



**HAL**  
open science

# Mécanismes d'infection de l'ectoparasite de poissons d'eau douce *Tracheliastes polycolpus*

Eglantine Mathieu-Bégné

► **To cite this version:**

Eglantine Mathieu-Bégné. Mécanismes d'infection de l'ectoparasite de poissons d'eau douce *Tracheliastes polycolpus*. Biologie animale. Université Paul Sabatier - Toulouse III, 2020. Français. NNT : 2020TOU30008 . tel-02972144

**HAL Id: tel-02972144**

**<https://theses.hal.science/tel-02972144>**

Submitted on 20 Oct 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# THÈSE

## En vue de l'obtention du DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Université Toulouse 3 - Paul Sabatier

---

Présentée et soutenue par  
**Eglantine MATHIEU-BEGNE**

Le 20 janvier 2020

**Mécanismes d'infection de l'ectoparasite de poissons d'eau douce  
Tracheliastes polycolpus**

---

Ecole doctorale : **SEVAB - Sciences Ecologiques, Vétérinaires, Agronomiques et  
Bioingenieries**

Spécialité : **Ecologie, biodiversité et évolution**

Unité de recherche :  
**EDB - Evolution et Diversité Biologique**

Thèse dirigée par  
**Géraldine LOOT et Olivier REY**

Jury

Mme Nathalie CHARBONNEL, Rapporteur  
Mme Nadia AUBIN-HORTH, Rapporteur  
M. Thierry RIGAUD, Examineur  
M. Guillaume MITTA, Examineur  
M. Philipp HEEB, Examineur  
Mme Géraldine LOOT, Directrice de thèse  
M. Olivier REY, Co-directeur de thèse



## Résumé

L'identification des mécanismes à l'origine des interactions entre espèces est essentielle afin de mieux appréhender les conséquences des changements globaux sur la biodiversité et le fonctionnement des écosystèmes. Au cours de ces travaux de thèse, je me suis ainsi attachée à identifier ces mécanismes en me basant sur les concepts théoriques de filtre de rencontre et de filtre de compatibilité. J'ai considéré les facteurs environnementaux qui favorisent le contact entre hôte et parasite, mais aussi les facteurs intrinsèques à l'hôte et/ou au parasite qui régissent leur compatibilité.

J'ai évalué le rôle de ces facteurs dans chacun de ces filtres théoriques en utilisant comme modèle l'interaction entre l'ectoparasite *Tracheliastes polycolpus* et les espèces piscicoles d'eau douce connues pour être des hôtes potentiels. Ma thèse s'articulait autour de quatre objectifs : i) tester à l'échelle micro-géographique le lien entre les caractéristiques environnementales et les patrons d'infection parasitaire via une approche expérimentale *in situ*, combinée à une approche descriptive des populations naturelles ; ii) tester à l'échelle macro-géographique (France) la part des facteurs environnementaux et du patrimoine génomique des populations hôtes dans la répartition actuelle de *T. polycolpus* ; iii) tester le rôle de la plasticité transcriptomique dans la capacité de *T. polycolpus* à exploiter différentes espèces hôtes et iv) évaluer le rôle du microbiote (associé à l'hôte, au parasite, et à l'environnement commun) dans le processus d'infection parasitaire au travers d'une approche expérimentale *in situ*.

J'ai d'abord montré qu'il existait des points chauds d'infection à l'échelle micro-géographique et que ces derniers dépendaient des conditions environnementales. La structure de l'environnement micro-géographique pourrait ainsi être un mécanisme général qui conditionne la rencontre entre hôtes et parasites. A contrario, à l'échelle macro-géographique, j'ai montré que la répartition actuelle de *T. polycolpus* en France est expliquée plutôt par le patrimoine génomique des populations hôtes que par les conditions environnementales. La compatibilité entre un hôte et son parasite pourrait ainsi limiter la dynamique d'invasion des parasites émergents. J'ai aussi montré que la plasticité transcriptomique jouait un rôle essentiel dans la capacité de *T. polycolpus* à exploiter différentes espèces hôtes (i.e., compatibilité multi-hôte). Enfin, j'ai montré que *T. polycolpus* induisait un changement de composition du microbiote de son hôte *L. burdigalensis*. De plus, les changements de composition du microbiote de l'hôte résultaient d'une dynamique de co-infection par *T. polycolpus* et son microbiote associé.

Au cours de ma thèse l'accent a été mis sur une vision intégrative des mécanismes de l'infection parasitaire. Par le déploiement d'approches variées et complémentaires (descriptive, expérimentale, génomique), de nouveaux éléments ont été amenés concernant la compréhension des interactions hôte-parasites et des mécanismes qui modulent leurs succès. Une meilleure compréhension des mécanismes sous-jacents aux différentes interactions est un prérequis pour prédire la dynamique de ces systèmes dans le contexte actuel des changements globaux.

## Abstract

The identification of the mechanisms underlying species interactions is essential to understand the consequences of global changes on biodiversity and ecosystem functioning. During this thesis I worked at identifying these mechanisms based on the theoretical concepts of encounter filter and compatibility filter. I considered the environmental factors favoring the contact between host and parasite, but also host- and parasite-related intrinsic factors that govern their compatibility.

I tested the role of these factors in each of these theoretical filters using as study model the interaction between the ectoparasite *Tracheliastes polycolpus* and freshwater fish species known to be potential hosts. My thesis was based on four objectives: i) testing at a microgeographical scale the link between environmental features and infection patterns via an experimental *in situ* approach, combined with a descriptive approach in natural populations; ii) testing at the macrogeographical scale (France) the part of environmental factors and host genomic background in the current distribution of *T. polycolpus*; iii) testing the role of transcriptomic plasticity in the ability of *T. polycolpus* to exploit different host species and iv) assessing the role of the microbiota (associated with host, parasite, and common environment) in the process of parasitic infection through an experimental approach *in situ*.

First, I identified hot spots of infection at the microgeographical scale in relation with environmental conditions. I suggest that the existence of microhabitats characterized by congruent environmental conditions could be a general mechanism that conditions the encountering rate of parasites. On the other hand, at the macrogeographical scale, I showed that the current distribution of *T. polycolpus* in France is more likely explained by host genomic background rather than the environmental conditions. The compatibility between hosts and parasites could hence limit the invasion potential of emerging parasites. I have also shown that transcriptomic plasticity plays a key role in the ability of *T. polycolpus* to exploit different host species (i.e., compatibility multi-host). Finally, I showed that *T. polycolpus* induces change in microbiota composition of its host *L. burdigalensis*. In addition, changes in host-associated microbiota composition likely reflected a co-infection dynamics between *T. polycolpus* and its associated microbiota.

During my thesis the emphasis was put on an integrative vision of the mechanisms of parasitic infection. Through the deployment of diverse and complementary approaches (descriptive, experimental, genomic), new elements have been brought to the understanding of host-parasite interactions and the mechanisms that modulate their success. A better understanding of the mechanisms underlying the different interactions is a prerequisite for predicting the dynamics of these systems in the current global changes context.

## Avant-propos

Cette thèse de l'école doctorale SEVAB de l'Université Toulouse III a été réalisée au laboratoire Evolution et Diversité Biologique à Toulouse, à la Station d'Ecologie Théorique et Expérimentale du CNRS de Moulis et au laboratoire Interaction Hôte-Pathogène-Environnement à Perpignan, sous la supervision de Géraldine LOOT, Simon BLANCHET et Olivier REY. Elle s'articule autour d'une introduction générale, de quatre chapitres sous forme d'articles scientifiques soumis ou en préparation, d'une discussion générale et d'une série d'annexes sous forme d'articles en préparation ou bien publiés. Cette thèse est rédigée en français à l'exception des chapitres et des annexes qui sont écrits en anglais.

Au total, cette thèse s'organise autour de 7 articles scientifiques:

- Article 1 (Chapitre I) : il s'agit d'une étude du rôle des facteurs environnementaux à l'échelle micro-géographique sur la rencontre entre un hôte et son parasite. Cet article est en préparation et fera l'objet d'une soumission dans une revue de parasitologie.
- Article 2 (Chapitre II) : il s'agit d'une étude des déterminants environnementaux et génomiques dans la distribution spatiale d'un parasite invasif. Ces travaux ont fait l'objet de deux communications orales : sous la forme d'une présentation orale au Réseau d'Etude des Interactions Durables à Toulouse en 2017 et sous la forme d'un poster lors du *Second Joint Congress on Evolutionary Biology* à Montpellier en 2018. Cet article fait l'objet d'une re-soumission dans *Ecography*.
- Article 3 (Chapitre III) : il s'agit d'une étude du rôle de la plasticité transcriptomique dans la capacité d'un parasite à exploiter différentes espèces hôtes. Cet article est en préparation et sera très prochainement soumis dans *Plos Pathogens*.
- Article 4 (Chapitre IV) : il s'agit d'une étude du rôle du microbiote (associé à l'hôte, au parasite et à l'environnement commun) dans le processus d'infection parasitaire. Cet article est en préparation. Il fera l'objet d'analyses complémentaires et d'un travail de synthétisation au niveau de l'écriture avant soumission dans *Nature Ecology and Evolution*.
- Article 5 (Annexe I) : il s'agit d'une étude intégrative de l'impact d'un parasite sur son hôte à différents niveaux d'organisation biologique (du niveau moléculaire à populationnel). Ces travaux ont fait l'objet d'une communication orale sous la forme d'une présentation lors du congrès de la Société Française d'Ecologie à Rennes en 2019. Cet article est en préparation et fera l'objet d'une soumission dans *Ecological Monograph*.
- Article 6 (Annexe II) : il s'agit d'une note technique sur l'assemblage du transcriptome de *Tracheliastes polycolpus*. Cet article est paru dans *Marine Genomics*.
- Article 7 (Annexe III) : Il s'agit d'un article portant sur le suivi temporel démographique et génétique sur dix ans de deux métapopulations de vandoises. Cet article est paru dans *Oikos*.



## Remerciements

Cette thèse fut un travail de longue haleine mais fort heureusement je ne fus pas seule sur la route et je tiens donc à remercier toutes les personnes qui m'ont aidé de près ou de moins près.

Mes premiers remerciements vont à Simon Blanchet et Géraldine Loot qui m'ont non seulement supervisée tout au long de la thèse mais qui avait déjà commencé lors de mes premiers stages au sein de l'équipe en M1. Je tiens à vous remercier pour le soutien que vous m'avez toujours apporté ainsi que pour m'avoir poussée à fournir le meilleur de moi-même et tout cela avec beaucoup d'humanité et d'humour. Je ne vous remercierai jamais assez pour la confiance que vous m'avez accordée ainsi que pour le bel exemple d'encadrement que vous m'avez fourni et que j'espère mettre en pratique dans le futur. Je souhaite ensuite remercier Olivier Rey qui m'a co-encadrée sur ce projet de thèse. Travailler et discuter avec toi a toujours été très stimulant et me donne une énergie folle pour aller de l'avant. Je suis vraiment très heureuse que nous puissions continuer à travailler ensemble sur un nouveau projet de parasitologie.

En second lieu je tiens à remercier toutes les personnes sans qui la réalisation des travaux présentés dans cette thèse n'aurait pas été possible. D'abord merci aux personnes de l'équipe au moment de ma thèse à savoir Charlotte Veyssière, Kéoni Saint-Pé, Allan Raffard, Jérôme Prunier, Éloïse Duval, Lisa Fourtune, Ivan Paz-Vinas, Lucie Di Gesu, qui m'ont aidé sur les sessions terrains et/ou sur certaines analyses. Un grand merci notamment à Charlotte Veyssière qui a accompli un travail de taille avec Sophie Manzi en réalisant la plupart des extractions. Je remercie également les stagiaires que j'ai eu le plaisir d'encadrer et qui ont travaillé soit sur l'analyse de mes données de thèse tels que Clément Le Potier et Orlane Scelsi, mais également ceux qui ont grandement contribué à la réalisation des manips de caging et de pêche microhabitat, Geoffrey Marselli, Camille Poesy, Yoann Buoro, Orlane (again) et Maxim Lefort. Mention spéciale à Camille et Geoffrey avec qui nous avons construit les près de 40 cages de la manip et à Orlane et Maxim avec qui nous avons récidivé l'année suivante. Enfin merci à Lisa Jacquin et Jessica Côte pour les conseils sur les swabs, à Murielle Richard, Michelle Huet pour l'aide apporté sur la manip bactériologique, à Antoine Gallard de l'INRA de Rennes pour nous avoir fourni les vandoises des expériences et Thierry Paulin pour avoir gentiment accepté de nous laisser conduire ces expériences de caging sur son terrain.

Je souhaite aussi remercier les personnes ayant participé à mes comités de thèse, qui m'ont fourni de judicieux avis et conseils et qui ont permis d'améliorer grandement les travaux présentés dans cette thèse. Je remercie donc Jean-Baptiste Ferdy, Carine Brouat et tout particulièrement Eve Toulza qui en outre m'a aidée sur les analyses transcriptomiques mais aussi qui s'est beaucoup investie sur le projet bactériologique.

Je tiens ensuite à dire un grand merci à mes colocataires de Montégut, Julien Cote, Serge (Laurane Winandy), Jeanne (Félix Pellerin), Mat (Mathilde Mousset), Elvire Bestion et les aoutats avec qui j'ai vécu pendant les saisons de terrain. On a passé plus que de bons moments et avec vous j'ai appris plein de choses que je n'aurais pas imaginées, comme se servir d'un lave-vaisselle ou d'une essoreuse à salade, faire de la salade fruit... C'est fou quand même la vie ! Je ne peux pas remercier mes colocataires de Montégut sans remercier Maryse, Christian et Gaby Delon. Merci pour votre générosité et l'accueil merveilleux que j'ai reçu là-bas.



De manière générale, je tiens aussi à remercier tous mes collègues Moulisssiens dont j'en connais certains depuis maintenant un bout de temps. Merci Nicolas Canto, Aisha Bruendl, Louis Sallé, Anne-Sophie Auguères, Elise Mazé-Guilmo, Jérémie Souchet, Hugo Le Chevalier, Rik Verdonck, Jonathan Remon, Elodie Darnet (j'espère que je n'oublie personne ...). Certains ont été mes premiers contacts avec cette étrange station du CNRS en Ariège et m'ont donné l'envie d'y revenir. Et avec tous j'ai passé de formidables moments d'humour et de partage que ce soit aux tablés de Moulis, à la fête de Montégut, ou à l'occasion de balades/randonnées, et évidemment aux nombreuses soirées que je ne risque pas d'oublier.

Ensuite je souhaite remercier mes co-galériens de Toulouse à EDB. D'abord mes anciens collègues de bureau Seb Cally et Marine Cambon, merci pour la bonne humeur et les délires dans le bureau 26 (y compris pour les insects farfelus). Je remercie aussi mes nouveaux collègues de bureau, Maëva Gabrielli et Sean Heighton qui assurent la relève. Je tiens à dire que maintenant nous avons la plus belle affichette de bureau du labo. Merci aussi à Maxime Pineaux, Luana Bourgeaud, Isabelle Cantera et Fabian Fisher j'ai vraiment senti qu'on s'est serré les coudes sur cette troisième année et ça c'était vraiment génial. Merci à Alexandre Réjaud pour avoir initié tant de soirées Dubli/fléchettes. Merci aussi à Kévin Cilleros, Céline Van de Paer, Jade Bruxaux, Julian Donald, E-Ping Rau, Juan Carvajal-Quintero, Quentin Petitjean, Iris Lang, Thibaut Rota. Merci à Dominique Pantalacci qui plus qu'assurer le travail de secrétaire de l'école doctorale de Sevab sait si bien soutenir et réconforter les thésards dans les moments difficiles et/ou administratifs de la thèse. De ces années passées à EDB (ou au 4R1 plutôt) je pense que je vais retenir la solidarité, la bonne humeur, les films de thèses et l'impression d'être un peu comme chez soi au boulot.

Enfin, je tiens à remercier mes amis et mes proches pour leur soutien. Notamment, merci à Claire et Priscilla Roberdeau pour les sessions D&D, merci à Eve Dussert pour les soirées Barakas et à Andreia Moreira pour les discussions autour d'un verre: vous avez été et continuez à être une bouffée d'air et de fun nécessaire pour moi. Merci aussi à ma famille qui maintenant commence à bien connaître les parasites et les vandoises à cause de moi. Un grand merci à ma sœur, Bérengère, qui m'a aidé sur tant d'aspects. Enfin, merci à Emmaus Goetz qui m'a soutenu durant les mois (parfois difficiles) de la fin de thèse et qui me rend plus heureuse chaque jour.

Enfin, je tiens à remercier les membres du jury qui ont accepté d'évaluer ces travaux de thèse.

## Table des matières

Résumé .....	1
Abstract .....	2
Avant-propos .....	3
Remerciements .....	5
Introduction .....	15
La place des interactions du vivant dans la biodiversité .....	15
Maintien de la biodiversité et rôles des interactions biologiques.....	15
Variabilité temporelle et spatiale des interactions.....	18
Appréhender les interactions biologiques par un retour aux mécanismes .....	19
Spécificité des interactions parasitaires.....	20
La première loi du parasitisme .....	20
Les impacts du parasitisme .....	22
Mécanismes d'infection et maladies émergentes .....	24
Les concepts de filtre de rencontre et de compatibilité : un cadre pour la compréhension des mécanismes d'infection.....	27
Le filtre de rencontre .....	27
Le filtre de compatibilité.....	29
Pourquoi les notions de filtres de rencontre et de compatibilité sont les clefs de la compréhension de l'agrégation parasitaire ?.....	31
Les facteurs pouvant moduler l'ouverture des filtres de rencontre et de compatibilité .....	31

Les facteurs environnementaux.....	32
Les facteurs liés à l'hôte.....	33
Les facteurs liés au parasite.....	34
De l'interaction triple à l'interaction multiple : le rôle du microbiome .....	36
Objectifs .....	38
(1) Rôle de l'environnement micro-géographique dans la rencontre entre un parasite et son hôte .....	38
(2) Rôle des déterminants environnementaux et génomiques dans la répartition d'un parasite émergent.....	39
(3) Rôle de la plasticité dans la capacité d'un parasite à exploiter plusieurs espèces hôtes	40
(4) Rôle des communautés bactériennes dans le processus d'infection parasitaire .....	40
Le modèle d'étude.....	41
Cycle de vie.....	41
Effets directs et indirects de <i>T. polycolpus</i> .....	43
Dynamique d'invasion des bassins versant français .....	45
Chapitre I : Rôle de l'environnement micro-géographique dans la rencontre entre un parasite et son hôte .....	47
Résumé .....	48
Abstract .....	50
Introduction .....	51

Material and methods .....	55
Microhabitat description .....	55
Experimental design and sampling .....	56
Statistical analysis .....	61
Results .....	64
Distribution of infection events .....	64
Spatial distribution of parasitized hosts .....	64
Effect environmental factors at the microgeographic scale on infection risks .....	65
Discussion .....	71
References .....	75
Supplementary.....	80
Chapitre II : Rôle des déterminants environnementaux et génomiques dans la répartition d'un parasite émergent.....	85
Résumé .....	86
Abstract .....	89
Introduction .....	90
Material and Method .....	93
Biological models.....	93
Testing the environmental suitability hypothesis.....	95
Testing the role of the genetic background. ....	101

Results .....	104
Testing the environmental suitability hypothesis.....	104
Testing the role of the genetic background. ....	107
Discussion .....	110
References .....	114
Supplementary.....	119
Chapitre III : Rôle de la plasticité dans la capacité d'un parasite à exploiter plusieurs espèces hôtes .....	
Résumé .....	131
Abstract .....	132
Introduction .....	135
Material and Method .....	136
Study model.....	139
Sampling design and sequencing .....	139
Gene expression analyses.....	141
Variant calling and FST outlier analysis .....	141
Results .....	144
Differentially expressed genes .....	145
Functional analyses .....	148
Detection of selection.....	150

Discussion .....	151
References .....	155
Supplementary.....	160
Chapitre IV : Rôle des communautés bactériennes dans le processus d'infection parasitaire.....	163
Résumé .....	164
Abstract .....	166
Introduction .....	167
Material and Method .....	170
Biological history of the interaction.....	170
Microbiota survey of fish fins and environment .....	171
Microbiota of parasites.....	172
DNA extraction and 16S rRNA gene amplicon sequencing.....	173
Bioinformatic analysis and filtering.....	174
Statistical analysis .....	175
Results .....	178
Hypothesis 1: Fin microbiota as a cause of parasite infection .....	178
Hypothesis 2: Host microbiota composition as a consequence of parasite infection.....	181
Potential source for host microbiota changes after infection .....	183
Discussion .....	185
References .....	189

Supplementary.....	193
Discussion générale et Perspectives .....	207
Les mécanismes de l'interaction parasitaire : effets environnementaux, génomiques et effets des communautés bactériennes. ....	207
Le rôle de l'environnement à l'échelle micro-géographique sur la probabilité de rencontre .....	207
Le rôle des facteurs génomiques de l'hôte et du parasite sur la compatibilité.....	209
Rôle des communautés bactériennes.....	212
Les mécanismes d'infection dans le contexte de l'agrégation parasitaire et des maladies émergentes.....	214
Mécanismes d'infection et agrégation parasitaire.....	214
Mécanismes d'infection et maladies émergentes .....	217
Echelle de travail et perspective holistique dans l'étude des interactions hôte-parasites .....	222
Les interactions hôte-parasite à différentes échelles spatiales .....	222
Les interactions hôte-parasite à différents niveaux d'organisation biologique.....	222
Les interactions hôte-parasite au-delà de l'interaction bipartite .....	223
Vers une vision holistique et intégrative des interactions hôte-parasites.....	224
Les notions de filtre de rencontre et de compatibilité par-delà les interactions hôte-parasite .....	225
Conclusion.....	229
Bibliographie Générale .....	231





*Que ce monde demeure,  
Que la feuille parfaite  
Ourle à jamais dans l'arbre  
L'imminence du fruit !*

*Que les huppes, le ciel  
S'ouvrant, à l'aube,  
S'envolent à jamais, de dessous le toit  
De la grange vide,*

*Puis se posent, là-bas  
Dans la légende,  
Et tout est immobile  
Une heure encore.*

Yves Bonnefoy *Les planches courbes*

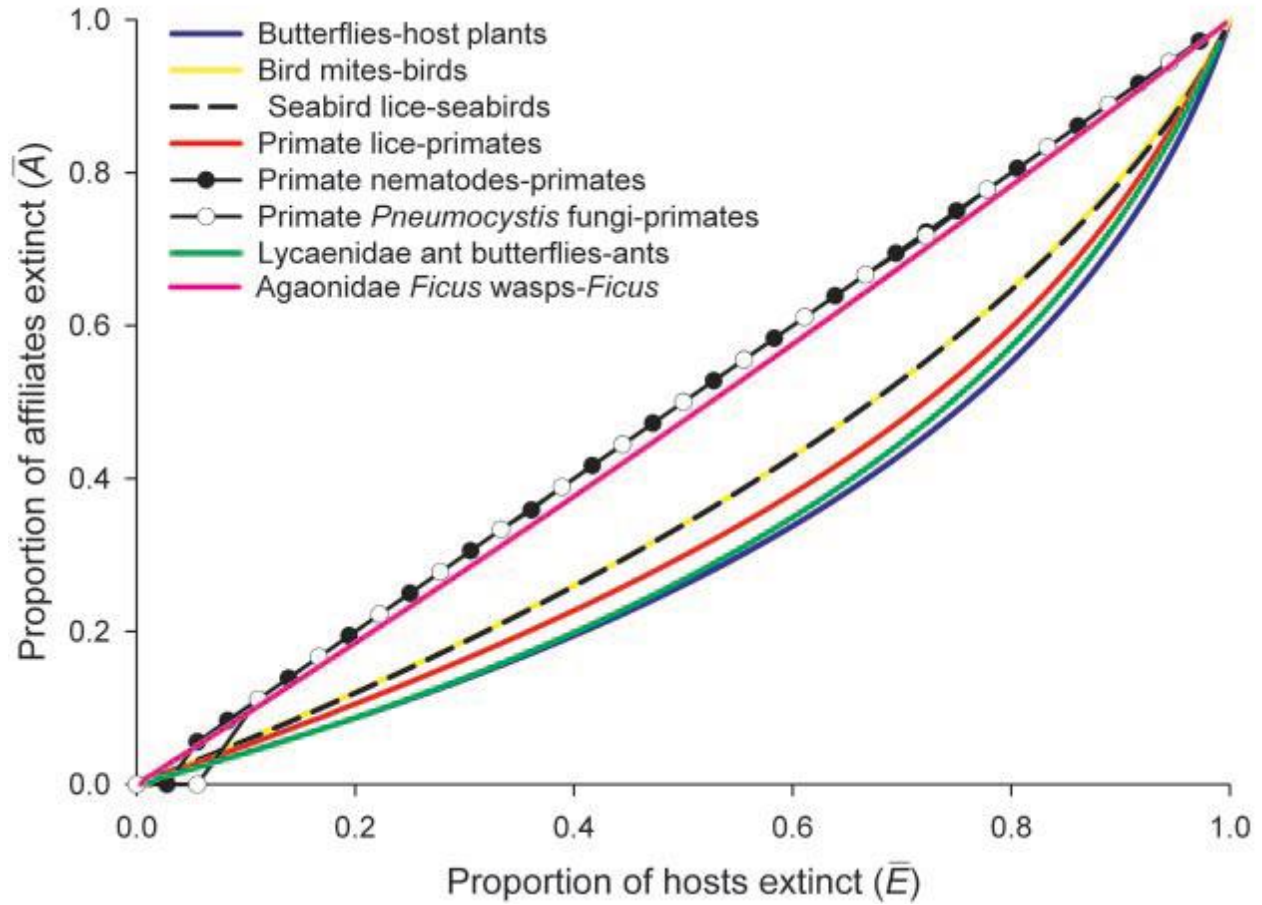
# **Introduction**

## **La place des interactions du vivant dans la biodiversité**

### **Maintien de la biodiversité et rôles des interactions biologiques**

Au-delà d'une compréhension fondamentale du monde vivant qui nous entoure, l'étude de la biologie se confronte actuellement à des enjeux majeurs qui découlent des profondes modifications qui s'opèrent à l'échelle planétaire et s'exacerbent depuis maintenant environ un siècle. Ces bouleversements globaux se mesurent, en terme de modification du climat, mais également en terme d'augmentation de l'urbanisation, de la fragmentation des habitats ou encore de la croissance des populations humaines qui conditionne en partie ces changements (Hansen et al. 2006, Grimm et al. 2008, Schwarzkopf and Ramaswamy 2008). Un des défis majeurs en biologie depuis plusieurs décennies se rapporte à notre capacité à établir des prédictions précises quant aux changements globaux à venir et leurs répercussions sur la biodiversité et le fonctionnement des écosystèmes (Guisan and Thuiller 2005, Parmesan 2006, Bellard et al. 2012, Ward et al. 2014). Pendant longtemps, les conséquences des changements globaux sur la biodiversité ont été évaluées en considérant les espèces de manière isolée (Parmesan and Yohe 2003, Parmesan 2006, Bellard et al. 2012). Cependant cette vision espèce centrée achoppe à prendre en compte un processus essentiel à la persistance des espèces : les interactions biologiques. En effet, les espèces ne sont pas des entités isolées mais sont au contraire en constante interaction les unes avec les autres (Koh et al. 2004, Araújo and Luoto 2007, Wisz et al. 2013). Un écosystème riche en interactions (compétition, prédation, mutualisme, commensalisme ou bien parasitisme) se caractérise par un niveau de diversité, taxonomique et fonctionnel plus élevé (Bascompte et al. 2006, Tylianakis et al. 2008), mais est aussi généralement plus stable et

plus résilient face à des conditions environnementales changeantes (Brooker 2006, Suttle et al. 2007, Ives and Carpenter 2007). Outre le fonctionnement des écosystèmes, la survie de nombreuses espèces dépend des interactions qu'elles entretiennent. Par exemple, en se basant sur un modèle probabiliste visant à estimer le nombre d'extinctions d'espèces associées à l'extinction d'une autre espèce pour une large gamme d'association du vivant, Koh et al. (2004) ont estimé que près de 6 300 espèces seraient menacées de co-extinction (Fig. 1). De manière similaire Valiente-Banuet et al. (2015), en établissant un modèle prédictif qui lie la diversité à la fois en terme d'espèces et d'interactions le long de gradients environnementaux, ont montré que la perte des interactions biologiques précéderait l'extinction des espèces. Cette dette d'extinction des interactions écologiques (*extinction debt of ecological interaction*, Valiente-Banuet et al. 2015) souligne l'importance des interactions entre espèces dans le maintien général de la biodiversité et par conséquent la nécessité de leur prise en compte dans les modèles prédictifs (Blois et al. 2013, Poisot et al. 2015, De Kort et al. 2018).



**Figure 1** : Proportion d'espèces susceptibles de s'éteindre via un processus de co-extinction pour une proportion donnée d'espèces hôtes éteintes, reproduit d'après Koh et al. (2004).

## Variabilité temporelle et spatiale des interactions

La complexité des interactions biologiques rend néanmoins difficile leur intégration dans des modèles prédictifs. D'une part, dans une communauté donnée, toutes les interactions potentielles ne se réalisent pas (Olesen et al. 2011, González-Varo and Traveset 2016). Par exemple, Olesen et al. (2011) ont montré que dans trois réseaux d'interactions mutualistes (i.e., un réseau plantes-pollinisateurs arctique et deux réseaux plantes-disperseurs de graines méditerranéens), 48% à 76% des interactions potentielles n'étaient pas observées, et 22% à 25% ne se réalisaient pas à cause de la non-concordance entre la phénologie des plantes (eg., période de floraison et de fructification) et des pollinisateurs/disperseurs (eg., période de visite et de nutrition). Ainsi, la co-occurrence entre espèces est une condition nécessaire mais insuffisante à la réalisation des interactions, de sorte qu'il est difficile de prédire si deux espèces présentes dans le même milieu entreront en interaction (Poisot et al. 2012). D'autre part, pour deux espèces dont l'interaction a été observée, il existe une variabilité temporelle et/ou spatiale, donnant un caractère labile à l'interaction (González-Varo and Traveset 2016). Par exemple, Trøjelsgaard et al. (2015) ont montré que dans les réseaux d'interactions plantes-pollinisateurs au sein des îles Canari, la fidélité à un partenaire donné diminue avec la distance géographique. De la même manière, les interactions proies-prédateurs peuvent montrer une variabilité temporelle à l'échelle de la vie des individus. Par exemple, dans l'interaction entre *Iphiseius degenerans* et *Neoseiulus cucumeris*, deux espèces de mites Méditerranéennes, les adultes et les stades juvéniles des deux espèces se nourrissent des juvéniles de l'autre espèce (Choh et al. 2012). Les stades les plus gros utilisent les plus petits comme proies, ce qui donne un caractère situationnel aux désignations de proies ou prédateurs pour les individus de chaque espèce au cours de leur vie (Choh et al. 2012). En conséquence, il est primordial de prendre en compte les interactions biologiques pour anticiper

les impacts des changements globaux sur la biodiversité, mais prédire si deux espèces ou plus entreront en interaction s'avère une tâche ardue.

### **Appréhender les interactions biologiques par un retour aux mécanismes**

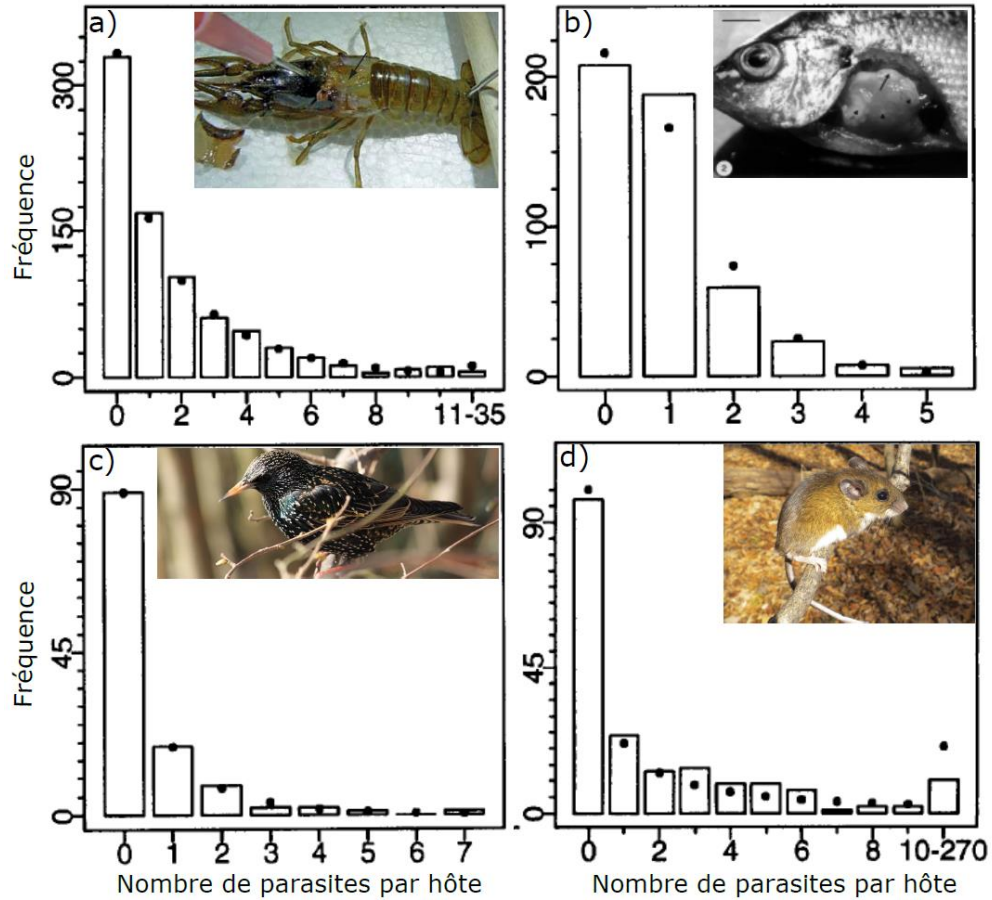
Comprendre les mécanismes écologiques et évolutifs à l'origine des interactions biologiques est nécessaire afin de mieux prédire les interactions futures. Autrement dit, pourquoi deux espèces à un endroit et à un moment donné entrent en interaction ? Une condition nécessaire, bien que non suffisante, à l'interaction directe de n'importe quel couple d'espèces est la concomitance spatio-temporelle des deux espèces (Poiso et al. 2015). Ainsi, Araujo et al. (2007) ont montré que les modèles de prédiction de la distribution du papillon Semi-Apollon en Europe étaient plus précis lorsque les distributions spatiales des plantes hôtes utilisées par les stades larvaires (*Corydalis solida*, *C. intermedia*, *C. cava* et *C. pumila*) étaient prises en compte. Toutefois la concomitance spatio-temporelle n'est pas le seul mécanisme permettant d'expliquer la réalisation d'une interaction. La réalisation des différents types d'interactions dépend également des exigences phénologiques, morphologiques, physiologiques et biochimiques des acteurs de l'interaction (González-Varo and Traveset 2016, Bartomeus et al. 2016). Par exemple Brose et al. (2006) ont étudié 5000 interactions proie-prédateur et montrent que les prédateurs sont en moyenne 42 fois plus grands que leurs proies, rendant ainsi l'ingestion possible. De même, les défenses chimiques d'une proie et la performance des enzymes digestives d'un prédateur, peuvent empêcher ou favoriser leur interaction (Feyereisen 1999, Eisner et al. 2005, Brousseau et al. 2018). Il est donc nécessaire d'étudier de manière approfondie l'ensemble des facteurs pouvant moduler le succès des interactions. Les interactions étroites telles que le mutualisme et le parasitisme offrent tout particulièrement un cadre propice à l'étude des mécanismes sous-jacents aux interactions biologiques en cela que les conditions pour la réalisation de l'interaction sont plus strictes. Au

cours de cette thèse je me suis ainsi focalisée sur une interaction hôte-parasite afin d'en étudier les mécanismes.

## **Spécificité des interactions parasitaires**

### **La première loi du parasitisme**

Le mode de vie parasitaire (i.e., l'exploitation durable par un organisme appelé parasite, d'un autre organisme appelé hôte, Crofton 1971, Poulin 2007) est extrêmement répandu dans le monde vivant (Price 1980). Une étude propose en effet que 30% des espèces au sein des eucaryotes sont des parasites et reconnaît très probablement sous-estimer cette proportion (de Meeûs and Renaud 2002). Tout organisme est amené à héberger une ou plusieurs espèces parasites au cours de sa vie (Combes 2005). Toutefois, les hôtes ne sont pas pour autant égaux face au parasitisme. Une variabilité tant en terme de probabilité d'infection (prévalence) qu'en terme de nombre de parasites par hôte (intensité parasitaire) est généralement observée. Plus particulièrement, dans les populations naturelles, peu d'hôtes hébergent un grand nombre de parasites alors que la grande majorité des hôtes ne portent que peu ou pas de parasites (Crofton 1971, Shaw et al. 1998, Poulin 2007, Fig. 2). Ce phénomène, désigné sous la terminologie d'agrégation parasitaire est même considéré comme la première loi du parasitisme puisqu'il s'agit d'une des rares caractéristiques communes à la grande majorité voire à l'ensemble des parasites (Shaw et al. 1998, Poulin 2007). Néanmoins, les causes de l'agrégation parasitaire, notamment les mécanismes permettant d'aboutir à de telles distributions chez les parasites, restent encore à ce jour difficiles à appréhender (Gourbière et al. 2015, Johnson and Wilber 2017, Rabajante et al. 2019).



**Figure 2 :** Figure illustrant l'agrégation parasitaire pour quatre interactions hôtes-parasites (d'après Shaw et al. 1998). Les barres représentent les fréquences observées et les points l'attendu selon une distribution binomiale négative (i.e., agrégée). a) *Orconectes rusticus* parasités par *Paragonimus kellicotti*, b) *Perca fluviatilis* parasitées par *Triaenophorus nodulosus*, c) *Strunurus vulgaris* parasités par *Porrocarcum ensicaudatum* et d) *Peromyscus maniculatus* parasités par *Hoplopleura hesperomydis*.



La compréhension des mécanismes sous-jacents à l'agrégation parasitaire reste toutefois d'une importance capitale car l'agrégation parasitaire conditionne la dynamique de transmission des parasites (Paull et al. 2012). Dans de nombreux cas d'épidémie, un hôte fortement parasité peut à lui seul contribuer à la propagation des cas d'infection ; cet hôte est alors généralement désigné sous le terme de super-propagateur (Lloyd-Smith et al. 2005, Paull et al. 2012). Un exemple souvent cité est celui de « Typhoïd Mary », une patiente atteinte par la typhoïde au début du XX<sup>ème</sup> siècle et qui a transmis la maladie à au moins 47 autres personnes tout en étant responsable de 28 foyers d'infection (Hudson et al. 2008). De même, Wang et al. (2006) répertorient un cas de Syndrome Respiratoire Aigu Sévère où 33 personnes ont été contaminées par un seul autre patient. Plus généralement, il est estimé que 20% des hôtes sont responsables de 80% des infections (Woolhouse et al. 1997). Ainsi, comprendre par quels mécanismes les parasites parviennent à parasiter de manière efficace un individu hôte est également une première étape vers la compréhension des dynamiques de transmission et par extension des phénomènes d'épidémies.

### **Les impacts du parasitisme**

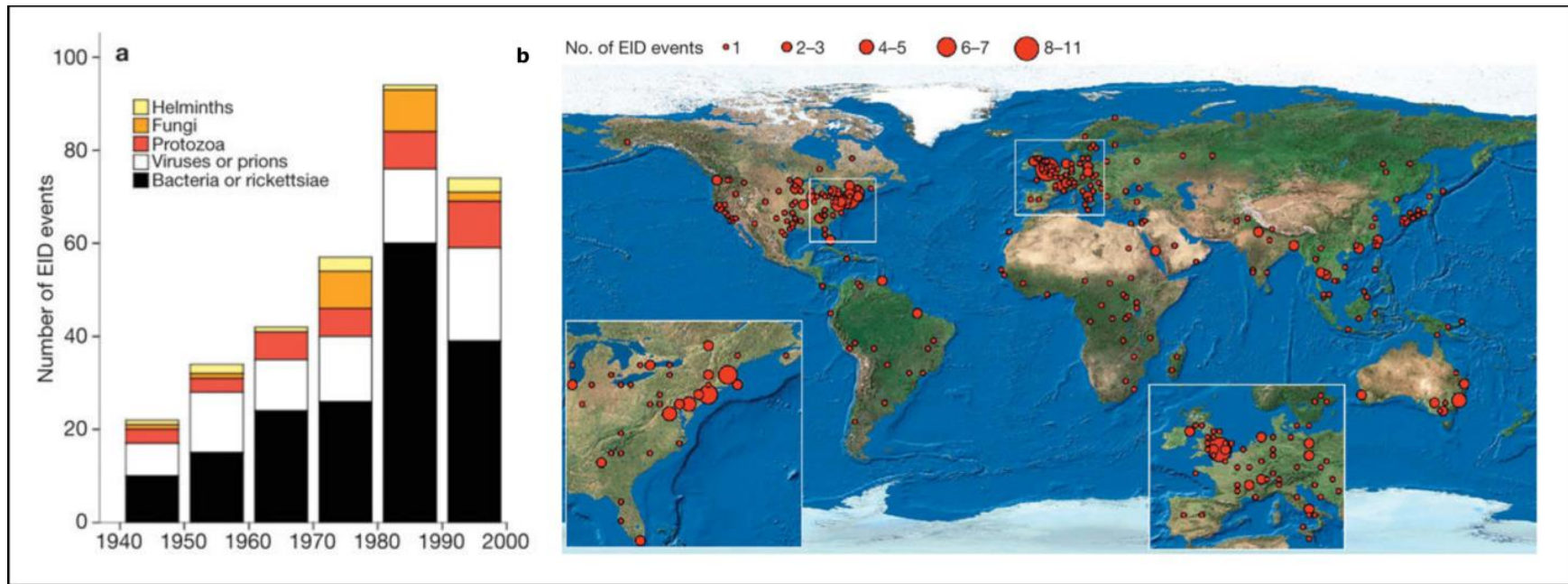
Les parasites sont des organismes pathogènes qui peuvent provoquer des modifications morphologiques, physiologiques et/ou comportementales de l'hôte et diminuer sa survie et son succès reproducteur (Crawley 2009). Ceci est d'autant plus vrai pour les parasites hétéroxènes (i.e., parasites comportant plusieurs hôtes au sein de leur cycle de vie), qui peuvent conduire à une issue fatale chez leurs hôtes intermédiaires afin de permettre la transmission à l'hôte définitif (surtout quand la transmission se fait par voie trophique, Adamo 2002, da Silva and Langoni 2009, Mikheev et al. 2015). Néanmoins, l'impact des parasites ne s'arrête pas à l'individu hôte. Les parasites sont notamment d'importants régulateurs démographiques des populations hôtes : il

a été montré par exemple que les fluctuations démographiques d'une population de Lagopèdes d'Écosse sont causées par les effets négatifs d'un nématode parasite sur la fécondité de ces oiseaux (Hudson 1998). De même, de plus en plus d'études mettent l'accent sur les effets indirects des parasites à des niveaux d'organisation plus larges telle que l'échelle des communautés et des écosystèmes (Wood et al. 2007, Hatcher et al. 2012, Sato et al. 2012). Il est à noter que certaines études mettent aussi en avant les impacts positifs des parasites à l'échelle des écosystèmes, tel qu'un accroissement de la stabilité, de la résilience ou bien encore de la diversité de l'écosystème (Hudson et al. 2006, Hatcher et al. 2012). Les infections parasitaires n'ont donc pas seulement des conséquences directes pour les hôtes infectés mais également des conséquences indirectes et profondes pouvant aller jusqu'à altérer le fonctionnement de tout un écosystème via un jeu complexe d'interactions entre différentes espèces (Dobson and Hudson 1986, Combes 1996, Hudson et al. 2006). Par exemple, Sato et al. (Sato et al. 2012) ont montré que les parasites nématomorphes, en poussant leurs hôtes criquets à se jeter dans l'eau pour parfaire leur cycle, occasionnaient ce faisant des effets en cascade sur le fonctionnement de l'écosystème. Notamment, cette « manipulation » de l'hôte à se jeter à l'eau permettait de relâcher la pression de prédation sur les invertébrés benthiques aquatiques ce qui conduisait à une diminution de la biomasse d'algues benthiques et à un accroissement du taux de destruction des feuilles (Sato et al. 2012). En cela, la compréhension des mécanismes de l'infection parasitaire soulève non seulement des enjeux relatifs aux effets directs des parasites sur leurs hôtes, souvent préjudiciables, mais aussi aux effets des parasites sur les autres espèces avec lesquelles ils interagissent de manière directe et indirecte et qui est souvent plus difficile à mesurer.

## Mécanismes d'infection et maladies émergentes

La deuxième moitié du XX<sup>ème</sup> siècle a connu une très large augmentation de cas de maladies émergentes (Jones et al. 2008, Fig. 3). Ces émergences de maladies infectieuses s'expliquent en partie par des facteurs socio-économiques, environnementaux et écologiques (Morse 1995, Patz 1996, Morens et al. 2004, Ogden and Gachon 2019). Non seulement les conditions environnementales changent rapidement, parfois de façon imprédictible, mais aussi l'accroissement des mouvements des populations humaines sont autant de facteurs qui contribuent à des modifications d'aire de distribution de nombreuses espèces dont des parasites (Patz et al. 2000, Morgan et al. 2009, Barrett et al. 2012, Frost et al. 2019). Ainsi, pour beaucoup de parasites, il est fort possible que les facteurs régissant le succès d'infection ont été affectés et vont être affectés par ces bouleversements globaux. Par exemple, les cas d'infection par le protozoaire parasite *Apicystis bombi* sur deux espèces locales de bourdon en Patagonie (*Bombus ruderatus* et *B. dahlbomii*) sont consécutifs au commerce et à l'introduction du Bourdon terrestre (*B. terrestris*) également porteur de ce parasite (Arbetman et al. 2013). L'exemple de *A. bombi* n'est pas anecdotique puisqu'à ce jour, il a déjà été fait état de nombreux cas où des parasites ont été introduits accidentellement (souvent via le déplacement de l'un de leurs hôtes ou un changement d'aire de distribution de leurs vecteurs) dans des endroits où ils n'étaient auparavant pas présents (Fig. 3, Bazsalovicsová et al. 2011, Rey et al. 2015, Kuchta et al. 2018). Ces parasites introduits, s'ils colonisent efficacement leur aire d'introduction, peuvent alors constituer des espèces dites invasives et donner lieu à l'émergence de maladies puisqu'ils sont susceptibles d'exploiter un nouvel éventail d'hôtes locaux (notion de *spill-over*, Thompson et al. 2009, Lymbery et al. 2014). Les parasites émergents doivent faire face à de nouvelles espèces hôtes locales et parfois à de nouvelles conditions environnementales ce qui peut freiner leur dynamique de colonisation (Lymbery et al. 2014). En contrepartie, les populations locales d'hôtes sont

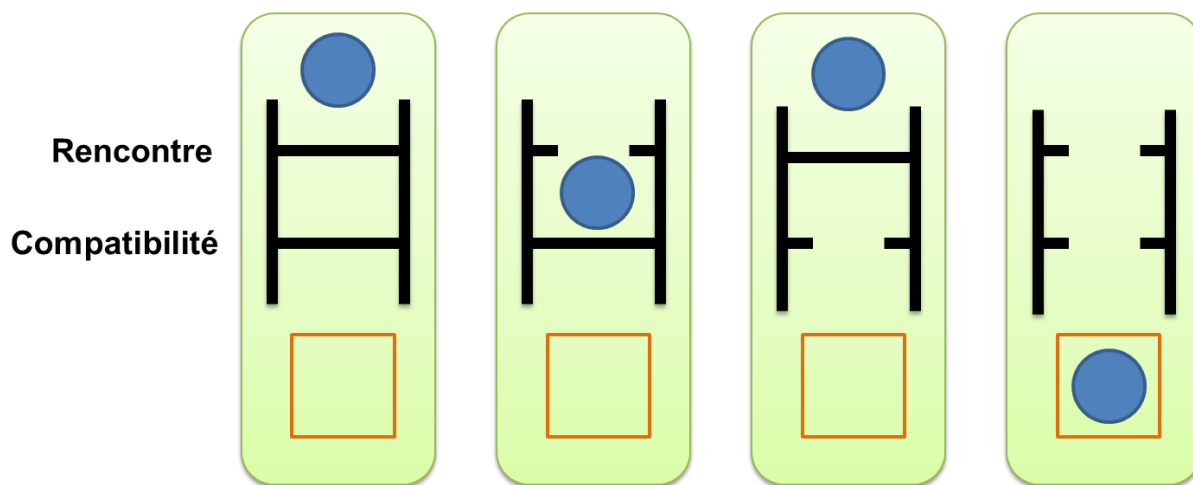
naïves face à ces parasites émergents et peuvent accuser des effets pathogènes d'autant plus importants (eg., l'extrême mortalité (80%) chez les lagomorphes causée par l'introduction du *Mixoma virus* (Fenner and Fantini 1999)). Dans ce cadre-là il est donc important d'identifier les facteurs clefs qui vont contribuer au succès de ces parasites, conditionner leur dynamique d'invasion mais aussi moduler leurs effets pathogènes.



**Figure 3** : Illustration de la progression des maladies infectieuses émergentes à l'échelle du globe. a) Nombre de maladies infectieuses émergentes de 1940 à 2000 en fonction du type de pathogène, b) Répartition des foyers de maladies émergentes de 1940 à 2004, le diamètre des cercles est proportionnel au nombre d'émergences de maladies infectieuses. Reproduit de Jones et al. (2008)

## Les concepts de filtre de rencontre et de compatibilité : un cadre pour la compréhension des mécanismes d'infection

Claude Combes a été parmi les premiers à donner un cadre conceptuel aux conditions de succès du mode de vie parasite en se basant sur deux notions de filtres que le parasite doit franchir afin de parvenir à infecter durablement son hôte (Combes 2005). Ainsi un parasite se doit de franchir d'une part un filtre dit de rencontre et d'autre part un filtre dit de compatibilité (Fig. 4).



**Figure 4 :** Représentation schématique des notions de filtres de rencontre et de compatibilité. Les ronds bleus symbolisent le parasite, les rectangles oranges les hôtes, et les rectangles verts l'environnement commun. L'infection ne peut se produire que lorsqu'à la fois le filtre de rencontre et de compatibilité sont ouverts.

### Le filtre de rencontre

Un des enjeux pour le parasite, afin d'assurer sa survie, est tout d'abord de rencontrer un hôte. En d'autres termes le parasite et l'hôte doivent se trouver au même endroit au même moment (« rendez-vous dans l'espace et dans le temps » (Combes 2005)). Ceci est d'autant plus compliqué pour les parasites que l'hôte constitue un support mobile. Les parasites ont donc

développé plusieurs stratégies adaptatives afin d'augmenter la probabilité de rencontre avec ce support malgré sa mobilité (Combes 2005).

Tout d'abord, un trait généralement attribué au mode de vie parasitaire est une forte fécondité (Crofton 1971, Turner et al. 1995, Mackinnon and Read 1999). La surfécondité des parasites est un mécanisme important par lequel les parasites accroissent passivement la probabilité de rencontre de leurs descendants en inondant le milieu de propagules infestants. Cette augmentation de la fécondité est très marquée pour les parasites hétéroxènes qui utilisent plusieurs hôtes pour accomplir leur cycle biologique. En effet, la route qui mène à l'hôte définitif est plus longue et les évènements de transmission représentent des étapes successives au cours desquelles les pertes sont énormes, souvent compensées par des phases de multiplication asexuée (Price 1974, Poulin 1995, Combes 2005).

Une autre stratégie déployée par les parasites, *la favorisation parasitaire*, consiste à augmenter la probabilité de rencontre avec l'hôte en réduisant le rôle du hasard. D'abord, certains parasites sont capables de détecter des stimuli indiquant la présence probable de l'hôte dans l'environnement. Par exemple, les moustiques, qui sont des ectoparasites en plus d'être souvent des vecteurs, localisent leurs hôtes via les émissions de dioxyde de carbone et en remontant les gradients de température (Lacey et al. 2014). D'autres parasites, notamment les parasites hétéroxènes, peuvent induire une modification du comportement de leur hôte intermédiaire afin d'augmenter le succès de passage à l'hôte définitif (Adamo 2002, da Silva and Langoni 2009, Mikheev et al. 2015). Par exemple les souris infectées par *Toxoplasma gondi* font preuve d'une activité accrue et occupent des espaces ouverts ce qui les rend plus vulnérables à la prédation par l'hôte définitif, à savoir un félin (Hay et al. 1984, Hrdá et al. 2000). Enfin, certains parasites tendent à synchroniser leurs pics d'émergence de larves infestantes avec le pic d'activité de leurs

hôtes dans le milieu (rendez-vous dans le temps). Par exemple, chez *Schistosoma mansoni*, parasite de l'homme, les cercaires sont émises dans le milieu aquatique en plein jour aux alentours de midi alors que chez *S. margrebowiei*, parasite de l'antilope, il y a deux pics, l'un au lever du jour et l'autre avant la tombée de la nuit (Theron and Combes 1995, Gautret et al. 1995).

La mise en évidence de ces différentes stratégies (écologique et comportementale) permettant d'accroître significativement les probabilités de rencontre entre hôtes et parasites a largement contribué à la compréhension des mécanismes sous-jacents à l'ouverture du filtre de rencontre. Toutefois, l'étude de ces mécanismes reste souvent difficile à appréhender, tout particulièrement en milieu naturel, du fait de la difficulté à dissocier les facteurs agissant sur le filtre de rencontre de ceux agissant sur le filtre de compatibilité.

### **Le filtre de compatibilité**

Si pendant longtemps la relation hôte-parasite a été comparée à la relation proie-prédateur, il existe tout de même une différence fondamentale entre ces deux modes de vie : le prédateur n'exploite qu'une seule fois sa proie ce qui implique que la rencontre suffit à son succès, (i.e., *interaction instantanée*). Le parasite en revanche exploite son hôte dans la durée (Rogers and Hassell 1974, Raffel et al. 2008), (i.e., *interaction durable*) (Combes 2005). Rencontrer son hôte n'est donc pas suffisant pour qu'un parasite infecte durablement son hôte. En effet, l'hôte doit convenir aux exigences métaboliques du parasite et le parasite doit survivre aux défenses immunitaires de l'hôte.

Ainsi le filtre de compatibilité correspond à la capacité d'un parasite à exploiter durablement son hôte, à y assurer son développement, se reproduire et infester un nouvel hôte. La compatibilité est supposée résulter de l'interaction entre le génotype de l'hôte et le génotype du parasite (*genotype-*



*by-genotype interaction*, Lambrechts et al. 2006). Cette interaction entre les génotypes respectifs de l'hôte et du parasite permet alors l'émergence du phénotype hôte susceptible (i.e., l'hôte peut contracter des parasites) ou résistant (i.e., l'hôte élimine des parasites) et du phénotype parasite infectieux (i.e., le parasite parvient à se développer) ou non infectieux (i.e., le parasite ne parvient pas à se développer) (Lambrechts et al. 2006). Le filtre de compatibilité est dit ouvert lorsque d'une part l'hôte est susceptible au génotype du parasite et que le parasite est infectieux vis-à-vis du génotype de l'hôte (Theron et al. 2008). Par exemple la plupart des schistosomes d'humain sont également capables de compléter leurs cycles chez la souris qui est le modèle d'étude privilégié en laboratoire pour ce parasite (i.e., filtre de compatibilité ouvert) ; en revanche les schistosomes d'oiseaux ne sont pas capables de se développer chez l'humain (i.e., filtre de compatibilité fermé) (Euzet and Combes 1980).

Quant aux mécanismes permettant l'émergence d'une part, de phénotypes hôtes susceptibles et d'autre part, de phénotypes parasites infectieux, ils ont été principalement étudiés du point de vue de la dynamique co-évolutive hôte-parasite (Dybdahl and Lively 1998, Lambrechts et al. 2006, Little et al. 2010). Il a été suggéré que suite à la dynamique co-évolutive hôte-parasite, des hôtes et des parasites sympatriques devraient être plus compatibles que des hôtes et des parasites allopatriques (Ebert 1994). Néanmoins, certaines études tendent à montrer que ce patron n'est pas toujours respecté, certaines populations de *Schistosoma mansoni* sont plus compatibles avec des populations de *Biomphalaria tenagophila* se situant dans des sites éloignés par rapport à des populations de *B. tenagophila* se situant dans des sites proches (Morand et al. 1996). Ceci suggère que la dynamique de coévolution entre un parasite et son hôte, à elle seule, ne permet pas de rendre compte de l'émergence de phénotypes compatibles. Certaines études suggèrent notamment que la plasticité (i.e., modifications phénotypiques qui ne sont pas associées à des

modifications du génome) pourraient également promouvoir la compatibilité entre hôtes et parasites (Little et al. 2006, Gervasi et al. 2015, De Fine Licht 2018).

### **Pourquoi les notions de filtres de rencontre et de compatibilité sont les clefs de la compréhension de l'agrégation parasitaire ?**

Dans la littérature, deux causes principales à l'agrégation parasitaire sont communément mentionnées. L'agrégation parasitaire peut résulter soit d'une hétérogénéité concernant l'exposition de l'hôte vis-à-vis du parasite (filtre de rencontre hétérogène à l'échelle de la population), soit d'une hétérogénéité de la réponse immunitaire au sein des populations hôtes au sein desquelles certains individus sont résistants et d'autres très susceptibles (filtre de compatibilité hétérogène à l'échelle de la population) (Poulin 2007, Morrill and Forbes 2012, Gourbière et al. 2015). Les hôtes plus exposés et ou plus susceptibles seraient ainsi le siège d'une agrégation parasitaire. Aussi, l'utilisation des notions de filtre de rencontre et de compatibilité semble être un cadre pertinent pour la compréhension de l'agrégation parasitaire. De plus, il est important de mentionner que se baser sur la notion de compatibilité plutôt que de susceptibilité apporte un cadre plus général, car la notion de compatibilité permet de considérer simultanément la variabilité de susceptibilité au sein des populations hôtes mais aussi la variabilité de virulence au sein de la population de parasites. Les concepts de filtres de rencontre et de compatibilité semblent donc être un cadre adéquat à la compréhension des mécanismes d'infection, de la distribution des parasites et de l'émergence des super-propagateurs.

### **Les facteurs pouvant moduler l'ouverture des filtres de rencontre et de compatibilité**

Le parasitisme est souvent décrit comme une interaction triple entre le parasite, l'hôte et l'environnement (Wolinska and King 2009). Ainsi, de manière générale, le succès d'un parasite

est influencé à la fois par des facteurs liés à l'hôte, qui constitue l'environnement biotique direct du parasite mais également par des facteurs se rapportant à l'environnement (Brunner and Ostfeld 2008). Enfin, la capacité d'un parasite à infecter dépend également de ses caractéristiques intrinsèques propres. Ces facteurs peuvent avoir un impact décisif sur les filtres de rencontre et de compatibilité, mais il est souvent difficile d'établir leurs influences relatives puisque dans les populations naturelles ces facteurs interagissent les uns avec les autres.

### **Les facteurs environnementaux**

Les facteurs environnementaux peuvent influencer la dynamique des populations de parasites (Patz et al. 2000, Browman et al. 2004, Cardon et al. 2011). Par exemple, des températures élevées sont associées à une transmission accrue de certains parasites en influençant le taux de développement (eg., le nématode *Umingmakstrongylus pallikuukensis*, parasite du bœuf musqué (Kutz et al. 2005)). De manière générale, de nombreuses études ont montré un lien important entre les conditions environnementales, telles que la température, le niveau de précipitation, la fragmentation de l'habitat (Gillespie and Chapman 2008, Lachish et al. 2013, Fourcade et al. 2014) et la distribution spatiale des parasites (Morgan et al. 2009, Barrett et al. 2012, Caminade et al. 2019). Ainsi, la variabilité de l'environnement dans l'espace peut contribuer à l'émergence de zones propices en terme d'exposition et de rencontre avec le parasite (*points chauds d'infection*, (Anacker et al. 2008, Paull et al. 2012)). Par exemple, Jennett et al. (2013) ont montré que les risques d'infection par des tiques dépend de l'hétérogénéité des paysages, les risques d'infection étant accrus dans les zones boisées et les lisières par rapport aux paysages ouverts (prairies) (Jennett et al. 2013). De tels patrons spatiaux de répartition des points chauds et des points froids sont particulièrement documentés à large échelle spatiale permettant ainsi de constituer des cartes de risques (Jones et al. 2008, Chaikaew et al. 2009, Jeefoo et al. 2010, Fig.

3-b). Par exemple, Loiseau et al. (2013) ont montré que la prévalence de la malaria aviaire en France, était fortement liée aux températures journalières et prédisent dans le futur, une expansion de l'aire de distribution de *Plasmodium relictum* vers le nord de la France en réponse au changement climatique. Toutefois, encore peu d'études ont pris en compte le rôle de la variabilité environnementale dans les risques d'infections à une fine échelle spatiale (i.e., échelle micro-géographique : quelques centaines de mètres (Richardson et al. 2014)). Or la variabilité de l'environnement à l'échelle micro-géographique pourrait être un mécanisme général pour moduler la probabilité de rencontre de nombreux parasites (Poulin and FitzGerald 1989, Thamm et al. 2009, Borer et al. 2010).

### **Les facteurs liés à l'hôte**

De nombreux facteurs liés à l'hôte, tels que la taille, l'âge, le sexe, le comportement ou le statut social peuvent moduler le taux d'infection par les parasites (Lafferty and Morris 1996, Cardon et al. 2011, Brommer et al. 2011). Par exemple, de forts taux de testostérone chez les lagopèdes d'Ecosse mâles sont associés à de plus fortes charges parasitaires, mais aussi ces mâles sont préférentiellement choisis par les femelles, ce qui pourrait avoir un impact important sur la transmission des parasites (Mougeot 2005). De même, Bartoli et al. (2000) ont montré que le trématode *Genitocotyle mediterranea*, qui infecte le tractus digestif du poisson Crénilabre ocellé (*Symphodus ocellatus*) est distribué en fonction de la hiérarchie sociale de cet hôte : la prévalence des mâles satellites (21%) est plus faible par rapport à la prévalence des mâles dominants (68%).

Au niveau génétique, le lien entre la diversité génétique de l'hôte et la prévalence parasitaire a fait l'objet d'une attention particulière en biologie évolutive (Bernatchez and Landry 2003). Par exemple, il existe une corrélation entre la diversité génétique de l'hôte mesurée sur des marqueurs neutres et la prévalence parasitaire chez de nombreux organismes, ce qui suggère un

lien entre le niveau de consanguinité des hôtes et leur susceptibilité aux pathogènes (Acevedo-whitehouse and Cunningham 2006, Luikart et al. 2008, Blanchet et al. 2010). De manière similaire, la recherche de lien causal entre la diversité génétique des gènes du Complexe d'Histocompatibilité Majeur et la prévalence ou l'intensité parasitaire a fait l'objet de nombreuses études (Wegner et al. 2003, Zhang et al. 2015, Aguilar et al. 2016). L'étude des facteurs génétiques liés à l'hôte et qui influent sur sa susceptibilité aux parasites a pris tout particulièrement de l'ampleur au cours des dernières années avec le développement des nouvelles techniques de séquençage haut débit qui rendent possible l'étude de millions de marqueurs neutres et non neutres. Cette nouvelle technologie permet de répondre à un enjeu actuel majeur, à savoir, la recherche des bases génomiques de la résistance/susceptibilité aux parasites (Wenzel et al. 2015, Benavides et al. 2016, Zueva et al. 2018).

### **Les facteurs liés au parasite**

Les caractéristiques des parasites régissent au même titre que les facteurs liés à l'hôte l'ouverture du filtre de rencontre et de compatibilité. Notamment, certains parasites ont développé des stratégies pour favoriser la rencontre avec leurs hôtes (Combes 2005) mais également des stratégies pour survivre aux défenses immunitaires de leurs hôtes (eg., camouflage antigénique, (Damian 1964, 1989, Das et al. 2018)). Par exemple Hebert et al. (Hebert et al. 2015) ont montré que le cestode *Schistocephalus solidus*, qui utilise comme hôte secondaire l'épinoche, synthétise des protéines mimétiques de son hôte qui sont susceptibles d'induire les changements comportementaux nécessaires à une favorisation parasitaire, soulignant ainsi à la fois une stratégie qui vise à accroître les probabilités de rencontre et une stratégie de camouflage vis-à-vis des défenses de l'hôte. La plasticité des parasites joue également un rôle important dans la réalisation d'une infection (eg., ajustements transcriptomiques du parasite lors des transitions

entre hôte intermédiaire et hôte définitif toujours chez *S. solidus*, (Hébert et al. 2017)) et représente un enjeu actuel dans le contexte des maladies émergentes (Agosta et al. 2010).

Les parasites ont pendant longtemps été considérés comme hautement spécifiques de certaines espèces hôtes (i.e., *spécificité parasitaire* (Adamson and Caira 1994)). Leur capacité à infecter leur hôte a principalement été étudiée du point de vue de la dynamique co-évolutive hôte-parasite (Ronquist 2003) et la sélection naturelle a ainsi été avancée en tant que mécanisme clef permettant d'expliquer comment certains parasites parviennent à infecter efficacement un hôte spécifique (Dybdahl and Lively 1998, Wegner et al. 2003). Par exemple, la sélection est supposée promouvoir une virulence optimale chez le parasite, qui maximise la transmission en modérant les coûts notamment en terme de mortalité chez les hôtes (par opposition à une virulence maximale non viable à long terme car occasionnant une trop forte mortalité chez les hôtes) (Alizon and Michalakis 2015, Ebert 2016).

Toutefois, de plus en plus d'études tendent à suggérer que les parasites ont un éventail d'espèces hôtes plus large que précédemment envisagé (i.e., *généralisme d'hôtes*) (Poulin and Keeney 2008, Agosta et al. 2010, Lootvoet et al. 2013). Cette observation soulève un questionnement référé sous le terme de *paradoxe du parasite* : en effet, si de fortes pressions de sélection sont censées favoriser des associations hautement spécifiques entre hôtes et parasites, notamment à cause du caractère étroit et durable de l'interaction, il faut tout de même expliquer que certains parasites conservent le potentiel (génétique) pour s'associer à de nouvelles espèces hôtes (*parasite paradox*, Agosta et al. 2010). Récemment, plusieurs études théoriques ont suggéré que la plasticité des parasites pourrait amener à une résolution de ce paradoxe notamment en permettant l'émergence de phénotypes adaptatifs pour de nouvelles espèces hôtes et cela avant même que la sélection ait pu moduler l'interaction hôte-parasite (Araujo et al. 2015, Nylin et al.

2018, De Fine Licht 2018). Cette plasticité adaptative serait d'autant plus avantageuse dans un environnement changeant, tel que décrit de nos jours.

### **De l'interaction triple à l'interaction multiple : le rôle du microbiome**

Outre les facteurs environnementaux, les facteurs relatifs à l'hôte et au parasite qui peuvent moduler l'ouverture du filtre de rencontre et/ou de compatibilité, depuis quelques années certaines études mettent l'accent sur le rôle des communautés bactériennes associées à l'hôte dans le processus d'infection (Hayes et al. 2010, Bär et al. 2015, Jenkins et al. 2018). Par exemple, Hayes et al. (2010) ont montré que la présence des bactéries *Escherichia coli* dans le tractus intestinal de souris infectées par le nématode *Trichuris muris*, favorise l'éclosion des œufs de *T. muris*. Le parasitisme ne serait donc plus seulement une interaction triple entre l'environnement, le parasite et l'hôte mais une interaction multiple au sein de laquelle il faudrait inclure les liens existants avec le microbiote de l'hôte, mais aussi avec le microbiote du parasite ainsi qu'avec les communautés bactériennes de l'environnement (*Holobiont-Holobiont interaction*, (Dheilly 2014).

