sorensen 2009; e Silva et al. 2013).

We also used ICCs to summarise the overall explanatory power of our models (ICC_{GLMM}), the magnitude of the three phylogenetic effects in link identity combined (ICC_{PHY}), and the relative magnitude of phylogenetic versus non-phylogenetic link identity effects ($ICC_{REL\text{-}PHY}$). For ICC_{GLMM}, following Nakagawa et al. (2017) the numerator for each of the 2000 posterior samples was the sum of variance parameters for all the fitted random effect terms and the denominator was the numerator plus a residual. For ICC_{PHY} the numerator was the sum for each set of posterior samples of the relevant phylogenetic link identity model terms (i.e., Parasitoid phylogenetic interaction, Host phylogenetic interaction and Cophylogenetic interaction) and the denominator was the sum of all the fitted random effect terms plus a residual. For $ICC_{REL\text{-}PHY}$ the numerator was the sum of the relevant phylogenetic link identity model terms and the denominator was the numerator plus the non-phylogenetic link identity term (i.e. the Parasitoid:host species interaction).

Sensitivity of Inference to Sample size

To assess the extent to which incorporation of sample size information influences the magnitude of model random effects, we compared the results of our full incidence models with alternatives without site-specific sample sizes as a covariate. This matters because in some systems, non-phylogenetic and/or phylogenetically-conserved variation in species abundance is thought to have causal impacts on associated link richness (Vazquez et al. 2005). It is thus unclear to what extent variation in gall sample sizes represents variation in sampling effort versus variation in a

biologically relevant host trait. Where the latter is true, controlling for sampling effort using link frequency could reduce support for MCMCglmm random effects capturing patterns in link richness (Hadfield et al. 2014).

Subsampling and Sample Completeness

All empirical trophic link datasets are likely to suffer from incomplete sampling, in that additional sampling could result in changed estimates of link incidence and/or frequency (Chao et al. 2020). Most common network metrics are sensitive to sampling completeness, though those incorporating interaction strength are usually less affected than those based on incidence (Nielsen and Bascompte 2007; Chacoff et al. 2012; Rivera-Hutinel et al. 2012; Vizentin-Bugoni et al. 2016; Falcão et al. 2016; Henriksen et al. 2019). A key question is then the extent to which results and inferences are sensitive to variation in sampling effort. We hypothesise that two phylogeny mixed models should show reduced ability to detect patterns in under-sampled networks (i.e. higher type II error) due to reduced host and parasitoid richness in incidence models, and that, as for network metrics, the inferences from count-based models should be less sensitive to reduced sampling than their incidence equivalents.

To quantify the impact of sample completeness on our inference, we generated datasets comprising 5%, 10%, 25%, 50%, and 75% subsets of the total number of sexual and asexual generation galls, based on a random draw process without replacement. These correspond to sample sizes of 590, 1180, 2948, 5895, and 8843 galls for the sexual generation datasets and 1342, 2685, 6712, 13424, and 20135 for the asexual generation datasets. To retain the methods and scope of the study but simulate a reduction in collecting effort, random draws treated each individual gall

as a distinct sample without regard to species, site, or year of collection. At each subsampling level we generated 60 subsampling replicates that were used to fit 30 replicate incidence models and 30 replicate frequency models for both the sexual generation and asexual generation gall datasets, applying the same model structures used for the full models. As previously, the only zeros included in the resulting dataset were those for host-parasitoid species pairs that were both present in the same site and generation in the sub-sample. To assess the completeness of our sampling and to summarise the impacts of reduced sampling on the datasets, we calculated the Chao-2 and first order Jackknife (Jack-1) richness estimators (Gotelli and Colwell 2011; Chao et al. 2020) for host gall types, parasitoid species and pairwise interactions at each level of subsampling. Sampling completeness of subsampled datasets was summarised as the value of the metric obtained as a proportion of the estimated total richness for the full dataset. Both estimators give very similar results and for brevity we present results for the Jack-1 estimator.

Results

Results for our four core analyses – incidence and frequency models for sexual and asexual generation gall communities – are summarised in table 1. Incidence and frequency model results for the same dataset were very similar, so we present them in parallel and highlight contrasts. All four models have substantial explanatory power on the latent scale, with median ICC_{GLMM} values (and 95% credible interval) ranging from 0.753 (0.671–0.860) for asexual generation frequency data to 0.829 (0.752–0.903) for sexual generation incidence data (table 1). In what follows, we identify the model term relevant to each inference in brackets, and in all cases inference is conditional on the other terms fitted in the model.

Patterns in Link Richness and Frequency

Neither non-phylogenetic (Host species effect, Parasitoid species effect) or phylogenetic (Host phylogeny effect, Parasitoid phylogeny effect) model terms for link richness and frequency were significant. While intraclass correlations (ICCs) were substantial for the host phylogeny effect in the sexual generation dataset (median ICC=0.179 for incidence data, 0.314 for frequency data; table 1), the lower 2.5% credible interval in both cases abutted zero and was non-significant by our threshold criterion. There was little evidence of among-site variation in the proportion of realised links or average link frequency (Site effect), or in parasitoid richness or average abundance (Site:parasitoid interaction in all four core models). However, host galls showed significant idiosyncratic among-site variation in richness for asexual generation galls (Site:host interaction, incidence data) and in average abundance for both generations (Site:host interaction, frequency data).

Patterns in Link Identity

Phylogenetic effects in link identity were substantial in all four core models, with median (and 95% credible interval) values of ICC_{PHY} for combined phylogenetic terms ranging from 0.118 (0.033–0.261) in the sexual generation frequency model to 0.457 (0.295–0.600) in the asexual generation incidence model (table 1). Point estimates for ICC_{PHY} (i.e. overall phylogenetic signal) were greater for asexual than for sexual generation models (p=0.051), and greater for incidence than for frequency models (though ICC values in these different models cannot be formally compared). The contribution of phylogenetic terms to the ICC for combined phylogenetic and non-phylogenetic link identity effects, $ICC_{REL\text{-}PHY}$, was always high but greater in the asexual

generation than in the sexual generation for both incidence models (0.832 (0.690-0.957) versus 0.561 (0.311–0.763), p=0.011) and frequency models (0.900 (0.807–0.987) versus 0.460 (0.186–0.699), p=0.0005) (table 1).

[Table 1 approximately here]

Cophylogenetic interaction effects were large and significant in the asexual generation, indicating that closely related parasitoids commonly associate with (incidence) or are similarly abundant on (frequency) closely related host galls. This contrasts markedly with the pattern in the sexual generation, in which the cophylogenetic interaction term was non-significant (point estimate < 0.01) in both incidence and frequency models. The trophic links that make a strong contribution to cophylogenetic effects in the asexual generation are identified in figs. 4C and 4E. Hotspots (red) and coldspots (blue) correspond to phylogenetically related sets of parasitoids whose presence (incidence models) or frequency (frequency models) is predicted to be high or low, respectively, on phylogenetically related sets of hosts. Examples in our asexual generation data of hotspots include sets of links between: (i) *Andricus* gall wasps and a clade of generalist *Sycophila*, *Eurytoma* and *Ormyrus* parasitoids; (ii) *Andricus* gall wasps and *Eupelmus annulatus* and *E. urozonus*, (iii) *Cynips* gall wasps and *Torymus* parasitoids, and (iv) *Pseudoneuroterus* gall wasps and *Aprostocetus* parasitoids. In examples (i) and (iii), the parasitoids diversified substantially before the clade of host galls they attack (Appendix 2, figs A1 and A2). The parasitoid clade containing *Sycophila*, *Eurytoma* and *Ormyrus* (median age of most recent common ancestor with 95% posterior credibility interval = 50.1 (35.0-59.7) million years, MY) is substantially older than the *Andricus* clade whose galls they attack (MRCA = 10.2 (8.3-12.7)MY). Similarly, the *Torymus*

clade (33.4 (26.3-42)MY) is substantially older than the *Cynips* clade whose galls they attack (5.7 (4.1-7.6) MY). In examples (ii) and (iv), the divergence of interacting host and parasitoid lineages is broadly contemporary. Divergence date estimates for *Eupelmus annulatus* and *E. urozonus* (14.8 (0-27) MY) encompass those for the *Andricus* clade they interact with (see above), and the same is true for *Aprostocetus* parasitoids (3.0 (0-12.0)MY) attacking *Pseudoneuroterus* galls (7.1 (4.1-11.0)). Cold spots (coloured blue) indicate links whose absence (incidence models) or low frequency (frequency models) shows a strong phylogenetic pattern. Examples (particularly visible for frequency data in fig. 4E) include low levels of interaction between *Andricus* gall wasps and each of (i) *Aprostocetus* parasitoids, (ii) *Pediobius lysis*, and (iii) a clade of pteromalid parasitoids in the genera *Cecidostiba* and *Mesopolobus*.

The two other phylogenetic effects in link identity (Parasitoid phylogenetic interaction and Host phylogenetic interaction) (table 1) were non-significant for all four core models. While ICCs were substantial for the Parasitoid phylogenetic interaction in the sexual generation incidence model (median ICC = 0.133) and for the Host phylogenetic interaction in the asexual generation frequency model (median ICC =0.099), in both cases the lower 2.5% credible interval abutted zero and thus neither was significant by our threshold criterion (table 1). The asexual generation frequency model including structural zeros for unsampled taxa (option b, see Methods) provided stronger support for a Host phylogenetic interaction (Supplement table S9), indicating (in addition to a significant Cophylogenetic interaction) a further tendency for related hosts to be attacked by similar sets of unrelated parasitoids. The lack of individually significant phylogenetic terms in the sexual generation models despite high and significant ICC_{PHY} values suggests that while there is substantial phylogenetic signal, it is not possible with our data to allocate it consistently to specific model terms.

Both generations showed additional patterning in link identity that was not associated with phylogenetic relationships in either trophic level (Parasitoid:host species interaction). This non-phylogenetic effect was supported in both incidence and frequency models and had higher ICC values for the sexual than for the asexual generation (table 1; the effect was marginally non-significant for the asexual generation frequency model by our threshold criterion). As we would expect for a non-phylogenetic effect, the strongly contributing links (dark red or blue in figs. 4A, B and D) are not phylogenetically clustered. We find many more strongly contributing links in models for the sexual generation (figs. 4A, B) than for the asexual generation (fig. 4D). The taxa involved in the strongly contributing links in the sexual generation are taxonomically diverse in both trophic levels, and again are very similar in incidence and frequency models (figs. 4A, B).

Sensitivity of Inference to Sampling Effort and Pooling of data across sites

Removal of sampling effort as a covariate in incidence models did not change the most strongly supported model terms in either generation dataset. Models without sampling effort had generally higher ICCs for link richness model terms (for example, in the sexual generation the median ICC and 95% credible interval for the Host phylogeny effect increased from 0.179 (0.00-0.47) to 0.316 (0.00-0.59)), though these terms remained non-significant (see Supplement section S3 and tables S8, S9).

For all four data sets, models in which data were pooled across sites showed elevated ICCs for most phylogenetic model terms relative to the equivalent full model (option 2; tables S4, S5), but no change in which of these were significant. Reallocation of site-associated variance to other terms in the pooled models was primarily apparent in elevated link richness terms. In the sexual

generation frequency model the ICC for the Parasitoid species effect increased from a mean/mode (and 95% credible interval) of 0.054 / 0.047 (0.008-0.116, non-significant) to 0.112 / 0.084 (0.026-0.227, significant). In the asexual generation frequency model the ICC for the Host species effect increased from 0.047 / 0.034 (0.000-0.098, non-significant) to 0.141 / 0.132 (0.077-0.227, significant).

Sensitivity of Model Results to Sampling Intensity

Our full datasets incorporated near complete sampling of the regional species pool for host galls (observed sexual and asexual generation richnesses are 100% and 98% respectively of the Jack-1 predicted values), but were less complete for parasitoid richness (86%) and interaction richness (75%) (See Supplement figure S4). We used analyses of *in silico* subsampled datasets to assess the possible impact of incomplete sampling on our inference. Subsampling to 75%, 50%, 25%, 10%, and 5% of the complete datasets had little effect on host gall richness, which even for the 5% subsamples exceeded 80% and 88% of Jack-1 estimates for the complete sexual and asexual generation datasets, respectively. The effects of subsampling were more pronounced for parasitoid richness and link richness, which in the 5% subsamples (590/1342 galls for the sexual/asexual generation datasets) fell to around 50% and 20% of Jack-1 estimates for the complete data, respectively (fig. S2).

For most model terms identified as significant in the complete dataset, increasing levels of data reduction were associated with increasing variance in modal ICC estimates across replicate subsamples (fig. 5B), lack of significance in a growing proportion of them (fig. 5A), and hence an increasing type II (false negative) error rate. While the impact of data reduction on ICC variance

was generally stronger in incidence than frequency models (as predicted), for most terms the consequences for significance, and hence inference, were similar for both data types (fig. 5A).

For both generations and data types, the same model terms identified as significant in the full datasets remained significant in > 80% of subsampling replicates with 50% data reduction (fig. 5A). For the sexual generation, power to detect a significant Parasitoid:host species interaction in both datasets remained >75% even down to 5% subsampling. For the asexual generation, power to detect a significant Co-phylogenetic interaction and Host-parasitoid species interaction was more sensitive to data reduction in frequency rather than incidence models; power to detect the Co-phylogenetic interaction with incidence data remained 100% down to 25% sampling. Power to detect the Site:host interaction term (variation in host richness or average abundance between sites), significant in three of the four core full data models, declined in the sexual generation frequency and asexual generation incidence models, but remained close to 100% even at 5% subsampling in the asexual generation frequency model. In contrast, while the Site:host interaction term was not significant in the sexual generation incidence model for the full dataset, the number of model replicates in which it was significant increased with reduced sampling, reaching 50% for the 5% dataset.

Discussion

Bipartite interaction networks are inherently parameter-rich datasets, varying in the richness and composition of interacting taxa and in the strength and distribution of links between these. Considerable effort continues to be invested in characterising patterns in these parameters, both to identify signatures of underlying drivers of network structure (Russo et al. 2018; Harmon et al.

2019; Hembry and Weber 2020; Blasco-Costa et al. 2021; Strydom et al. 2021; Dismukes et al. 2022; Perez-Lamarque and Morlon 2024) and to develop predictive models of link properties for unsampled species (Pearse and Altermatt 2013, 2015; Heimpel et al. 2021; Strydom et al. 2021, 2022, 2023). Two phylogeny GLMMs (2PhyGLMMs) provide a powerful and flexible framework for co-estimation of these multiple-parameter patterns. In particular, when applied to multi-site data, 2PhyGLMMs allow partitioning of patterns into components with and without phylogenetic structure, and hence the extent to which link properties are predictable from phylogenetic information alone. To date 2PhyGLMMs have been used in only a small number of studies encompassing host-parasite interactions (mammals and fleas (Hadfield et al. 2014), sawflies and their host plants (Endara et al. 2018), birds and malaria (Galen et al. 2019)), plant-bacteria associations (Lajoie and Kembel 2021), and insect visits to flowers (Rafferty and Ives 2013). Our analysis provides the first application to host-parasitoid interactions, and joins Hadfield et al. (2014) and Endara et al. (2018) in using spatially structured sampling to separate phylogenetic and non-phylogenetic effects.

Our overarching question is whether associations between cynipid galls and chalcid parasitoids have a strong phylogenetic signal (in host galls, parasitoids, or both) or show strong spatially consistent structure independent of phylogeny in either group. Our rationale is that these patterns reflect alternative distributions of the phenotypic traits that structure parasitoid-gall associations, and hence suggest alternative selective and/or sorting processes acting on these traits. Separate models for communities associated with the two generations in the gallwasp lifecycle had substantial explanatory power, suggesting that some network properties in this system are statistically predictable. Phylogenetic signal was substantial, strongest in the asexual generation community (expressed using the summary statistic ICC_{PHY}), and much stronger for link identity

than for link richness or frequency. Signal for significant model terms was contributed by specific subsets of gall-parasitoid associations, facilitating future targeted research on structuring traits. Our results are robust to whether we use incidence (link presence/absence) or frequency (link abundance) data, whether or not we include sample size as a measure of sampling effort, and to alternative approaches to modelling available pools of interacting species.

Interpreting patterns in association networks

The association networks in this study represent interactions between parasitoids and unknown concealed hosts in three guilds: cynipid gall inducers, herbivorous cynipid inquilines, and other parasitoids (Askew 1961a, Askew 1965; Askew 1975). Some parasitoid species that contribute signal to significant model terms in our study (red cells in fig. 4) are only known to attack the gall inducer (*Aulogymnus* and *Pediobius* species, *Sycophila* species, *Torymus cyaneus*) while others feed on two or all three host guilds (*Bootanomyia dorsalis*, *Eupelmus urozonus*, *Eurytoma brunniventris*, *Hobbya stenonota*, *Mesopolobus fasciiventris*, *M. jucundus*, *Torymus auratus*, *T. erucarum* and *T. geranii*) (Askew 1961a; Askew 1961b; Askew 1961c; Askew 1965; Askew 1975; Schönrogge et al. 1995). We do not know the relative contributions of links with alternative host guilds to the associations in our data. Observed patterns in link properties could thus reflect direct impacts of the inducer's extended gall phenotype on gallwasp-parasitoid links, or indirect impacts on links between parasitoids and non-inducer hosts. Thus a cophylogenetic interaction in our data could indicate either that related parasitoids feed on related gall inducing wasps, or that related parasitoids feed on similar sets of non-inducer hosts that occupy the galls of related gall wasps. Similarly, low support for a link identity term could arise either because no pattern exists or because parasitoid links with galls are a poor proxy for links with non-inducer

hosts within the galls. The main alternative hosts for parasitoids in this system are cynipid inquilines (Askew 1961; Schönrogge et al. 1995). Discovery of multiple cryptic inquiline taxa means that much of the western palaearctic data on inquiline-gall inducer links needs to be verified (Ács et al. 2010). In American communities, cynipid inquilines show a significant (but weak and phylogenetically patchy) signal of co-phylogeny with inducing gall wasps (Ward et al. 2024). It is thus possible that indirect effects could in principle contribute to the patterns observed in our study. A key aspect of this system is that, regardless of host taxon, all parasitoid attack takes place through gall tissues, which are an extended phenotype of the inducing gall wasp. It is thus of interest to consider impacts of gall inducer identity and phylogeny on parasitoid link properties. Similar logic would apply to analysis of species associations in other phenotypic microcosms harbouring specialised communities, such as pitcher plants (Bittleston et al. 2018) or the nests of ants (Parmentier et al. 2020) and birds (Hanzelka et al. 2023).

Link richness and frequency

Link richness and frequency are influenced both by species traits (Ollerton et al. 2007; Stokke et al. 2018; Pichler et al. 2019) and by temporal and spatial variation that can obscure trait-related patterns (Askew 1980; Berlow et al. 2004; Tylianakis and Morris 2017). Some mutualistic pollination and fruit dispersal networks (Rezende et al. 2007a; Tedersoo et al. 2013) and antagonist parasite-host networks (Poulin et al. 2011) show phylogenetic signal in link richness, but much less is known in general about patterns in link frequency (Bailey et al. 2009). We found significant non-phylogenetic spatial patterns in link richness (species degree) and average link frequency (interaction strength) for host galls, but no significant tendency for related host galls to be attacked

by a similar richness of parasitoids, or to experience a similar mean parasitoid attack rate. Parasitoids showed neither significant spatial nor phylogenetic patterns, with little evidence that related parasitoids are similarly specialist or generalist, or attack hosts at similar average rates. Our results suggest that gall traits have a stronger impact on variation in link richness and frequency than parasitoid traits, and/or that any impact of parasitoid traits is more strongly masked by incomplete host information or other ecological and population dynamic effects (Washburn and Cornell 1981; Tylianakis and Morris 2017). Either way, our dataset shows limited potential for phylogeny-based prediction of these parameters.

Link identity

We found strong phylogenetic and non-phylogenetic patterns in link identity (table 1). Phylogenetic signal was greater than non-phylogenetic signal overall, but only significantly so (i.e. lower confidence limits for $ICC_{REL\text{-}PHY} > 0.5$) in the asexual generation dataset. Phylogenetic signal (ICC_{PHY}) was greater for the asexual generation network but was also substantial and significantly non-zero for the sexual generation dataset. Lack of significance for any single phylogenetic effect in the sexual generation dataset is likely a consequence of insufficient informativeness (lack of orthogonality) in the data for separating out the many random effect predictors in the model.

The dominant co-phylogenetic signature in the asexual generation network parallels similar signatures in other host-parasitoid networks (Ives and Godfray 2006; Nyman et al. 2007; Leppänen et al. 2013; Wang et al. 2023), host-pathogen interactions (Clark and Clegg 2017), marine food webs (Eklöf and Stouffer 2016) and mutualistic plant-pollinator and plant-frugivore networks (Rezende et al. 2007b). The strongest pattern in the sexual generation network, however, was non-phylogenetic: similar sets of unrelated parasitoids tend to attack (and show similar attack rates on)

similar sets of unrelated host galls across sites. This same non-phylogenetic effect was also present, though less strongly, in asexual generation galls. To our knowledge this is the first study of herbivore-parasitoid communities to reveal such spatially consistent patterns in link identity, having controlled for phylogenetic and sample size effects.

