

# The role of phylogeny and ecological opportunity in host-parasite interactions: network metrics, host repertoire, and network link prediction.

Armando Cruz-Laufer<sup>1</sup>, Tom Artois<sup>1</sup>, Stephan Koblmüller<sup>2</sup>, Antoine Pariselle<sup>3</sup>, Karen Smeets<sup>1</sup>, Maarten Van Steenberge<sup>4</sup>, and Maarten Vanhove<sup>1</sup>

<sup>1</sup>Hasselt University

<sup>2</sup>Karl-Franzens-Universität Graz

<sup>3</sup>ISEM

<sup>4</sup>KU Leuven

September 27, 2021

## Abstract

Hosts and parasites have often intimate associations. Therefore, the evolution of their interactions is crucial for understanding species-rich host-parasite communities. Yet relatively few studies investigate eco-evolutionary feedbacks in these systems as large datasets remain scarce. Here, we explore African cichlid fishes and their flatworm gill parasites (*Cichlidogyrus* spp.) including 9901 reported infections and 473 different host-parasite combinations collected through a survey of peer-reviewed literature. We apply network metrics, estimate host repertoires, and use network link prediction (NLP) algorithms to investigate meta-community structures and their predictors including evolutionary, ecological, and morphological parameters. Host repertoire was mostly determined by the hosts' evolutionary history. Both ecological and evolutionary parameters predicted host parasite associations but many interactions remain undetected according to NLP. We conclude that ecological opportunity paired with ecological fitting has shaped interactions. The cichlid-*Cichlidogyrus* network is a suitable study system for eco-evolutionary feedbacks but taxonomic research remains key to finding undetected interactions.

The role of phylogeny and ecological opportunity in host-parasite interactions: network metrics, host repertoire, and network link prediction

Armando J. Cruz-Laufer<sup>1</sup>, Tom Artois<sup>1</sup>, Stephan Koblmüller<sup>2</sup>, Antoine Pariselle<sup>3,4</sup>, Karen Smeets<sup>1</sup>, Maarten Van Steenberge<sup>5,6</sup>, Maarten P. M. Vanhove<sup>1,5,7</sup>

*1 UHasselt – Hasselt University, Faculty of Sciences, Centre for Environmental Sciences, Research Group Zoology: Biodiversity and Toxicology, Agoralaan Gebouw D, 3590 Diepenbeek, Belgium.*

*2 Institute of Biology, University of Graz, Universitätsplatz 2, 8010, Graz, Austria.*

*3 ISEM, Université de Montpellier, CNRS, IRD, Montpellier, France.*

*4 Faculty of Sciences, Laboratory “Biodiversity, Ecology and Genome”, Research Centre “Plant and Microbial Biotechnology, Biodiversity and Environment”, Mohammed V University, Rabat, Morocco.*

*5 Laboratory of Biodiversity and Evolutionary Genomics, KU Leuven, Charles Deberiotstraat 32, B-3000, Leuven, Belgium.*

*6 Operational Directorate Taxonomy and Phylogeny, Royal Belgian Institute of Natural Sciences, Vautierstraat 29, B-1000 Brussels, Belgium.*

7 Department of Botany and Zoology, Faculty of Science, Masaryk University, Kotlářská 2, CZ-611 37, Brno, Czech Republic.

Running title

Host parasite interaction networks

Key words

*Cichlidogyrus*, Cichlidae, fish parasites, flatworms, functional-phylogenetic distances, host niche, Monogenea, phylogenetic specificity, species interactions, structural specificity.

Type of article

Letter

Number of words in abstract

150

Number of words in text

4995

Number of references

132

Number of figures

5

Number of tables

1

Corresponding author

Armando J. Cruz-Laufer, armando.cruzlaufer@uhasselt.be

Statement of authorship

AJCL conceptualised the study and conducted the literature survey. AJCL performed all analyses and produced tables and graphs. AJCL produced host phylogenies with input from SK. AJCL and MPMV wrote the manuscript with input from TA, SK, AP, KS, and MVS.

Conflict of interest

The authors declare that they have no conflict of interest.

Data accessibility statement

Species interaction, host ecological, and community membership data as well as DNA sequence alignments underlying this article are available in Zenodo at [www.zenodo.org](http://www.zenodo.org), at <https://dx.doi.org/XXXXXXXXX>.

Abstract

Hosts and parasites have often intimate associations. Therefore, the evolution of their interactions is crucial for understanding species-rich host-parasite communities. Yet relatively few studies investigate eco-evolutionary feedbacks in these systems as large datasets remain scarce. Here, we explore African cichlid fishes and their flatworm gill parasites (*Cichlidogyrus* spp.) including 9901 reported infections and 473 different host-parasite combinations collected through a survey of peer-reviewed literature. We apply network metrics, estimate host repertoires, and use network link prediction (NLP) algorithms to investigate meta-community structures and their predictors including evolutionary, ecological, and morphological parameters.

Host repertoire was mostly determined by the hosts' evolutionary history. Both ecological and evolutionary parameters predicted host parasite associations but many interactions remain undetected according to NLP. We conclude that ecological opportunity paired with ecological fitting has shaped interactions. The cichlid-*Cichlidogyrus* network is a suitable study system for eco-evolutionary feedbacks but taxonomic research remains key to finding undetected interactions.

## Graphical Abstract

Hosts-parasite interactions are shaped by ecological and evolutionary processes. We investigate interactions of African cichlids and their flatworm parasites belonging to *Cichlidogyrus* (a) through network analyses (b), host repertoire estimation, and network link prediction (c). We show that the hosts' evolutionary history and environment determine observed host repertoires and network structure. Cichlid-*Cichlidogyrus* interactions are, thus, shaped by host phylogeny and ecological opportunity.

## Introduction

Many species on the planet are parasites at least during a portion of their lifetimes (Poulin 2014). Host-parasite interactions are often intimate associations that can profoundly affect host fitness (Kutzer & Armitage 2016) and, thus, shape biological communities (Gómez & Nichols 2013). However, host-parasite interactions, like other interactions, are not fixed in time. Ancient (Algar *et al.* 2009) and recent (Fussmann *et al.* 2007) evolutionary processes have produced present-day communities (Toju *et al.* 2017). To investigate this interplay of evolutionary history and community structure, integrative analyses of ecological and evolutionary patterns are crucial (Segare *et al.* 2020). For instance, host range, a key characteristic of parasite ecology (Poulin *et al.* 2011), is influenced not only by environmental factors but also the evolutionary history of the hosts (Poulin *et al.* 2011). More integrative measures, e.g. functional-phylogenetic distance metrics (FPDist) (Cadotte *et al.* 2013), can account for host ecology as well as evolutionary history (Clark & Clegg 2017). But do these metrics fully grasp the niche limitations of the parasites? The frequency of host switches recorded in the past (see Agosta *et al.* 2010) suggests otherwise. Host repertoires observed today have likely resulted from alternating phases of host range expansions and isolation (*oscillation hypothesis*) (Janz & Nylin 2008). Parasites expand their host range through their capacity to access novel resources (*ecological fitting*) (Agosta *et al.* 2010), i.e. host species, and through the opportunity emerging from the rise and fall of ecological barriers (D'Bastiani *et al.* 2020), e.g. after anthropogenic introductions (Brooks *et al.* 2021). Therefore, the *realised host repertoire*, which is approximated through FPDist, does not equate the full repertoire of host species that can potentially be infected (*fundamental host repertoire*) (Braga *et al.* 2020). The oscillation of host ranges resulting from ecological fitting and opportunity has been termed the *Stockholm Paradigm* (Brooks *et al.* 2019) and is considered one of the main sources of parasite biodiversity (Agosta & Brooks 2020).

One of the aspects highlighted by the *Stockholm paradigm* is the potential of predicting future host-parasite interactions in the context of emerging parasitic diseases. Understanding the mechanisms behind these diseases is increasingly relevant in a world where environmental degradation promotes host switches between previously unconnected hosts (Brooks *et al.* 2019). Host switches may present threats to human health and food security (Fitzpatrick 2013; Jenkins *et al.* 2015; Ekroth *et al.* 2019; Brooks *et al.* 2021). To understand parasitic interactions (Bogich *et al.* 2013; Bordes *et al.* 2017), ecological research has put forward network theory (Poulin 2010) through which species are represented as discrete interacting units, e.g. in plant-pollinator (Soares *et al.* 2017; Vizentin-Bugoni *et al.* 2018), predator-prey (Allesina & Pascual 2008), and plant-mycorrhiza systems (Simard *et al.* 2012). Ecologists widely employ this approach to characterise and visualise species interactions (Pocock *et al.* 2016). Moreover, rising computational capacities have promoted the use of network link prediction (NLP) algorithms to model undetected interactions. These methods originating in social network studies (Wang *et al.* 2015), have lately been optimised for biological systems (Martínez *et al.* 2016) including ecological networks (Dallas *et al.* 2017; Zhao *et al.* 2017; Fu *et al.* 2019).

Few recent studies on the *Stockholm paradigm* have integrated network analyses (but see D'Bastiani *et al.* 2020; Braga *et al.* 2021). Instead, the focus has remained on inferring ancestral host-parasite interactions

(Braga *et al.* 2020, 2021) rather than predicting undetected links. The distinction between undetected and unrealised links remains a major hurdle for network studies as observed interactions will often present an underestimation of the real interaction diversity (Fu *et al.* 2019). Furthermore, previous studies (Braga *et al.* 2020, 2021) treated interactions as discrete states, e.g. as non-hosts, potential hosts, and real hosts, despite the literature on network analyses substantiating that some host-parasite interactions are more prevalent than others (Blüthgen *et al.* 2008; Poulin *et al.* 2011). Many of the metrics describing the structure of species networks, like nestedness, connectance, and specialisation, have been optimised to account for interaction strength, i.e. the frequency of an observed interaction (see Blüthgen *et al.* 2008). Undetected links and interactions strength can be addressed through NLP as the algorithms account for both of these issues (Dallas *et al.* 2017; Fu *et al.* 2019).