Plusieurs études s'accordent notamment à montrer que les compositions des communautés bactériennes entre hôtes sains et parasités sont différentes (Ras et al. 2015, Guernier et al. 2017, Afrin et al. 2019). Par exemple il a été montré que les genres *Janthinobacterium*, *Pseudomonas* et *Flavobacterium* étaient bien plus abondants dans le microbiote de la peau des Grenouilles Italiennes (*Rana italica*) saines que dans celui de grenouilles infectées par des parasites tels qu'*Amphibiocystidium sp.* (Federici et al. 2015). Toutefois, à ce jour il reste difficile de savoir si ces différences de communautés bactériennes chez l'hôte sont la cause ou la conséquence des infections parasitaires.

En effet, les différences de communautés bactériennes chez l'hôte pourraient avoir une influence considérable sur le taux d'infection. Par exemple, si des différences de communautés bactériennes préexistent à l'infection parasitaire, il est possible d'envisager que certaines bactéries favorisent l'installation du parasite en créant des conditions propices (le microbiote constitue une source de nourriture qui satisfait les exigences du parasite, (Hayes et al. 2010)) ou inversement protègent l'hôte contre d'éventuelles infections parasitaires (eg., certaines bactéries peuvent sécréter des substances antibiotiques ou antifongiques, (Buffie and Pamer 2013, Britton and Young 2014)). Alternativement, les changements de communautés bactériennes pourraient être une conséquence des infections parasitaires (Lee et al. 2014, Ras et al. 2015, Beatty et al. 2017). Dans ce cas, l'infection parasitaire est supposée pouvoir faciliter la colonisation par de nouvelles bactéries selon différents types de mécanismes. D'une part les parasites, par leur activité de nutrition par exemple, peuvent avoir un effet direct qui perturbe les bactéries déjà établies et modifient ainsi les abondances relatives des bactéries associées à l'hôte (Lee et al. 2014, Beatty et al. 2017). D'autre part, la présence de parasites peut perturber l'immunité de l'hôte est ainsi décroître sa résistance face à des colonisations ultérieures par de nouvelles bactéries opportunistes (Mooney et al. 2015, Portet et al. 2018). Enfin, à l'image du système *Wolbachia*-nematodes, certaines bactéries pourraient être directement issues du parasite et faciliter l'infection dans une dynamique de co-infection (Tamarozzi et al. 2011, Swe et al. 2014). De manière générale, le rôle du microbiote, au même titre que les facteurs environnementaux, les facteurs liés à l'hôte et au parasite pourrait être décisif dans la compréhension des mécanismes d'infection parasitaire.



## Objectifs

Au cours de cette thèse mon objectif a été d'amener une meilleure compréhension du processus complexe d'infection parasitaire en me basant très largement sur les concepts de filtres de rencontre et de compatibilité. A la lumière d'une vision intégrative, j'ai cherché à mettre en évidence des facteurs déterminants dans les interactions hôtes-parasites mais parfois encore négligés. Pour ce faire, je me suis focalisée sur un ectoparasite de poisson d'eau douce, *Tracheliastes polycolpus*. En France, *T. polycolpus* parasite principalement la vandoise (*Leuciscus sp.*) mais est aussi capable d'infecter plusieurs autres espèces de cyprinidés d'eau douce de manière occasionnelle (Lootvoet et al. 2013). *Tracheliastes polycolpus* est une espèce invasive dans les rivières françaises, et en cela constitue un modèle intéressant en ce qui concerne les enjeux actuels relatifs aux maladies émergentes et dont les mécanismes d'infection restent encore difficiles à appréhender. Afin d'étudier les mécanismes d'infection chez *T. polycolpus*, j'ai combiné des approches expérimentales et de modélisation basées sur l'analyse de données empiriques ainsi que diverses approches génomiques afin de considérer le rôle des facteurs environnementaux, des facteurs liés à l'hôte, au parasite et à leur microbiote respectifs dans le processus d'infection parasitaire.

### **(1) Rôle de l'environnement micro-géographique dans la rencontre entre un parasite et son hôte**

Le premier chapitre de cette thèse porte sur l'étude du rôle des facteurs environnementaux dans le filtre de rencontre. Plus particulièrement, je me suis focalisée sur l'impact de l'environnement à l'échelle micro-géographique sur les probabilités de rencontre entre l'hôte et le parasite. En effet, les effets environnementaux sur la distribution des parasites ont été bien documentés à de larges échelles spatiales, mais peu d'études à ce jour ont considéré le rôle de la variabilité locale de

l'environnement sur ces distributions (Poulin and FitzGerald 1989, Thamm et al. 2009, Jennett et al. 2013). Cet axe repose ainsi sur l'hypothèse majeure que les probabilités d'infection d'un parasite peuvent être structurées par l'environnement à l'échelle micro-géographique où certains microhabitats contribuent de manière disproportionnée à la transmission du parasite (points chauds d'infection) (Paull et al. 2012). En m'appuyant sur une approche semi-expérimentale en milieu naturel (i.e., encagement d'hôtes dans une rivière), l'objectif a d'abord été de tester si les événements d'infections se produisaient préférentiellement dans certains microhabitats et cela en contrôlant les facteurs pouvant agir sur le filtre de compatibilité. Dans un second temps, j'ai testé, cette fois en population naturelle, si les hôtes parasités s'agrégeaient au sein de points chauds d'infection. Dans les deux cas (approche semi-expérimentale et descriptive) j'ai testé le lien entre les risques d'infection et les facteurs environnementaux à l'échelle micro-géographique.

## **(2) Rôle des déterminants environnementaux et génomiques dans la répartition d'un parasite émergent.**

La plupart des modèles actuels qui visent à prédire l'aire de répartition des parasites émergents dans le contexte du changement global sont très largement basés sur les facteurs environnementaux et négligent bien souvent le rôle de la génomique de l'hôte (Lachish et al. 2013, Pérez-Rodríguez et al. 2013, Pickles et al. 2013). Dans ce second chapitre, je me suis attachée à tester l'influence relative des effets environnementaux et des effets génomiques de l'hôte sur la récente dynamique d'invasion de *T. polycolpus* en France. En effet, ce parasite fait preuve d'une large aire de distribution à l'échelle de la France, cependant aucune infection de *T. polycolpus* n'est répertoriée à ce jour dans le Nord-Est. L'objectif a donc été de combiner une approche de modélisation de distribution d'espèces basée sur des facteurs environnementaux et une approche de génomique des populations afin de déterminer quels sont les facteurs

(environnementaux et/ou génomiques) qui expliquent la présence de cette zone où l'infection ne se réalise pas.

### **(3) Rôle de la plasticité dans la capacité d'un parasite à exploiter plusieurs espèces hôtes**

De plus en plus d'études suggèrent que les différentes espèces de parasites sont capables d'exploiter une large gamme d'hôtes (i.e., *généralisme d'hôtes*) (Poulin and Keeney 2008, Jones et al. 2008, Agosta et al. 2010). Ainsi plusieurs études théoriques récentes suggèrent que la plasticité pourrait contribuer à accroître la compatibilité d'une espèce parasite vis-à-vis de nouvelles espèces hôtes (Araujo et al. 2015, Nylin et al. 2018, De Fine Licht 2018). La plasticité pourrait notamment être un mécanisme important pour comprendre les dynamiques d'invasion de parasites accidentellement introduits et qui ont réussi à s'adapter à de nouvelles espèces hôtes locales (notion de *spill-over*, Daszak 2000, Nylin et al. 2018, Clark et al. 2018). Toutefois, à ce jour, peu d'études empiriques ont mis en évidence le rôle de la plasticité dans la capacité d'un parasite à exploiter des hôtes alternatifs. Dans ce troisième chapitre, je me suis appuyée sur une approche transcriptomique comparative afin d'étudier le rôle de la plasticité transcriptomique dans le généralisme d'hôtes. J'ai en particulier comparé les patrons d'expression génique chez *T. polycolpus* en fonction de l'espèce hôte utilisée (i.e., l'hôte principale et deux espèces hôtes alternatives).

### **(4) Rôle des communautés bactériennes dans le processus d'infection parasitaire**

Le dernier chapitre de ma thèse s'intéresse au rôle du microbiote dans le processus d'infection. Plus particulièrement l'objectif a été de tester si les différences de microbiotes entre hôtes parasités et hôtes sains telles que couramment observées représentent une cause ou une conséquence de l'infection parasitaire (Lee et al. 2014, Bär et al. 2015, Newbold et al. 2017, Portet et al. 2018). Pour ce faire, j'ai utilisé une approche semi-expérimentale d'encagement

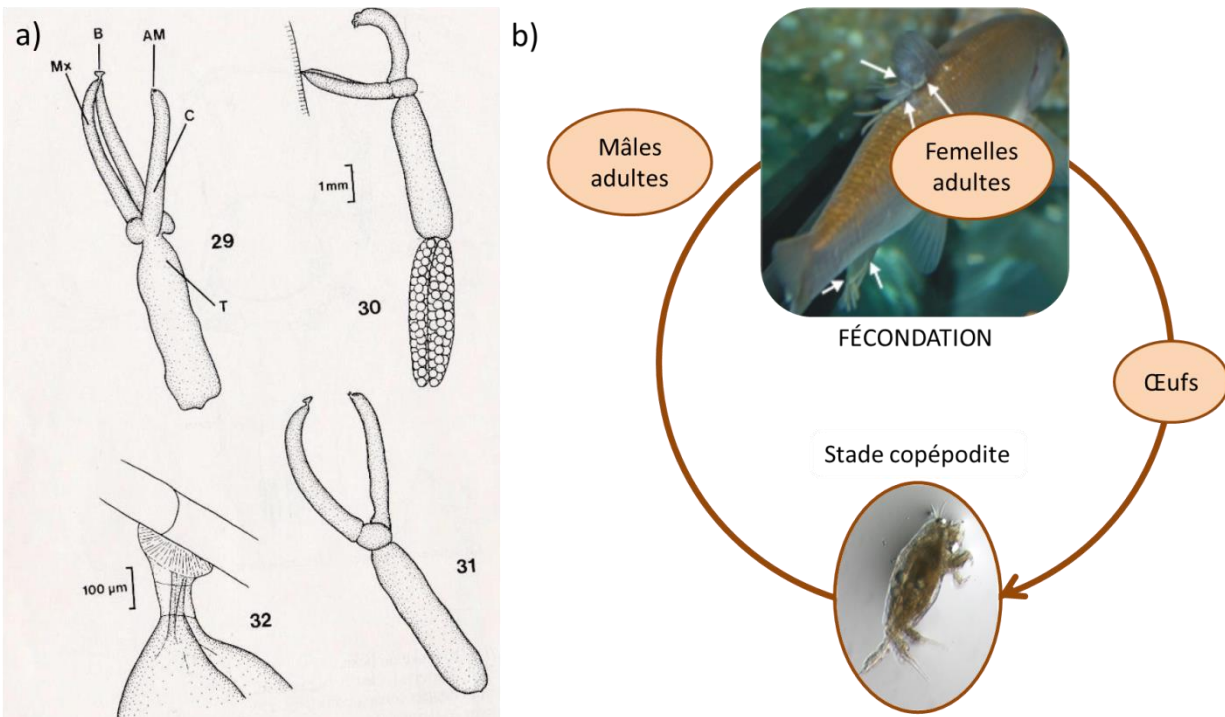
d'hôtes en milieu naturel afin de mesurer la composition des communautés bactériennes avant et après infection par *T. polycolpus* pour chaque individu hôte. J'ai également étudié les communautés bactériennes de l'environnement, des parasites adultes fixés et de larves infectieuses naïves (i.e., jamais exposées à l'environnement des hôtes). Ainsi, dans un premier temps j'ai testé dans quelle mesure les différences de communautés bactériennes peuvent favoriser (ou non) l'infection et potentiellement influencer la susceptibilité de l'hôte vis-à-vis du parasite (cause de l'infection). Dans un second temps, j'ai testé si des changements de communautés de bactéries induits par l'infection parasitaire (conséquence de l'infection) se produisaient, et le cas échéant j'ai identifié le set d'Unité Taxonomique Opérationnelle (OTU) bactériens caractéristiques de l'infection par *T. polycolpus*. Enfin, pour ces OTUs caractéristiques de l'infection j'ai déterminé la source la plus probable, à savoir, l'environnement, le parasite adulte et/ou les larves de parasite.

## **Le modèle d'étude**

### **Cycle de vie**

*Tracheliastes polycolpus* (Copepoda, Neocopepoda, Podoplea, Siphonostomatoida, Lerneapodidae) est un crustacé d'eau douce. Les femelles sont des ectoparasites de quelques de millimètres ( $5.49 \pm 0.36$  mm) qui parasitent plusieurs espèces de poissons d'eau douce et ne requièrent qu'un seul hôte pour la réalisation de leur cycle de vie (i.e., parasite monoxène, Fig. 5). En France, les hôtes de *T. polycolpus* sont principalement deux espèces de vandoises (*Leuciscus leuciscus* et *Leuciscus burdigalensis*) mais aussi dans une moindre mesure le toxostome (*Parachondrostoma toxostoma*), deux espèces de goujons (*Gobio gobio* et *Gobio occitaniae*), le vairon commun (*Phoxinus phoxinus*), le gardon (*Rutilus rutilus*) et le chevesne (*Squalius cephalus*) (Lootvoet et al. 2013). Les *T. polycolpus* femelles s'attachent aux nageoires

des poissons (mais aussi de manière occasionnelle sur leur corps) par un bouton fixateur appelé *bulla*. Elles s'y nourrissent de mucus et de cellules épithéliales via l'utilisation d'une paire de maxilles (Fryer 1982, Loot et al. 2004). Leur activité de nutrition soutient leur croissance ainsi que le développement de deux sacs ovigères contenant jusqu'à une centaine d'œufs chacun (Loot et al. 2004, Mazé-Guilmo 2016). Une fois matures, les œufs éclosent et les larves *nauplius* sont alors relâchées dans le milieu environnant (Fryer 1982). Au stade *nauplius* succède le stade copépodite (qui est le stade infestant) et enfin le stade adulte. Les mâles sont des organismes libres et microscopiques qui sont supposés féconder la femelle avant ou après que celle-ci ne s'attache à un nouvel hôte (Piasecki 1839, Fryer 1982).

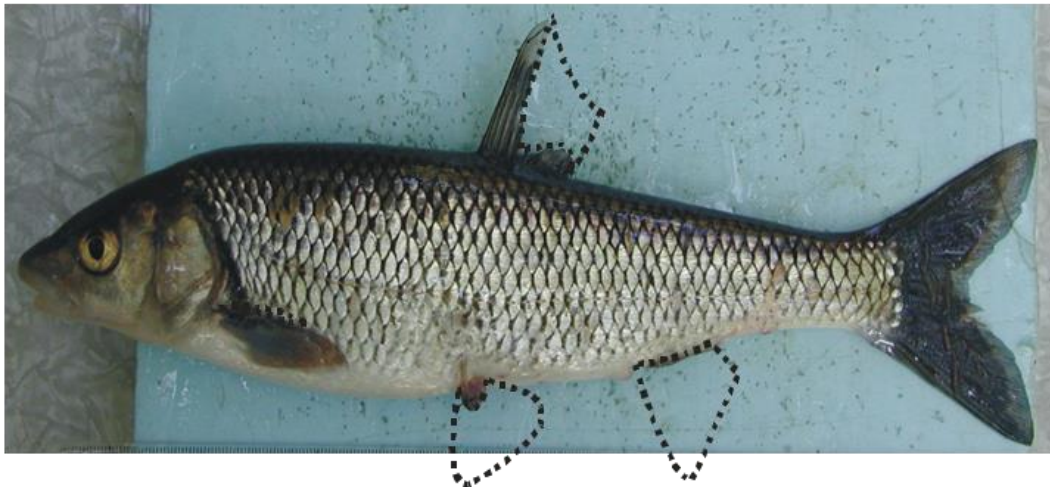


**Figure 5** : Présentation du modèle d'étude, *Tracheliastes polycolpus* et de son hôte, la vandoise. a) Dessin d'observation de *T. polycolpus* sous sa forme parasite (femelle adulte) 29 et 31 : femelles non matures ; 30 : femelle mature portant deux sacs ovigères. 32 : bulla fixée à un rayon de la nageoire d'un poisson hôte. AM : région buccale antérieure ; B : bulla ; C : cephalothorax ; Mx : maxilles et T : tronc, reproduit de Fryer (1982). b) Représentation schématique du cycle de vie de *T. polycolpus*, l'hôte représenté est la vandoise rostrée (*Leuciscus burdigalensis*), les flèches blanches indiquent des femelles *T. polycolpus* fixées sur les nageoires dorsale et caudale de l'hôte.

### Effets directs et indirects de *T. polycolpus*

*T. polycolpus* induit des effets pathogènes directs et indirects sur son hôte. De manière directe, par son activité de nutrition *T. polycolpus* cause des dégradations partielles à totales des nageoires de son hôte (Loot et al. 2004, Fig. 6). Il est également fréquent d'observer des inflammations des tissus autour du site d'ancrage dues à l'insertion de la *bulla*. Ces inflammations sont supposées favoriser d'éventuelles infections secondaires par des bactéries, des virus et/ou des champignons (Loot et al. 2004). L'ampleur de ces effets pathogènes directs dépend de l'intensité de l'infection (i.e., de la charge parasitaire) ; jusqu'à 80 parasites ont pu être dénombrés sur un seul individu

hôte (Lootvoet et al. 2013). Les effets pathogènes peuvent également dépendre de la sensibilité de la nageoire parasitée à la dégradation (les nageoires anale et pelviennes étant particulièrement vulnérables (Loot et al. 2004). A ces effets directs, s'ajoutent plusieurs effets indirects. Blanchet et al. (2009b) ont montré que les poissons parasités par *T. polycolpus*, et plus particulièrement les juvéniles, avaient une croissance ralentie et une survie plus faible que les individus sains. A l'échelle de la population hôte, il a également été montré que ces parasites tendaient à accroître la variance globale de l'hétérozygotie génétique via des processus de sélection diversifiante (Blanchet et al. 2009a).

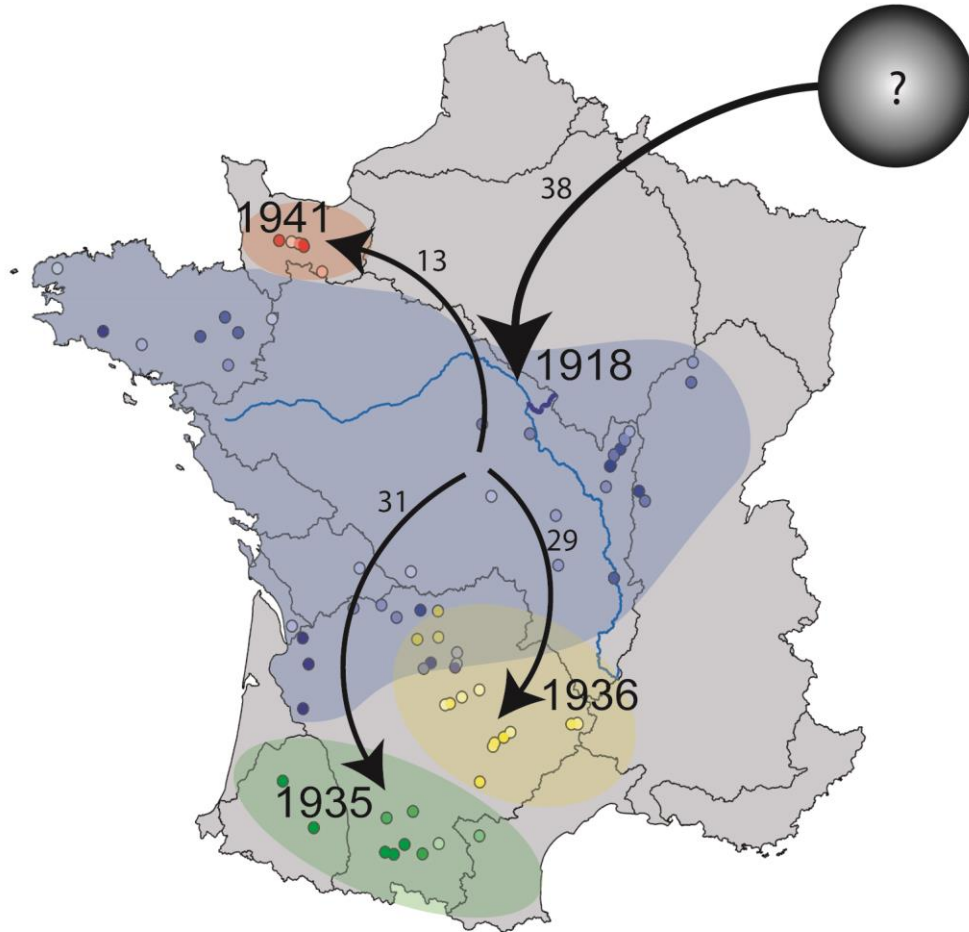


**Figure 6** : Photographie illustrant les dégradations des nageoires induites par *T. polycolpus* sur une vandoise rostrée. Les pointillés signalent les portions manquantes des nageoires. Reproduit de Loot et al. (2004).

## **Dynamique d'invasion des bassins versant français**

En France, *T. polycolpus* est une espèce non-native originaire d'Europe de l'Est. *T. polycolpus* a été introduit accidentellement en France autour des années 1920 dans le bassin versant de la Loire via l'importation de son hôte originel (l'Idé commune, *Leuciscus idus*) pour l'aquaculture (Rey et al. 2015, Fig. 7). L'idé commune étant très rare en France, *T. polycolpus* a réalisé un transfert vers de nombreuses nouvelles espèces hôtes telles que la vandoise rostrée (*L. burdigalensis*) dans la partie Sud-Ouest de la France et la vandoise commune (*L. leusiscus*) dans la partie Nord-Est de la France (Costedoat et al. 2006, Lootvoet et al. 2013). Dans les années 1940s, *T. polycolpus* a connu une colonisation rapide en trois phases s'étendant à la quasi-totalité des bassins versants Français (Rey et al. 2015). Le parasite a tout d'abord colonisé le bassin versant de l'Adour (1935), puis le bassin versant de la Garonne (1936) et enfin le bassin versant englobant les cours d'eau Normands (1941). Enfin, aujourd'hui des infections sont également répertoriées dans le bassin versant du Rhône couvrant la partie Est et Sud-Est de la France. Toutefois, de manière surprenante, dans les bassins versants Nord-Est de la France (Rhin, Seine et Meuse) aucune infection par *T. polycolpus* n'a été à ce jour documentée en dépit de suivis réguliers depuis une dizaine d'années.





**Figure 7** : Carte représentant le scénario d'introduction et de colonisation de *T. polycolpus* en France tel que reconstruit par Rey et al. (2015). Les dates indiquées correspondent aux années estimées d'introduction de *T. polycolpus* dans chaque bassin versant et les flèches indiquent l'origine et la direction des vagues de colonisation. Les nombres associés à chaque flèche indiquent les tailles effectives supposées des populations d'origine de *T. polycolpus*. Les couleurs représentent les quatre populations génétiquement différenciées de parasites. Reproduit de Rey et al. (2015).

## **Chapitre I :**

# **Rôle de l'environnement micro-géographique dans la rencontre entre un parasite et son hôte**

## Résumé

Que les parasites parviennent à trouver leurs hôtes dans un vaste environnement peut susciter des interrogations. Notamment pour les espèces parasites qui ne s'appuient pas sur les stratégies communément documentées permettant d'augmenter le taux de rencontre avec l'hôte - telles que l'utilisation de signaux environnementaux, les modifications comportementales de l'hôte et la synchronicité avec l'hôte - trouver un hôte revient à chercher une aiguille dans une botte de foin. Dans cette étude, nous avons étudié les facteurs permettant à *Tracheliastes polycolpus*, un ectoparasite d'eau douce de trouver et d'infecter plusieurs espèces de poissons d'eau douce. Nous avons combiné deux approches de terrain à l'échelle micro-géographique, consistant à identifier les points chauds d'infection et les caractéristiques environnementales de ces points chauds. Dans une expérience d'encagement en milieu naturel, nous avons d'abord vérifié si les événements d'infection étaient distribués de manière aléatoire ou agrégés dans l'espace, puis dans une étude descriptive, nous avons vérifié si, dans des populations hôtes naturelles, les hôtes parasités étaient agrégés dans des points chauds d'infection. Enfin dans les deux approches, nous avons testé si les risques d'infection étaient liés à des facteurs environnementaux. Nos résultats indiquent qu'à la fois, les événements d'infection chez des hôtes encagés, mais aussi les hôtes infectés en populations naturelles, s'agrègent dans l'espace et cela dans certains contextes environnementaux. Notamment, les zones d'eau profondes et les zones où la vitesse du courant est élevée sont associées à un plus grand risque d'infection pour les poissons encagés mais aussi pour les poissons sauvages. Nous suggérons que ces microhabitats pourraient bénéficier localement au cycle de vie des parasites, ce qui tend à augmenter la densité des stades infestants ou bien pourraient accroître le taux de contact entre parasites et hôtes, favorisant *in fine* le taux de rencontre. Nous suggérons donc qu'une structure spatiale en microhabitats caractérisés par des conditions environnementales congrues à l'échelle micro-géographique pourrait être un mécanisme général qui tend à expliquer comment de nombreuses infections parasitaires peuvent se produire et cela même dans un environnement labile tel que des rivières.

# **Finding the needle in the haystack: The role of environmental factor at microgeographic scale on parasite opportunity to find hosts**

---

Mathieu-Bégné Eglantine<sup>1,2,3\*</sup>, Blanchet Simon<sup>1,2</sup>, Rey Olivier<sup>3</sup>, Scelsi Orlane<sup>1</sup> and Loot Géraldine<sup>1</sup>,

<sup>1</sup>Université Paul Sabatier, Centre National pour la Recherche Scientifique, Ecole Nationale de Formation Agronomique; Laboratoire Evolution et Diversité Biologique (UMR5174), 118 route de Narbonne, F-31062 Toulouse, France

<sup>2</sup>Université Paul Sabatier, Centre National pour la Recherche Scientifique; Station d'Ecologie Théorique et Expérimentale (UMR5321), 2 route du CNRS, F-09200 Moulis, France

<sup>3</sup>Université Perpignan Via Domitia, UMR5244 Interaction Hôte-Parasite-Environnement, 58 Avenue Paul Alduy, F-66860, Perpignan, France

**Key-words:** host-parasite interactions, encountering filter, microhabitat, freshwater, aggregation

## Abstract

How parasites find their hosts in a vast environment is an intriguing question. Especially for parasite species that do not rely on documented strategies to increase their encountering rate with their host –such as the use of environmental and host cues or behavioral modification of host and synchronicity with host– finding their hosts ends up finding a needle in a haystack. In this study we studied the factors allowing *Tracheliastes polycolpus*, a freshwater ectoparasite to find and infect several freshwater fish species. We combined two field approaches at the microgeographic scale, consisting in identifying infection hotspots and the environmental characteristics of these hotspots. First in a semi-controlled caging experiment we tested whether infection events were randomly distributed or aggregated in space, then in a field study we tested whether in wild host populations parasitized hosts aggregated in hotspots of infection. Finally in both approaches we tested whether these spatial patterns of infection risks were related to environmental factors. Our results indicate that both infection events in experimentally housed hosts and infected wild hosts were aggregated in space and preferentially occurred in some specific environmental contexts. Notably, among all variable tested, deep water and high stream velocity are those that were found to be associated with higher infection risk for both wild and housed fish. We suggest that such microhabitats either favor locally parasite life cycle which tend to increase parasite density and/or favor contact rate between parasites and hosts, overall favoring encountering rate. We argue that a spatial structure in microhabitats characterized by congruent environmental features at microgeographic scale could be a general mechanism that tends to explain how infection by many parasites can be achieved even in labile environment such as rivers.

## Introduction

Parasitism is a widespread lifestyle (de Meeûs and Renaud 2002, Dobson et al. 2008, Hechinger et al. 2011). Due to the close interaction between hosts, parasites and their common environment, parasites strongly rely on both environmental- and host-related factors to successfully complete their life cycle (Wolinska and King 2009, Hall and Ebert 2012). Moreover, once parasites have found their hosts they still have to overcome host defenses to eventually settle durably and reproduce (Combes 2001). Combes et al. (2001) conceptualized parasite conditional success into two main steps. The first step, refers to the spatial and temporal overlap between parasites and hosts that is required for parasites to find and meet their (future) hosts, the so called “encounter filter” (also known as the *Opportunity concept*, Araujo et al. 2015). The second step –“the compatibility filter”– occurs once parasites have met their hosts and refers to the conditions required for parasites to durably exploit their hosts (Combes 2001). Since parasites can have drastic outcome on their hosts and more generally on the overall food webs, understanding the determinants of the parasites ability to successfully complete their life cycle is a critical issue that has raised a lot of interest in ecology and evolution (Morand and Deter 2007, Hoberg et al. 2015).

Host-parasite interactions have been extensively studied from a co-evolutionary perspective (Dybdahl and Lively 1998, Lambrechts et al. 2006, Little et al. 2010). Co-evolution is indeed a key process explaining how one parasite genotype and one host genotype can reach compatibility (Webster and Davies 2001, Theron et al. 2014). However, comparatively less studies have focused on the factors involved in the encounter filter, despite encountering being the first step during infection process (Ward et al. 1998, Kuris et al. 2007, Lacey et al. 2014). Switching from one host to another for a parasite or finding a new host for a parasite infectious larvae is looking

for a needle in a haystack and how parasites (often very small) are able to encounter their hosts (sometimes rare) in opened environments hence remains an intriguing question.

So far, several parasite strategies aiming at increasing their encountering rate with hosts have been documented and sometimes encompass the use of environmental or host cues, behavioral modification of hosts and synchronicity with hosts. The use of host or environmental cues allows the parasite to actively track hosts (Lawrence 1981, Pasternak 2004, Lacey et al. 2014). For instance, the parasitic barnacle *Heterosaccus dollfusi* uses chemical cues to efficiently localize its crab host (Pasternak 2004). Similarly a considerable amount of studies have focused on the intriguing ability of some parasites to alter their host behavior in order to promote infestation especially regarding parasites having complex life cycle requiring trophic transmission from one host species to another (Adamo 2002, da Silva and Langoni 2009, Mikheev et al. 2015). Alternatively, several parasites have been shown to synchronize their circadian shedding rhythms with host activity (Theron and Combes 1995, Gautret et al. 1995). Finally, high fecundity is an adaptive trait widely observed among parasite species that passively increase the encountering rate between parasite infectious stages and hosts hence enhancing the fitness of parasites (Rea and Irwin 1994). Many parasites display numerous replications within their hosts with one individual sometime producing hundred to millions of individuals of the next parasitic stage in order to increase transmission rate (Turner et al. 1995, Mackinnon and Read 1999, Ebert et al. 2000).

Surprisingly, some parasites species maintain themselves without exhibiting any of those strategies. In these cases, environmental factors might facilitate the process of infection. In this context, if parasite infectious stages are randomly distributed in the ecosystem, then the encountering rate is expected to be very low because hosts are diluted in the environment.

Alternatively, if parasite infectious stages are not randomly distributed in the ecosystem, but rather concentrated/clustered in some specific microhabitats then the encountering rate is expected to be higher provided that hosts also use (at least partly) these specific areas (Benavides et al. 2012, Paull et al. 2012). In this case, infection is expected to occur specifically at some micro-geographical “hotspots”, which can be seen as a strategy for parasite to optimize infection rate (Zimmermann et al. 2016). Some microhabitats could indeed provide particularly favorable conditions for parasite life cycle, hence concentrating locally the number of infesting stages (Browman et al. 2004, Loot et al. 2011, Marcogliese 2016). If these hotspots of infection do exist, this implies that the probability for a host of being infected is not homogeneous at the microgeographic scale, and that –if hosts are relatively sedentary– the spatial distribution of parasitized hosts might itself be spatially heterogeneous and might actually overlap the distribution of infection hotspots.

So far, some studies have tested for the existence of hotspots of infection at large scale (Ramirez-Sierra et al. 2010, Bousema et al. 2010, Murphy et al. 2012) and at local scale (Poulin and FitzGerald 1989, Thamm et al. 2009, Jennett et al. 2013), but very few studies have tested for the possibility of microgeographic heterogeneity in infection risk (but see Zimmermann et al. 2016). Nonetheless, this could improve our understanding of the factors that favor encountering between parasite and host even for parasites not displaying any common strategies to increase such encountering events. Such studies are also crucial to better predict potential infection hotspots, those being key targets for disrupting parasites life cycle and thus eradicating pathogenic parasites for Humans and/or animals.

Here we aimed to test at microgeographic scale i) whether the probability of being infected is heterogeneous rather than random ii) whether this translates in aggregated spatial distribution of



parasitized hosts and finally iii) whether environmental factors drive infection risks. We focused on an ectoparasite with a direct lifecycle and that infects several freshwater fish species, the crustacean *Tracheliastes polycolpus*. Up to our knowledge, this parasite does not actively modulate the behavior of its hosts for increasing the encountering rate (as most fish ectoparasite, but see Mikheev 2011). Moreover, each female lay a relatively low number of eggs compared to other parasitic copepods (usually less than a hundred, Loot et al. 2004, Mazé-Guilmo 2016 and personal observations) that gives birth to a single infective and free-living stage. This parasite species infects host fish species inhabiting (sometimes large) rivers with a relatively high flow, which suggests that on average, the concentration of infective stages per liter is extremely low in the river. Moreover, infesting stages of *T. polycolpus* exhibit relatively poor locomotion abilities preventing larvae from an active search of the host (Fryer 1982) and display short lifetime (Maze-Guilmo). However, *T. polycolpus* is still able to encounter efficiently its host, since it is able to infect a relatively rare host species (*Leuciscus sp.*, in general 0.02 to 0.12 individuals per m<sup>2</sup> of rivers in good areas) with some individuals sometimes harboring up to 80 parasites (Loot et al. 2004, Lootvoet et al. 2013). This system hence constitutes an ideal opportunity to test our predictions.

Specifically, we first conducted an *in-situ* caging experiment to quantify the level of spatial aggregation of infection events in experimental cages. In this experiment we used F1 hosts from a common and controlled origin so as to control for the confounding effects of host-related factors linked to the compatibility of hosts and parasites, an important prerequisite that is yet rarely taken into account. Second, we tested whether parasitized wild hosts from the same river were spatially aggregated in hotspots through an exhaustive electric fishing campaign. Finally we compared potential microhabitat characteristics that partly explain the distribution of infected hosts in our

housing and wild approach. Our specific predictions are that primary infection events of experimental hosts will be aggregated in some “hotspot” cages. If so and if wild parasitized hosts are poorly mobile, we expect patterns of spatial aggregation of wild parasitized hosts matching hotspots of infection. Finally in the experimental and the descriptive approaches we expect that environmental factors at the microgeographic scale will be related to infection risk (i.e., number of infection events, and aggregated hosts in hotspots of infection respectively).

## **Material and methods**

### **Microhabitat description**

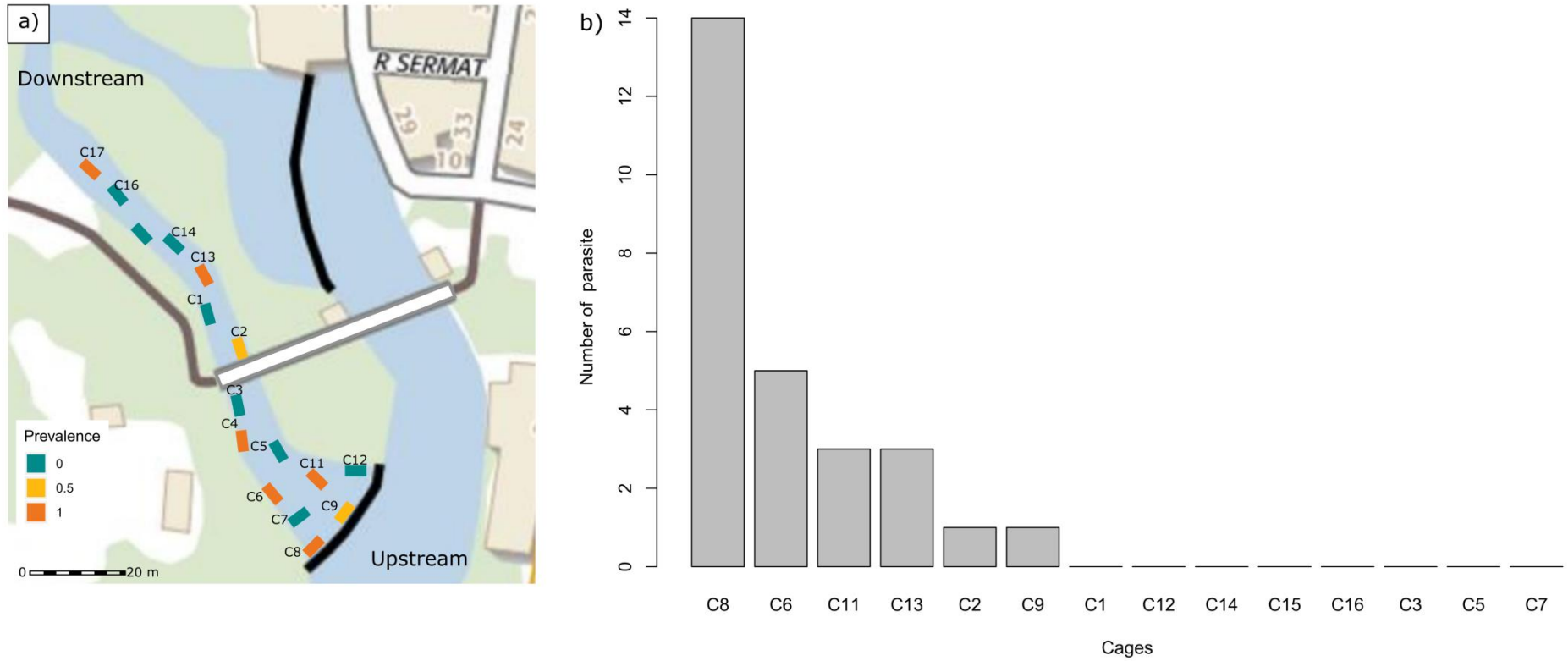
In this study microhabitat is defined by an area of 1m<sup>2</sup> matching either the area where a cage was set up or a point in which an exhaustive electro-fishing sampling of *T. polycolpus* hosts was conducted. Microhabitats were characterized by measuring the geographic coordinates (WGS 84 / UTM zone 31N), the distance (m) from the closest river bank, the water column depth (cm), the stream velocity (cm.s<sup>-1</sup>), (measured 7 cm above substrate level using the OTT Z400 counter associated with the current meter OTT C2) and the substrate composition in a one meter circle around the cage or sampling point (i.e., percentage of each substrate categories defined on substrate size: <2 cm, 2–10 cm, 10–40 cm, over 40 cm and slab). Substrate categories were then summarized in two synthetic variables using Principal Component Analysis (PCA). We kept the two first components as synthetic variables hence summarizing 68% of the variance. The first component, synthetizing 37% of the total variance, distinguished between substrate composed of medium sized substrate (10-40 cm) from substrates composed of slab and smallest substrate such as sand and silt. The second component, synthetizing 31% of the total variance, distinguished between coarse substrate (>40 cm) from small substrate (2-10 cm). We named these synthetic variables granulometry 1 and granulometry 2 respectively (Fig. S1).

## Experimental design and sampling

### *Testing heterogeneity in infection events*

We first design a semi-experimental design to test for aggregation in infection events between *T. polycolpus* and its main host (*Leuciscus burdigalensis*) in relation to microhabitat environmental features. Specifically, we performed a caging experiment during summer 2017 in a stretch of the Arize River (South-Western France, Mas d’Azil, X= 366422, Y= 4770813, WGS 84 / UTM zone 31N). More specifically, 16 cages (30x50x80 cm, and 1 cm mesh) containing two fish each were settled into various microhabitats along a river section of about 200 m long (Fig. 1-a). We selected microhabitats based on the visual inspection of stream velocity, depth and distance from the bank, and we ought to spread the cages in the main microhabitats available along the river stretch. As the infective stage of *T. polycolpus* is tiny and has a poor swimming ability, we suspected low stream velocity habitats to be better for infection and we hence aimed at contrasting these habitat types to more rapid habitats. Fish belonged to the species *L. burdigalensis* which is the main host species of *T. polycolpus* in southwestern France. We used dace of similar body size, weight and genetic background to control for potential confounding host-related factors. We therefore used F1 dace that were bred at the U3E experimental platform (INRA, Rennes, France Unité Expérimentale d'Ecologie et d'Ecotoxicologie aquatique) from wild genitors from the Oir River (Ille-et-Villaine). We selected individuals of the same age (i.e., three years old dace) and with similar body weight and size (size of 163mm on average ranging from 213mm to 155mm and weight of 45g on average ranging from 44g to 46g). All fish used in this experiment were naïve regarding previous exposure to *T. polycolpus*, although the genitors used for generated offspring have been exposed –prior to being brought to the facility– to the parasite in the wild. The caging experiment lasted 45 days during which fish were visually checked and

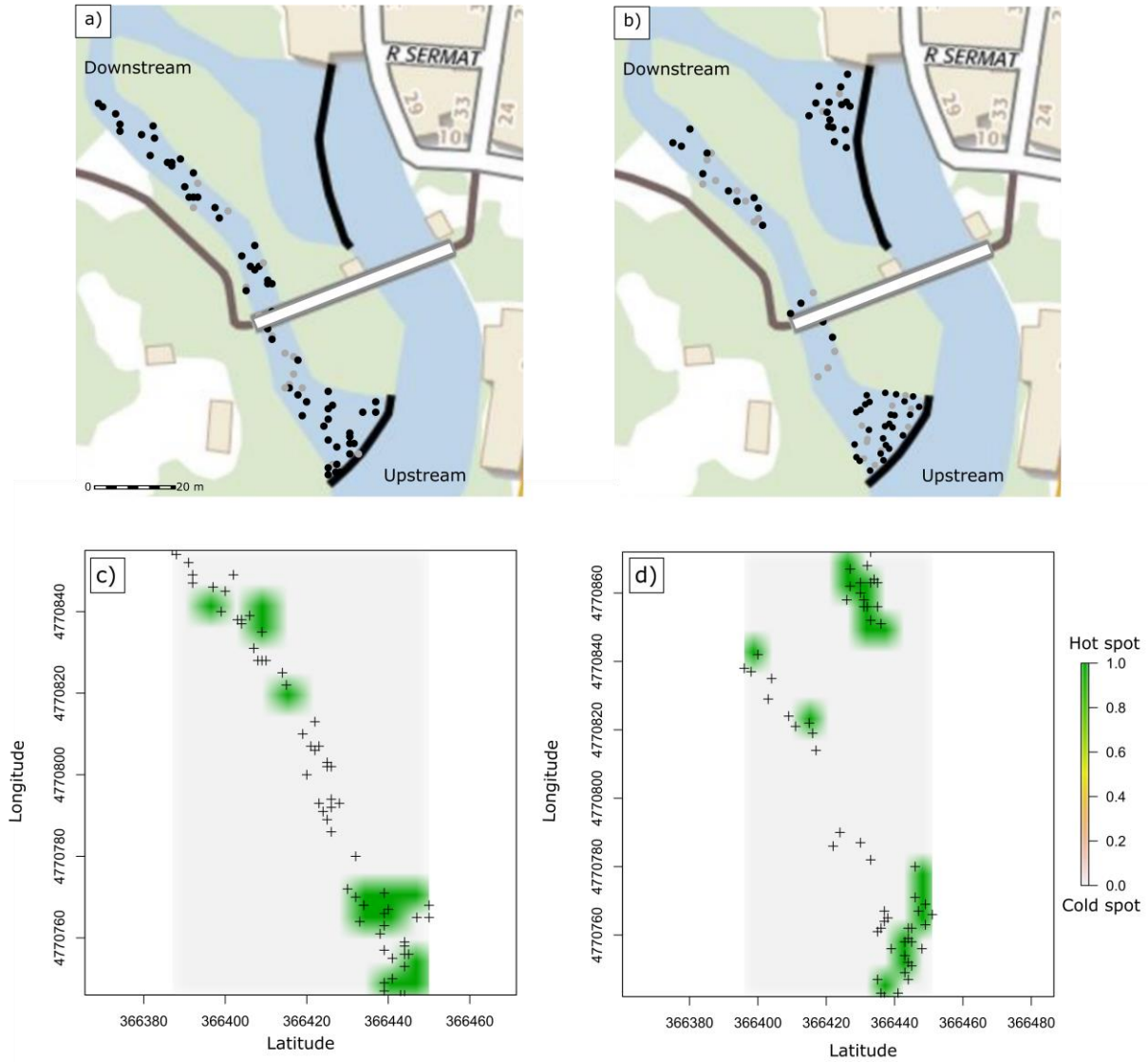
cages were cleaned every two days so as to ensure the well-being of fish and an appropriate environment for infection. After 45 days, fish were removed from the cage, they were euthanized using an overdose of benzocaine according to French legislation, and the number of parasites of each fish was recorded.



**Figure 1:** Panel showing, a) the position of each cage (rectangles) along with the prevalence recorded at the end of the caging experiment and b) the distribution of the total number of *Tracheliastes polycolpus* per cages.

### ***Testing for spatial aggregation of parasitized hosts in wild populations***

We relied on a descriptive approach to test for spatial aggregation of parasitized host in wild host populations (i.e., when not controlling for host related factors) in relation to microhabitat environmental features. Specifically, we conducted an electro-fishing sampling at the same site as for the caging experiment (Mas d'Azil in the Arize River) but replicated the sampling during summer 2017 and summer 2018. We randomly selected ~80 microhabitats along a river section of about 200 m (80 microhabitats in 2017, Fig. 2-a, and 82 microhabitats in 2018, Fig2-b) that we sampled exhaustively for potential hosts using electric-fishing. Briefly, for each microhabitat, we immersed the anode for a few seconds in the microhabitat and sampled with a net all visible fish. We specifically focused on the potential and most abundant hosts of *T. polycolpus*; the occitanian gudgeon (*Gobio occitanea*) and the common minnow (*Phoxinus phoxinus*). All individuals from a microhabitat were kept isolated, and they were anesthetized, measured and parasites were counted before fish being gently returned to the river. Overall, 977 fish were sampled in 2017 including 52 parasitized fish and 270 fish were sampled in 2018 including 34 parasitized fish.



**Figure 2:** Distribution of the sampling points for the descriptive approach *in natura*. Panels a) and b) show sampling points with hosts (black points) and without host (grey points) in 2017 a) and 2018 b) respectively, as localized in the Mas d’Azil in the Arize River. Large black lines refer to weirs. Panels c) and d) show the results of the joint count analysis in 2017 c) and 2018 d). Black crosses refer to sampling points (with hosts only) and green areas are predicted hotspots of infections (i.e., zones where parasitized hosts are aggregated in space).

## Statistical analysis

### *Parasite distribution assessment*

We first aimed to test whether infection events were randomly distributed between cages. We excluded from this analysis two cages, in which one of the two fish died, so as to control for host density. We then computed the observed aggregation coefficient  $k_{obs}$  which refers to the variance to mean ratio of *T. polycolpus* number per cage (Poulin 1993). If parasites are randomly distributed between the cages,  $k$  is expected to be close to 1. On the contrary if parasites are aggregated in a few cages (i.e. hotspots of infection), then  $k$  is supposed to be greater than 1 (Poulin 1993). We further simulated the expected aggregation coefficient  $k_{exp}$  in the case of a random distribution of *T. polycolpus*: in such a case one would expect that the number of parasites would follow a Poisson distribution of mean matching the observed mean of our dataset (i.e., about 2 parasites per cage). We finally compared  $k_{obs}$  to the distribution of 10 000 simulated  $k_{exp}$ . This test assumes that aggregation is due to the microhabitat in which the cage has been settled, and not to host-related factors (which is likely since hosts were homogenized). Nevertheless, we tested whether *T. polycolpus* distribution across cages was aggregated while taking into account potential remaining host effects. To do so, we simulated  $k_{exp}$  when reshuffling cages identity but keeping the number of parasite overserved per individuals. We then compared  $k_{obs}$  to the distribution of  $k_{exp}$  after 10 000 simulations.

We secondly tested whether parasitized hosts were spatially aggregated or on the contrary homogeneously distributed at the microgeographic scale in wild host populations. We conducted a join count analysis correlating the parasite occurrence (binary variable: 1 corresponds to the presence of at least a parasite among captured hosts in one microhabitat, 0 corresponds to the absence of parasite among all hosts captured in one microhabitat) to geographical coordinates



using the function *joincount.mc* from the R package *spdep* (Bivand 2010). This analysis is based on a theoretical distribution of the object of interest (here parasitized hosts) in a grid of cells. We defined a grid of 20x10 cells (so as a cell matches a microhabitat of 1m<sup>2</sup>) defined by actual geographical coordinates. This analysis aims to test whether occurrences or absences of the parasite are aggregated (i.e., occur in neighboring cells) in comparison to expectations if the parasitized hosts are randomly distributed in the grid (Lawson 2010). Neighbor relationships between cells were defined by queen movements. Finally, significance was assessed from 999 random permutations and separately for the sampling realized in 2017 and for the sampling realized in 2018.

### ***Testing the effect environmental factors at the microgeographic scale on infection risks***

We first aimed to test whether microhabitat environmental features were significantly linked to parasitic load observed for each individual when kept in cage. We built a generalized linear model that linked the parasitic load to microhabitat environmental features (i.e., distance from the closest rim, depth, stream velocity, granulometry 1 and granulometry 2) as well as interactions between stream velocity and depth and between stream velocity and granulometry 1 and granulometry 2 respectively. We focused on these specific interactions terms because synergetic effects between stream velocity and depth are expected to favor parasite development (i.e., both calm and deep environment, Loot et al. 2004) and because stream velocity can greatly influence substrate particles according to their size (Hjulstrom 1935) and consequently *T. polycolpus* larvae dispersion. For this analysis, all cages were considered and we therefore included the number of fish per cage as a covariate. Finally, a Poisson error term distribution was used to account for a count dependent variable. The model described and built in this section is referred hereafter as a full-model. Importantly, we did not included a random effect on cage identity since the variance

estimated for such a random effect was null (and thus negligible, Pasch et al. 2013), and we consequently preferred a more parsimonious model. Similarly, we did not take into account spatial auto-correlation in the full model since we did not detect significant spatial auto-correlation in model residuals (Moran test,  $p$ -value=0.655), and hence hypothesized that if any spatial-autocorrelation exists in the raw parasitic load, it is explained by microhabitat environmental features. From the full-model, we then conducted a model selection approach based on Akaike Information Criterion (AIC) ranking in order to select the best models using the function *dredge* from the R package *MuMin* (Barton 2016). Finally, we computed variable importance (i.e., sum of model weights over all models including each explanatory variable) when considering models having the lowest AIC and no more than four points of differences in AIC (i.e., best models).

We then tested whether the environmental features of microhabitats were linked to the parasite prevalence at the microhabitat level in wild host populations. We built a generalized linear mixed model that linked parasite prevalence (i.e., the number of parasitized fish over the total number of fish in a sampled microhabitat) to microhabitat environmental features (i.e., distance from the closest rim, depth, stream velocity, granulometry 1 and granulometry 2 as well as interactions between stream velocity and depth and between stream velocity and granulometry 1 and granulometry 2 respectively). We did not account for the total number of fish since it was included in the dependent variable coded as success/failures corresponding to parasitized/unparasitized fish and consequently used a binomial error term. We included as random effect both the year of sampling (2017 or 2018) and the fish species (gudgeon or minnow). The model described and built in this section is referred hereafter as a full-model. Similarly to caging experiment, we did not detect spatial autocorrelation in the full model

residuals (Moran test, p-value=0.865). We then selected the best models based on AIC (i.e., models having the lowest AIC and no more than four points of differences in AIC) and computed variable importance in the best models.

All statistics were run using R software (version 3.5.3, R Core Team 2017).

## **Results**

### **Distribution of infection events**

In the caging experiment, the aggregation coefficient observed ( $k_{obs}$ ) of *T. polycolpus* was of 7.5, which denoted a clear pattern of spatial aggregation in infection events by *T. polycolpus* (Poulin 1993). In other words, the probability of being infected by *T. polycolpus* was not homogeneous among cages (microhabitats), which resulted in a few cages with a high number of parasites on hosts (up to 14 parasites per cage) and half of the cages where all fish remain uninfected (Fig. 1-b). We found that  $k_{obs}$  was significantly different from  $k_{exp}$  computed from a random distribution and irrespectively from host related factors (p-value <0.001, Fig. S2-a). Similarly, we found that  $k_{obs}$  significantly differed from  $k_{exp}$  computed from a random distribution that took into account host related factors (p-value=0.002 Fig. S2-b).

### **Spatial distribution of parasitized hosts**

In the descriptive approach, the joint count analysis revealed a significant spatial aggregation of parasitized hosts in 2017 (rank of observed statistic = 956, p-value = 0.045) but the absence of parasitized hosts was not significantly aggregated (rank of observed statistic = 874.5, p-value = 0.126, see Fig. 2-c). Regarding the sampling realized in 2018, the joint count analysis revealed a significant spatial aggregation of both absence and occurrence of parasitized hosts (rank of observed statistic = 996.5, p-value = 0.004 and rank of observed statistic = 1000.5, p-value =

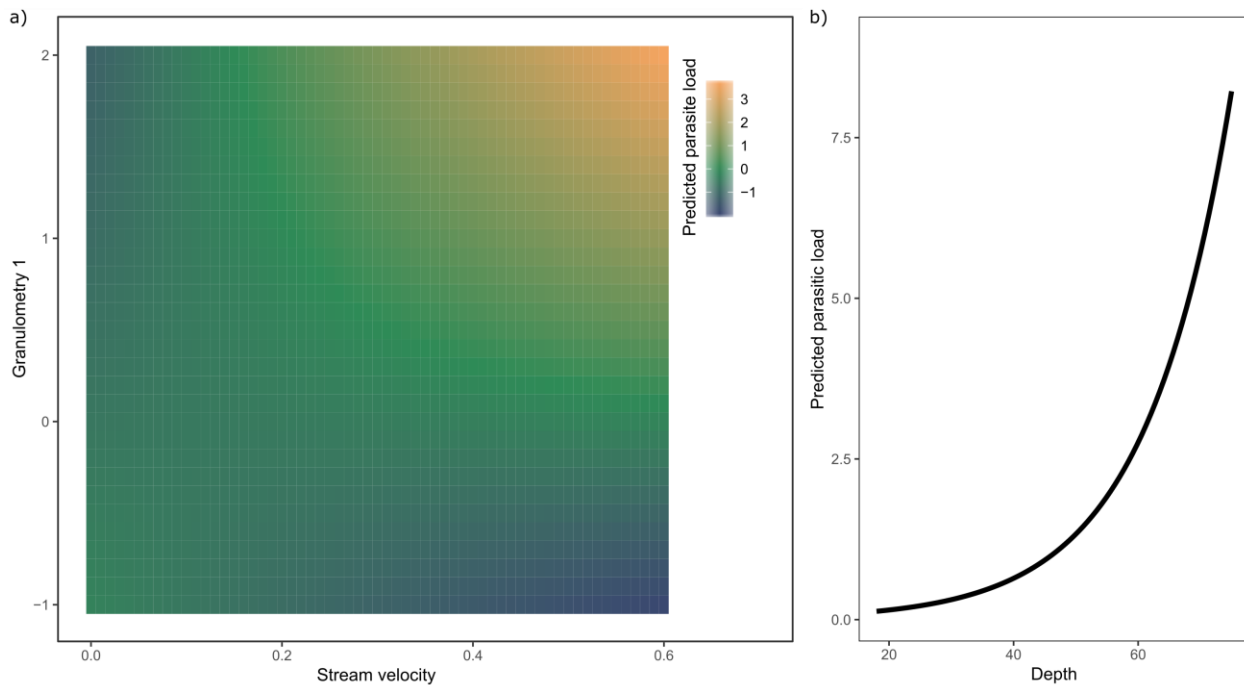
0.001 for absence and presence respectively, see Fig. 2-d). Hence, for the two years of sampling we detected hotspots in which parasitized hosts occurred more frequently than expected by chance. Interestingly, the position of these hotspots was spatially consistent between 2017 and 2018 sampling, with hotspots localized near weirs and downstream of the sampled stream section (Fig. 2-c and Fig. 2-d).

### **Effect environmental factors at the microgeographic scale on infection risks**

When investigating the relation between microhabitat features and parasitic load observed for each individual in the caging experiment, our model selection approach selected up to eight models not differing for more than four points of AIC (Table 1). These eight models (referred hereafter as best models) all retained as explicative variables depth, stream velocity, granulometry 1 and the interaction term between stream velocity speed and granulometry 1 (importance= 100%, Table 1 and Fig. S3). To a lesser extent, the variable distance from the rim was also retained in six of the eight best models having an importance of 87% (Table 1 and Fig. S3). When using the most influential variables (i.e., arbitrarily defined as the ones with an importance higher than 70%) for predictions, we found that the highest parasitic loads are expected for microhabitats characterized by low stream velocity and medium sized substrate on the one side, and for microhabitats characterized by high stream speed and substrate composed of smallest sized substrate (such as sand and silt) and slab on the other side (Fig. S1 and Fig. 3-a). Similarly, highest parasite loads are expected for microhabitats characterized by deep water column and localized far from the rim (Fig. 3-b and Fig. S4).

**Table 1:** Results of the model selection for the caging experiment. Best models linking parasitic load to microhabitat variables are displayed in AIC decreasing order. Models are displayed until a delta AIC of four. For each variable the estimate is shown. *NA* is denoted in place of the estimate when a variable is not retained in the model.

<b>Intercept</b>	<b>Depth</b>	<b>Rim distance</b>	<b>Stream velocity</b>	<b>Granul. 1</b>	<b>Granul. 2</b>	<b>Host number</b>	<b>Depth:Stream velocity</b>	<b>Stream velocity: Granul. 1</b>	<b>Stream velocity :Granul.2</b>	<b>AICc</b>
<b>-5.025</b>	0.073	0.675	3.908	-1.224	NA	NA	NA	14.461	NA	68
<b>-0.167</b>	0.068	0.607	1.725	-1.333	NA	-2.282	NA	17.097	NA	68
<b>-1.307</b>	0.092	0.679	2.035	-1.732	0.633	-2.381	NA	18.391	NA	70
<b>-5.720</b>	0.088	0.671	3.551	-1.468	0.488	NA	NA	14.892	NA	70
<b>2.434</b>	0.035	NA	-1.170	-1.361	NA	-2.163	NA	14.029	NA	70
<b>-5.038</b>	0.073	0.676	4.189	-1.219	NA	NA	-0.009	14.454	NA	71
<b>-1.903</b>	0.037	NA	0.371	-1.225	NA	NA	NA	11.341	NA	71
<b>0.215</b>	0.068	0.590	3.625	-1.340	NA	-2.457	-0.071	17.580	NA	71



**Figure 3:** Predicted parasitic load (log scaled) from the average model (based on models having the lowest AIC at a delta AIC of four) for the effect of a) the interaction between stream velocity and the first synthetic variable of substrate composition in the caging experiment and b) for the effect of depth. Highest parasitic load are expected for low stream velocity and medium sized substrate but also for high stream velocity and substrate composed of smallest sized substrate (such as sand and silt) and slab. Similarly, parasitic load is predicted to increase with the depth of the water column.

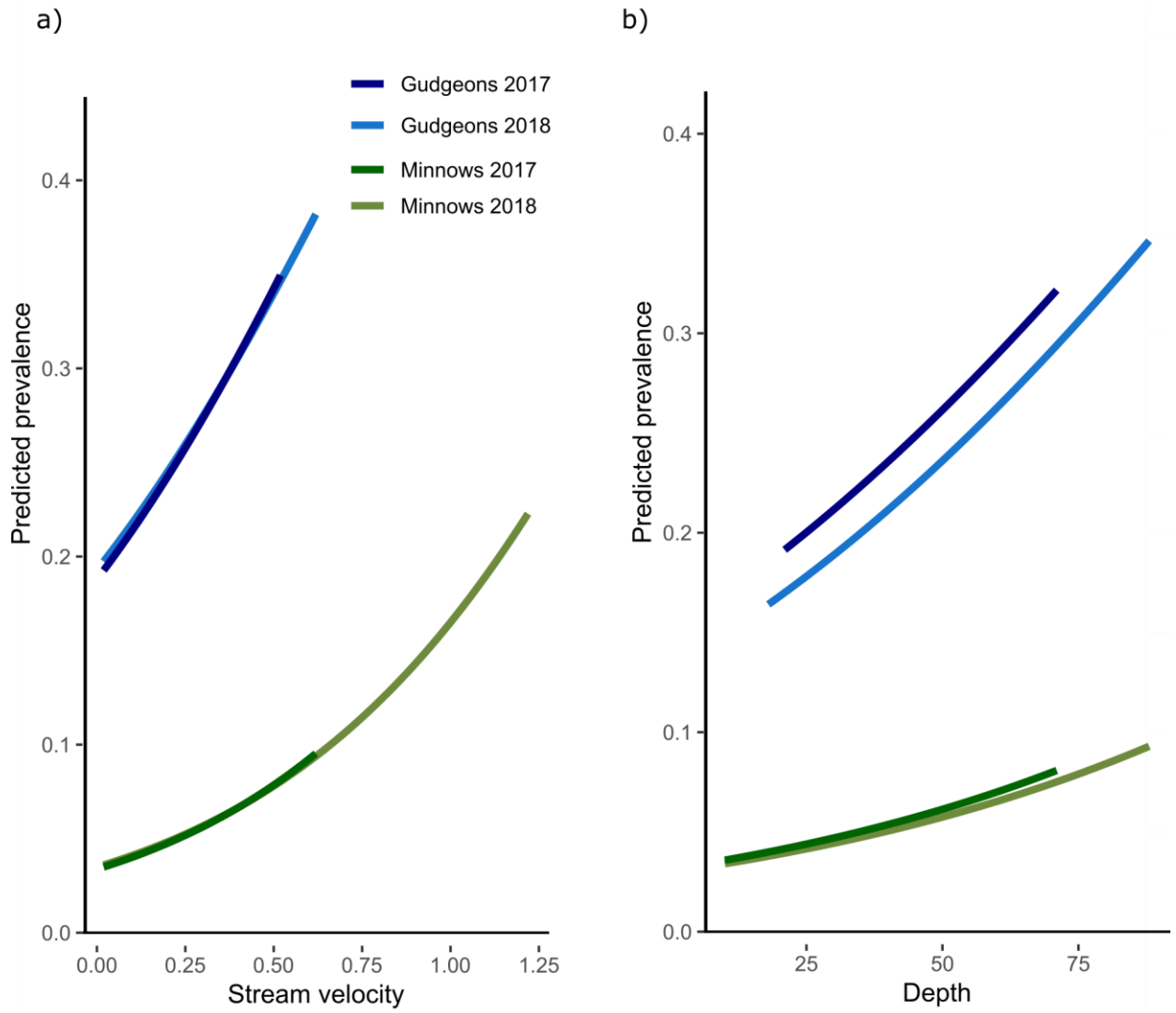
When investigating the relation between microhabitat features and parasite prevalence in the descriptive approach, our model selection approach selected up to 29 models not differing for more than four points of AIC (Table 2). The 29 best models consistently retained stream velocity variable (importance= 100%) and to a lesser extend depth (importance =79%, Fig. S5 and Table 2). Other variables were retained in some of the 29 models, and display importance always below 40% (Fig.S5). When focusing on the most influential variables (i.e., variables having an importance higher than 70%), predictions revealed that highest prevalence is expected for microhabitats characterized by high stream velocity and deep water column (Fig. 4).

**Table 2:** Results of the model selection for the descriptive approach *in natura*. Best models linking prevalence to microhabitat variables are displayed in AIC decreasing order. Models are displayed until a delta AIC of four. For each variable the estimates is shown. *NA* is denoted in place of the estimate when a variable is not retained in the model.

Intercept	Depth	Rim distance	Stream velocity	Granul. 1	Granul. 2	Depth:Stream velocity	Stream velocity:Granul. 1	Stream velocity:Granul. 2	AICc
-2.090	0.300	NA	0.442	NA	NA	NA	NA	NA	263
-2.095	0.332	0.189	0.409	NA	NA	NA	NA	NA	263
-2.159	0.229	NA	0.367	NA	NA	-0.178	NA	NA	264
-2.081	0.366	NA	0.412	-0.094	NA	NA	NA	NA	264
-2.162	0.292	NA	0.319	-0.115	NA	-0.216	NA	NA	264
-2.151	0.275	0.167	0.353	NA	NA	-0.142	NA	NA	264
-2.086	0.379	0.167	0.391	-0.074	NA	NA	NA	NA	265
-2.094	0.276	0.210	0.442	NA	-0.143	NA	NA	NA	265
-2.088	0.260	NA	0.466	NA	-0.096	NA	NA	NA	265
-2.058	NA	NA	0.336	NA	NA	NA	NA	NA	265
-2.176	NA	NA	0.276	NA	-0.158	NA	NA	0.194	265
-2.061	NA	NA	0.417	NA	-0.204	NA	NA	NA	265
-2.065	NA	0.200	0.388	NA	-0.253	NA	NA	NA	265
-2.153	0.319	0.132	0.317	-0.097	NA	-0.179	NA	NA	265
-2.184	0.227	NA	0.338	NA	-0.068	NA	NA	0.169	265
-2.159	0.186	NA	0.389	NA	-0.097	-0.182	NA	NA	265
-2.195	0.307	NA	0.274	-0.118	-0.056	NA	NA	0.208	266
-2.059	NA	0.151	0.297	NA	NA	NA	NA	NA	266
-2.079	0.328	NA	0.435	-0.093	-0.092	NA	NA	NA	266
-2.051	0.375	NA	0.446	0.019	NA	NA	0.158	NA	266
-2.162	0.250	NA	0.339	-0.114	-0.094	-0.220	NA	NA	266
-2.160	NA	0.165	0.272	NA	-0.208	NA	NA	0.160	266
-2.149	0.220	0.187	0.383	NA	-0.138	-0.141	NA	NA	266
-2.168	0.249	0.182	0.342	NA	-0.117	NA	NA	0.130	266

<b>-2.153</b>	0.296	NA	0.328	-0.091	NA	-0.210	0.033	NA	266
<b>-2.085</b>	0.324	0.188	0.424	-0.070	-0.135	NA	NA	NA	266
<b>-2.179</b>	NA	NA	0.227	-0.069	-0.169	NA	NA	0.221	266
<b>-2.085</b>	0.379	0.167	0.393	-0.071	NA	NA	0.005	NA	267
<b>-2.053</b>	NA	NA	0.324	-0.021	NA	NA	NA	NA	267





**Figure 4:** Predicted prevalence from the average model (based on models having the lowest AIC at a delta AIC of four) for the effect of stream velocity a) and depth b) respectively in the descriptive approach realized *in natura*.

## Discussion

In this study we aimed to test whether local environment features were related to infection risks by an ectoparasite of freshwater fish that is not relying on common strategies to increase its encounter rate with its host (i.e., use of environmental or host cues, behavioral modification of host, synchronicity with hosts and high fecundity).