Previous studies have found greater phylogenetic conservatism in link identity for hosts/prey than for their parasites/predators (Ives and Godfray 2006; Naisbit et al. 2012; Peralta 2016; Cruz-Laufer et al. 2022), perhaps reflecting more rapid evolution of traits determining enemy ability to exploit hosts/prey than of traits determining host/prey vulnerability (Peralta 2016). In our particular case, we might also expect stronger phylogenetic effects for host galls than parasitoids because the parasitoids attacking each gall comprise representatives of phylogenetically diverse chalcidoid families that diverged before the origin of oak gall wasps and associated inquiline cynipids (see below). If there were stronger phylogenetic signal in link identity for galls than for parasitoids we would expect ICC support for the Host phylogenetic interaction having controlled for the Cophylogenetic interaction. While we found no evidence of this in models for associations at the site level (option 1), this prediction was supported for models in which gall-parasitoid associations at any single site were assumed to be possible at all sites (option 2; Supplement table S9). There is thus some evidence that the pattern in oak gall communities parallels that seen in other systems.

Host-Parasitoid Community Assembly

The patterns discussed above suggest that parasitoid-gall associations are shaped at least in part by the distributions across each trophic level of phenotypic traits predicting species overlap in space and time (such as habitat, host plant, phenology, population size and geographic

distribution) (Plantard and Hochberg 1998; Lindenfors et al. 2007; Slove and Janz 2011; Slatyer et al. 2013; Nicholls et al. 2018; Warren et al. 2022) and the outcome of encounters between species. For parasitoids, candidate traits for successful exploitation of a host within a gall include behavioural and physiological traits (such as chemosensory systems) that allow detection of the gall and of hosts concealed within it, morphological traits (such as ovipositor length) that allow successful oviposition through gall and non-gall plant tissues (fig. 2 I-K), and the ability to develop on the resources a host provides (Askew 1965, 1980; Quicke et al. 1998; Egan and Ott 2007; Bailey et al. 2009). Candidate traits in host galls that are under inducer control and influence the probability of parasitoid attack include morphological (Bailey et al. 2009; Egan et al. 2011) and chemical defences (Guiguet et al. 2023), and nectar-based recruitment of ant bodyguards (Abe 1992; Inouye and Agrawal 2004; Warren et al. 2022). Our contrasting results for sexual and asexual generation datasets could indicate structuring by different sets of gall and parasitoid traits, contrasting evolutionary histories of the same (or overlapping) sets of traits in each gall generation, or contrasting impacts of trait-independent population processes.

All of the inducer-controlled gall traits above thought to mediate gall-parasitoid interactions appear able to evolve independently in sexual and asexual gall generations, and show examples of both convergent evolution and phylogenetic conservatism within gall wasp lineages (Stone and Cook 1998; Cook et al. 2002; Stone et al. 2009; Nicholls et al. 2017; Ward et al. 2022). While little is known about patterns of evolution in traits associated with host location behaviour in chalcidoid parasitoids, ovipositor length shows both phylogenetic conservatism and convergent evolution (Al Khatib et al. 2016; Maletti et al. 2021). The significant patterns in link identity in both gall generations that are uncorrelated with either phylogeny are compatible with ecological sorting of associations by convergently evolved traits in galls and parasitoids. In

contrast, the dominant cophylogenetic interaction in asexual generation galls confirms a major role for structuring by traits that are phylogenetically conserved in both trophic levels. Cophylogenetic patterns can arise through (a) coevolutionary codiversification, driven by arms race-type reciprocal adaptive change in gall defences and parasitoid countermeasures (Currie et al. 2003); (b) phylogenetic host tracking, in which parasitoids radiate across an existing diversity of hosts; or (c) trait-based sorting of existing parasitoid lineages over a later radiation of hosts (ecological sorting) (Janz 2011, Althoff et al. 2014). The chalcid parasitoid lineages in our study diversified over 125 million years ago (Cruaud et al. 2024), long before cynipid gall wasp or associated inquiline hosts were available (Blaimer et al. 2020). Assembly of this community has thus involved independent shifts onto gall wasp hosts by multiple parasitoid lineages over tens of millions of years, implying an initial role for ecological sorting (the same is true for parasitoids attacking sawfly hosts (Leppänen et al. 2013)). Concordant ages for some gall wasp and parasitoid divergence events (e.g. for *Eupelmus* parasitoids attacking *Andricus* galls and for *Aprostocetus* parasitoids attacking *Pseudoneuroterus* galls) are more compatible with simultaneous diversification (codiversification).

One way to improve our models and better understand underlying processes would be to incorporate candidate traits for both trophic levels as additional variables in our models (Ives 2022). Our heat maps of predicted model solutions (fig. 4) identify which species associations contribute most to observed patterns, providing a hypothesis for the cross-species distribution of structuring traits. A strength of MCMCglmm and similar model-based approaches (Rafferty and Ives 2013; Ives 2022) is that incorporation of trait data is straightforward in principle, though the models may take a very long time to run. Persistently strong phylogenetic or non-phylogenetic model terms after inclusion of known candidate traits would imply either that the candidate traits

have been correctly identified but incorrectly quantified, or that additional important structuring traits remain to be discovered. Endara et al. (2018) provide an example of this approach, incorporating plant defensive chemistry traits into a 2PhyGLMM analysis of interactions between *Inga* tree and insect herbivores across multiple sites in South America.

Food web prediction in the absence of trait data

High values of ICC_{PHY} in our models indicate high power to predict associations involving specific gall types or parasitoid species based on phylogenetic position in the absence of trait data (Ives and Godfray 2006; Poisot and Stouffer 2018; Braga et al. 2020, 2021; Strydom et al. 2022). High phylogenetic signal in our models could reflect the fact that our network involves a single clade of gall inducers and representatives of a single parasitoid superfamily. Other trophic networks (and many flower visitation networks) involve sets of species in much more distantly related lineages. In MCMCglmm, phylogenetic covariance between a pair of taxa is modelled as the proportion of total tree height that is shared from the root of the phylogeny to their most recent common ancestor. The greater this proportion, the greater the expected covariance (Hadfield 2010; Hadfield and Nakagawa 2010). One potential consequence of incorporating sets of species in phylogenetically distant lineages is that long basal branches compress the informative component of branch length information towards the tips, resulting in very high expected covariance in each major lineage, which gives lower power to estimate phylogenetic effects. This is an issue amenable to exploration using simulated datasets. One potential solution would be to estimate model terms separately for each phylogenetically distant lineage.

For our data, ICC_{PHY} and $ICC_{REL\text{-}PHY}$ values also indicate greater power to predict link identity in models based on incidence data rather than frequency data. Given the substantial effort

required to sample interaction data, the extent to which interactions can be accurately predicted by trait-free phylogenetic models is an important question (Pearse and Altermatt 2013; Strydom et al. 2021). In networks with strong phylogenetic signal, such an approach has high potential applied value in relatively low cost prediction of interactions involving invasive species, introduced control agents, and also species of conservation concern. A potential focus for such approaches includes prediction of the native natural enemies attacking the pest chestnut gall wasp *Dryocosmus kuriphilus*, and prediction of risk to non-target gall wasp hosts associated with widespread release of its biological control agent, *Torymus sinensis* (Gil-Tapetado et al. 2023). The strong non-phylogenetic effects we observed can also have predictive power, but only for the specific species and sites present in the tested dataset. An interesting avenue for future research is to survey the variation in ICC_{PHY} and $ICC_{REL\text{-}PHY}$ values across different types of interaction network. Such a comparative approach should be helpful in identifying key drivers of phylogenetic signal, having controlled for other effects.

Critique of our data and approach

Inference from two phylogeny mixed models is potentially sensitive to multiple aspects of the data used to construct the network and phylogenies, and the sampling design. We therefore assessed potential impacts of sampling completeness (Goldwasser and Roughgarden 1997; Jordano 2016), modelling of sample size (Gotelli and Colwell 2011; Chao et al. 2020), available species pools (Hadfield et al. 2014) and spatial structure (Thompson and Townsend 2005; Brimacombe et al. 2023). Here we consider these in turn, together with potential impacts of phylogenetic uncertainty (Perez-Lamarque et al. 2022).

Sampling completeness

Analyses of simulated networks show that the sampling effort required to adequately estimate network properties is strongly dependent on sampling design and underlying network topology (De Aguiar et al. 2019). The same considerations are likely to influence MCMCglmm model-based inference. Our analyses of replicate subsampled datasets showed that power to identify the same sets of significant model terms was largely resilient to at least 50% data reduction (fig. 5A). Such stability of inference implies that the patterns we observe in the full datasets are unlikely to be artefacts of undersampling. While this resilience may reflect the large size of our multi-site dataset, it also suggests that robust inference may be possible from analyses of other networks with substantially lower sampling than this study.

Sample size

We incorporated sample size as a covariate in incidence models to capture the potential impact of the numbers of each species sampled on the detection of links involving them (Nyman et al. 2007; Bailey et al. 2009; Hadfield et al. 2014; Endara et al. 2018). However, interspecific variation in sample size likely reflects underlying variation in population size, a biological trait that can covary with phylogeny or niche (Vazquez et al. 2005). Incorporation of sampling effort thus has the potential to reduce support for other model terms (Hadfield et al. 2014). We assessed potential for this effect by comparing models with and without incorporation of sample size. For both generation incidence datasets, exclusion did not change the dominant (i.e. highest ICC) model term, though a significant Parasitoid:host species interaction was lost for the asexual generation (table S5).

Phylogenetic Uncertainty

We expect ability to accurately resolve phylogenetic patterns in link properties to depend on the strength of those patterns, the topology of the true phylogeny, and the accuracy with which empirically derived phylogenies capture true relationships (Ives and Godfray 2006; Hadfield et al. 2014; Ives 2022; Perez-Lamarque et al. 2022). Errors that distort the relative phylogenetic distances between taxa distort the covariance matrix used by MCMCglmm, and hence potentially influence model results and inference. Uncertainty over relationships basal to the most recent common ancestor of a pair of taxa does not alter the branch length they share, and so has no direct impact on MCMCglmm models. We have high confidence in our gall wasp phylogeny (Appendix 2, fig. A1), which has high congruence in topology with a previous analysis using larger samples of taxa and hundreds of genes (Blaimer et al. 2020). Resolving chalcid parasitoid relationships is much more challenging, due to a widely recognised signature of rapid radiation (short internal branch lengths) towards the root of the phylogeny (Munro et al. 2011; Cruaud et al. 2024). We therefore used information from recent phylogenomic analyses to inform construction of our parasitoid phylogeny (Appendix 2, fig. A2). Our phylogeny resolves the same monophyletic families as a recent genome-level analysis with high support (Cruaud et al. 2024). Because our phylogenetic uncertainty for parasitoids largely concerns support for short internal branches deep in the phylogeny, we suggest that impacts on estimates of variance covariance, and hence on our inference, will be small. One could quantify the impact of phylogenetic uncertainty by fitting MCMCglmm models to sets of alternative phylogenies for one or both trophic levels (Healy et al. 2014). However, the computational effort required to do this in MCMCglmm would be prohibitive – it would take over 25 core years on the machine used in this study to do only a hundred replicates of each of the core models.

Available Species Pools

A potentially important issue in modelling of species interactions concerns the interpretation of links that are present at some sites but not others. Each unobserved link could be genuinely absent, or be present but undetected (Olesen et al. 2011; Terry and Lewis 2020). We fitted alternative models for these situations in MCMCglmm (our options 1 and 2, respectively). For our data the two options gave the same dominant effects – the main difference being the significant support for a Host phylogenetic interaction in addition to a significant Cophylogenetic interaction only for option 2 in the Asexual generation frequency model. In the Western Palaearctic, the oak gall wasp system is characterised both by short-term patchiness in distributions and frequencies of interaction of individual species (captured by our option 1) (Askew 1980; Washburn and Cornell 1981), and by geographically wide distributions and records of species interaction (captured by option 2) (Askew et al. 2013; Bunnefeld et al. 2018). In each study system the truth is likely to lie somewhere between these two models, and the extent to which their results are concordant provides an indication of how sensitive inference is to model choice.

Spatial Structure

Incorporation of site level information is fundamental to estimating non-phylogenetic patterns in link properties in our full MCMCglmm models, because this effect is estimated through consistency of interactions across replicate samples. Incorporation of site level information was particularly important for our Sexual generation models – we would have detected no significant patterns in link identity without it. Incorporation of site level information also allows controlling for spatial variation in abundance in both trophic levels, with the possibility of scoring unsampled

interactions as structural zeros rather than real data zeros (see (d), above). The effect this can have on inference suggests that spatial structure should be incorporated where possible (Hadfield et al. 2014).

Conclusions

This study shows how a rigorous statistical methodology can disentangle multiple patterns in link properties. Oak gall-parasitoid associations show substantial cophylogenetic signal in link identity, suggesting structuring by phylogenetically conserved traits in both galls and parasitoids, and utility in predicting associations for unsampled species. Strong link identity patterns that are not associated with either the gall or parasitoid phylogeny are potentially attributable to convergent evolution in traits associated with gall defence and parasitoid attack. Further analyses, such as simulations of coevolution followed by examination of resulting patterns, are needed to more closely examine links between pattern and process. It remains to be seen how strong these and other patterns are in other antagonistic and mutualistic interaction networks.

Acknowledgements

This research was funded by NERC grants NE/T000120/1 to G.N.S., A.B.P. and K.S., NE/E014453/1 to G.N.S. and J.A.N., and GR/12847 to G.N.S. and K.S. Y.M.Z was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 101024056. Alex Reiss is funded by a PhD studentship from the UKRI-funded EASTBIO doctoral training program. The authors thank Jarrod Hadfield for his help with MCMCglmm. For the purpose of open access, the authors have applied a

Statement of Authorship

G.N.S., A.B.P., and K.S. conceptualized the research and raised the funding. R.I.B and G. L.C. designed the field sampling, collected the field samples, identified the galls and reared the gall inhabitants. G.M. and J.L.N.A. identified gall inhabitants and provided samples for DNA extraction. C.T.T., J.A.N., and Y.M.Z. generated and analysed the DNA sequence data. F.H.S. led the statistical modelling, data visualisation and data management, with assistance from A.B.P., K.S., and R.I.B. The lead writers of the article were G.N.S. and F.H.S. with editorial input from R.I.B., A.B.P., A.R., K.S., C.T.T., and Y.M.Z.

Data and Code Accessibility

R code for our analyses (including intraclass correlation calculations and figure plotting) and the link datasets are available on the Edinburgh DataArchive (https://doi.org/10.7488/ds/7760; Sinclair et al. 2025)

Appendix A

Table A1 provides interpretations and potential causal processes for significant terms in MCMCglmm models for incidence and frequency data. A link is Table A1 is defined as a bipartite association between specific host gall and parasitoid taxa. Host here refers to the gall from which a parasitoid emerged, rather than its trophic host. In incidence models, link richness (how many taxa a focal taxon is linked to) is also termed the degree, specialisation or generality of a parasitoid

species, and the vulnerability of a host. In frequency models the equivalent terms capture variation in the mean frequency with which a focal taxon interacts with others, and are measures of interaction strength. Link identity (which species interact) is also termed host range or host repertoire. For the Site random effect, which links are considered possible depends on whether unsampled links are modelled as totally absent (structural zeros, option 1), or present but unrecorded (data zeros, option 2) (see Methods).

Table A1. Model terms and the corresponding interpretations if substantiated in incidence or frequency-based MCMCglmm models. Note: Abbreviations in brackets after a term name are as used by Hadfield et al. (2014). Non-phylogenetic random effect terms that are marked with † incorporate an identity matrix for variance covariance between levels. Where present, a superscripts after a model term refers to a footnote listing a process or processes that could generate significant signal for that term.

Term	Description	Interpretation (Incidence models)	Interpretation (Frequency models)
Fixed effect terms			
ln (# hosts sampled)	Natural log of number of galls of a particular type from a site, fitted as a fixed effect	Controls for inter-species and inter-site variation in the number of host galls sampled	Not fitted
ln (# parasitoids sampled)	Natural log of number of parasitoids of a particular species from a site, fitted as a fixed effect	Controls for inter-species and inter-site variation in the number of parasitoids sampled	Not fitted
Random effect terms			
Site	A random factor with a level for each sampling site. †	Estimates among site variation in the proportion of realised links between parasitoids and hosts (of those considered possible in the model)	Estimates among site variation in the average frequency of parasitoids per host (of those considered possible in the model).
Site:host interaction	A random factor with a level for each observed site/host taxon	Estimates among site variation in host richness	Estimate among site variation in mean link frequency of hosts

	combination. †		
Site:parasitoid interaction	A random factor with a level for each observed site /parasitoid taxon combination. †	Estimates among site variation in parasitoid richness	Estimates among site variation in mean link frequency of parasitoids
Host species (r_h)	A random factor with a level for each host taxon. †	Estimates among host species (non-phylogenetic) variation in link richness	Estimates among host species (non-phylogenetic) variation in average link frequency
Parasitoid species (r_p)	A random factor with a level for each parasitoid taxon. †	Estimates among parasitoid species (non-phylogenetic) link richness	Estimates among parasitoid species (non-phylogenetic) variation in average link frequency
Host phylogeny ($r_{[h]}$)[1]	A random factor with a level for each host taxon and variance covariance between levels based on the inverse of the host phylogeny	Estimates among host species (phylogenetic) variation in link richness. Tests whether related hosts show similar link richness	Estimates among host species (phylogenetic) variation in link frequency. Tests whether related hosts show similar average link frequency.
Parasitoid phylogeny ($r_{[p]}$)[2]	A random factor with a level for each host taxon and variance covariance between levels based on the inverse of the parasitoid phylogeny	Estimates among parasitoid species (phylogenetic) variation in link richness. Tests whether elated parasitoids show similar link richness	Estimates among parasitoid species (phylogenetic) variation in link frequency. Tests whether related parasitoids show similar average link frequency
Parasitoid-host Species interaction (r_{ph})[3]	A random factor with a level for each observed host- parasitoid link †	Estimates the variation in link probability across host (non-phylogenetic) and parasitoid (non-phylogenetic) pairs. Tests whether parasitoids have sets of hosts that are not explained by either phylogeny but are consistent across sites	Estimates the variation in link frequency across host (non-phylogenetic) and parasitoid (non-phylogenetic) pairs. Tests whether parasitoids have similar frequencies on particular sets of hosts that are not explained by either phylogeny but are consistent across sites
Host phylogenetic interaction ($r_{p[h]}$)[4]	A random factor with a level for each observed host-parasitoid link, with covariance between levels based on the Kronecker product of the inverted host phylogeny and a parasitoid identity matrix	Estimates the variation in link probability across host (phylogenetic) and parasitoid (non-phylogenetic) pairs. Tests whether related hosts have similar sets of parasitoids (independent of parasitoid phylogeny)	Estimates the variation in link frequency across host (phylogenetic) and parasitoid (non-phylogenetic) pairs. Tests whether related hosts have similar frequencies of particular parasitoids (independent of parasitoid phylogeny)
Parasitoid phylogenetic interaction ($r_{[p]h}$)[5]	A random factor with a level for each observed host-parasitoid link, with covariance between	Estimates the variation in link probability across host (non-phylogenetic) and parasitoid (phylogenetic) pairs. Tests	Estimates the variation in link frequency across host (non-phylogenetic) and parasitoid (phylogenetic)

	levels based on the Kronecker product of the inverted parasitoid phylogeny and a host identity matrix	whether related parasitoids have similar sets of hosts (independent of host phylogeny)	pairs. Tests whether related parasitoids have similar frequencies on particular hosts (independent of host phylogeny)
Cophylogenetic interaction ($r_{[ph]}$)[6]	A random factor with a level for each observed host-parasitoid link, with covariance between levels based on the Kronecker product of the inverted host and parasitoid phylogenies	Estimates the variation in link probability across host (phylogenetic) and parasitoid (phylogenetic) pairs. Tests whether related parasitoids attack related hosts	Estimates the variation in link frequency across host (phylogenetic) and parasitoid (phylogenetic) pairs. Test whteher Related parasitoids are similarly frequent on related hosts
Residual	A random factor with a level for each unique combination of host and parasitoid taxa and site, plus an additional distribution specific variance	Term estimates variation in incidence across observations (i.e an interaction at a site) . Note: for incidence models the random factor not fitted but is fixed at 1, and the distribution specific variance is $\pi^{2/3}$	Term estimates variation in parasitoid frequency across observations (i.e. an interaction at a site). Note: the distribution specific variance is estimated as ln(1/exp(intercept))+1.

Footnotes: Process or processes consistent with host and/or parasitoid structuring traits, i.e. traits contributing to link properties. NB. A significant model term does not imply that all host and/or parasitoid links contribute to signal for that term, but that at least a subset do.

1. A significant Host phylogeny effect implies that host traits substantially correlated with link richness or average link frequency are phylogenetically conserved.

2. A significant Parasitoid phylogeny effect implies that parasitoid traits substantially correlated with link richness or average link frequency are phylogenetically conserved.

3. A significant Parasitoid-host species interaction implies that traits substantially correlated with link identity are phylogenetically labile in both hosts and parasitoids.

4. A significant Host phylogenetic interaction implies that traits substantially correlated with link identity are phylogenetically conserved in hosts but are more labile in parasitoids.

5. A significant Parasitoid phylogenetic interaction implies that traits substantially correlated with link identity are phylogenetically conserved in parasitoids but are more labile in hosts.

6. A significant Cophylogenetic interaction has multiple possible causes, not all of which are trait related:

(a) Simultaneous codiversification of hosts and parasitoids.