Here, we investigate host-parasite interactions in a species-rich network using network theory and NLP. As model system, we selected one of the best known examples for explosive speciation: African cichlid fishes. Approximately 2000 species reside in the East African Great Lakes alone, many of which are endemic (Salzburger *et al.* 2014). Cichlid science has been at the forefront of evolutionary (e.g. Salzburger 2018; Ronco *et al.* 2021) and behavioural (see Koblmüller *et al.* 2019) research. Outside of feeding behaviour (e.g. Cooper *et al.* 2010; Hulsey *et al.* 2019), and fish-fish (e.g. Blažek *et al.* 2018; Marshall 2018) and human-fish interactions (Irvine *et al.* 2019), studies on interactions of cichlids with non-cichlid organisms have focused mostly on parasitic interactions (Cruz-Laufer *et al.* 2021a). One parasite lineage infecting African cichlids, the monogenean flatworms belonging to *Cichlidogyrus* Paperna, 1960 sensu Wu *et al.* (2007) (including *Scutogyrus* Pariselle & Euzet, 1995) is particularly species-rich. Currently, 137 species are described that infect the gills of 135 cichlid and five non-cichlid species (see Cruz-Laufer *et al.* 2021a). For this reason, the monogenean gill parasites of cichlids were proposed as model system for host-parasite interaction studies (Pariselle *et al.* 2003; Vanhove *et al.* 2016) (Fig. 1).

We explore cichlid-*Cichlidogyrus* interactions at a global scale. First, we use network metrics and community detection to characterise the structure of the observed network and meta-communities. Second, we assess the observed host range considering both functional and phylogenetic host diversity (Poulin *et al.* 2011; Esser *et al.* 2016) and discuss the limitations of this traditional approach to host repertoires. Third, we assess the performance of two recently proposed NLP models. We aim to address the following questions on the ecology and evolution of parasites using the cichlid-*Cichlidogyrus* network as a model system: (i) Do cichlid-*Cichlidogyrus* meta-communities reflect the evolutionary history of hosts and parasites, (ii) does the observed host repertoire correlate with functional or phylogenetic host diversity, and (iii) what can network link prediction models reveal about predictors of host-parasite interactions?

## Materials & methods

### *Data assembly: Infection data, niche modelling, phylogenies*

We assembled infection data through a survey of peer-reviewed literature. This survey resulted in an updated version (Supporting Information) of the list published by Cruz-Laufer *et al.* (2021a). For abundance weighting in downstream analyses, we also assembled infection parameters including the number of examined hosts, infected hosts, and parasites. If no infection parameters were reported, we considered a report as a single infected specimen.

We built host niche dendrograms based on ecological, geographical, and morphological data (Table 1) available in FishBase (Froese & Pauly 2000) and accessed through the *R* package *rfishbase* (Boettiger *et al.* 2012). Missing trophic level and habitat data were added through a literature survey (see Supporting Information). Dendrograms were built through hierarchical clustering in *R* (Pavoine *et al.* 2009) based on a Gower's distance matrix (Gower 1971). Gower's distances were calculated using the function *dist.ktab* in the *R* package *ade4* v1.7.16 (Pavoine *et al.* 2009). As suggested by Clark & Clegg (2017), we accounted for uncertainty of the host niche by implementing a range of clustering algorithms implemented in the *hclust* function in *R* (incl. *ward.D2*, *single*, *complete*, *average*, *mcquitty*, *median*, and *centroid*) (R Core Team 2021). We tested for topological congruence of the resulting dendrograms using the congruence among distance matri-

ces (CADM) test (Legendre & Lapointe 2004; Campbell *et al.* 2011) in the *R* package *apev5.4* (Paradis & Schliep 2019).

As no previous phylogenetic study on fishes covers all the species known to host members of *Cichlidogyrus*, we conducted a new analysis (see Appendix S1.1) based on DNA sequence data accessed on GenBank (Appendix S2) to infer phylogenetic distances between hosts. For the parasites, we included morphometric and phylogenetic data from Cruz-Laufer *et al.* (2021b), i.e. morphological measurements and 100 randomly sampled Bayesian tree topologies from the post-burn in fraction.

### *Network metrics and meta-community structure*

The infection data assembled here originate from different ecosystems. Therefore, we considered all communities inferred from these data as meta-communities of cichlids and species of *Cichlidogyrus*. We inferred meta-communities through the Louvain community detection algorithm, an approach based on optimisation of network modularity (see Blondel *et al.* 2008) implemented in the *R* package *igraph* v1.2.5 (Csardi & Nepusz 2006). The algorithm was applied to the entire (natural and invasive) documented host ranges with hosts and parasites treated equally as nodes. To characterise meta-community structure, we calculated a range of widely used network metrics including the weighted nestedness based on overlap and decreasing fill (NODF<sub>w</sub>) (Almeida-Neto & Ulrich 2011), weighted connectance (C<sub>w</sub>) (Bersier *et al.* 2002), specialisation asymmetry (SA) (Blüthgen *et al.* 2007), interaction evenness (E<sub>i</sub>) (Bersier *et al.* 2002), and the standardised interaction diversity (H<sub>2</sub>') (Blüthgen *et al.* 2006) using *R* package *bipartite* v2.15 (Dormann *et al.* 2008, 2009; Dormann 2011). We tested if meta-community membership was correlated to the parasite phylogeny through phylogenetic signal detection using the *R* package *geiger* (Pennell *et al.* 2014) as previously reported (Cruz-Laufer *et al.* 2021b).

We calculated network metrics for the ten most species-rich meta-communities (number of species > 10) (Fig. 2) separately for the full realised host repertoire and geographical distribution (including the result of anthropogenic translocations) and the natural ranges. To correct for varying sampling intensity, we produced two null distributions (NM): Patefield's algorithm (Patefield 1981), which randomly redistributes rows and columns of the interaction matrix (NM<sub>1</sub>) and the redistribution algorithm proposed by Vázquez *et al.* (2007) (NM<sub>2</sub>), which maintains the network connectance, i.e. only realised interactions are redistributed. We generated 1000 null matrices through the function *nullmodel* in *bipartite* and assessed significance as proportion of null estimates greater than the observed estimates.

### *Host repertoire*

We estimated the realised host repertoire as the structural and phylogenetic specificity of species of *Cichlidogyrus* (Esseret *et al.* 2016). Structural specificity was quantified through the specialisation index  $d_i'$  with significance levels inferred from NM<sub>1</sub> and NM<sub>2</sub>. This index measures the deviation from a perfectly nested network and is derived from Shannon's entropy, which takes diversity, abundance, and evenness of interactions into account (Blüthgen *et al.* 2006, 2008). Phylogenetic specificity was quantified as mean pairwise distance (MPD) and mean nearest taxon distance (MNTD) of host species to investigate ancient and recent relationships respectively (Clark & Clegg 2017). For the null model (NM<sub>3</sub>), we randomly redistributed the species labels of the host phylogenetic distance matrix to test if phylogenetic distances differed significantly from a random distribution. We calculated MPD, MNTD, and 1000 null estimates through the functions *mpd*, *mntd*, and *taxaShuffle* in the *R* package *picante* v1.8.2 (Kembel *et al.* 2010) for each parasite tree. We assessed the MPD and MNTD through the Z-scores provided by these functions. Negative Z-scores indicate a greater phylogenetic distance than expected under NM<sub>3</sub>, positive scores a smaller distance. Significance was assessed as proportion of MPD and NTD values smaller (negative Z-scores) or greater (positive Z-scores) than for the observed network.

Next, we investigated the effects of host environment and phylogeny on the host repertoire of the parasites. We calculate MPD and MNTD from functional-phylogenetic distance (FPDist) matrices. These matrices are derived from functional (FDist) and phylogenetic (PDist) distance matrices of the host species infected by each parasite species. We inferred the FDist matrices from the host niche dendrograms and the PDist matrices

from the host phylogenetic trees and scaled the matrices by dividing the values through the respective maximum distance. We accounted for uncertainty in the host niche/tree topology by drawing random samples of dendrograms/trees every time we calculated  $FPDist = (aPDist^p + (1 - a)FDist^p)^{1/p}$  (Cadotte *et al.* 2013) with  $p = 2$  to calculate Euclidean distances (see Cadotte *et al.* 2013; Burbrink *et al.* 2017; Clark & Clegg 2017). Finally, we applied 100 incremental increases from 0 to 1 to the weighting parameter  $a$ . We calculated  $FPDist$  through the function *FPDist* in the R package *funphylocom* v1.1 (Walker 2014). We generated null distributions by randomly redistributing the labels of the  $FPDist$  matrices resulting from 1000 random draws from  $FDist$  and  $PDist$  matrices. We applied the same redistribution algorithm as for  $NM_3$ . For interpretation of the  $FPDist$  plots, we followed Cadotte *et al.* (2013).

### Network link prediction

We applied the following network link prediction (NLP) algorithms:

The *plug-and-play* algorithm (Dallas *et al.* 2017) predicts missing links based on conditional probability estimation. This model was developed to infer the probabilities of unobserved links being undetected through a set of input parameters.