While controlling for host related factors, we first showed that *T. polycolpus* infections did not occur randomly at the microgeographic scale, which suggests the existence of hotspots of infection at a very fine spatial extent. In our caging experiment we used standardized hosts regarding individual characteristics that are likely to impact host susceptibility to parasitism such as body size, weight, age or genetic background (Blanchet et al. 2009a, b, Sutherland et al. 2011), which permitted to take into account potential confounding effects related to the conditions of hosts. Furthermore, even when we accounted for (remaining) potential host variability we demonstrated that parasite infections were very abundant in few cages and low to null in most cages, hence suggesting that variability in exposure through space led to this aggregation pattern. It is worth noting that aggregation could result not only from host variability in term of susceptibility but also from variability associated with parasite ability to infest its host (encompassed by the notion of compatibility filter, Combes 2001). However, we conducted our experiment in a single locality in which the parasite population was shown to be genetically homogeneous and thus similar regarding their level of virulence (Mazé-Guilmo 2016). We consequently argue that parasite aggregation as observed here for *T. polycolpus* most likely results from heterogeneous infection risk through space.

We further showed that—even when host related factors are not controlled, and notably when host were allowed to move freely across the environment (i.e., in wild host populations)—a non-

random spatial distribution of parasitized hosts emerged at the microgeographic scale. Using a descriptive approach aiming to screen infection along a river section we indeed were able to identify significant aggregation of parasitized hosts. Interestingly parasitized hosts seem to aggregate in previously identified hotspots (Fig.1-a, Fig. 2-a and Fig. 2-b) which suggests that parasitized hosts tend to stay in infection hotspots or that host dispersion is affected by parasites (Fellous et al. 2011, van Dijk et al. 2015, Binning et al. 2017). We hence suggest that encountering between parasite and host can be favored in some specific areas and this even at very small spatial scale (i.e., microgeographic scale) since our study site was constituted of an only 200 m river section. Some other studies tend as well to suggest similar results, notably showing that hosts are able to develop behavioral resistance to parasite based on habitat choice in order to avoid such areas in which encountering rate with the parasite is increased (Decaestecker et al. 2002, Karvonen et al. 2004b). However, studies focusing on the role of the local context on infection risk so far have not given evidence of parasite hotspots at such a small spatial scale as we do here (Poulin and FitzGerald 1989, Thamm et al. 2009, Borer et al. 2010, Jennett et al. 2013). Habitat choice by hosts, and more precisely microhabitat choice and occupancy, is thus likely a key component of infection risk.

More precisely, we showed that local environmental features of the microhabitat were significantly associated with infection risk. Our model selection approach identified environmental variables at the microhabitat scale strongly associated to parasitic intensity and parasite prevalence in the caging experiment and in the descriptive approach respectively. Other factors than environmental features of microhabitat could have explained *T. polycolpus* aggregated distribution. For instance (Gourbière *et al.* 2015) showed that random environmental variability is enough to generate aggregation pattern, and in our descriptive approach hosts could

also have chosen specific habitats depending on their susceptibility or previous infections. However we were able to identify specific microhabitat features such as high stream velocity associated with small sized substrate, low stream velocity associated with medium sized substrate and deep water column significantly related to parasite infection. Furthermore, the environmental features of microhabitat favoring infection were similar in the descriptive approach in wild populations and in the caging experiment (notably regarding the effect of river depth) and were repeatable over time (Figure 2-c and Fig. 2-d). We consequently suggest that not only encountering rate between hosts and parasites is not homogeneous in space, but also that encountering rate is conditioned by particular environmental features of such microhabitats.

Finally we suggest that environmental features of microhabitats could favor the encountering between *T. polycolpus* and its hosts by either increasing contact rate between *T. polycolpus* and its hosts or by offering conditions that promote a higher production of infesting stages. Microhabitats favoring *T. polycolpus* infections were indeed characterized overall by high stream velocity (about,  $0.4\text{--}1\text{m}\cdot\text{s}^{-1}$ , but potentially modulated by substrate size) and deep water column (about 60–100 cm). We actually would have expect that low stream velocity zones would have favor *T. polycolpus* since such calm zone were shown to benefit the development and the density of larvae of both copepod and monogean parasites (Barker and Cone 2000a, Bjork and Bartholomew 2009a). However, in another study, Samsing et al. (2015) showed that intermediate stream velocity was associated with higher levels of parasitism by a nematode infecting rainbow trout. The authors suggested that lower velocity ease the anchoring process for infesting larvae but also that higher velocity tend to increase contact rate between host and parasite hence leading to an optimum at medium stream velocity (Samsing et al. 2015). Here we found that higher stream velocity was generally associated with higher infections risk by *T. polycolpus* but also that

low stream velocity could favor infections when associated to medium sized substrate. However, the effect of the interaction between stream velocity and substrate size on parasite infection was only detected in the caging experiment. In the caging experiment, since hosts could not escaped microhabitats, environmental effects on encountering rate could have been exacerbated compared to natural systems, thus allowing detecting some subtle environmental effects. Another main driver of *T. polycolpus* infection success was water column depth which was found in both the caging experiment and in the descriptive approach. We suggest that deep water column could first offer favored habitat with more surface hence increasing exposure between the parasite and potential host. Besides, deep water columns could also possibly increase locally the number of infesting stages of *T. polycolpus* by offering calm water zones (Loot et al. 2004). Overall, we suggest that stream zones characterized by deep water could offer good hatching and developing condition for *T. polycolpus* whereas substantial stream velocity could allow the quite movement limited infesting larvae to be re-suspended which increase contact rate with host and hence lead to microhabitats particularly favorable for *T. polycolpus* transmission.

To conclude, in this study we combined both semi-experimental and descriptive approaches in order to test the role of environmental features at microgeographic scale on the encountering rate between a parasite and its host. We consistently found that *T. polycolpus* distribution was aggregated through space, and notably identified hotspots of infection at the microgeographic scale. We highlighted a key role of environmental features of some microhabitats, mainly characterized by deep water column and high stream velocity, on infection risk by a parasite that does not rely on any documented strategy to increase its encountering rate. We hence propose that the spatial structure in microhabitats characterized by environmental features could be a

general mechanism allowing to explain how encountering between host and parasites could be achieved even in labile and opened environments such as river streams.

## References

- Adamo, S. A. 2002. Modulating the Modulators: Parasites, Neuromodulators and Host Behavioral Change. - *Brain, Behavior and Evolution* 60: 370–377.
- Araujo, S. B. L. et al. 2015. Understanding Host-Switching by Ecological Fitting. - *PLOS ONE* 10: e0139225.
- Barker, D. E., and D. K. Cone. 2000. Occurrence of *Ergasilus celestis* (Copepoda) and *Pseudodactylogyrus anguillae* (Monogenea) among wild eels (*Anguilla rostrata*) in relation to stream flow, pH and temperature and recommendations for controlling their transmission among captive eels. *Aquaculture* 187: 261–274.
- Barton, K. 2016. MuMIn: Multi-Model Inference.
- Benavides, J. et al. 2012. Transmission of Infectious Diseases En Route to Habitat Hotspots (WM Getz, Ed.). - *PLoS ONE* 7: e31290.
- Binning, S. A. et al. 2017. Parasites and Host Performance: Incorporating Infection into Our Understanding of Animal Movement. - *Integrative and Comparative Biology* 57: 267–280.
- Bivand, R. 2010. Package ‘spdep’.
- Bjork, S. J. and Bartholomew, J. L. 2009. The effects of water velocity on the *Ceratomyxa shasta* infectious cycle. - *Journal of Fish Diseases* 32: 131–142.
- Blanchet, S. et al. 2009a. Evidence of parasite-mediated disruptive selection on genetic diversity in a wild fish population. - *Molecular Ecology* 18: 1112–1123.
- Blanchet, S. et al. 2009b. Why do parasitized hosts look different? Resolving the “chicken-egg” dilemma. - *Oecologia* 160: 37–47.
- Borer, E. T. et al. 2010. Local context drives infection of grasses by vector-borne generalist viruses: Local vs. regional context and infection risk. - *Ecology Letters* 13: 810–818.
- Bousema, T. et al. 2010. Identification of Hot Spots of Malaria Transmission for Targeted Malaria Control. – *The Journal of Infectious Diseases*. 201: 1764–1774.
- Browman, H. I. et al. 2004. The effect of light on the settlement of the salmon louse, *Lepeophtheirus salmonis*, on Atlantic salmon, *Salmo salar* L. - *Journal of fish diseases* 27: 701–708.
- Combes, C. 2001. Parasitism: The Ecology and Evolution of Intimate Interactions. - University of Chicago Press.

- da Silva, R. C. and Langoni, H. 2009. *Toxoplasma gondii*: host–parasite interaction and behavior manipulation. - Parasitology Research 105: 893–898.
- de Meeûs, T. and Renaud, F. 2002. Parasites within the new phylogeny of eukaryotes. - Trends in Parasitology 18: 247–251.
- Decaestecker, E. et al. 2002. In deep trouble: Habitat selection constrained by multiple enemies in zooplankton. - Proceedings of the National Academy of Sciences 99: 5481–5485.
- Dobson, A. et al. 2008. Homage to Linnaeus: How many parasites? How many hosts? - Proceedings of the National Academy of Sciences 105: 11482–11489.
- Dybdahl, M. F. and Lively, C. M. 1998. Host-Parasite Coevolution: Evidence for Rare Advantage and Time-Lagged Selection in a Natural Population. - Evolution 52: 1057–1066.
- Ebert, D. et al. 2000. Dose effects and density-dependent regulation of two microparasites of *Daphnia magna*. - Oecologia 122: 200–209.
- Fellous, S. et al. 2011. Parasitic infection reduces dispersal of ciliate host. - Biol. Lett. 7: 327–329.
- Fryer, G. 1982. The parasitic copepoda and branchiura of British freshwater fishes. A handbook and key. - Freshwater Biological Association.
- Gautret, P. et al. 1995. The adjustment of the schizogonic cycle of *Plasmodium chabaudi chabaudi* in the blood to the circadian rhythm of the host. - Parasite 2: 69–74.
- Gourbière, S. et al. 2015. Fundamental Factors Determining the Nature of Parasite Aggregation in Hosts. - PLoS ONE 10: e0116893.
- Hall, M. D. and Ebert, D. 2012. Disentangling the influence of parasite genotype, host genotype and maternal environment on different stages of bacterial infection in *Daphnia magna*. - Proceedings of the Royal Society B: Biological Sciences 279: 3176–3183.
- Hechinger, R. F. et al. 2011. A Common Scaling Rule for Abundance, Energetics, and Production of Parasitic and Free-Living Species. - Science 333: 445–448.
- Hjulstrom, F. 1935. The Morphological Activity of Rivers as Illustrated by River Fyris. - Bull. Geol. Inst. Univ. Uppsala 25: 221-527.
- Hoberg, E. P. et al. 2015. An integrated parasitology: revealing the elephant through tradition and invention. - Trends in Parasitology 31: 128–133.
- Jennett, A. L. et al. 2013. Tick infestation risk for dogs in a peri-urban park. - Parasites & vectors 6: 358–368.
- Karvonen, A. et al. 2004. Parasite resistance and avoidance behaviour in preventing eye fluke infections in fish. - Parasitology 129: 159–164.

- Kuris, A. M. et al. 2007. An experimental evaluation of host specificity: The role of encounter and compatibility filters for a rhizocephalan parasite of crabs. - *International Journal for Parasitology* 37: 539–545.
- Lacey, E. S. et al. 2014. Close encounters: contributions of carbon dioxide and human skin odour to finding and landing on a host in *Aedes aegypti*: Response of *Aedes aegypti* to CO<sub>2</sub> and human foot odour. - *Physiological Entomology* 39: 60–68.
- Lambrechts, L. et al. 2006. Coevolutionary interactions between host and parasite genotypes. - *Trends in Parasitology* 22: 12–16.
- Lawrence, P. O. 1981. Host vibration? A cue to host location by the parasite, *Biosteres longicaudatus*. - *Oecologia* 48: 249–251.
- Lawson, A. B. 2010. Hotspot detection and clustering: ways and means. - *Environmental and Ecological Statistics* 17: 231–245.
- Little, T. J. et al. 2010. The Coevolution of Virulence: Tolerance in Perspective. - *PLoS Pathogens* 6: e1001006.
- Loot, G. et al. 2004. The effects of the ectoparasite *Tracheliastes polycolpus* (Copepoda: Lernaepodidae) on the fins of rostrum dace (*Leuciscus leuciscus burdigalensis*). *Parasitology Research* 94: 16–23.
- Loot, G. et al. 2011. Determinants of life-history traits in a fish ectoparasite: a hierarchical analysis. - *Parasitology* 138: 848–857.
- Lootvoet, A. et al. 2013. Patterns and processes of alternative host use in a generalist parasite: insights from a natural host-parasite interaction. - *Functional Ecology* 27: 1403–1414.
- Mackinnon, M. J. and Read, A. F. 1999. Genetic relationships between parasite virulence and transmission in the rodent malaria *plasmodium chabaudi*. - *Evolution* 53: 689–703.
- Marcogliese, D. J. 2016. The Distribution and Abundance of Parasites in Aquatic Ecosystems in a Changing Climate: More than Just Temperature. - *Integr. Comp. Biol.* 56: 611–619.
- Mazé-Guilmo, E. 2016. Etude du potentiel de colonisation des parasites: une approche intégrative.
- Mikheev, V. N. 2011. Monoxenous and heteroxenous fish parasites manipulate the behavior of their hosts in different ways. - *Biol Bull Rev* 1: 446–457.
- Mikheev, V. N. et al. 2015. Behavioural adaptations of argulid parasites (Crustacea: Branchiura) to major challenges in their life cycle. - *Parasites & Vectors* 8: 394–404
- Morand, S. and Deter, J. 2007. Parasitisme et régulation des populations hôtes. - *Ecologie et évolutin des systèmes parasités*. De Boeck, Bruxelles: 13–56.
- Murphy, T. M. et al. 2012. The prevalence and distribution of *Alaria alata*, a potential zoonotic parasite, in foxes in Ireland. - *Parasitol Res* 111: 283–290.



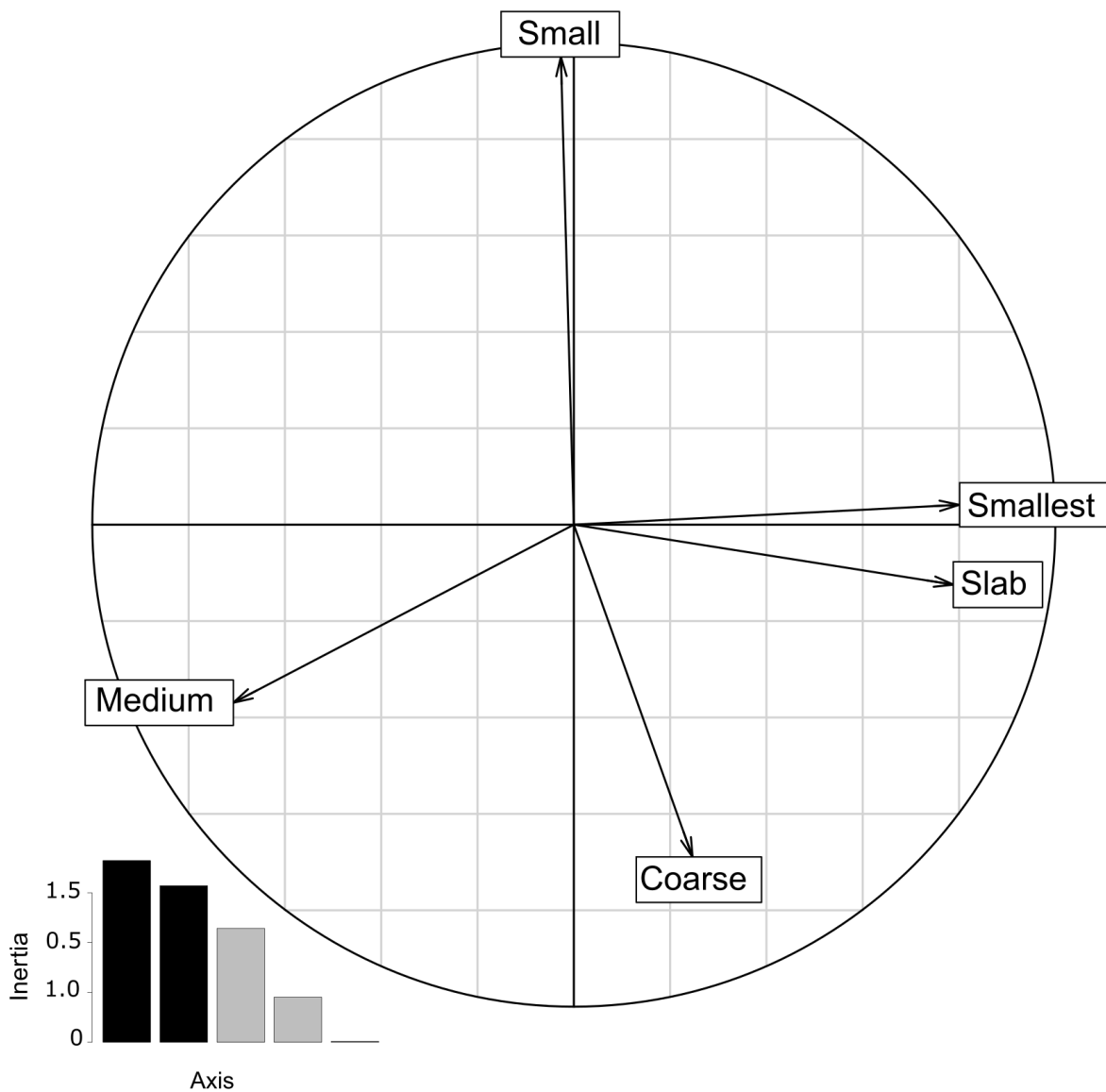
- Pasch, B. et al. 2013. Interspecific Dominance Via Vocal Interactions Mediates Altitudinal Zonation in Neotropical Singing Mice. - *The American Naturalist* 182: 161–173.
- Pasternak, Z. 2004. Host location by larvae of a parasitic barnacle: larval chemotaxis and plume tracking in flow. - *Journal of Plankton Research* 26: 487–493.
- Paull, S. H. et al. 2012. From superspreaders to disease hotspots: linking transmission across hosts and space. - *Frontiers in Ecology and the Environment* 10: 75–82.
- Poulin, R. 1993. The disparity between observed and uniform distributions: A new look at parasite aggregation. - *International Journal for Parasitology* 23: 937–944.
- Poulin, R. and FitzGerald, G. J. 1989. Risk of parasitism and microhabitat selection in juvenile sticklebacks. - *Can. J. Zool.* 67: 14–18.
- R Core Team 2017. R: A language and environment for statistical computing.
- Ramirez-Sierra, M. J. et al. 2010. Patterns of house infestation dynamics by non-domiciliated *Triatoma dimidiata* reveal a spatial gradient of infestation in rural villages and potential insect manipulation by *Trypanosoma cruzi*. - *Tropical Medicine & International Health* 15: 77–86.
- Rea, J. G. and Irwin, S. W. B. 1994. The ecology of host-finding behaviour and parasite transmission: past and future perspectives. - *Parasitology* 109: S31–S39.
- Samsing, F. et al. 2015. Gone with the flow: current velocities mediate parasitic infestation of an aquatic host. - *International Journal for Parasitology* 45: 559–565.
- Sutherland, B. J. G. et al. 2011. Differentiating size-dependent responses of juvenile pink salmon (*Oncorhynchus gorbuscha*) to sea lice (*Lepeophtheirus salmonis*) infections. - *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics* 6: 213–223.
- Thamm, S. et al. 2009. Ectoparasite Infestations of Hedgehogs (*Erinaceus europaeus*) are Associated with Small-Scale Landscape Structures in an Urban–Suburban Environment. - *EcoHealth* 6: 404–413.
- Theron, A. and Combes, C. 1995. Asynchrony of Infection Timing, Habitat Preference, and Sympatric Speciation of Schistosome Parasites. - *Evolution* 49: 372–375.
- Theron, A. et al. 2014. Multi-parasite host susceptibility and multi-host parasite infectivity: A new approach of the *Biomphalaria glabrata*/*Schistosoma mansoni* compatibility polymorphism. - *Infection, Genetics and Evolution* 26: 80–88.
- Turner, C. M. R. et al. 1995. Replication, differentiation, growth and the virulence of *Trypanosoma brucei* infections. - *Parasitology* 111: 289–300.
- van Dijk, J. G. B. et al. 2015. Weak negative associations between avian influenza virus infection and movement behaviour in a key host species, the mallard *Anas platyrhynchos*. - *Oikos* 124: 1293–1303.
- Ward, S. A. et al. 1998. Mortality during dispersal and the cost of host-specificity in parasites: how many aphids find hosts? - *Journal of Animal Ecology* 67: 763–773.

Webster, J. P. and Davies, C. M. 2001. Coevolution and compatibility in the snail–schistosome system. - *Parasitology* 123: 41–56.

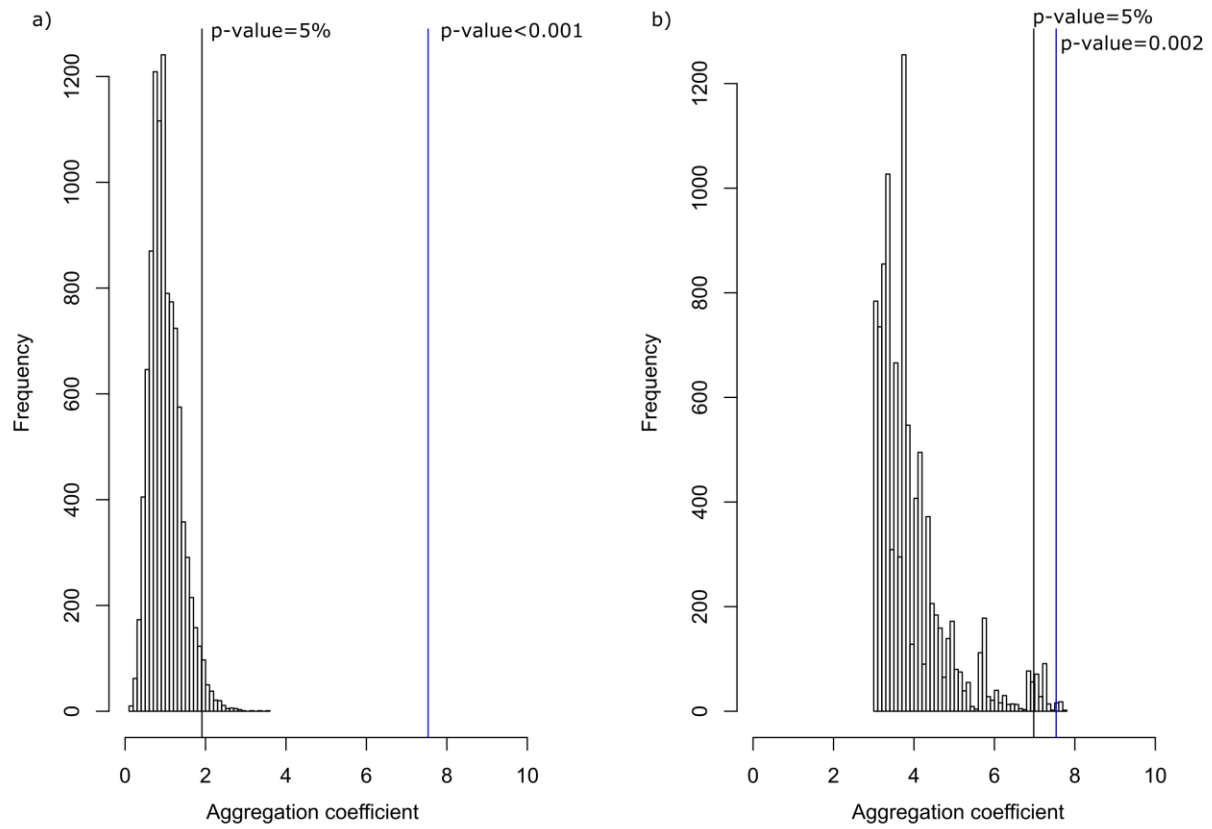
Wolinska, J. and King, K. C. 2009. Environment can alter selection in host–parasite interactions. - *Trends in Parasitology* 25: 236–244.

Zimmermann, M. R. et al. 2016. Microhabitat Differences in the Benthic Substrata Affect Parasitism in a Pulmonate Snail Host, *Helisoma anceps*. - *Journal of Parasitology* 102: 306–311.

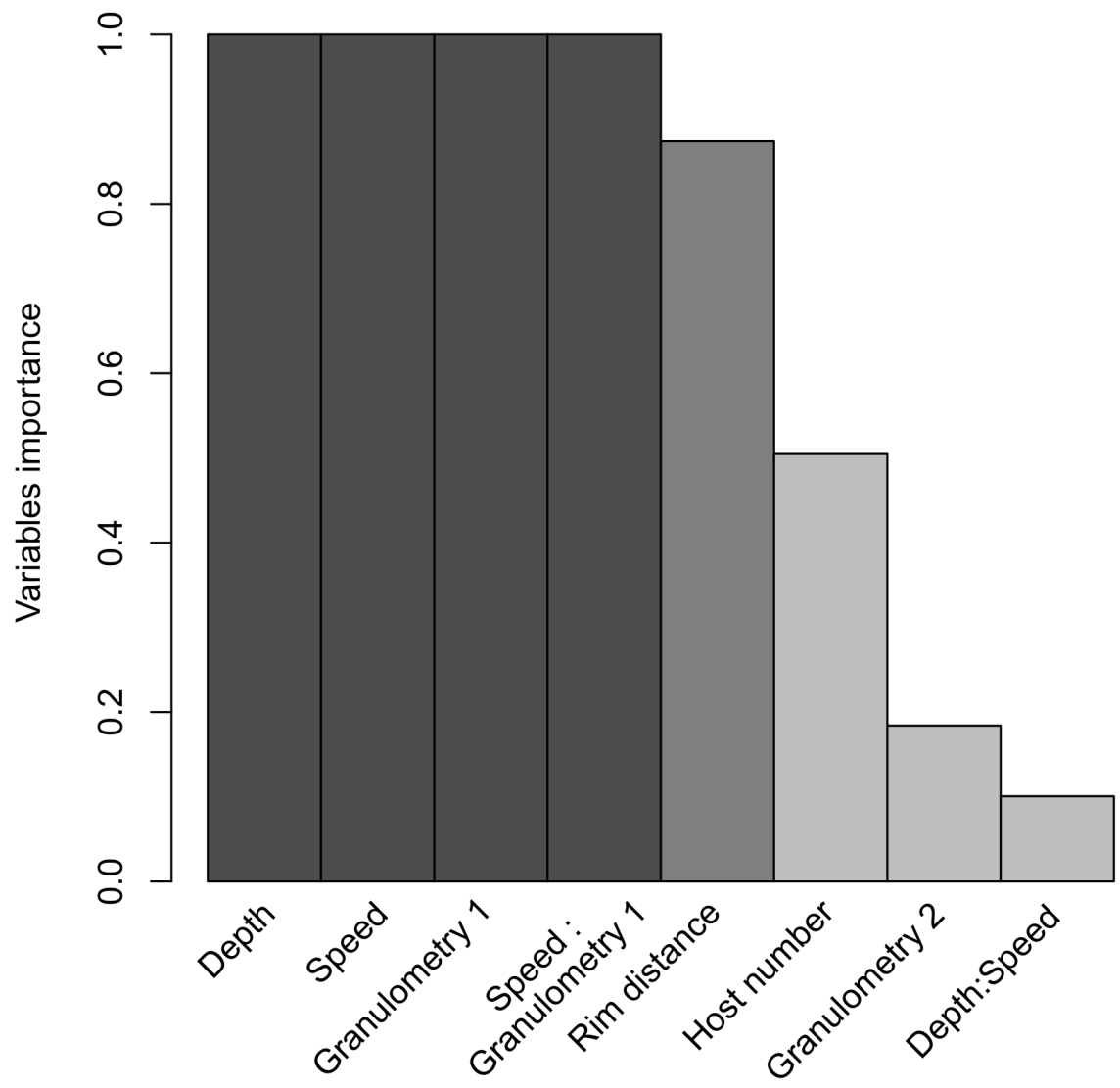
## Supplementary



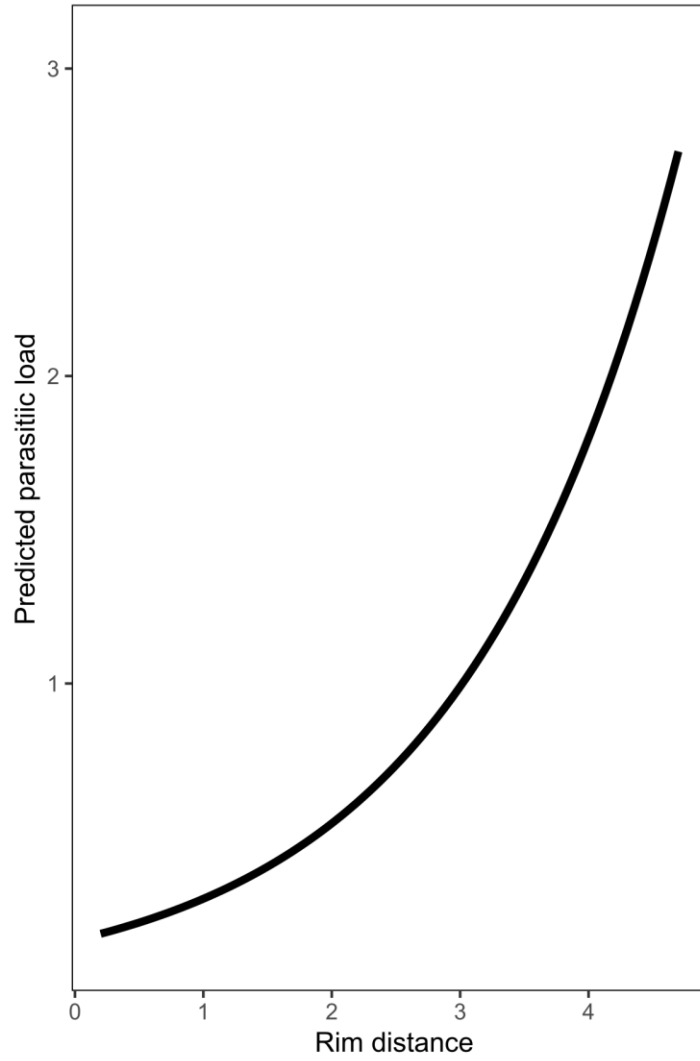
**Figure S1:** Correlation circle of substrate variables realized on the two first PCA components. Barplots indicate how much of the variance is explained by each PCA component and black bars indicate retained PCA component as synthetic variables in subsequent analyses (Granulometry 1 and Granulometry 2). Granulometry 1 distinguished substrate composed of medium sized substrate (10-40 cm) from substrates composed of slab and smallest substrate such as sand and silt. Granulometry 2 distinguished coarse substrate (>40 cm) from small substrate (2-10 cm).



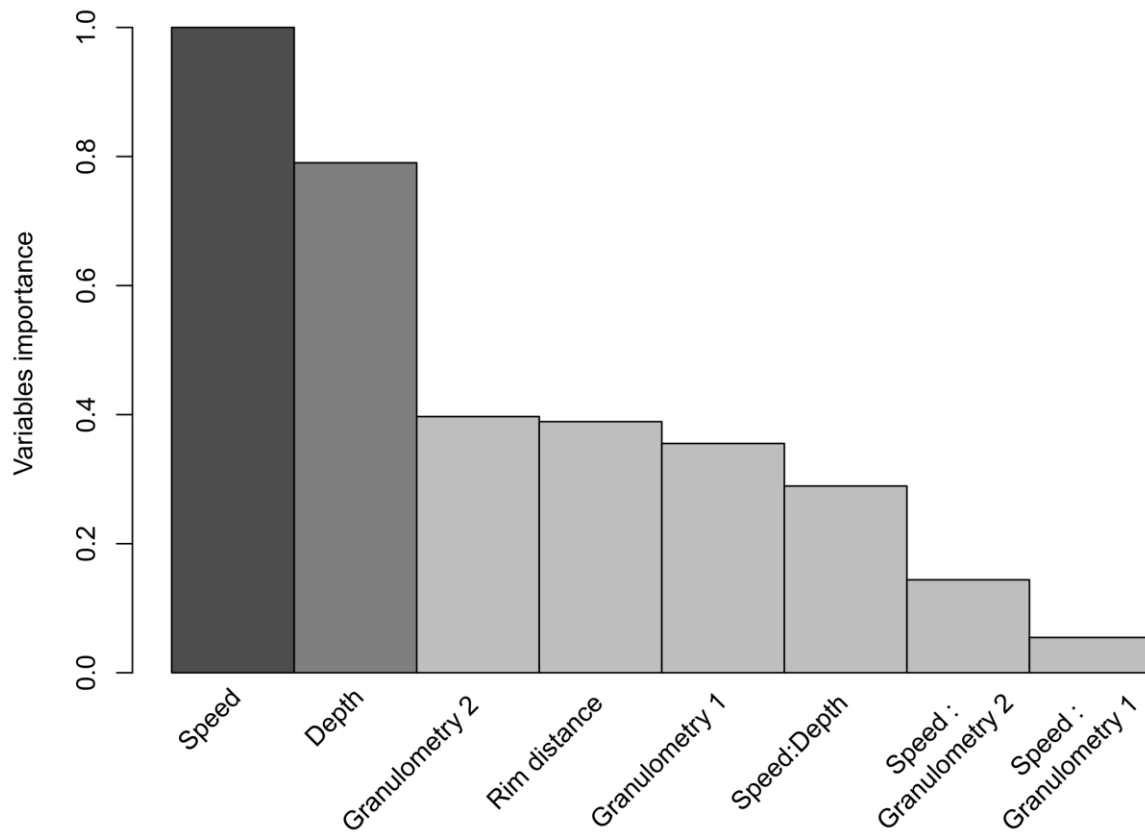
**Figure S2:** Distribution of the coefficient of aggregation ( $k$ ) simulated for a random distribution of *T. polycolpus* without accounting for host effects a) and when accounting for host effect b). Black lines stand for  $k$  at 95% of the distribution (corresponding to a p-value of 5%) and blue lines stand for observed  $k$ .



**Figure S3:** Variable importance (i.e., frequency) computed as variable occurrence in the best models linking parasitic load to microhabitat variables in caging experiment. Best models refer to models having the lowest AIC and with no more than 4 points of differences in AIC. The darker bars are the most important corresponding variables are.



**Figure S4:** Predicted parasitic load (log scaled) from the average model (based on models having the lowest AIC at a delta AIC of four) for the effect of rim distance in the caging experiment.



**Figure S5:** Variable importance (i.e., frequency) computed as variable occurrence in the best models linking parasite prevalence (measured as the number of parasitized fish over the total number of fish) to microhabitat variables in the descriptive approach realized *in natura*. Best models refer to models having the lowest AIC and with no more than 4 points of differences in AIC. The darker bars are the most important corresponding variables are.

## **Chapitre II :**

# **Rôle des déterminants environnementaux et génomiques dans la répartition d'un parasite émergent**



## Résumé

Apporter une meilleure compréhension des facteurs qui conditionnent la répartition des espèces est nécessaire afin de prévoir correctement les aires de répartition futures des espèces colonisatrices. Cependant, cette tâche s'avère difficile et tout particulièrement pour les espèces impliquées dans des interactions étroites telles que les parasites. En effet, la distribution des parasites est vraisemblablement façonnée par une interaction complexe entre des facteurs liés à l'hôte, au parasite et à leur environnement commun. Nous avons développé dans cette étude une approche originale combinant un modèle de distribution d'espèce (SDM) et des outils de génomique des populations afin de déterminer si les conditions environnementales locales ou bien le patrimoine génomique de l'hôte limitaient la colonisation d'un ectoparasite de poisson d'eau douce émergent (*Tracheliastes polycolpus*). Nous avons émis l'hypothèse que l'absence de *T. polycolpus* dans certaines zones pourrait être due à un environnement défavorable dans ces zones («hypothèse de l'environnement favorable») ou à la présence d'hôtes résistants dans ces zones («hypothèse du patrimoine génomique hôte favorable»). En utilisant un SMD à l'échelle spatiale de la France, nous avons d'abord constaté que les conditions environnementales des zones non infectées étaient aussi favorables pour le parasite que celles des zones infectées. Par ailleurs, en utilisant des données de polymorphisme nucléotidique (SNP) à l'échelle du génome hôte, nous avons démontré qu'il existait une forte association entre les occurrences spatiales du parasite et le patrimoine génomique de l'hôte. En particulier, les zones dans lesquelles le parasite était absent abritaient des populations hôtes uniques du point de vue génomique et quatre-vingt-dix SNP étaient associés de manière significative à la prévalence du parasite au niveau individuel. La distribution spatiale de *T. polycolpus* (et son potentiel de colonisation) s'avère donc plus probablement due aux caractéristiques intrinsèques de l'hôte associées à la résistance contre ce parasite, plutôt qu'aux conditions environnementales locales. Cette étude illustre l'utilité de combiner des approches de modélisation et de génomique des populations pour identifier les facteurs déterminants dans la distribution des espèces et cela en vue d'améliorer les prédictions concernant leurs futures aires de répartition.

# Combining species distribution models and population genomics underlines the determinants of range limitation in an emerging parasite

---

**Running Head:** Determinants of parasite range limitation

Mathieu-Bégné Eglantine<sup>1\*</sup>, Loot Géraldine<sup>1,2</sup>, Mazé-Guilmo Elise<sup>1,3</sup>, Mullet Valentin<sup>1,3</sup> and  
Blanchet Simon<sup>1,3</sup>

<sup>1</sup> Evolution et Diversité Biologique, Université Paul Sabatier (UPS), Centre National de la  
Recherche Scientifique, Ecole Nationale pour la Formation Agronomique (ENFA); UMR5174  
EDB, 118 route de Narbonne, F-31062 Toulouse, France

<sup>2</sup> Institut Universitaire de France (IUF), Paris, France

<sup>3</sup> Station d'Ecologie Théorique et Expérimentale, CNRS, UPS; UMR5321 SETE, 2 route du  
CNRS, F-09200 Moulis, France

**Corresponding author:**

\*Mathieu-Bégné Eglantine

UMR5174 EDB

118 route de Narbonne

31062 Toulouse

E-mail:[eglantine.mb@gmail.com](mailto:eglantine.mb@gmail.com)

## **Declarations**

### **Authors' contributions:**

GL and SB coordinated the study, supervised the data acquisition and the molecular analyses. EMB, EMG and VM acquired complementary data. EMB ran statistical and bioinformatics analyses. EMB, SB and GL wrote the manuscript. All authors participated to the conception of the ideas, and they all agree to be held accountable for the content therein and approve the final version of the manuscript.

### **Competing interest statement:**

We have no competing interest.

### **Acknowledgments:**

We would like to thank all the people that contributed to this study, notably people who helped collecting data on the field such as people from our own team during that period (V. Dubut, I. Paz-Vinas, C. Veyssi re ...) but also people from The Agence National pour la Biodiversit  (AFB) and the F d rations D partementales de P che et de Protection des Milieux Aquatiques (FDPPMA). We also are grateful to people who gave advice and feedback on this work such as M. Bernard, M. Chevalier, J-B. Ferdy, L. Fourtune, E. Goetz, C. Mathieu and J. Prunier.

**Key words:** Species Distribution Model, Population genomics, Genome Wide Association Study, Invasive species, Integrative approaches, Freshwater

## Abstract

Understanding the drivers of species distribution is required to properly predict the future geographic ranges of colonizing species. Yet, this task is challenging, notably for species involved in intimate interactions such as parasites, since their distribution is likely shaped by a complex interplay between environmental-, host- and parasite-related factors. Here, we developed an original approach combining species distribution models (SDMs) and population genomics to test whether local environmental conditions or the host genomic background most likely limit the colonization of an emerging freshwater fish ectoparasite (*Tracheliastes polycolpus*). We hypothesized that the absence of *T. polycolpus* in some areas may be due to an unsuitable environment in these areas (the “environmental suitability hypothesis”) or to the presence of resistant hosts in these areas (the “genomic background hypothesis”). Using a SDM set at the spatial extent of France, we first found that the environmental conditions of uninfected areas were as suitable for the parasite as those of infected areas. Then, using Single Nucleotide Polymorphisms (SNPs) data at the genome scale, we demonstrated that there was a strong association between the spatial occurrence of parasites and the host genomic background. In particular, areas in which the parasite was absent sustained unique host populations from a genomic standpoint, and ninety SNPs were significantly associated to parasite prevalence at the individual level. We concluded that the spatial distribution of *T. polycolpus* (and its colonization potential) was more likely due to intrinsic host characteristics associated to parasite resistance, rather than to the local environmental conditions. This study illustrates the usefulness of combining modeling and genomic approaches to reveal the determinants of species distribution and for improving predictions about their future ranges.

## **Introduction**

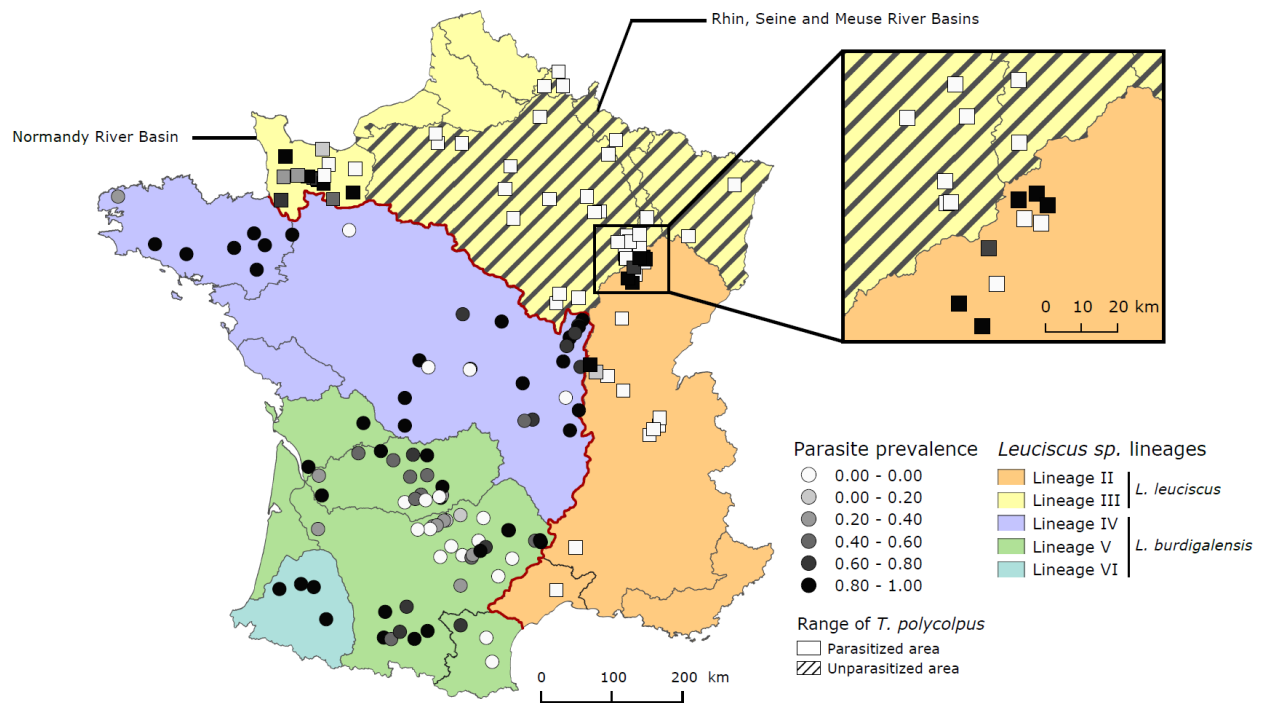
Unrevealing the factors limiting species ranges has long fascinated ecologists (Gaston 2009, Wiens et al. 2010). This issue has recently gained particular attention because of the major implications of assessing the future geographic ranges of species facing global change (Guisan and Thuiller 2005, Moor et al. 2015, Narouei-Khandan et al. 2016). This is particularly true for emerging parasites that have already been observed to expand (or shift) their ranges due to human activities, which can have direct negative feedback on human health and biodiversity (Kutz et al. 2005, Morgan et al. 2009, Escobar et al. 2017). Nonetheless, parasites are involved in a triple interaction between the parasite, the host and their common environment (Wolinska and King 2009). Hence, both the ecological conditions of the surrounding environment and factors related to host defense or parasite virulence are likely to determine parasite spatial distribution (Grosholz 1994, Bozick and Real 2015).

The influence of environment, host- and parasite-related factors on parasite spatial distribution has often been assessed separately. Studies on the environmental determinants of parasite distribution generally focus on the abiotic features that are likely affecting the demographic dynamics of parasites (Lachish et al. 2013, Pérez-Rodríguez et al. 2013, Pickles et al. 2013). For instance, for many parasites, a higher temperature has been shown to be associated to a broader seasonal window for an efficient transmission to hosts (including transmission via vectors), which facilitates the colonization of new areas (Kutz et al. 2005, Paull et al. 2015, Caminade et al. 2019). Recent modeling approaches tend nonetheless to integrate both abiotic factors and biotic factors (notably host distribution) as potential determinants of the spatial distribution of parasites (Giannini et al. 2013, Gehman et al. 2018, Byers et al. 2019).

Alternatively, a surge of studies have aimed to understand the distribution of parasites by focusing on the individual characteristics of hosts (physiology, behavior, immunity...) or parasites (virulence factors) underlying infection levels. Notably, host genomic background, for instance the diversity of genes involved in the Major Histocompatibility Complex, seems to be a key determinant of individual infection level (Bernatchez and Landry 2003, Wegner et al. 2003, Aguilar et al. 2016). Consequently, in geographic areas in which founder effects and/or bottlenecks have led to a heterogeneous spatial distribution of alleles, parasite distribution could be spatially limited because of the presence of specific alleles conferring resistance (Zhang et al. 2015, Charbonnel et al. 2019). Next-Generation Sequencing approaches have recently permitted to go further by considering thousands of both non-coding and coding markers related to host defense at the genome-wide level (Wenzel et al. 2015, Benavides et al. 2016, Zueva et al. 2018). Although these studies are usually set at small spatial scales (but see Schwabl et al. 2017, Vajana et al. 2018), they provide crucial complementary information to that gathered at larger spatial scales from approaches linking environmental features to parasite distribution (Cardon et al. 2011, Fourcade et al. 2014, Vajana et al. 2018).

We focused on the fish ectoparasite, *Tracheliastes polycolpus* that has rapidly emerged in French watersheds from northwestern Europe (Rey et al. 2015). *Tracheliastes polycolpus* is a non-native species in France that was accidentally introduced in the Loire River Basin in the 1920's, most probably from Central Europe through the introduction of its main fish host, *Leuciscus idus*, for aquaculture (Rey et al. 2015). From this location, it expanded its range within 15–20 years by invading almost all French river basins (Rey et al. 2015). However, repeated surveys by local managers never recorded infection by *T. polycolpus* in the northeastern river basins of France such as the Seine, the Rhin and the Meuse River Basins (Fig. 1). This is highly surprising given

that (i) the parasite has been able to rapidly colonize all other river basins in France (which suggests a high propensity to disperse), and (ii) it is found in sites directly adjacent to these uninvaded river basins (see the inset Fig.1) that are connected through man-made channels. Teasing apart environmental- and host-related characteristics limiting the parasite's range expansion is an important, yet challenging, objective that could be fulfilled using integrative approaches.



**Figure 1:** Map showing the distribution of sampled sites in the SDM approach for the two *Leuciscus* species (dots, *Leuciscus burdigalensis*; squares, *Leuciscus leuciscus*), together with their levels of parasite prevalence (expressed as percentage of parasitized hosts). The main *Leuciscus* lineages are also displayed. Grey lines delimitate the main River Basin (and some are indicated). The red line shows range delimitation between the two *leuciscus* species. Finally, hatched area refers to area where hosts are consistently unparasitized. The *L. leuciscus* Lineage I is not represented on this map since its range covers central Europe Rivers and is not present in France (see Costedoat et al. 2014).

Here, we aimed to test environmental and host-related factors that underlie the spatial distribution of *T. polycolpus*, and that hence explain why its range expansion is currently limited. Specifically, we tested two non-exclusive hypotheses. First, the “environmental suitability hypothesis” states that environmental features (*sensu largo*) are driving the distribution of *T. polycolpus* in France, and it can be predicted that environmental features of the unparasitized area (Fig. 1) are unsuitable for the *T. polycolpus* life-cycle. Secondly, the “genomic background hypothesis” states that the spatial distribution of *T. polycolpus* is driven by individual host characteristics, and it can be predicted that the genomic background of hosts from the unparasitized area are different from that of hosts from parasitized areas and likely confers a resistance to *T. polycolpus*. To test these two, non-mutually exclusive, hypotheses, we combined Species Distribution Modelling (SDM) approach and population genomic approach respectively. The first approach aimed to statistically test whether or not the environment is suitable for parasites in the unparasitized area. In contrast, the second approach aimed to test whether the genomic structure of host populations underlines the difference in parasite prevalence among geographic areas and whether some genomic regions are statistically related to parasite prevalence measured at the individual level.

## **Material and Method**

### **Biological models**

*Tracheliastes polycolpus* is a crustacean copepod in which females are ectoparasite of freshwater fishes (males are dwarf and free-living, Piasecki, 1839). Females anchor to the fins of fish where they are fecundated by a single male and rapidly develop two egg sacs carrying about a hundred of eggs (Loot et al., 2011). After hatching, free-living copepodids are released in the water; this represents the infective stage of the parasite (Mazé-Guilmo, 2016). The feeding activity of the



parasite directly causes fin damage and inflammation which indirectly affect host fitness (growth and survival, Blanchet et al. 2009b).

In French watersheds, *Leuciscus burdigalensis* and *Leuciscus leuciscus* (rostrum dace and common dace respectively) are the main host species of *T. polycolpus*. *Leuciscus leuciscus* inhabits rivers from the northern and eastern parts of France and extends its range far east in Eurasia, whereas *L. burdigalensis* is endemic to rivers of southwestern France (Fig. 1). Previous studies based on microsatellite and mitochondrial markers demonstrated that these two species are structured into five lineages in France (i.e., three and two lineages for *L. burdigalensis* and *L. leuciscus* in France respectively, Costedoat et al., 2014), that are well geographically structured by French river basins (Fig. 1). Notably, the unparasitized area encompasses three river basins (Rhin, Seine and Meuse, Fig. 1 and Fig. 2) that are inhabited by a single lineage of *L. leuciscus* (hereafter referred to as “lineage III-unparasitized”, see Fig. 1), and this lineage also inhabits the Normandy River Basin where the parasite is present (hereafter referred to as “lineage III-parasitized”, see Fig. 1). Despite strong genetic differentiation, the two *Leuciscus* species have very similar ecological requirements. They both live in running waters with moderate water temperature (from 14°C to 24°C in warmer months), and they are gregarious species usually feeding on benthic invertebrates. *Tracheliastes polycolpus* also has several alternative host species including *Parachondrostoma toxostoma*, *Gobio sp.*, *Phoxinus phoxinus*, *Rutilus rutilus* and *Squalius cephalus* (Lootvoet et al. 2013).



**Figure 2:** Map showing the distribution of sampled sites in the population genomic approach. Grey lines delimitate the main river basins and their names are indicated in white rectangles for relevant ones. Hatched area refers to area where hosts are consistently unparasitized.

### Testing the environmental suitability hypothesis.

We aimed to test whether environmental conditions could limit *T. polycolpus* distribution notably in the host lineage III-unparasitized. We built a SDM at the spatial extent of France linking environmental features to parasite prevalence measured at the sampling site level. Prevalence always refers to the number of parasitized host over the number of total host. In this section we considered prevalence measured at the site level, which corresponds to the number of parasitized dace at one site over the number of dace caught at this site. We used the SDM to test which of the environmental variables affect the distribution of *T. polycolpus* in France, and then used it to predict prevalence as expected based on the actual environmental features. If the environment

suitability hypothesis is verified, we expect the SDM to predict low parasite prevalence in the lineage III-unparasitized.

**Field data.** 158 sites were sampled in French rivers from 2009 to 2011 (Fig. 1). Sites were sampled using electric-fishing by the the Agence Française de la Biodiversité (AFB), the Fédérations Départementales de Pêche et de Protection des Milieux Aquatiques (FDPPMA) or directly by our own team. At each site, the number of dace sampled was recorded (n=1935, on average 12 individuals per site), as well as their individual body length (mm) and the prevalence of *T. polycolpus* measured at the site level (number of daces carrying at least one parasite on the total number of daces sampled). In most sites, a piece of pelvic fin of each dace (5mm<sup>2</sup>) was sampled and preserved in 95° alcohol.

**Environmental data.** We focused on thirteen abiotic and biotic variables that may explain the spatial distribution of *T. polycolpus*. We focused on biotic variables measured at the individual host level, biotic variables measured at the site level and abiotic variables measured at the site level (see Table 1 for a detailed list).

**Table 1:** Summary of the variables used in the Species Distribution Model aiming to account for *T. polycolpus* distribution in France. For each variable are indicated the organization level at which the variable has been measured, the name of the variable and associated biological hypotheses.

Level	Explicative variables	Hypotheses
<b>Individual level</b>	Expected heterozygosity	Global level of individual heterozygosity has been shown to impact parasite resistance in many host species, including dace (Blanchet et al. 2009a, Cardon et al. 2011)

	Dace body length	Larger hosts are expected to harbor more parasites because they are older (and less resistant) and have a higher body surface (Cardon et al. 2011)
	Species identity	We expect that the two main hosts species (i.e., <i>L. leuciscus</i> and <i>L. burdigalensis</i> ) on which we measured prevalence may differ in their intrinsic ability to resist the parasite
<b>Site level (biotic)</b>	Number of hosts species	<i>T. polycolpus</i> is a generalist species (i.e. it can be found at lower prevalence on other cyprinid species than <i>Leuciscus sp.</i> , Lootvoet et al. 2013) and may be affected by the dilution effect: the higher the host species richness, the lowest the prevalence (Johnson et al. 2008)
	Fish species richness	“Fish species richness” refers to the same hypothesis than above (the dilution effect) except than it focuses on all species of the fish community since it can include potential host species that are fully resistant (Roche et al. 2012)
<b>Site level (abiotic):</b>	<b>Climatic variables</b>	
	Temperature (PC1)	Temperature affects several aspects of the parasite life-cycle (Cardon et al. 2011, Mazé-Guilmo et al. 2016) and may affect host resistance
	<b>Hydrographic variables</b>	
	Hydrology (PC1)	We expect that the weakly mobile infective larvae may be favored in waters with low velocities/discharges
	Topology (PC2)	Different topologic characteristics (slope, altitude) could define habitats that are more or less suitable to the parasite: for instance sites at higher altitude may be less suitable

<b>Landscape variables</b>	
Landscape 1 and Landscape 2 (PC1 and PC2)	Landscape composition around the site describe a particular level of anthropogenic pressures that can affect the success of non-native species
<b>Fragmentation variables</b>	
Dam density	Dams affect water velocity/discharge and may influence the survival of the infective larva
Weirs density	In the same way, weirs affect water velocity/discharge and may also influence the survival of the infective larvae
Weirs position	Weirs position relates to the same hypothesis than above

At the individual level, biotic variables were host body length (mm), expected heterozygosity ( $H_e$ ) computed from a set of microsatellite data (Mazé-Guilmo, 2016; Rey et al., 2015) and the identity of the main host species of *T. polycolpus* on which we recorded infestations (i.e., either *L. burdigalensis* or *L. leuciscus*, Table 1). These variables measured at the individual level were then averaged at the site level before being included in the SDM.

At the site level, we first extracted biotic variables related to the fish community composition using the AFB database (Poulet 2007, Blanchet et al. 2014). This database contains fish occurrence data from yearly fish sampling in many sites evenly distributed across France. For each sampling site we drew a list of fish species occurrence over the period 2009-2011, from which we summed the total number of fish species (“fish species richness”) and the total number of potential host species for *T. polycolpus* (“number of host species”, Table 1) defined according to Lootvoet et al. (2013).

We then extracted abiotic variables belonging to four main categories namely climatic variables, hydrographic variables, landscape variables and habitat fragmentation variables (Table 1). Climatic and hydrographic variables at each site were extracted from the Hydrographique Théorique databases (Pella et al. 2012). Data on water temperature was not directly available at this spatial scale. We therefore used four different proxies for water climatic conditions (Le Moine 2002). Three of these proxies were air temperature extrapolated at the site level, averaged at the annual level, for the months of January and July respectively (gathered using the RHT database). The last proxy was a mean annual water temperature statistically inferred (and extrapolated for each site) from annual air temperature (Chevalier et al. unpublished). Regarding hydrographic variables, we extracted the following variables from the RHT database at each site: altitude (m), river slope (%), riverine distance from the source (km), area of the upstream watershed (km<sup>2</sup>), Strahler index (i.e., the hierarchical level in the hydrographic network of each river), river width (m), river depth (m), average size of bed sediment (mm), minimum average monthly river flow (m<sup>3</sup>.s<sup>-1</sup>) and average natural river flow (m<sup>3</sup>.s<sup>-1</sup>). We extracted Landscape variables from the Corine land cover 2012 database (<https://www.statistiques.developpement-durable.gouv.fr/corine-land-cover-0>) as percentages of main land cover types within a 10 km radius around each site using the first level of land cover description (Agricultural land, Artificialized lands, Forest and semi-natural lands and Wet lands grouped with Water bodies). We then synthesized climatic, hydrographic and landscape variables using Principal Component Analysis (one PCA per category) in order to avoid both collinearity and over-parametrization in subsequent linear models (Figure S1 and Table S1). We conserved the first PCA axis (69% of the total variance, see Fig. S1-a and Table S1) for climatic variables (hereafter called “Temperature”). For the hydrographic variables, we conserved the first two axes (58 % and 15 % of the variance respectively, Fig. S1-B and Table S1) that we named “Hydrology” and

“Topology” respectively. For landscape variables we conserved the first two axes (50 % and 28% of the variance respectively, Fig. S1-c and Table S1) that we arbitrarily named “Landscape 1” and “Landscape 2”. Finally, we extracted fragmentation variables from the “Référentiel des Obstacles à l'Écoulement” database (ROE, <http://www.onema.fr/REFERENTIEL-DES-OBSTACLES-A-L>) to quantify habitat fragmentation by weirs (number of obstacles being < 3m height) or dams (number of obstacles being > 3m height) within a 10 km radius around each sampling site. We further informed whether a weir was present directly upstream, downstream or both upstream and downstream for each sampling site (within a 2km radius) to generate a more direct estimate of fragmentation at the sampling site level (categorical variable: upstream, downstream, both upstream and downstream or no weir at all). All abiotic characteristics were extracted using a Geographic Information System (QGIS Development Team 2008) and PCA were run using the R package *ade4* (Dray et al. 2007).

**Statistical analyses.** We developed the SDM for *T. polycolpus* distribution at the site level by relying on a spatially explicit model as we detected a strong signal of spatial autocorrelation for *T. polycolpus* prevalence (Moran test, p-value<0.001). We thus used the R package *spaMM* (Rousset and Ferdy 2014) and its function *corrHLfit* to specify geographic coordinates of each sampling site into continuous random terms.

We first built a full model fitted on the dataset covering the current range of *T. polycolpus* (i.e., parasitized areas, Fig. 1, n=128) to test the significance of each variable on parasite prevalence. It included all explicative variables (Table 1) as well as the quadratic term for body length because preliminary visual inspection suggested a non-linear relationship between prevalence and body length. A binomial error terms distribution with a logit link function was fitted to the prevalence data. Then, likelihood ratio tests based on PQL/L fits were performed by contrasting the full

model to a model excluding the explicative variable of interest, so as to generate a  $\chi^2$  statistic and its associated p-value for each explicative variable.

Secondly, we parametrized the model on a restricted dataset including 75% of the sampling sites covering the current range of *T. polycolpus* (i.e., parasitized areas, Fig. 1). The parameterized model allowed predicting *T. polycolpus* prevalence both on the remaining 25% of the dataset covering the current range of *T. polycolpus* and in sites out of the range limit of *T. polycolpus* (Lineage III-unparasitized, Fig. 1). This procedure was reiterated 100 times to increase prediction accuracy. We finally compared predicted prevalence between lineages using contrast tests using the R package *lsmeans* and a binomial generalized linear mixed model (Lenth 2016).

### **Testing the role of the genetic background.**

We aimed to test whether genomic background of hosts could explain prevalence pattern of *T. polycolpus* and hence its range limitation in the lineage III-unparasitized (Fig. 1). In this section prevalence is measured at the individual level and hence refers to either parasitized or not parasitized (i.e., total number of host equals one). We investigated whether some genomic structure underlined the absence of *T. polycolpus* in the lineage III-unparasitized. Secondly we investigated genomic variation associated to *T. polycolpus* prevalence at the individual level.

**Data collection.** DNA was extracted from the piece of pelvic fin sampled on 96 individual hosts belonging to the species *L. leuciscus* so as to reflect the variability of parasite prevalence and the different lineages (lineage II, lineage III-parasitized and lineage III-unparasitized). Overall, 28 individuals from the lineage II were sampled at four sampling sites belonging to the Rhone River Basin (Fig. 2), and we selected half of them with parasites and the other half unparasitized. We then sampled 28 individuals from the lineage III-parasitized at five sampling sites belonging to



the Normandy River Basin (Fig. 2), and we selected half of them with parasites and the other half unparasitized. Finally 40 individuals from the lineage III-unparasitized were sampled at five sampling sites localized in the Rhin, Seine and Meuse Rivers Basins (Fig. 2).

We used a paired-end restriction site-associated DNA sequencing approach (RAD-seq, Baird et al., 2008) to sequence a reduced representation of the host genome into reads of 145 bp. The sequencing was performed on one lane of an Illumina HiSeq3000 containing the 96 independent banks (one per individual), with a minimum of 200ng of DNA per bank and using the chemistry V4 from kits TruSeq (2x125 pb).

***Data filtering and catalogue construction.*** We relied on the *Stacks* (version 2.3.4) pipeline to both filter the raw dataset and identify SNPs genotyped at the individual level (Catchen et al. 2013). First, we quality filtered reads with unidentified bases and/or with low quality score (below 20 phred33 quality score) using *process\_radtags*. PCR duplicates were also removed using *clone\_filter*. To build our catalogue of loci, alleles and ultimately SNPs from the filtered reads, we run the *denovo\_map.pl* using as parameters value:  $m=3$ ,  $M=2$ ,  $n=4$  (i.e., set of parameters maximizing the number of SNPs having high likelihood according to preliminary tests). Finally, *populations* was used to select loci with a minimum allele frequency above 3%, with a likelihood over -20, with a minimum coverage depth greater than 3, occurring in the three populations and in at least 40% of the individuals for each population. We also kept only one SNP per locus to limit linkage disequilibrium between SNPs. When applied to the whole database (i.e., 95 individuals because one individual was discarded as the sequencing quality was poor), this parameter set resulted in a catalog of 14 255 SNPs. A final filtering excluded from further analyzes SNPs with more than 30% of missing values using the library *poppr* in R (Kamvar et al. 2014). This resulted in a final dataset of 2 543 SNPs on 95 individuals.

**Population genetic structure.** We investigated population structure and whether this structure underlines prevalence patterns. A Discriminant Analysis of Principal Component (DAPC) was conducted using *adegenet* (Jombart 2008) in R to first assess genetic structure without *a priori* with the *find.cluster* function. All principal components were kept and the best number of clusters was selected using the Bayesian Information Criterion (BIC). A second DAPC was then run using 30 principal components (see Fig. S2) and the  $k$  number of clusters selected previously. The  $k-1$  discriminant functions (i.e., individuals' scores extracted along the  $k-1$  discriminant axes) were then extracted. We finally constructed a generalized linear model (GLM) with a quasibinomial error term linking parasite prevalence at the individual level to the  $k-1$  discriminant functions of the DAPC to test their significance. In this model, we also controlled for host body length and individual level of genomic diversity (measured as expected heterozygosity,  $H_e$ , calculated from allelic frequencies over all 2 543 SNPs using the *adegenet* R package). These later variables were included as continuous explicative variables (together with their respective quadratic terms) since they have been shown to be (sometimes non-linearly) related to individual parasite prevalence in this host-parasite system (Blanchet et al. 2009a).

**Genome wide association study.** We used a genome wide association study (GWAS) approach to test whether prevalence at the individual level was significantly related to some SNPs. We relied on GLMs with a quasibinomial error term linking parasite prevalence at the individual level (parasitized or unparasitized, response variable) to the genotype (coded as a categorical factor: homozygote for allele  $a$  or  $b$  and heterozygote  $ab$ ) of each SNP. One model per SNP marker (i.e., 2 543) was run in which we controlled for population structure using the  $k-1$  discriminant functions previously established. We also included host body length and  $H_e$  as continuous co-variable with their respective quadratic term. Type I Anovas with a Fisher statistic were used to

extract the p-value associated to each SNP marker. P-values were then corrected for multiple comparisons using a false discovery rate approach (FDR) and using the *fdrtool* R package (Strimmer 2008).

The R software (version 3.5.3, R Core Team 2017) was used to run statistical analyses when no other software is mentioned.

## Results

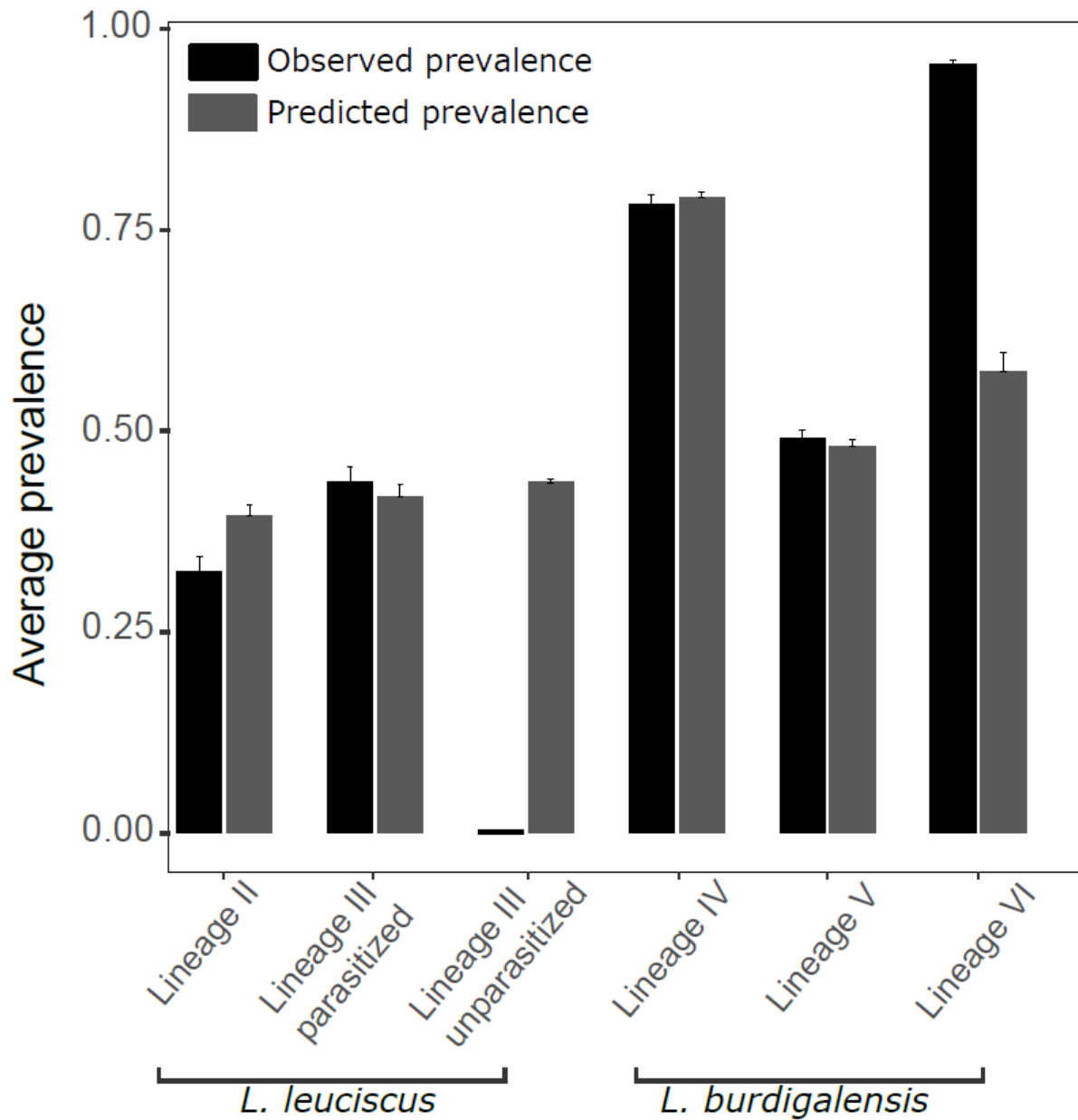
### Testing the environmental suitability hypothesis.