(b) Delayed host tracking: radiation of parasitoids across an existing radiation of hosts, where the ability to utilise a host depends on host and parasitoid traits, both of which are phylogenetically conserved.

(c) Geographic vicariance-driven speciation of hosts and their parasitoids, followed by secondary sympatry in both trophic levels.

Literature Cited

Abe, Y. 1992. The advantage of attending ants and gall aggregation for the gall wasp Andricus symbioticus (Hymenoptera: Cynipidae). Oecologia 89:166–167.

Ács, Z., R. J. Challis, P. Bihari, M. Blaxter, A. Hayward, G. Melika, G. Csóka, et al. 2010. Phylogeny and DNA barcoding of inquiline oak gallwasps (Hymenoptera: Cynipidae) of the Western Palaearctic. Molecular Phylogenetics and Evolution 55:210–225.

Al Khatib, F., A. Cruaud, L. Fusu, G. Genson, J.-Y. Rasplus, N. Ris, and G. Delvare. 2016. Multilocus phylogeny and ecological differentiation of the "*Eupelmus urozonus* species group" (Hymenoptera, Eupelmidae) in the West-Palaearctic. BMC Evolutionary Biology 16:13.

Albert, C. H., F. Grassein, F. M. Schurr, G. Vieilledent, and C. Violle. 2011. When and how should intraspecific variability be considered in trait-based plant ecology? Perspectives in Plant Ecology, Evolution and Systematics 13:217–225.

Althoff, D. M., K. A. Segraves, and M. T. J. Johnson. 2014. Testing for coevolutionary diversification: linking pattern with process. Trends in Ecology and Evolution 29:82–89.

Anderson, B., and S. D. Johnson. 2008. The geographical mosaic of coevolution in a plant–pollinator mutualism. Evolution 62:220–225.

Askew, R. R. 1961a. On the biology of the inhabitants of oak galls of Cynipidae (Hymenoptera) in Britain. Transactions of the Society for British Entomology 14:237–268.

Askew, R. R. 1961b. *Eupelmus urozonus* Dalman (Hym., Chalcidoidea) as a parasite in cynipid oak galls. The Entomologist 1:196-201.

Askew, R. R. 1961c. Some biological notes on the pteromalid (Hym. Chalcidoidea) genera *Caenacis* Förster, *Cecidostiba* Thomson and *Hobbya* Delucchi, with descriptions of two new species. Entomophaga 6:58-67.

Askew, R. R. 1965. The biology of the British species of the genus *Torymus* Dalman (Hymenoptera: Torymidae) associated with galls of Cynipidae (Hymenoptera) on oak, with special reference to alternation of forms. Transactions of the Society for British Entomology 16:217–232.

Askew, R. R. 1975. The organisation of chalcid-dominated communities centred upon endophytic hosts. Pages 130-153 *in* P. W. Price, ed. Evolutionary strategies of parasitic insects and mites. Plenum Press, New York & London.

Askew, R. R. 1980. The diversity of insect communities in leafmines and plant galls. The Journal of Animal Ecology 49:817.

Askew, R. R., G. Melika, J. Pujade-Villar, K. Schönrogge, G. N. Stone, and J. L. Nieves-Aldrey. 2013. Catalogue of parasitoids and inquilines in cynipid oak galls in the West Palaearctic. Zootaxa 3643:1–133.

Bailey, R., K. Schönrogge, J. M. Cook, G. Melika, G. Csóka, C. Thuróczy, and G. N. Stone. 2009. Host niches and defensive extended phenotypes structure parasitoid wasp communities. PLoS Biology 7:e1000179.

Ballantyne G., K. C. R. Baldock and P. G. Willmer. 2015. Constructing more informative plant-pollinator networks: visitation and pollen deposition networks in a heathland plant community. Proceedings of the Royal Society B: Biological Sciences 282 (1814):20151130.

Bascompte, J., P. Jordano, and J. M. Olesen. 2006. Asymmetric coevolutionary networks facilitate biodiversity maintenance. Science 312:431–433.

Belshaw, R., A. Grafen, and D. L. J. Quicke. 2003. Inferring life history from ovipositor morphology in parasitoid wasps using phylogenetic regression and discriminant analysis. Zoological Journal of the Linnean Society 139:213–228.

Benadi, G., C. F. Dormann, J. Fründ, R. Stephan, and D. P. Vázquez. 2022. Quantitative prediction of interactions in bipartite networks based on traits, abundances, and phylogeny. The American Naturalist 199:841–854.

Berlow, E. L., A. Neutel, J. E. Cohen, P. C. De Ruiter, B. Ebenman, M. Emmerson, J. W. Fox, V. A. A. Jansen, J. Iwan Jones, G. D. Kokkoris, D. O. Logofet, A. J. McKane, J. M. Montoya, and O. Petchey. 2004. Interaction strengths in food webs: issues and opportunities. Journal of Animal Ecology 73:585–598.

Bersier, L. F., C. Banasek-Richter, and M. F. Cattin. 2002. Quantitative descriptors of food-web matrices. Ecology 83:2394–2407.

Bittleston, L. S., C. J. Wolock, B. E. Yahya, X. Y. Chan, K. G. Chan, N. E. Pierce, and A. pringle. (2018). Convergence between the microcosms of southeast Asian and north American pitcher plants. eLife 7:e36741.

Blaimer, B. B., D. Gotzek, S. G. Brady, and M. L. Buffington. 2020. Comprehensive phylogenomic analyses re-write the evolution of parasitism within cynipoid wasps. BMC Evolutionary Biology 20:155.

Blasco-Costa, I., A. Hayward, R. Poulin, and J. A. Balbuena. 2021. Next-generation cophylogeny: unravelling eco-evolutionary processes. Trends in Ecology & Evolution 36:10.

Boucher, F. C., and V. Démery. 2016. Inferring bounded evolution in phenotypic characters from phylogenetic comparative data. Systematic Biology 65:651–661.

Boucher, F. C., V. Démery, E. Conti, L. J. Harmon, and J. Uyeda. 2018. A general model for estimating macroevolutionary landscapes. Systematic Biology 67:304–319.

Braga, M. P., N. Janz, S. Nylin, F. Ronquist, and M. J. Landis. 2021. Phylogenetic reconstruction of ancestral ecological networks through time for pierid butterflies and their host plants. Ecology Letters 24:2134–2145.

Braga, M. P., M. J. Landis, S. Nylin, N. Janz, and F. Ronquist. 2020. Bayesian inference of ancestral host–parasite interactions under a phylogenetic model of host repertoire evolution. Systematic Biology 69:1149–1162.

Bramon Mora, B., E. Shin, P. J. CaraDonna, and D. B. Stouffer. 2020. Untangling the seasonal dynamics of plant-pollinator communities. Nature Communications 11:4086.

Brimacombe, C., K. Bodner, M. Michalska-Smith, T. Poisot, and M.-J. Fortin. 2023. Shortcomings of reusing species interaction networks created by different sets of researchers. PLOS Biology 21:e3002068.

Bunnefeld, L., J. Hearn, G. N. Stone, and K. Lohse. 2018. Whole-genome data reveal the complex history of a diverse ecological community. Proceedings of the National Academy of Sciences 115:E6507-E6515.

Butler, D. G., B. R. Cullis, A. R. Gilmour, B. J. Gogel, and R. Thompson. 2017. ASReml-R reference manual version 4. VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.

Cagnolo, L., A. Salvo, and G. Valladares. 2011. Network topology: patterns and mechanisms in plant-herbivore and host-parasitoid food webs. Journal of Animal Ecology 80:342–351.

Carvalheiro, L. G., J. C. Biesmeijer, G. Benadi, J. Fründ, M. Stang et al. 2014. The potential for indirect effects between co-flowering plants via shared pollinators depends on resource abundance, accessibility and relatedness. Ecology Letters 17:1389-1399.

Cattin, M.-F., L.-F. Bersier, C. Banašek-Richter, R. Baltensperger, and J.-P. Gabriel. 2004. Phylogenetic constraints and adaptation explain food-web structure. Nature 427:835–839.

Cavender-Bares, J., K. H. Kozak, P. V. A. Fine, and S. W. Kembel. 2009. The merging of community ecology and phylogenetic biology. Ecology Letters 12:693–715.

Caves, E. M. 2021. The behavioural ecology of marine cleaning mutualisms. Biological Reviews 96:2584–2601.

Chacoff, N. P., D. P. Vázquez, S. B. Lomáscolo, E. L. Stevani, J. Dorado, and B. Padrón. 2012. Evaluating sampling completeness in a desert plant–pollinator network. Journal of Animal Ecology 81:190–200.

Chao, A., Y. Kubota, D. Zelený, C.-H. Chiu, C.-F. Li, B. Kusumoto, M. Yasuhara, et al. 2020. Quantifying sample completeness and comparing diversities among assemblages. Ecological Research 35:292–314.

Clark, N. J., and S. M. Clegg. 2017. Integrating phylogenetic and ecological distances reveals new insights into parasite host specificity. Molecular Ecology 26:3074–3086.

Cook, J. M., A. Rokas, M. Pagel, and G. N. Stone. 2002. Evolutionary shifts between host oak sections and host-plant organs in *Andricus* gall wasps. Evolution 56:1821–1830.

Cruaud, A., J.-Y. Rasplus, J. Zhang, R. Burks, G. Delvare, L. Fusu, A. Gumovsky, J. T. H., et al. 2024. The Chalcidoidea bush of life: evolutionary history of a massive radiation of minute wasps. Cladistics 40:34–63.

Cruz-Laufer, A. J., T. Artois, S. Koblmüller, A. Pariselle, K. Smeets, M. Van Steenberge, and M. P. M. Vanhove. 2022. Explosive networking: the role of adaptive host radiations and ecological opportunity in a species-rich host–parasite assembly. Ecology Letters 25:1795–1812.

Currie, C. R., B. Wong, A. E. Stuart, T. R. Schultz, S. A. Rehner, U. G. Mueller, G.-H. Sung, J. W. Spatafora, and N. A. Straus. 2003. Ancient tripartite coevolution in the attine ant-microbe symbiosis. Science 299:386–388.

Darwin, C. 1862. On the various contrivances by which British and foreign orchids are fertilised by insects, and on the good effects of intercrossing. John Murray, London.

Davies, T. J. 2021. Ecophylogenetics redux. Ecology Letters 24:1073–1088.

De Aguiar, M. A. M., E. A. Newman, M. M. Pires, J. D. Yeakel, C. Boettiger, L. A. Burkle, D. Gravel, et al. 2019. Revealing biases in the sampling of ecological interaction networks. PeerJ 7:e7566.

De Araújo, W. S., and V. C. Maia. 2021. Topological structure of a tritrophic network composed of host plants, gall-inducing insects and parasitoids in a restinga area in Brazil. Entomological Science 24:201–216.

Dismukes, W., M. P. Braga, D. H. Hembry., T. A. Heath, and M. J. Landis. 2022. Cophylogenetic methods to untangle the evolutionary history of ecological interactions. Annual Review of Ecology, Evolution, and Systematics 53:13.1-13.24.

Egan, S. P., G. R. Hood, and J. R. Ott. 2011. Natural selection on gall size: variable contributions of individual host plants to population-wide patterns. Evolution 65:3543–3557.

Egan, S. P., and J. R. Ott. 2007. Host plant quality and local adaptation determine the distribution of a gall-forming herbivore. Ecology 88:2868–2879.

Ehrlich, P. R., and P. H. Raven. 1964. Butterflies and plants: a study in coevolution. Evolution 18:586–608.

Eklöf, A., and D. B. Stouffer. 2016. The phylogenetic component of food web structure and intervality. Theoretical Ecology 9:107–115.

Elias, L. G., F. Kjellberg, F. H. A. Farache, E. A. B. Almeida, J.-Y. Rasplus, A. Cruaud, Y.-Q. Peng, et al. 2018. Ovipositor morphology correlates with life history evolution in agaonid fig wasps. Acta Oecologica 90:109–116.

Endara, M. J., J. A. Nicholls, P. D. Coley, D. L. Forrister, G. C. Younkin, K. G. Dexter, C. A. Kidner, R. T., et al. 2018. Tracking of host defenses and phylogeny during the radiation of neotropical inga-feeding sawflies (Hymenoptera; Argidae). Frontiers in Plant Science 9:1237.

Endara, M.-J., P. D. Coley, G. Ghabash, J. A. Nicholls, K. G. Dexter, D. A. Donoso, G. N. Stone, et al. 2017. Coevolutionary arms race versus host defense chase in a tropical

herbivore–plant system. Proceedings of the National Academy of Sciences 114:E7499–E7505.

Espírito-Santo, M. and G. W. Fernandes. 2007. How many species of gall inducing insects are there on earth, and where are they? Annals of the Entomological Society of America 100:95–99.

Falcão, J. C. F., W. Dáttilo, and V. Rico-Gray. 2016. Sampling effort differences can lead to biased conclusions on the architecture of ant–plant interaction networks. Ecological Complexity 25:44–52.

Farache, F. H. A., A. Cruaud, J. -Y. Rasplus, M. T. Cerezini, L. Rattis, F. Kjellberg, and R. A. S. Pereira. 2018. Insights into the structure of plant-insect communities: specialism and generalism in a regional set of non-pollinating fig wasp communities. Acta Oecologica 90:49-59.

Fontaine, C., and E. Thébault. 2015. Comparing the conservatism of ecological interactions in plant–pollinator and plant–herbivore networks. Population Ecology 57:29–36.

Forister, M. L. and Feldman, C. R. 2010. Phylogenetic cascades and the origins of tropical diversity. Biotropica 43: 270-278.

Galen, S. C., K. A. Speer, and S. L. Perkins. 2019. Evolutionary lability of host associations promotes phylogenetic overdispersion of co-infecting blood parasites. Journal of Animal Ecology 88:1936–1949.

Gallinat, A. S., and W. D. Pearse. 2021. Phylogenetic generalized linear mixed modeling presents novel opportunities for eco-evolutionary synthesis. Oikos 130:669–679.

Gilmour, A. R., R. D. Anderson, and A. L. Rae. 1985. The analysis of binomial data by a generalized linear mixed model. Biometrika 72:593–599.

Gil-Tapetado, D., E. K. López-Estrada, Y. Jiménez Ruiz, F. J. Cabrero-Sañudo, J. F. Gómez, P. Durán Montes, C. Rey del Castillo, et al. 2023. *Torymus sinensis* against the invasive chestnut gall wasp: Evaluating the physiological host range and hybridization risks of a classical biological control agent. Biological Control 180:105187.

Giron, D., E. Huguet, G. N. Stone, and M. Body. 2016. Insect-induced effects on plants and possible effectors used by galling and leaf-mining insects to manipulate their host-plant. Journal of Insect Physiology 84:70-89.

Goldwasser, L., and J. Roughgarden. 1997. Sampling effects and the estimation of food-web properties. Ecology 78:41–54.

Gotelli, N. J., and R. K. Colwell. 2011. Estimating species richness. Pages 39–54 *in* Biological diversity: frontiers in measurement and assessment. Oxford University Press, Oxford.

Gripenberg, S., Y. Basset, O. T. Lewis, J. C. D. Terry, S. J. Wright, I. Simón, D. C. Fernández, et al. 2019. A highly resolved food web for insect seed predators in a species-rich tropical forest. Ecology Letters 22:1638–1649.

Guiguet, A., N. B. McCartney, K. J. Gilbert, J. F. Tooker, A. R. Deans, J. G. Ali, and H. M. Hines. 2023. Extreme acidity in a cynipid gall: a potential new defensive strategy against natural enemies. Biology Letters 19:20220513.

Hadfield, J. D. 2017. *MCMCglmm course notes.* Retrieved from: https://github.com/jarrodhadfield/MCMCglmm/blob/master/vignettes/CourseNotes.pdf

Hadfield, J. D. 2010. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. Journal of Statistical Software 33:1–22.

Hadfield, J. D., B. R. Krasnov, R. Poulin, and S. Nakagawa. 2014. A tale of two phylogenies: comparative analyses of ecological interactions. The American Naturalist 183:174–87.

Hadfield, J. D., and S. Nakagawa. 2010. General quantitative genetic methods for comparative biology: phylogenies, taxonomies and multi-trait models for continuous and categorical characters. Journal of Evolutionary Biology 23:494–508.

Hanzelka, J., D. Baroni, P. Martikainen, T. Eeva, and T. Laaksonen. 2023. Cavity-breeding birds create specific microhabitats for diverse arthropod communities in boreal forests. Biodiversity and Conservation 32:3845–3874.

Harmon, L. J., C. S. Andreazzi, F. Débarre, J. Drury, E. E. Goldberg, A. B. Martins, C. J. Melián, et al. 2019. Detecting the macroevolutionary signal of species interactions. Journal of Evolutionary Biology 32:769–782.

Harris, M. O. and A. Pitzschke. 2020. Plants make galls to accommodate foreigners: some are friends, most are foes. New Phytologist 225:1852-1872.

Healy, K., T. Guillerme, S. Finlay, A. Kane, S. B. A. Kelly, D. McClean, D. J. Kelly, I., et al. 2014. Ecology and mode-of-life explain lifespan variation in birds and mammals. Proceedings of the Royal Society B 281:20140298.

Heimpel, G. E., P. K. Abram, and J. Brodeur. 2021. A phylogenetic perspective on parasitoid host ranges with implications for biological control. Current Opinion in Insect Science 44:95–100.

Hembry, H. H., and M. G. Weber. 2020. Ecological interactions and macroevolution: a new field with old roots. Annual Review of Ecology, Evolution, and Systematics. 51:215-243.

Henriksen, M. V., D. G. Chapple, S. L. Chown, and M. A. McGeoch. 2019. The effect of network size and sampling completeness in depauperate networks. Journal of Animal Ecology 88:211–222.

Hommola, K., J. E. Smith, Y. Qiu, and W. R. Gilks. 2009. A permutation test of host-parasite cospeciation. Molecular Biology and Evolution 26:1457–1468.

Hood, G. R. and J. Ott. 2017. Independent life history evolution between generations of bivoltine species: a case study of cyclical parthenogenesis. Oecologia 183:1053–1064.

Inouye, B. D., and A. A. Agrawal. 2004. Ant mutualists alter the composition and attack rate of the parasitoid community for the gall wasp *Disholcaspis eldoradensis* (Cynipidae). Ecological Entomology 29:629–696.

Ives, A. R. 2022. Random errors are neither: on the interpretation of correlated data. Methods in Ecology and Evolution 13:2092–2105.

Ives, A. R., and H. C. J. Godfray. 2006. Phylogenetic analysis of trophic associations. The American Naturalist 168:E1–E14.

Ives, A. R., and M. R. Helmus. 2011. Generalized linear mixed models for phylogenetic analyses of community structure. Ecological Monographs 81:511–525.

Janz, N. 2011. Ehrlich and Raven revisited: mechanisms underlying codiversification of plants and enemies. Annual Review of Ecology Evolution and Systematics 42:71–89.

Jordano, P. 2016. Sampling networks of ecological interactions. Functional Ecology 30:1883–1893.

Lajoie, G., and S. W. Kembel. 2021. Plant-bacteria associations are phylogenetically structured in the phyllosphere. Molecular Ecology 30:5572–5587.

Legendre, P., Y. Desdevises, and E. Bazin. 2002. A statistical test for host-parasite coevolution. Systematic Biology 51:217–234.

Lemos-Costa, P., Miller, Z. R. and Allesina, S. (2024) Phylogeny structures species' interactions in experimental ecological communities. Ecology Letters 27:e14490.

Leppänen, S. A., E. Altenhofer, A. D. Liston, and T. Nyman. 2013. Ecological versus phylogenetic determinants of trophic associations in a plant-leafminer-parasitoid food web. Evolution 67:1493–1502.

Li, D., R. Dinnage, L. A. Nell, M. R. Helmus, and A. R. Ives. 2020. phyr: An r package for phylogenetic species-distribution modelling in ecological communities. Methods in Ecology and Evolution 11:1455–1463.

Lindenfors, P., C. L. Nunn, K. E. Jones, A. A. Cunningham, W. Sechrest, and J. L. Gittleman. 2007. Parasite species richness in carnivores: effects of host body mass, latitude, geographical range and population density. Global Ecology and Biogeography 16:496–509.

Losos, J. B. 2008. Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. Ecology Letters 11:995–1003.

Losos, J. B. 2011. Seeing the forest for the trees: the limitations of phylogenies in comparative biology. The American Naturalist 177: 709-727.

Maia, K. P., C. Rasmussen, J. M. Olesen, and P. R. Guimarães. 2019. Does the sociality of pollinators shape the organisation of pollination networks? Oikos 128:741–752.

Maletti, S., O. Niehuis, C. Mayer, M. Sann, S. Klopfstein, G. Nottebrock, H. Baur, and R. S. Peters. 2021. Phylogeny, taxonomics, and ovipositor length variation of the *Pteromalus albipennis* species group (Hymenoptera: Chalcidoidea: Pteromalidae: Pteromalinae). Journal of Zoological Systematics and Evolutionary Research 59:349–358.

Marjakangas, E., G. Muñoz, S. Turney, J. Albrecht, E. L. Neuschulz, M. Schleuning, and J. Lessard. 2022. Trait-based inference of ecological network assembly: A conceptual framework and methodological toolbox. Ecological Monographs 92:e1502.