The *Poisson N-mixture* link prediction model (Fu *et al.* 2019) combines the Poisson N-mixture model used in ecological research with a low-rank collaborative filtering approach. Poisson N-mixture models are used in ecological research to account for imperfect detection in field observations (Royle 2004). Meanwhile, low-rank matrix completion-based collaborative filtering methods are a popular approach for NLP in social network studies. Missing entries in a data matrix are completed based on a low number of known entries (low rank matrix), e.g. to predict consumer preferences (Candes & Plan 2010).

We provided ecological, morphological, and phylogenetic input parameters to these models (Table 1). Both NLP models do currently not allow to account for phylogenetic uncertainty. Therefore, we included only the majority-rule consensus host and parasite BI phylogenies and the dendrograms calculated through the algorithm *ward.D2* (Murtagh & Legendre 2014), one of the most widely used clustering algorithm (Murtagh & Legendre 2014). To avoid overfitting, we reduced the number of input variables per parameter through principal coordinate analyses (PCoA) of the distance matrices of each parameter. Distance matrices of some parameters (Table 1) were inferred from dendrograms built through clustering methods employed for the host niche dendrograms. Distance matrices were computed through the *cophenetic* function in R v4.0.0 (R Core Team 2021). To address missing data, we imputed the data matrix (see Dallas *et al.* 2017) through the expectation-maximisation with bootstrapping as implemented in the R package *amelia* (Honaker *et al.* 2011). Overall, we provided 9 input parameters consisting of 25 variables (Table 1).

We determined model accuracies as the Area Under the Receiver Operating Characteristic curve (AUROC) statistic through 10-fold cross validation. Each time, the algorithms were trained on 80% of the interaction matrix to predict the remaining 20%. We implemented the models in R v4.0.0 (R Core Team 2021) and *MATLAB* v9.9.0 (MathWorks, Natick, USA) using the provided codes (doi: 10.6084/m9.figshare.4965038; <https://github.com/Hutchinson-Lab/Poisson-N-mixture>). Following Dallas *et al.* (2017), we assessed variable importance of the *plug-and-play* algorithm by measuring the reduction in model performance resulting from 500-fold permutation of each of the input variables. For variable assessment and host-parasite link prediction, the algorithm was trained on the full dataset. This assessment was not performed for the *Poisson N-mixture* model due to lacking implementation.

### Graphing

We plotted the species network through the R package *igraph* v1.2.5 (Csardi & Nepusz 2006), chord diagrams through *circlize* v0.4.12 (Gu *et al.* 2014), and dendrograms and phylogenetic trees through *ggplot2* v3.3.3 (Wickham 2016) and *ggtree* v2.2.1 (Yu *et al.* 2017, 2018).

## Results

### Data assembly

Across their natural and expanded host and geographical ranges, we found reports of 9901 infected host specimens and 473 different host-parasite interactions in 166 publications. The resulting species network comprises 141 host and 142 parasite species (described and undescribed with cheironyms). We assembled environmental and morphometric data for 133 and DNA sequence data for 117 host species (Appendix S1.1). Niche dendrograms resulting from the different clustering algorithms (Appendix S3) produced similar topologies (CADM, Kendall's  $W = 0.66$ ,  $\chi^2 = 34944$ ,  $p < 0.01$ ). The BI consensus tree topology showed no deviation from recent phylogenetic hypotheses likely to affect downstream analyses (see discussion in Appendix S1.1).

#### *Network metrics and community structure*

We identified 41 meta-communities in the cichlid-*Cichlidogyrus* network through the Louvain algorithm. A majority of these meta-communities included only a few species and were completely unconnected to the other communities (Fig. 2a). Six of these meta-communities included 10 or more species (Fig. 2b, c; Appendix S4): the 'Coptodon zillii' (CZ), 'Oreochromis niloticus' (ON), 'Hemichromis' (He), 'Ophthmotilapia' (Op), and 'Tilapia sparrmanii' (TS), and Lake Victoria (LV) clusters. These names were chosen to reflect the host taxa accounting for a majority of reported infections except for the LV community (Appendix S1.2). The network structure differed between the meta-communities but anthropogenic introductions had only a limited effect (Fig. 3; Appendix S1.2). Parasite community memberships showed no significant phylogenetic signal ( $\Delta AICc = -0.28$ ,  $p > 0.05$ ). However, some parasites infecting coptodonine and oreochromine cichlids are closely related as previously indicated by Cruz-Laufer *et al.* (2021b).

#### *Host repertoire*

Most species of *Cichlidogyrus* had a higher structural specificity than expected at random as their  $d_i'$  estimates often significantly exceeded the respective null distributions (Appendix S5). For the phylogenetic specificity, differences to the null models were less often significant but in case of significant differences, host meta-communities were clustered at the level of mean pairwise distance (MPD) or MPD and mean nearest taxon distance (MNTD) (Appendix S5).

Functional-phylogenetic distances (FPDist) were calculated for all species of *Cichlidogyrus* with more than one host species (Appendix S6) with the weighting parameter  $a$  between 0 and 1. Plots for some species are shown in Fig. 4 to illustrate the main patterns observed across all species. Generally, FPDist decreased with increasing phylogenetic weight (MPD or MNTD) as indicated by the falling trend of FPDist estimates (Fig. 4). Observed MPD and MNTD showed a significant pattern of underdispersion compared to 1000 random simulations in some cases. Specifically, MPD metrics of FPDist became mostly significant if the weight of phylogenetic distances (PDist) increased (i.e.  $a - 1$ ) (Fig. 4). In contrast, MNTD estimates were largely non-informative regardless of  $a$ . One exception to these observation is *Cichlidogyrus* sp. 'nyanza' (Gobbin *et al.* 2021) with MNTD estimates being lowest for  $a = 0$  (Fig. 4).

#### *Network link prediction*

The *plug-and-play* algorithm performed solidly (mean AUROC = 0.72, full model AUROC = 0.85) despite the missing data (Fig. 5a) and outperformed the *Poisson N-mixture* model (mean AUROC = 0.59). According to the former, a substantial amount of species interactions likely remains undetected (Fig. 5b; Appendix S7). All input parameters improved performance of the *plug-and-play* algorithm (Fig. 5c). However, the ecosystem, host and parasite phylogenies, and trophic level parameters were significantly more important than morphological parameters including the host length and attachment and reproductive organ measurements.

#### Discussion

##### *Diversity of host-parasite meta-communities*

We investigated the patterns of host-parasite interaction of African cichlid fishes and their gill parasites belonging to *Cichlidogyrus*, a proposed model system for macroevolutionary research (Pariselle *et al.* 2003;

Vanhove *et al.* 2016). The size of this species network (9901 infections, 473 interactions) is comparable to widely used host-parasite datasets in terms of species richness, e.g. the Global Mammal Parasite Database (GMPD) (Nunn & Altizer 2005), the Sevillata interaction network (Dallas & Presley 2014), and other fish-parasite (Lima Jr *et al.* 2012; Bellay *et al.* 2015) and plant-arthropod systems (López-Carretero *et al.* 2014; de Araújo *et al.* 2020; Oliveira *et al.* 2020; de Araújo & Maia 2021). Additionally, the dataset here offers the opportunity to investigate the effects of evolutionary mechanisms, especially adaptive radiation events, on species interactions. This system is also the first to encompass closely related parasite species infecting hosts that are a model system for speciation research (Seehausen 2006). Therefore, the cichlid-*Cichlidogyrus* dataset could be a valuable asset for studies on network ecology and ecological parasitology.

Some of the host-parasite meta-communities inferred from this network (Fig. 2b, c) could be of particular interest for future studies due to their sample size and species richness. Cichlid-*Cichlidogyrus* interactions are diverse regarding their meta-community structures (Appendix S1.2). For instance, only one recorded monogenean species infects deepwater cichlids from the tribe Bathybatini (Kmentová *et al.* 2016) (not in Fig. 2c as less than 10 species) whereas 27 host species in the LV community share multiple parasite species (Fig. 2b, Appendix S1.2). Some communities show a reversed specialisation asymmetry (Appendix S1.2) with the host communities being more diverse than their parasites unlike most host-parasite (Vázquez *et al.* 2005) or mutualistic (Thébault & Fontaine 2008) networks possibly due to the young age of the host radiation (see Appendix S1.2). Furthermore, anthropogenic introductions have expanded the realised host repertoire of some meta-communities (Appendix S1.2) albeit with little changes to their network structure (Fig. 3). The diversity of meta-communities could lay the ground for future comparative studies, e.g. on the communities from the East African lakes mirroring the research conducted on cichlids from the same region (e.g. Duponchelle *et al.* 2008).