In the parasitized area, *T. polycolpus* prevalence was significantly related to host body length, host species identity, temperature, hydrology, the first synthetic variable of landscape composition and the position of weirs around the sampling site (Table 2). These relationships suggested that *T. polycolpus* prevalence was higher in *L. burdigalensis* than in *L. leuciscus* and that *T. polycolpus* prevalence was higher for hosts with intermediate body length than for smallest and largest hosts (Table S2). Furthermore, *T. polycolpus* prevalence is higher in cold sites localized in smaller rivers and preferentially associated to agricultural lands rather than forest or semi-natural lands (Table 2, Table S2 and Fig. S1). Finally, fragmentation by weirs appeared to have a significant impact on *T. polycolpus* prevalence, with highest prevalence expected in site in which a weir is present downstream (Table 2 and Table S2).

**Table 2:** Likelihood ratio tests [Chi-squared statistics ( $\chi^2$ ), degrees of freedom (d.f.) and p-values] testing the significance of each explicative variables of the spatial linear model used to explain and to predict the probability of *T. polycolpus* prevalence in French rivers.

	<b>X<sup>2</sup> statistics</b>	<b>d.f.</b>	<b>p-values</b>
Expected Heterozygosity	0.128	1	0.721
<b>Dace body length</b>	<b>14.704</b>	<b>1</b>	<b>&lt;0.001</b>
<b>(Dace body length)<sup>2</sup></b>	<b>5.980</b>	<b>1</b>	<b>0.014</b>
<b>Species identity</b>	<b>22.686</b>	<b>1</b>	<b>&lt;0.001</b>
Number of hosts species	-0.042	1	1.000
Fish species Richness	3.267	1	0.071
<b>Temperature</b>	<b>18.205</b>	<b>1</b>	<b>&lt;0.001</b>
<b>Hydrology</b>	<b>4.527</b>	<b>1</b>	<b>0.033</b>
Topology	2.869	1	0.090
<b>Landscape 1</b>	<b>11.106</b>	<b>1</b>	<b>0.001</b>
Landscape 2	0.633	1	0.426
Dams density	0.659	1	0.417
Weirs density	0.743	1	0.389
<b>Weirs position</b>	<b>9.817</b>	<b>3</b>	<b>0.020</b>

In the five host lineages localized in the current range of *T. polycolpus*, prevalence predicted from our model were very close to the observed prevalence (except in the lineage VI), which indicated an overall good predictive power of the model (Figure 3). In the unparasitized area of the lineage III (i.e., Lineage III-unparasitized, Fig. 3), the predicted parasite prevalence was high and similar to prevalence values observed and predicted in areas and lineages in which *T. polycolpus* is actually present (i.e., Lineage II, Lineage III-parasitized, Lineage IV, Lineage V, Lineage VI, Figure 3). More specifically, in Lineage III-unparasitized, the predicted parasite prevalence was not significantly different from the prevalence predicted in the Lineage III-parasitized (Contrast test,  $p=0.875$ , Fig. S3) and significantly higher than the prevalence predicted in the lineage II (Contrast test,  $p=0.019$ , Fig. S3).



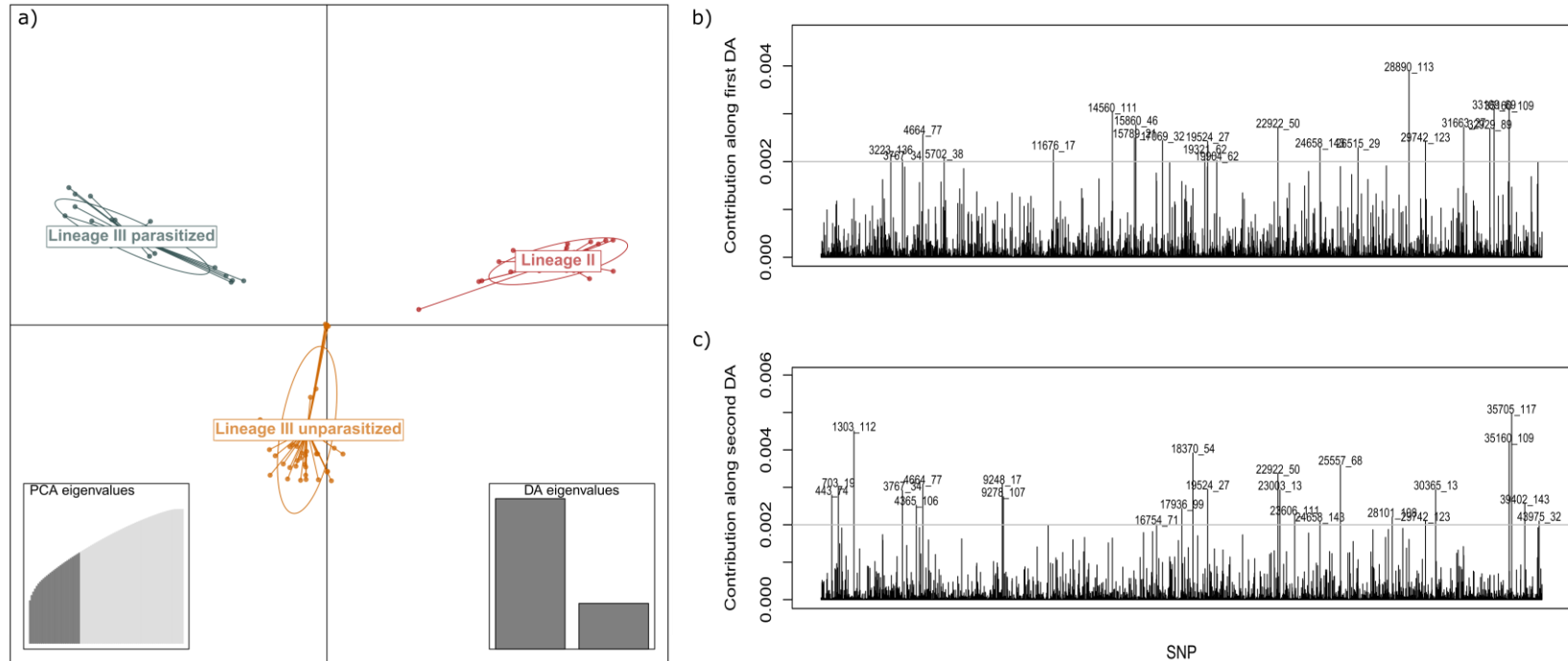
**Figure 3:** Barplot showing the average prevalence for each dace lineage. Black bars correspond to data measured in the wild. Grey bars correspond to values predicted by the SDM model (averaged over 100 iterations). Errors-bars are standard deviation from the mean.

**Testing the role of the genetic background.**

*Population genomic structure and GWAS.* The clustering analysis based on 2 543 SNPs revealed that the best clustering was achieved for  $k = 3$  (Fig. S4). The subsequent DAPC realized with 3

groups and based on 30 PCA components summarized 67.3% of the total variance along the two axes (Figure 4). The first axis discriminated the two dace lineages (i.e., lineage II and lineage III), as well as individuals of the lineage III belonging to the parasitized area from individuals of the same lineage but belonging to the unparasitized area (Fig. 4 and Fig. S5). Interestingly the second axis discriminated individuals localized in the parasitized areas from individuals localized in the unparasitized area (lineages II and lineage III-parasitized vs. lineage III-unparasitized, Fig. 4). This suggests that the genomic background of dace localized in the unparasitized area is different from both the genomic background of dace belonging to the lineage II and of dace belonging to the lineage III but localized in the parasitized area. Furthermore, we found that both the first and second discriminant function had a significant effect on individual prevalence ( $F=9.750$ ;  $df=88,1$ ;  $p=0.002$  and  $F=77.830$ ;  $df=88,1$ ;  $p<0.001$  respectively).

Finally, when controlling for population structure, the GWAS revealed 90 SNPs significantly associated to *T. polycolpus* prevalence measured at the individual level for *L. leuciscus* (at a FDR threshold of 5%, Table S3).



**Figure 4:** a) Scatter plot from a DAPC based on genomic data (2 543 SNP markers) and showing population differentiation between the three genomic clusters previously inferred. The 30 first axes of a PCA were retained and two axes were retained for the DAPC as eigenvalues plots show. b) and c) Contribution of each SNP to the multivariate DAPC analysis. These plots shows the extent to which each allele is driving genomic population differentiation along the first b) and c) second axis of the DAPC. Each vertical line corresponds to the contribution of a unique SNP. SNP whose value is above the grey horizontal line are SNP contributing to more than 2‰ to the genomic differentiation between clusters. The SNPs are identified as follows: loci number\_position of the SNP.



## Discussion

In this study we aimed to tease apart the factors limiting the spatial distribution of an emerging parasite in its invasive range. Combining Species Distribution Model and population genomics approaches we suggest that host characteristics, notably the genomic background of individuals living in the unparasitized area, are more likely than environmental features to explain the spatial distribution of *T. polycolpus*, and hence why it has not invaded some areas.

Using a Species Distribution Modelling approach we first highlighted a complex role of environmental factors involving multifactorial drivers of *T. polycolpus* spatial distribution. Climate is one of the most studied driver in parasite distribution (Kutz et al. 2005, Barrett et al. 2012, Caminade et al. 2019). We consistently found a significant effect of temperature. In particular, colder sites were associated with higher prevalence, which matches what we know from the biology of this parasite species and also the host habitat preferences (see Mazé-Guilmo et al. 2016, Franke et al. 2019). However, we also found significant effects of various environmental drivers such as host related factors (host body length and host species), river topography, surrounding landscape composition and fragmentation. Overall, these later findings corroborate previous findings on this species. For instance, as in Blanchet et al. (2009a), we found that intermediate body length favor parasite attachment, probably because of a trade-off between the surface available to the parasite to anchor and age-related ability of hosts to resist/tolerate the parasite (see also Cardon et al. 2011). Similarly, we found that sites localized in small rivers with low stream velocity (negative effect of the “Hydrology” variable, see Fig. S1) and surrounded by weirs were associated to higher prevalence. These type of effects were expected as the free-living infective stage of this parasite requires low water velocity to develop (Loot et al., 2004). More generally, such a multifactorial-environmental determinism of *T.*

*polycolpus* distribution strongly supports the growing view that integrative models accounting for multiple environmental drivers, including biological interactions, are particularly powerful for predicting parasite spatial distribution (Lafferty 2009, Giannini et al. 2013, Marcogliese 2016).

Based on Species Distribution Model predictions, we further showed that environmental suitability was unlikely to explain *T. polycolpus* range limitation. Prevalence predicted from environmental conditions encountered in the area in which the parasite is currently absent were similar to that reached in areas in which the parasite is recurrently observed, hence suggesting suitable habitats for *T. polycolpus*. It could be argued that we may have missed some key environmental variables actually predicting the occurrence of the parasite. Nonetheless, our species distribution model reached very accurate predictions for all other host lineages highlighting a very good predictive power. It is noteworthy that predicted parasite prevalence was somewhat underestimated for the host lineage VI. Interestingly, the lineage VI is the most phylogenetically differentiated lineage of *L. burdigalensis* and is actually considered as a separate endemic species by some authors (*L. bearnensis*, Kottelat & Freyhof, 2007 but see Costedoat et al., 2014). Discrepancy between model predictions and observed prevalence in this phylogenetically divergent lineage and in the host lineage III-unparasitized suggest that other factors than environmental suitability, notably host genomic background, could determine parasite prevalence patterns and consequently their range limitation.

Using a population genomic approach based on 2 543 SNPs we were able to show that host genomic background is associated with *T. polycolpus* distribution, both at the population level and at the individual level. We first revealed an underlying host population structure in *L. leuciscus* that spatially matched the distribution of *T. polycolpus*. We showed that the lineage III proposed by Costedoat et al. (2014) actually splits into a parasitized “sub-lineage” localized in

the Normandy River Basin (Fig. 2) and unparasitized “sub-lineage” in the Rhin-Seine-Meuse River Basins. Previous studies on the genetic structure of the *Leuciscus* species complex were based on microsatellite, allozyme, and mitochondrial markers and covered a more extended spatial scale but with fewer sampled individuals per location (Costedoat et al. 2006, 2014). Here, the finer genetic structure that we were able to detect probably results from the higher resolution of the SNP panel we uncovered, as well as from the more intensive sampling of the *Leuciscus* species within lineages. Second, using a Genome Wide Association Study approach, we found that 90 SNPs were associated with parasite prevalence measured at the individual level, even when controlling for population structure. Both the genetic uniqueness of host populations localized in the unparasitized area and the association of SNPs with individual prevalence suggest that *T. polycolpus* range limitation is due to local host resistance, hence providing a strong support to the “host genomic background hypothesis.”

The emergence of resistance against *T. polycolpus* in northeastern river basins is an intriguing question and two main hypotheses can be formulated: one related to macro-evolutionary processes and the other one related to recent and rapid adaptation to this invasive parasite. The “macro-evolutionary” hypothesis assumes that resistance to *T. polycolpus* emerged during the differentiation of *Leuciscus* lineages over time. Costedoat et al. (2006) dated the differentiation of the lineage III back to the Pleistocene (~500 000 years ago), with the colonization of river basins of Northern France from Eastern Europe. The differentiation between the Lineage III-parasitized and Lineage III-unparasitized (and hence the emergence of resistance) may have consequently occurred during late Pleistocene, probably after the former Channel River has retracted hence causing a loss of connectivity between Normandy River Basin and Rhin-Seine-Meuse River Basins (which occurred ~10-12 000 years ago, Antoine et al. 2007, Hugueny et al. 2011, Dias et

al. 2014). Alternatively, the recent and rapid adaptation scenario assumes rapid adaptive changes in hosts since the introduction of *T. polycolpus* in France during the 1920's (Rey et al. 2015). Some examples of rapid evolution to disease resistance in wild host populations have been documented (Duffy and Sivars-Becker 2007, Bonneaud et al. 2011, Epstein et al. 2016), but these examples remain rare. In the *L. leuciscus*-*T. polycolpus* system, adaptive evolution toward resistance cannot be excluded since we found significant SNP-prevalence associations, but whether emergence of resistance is either due to drift or to adaptive evolution, and whether this occurred either far in the past or very recently are questions that would need to be fully tested.

Overall, our results demonstrated the importance of combining approaches for predicting the distribution range of co-evolving and interacting species. Species Distribution Modelling approaches focusing on parasite species have, so far, purely focused on environmental factors (Ostfeld et al. 2005, Caminade et al. 2019). However, our results suggest that predictions of future parasite distribution (notably northern range expansion as often predicted in climate change contexts, eg., Altizer et al. 2013, Carter 2018), when only based on environmental drivers should be carefully interpreted. Here, we indeed provided an empirical example in which the presence of a parasite would have been expected in an area where the parasite never actually achieved to parasitize local hosts. This suggests that ecological features alone are not sufficient to explain the range and potential spread of emerging parasites (Pérez-Rodríguez et al. 2013, Anderson 2017). Furthermore, the genomic-based approach allowed demonstrating that the host genomic background was likely a critical factor for explaining *T. polycolpus* distribution both at the population and the individual levels. Population genomic approaches allow accounting for evolutionary processes leading to parasite resistance such as local adaptation, which, from our results, appears to be a crucial process to take into account if improvements to predictions

regarding the potential spread of emerging pathogens are to be made. To conclude, we aimed to consider an extensive range of explicative factors related to both ecological and evolutionary dynamics of host-parasite interactions. We thus illustrated and argue the usefulness of integrative study to reveal the determinants of emerging pathogens at large spatial scales as well as to predict future spread.

## References

- Aguilar, J. R. et al. 2016. MHC-I provides both quantitative resistance and susceptibility to blood parasites in blue tits in the wild. - *J Avian Biol* 47: 669–677.
- Altizer, S. et al. 2013. Climate Change and Infectious Diseases: From Evidence to a Predictive Framework. - *Science* 341: 514–519.
- Anderson, R. P. 2017. When and how should biotic interactions be considered in models of species niches and distributions? - *Journal of Biogeography* 44: 8–17.
- Antoine, P. et al. 2007. Pleistocene fluvial terraces from northern France (Seine, Yonne, Somme): synthesis, and new results from interglacial deposits. - *Quaternary Science Reviews* 26: 2701–2723.
- Baird, N. A. et al. 2008. Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. - *PLoS ONE* 3: e3376.
- Barrett, T. M. et al. 2012. Host–parasite distributions under changing climate: *Tsuga heterophylla* and *Arceuthobium tsugense* in Alaska. - *Canadian Journal of Forest Research* 42: 642–656.
- Benavides, M. V. et al. 2016. Genomic Regions Associated with Sheep Resistance to Gastrointestinal Nematodes. - *Trends in Parasitology* 32: 470–480.
- Bernatchez, L. and Landry, C. 2003. MHC studies in non-model vertebrates: what have we learned about natural selection in 15 years? - *Journal of Evolutionary Biology* 16: 363–377.
- Blanchet, S. et al. 2009a. Evidence of parasite-mediated disruptive selection on genetic diversity in a wild fish population. - *Molecular Ecology* 18: 1112–1123.
- Blanchet, S. et al. 2009b. Why do parasitized hosts look different? Resolving the “chicken-egg” dilemma. - *Oecologia* 160: 37–47.
- Blanchet, S. et al. 2014. Regional vs local drivers of phylogenetic and species diversity in stream fish communities. - *Freshw Biol* 59: 450–462.
- Bonneaud, C. et al. 2011. Rapid evolution of disease resistance is accompanied by functional changes in gene expression in a wild bird. - *Proceedings of the National Academy of Sciences* 108: 7866–7871.

- Bozick, B. A. and Real, L. A. 2015. Integrating Parasites and Pathogens into the Study of Geographic Range Limits. - *The Quarterly Review of Biology* 90: 361–380.
- Byers, J. E. et al. 2019. What factors explain the geographical range of mammalian parasites? - *Proc. R. Soc. B* 286: 20190673.
- Caminade, C. et al. 2019. Impact of recent and future climate change on vector-borne diseases: Climate change and vector-borne diseases. - *Ann. N.Y. Acad. Sci.* 1436: 157–173.
- Cardon, M. et al. 2011. Host characteristics and environmental factors differentially drive the burden and pathogenicity of an ectoparasite: a multilevel causal analysis: Causalities in host-parasite interactions. - *Journal of Animal Ecology* 80: 657–667.
- Carter, N. 2018. Climate change, disease range shifts, and the future of the Africa lion. - *Conservation Biology* 32: 1207–1210.
- Catchen, J. et al. 2013. Stacks: an analysis tool set for population genomics. - *Molecular Ecology* 22: 3124–3140.
- Charbonnel, N. et al. 2019. Differential immune gene expression associated with contemporary range expansion of two invasive rodents in Senegal. - *bioRxiv* in press.
- Costedoat, C. et al. 2006. Heterogeneous colonization pattern of European Cyprinids, as highlighted by the dace complex (*Teleostei: Cyprinidae*). - *Molecular Phylogenetics and Evolution* 41: 127–148.
- Costedoat, C. et al. 2014. Révision de la taxonomie ichtyologique en métropole : validation moléculaire du guide de détermination et investigation des aires de répartition.
- Dias, M. S. et al. 2014. Global imprint of historical connectivity on freshwater fish biodiversity. - *Ecology Letters* 17: 1130–1140.
- Dray, S. et al. 2007. The ade4 package: implementing the duality diagram for ecologists. - *Journal of statistical software* 22: 1–20.
- Duffy, M. A. and Sivars-Becker, L. 2007. Rapid evolution and ecological host-parasite dynamics: Rapid evolution and disease dynamics. - *Ecology Letters* 10: 44–53.
- Epstein, B. et al. 2016. Rapid evolutionary response to a transmissible cancer in Tasmanian devils. - *Nature Communications* 7: 12684.
- Escobar, L. E. et al. 2017. Inferring the Ecological Niche of *Toxoplasma gondii* and *Bartonella spp.* in Wild Felids. - *Front. Vet. Sci.* 4: 172.
- Fourcade, Y. et al. 2014. Continental-scale patterns of pathogen prevalence: a case study on the corncrake. - *Evolutionary Applications* 7: 1043–1055.
- Franke, F. et al. 2019. Consequences of divergent temperature optima in a host–parasite system. - *Oikos* 128: 869–880.

- Gaston, K. J. 2009. Geographic range limits: achieving synthesis. - *Proceedings of the Royal Society B: Biological Sciences* 276: 1395–1406.
- Gehman, A.-L. M. et al. 2018. Host and parasite thermal ecology jointly determine the effect of climate warming on epidemic dynamics. - *Proceedings of the National Academy of Sciences* 115: 744–749.
- Giannini, T. C. et al. 2013. Improving species distribution models using biotic interactions: a case study of parasites, pollinators and plants. - *Ecography* 36: 649–656.
- Grosholz, E. D. 1994. The Effects of Host Genotype and Spatial Distribution on Trematode Parasitism in a Bivalve Population. - *Evolution* 48: 1514.
- Guisan, A. and Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. - *Ecology Letters* 8: 993–1009.
- Hugueny, B. et al. 2011. Habitat fragmentation and extinction rates within freshwater fish communities: a faunal relaxation approach: Extinction rates within freshwater fish communities. - *Global Ecology and Biogeography* 20: 449–463.
- Johnson, P. T. J. et al. 2008. Diversity and disease: community structure drives parasite transmission and host fitness. - *Ecology Letters* 11: 1017–1026.
- Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. - *Bioinformatics* 24: 1403–1405.
- Kamvar, Z. N. et al. 2014. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. - *PeerJ* 2: e281.
- Kottelat, M. and Freyhof, J. 2007. Handbook of European freshwater fishes. - Publications Kottelat.
- Kutz, S. . et al. 2005. Global warming is changing the dynamics of Arctic host-parasite systems. - *Proceedings of the Royal Society B: Biological Sciences* 272: 2571–2576.
- Lachish, S. et al. 2013. Spatial determinants of infection risk in a multi-species avian malaria system. - *Ecography* 36: 587–598.
- Lafferty, K. D. 2009. The ecology of climate change and infectious diseases. - *Ecology* 90: 888–900.
- Le Moine, P. 2002. Description de l'analyse des champs de surface sur la France par le système SAFRAN.
- Lenth, R. V. 2016. Least-squares means: the R package lsmeans. - *J Stat Softw* 69: 1–33.
- Loot, G. et al. 2004. The effects of the ectoparasite *Tracheliastes polycolpus* (Copepoda: Lernaepodidae) on the fins of rostrum dace (*Leuciscus leuciscus burdigalensis*). - *Parasitology Research* 94: 16–23.

Loot, G. et al. 2011. Determinants of life-history traits in a fish ectoparasite: a hierarchical analysis. - *Parasitology* 138: 848–857.

Lootvoet, A. et al. 2013. Patterns and processes of alternative host use in a generalist parasite: insights from a natural host-parasite interaction. - *Functional Ecology* 27: 1403–1414.

Marcogliese, D. J. 2016. The Distribution and Abundance of Parasites in Aquatic Ecosystems in a Changing Climate: More than Just Temperature. - *Integr. Comp. Biol.* 56: 611–619.

Mazé-Guilmo, E. 2016. Etude du potentiel de colonisation des parasites : une approche intégrative.

Mazé-Guilmo, E. et al. 2016. Local adaptation drives thermal tolerance among parasite populations: a common garden experiment. - *Proceedings of the Royal Society B: Biological Sciences* 283: 20160587.

Moor, H. et al. 2015. Predicting climate change effects on wetland ecosystem services using species distribution modeling and plant functional traits. - *Ambio* 44: 113–126.

Morgan, E. R. et al. 2009. Canine pulmonary angiostrongylosis: The influence of climate on parasite distribution. - *Parasitology International* 58: 406–410.

Narouei-Khandan, H. A. et al. 2016. Global climate suitability of citrus huanglongbing and its vector, the Asian citrus psyllid, using two correlative species distribution modeling approaches, with emphasis on the USA. - *European Journal of Plant Pathology* 144: 655–670.

Ostfeld, R. et al. 2005. Spatial epidemiology: an emerging (or re-emerging) discipline. - *Trends in Ecology & Evolution* 20: 328–336.

Paull, S. H. et al. 2015. How temperature shifts affect parasite production: testing the roles of thermal stress and acclimation. - *Functional Ecology* 29: 941–950.

Pella, H. et al. 2012. Le réseau hydrographique théorique (RHT) français et ses attributs environnementaux. - *Géomorphologie : relief, processus, environnement*: 317–336.

Pérez-Rodríguez, A. et al. 2013. Finding the appropriate variables to model the distribution of vector-borne parasites with different environmental preferences: climate is not enough. - *Glob Change Biol* 19: 3245–3253.

Piasecki, W. 1839. Life cycle of *Tracheliastes maculatus* Kollar, 1835 (*Copepoda*, *Siphonostomatoida*, *Lernaeopodidae*). - *Wiadomości Parazytologiczne*.

Pickles, R. S. A. et al. 2013. Predicting shifts in parasite distribution with climate change: a multitrophic level approach. - *Global Change Biology* 19: 2645–2654.

Poulet, N. 2007. Impact of weirs on fish communities in a piedmont stream. - *River Res. Applic.* 23: 1038–1047.

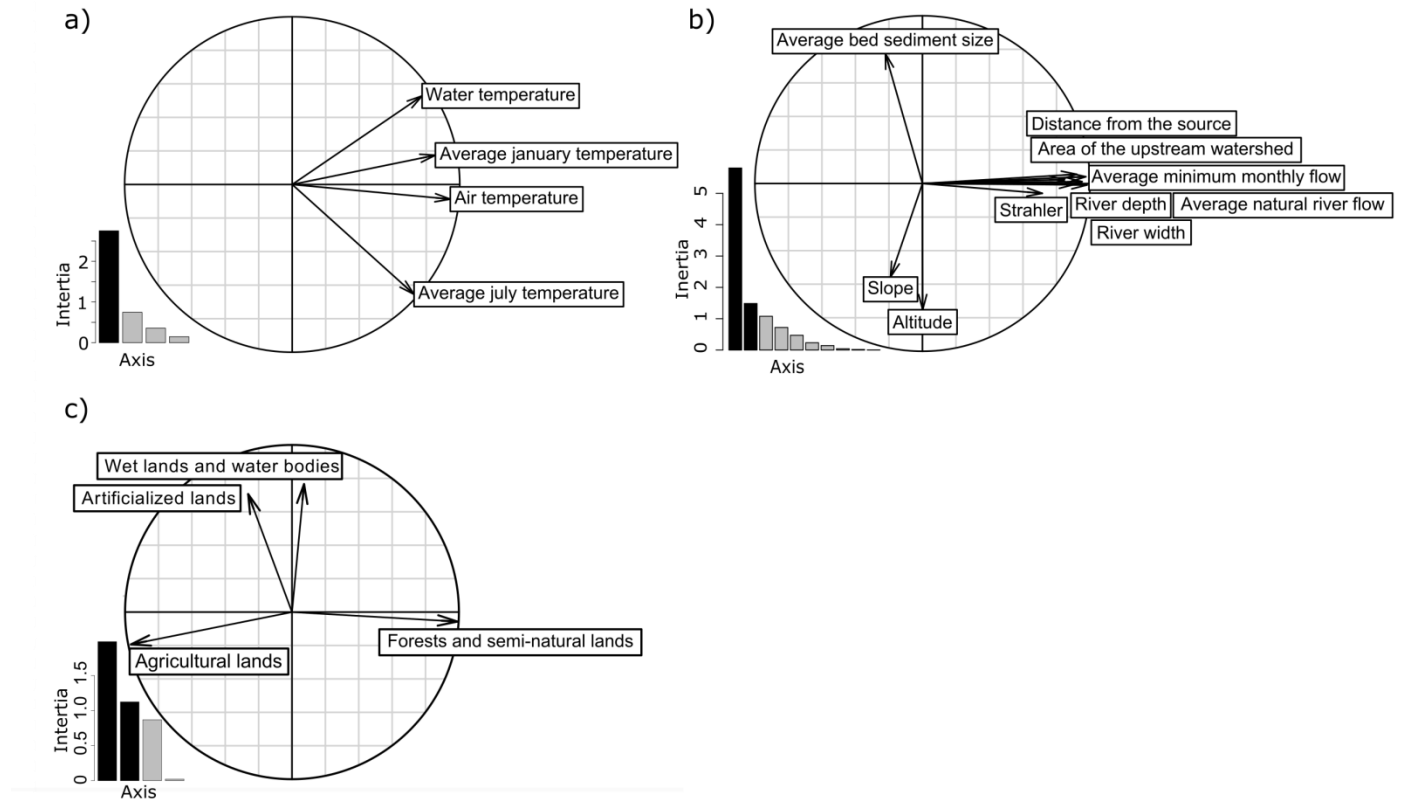
QGIS Development Team 2008. QGIS Geographic Information System.

R Core Team 2017. R: A language and environment for statistical computing.

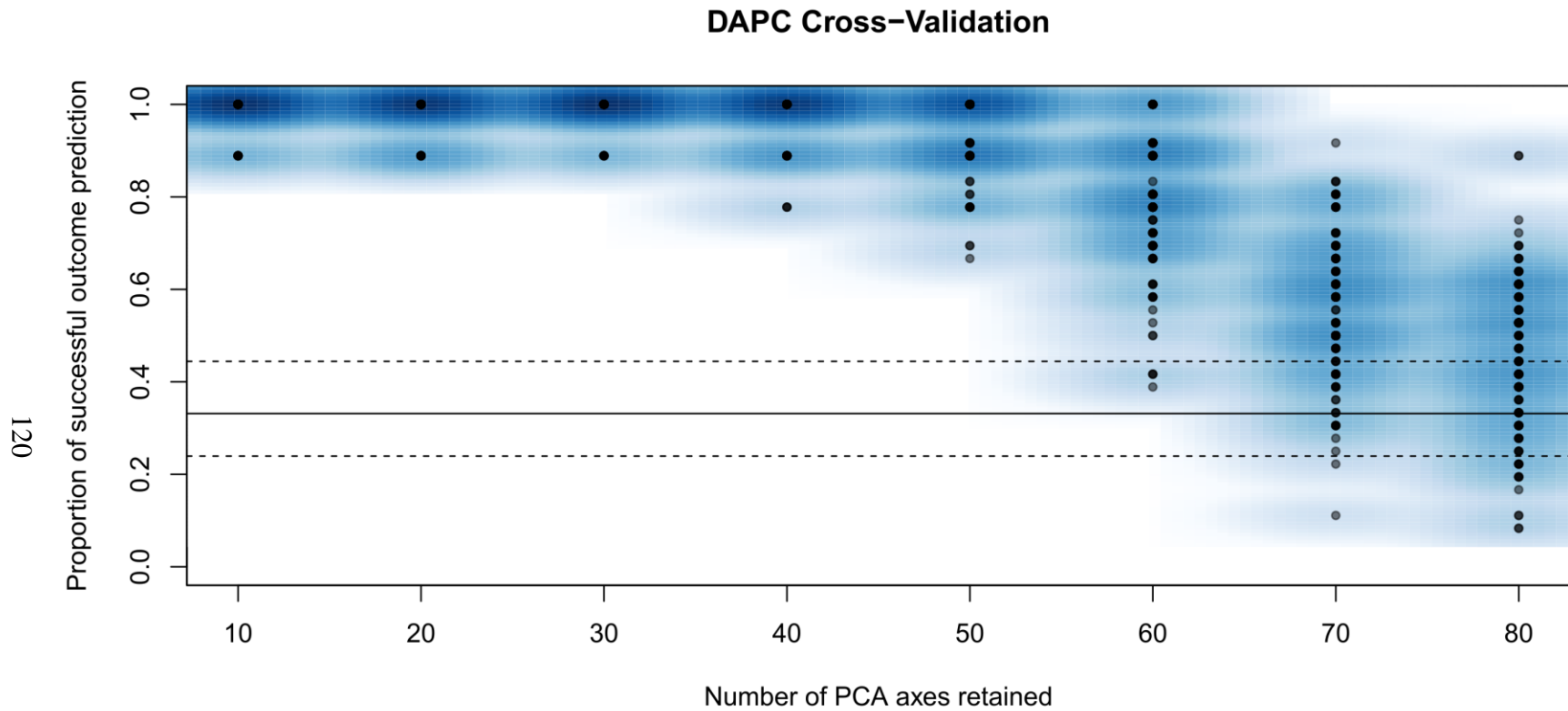


- Rey, O. et al. 2015. Elucidating the spatio-temporal dynamics of an emerging wildlife pathogen using approximate Bayesian computation. - *Molecular Ecology* 24: 5348–5363.
- Roche, B. et al. 2012. Linking community and disease ecology: the impact of biodiversity on pathogen transmission. - *Philosophical Transactions of the Royal Society B: Biological Sciences* 367: 2807–2813.
- Rousset, F. and Ferdy, J.-B. 2014. Testing environmental and genetic effects in the presence of spatial autocorrelation. - *Ecography* 37: 781–790.
- Schwabl, P. et al. 2017. Prediction and Prevention of Parasitic Diseases Using a Landscape Genomics Framework. - *Trends in Parasitology* 33: 264–275.
- Strimmer, K. 2008. fdrtool: a versatile R package for estimating local and tail area-based false discovery rates. - *Bioinformatics* 24: 1461–1462.
- Vajana, E. et al. 2018. Combining Landscape Genomics and Ecological Modelling to Investigate Local Adaptation of Indigenous Ugandan Cattle to East Coast Fever. - *Front. Genet.* 9: 385.
- Wegner, K. M. et al. 2003. Parasite Selection for Immunogenetic Optimality. - *Science* 301: 1343–1343.
- Wenzel, M. A. et al. 2015. Genome-wide association and genome partitioning reveal novel genomic regions underlying variation in gastrointestinal nematode burden in a wild bird. - *Molecular Ecology* 24: 4175–4192.
- Wiens, J. J. et al. 2010. Niche conservatism as an emerging principle in ecology and conservation biology: Niche conservatism, ecology, and conservation. - *Ecology Letters* 13: 1310–1324.
- Wolinska, J. and King, K. C. 2009. Environment can alter selection in host–parasite interactions. - *Trends in Parasitology* 25: 236–244.
- Zhang, L. et al. 2015. Major histocompatibility complex alleles associated with parasite susceptibility in wild giant pandas. - *Heredity* 114: 85–93.
- Zueva, K. J. et al. 2018. Genomic signatures of parasite-driven natural selection in north European Atlantic salmon (*Salmo salar*). - *Marine Genomics* 39: 26–38.

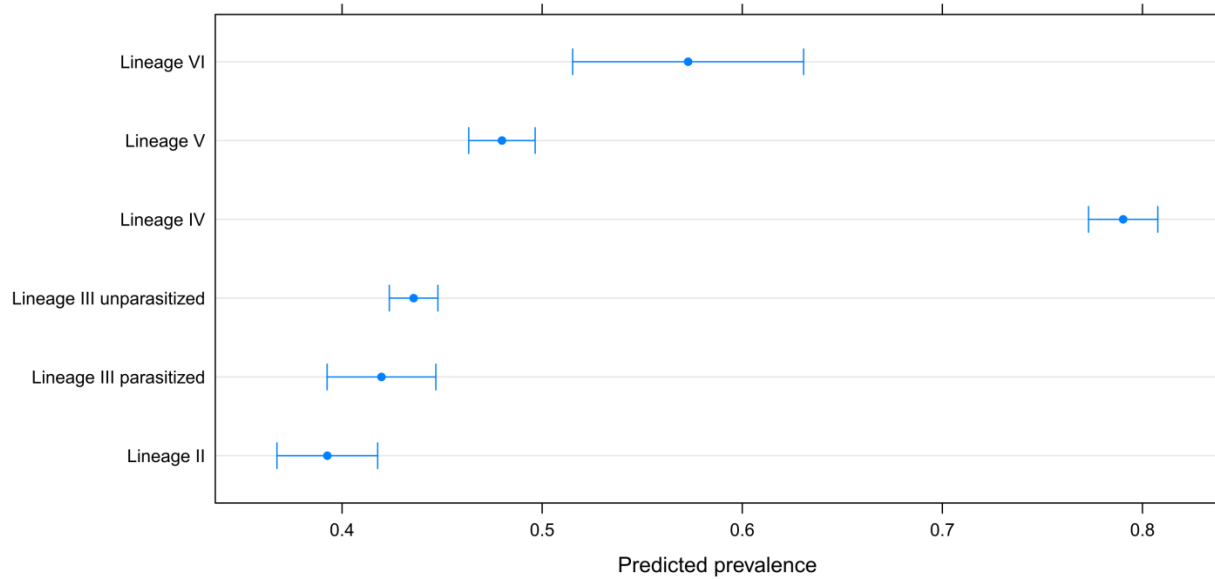
## Supplementary



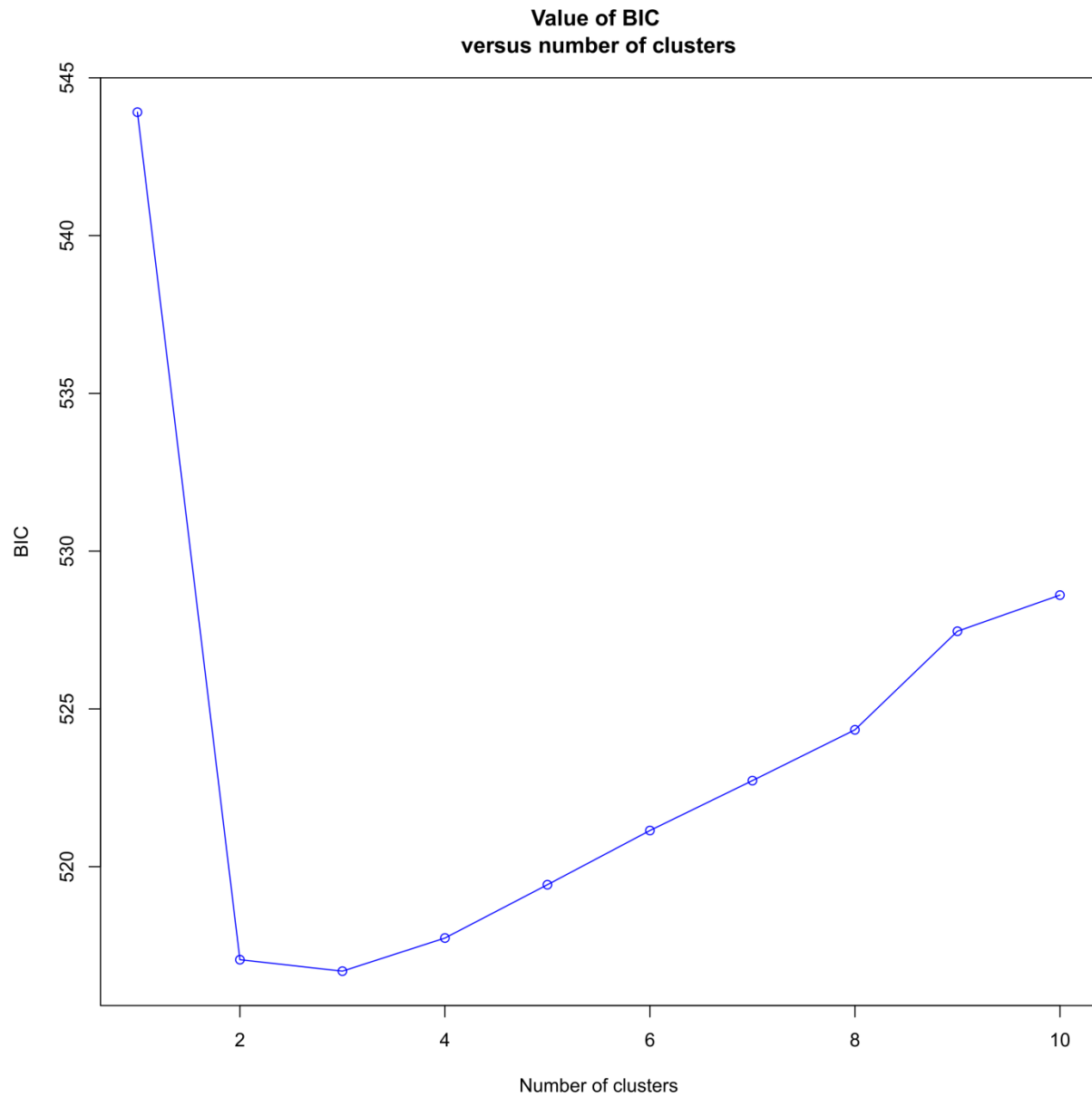
**Figure S1:** Principal Component Analyses for climatic variables (a), hydrographic variables (b) and landscape variables (c) respectively. Bar plots indicate how much of the variance is explained by each PCA axis and black bars indicate retained PCA axis as synthetic variables in subsequent analysis.



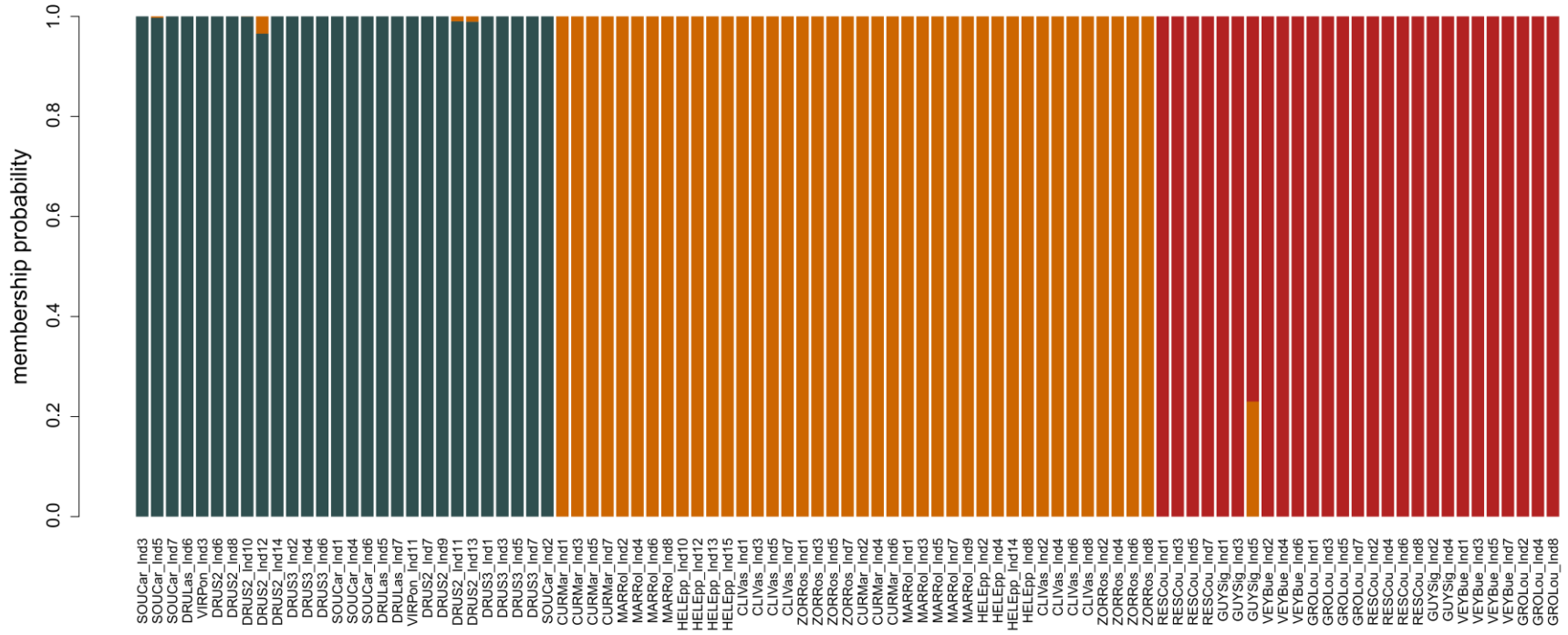
**Figure S2:** Graphic showing the result of the cross-validation procedure that aimed to identify the number of PCA components that should be retained for the DAPC (i.e., the number of components that maximize successful individual assignment to the  $k$  clusters based on 150 iterations). The cross-validation here suggests that 30 PCA components should be retained.



**Figure S3:** Forest plot showing the results of the contrast tests between predicted prevalence in each *Leuciscus* lineage. Dots refer to mean predicted prevalence and bar shows confident intervals. Non overlapping confident intervals mean that the difference between predicted prevalence in different lineages is significant.



**Figure S4:** Evolution of the Bayesian Information Criterion (BIC) according to the number of cluster retained in the Discriminant Analysis of Principal Component (DAPC). The lowest BIC (here  $k=3$ ) indicates the number of cluster that is the more likely to describe the input dataset (here allelic frequencies at 2543 Single Nucleotide Polymorphisms).



**Figure S5:** Graphic showing the individual assignment to the three clusters revealed by the DAPC. Each individual is represented by a vertical line that is partitioned into three colors that represent its estimated membership fractions in  $K=3$  clusters. Labels above the plot provide for each individual its sampling site (eg, SOUCar) and its own ID (eg., Ind3).

**Table S1:** Summary of the three Principal Component Analysis realized on climatic variables, hydrographic variables and landscape variables. Correlation of each raw variable with the two first components is displayed (PC1 and PC2) as well as the range of each raw variable (Range). Higher correlations are in bold and refer to the component associated to each variable.

Category	Variables	PC1	PC2	Range
Climatic	Average January temperature	<b>0.855</b>	NA	0–7 C°
	Average July temperature	<b>0.727</b>	NA	16–23 C°
	Water temperature	<b>0.775</b>	NA	8–15 C°
	Air temperature	<b>0.944</b>	NA	8-15 C°
Hydrographic	Altitude	0.002	<b>-0.752</b>	3–571 m
	Slope	-0.191	<b>-0.557</b>	0–48 ‰
	Distance from the source	<b>0.920</b>	0.055	3–394 km
	Area of the upstream watershed	<b>0.975</b>	0.041	7–20 740 km <sup>2</sup>
	Strahler index	<b>0.715</b>	-0.059	1–6
	River width	<b>0.982</b>	-0.008	1–153 m
	River depth	<b>0.908</b>	-0.005	0–2 m
	Average size of the bed sediments	-0.221	<b>0.774</b>	7–3
	Minimum average monthly river flow	<b>0.848</b>	0.023	3 436 m <sup>3</sup> /s
Average natural river flow	<b>0.953</b>	0.011	0–511 m <sup>3</sup> /s	
Landscape	Artificialized lands	-0.262	<b>0.706</b>	0%–25%
	Agricultural lands	<b>-0.964</b>	-0.193	0%–96%
	Forest and semi-natural lands	<b>0.991</b>	-0.057	0%–90%
	Wet lands and water bodies	0.072	<b>0.765</b>	0%–17%

**Table S2:** Table summarizing coefficient estimates, standard errors (SE) and associated t-value for the SDM model adjusted on the dataset covering the current range of *T. polycolpus* (n=128 sites).

	Estimate	Cond. SE	t-value
(Intercept)	2.565	2.23	-1.149
Hexp	0.725	1.35	0.536
Dace size	0.056	0.018	3.031
(Dace size) <sup>2</sup>	-1.09E-04	4.74E-05	-2.307
Species Richness	-0.114	0.068	-1.678
Number of hosts	-0.044	0.263	-0.167
Temperature	-0.677	0.171	-3.954
Hydrology	-0.26	0.156	-1.674
Topology	0.315	0.204	1.543
Landscape 1	-0.523	0.164	-3.189
Landscape 2	-0.181	0.237	-0.763
Dams	-0.013	0.016	-0.808
Weirs (number)	0.011	0.015	0.739
species ( <i>L. leuciscus</i> )	-3.143	0.689	-4.56
weir (none)	-0.889	0.641	-1.386
weirs (upstream)	-1.562	0.648	-2.409
weirs (upstream and downstream)	-1.621	0.672	-2.414



**Table S3:** Table synthesizing information for SNPs showing a significant association ( $\alpha = 0.05$ ) with parasite prevalence measured at the individual scale (GWAS approach). We displayed for each SNP, its name (number of the loci\_position of the SNP) and the outputs from an analysis of deviance: F statistic, degree of freedom, raw p-value and adjusted p-value using a false discovery rate approach (FDR).

<b>Loci</b>	<b>F value</b>	<b>Df</b>	<b>pvalue</b>	<b>FDR</b>
<b>30365_13</b>	18.07E8	61,2	<0.001	<0.001
<b>31247_45</b>	39.792	61,2	<0.001	<0.001
<b>20735_106</b>	24.223	59,2	<0.001	<0.001
<b>6830_17</b>	33.182	60,1	<0.001	<0.001
<b>9706_73</b>	17.988	63,2	<0.001	<0.001
<b>9473_86</b>	30.338	60,1	<0.001	<0.001
<b>14042_124</b>	16.998	60,2	<0.001	<0.001
<b>15789_21</b>	16.215	61,2	<0.001	0.001
<b>23160_22</b>	15.267	59,2	<0.001	0.001
<b>14397_60</b>	15.123	59,2	<0.001	0.001
<b>3163_102</b>	14.005	63,2	<0.001	0.002
<b>31547_125</b>	13.688	66,2	<0.001	0.002
<b>18121_75</b>	13.444	68,2	<0.001	0.002
<b>12378_47</b>	13.293	65,2	<0.001	0.002
<b>3881_89</b>	13.296	59,2	<0.001	0.002
<b>18955_50</b>	13.033	65,2	<0.001	0.002
<b>451_130</b>	12.841	59,2	<0.001	0.003
<b>32965_13</b>	12.119	60,2	<0.001	0.004

<b>19792_54</b>	11.949	65,2	<0.001	0.004
<b>6084_110</b>	11.866	60,2	<0.001	0.004
<b>22427_103</b>	11.781	62,2	<0.001	0.004
<b>15517_29</b>	11.734	60,2	<0.001	0.005
<b>2126_86</b>	11.676	59,2	<0.001	0.005
<b>20777_89</b>	11.505	59,2	<0.001	0.005
<b>9391_96</b>	10.832	71,2	<0.001	0.006
<b>26600_107</b>	17.674	67,1	<0.001	0.006
<b>26536_32</b>	10.694	62,2	<0.001	0.007
<b>23585_10</b>	10.423	63,2	<0.001	0.008
<b>14791_89</b>	16.833	61,1	<0.001	0.008
<b>6225_83</b>	10.435	60,2	<0.001	0.008
<b>8863_98</b>	10.217	65,2	<0.001	0.009
<b>567_75</b>	10.281	60,2	<0.001	0.009
<b>32434_98</b>	10.196	59,2	<0.001	0.009
<b>13572_22</b>	10.083	62,2	<0.001	0.01
<b>25567_94</b>	10.047	63,2	<0.001	0.01
<b>4249_123</b>	9.975	64,2	<0.001	0.01
<b>33337_20</b>	9.830	66,2	<0.001	0.01
<b>41022_86</b>	9.595	65,2	<0.001	0.012
<b>31207_90</b>	9.617	60,2	<0.001	0.012
<b>18497_117</b>	9.490	64,2	<0.001	0.012

<b>1055_61</b>	9.440	63,2	<0.001	0.013
<b>10039_82</b>	9.232	67,2	<0.001	0.014
<b>25585_44</b>	9.324	62,2	<0.001	0.014
<b>17689_63</b>	9.328	61,2	<0.001	0.014
<b>16758_77</b>	9.109	65,2	<0.001	0.015
<b>41153_8</b>	8.982	63,2	<0.001	0.016
<b>28491_38</b>	8.500	66,2	0.001	0.021
<b>28298_104</b>	8.503	64,2	0.001	0.021
<b>2909_66</b>	8.524	61,2	0.001	0.021
<b>13326_67</b>	13.195	60,1	0.001	0.022
<b>12878_52</b>	8.421	60,2	0.001	0.023
<b>17321_81</b>	8.388	60,2	0.001	0.023
<b>644_69</b>	8.332	62,2	0.001	0.023
<b>25467_120</b>	13.003	61,1	0.001	0.024
<b>32865_92</b>	8.332	59,2	0.001	0.024
<b>11384_48</b>	8.242	60,2	0.001	0.025
<b>13842_7</b>	12.781	61,1	0.001	0.025
<b>20915_81</b>	8.190	61,2	0.001	0.025
<b>13913_80</b>	12.550	61,1	0.001	0.026
<b>6311_24</b>	7.976	63,2	0.001	0.027
<b>11820_144</b>	7.907	63,2	0.001	0.028
<b>19016_53</b>	12.279	61,1	0.001	0.028

<b>7497_72</b>	7.903	61,2	0.001	0.029
<b>34614_12</b>	12.074	66,1	0.001	0.029
<b>33082_120</b>	7.768	62,2	0.001	0.03
<b>16630_129</b>	7.777	61,2	0.001	0.03
<b>1972_142</b>	7.726	63,2	0.001	0.03
<b>19731_124</b>	7.768	60,2	0.001	0.03
<b>33374_128</b>	7.625	67,2	0.001	0.031
<b>725_116</b>	7.667	59,2	0.001	0.032
<b>9118_144</b>	7.647	59,2	0.001	0.033
<b>25353_20</b>	7.579	59,2	0.001	0.034
<b>9271_55</b>	7.524	62,2	0.001	0.034
<b>4594_43</b>	7.367	60,2	0.001	0.037
<b>28712_59</b>	7.350	59,2	0.001	0.038
<b>25438_128</b>	7.238	63,2	0.001	0.039
<b>15286_61</b>	7.236	63,2	0.001	0.039
<b>29363_50</b>	7.188	63,2	0.002	0.04
<b>1276_70</b>	7.216	59,2	0.002	0.04
<b>5853_56</b>	7.123	65,2	0.002	0.041
<b>12705_36</b>	7.116	65,2	0.002	0.041
<b>2752_120</b>	7.138	60,2	0.002	0.042
<b>3737_143</b>	6.991	71,2	0.002	0.042
<b>10393_72</b>	7.044	59,2	0.002	0.044

<b>33235_100</b>	6.979	64,2	0.002	0.044
<b>29939_90</b>	10.594	64,1	0.002	0.044
<b>16343_28</b>	6.943	61,2	0.002	0.045
<b>11938_12</b>	6.748	65,2	0.002	0.049
<b>27398_144</b>	10.239	61,1	0.002	0.049
<b>26335_72</b>	6.748	64,2	0.002	0.05

---

## **Chapitre III :**

# **Rôle de la plasticité dans la capacité d'un parasite à exploiter plusieurs espèces hôtes**

## Résumé

La capacité d'un parasite à exploiter et à s'adapter à de nouvelles espèces hôtes conditionne fortement sa dynamique écologique et évolutive mais aussi ses effets pathogènes sur les communautés hôtes. Bien que chez de nombreux parasites il a été fait état de transfert vers de nouvelles espèces hôtes, les processus permettant aux parasites de s'adapter à ces nouveaux hôtes restent encore à ce jour méconnus. Ici, nous avons testé l'hypothèse selon laquelle la plasticité serait un mécanisme qui peut promouvoir le transfert d'hôte chez les parasites. Nous avons considéré le cas d'un parasite émergent (*Tracheliastes polycolpus*) capable d'infecter plusieurs espèces de poissons d'eau douce et nous avons réalisé une étude transcriptomique comparative entre les parasites infectant l'espèce hôte principale et ceux infectant deux espèces hôtes alternatives. Nous avons trouvé 120 gènes codant pour des protéines qui étaient exprimés de manière différentielle (DEG) entre des parasites infectant différentes espèces hôtes. Le plus grand nombre de DEG a été trouvé pour les comparaisons entre des parasites utilisant l'espèce hôte principale par rapport à des parasites utilisant les deux espèces hôtes alternatives, alors que seuls quelques DEG ont été identifiés lors de la comparaison entre parasites provenant des deux espèces hôtes alternatives (7 DEG). Nous avons aussi montré que les principaux processus biologiques associés à l'exploitation de différentes espèces hôtes étaient impliqués dans la machinerie cellulaire, le métabolisme énergétique, l'activité musculaire et le stress oxydatif. En contrepartie nous n'avons trouvé aucune trace de sélection associée à la spécificité d'hôte parmi les 55 645 SNP identifiés, ce qui suggère que l'adaptation à de nouvelles espèces hôtes résulte plus probablement de la plasticité que de processus sélectifs. Cette étude fournit un exemple empirique unique concernant le rôle de la plasticité comme mécanisme permettant aux parasites d'utiliser et de s'adapter à des espèces hôtes alternatives ce qui, pour cette espèce parasite en particulier, faciliterait probablement sa dynamique d'invasion.

# Investigating the role of parasite plasticity in parasite host shift: evidences from a transcriptomic approach

---

Mathieu-Bégné Eglantine<sup>1,2,3\*</sup>, Loot Géraldine<sup>1</sup>, Blanchet Simon<sup>1,2</sup>, Mitta Guillaume<sup>3</sup>, Le Potier Clément<sup>1</sup> and Rey Olivier<sup>3</sup>

<sup>1</sup>Université Paul Sabatier, Centre National pour la Recherche Scientifique, Ecole Nationale de Formation Agronomique; Laboratoire Evolution et Diversité Biologique (UMR5174), 118 route de Narbonne, F-31062 Toulouse, France

<sup>2</sup>Université Paul Sabatier, Centre National pour la Recherche Scientifique; Station d'Ecologie Théorique et Expérimentale (UMR5321), 2 route du CNRS, F-09200 Moulis, France

<sup>3</sup>Université Perpignan Via Domitia, UMR5244 Interaction Hôte-Parasite-Environnement, 58 Avenue Paul Alduy, F-66860, Perpignan, France

## Corresponding author:

\*Mathieu-Bégné Eglantine

UMR5174 EDB

118 route de Narbonne

31062 Toulouse

E-mail:[eglantine.mb@gmail.com](mailto:eglantine.mb@gmail.com)



**Authors' contributions:**

OR, GL and SB coordinated the study, supervised the data acquisition. OR performed the RNA extractions and other molecular analyses. EMB ran the bioinformatics analyses. EMB and CLP run the differential gene expression analysis. EMB performed the selection detection analysis. GM, OR and EMB interpreted the results. EMB, OR and SB wrote the manuscript. All authors participated to the conception of the ideas, and they all agree to be held accountable for the content therein and approve the final version of the manuscript.

**Competing interest statement:**

We have no competing interest.

**Key-words:** Plasticity, Alternative host, Parasite specificity, RNA-seq, Variant calling

## Abstract

The ability of parasites to use and adapt to new host species strongly affects their ecological and evolutionary dynamics, as well as their pathogenic effects on host communities. Although host shift has been reported in many parasite species, very few are known about the processes permitting parasites to adapt to alternative host species. Here, we tested the hypothesis that plasticity is a key mechanism allowing host shift in parasites. Focusing on an emerging parasite (*Tracheliastes polycolpus*) infecting freshwater fish species, we used a transcriptomic approach to compare patterns of gene expression between parasites infecting their principal host species to those of parasites infecting two alternative host species. We found 120 protein-coding genes that were differentially expressed (DEGs) between parasites infecting different host species. Most DEGs were found between parasites using the principal host species compared to parasites using the two alternative host species, whereas only a few DEGs were identified when comparing parasites from the two alternative host species (7 DEGs). The main biological processes associated with the exploitation of different host species were related to cellular machinery, energetic metabolism, muscle activity and oxidative stress. We found no evidence for selection associated with host specificity among the 55 645 identified SNPs, which rather suggests that adaptation to alternative host species likely results from plasticity. This study provides unique empirical evidence that plasticity is a key mechanism for parasites to use and adapt to alternative host species, which probably facilitates invasion in this particular parasite species.

## Introduction

The degree of specificity of parasites and their potential ability to shift and use new alternative host species strongly impact their ecological and evolutionary dynamics (Heath 1981, Hoberg et al. 1997). Until recently, parasites were considered as highly specific to a given host species, and the use of alternative host species was thus considered an exception rather than a rule (Ronquist 2003). However, accruing evidence indicate that many parasites display a larger host range than previously thought, and that invasive and emerging diseases can rise from new species associations, the so called *parasite paradox* (Poulin and Keeney 2008, Jones et al. 2008, Agosta et al. 2010). Parasite ability to exploit another host species in addition to the principal host has important consequences for parasite population sustainability and dynamics, but also for their potential to establish and spread into new areas. This have recently raised public health and agronomic concerns (eg., epizootic outbreaks, parasite spill-over and spill-back, (Thompson et al. 2009)) due to the strong impact of parasites on host populations and communities (Daszak 2000, Dunn 2009, Moir et al. 2010). As such, although selection is expected to favor host-parasite specificity, parasites seem to have maintained the potential for infecting new host species. Elucidating this paradox remains an enigmatic question.

At the beginning of the 21th century, Claude Combes (Combes 2001) theoretically defined the conditions that allow a parasite to infest a host based on the idea that parasites have to cross “filters” for being able to infest a particular host species. First, the *encountering* filter (also known as the *concept of opportunity*, (Araujo et al. 2015)) consists in the temporal and spatial overlap between parasites and hosts that make their encountering possible. The second filter relies on the *compatibility* between parasites and hosts, which can be defined as the ability of one parasite lineage to durably exploit one host lineage (Combes 2001). Based on this concept,

parasite host shift is expected to result from a relaxation of the encounter and/or the compatibility filter (Lootvoet et al. 2013).

Moreover, according to the “ecological fitting” concept, parasite host shift also relies on the physiological capacity of parasites to establish on the new host (Janzen 1985, Agosta et al. 2010). More generally, when facing new biotic and abiotic environmental conditions (in this case a new host species), organisms (here parasites) are expected to survive and reproduce, provided that these new conditions belong to their reaction norms (Agosta et al. 2010). In this regard, two main mechanisms of ecological fitting are likely to favor parasites’ host shift. First the ecological fitting via “resource tracking” relies on the idea that parasite shifts from one host species to another is mainly determined by the ability of parasites to cope with defense mechanisms and resources that are similar. This process is associated with the relaxation of the compatibility filter (Ricklefs and Fallon 2002). Second the ecological fitting via the “sloppy fitness space” hypothesis relies on the idea that a plastic response of parasites infecting a new host species would favor the emergence of an adaptive phenotype on this new host (De Fine Licht 2018). Potential host range of parasites would hence be more likely determined by the opportunity to found new potential host species in a common habitat (relaxation of the encountering filter) (Timms, and Read, 1999, Cooper et al. 2012, Lootvoet et al. 2013). So far, studies have mainly focused on demonstrating the existence of alternative host use in many parasite species (Edwards and Vidrine 2006, Kvach and Sasal 2010, Poulin et al. 2011), and only a few of them have investigated the factors facilitating/allowing host shift (Little et al. 2006, Lootvoet et al. 2013).

In particular, the role of phenotypic plasticity in facilitating parasite shifts from one main host species to alternative host species has been yet empirically overlooked. Recent theoretical studies have proposed an appealing theoretical framework for the understanding parasite host shift

(Araujo et al. 2015, Nylin et al. 2018, De Fine Licht 2018). In particular, these studies have extended the concept of ecological fitting by sloppy fitness space to parasitism and introduced the role of plasticity as a mechanism allowing exploitation of new host species (Araujo et al. 2015, Nylin et al. 2018, De Fine Licht 2018). For instance plasticity is expected to allow marginal exploitation of new host species by parasite which at longer term can lead to adaptation through selective processes and/or (epi-)genetic assimilation (De Fine Licht 2018, Danchin et al. 2019). However, so far such a theoretical framework is lacking of empirical evidences (Little et al. 2006). This might be –at least partly– explained by the difficulty to measure plasticity related to traits associated to host use. With the advance of Next Generation Sequencing approaches, we have now the opportunity to more easily and accurately quantify parasite plastic responses at the molecular level, notably at the gene expression level using transcriptomic approaches (De Fine Licht 2018). Thus, transcriptomic approaches offer a new opportunity i) to assess the role of parasite gene expression plasticity associated with the use of alternative host species and ii) to identify the functions of the genes involved in host shift. Together, these two improvements are likely to provide important insights for better understanding the evolutionary dynamics of parasites. Within the current context of accruing emerging infectious diseases in both natural and cultivar systems, this is of prime interests for predicting and addressing the general impact of pathogens on host communities in both wild and domestic populations.

Here, we tested whether the ability to exploit different host species of an emerging parasite of fish is associated to a plastic response, relying on a transcriptomic approach. We considered an homogeneous parasite population regarding genetic differentiation (Mazé-Guilmo 2016) and thus assessed plastic response of parasites exploiting different hosts species by measuring their gene expressions. More specifically, we compared gene expression patterns between parasites from

natural populations sampled on their principal host species and on two alternative hosts. We investigated the molecular processes and searched for possible trace of selection associated with these host shifts at the transcribed genome scale. We built our study on *Tracheliastes polycolpus*, a recently introduced and emerging parasite that has successfully shifts on several local host species/lineages over a relatively short period (i.e., less than 200 generations (Lootvoet et al. 2013, Rey et al. 2015, Mazé-Guilmo 2016)). Since *T. polycolpus* was shown to infect host species having similar habitat and diet rather than similar defense mechanisms (Lootvoet et al. 2013), we predict that host shifts will rely on a plastic response rather than on selective process (accordingly to the ecological fitting via “sloppy fitness space” hypothesis). More precisely, we expect i) to identify differentially expressed genes in parasites exploiting different host species and ii) that parasite transcriptomic adjustments will mainly differentiate parasites exploiting the principal host from parasites using alternative hosts since parasites are assumed to be better adapted to their principal host on which they perform better (Lootvoet et al. 2013, Araujo et al. 2015). Consequently we also expect to identify gene functions associated with the parasite exploitation of either their main or alternative hosts, which include energy acquisition, metabolism, growth and reproduction. In contrast, we expect to detect no genomic variation under selection between parasites from different host species i) given we sampled a homogeneous parasite population and ii) given the recent establishment of *T. polycolpus* populations in France, which means that transcriptomic adjustments will indeed indicate plasticity.