McGill, B. J., B. J. Enquist, E. Weiher, and M. Westoby. 2006. Rebuilding community ecology from functional traits. Trends in Ecology and Evolution 21: 178–185.

Memmott, J., H. C. J. Godfray, and I. D. Gauld. 1994. The structure of a tropical host-parasitoid community. The Journal of Animal Ecology 63:521.

Moran, N. A. 1994 Adaptation and constraint in the complex life cycles of animals. Annual Review of Ecology and Systematics 25:573–600.

Munro, J. B., J. M. Heraty, R. A. Burks, D. Hawks, J. Mottern, A. Cruaud, J.-Y. Rasplus, and P. Jansta. 2011. A molecular phylogeny of the Chalcidoidea (Hymenoptera). PLOS ONE 6:e27023.

Naisbit, R. E., R. P. Rohr, A. G. Rossberg, P. Kehrli, and L.-F. Bersier. 2012. Phylogeny versus body size as determinants of food web structure. Proceedings of the Royal Society B 279:3291–3297.

Nakagawa, S., P. C. D. Johnson, and H. Schielzeth. 2017. The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. Journal of The Royal Society Interface 14:20170213.

Nakagawa, S., and H. Schielzeth. 2010. Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. Biological Reviews 85:935–956.

Nicholls, J. A., G. Melika, and G. N. Stone. 2017. Sweet tetra-trophic interactions: multiple evolution of nectar secretion, a defensive extended phenotype in Cynipid gall wasps. The American Naturalist 189:67–77.

Nicholls, J. A., K. Schönrogge, S. Preuss, and G. N. Stone. 2018. Partitioning of herbivore hosts across time and food plants promotes diversification in the *Megastigmus dorsalis* oak gall parasitoid complex. Ecology and Evolution 8:1300–1315.

Nielsen, A., and J. Bascompte. 2007. Ecological networks, nestedness and sampling effort. Journal of Ecology 95:1134–1141.

Novotny, V., S. E. Miller, L. Baje, S. Balagawi, Y. Basset, L. Cizek, K. J. Craft, et al. 2010. Guild-specific patterns of species richness and host specialization in plant-herbivore food webs from a tropical forest. Journal of Animal Ecology 79:1193–1203.

Nyman, T., F. Bokma, and J.-P. Kopelke. 2007. Reciprocal diversification in a complex plant-herbivore-parasitoid food web. BMC Biology 5:49.

Olesen, J. M., J. Bascompte, Y. L. Dupont, H. Elberling, C. Rasmussen, and P. Jordano. 2011. Missing and forbidden links in mutualistic networks. Proceedings of the Royal Society B 278:725–732.

Oliveira, J. B. B. S., M. L. Faria, M. A. Z. Borges, M. Fagundes, and W. S. de Araújo. 2020. Comparing the plant–herbivore network topology of different insect guilds in neotropical savannas. Ecological Entomology 45:406–415.

Ollerton, J., Killick, A., Lamborn, E., Watts, S. and Whiston, M. 2007. Multiple meanings and modes: On the many ways to be a generalist flower. Taxon 56:717-728.

Pakeman, R. J. 2014. Functional trait metrics are sensitive to the completeness of the species' trait data? Methods in Ecology and Evolution 5: 9–15.

Parmentier, T., F. De Laender, and D. Bonte. 2020. The topology and drivers of ant–symbiont networks across Europe. Biological Reviews 95:1664-1688.

Pearse, I. S., and F. Altermatt. 2013. Predicting novel trophic interactions in a non-native world. Ecology Letters 16:1088–1094.

Pearse, I. S., and F. Altermatt. 2015. Out-of-sample predictions from plant–insect food webs: robustness to missing and erroneous trophic interaction records. Ecological Applications 25:1953–1961.

Pearse, I. S., and A. L. Hipp. 2009. Phylogenetic and trait similarity to a native species predict herbivory on non-native oaks. Proceedings of the National Academy of Sciences 106:18097–18102.

Penone, C., A. D. Davidson, K. T. Shoemaker, M. Di Marco, C. Rondinini, T. M. Brooks, B. E. Young, et al. 2014. Imputation of missing data in life-history trait datasets: which approach performs the best? Methods in Ecology and Evolution 5: 961–970.

Peralta, G. 2016. Merging evolutionary history into species interaction networks. Functional Ecology 30:1917–1925.

Perez-Lamarque, B., O. Maliet, B. Pichon, M.-A. Selosse, F. Martos, and H. Morlon. 2022. Do closely related species interact with similar partners? Testing for phylogenetic signal in bipartite interaction networks. Peer Community Journal 2:e59.

Perez-Lamarque, B. and H. Morlon. 2024. Distinguishing cophylogenetic signal from phylogenetic congruence clarifies the interplay between evolutionary history and species interactions. Systematic Biology 73:613-622.

Pichler, M., V. Boreux, A. Klein, M. Schleuning, and F. Hartig. 2019. Machine learning algorithms to infer trait-matching and predict species interactions in ecological networks. Methods in Ecology and Evolution 11:281-293.

Pinilla-Gallego, M. S., W. H. Ng, V. E. Amaral, and R. E. Irwin. 2022. Floral shape predicts bee–parasite transmission potential. Ecology 103:e3730.

Plantard, O., and M. E. Hochberg. 1998. Factors affecting parasitism in the oak-galler *Neuroterus quercusbaccarum* (Hymenoptera: Cynipidae). Oikos 81:289.

Poisot, T., and D. B. Stouffer. 2018. Interactions retain the co-phylogenetic matching that communities lost. Oikos 127:230–238.

Poulin, R., B. R. Krasnov, D. Mouillot, and D. W. Thieltges. 2011. The comparative ecology and biogeography of parasites. Philosophical Transactions of the Royal Society B 366:2379–2390.

Quicke, D. L. J., and R. Belshaw. 1999. Incongruence between morphological data sets: an example from the evolution of endoparasitism among parasitic wasps (Hymenoptera: Braconidae). Systematic Biology 48:436–454.

Quicke, D. L. J., P. Wyeth, J. D. Fawke, H. H. Basibuyuk, and J. F. V. Vincent. 1998. Manganese and zinc in the ovipositors and mandibles of hymenopterous insects. Zoological Journal of the Linnean Society 124:387–396.

R Core Team. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rafferty, N. E., and A. R. Ives. 2013. Phylogenetic trait-based analyses of ecological networks. Ecology 94:2321–2333.

Raman, A., C. W. Schaeffer, and T. M. Withers (eds.). 2005. Biology, ecology and evolution of gall-inducing arthropods. Science Publishers Inc., Enfield, New Hampshire, USA.

Rezende, E. L., P. Jordano, and J. Bascompte. 2007a. Effects of phenotypic complementarity and phylogeny on the nested structure of mutualistic networks. Oikos 116:1919–1929.

Rezende, E. L., J. E. Lavabre, P. R. Guimarães, P. Jordano, and J. Bascompte. 2007b. Non-random coextinctions in phylogenetically structured mutualistic networks. Nature 448:925–928.

Rivera-Hutinel, A., R. O. Bustamante, V. H. Marín, and R. Medel. 2012. Effects of sampling completeness on the structure of plant–pollinator networks. Ecology 93:1593–1603.

Roskam, J. C. 2019. Plant galls of Europe. KNNV Uitgeverij, Amsterdam.

Rossberg, A. G., H. Matsuda, T. Amemiya, and K. Itoh. 2006. Food webs: experts consuming families of experts. Journal of Theoretical Biology 241:552–563.

Russo, L., A. D. Miller, J. Tooker, O. N. Bjornstad, and K. Shea. 2018. Quantitative evolutionary patterns in bipartite networks: vicariance, phylogenetic tracking or diffuse co-evolution? Methods in Ecology and Evolution 9:761–772.

Santos, N., J. F. D. Andrade, R. A. S. Pereira, and F. H. A. Farache. 2022. Community structure and specialization in fig wasps (Hymenoptera: Chalcidoidea) in a region of Cerrado. Revista Brasileira de Entomologia 66:e20210101.

Schoener, T. W. 1989. Food webs from the small to the large: The Robert H. MacArthur award lecture. Ecology 70:1559–1589.

Schönrogge, K., G. N. Stone, and M. J. Crawley. 1995. Spatial and temporal variation in guild structure: parasitoids and inquilines of *Andricus quercuscalicis* (Hymenoptera: Cynipidae) in its native and alien ranges. Oikos 72:51-60.

Silva, F. F. e, J. M. S. Viana, V. R. Faria, and M. D. V. de Resende. 2013. Bayesian inference of mixed models in quantitative genetics of crop species. Theoretical and Applied Genetics 126:1749–1761.

Sinclair, F. H., G. N. Stone, J. A. Nicholls, S. Cavers, M. Gibbs, P. Butterill, S. Wagner, et al. 2015. Impacts of local adaptation of forest trees on associations with herbivorous insects: implications for adaptive forest management. Evolutionary Applications 8:972–987.

Sinclair, F. H., C.-T. Tang, R. A. Bailey, G. L., Csóka, G. Melika, J. A. Nicholls, J.-L. Nieves-Aldrey, et al. 2025. Data and R code from: Phylogenetic and non-phylogenetic patterns in the richness, frequency and composition of links in a herbivore-parasitoid interaction network. The American Naturalist. Edinburgh DataShare Repository, https://doi.org/10.7488/ds/7760

Singer, M. S., and J. O. Stireman. 2005. The tri-trophic niche concept and adaptive radiation of phytophagous insects. Ecology Letters 8:1247–1255.

Slatyer, R. A., M. Hirst, and J. P. Sexton. 2013. Niche breadth predicts geographical range size: a general ecological pattern. Ecology Letters 16:1104–1114.

Slove, J., and N. Janz. 2011. The relationship between diet breadth and geographic range size in the butterfly subfamily Nymphalinae – a study of global scale. PLOS ONE 6:e16057.

Smith, M. A., J. J. Rodriguez, J. B. Whitfield, A. R. Deans, D. H. Janzen, W. Hallwachs, and P. D. N. Hebert. 2008. Extreme diversity of tropical parasitoid wasps exposed by iterative integration of natural history, DNA barcoding, morphology, and collections. Proceedings of the National Academy of Sciences 105:12359–12364.

Sorensen, D. 2009. Developments in statistical analysis in quantitative genetics. Genetica 136:319–332.

Stireman, J. O., and M. S. Singer. 2003. Determinants of parasitoid-host associations: insights from a natural Thacinid-Lepidopteran community. Ecology 84:296–310.

Stokke, B. G., I. I. Ratikainen, A. Moksnes, E. Røskaft, K. Schulze-Hagen, D. I. Leech, A. Pape Møller, and F. Fossøy. 2018. Characteristics determining host suitability for a generalist parasite. Scientific Reports 8:6285.

Stone, G. N., and J. M. Cook. 1998. The structure of cynipid oak galls: patterns in the evolution of an extended phenotype. Proceedings of the Royal Society B 265:979–988.

Stone, G. N., A. Hernandez-Lopez, J. A. Nicholls, E. di Pierro, J. Pujade-Villar, G. Melika, and J. M. Cook. 2009. Extreme host plant conservatism during at least 20 million years of host plant pursuit by gall wasps. Evolution 63:854–869.

Stone, G. N., and K. Schönrogge. 2003. The adaptive significance of insect gall morphology. Trends in Ecology & Evolution 18:512–522.

Stouffer, D. B., M. Sales-Pardo, M. I. Sirer, and J. Bascompte. 2012. Evolutionary conservation of species' roles in food webs. Science 335:1489–1492.

Strydom, T., S. Bouskila, F. Banville, C. Barros, D. Caron, M. J. Farrell, M. Fortin, V. Hemming, et al. 2022. Food web reconstruction through phylogenetic transfer of low-rank network representation. Methods in Ecology and Evolution 13:2838–2849.

Strydom, T., S. Bouskila, F. Banville, C. Barros, D. Caron, J. Maxwell, M. Fortin, B. Mercier, et al. 2023. Graph embedding and transfer learning can help predict potential species interaction networks despite data limitations. Methods in Ecology and Evolution 14:2917-2930.

Strydom, T., M. D. Catchen, F. Banville, D. Caron, G. Dansereau, P. Desjardins-Proulx, N. R. Forero-Muñoz, et al. 2021. A roadmap towards predicting species interaction networks (across space and time). Philosophical Transactions of the Royal Society B 376:20210063.

Tedersoo, L., M. Mett, T. A. Ishida, and M. Bahram. 2013. Phylogenetic relationships among host plants explain differences in fungal species richness and community composition in ectomycorrhizal symbiosis. New Phytologist 199:822–831.

Terry, J. C. D., and O. T. Lewis. 2020. Finding missing links in interaction networks. Ecology 101:e03047.

Thompson, J. N. 2005. The geographic mosaic of coevolution. University of Chicago Press, Chicago.

Thompson, R. M., and C. R. Townsend. 2005. Food-web topology varies with spatial scale in a patchy environment. Ecology 86:1916–1925.

Truitt, L. L., S. H. McArt, A. H. Vaughn, and S. P. Ellner. 2019. Trait-based modelling of multihost pathogen transmission: plant-pollinator networks. The American Naturalist 193:E149–E167.

Tylianakis, J. M. and Morris, R. J. 2017. Ecological networks across environmental gradients. Annual Review of Ecology, Evolution, and Systematics 48:25–48.

Vazquez, D. P., R. Poulin, B. R. Krasnov, and G. I. Shenbrot. 2005. Species abundance and the distribution of specialization in host-parasite interaction networks. Journal of Animal Ecology 74:946–955.

Vizentin-Bugoni, J., P. K. Maruyama, V. J. Debastiani, L. da S. Duarte, B. Dalsgaard, and M. Sazima. 2016. Influences of sampling effort on detected patterns and structuring processes of a Neotropical plant–hummingbird network. Journal of Animal Ecology 85:262–272.

Wang, A., Y. Peng, J. M. Cook, D. Yang, D. Zhang, and W. Liao. 2023. Host insect specificity and interspecific competition drive parasitoid diversification in a plant–insect community. Ecology 104:e4062.

Ward, A. K. G., R. K. Bagley, S. P. Egan, G. R. Hood, J. R. Ott, K. M. Prior, S. I. Sheikh, et al. 2022. Speciation in Nearctic oak gall wasps is frequently correlated with changes in host plant, host organ, or both. Evolution 76:1849–1867.

Ward, A. K. G., Y. M. Zhang, G. E. Brown, A. C. Hippee, K. M. Prior, S. Rollins, N. Sierra, et al. 2024. Speciation in kleptoparasites of oak gall wasps often correlates with shifts into new tree habitats, tree organs, or gall morphospace. Evolution 78:174–187.

Warren, R. J., A. Guiguet, C. Mokadam, J. F. Tooker, and A. R. Deans. 2022. Oak galls exhibit ant dispersal convergent with Myrmecochorous seeds. The American Naturalist 200:292–301.

Washburn, J. O., and H. V. Cornell. 1981. Parasitoids, patches, and phenology: their possible role in the local extinction of a cynipid gall wasp population. Ecology 62:1597–1607.

Whittall, J. B., and S. A. Hodges. 2007. Pollinator shifts drive increasingly long nectar spurs in columbine flowers. Nature 447:706–709.

Wong, M. K. L., and C. P. Carmona. 2021. Including intraspecific trait variability to avoid distortion of functional diversity and ecological inference: Lessons from natural assemblages. Methods in Ecology and Evolution 12:946–957.

Yeakel, J. D., P. R. Guimarães, M. Novak, K. Fox-Dobbs, and P. L. Koch. 2012. Probabilistic patterns of interaction: the effects of link-strength variability on food web structure. Journal of The Royal Society Interface 9:3219–3228.

References Cited Only in the Online Enhancements

Barido-Sottani, J., V. Bošková, L. du Plessis, D. Kühnert, C. Magnus, V. Mitov, N. F. Müller, et al. 2018. Taming the BEAST – A community teaching material resource for BEAST 2. Systematic Biology 67:170–174.

Burks, R. A., J. M. Heraty, M. Gebiola, and C. Hansson. 2011. Combined molecular and morphological phylogeny of Eulophidae (Hymenoptera: Chalcidoidea), with focus on the subfamily Entedoninae. Cladistics 27:581–605.

Bouckaert, R., T. G. Vaughan, J. Barido-Sottani, S. Duchêne, M. Fourment, Gavryushkina A., Heled, J., et al. 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. PLoS Computational Biology 15:e1006650.

Gil-Tapetado, D., P. Durán-Montes, M. García-París, E. K. López-Estrada, A. Sánchez-Vialas, Y. Jiménez-Ruiz, J. F. Gómez, and J.-L. Nieves-Aldrey. 2022. Host specialization is ancestral in *Torymus* (Hymenoptera, Chalcidoidea) cynipid gall parasitoids. Zoologica Scripta 51:91–118.

Heraty, J., D. Hawks, J. S. Kostecki, and A. Carmichael. 2004. Phylogeny and behaviour of the Gollumiellinae, a new subfamily of the ant-parasitic Eucharitidae (Hymenoptera: Chalcidoidae). Systematic Entomology 29:544-559.

Ho, S. Y. W., and M. J. Phillips. 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. Systematic Biology 58:367–380.

Jermiin, L. S. and R. H. Crozier. 1994. The cytochrome-b region in the mitochondrial DNA of the ant *Tetraponera rufoniger*–sequence divergence in Hymenoptera may be associated with nucleotide content. Journal of Molecular Evolution 38:282–294.

Kaartinen, R., G. N. Stone, J. Hearn, K. Lohse, and T. Roslin. 2010. Revealing secret liaisons: DNA barcoding changes our understanding of food webs. Ecological Entomology 35:623–638.

Kalyaanamoorthy, S., B.Q. Minh, T.K. Wong, A. Von Haeseler, and L.S. Jermiin 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nature Methods 14:587–589.

Kass, R. E. and Raftery, A. E. 1995. Bayes factors. Journal of the American Statistical Association 90:773–795.

Katoh, K., Rozewicki, J., and Yamada, K. D. 2019. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Briefings in Bioinformatics 20:1160–1166.

Maturana R., P., B. J. Brewer, S. Klaere, and R. Bouckaert. 2018. Model selection and parameter inference in phylogenetics using Nested Sampling. Systematic Biology 68:219–233.

Minh, B. Q., H. A. Schmidt, O. Chernomor, D. Schrempf, M. D. Woodhams, A. von Haeseler, and R. Lanfear. 2020. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. Molecular Biology and Evolution 37:1530–1534.

Nicholls, J.A., Preuss, S., Hayward, A., Melika, G., Csóka, G., Nieves-Aldrey, J.-L., Askew, R. R. et al. 2010. Concordant phylogeography and cryptic speciation in two Western Palaearctic oak gall parasitoid species complexes. Molecular Ecology 19:592–609.

Ogilvie, H. A., R. R. Bouckaert, and A. J. Drummond. 2017. Starbeast2 brings faster species tree inference and accurate estimates of substitution rates. Molecular Biology and Evolution 34:2101–2114.

Peters, R. S., O. Niehuis, S. Gunkel, M. Bläser, C. Mayer, L. Podsiadlowski, A. Kozlov, et al. 2018. Transcriptome sequence-based phylogeny of chalcidoid wasps (Hymenoptera: Chalcidoidea) reveals a history of rapid radiations, convergence, and evolutionary success. Molecular Phylogenetics and Evolution 120:286–296.

Pringle, R.M. & Hutchinson, M.C. (2020) Resolving food-web structure. Annual Review of Ecology, Evolution, and Systematics 51:55–80.

Rambaut A., Drummond, A. J., Xie, D., Baele, G. and Suchard, M. A. 2018. Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. Systematic Biology 67:901–904.

Rasplus, J.-Y., B. B. Blaimer, S. G. Brady, R. A. Burks, G. Delvare, N. Fisher, M. Gates, et al. 2020. A first phylogenomic hypothesis for Eulophidae (Hymenoptera, Chalcidoidea). Journal of Natural History 54:597–609.

Simon, C., F. Frati, A. Beckenbach, B. Crespi, H. Liu, and P. Flook. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Annals of the Entomological Society of America 87:651–701.

Skilling, J. 2006. Nested sampling for general Bayesian computation. Bayesian Analysis 1:833–860.

Swofford, D. L. 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Tables

Table 1: Intraclass correlations (ICCs) for terms in the 4 core models (incidence and frequency-based models of sexual and asexual generation gall datasets). Values are presented as median/mode (and 95% credible interval) over 2000 MCMC point estimates. Significant terms (i.e. where the lower 95% credible interval >0.01) are highlighted in **bold** for each model. ICC_{GLMM} is the ICC for all fitted model terms (i.e. excluding the residual), analogous to an R^2 for the model. ICC_{PHY} is the combined ICC for phylogenetic interaction terms in link identity, and $ICC_{REL\text{-}PHY}$ is the relative contribution of these terms to the ICC for all interaction terms in link identity.