#### *Host repertoire: Community structure shaped by host evolution*

The evolutionary history of the hosts can have a significant impact on the community structure in host-parasite networks (e.g. Mouillot *et al.* 2008a; Braga *et al.* 2014, 2020). Here, we show that the realised host repertoire of most species of *Cichlidogyrus* is determined more by the host evolutionary history than the host environment. First, host species differed more regarding the ecological niche than the phylogenetic relationships as evidenced by a reduction of the mean functional-phylogenetic distances (FPDist) with increasing phylogenetic weight ( $a - 1$ ) (Fig. 4). Second, FPDist estimates rarely differed from the null distribution for increasing functional weight ( $a - 0$ ) (but see *C. sp.* ‘nyanza’ in Fig. 4). Third, estimates outside the null distribution were always underdispersed (clustered), i.e. lower than expected at random at both the ancient and recent evolutionary scale (measured as MPD and MNTD respectively). The strong phylogenetic influence and underdispersion have previously been associated with co-divergent evolution (Clark & Clegg 2017). Co-divergence assumes that host and parasite phylogenies are phylogenetically congruent (Page 2003; Hoyal Cuthill & Charleston 2012), a pattern that has already been observed for species of *Cichlidogyrus* (Vanhove *et al.* 2015). In fact, host repertoires observed here frequently coincide with related groups of host species. According to the *Stockholm Paradigm*, congruence might arise especially in younger lineages that have experienced a phase of isolation resulting in co-diverging host and parasite lineages (Agosta & Brooks 2020), e.g. species of *Cichlidogyrus* infecting trophic cichlids (Vanhove *et al.* 2015), which arose 7–11 MYA (Schedel *et al.* 2019).

Another explanation for the important role of the host phylogeny on the host range of the parasites might be phylogenetic constraints. Within the same ecosystem, parasites cannot simply infect any host but are often limited to compatible host lineages. For instance, phylogenetic relationships are reportedly determinants of neotropical (Braga *et al.* 2014) or Mediterranean (Desdevises *et al.* 2002) monogenean-fish communities but also in plant-pollinator, plant-frugivore (Rezende *et al.* 2007), plant-mycorrhiza (Jacquemyn *et al.* 2011), and other host-parasite networks (Mouillot *et al.* 2008). Yet natural (Birgi & Euzet 1983; Birgi & Lambert 1986) and invasion-induced (Jimenez-Garcia *et al.* 2001; Šimková *et al.* 2019; Jorissen *et al.* 2020) host switches of species of *Cichlidogyrus* highlight that fundamental host repertoires might differ considerably from the realised host repertoires estimated here. Monogeneans have the capacity to infect

new hosts and adapt to new environments (Braga *et al.* 2014). For instance, the attachment organ, as the main physical connection of the parasite to its host, is often used as a proxy for evolutionary processes in monogenean flatworms (e.g. Jarkovský *et al.* 2004; Vignon *et al.* 2011). In some species of *Cichlidogyrus* that underwent recent host switches (Messu Mandeng *et al.* 2015), the organ’s morphology diverges considerably from those of close relatives suggesting that several species were able to expand their host ranges and make subsequent morphological changes. Consequently, phylogenetic constraints might play less of a role in cichlid-*Cichlidogyrus* than the observed host repertoires suggest at the first glance.

#### *Network link prediction: Community structure shaped by ecological factors*

Despite the uncertainty in the host range estimation, network link prediction (NLP) models confirmed the influence of the host evolutionary history on the structure of fish-monogenean communities. The host phylogeny contributed considerably to the acceptable performance of the *plug-and-play* algorithm (AUROC: 0.72), which outperformed the more complex *Poisson N-mixture* model. However, host-parasite links appear to be mostly predicted by ecological parameters as the ecosystem variable (Table 1) contributed the most to model performance (Fig. 5c). Therefore, ecological opportunity might play a major role in the assembly of cichlid-*Cichlidogyrus* communities similar to neotropical fish-monogenean communities (Braga *et al.* 2014), and these opportunities are likely created by host geographical and habitat distribution.

The uncovered significance of opportunity is highly relevant for aquaculture and fish conservation efforts. Introductions of infectious diseases can have devastating effects on native ecosystems (Thompson 2013). For instance, economic consequences for tilapia aquaculture were felt in some countries in association with the co-introduction of the tilapia-lake virus (Eyngor *et al.* 2014; Fathi *et al.* 2017). Moreover, introductions of Nile tilapia (*Oreochromis niloticus* L.) and other large cichlids have led to co-introductions of their monogenean parasites, e.g. in continental Africa (Jorissen *et al.* 2020), Madagascar (Šimková *et al.* 2019), Asia (Paperna 1960; Duncan 1973; Wu *et al.* 2006), Australia (Wilson *et al.* 2019), and the Americas (Jiménez-García *et al.* 2001; de Azevedo *et al.* 2006), and to occasional host switches to native fishes (Jiménez-García *et al.* 2001; Šimková *et al.* 2019). Our results suggest that more of these host range expansions might occur through anthropogenic introductions. Therefore, introduced populations and their surrounding environments should continue to be monitored.

Our results show that NLP can be a useful tool to verify traditional statistical analyses and to gain further insight into ecological and evolutionary mechanisms shaping host-parasite interactions. For instance, we inferred that the trophic level of the host is one of the more informative predictors of cichlid-*Cichlidogyrus* interactions. Host size, life style, and parasite phylogenetic and attachment organ morphological parameters also improved model performance (Fig. 5b). In contrast, previous studies on fish parasites have delivered inconclusive results for the role of host and parasite traits on parasite community composition. Parasite community composition correlated with the host *trophic level* in some cases, e.g. for shelf fish off Buenos Aires (Timi *et al.* 2011), but not in others, e.g. for freshwater fish in Canada (Locke *et al.* 2013) and marine fish in Finland (Locke *et al.* 2014). No studies investigated the effects of *life style* as coded here (Table 1) but other studies suggest that host habitat preference can affect parasite communities (Locke *et al.* 2013). *Host size* was suggested as important predictor for the community composition of ectoparasitic monogeneans (Sasal & Morand 1998; Sasal *et al.* 1999; Šimková *et al.* 2001; Desdevises *et al.* 2002; Morand *et al.* 2002). However, these correlations might reflect phylogenetic patterns of host size (Poulin 2002) explaining the variable importance of host size here. Lastly, no correlation of *attachment* or *reproductive organ morphology* with community composition was found for species of *Cichlidogyrus* unlike for other monogeneans, e.g. *Dactylogyrus* (Šimková *et al.* 2001; Jarkovský *et al.* 2004). Instead, the morphology mostly reflects phylogenetic relationships of the parasites (Vignon *et al.* 2011; Cruz-Laufer *et al.* 2021b). The results of these studies highlight the challenge of linking host and parasite traits with community composition parameters and generalising observed patterns as sampling biases (Fründ *et al.* 2016) (Fig. 5a) and character coding (Pavoine *et al.* 2009) can influence the results. NLP can complement these analyses by indicating possibly undetected interactions (Fig. 5b) and assessing the predictive power of the ecological, evolutionary, and morphological parameters (Fig. 5c).

### *Limitations and opportunities: Sampling bias, missing data, and databases*

Through a literature survey, we produced the most extensive study on species interaction of cichlid fishes or any other lineage of adaptive radiations to date. We also inferred patterns in the community structure through a series of network analytical methods ranging from more traditional to new approaches. However, this study has also limitations. We suggest addressing these limitations through the following measures:

- Because of the sampling bias in cichlid-*Cichlidogyrus* interactions studies towards economically relevant hosts (Cruz-Laufer *et al.* 2021a), the present data likely give an incomplete picture as confirmed by the *plug-and-play* algorithm (Fig. 5b). Null models can account for this issue but, ultimately, taxonomic research remains essential for closing knowledge gaps on cichlid-*Cichlidogyrus* interactions. Data generated from such studies should be gathered in online databases, e.g. the Global Interaction Database GLOBI (Poelen *et al.* 2014), to increase accessibility of this study system for research communities worldwide (Molloy 2011; Upham *et al.* 2021).
- The host-parasite interaction analysed here are pooled from different geographic regions and climatic conditions. Future studies should also account for geographical distribution as geographic coordinates to infer local interaction patterns as we expect environmental parameters to vary considerably across species ranges.
- New models for NLP are continuously being developed and employed in an increasing number of fields (Martínez *et al.* 2016). We suggest that a streamlined software package or library targeted at ecological research could simplify model implementation for ecologists.
- The NLP algorithms applied here differentiate between true (impossible or ‘forbidden’ links) and false negatives (undetected links) (Dallas *et al.* 2017; Fu *et al.* 2019) among unobserved interactions. However, wildlife host-parasite infection data often include additional information in the form of prevalence data, i.e. ratios of uninfected host specimens. This information could be included in future models.

Addressing these limitations will increase the potential of cichlids and their gill parasites belonging to *Cichlidogyrus* as an upcoming model system for eco-evolutionary studies in host-parasite systems. We were able to detect key mechanisms of ecology and evolution. First, the realised host repertoire is phylogenetically constrained as host range parameters are determined more by the host evolutionary history than by ecological parameters. However, recent host switches indicate that fundamental host repertoire might be more extensive than the present data suggest. Second, network link prediction algorithms show that network structure is shaped by ecological opportunity induced by habitat sharing but host and parasite evolution and host trophic level are also influential factors. Our results demonstrate that cichlid-*Cichlidogyrus* data can be utilised for a range of network analyses because of a substantial amount of interaction data, and molecular and morphological information for hosts and parasites. We encourage researchers to reuse the data provided here to diversify the portfolio of host-parasite interaction research in the future.

### Acknowledgements

We would like to thank Walter A. Boeger for his extensive comments and thoughts on the manuscript. Data collection started within the BRAIN-be Pioneer Project BR/132/PI/TILAPIA (Belgian Federal Science Policy Office) under the supervision of Tine Huyse and Jos Snoeks and the Knowledge Management Centre project CiMonoWeb (Royal Museum for Central Africa) under the supervision of Tine Huyse with the kind help of Wouter Fannes. Part of the research leading to results presented in this publication was carried out with infrastructure funded by the European Marine Biological Research Centre (EMBRC) Belgium, Research Foundation – Flanders (FWO) project GOH3817N. AJCL is funded by Hasselt University (BOF19OWB02) and MPMV receives support from the Special Research Fund of Hasselt University (BOF20TT06).