## **Material and Method**

### **Study model**

*Tracheliastes polycolpus* is an Crustacean ectoparasite of freshwater fish that belongs to the *Copepoda* order (Fryer 1982). This parasite originates from eastern Eurasia where it is mainly

associated with the common ide (*Leuciscus idus*) but also occasionally with some other cyprinids species (von Nordmann 1832, Monod and Vladykov 1931). During the 1920s' due to fish trades, *T. polycolpus* has reached western Europe and has rapidly spread over several watersheds in England, France and Spain where it now constitutes an invasive species (Rey et al. 2015). This invasive dynamics likely results from the ability of *T. polycolpus* to infect a large range of new host species. For instance in France, the principal hosts of *T. polycolpus* are two *Leuciscus* species (i.e., the common dace *Leuciscus leuciscus* in the northeastern part of France and the rostrum dace *L. burdigalensis* in the southwestern part of France) but *T. polycolpus* is also regularly found on five alternative host species: the toxostome (*Parachondrostoma toxostoma*), the gudgeon (*Gobio gobio* and *Gobio occitaniae*), the common minnow (*Phoxinus phoxinus*), the roach (*Rutilus rutilus*) and the chub (*Squalius cephalus*) (Lootvoet et al. 2013).

*Tracheliastes polycolpus* life cycle includes three development stages namely, the *nauplius* and copepodite stages that are larvae stages and the adult stage (Fryer 1982). Males and larvae are free-living forms whereas females are parasitic forms that complete their life cycle on one single host (Fryer 1982). As many copepod parasites, only apparatus involved in anchoring (a disk-shaped bulla), reproduction (two eggs sacs) and feeding (two maxilla) have been maintained in females *T. polycolpus*. They usually attach to fish fins and feed on epithelial and mucus cells in order to sustain their growth and the development of their egg sacs. This feeding activity results in direct damages on their host such as partial to total fin degradation and secondary inflammations (Loot et al. 2004), which can result on a decrease in host growth and a poorer survival (Blanchet et al. 2009b).

## **Sampling design and sequencing**

The sampling design and molecular analysis up to the transcriptome assembly are detailed in Mathieu-Bégné *et al.*(2019). Briefly, five parasitized dace, four parasitized gudgeon and four parasitized minnow were caught by electro-fishing the 11<sup>th</sup> of July 2013 in a single locality in the Salat River (southwestern France, 43°04'43.0"N; 0°57'29.0"E) so as to avoid confounding environmental and temporal effects. Five parasites on each host species were collected with sterile forceps resulting in fifteen parasite samples directly stored in RNAlater for 24 hours and then conditioned at -80°C before RNA extraction. To limit contamination with fish tissues, total RNAs were extracted from the parasite trunk only (a part of the parasite that is not physically in contact with the host). Individual libraries were obtained using the RNeasy Plus Mini Kit (Qiagen) with a final elution volume of 40 µLRNase-free water. A nanodrop ND-8000 (Thermo Scientific) and a BioAnalyser (Agilent Technologies) were then used to assess the quality and quantity of RNA extractions respectively. Finally, individual libraries were paired-end sequenced using Illumina Hiseq 2000 technology. The sequencing resulted in about 420 million 2x100 bp paired-end reads, with an average of 28 million paired-end reads per sample (see Mathieu-Bégné *et al.* 2019 for more details).

## **Gene expression analyses**

### ***Quantification of genes expression levels***

For each of the fifteen samples, the raw sequenced reads were first quality trimmed using Trimmomatic (Bolger *et al.* 2014). For each gene, the number of transcripts was then quantified using the reference transcriptome of *T. polycolpus* which includes 17 157 non-redundant protein-coding genes (available on Genebank under Bioproject PRJNA476682, (Mathieu-Bégné *et al.* 2019a)) and using the script *align\_and\_estimate\_abundances.pl* from the Trinity platform



(version 2014-07-17 (Haas et al. 2013)) which calls successively the aligner Bowtie2 (version 2.0 (Langmead and Salzberg 2012)) and the software RSEM (version 2.3.1 (Li and Dewey 2011)). All parameters were set as default except for Bowtie2 in which we tolerated one nucleotide mismatched and set the mismatch penalty to 2. These parameters were used in order to better account for the potential genetic diversity existing in wild populations and improving the mapping rate for each individual. Finally, the Trinity script *abundance\_estimates\_to\_matrix.pl* was used to set-up all transcript counts from all of the 15 samples into a single matrix for subsequent analyses.

### ***Identification of differentially expressed genes***

We used the R package *EdgeR* (R version 3.4.2 (Robinson et al. 2010, R Core Team 2017)) to identify significant differentially expressed genes between parasites associated to the three different host species (i.e., *L. burdigalensis*, *G. occitaniae* and *P. phoxinus*). First, we filtered out genes with too low expression levels based on their transcripts counts (i.e., with count of 5–10 in each library according to *EdgeR* user guide, (Robinson et al. 2010)) using a copy per million threshold of 0.4 in at least the five samples out of the fifteens thus resulting in a set of 12 357 genes. We then calculated the normalization factor using the trimmed mean of M-values (TMM) normalization (Robinson and Oshlack 2010) to control for potential heterogeneity in library sizes among individuals. Finally, common dispersion and tagwise dispersion (i.e., two dispersion measures that allow accounting for heterogeneity of genes expression levels between genes) were estimated and included in the subsequent models.

In order to compare gene expression levels between parasites from different host species we first performed a multi-group comparison (i.e., an ANOVA-like test that detects genes differentially expressed between any groups). A generalized linear model was adjusted using the *glmFit*

function in *EdgeR* and likelihood ratio tests were performed using the *EdgeR* function *glmLRT*. Additionally, pairwise comparisons were conducted between parasites from different host species to refine differentially expressed genes between groups of parasites. No intercept was used, and contrasts between groups were set so as to perform each pairwise comparison using the *EdgeR* function *exactTest* (i.e., equivalent of Fisher exact tests but adapted for negative binomial distributed data). For both the multi-group test and the pairwise comparisons, p-values were corrected for multiple comparisons and significant differentially expressed genes were selected at a False Discovery Rate (FDR) of 0.05 (Storey 2011). Finally, common significant DEGs identified from either the multi-group or the pairwise comparisons were used in a clustering approach analysis (per genes and per sample) on genes moderated log-counts-per-million. The results obtained from this clustering approach were visualized through a heat map.

### ***Functional analyses***

We conducted a functional analysis at the whole transcriptome level in order to investigate the relevant biological processes impacted during *T. polycolpus* host shift from its principal host *L. burdigalensis* to alternative host species (i.e., *G. occitaniae* and *P. Phoxinus*). To do so, we performed a Rank Based Gene Ontology Analysis (RBGOA, (Wright et al. 2015)) between parasites from *L. burdigalensis* and parasites from *G. occitanea* and *P. phoxinus* respectively. The RBGOA leverages on gene ontologies (GOs, i.e., functional categories associated to genes and resulting from annotation) to investigate the most represented biological processes at the whole transcriptome level and their content in differentially expressed genes. The RBGOA was based on gene annotation from Mathieu-Bégné *et al.* (2019) and the differential expression of the 12 357 genes measured as Log Fold Change (LFC). Briefly, RBGOA first clusters GOs according to their representative genes in order to identify the most meaningful GOs, and then

ranks the identified biological processes according to their average expression levels (over all representative genes). Finally, biological processes significantly enriched in differentially expressed genes are identified through a Mann-Whitney rank test and applying a FDR correction (FDR<0.001).

### **Variant calling and FST outlier analysis**

In order to investigate signature of selection between parasites collected from different host species we aimed to identify Single Nucleotide Polymorphism (SNPs) from RNA sequencing. To do so we used the Genome Analysis Toolkit (GATK version 4.0.1.2 (McKenna et al. 2010)) and followed the recommendations regarding variant calling based on transcriptome assembly of non-model organisms presented in Eldem et al. (2017). Succinctly, quality trimmed reads were first aligned on *T. polycolpus* transcriptome using the aligner Bowtie2 using the same parametrization used for the transcript quantification (Langmead and Salzberg 2012). Picard (version 2.1.1., <http://picard.sourceforge.net/>) and SAMtools (version 1.3.1 (Li et al. 2009)) were then used to conduct the following pre-processing steps: 1) Assign a read group to each sample, 2) sort aligned reads, 3) remove duplicates and 4) merge ‘.bam’ alignment results in fifteen ready reads files. After these pre-processing steps, we used the GATK tools to call potential variants. First, the fifteen ready reads files were parsed to *HaplotypeCaller* to call variants in each sample. Each genomic Variant Call Format (gVCF) files were then combined in one single gVCF file using the *CombineGVCFs* function. The *GenotypeGVCFs* function was next used to perform the joint genotyping. Variants were filtered out when having a phred-scaled confidence threshold below 30. At this stage, since we were only interested in SNPs, we used the function *SelectVariant* to sort SNPs and stored them in a raw SNPs gVCF file. We then filtered the variants using the GATK hard filtering method (*VariantFiltration* function) and self-defined parameters (QD < 5.0,

FS > 30.0, MQ < 10.0, MQRankSum < -12.5, ReadPosRankSum < -8.0) resulting in a set of 94 403 SNPs.

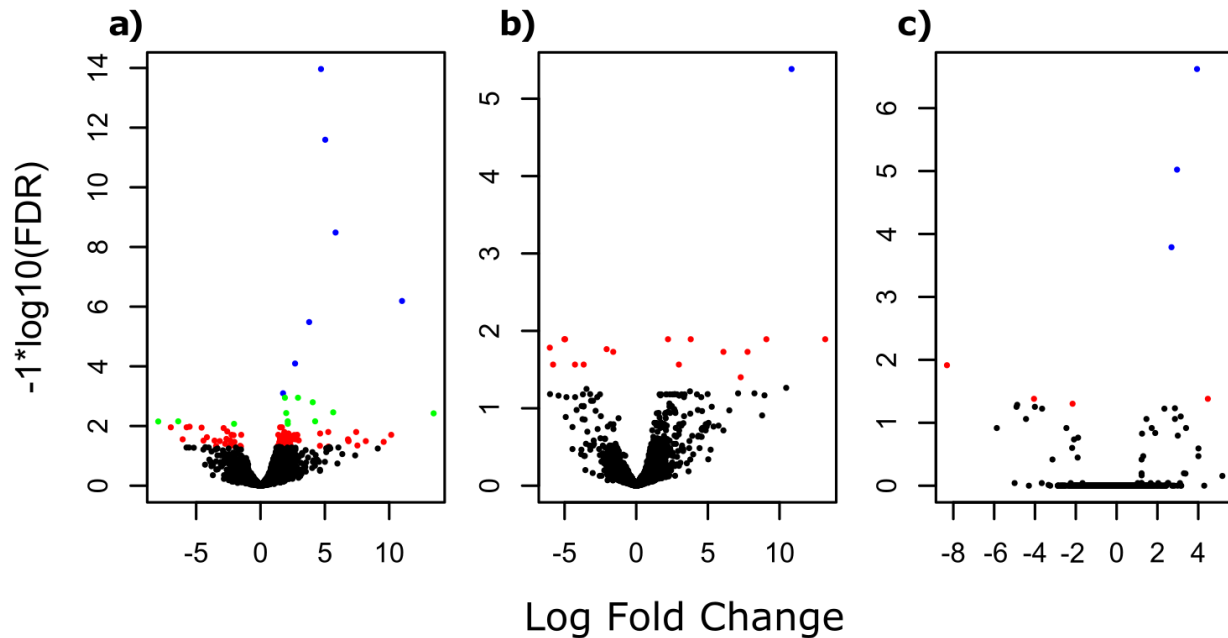
To detect variants possibly under selection and that could differentiate parasites from different host species we conducted a FST outlier analysis using the software Bayescan (version 2.0 (Foll and Gaggiotti 2008)). First, the script *make\_bayescan\_input.py* provided in De Wit et al. (2012) was run to format our set of SNPs in Bayescan format with the number of individual required per population set to 4 and resulting in a set of 61 880 SNPs. Finally we removed from this set of SNPs those having a low allelic frequency (i.e., minor alleles must be at least recorded twice otherwise they likely corresponds to sequencing errors) using the option *-d* in Bayescan that allows discarding any listed SNPs and finally run the FST outlier test on the resulting 53 645 high quality SNPs.

## Results

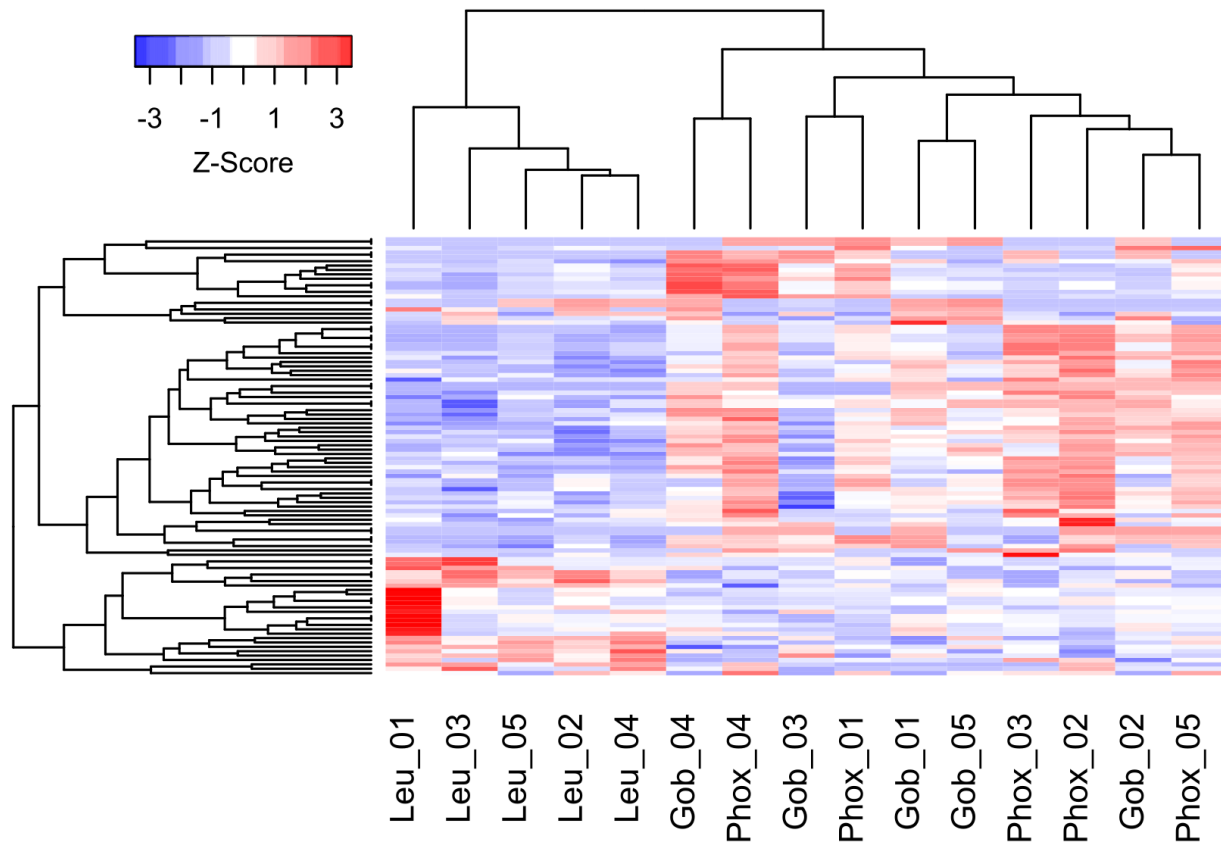
### Differentially expressed genes

The multi-group comparison (i.e., comparison of parasite genes expression between dace, minnow and gudgeon parasites) of gene expression levels revealed a total of 120 significant differentially expressed genes (DEGs) out of the 12 357 studied genes. Moreover, according to our pairwise comparison approach, 92 significant DEGs were found between parasites infecting minnow and dace, among which 27 were under-expressed and 65 were over-expressed genes in parasites infecting minnow compared to parasites infecting dace (Fig. 1-a). When comparing parasites from gudgeon and dace, 17 significant DEGs were found, including 8 under-expressed and 9 over-expressed genes in parasites from gudgeon compared to parasites from dace (Fig. 1-b). Finally, only seven significant DEGs were found when comparing parasites from minnow and

from gudgeon including three under-expressed and four over-expressed genes in parasites from minnow compared to parasites from gudgeon (Fig.1-c). Overall, 100 DEGs were found in common between these two approaches (i.e., the multi-group and the pairwise analyses). The clustering approach conducted on these common 100 DEGs allowed segregating parasites infecting daces from those associated with alternative host species, but failed to correctly discriminate parasites infecting gudgeon from parasites infecting minnow (Fig. 2).



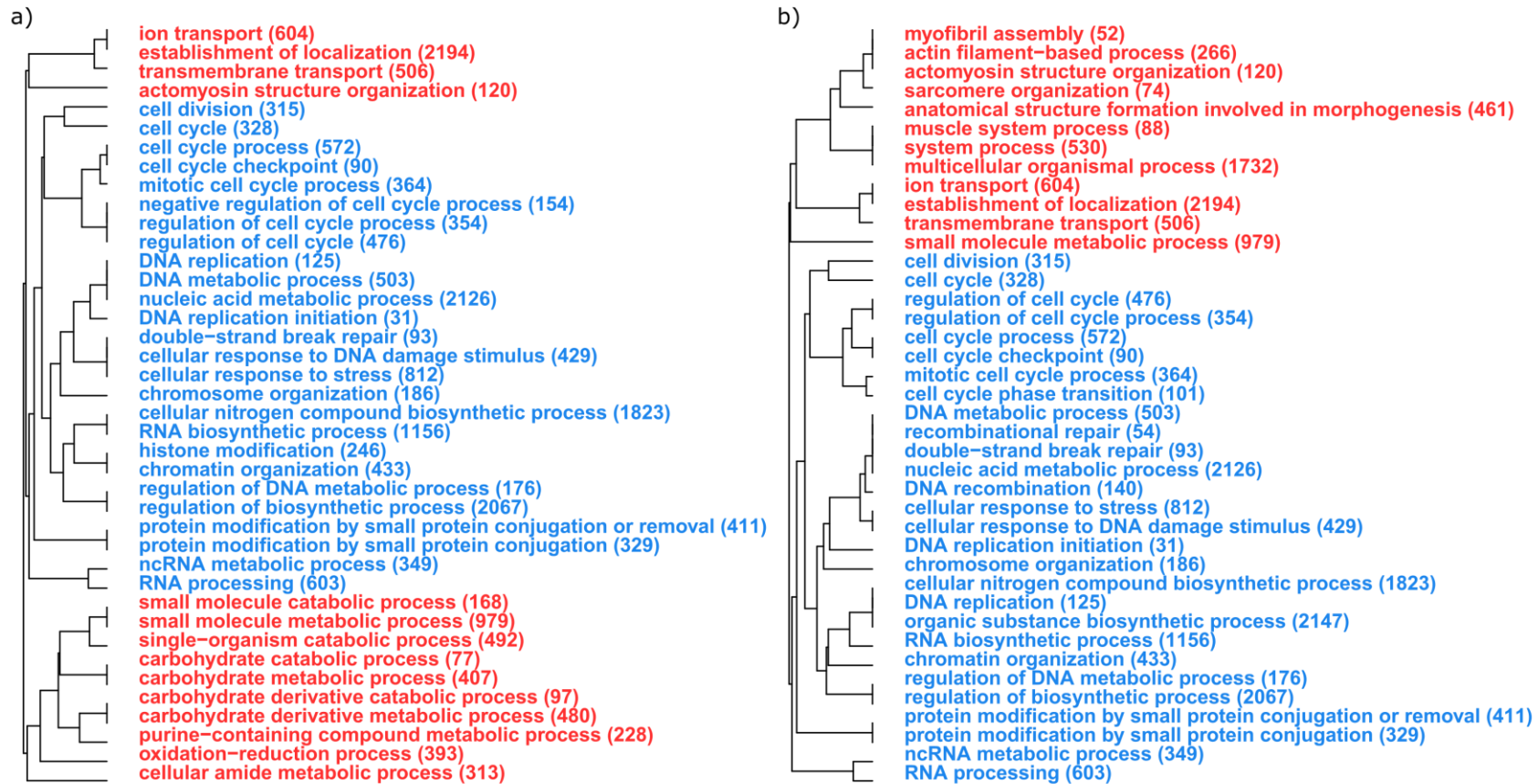
**Figure 1:** Volcano plot showing the log transformed adjusted p-values (i.e., FDR) and the log fold changes for the 17157 protein-coding genes of *T. polycolpus* transcriptome regarding a) the comparison between parasites infecting dace and minnow respectively, b) the comparison between parasites infecting dace and gudgeon respectively and c) the comparison between parasites infecting minnow and gudgeon respectively. Black dots refer to differentially expressed genes below a FDR of 5%. Red dots refer to differentially expressed genes at a FDR of 5%, green dots refer to differentially expressed genes at a FDR of 1% and blue dots refer to differentially expressed genes at a FDR of 1 %.



**Figure 2:** Heat map showing gene expression profiles (higher Z-score refer to higher expression) along with clustering between the different parasite genes and samples for the 100 DEG commonly identified from the multi-group and the pairwise comparisons. “Leu” refers to parasites infecting dace, “Gob” to parasites infecting gudgeon and “Phox” to parasites infecting minnow.

## Functional analyses

According to our functional analyses, 101 and 125 GO significantly enriched (at a 0.001 FDR threshold) with DEGs (either over- or under-expressed) were found for the pairwise comparisons between parasites infecting minnow and dace and between parasites infecting gudgeon and dace respectively (Fig. S1). In order to ease interpretation we focused on the 40 most significant GOs for each comparison. The 40 GO most significantly enriched with over- or under-expressed genes for the two comparisons between parasites from the principal host (i.e., dace) and the two alternative hosts (minnow and gudgeon) were very similar (Fig. 3). For example, GO enriched with under-expressed genes in parasites from minnow or gudgeon (i.e., alternative hosts) compared to dace parasites (i.e., principal host) are both involved in cellular machinery (e.g., *cell division*, *DNA replication* or *RNA processing*, Fig. 3-a and Fig. 3-b), whereas GO enriched with over-expressed genes in parasites infecting minnow or gudgeon compared to parasites infecting dace are both related to energetic metabolism (e.g., GO *carbohydrate catabolic process*, *carbohydrate derivative metabolic processor*, *small molecule catabolic process*; Fig. 3-a and Fig. 3-b), to muscle activity (e.g., GO *ion transport*, *transmembrane transport*, *actomyosin structure organization*; Fig. 3-a and Fig. 3-b) and to a lesser extent to oxidative stress (e.g., GO *oxidation-reduction process*; Fig. 3-a and Fig. 3-b).

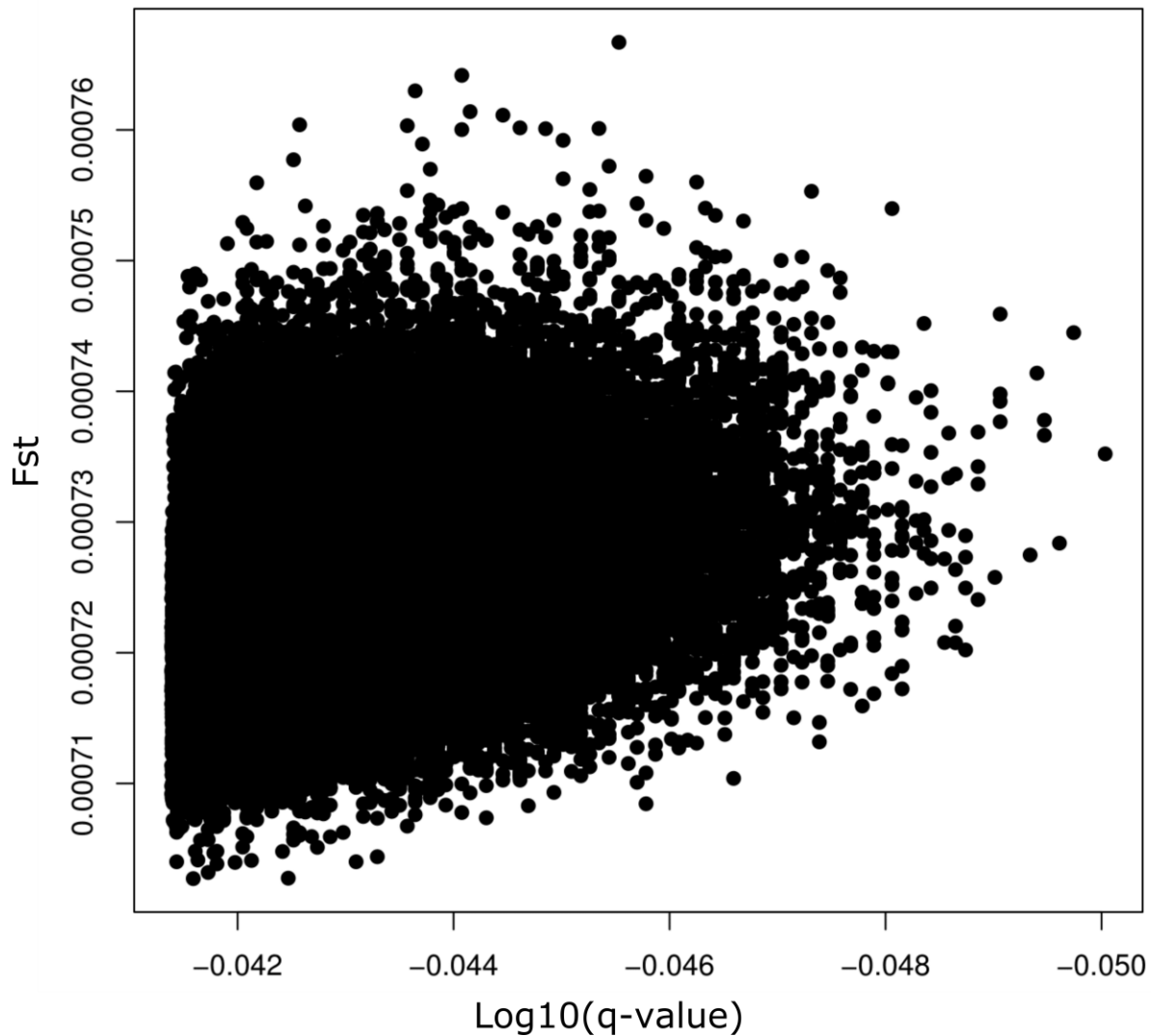


**Figure 3:** Top 40 of the most enriched in differentially expressed genes Gene Ontologies (GOs) for the comparison between a) parasites infecting minnow and dace and b) parasites infecting gudgeon and dace. Blue GOs refer to GOs enriched in under-expressed genes between alternative host parasites compared to principal host parasites and red GOs refer to GOs enriched in over-expressed genes between alternative host parasites compared to principal host parasites. The total number of genes involved in each GO is given in brackets.



### Detection of selection

The outlier  $F_{st}$  analysis based on the 53 645 identified SNPs to compare the level of genetic differentiation between parasites from the three host species revealed very low  $F_{st}$  values with an average of  $0.001 \pm 0.001$  ( $\pm SE$ ). No  $F_{st}$  outliers –and consequently no trace of selection–were identified among all the investigated protein-coding genes (Fig. 4).



**Figure 4:** Plot showing  $F_{st}$  values and corresponding log transformed adjusted p-values for the 53 645 high quality SNPs identified from *T. polycolpus* RNA-sequencing. Black dots refer to non-significant outliers.

## Discussion

In this study we aimed to test whether plasticity could underline the ability of an emerging parasite to exploit new host species. Overall we provided strong support for the role of transcriptomic plasticity in parasite ability to exploit alternative host species. We indeed showed that parasites exploitation of either the principal host or alternative host was associated with transcriptomic adjustments that likely reflect a plastic response rather than the result of selective processes.

Using a variant calling approach we indeed showed that *T. polycolpus* ability to exploit alternative host species was not the result of selection acting on protein-coding genes, and thus transcriptomic adjustments would more likely reflect plasticity. However, the hypothesis of an ongoing selection for host species in *T. polycolpus* cannot totally be excluded since we only focused on protein-coding regions and not on the entire genome. As we observed differences in genes expression between parasites from different host species, we might have missed some (adaptive) genomic polymorphism occurring in non-coding regulatory regions (Chen and Rajewsky 2007). Though, a weak genomic differentiation between parasites infesting dace, minnow or gudgeon was expected in our case since i) all the parasites sampled in the present study were collected at the same sampling site and ii) a previous microsatellite-based study showed that southwestern *T. polycolpus* constitute an homogeneous population with no differentiation between parasites feeding on different host species (which the average  $F_{st}$  value under 0.001 we measured confirmed) (Mazé-Guilmo 2016). It is worth noting that, plasticity and selection are actually not antagonist mechanisms regarding their theoretical role in parasite host shift (De Fine Licht 2018). Notably, plasticity allows some variation in parasite phenotypes that is expected to provide the raw material for natural selection to shape the evolution of parasite

specificity at longer timescales (Chevin et al. 2010, De Fine Licht 2018). Since *T. polycolpus* recently emerged in France (Rey et al. 2015), it is likely that selection did not have enough time or that there is not enough base genomic variation to lead *T. polycolpus* to specialization on the alternative host.

In this context, the total of 120 significant differentially expressed genes we identified between parasite infecting different host species give a strong support for the role of transcriptomic plasticity in parasite ability to exploit alternative host species. Such a plastic response to different host species was expected for *T. polycolpus*, and this meets the theoretical expectation of the ecological fitting via sloppy fitness space hypothesis (Araujo et al. 2015, Nylin et al. 2018, De Fine Licht 2018). Overall, our results are in line with previous studies suggesting that parasites are able to cope with new environmental conditions through plasticity (Poulin and Hamilton 2000, Loot et al. 2008). However, contrary to previous studies that have investigated parasite plasticity in response to environmental changes at the within-host scale (Cameron et al. 2013, Birget et al. 2017, Shocket et al. 2018, but see Hébert et al. 2017 for an example of transcriptomic adjustments during host transition within a heteroxenous parasite life cycle), we here provide evidence that transcriptomic plasticity constitutes a key mechanism permitting parasites to shift from their principal hosts to alternative hosts.

When specifically studying pairwise differences in gene expression between parasites sampled from the three host species, our results revealed that more transcriptomic adjustments are required from *T. polycolpus* when shifting from its main host to minnow than to gudgeon. The largest amount of DEG was indeed found for the comparison between dace and minnow parasites. Interestingly, *T. polycolpus* infecting minnow also display lower fitness (a reduced body size and a fewer number of eggs) than those infecting dace, whereas *T. polycolpus* maintains a fitness similar to that observed on dace when infecting gudgeon (Lootvoet et al.

2013). We thus suggest that transcriptomic plasticity allows *T. polycolpus* shifting from its principal host to the gudgeon better than to the minnow. We also argue that consistently to theoretical expectations, plasticity can allow to cope with a new host resource without necessarily being yet optimally adapted (Lootvoet et al. 2013, De Fine Licht 2018).

We further demonstrated that *T. polycolpus* is likely developing similar molecular strategies to exploit alternative hosts irrespectively of the host species. Our results indeed show that when accounting for the 100 DEGs commonly found in multi-group and in pairwise analyses, the transcriptomic profiles between parasites infecting either one alternative host or the other were indistinguishable. Moreover, very similar functions enriched in genes differentially expressed between the main host and either of the two alternative hosts were identified (i.e., cellular machinery, energetic metabolism, muscle activity and oxidative stress). These two results support a similar–although not identical– plastic response to both alternative host species when compared to the main host. We consequently propose that transcriptomics plasticity in *T. polycolpus* allows exploitation of alternative host species but in a limited fashion, since *T. polycolpus* is constrained to one similar strategy on both gudgeon and minnow. Such limitation could be explained by a still standing compatibility filter when exploiting common alternative hosts, but that might display different immune defenses compared to the principal host that is phylogenetically closer to the original host *T. polycolpus* have co-evolved with (Kottelat and Freyhof 2007, Lootvoet et al. 2013). Consequently, although minnow apparently constitute the host that requires the highest plastic adjustment for *T. polycolpus*, parasite switching to minnow seem nonetheless constrained to the expression of biological functions involved in a general host shift process (similar in minnow and gudgeon) hence not yet specialized. In this regard, *T. polycolpus* only recently settled in French River Basins (Rey et al. 2015), and the limited specific response to each host

species of *T. polycolpus* suggests that a longer interaction between parasite and alternative host species would eventually be necessary to achieve more specific strategies.

The identified biological processes involved in parasite responses to host shift also suggest that alternative host exploitation by the parasite is challenging and might turn out less efficient compared to principal host exploitation (which is more manifest on minnow parasites). Parasites using alternative host species displayed over-expression (compared to the main host) of genes related to energetic metabolism, muscle activity and oxidative stress, and under-expression of genes related to cellular machinery. Specifically, parasites infecting alternative host species seem to invest more in energy intake but without being able to translate this energy to cellular machinery. For instance, alternative host parasites displayed biological processes enriched in over-expressed genes such as energetic metabolism but also muscle activity, which may refer to maxilla movement during the *T. polycolpus* feeding activity (Fryer 1982). At the same time, alternative host parasites display a lower investment in cell machinery, which is at the core of their development. Fewer cell cycles in parasite using alternative host species is also congruent with the smaller body size and a lower egg number that was observed in parasites infecting minnow (Lootvoet et al. 2013). Interestingly on gudgeons *T. polycolpus* are not smaller than on dace, so the under-investment in cell machinery is somehow overcome by the investment in energy intake (Lootvoet et al. 2013). Finally, parasite using alternative hosts also over-expressed genes involved in oxidative stress, which is a common response of many organisms to stressful environments (Storz and Imlay 1999). Hence, exploiting alternative hosts seems more challenging for *T. polycolpus* than exploiting its main host. We consequently argue that parasites using alternative host species are more prone to invest in energy intake, probably to overcome the costs associated to a poorer exploitation of the resources provided by an alternative host species.

Overall, this study is one of the first empirical essays strongly suggesting that plasticity has a critical role in parasite ability to exploit new alternative host species. In particular, we showed that adjusting the expression level of some specific genes allows the parasite to exploit alternative hosts that might challenge the parasite with new defenses or resources, yet in a suboptimal fashion. The suboptimal exploitation of alternative hosts might be explained by a limited range of plastic response of *T. polycolpus*. Such a lower adaptation of parasite infecting alternative host is in agreement with the ecological fitting via sloppy fitness space theory that expects marginal adjustment which lead, at first, to suboptimal parasite fitness on alternative hosts (Araujo et al. 2015, De Fine Licht 2018). Hence, we argue that parasite plastic response enabling exploitation of suboptimal alternative hosts could be critical at short time scale for population sustainability making alternative hosts some potential reservoirs for the parasite when the principal host is not (enough) encountered. Such evolutionary short-term process might provide the opportunity for the emergence of more sustainable genetically-based phenotypic change in the parasites, and hence could constitute the first step of a new co-evolutionary dynamics between parasites and new host species at broader timescales.

## References

- Agosta, S. J. et al. 2010. How specialists can be generalists: resolving the “parasite paradox” and implications for emerging infectious disease. - *Zoologia* 27: 151–162.
- Araujo, S. B. L. et al. 2015. Understanding Host-Switching by Ecological Fitting. - *PLOS ONE* 10: e0139225.
- Birget, P. L. G. et al. 2017. Phenotypic plasticity in reproductive effort: malaria parasites respond to resource availability. - *Proc. R. Soc. B* 284: 20171229.
- Blanchet, S. et al. 2009. Why do parasitized hosts look different? Resolving the “chicken-egg” dilemma. - *Oecologia* 160: 37–47.
- Bolger, A. M. et al. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. - *Bioinformatics* 30: 2114–2120.

- Cameron, A. et al. 2013. Plasticity in transmission strategies of the malaria parasite, *Plasmodium chabaudi* : environmental and genetic effects. - *Evol Appl* 6: 365–376.
- Chen, K. and Rajewsky, N. 2007. The evolution of gene regulation by transcription factors and microRNAs. - *Nat Rev Genet* 8: 93–103.
- Chevin, L.-M. et al. 2010. Adaptation, Plasticity, and Extinction in a Changing Environment: Towards a Predictive Theory. - *PLoS Biol* 8: e1000357.
- Combes, C. 2001. Parasitism: The Ecology and Evolution of Intimate Interactions. - University of Chicago Press.
- Cooper, N. et al. 2012. Phylogenetic host specificity and understanding parasite sharing in primates. - *Ecology Letters* 15: 1370–1377.
- Danchin, E. et al. 2019. Epigenetically facilitated mutational assimilation: epigenetics as a hub within the inclusive evolutionary synthesis: Epigenetics as a hub for genetic assimilation. - *Biol Rev* 94: 259–282.
- Daszak, P. 2000. Emerging Infectious Diseases of Wildlife-- Threats to Biodiversity and Human Health. - *Science* 287: 443–449.
- De Fine Licht, H. H. 2018. Does pathogen plasticity facilitate host shifts? - *PLOS Pathogens* 14: e1006961.
- De Wit, P. et al. 2012. The simple fool’s guide to population genomics via RNA-Seq: an introduction to high-throughput sequencing data analysis. - *Molecular Ecology Resources* 12: 1058–1067.
- Dunn, A. M. 2009. Chapter 7 Parasites and Biological Invasions. - In: *Advances in Parasitology*. Elsevier, pp. 161–184.
- Edwards, D. D. and Vidrine, M. F. 2006. Host Specificity among *Unionicola* spp. (*Acari: Unionicolidae*) Parasitizing Freshwater Mussels. - *The Journal of Parasitology* 92: 977–983.
- Eldem, V. et al. 2017. Transcriptome Analysis for Non-Model Organism: Current Status and Best-Practices. - *Applications of RNA-Seq and Omics Strategies - From Microorganisms to Human Health*.
- Foll, M. and Gaggiotti, O. 2008. A Genome-Scan Method to Identify Selected Loci Appropriate for Both Dominant and Codominant Markers: A Bayesian Perspective. - *Genetics* 180: 977–993.
- Fryer, G. 1982. The parasitic copepoda and branchiura of British freshwater fishes. A handbook and key. - Freshwater Biological Association.
- Haas, B. J. et al. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. - *Nature Protocols* 8: 1494–1512.
- Heath, M. C. 1981. A generalized concept of host-parasite specificity. - *Phytopathology* 71: 1121–1123.

- Hébert, F. O. et al. 2017. Major host transitions are modulated through transcriptome-wide reprogramming events in *Schistocephalus solidus*, a threespine stickleback parasite. - *Molecular Ecology* 26: 1118–1130.
- Hoberg, E. P. et al. 1997. Host-parasite co-speciation: history, principles, and prospects. - *Host-Parasite Evolution: General Principles and Avian Models*: 212–235.
- Janzen, D. H. 1985. On ecological fitting. - *Oikos* 45: 308–310.
- Jones, K. E. et al. 2008. Global trends in emerging infectious diseases. - *Nature* 451: 990–993.
- Kottelat, M. and Freyhof, J. 2007. *Handbook of European freshwater fishes*. - Publications Kottelat.
- Kvach, Y. and Sasal, P. 2010. *Telosentis exiguus* (von Linstow, 1901) (*Palaeacanthocephala: Illiosentidae*), a generalist parasite of fishes in the Mediterranean basin. - *Systematic Parasitology* 76: 9–18.
- Langmead, B. and Salzberg, S. L. 2012. Fast gapped-read alignment with Bowtie 2. - *Nature Methods* 9: 357–359.
- Li, B. and Dewey, C. N. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. - *BMC bioinformatics* 12: 323.
- Li, H. et al. 2009. The Sequence Alignment/Map format and SAMtools. - *Bioinformatics* 25: 2078–2079.
- Little, T. J. et al. 2006. Parasite-host specificity: experimental studies on the basis of parasite adaptation. - *Evolution* 60: 31–38.
- Loot, G. et al. 2004. The effects of the ectoparasite *Tracheliastes polycolpus* (*Copepoda: Lernaepodidae*) on the fins of rostrum dace (*Leuciscus leuciscus burdigalensis*). - *Parasitology Research* 94: 16–23.
- Loot, G. et al. 2008. Evidence of Plasticity in the Reproduction of a Trematode Parasite: The Effect of Host Removal. - *Journal of Parasitology* 94: 23–27.
- Lootvoet, A. et al. 2013. Patterns and processes of alternative host use in a generalist parasite: insights from a natural host-parasite interaction. - *Functional Ecology* 27: 1403–1414.
- Mathieu-Bégné, E. et al. 2019. De novo transcriptome assembly for *Tracheliastes polycolpus*, an invasive ectoparasite of freshwater fish in western Europe. - *Marine Genomics* 46: 58-61.
- Mazé-Guilmo, E. 2016. Etude du potentiel de colonisation des parasites : une approche intégrative.
- McKenna, A. et al. 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. - *Genome Research* 20: 1297–1303.
- Moir, M. L. et al. 2010. Current Constraints and Future Directions in Estimating Coextinction: Coextinction: Constraints and Directions. - *Conservation Biology* 24: 682–690.



- Monod, T. and Vladykov, V. 1931. Sur quelques copépodes parasites provenant de la Russie sous-carpathique (Tchécoslovaquie). - *Annales de Parasitologie Humaine et Comparée* 9: 202–224.
- Nylin, S. et al. 2018. Embracing Colonizations: A New Paradigm for Species Association Dynamics. - *Trends in Ecology & Evolution* 33: 4–14.
- Poulin, R. and Hamilton, W. J. 2000. Egg size variation as a function of environmental variability in parasitic trematodes. *78*: 6.
- Poulin, R. and Keeney, D. B. 2008. Host specificity under molecular and experimental scrutiny. - *Trends in Parasitology* 24: 24–28.
- Poulin, R. et al. 2011. Host specificity in phylogenetic and geographic space. - *Trends in Parasitology* 27: 355–361.
- R Core Team 2017. R: A language and environment for statistical computing.
- Rey, O. et al. 2015. Elucidating the spatio-temporal dynamics of an emerging wildlife pathogen using approximate Bayesian computation. - *Molecular Ecology* 24: 5348–5363.
- Ricklefs, R. E. and Fallon, S. M. 2002. Diversification and host switching in avian malaria parasites. - *Proceedings of the Royal Society of London. Series B: Biological Sciences* 269: 885–892.
- Robinson, M. D. and Oshlack, A. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. - *Genome biology* 11: 25–34.
- Robinson, M. D. et al. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. - *Bioinformatics* 26: 139–140.
- Ronquist, F. 2003. Parsimony analysis of coevolving species associations. - *Tangled trees: phylogeny, cospeciation and coevolution*: 22–64.
- Shocket, M. S. et al. 2018. Parasite rearing and infection temperatures jointly influence disease transmission and shape seasonality of epidemics. - *Ecology* 99: 1975–1987.
- Storey, J. D. 2011. False discovery rate. - *International encyclopedia of statistical science*. Springer. 504–508.
- Storz, G. and Imlay, J. A. 1999. Oxidative stress. - *Current opinion in microbiology* 2: 188–194.
- Thompson, R. C. A. et al. 2009. Parasite Zoonoses and Wildlife: Emerging Issues. - *IJERPH* 6: 678–693.
- Timms, R. and Read, A. F. 1999. What makes a specialist special? - *Trends in Ecology & Evolution* 14: 333–334.
- von Nordmann, A. 1832. Mikrographische beiträge zur naturgeschichte der wirbellosen thiere. - Reimer.

Wright, R. M. et al. 2015. Gene expression associated with white syndromes in a reef building coral, *Acropora hyacinthus*. - BMC Genomics 16: 371–383.

## Supplementary



**Figure S1:** Differentially expressed genes Gene Ontologies (FDR<0.001) for the comparison between a) parasites infecting minnow and dace and b) parasites infesting gudgeon and dace. Blue GOs refer to GOs enriched in under-expressed genes between alternative host parasites compared to principal host parasites and red GOs refer to GOs enriched in over-expressed genes between alternative host parasites compared to principal host parasites. The total number of genes involved in each GO is given in brackets.



## **Chapitre IV :**

# **Rôle des communautés bactériennes dans le processus d'infection parasitaire**

## Résumé

Selon plusieurs études récentes, les communautés microbiennes hôtes se modifient après et/ou au cours d'une infection parasitaire, suggérant une interaction complexe et multipartite. Cependant, nous ignorons à ce jour si les modifications du microbiote de l'hôte sont une cause d'infection ou bien une conséquence. Dans cette étude, nous avons cherché à déterminer si les modifications du microbiote associé à l'hôte au cours de l'infection parasitaire par *T. polycolpus*, ectoparasite de poissons d'eau douce, sont une cause ou une conséquence des infections chez son hôte principal, *Leuciscus burdigalensis* (un cyprinidé). Pour ce faire, nous avons mené un suivi individuel du microbiote associé à l'hôte, au parasite (adulte et larve) et à l'environnement avant et après infection parasitaire et ceci en conditions naturelles via une expérience d'encagement. Nous avons d'abord montré qu'avant l'infection, il n'existait aucune différence significative entre le microbiote des nageoires des hôtes susceptibles de contracter des parasites et les hôtes restant en bonne santé. Cependant, nous avons trouvé que l'infection parasitaire provoquait des changements importants au sein du microbiote associé à l'hôte au niveau des nageoires parasitées. Nous avons en outre montré que les nageoires parasitées avaient une composition de microbiote similaire à celle des larves de parasite, ce qui suggère une dynamique de co-infection avec les bactéries associées aux parasites lors de l'infection parasitaire. La famille bactérienne des *Burkholderiaceae* s'est avérée particulièrement représentée au sein du microbiote associé à la fois aux nageoires infectées et aux parasites, suggérant ainsi un rôle clé de cette famille dans le déroulement de l'infection. Nous suggérons donc que la composition du microbiote associé à l'hôte serait plutôt une conséquence qu'une cause de l'infection parasitaire pour le système *T. polycolpus*-*L. burdigalensis* et que certaines bactéries associées au parasite pourraient avoir un rôle déterminant dans le déroulement de l'infection.

# Parasites modulate rather than benefits from pre-established host microbiota.

---

Mathieu-Bégné Eglantine<sup>1,2,3\*</sup>, Blanchet Simon<sup>1,2</sup>, Rey Olivier<sup>3</sup>, Toulza Eve<sup>3</sup>, Veysière Charlotte<sup>1</sup>, Manzi Sophie<sup>1</sup> and Loot Géraldine<sup>1</sup>

<sup>1</sup>Université Paul Sabatier, Centre National pour la Recherche Scientifique, Ecole Nationale de Formation Agronomique; Laboratoire Evolution et Diversité Biologique (UMR5174), 118 route de Narbonne, F-31062 Toulouse, France

<sup>2</sup>Université Paul Sabatier, Centre National pour la Recherche Scientifique; Station d'Ecologie Théorique et Expérimentale (UMR5321), 2 route du CNRS, F-09200 Moulis, France

<sup>3</sup>Université Perpignan Via Domitia, UMR5244 Interaction Hôte-Parasite-Environnement, 58 Avenue Paul Alduy, F-66860, Perpignan, France

**Keywords:** Host-parasite interaction, Co-infection, Bacteria community, Individual survey, Metabarcoding.



## Abstract

Many studies have recently demonstrated that host microbial communities change after and/or during parasite infection, hence suggesting a complex multi-partite interaction. However, whether changes in host microbiota are a cause of infection or a consequence has yet to be addressed. In this study we aimed to decipher whether changes in host microbiota during parasite infection by the ectoparasite of freshwater fish *Tracheliastes polycolpus* are a cause or a consequence of infections in its main host *Leuciscus burdigalensis* (a cyprinid fish). To do so, we conducted an individual survey of host, parasite (adult and larvae), and environmental associated microbiota before and after parasite infection in natural conditions relying on a caging experiment. We first showed that before infection there were no significant differences in host fin associated microbiota between hosts that will contract parasites and host that will remain healthy. However, we found that parasite infection triggered important changes in host-associated microbiota at the fin level. We further showed that parasitized fins had similar microbiota composition to the one of parasite larvae suggesting a co-infection dynamics with parasites associated bacteria during parasite infection. The *Burkholderiaceae* bacteria family was particularly represented in parasitized fin and parasite larvae microbiota, hence suggesting a key role of this family in infection progress. We thus conclude that microbiota changes are rather a consequence than a cause of parasite infection for the *T. polycolpus-L.burdigalensis* system and particularly emphasize the potential role of some parasite-associated bacteria in infection outcomes and progression.

## Introduction

Understanding the factors that shape host susceptibility to pathogens is of prime importance to anticipate and predict disease risks. For a long time, infection risk was thought to mainly result from complex interactions between parasites, hosts and environmental factors. For instance, infection can result from intrinsic factors related to the interplay between parasite virulence and host resistance, but at the same time is shaped by extrinsic environmental factors since host and parasite share a common environment (Wolinska and King 2009). Recent studies demonstrated that the occurrence of certain parasites was also associated to specific community of host microbiota, hence suggesting that the interaction between parasites, hosts and the microbiota may also be an important determinant of host pathology and disease progression (Reynolds et al. 2015, Bär et al. 2015, Aivelo and Norberg 2018, Portet et al. 2018). Differences in microbiota among potential host individuals –before infection– provide a novel mechanistic framework to understand and predict infection risk within host populations. More than a triple interaction between hosts, parasites and their common environment, parasitism could rather be seen as a multiple interaction including the role of host- and parasite-associated microbiota (Dheilly 2014, Bordenstein and Theis 2015).

A key question regarding the role of host-associated microbiota in host-parasite interactions relates to the role of the microbiota between parasites and their hosts, and in particular whether differences in host microbiota are a cause of infection or a consequence (Bär et al. 2015). On the one hand, it has been argued that difference in host microbiota *before* infection can cause the host susceptibility to parasite infection, making the host microbiota central to predict individual infection risks. For instance, host microbiota can increase host susceptibility by providing optimum condition for parasite development (Hayes et al. 2010, but see Buffie and Pamer 2013,

Britton and Young 2014 for the inverse). On the other hand, differences in host microbial communities can occur *after* parasite infection, making differences in host microbiota a consequence of parasite infection. For instance, parasites can induce changes in the relative abundance of certain microbial species that were already present within hosts before infection (Lee et al. 2014, Beatty et al. 2017). Parasites can also favor the colonization of new species of the microbiota coming from extrinsic sources, for instance because of a global decline in host resistance that facilitate colonization by opportunist pathogens from the surrounding environment (*Decreasing colonization resistance strategy*, Mooney et al. 2015). Changes in host microbiota after infection can also be due to the colonization of new species harbored by the parasites itself, leading to patterns of co-infection, which can ultimately increase the global pathogenic effects (*Weapon strategy*, Reynolds, Finlay, and Maizels 2015; Dheilly, Poulin, and Thomas 2015). Specific association between specific microbiota communities and parasites are common patterns in the wild (Andre 2002, Boissière et al. 2012, Stensvold and van der Giezen 2018), and a key contemporary challenge is to tease apart between the different strategies (cause or consequence) leading to these associations in wild populations.

A powerful approach to decipher whether host-associated microbiota is a cause or a consequence of infection, is to survey individual hosts for their microbiota longitudinally, that is before and after infection (Blanchet et al. 2009c, Jenkins et al. 2018, Afrin et al. 2019, Gaulke et al. 2019). Nonetheless, most longitudinal surveys implying changes in host microbiota over time have focused on changes occurring during or after parasite exposure, which strikingly limits our understanding of mechanisms sustaining relationships between parasites, hosts and the microbiota (but see Jenkins et al. 2018). Furthermore, in order to investigate the mechanisms leading to changes in host microbiota after parasite infection (change in abundance of already

existing bacteria within hosts, colonization of opportunistic bacteria from the environment, and colonization of bacteria coming from the parasite microbiota), both host-associated microbiota and microbiota from potential sources over the course of infection have to be investigated. Hence, tackling the question of host-associated microbiota as a cause or a consequence of infection is a challenging task but that could be achieved through a proper and original design. This challenge is obviously trickier in wild setting in which the level of control is lower, although this undoubtedly constitutes the only way to provide realistic and novel insights into the complex links between parasites, hosts, the environments and the microbiota.

In this study, we tested for variation in host-associated microbiota along the process of parasite infection, relying on a longitudinal and semi-natural experiment. We focused on a freshwater fish ectoparasite (*Tracheliastes polycolpus*) that has recently emerged in Western Europe and that infects several native fish species. The parasite anchored on fins where secondary infections by bacteria or fungi are often observed (Blanchet et al. 2009a). These secondary infections could be purely opportunistic and neutral for the parasite or on the contrary could be non-random and actually favor –either before or after infection– the parasite. To tackle this important issue, we tracked the microbiota of fin mucus of its main host species, the rostrum dace (*Leuciscus burdigalensis*), both before and after infection at the individual (fin) host level. We first tested the hypothesis that fin microbiota is a cause of *T. polycolpus* infection by comparing fin microbiota before parasite infection between hosts that will become infected and others that will remain healthy. We assumed that –if fin microbiota is a cause of infection– specific bacteria before infection should be associated (or counter-associated) to future infection by *T. polycolpus*. Secondly, we investigated the temporal dynamic of microbial communities during infection to test whether the infection by *T. polycolpus* resulted in changes in fin microbiota composition, i.e.,

whether fin microbiota is a consequence of infection. If so, we expect that, during infection, the microbiota composition of infected host will strikingly shift away from that of healthy host, i.e., that the temporal evolution of the fin microbiota will differ between infected and healthy hosts. Finally –and assuming that fin microbiota changes is a consequence of infection–, we will test whether these changes after infection occurred because of the colonisation of opportunistic bacteria from the environment (*Opportunistic colonisation hypothesis*) and/or because of the colonisation of bacteria associated with the parasite (*Co-infection hypothesis*). We expect a mixture of these two hypotheses to sustain potential changes in fin microbiota after infection.

## **Material and Method**

### **Biological history of the interaction**

In this study we considered the interaction between the ectoparasite of freshwater fish *T. polycolpus* and its main host in southwestern France the cyprinid *L. burdigalensis* (the rostrum dace). *Tracheliastes polycolpus* is a parasitic copepod that anchors to fish fin (and occasionally to their body) causing fin degradation and secondary inflammation due to infections by bacteria or fungi (Blanchet et al. 2009a). Only females are parasitic whereas males are microscopic free living forms. Once anchored on fish fin the females *T. polycolpus* develop two egg sacs containing up to one hundred of eggs (Mazé-Guilmo 2016). After few weeks eggs hatch and release each one a single free living larvae that go through two larvae stages (*nauplius* stage and then *copepodite* stage) before reaching the adult stage (Fryer 1982). Interestingly larvae are not in direct contact with the environment until hatching and hence the microbiota of *T. polycolpus* larvae likely reflects their own microbiota probably transmitted directly from the mother, which makes this model particularly powerful to study parasite associated microbiota and its role during parasitic infection of hosts.

## **Microbiota survey of fish fins and environment**

We used a semi-natural caging experiment to allow for an individual “before-after infection” survey of fin microbiota of potential hosts. Twenty cages (30x50x80 cm, and mesh of 1 cm), each containing four rostrum dace (80 dace on total at the onset of the experiment) were set up along a 200-m long section in the Arize River in southwestern France (X= 366422, Y= 4770813) during summer 2018. Dace were F2 individuals from wild ancestors bred at the INRA hatchery facility in Rennes (IE3, France). All experimental individuals were of the same age (2+ fish) and genetic background, and were of similar in weight and body size ( $15\pm 3$  g and  $110\pm 8$  mm) so as to limit confounding factors from the host status. They were brought back from Rennes in early spring, raised in large outdoor tanks at the SETE laboratory (Moulis, France) and fed *ad libitum* before the onset of the experiment. None of them have previously been exposed to *T. polycolpus*. Fish were introduced in cages in early summer (9<sup>th</sup> of July 2018). Once every two days, cages were brushed (to avoid mesh clogging) and daces were visually inspected for potential infections. After ~3 weeks, dace were collected from their cages and a sampling of fin mucus was performed to characterize fin microbiota *before infection*. This amount of time ensured their fin microbiota shift toward to the local microbiota communities, while infection by *T. polycolpus* did not occur (Schmidt et al. 2015, Mohammed and Arias 2015, Lokesh and Kiron 2016 and Chapter I). All individuals were anesthetized (using benzocaine) and a swab was sampled on each fin (anal, caudal, dorsal, pelvic and pectoral) and on the body. Fish were also weighted, measured and individually tagged using colored elastomers injected at fin insertions. Fish were then gently replaced in their cage. From that period, cages were checked daily for infection, and a second swab sampling was performed on surviving daces (68 dace including 21 infected dace) one to three weeks after the first infections were observed (i.e. about one month and a half after fish were introduced into the cages). Fish were collected from their cages, measured, weighted,

closely checked for infection, and a swab was sampled on each fin and on the body (whether or not they were parasitized) to characterize fin microbiota communities *after infection*. It is noteworthy that fin microbiota from fish that remain healthy over time were also characterized, and were used as “temporal control”. When a parasite was found (34 parasites in total), it was gently removed and stored in a 50-ml vial containing water from the river for future inspection of the microbiota of the parasite and of the infecting larvae (see hereafter) as they can be sources of bacteria colonization during infection. We also collected a water sample at each cage to characterize the environmental microbiota that can be a source of opportunistic colonization after infection. All swab samples were stored at -40°C until extraction.

To sum up, the microbiota was sampled on each fin (and the body) before and after infection occurred, and four main fin categories can be defined according to their status *after infection*: (i) microbiota from infected fins (n = 19–22 fins from 14–16 fish), (ii) microbiota from healthy fins sampled on an infected fish (n = 82–95 fins from 18–20 fish), (iii) microbiota from healthy fins sampled on healthy fish in a cage where at least one infected fish was observed (n = 132–154 fins from 21–24 fish), and (iv) microbiota from healthy fins sampled on healthy fish in cages where no infected fish was observed (n = 60–71 fins from 12 fish). These four categories were surveyed for their microbiota before and after infection (hence leading to a total of eight groups).

### **Microbiota of parasites**

The microbiota of all parasites was sampled by sequencing bacteria found in the trunk of adult parasite and in reared parasite larvae. When parasites carried mature egg sacs (7 parasites, corresponding to 14 egg sacs), they were brought alive to the laboratory in order to obtain parasite larvae. The two egg sacs were separated from each female parasite using sterile scalpels. Trunks of adult parasites were stored at -40°C until extraction and egg sacs were placed in single

2mL well from a sterile 24-well plate (see Mazé-Guilmo et al. 2016 for details). Wells were filled with 2 mL natural mineral water (changed every day) until hatching. Well plates were stored in a thermostat incubator at a temperature close to the one measured *in natura* during sampling ( $16^{\circ}\text{C} \pm 1^{\circ}\text{C}$ ). The thermostat incubator was enlightened every day from 8 am to 8 pm to approximate natural hatching condition. Once egg sacs hatched, larvae were carefully collected, stored in 2mL Eppendorf® and kept at  $-40^{\circ}\text{C}$  until extraction (~225 larvae, pooled by egg sac). For each hatching day, a swab sample of mineral water unexposed to the parasites was also collected.

### **DNA extraction and 16S rRNA gene amplicon sequencing**

We followed the same DNA extraction and amplification protocol as Bestion et al. (2017). Briefly, bacteria DNA of each samples was extracted using a standard protocol for the purification of total DNA from gram-positive bacteria with the Qiagen DNeasy Blood & Tissue Kit (Qiagen GmbH, Hilden, Germany). After extraction, the concentration of the resulting genomic DNA was 10-20 ng/ $\mu\text{l}$ . DNA was then amplified using PCR realized in a final volume of 20  $\mu\text{L}$  including 2  $\mu\text{L}$  of DNA, 10  $\mu\text{L}$  of AmpliTaq Gold™ 360 Master Mix (ThermoFisher Scientific), 0.5  $\mu\text{M}$  of each primer and 3,2  $\mu\text{g}$  of bovine serum albumin. Initial denaturation was realized at  $95^{\circ}\text{C}$  for 10 minutes followed by 30 cycles of denaturation ( $95^{\circ}\text{C}$  for 30 s), annealing ( $57^{\circ}\text{C}$  for 30 s) and elongation ( $72^{\circ}\text{C}$  for 90 s) and finally followed by a 7 min final elongation step at  $72^{\circ}\text{C}$ . We specifically amplified the V5-V6 region of the 16S rRNA gene (295 bp) using a pair of universal bacteria primers (BACTB-F: GGATTAGATACCCTGGTAGT; and BACTB-R: CACGACACGAGCTGACG, Fliegerova et al. 2014). Sample identification was achieved using unique combination of 8-base tags added in 5' end of each primer. PCR products were purified (QIAquick PCR purification Kit, Qiagen GmbH, Hilden, Germany) and pooled. A single library of the amplicon multiplex was prepared accordingly to the MetaFast protocol. Finally sequencing



was performed at the GeT-PlaGe facility (Toulouse, France) using the sequencer Illumina MiSeq and 2x250 bp pair-end sequencing technology. Importantly, positive controls (Zymobiomics Microbial Community DNA Stan) and negative controls (blanks, and air samples realized during the sampling campaign and during extraction) were included in the sequenced multiplex in order to investigate extraction and amplification success as well as marker universality, and potential contamination, respectively. Overall, 842 samples were sequenced (see Table S1 for an overview of the different type of samples and their number).

### **Bioinformatic analysis and filtering**

Illumina sequencing data were first demultiplexed using the OBITools suite of scripts (Boyer et al. 2016). Once one file of assembled reads per sample constituted, data were processed using FROGS (Find Rapidly OTUs with Galaxy Solution, Escudié et al. 2017). Briefly, FROGS first include one step of pre-processing that de-replicates and filters amplicons according to their length. We set minimum amplicon size to 200 and maximum amplicon size to 350 so as to target amplicons of size expected for the V5-V6 region (i.e., 295 bp length) of the bacterial 16S rRNA gene. Clustering in Operation Taxonomic Units (OTU) was then realized using the swarm algorithm implemented in FROGS with aggregation distance set to three (Mahé et al. 2014). Chimera and singletons were removed. Finally, the taxonomic affiliation was performed per OTU with blast against Silva database (v132, Quast et al. 2012). At the end of FROGS processing, our OTUs dataset included 38 584 OTUs covering the 842 samples. Finally, using the R package *Phyloseq* (McMurdie and Holmes 2013), OTU composition of each sample was rarefied at 2 000 sequences per OTUs to allow samples comparisons and hence discarding 86 samples from further analysis due to too low OTUs abundances (including all the negative and blank controls). Finally,

we filtered out OTUs being present in less than 3 samples and being under a frequency of 1%. The final dataset hence resulted in 4 987 OTUs covering 738 samples.

## **Statistical analysis**

### ***Hypothesis 1: Fin microbiota as a cause of parasite infection***

In order to test whether differences in microbiota associated with fin mucus existed prior to parasite infection (fin microbiota as a cause of infection), we focused on fin microbiota sampled before infection, and we discarded samples from cages in which all fish remained healthy to avoid confounding effects due to differences in microhabitat (i.e., cages in which all fish remained healthy could be in particular habitat with particular microbiota, see Chapter I). We expected that the microbiota of fins that will be infected should be different from the microbiota of fins that remain healthy (differences occur at the fin level), or that the microbiota of fins from an infected host (regardless whether all fins were indeed infected) should be different from the microbiota of fins from a healthy host (differences occur at the host level). We used a random forest classifier model to test these predictions.

We used the R package *randomForest* (Liaw and Wiener 2002) to build two specific models. In the first model, we considered only fish that will be parasitized, and we compared the microbiota of fins that will become infected to those of fins that will remain healthy. In the second model, we considered –within a specific cage– both fish that will become parasitized and those that will remain healthy, and we compared the microbiota of fins from a fish that will be infected to that of fins from a fish that will remain healthy. In the first model we expect difference at the *fin level* (within infected fish), whereas in the second model we expect differences at the *individual fish level* (between fish that will or not be infected). In the model testing for microbiota differences at

the fin level, we set the tree number to 4000 and best mtry value was estimated to 35 (assessed using the *tuneRF* function). In the model testing for microbiota differences at the individual fish level, we set tree number to 4000 and best mtry value was estimated to 70. Since each fin category had heterogeneous number of samples, we constrained in each random forest model the number of samples used to achieve classification to two third of the minimum sample size (i.e., 11), and the remaining third was used to estimate error rate. Each time the model aimed to classify samples in two classes (i.e., infected and healthy) based on OTU relative abundances and difference extent was estimated base on model Out Of Bag (OOB) error. Finally, microbiota differences were graphically illustrated using the first axis of a Discriminant Analysis of Principal Components (DAPC).

### ***Hypothesis 2: Fin microbiota composition as a consequence of parasite infection***

In order to test whether parasite infection lead to changes in fin microbiota, we investigated the temporal dynamics of fin microbial communities during infection. We expected that temporal shifts (from *before* to *after* infection) in fin microbiota should be different between infected and healthy fins, and that these shifts should be stronger in magnitude in infected- than in healthy fins. To test these predictions, we considered both before and after infection samples, and healthy fins from fish that were infected, from fish that remained healthy in cages with at least one infected fish and from fish that remained healthy in cages where all fish remained healthy (categories (ii), (iii) and (iv) as described above). Considering these three categories of healthy fins allowed investigating several temporal “controls”. We conducted a Community Trajectory Analysis (CTA, De Cáceres et al. 2019) based on a Discriminant Analysis of Principal Component ran on OTU relative abundances using the R package *adeget* (Jombart 2008). The four categories of fins microbiota samples described above (infected fins and the three healthy

fins categories), both before and after infection were used as grouping variables. Trajectories through time for each fin category were then drawn using the R package *vegclust* (De Caceres and De Caceres 2010) and using the DAPC as projection plan. For each trajectory we extracted their length and computed pairwise Euclidean distances between trajectories. We tested differences in trajectory length using a Kruskal-Wallis test combined with a post-hoc test correcting for false discovery rate. Similarly, we tested for differences in trajectory direction (relying on a comparison of pairwise distances between trajectories) using a manova followed by a pairwise permutation manova (10 000 permutations) and correcting for false discovery rate using the R packages *vegan* and *RVaideMemoire* (Oksanen et al. 2011, Hervé 2019).

### ***Potential source for host microbiota changes after infection***

First we aimed to identify OTUs that most discriminated healthy and infected fins after infection (i.e., healthy fins from fish in a cage where no infection occurred, and infected fins, in order to compare the two most contrasted groups). To do so we built a random forest model that classified healthy fins and infected fins according to their respective OTU relative abundances. We set the number of tree to 4 000 and estimated the best *mtry* value to 1 120. As above, we constrained in the random forest model the number of samples used to achieve classification to two third of the minimum sample size (i.e., 14) and the remaining third was used to estimate OOB error. We then assessed OTUs importance to the random forest classifier using the *rfPermute* R package and computed p-values on a 1 000 permutations base (Archer 2019). Significant OTUs (i.e., candidate OTU hereafter) were selected according to the mean decrease of Gini index and at a p-value threshold of 5%. We further evaluated OTUs significance by building a second random forest model (number of tree=4 000 and *mtry*=52) relying only on candidate OTUs in order to compare

its OOB error to the first random forest including all OTUs. We kept this later model for further analysis.

Secondly we tested for potential source for OTUs characterizing healthy and infected fins associated microbiota, namely environmental bacteria (i.e., collected from water samples after infections), bacteria associated to adult parasites and bacteria associated to parasite larvae. We classified samples belonging to these three potential sources according to their relative abundances observed on the set of candidate OTUs using the random forest model previously established on candidate OTUs. If a source (environmental, adult parasite or parasite larvae) significantly contributed to either the healthy or infected fin microbiota, we expected that the random forest model will classify these source samples more often than randomness to the same class as either healthy or infected fins (meaning they are similar in term of microbiota composition). We performed a Chi-squared test to test whether these sources were randomly attributed either to the same class than healthy or infected fins respectively by the random forest model. We visualized differences between random expectation and observed attribution from the random forest model using barplot.