Term	Sexual generation galls		Asexual generation galls	
	Incidence model	Frequency model	Incidence model	Frequency model
Site	.014 / .000 (.000-.096)	.027 / .016 (.000-.134)	.021 / .001 (.000-.126)	.037 / .023 (.002-.159)
Site:host interaction	.030 / .031 (.006-.069)	**.064 / .055 (.025-.122)**	**.033 / .029 (.012-.059)**	**.122 / .127 (.072-.175)**
Site:parasitoid interaction	.001 / .000 (.000-.006)	.002 / .000 (.000-.009)	.001 / .000 (.000-.005)	.003 / .000 (.000-.007)
Host species	.031 / .001 (.000-.142)	.024 / .001 (.000-.157)	.030 / .030 (.000-.075)	.061 / .069 (.000-.122)
Parasitoid species	.007 / .000 (.000-.049)	.040 / .039 (.000-.099)	.008 / .000 (.000-.052)	.032 / .001 (.000-.093)
Host phylogeny	.179 / .002 (.000-.471)	.314 / .340 (.000-.579)	.042 / .001 (.000-.202)	.061 / .002 (.000-.306)
Parasitoid phylogeny	.019 / .000 (.000-.106)	.015 / .001 (.000-.099)	.011 / .001 (.000-.086)	.013 / .000 (.000-.104)
Parasitoid:host species interaction	**.207 / .224 (.095-.329)**	**.143 / .159 (.068-.230)**	**.092 / .080 (.025-.164)**	.035 / .040 (.009-.068)
Host phylogenetic interaction	.013 / .000 (.000-.077)	.009 / .000 (.000-.070)	.024 / .001 (.000-.175)	.099 / .001 (.000-.231)
Parasitoid phylogenetic interaction	.133 / .123 (.000-.268)	.047 / .000 (.000-.110)	.040 / .000 (.000-.106)	.022 / .014 (.000-.049)
Cophylogenetic interaction	.089 / .002 (.000-.281)	.038 / .001 (.000-.192)	**.367 / .392 (.178-.537)**	**.185 / .142 (.026-.329)**
Residual	.171 / .161 (.097-.248)	.185 / .237 (.052-.299)	.244 / .258 (.172-.319)	.246 / .268 (.140-.333)
ICC denominator	25.157 / 22.228 (16.192-40.403)	30.150 / 27.075 (18.451-49.273)	17.549 / 16.641 (12.996-24.187)	26.895 / 24.711 (19.63-37.80)
ICC_{GLMM}	**.829 / .839 (.752-.903)**	**.815 / .763 (.701-.948)**	**.756 / .742 (.681-.828)**	**.754 / .732 (.667-.860)**
ICC_{PHY}	**.263 / .291 (.098-.443)**	**.118 / .076 (.033-.261)**	**.457 / .450 (.295-.600)**	**.315 / .335 (.195-.432)**
$ICC_{REL\text{-}PHY}$	**.561 / .581 (.311-.763)**	**.460 / .445 (.186-.699)**	**.832 / .845 (.690-.957)**	**.900 / .912 (.807-.976)**

Figure legends

Figure 1. Conceptual representations of variation among taxa in link richness (A & B – species vary in how many taxa they interact with) and link identity (C, D, & E – species (co)vary in which taxa they interact with) in a bipartite interaction network. Each diagram shows links (coloured lines) between an upper parasitoid trophic level and a lower herbivore trophic level. Patterns B, C and D incorporate phylogenetic information for one or both trophic levels. Patterns A to C are illustrated in terms of the upper (parasitoid) trophic level but analogous patterns also exist for the herbivore trophic level. Colours show variation in link richness from the parasitoid trophic level perspective (see key). The title in bold text for each pattern is the name of the relevant model term in our models. For clarity, links for some parasitoid taxa in panels A and E have been omitted.

Figure 2. Examples of the diverse gall structures induced by sexual and asexual generations of oak gall wasps in this study. A-H. Sexual generation galls (A-D, above) and asexual generation galls (E-H, below) of the same set of species (A, E: *Andricus conificus.* B, F: *A. coriarius.* C, G: *A. grossulariae.* D, H: *A. lignicolus*). I. A female *Torymus* parasitoid using its long ovipositor to drill through a leaf midrib and gall tissues to lay an egg on a concealed insect host in an asexual generation gall of *Neuroterus anthracinus.* J. A female *Bootanomyia stigmatizans* ovipositing directly into an asexual generation gall of *Andricus infectorius*. K. The same female *B. stigmatizans,* showing the length of the withdrawn ovipositor. The scale bar in all images is 5mm. Images A-H, J, K: G. Csóka. I: Darren Obbard.

Figure 3. Heat maps showing bipartite link frequencies between parasitoids and sexual generation (top) and asexual generation (bottom) cynipid galls, pooled across sampling sites. Bar plots summarise sample sizes for host galls (right) and parasitoids (top) in the datasets for the sexual

generation (blue) and asexual generation (orange). Phylogenies with node ages and branch support information are shown in Appendix 2.

Figure 4. Heatmaps showing the extent to which specified model terms increase (red) or decrease (blue) link probability (incidence model) or frequency (frequency model). Grey-filled cells indicate host and parasitoid links whose solution was not estimated because the two species were not present in the same site and generation (i.e. structural zeros; see Methods). Cell values are posterior modes of the predicted MCMCglmm solutions (see Methods). In all plots, parasitoid taxa are arranged across the top, and cynipid gall generations down the side. A. Sexual generation incidence model Parasitoid:host species interaction term. B. Sexual generation frequency model Parasitoid:host species interaction term. C. Asexual generation incidence model Cophylogenetic interaction term. D. Asexual generation incidence model Parasitoid:host species interaction term. E. Asexual generation frequency model Cophylogenetic interaction term. Gall taxon lists omitted from B and D are the same as those in panels to the left, and the parasitoid taxon list omitted from E is the same as for the panel above.

Figure 5. A. Proportion of models for each generation and data type in which a given model term was significant (black bars: i.e. the lower bound of the 95% credible interval for the ICC >0.01) or nonsignificant (red bars) for different levels of data reduction. Filled circles to the left of each panel indicate the outcome for the model of the full dataset. Bars in each panel represent the proportion of significant and nonsignificant outcomes for each model term over 30 replicate data subsets for levels of data completeness from 75% (top), 50%, 25%, 10%, & 5% (bottom). B. Intraclass correlations (ICCs) for models of parasitoid incidence and frequency on sexual and asexual generation galls. For models of full datasets, black dots and bars represent the posterior

mode and 95% credible intervals for ICC estimates. For subsampled datasets, coloured dots represent the posterior modes for each of 30 replicate data subsets.

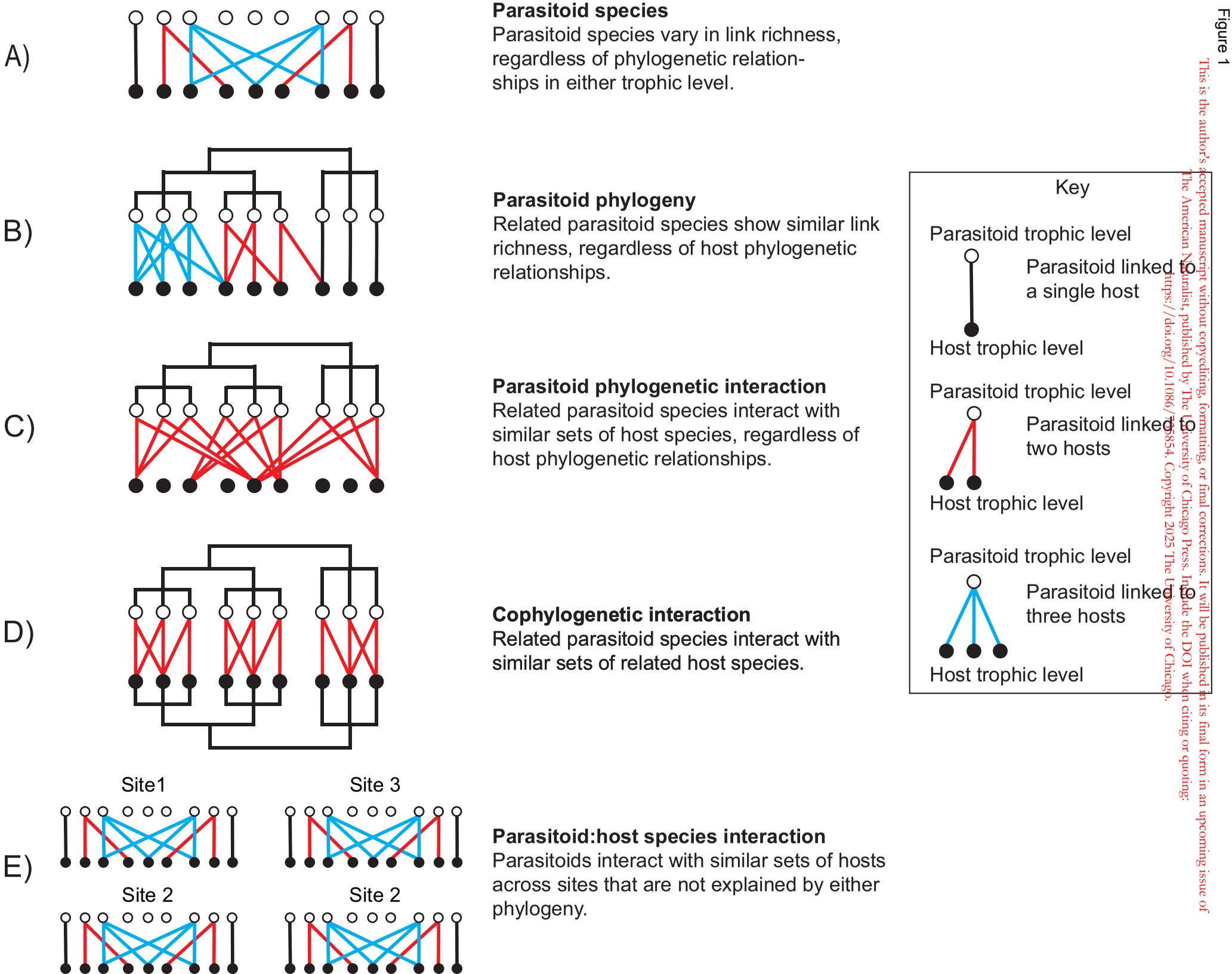


Figure 1

Figure 2

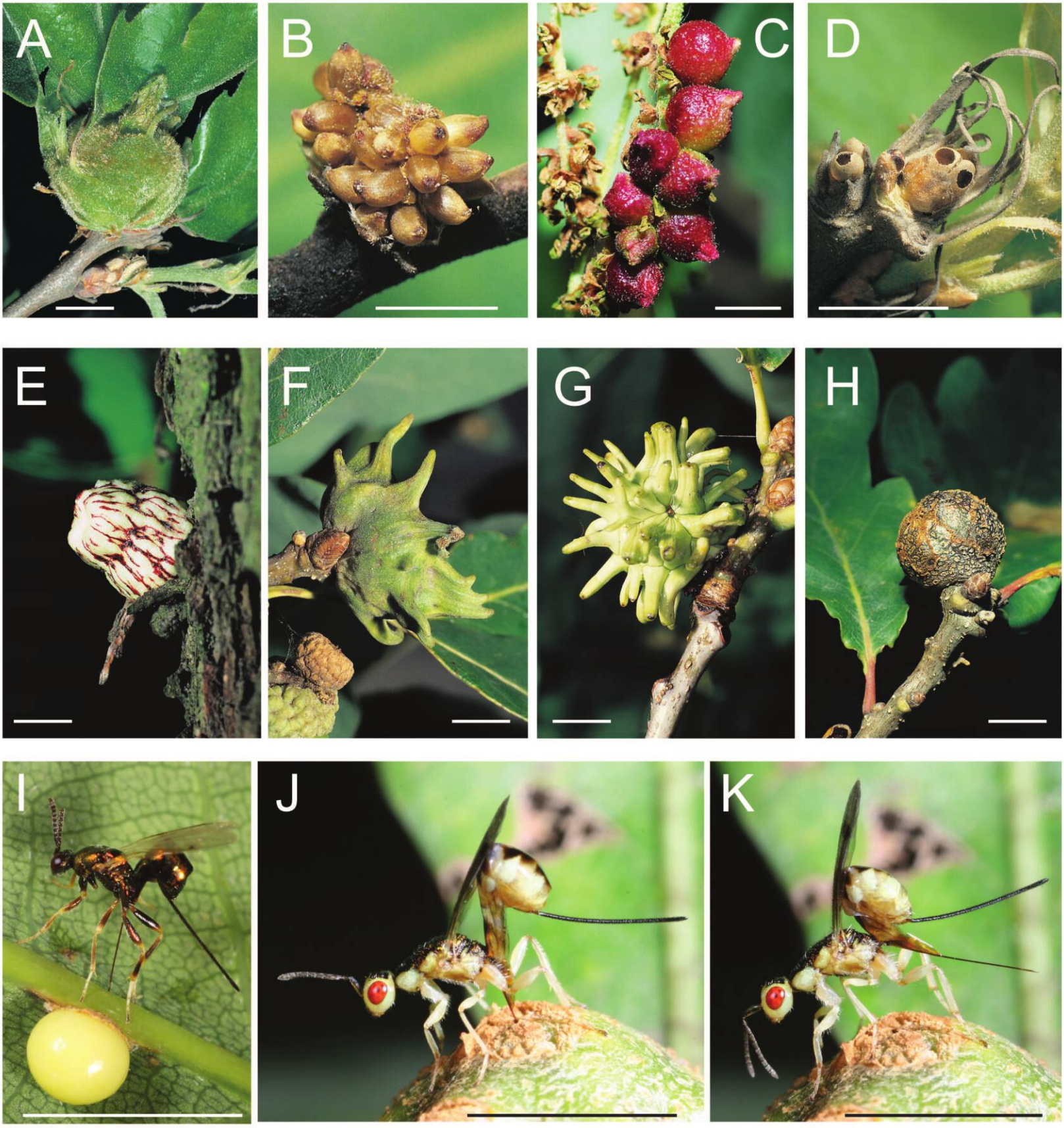


Figure 3

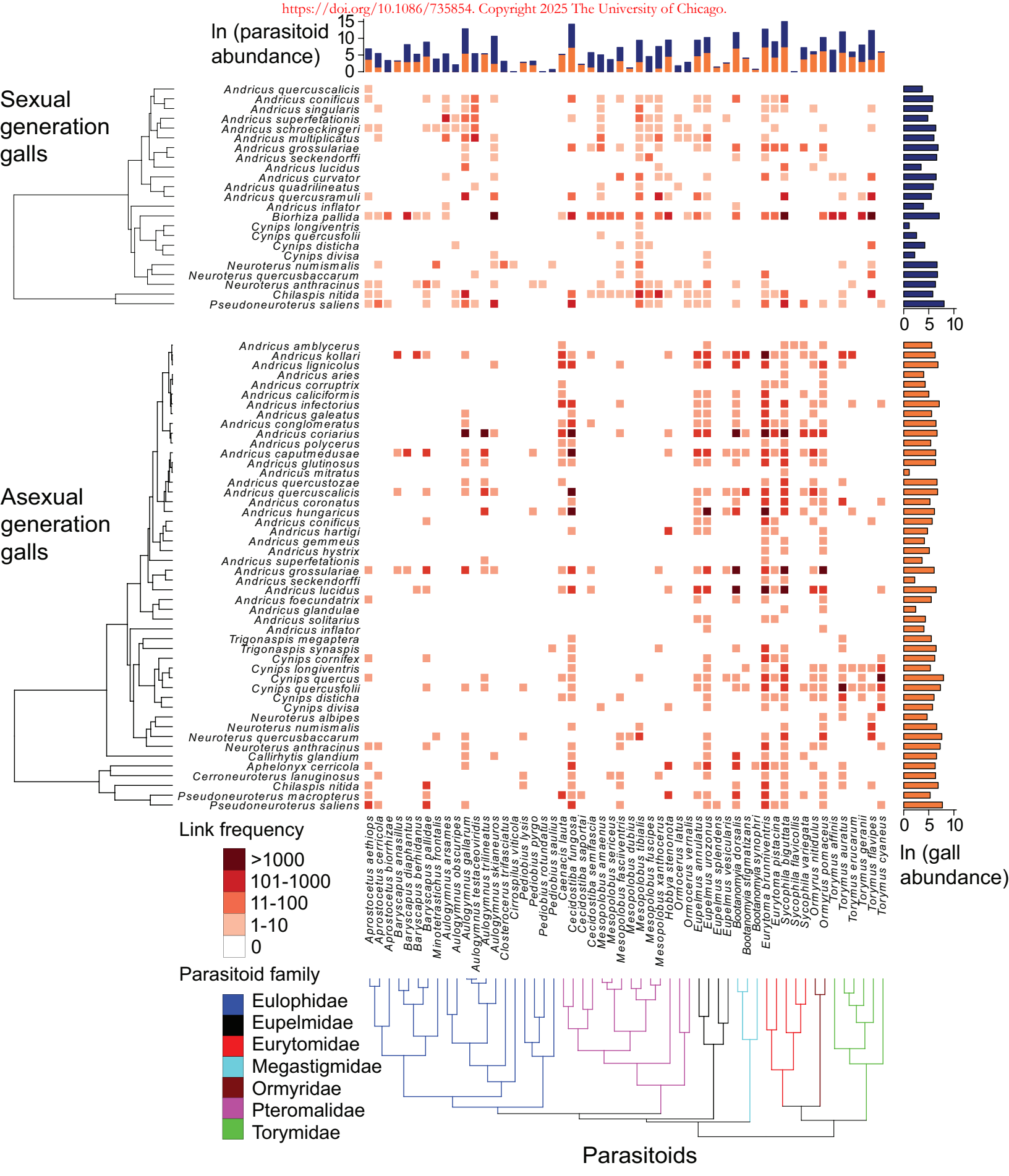


Figure 4

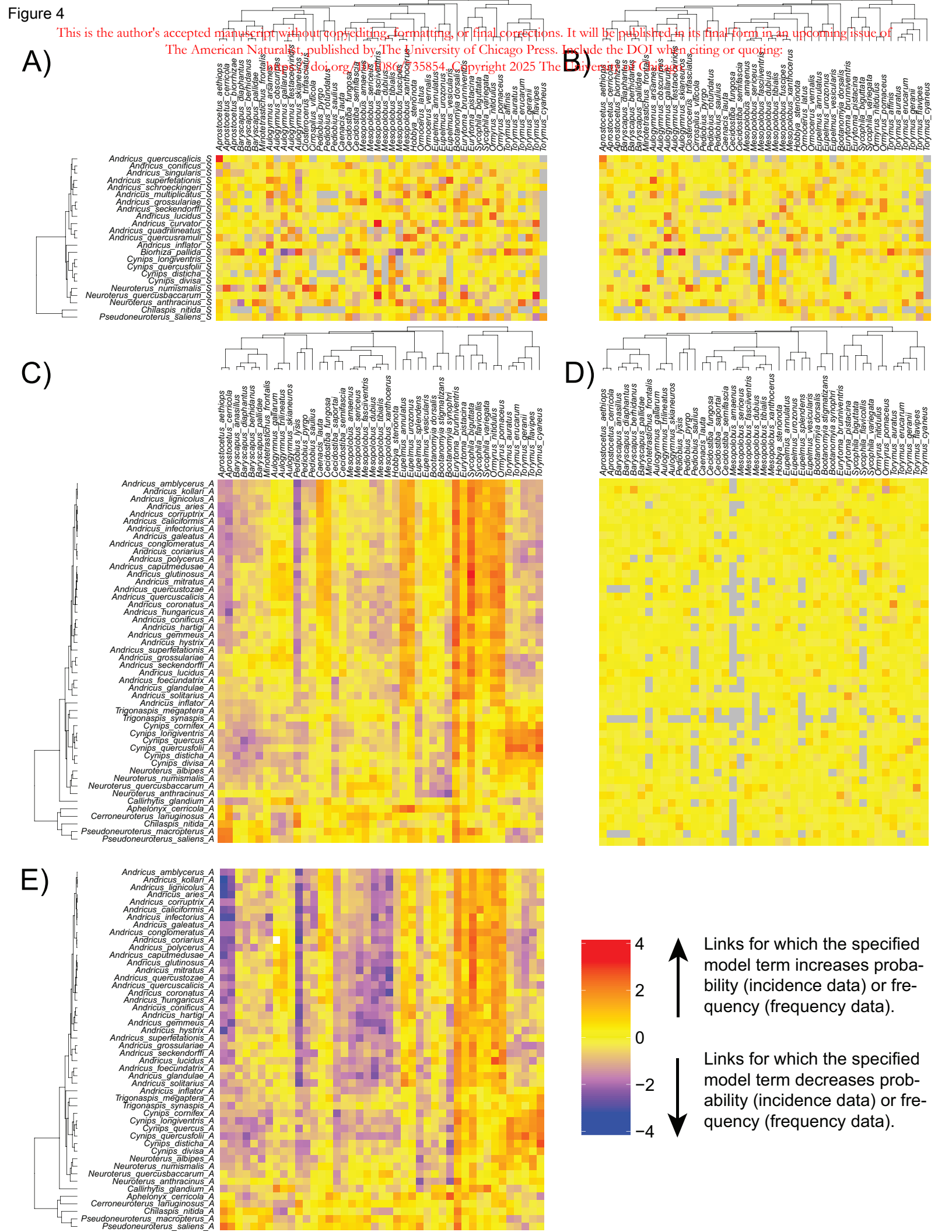
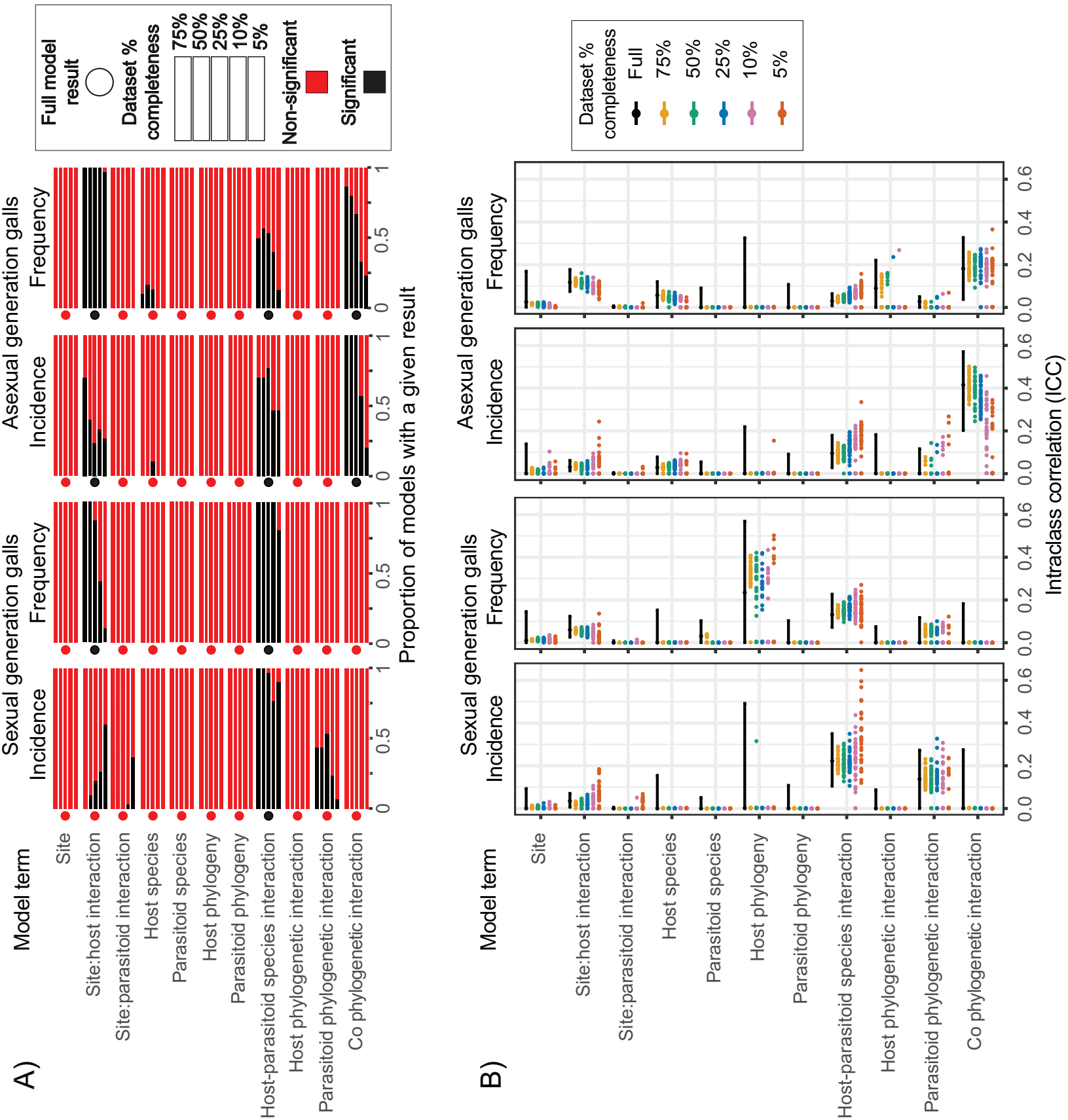