### References

Agosta, S.J. & Brooks, D.R. (2020). Chapter 10: The Stockholm Paradigm. In: *The major metaphors of evolutionary: Darwinism then and now*, Evolutionary biology – new perspectives on its development, vol. 2 (eds. Agosta, S.J. & Brooks, D.R.). Springer, Cham, Germany, pp. 219–242.

- Agosta, S.J., Janz, N. & Brooks, D.R. (2010). How specialists can be generalists: resolving the and “parasite paradox” and implications for emerging infectious disease. *Zoologia* , 27, 151–162.
- Algar, A.C., Kerr, J.T. & Currie, D.J. (2009). Evolutionary constraints on regional faunas: whom, but not how many. *Ecol. Lett.* , 12, 57–65.
- Allesina, S. & Pascual, M. (2008). Network structure, predator–prey modules, and stability in large food webs. *Theor. Ecol.* , 1, 55–64.
- Almeida-Neto, M. & Ulrich, W. (2011). A straightforward computational approach for measuring nestedness using quantitative matrices. *Environ. Model. Softw.* , 26, 173–178.
- de Araújo, W.S., de Freitas, É.V.D., Silveira, L.T. & Daud, R.D. (2020). Network structure of interactions between phytophagous mites and their host-plants in natural ecosystems in Brazil. *Syst. Appl. Acarol.* , 25, 821–832.
- de Araújo, W.S. & Maia, V.C. (2021). Topological structure of a tritrophic network composed of host plants, gall-inducing insects and parasitoids in a restinga area in Brazil. *Entomol. Sci.* , 24, 201–216.
- de Azevedo, T.M.P., Martins, M.L., Bozzo, F.R. & de Moraes, F.R. (2006). Haematological and gill responses in parasitized tilapia from valley of Tijucas River, SC, Brazil. *Sci. Agric.* , 63, 115–120.
- Bellay, S., de Oliveira, E.F., Almeida-Neto, M., Abdallah, V.D., de Azevedo, R.K., Takemoto, R.M., *et al.* (2015). The patterns of organisation and structure of interactions in a fish-parasite network of a neotropical river. *Int. J. Parasitol.* , 45, 549–557.
- Bersier, L.-F., Banašek-Richter, C. & Cattin, M.-F. (2002). Quantitative descriptors of food-web matrices. *Ecology* , 83, 2394–2407.
- Birgi, E. & Euzet, L. (1983). Monogènes parasites des poissons des eaux douces du Cameroun. Présence des genres *Cichlidogyrus* et *Dactylogyrus* chez *Aphyosemion* (Cyprinodontidae). *Bull. la Soc. Zool. Fr.* , 108, 101–106.
- Birgi, E. & Lambert, A. (1986). Présence chez un Nandidae (Téléostéen), *Polycentropsis abbreviata* Boulenger, 1901, du genre *Cichlidogyrus* (Monogenea, Monopisthocotylea, Ancyrocephalidae). *Ann. Parasitol. Hum. Comp.* , 61, 521–528.
- Blažek, R., Polačik, M., Smith, C., Honza, M., Meyer, A. & Reichard, M. (2018). Success of cuckoo catfish brood parasitism reflects coevolutionary history and individual experience of their cichlid hosts. *Sci. Adv.* , 4, eaar4380.
- Blondel, V.D., Guillaume, J.L., Lambiotte, R. & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* , 2008.
- Blüthgen, N., Fründ, J., Vazquez, D.P. & Menzel, F. (2008). What do interaction network metrics tell us about specialization and biological traits? *Ecology* , 89, 3387–3399.
- Blüthgen, N., Menzel, F. & Blüthgen, N. (2006). Measuring specialization in species interaction networks. *BMC Ecol.* , 6, 9.
- Blüthgen, N., Menzel, F., Hovestadt, T., Fiala, B. & Blüthgen, N. (2007). Specialization, constraints, and conflicting interests in mutualistic networks. *Curr. Biol.* , 17, 341–346.
- Boettiger, C., Lang, D.T. & Wainwright, P.C. (2012). rfishbase: exploring, manipulating and visualizing FishBase data from R. *J. Fish Biol.* , 81, 2030–2039.
- Bogich, T.L., Funk, S., Malcolm, T.R., Chhun, N., Epstein, J.H., Chmura, A.A., *et al.* (2013). Using network theory to identify the causes of disease outbreaks of unknown origin. *J. R. Soc. Interface* , 10.

- Bordes, F., Caron, A., Blasdell, K., de Garine-Wichatitsky, M. & Morand, S. (2017). Forecasting potential emergence of zoonotic diseases in South-East Asia: network analysis identifies key rodent hosts. *J. Appl. Ecol.* , 54, 691–700.
- Braga, M.P., Araújo, S.B.L. & Boeger, W.A. (2014). Patterns of interaction between Neotropical freshwater fishes and their gill Monogenoidea (Platyhelminthes). *Parasitol. Res.* , 113, 481–490.
- Braga, M.P., Janz, N., Nylin, S., Ronquist, F. & Landis, M.J. (2021). Phylogenetic reconstruction of ancestral ecological networks through time for pierid butterflies and their host plants. *Ecol. Lett.*
- Braga, M.P., Landis, M.J., Nylin, S., Janz, N. & Ronquist, F. (2020). Bayesian inference of ancestral host-parasite interactions under a phylogenetic model of host repertoire evolution. *Syst. Biol.* , 69, 1149–1162.
- Brooks, D.R., Hoberg, E.P. & Boeger, W.A. (2019). *The Stockholm Paradigm: climate change and emerging disease* . University of Chicago Press, Chicago, USA.
- Brooks, D.R., Hoberg, E.P., Boeger, W.A. & Trivellone, V. (2021). Emerging infectious disease: An underappreciated area of strategic concern for food security. *Transbound. Emerg. Dis.*
- Burbrink, F.T., Lorch, J.M. & Lips, K.R. (2017). Host susceptibility to snake fungal disease is highly dispersed across phylogenetic and functional trait space. *Sci. Adv.* , 3, e1701387.
- Cadotte, M., Albert, C.H. & Walker, S.C. (2013). The ecology of differences: Assessing community assembly with trait and evolutionary distances. *Ecol. Lett.* , 16, 1234–1244.
- Campbell, V., Legendre, P. & Lapointe, F.J. (2011). The performance of the Congruence Among Distance Matrices (CADM) test in phylogenetic analysis. *BMC Evol. Biol.* , 11.
- Candes, E.J. & Plan, Y. (2010). Matrix completion with noise. *Proc. IEEE* , 98, 925–936.
- Clark, N.J. & Clegg, S.M. (2017). Integrating phylogenetic and ecological distances reveals new insights into parasite host specificity. *Mol. Ecol.* , 26, 3074–3086.
- Cooper, W.J., Parsons, K., McIntyre, A., Kern, B., McGee-Moore, A. & Albertson, R.C. (2010). Benthopelagic divergence of cichlid feeding architecture was prodigious and consistent during multiple adaptive radiations within African Rift-Lakes. *PLoS One* , 5, e9551.
- Cruz-Laufer, A.J., Artois, T., Smeets, K., Pariselle, A. & Vanhove, M.P.M. (2021a). The cichlid–*Cichlidogyrus* network: a blueprint for a model system of parasite evolution. *Hydrobiologia* , 848, 3847–3863.
- Cruz-Laufer, A.J., Pariselle, A., Jorissen, M.W.P., Muterezi Bukinga, F., Al Assadi, A., Van Steenberge, M., *et al.* (2021b). Somewhere I belong: phylogenetic comparative methods and machine learning to investigate the evolution of a species-rich lineage of parasites [preprint]. *bioRxiv* , 10.1101/2021.03.22.435939.
- Csardi, G. & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Syst.* 1695 .
- D’Bastiani, E., Campião, K.M., Boeger, W.A. & Araújo, S.B.L. (2020). The role of ecological opportunity in shaping host-parasite networks. *Parasitology* , 147, 1452–1460.
- Dallas, T., Park, A.W. & Drake, J.M. (2017). Predicting cryptic links in host-parasite networks. *PLoS Comput. Biol.* , 13, e1005557.
- Dallas, T. & Presley, S.J. (2014). Relative importance of host environment, transmission potential and host phylogeny to the structure of parasite metacommunities. *Oikos* , 123, 866–874.
- Desdevises, Y., Morand, S. & Legendre, P. (2002). Evolution and determinants of host specificity in the genus *Lamellodiscus* (Monogenea). *Biol. J. Linn. Soc.* , 77, 431–443.
- Dormann, C.F. (2011). How to be a specialist? Quantifying specialisation in pollination networks. *Netw. Biol.* , 1, 1–20.