Finally, in order to investigate the taxonomic composition of candidate OTUs, we represented their relative abundances per bacteria family over the whole community for the different potential source microbiota as well as for healthy fin and infected fins.

## **Results**

### **Hypothesis 1: Fin microbiota as a cause of parasite infection**

Random forest classifiers assuming microbiota differences either at the fin or at the host levels achieved a low assignment success (OOB error rate, 40% and 66% respectively, Table 1). These

results strongly suggest that there is no significant difference in host-associated microbiota before infection between fins that will or not become infected. Based on the models, it is indeed almost impossible to distinguish microbial communities from fins that will be infected from microbial communities from fins that will remain healthy (model 1), or microbial communities from fins of a fish that will be infected from microbial communities from fins belonging to a fish that will remain healthy (model 2) (Table 1, Fig.1).

**Table 1:** Table summarizing the error rate per class and the OOB error for the two random forest models testing microbiota differences before exposure at the fin and the host levels.

Hierarchical level of infection	Contrasted groups	Class error	OOB error
<b>Model 1 ("fin" level)</b>	Infected fins from infected fish	0.653	0.663
	Healthy fins from infected fish	0.706	
<b>Model 2 ("fish" level)</b>	Fins from infected fish	0.310	0.399
	Fins from healthy fish	0.511	

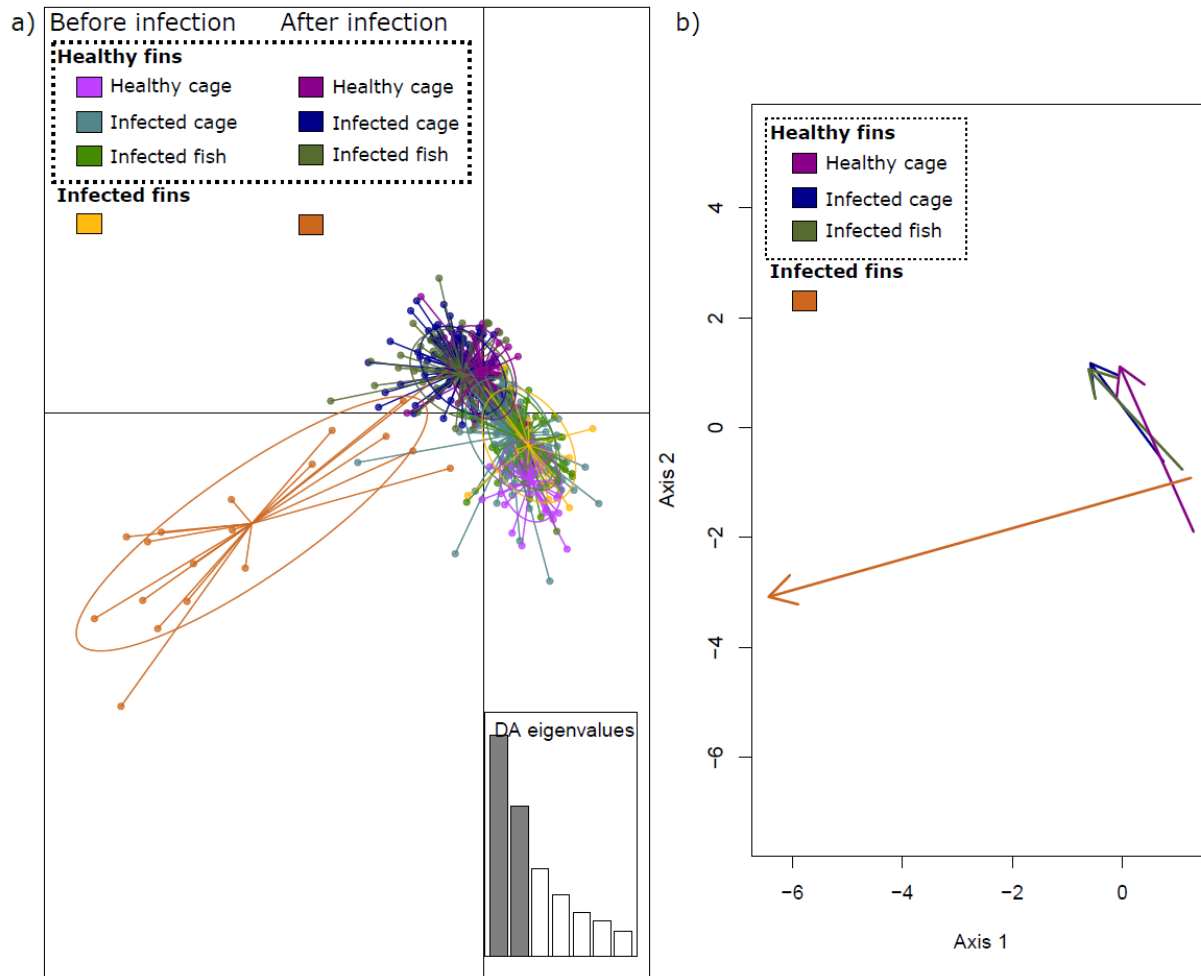


**Figure 1:** Violin plot showing for microbiota collected on fins before infection the corresponding scores on the first axis of a DAPC. NS stands for non-significant differences in microbiota based on random forest models.

## **Hypothesis 2: Host microbiota composition as a consequence of parasite infection**

We found significant differences between microbiota trajectories associated with infected fins and microbiota associated with the different categories of healthy fins, in term of trajectory length (Kruskal-Wallis chi-squared = 61.067, df = 3, p-value <0.001) and in term of trajectory direction (Manova F=115.6; df=239,3; p-value=0.001). Specifically, we found that changes –from *before* to *after* infection– in microbial communities of infected fins were significantly stronger in magnitude (i.e., difference in term of trajectory lengths) than changes measured in all other healthy fin categories (Fig. 2, Table S2). In the same way, the trajectory direction observed for the microbial communities of infected fins was significantly different from that of all other categories of healthy fins (Fig. 2, Table S2). Overall these results suggest that parasite infection strikingly modify the microbial communities of fins on which it anchored, and that these changes are different from those that naturally occur over time in the wild (Fig 2).



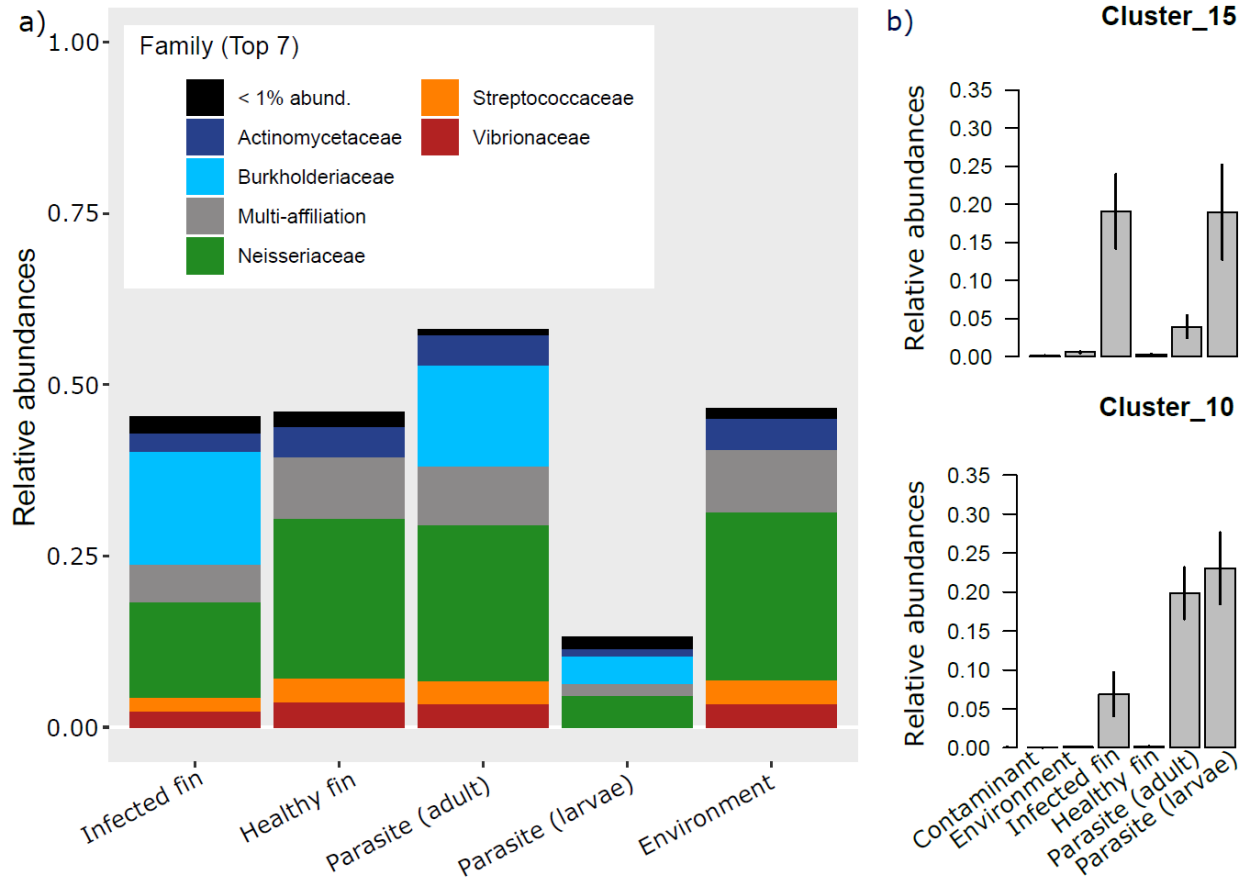


**Figure 2:** a) Scatter plot from the DAPC realized on OTU relative abundances and used as projection plan to draw microbiota trajectories. b) Mean trajectory over the course of infection for microbiota associated with infected fins and healthy fins (collected either from an infected fish, from healthy fish in an infected cage or from a healthy fish in a healthy cage).

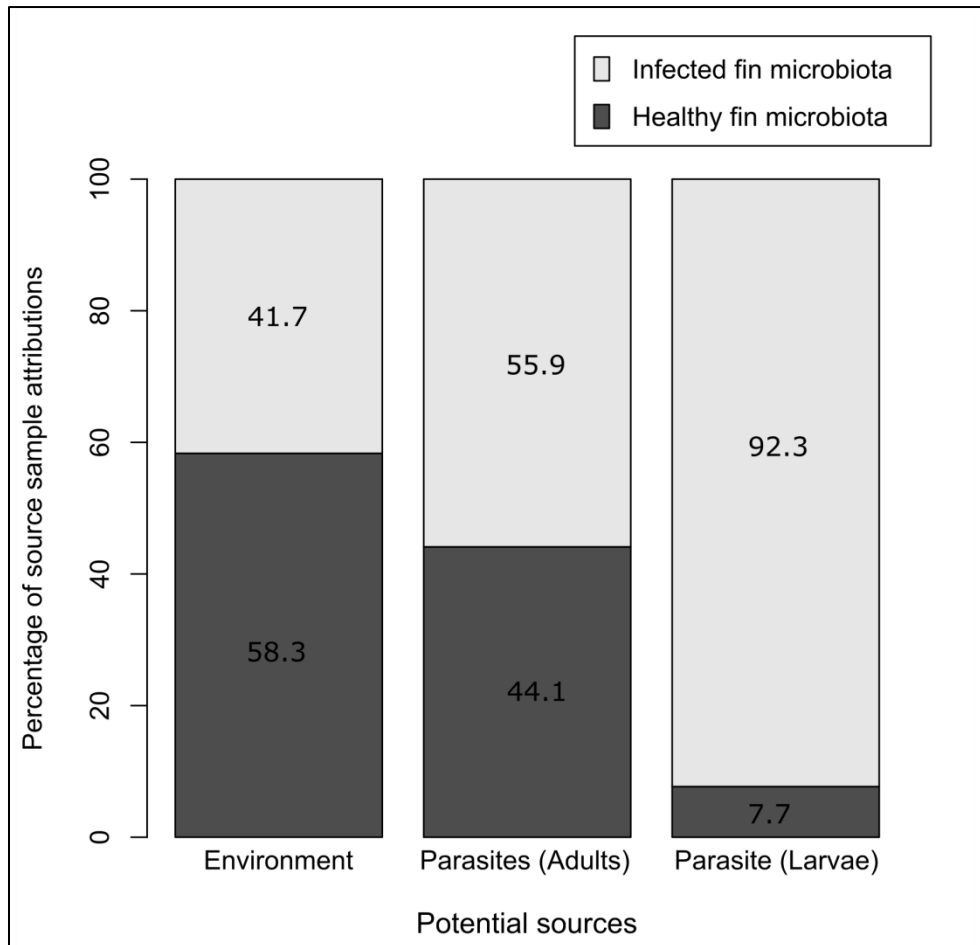
### **Potential source for host microbiota changes after infection**

The first random model classifier we built achieved a very high assignment success (OOB was 3.33% and class errors were 2.89% and 4.76% for healthy and infected fins respectively). Up to 172 OTUs significantly contributed to the decrease of Gini index, and hence to the purity of the random forest model (Table S3). The significance of these candidate OTUs was further validated by the second random forest model relying only on these 172 candidate OTUs, since the model achieved an even lower OOB error (OOB was 2.22% and class errors were 1.45% and 4.76% for healthy and infected fins respectively).

When considering candidate OTUs, we found that the *Burkholderiaceae* family was abundant in both infected fin associated microbiota and parasite associated microbiota (16%, 15% and 4% in relative abundance in infected fins, adult parasites and parasite larvae respectively), whereas this family was rare (0.6% and 0.4% in relative abundance in the environment and healthy fins respectively) in other samples (Figure 3-a). We further investigated the similarity of potential source of bacteria (i.e., environmental microbiota and/or parasite associated microbiota including adults and larvae) with either healthy fin or with infected fins using the random forest classifier based on candidate OTUs. Attributions resulting from the random forest classifier were significantly different from randomness (Chi-squared=8.956, df=2, p-value=0.011). Specifically, we found that microbiota associated with parasite larvae was classified more often than chance in the same class as infected fins (Chi-squared = 9.308, df = 1, p-value = 0.002, Fig. 4). The two OTUs that mostly contributed to the purity of the random forest classifier indeed showed higher relative abundances in infected fins and in parasites larvae and both belonged to the *Burkholderiaceae* family (Fig. 3-b, Table S3).



**Figure 3:** a) Microbiota composition, associated with infected fins, healthy fins, adult parasite, parasite larvae and surrounding environment. Microbiota composition is drawn at the family level and for candidates OTUs significantly discriminating healthy fin associated microbiota from infected fin associated microbiota. Relative abundances over the total community are displayed. b) Relative abundances (over candidate OTUs) of the two OTUs that had the most significant effect in term of decrease of Gini index in the random forest classifying infected fin microbiota and healthy fin microbiota.



**Figure 4:** Barplot showing the percentage of samples belonging to each potential source (environment, adult parasite and parasite larvae) attributed either to the same class as microbiota associated with healthy fin or as microbiota associated with infected fin. Attributions are based on the random forest model built on candidate OTUs and classifying microbiota associated with healthy or infected fins.

## Discussion

In this study we aimed to test whether host-associated microbiota was a cause or a consequence of parasite infection. We relied on a semi-natural experimental approach to monitor host-associated microbiota over the course of infection. We showed that for the system *T. polycolpus-L. burdigalensis* differences in fin associated microbiota did not exist prior to infection, neither at the fin nor at the host level. On the contrary, parasitic infections triggered important changes in

host-associated microbiota at the parasite anchoring site level (i.e., fin level). Furthermore, we were able to demonstrate that the microbiota of infected fins was similar to the microbiota of parasite larvae, hence suggesting that co-infection dynamics between parasite and their bacteria could play an important role during the process of infection.

We demonstrated that differences in host-associated microbiota did not exist prior to infection. Some studies suggested that specific host-associated microbiota prior to infection could favor parasitic infection, for instance by providing suitable conditions for parasite development (Hayes et al. 2010) or by impacting host resistance (Buffie and Pamer 2013, Britton and Young 2014). However, in our case our random model classifiers were unable to properly distinguish microbiota from either fin or host before exposure depending on whether or not they will be infected later on. Since we have done comparisons within cages in which at least one fish experienced parasitic infection over time, environmental conditions were controlled and homogeneous at the cage level, hence limiting confounding environmental factors. Several studies have suggested a link between environmental conditions or individual diet and their associated microbiota (Schwab et al. 2011, Yun et al. 2014, Bergmann et al. 2015). It is thus likely that microbiota was conditioned by the surrounding environment as defined by the cage extent hence resulting in similar microbiota among all individuals from a cage and without this microbiota specifically benefits future *T. polycolpus* infection. We consequently suggest that host-associated microbiota is not a direct cause of future infection by *T. polycolpus* and thus *T. polycolpus* infections are not favoured by pre-established host microbiota.

On the contrary, we showed that important changes in microbiota communities occurred consequently to parasite infection at the fin level. Our Community Trajectory Analysis indeed underlined significant microbial community changes through the infection course, notably at the

fin level that is the anchoring site of the parasite. In the studies that have monitored host microbiota over infection course, changes of host microbiota over time is a typical finding (Afrin et al. 2019, Gaulke et al. 2019). Such microbiota changes could result from different strategies notably the *decreasing colonization resistance strategy* or the *weapon strategy* (Swe et al. 2014, Mooney et al. 2015, Reynolds et al. 2015). First, *T. polycolpus* could have weakened dace for instance by triggering and/or disrupting host immune response (Lochmiller and Deerenberg 2000, Sandland and Minchella 2003) which could have eased colonization by opportunistic bacteria from the environment (i.e., *Decreasing colonization strategy*) (Mooney et al. 2015). In this case, if immune response decreases or is too costly in energetic demand, the general state of the host might be impacted and bacteria colonization should be observed at all infected host fins. Alternatively, *T. polycolpus* could have disrupted fin microbiota by feeding on fin mucus, and notably while damaging on host tissues co-infection with *T. polyvolpus* bacteria could have occurred (i.e. *Weapon strategy*) (Fryer 1982, Loot et al. 2004). In this case, local changes are expected at the parasite anchoring site level (i.e., the infected fin only). Since, we observed that parasite infection drove the composition of host microbiota in a singular direction at the infected fin level rather than at the host level, we expect that host microbiota changes observed during *T. polycolpus* infection would more likely follow the *weapon strategy* rather than *decreasing colonization resistance strategy*.

We went further by showing that parasite larvae is the most likely source for the bacteria found on parasitized fins. When investigating the potential sources of host-associated microbiota after infection we indeed showed that infected fin displayed a microbiota similar to the microbiota associated with parasite larvae. Such similarities between parasite microbiota and infected fin microbiota could indicate a potential co-infection dynamics between parasite and its associated

microbiota (Dheilly et al. 2015, Reynolds et al. 2015). An alternative hypothesis could be that it is actually the bacteria on parasitized fins that colonize the parasite during infection hence explaining the similarity between the two microbiota. However, interestingly, we showed that parasites larvae but not adult parasites had a similar microbiota to those of parasitized fin. Since parasite larvae were reared in a different and control environment and without being in contact with host tissue, it is thus likely that parasite associated bacteria colonized infected fins rather than the opposite. A taxonomic investigation of the most abundant bacteria families of candidate OTU discriminating healthy and infected fin microbiota further revealed that the *Burkholderiaceae* family was quite abundant in both infected fin microbiota and parasite (adult and larvae) microbiota but not in healthy fin and environmental microbiota (Fig. 3). We consequently suggest that *T. polycolpus* likely rely on a *weapon strategy* (Reynolds et al. 2015) during infection process with notably a critical role of *Burkholderiaceae* bacteria. Future works aiming to identify candidate bacteria species that are likely facilitating *T. polycolpus* infection would nevertheless be of prime interest to better understand the role of parasite associated microbiota in infection process.

To conclude we conducted one of the first studies providing an individual survey of host microbiota over the course of parasitic infection in natural conditions. We showed that for the system *T. polycolpus*-*L. burdigalensis* host microbiota composition was rather a consequence than a cause of parasitic infection. We notably identified important host microbiota changes at the fin level triggered by parasite infection and following a possible co-infection dynamics between parasite and its associated bacteria. We thus suggest that a complex interaction between host, parasite, environment and respective microbiota is shaping the infection process. We also argue for further investigations of these complex interactions in natural setting in order to

improve our knowledge of relative role of parasite-, host-, environment- and microbiota- related factors in parasite infection process.

## References

- Afrin, T. et al. 2019. Sequential Changes in the Host Gut Microbiota During Infection With the Intestinal Parasitic Nematode *Strongyloides venezuelensis*. - Front. Cell. Infect. Microbiol. 9: 217–228.
- Aivelo, T. and Norberg, A. 2018. Parasite-microbiota interactions potentially affect intestinal communities in wild mammals. - J Anim Ecol 87: 438–447.
- Andre, A. v. S. 2002. The Role of Endosymbiotic Wolbachia Bacteria in the Pathogenesis of River Blindness. - Science 295: 1892–1895.
- Archer, E. 2019. Package ‘rfPermute’.
- Bär, A.-K. et al. 2015. The Interplay of Host Microbiota and Parasitic Protozoans at Mucosal Interfaces: Implications for the Outcomes of Infections and Diseases. - PLoS Negl Trop Dis 9: e0004176.
- Beatty, J. K. et al. 2017. *Giardia duodenalis* induces pathogenic dysbiosis of human intestinal microbiota biofilms. - International Journal for Parasitology 47: 311–326.
- Bergmann, G. T. et al. 2015. Seasonal Shifts in Diet and Gut Microbiota of the American Bison (*Bison bison*). - PLoS ONE 10: e0142409.
- Bestion, E. et al. 2017. Climate warming reduces gut microbiota diversity in a vertebrate ectotherm. - Nat Ecol Evol 1: 0161.
- Blanchet, S. et al. 2009a. Evidence of parasite-mediated disruptive selection on genetic diversity in a wild fish population. - Molecular Ecology 18: 1112–1123.
- Blanchet, S. et al. 2009b. Reciprocal effects between host phenotype and pathogens: new insights from an old problem. - Trends in Parasitology 25: 364–369.
- Boissière, A. et al. 2012. Midgut Microbiota of the Malaria Mosquito Vector *Anopheles gambiae* and Interactions with *Plasmodium falciparum* Infection. - PLoS Pathog 8: e1002742.
- Bordenstein, S. R. and Theis, K. R. 2015. Host Biology in Light of the Microbiome: Ten Principles of Holobionts and Hologenomes. - PLOS Biology: 23 e1002226.
- Boyer, F. et al. 2016. OBITOOLS : a UNIX -inspired software package for DNA metabarcoding. - Mol Ecol Resour 16: 176–182.
- Britton, R. A. and Young, V. B. 2014. Role of the Intestinal Microbiota in Resistance to Colonization by *Clostridium difficile*. - Gastroenterology 146: 1547–1553.



- Buffie, C. G. and Pamer, E. G. 2013. Microbiota-mediated colonization resistance against intestinal pathogens. - *Nat Rev Immunol* 13: 790–801.
- De Cáceres, M. and De Cáceres, M. M. 2010. Package ‘vegclust’.
- De Cáceres, M. et al. 2019. Trajectory analysis in community ecology. - *Ecol Monogr* 89: e01350.
- Dheilly, N. M. 2014. Holobiont–Holobiont Interactions: Redefining Host–Parasite Interactions. - *PLoS Pathog* 10: e1004093.
- Dheilly, N. M. et al. 2015. Biological warfare: Microorganisms as drivers of host–parasite interactions. - *Infection, Genetics and Evolution* 34: 251–259.
- Escudié, F. et al. 2017. FROGS: Find Rapidly OTU with Galaxy Solution. - *Bioinformatics* 34: 1287–1294.
- Fliegerova, K. et al. 2014. Effect of DNA extraction and sample preservation method on rumen bacterial population. - *Anaerobe* 29: 80–84.
- Fryer, G. 1982. The parasitic *copepoda* and *branchiura* of British freshwater fishes. A handbook and key. - Freshwater Biological Association.
- Gaulke, C. A. et al. 2019. A longitudinal assessment of host-microbe-parasite interactions resolves the zebrafish gut microbiome’s link to *Pseudocapillaria tomentosa* infection and pathology. - *Microbiome* 7: 10–26.
- Hayes, K. S. et al. 2010. Exploitation of the Intestinal Microflora by the Parasitic Nematode *Trichuris muris*. - *Science* 328: 1391–1394.
- Hervé, M. 2019. Package ‘RVAideMemoire’.
- Jenkins, T. P. et al. 2018. *Schistosoma mansoni* infection is associated with quantitative and qualitative modifications of the mammalian intestinal microbiota. - *Sci Rep* 8: 12072.
- Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. - *Bioinformatics* 24: 1403–1405.
- Lee, S. C. et al. 2014. Helminth Colonization Is Associated with Increased Diversity of the Gut Microbiota. - *PLoS Negl Trop Dis* 8: e2880.
- Liaw, A. and Wiener, M. 2002. Classification and Regression by randomForest. 2: 18–22.
- Lochmiller, R. L. and Deerenberg, C. 2000. Trade-offs in evolutionary immunology: just what is the cost of immunity? - *Oikos* 88: 87–98.
- Lokesh, J. and Kiron, V. 2016. Transition from freshwater to seawater reshapes the skin-associated microbiota of Atlantic salmon. - *Sci Rep* 6: 19707.

- Loot, G. et al. 2004. The effects of the ectoparasite *Tracheliastes polycolpus* (Copepoda: Lernaepodidae) on the fins of rostrum dace (*Leuciscus leuciscus burdigalensis*). - Parasitology Research 94: 16–23.
- Mahé, F. et al. 2014. Swarm: robust and fast clustering method for amplicon-based studies. - PeerJ 2: e593.
- Mazé-Guilmo, E. 2016. Etude du potentiel de colonisation des parasites : une approche intégrative.
- McMurdie, P. J. and Holmes, S. 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. - PLoS one 8: e61217.
- Mohammed, H. H. and Arias, C. R. 2015. Potassium permanganate elicits a shift of the external fish microbiome and increases host susceptibility to columnaris disease. - Vet Res 46: 82–95.
- Mooney, J. P. et al. 2015. Inflammation-associated alterations to the intestinal microbiota reduce colonization resistance against non-typhoidal *Salmonella* during concurrent malaria parasite infection. - Scientific Reports 5: 14603.
- Oksanen, J. et al. 2011. vegan: Community Ecology Package. - R Foundation for Statistical Computing, Vienna. Available: CRAN. R-project. org/package= vegan.
- Portet, A. et al. 2018. Infection of the *Biomphalaria glabrata* vector snail by *Schistosoma mansoni* parasites drives snail microbiota dysbiosis. – bioRxiv: 386623.
- Quast, C. et al. 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. - Nucleic Acids Research 41: 590–596.
- Reynolds, L. A. et al. 2015. Cohabitation in the Intestine: Interactions among Helminth Parasites, Bacterial Microbiota, and Host Immunity. - J.I. 195: 4059–4066.
- Sandland, G. J. and Minchella, D. J. 2003. Costs of immune defense: an enigma wrapped in an environmental cloak? - Trends in Parasitology 19: 571–574.
- Schmidt, V. T. et al. 2015. Community assembly of a euryhaline fish microbiome during salinity acclimation. - Mol Ecol 24: 2537–2550.
- Schwab, C. et al. 2011. Diet and Environment Shape Fecal Bacterial Microbiota Composition and Enteric Pathogen Load of Grizzly Bears. - PLoS ONE 6: e27905.
- Stensvold, C. R. and van der Giezen, M. 2018. Associations between Gut Microbiota and Common Luminal Intestinal Parasites. - Trends in Parasitology 34: 369–377.
- Swe, P. M. et al. 2014. Parasitic scabies mites and associated bacteria joining forces against host complement defence. - Parasite Immunol 36: 585–593.
- Wolinska, J. and King, K. C. 2009. Environment can alter selection in host–parasite interactions. - Trends in Parasitology 25: 236–244.

Yun, J.-H. et al. 2014. Insect Gut Bacterial Diversity Determined by Environmental Habitat, Diet, Developmental Stage, and Phylogeny of Host. - *Appl. Environ. Microbiol.* 80: 5254–5264.

## Supplementary

**Table S1:** Number of samples per sample type.

Sample type	Control	Post-infection	Pre-infection	Replicates	Total
Water (parasite rearing environment)	6				6
Water (natural environment)		24	18		42
Healthy fins (from a healthy fish in an infected cage)		144	125	17	286
Healthy fins (from an infected fish)		89	79	9	177
Infected fins		21	19	1	41
Parasites (larvae)		13		2	15
Parasites (adults)		34		7	41
Negative control	18				18
Mock communities	9				9
Healthy fins (from a healthy fish in a healthy cage)		71	54	6	131
Extraction control	35				35
Experimenter	2				2
Blank	29				29
Air (parasite rearing environment)	1				1
Air (during sampling campaign)	9				9

**Table S2:** P-values for the pairwise comparisons between microbial community trajectories during parasitic infection in term of trajectory length and of pairwise distances

<b>Trajectory length</b>			
	Healthy fin From a healthy fish In a healthy cage	Healthy fin From a healthy fish In an infected cage	Healthy fin From an infected fish
Healthy fin From a healthy fish In an infected cage	<0.001		
Healthy fin from an infected fish	<0.001	0.169	
Infected fin	<0.001	<0.001	<0.001
<b>Trajectory pairwise distances</b>			
	Healthy fin From a healthy fish In a healthy cage	Healthy fin From a healthy fish In an infected cage	Healthy fin From an infected fish
Healthy fin From a healthy fish In an infected cage	<0.001	-	-
Healthy fin From an infected fish	<0.001	0.168	-
Infected fin	<0.001	<0.001	<0.001

**Table S3:** Candidates OTUs from the random forest model aiming at classifying healthy fin associated microbiota and infected fin associated microbiota. For each candidate is displayed the mean decrease of Gini index, associated p-value, and taxonomic affiliation.

OTUs	Mean Decrease Gini	p-value	Kingdom	Phylum	Class	Order	Family	Genus	Species
Cluster_10	1.041	0.001	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacterales	Burkholderiaceae	Ideonella	Multi-affiliation
Cluster_15	3.609	0.001	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacterales	Burkholderiaceae	Multi-affiliation	Multi-affiliation
Cluster_15_37	0.004	0.001	Bacteria	Armatimonadetes	Armatimonadetes	Armatimonadales	Armatimonadaceae	Armatimonas	unknown species
Cluster_17_07	0.014	0.001	Bacteria	Acidobacteria	Subgroup 17	unknown order	unknown family	unknown genus	unknown species
Cluster_17_9	0.014	0.001	Bacteria	Actinobacteria	Actinobacteria	Pseudonocardiales	Pseudonocardiaceae	Alloactinosynnema	Actinosynnemataceae bacterium R19-10
Cluster_18_20	0.002	0.001	Bacteria	Bacteroidetes	Bacteroidia	Chitinophagales	Chitinophagaceae	unknown genus	unknown species
Cluster_21_69	0.015	0.001	Bacteria	Planctomycetes	Planctomycetia	Pirellulales	Pirellulaceae	unknown genus	Multi-affiliation
Cluster_25_22	0.010	0.001	Bacteria	Armatimonadetes	Armatimonadetes	Armatimonadales	Armatimonadaceae	Armatimonas	Multi-affiliation
Cluster_32_45	0.001	0.001	Bacteria	Proteobacteria	Deltaproteobacteria	Desulfobacterales	Desulfobulbaceae	Multi-affiliation	Multi-affiliation
Cluster_33_33	0.002	0.001	Bacteria	Verrucomicrobia	Verrucomicrobiae	Pedosphaerales	Pedosphaeraceae	unknown genus	Oikopleuradioica
Cluster_38_64	0.004	0.001	Bacteria	Acidobacteria	Subgroup 17	unknown order	unknown family	unknown genus	unknown species
Cluster_71_46	0.002	0.001	Bacteria	Proteobacteria	Deltaproteobacteria	Myxococcales	Sandaracinaceae	Sandaracinus	unknown species
Cluster_11_97	0.019	0.002	Bacteria	Gemmatimonadetes	Gemmatimonadetes	Gemmatimonadales	Gemmatimonadaceae	unknown genus	unknown species

Cluster_2	0.393	0.002	Bacteria	Multi-affiliation	Multi-affiliation	Multi-affiliation	Multi-affiliation	Multi-affiliation	Multi-affiliation
Cluster_2200	0.002	0.002	Bacteria	Bacteroidetes	Bacteroidia	Chitinophagales	Saprospiraceae	unknown genus	unknown species
Cluster_276	0.279	0.002	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Leptothrix	Multi-affiliation
Cluster_2975	0.003	0.002	Bacteria	Bacteroidetes	Bacteroidia	Cytophagales	Spirosomaceae	Arsenicibacter	Cenchrus americanus
Cluster_1261	0.006	0.003	Bacteria	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Arenimonas	Multi-affiliation
Cluster_1919	0.003	0.003	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Comamonas	unknown species
Cluster_2617	0.002	0.003	Bacteria	Actinobacteria	Thermoleophilia	Gaiellales	unknown family	unknown genus	unknown species
Cluster_5645	0.001	0.003	Bacteria	Bacteroidetes	Bacteroidia	Cytophagales	Hymenobacteraceae	Hymenobacter	unknown species
Cluster_623	0.003	0.003	Bacteria	Actinobacteria	Actinobacteria	Streptomycetales	Streptomycetaceae	E1B-B3-114	Rhacophorus dennysi
Cluster_8582	0.001	0.003	Bacteria	Proteobacteria	Gammaproteobacteria	Gammaproteobacteria Incertae Sedis	unknown family	unknown genus	unknown species
Cluster_946	0.017	0.003	Bacteria	Bacteroidetes	Bacteroidia	Cytophagales	Microscillaceae	unknown genus	Multi-affiliation
Cluster_1	0.379	0.004	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Neisseriaceae	Uruburuella	Multi-affiliation
Cluster_1453	0.008	0.004	Bacteria	Bacteroidetes	Bacteroidia	Cytophagales	Spirosomaceae	Arcicella	unknown species
Cluster_2233	0.002	0.004	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodobacteriales	Rhodobacteraceae	Paracoccus	Multi-affiliation
Cluster_2254	0.002	0.004	Bacteria	Fusobacteriia	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Hypnocycli	unknown species
Cluster_5959	0.002	0.004	Bacteria	Actinobacteria	Actinobacteria	Bifidobacteriales	Bifidobacteriaceae	Bifidobacterium	Multi-affiliation

Cluster_1230	0.019	0.005	Bacteria	Proteobacteria	Gammaproteobacteria	Salinisphaerales	Solimonadaceae	unknown genus	unknown species
Cluster_1368	0.008	0.005	Bacteria	Proteobacteria	Deltaproteobacteria	Bdellovibrionales	Bacteriovoracaceae	Bacteriovorax	unknown species
Cluster_46	0.386	0.005	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	Multi-affiliation
Cluster_5460	0.001	0.005	Bacteria	Verrucomicrobia	Verrucomicrobiae	Pedosphaerales	Pedosphaeraceae	unknown genus	unknown species
Cluster_729	0.010	0.005	Bacteria	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Rhodanobacteraceae	unknown genus	unknown species
Cluster_1776	0.004	0.006	Bacteria	Proteobacteria	Deltaproteobacteria	Bdellovibrionales	Bdellovibrionaceae	Bdellovibrion	unknown species
Cluster_3774	0.002	0.006	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingopyxis	unknown species
Cluster_512	0.006	0.006	Bacteria	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Rhodanobacteraceae	unknown genus	unknown species
Cluster_6490	0.001	0.006	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaaceae	Microvirga	Bradyrhizobium sp.
Cluster_672	0.043	0.006	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Rhodocyclaceae	C39	Multi-affiliation
Cluster_2053	0.007	0.007	Bacteria	Bacteroidetes	Bacteroidia	Chitinophagales	Chitinophagaceae	unknown genus	unknown species
Cluster_2483	0.001	0.007	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Nitrosomonadaceae	mle1-7	unknown species
Cluster_2489	0.001	0.007	Bacteria	Proteobacteria	Deltaproteobacteria	Syntrophobacteriales	Syntrophacaceae	Desulfobacca	unknown species
Cluster_4951	0.001	0.007	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiales Incertae Sedis	Phreatobacter	unknown species
Cluster_6833	0.001	0.007	Bacteria	Patescibacteria	Saccharimonadia	Saccharimonadales	unknown family	unknown genus	unknown species
Cluster_736	0.006	0.007	Bacteria	Actinobacteria	Actinobacteria	Corynebacteriales	Corynebacteriaceae	Corynebacterium	Multi-affiliation



Cluster_1125	0.001	0.008	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Alkanindiges	unknown species
Cluster_1929	0.003	0.008	Bacteria	Bacteroidetes	Bacteroidia	Cytophagales	Spirosomaceae	Arcicella	Arcicella sp.
Cluster_48	0.215	0.008	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	unknown genus	unknown species
Cluster_536	0.002	0.009	Bacteria	Acidobacteria	Thermoanaerobaculia	Thermoanaerobaculales	Thermoanaerobaculaceae	Subgroup 10	unknown species
Cluster_7004	0.001	0.009	Bacteria	Proteobacteria	Alphaproteobacteria	Caulobacterales	Parvularculaceae	Amhiplicatus	unknown species
Cluster_3671	0.005	0.010	Bacteria	Multi-affiliation	Multi-affiliation	Multi-affiliation	Multi-affiliation	Multi-affiliation	Multi-affiliation
Cluster_457	0.002	0.010	Bacteria	Actinobacteria	Actinobacteria	Corynebacteriales	Mycobacteriaceae	Mycobacterium	Multi-affiliation
Cluster_8662	0.001	0.010	Bacteria	Proteobacteria	Deltaproteobacteria	Bdellovibrionales	Bdellovibrionaceae	Bdellovibrionaceae	unknown species
Cluster_875	0.009	0.010	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Novosphingobium	unknown species
Cluster_9264	0.001	0.010	Bacteria	Bacteroidetes	Bacteroidia	Sphingobacteriales	AKYH767	unknown genus	metagenome
Cluster_2812	0.001	0.011	Bacteria	Bacteroidetes	Bacteroidia	Cytophagales	Microscillaceae	unknown genus	unknown species
Cluster_298	0.030	0.011	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Massilia	Multi-affiliation
Cluster_5066	0.001	0.011	Bacteria	Bacteroidetes	Bacteroidia	Flavobacteriales	Weeksellaceae	Chryseobacterium	Chryseobacterium sp.
Cluster_1454	0.002	0.012	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Nitrosomonadaceae	mle1-7	unknown species
Cluster_1627	0.002	0.012	Bacteria	Nitrospirae	Nitrospira	Nitrospirales	Nitrospiraceae	Nitrospira	unknown species
Cluster_2368	0.001	0.012	Bacteria	Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	Verrucomicrobiaceae	unknown genus	unknown species

Cluster_44	0.271	0.012	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacterales	Burkholderiaceae	Sphaerotilus	unknown species
Cluster_9	0.251	0.012	Bacteria	Proteobacteria	Gammaproteobacteria	Multi-affiliation	Multi-affiliation	Multi-affiliation	Multi-affiliation
Cluster_29_84	0.001	0.013	Bacteria	Bacteroidetes	Bacteroidia	Cytophagales	Microscillaceae	unknown genus	Sedum alfredii
Cluster_47_96	0.001	0.013	Bacteria	Actinobacteria	Actinobacteria	Corynebacteriales	Corynebacteriaceae	Corynebacterium 1	Multi-affiliation
Cluster_89_29	0.001	0.013	Bacteria	Planctomyces	Planctomycetia	Pirellulales	Pirellulaceae	Blastopirellula	unknown species
Cluster_98_4	0.004	0.013	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Devosiaceae	Devosia	unknown species
Cluster_19_6	0.002	0.014	Bacteria	Proteobacteria	Deltaproteobacteria	Oligoflexales	Oligoflexaceae	Silvanigrella	Multi-affiliation
Cluster_28_74	0.002	0.014	Bacteria	Actinobacteria	Actinobacteria	Propionibacteriales	Nocardioidaceae	Aeromicrobium	Multi-affiliation
Cluster_42_7	0.003	0.014	Bacteria	Proteobacteria	Deltaproteobacteria	Myxococcales	Haliangiaceae	Haliangium	unknown species
Cluster_74_7	0.017	0.014	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacterales	Burkholderiaceae	Variovorax	unknown species
Cluster_15_586	0.001	0.016	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacterales	Burkholderiaceae	unknown genus	unknown species
Cluster_19_16	0.001	0.016	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacterales	Nitrosomonadaceae	MND1	unknown species
Cluster_22_68	0.007	0.016	Bacteria	Bacteroidetes	Bacteroidia	Flavobacteriales	Weeksellaceae	Chryseobacterium	Multi-affiliation
Cluster_26_95	0.005	0.016	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Rhodospirillaceae	unknown genus	Multi-affiliation
Cluster_29_39	0.001	0.016	Bacteria	Actinobacteria	Acidimicrobiia	unknown order	unknown family	unknown genus	metagenome
Cluster_48_9	0.011	0.018	Bacteria	Bacteroidetes	Bacteroidia	Chitinophagales	unknown family	unknown genus	unknown species
Cluster_49_6	0.006	0.018	Bacteria	Planctomyces	vadinHA49	unknown order	unknown family	unknown genus	unknown species

Cluster_10 317	0.001	0.019	Bacteria	Dependentia	Babeliae	Babeliales	Babeliaceae	unknown genus	Multi- affiliation
Cluster_17 60	0.001	0.019	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Nitrosomonadaceae	Ellin6067	unknown species
Cluster_6	0.217	0.019	Bacteria	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Streptococcus	Streptococcus varani
Cluster_69 9	0.013	0.019	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Mesorhizobium	Multi- affiliation
Cluster_12 41	0.001	0.020	Bacteria	Proteobacteria	Alphaproteobacteria	Reyranelles	Reyranelleae	Reyranelle	Multi- affiliation
Cluster_17 75	0.001	0.020	Bacteria	Acidobacteria	Blastocatellia (Subgroup 4)	Blastocatellales	Blastocatellaceae	unknown genus	unknown species
Cluster_18 79	0.010	0.020	Bacteria	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	Photobacterium	unknown species
Cluster_20 63	0.001	0.020	Bacteria	Proteobacteria	Gammaproteobacteria	Beggiatoales	Beggiatoaceae	Beggiatoa	metagenome
Cluster_45 89	0.002	0.020	Bacteria	Actinobacteria	Thermoleophilina	Solirubrobacteriales	Solirubrobacteraceae	Conexibacter	unknown species
Cluster_75 3	0.002	0.020	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	Paracoccus	Multi- affiliation
Cluster_17 66	0.003	0.021	Bacteria	Acidobacteria	Subgroup 6	unknown order	unknown family	unknown genus	unknown species
Cluster_37 1	0.036	0.021	Bacteria	Gemmatimonadetes	Gemmatimonadetes	Gemmatimonadales	Gemmatimonadaceae	unknown genus	Multi- affiliation
Cluster_55 68	0.001	0.021	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	Multi- affiliation
Cluster_21 16	0.002	0.022	Bacteria	Proteobacteria	Deltaproteobacteria	Bdellovibrionales	Bdellovibrionaceae	Bdellovibrio	Multi- affiliation
Cluster_33 32	0.004	0.022	Bacteria	Proteobacteria	Alphaproteobacteria	Rickettsiales	SM2D12	unknown genus	unknown species
Cluster_83 1	0.004	0.023	Bacteria	Proteobacteria	Alphaproteobacteria	Rickettsiales	Rickettsiaceae	Candidatus Megaira	unknown species
Cluster_32 23	0.002	0.024	Bacteria	Actinobacteria	Actinobacteria	Corynebacteriales	Corynebacteriaceae	Corynebacterium 1	Multi- affiliation

Cluster_3856	0.00	0.024	Bacteria	Proteobacteria	Alphaproteobacteria	unknown order	unknown family	unknown genus	unknown species
Cluster_1560	0.004	0.025	Bacteria	Proteobacteria	Deltaproteobacteria	Bdellovibrionales	Bdellovibrionaceae	Bdellovibrionaceae	unknown species
Cluster_697	0.008	0.025	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	SC-I-84	unknown genus	unknown species
Cluster_1693	0.001	0.026	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Chitinimonadaceae	Chitinivora	Multi-affiliation
Cluster_1477	0.001	0.027	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Multi-affiliation	Multi-affiliation
Cluster_2560	0.002	0.027	Bacteria	Planctomyces	Planctomycetacia	Gemmatales	Gemmataceae	unknown genus	unknown species
Cluster_4820	0.001	0.027	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	unknown genus	unknown species
Cluster_1691	0.002	0.028	Bacteria	Actinobacteria	Actinobacteria	Micrococcales	Multi-affiliation	Multi-affiliation	Multi-affiliation
Cluster_2466	0.000	0.028	Bacteria	Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	Caulobacter	Multi-affiliation
Cluster_2581	0.000	0.028	Bacteria	Proteobacteria	Gammaproteobacteria	Gammaproteobacteria Incertae Sedis	unknown family	Candidatus Ovatusbacter	unknown species
Cluster_3260	0.000	0.028	Bacteria	Planctomyces	Planctomycetacia	Pirellulales	Pirellulaceae	unknown genus	metagenome
Cluster_5754	0.001	0.028	Bacteria	Actinobacteria	MB-A2-108	unknown order	unknown family	unknown genus	unknown species
Cluster_629	0.009	0.028	Bacteria	Bacteroidetes	Ignavibacteria	OPB56	unknown family	unknown genus	unknown species
Cluster_1956	0.002	0.029	Bacteria	Planctomyces	Planctomycetacia	Pirellulales	Pirellulaceae	Pirellula	metagenome
Cluster_2707	0.000	0.030	Bacteria	Verrucomicrobia	Verrucomicrobiae	Pedosphaerales	Pedosphaeraceae	unknown genus	unknown species
Cluster_1127	0.003	0.031	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas	unknown species

Cluster_16 52	0.003	0.031	Bacteria	Bacteroidetes	Bacteroidia	Chitinophagales	Saprospiraceae	unknown genus	unknown species
Cluster_5	0.178	0.031	Bacteria	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	Salinivibrio	Multi-affiliation
Cluster_66 38	0.001	0.031	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiaceae 1	Multi-affiliation	Multi-affiliation
Cluster_69 3	0.002	0.031	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	T34	unknown genus	metagenome
Cluster_63 77	0.000	0.032	Bacteria	Proteobacteria	Alphaproteobacteria	Rickettsiales	Mitochondria	unknown genus	unknown species
Cluster_96 6	0.010	0.032	Bacteria	Actinobacteria	Actinobacteria	Multi-affiliation	Multi-affiliation	Multi-affiliation	Multi-affiliation
Cluster_13 90	0.001	0.033	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Hyphomicrobiaceae	Hyphomicrobium	Multi-affiliation
Cluster_35 39	0.001	0.033	Bacteria	Proteobacteria	Alphaproteobacteria	Acetobacterales	Acetobacteraceae	unknown genus	unknown species
Cluster_11 39	0.002	0.034	Bacteria	Actinobacteria	Thermoleophilina	Solirubrobacteriales	Solirubrobacteraceae	Patulibacter	Patulibacter brassicae
Cluster_11 957	0.000	0.034	Bacteria	Planctomycetes	Planctomycetacia	Gemmatales	Gemmataceae	unknown genus	unknown species
Cluster_14 11	0.007	0.034	Bacteria	Verrucomicrobia	Verrucomicrobiae	Pedosphaerales	Pedosphaeraceae	unknown genus	unknown species
Cluster_25 49	0.003	0.034	Bacteria	Acidobacteria	Acidobacteriia	Solibacterales	Solibacteraceae (Subgroup 3)	Bryobacter	unknown species
Cluster_39 0	0.001	0.034	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Escherichia-Shigella	unknown species
Cluster_55 98	0.001	0.034	Bacteria	Proteobacteria	Deltaproteobacteria	Myxococcales	Sandaracinaceae	unknown genus	unknown species
Cluster_72 20	0.000	0.034	Bacteria	Proteobacteria	Deltaproteobacteria	Oligoflexales	Oligoflexaceae	unknown genus	unknown species
Cluster_11 79	0.001	0.036	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Devosiaceae	Devosia	Multi-affiliation

Cluster_13 26	0.004	0.036	Bacteria	Proteobact eria	Gammaproteo bacteria	Betaproteobact eriales	Rhodocycla ceae	unknown genus	unknown species
Cluster_19 66	0.003	0.036	Bacteria	Bacteroidet es	Bacteroidia	Sphingobacteria les	env.OPS 17	unknown genus	unknown species
Cluster_31 66	0.001	0.036	Bacteria	Proteobact eria	Gammaproteo bacteria	Oceanospirillale s	Pseudohon giellaceae	Blyi10	metagenome
Cluster_40 58	0.000	0.036	Bacteria	Proteobact eria	Gammaproteo bacteria	Enterobacterial es	Enterobact eriaceae	Yersinia	unknown species
Cluster_79 83	0.001	0.036	Bacteria	Actinobact eria	Actinobacteria	Kineosporiales	Kineosporia ceae	Quadrispha era	unknown species
Cluster_94 8	0.001	0.036	Bacteria	Bacteroidet es	Bacteroidia	Bacteroidales	Prolixibacte raceae	unknown genus	unknown species
Cluster_15 50	0.002	0.037	Bacteria	Firmicutes	Bacilli	Bacillales	Paenibacill aceae	Paenibacill us	Multi- affiliation
Cluster_24 88	0.001	0.037	Bacteria	Proteobact eria	Alphaproteoba cteria	Rhizobiales	Hyphomicro biaceae	Pedomicro bium	unknown species
Cluster_25 1	0.022	0.037	Bacteria	Proteobact eria	Gammaproteo bacteria	Betaproteobact eriales	Burkholderi aceae	Massilia	Multi- affiliation
Cluster_15 53	0.012	0.038	Bacteria	Proteobact eria	Gammaproteo bacteria	Betaproteobact eriales	Multi- affiliation	Multi- affiliation	Multi- affiliation
Cluster_17 50	0.002	0.038	Bacteria	Verrucomicro bia	Verrucomicrob iae	Verrucomicrob iales	Verrucomicro biaceae	unknown genus	unknown species
Cluster_19 49	0.007	0.038	Bacteria	Proteobact eria	Gammaproteo bacteria	Oceanospirillale s	Halomonad aceae	Halomonas	Multi- affiliation
Cluster_22 07	0.003	0.038	Bacteria	Proteobact eria	Gammaproteo bacteria	Ga0077536	unknown family	unknown genus	unknown species
Cluster_43 1	0.002	0.038	Bacteria	Gemmatim onadetes	Longimicrobia	Longimicrobiale s	Longimicro biaceae	unknown genus	unknown species
Cluster_48 72	0.001	0.038	Bacteria	Proteobact eria	Gammaproteo bacteria	Betaproteobact eriales	Neisseriace ae	unknown genus	unknown species
Cluster_11 84	0.003	0.039	Bacteria	Bacteroidet es	Bacteroidia	Sphingobacteria les	Sphingobac teriaceae	Nubsella	unknown species
Cluster_39 72	0.000	0.039	Bacteria	Planctomyc etes	OM190	unknown order	unknown family	unknown genus	unknown species

Cluster_408	0.065	0.039	Bacteria	Proteobacteria	Alphaproteobacteria	Caulobacteriales	Caulobacteraceae	Brevundimonas	Multi-affiliation
Cluster_557	0.011	0.040	Bacteria	Bacteroidetes	Bacteroidia	Cytophagales	Microscillaceae	unknown genus	unknown species
Cluster_1898	0.004	0.042	Bacteria	Proteobacteria	Deltaproteobacteria	Myxococcales	mle1-27	unknown genus	unknown species
Cluster_2085	0.002	0.042	Bacteria	Firmicutes	Clostridia	Clostridiales	Peptostreptococaceae	Acetoanaerobium	Multi-affiliation
Cluster_589	0.008	0.042	Bacteria	Proteobacteria	Alphaproteobacteria	Rickettsiales	Anaplasmataceae	Neorickettsia	unknown species
Cluster_2023	0.003	0.043	Bacteria	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Lysobacter	unknown species
Cluster_2839	0.002	0.043	Bacteria	Bacteroidetes	Bacteroidia	Sphingobacteriales	NS11-12 marine group	unknown genus	Multi-affiliation
Cluster_2944	0.000	0.043	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Dysgonomonadaceae	Dysgonomonas	Dysgonomonas sp.
Cluster_3	0.149	0.043	Bacteria	Actinobacteria	Actinobacteria	Actinomycetales	Actinomycetaceae	Actinomyces	Actinomyces marimammali
Cluster_2633	0.002	0.044	Bacteria	Proteobacteria	Deltaproteobacteria	MBNT15	unknown family	unknown genus	unknown species
Cluster_647	0.001	0.044	Bacteria	Actinobacteria	Thermoleophila	Solirubrobacterales	Solirubrobacteraceae	Solirubrobacter	Multi-affiliation
Cluster_7709	0.000	0.044	Bacteria	Actinobacteria	Acidimicrobiia	IMCC26256	unknown family	unknown genus	unknown species
Cluster_6301	0.001	0.045	Bacteria	Bacteroidetes	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Ulvibacter	unknown species
Cluster_12384	0.001	0.046	Bacteria	Chloroflexi	Ktedonobacteriia	C0119	unknown family	unknown genus	unknown species
Cluster_1689	0.000	0.046	Bacteria	Bacteroidetes	Bacteroidia	Chitinophagales	Chitinophagaceae	Dinghuibacter	unknown species

Cluster_577	0.021	0.046	Bacteria	Bacteroidetes	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Flavobacterium	unknown species
Cluster_626	0.002	0.046	Bacteria	Proteobacteria	Gammaproteobacteria	Aeromonadales	Aeromonadaceae	Tolumonas	bacterium YT0325
Cluster_2497	0.002	0.047	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	unknown genus	metagenome
Cluster_6034	0.001	0.047	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiales Incertae Sedis	unknown genus	unknown species
Cluster_7271	0.000	0.047	Bacteria	Bacteroidetes	Bacteroidia	Chitinophagales	unknown family	unknown genus	metagenome
Cluster_1732	0.000	0.048	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Novosphingobium	unknown species
Cluster_199	0.001	0.048	Bacteria	Firmicutes	Bacilli	Lactobacillales	Enterococcaceae	Enterococcus	Enterococcus lactis
Cluster_1317	0.001	0.049	Bacteria	Bacteroidetes	Rhodothermia	Rhodothermale s	Rhodothermaceae	unknown genus	unknown species
Cluster_77	0.114	0.049	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Pelomonas	Multi-affiliation
Cluster_847	0.001	0.049	Bacteria	Actinobacteria	Actinobacteria	Actinomycetales	Actinomycetaceae	Actinomycetes	Multi-affiliation
Cluster_8726	0.001	0.049	Bacteria	Proteobacteria	Deltaproteobacteria	Oligoflexales	Oligoflexaceae	unknown genus	metagenome
Cluster_75	0.003	0.050	Bacteria	Proteobacteria	Gammaproteobacteria	Legionellales	Legionellaceae	Legionella	Multi-affiliation
Cluster_900	0.010	0.050	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Pseudohongiellaceae	Blyi10	unknown species





## **Discussion générale et Perspectives**

Au cours de ma thèse, je me suis attachée à mieux comprendre le processus d'infection parasitaire en milieu naturel. L'objectif était d'identifier le rôle relatif des facteurs environnementaux abiotiques, liés à l'hôte, au parasite et au microbiote dans le succès d'une infection parasitaire. De manière plus globale, j'espère avoir amené des éléments nouveaux et des pistes de réflexion qui permettront de mieux appréhender les interactions entre espèces, notamment quant aux mécanismes évolutifs et écologiques les rendent possibles.

### **Les mécanismes de l'interaction parasitaire : effets environnementaux, génomiques et effets des communautés bactériennes.**

#### **Le rôle de l'environnement à l'échelle micro-géographique sur la probabilité de rencontre**

Un des apports originaux de ces travaux de thèse porte sur l'étude du processus d'infection à l'échelle micro-géographique, qui dans ce cas correspond à un tronçon de rivière de quelques centaines de mètres structurée en plusieurs faciès (radiés, zones calmes, profondes ou peu profondes etc.). Une échelle spatiale aussi fine est très rarement considérée que ce soit en ce qui concerne l'étude des interactions hôte-parasite ou bien plus généralement en écologie évolutive (Richardson et al. 2014). Dans le cadre de cette thèse, j'ai pu mettre en évidence que certains microhabitats au sein d'un tronçon de rivière constituent des points chauds d'infection pour *T. polycolpus* alors que d'autres, au contraire constituent des points froids (Chapitre I). Des patrons spatiaux d'infections en points chauds et froids avaient déjà été documentés à large échelle spatiale (Ramirez-Sierra et al. 2010, Bousema et al. 2010, Murphy et al. 2012) ainsi qu'à l'échelle locale (eg., à l'échelle d'un parc, (Jennett et al. 2013) ou dans un rayon de 12 km (Thamm et al. 2009)), mais je documente ici un des rares exemple d'une telle distribution à

l'échelle micro-géographique. Il est intéressant de souligner que les études qui mettent en avant des points chauds d'infections à l'échelle locale portent elles aussi sur des ectoparasites monoxènes tels que les tiques, les puces (Thamm et al. 2009, Jennett et al. 2013) ou d'autres ectoparasites de poisson tels que les *Argulus* (Poulin and FitzGerald 1989). Ces parasites se caractérisent, comme *T. polycolpus*, par de faible fécondité lorsque comparés à certains parasites hétéroxènes ayant plusieurs phases multiplicatives (Turner et al. 1995, Mackinnon and Read 1999, Ebert et al. 2000). Aussi particulièrement pour ces ectoparasites, la répartition des stades infestants dans des (micro-)habitats favorables semble être un mécanisme facilitateur de la rencontre et qui pourrait compenser l'absence d'autres stratégies.

Chez *T. polycolpus* il apparaît que les microhabitats favorables à la rencontre sont des zones d'eau profondes (entre soixante centimètres et un mètre) mais sont aussi des microhabitats qui assurent un brassage de l'eau (vitesse du courant de 0.4 m/s à 1m/s). Les zones d'eau calmes (généralement caractérisées par de faibles vitesses de courant mais ce peut être aussi le cas des zones profondes) sont supposées favoriser la survie mais également la transmission des stades larvaires chez les parasites monogènes ou copépodes (Barker and Cone 2000b, Bjork and Bartholomew 2009b). Toutefois il apparaît qu'un certain brassage de l'eau (tel que reflété par l'effet de la vitesse du courant sur les risques d'infections) pourrait être nécessaire, afin par exemple d'accroître les fréquences de contact entre *T. polycolpus* et ses hôtes, notamment en remettant en suspension les larves infestantes qui sont peu mobiles par elles-mêmes (Samsing et al. 2015). Quoi qu'il en soit, il est intéressant de souligner que les points chauds d'infection se caractérisent par des facteurs environnementaux congrus dans le temps et l'espace. Ceci est un apport nouveau par rapport aux attendus théoriques qui suggéraient jusqu'à lors qu'une variabilité aléatoire de l'environnement pouvait suffire à ménager des zones de rencontre

préférentielles (Gourbière et al. 2015). Une congruence des caractéristiques environnementales des microhabitats favorables à l'infection sous-entend notamment qu'il serait possible d'établir des prédictions quant aux risques d'infection à l'échelle micro-géographique basées sur la mesure de variables environnementales.

Une perspective intéressante afin d'étudier le rôle des facteurs environnementaux sur le processus d'infection serait de replacer cette étude à l'échelle micro-géographique dans une dynamique temporelle. J'ai montré que d'une année sur l'autre les points chauds d'infections se situaient dans les mêmes zones du tronçon de rivière considéré (notamment en aval des chaussées, Chapitre I). Toutefois, les rivières sont des environnements particulièrement dynamiques du fait des variations du niveau d'eau qui peuvent affecter non seulement l'abondance des espèces localement (Taylor and Warren 2001) mais aussi la structuration spatiale en microhabitats. Ainsi, une prochaine étape pourrait être de tester comment la probabilité de rencontre entre hôte et parasite peut être modulée par la dynamique spatiale et temporelle de l'environnement micro-géographique. Ceci est particulièrement pertinent dans le cadre des changements globaux actuels qui amènent à des perturbations environnementales de plus grande amplitude mais aussi plus difficiles à anticiper (Alley et al. 2003, Stireman et al. 2005, Bauder 2005).

### **Le rôle des facteurs génomiques de l'hôte et du parasite sur la compatibilité**

Au cours de mes travaux de thèse, j'ai par ailleurs considéré le rôle des bases moléculaires de la compatibilité entre hôte et parasite dans le processus d'infection. J'ai abordé ce thème sous deux aspects : d'une part j'ai évalué le rôle de la structure et de la composition génomique des populations hôtes à large échelle spatiale sur le processus d'infection et plus particulièrement sur l'acquisition de résistance, et d'autre part le rôle de la plasticité transcriptomique des parasites sur leur capacité à exploiter de nouvelles espèces hôtes alternatives (i.e., généralisme d'hôte).

*Tracheliastes polycolpus* présente une distribution hétérogène au sein des bassins versants français. En dépit d'un fort potentiel colonisateur du parasite (Rey et al. 2015), les vandoises situées dans les bassins versant Nord-Est ne sont pas infectées. En combinant un modèle de distribution d'espèces basé sur des variables environnementales abiotiques (eg., température, composition paysagère, fragmentation ...) ainsi que biotiques (i.e., se rapportant à l'hôte du parasite, eg., taux d'hétérozygotie, taille ...) et une approche de génomique des populations, j'ai ainsi montré que la distribution actuelle de *T. polycolpus* était vraisemblablement limitée par le patrimoine génomique de ses populations hôtes plutôt que par les facteurs environnementaux (Chapitre II). En effet, les bassins versants où aucune infection n'a été enregistrée à ce jour sont occupés par une population de vandoises génomiquement différente des populations de vandoises présentant des infections. D'autre part, en contrôlant pour la structure génomique des populations hôtes, 90 variations nucléotidiques présentes chez les hôtes et directement associées aux patrons de prévalence de *T. polycolpus* ont été identifiées. Ce résultat suggère que la structure et la composition génomique des populations hôtes peuvent conférer à certaines populations hôtes une résistance face à un parasite émergent, ce qui peut contraindre en retour la dynamique d'expansion du parasite.

En ce qui concerne les facteurs génomiques du parasite, j'ai mis en avant le rôle de la plasticité transcriptomique dans la capacité d'un parasite à exploiter plusieurs espèces hôtes (Chapitre III). Cent vingt gènes différentiellement exprimés ont été identifiés chez *T. polycolpus* selon l'espèce hôte exploitée. Une analyse des processus biologiques sous-jacents a également permis de mettre en avant deux stratégies d'exploitation mises en place par le parasite suivant qu'il infecte une espèce hôte principale ou bien une espèce hôte alternative. Lorsque *T. polycolpus* infecte son hôte principal, les processus biologiques impliqués dans la machinerie cellulaire sont surexprimés, ce

qui pourrait expliquer à terme que les parasites de vandoises sont généralement de taille plus importante et produisent plus d'œufs que les parasites trouvés sur un hôte alternatif tel que le vairon (Lootvoet et al. 2013). Sur les hôtes alternatifs, en revanche *T. polycolpus* semble investir dans l'acquisition d'énergie en sur-exprimant des fonctions liées soit au métabolisme énergétique soit à l'activité musculaire qui conditionne son activité de nutrition. L'investissement dans l'acquisition énergétique des parasites présents sur les hôtes alternatifs reflète vraisemblablement un mécanisme compensatoire face à une ressource différente (Agosta et al. 2010). Ainsi l'exploitation d'hôtes alternatifs bien que permise par des ajustements plastiques resterait néanmoins sub-optimale par rapport à l'exploitation de l'espèce hôte principale (Araujo et al. 2015, De Fine Licht 2018).

Le rôle de la plasticité dans la capacité d'un parasite à faire face à de nouvelles conditions biotiques telle que représentées par une nouvelle espèce hôte est un résultat en accord avec plusieurs études théoriques (Agosta et al. 2010, Araujo et al. 2015, De Fine Licht 2018). La plasticité est un mécanisme facilitateur qui a déjà été documenté chez des parasites faisant face à de nouvelles conditions biotiques et abiotique (eg., la condition de l'hôte ou bien la température ambiante, Cameron et al. 2013, Birget et al. 2017, Shocket et al. 2018). Par exemple, Hébert et al. (Hébert et al. 2017) ont montré que chez un parasite hétéroxène (*Schistocephalus solidus*, qui utilise successivement l'épinoche et un oiseau piscivore), d'importants ajustements transcriptomiques accompagnent la transition entre hôte intermédiaire et hôte définitif. En particulier, des gènes impliqués dans les voies neurales et la perception sensorielle sont surexprimés chez l'hôte intermédiaire, vraisemblablement pour soutenir les modifications comportementales qu'occasionne ce parasite, tandis que chez l'hôte définitif, une activation des gènes impliqués dans la reproduction est observée (Hébert et al. 2017). L'étude présentée en

Chapitre III de cette thèse constitue quant à elle un des premiers exemples mettant en évidence le rôle de la plasticité dans la capacité d'un parasite à utiliser des espèces hôtes alternatives. Or, la capacité d'un parasite à exploiter différentes espèces hôtes est à la base des phénomènes de transfert d'hôtes qui conditionnent fortement la dynamique d'adaptation de la plupart des parasites émergents (*spill-over*, *spill-back*, (Thompson et al. 2009)).

### **Rôle des communautés bactériennes**

Enfin, ces travaux de thèse ont apporté de nouveaux éléments quant au rôle des communautés bactériennes (associées à l'hôte, au parasite et / ou à l'environnement) dans le processus d'infection parasitaire (Chapitre IV).

Certaines études suggéraient que la composition du microbiote de l'hôte pouvait influencer les infections parasitaires en fournissant des conditions favorables au développement du parasite (Hayes et al. 2010) ou bien en accroissant la résistance de l'hôte contre les parasites (Buffie and Pamer 2013, Britton and Young 2014). Dans le Chapitre IV de cette thèse, j'ai considéré le microbiote du mucus des vandoises, qui constitue la première barrière de défense contre les pathogènes chez les poissons et les amphibiens (Subramanian et al. 2005, Alvarez-Pellitero 2008). Toutefois par le déploiement d'un suivi individuel, j'ai pu explicitement tester le rôle du microbiote associé à l'hôte avant infection dans le processus d'infection et montrer que le microbiote des vandoises n'était pas une cause directe de l'infection par *T. polycolpus*. En revanche les conditions environnementales à l'échelle du microhabitat qui elles influencent de manière directe les probabilités d'infection (chapitre II) pourraient également influencer la composition du microbiote qui s'établit chez les hôtes sans pour autant que celui-ci ne présage des potentielles infections.

Par ailleurs, j'ai mis en évidence que les infections parasitaires par *T. polycolpus* provoquent d'importants changements au sein du microbiote associé à l'hôte et ceci de manière très localisée, à savoir au niveau du site d'ancrage du parasite. J'ai également montré que la composition des communautés bactériennes des larves de *T. polycolpus* est proche de celle des communautés bactériennes associées aux nageoires parasitées. Ceci suggère que le microbiote du parasite pourrait jouer un rôle important dans le processus d'infection probablement via une dynamique de co-infection qui faciliterait l'exploitation de l'hôte par le parasite (*Weapon strategy*, (Dheilly et al. 2015, Reynolds et al. 2015)). J'ai notamment mis en évidence que la famille bactérienne des *Burkholderiaceae* était abondante (4–16% du microbiote) à la fois dans les communautés de nageoires parasitées et chez le parasite tant au stade larvaire qu'au stade adulte. Une prochaine étape pour le futur serait donc de caractériser plus précisément certaines de ces bactéries, qui semblent avoir un rôle important dans les infections par *T. polycolpus*, à la fois au niveau taxonomique mais également au niveau fonctionnel.