Figure 5



Supplemental PDF for

Phylogenetic and nonphylogenetic patterns in the richness, frequency, and identity of links in a herbivore-parasitoid interaction network

Frazer H. Sinclair,[1,2] Chang-Ti Tang,[1] Richard I. Bailey,[3] György L. Csóka,[4] George Melika,[5] James A. Nicholls,[6] José-Luis Nieves-Aldrey,[7] Alex Reiss,[1] Y. Miles Zhang,[1,8] Albert B. Phillimore,[1] Karsten Schönrogge,[9] & Graham N Stone[1]

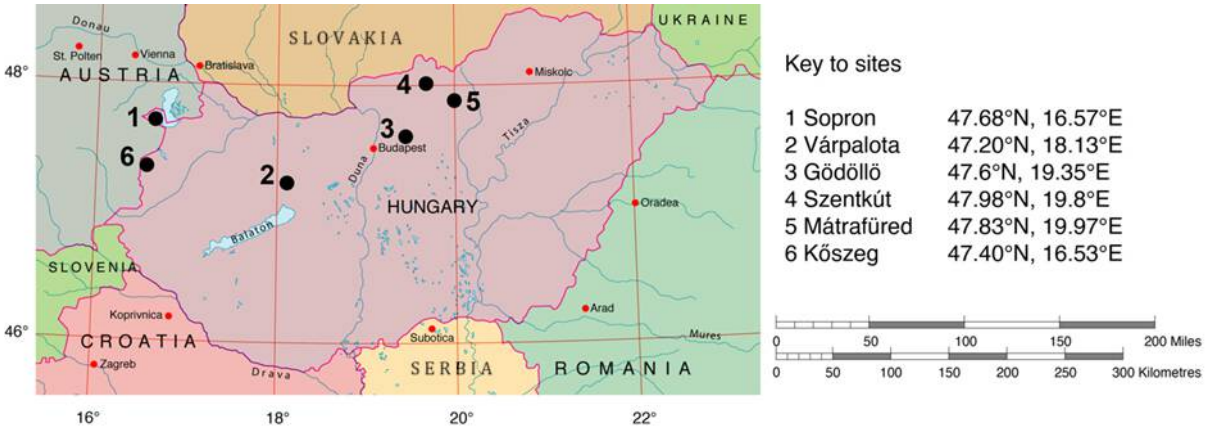
1. Institute of Ecology and Evolution, University of Edinburgh, Charlotte Auerbach Road, Edinburgh, EH9 3FL, Scotland, United Kingdom;
2. Current address: Tropical Biology Association, David Attenborough Building, Pembroke Street, Cambridge, CB2 3QZ, United Kingdom.
3. Department of Ecology and Vertebrate Zoology, University of Lodz, Faculty of Biology and Environmental Protection, 12/16 Banacha St., 90-237 Łódź, Poland;
4. University of Sopron, Forest Research Institute, 3232 Mátrafüred, Hungary;
5. Plant Health Diagnostic National Reference Laboratory, National Food Chain Safety Office, Budaörsi u. 141-145, 1118 Budapest, Hungary;
6. CSIRO, Australian National Insect Collection, Clunies Ross Street, Acton, ACT 2601, Australia. Current address: Royal Botanic Garden Edinburgh, Inverleith Row, Edinburgh, United Kingdom;
7. Museo Nacional de Ciencias Naturales (CSIC), Departamento de Biodiversidad y Biología Evolutiva, C/ José Gutiérrez Abascal 2, ES-28006 Madrid, Spain;
8. Current address: Daniel K. Inouye U.S. Pacific Basin Agricultural Research Center, Hilo, Hawaii, USA.
9. UK Centre for Ecology & Hydrology, Maclean Building, Benson Lane, Wallingford, OX10 8BB, UK

Corresponding author: graham.stone@ed.ac.uk

Contents

Supplementary Methods

Figure S1. Sample site locations.



Section S1 Taxonomic resolution

The Linnaean (morphology-based) taxonomy of both trophic levels in our datasets is well established (Richard R. Askew et al. 2013; Roskam 2019). However, since our datasets were generated, cryptic taxa have been demonstrated in four of the parasitoid species: Bootanomyia (Megastigmus) dorsalis and B. synophri (Nicholls et al. 2018; 2010), and Torymus cyaneus and T. flavipes (Kaartinen et al. 2010; Gil-Tapetado et al. 2022). All four Linnaean species attack multiple host galls, and B. dorsalis and T. flavipes are two of the more generalist parasitoids in Western Palaearctic oak cynipid galls (Richard R. Askew et al. 2013)). The impact of lumping cryptic species on our inference depends critically on how links recorded for the Linnaean species are divided among them. If cryptic sister parasitoid species attack similar sets of related host galls, we would expect increased support for the Cophylogenetic interaction. DNA barcode-based analyses have shown cryptic species within Bootanomyia (Megastigmus) dorsalis and T. flavipes to still be generalists (Kaartinen et al. 2010; Nicholls et al. 2018; 2010). In the absence of directional predictions for potential errors associated with cryptic taxa, we interpret our results at face value while noting that consideration of cryptic taxa is an issue in all trophic network analyses (Pringle and Hutchinson 2020).

Table S1. Sampled gall-types in Spring (S) and Autumn (A) datasets, with details of the number of galls reared, the number of parasitoid individuals that emerged from the reared galls, and the number of different parasitoid links.

Cynipid species and generation	# galls	# parasitoids	# links
Sexual generation dataset			
Andricus conificus (S) (Hartig, 1843)	340	113	18
Andricus curvator (S) Hartig, 1840	644	146	12
Andricus grossulariae (S) Giraud, 1859	955	224	16
Andricus inflator (S) (Hartig, 1840)	51	2	2
Andricus lucidus (S) (Hartig, 1843)	32	85	4
Andricus multiplicatus (S) Giraud, 1859	418	315	14
Andricus quadrilineatus (A) (Hartig, 1840)	368	6	4
Andricus quercuscalicis (S) (Burgsdorf 1783)	41	3	1
Andricus quercusramuli (S) (Linnaeus, 1761)	248	1940	14
Andricus schroeckingeri (S) Wachtl, 1876	600	64	18
Andricus seckendorffi (S) Wachtl (=*A. vindobonensis* Müllner, 1901)	721	27	7
Andricus singularis (S) Mayr, 1870	299	95	14
Andricus superfetationis (S) (Giraud, 1859) (=*A. crispator* Tschek, 1871)	122	261	12
Biorhiza pallida (S) (Olivier, 1791)	1185	12508	29
Chilaspis nitida (S) (Giraud, 1882)	330	1297	23
Cynips disticha (S) Hartig, 1840	67	33	5
Cynips divisa (S) Hartig, 1840	9	6	3
Cynips longiventris (S) Hartig, 1840	3	2	1
Cynips quercusfolii (S) (Linnaeus, 1758)	13	3	2
Dryocosmus cerriphilus (S) (Giraud, 1859)	2	0	0
Neuroterus albipes (S) (Schenck, 1863)	6	0	0
Neuroterus anthracinus (S) (Curtis, 1838)	565	195	18
Neuroterus numismalis (S) Geoffroy in Fourcroy, 1785	777	121	11
Neuroterus politus (S) Hartig, 1840	8	0	0
Neuroterus quercusbaccarum (S) (Linnaeus, 1758)	846	129	5
Pseudoneuroterus saliens (S) (Kollar, 1857)	3141	635	22
Asexual generation dataset			
Andricus amblycerus (A) (Giraud, 1859)	271	18	6
Andricus aries (A) (Giraud, 1859)	54	2	2
Andricus caliciformis (A) (Giraud, 1859)	151	45	7
Andricus caputmedusae (A) (Hartig, 1843)	592	478	20
Andricus conglomeratus (A) (Giraud, 1859)	624	124	12
Andricus conificus (A) (Hartig, 1843)	293	52	7
Andricus coriarius (A) (Hartig, 1843)	775	2332	17
Andricus coronatus (A) (Giraud, 1859)	189	129	9
Andricus corruptrix (A) (Schlechtendal, 1870)	72	8	5
Andricus dentimitratus (A) (Rejtő, 1887)	3	0	0
Andricus foecundatrix (A) (Hartig, 1840)	243	9	5
Andricus galeatus (A) (Giraud, 1859)	260	32	8

Andricus gemmeus (A) (Giraud, 1859)	62	7	3
Andricus glandulae (A) (Hartig, 1840)	11	2	2
Andricus glutinosus (A) (Giraud, 1859)	576	108	11
Andricus grossulariae (A) Giraud, 1859	435	1292	20
Andricus hartigi (A) (Hartig, 1843)	116	36	7
Andricus hungaricus (A) Hartig, 1843	462	607	13
Andricus hystrix (A) Kieffer, 1897	166	8	3
Andricus infectorius (A) (Hartig, 1843)	1197	176	12
Andricus inflator (A) (Hartig, 1840)	57	4	2
Andricus kollari (A) (Hartig, 1843)	532	473	20
Andricus lignicolus (A) (Hartig, 1840)	928	251	12
Andricus lucidus (A) (Hartig, 1843)	647	662	15
Andricus mitratus (A) (Mayr, 1870)	3	1	1
Andricus paradoxus (A) (Radoszkowski, 1866)	31	0	0
Andricus polycerus (A) (Giraud, 1859)	219	17	4
Andricus quercuscalicis (A) (Burgsdorf 1783)	858	414	16
Andricus quercustozae (A) (Bosc, 1792)	769	54	7
Andricus seckendorffi (A) Wachtl	9	6	2
Andricus solitarius (A) (Fonscolombe, 1832)	75	16	4
Andricus superfetationis (A) (Giraud, 1859)	39	2	2
Andricus truncicolus (A) (Giraud, 1859)	1	0	0
Aphelonyx cerricola (A) (Giraud, 1859)	532	153	13
Callirhytis glandium (A) (Giraud, 1859)	697	61	8
Cerroneuroterus lanuginosus (A) (Giraud, 1859)	574	37	11
Chilaspis nitida (A) (Giraud, 1882)	959	88	11
Cynips cornifex (A) Hartig, 1843	492	49	8
Cynips disticha (A) Hartig, 1840	426	43	12
Cynips divisa (A) Hartig, 1840	316	40	7
Cynips longiventris (A) Hartig, 1840	201	61	12
Cynips quercus (A) (Fourcroy, 1785)	2800	438	17
Cynips quercusfolii (A) (Linnaeus, 1758)	1519	363	19
Dryocosmus cerriphilus (A) (Giraud, 1859)	8	0	0
Neuroterus albipes (A) (Schenck, 1863)	106	6	3
Neuroterus anthracinus (A) (Curtis, 1838)	1438	77	12
Neuroterus numismalis (A) Geoffroy in Fourcroy, 1785	726	35	5
Neuroterus quercusbaccarum (A) (Linnaeus, 1758)	2001	97	11
Pseudoneuroterus macropterus (A) (Hartig, 1843)	195	181	13
Pseudoneuroterus saliens (A) (Kollar, 1857)	2247	117	14
Trigonaspis megaptera (A) (Panzer, 1801)	246	2	2
Trigonaspis synaspis (A) (Hartig, 1841)	644	22	7

Table S2. Alphabetic list of recorded parasitoid species, with details of the number of individuals and the number of different host gall-type links from each host gall generation. Parasitoid family names are abbreviated as follows: Eulophidae (Eul), Eupelmidae (Eup), Eurytomidae (Eury), Megastigmidae (Meg), Ormyridae (Orm), Pteromalidae (Pter), Torymidae (Tor). Species previously placed in the genus *Megastigmus* are here given their currently recognised names in the genus *Bootanomyia*. Since the datasets in our study were generated, molecular studies have shown that three of the morphologically defined parasitoid species in our data (*Megastigmus dorsalis, Torymus cyaneus and T. flavipes*) comprise sets of cryptic sister taxa (Kaartinen et al. 2010; Nicholls et al. 2010, 2018). Our analyses retain the original morphological classification as most of the original specimens are unavailable for sequence-based verification.

Parasitoid species	Family	Sexual generation		Asexual generation	
		# individuals	# links	# individuals	# links
Aprostocetus aethiops (Zetterstedt, 1838)	Eul	23	8	41	9
Aprostocetus biorrhizae (Szelényi, 1941)	Eul	28	2	0	0
Aprostocetus cerricola (Erdös, 1954)	Eul	59	8	4	3
Aulogymnus arsames (Walker, 1838)	Eul	209	8	0	0
Aulogymnus gallarum (Linnaeus, 1761)	Eul	1498	11	252	15
Aulogymnus obscuripes (Mayr, 1877)	Eul	8	5	0	0
Aulogymnus skianeuros (Ratzeburg, 1844)	Eul	3399	11	12	5
Aulogymnus testaceoviridis (Erdös, 1961)	Eul	226	8	0	0
Aulogymnus trilineatus (Mayr, 1877)	Eul	0	0	203	10
Baryscapus anasillus Graham, 1991	Eul	0	0	25	4
Baryscapus berhidanus Erdös, 1954	Eul	9	2	21	2
Baryscapus diaphantus (Walker, 1939)	Eul	167	1	20	2
Baryscapus pallidae Graham, 1991	Eul	66	6	106	12
Bootanomyia dorsalis (Fabricius, 1798)	Meg	108	7	1148	23
Bootanomyia stigmatizans (Fabricius, 1798)	Meg	0	0	57	6
Bootanomyia synophri (Mayr, 1874)	Meg	0	0	2	1
Caenacis lauta (Walker, 1835)Walker	Pter	1	1	180	18
Cecidostiba fungosa Geoffr. in Fourcroy, 1785	Pter	1019	7	1485	30
Cecidostiba saportai Graham, 1984	Pter	0	0	9	1
Cecidostiba semifascia (Walker, 1835)	Pter	77	3	4	4
Cirrospilus viticola (Rondani, 1877)	Pter	1	1	0	0
Closterocerus trifasciatus Westwood, 1833	Pter	23	1	0	0
Eupelmus annulatus Nees, 1834	Eup	107	8	118	22
Eupelmus splendens Giraud	Eup	0	0	4	1
Eupelmus urozonus Dalman, 1820	Eup	90	13	312	31
Eupelmus vesicularis (Retzius, 1783)	Eup	1	1	14	6
Eurytoma brunniventris Ratzeburg, 1852	Eury	210	12	1631	40
Eurytoma pistacina Rondani, 1877	Eury	101	6	78	15
Hobbya stenonota (Ratzeburg, 1848)	Pter	128	4	101	8
Mesopolobus amaenus (Walker, 1834)	Pter	155	12	1	1
Mesopolobus dubius (Walker, 1834)	Pter	1	1	3	1
Mesopolobus fasciiventris Westwood, 1833	Pter	53	8	27	9
Mesopolobus fuscipes (Walker, 1834)	Pter	94	9	0	0

Mesopolobus sericeus (Förster, 1770	Pter	38	2	1	1
Mesopolobus tibialis (Westwood, 1833)	Pter	613	20	20	3
Mesopolobus xanthocerus (Thomson, 1878)	Pter	638	13	3	3
Minotetrastichus frontalis (Nees, 1834)	Eul	45	3	1	1
Ormocerus latus Walker, 1834	Pter	6	5	0	0
Ormocerus vernalis Walker, 1834	Pter	17	6	0	0
Ormyrus nitidulus (Fabricius, 1804)	Orm	2	2	202	19
Ormyrus pomaceus (Geoffr. in Fourcroy, 1785)	Orm	56	5	507	33
Pediobius lysis (Walker, 1839)	Eul	0	0	17	4
Pediobius pyrgo (Walker, 1839)	Eul	3	1	8	2
Pediobius rotundatus (Fonscolombe, 1832)	Eul	1	1	0	0
Pediobius saulius (Walker, 1839)	Eul	2	1	1	1
Sycophila biguttata (Swederus, 1795)	Eury	1927	9	1717	39
Sycophila flavicollis (Walker, 1834)	Eury	0	0	1	1
Sycophila variegate (Curtis, 1831)	Eury	31	2	44	11
Torymus affinis (Fonscolombe, 1832)	Tor	629	2	0	0
Torymus auratus (Müller, 1764)	Tor	465	6	330	15
Torymus cyaneus Walker, 1847	Tor	1	1	373	11
Torymus erucarum (Schrank, 1781)	Tor	4	1	91	5
Torymus flavipes (Walker, 1833)	Tor	5730	9	40	7
Torymus geranii (Walker, 1833)	Tor	141	2	21	4

Section S2 Detailed molecular and phylogenetic methods

(a) *Sequence generation and alignment*

Phylogenetic relationships were reconstructed using fragments of one mitochondrial coding gene [698bp of cytochrome c oxidase subunit I (COI) for parasitoids, 433 bp of cytochrome b (cytb) for gallwasps] and a 560–580 bp long fragment of the non-coding 28S D2 region (both trophic levels).

Parasitoid COI sequences were amplified using primers COI_pf1: 5’ AGG RGY YCC WGA TAT AGC WTT YCC 3’ (designed by James Nicholls) and COI_2413d: 5’ GCT ADY CAI CTA AAA ATY TTR ATW CCD GT 3’ (modified from primer C1-J-2441 in Simon et al. 1994) using PCR conditions in Kaartinen et al. (2010). Gallwasp cytb sequences were amplified using primers CB1/CB2 (Jermiin and Crozier, 1994), using PCR conditions in Stone et al. (2009). For both parasitoids and gallwasps, the 28S D2 region was amplified using primers D2F/D2R (Heraty et al. 2004), using PCR conditions in Stone et al. (2009). PCRs were performed in 20µL reactions with the following final concentrations of each reagent: 1x PCR buffer; 2mM $MgCl_2$ for cytb,1.5mM $MgCl_2$ for COI and D2; 0.2µM of each primer; 125µM of each dNTP for COI, 200µM of each dNTP for cytb, and 250µM of each dNTP for D2; 1mg/mL BSA and 0.5U/µL polymerase (Bioline). Post-PCR cleanup used a SAP/ExoI protocol, and fragments were sequenced in both directions using BigDye v3.1 terminator chemistry on an ABI 3731XL capillary sequencer (Life Technologies/Thermofisher). Base calls were confirmed by eye in Sequencher 5.4.6 (Gene Codes Corporation).