- Dormann, C.F., Fründ, J., Blüthgen, N. & Gruber, B. (2009). Indices, graphs and null models: analyzing bipartite ecological networks. *Open Ecol. J.* , 2, 7–24.
- Dormann, C.F., Gruber, B. & Fründ, J. (2008). Introducing the bipartite package: analysing ecological networks. *R News* , 8, 8–11.
- Duncan, B.L. (1973). *Cichlidogyrus sclerosus* Paperna and Thurston from cultured *Tilapia mossambica* . *Philipp. J. Biol.* , 2, 154–158.
- Duponchelle, F., Paradis, E., Ribbink, A.J. & Turner, G.F. (2008). Parallel life history evolution in mouth-brooding cichlids from the African Great Lakes. *Proc. Natl. Acad. Sci. U. S. A.* , 105, 15475–15480.
- Ekroth, A.K.E., Rafaluk-Mohr, C. & King, K.C. (2019). Host genetic diversity limits parasite success beyond agricultural systems: a meta-analysis. *Proc. R. Soc. B Biol. Sci.* , 286.
- Esser, H.J., Herre, E.A., Blüthgen, N., Loaiza, J.R., Bermúdez, S.E. & Jansen, P.A. (2016). Host specificity in a diverse Neotropical tick community: an assessment using quantitative network analysis and host phylogeny. *Parasit. Vectors* , 9, 372.
- Eyngor, M., Zamostiano, R., Tsosack, J.E.K., Berkowitz, A., Bercovier, H., Tinman, S., *et al.* (2014). Identification of a novel RNA virus lethal to tilapia. *J. Clin. Microbiol.* , 52, 4137–4146.
- Fathi, M., Dickson, C., Dickson, M., Leschen, W., Baily, J., Muir, F., *et al.* (2017). Identification of Tilapia Lake Virus in Egypt in Nile tilapia affected by ‘summer mortality’ syndrome. *Aquaculture* , 473, 430–432.
- Fitzpatrick, J.L. (2013). Global food security: the impact of veterinary parasites and parasitologists. *Vet. Parasitol.* , 195, 233–248.
- Froese, R. & Pauly, D. (Eds.). (2000). *FishBase 2000: Concepts, designs and data sources* . ICLARM, Los Baños, Laguna, Philippines.
- Fründ, J., Mccann, K.S. & Williams, N.M. (2016). Sampling bias is a challenge for quantifying specialization and network structure: Lessons from a quantitative niche model. *Oikos* , 125, 502–513.
- Fu, X., Seo, E., Clarke, J. & Hutchinson, R.A. (2019). Link prediction under imperfect detection: collaborative filtering for ecological networks. *IEEE Trans. Knowl. Data Eng.* , 33, 3117–3128.
- Fussmann, G.F., Loreau, M. & Abrams, P.A. (2007). Eco-evolutionary dynamics of communities and ecosystems. *Funct. Ecol.* , 21, 465–477.
- Gobbin, T.P., Vanhove, M.P.M., Seehausen, O., Maan, M.E. & Pariselle, A. (2021). Four new species of *Cichlidogyrus* (Platyhelminthes, Monogenea, Dactylogyridae) from Lake Victoria haplochromine cichlid fishes, with the redescription of *C. bifurcatus* and *C. longipenis* [preprint]. *bioRxiv* , 2021.01.29.428376.
- Gómez, A. & Nichols, E. (2013). Neglected wild life: parasitic biodiversity as a conservation target. *Int. J. Parasitol. Parasites Wildl.* , 2, 222–227.
- Gower, J.C. (1971). A general coefficient of similarity and some of its properties. *Biometrics* , 27, 857.
- Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. (2014). Circlize implements and enhances circular visualization in R. *Bioinformatics* , 30, 2811–2812.
- Honaker, J., King, G. & Blackwell, M. (2011). Amelia II: a program for missing data. *J. Stat. Softw.* , 45, 1–45.
- Hoyal Cuthill, J. & Charleston, M. (2012). Phylogenetic codivergence supports coevolution of mimetic *Heliconius* butterflies. *PLoS One* , 7, e36464.
- Hulsey, C.D., Alfaro, M.E., Zheng, J., Meyer, A. & Holzman, R. (2019). Pleiotropic jaw morphology links the evolution of mechanical modularity and functional feeding convergence in Lake Malawi cichlids. *Proc. R. Soc. B Biol. Sci.* , 286.

- Irvine, K., Etiegni, C.A. & Weyl, O.L.F. (2019). Prognosis for long-term sustainable fisheries in the African Great Lakes. *Fish. Manag. Ecol.* , 26, 413–425.
- Jacquemyn, H., Merckx, V., Brys, R., Tyteca, D., Cammue, B.P.A., Honnay, O., *et al.* (2011). Analysis of network architecture reveals phylogenetic constraints on mycorrhizal specificity in the genus *Orchis* (Orchidaceae). *New Phytol.* , 192, 518–528.
- Janz, N. & Nylin, S. (2008). The oscillation hypothesis of host-plant range and speciation. In: *Specialization, speciation, and radiation: the evolutionary biology of herbivorous insects* . University of California Press, pp. 203–215.
- Jarkovský, J.J., Morand, S., Šimková, A. & Gelnar, M. (2004). Reproductive barriers between congeneric monogenean parasites (Dactylogyrus: Monogenea): attachment apparatus morphology or copulatory organ incompatibility? *Parasitol. Res.* , 92, 95–105.
- Jenkins, E.J., Simon, A., Bachand, N. & Stephen, C. (2015). Wildlife parasites in a One Health world. *Trends Parasitol.* , 31, 174–180.
- Jiménez-García, M.I., Vidal-Martínez, V.M., López-Jiménez, S., Jiménez-García, M.I., Vidal-Martínez, V.M. & López-Jiménez, S. (2001). Monogeneans in introduced and native cichlids in México: Evidence for transfer. *J. Parasitol.* , 87, 907.
- Jorissen, M.W.P., Huyse, T., Pariselle, A., Wamuini Lunkayilakio, S., Muterezi Bukinga, F., Chocha Manda, A., *et al.* (2020). Historical museum collections help detect parasite species jumps after tilapia introductions in the Congo Basin. *Biol. Invasions* , 11, 1123.
- Kembel, S.W., Cowan, P.D., Helmus, M.R., Cornwell, W.K., Morlon, H., Ackerly, D.D., *et al.* (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* , 26, 1463–1464.
- Kmentová, N., Gelnar, M., Mendlová, M., Van Steenberge, M., Koblmüller, S. & Vanhove, M.P.M. (2016). Reduced host-specificity in a parasite infecting non-littoral Lake Tanganyika cichlids evidenced by intraspecific morphological and genetic diversity. *Sci. Rep.* , 6, 39605.
- Koblmüller, S., Albertson, R.C., Genner, M.J., Sefc, K.M. & Takahashi, T. (2019). Preface: advances in cichlid research III: behavior, ecology, and evolutionary biology. *Hydrobiologia* , 832, 1–8.
- Kutzer, M.A.M. & Armitage, S.A.O. (2016). Maximising fitness in the face of parasites: a review of host tolerance. *Zoology* , 119, 281–289.
- Legendre, P. & Lapointe, F.-J. (2004). Assessing congruence among distance matrices: single-malt scotch whiskies revisited. *Aust. N. Z. J. Stat.* , 46, 615–629.
- Lima Jr, D.P., Giacomini, H.C., Takemoto, R.M., Agostinho, A.A. & Bini, L.M. (2012). Patterns of interactions of a large fish-parasite network in a tropical floodplain. *J. Anim. Ecol.* , 81, 905–913.
- Locke, S.A., Marcogliese, D.J. & Tellervo Valtonen, E. (2014). Vulnerability and diet breadth predict larval and adult parasite diversity in fish of the Bothnian Bay. *Oecologia* , 1, 253–262.
- Locke, S.A., McLaughlin, J.D. & Marcogliese, D.J. (2013). Predicting the similarity of parasite communities in freshwater fishes using the phylogeny, ecology and proximity of hosts. *Oikos* , 122, 73–83.
- López-Carretero, A., Díaz-Castelazo, C., Boege, K. & Rico-Gray, V. (2014). Evaluating the spatio-temporal factors that structure network parameters of plant-herbivore interactions. *PLoS One* , 9, e110430.
- Marshall, B.E. (2018). Guilty as charged: Nile perch was the cause of the haplochromine decline in Lake Victoria. *Can. J. Fish. Aquat. Sci.* , 75, 1542–1559.
- Martínez, V., Berzal, F. & Cubero, J.C. (2016). A survey of link prediction in complex networks. *ACM Comput. Surv.* , 49, 69.