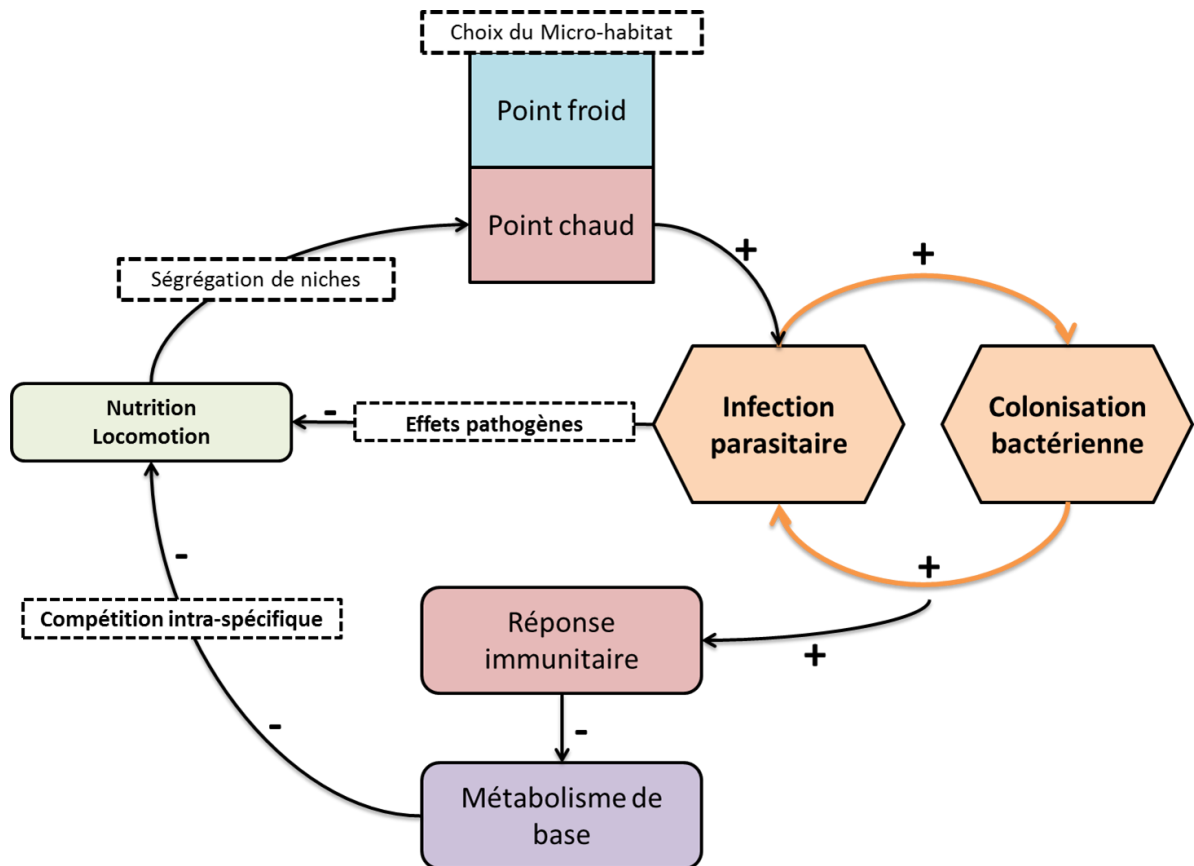
De manière générale, la prise en compte du microbiote associé à l'hôte, au parasite ou bien à leur environnement commun est une thématique de recherche qui connaît un essor récent. Nous connaissons actuellement un changement de perspective quant aux interactions hôte-parasites qui tendent à non plus seulement être considérées comme une interaction bipartite mais plutôt comme une interaction faisant intervenir un cortège de bactéries en plus du parasite (Dheilly 2014, Bordenstein and Theis 2015). Cette nouvelle direction s'avère stimulante et prometteuse et amènera vraisemblablement à une meilleure compréhension des facteurs qui modulent les interactions parasitaires.



## **Les mécanismes d'infection dans le contexte de l'agrégation parasitaire et des maladies émergentes.**

### **Mécanismes d'infection et agrégation parasitaire**

Les distributions agrégées au sein des populations hôtes sont une caractéristique commune à la grande majorité des parasites (Crofton 1971, Shaw et al. 1998, Poulin 2007). En d'autres termes, pour certains hôtes les conditions permettant l'infection parasitaire (i.e., ouverture du filtre de rencontre et du filtre de compatibilité) sont particulièrement favorables. Aussi il est intéressant de se questionner sur la manière dont les mécanismes présentés précédemment pourraient permettre d'aboutir à ce phénomène d'agrégation parasitaire. En considérant les différents résultats amenés au cours de ces travaux de thèse dans leur ensemble il est possible d'envisager des liens causaux et/ou synergétiques entre plusieurs mécanismes qui favorisent l'infection parasitaire et qui pourraient ainsi aboutir à de l'agrégation parasitaire (Fig. 8).



**Figure 8 :** Représentation d'un scénario possible pouvant amener à de l'agrégation parasitaire basé sur les résultats décrits dans cette thèse. Les flèches noires représentent les liens causaux et les flèches orange les effets synergétiques. Les hexagones représentent les événements de colonisation de l'hôte. Les rectangles arrondis représentent les réponses de l'hôte et les signes (+ et -) indiquent de quelle manière cette réponse est affectée. Les rectangles entourés de pointillés représentent selon quel processus les liens causaux peuvent se faire.

D'abord il est probable que les infections passées des hôtes puissent influencer leur probabilité d'être réinfectés dans le futur et ainsi amener à des patrons d'agrégation parasitaire. Dans une étude complémentaire à ces travaux de thèse (Annexe I), j'ai notamment montré que *T. polycolpus* avait très probablement des effets en cascade sur son hôte. Les vandoises parasitées par *T. polycolpus* sont de moins bonnes compétitrices face à des congénères sains et ceci se traduit au niveau de la niche trophique qui est différente entre vandoises saines et parasitées

(Annexe I). Bien que la niche trophique ne se superpose pas à l'habitat effectivement occupé, il est néanmoins probable que les poissons parasités occupent alors des habitats de moins bonnes qualités –à savoir les microhabitats identifiés comme favorables à l'infection parasitaire– du fait de leurs plus faibles performances dans un contexte de compétition (Chapitre I, Annexe I). Cette hypothèse est renforcée par le fait que des points chauds d'infections ont été identifiés en milieu naturel alors que dans ce cas le mouvement des hôtes n'était pas entravé (Chapitre I). Une telle agrégation spatiale des hôtes parasités en ces points chauds d'infections suggère que les hôtes parasités sont moins mobiles que les individus sains et se retrouvent alors confinés dans des zones particulièrement favorables aux infections (Poulin 2007, Paull et al. 2012, Gourbière et al. 2015). Or les vandoises fortement parasitées font justement preuve d'une locomotion ralentie lorsqu'elles sont mises en compétition avec des congénères moins parasités (Annexe I), vraisemblablement à cause des dégradations de nageoires causées par *T. polycolpus* (Fig. 6, Loot et al. 2004). Ainsi l'agrégation parasitaire pourrait résulter des modifications induites sur l'hôte une fois un premier parasite acquis ce qui favoriserait les conditions nécessaires à l'acquisition de parasites supplémentaires dans une dynamique de cercle vicieux (Fig. 8, (Beldomenico and Begon 2010)).

D'autre part il pourrait exister des effets synergétiques notamment entre les infections parasitaires à proprement parler et les infections par les bactéries associées au parasite (Chapitre IV, Fig. 8). Bien que nous ignorions à ce jour si certaines de ces bactéries sont pathogènes pour les hôtes de *T. polycolpus*, cette hypothèse reste néanmoins fortement probable. En effet, une analyse de l'expression des gènes des vandoises parasitées par rapport à des vandoises saines (Annexe I), a révélé que les vandoises parasitées exprimaient au niveau de la nageoire des fonctions impliquées dans la réponse immunitaire dirigées à l'encontre de bactéries (*response to bacterium*, Annexe I

Fig. 4-a). En conséquence la dynamique de co-infection entre *T. polycolpus* et certaines bactéries potentiellement pathogènes pourrait contribuer à fragiliser les hôtes parasités (notamment en contribuant au coût de la réponse immunitaire) par rapport aux infections parasitaires futures. Toutefois un examen plus approfondi du potentiel pathogène des bactéries qui sont retrouvées à la fois chez *T. polycolpus* et dans le mucus des nageoires des vandoises parasitées serait nécessaire pour tester cet effet synergétique.

### **Mécanismes d'infection et maladies émergentes**

*Tracheliastes polycolpus* est un parasite émergent et invasif en France (Rey et al. 2015). Aussi par rapport aux mécanismes modulant le succès d'infection que j'ai documenté dans cette thèse, il est important de s'interroger sur la part des facteurs qui pourraient plus particulièrement expliquer la dynamique d'invasion de certains parasites émergents. Les parasites émergents résultent bien souvent de leur introduction dans de nouveaux endroits via la translocation de quelques hôtes originels (Bazsalovicsová et al. 2011, Arbetman et al. 2013, Rey et al. 2015, Kuchta et al. 2018). Toutefois, le potentiel invasif de ces parasites requiert que le parasite parvienne à exploiter de nouvelles espèces hôtes locales avec lesquelles il ne dispose pas d'une longue histoire co-évolutive (Thompson et al. 2009). Ainsi, bien souvent si le filtre de rencontre se trouve ouvert par l'introduction du parasite, les facteurs agissant sur le filtre de compatibilité quant à eux sont déterminants pour la dynamique de colonisation du parasite.

Le patrimoine génomique des populations hôtes est par exemple un facteur agissant sur le succès d'infection qui se trouve être particulièrement déterminant pour un parasite invasif (Chapitre II). Le fait que les différences génomiques entre hôtes puissent influencer le processus d'infection parasitaire à l'échelle individuelle (de l'hôte), n'est pas en soit un résultat nouveau (Bernatchez and Landry 2003, Wegner et al. 2003, Aguilar et al. 2016). Pourtant la plupart des modèles de

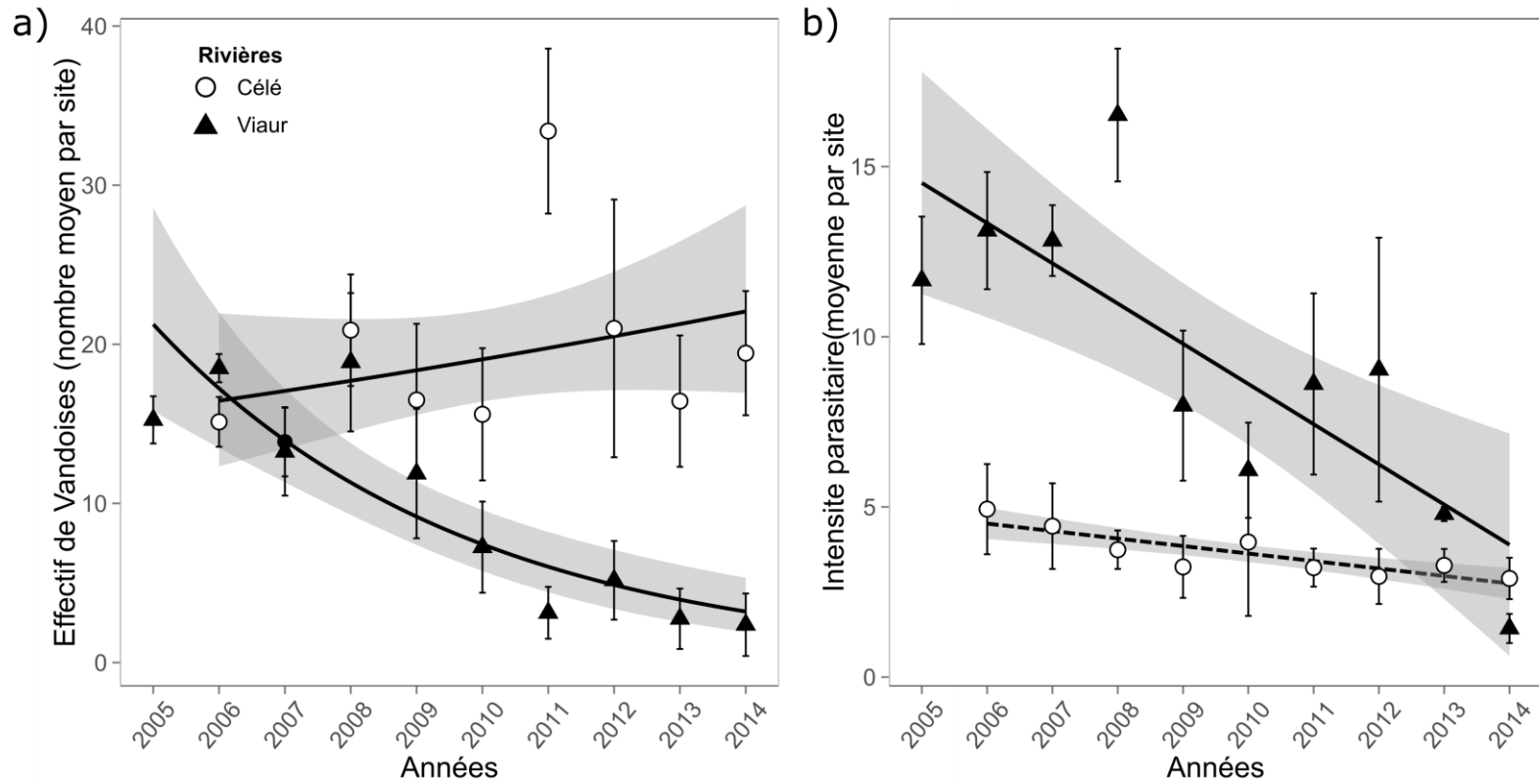
prédiction des aires de distribution futures de parasites à ce jour se basent uniquement sur des variables environnementales et tardent à prendre en compte les facteurs qui peuvent influencer la compatibilité entre hôte et parasite tel que le patrimoine génomique de l'hôte (Peterson et al. 2019, Razgour et al. 2019, Byers et al. 2019). De nombreux modèles de distributions d'espèces appliqués à des parasites à ce jour prédisent par exemple des expansions des aires de distribution des parasites ou de leurs vecteurs vers le Nord, conséquemment aux évolutions des températures (Loiseau et al. 2013, Altizer et al. 2013, Carter 2018). Il est intéressant de noter qu'en se référant au modèle de distribution d'espèce pour *T. polycolpus* établi dans le chapitre II, lui-même basé sur des variables environnementales, une expansion dans les bassins versants Nord-Est aurait été prédite à tort, et cela bien que ce modèle prenne déjà en compte un certain nombre de variables décrivant l'hôte en tant qu'environnement biotique du parasite (tel que la taille ou le niveau d'hétérozygotie). Ainsi, la structure et la composition génomique des populations hôtes parce qu'elles conditionnent en partie la compatibilité d'un parasite vis-à-vis de son hôte sont des éléments à prendre en compte et particulièrement pour considérer le succès des parasites émergents.

De manière analogue le rôle de la plasticité transcriptomique mis en avant chez *T. polycolpus* dans sa capacité à exploiter plusieurs espèces hôtes (Chapitre III) fait également partie des mécanismes qui peuvent influencer le succès des interactions parasitaires en général mais aussi tout particulièrement le succès des parasites émergents. En effet pour les parasites natifs, une longue histoire co-évolutive vis-à-vis de leurs espèces hôtes peut favoriser par sélection naturelle une adaptation du parasite à ses hôtes (Ebert 1994). Pour les parasites émergents à contrario si leur espèce hôte n'est pas présente dans leur aire d'introduction alors leur survie et leur potentiel colonisateur reposent sur leurs capacités à mettre en place des mécanismes à court terme leur

permettant d'exploiter de nouvelles espèces hôtes locales (Agosta et al. 2010, De Fine Licht 2018). *Tracheliastes polycolpus* a été introduit dans les bassins versants français seulement depuis les années 1920 (Rey et al. 2015), ce qui suppose une courte histoire évolutive, tout particulièrement préjudiciable vis-à-vis de ses hôtes alternatifs qui constituent une ressource éloignée de l'hôte originel de *T. polycolpus* dans son aire native (i.e., *Leuciscus idus*, Galicka and Penczak 1989, Sobocka et al. 2004, Lootvoet et al. 2013). Or, j'ai justement montré que *T. polycolpus* parvenait à exploiter de nouvelles espèces hôtes en s'appuyant sur la plasticité sans pour autant que du polymorphisme génomique sous-sélection selon l'espèce hôte parasitée soit détecté (Chapitre III). La plasticité constituerait donc un mécanisme que les parasites émergents peuvent déployer à court terme et de prime abord face à l'exploitation des espèces hôtes locales qui sont présentes dans l'aire d'introduction. Enfin il est important de mentionner que *T. polycolpus* est un parasite généraliste dans son aire native (des cas d'infection par *T. polycolpus* ont été également répertoriés chez *L. walekii* (Yamaguti 1940) ou encore *L. cephalus* (Barzegar and Jalali 2009)). Il est donc possible que des parasites qui présentent déjà un certain degré de plasticité pour a minima exploiter plusieurs espèces hôtes soient plus à même de s'établir et de se propager en dehors de leurs aires natives.

Une prochaine étape stimulante pourrait concerner la conciliation du rôle du patrimoine génomique des populations hôtes et de la plasticité chez le parasite dans les dynamiques d'invasion. En effet, *T. polycolpus* s'avère limité dans sa dynamique d'invasion de certains bassins versants français à cause de la divergence génomique d'une population hôte locale (population résistante) (Chapitre II). Cependant, par plasticité ce parasite s'avère capable d'exploiter des hôtes appartenant à des espèces éloignées (Chapitre III, (Lootvoet et al. 2013)). Une perspective, notamment pour le cas de *T. polycolpus*, pourrait ainsi être de s'interroger sur

les bénéfiques (i.e., colonisation favorisée en s'appuyant sur plus d'espèces hôtes potentielles) et les coûts (i.e., une exploitation sub-optimale de chaque espèce hôte) associés à l'exploitation des espèces hôtes alternatives et leurs rôles dans la persistance et le processus de colonisation d'un parasite émergent. Dans des travaux présentés en annexe (Annexe III), je renseigne en particulier un déclin démographique et génétique drastique au sein d'une métapopulation de vandoises présentant jusqu'à lors de très forts taux de parasitismes par *T. polycolpus* (Lootvoet et al. 2013a, Mathieu-Bégné et al. 2019b). Ce déclin démographique des populations hôtes a également occasionné une forte baisse de l'intensité parasitaire au sein des vandoises de cette métapopulation de 2004 à 2014 (Fig. 9). Il serait particulièrement intéressant de s'interroger sur l'intensité parasitaire de *T. polycolpus* sur ses espèces alternatives suite à la diminution drastique de l'abondance de son hôte principal. Deux attendus peuvent être formulés : soit la plasticité de *T. polycolpus* n'est pas suffisante pour lui permettre de se maintenir sur les hôtes alternatifs et dans ce cas l'intensité parasitaire sur les hôtes alternatifs devrait également décroître au cours du temps (attendu 1) ; soit la plasticité permet de soutenir les effectifs de *T. polycolpus* sur les espèces hôtes alternatives et dans ce cas à mesure que l'intensité parasitaire sur l'hôte principal diminue, celle sur les espèces alternatives devrait s'accroître (attendu 2). L'attendu 1 pourrait être plus probable puisque j'ai montré que lorsque la population locale hôte principale est résistante (i.e., pas ou peu d'hôtes principaux exploitables pour *T. polycolpus*) la distribution de *T. polycolpus* est limitée (Chapitre II) mais aussi parce que sa réponse plastique est sub-optimale sur les hôtes alternatifs (Chapitre III). Quoi qu'il en soit, une telle étude sur l'exploitation des hôtes alternatifs suite à une diminution démographique drastique de l'hôte principal pourrait définitivement éclairer la compréhension générale du rôle du filtre de compatibilité dans la persistance et la dynamique d'invasion de certains parasites.



**Figure 9 :** Illustration a) du déclin démographique observé entre 2005 et 2014 de la métapopulation de vandoises située dans la rivière Viaur (par rapport à l'expansion démographique de la métapopulation de la rivière Célé) et b) du déclin de l'intensité parasitaire du parasite *T. polycolpus* sur les vandoises du Viaur sur la même période de temps. Les lignes pleines signifient des tendances temporelles significatives et celle en pointillé une tendance non significative.



## **Echelle de travail et perspective holistique dans l'étude des interactions hôte-parasites**

Une des constatations plus transversale qui ressort également de ces travaux de thèse est que différentes échelles de travail révèlent différentes facettes du processus d'infection parasitaire. Dès lors, une vision holistique s'avère appropriée à l'étude des interactions hôtes-parasites.

### **Les interactions hôte-parasite à différentes échelles spatiales**

D'abord il est intéressant de remarquer que si l'interaction hôte-parasite est considérée à l'échelle micro-géographique (Chapitre I) ou bien à l'échelle macro-géographique (Chapitre II) différents facteurs influençant le succès de l'infection peuvent être mis en évidence. A l'échelle micro-géographique, j'ai mis en évidence un rôle important des facteurs environnementaux (tels que la profondeur de la colonne d'eau et la vitesse du courant, Chapitre I) sur les patrons d'infection spatiaux de *T. polycolpus*. A contrario à l'échelle macro-géographique, cette fois les facteurs environnementaux ne permettaient pas de rendre compte de la distribution spatiale de *T. polycolpus* et j'ai plutôt mis en évidence le rôle du patrimoine génomique des populations hôtes notamment dans la limite d'aire de distribution de *T. polycolpus*. L'étude du processus d'infection parasitaire à plusieurs niveaux de résolution est donc primordiale afin de capturer les différents facteurs qui conditionnent le succès de l'infection parasitaire et qui sont sensibles à l'échelle spatiale.

### **Les interactions hôte-parasite à différents niveaux d'organisation biologique**

Une meilleure compréhension du processus d'infection parasitaire nécessite également une vision globale pour mieux appréhender les interactions parasites dans un contexte écosystémique. Dans une des annexes de cette thèse (Annexe I) je présente une étude intégrative des impacts de *T. polycolpus* sur la vandoise rostrée de l'échelle moléculaire à l'échelle populationnelle. J'ai

notamment mis en évidence que *T. polycolpus* déclenchait une réponse immunitaire générale chez son hôte très probablement au détriment des processus métaboliques basaux. D'autre part, à l'échelle populationnelle, les individus parasités présentaient également de plus faibles performances dans un contexte de compétition mais aussi des niches trophiques différentes, sans pour autant que les compétences à l'échelle strictement individuelle ne soient influencées par le parasite. En donnant ainsi une vision de l'interaction hôte-parasite à différents niveaux d'organisation biologique, d'une part il a pu être montré que les impacts des parasites sur leurs hôtes pouvaient être dépendants des interactions que l'hôte entretient avec ses congénères ou bien avec ses proies. D'autre part une hypothèse, quant aux liens causaux existant entre les effets proximaux chez l'hôte à l'échelle moléculaire et leurs répercussions en effets indirects à plus large échelle biologique, a pu être formulée (i.e., des effets en cascade provoqués par le coût de la réponse immunitaire, Annexe I).

### **Les interactions hôte-parasite au-delà de l'interaction bipartite**

Les interactions hôte-parasites sont des interactions complexes qui, dans les populations naturelles, s'intègrent à un contexte d'interactions plus larges. Plusieurs études ont déjà suggéré l'importance de prendre en compte les interactions hôte-parasites en écologie des communautés de par l'important rôle régulateur que les parasites peuvent avoir sur les autres interactions (Bernstein et al. 2002, Wood et al. 2007, Sato et al. 2012). Toutefois, les interactions hôte-parasites sont également influencées par les autres interactions qu'entretiennent les hôtes (Møller 2008). Au cours de ces travaux de thèse j'ai par exemple montré que les effets d'un parasite sur son hôte peuvent être exacerbés par les interactions de compétition que les hôtes entretiennent entre eux et se traduire au niveau de leurs interactions trophiques (Annexe I) avec de possibles boucles de rétroaction sur le succès des infections futures (Fig. 8). D'autre part, dans le chapitre

IV il a été mis en évidence que les interactions de l'hôte et du parasite vis-à-vis de leurs communautés bactériennes respectives jouaient un rôle important par rapport au déroulement du processus d'infection parasitaire. Plusieurs études récentes proposent également de s'appuyer sur la notion d'holobionte afin de considérer non seulement l'interaction entre l'hôte et son parasite mais plus généralement d'y inclure l'interaction entre le cortège de bactéries qui leur sont associées (Dheilly 2014, Bordenstein and Theis 2015). Deux perspectives intéressantes concernant l'étude du processus d'infection seraient ainsi d'une part la prise en compte des interactions entre interactions (Møller 2008) et d'autre part d'intégrer pleinement la notion d'holobionte hôte et holobionte parasite dans l'étude des interactions parasitaires.

### **Vers une vision holistique et intégrative des interactions hôte-parasites**

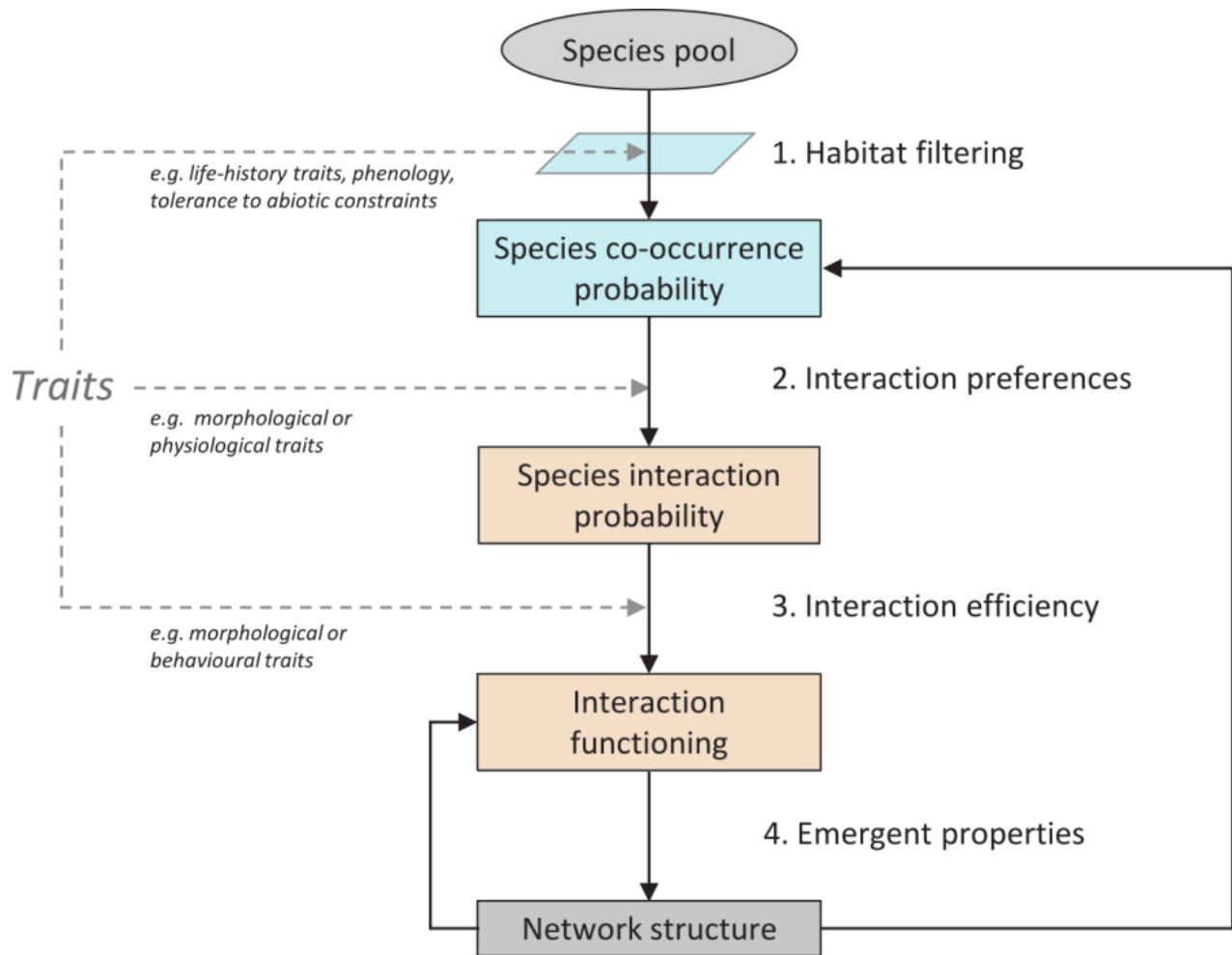
De manière générale, l'étude des interactions parasitaires est un champ disciplinaire qui se caractérise par la pluralité des approches qui y sont déployées afin de capturer la complexité de ces interactions (Altizer et al. 2003, Schmid Hempel 2011, Clark and Clegg 2017). C'est pourquoi au cours de ces travaux de thèse des approches variées telles que des approches purement expérimentales (Annexe I), des approches semi-expérimentales en populations naturelles (Chapitre I et IV), des approches descriptives (Chapitre II et III) et l'utilisation de divers outils de génomiques (Rad-seq, RNAs-seq et Métabarcoding, Chapitre II, III et IV) ont été combinées. L'étude des déterminants environnementaux et génomiques sous-jacents à la distribution spatiale de *T. polycolpus* (Chapitre II) comme l'étude des impacts de *T. polycolpus* sur son hôte à différents niveaux d'organisations biologiques (Annexe I) sont deux exemples notables qui illustrent le pouvoir des approches intégratives à apporter une compréhension approfondie du processus d'infection parasitaire. De plus en plus, la recherche en biologie s'engage dans une vision plus intégrative nécessitant le déploiement d'approches holistiques qui

tendent à concilier plusieurs champs disciplinaires (Borer et al. 2011, Matthews et al. 2011, Gunderson et al. 2017). Dans les années à venir, ceci constitue un enjeu majeur de la recherche pour la compréhension des interactions hôte-parasites (Restif et al. 2012), mais également pour la compréhension de nombreux autres processus évolutifs et écologiques.

## **Les notions de filtre de rencontre et de compatibilité par-delà les interactions**

### **hôte-parasite**

Les notions de filtres de rencontre et de compatibilité ont été centrales aux travaux exposés au sein de cette thèse. Ces notions ont permis en particulier de comprendre à quel niveau du processus d'infection parasitaire chaque facteur envisagé pouvait moduler son succès (eg., l'environnement micro-géographique qui agit sur le filtre de rencontre, la plasticité qui agit sur le filtre de compatibilité). Les notions de filtres de rencontre et de compatibilité ont été développées pour être appliquées à la compréhension des interactions étroites et durables (Combes 2001). Toutefois je pense que dans une certaine mesure, ces notions et les mécanismes sous-jacents qui en découlent, peuvent également s'appliquer à la compréhension des conditions qui favorisent bon nombre d'interactions du vivant (Fig. 10).



**Figure 10** : Schéma adaptée de Bartomeus et al. (Bartomeus et al. 2016) qui illustre la place des notions de filtre de rencontre et de filtre de compatibilité dans l'étude des réseaux d'interaction en général. En bleu sont indiqués les processus qui se réfèrent au filtre de rencontre (la filtration par l'habitat qui conditionne l'abondance et la co-occurrence des espèces) et en orange les processus qui se réfèrent au filtre de compatibilité (concordance des traits morphologiques, physiologiques et comportementaux). L'ensemble de ces processus conditionne la possibilité de chaque interaction et donc la structure du réseau d'interaction qui elle-même peut avoir des effets rétroactifs sur la dynamique de la communauté d'espèces.

Une question fondamentale en écologie des communautés est par exemple de comprendre pourquoi les espèces entrent ou non en interaction. Quel que soit le type d'interaction considéré, plusieurs auteurs s'accordent notamment sur le fait que la co-occurrence entre les espèces n'est pas une condition suffisante (Poisot et al. 2015, González-Varo and Traveset 2016, Bartomeus et al. 2016). La co-occurrence entre les espèces n'en est pas moins une première condition nécessaire à l'interaction, et dans une certaine mesure, elle se superpose à la notion de filtre de rencontre (Fig. 10). A ce jour en ce qui concerne les interactions du vivant au sens large, le filtre de rencontre est surtout pris en compte du point de vue de la densité entre les individus de deux espèces amenées ou non à interagir. Par exemple plus deux espèces sont abondantes plus leurs probabilités d'interaction seraient élevées (Altizer et al. 2003, Paull et al. 2012, Bartomeus et al. 2016). Aussi une première étape pour mieux prédire l'interaction entre des espèces serait d'affiner notre vision des facteurs qui vont moduler leur rencontre, par exemple en prenant en compte les effets environnementaux (notion d'*habitat filtering*, (Bartomeus et al. 2016), Fig. 10) et tout particulièrement à différentes échelles spatiales comme j'ai pu le montrer au cours de cette thèse.

D'autre part, certains auteurs suggèrent que pour qu'une interaction entre espèces ou individus se produise, leurs traits morphologiques, physiologiques et comportementaux doivent être compatibles (Fig. 10, (González-Varo and Traveset 2016, Bartomeus et al. 2016)). Cette notion de *trait matching* se superpose grandement à la notion de filtre de compatibilité et est particulièrement invoquée pour expliquer les liens manquants dans un réseau d'interaction (i.e., *forbidden links*, (Bartomeus et al. 2016), c'est-à-dire pourquoi bien que deux espèces puissent se rencontrer celles-ci n'interagissent pas. Par exemple si le diamètre d'un fruit est trop grand, alors la capacité d'un prédateur à l'ingérer pourra être entravée et ainsi empêcher l'interaction

(Wheelwright 1985, Olesen et al. 2011). Dans ces travaux de thèse, j'ai tout particulièrement mis en avant les facteurs génomiques qui pouvaient moduler la compatibilité entre les acteurs d'une interaction. Cette composante du filtre de compatibilité n'est pas incluse à ce jour dans la notion de *trait matching*. Une prise en compte notamment de la variabilité et de la structure génomique des populations, mais aussi de la plasticité transcriptomique au sein des espèces susceptibles d'interagir pourrait être également un élément prometteur pour améliorer notre compréhension des conditions qui amènent certaines espèces, populations ou individus à interagir.

Enfin quel que soit le type d'interaction auquel les notions de filtre de rencontre et de compatibilité s'appliquent, il pourrait être pertinent de tendre vers une vision moins binaire de ces filtres. L'exemple du rôle de la plasticité dans l'exploitation sub-optimale des espèces hôtes alternatives de *T. polycolpus* (Chapitre III) suggère notamment que le filtre de compatibilité, plutôt qu'un filtre ouvert ou fermé, pourrait être envisagé comme un goulot d'étranglement qui permet l'exploitation d'espèces hôtes de manière plus ou moins efficace selon un gradient de compatibilité. A un certain point le goulot d'étranglement pourrait se fermer lorsque la population d'hôte locale est totalement résistante au parasite (Chapitre II) ou bien l'espèce hôte trop éloignée de la ressource principale du parasite pour lui permettre de l'exploiter. Bien que plus complexe cette vision graduelle des notions de filtres de rencontre et de compatibilité pourrait permettre une compréhension plus fine des interactions du vivant telles qu'elles se produisent dans la nature.

En conclusion, les notions de filtres de rencontre et de compatibilité sont des notions importantes car elles donnent un cadre propice à l'identification des mécanismes qui sous-tendent certaines interactions et tout particulièrement les interactions étroites et durables. Toutefois, ces notions pourraient également être affinées et intégrées à l'étude des interactions du vivant en général,

permettant ainsi d'améliorer notre compréhension des facteurs écologiques et évolutifs qui conditionnent la propension des espèces à interagir. Or, comprendre les mécanismes qui permettent aux interactions du vivant de se produire est un enjeu majeur pour tendre vers une meilleure anticipation des conséquences des changements globaux sur la biodiversité et le fonctionnement des écosystèmes.

## **Conclusion**

Les interactions du vivant sont une composante essentielle de la biodiversité. Comprendre les mécanismes qui sous-tendent les interactions entre espèces est au cœur d'une anticipation éclairée de la réponse des individus, des populations et des communautés face aux changements globaux actuels. Au cours de ces travaux de thèse, je me suis plus particulièrement penchée sur les interactions hôtes-parasites afin d'en étudier les mécanismes. Il ressort que le déploiement d'approches intégratives est essentiel afin de considérer à la fois les facteurs écologiques mais aussi évolutifs qui façonnent les interactions hôtes-parasites au travers différents niveaux de complexité, à la fois en terme d'échelle spatiale, de niveaux d'organisation biologique, mais aussi en terme de contexte interactif. J'espère par ces travaux avoir amené des éléments qui permettront de mieux appréhender le processus d'infection parasitaire et tout particulièrement de mieux anticiper les risques et les conséquences liés à l'émergence de pathogènes qui sont au cœur des enjeux sociétaux actuels.





## Bibliographie Générale

- Acevedo-whitehouse, K. and Cunningham, A. 2006. Is MHC enough for understanding wildlife immunogenetics? - Trends in Ecology & Evolution 21: 433–438.
- Adamo, S. A. 2002. Modulating the Modulators: Parasites, Neuromodulators and Host Behavioral Change. - Brain, Behavior and Evolution 60: 370–377.
- Adamson, M. L. and Cairn, J. N. 1994. Evolutionary factors influencing the nature of parasite specificity. - Parasitology 109: 85–95.
- Afrin, T. et al. 2019. Sequential Changes in the Host Gut Microbiota During Infection With the Intestinal Parasitic Nematode *Strongyloides venezuelensis*. - Front. Cell. Infect. Microbiol. 9: 217.
- Agosta, S. J. et al. 2010. How specialists can be generalists: resolving the “parasite paradox” and implications for emerging infectious disease. - Zoologia 27: 151–162.
- Aguilar, J. R. et al. 2016. MHC-I provides both quantitative resistance and susceptibility to blood parasites in blue tits in the wild. - J Avian Biol 47: 669–677.
- Alizon, S. and Michalakis, Y. 2015. Adaptive virulence evolution: the good old fitness-based approach. - Trends in Ecology & Evolution 30: 248–254.
- Alley, R. B. et al. 2003. Abrupt climate change. - science 299: 2005–2010.
- Altizer, S. et al. 2003. Social organization and parasite risk in mammals: integrating theory and empirical studies. - Annual Review of Ecology, Evolution, and Systematics 34: 517–547.
- Altizer, S. et al. 2013. Climate Change and Infectious Diseases: From Evidence to a Predictive Framework. - Science 341: 514–519.
- Alvarez-Pellitero, P. 2008. Fish immunity and parasite infections: from innate immunity to immunoprophylactic prospects. - Veterinary Immunology and Immunopathology 126: 171–198.
- Anacker, B. L. et al. 2008. Susceptibility to *Phytophthora ramorum* in a key infectious host: landscape variation in host genotype, host phenotype, and environmental factors. - New Phytologist 177: 756–766.
- Araújo, M. B. and Luoto, M. 2007. The importance of biotic interactions for modelling species distributions under climate change. - Global Ecology and Biogeography 16: 743–753.
- Araujo, S. B. L. et al. 2015. Understanding Host-Switching by Ecological Fitting. - PLOS ONE 10: e0139225.
- Arbetman, M. P. et al. 2013. Alien parasite hitchhikes to Patagonia on invasive bumblebee. - Biol Invasions 15: 489–494.

Bär, A.-K. et al. 2015. The Interplay of Host Microbiota and Parasitic Protozoans at Mucosal Interfaces: Implications for the Outcomes of Infections and Diseases. - PLoS Negl Trop Dis 9: e0004176.

Barker, D. E. and Cone, D. K. 2000. Occurrence of *Ergasilus celestis* (Copepoda) and *Pseudodactylogyrus anguillae* (Monogenea) among wild eels (*Anguilla rostrata*) in relation to stream flow, pH and temperature and recommendations for controlling their transmission among captive eels. - Aquaculture 187: 261–274.

Barrett, T. M. et al. 2012. Host–parasite distributions under changing climate: *Tsuga heterophylla* and *Arceuthobium tsugense* in Alaska. - Canadian Journal of Forest Research 42: 642–656.

Bartoli, P. et al. 2000. Acquisition of parasites correlated with social rank and behavioural changes in a fish species. - Journal of Helminthology 74: 289–293.

Bartomeus, I. et al. 2016. A common framework for identifying linkage rules across different types of interactions. - Funct Ecol 30: 1894–1903.

Barzegar, M. and Jalali, B. 2009. Crustacean parasites of fresh and brackish (Caspian Sea) water fishes of Iran. - J. Agric. Sci. Technol 11: 161–171.

Bascompte, J. et al. 2006. Asymmetric Coevolutionary Networks Facilitate Biodiversity Maintenance. - Science 312: 431–433.

Bauder, E. T. 2005. The effects of an unpredictable precipitation regime on vernal pool hydrology. - Freshwater Biology 50: 2129–2135.

Bazsalovicsová, E. et al. 2011. Population study of *Atractolytocestus huronensis* (Cestoda: Caryophyllidea), an invasive parasite of common carp introduced to Europe: mitochondrial cox1 haplotypes and intragenomic ribosomal ITS2 variants. - Parasitol Res 109: 125–131.

Beatty, J. K. et al. 2017. *Giardia duodenalis* induces pathogenic dysbiosis of human intestinal microbiota biofilms. - International Journal for Parasitology 47: 311–326.

Beldomenico, P. M. and Begon, M. 2010. Disease spread, susceptibility and infection intensity: vicious circles? - Trends in Ecology & Evolution 25: 21–27.

Bellard, C. et al. 2012. Impacts of climate change on the future of biodiversity: Biodiversity and climate change. - Ecology Letters 15: 365–377.

Benavides, M. V. et al. 2016. Genomic Regions Associated with Sheep Resistance to Gastrointestinal Nematodes. - Trends in Parasitology 32: 470–480.

Bernatchez, L. and Landry, C. 2003. MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? - Journal of Evolutionary Biology 16: 363–377.

Bernstein, C. et al. 2002. Intraspecific competition between healthy and parasitised hosts in a host-parasitoid system: consequences for life-history traits. - Ecological Entomology 27: 415–423.

- Birget, P. L. G. et al. 2017. Phenotypic plasticity in reproductive effort: malaria parasites respond to resource availability. - *Proc. R. Soc. B* 284: 20171229.
- Bjork, S. J. and Bartholomew, J. L. 2009. The effects of water velocity on the *Ceratomyxa shasta* infectious cycle. - *Journal of Fish Diseases* 32: 131–142.
- Blanchet, S. et al. 2009a. Evidence of parasite-mediated disruptive selection on genetic diversity in a wild fish population. - *Molecular Ecology* 18: 1112–1123.
- Blanchet, S. et al. 2009b. Why do parasitized hosts look different? Resolving the “chicken-egg” dilemma. - *Oecologia* 160: 37–47.
- Blanchet, S. et al. 2010. Evidence for host variation in parasite tolerance in a wild fish population. - *Evolutionary Ecology* 24: 1129–1139.
- Blois, J. L. et al. 2013. Climate Change and the Past, Present, and Future of Biotic Interactions. - *Science* 341: 499–504.
- Bordenstein, S. R. and Theis, K. R. 2015. Host Biology in Light of the Microbiome: Ten Principles of Holobionts and Hologenomes. - *PLOS Biology*: 13: e1002226.
- Borer, E. T. et al. 2010. Local context drives infection of grasses by vector-borne generalist viruses: Local vs. regional context and infection risk. - *Ecology Letters* 13: 810–818.
- Borer, E. T. et al. 2011. Bridging Taxonomic and Disciplinary Divides in Infectious Disease. - *EcoHealth* 8: 261–267.
- Bousema, T. et al. 2010. Identification of Hot Spots of Malaria Transmission for Targeted Malaria Control. - *J INFECT DIS* 201: 1764–1774.
- Britton, R. A. and Young, V. B. 2014. Role of the Intestinal Microbiota in Resistance to Colonization by *Clostridium difficile*. - *Gastroenterology* 146: 1547–1553.
- Brommer, J. E. et al. 2011. Body Size and Immune Defense of Nestling Blue Tits (*Cyanistes caeruleus*) in Response to Manipulation of Ectoparasites and Food Supply. - *The Auk* 128: 556–563.
- Brooker, R. W. 2006. Plant-plant interactions and environmental change. - *New Phytol* 171: 271–284.
- Brousseau, P.-M. et al. 2018. Trait matching and phylogeny as predictors of predator–prey interactions involving ground beetles. - *Functional ecology* 32: 11–202.
- Browman, H. I. et al. 2004. The effect of light on the settlement of the salmon louse, *Lepeophtheirus salmonis*, on Atlantic salmon, *Salmo salar* L. - *Journal of fish diseases* 27: 701–708.
- Brunner, J. L. and Ostfeld, R. S. 2008. Multiple causes of variable tick burdens on small-mammal hosts. - *Ecology* 89: 2259–2272.

- Buffie, C. G. and Pamer, E. G. 2013. Microbiota-mediated colonization resistance against intestinal pathogens. - *Nat Rev Immunol* 13: 790–801.
- Byers, J. E. et al. 2019. What factors explain the geographical range of mammalian parasites? - *Proc. R. Soc. B* 286: 20190673.
- Cameron, A. et al. 2013. Plasticity in transmission strategies of the malaria parasite, *Plasmodium chabaudi*: environmental and genetic effects. - *Evol Appl* 6: 365–376.
- Caminade, C. et al. 2019. Impact of recent and future climate change on vector-borne diseases: Climate change and vector-borne diseases. - *Ann. N.Y. Acad. Sci.* 1436: 157–173.
- Cardon, M. et al. 2011. Host characteristics and environmental factors differentially drive the burden and pathogenicity of an ectoparasite: a multilevel causal analysis: Causalities in host-parasite interactions. - *Journal of Animal Ecology* 80: 657–667.
- Carter, N. 2018. Climate change, disease range shifts, and the future of the Africa lion. - *Conservation Biology* 32: 1207–1210.
- Chaikaew, N. et al. 2009. Exploring spatial patterns and hotspots of diarrhea in Chiang Mai, Thailand. - *Int J Health Geogr* 8: 36–46.
- Choh, Y. et al. 2012. Predator-prey role reversals, juvenile experience and adult antipredator behaviour. - *Sci Rep* 2: 728–737.
- Clark, N. J. and Clegg, S. M. 2017. Integrating phylogenetic and ecological distances reveals new insights into parasite host specificity. - *Molecular Ecology* 26: 3074–3086.
- Clark, N. J. et al. 2018. Parasite spread at the domestic animal - wildlife interface: anthropogenic habitat use, phylogeny and body mass drive risk of cat and dog flea (*Ctenocephalides spp.*) infestation in wild mammals. - *Parasites & Vectors* 11: 8–19.
- Combes, C. 1996. Parasites, biodiversity and ecosystem stability. - *Biodivers Conserv* 5: 953–962.
- Combes, C. 2001. *Parasitism: The Ecology and Evolution of Intimate Interactions*. - University of Chicago Press.
- Combes, C. 2005. *The art of being a parasite*. - University of Chicago Press.
- Costedoat, C. et al. 2006. Heterogeneous colonization pattern of European Cyprinids, as highlighted by the dace complex (*Teleostei: Cyprinidae*). - *Molecular Phylogenetics and Evolution* 41: 127–148.
- Crawley, M. J. 2009. *Natural enemies: the population biology of predators, parasites and diseases*. - John Wiley & Sons.
- Crofton, H. D. 1971. A quantitative approach to parasitism. - *Parasitology* 62: 179–193.
- da Silva, R. C. and Langoni, H. 2009. *Toxoplasma gondii*: host–parasite interaction and behavior manipulation. - *Parasitology Research* 105: 893–898.

- Damian, R. T. 1964. Molecular mimicry: antigen sharing by parasite and host and its consequences. - *The American Naturalist* 98: 129–149.
- Damian, R. T. 1989. Molecular mimicry: parasite evasion and host defense. - In: *Molecular Mimicry*. Springer, pp. 101–115.
- Das, M. et al. 2018. Exploration of Molecular/Antigenic Mimicry between *Trichuris ovis* and its Ovine Host. - *Int. J. Curr. Microbiol. App. Sci.* 7: 569–582.
- Daszak, P. 2000. Emerging Infectious Diseases of Wildlife-- Threats to Biodiversity and Human Health. - *Science* 287: 443–449.
- De Fine Licht, H. H. 2018. Does pathogen plasticity facilitate host shifts? - *PLOS Pathogens* 14: e1006961.
- De Kort, H. et al. 2018. Interacting grassland species under threat of multiple global change drivers. - *J Biogeogr* 45: 2133–2145.
- de Meeûs, T. and Renaud, F. 2002. Parasites within the new phylogeny of eukaryotes. - *Trends in Parasitology* 18: 247–251.
- Dheilly, N. M. 2014. Holobiont–Holobiont Interactions: Redefining Host–Parasite Interactions. - *PLoS Pathog* 10: e1004093.
- Dheilly, N. M. et al. 2015. Biological warfare: Microorganisms as drivers of host–parasite interactions. - *Infection, Genetics and Evolution* 34: 251–259.
- Dobson, A. P. and Hudson, P. J. 1986. Parasites, disease and the structure of ecological communities. - *Trends in Ecology & Evolution* 1: 11–15.
- Dybdahl, M. F. and Lively, C. M. 1998. Host-Parasite Coevolution: Evidence for Rare Advantage and Time-Lagged Selection in a Natural Population. - *Evolution* 52: 1057–1066.
- Ebert, D. 1994. Virulence and Local Adaptation of a Horizontally Transmitted Parasite. - *Science* 265: 1084–1086.
- Ebert, D. 2016. Evolution of Parasite Virulence. - *How Evolution Shapes Our Lives: Essays on Biology and Society*: 134.
- Ebert, D. et al. 2000. Dose effects and density-dependent regulation of two microparasites of *Daphnia magna*. - *Oecologia* 122: 200–209.
- Eisner, T. et al. 2005. Secret weapons: defenses of insects, spiders, scorpions, and other many-legged creatures. - Harvard University Press.
- Euzet, L. and Combes, C. 1980. Species problems in parasitic animals. - *Species problems in parasitic animals.*: 239–285.
- Federici, E. et al. 2015. Characterization of the Skin Microbiota in Italian Stream Frogs (*Rana italica*) Infected and Uninfected by a Cutaneous Parasitic Disease. - *Microbes and environments* 30: 262–269.

- Fenner, F. and Fantini, B. 1999. Biological control of vertebrate pests: the history of myxomatosis, an experiment in evolution. - CABI publishing.
- Feyereisen, R. 1999. Insect P450 enzymes. - Annual review of entomology 44: 507–533.
- Fourcade, Y. et al. 2014. Continental-scale patterns of pathogen prevalence: a case study on the corncrake. - Evolutionary Applications 7: 1043–1055.
- Frost, C. M. et al. 2019. Using Network Theory to Understand and Predict Biological Invasions. - Trends in Ecology & Evolution 34: 831–843.
- Fryer, G. 1982. The parasitic copepoda and branchiura of British freshwater fishes. A handbook and key. - Freshwater Biological Association.
- Galicka, W. and Penczak, T. 1989. *Tracheliastes polycolpus* Nordm. 1832 (*Lernaeopodidae*) in the Narew river. - Wiadomosci parazytologiczne 35: 247–250.
- Gaulke, C. A. et al. 2019. A longitudinal assessment of host-microbe-parasite interactions resolves the zebrafish gut microbiome's link to *Pseudocapillaria tomentosa* infection and pathology. - Microbiome 7: 10–26.
- Gautret, P. et al. 1995. The adjustment of the schizogonic cycle of *Plasmodium chabaudi chabaudi* in the blood to the circadian rhythm of the host. - Parasite 2: 69–74.
- Gervasi, S. S. et al. 2015. The context of host competence: a role for plasticity in host–parasite dynamics. - Trends in Parasitology 31: 419–425.
- Gillespie, T. R. and Chapman, C. A. 2008. Forest fragmentation, the decline of an endangered primate, and changes in host–parasite interactions relative to an unfragmented forest. - American Journal of Primatology 70: 222–230.
- González-Varo, J. P. and Traveset, A. 2016. The Labile Limits of Forbidden Interactions. - Trends in Ecology & Evolution 31: 700–710.
- Gourbière, S. et al. 2015. Fundamental Factors Determining the Nature of Parasite Aggregation in Hosts. - PLoS ONE 10: e0116893.
- Grimm, N. B. et al. 2008. Global Change and the Ecology of Cities. - Science 319: 756–760.
- Guernier, V. et al. 2017. Gut microbiota disturbance during helminth infection: can it affect cognition and behaviour of children? - BMC Infect Dis 17: 58–64.
- Guisan, A. and Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. - Ecology Letters 8: 993–1009.
- Gunderson, A. R. et al. 2017. Indirect Effects of Global Change: From Physiological and Behavioral Mechanisms to Ecological Consequences. - Integrative and Comparative Biology 57: 48–54.
- Hansen, J. et al. 2006. Global temperature change. - Proceedings of the National Academy of Sciences 103: 14288–14293.

- Hatcher, M. J. et al. 2012. Diverse effects of parasites in ecosystems: linking interdependent processes. - *Frontiers in Ecology and the Environment* 10: 186–194.
- Hay, J. et al. 1984. The effect of congenital *Toxoplasma* infection on mouse activity and relative preference for exposed areas over a series of trials. - *Annals of Tropical Medicine & Parasitology* 78: 611–618.
- Hayes, K. S. et al. 2010. Exploitation of the Intestinal Microflora by the Parasitic Nematode *Trichuris muris*. - *Science* 328: 1391–1394.
- Hebert, F. O. et al. 2015. Identification of candidate mimicry proteins involved in parasite-driven phenotypic changes. - *Parasites & vectors* 8: 225.
- Hébert, F. O. et al. 2017. Major host transitions are modulated through transcriptome-wide reprogramming events in *Schistocephalus solidus*, a threespine stickleback parasite. - *Molecular Ecology* 26: 1118–1130.
- Hrdá, Š. et al. 2000. Transient nature of *Toxoplasma gondii*-induced behavioral changes in mice. - *Journal of Parasitology* 86: 657–664.
- Hudson, P. J. 1998. Prevention of Population Cycles by Parasite Removal. - *Science* 282: 2256–2258.
- Hudson, P. J. et al. 2006. Is a healthy ecosystem one that is rich in parasites? - *Trends in Ecology & Evolution* 21: 381–385.
- Hudson, P. J. et al. 2008. The emergence of wildlife disease and the application of ecology. - *Infectious disease ecology: the effects of ecosystems on disease and of disease on ecosystems*. 11th Cary Conference, Millbrook, New York, USA, May 2005.: 347–367.
- Ives, A. R. and Carpenter, S. R. 2007. Stability and Diversity of Ecosystems. - *Science* 317: 58–62.
- Jeefoo, P. et al. 2010. Spatio-Temporal Diffusion Pattern and Hotspot Detection of Dengue in Chachoengsao Province, Thailand. - *IJERPH* 8: 51–74.
- Jenkins, T. P. et al. 2018. *Schistosoma mansoni* infection is associated with quantitative and qualitative modifications of the mammalian intestinal microbiota. - *Sci Rep* 8: 12072.
- Jennett, A. L. et al. 2013. Tick infestation risk for dogs in a peri-urban park. - *Parasites & vectors* 6: 358.
- Johnson, P. T. J. and Wilber, M. Q. 2017. Biological and statistical processes jointly drive population aggregation: using host–parasite interactions to understand Taylor’s power law. - *Proc. R. Soc. B* 284: 20171388.
- Jones, K. E. et al. 2008. Global trends in emerging infectious diseases. - *Nature* 451: 990–993.
- Koh, L. P. et al. 2004. Species Coextinctions and the Biodiversity Crisis. 305: 1632–1634.



- Kuchta, R. et al. 2018. Asian Fish Tapeworm: The Most Successful Invasive Parasite in Freshwaters. - *Trends in Parasitology* 34: 511–523.
- Kutz, S. . et al. 2005. Global warming is changing the dynamics of Arctic host-parasite systems. - *Proceedings of the Royal Society B: Biological Sciences* 272: 2571–2576.
- Lacey, E. S. et al. 2014. Close encounters: contributions of carbon dioxide and human skin odour to finding and landing on a host in *Aedes aegypti*: Response of *Aedes aegypti* to CO<sub>2</sub> and human foot odour. - *Physiological Entomology* 39: 60–68.
- Lachish, S. et al. 2013. Spatial determinants of infection risk in a multi-species avian malaria system. - *Ecography* 36: 587–598.
- Lafferty, K. D. and Morris, A. K. 1996. Altered Behavior of Parasitized Killifish Increases Susceptibility to Predation by Bird Final Hosts. - *Ecology* 77: 1390–1397.
- Lambrechts, L. et al. 2006. Coevolutionary interactions between host and parasite genotypes. - *Trends in Parasitology* 22: 12–16.
- Lee, S. C. et al. 2014. Helminth Colonization Is Associated with Increased Diversity of the Gut Microbiota. - *PLoS Negl Trop Dis* 8: e2880.
- Little, T. J. et al. 2006. Parasite-host specificity: experimental studies on the basis of parasite adaptation. - *Evolution* 60: 31–38.
- Little, T. J. et al. 2010. The Coevolution of Virulence: Tolerance in Perspective. - *PLoS Pathogens* 6: e1001006.
- Lloyd-Smith, J. O. et al. 2005. Superspreading and the effect of individual variation on disease emergence. - *Nature* 438: 355–359.
- Loiseau, C. et al. 2013. Predictions of avian *Plasmodium* expansion under climate change. - *Scientific Reports* 3: 1126–1132.
- Loot, G. et al. 2004. The effects of the ectoparasite *Tracheliastes polycolpus* (Copepoda: *Lernaeopodidae*) on the fins of rostrum dace (*Leuciscus leuciscus burdigalensis*). - *Parasitology Research* 94: 16–23.
- Lootvoet, A. et al. 2013. Patterns and processes of alternative host use in a generalist parasite: insights from a natural host-parasite interaction. - *Functional Ecology* 27: 1403–1414.
- Luikart, G. et al. 2008. Candidate gene microsatellite variation is associated with parasitism in wild bighorn sheep. - *Biology Letters* 4: 228–231.
- Lymbery, A. J. et al. 2014. Co-invaders: The effects of alien parasites on native hosts. - *International Journal for Parasitology: Parasites and Wildlife* 3: 171–177.
- Mackinnon, M. J. and Read, A. F. 1999. Genetic relationships between parasite virulence and transmission in the rodent malaria *plasmodium chabaudi*. - *Evolution* 53: 689–703.

- Mathieu-Bégné, E. et al. 2019. Demographic and genetic collapses in spatially structured populations: insights from a long-term survey in wild fish metapopulations. - *Oikos* 128: 196–207.
- Matthews, B. et al. 2011. Toward an integration of evolutionary biology and ecosystem science: Integration of evolutionary biology and ecosystem science. - *Ecology Letters* 14: 690–701.
- Mazé-Guilmo, E. 2016. Etude du potentiel de colonisation des parasites : une approche intégrative.
- Mikheev, V. N. et al. 2015. Behavioural adaptations of argulid parasites (*Crustacea: Branchiura*) to major challenges in their life cycle. - *Parasites & Vectors* 8: 394–404.
- Møller, A. P. 2008. Interactions between Interactions. - *Annals of the New York Academy of Sciences* 1133: 180–186.
- Mooney, J. P. et al. 2015. Inflammation-associated alterations to the intestinal microbiota reduce colonization resistance against non-typhoidal *Salmonella* during concurrent malaria parasite infection. - *Scientific Reports* 8: 394–404.
- Morand, S. et al. 1996. Parasite—host coevolution and geographic patterns of parasite infectivity and host susceptibility. - *Proc. R. Soc. Lond. B* 263: 119–128.
- Morens, D. M. et al. 2004. The challenge of emerging and re-emerging infectious diseases. - *Nature* 430: 242.
- Morgan, E. R. et al. 2009. Canine pulmonary angiostrongylosis: The influence of climate on parasite distribution. - *Parasitology International* 58: 406–410.
- Morrill, A. and Forbes, M. R. 2012. Random parasite encounters coupled with condition-linked immunity of hosts generate parasite aggregation. - *International Journal for Parasitology* 42: 701–706.
- Morse, S. S. 1995. Factors in the Emergence of Infectious Diseases. - *Emerging Infectious Diseases* 1: 9.
- Mougeot, F. 2005. Elevated spring testosterone increases parasite intensity in male red grouse. - *Behavioral Ecology* 17: 117–125.
- Murphy, T. M. et al. 2012. The prevalence and distribution of *Alaria alata*, a potential zoonotic parasite, in foxes in Ireland. - *Parasitol Res* 111: 283–290.
- Newbold, L. K. et al. 2017. Helminth burden and ecological factors associated with alterations in wild host gastrointestinal microbiota. - *ISME J* 11: 663–675.
- Nylin, S. et al. 2018. Embracing Colonizations: A New Paradigm for Species Association Dynamics. - *Trends in Ecology & Evolution* 33: 4–14.
- Ogden, N. and Gachon, P. 2019. Climate change and infectious diseases: What can we expect? - *CCDR* 45: 76–80.

Olesen, J. M. et al. 2011. Missing and forbidden links in mutualistic networks. - Proc. R. Soc. B 278: 725–732.

Parnesan, C. 2006. Ecological and Evolutionary Responses to Recent Climate Change. - Annual Review of Ecology, Evolution, and Systematics 3: 637–669.

Parnesan, C. and Yohe, G. 2003. A globally coherent fingerprint of climate change impacts across natural systems. - Nature 421: 37–42.

Patz, J. A. 1996. Global Climate Change and Emerging Infectious Diseases. - JAMA 275: 217.

Patz, J. A. et al. 2000. Effects of environmental change on emerging parasitic diseases. - International Journal for Parasitology 30: 1395–1405.

Paull, S. H. et al. 2012. From superspreaders to disease hotspots: linking transmission across hosts and space. - Frontiers in Ecology and the Environment 10: 75–82.

Pérez-Rodríguez, A. et al. 2013. Finding the appropriate variables to model the distribution of vector-borne parasites with different environmental preferences: climate is not enough. - Glob Change Biol 19: 3245–3253.

Peterson, M. L. et al. 2019. Incorporating local adaptation into forecasts of species' distribution and abundance under climate change. - Glob Change Biol 25: 775–793.

Piasecki, W. 1839. Life cycle of *Tracheliastes maculatus* Kollar, 1835 (*Copepoda*, *Siphonostomatoida*, *Lernaeopodidae*). - Wiadomości Parazytologiczne.

Pickles, R. S. A. et al. 2013. Predicting shifts in parasite distribution with climate change: a multitrophic level approach. - Global Change Biology 19: 2645–2654.

Poisot, T. et al. 2012. The dissimilarity of species interaction networks. - Ecol Lett 15: 1353–1361.

Poisot, T. et al. 2015. Beyond species: why ecological interaction networks vary through space and time. - Oikos 124: 243–251.

Portet, A. et al. 2018. Infection of the *Biomphalaria glabrata* vector snail by *Schistosoma mansoni* parasites drives snail microbiota dysbiosis. – bioRxiv: 386623.

Poulin, R. 1995. Evolution of parasite life history traits: myths and reality. - Parasitology Today 11: 342–345.

Poulin, R. 2007. Are there general laws in parasite ecology? - Parasitology 134: 763–776.

Poulin, R. and FitzGerald, G. J. 1989. Risk of parasitism and microhabitat selection in juvenile sticklebacks. - Can. J. Zool. 67: 14–18.

Poulin, R. and Keeney, D. B. 2008. Host specificity under molecular and experimental scrutiny. - Trends in Parasitology 24: 24–28.

Price, P. W. 1974. Strategies for egg production. - Evolution: 76–84.

- Price, P. W. 1980. *Biology of parasites*. - Princeton University Monograph.
- Rabajante, J. F. et al. 2019. On a mechanistic process of macroparasite aggregation. - bioRxiv, 680041.
- Raffel, T. R. et al. 2008. Parasites as predators: unifying natural enemy ecology. - *Trends in Ecology & Evolution* 23: 610–618.
- Ramirez-Sierra, M. J. et al. 2010. Patterns of house infestation dynamics by non-domiciliated *Triatoma dimidiata* reveal a spatial gradient of infestation in rural villages and potential insect manipulation by *Trypanosoma cruzi*. - *Tropical Medicine & International Health* 15: 77–86.
- Ras, R. et al. 2015. Perturbation of the intestinal microbiota of mice infected with *Cryptosporidium parvum*. - *International Journal for Parasitology* 45: 567–573.
- Razgour, O. et al. 2019. Considering adaptive genetic variation in climate change vulnerability assessment reduces species range loss projections. - *Proc Natl Acad Sci USA* 116: 10418–10423.
- Restif, O. et al. 2012. Model-guided fieldwork: practical guidelines for multidisciplinary research on wildlife ecological and epidemiological dynamics. - *Ecol Lett* 15: 1083–1094.
- Rey, O. et al. 2015. Elucidating the spatio-temporal dynamics of an emerging wildlife pathogen using approximate Bayesian computation. - *Molecular Ecology* 24: 5348–5363.
- Reynolds, L. A. et al. 2015. Cohabitation in the Intestine: Interactions among Helminth Parasites, Bacterial Microbiota, and Host Immunity. - *J.I.* 195: 4059–4066.
- Richardson, J. L. et al. 2014. Microgeographic adaptation and the spatial scale of evolution. - *Trends in Ecology & Evolution* 29: 165–176.
- Rogers, D. J. and Hassell, M. P. 1974. General Models for Insect Parasite and Predator Searching Behaviour: Interference. - *The Journal of Animal Ecology* 43: 239.
- Ronquist, F. 2003. Parsimony analysis of coevolving species associations. - *Tangled trees: phylogeny, cospeciation and coevolution*: 22–64.
- Samsing, F. et al. 2015. Gone with the flow: current velocities mediate parasitic infestation of an aquatic host. - *International Journal for Parasitology* 45: 559–565.
- Sato, T. et al. 2012. Nematomorph parasites indirectly alter the food web and ecosystem function of streams through behavioural manipulation of their cricket hosts. - *Ecology Letters* 15: 786–793.
- Schmid Hempel, P. 2011. *Evolutionary parasitology: The integrated study of infections, immunology, ecology, and genetics*.
- Schwarzkopf, M. D. and Ramaswamy, V. 2008. Evolution of stratospheric temperature in the 20th century. - *Geophys. Res. Lett.* 35: L03705.
- Shaw, D. J. et al. 1998. Patterns of macroparasite aggregation in wildlife host populations. - *Parasitology* 117: 597–610.