Sequences were aligned using MAFFT online (Katoh et al. 2019) with G-INS-i specified as the iterative refinement method. Sequences for the coding COI and cytb mitochondrial genes were unambiguously aligned at 698 bp and 433 bp, respectively. However,

because some previously published COI sequences were not amplified using the same COI_pf1/COI_2413d primer pair, CO1 sequences in the alignment ranged in length from 451 to 698 bp. D2 alignments were edited by removal of indels that appeared in a single specimen and of regions that were difficult to align unambiguously, giving final alignment lengths of 566 bp for gallwasps and 606 bp for parasitoids. Genbank accession numbers for all sequences are provided in table S7.

(b) *Selection of partition substitution models*

Initial substitution models were estimated by analysis of a concatenated data matrix totalling 999 bp for gallwasps and 1304bp for parasitoids using the -m TEST command (Kalyaanamoorthy et al. 2017) in IQ-TREE v2.1.2 (Minh et al. 2020). The concatenated data matrix was partitioned by codon position for COI and cytb, with a separate partition for 28S D2. The best BIC models of respective partitions are listed in table S3.

We used the StarBEAST2 module (Ogilvie et al., 2017) implemented in BEAST2 (Bouckaert et al., 2019) to generate a species tree for each trophic level. Because the substitution models available in StarBEAST2 are limited to JC69, HKY, TN93 and GTR, we adopted the following procedure to select the closest suitable option to the IQ-TREE-inferred model (table S3). First, we reduced the complexity of the rate matrix for specific types of transition and transversion in each partition. We examined the empirical frequencies of transition and transversion types present using the pairwise base frequencies command in PAUP* 4.0 (Swofford 2003), and where specific transition or transversion types had very low frequencies, we assigned them the same rates as other types with higher frequencies, following recommendations in the MrBayes manual v3.2. Second, since the base frequencies parameter F is not an option in StarBEAST2, we excluded it from simplified models (table S3). Third, because partitions for gallwasp cytb codon positions 1 and 2 and parasitoid 28S D2 show a proportion of conserved sites (table S4), we incorporated proportion of invariant sites (I) as an additional parameter in the substitution models for these partitions. These substitution models were used in initial runs in StarBEAST2 with the following specifications: mitochondrial gene ploidy set to 0.5, population model: constant population size, phylogenetic model: birth-death, chain length set to 100 million generations and sampled every 12,500 generations. Because MCMCglmm requires ultrametric phylogenies (see below) we used a strict clock model.

Initial analysis of the 28S D2 gene tree of oak gallwasps resulted in multiple unresolved polytomies, with knock-on effects on the resolution of the two-gene species tree. Because the D2 gene alignment of oak gallwasps has 87.9% identical sites (table S4), we judged that inclusion only of the parameter I for rate heterogeneity across sites might not adequately account for between-site variability. We therefore respecified the substitution model for this gene to include gamma-distributed rate variation (G), with 4 gamma categories. This model yielded a better-resolved bifurcating tree, and the G parameter was therefore retained for downstream analyses. The final substitution models and rate parameters for all genes and partitions are shown in table S3.

(c) *Imposition of phylogenetic constraints and clock calibration*

Initial analysis of the parasitoid COI data found the family Eulophidae to be paraphyletic, while the D2 gene tree for the same taxa supports monophyly. This discordance of gene tree topologies results in Eulophidae being paraphyletic in the species tree. Because more in-depth phylogenetic analyses support monophyly of Eulophidae (e.g. Burks et al. 2011; Munro et al. 2011; Rasplus et al. 2020), we constrained Eulophidae to be monophyletic in downstream analyses (see descriptions of monophyly settings below). All other parasitoid families in our dataset were recovered as monophyletic.

We calibrated molecular clocks for the gallwasp and parasitoid species trees to identify the timescales involved for phylogenetic patterns inferred in MCMCglmm. Node ages were calibrated in StarBEAST2 using estimates of node ages and 95% confidence intervals (CI) inferred in previous more extensive analyses by Peters et al. (2018) for parasitoids and by Blaimer et al. (2020) for gallwasps, following guidance for most recent common ancestor (MRCA) prior specification and setting of monophyletic group constraints in the Taming the BEAST online resource (https://taming-the-beast.org/tutorials/Introduction-to-BEAST2/; Barrido-Sottani et al. 2018). To allow date calibration using shared nodes in Peters et al. (2018) for the parasitoid species tree, we selected an MRCA prior and imposed monophyly for Eulophidae, Eupelmidae, Eurytomidae and Torymidae. To allow date calibration using shared nodes in Blaimer et al. (2020) for the oak gallwasp species tree, the same MRCA prior and monophyletic group constraint approach was applied to species in the lucidus clade (*Andricus grossulariae*, *A. lucidus*, *A. seckendorffi*) as defined in Stone et al. (2009), and to all species of oak gallwasps. A lognormal distribution (Ho and Phillips 2009) was applied for node age inference. The specific shared node ages in Peters et al. (2018) used in our clock calibration were node 18 (86 mya; CI: 62–146 mya) for the MRCA of parasitoids in our dataset, node 25 (40 mya; CI: 21–71 mya) for Torymidae, node 28 (48 mya; CI: 25–86 mya) for Eupelmidae, and node 36 (23 mya; CI: 11–40 mya) for Eurytomidae. The specific shared node ages in Blaimer et al. (2020) for gallwasps were node 129 (75.6 mya; CI: 50.5–106.9 mya) for Cynipini, and node 150 (17.2 mya; CI: 6.3–30.1 mya) for the lucidus clade. In StarBEAST2 monophyletic group priors, the minimum age CI (see above) was entered as the offset value for each monophyletic group, M was entered as the mean difference between the maximum CI and the minimum CI, and S was tuned such that the 2.5%–97.5% interquantile range matched the estimated CI values in Peters et al. (2018) and Blaimer et al. (2020).

(d) *Selection of appropriate population and clock models*

Our MCMCglmm analyses require an ultrametric tree of each trophic level, which requires a strict clock model. To assess the suitability of such an assumption we compared the strict clock model to alternatives for the parasitoid and gallwasp sequence data. We used nested sampling (Skilling, 2006; Maturana et al., 2018) to estimate relative support for the best combination of the two population models (Analytical Population Size Integration *versus* Constant populations) and clock models (Strict Clock *versus* Uncorrelated Lognormal) available in BEAST2, following the online instructions in Taming the BEAST (Barido-Sottani et al., 2018; https://taming-the-beast.org/tutorials/NS-tutorial/). We first set up the substitution models for each partition as well as the values for time calibration, as described above, and ran models with the four possible population and clock model combinations, with chainLength=500,000, particleCount=1 and subChainLength=5,000. The difference in log of marginal likelihood between models was used to calculate Bayes Factors (BF) following Kass and Raftery (1995). We first compared the log of marginal likelihood between models to find the best two models, and then further compared the best two models using BF. Where the BF $< 2*\sqrt{(SD1^2+SD2^2)}$ between the best two models, we increased the particleCount for a longer sampling, following the aforementioned online instructions for nested sampling in Taming the BEAST (Barido-Sottani et al., 2018).

For the gallwasp data, the alternative population models combined with the Uncorrelated Lognormal clock model are indistinguishable from each other, but both have higher log marginal likelihood than either of the two population models combined with the Strict Clock model. In parasitoids, the two clock models combined with Analytical Population Size Integration population model could not be differentiated from each other but were better than the two clock models combined with Constant populations. We carried out additional longer analyses to further resolve log marginal likelihood estimates for the two better supported

model combinations for each dataset (two population models with Uncorrelated Lognormal clock model for gallwasps, and the combinations of two clock models with Analytical Population Size Integration population model for parasitoids). We used the formula SD = sqrt(H/N) where SD is the standard deviation of the log marginal likelihood estimate for a model, H was taken from the information value in the first run of the model combination with the best log of marginal likelihood, and N is the number of particles (Barido-Sottani et al., 2018). The target SD values using this approach were 6 and 5 for gallwasps and parasitoids respectively, which made N = 23 in both trophic levels for further model selection processes. chainLength and subChainLength were as for the initial runs. Model comparison using the procedure above then favored Analytical Population Size Integration + Uncorrelated Lognormal for gallwasps (table S5), and Analytical Population Size Integration + Strict Clock for parasitoids (table S6).

In the final phylogenetic analyses we applied all the estimated models and divergence times for species tree inference in StarBEAST2. Analyses were run for 100 million generations and sampled every 12,500 generations for both parasitoids and gallwasps. Data convergence was examined in Tracer 1.7.2 (Rambaut et al., 2018) using the log file. The maximum clade credibility trees used in the MCMCglmm analyses were generated from the last 20% of generations using common ancestor height in TreeAnnotator (included in BEAST 2.6.6), and the tree was visualised in FigTree 1.4.4 (https://github.com/rambaut/figtree/releases).

Table S3. Substitution models for each gene and codon position estimated in IQ-TREE, and the simplified nearest equivalents used in StarBEAST2 analyses. N/A for partition indicates that a gene fragment was treated as a single partition.

Locus	Alignment length (bp)	Codon position	IQ-TREE	StarBEAST2
COI in parasitoids	698	1	TIM3+F+I+G4	HKY+I+G
		2	TVM+F+I	TN93+I
		3	TIM2+I+G4	HKY+G
Cytb in gallwasps	433	1	TIM+F+G4	HKY+I+G
		2	HKY+F+G4	HKY+I+G
		3	TIM3+F+G4	TN93+G
D2 in parasitoids	606	N/A	TIM3e+G4	TN93+I+G
D2 in gallwasps	566	N/A	TNe+I	TN93+I+G

Table S4. Partition sizes by genes and codon position, and the number of conserved and variable sites within each partition.

Locus	Codon position	Length (bp)	Conserved	Variable	% Invariant sites
COI in parasitoids	1	232	154	78	66.3
	2	233	199	34	85.4
	3	233	11	222	4.7
Cytb in gallwasps	1	144	98	46	68.0
	2	144	121	23	84.0
	3	145	18	127	12.4
D2 in parasitoids	N/A	606	426	180	70.2
D2 in gallwasps	N/A	566	498	68	87.9

Table S5. Model log marginal likelihoods used in Bayes Factor (BF)-based model selection for the gallwasp sequence dataset. In each run, the models are ranked by log marginal likelihood. Support for the top two ranked models is compared using BF, with the assessment of meaningful difference based on the BF ≥ 2*(SD1+SD2) where SD1 and SD2 are the standard deviations of the log marginal likelihoods for the 2 models being compared. N is the number of particles and H is the information content.

	Model	N	H	SD	log marginal likelihood	BF (ML1−ML2)	$2*\sqrt{(SD1^2+SD2^2)}$	Model selection result
First run	Analytical Population Size Integration + Uncorrelated clock lognormal	1	818.9	28.6 (SD1)	-7545.1	38.6	83.62	BF < $2*\sqrt{(SD1^2+SD2^2)}$ Models are indistinguishable
	Constant populations + Uncorrelated clock lognormal	1	928.3	30.5 (SD2)	-7583.7			
	Analytical Population Size Integration + Strict clock	1	834.3	28.9 (SD3)	-7789.5			
	Constant populations + Strict clock	1	874.8	29.6 (SD4)	-7683.1			
Second run	Analytical Population Size Integration + Uncorrelated clock lognormal	23	643.5	5.3 (SD1)	-7307.5	167.9	15.86	BF > $2*\sqrt{(SD1^2+SD2^2)}$ Analytical Population Size Integration + Uncorrelated lognormal clock is overwhelming supported
	Constant populations + Uncorrelated clock lognormal	23	812.0	5.9 (SD2)	-7475.3			

Table S6. Model log marginal likelihoods used in Bayes Factor (BF)-based model selection for the parasitoid sequence dataset. In each run, the models are ranked by log marginal likelihood. Support for the top two ranked models is compared using BF, with the assessment of meaningful difference based on the BF ≥ 2*(SD1+SD2) where SD1 and SD2 are the standard deviations of the log marginal likelihoods for the 2 models being compared. N is the number of particles and H is the information content.

	Model	N	H	SD	log marginal likelihood	BF (ML1–ML2)	$2*\sqrt{(SD1^2+SD2^2)}$	Model selection result
First run	Analytical Population Size Integration + Strict clock	1	556.2	23.6 (SD1)	-15134.1 (ML1)	11.0	65.63	$BF < 2*\sqrt{(SD1^2+SD2^2)}$ Models are indistinguishable
	Analytical Population Size Integration + Uncorrelated clock lognormal	1	519.4	22.8 (SD2)	-15145.1 (ML2)			
	Constant populations + Strict clock	1	456.3	21.4 (SD3)	-15352.28 (ML3)			
	Constant populations + Uncorrelated clock lognormal	1	517.7	30.5 (SD4)	-15311.66 (ML4)			
Second run	Analytical Population Size Integration + Strict clock	23	441.6	4.4 (SD1)	-15002.1 (ML1)	41.5	12.44	$BF > 2*\sqrt{(SD1^2+SD2^2)}$ Analytical Population Size Integration+Strict clock is overwhelmingly supported
	Analytical Population Size Integration + Uncorrelated clock lognormal	23	455.3	4.4 (SD2)	-15043.6 (ML2)			

Figure S2. Maximum clade credibility tree for the cynipid gallwasps in this study used in MCMCglmm analyses. Numbers in black indicate node age in millions of years (with 95% posterior credibility interval). To preserve clarity of the figure, some node ages have been excluded. Posterior probability for all nodes is 1.0, with the exception of 3 nodes whose posterior probability is labelled in red, and for very recent divergences in the topmost clade containing *Andricus kollari*. All nodes marked with a red circle are shared with the best supported relaxed clock model, which also shows very similar branch length and node dates (data not shown).

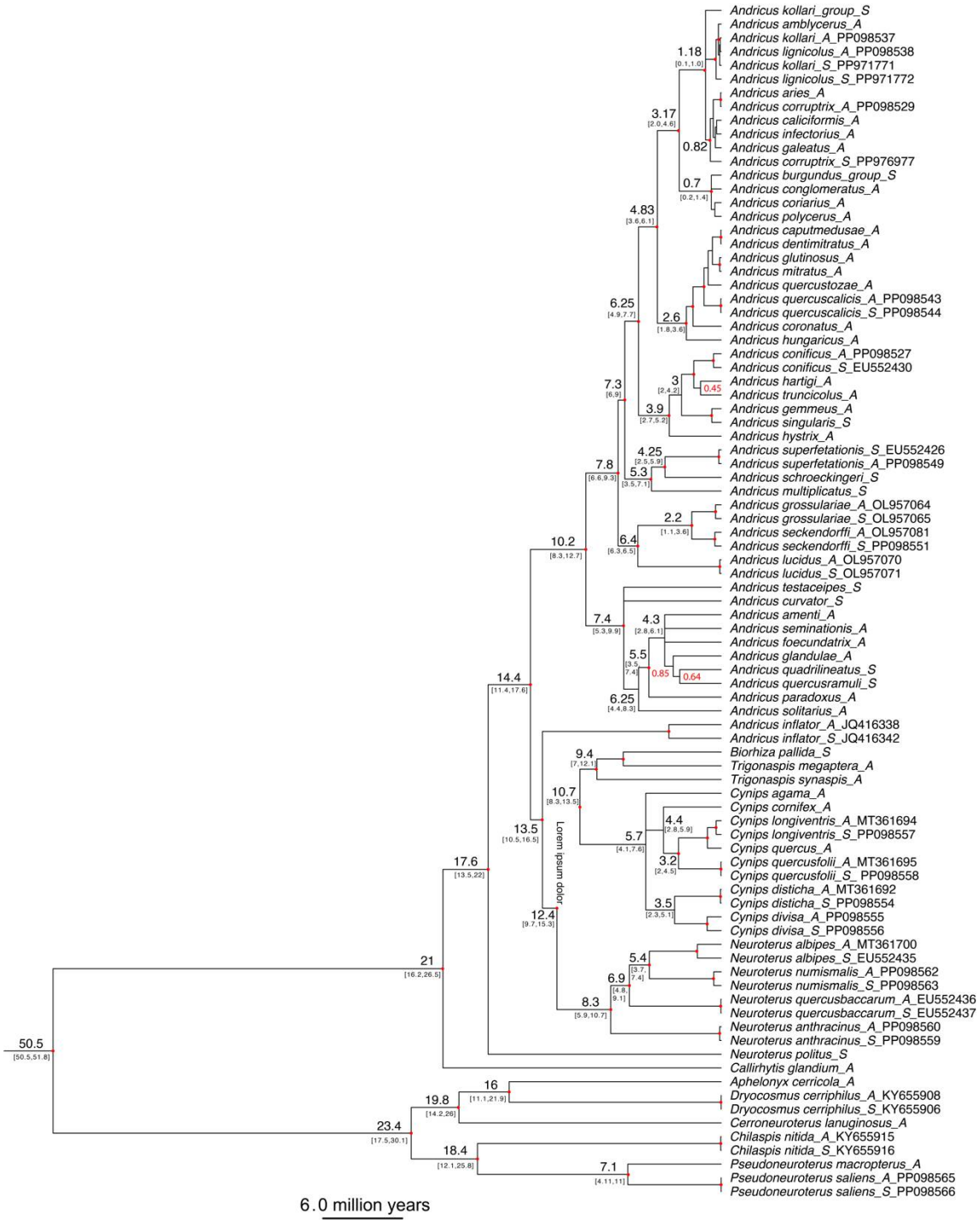


Figure S3. Maximum clade credibility tree for the chalcid parasitoids in this study used ibn MCMCglmm analyses. Numbers in black at selected nodes indicate node age in millions of years (with 95% posterior credibility interval). Our estimates for family common ancestors are close to the youngest inferred in a recent extensive analysis by Cruaud et al. (2024). Posterior probabilities of 0.90-1.0 are labelled with a red asterisk, and of 0.80-0.89 are labelled with a black asterisk. The family Eulophidae were constrained to be monophyletic in the analysis; monophyly was inferred for all other families with high posterior support.

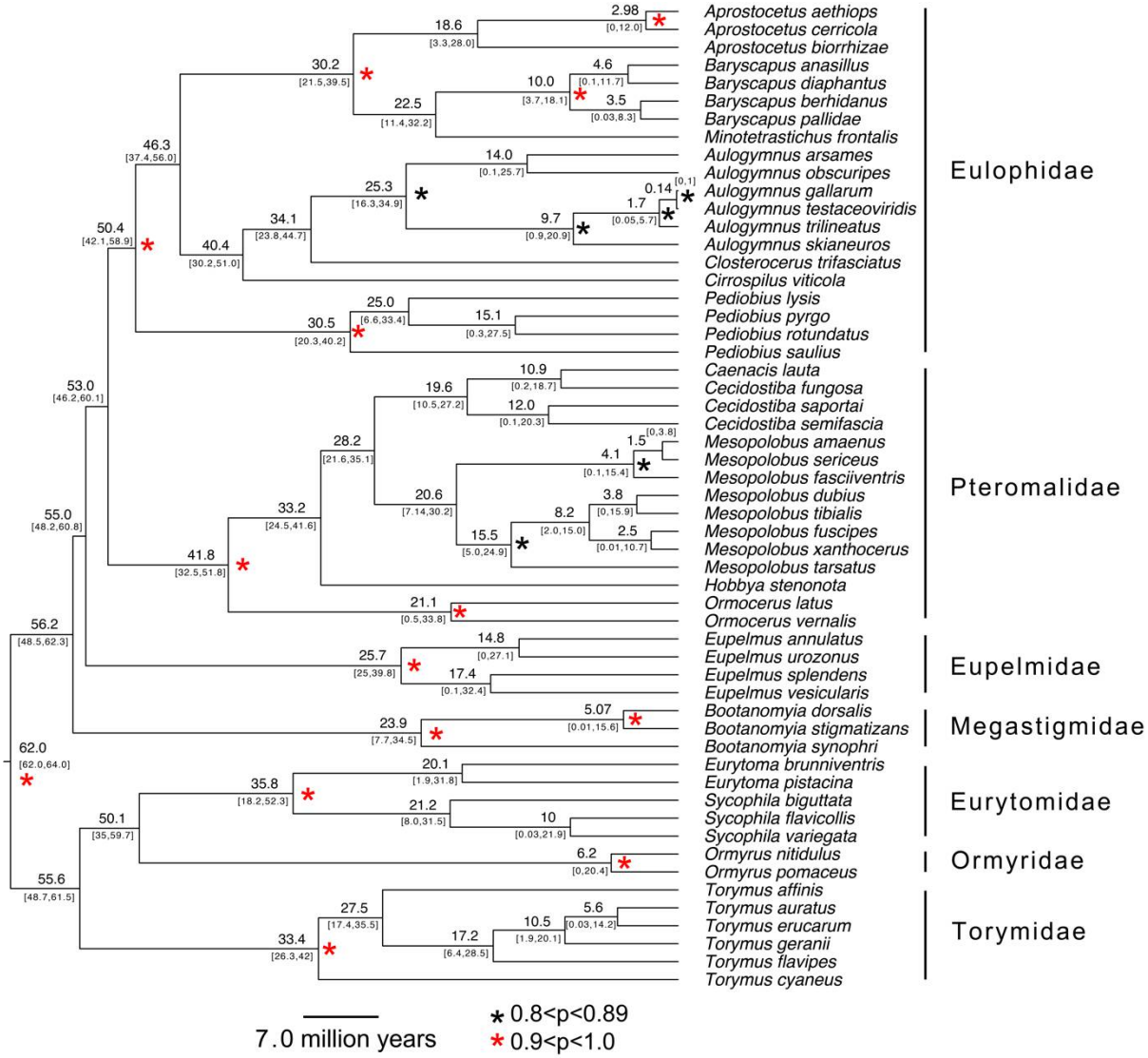


Table S7. GenBank accession numbers and internal specimen codes (where applicable) for (a) cynipid gall inducers and (b) parasitoid species used in phylogeny construction. New sequences are indicated with an asterisk.