- Messu Mandeng, F.D., Bilong Bilong, C.F., Pariselle, A., Vanhove, M.P.M., Bitja Nyom, A.R. & Agnès, J.F. (2015). A phylogeny of *Cichlidogyrus* spp. (Monogenea, Dactylogyridea) clarifies a host-switch between fish families and reveals an adaptive component to attachment organ morphology of this parasite genus. *Parasit. Vectors* , 8, 582.
- Molloy, J.C. (2011). The Open Knowledge Foundation: open data means better science. *PLoS Biol.* , 9, e1001195.
- Morand, S., Simková, A., Matejusová, I., Plaisance, L., Verneau, O. & Desdevises, Y. (2002). Investigating patterns may reveal processes: evolutionary ecology of ectoparasitic monogeneans. *Int. J. Parasitol.* , 32, 111–119.
- Mouillot, D., Krasnov, B.R. & Poulin, R. (2008). High intervality explained by phylogenetic constraints in host-parasite webs. *Ecology* , 89, 2043–2051.
- Murtagh, F. & Legendre, P. (2014). Ward’s hierarchical agglomerative clustering method: Which algorithms implement Ward’s criterion? *J. Classif.* , 31, 274–295.
- Nunn, C.L. & Altizer, S.M. (2005). The global mammal parasite database: an online resource for infectious disease records in wild primates. *Evol. Anthropol. Issues, News, Rev.* , 14, 1–2.
- Oliveira, J.B.B.S., Faria, M.L., Borges, M.A., Fagundes, M. & Araújo, W.S. (2020). Comparing the plant–herbivore network topology of different insect guilds in Neotropical savannas. *Ecol. Entomol.* , 45, 406–415.
- Page, R.D.M. (2003). Introduction. In: *Tangled trees. Phylogeny, cospeciation, and coevolution* (ed. Page, R.D.M.). The University of Chicago Press, Chicago, USA & London, UK, pp. 1–21.
- Paperna, I. (1960). Studies on monogenetic trematodes in Israel. 2. Monogenetic trematodes of cichlids. *Bamidgeh* , 12, 20–33.
- Paradis, E. & Schliep, K. (2019). Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* , 35, 526–528.
- Pariselle, A., Morand, S., Deveney, M.R. & Pouyau, L. (2003). Parasite species richness of closely related hosts: historical scenario and “genetic” hypothesis. In: *Taxonomie, écologie et évolution des métazoaires parasites: Livre-hommage à Louis Euzet* (eds. Combes, C., Jourdane, J., Ducreux-Modat, A. & Pages, J.-R.). Presses Universitaires de Perpignan, Perpignan, France, pp. 147–166.
- Patefield, W.M. (1981). Algorithm AS 159: An efficient method of generating random  $R \times C$  tables with given row and column totals. *Appl. Stat.* , 30, 91–97.
- Pavoine, S., Vallet, J., Dufour, A.-B., Gachet, S. & Daniel, H. (2009). On the challenge of treating various types of variables: application for improving the measurement of functional diversity. *Oikos* , 118, 391–402.
- Pennell, M.W., Eastman, J.M., Slater, G.J., Brown, J.W., Uyeda, J.C., Fitzjohn, R.G., *et al.* (2014). Geiger v2.0: an expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. *Bioinformatics* , 30, 2216–2218.
- Pocock, M.J.O., Evans, D.M., Fontaine, C., Harvey, M., Julliard, R., McLaughlin, Ó., *et al.* (2016). The visualisation of ecological networks, and their use as a tool for engagement, advocacy and management. *Adv. Ecol. Res.* , 54, 41–85.
- Poelen, J.H., Simons, J.D. & Mungall, C.J. (2014). Global biotic interactions: an open infrastructure to share and analyze species-interaction datasets. *Ecol. Inform.* , 24, 148–159.
- Poulin, R. (2002). The evolution of monogenean diversity. *Int. J. Parasitol.* , 32, 245–254.
- Poulin, R. (2010). Network analysis shining light on parasite ecology and diversity. *Trends Parasitol.* , 26, 492–498.

- Poulin, R. (2014). Parasite biodiversity revisited: frontiers and constraints. *Int. J. Parasitol.* , 44, 581–589.
- Poulin, R., Krasnov, B.R. & Mouillot, D. (2011). Host specificity in phylogenetic and geographic space. *Trends Parasitol.* , 27, 355–361.
- R Core Team. (2021). *R: a language and environment for statistical computing* . R Foundation for Statistical Computing, Vienna, Austria. Available at: <https://www.r-project.org/>. Last accessed 5 July 2021.
- Rezende, E.L., Lavabre, J.E., Guimarães, P.R., Jordano, P. & Bascompte, J. (2007). Non-random coextinctions in phylogenetically structured mutualistic networks. *Nature* , 448, 925–928.
- Ronco, F., Matschiner, M., Böhne, A., Boila, A., Büscher, H.H., El Taher, A., *et al.* (2021). Drivers and dynamics of a massive adaptive radiation in cichlid fishes. *Nature* , 589, 76–81.
- Royle, J.A. (2004). *N* -Mixture models for estimating population size from spatially replicated counts. *Biometrics* , 60, 108–115.
- Salzburger, W. (2018). Understanding explosive diversification through cichlid fish genomics. *Nat. Rev. Genet.* , 19, 705–717.
- Salzburger, W., van Bocxlaer, B. & Cohen, A.S. (2014). Ecology and evolution of the African Great Lakes and their faunas. *Annu. Rev. Ecol. Evol. Syst.* , 45, 519–545.
- Sasal, P. & Morand, S. (1998). Comparative analysis: a tool for studying monogenean ecology and evolution. *Int. J. Parasitol.* , 28, 1637–1644.
- Sasal, P., Trouve, S., Muller-Graf, C. & Morand, S. (1999). Specificity and host predictability: a comparative analysis among monogenean parasites of fish. *J. Anim. Ecol.* , 68, 437–444.
- Schedel, F.D.B., Musilova, Z. & Schliewen, U.K. (2019). East African cichlid lineages (Teleostei: Cichlidae) might be older than their ancient host lakes: new divergence estimates for the East African cichlid radiation. *BMC Evol. Biol.* , 19, 1–25.
- Seehausen, O. (2006). African cichlid fish: a model system in adaptive radiation research. *Proc. R. Soc. B Biol. Sci.* , 273, 1987–1998.
- Segar, S.T., Fayle, T.M., Srivastava, D.S., Lewinsohn, T.M., Lewis, O.T., Novotny, V., *et al.* (2020). The role of evolution in shaping ecological networks. *Trends Ecol. Evol.* , 35, 454–466.
- Simard, S.W., Beiler, K.J., Bingham, M.A., Deslippe, J.R., Philip, L.J. & Teste, F.P. (2012). Mycorrhizal networks: mechanisms, ecology and modelling. *Fungal Biol. Rev.* , 26, 39–60.
- Šimková, A., Desdevises, Y., Gelnar, M. & Morand, S. (2001). Morphometric correlates of host specificity in *Dactylogyrus* species (Monogenea) parasites of European Cyprinid fish. *Parasitology* , 123, 169–177.
- Šimková, A., Řehulková, E., Rasoloariniaina, J.R., Jorissen, M.W.P., Scholz, T., Faltýnková, A., *et al.* (2019). Transmission of parasites from introduced tilapias: a new threat to endemic Malagasy ichthyofauna. *Biol. Invasions* , 21, 803–819.
- Soares, R.G.S., Ferreira, P.A. & Lopes, L.E. (2017). Can plant-pollinator network metrics indicate environmental quality? *Ecol. Indic.* , 78, 361–370.
- Thébaud, E. & Fontaine, C. (2008). Does asymmetric specialization differ between mutualistic and trophic networks? *Oikos* , 117, 555–563.
- Thompson, R.C.A. (2013). Parasite zoonoses and wildlife: One Health, spillover and human activity. *Int. J. Parasitol.* , 43, 1079–1088.
- Timi, J.T., Rossin, M.A., Alarcos, A.J., Braicovich, P.E., Cantatore, D.M.P. & Lanfranchi, A.L. (2011). Fish trophic level and the similarity of non-specific larval parasite assemblages. *Int. J. Parasitol.* , 41, 309–316.

- Toju, H., Yamamichi, M., Guimarães, P.R., Olesen, J.M., Mougi, A., Yoshida, T., *et al.* (2017). Species-rich networks and eco-evolutionary synthesis at the metacommunity level. *Nat. Ecol. Evol.* , 1, 0024.
- Upham, N., Poelen, J.H., Paul, D.L., Groom, Q., Simmons, N.B., Vanhove, M.P.M., *et al.* (2021). Liberating host-virus knowledge from COVID-19 lockdown [preprint]. *EcoEvoRxiv* , 10.32942/OSF.IO/TXEKQ.
- Vanhove, M.P.M., Hablützel, P.I., Pariselle, A., Šimková, A., Huysel, T. & Raeymaekers, J.A.M. (2016). Cichlids: a host of opportunities for evolutionary parasitology. *Trends Parasitol.* , 32, 820–832.
- Vanhove, M.P.M., Pariselle, A., Van Steenberge, M., Raeymaekers, J.A.M., Hablützel, P.I., Gillardin, C., *et al.* (2015). Hidden biodiversity in an ancient lake: phylogenetic congruence between Lake Tanganyika tropheine cichlids and their monogenean flatworm parasites. *Sci. Rep.* , 5, 13669.
- Vázquez, D.P., Melián, C.J., Williams, N.M., Blüthgen, N., Krasnov, B.R. & Poulin, R. (2007). Species abundance and asymmetric interaction strength in ecological networks. *Oikos* , 116, 1120–1127.
- Vázquez, D.P., Poulin, R., Krasnov, B.R. & Shenbrot, G.I. (2005). Species abundance and the distribution of specialization in host-parasite interaction networks. *J. Anim. Ecol.* , 74, 946–955.
- Vignon, M., Pariselle, A. & Vanhove, M.P.M. (2011). Modularity in attachment organs of African *Cichlidogyrus* (Platyhelminthes: Monogenea: Ancyrocephalidae) reflects phylogeny rather than host specificity or geographic distribution. *Biol. J. Linn. Soc.* , 102, 694–706.
- Vizentin-Bugoni, J., Maruyama, P.K., de Souza, C.S., Ollerton, J., Rech, A.R. & Sazima, M. (2018). Plant-pollinator networks in the tropics: a review. In: *Ecological networks in the tropics* . Springer International Publishing, pp. 73–91.
- Walker, S. (2014). *funphylocom: Functional traits, phylogenies, communities, simulations. R package version 1.1/r211* . Available at: <https://r-forge.r-project.org/projects/multitable>. Last accessed 28 March 2021.
- Wang, P., Xu, B.W., Wu, Y.R. & Zhou, X.Y. (2015). Link prediction in social networks: the state-of-the-art. *Sci. China Inf. Sci.* , 58, 011101:1-011101:38.
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis* . Use R! 2nd edn. Springer, New York, USA.
- Wilson, J.R., Saunders, R.J. & Hutson, K.S. (2019). Parasites of the invasive tilapia *Oreochromis mossambicus* : evidence for co-introduction. *Aquat. Invasions* , 14, 332–349.
- Wu, X.Y., Zhu, X.Q., Xie, M.Q. & Li, A.X. (2006). The radiation of *Haliotrema* (Monogenea: Dactylogyridae: Ancyrocephalinae): molecular evidence and explanation inferred from LSU rDNA sequences. *Parasitology* , 132, 659–668.
- Wu, X.Y., Zhu, X.Q., Xie, M.Q. & Li, A.X. (2007). The evaluation for generic-level monophyly of Ancyrocephalinae (Monogenea, Dactylogyridae) using ribosomal DNA sequence data. *Mol. Phylogenet. Evol.* , 44, 530–544.
- Yu, G., Lam, T.T.-Y., Zhu, H. & Guan, Y. (2018). Two methods for mapping and visualizing associated data on phylogeny using ggtree. *Mol. Biol. Evol.* , 35, 3041–3043.
- Yu, G., Smith, D.K., Zhu, H., Guan, Y. & Lam, T.T.-Y. (2017). ggtree : An R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* , 8, 28–36.
- Zhao, Y., Wu, Y.-J., Levina, E. & Zhu, J. (2017). Link prediction for partially observed networks. *J. Comput. Graph. Stat.* , 26, 725–733.

List of table headers

Table 1. Selection of evolutionary, ecological, and morphological parameters of hosts and parasites used for calculation of host habitat niche dendrogram and network link prediction (NLP) models. Host parameters

were accessed in FishBase (Froese & Pauly 2000) and parasite parameters were reused from Cruz-Laufer *et al.* (2021b). To avoid overfitting NLP models, variable numbers per parameter were reduced through principal coordinate analyses (PCoA) based on distance matrices of phylogenetic trees or *dendrograms* built through clustering methods (see number of PCoA axes used for NLP and their proportion of parameter variation in brackets).

#### List of figure captions

Figure 1. Ecological and evolutionary processes shape the structure of the cichlid-*Cichlidogyrus* network consisting of cichlid fishes, a model system for explosive speciation research, and the parasitic flatworms belonging to *Cichlidogyrus* infecting the gills of cichlid and few non-cichlid fishes. Species presented in the graph are *Coptodon guineensis* (Gunther, 1862) and *Cichlidogyrus gallus* Pariselle & Euzet, 1995.

Figure 2. Cichlid-*Cichlidogyrus* species network. (A) Whole network with unweighted links and for the most species-rich communities ( $n > 10$ ) highlighted with colours. Circles indicate host species and squares species of *Cichlidogyrus*. Meta-communities were detected using the Louvain cluster algorithm including the Lake Victoria (LV), ‘*Coptodon zillii*’ (CZ), ‘*Oreochromis niloticus*’ (ON), ‘*Hemichromis*’ (He), ‘*Ophthalmotilapia*’ (Op) and ‘*Tilapia sparrmanii*’ (TS) cluster. Many small (meta-)communities that fall outside these four groups (C-shaped cluster) are completely unconnected to these large groups or each other. (B) Chord diagrams of the LV cluster, the most species-rich meta-community in the network (in terms of host and parasite species). (C) Five other species-rich meta-communities involving species of *Cichlidogyrus* and *Scutogyrus* with links weighted by number of observed infections communities. Unlike the LV cluster, communities CZ, ON, He, and TS are characterised by sampling bias towards few, economically relevant host species, e.g. *Coptodon zillii*, *Oreochromis niloticus*, *Hemichromis fasciatus*, and *Tilapia sparrmanii*. Species names were omitted from (B) and (C) but are included in Appendix S4.

Figure 3. Changes of network metrics when only including natural host repertoires and geographical ranges of cichlid-*Cichlidogyrus* meta-communities including ‘*Oreochromis niloticus*’ (ON), ‘*Hemichromis*’ (He), and ‘*Coptodon zillii*’ (CZ). Most values that differed significantly from the null distributions (NM<sub>1</sub>, NM<sub>2</sub>) remained unchanged (see Appendix S1.2 for detailed discussion).

Figure 4. Functional-phylogenetic distances (FPDist) inferred from host repertoires of selected species of *Cichlidogyrus* calculated as mean pairwise distance (MPD) and mean nearest taxon distance (MNTD) weighted by abundance of interactions (blue). FPDist matrices are a function of functional (FDist) and phylogenetic (PDist) distance matrices of the host species weighted by the parameter  $a$ . Shaded areas (grey) indicate 5% and 95% quantiles of 1000 null distributions resulting from taxon shuffling. If estimates fall outside the null distribution, they can be considered informative. Smaller values indicated higher functional-phylogenetic similarities of host repertoires. A decreasing trend for FPDist estimates indicates that host communities are more phylogenetically than ecologically similar. For plots of other species infecting at least two host species, see Appendix S6.

Figure 5. Network link prediction based on host [H] and parasite [P] data in the cichlid-*Cichlidogyrus* network including missingness map of input variables (a), and heat map of host-parasite links with rows and columns order by numbers of observed interactions (b) and bar plot of variable importance (c) predicted by the *plug-and-play* algorithm Dallas *et al.* (2017). Missingness map illustrates significant gaps in the taxon coverage of phylogenetic data and host standard lengths. Heat map shows that a large proportion of host-parasite interactions likely remain undetected (highlighted in colour) (for taxon labels, see Appendix S7). Variable importance graph indicates that the ecosystem of the hosts is the most important predictor of cichlid-*Cichlidogyrus* interactions.

#### Supporting information

Appendix S1. Other methods and results including phylogenetic reconstruction and structure of species-rich meta-communities in the cichlid-*Cichlidogyrus* system.

Appendix S2. GenBank accession numbers of DNA sequences used to render host phylogenetic distances.

Appendix S3. Host niche dendrograms resulting from different clustering algorithms.

Appendix S4. Chord diagrams of six most species-rich meta-communities presented in Fig. 2 with additional species labels. Host species names are abbreviated with the first three letters of the genus name and the first four letters of the species epithet. Parasite species names are abbreviated with the first and first six letters respectively.

Appendix S5. Structural and phylogenetic host specificity indices of species of *Cichlidogyrus* with more than four infected hosts reported in peer-reviewed literature. Significance of these indices was tested against null models NM<sub>1</sub>, NM<sub>2</sub>, and NM<sub>3</sub>. Structural specificity is measured as specialisation index  $d_i'$ , phylogenetic specificity is measure as z-scores (standardised effect size) of average mean pairwise distance (MPD) and mean nearest taxon distance (MNTD) of 100 parasite BI tree topologies randomly selected from the post-burn in fraction.

Appendix S6. Functional phylogenetic distance (FPDist) plots of host repertoires of all species of *Cichlidogyrus* not included in Fig. 4.

Appendix S7. Heat map of links predicted by the *plug-and-play* algorithm with complete taxon labels. See Fig. 5c for simplified version.

Table 1. Selection of evolutionary, ecological, and morphological parameters of hosts and parasites used for calculation of host habitat niche dendrogram and network link prediction (NLP) models. Host parameters were accessed in FishBase (Froese & Pauly 2000) and parasite parameters were reused from Cruz-Laufer *et al.* (2021). To avoid overfitting NLP models, variable numbers per parameter were reduced through principal coordinate analyses (*PCoA*) based on distance matrices of phylogenetic trees or *dendrograms* built through clustering methods (see number of *PCoA* axes used for NLP and their proportion of parameter variation in brackets).

Organism	Variable	Type	Values	Host niche dendrogram	Network link prediction	Source
Host	Phylogeny	distance matrix	–		PCoA+ (5 axes: 88.5%)	Appendix S1
	Ecosystem	binary, multi-level*	Presence/absence in different basins and lakes		dendrogram + PCoA (including ‘Habitat’) (5 axes: 95.6%)	FishBase
	Habitat	binary, multi-level*	Presence/absence in stream and lake habitats		Included with ‘Ecosystem’	FishBase
	Life style	categorical	‘benthopelagic’ ‘semipelagic’ ‘pelagic’			FishBase
	Trophic level	categorical	‘mainly plant/detritus-feeding’ ‘plant/detritus and animal-feeding’ ‘mainly animal-feeding’			FishBase

Organism	Variable	Type	Values	Host niche dendrogram	Network link prediction	Source
Parasite	Standard length	continuous	> 0			FishBase
	Phylogeny	distance matrix	–		PCoA+ (5 axes: 78.0%)	Cruz-Laufer <i>et al.</i> (2021)
	Attachment organ morphology	multiple continuous variables	> 0		dendrogram + PCoA (3 axes: 88.3%)	Cruz-Laufer <i>et al.</i> (2021)
	Reproductive organ morphology	multiple continuous variables	> 0		dendrogram + PCoA (3 axes: 87.3%)	Cruz-Laufer <i>et al.</i> (2021)

\* more than level can be true at the same time, + of Bayesian majority-rule consensus tree

**Hosted file**

Fig1\_overview.eps available at <https://authorea.com/users/438073/articles/539356-the-role-of-phylogeny-and-ecological-opportunity-in-host-parasite-interactions-network-metrics-host-repertoire-and-network-link-prediction>



figures/Fig3-NetworkMetrics/Fig3-NetworkMetrics-eps-converted-to.pdf

figures/Fig4-FPDist/Fig4-FPDist-eps-converted-to.pdf

figures/Fig5-NLP/Fig5-NLP-eps-converted-to.pdf