(a) Cynipid gall inducers

Gall species	cytb specimen	cytb GB number	D2 specimen	D2 GB number
Andricus amblycerus	Andambl11	PP098522*	Andambl11	PP099067*
Andricus aries	Andari4	PP098523*	NA	DQ201458
Andricus caliciformis	Andcali1	PP098524*	NA	DQ201448
Andricus caputmedusae	Andcap10	PP098525*	NA	DQ201496
Andricus conglomeratus	Andcong3	PP098526*	Andcong3	PP099068*
Andricus conificus	Andcyd2	EU552430	Andcyd2	PP099070*
	Andconi1	PP098527*	NA	DQ201470
Andricus coriarius	Andcori37	PP098528*	NA	EF030037
Andricus coronatus	Andcoro2	EU552425	Andcoro2	EU552440
Andricus corruptrix	Andcorr15	PP098529*	NA	DQ201461
	Andcorr17	PP976977*	NA	DQ201461
Andricus curvator	Andcurv38	JQ416322	Andcurv38	PP099069*
Andricus dentimitratus	Andden8	PP098530*	Andden8	PP099071*
Andricus foecundatrix	Andfec1	PP098531*	Andfec1	EU552442
Andricus galeatus	Andgal1	PP098532*	Andgal1	PP099072*
Andricus gemmeus	Andgem1	EU552431	Andgem1	EU552443
Andricus glandulae	Andgla4	PP098533*	Andgla1	KX683438
Andricus glutinosus	Andglu1	EU552432	Andglu1	EU552444
Andricus grossulariae	Andgro427	OL957065	NA	DQ012582
	Andgro319	OL957064	NA	DQ012582
Andricus hartigi	Andhar3	PP098534*	Andhar3	PP099073*
Andricus hungaricus	Andhun2	PP098535*	Andhun2	PP099074*
Andricus hystrix	Andhys1	EU100370	Andhys1	EU552446
Andricus infectorius	Andinfe119	PP098536*	Andinfe119	PP099075*
Andricus inflator	Andinfl6	JQ416338	Andinfl6	PP099076*
	Andinfl3	JQ416342	Andinfl3	KX683439
Andricus kollari	Andkol37	PP098537*	Andkol87	PP099077*
	Andkol98	PP971771*	Andkol98	PP962513*
Andricus lignicolus	Andlig15	PP098538*	Andlig15	PP099078*
	Andlig129	PP971772*	Andlig129	PP962514*
Andricus lucidus	Andluc99	OL957071	Andluc5	PP099080*
	Andluc98	OL957070	Andluc2	PP099079*
Andricus mitratus	Andmit2	PP098539*	Andmit2	PP099081*
Andricus multiplicatus	Andmul70	PP098540*	Andmul20	PP099082*
Andricus paradoxus	Andpara5	PP098541*	Andpara5	PP099083*
Andricus polycerus	Andpol8	PP098542*	Andpol1	EU552449
Andricus quadrilineatus	Andqua4	PP098545*	Andqua3	PP099087*
Andricus quercuscalicis	Andqca1361	PP098544*	Andqca1361	PP099085*
	Andqca1359	PP098543*	Andqca1359	PP099084*
Andricus quercusramuli	Andqra60	JQ416374	Andqra60	PP099086*
Andricus quercustozae	Andqtz104	EU552433	Andqtz104	EU552445
Andricus schroeckingeri	Andschr3	PP098546*	Andschr3	PP099088*
Andricus seckendorffi	Andvin1	PP098551*	Andsec30	PP099089*
	Andsec58	OL957081	Andsec59	PP099090*
Andricus singularis	Andsin22	PP098547*	Andsin22	PP099091*
Andricus solitarius	Andsol1	PP098548*	Andsol1	PP099092*
Andricus superfetationis	Andsup1	PP098549*	Andsup1	PP099093*
Andricus truncicolus	Andtru2	PP098550*	NA	DQ201464
Aphelonyx cerricola	Aphcer20	PP098552*	Aphcer20	PP099094
Biorhiza pallida	Biopal1637	PP098553*	Biopal1	MZ969835
Callirhytis glandium	Calgla1	MT152286	Calgla1	PP099095*
Cerroneuroterus lanuginosus	Neulan7	PP098561*	Neucer1	PP099112*
Chilaspis nitida	Chinit2	KY655915	Chinit2	PP099096*
	Chinit6	KY655916	Chinit6	PP099097*
Cynips cornifex	Cyncorn1	MT361691	Cyncorn1	PP099098*
Cynips disticha	Cyndis6	PP098554*	Cyndis6	PP099100*
	Cyndis4	MT361692	Cyndis4	PP099099*
Cynips divisa	Cyndiv35	PP098556*	Cyndiv34	KX683464

	Cyndiv21	PP098555*	Cyndiv25	PP099101*
Cynips longiventris	Cynlon15	PP098557*	Cynlon15	PP099103*
	Cynlon14	MT361694	Cynlon14	PP099102*
Cynips quercus	Cynque20	JQ416464	Cynque20	PP099106*
Cynips quercusfolii	Cynqfo219	PP098558*	Cynqfo219	PP099105*
	Cynqfo154	MT361695	Cynqfo154	PP099104*
Dryocosmus cerriphilus	Drycer2	KY655906	Drycer2	PP099107*
	Drycer4	KY655908	Drycer4	PP099108*
Neuroterus albipes	Neualb3	MT361700	Neualb3	PP099109*
	Neualb1	EU552435	Neualb1	EU552451
Neuroterus anthracinus	Neuant44	PP098560*	Neuant44	PP099111*
	Neuant37	PP098559*	Neuant37	PP099110*
Neuroterus numismalis	Neunum11	PP098562*	Neunum11	PP099113*
	Neunum5	PP098563*		
Neuroterus politus	Neupol2	PP098564*	Neupol2	PP099114*
Neuroterus quercusbaccarum	Neuqba2	EU552437	Neuqba115	PP099115*
	Neuqba1	EU552436		
Pseudoneuroterus macropterus	Psemac7	PP098567*	Psemac15	PP099117*
Pseudoneuroterus saliens	Neusal161	PP098565*	NA	DQ201483
	Neusal39	PP098566*	Neusal39	PP099116*
Trigonaspis megaptera	Trgmeg1	MT361701	Trgmeg1	PP099118*
Trigonaspis synaspis	Trgsyn205	MT361703	Trgsyn205	PP099119*

(b) Parasitoids

Species	Specimen_code	COI	D2
Aprostocetus aethiops	Aaet46	PP826939*	PP115595*
Aprostocetus biorrhizae	Abio04	PP826462*	PP115597*
Aprostocetus cerricola	Acer11	PP829062*	PP115598*
Aulogymnus arsames	Aars02	PP829061*	PP115596*
Aulogymnus gallarum	Agal189	PP826463*	PP115599*
Aulogymnus obscuripes	Aobs02	PP826464*	PP115600*
Aulogymnus skianeuros	Aski133	PP826465*	PP115601*
Aulogymnus testaceoviridis	Ates21	PP826466*	PP115602*
Aulogymnus trilineatus	Atri130	PP826467*	PP115603*
Baryscapus anasillus	Bana03	PP826468*	PP115604*
Baryscapus berhidanus	Bber15	PP826469*	PP115605*
Baryscapus diaphantus	Bdia09	PP826470*	PP115606*
Baryscapus pallidae	Bpal006	PP829063*	PP115607*
Caenacis lauta	Clau01	PP829065*	PP115610*
Cecidostiba fungosa	Cfun0079	PP829064*	PP115609*
Cecidostiba saportai	Csap05	PP826471*	PP115611*
Cecidostiba semifascia	Csem60	PP826472*	PP115612*
Cirrospilus viticola	Cdia17	PP829066*	PP115608*
Closterocerus trifasciatus	Ctri24	PP826473*	PP115613*
Eupelmus annulatus	Eann85	PP826474*	PP115614*
Eupelmus splendens	Espl01	PP826475*	PP115617*
Eupelmus urozonus	Euro3	PP829067*	PP115618*
Eupelmus vesicularis	Eves1	PP826942*	PP115619*
Eurytoma brunniventris	Ebru2231	PP826940*	PP115615*
Eurytoma pistacina	Epis49	PP826941*	PP115616*
Hobbya stenonota	Hste13	PP829068*	PP115620*
Megastigmus dorsalis	Mdor1445	PP826476*	PP115622*
Megastigmus stigmatizans	Msti0082	PP826477*	PP115628*
Megastigmus synophri	Msyn0028	PP826945*	PP115629*

Mesopolobus amaenus	Mama52	PP826943*	PP115621*
Mesopolobus dubius	Mdub39	PP826944*	PP115623*
Mesopolobus fasciiventris	Mfas32	PP829069*	PP115624*
Mesopolobus fuscipes	Mfus01	PP829070*	PP115626*
Mesopolobus sericeus	Mser70	PP829071*	PP115627*
Mesopolobus tarsatus	Mtar06	PP826946*	PP115630*
Mesopolobus tibialis	Mtib02	PP946230*	PP115631*
Mesopolobus xanthocerus	Mxan10	PP829072*	PP115632*
Minotetrastichus frontalis	Mfro01	PP826478*	PP115625*
Ormocerus latus	Olat04	PP826479*	PP115633*
Ormocerus vernalis	Over03	PP826947*	PP115636*
Ormyrus nitidulus	Onit192	PP829073*	PP115634*
Ormyrus pomaceus	Opom182	PP829074*	PP115635*
Pediobius lysis	Plys06	PP829075*	PP115637*
Pediobius pyrgo	NA	MG836443.1	MH169076.1
Pediobius rotundatus	Prot26	PP829076*	PP115638*
Pediobius saulius	NA	GU087095.1	
Sycophila biguttata	Sbig1340	PP826480*	PP115639*
Sycophila flavicollis	Sfla1	PP826481*	PP115640*
Sycophila variegata	Svar0033	PP826482*	PP115641*
Torymus affinis	Taff13	PP829077*	PP115642*
Torymus auratus	Taur54	PP829078*	PP115643*
Torymus cyaneus	Tcya1	PP829079	PP115644*
Torymus erucarum	Teru26	PP826948*	PP115645*
Torymus flavipes	Tfla5	PP829080*	PP115646*
Torymus geranii	Tger93	PP829081*	PP115647*

Section S3 Supplementary Results

Tables S8 and S9 show intraclass correlations (ICCs) for terms in alternate versions of the 4 core models (incidence and frequency-based models of sexual and asexual generation gall datasets). Each table shows results for the full model and one or two alternate models in which specific model terms are excluded. ICC values are presented as median/mode (and 95% credible interval) over 2000 MCMC point estimates. NA indicates a model term that cannot be fitted in the specified model. For ease of comparison across models, significant terms (i.e. where the lower 95% credible interval >0.01) are highlighted in each model. ICC_{GLMM} is the ICC for all fitted model terms (i.e. excluding the residual), analogous to an R^2 for the model. In the Full model (option 1), interactions between taxa missing from an individual site are omitted as structural zeros. In the Full model (option 2) and Pooled sites models, we assumed that the full set of parasitoid species recorded in a given generation across all six sites was available to interact at each site, and all links absent from a single site and generation were given a value of zero.

Table S8. Sexual generation models.

	Incidence models				Frequency models		
Model term	Full model (option 1)	Without sample size	Full model (option 2)	Pooled sites (option 2)	Full model (option 1)	Full model (option 2)	Pooled sites (option 2)
Site	.014 / .000 (.000-.096)	.028 / .016 (.000-.147)	.012 / .000 (.000-.090)	NA	.027 / .016 (.000-.134)	.092 / .063 (.008-.374)	NA
Site:host interaction	.030 / .031 (.006-.069)	.073 / .066 (.028-.143)	.054 / .046 (.016-.104)	NA	.064 / .055 (.025-.122)	.139 / .119 (.062-.243)	NA
Site:parasitoid interaction	.001 / .000 (.000-.006)	.001 / .000 (.000-.008)	.001 / .000 (.000-.010)	NA	.002 / .000 (.000-.009)	.002 / .000 (.000-.008)	NA
Host species	.031 / .001 (.000-.142)	.020 / .001 (.000-.141)	.046 / .001 (.000-.140)	.055 / .001 (.000-.163)	.024 / .001 (.000-.157)	.055 / .001 (.000-.231)	.073 / .002 (.000-.270)
Parasitoid species	.007 / .000 (.000-.049)	.029 / .001 (.000-.096)	.023 / .000 (.000-.096)	.030 / .001 (.000-.135)	.040 / .039 (.000-.099)	.054 / .047 (.008-.116)	.112 / .084 (.026-.227)
Host phylogeny	.179 / .002 (.000-.471)	.316 / .258 (.000-.595)	.054 / .001 (.000-.351)	.086 / .002 (.000-.424)	.314 / .340 (.000-.579)	.288 / .003 (.000-.621)	.380 / .003 (.000-.688)
Parasitoid phylogeny	.019 / .000 (.000-.106)	.014 / .000 (.000-.098)	.043 / .001 (.000-.162)	.053 / .001 (.000-.212)	.015 / .001 (.000-.099)	.011 / .000 (.000-.081)	.027 / .001 (.000-.158)
Parasitoid:host species interaction	.207 / .224 (.095-.329)	.132 / .138 (.056-.227)	.209 / .211 (.103-.320)	NA	.143 / .159 (.068-.230)	.057 / .064 (.021-.100)	NA
Host phylogenetic interaction	.013 / .000 (.000-.077)	.022 / .001 (.000-.104)	.024 / .001 (.000-.115)	.084 / .002 (.000-.255)	.009 / .000 (.000-.070)	.005 / .000 (.000-.038)	.011 / .001 (.000-.077)
Parasitoid phylogenetic interaction	.133 / .123 (.000-.268)	.072 / .046 (.000-.173)	.143 / .125 (.000-.261)	.185 / .215 (.000-.358)	.047 / .000 (.000-.110)	.027 / .024 (.000-.061)	.053 / .040 (.000-.122)
Cophylogenetic interaction	.089 / .002 (.000-.281)	.065 / .002 (.000-.282)	.052 / .001 (.000-.251)	.054 / .002 (.000-.332)	.038 / .001 (.000-.192)	.012 / .000 (.000-.086)	.021 / .001 (.000-.177)
Residual	.171 / .161 (.097-.248)	.114 / .105 (.060-.178)	.201 / .207 (.121-.275)	.292 / .294 (.128-.457)	.185 / .237 (.052-.299)	.159 / .164 (.032-.254)	.250 / .257 (.092-.407)
ICC denominator	25.157 / 22.228 (16.192-40.403)	37.743 / 32.890 (22.131-68.044)	21.299 / 20.794 (14.473-32.140)	14.694 / 12.945 (8.108-27.657)	30.150 / 27.075 (18.451-49.273)	74.146 / 63.437 (40.433-143.656)	34.182 / 30.792 (18.500-61.745)
ICC_{GLMM}	.829 / .839 (.752-.903)	.886 / .895 (.822-.940)	.799 / .793 (.725-.879)	.708 / .706 (.543-.872)	.815 / .763 (.701-.948)	.841 / .836 (.746-.968)	.750 / .743 (.593-.908)
ICC_{PHY}	.263 / .291 (.098-.443)	.194 / .173 (.060-.368)	.253 / .221 (.119-.435)	.383 / .409 (.157-.631)	.118 / .076 (.033-.261)	.057 / .050 (.014-.130)	.113 / .094 (.028-.253)
$ICC_{REL-PHY}$	.561 / .581 (.311-.763)	.597 / .609 (.345-.834)	.550 / .563 (.302-.752)	NA	.460 / .445 (.186-.699)	.503 / .486 (.269-.768)	NA

Table S9. Asexual generation models

	Incidence models				Frequency models		
Model term	Full model (option 1)	Without sample size	Full model (option 2)	Pooled sites (option 2)	Full model (option 1)	Full model (option 2)	Pooled sites (option 2)
Site	.021 / .001 (.000-.126)	.041 / .025 (.000-.218)	.026 / .015 (.000-.141)	NA	.037 / .023 (.002-.159)	.093 / .066 (.013-.349)	NA
Site:host interaction	.033 / .029 (.012-.059)	.152 / .157 (.079-.214)	.038 / .038 (.015-.067)	NA	.122 / .127 (.072-.175)	.175 / .175 (.094-.242)	NA
Site:parasitoid interaction	.001 / .000 (.000-.005)	.001 / .000 (.000-.004)	.001 / .000 (.000-.008)	NA	.003 / .000 (.000-.007)	.003 / .002 (.000-.006)	NA
Host species	.030 / .030 (.000-.075)	.049 / .053 (.000-.118)	.032 / .025 (.000-.075)	.058 / .060 (.000-.117)	.061 / .069 (.000-.122)	.047 / .034 (.000-.098)	.141 / .132 (.077-.227)
Parasitoid species	.008 / .000 (.000-.052)	.043 / .001 (.000-.113)	.009 / .000 (.000-.054)	.005 / .000 (.000-.038)	.032 / .001 (.000-.093)	.067 / .069 (.000-.141)	.122 / .130 (.000-.236)
Host phylogeny	.042 / .001 (.000-.202)	.091 / .002 (.000-.352)	.036 / .002 (.000-.203)	.043 / .002 (.000-.234)	.061 / .002 (.000-.306)	.033 / .001 (.000-.218)	.023 / .001 (.000-.184)
Parasitoid phylogeny	.011 / .001 (.000-.086)	.017 / .001 (.000-.133)	.010 / .001 (.000-.084)	.013 / .001 (.000-.096)	.013 / .000 (.000-.104)	.017 / .001 (.000-.135)	.030 / .001 (.000-.233)
Parasitoid:host species interaction	.092 / .080 (.025-.164)	.035 / .027 (.000-.070)	.095 / .093 (.031-.178)	NA	.035 / .040 (.009-.068)	.019 / .019 (.003-.040)	NA
Host phylogenetic interaction	.024 / .001 (.000-.175)	.075 / .001 (.000-.224)	.031 / .001 (.000-.186)	.019 / .001 (.000-.170)	.099 / .001 (.000-.231)	.094 / .092 (.014-.193)	.175 / .164 (.011-.339)
Parasitoid phylogenetic interaction	.040 / .000 (.000-.106)	.017 / .000 (.000-.051)	.036 / .001 (.000-.100)	.124 / .130 (.000-.254)	.022 / .014 (.000-.049)	.017 / .016 (.000-.036)	.038 / .039 (.005-.076)
Cophylogenetic interaction	.367 / .392 (.178-.537)	.232 / .253 (.068-.392)	.337 / .336 (.156-.508)	.302 / .315 (.096-.508)	.185 / .142 (.026-.329)	.110 / .093 (.014-.224)	.180 / .170 (.029-.358)
Residual	.244 / .258 (.172-.319)	.140 / .136 (.083-.188)	.253 / .270 (.173-.320)	.355 / .350 (.213-.513)	.247 / .268 (.140-.329)	.228 / .226 (.120-.304)	.218 / .234 (.128-.306)
ICC denominator	17.549 / 16.641 (12.996-24.187)	30.693 / 29.446 (20.352-45.859)	16.945 / 15.834 (12.529-23.094)	12.094 / 10.864 (7.669-18.880)	26.786 / 26.672 (19.835-38.198)	39.950 / 38.639 (29.296-59.224)	22.247 / 21.827 (17.041-30.429)
ICC_{GLMM}	.756 / .742 (.681-.828)	.860 / .864 (.812-.917)	.747 / .730 (.680-.827)	.645 / .650 (.487-.787)	.753 / .732 (.671-.860)	.772 / .774 (.696-.880)	.782 / .766 (.694-.872)
ICC_{PHY}	.457 / .450 (.295-.600)	.344 / .350 (.193-.505)	.430 / .426 (.281-.578)	.476 / .484 (.281-.636)	.318 / .309 (.196-.437)	.230 / .227 (.130-.333)	.404 / .406 (.262-.533)
$ICC_{REL-PHY}$	.832 / .845 (.690-.957)	.905 / .904 (.797-1.000)	.817 / .827 (.673-.960)	NA	.900 / .889 (.807-.987)	.921 / .929 (.839-.990)	NA

Sensitivity of model results to reduced sampling intensity

Figure S4. Accumulation curves in the sexual generation (A-C, top row) and asexual generation (D-F, bottom row) datasets for gall type richness (A, D), parasitoid species richness (B, E), and interaction richness (C, F). Horizontal lines indicate the Chao-2 (orange) and first order jackknife (blue) estimates of richness (with +/- 1 standard deviation) from the full datasets. From left to right the red points indicate the mean observed richness for 5%, 10%, 25%, 50%, and 75% subsets of the full data with bars for +/- 1 standard deviation, and the observed value for the full dataset. Values below points show their percentage relative to the Chao-2 (orange) and first order jackknife (blue) richness estimates.