- Shocket, M. S. et al. 2018. Parasite rearing and infection temperatures jointly influence disease transmission and shape seasonality of epidemics. - *Ecology* 99: 1975–1987.
- Sobecka, E. et al. 2004. Parasite fauna of ide, *Leuciscus idus* in Lake Dabie, Poland. - *Acta ichthyologica et Piscatoria*. 34: 33–42.
- Stireman, J. O. et al. 2005. Climatic unpredictability and parasitism of caterpillars: Implications of global warming. - *Proceedings of the National Academy of Sciences* 102: 17384–17387.
- Subramanian, A. et al. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. - *Proceedings of the National Academy of Sciences* 102: 15545–15550.
- Suttle, K. B. et al. 2007. Species Interactions Reverse Grassland Responses to Changing Climate. - *Science* 315: 640–642.
- Swe, P. M. et al. 2014. Parasitic scabies mites and associated bacteria joining forces against host complement defence. - *Parasite Immunol* 36: 585–593.
- Tamarozzi, F. et al. 2011. Onchocerciasis: the Role of Wolbachia Bacterial Endosymbionts in Parasite Biology, Disease Pathogenesis, and Treatment. - *Clinical Microbiology Reviews* 24: 459–468.
- Taylor, C. M. and Warren, M. L. 2001. Dynamics in species composition of stream fish assemblages: environmental variability and nested subsets. - *Ecology* 82: 2320–2330.
- Thamm, S. et al. 2009. Ectoparasite Infestations of Hedgehogs (*Erinaceus europaeus*) are Associated with Small-Scale Landscape Structures in an Urban–Suburban Environment. - *EcoHealth* 6: 404–413.
- Theron, A. and Combes, C. 1995. Asynchrony of Infection Timing, Habitat Preference, and Sympatric Speciation of Schistosome Parasites. 49: 372–375.
- Theron, A. et al. 2008. Effects of laboratory culture on compatibility between snails and schistosomes. - *Parasitology* 135: 1179–1188.
- Thompson, R. C. A. et al. 2009. Parasite Zoonoses and Wildlife: Emerging Issues. - *IJERPH* 6: 678–693.
- Trøjelsgaard, K. et al. 2015. Geographical variation in mutualistic networks: similarity, turnover and partner fidelity. - *Proc. R. Soc. B* 282: 20142925.
- Turner, C. M. R. et al. 1995. Replication, differentiation, growth and the virulence of *Trypanosoma brucei* infections. - *Parasitology* 111: 289–300.
- Tylianakis, J. M. et al. 2008. Global change and species interactions in terrestrial ecosystems. - *Ecology Letters* 11: 1351–1363.
- Valiente-Banuet, A. et al. 2015. Beyond species loss: the extinction of ecological interactions in a changing world. - *Funct Ecol* 29: 299–307.

- Wang, S. X. et al. 2006. The SARS outbreak in a general hospital in Tianjin, China – the case of super-spreader. - *Epidemiol. Infect.* 134: 786–791.
- Ward, E. J. et al. 2014. Complexity is costly: a meta-analysis of parametric and non-parametric methods for short-term population forecasting. - *Oikos* 123: 652–661.
- Wegner, K. M. et al. 2003. Parasite Selection for Immunogenetic Optimality. - *Science* 301: 1343–1343.
- Wenzel, M. A. et al. 2015. Genome-wide association and genome partitioning reveal novel genomic regions underlying variation in gastrointestinal nematode burden in a wild bird. - *Molecular Ecology* 24: 4175–4192.
- Wheelwright, N. T. 1985. Fruit-Size, Gape Width, and the Diets of Fruit-Eating Birds. - *Ecology* 66: 808–818.
- Wisz, M. S. et al. 2013. The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. - *Biological Reviews* 88: 15–30.
- Wolinska, J. and King, K. C. 2009. Environment can alter selection in host–parasite interactions. - *Trends in Parasitology* 25: 236–244.
- Wood, C. L. et al. 2007. Parasites alter community structure. - *Proceedings of the National Academy of Sciences* 104: 9335–9339.
- Woolhouse, M. E. J. et al. 1997. Heterogeneities in the transmission of infectious agents: Implications for the design of control programs. - *Proceedings of the National Academy of Sciences* 94: 338–342.
- Yamaguti, S. 1940. *Trachealiastes polycolpus* von Nordmann, 1832, parasitic on *Leuciscus waleckii* (Dybowski) from Manchoukuo and Sakhalin. - *日本動物学彙報* 19: 39–42.
- Zhang, L. et al. 2015. Major histocompatibility complex alleles associated with parasite susceptibility in wild giant pandas. - *Heredity* 114: 85–93.
- Zueva, K. J. et al. 2018. Genomic signatures of parasite-driven natural selection in north European Atlantic salmon (*Salmo salar*). - *Marine Genomics* 39: 26–38.



## Annexes

En annexe trois travaux complémentaires à ma thèse sont présentés.

**L'Annexe I** correspond à une étude à différents niveaux d'organisation de l'infection parasitaire et de ses conséquences pour l'hôte. Cette étude rassemble près de dix ans de travaux sur le système *T. polycolpus*-*L. burdigalensis* coordonnés par Géraldine Loot et Simon Blanchet. Cette étude met particulièrement en exergue les potentiels effets en cascade d'un parasite sur son hôte et tente de faire le lien entre les effets directs et indirects des parasites sur leurs hôtes en proposant une vision englobante.

**L'Annexe II** correspond à une note technique sur l'assemblage du transcriptome de *T. polycolpus*, qui constitue la ressource de base sur laquelle s'appuie le chapitre II de cette thèse.

**L'Annexe III** correspond à un suivi démographique et génétique de deux métapopulations de vandoises rostrées sur une période de dix ans (de 2004 à 2014). Cette étude renseigne notamment l'association entre un déclin démographique dans l'une des métapopulations de vandoises et la perte de diversité génétique mesurée sur des marqueurs neutres. Enfin cette étude met l'accent sur l'impact de la structure en métapopulation et le rôle des populations locales –en tant que source ou puit d'individus ou bien d'allèles– dans les déclins démographiques et génétiques.





# Parasite impacts goes from host gene expression to host performance in ecosystems

---

Mathieu-Bégné Eglantine<sup>1,2,3\*§</sup>, Blanchet Simon<sup>1,2§</sup> Rey Olivier<sup>3</sup>, Cucherousset Julien<sup>1</sup>, Méjean Lionel<sup>1</sup>, Mitta Guillaume<sup>3</sup> and Loot Géraldine<sup>1</sup>

<sup>1</sup>Université Paul Sabatier, Centre National pour la Recherche Scientifique, Ecole Nationale de Formation Agronomique; Laboratoire Evolution et Diversité Biologique (UMR5174), 118 route de Narbonne, F-31062 Toulouse, France

<sup>2</sup>Université Paul Sabatier, Centre National pour la Recherche Scientifique; Station d'Ecologie Théorique et Expérimentale (UMR5321), 2 route du CNRS, F-09200 Moulis, France

<sup>3</sup>Université Perpignan Via Domitia, UMR5244 Interaction Hôte-Parasite-Environnement, 58 Avenue Paul Alduy, F-66860, Perpignan, France

§: Denotes co-first authors

## Corresponding author:

\*Mathieu-Bégné Eglantine

UMR5174 EDB

118 route de Narbonne

31062 Toulouse

E-mail:[eglantine.mb@gmail.com](mailto:eglantine.mb@gmail.com)

**Authors' contributions:**

GL and SB coordinated the study and supervised the data acquisition. OR, GL and SB acquired molecular data for the objective O1. LM, GL and SB conducted the experiments of the objectives O2 and O3 and acquired primary data. JC, GL and SB acquired isotopic data for the objective O4. EMB ran bioinformatics and statistical analyses with essential feedback from GM and JC from transcriptomic and isotopic related results respectively. EMB, SB, GL and OR wrote the manuscript. All authors participated to the conception of the ideas, and they all agree to be held accountable for the content therein and approve the final version of the manuscript.

**Key-words:** Host-parasite interactions, Multi-level study, biological level of organization, parasite-mediated effects, host response

## Abstract

Ecological interactions can be studied at different levels of biological organization, from molecular to ecosystem levels. However, studies generally focus on one or a few levels of organization at a time, hence limiting our ability to generate integrative predictions. Here, we aimed at measuring the responses of a host to an ectoparasite, from the molecular level to the population level using a set of complementary approaches. We used the ectoparasitic copepod *Tracheliastes polycolpus* and its main host the rostrum dace (*Leuciscus burdigalensis*, a freshwater fish species) as model species, and we tested whether –and how– this parasite generates cascading effects from molecules to populations. Specifically we tested whether *T. polycolpus* altered the whole gene expression (in three different tissues), the functional responses, the competitive ability and the trophic niche of *L. burdigalensis*. We showed that parasitized hosts –compared to healthy ones– over-expressed genes related to immune processes at the expense of genes related to metabolism, suggesting an energetic trade-off imposed by the parasite. We further showed that parasitized hosts were poorer competitors (i.e., they had a lower access to feeding resources) compared to healthy hosts, although the functional response of hosts was not affected. Finally, the parasite modified the trophic niche of its hosts, which suggests potential impacts of the parasite at biological levels broader than the population. We concluded that *T. polycolpus* has a strong effect at the molecular level (with a costly immune response expressed in parasitized hosts) that generated cascading effects on other biological levels, notably those concerning interactions with the prey community of hosts. By using complementary tools, this study is one of the first at demonstrating how a parasite can affect the host biology at multiple biological scales, which illustrates the usefulness of integrative studies for dissecting the complexity of ecological interactions.

## **Introduction**

Parasites have strong impacts on their host at several biological levels of organization, from the molecular to the ecosystem levels. These levels are hierarchically structured, as within host processes at the lowest level underlie among-host processes in populations, and host populations interact with non-host populations in ecosystems. Depending on the biological level host-parasite interactions are studied, host response to parasite usually refer to different outputs (Schmid Hempel 2011). Having an overview of these interactions is a complex task but essential to provide a deeper understanding of evolutionary and ecological processes governing the conflicts between hosts and parasites.

At the molecular level, studies have mainly focused on how parasites can alter host gene expression and notably induce activation of genes and molecular pathways related to host immune defenses (Thompson and Kavaliers 1994, Richman 1997, Adamo 2002). Also, some parasites were shown to directly regulate gene expression in host brain underlying a parasite strategy that induces behavioral changes in the host to promote parasite transmission (Adamo 2013, Feldmeyer et al. 2016, Ontoria et al. 2018). At the individual level, parasite impacts on host usually refer to direct phenotypic alterations that are eventually associated with a fitness cost for the host (Lehmann 1993, Lafferty and Morris 1996, Wolinska et al. 2004). Some studies have also focus on how parasites could alter the life-history traits of their hosts such as growth, fecundity, behavior or dispersal (Agnew et al. 2000, Loot et al. 2004, Horký et al. 2014). Finally, these direct effects on host phenotype and life-history traits are supposed to possibly modify interactions between parasitized hosts and healthy conspecifics through 'parasitic arbitration', which ultimately modulate the community structure and the whole ecosystem functioning (through top-down or bottom-up effects, Bernstein et al. 2002, Wood et al. 2007, Sato et al.

2012). However, even if there are evidences of parasite impacts at the different biological levels of organization, previous studies have mainly focused on one or a few levels at a time, hence limiting our ability to reach a global overview of interactions between parasites and their hosts.

There are still some examples of multi-level studies that have brought a broader vision of the role of parasites than the one restricted to pathogenic effects of parasites on hosts, and that stands at the core of many definitions of parasitism (Hatcher et al. 2012). Multi-levels studies especially highlighted trait-mediated effects of parasites at the community level, in addition to direct effects on their hosts (Mouritsen and Poulin 2005, Wood et al. 2007, Sato et al. 2012). For instance, the impact of some trematode parasites on the burying behavior of their cockle hosts have been shown to positively affect the species richness of the whole community (Mouritsen and Poulin 2005). In another study, Sato et al. (2012) have shown that nematomorph parasites, by manipulating their cricket hosts into committing suicides by jumping into water, released the predation pressure of fish predators on benthic aquatic invertebrates which consequently decreased the benthic algae biomass and increased the leaf break-down rate. Such multi-levels and integrative studies have thus greatly improved our understanding of the role of parasites, notably by identifying parasite-mediated effects on ecosystem functioning (Hatcher et al. 2012).

However, so far multi-level studies integrated parasite impacts on their hosts from the individual level to the population level or to the community level, whereas parasite effects occurring at the molecular level remain mainly studied independently (Bankers et al. 2017). At the molecular level the impact of parasite is usually measured on the expression of genes or their resulting proteins (Chetouhi et al. 2015, McTaggart et al. 2015). With the development of advanced genomic-based approaches, more candidates (proteins and genes) not necessarily related to immunity can now be addressed, which opens new avenue to study the proximate impacts of

parasites on their hosts (Robledo et al. 2014, Feldmeyer et al. 2016). Nevertheless, relating these proximate effects of parasites occurring at the molecular level to the consequences for host fitness at broader levels remain difficult since it would require i) a reciprocal measure of host response at the molecular level and of host fitness at other levels of organization, and ii) to combine a set of appropriate experimental and descriptive approaches dedicated to each level of biological organization. Usually molecular studies rely on the inference of the functional processes in which specific genes or protein candidates are involved without measuring parasite impacts on hosts at broader biological levels of organization (but pathogenic effects of parasite on host, Chetouhi et al. 2015, McTaggart et al. 2015, Ronza et al. 2016). Hence, although challenging, empirical multi-level studies are required to formulate solid hypotheses regarding the causal links between proximate processes occurring at the molecular and/or individual levels and trait-mediated processes that occur at the population or the community levels.

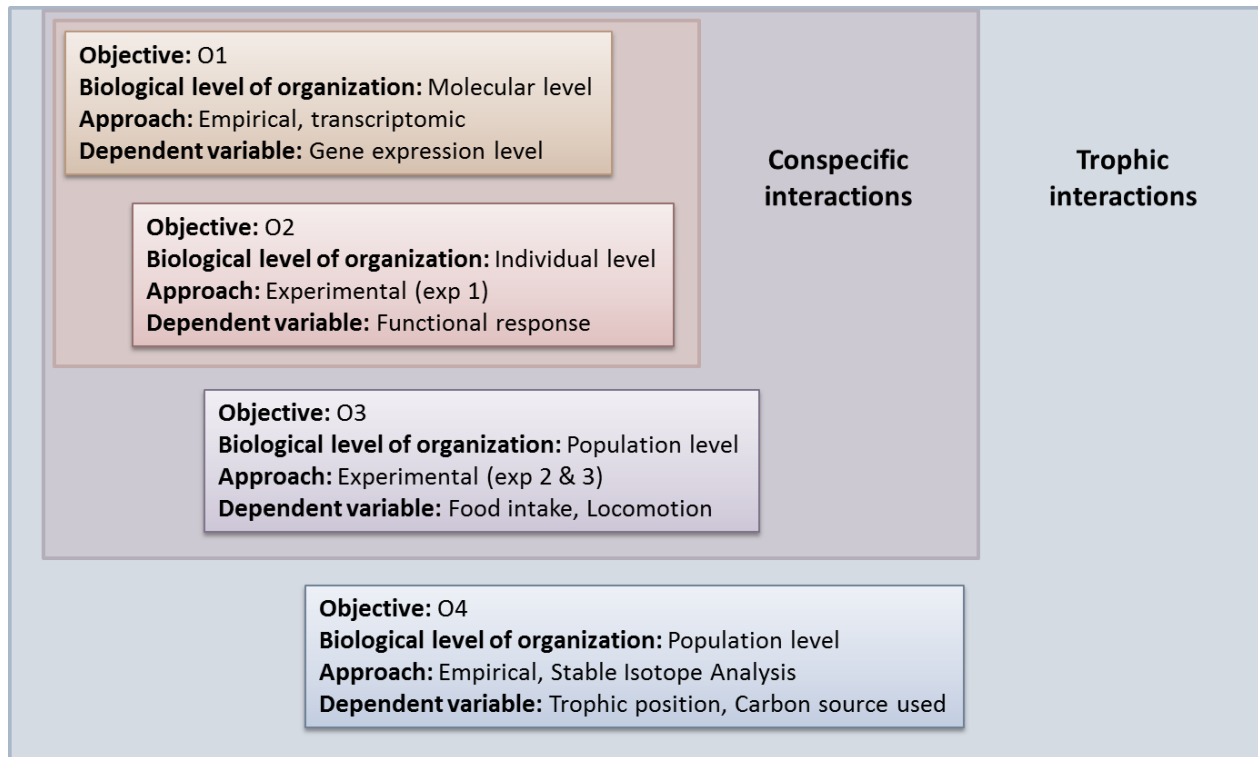
We aimed to conduct an integrative study in order to address the impacts of parasite on host at different biological levels of organization. To do so we focused on the ectoparasitic copepod *Tracheliastes polycolpus* and its main host in southwestern France, the cyprinid *Leuciscus burdigalensis* (i.e., the rostrum dace). *Tracheliastes polycolpus* usually anchors on the fins of its host where it feeds on epithelium and mucus, hence causing fin degradation (Loot et al. 2004). The *T. polycolpus*-*L. burdigalensis* system fits well our purpose since *T. polycolpus* is not directly lethal for its host but have important pathogenic impacts (Loot et al. 2004, Blanchet et al. 2009a). Furthermore, the main host of *T. polycolpus*, the rostrum dace, is a gregarious fish species (small shoal of 3-5 individuals usually) feeding on invertebrates. With such a model it is thus possible to measure how both conspecific interactions and trophic interactions influence the impact of the parasite on its hosts. Hence, both direct effects and indirect trait-mediated effects of

*T. polycolpus* on its host can be investigated. Finally, *Tracheliastes polycolpus* is also an emerging parasite in French River Basins (Rey et al. 2015a) hence constituting an interesting model for the integrative study of emerging parasite impacts on local host populations.

More precisely, in this study we aimed to measure the impact of *T. polycolpus* on *L. burdigalensis* through four main objectives each related to one level of organization (Fig. 1). At the molecular level, we used a transcriptomic approach in order to test whether parasitized dace had a different pattern of gene expression and use different molecular pathways compare to healthy ones (O1). At the individual level we used an experimental approach to test whether *T. polycolpus* could alter the functional response (i.e., the intake rate of a consumer as a function of food density) of its hosts (O2). At the population level, we used an experimental approach to test whether *T. polycolpus* can affect the competitive performances of its hosts (compared to healthy conspecifics), notably in term of food exploitation and locomotion abilities (O3). Finally, we went further by including potential interspecific interactions, using a Stable Isotope Analyses (SIA) to test whether *T. polycolpus* can modulate the trophic niche of its hosts (O4). Since *T. polycolpus* has strong pathogenic effects we expect to highlight an over-expression of immune-related genes in parasitized dace compared to healthy conspecifics. We also expect this immune response to be strong and costly since *T. polycolpus* is an emerging parasite, and *L. burdigalensis* is still quite naïve toward this parasite (Lee and Klasing 2004, Rey et al. 2015a). At the other organization levels, we more likely expect indirect negative effects of *T. polycolpus* on its hosts, due to a reallocation of energy from development toward immunity (Lochmiller and Deerenberg 2000). More precisely we expect to measure less efficient functional responses, lower competitive performances and a different trophic niche in parasitized hosts compared to healthy



conspecifics. This pioneer multi-scale study of host response should likely improve our ability to formulate insightful integrative predictions regarding host-parasite relationship outcomes.



**Figure 1:** Graphical representation of the different objectives addressed in this study that aimed to investigate the impacts of *T. polycolpus* on its main host at different biological levels of organization. Each objective is set at a particular biological level (from molecular level to population level) following a nested structure which include or not other interactions such as conspecific or trophic interactions.

## Material and Method

### Study model and sampling design

We focused on a crustacean ectoparasite *T. polycolpus* and its main host in southwestern France, the rostrum dace (*L. burdigalensis*). *Tracheliastes polycolpus* is a monoxene parasite, meaning that only one host is needed to fulfil the parasite life cycle. Only females *T. polycolpus* are parasitic (males are dwarf and free-living). More specifically, *T. polycolpus* anchors to host fins (sometimes to the body) where it feeds on epithelium and mucus hence causing fins degradation, inflammations and favoring secondary infections by other pathogens such as bacteria, fungi and viruses (Loot *et al.* 2004). *Tracheliastes polycolpus* has thus strong pathogenic impacts on their hosts that were notably shown to decrease host growth rate (Loot *et al.* 2004, Blanchet *et al.* 2009a). *Leuciscus burdigalensis* is a cyprinid freshwater fish endemic from the southwestern France, feeding on benthic or drifting invertebrates, and usually encountered in small shoals of 3-5 individuals. Rostrum dace can reach 30 cm in length, and can live up to 10-12 years. They reach maturity at ~2–3 years after which they can reproduce during spring and several times along their life time (Keith *et al.* 2011).

Our study is based on two rivers belonging to the Garonne-Dordogne River Basin in southwestern France, the Viaur and the Célé Rivers respectively (see map displayed in Fig. 1 in Mathieu-Bégné *et al.* 2019). Parasite and dace populations from these two rivers are monitored since 2005–2006. Here, we combined empirical and experimental approaches involving parasites and hosts from the two rivers simultaneously or from one of the two rivers to reach several objectives (i.e., O1, O2, O3 and O4, Fig. 1). For both empirical and experimental approaches, fish were caught using electric-fishing, they were anesthetized using oil clove when necessary,

they were measured to the nearest mm (fork length) and their parasite load (intensity) was retrieved. Sampling and experiments were carried out according to national laws and appropriate permits. We detailed hereafter the sampling details for each specific objective (see Table 1 for an overview of sampling effort for each objective).

**Table 1:** Summary of the sampling effort regarding each objectives

Objective	Year	River	Site	X L93	Y L93	Parasitic load	Health status	Nbr of dace	Size (mm)	Weight (g)
<b>O1</b>	2012	Célé	Bagnac	633254.96	6396550.58	0–0	Healthy	4	140-183	26.9–63.12
<b>O1</b>	2012	Célé	Bagnac	633254.96	6396550.58	5–18	Parasitized	4	165-217	43.05- 106.07
<b>O2</b>	2009	Célé	Ami du Célé	595010.93	6381008.37	0–0	Healthy	6	147-245	NA
<b>O2</b>	2009	Célé	Ami du Célé	595010.93	6381008.37	1–8	Parasitized	3	183-225	NA
<b>O2</b>	2009	Célé	Bagnac	633254.96	6396550.58	1–80	Parasitized	3	145-240	NA
<b>O2</b>	2009	Viaur	Saint Juste	648686.99	6336533.14	0–0	Healthy	5	174-216	NA
<b>O2</b>	2009	Viaur	Saint Juste	648686.99	6336533.14	1–5	Parasitized	9	167-233	NA
<b>O3</b>	2006	Viaur	Bannes	672895.03	6352109.02	7–20	Highly parasitized	10	144.5- 222.8	25.3–95.4
<b>O3</b>	2006		Bannes	672895.03	6352109.02	0–4	weakly parasitized	10	148-216	24.5–92.7

<b>O4</b>	2009	Célé	Bagnac	633254.96	6396550.58	0–0	Healthy	8	168-245	NA
<b>O4</b>	2009	Célé	Bagnac	633254.96	6396550.58	1–18	Parasitized	11	195-250	NA
<b>O4</b>	2009	Viaur	La Capelle	668472.06	6350728.12	2–15	Parasitized	21	166-265	NA
<b>O4</b>	2009	Viaur	Navech	651160.22	6340130.3	0–0	Healthy	5	181-280	NA
<b>O4</b>	2009	Viaur	Navech	651160.22	6340130.3	1–29	Parasitized	13	172-300	NA

## **Transcriptomic study: Unraveling the molecular pathways involved in *L. burdigalensis* responses to *T. polycolpus***

### ***Sampling design and data acquirement***

In order to unravel the molecular pathways involved in *L. leuciscus* infestation by *T. polycolpus* we aimed to compare gene expression levels between healthy and parasitized dace in three different tissues. We based this analysis on the transcriptome assembly of *L. Leuciscus* previously published in Rey et al. (2015b). Briefly, eight adult fish, similar in size and weight (26.9–106.7 g and 140–217 mm, Table 1), were sampled in one site (Bagnac, Table 1) in the Célé River in 2012. Four individuals were free from parasites and refer hereafter as healthy regarding *T. polycolpus* infestation, and four individuals were parasitized (5, 7, 14 and 18 parasites). Individuals were ethically sacrificed and three tissues, namely the cephalic kidney, the spleen (two tissues known to be involved in fish immune response, Alvarez-Pellitero 2008) and a piece of dorsal or anal fin (anchorage areas of *T. polycolpus*), were carefully collected and stored at -80°C. RNA was extracted using the RNeasy Plus Mini Kit (Qiagen reference: 7413) and then sequenced in short sequences of 100bp called reads through Illumina sequencing. Finally, RNA extraction was dosed (to estimate concentration and contamination) using a nanodrop ND-8000 (Thermo Scientific) and quality of each extraction was recorded using a BioAnalyser (Agilent Technologies). All details are provided in (Rey et al. 2015b, Genomic Resources Development Consortium et al. 2015).

### ***Transcriptome assembly improvement and RNA quantification***

The transcriptome was previously assembled (meaning short sequences elaborated from reads were assembled in longer sequences called contigs) and is available from Rey et al. (2015,

Genomic Resources Development Consortium et al. 2015). However, the transcriptome was rather large (~659 364 contigs), notably due to the presence of non-coding RNA as well as some contigs being redundant with others (Eldem et al. 2017). Consequently, we first aimed at reducing the size of the transcriptome.

First, only putative protein-coding contigs were conserved as we aimed to compare genes expression levels. To do so, the Trinity plugging TransDecoder version 2.0.1 (Haas et al. 2013) was used with default parameters in order to target contigs containing Open Reading Frames (ORF). This step resulted in 318 812 contigs likely corresponding to protein-coding genes. Secondly, in order to reduce the transcriptome redundancy, we used CD-HIT-EST version 4.6 (Huang et al. 2010) to cluster contigs according to their similarity (cutoff of 0.95) and retain only the longest exemplary of each contig (i.e., the most representative contig). This later step resulted in a transcriptome assembly of 55 545 contigs length, likely corresponding to protein-coding and non-redundant genes.

Finally, we quantified genes expression level, through estimation of transcripts abundances. Basically, raw reads were mapped back to genes of the transcriptome assembly (here our improved version of the assembly), and then counted in order to evaluate gene expression levels as transcript abundances. Gene expression levels were first estimated per individual using the Trinity perl script *align\_and\_estimate\_abundances.pl* that calls the short sequence aligner Bowtie2 version 2.0 (Langmead and Salzberg 2012) and RSEM version 2.3.1 (Li and Dewey 2011) successively. Parameters were kept as default except for Bowtie 2 in which more mismatch between reads and transcriptome was tolerated due to the previous clustering step (mismatch penalty=2, allowed mismatch=1). Transcripts abundances of each individual were then gathered

into expected raw count matrices (one per tissue) using the Trinity perl script *abundance\_estimates\_to\_matrix.pl*.

### ***Identification of genes differentially expressed (DEG)***

We used the R package *EdgeR* to test whether some genes were differentially expressed genes (DEG) between healthy and parasitized fish and following *EdgeR* recommendations (Robinson et al. 2010). We excluded genes for which transcripts had less than 0.01 copy per million in at least four samples (or three samples regarding the spleen because an outlier individual had to be removed) to exclude low expressed genes. To account for heterogeneity of library size and heterogeneity of expression level between genes respectively, a normalization factor was computed (trimmed mean of M-values method, Robinson and Oshlack 2010), common and tagwise dispersions were estimated. Finally, exact tests using a negative binomial distribution were performed to contrast the two conditions (healthy vs. parasitized). Differences in numbers of expressed transcripts between the two conditions were considered as significant at the level  $\alpha = 0.05$  after a False Discovery Rate correction for multiple tests (Storey 2011).

### ***Annotation and biological processes investigation***

Annotations are functional information related to a set of genes or proteins, and that are usually regrouped into Gene Ontologies (GO, i.e., referenced functional categories, The Gene Ontology Consortium 2015). Here, we conducted a blastp query (Camacho et al. 2009) of the translated gene sequences from our reduced transcriptome assembly against the SwissProt database of referenced proteins (update of the 05/02/2015) so as to retrieve GO associated to each gene.

In order to investigate the biological processes related to health status we used a Rank Based Gene Ontology Analysis (RBGOA, Voolstra et al. 2011, Wright et al. 2015). This analysis aimed



to first cluster GO according to shared genes in order to gather the most specific GO for a set of genes, and secondly test -through Mann and Whitney U tests- whether each specific GO is significantly enriched in either over- or under-expressed genes based on their log fold change. The main advantage of RBGOA is to search for GO enriched in DEG not only in significant genes but at the whole transcriptome scale.

## **Laboratory experiment 1: Testing whether *T. polycolpus* alters the functional response of *L. burdigalensis***

### ***Sampling method***

A total of fifteen parasitized dace and eleven healthy dace of similar body size (145–245 mm, Table 1) were sampled in the Viaur and Célé rivers, and brought back to the laboratory. More specifically, five healthy dace and nine parasitized dace with an average of four parasites (range 1–9) were sampled from a single site in the Viaur River (“Saint-Just”, Table 1), and six healthy dace and six parasitized dace with an average of sixteen parasites (range 1–80) were sampled from two sites in the Célé River (“Bagnac” and “Ami du Célé”, Table 1). All fish were individually acclimatized in holding tanks at 16°C for twenty days before starting the experiments.

### ***Measure of functional responses***

Functional responses of predators correspond to the ability of an individual to deplete resources (prey) as the density of the resource increases (Holling 1959b). Here we measured the functional response of each fish as the relationship between the feeding rate (measured as the time needed to ingest the first four preys) and the density of available prey. We manipulated the density of

available prey by increasing the number of prey (i.e., from 4 to 80 prey items) provided to each fish in a given time slot.

The experiment consisted in placing a single fish in a well and continuously aerated 300 L tank filled with 200 L of dechlorinated freshwater at 16°C. In each aquarium, a refuge was placed in a corner to provide the opportunity for the target individual to hide, hence limiting potential stress during the experiment. For each fish, the experiment lasted eighteen days. Once every two days, we introduced into the aquarium 4, 8, 12, 16, 20, 30, 50, 60 or 80 alive maggots (the prey items) in a random fashion regarding the prey density so as to measure the feeding rate of each fish over a large range of prey density while taking into account the effect of satiety. Each trial was video recorded, and we measured (from the video) the time –for each prey density and each individual– needed to ingest the first four prey items, which was our measure of feeding rate.

### ***Statistical analyses***

To test whether the functional response of dace was related to the infection status of each individual, we first assessed the type of functional response of dace across the whole dataset. We compared the three most common types of functional response (types I, II & III respectively, Holling 1959a). The type I functional response is characterized by a linear relationship between the feeding rate and prey density (Holling 1959a, Jeschke et al. 2004):

$$FR = a \times D + \beta$$

Where  $FR$  stands for feeding rate,  $D$ , the prey density,  $a$ , the attack rate and  $\beta$  the intercept. The type II functional response is characterized by a logarithmic relationship between the feeding rate and prey density; for low prey densities the feeding rate increases sharply and linearly, while the feeding rate saturates to attain a plateau for high prey densities (Holling 1959b):

$$FR = \frac{a \times D}{1 + a \times h \times D}$$

Parameters estimated from the equation are the same than for type I functional response, with in addition the handling time parameter ( $h$ ). Finally, the type III functional response is similar to a type II functional response, except that for low prey densities the feeding rate increases less than linearly (Holling 1959a, Deville et al. 2013):

$$FR = \frac{a \times D^s}{1 + a \times h \times D^s}$$

Parameters are the same than for the type II functional response, except for the “shape parameter”  $s$  allowing a nonlinear relationship between  $FR$  and  $D$  for low prey density.

We fitted a parametric model linking the feeding rate to the different densities of available prey for each type of functional response (type I, II or III), and compared these models using the Akaike Information Criteria (AIC) so as to select the model that best supported the data. For the type I functional response, we adjusted a linear mixed model using the R package *nlme* (Tornøe et al. 2004). The identity of each dace was included as a random effect to account for repeated measures (one measure per prey density). To test the type II and type III functional responses, we also used the *nlme* package to fit the nonlinear relationships corresponding to type II and III functional responses respectively. More specifically, *nlme* allowed fitting functional responses (type II and III) by specifying their formula and setting each parameter ( $a$ ,  $D$ ,  $h$ , and  $s$ ) as fixed effects in addition to prey density. The identity of each dace was also included as a random effect in these two later models. Once the most likely functional response (i.e., *nlme* model) identified, we fitted this model for parasitized and healthy dace respectively and estimated the parameters (e.g. attack rate, handling time...) of the functional response of each group of dace (i.e., healthy

vs parasitized). We compared these estimates between the two groups by testing to which extent their 95% confidence intervals (CIs) overlapped.

## **Laboratory experiment 2: Testing whether *T. polycolpus* alters the competitive ability for food resources of *L. burdigalensis***

Twenty adult rostrum dace from a single site in the Viaur River (Bannes, Table 1) were sampled, brought back to the laboratory and acclimatized as described above. Our initial objective was to compare competitive ability between healthy and parasitized hosts, but due to the difficulty to sample enough healthy dace (prevalence at this period was over 90%), we compared competitive ability between weakly and highly parasitized dace. We therefore constituted ten pairs of dace, each of which comprising one healthy (or weakly) parasitized dace (average = 1.6; rang: 0–4 parasites per fish, Table 1) and one highly parasitized dace (average = 10.6; range: 7–20 parasites per fish, Table 1). Pairs of dace were constituted so that the difference in parasite number within each pair was always higher than six (range: 6–20). Each pair was introduced in a 200 L aquarium filled with 170L of dechlorinated fresh water for four days before the experiment. On the fifth day, five living maggots were introduced successively in the aquariums, and successful feeding attempts were recorded during four minutes. The experiment was repeated twice per pair of dace (four days after the first trial), and results of the two sessions were pooled per dace.

### ***Statistical analysis***

In order to test the effect of parasitic load on competitive ability for food resources, we built a Generalized Linear Mixt Model (GLMM) using the R package *glmmADMB* (Bolker et al. 2012) where the dependent variable was the food intake, the fixed effect was the health status of the fish

(highly vs weakly parasitized), and we set body size as controlled variable. The identity of the aquarium used was set as random effect. A Poisson error term distribution was used.

### **Laboratory experiment 3: Testing whether *T. polycolpus* alters the locomotion ability of *L. burdigalensis***

The same pairs of dace used for testing competitive ability for food resources were used to test for differences in locomotion performance and behavior between weakly and highly parasitized dace. Two days after the second feeding trial, the locomotion behavior of each fish was video-recorded for thirty minutes in a three-dimensional coordinate system using two cameras JVC GZ-MG20e. Over the thirty minutes of recording, only ten minutes (from the fifteenth and the twentieth minutes) were analyzed so as to minimize stress due to the settling of the camera. We recorded four parameters to describe fish locomotion behavior: the average and maximal speed, as well as maximal accelerations and decelerations (Harper and Blake 1990, Higham 2007). The software RAD Video Tools Blink (Version 1.7d) was first used to extract one image per second from the initial movie, thus resulting in six-hundred consecutive images. Then, a tracking system elaborated by Dr. Maude Combe (Centre de Recherches sur la Cognition Animal, Paul Sabatier University) was used to extract dace coordinates ( $X_t$ ,  $Y_t$ ,  $Z_t$ ) for each image. We used these coordinates to calculate the average and the maximal speed. Maximal accelerations and decelerations were measured by the maximal difference between two successive swimming speeds.

#### ***Statistical analysis***

In order to test the effect of parasitic load on locomotion abilities we first performed a Principal Component Analysis (PCA) using the R package *ade4* to describe collinearity among the four

locomotion parameters (Dray et al. 2007, Fig.S1). All four parameters were highly collinear and strongly loaded on the first component the PCA (that explain 79% of the total variance, see Fig. S1), and we extracted the coordinated of each dace on this first component as a synthetic dependent variable. We then built a GLMM (with a Gaussian error term distribution) linking this synthetic locomotion variable to dace body size and its health status (highly vs weakly parasitized) as fixed continuous effects. The aquarium identity was included as a random term.

### **Stable isotope analysis: Testing whether *T. polycolpus* modifies isotopic niche characteristics of *L. burdigalensis***

In order to compare trophic niche characteristics between parasitized and healthy *L. burdigalensis* we conducted a Stable Isotope Analysis (SIA, Newsome et al. 2009, Layman et al. 2012). We considered stable isotopes of nitrogen ( $\delta^{15}\text{N}$ ) and carbon ( $\delta^{13}\text{C}$ ) that respectively describe the trophic position of consumers (i.e., *L. burdigalensis* in this study) in the food web and the origin of the trophic resources between allochthonous and autochthonous carbon sources (i.e., litter and primary producers such as *algae* in stream ecosystems, Doucett et al. 1996, Post 2002, Fry 2006)

Sampling was carried-out during September 2009 in three sites located in the Viaur River (two sites, Navech and La Capelle, Table 1) and in the Célé River (one site, Bagnac, Table 1). Respectively, eighteen, twenty-one and nineteen dace were sampled by electro-fishing in each site along a similar river length (see Table 1 for details). Number of *T. polycolpus* on each dace as well as fin degradation level (from 0 to 4 on each fin corresponding to 0, 25, 50, 75 an 100% of the fin surface eroded by the parasite and them summed at the host level to get a total level of fin degradation, Blanchet et al. 2009) and inflammation level were recorded (from 0 to 3 on each fin and them summed at the host level to get a total level of inflammation, see Loot et al. 2004

and Fig. S2 for a picture of a damaged rostrum dace). Secondary inflammations due to bacterial infections at the site of anchorage of the parasite are often visible on fins as red spots varying in diameter (~from 1 to 7 mm). The higher the score for each spot, the redder and the larger it was. Inflammations are always scored by the same person (SB) as it is a rather objective measure. A piece of pelvic fin was collected, store in ice and frozen for SIA. Fishes were then gently returned alive in their original site. In order to control for site variation in isotopic signature in subsequent analysis (notably to assess the trophic position of dace), we also collected invertebrates generally consumed by dace. More specifically, three *Hydropsychidae* and three *Ephemeridae* were randomly collected along the stream in each site.

### ***Statistical analysis***

First values for the  $\delta^{15}\text{N}$  were corrected in order to get the trophic position of dace as followed:

$$TP=[(\delta^{15}\text{N}_{\text{dace}}-\delta^{15}\text{N}_{\text{hydropsychidae}})/3.4]+2$$

With *TP* corresponding to the trophic position of dace, 3.4 to the assimilation factor (from prey item to consumer, Post 2002) and 2 corresponding to the trophic position of *Hydropsychidae*. We then fitted a linear model (Gaussian error terms) with the trophic position of dace as the response variable. The sampling site identity (categorical), the parasitic load (number of *T. polycolpus*), a measure of fin alteration (sum of the degradation and inflammation scores at the individual level) and the two-term interactions between the sampling site identity and each of the parasite-related variables (load and alteration respectively) were included as explicative variables. Additionally, dace body size was set as covariate since body size generally highly affects trophic position in fish. We included a measure of fin alteration in addition to the parasite load in this model because isotope signatures can reflect diet up to several weeks in fish tissues (Boecklen et al. 2011). Fin

alteration is a good proxy of past exposures to *T. polycolpus* (whereas parasitic load accounts for current exposures), and we expected that it may also reflect the isotope signatures. Regarding the  $\delta^{13}\text{C}$  signatures, we used the same modelling framework than the one used for the trophic position analysis (sampling site identity (categorical), the parasitic load, fin alteration and the two-term interactions between the sampling site identity and each of the parasite-related variables (load and alteration respectively) were included as explicative variables and body size as control variable), but we used  $\delta^{13}\text{C}$  as the response variable.

The R software (version 3.4.2, R Core Team 2017) was used to run statistical analyses when no other software is mentioned.

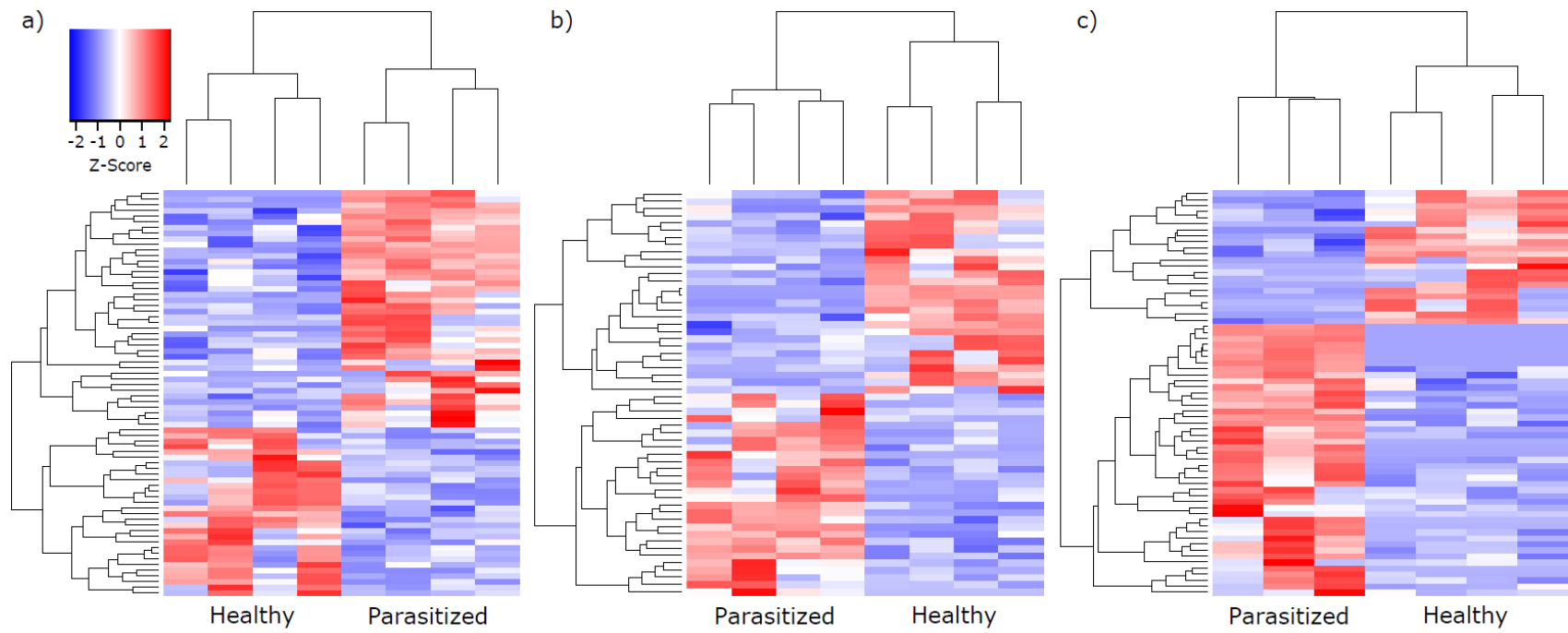
## **Results**

### **Transcriptomic study: Unraveling the molecular pathways involved in *L. burdigalensis* response to *T. polycolpus***

#### ***Genes differentially expressed and tissue specificity***

We worked with a non-redundant transcriptome of 55 545 contig length. We found respectively in the fin, the kidney and the spleen 72, 56 and 67 DEG (FDR threshold at 5%) between healthy and parasitized dace (Fig. S3). Only four DEG were shared between the three tissues (Fig. S3). More precisely, 30 genes were over-expressed (42 under-expressed) in the fin, 28 genes were over-expressed (28 under-expressed) in the cephalic kidney and 22 genes were over-expressed (45 under-expressed) in the spleen in parasitized dace compared to healthy dace (Figs. 2-a-c).





**Figure 2:** Heatmap showing gene expression profiles (higher Z-score refer to higher expression) found in the a) the fin, b) the cephalic kidney, and c) the spleen of parasitized and healthy dace.

### ***Biological processes impacted during infection***

We conducted a RBGOA separately for each tissue. Overall, 228, 74 and 632 GO significantly enriched in over or under-expressed genes were identified in the fin, the cephalic kidney and the spleen respectively (FDR threshold at 5%). To ease the biological interpretation we arbitrarily focused on the 40 most significant GO in each tissue.

In the fin, two distinct groups were identified from the top 40 significant GO (Fig.4-a). On the one hand, GO enriched in over-expressed genes in parasitized dace (compared to healthy dace) were mainly involved in immunity (eg, *innate immune response*, *defense response*, *antigen processing and presentation of exogenous antigen*, Fig. 3-a). On the other hand, GO enriched in under-expressed genes (in parasitized dace compared to healthy dace) were involved in various maintenance functions such as RNA processing (eg., *RNA splicing*, *RNA processing*, Fig. 3-a) or morphogenesis (eg, *Animal organ morphogenesis*, Fig. 3-a).

A similar dichotomy was observed for the cephalic kidney (Fig.3-b). GO enriched in over-expressed genes (in parasitized dace compared to healthy dace) were involved in processes related to cell motility and transposition (eg., *transposition*, *microtubule-based movement*, Fig. 3-b). However, surprisingly knowing that cephalic kidney is involved in immune response, GO related to immunity regulation (eg., *positive regulation of immune system process*, *regulation of cytokine production* or *regulation of programmed cell death*, Fig. 3-b) were enriched in under-expressed genes (in parasitized dace compared to healthy one, Fig. 3-b). Some other processes were also under-expressed in parasitized fishes such as *cell-division* or *positive regulation of blood circulation*.



**Figure 3:** Top 40 of the most enriched in differentially expressed genes Gene Ontologies for the comparison between parasitized and healthy dace investigated in a) their fin, b) their cephalic kidney and c) their spleen. Blue GO refer to GO enriched in under-expressed genes in parasitized dace compared to healthy dace and red GO refer to GO enriched in over-expressed genes in parasitized dace compared to healthy dace. The total number of genes involved in each GO is given in brackets.

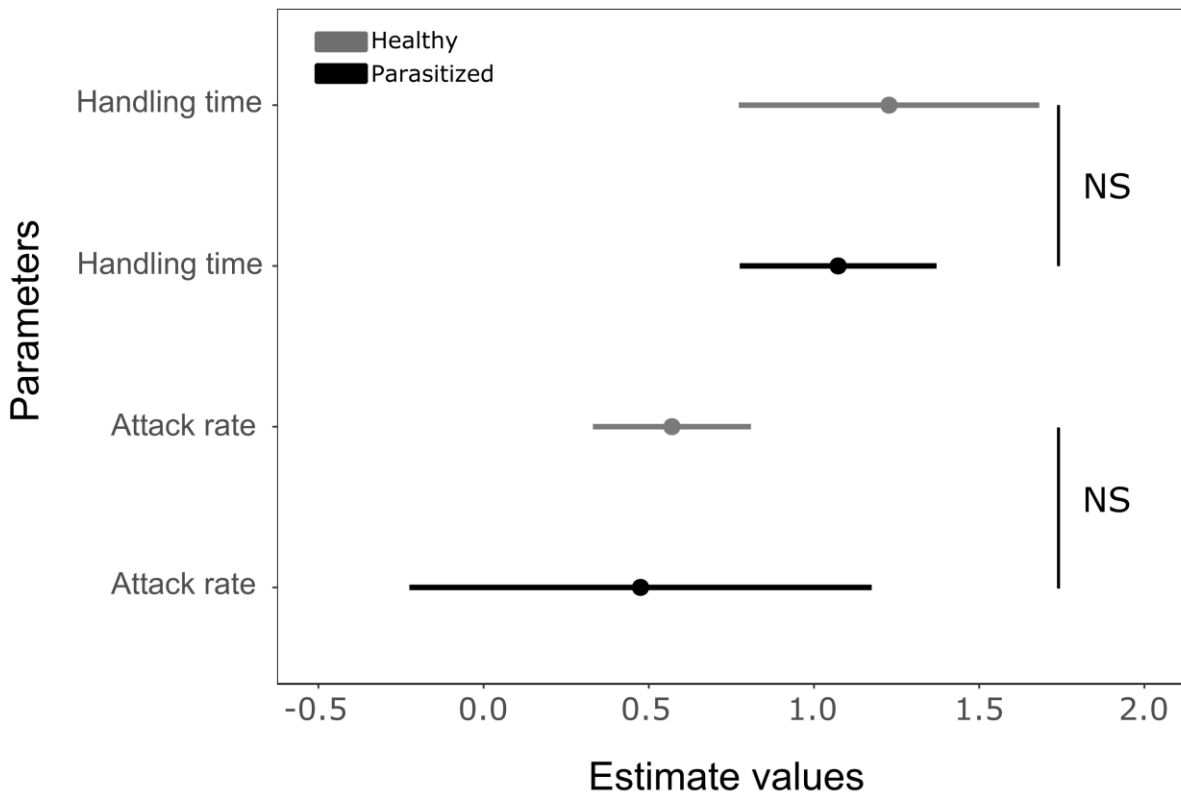
Finally, in the spleen, the top 40 most significant GO divided in two groups of similar size (Fig. 3-c). GO enriched in over-expressed genes (in parasitized dace compared to healthy dace) were related to immunity (eg., *humoral and innate immune response, regulation of immune response, antigen receptor-mediated signaling pathway*, Fig. 3-c). GO categories enriched in under-expressed genes (in parasitized dace compared to healthy dace) were mainly involved in developmental processes (eg., *regionalization, system development or regulation of cellular response to growth factor stimulus*, Fig. 3-c) including morphogenesis (*animal organ morphogenesis*, Fig. 3-c).

### **Laboratory experiment 1: Testing whether *T. polycolpus* alters the functional response of *L. burdigalensis***

Independently from the health status, the more likely functional response in dace was the type II functional response (Table 2). Thus, dace feeding rate saturates for high levels of prey density, but for low level of prey density the relation between feeding rate and prey density is linear. In the whole dataset (i.e., all dace combined) we estimated attack rate ( $a$ ) at 0.320 ( $\pm 0.139$ , p-value=0.023) and handling time ( $h$ ) at 1.068 ( $\pm 0.124$ , p-value<0.001). However we found overlapping attack rate and handling time estimates when adjusting on one hand on parasitized dace and on other hand on healthy dace (Fig. 4) thus implying similar functional responses between parasitized and healthy dace.

**Table 2:** Table showing for each functional response the formula used to model the feeding rate of dace (paratized and healthy), as well as the resulting AIC when each model was adjusted to the data.

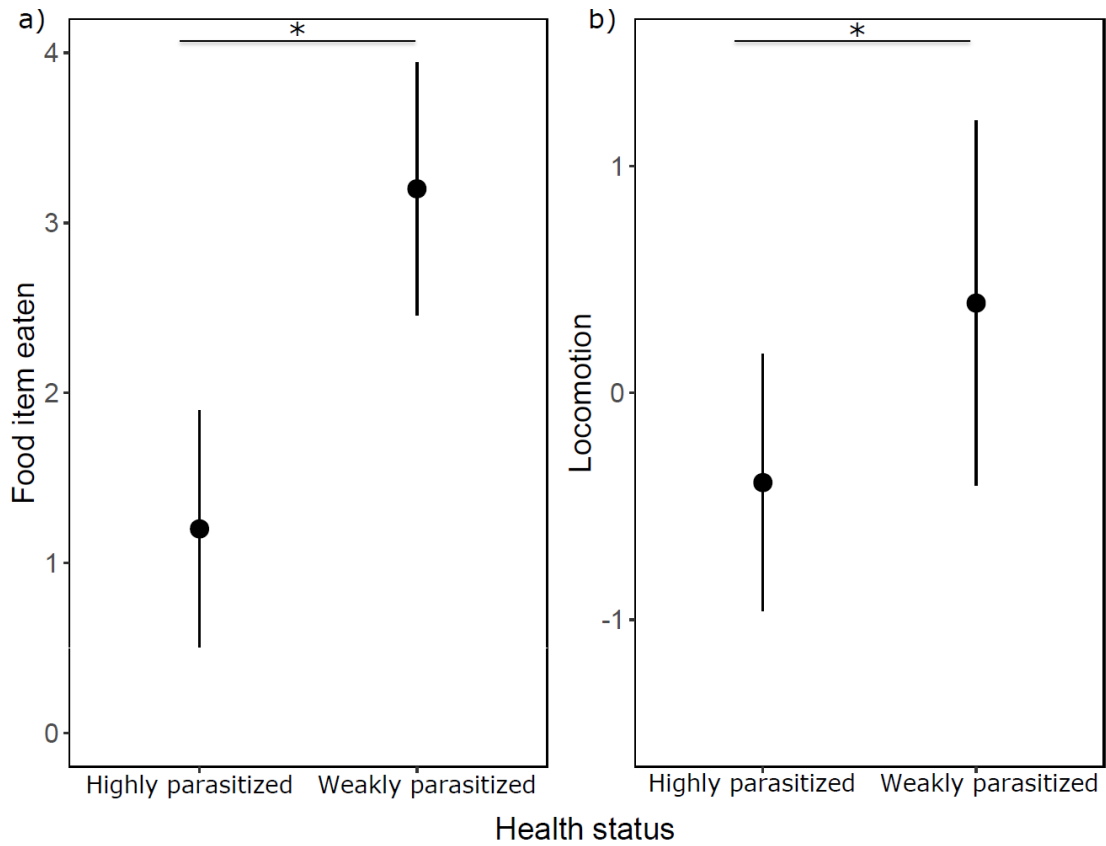
Functional response	Feeding rate formula	AIC
Type I	$a \times D + \beta$	288
Type II	$\frac{a \times D}{1 + a \times h \times D}$	270
Type III	$\frac{a \times D^s}{1 + a \times h \times D^s}$	276



**Figure 4:** Forest plot showing average estimates as well as 95% confidence intervals for the parameters of the type II functional response of parasitized and healthy dace respectively. NS stands for Non-Significant at the level  $\alpha = 0.05$ .

## **Laboratory experiments 2 and 3: Testing whether *T. polycolpus* alters competitive performances of *L. burdigalensis***

Regarding feeding ability in competing hosts, we found a significant effect of parasite load on the food intake (Chisq= 9.1346; df=1,16; p-value=0.003 and Table S1). Highly parasitized dace ate significantly less food items compared to weakly parasitized dace when competing for a limited amount of food items (Fig. 5-a). Similarly, we found that locomotion behavior in host (speed component) was significantly impacted by parasite load (F= 5.3256; df=1,15; p-value=0.036 and Table S2). More precisely, highly parasitized dace had a slower locomotion compared to weakly parasitized dace (lower speed and acceleration and higher decelerations, Fig. 5-b and Fig. S1).



**Figure 5:** Plot showing estimates along with standard errors of a) the feeding rate and b) the synthetic variable of locomotion between highly and weakly parasitized dace respectively and extracted from the corresponding GLMM. Stars indicate significant differences at the level  $\alpha = 0.05$ .

## Stable isotope analysis: Testing whether *T. polycolpus* modifies isotopic niche characteristics of *L. burdigalensis*

Regarding the trophic position, we found a significant and positive relationship between fin alteration and the trophic position of dace (Table 3 and Fig. 6-a). Dace with high fin alteration showed a significantly higher trophic position than dace with a weak or no fin alteration. We did not detect significant relationship between parasitic load and the trophic position of dace, and none of the interaction terms was significant (Table 3). Regarding the  $\delta^{13}\text{C}$ , we detected a significant interaction between the parasitic load of the dace and the site where they were sampled (Table 4 and Fig. 6-b). However, we did not detect significant relationship between fin alteration and the  $\delta^{13}\text{C}$  signature of dace (Table 4). These results suggest that the origin of the carbon source use by *L. burdigalensis* tended to vary according to the parasitic load of dace but also accordingly to the site (likely due to site specific physicochemical characteristics influencing dace carbon signature).

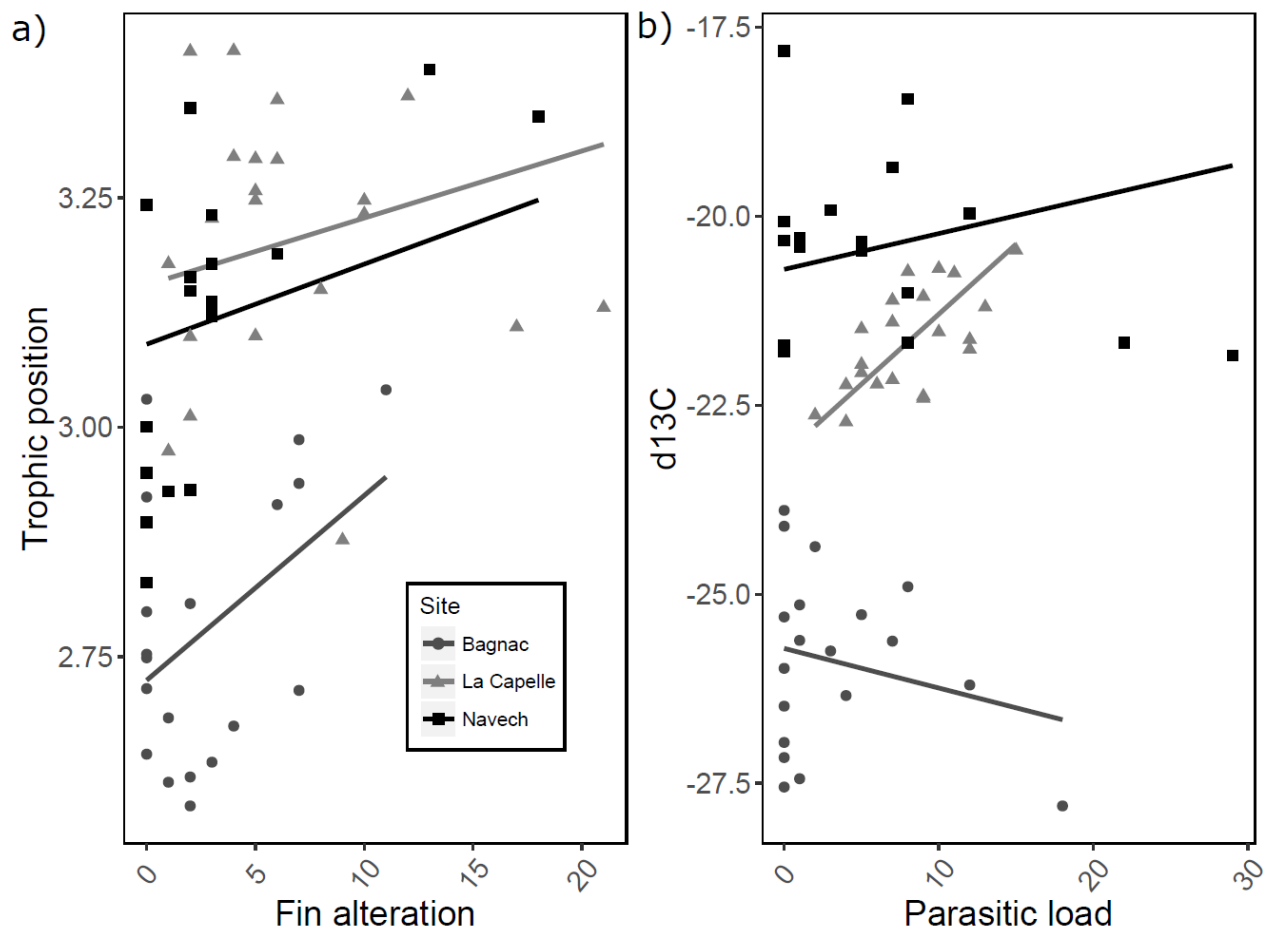
**Table 3:** Output of the linear model testing the relationship between the trophic position of dace and their parasitic load and their fin alteration caused by *T. polycolpus* and controlling for host size and site effects. Degree of freedom (Df), Fisher (F) statistics and p-values are displayed for each fixed effect.

	F	Df	P-value
Size	4.575	1,48	0.038
Parasitic load	0.107	1,48	0.745
Fin alteration	4.877	1,48	0.032
Site	43.512	2,48	<0.001
Parasitic load : Site	2.592	2,48	0.085
Fin alteration : Site	0.620	2,48	0.542



**Table 4:** Output of the linear model testing the relationship between the  $\delta^{13}\text{C}$  of dace and their parasitic load and their fin alteration caused by *T. polycolpus* and controlling for host size and site effects. Degree of freedom (Df), Fisher (F) statistics and p-values are displayed for each fixed effect.

	F	Df	P-value
Size	5.762	1,48	0.020
Parasitic load	0.649	1,48	0.424
Fin alteration	0.435	1,48	0.513
Site	188.677	2,48	<0.001
Parasitic load : Site	4.349	2,48	0.018
Fin alteration : Site	2.835	2,18	0.069



**Figure 6:** Graphical representation of the relationships between the a) trophic position and fin alteration of dace and b) the  $\delta^{13}\text{C}$  and the parasitic load of dace, in the three sites the SIA was conducted. Points, squares and triangles refer to observations, lines refer to predictions from the associated GLMMs while keeping other variables than fin alteration or parasitic load fixed to their mean respectively.

## Discussion

In this study we aimed to conduct an integrative assessment of the impact of a parasite on its main host from a molecular to an ecosystemic perspective. Overall, our study suggests that the ectoparasite *T. polycolpus* leads to altered patterns of gene expression in parasitized hosts compared to healthy hosts, with –as expected– immune-related genes being strongly over-expressed in two out of the three tissues. This parasite-driven modulation of host transcriptomic was not associated with notable change in host feeding behavior when hosts were isolated. However, we detected a significant reduction of feeding and locomotion abilities of infected hosts in a context in which intraspecific competition for food was elicited. In agreement with this latter result, we also identified shifts in the isotope signature of parasitized hosts compared to that of healthy hosts *in natura*. This potential shift in trophic niche of parasitized hosts in the wild may scale up at the ecosystem level. In the following paragraphs, we develop further these parasite effects on hosts at difference levels of organization, from molecular to ecosystems.

### **Molecular level: proximate response of hosts to parasitism**

The first expected proximate response of hosts to parasitism is a modulation of their gene expression. Accordingly, we showed that infection by *T. polycolpus* leads to general transcriptomic adjustments in parasitized hosts. In particular, parasitized hosts differentially expressed several genes compared to healthy hosts not only in the tissue in which the parasite is anchored (i.e., the fin) but also in two tissues devoted to immunity in fish (the cephalic kidney and the spleen, Alvarez-Pellitero 2008). This suggests that *T. polycolpus* induce a general rather than a localized immune response of their hosts. Interestingly, we did find a tissue specific response with few common significant DEGs between tissues but these genes were shown to be

involved in similar biological processes (i.e., immunity and various maintenance functions such as morphogenesis and development). We thus suggest that despite being an ectoparasite, *T. polycolpus* have a profound impact on its host at the molecular level: *Tracheliastes polycolpus* triggers a generalized response which suggests strong pathogenic effects, probably also exacerbated by the naivety of local hosts facing an invasive parasite (Lee and Klasing 2004, Charbonnel et al. 2019).

More precisely, we suggest that parasitized hosts mostly display an activation of immunity associated to a repression of maintenance processes comparatively to healthy hosts. A consistent pattern was indeed observed in the fin and the spleen of parasitized dace with GO enriched in over-expressed genes involved in immune response, whereas GO enriched in under-expressed genes were involved in morphogenesis and development. For instance, GO enriched in over-expressed genes in the fin suggested a role of interleukin-1, I-kappa B and interferons-gamma pathways (Fig. 3-a and Fig. 3 -c) that are known to be involved in inflammatory responses and acute phase of infection (Verma et al. 1995, Jéru and Amselem 2011, Miller et al. 2014). Similarly, GO enriched in over-expressed genes in the spleen of parasitized dace referred to the activation of the complement that participates in inflammation but also enhances the adaptive immune response (Carroll 2004). A systemic inflammatory response in parasitized hosts is particularly relevant in *T. polycolpus*-*L. leuciscus* system since this parasite can cause virulent inflammations when anchoring on fins, usually because of the co-development of opportunistic bacteria (Loot et al. 2004). In parallel, maintenance processes (morphogenesis and development) were enriched in under-expressed genes in the fin and the spleen of parasitized dace. Although, processes enriched in over or under-expressed genes not necessary means that this related process is up or down-regulated (du to potential repressor genes being involved), maintenance processes

being enriched in under-expressed genes in parasitized dace echoes previous results showing that parasitized dace have a reduced growth rate compared to healthy ones (Blanchet et al. 2009a). Several studies suggest that activating an immune response is costly regarding energetic demands (Soler et al. 2003, Chetouhi et al. 2015), and even more costly for naïve hosts facing a new pathogen such as in this case (Lee and Klasing 2004), although potential negative consequences on other traits or processes are still debated (Lochmiller and Deerenberg 2000). We consequently suggest that over-expression of immune related genes associated to down-expression of maintenance related genes in the fin and spleen of parasitized dace could indicate a re-allocation of energy from maintenance functions associated with growth (i.e., development, morphogenesis, RNA processing) toward immunity (Lochmiller and Deerenberg 2000).

Interestingly, an opposite pattern was observed in the cephalic kidney of parasitized fish with immune-related processes being enriched in under-expressed genes. At a first sight, such a down expression of genes involved in immunity in an immune devoted organ during parasite infection might seem surprising. However, a similar repression of immunity in the cephalic kidney was already observed in turbot suffering enteromyxosis (Robledo et al. 2014). Robledo et al (2014) interpreted such a repression of immunity in infested hosts as an evidence of immune evasion by the parasite. In our case, repression of immune processes in the cephalic kidney of parasitized dace could as well reflect some immunity disruption caused by the parasite itself or potentially by some associated pathogens such as virus or bacteria (see GO referring to *response to bacterium* and *regulation of viral release from host cell*, Fig. 3-a and Loot et al. 2004). We hence argue that *T. polycolpus* could either directly or indirectly lead to a disruption of immune response or inadequate immune response in the cephalic kidney of parasitized hosts.

Overall, regarding our first objective (Fig. 1) aiming to unravel host response to *T. polycolpus* at the molecular level, we suggest that infection by *T. polycolpus* leads to global transcriptomic adjustments by its host. Notably, we argue that parasitized hosts display a strong immune response and a possible disruption of this immune response in some tissue. Such a global immunity related response could have potential cost associated to maintenance processes that we expect to detect at other level of biological organization.

### **Individual and population levels: direct vs indirect parasite-mediated impacts on host**

We did not detect a direct impact of *T. polycolpus* on its host at the strict individual level but we showed an impact of the parasite on its host when this later was confronted to competition with a conspecific and under natural (and hence variable) conditions. Together these results suggest that *T. polycolpus* induces an indirect rather than direct impact on its individual hosts. Our first experimental approach (objective O2, Fig. 1) indeed revealed that *T. polycolpus* did not significantly alter the functional response of isolated dace. More precisely, neither the attack rate –usually interpreted as the volume of water inspected for a certain amount of time (Murray et al. 2013)–, nor the handling time –which refers to the time to process a prey (handling, ingestion, digestion and resting during satiated condition, Jeschke et al. 2002)– appeared to be impacted by parasitism. Thus, *T. polycolpus* does not impact individual feeding performance of its hosts at least for these specific traits. These results differ from previous empirical observations that parasitized hosts are less efficient at feeding, notably by displaying lower attack rate and higher handling time (Haddaway et al. 2012, Toscano et al. 2014, Leaphart and Zelmer 2017). Conversely, when accounting for intra-specific competition at the population level, we did show that highly parasitized hosts displayed lower feeding competitive ability (i.e., lower feeding rate)

as well as a slower locomotion compared to weakly parasitized hosts (objective O3, Fig. 1). We thus suggest that *T. polycolpus* more likely impact host competitive performances rather than individual host performances *per se*. Sandland and Minchella (2003) suggested that the impacts of parasites on host traits, notably associated to the cost of host defense against parasite, could depend not only on abiotic but also on biotic environmental factors, including competition and predation (i.e., interactions occurring at the population or the community level). Here we demonstrate and argue that the effect of an emerging parasite on the feeding behavior and the locomotion of their hosts depends on intra-specific interactions at the population level (Bedhomme et al. 2005, Vivas Muñoz et al. 2017).

Finally, our results also suggest that *T. polycolpus* induces some effects beyond the population scale and that could affect the ecosystem by modifying the trophic niche of their hosts (objective O4, Fig. 1). In particular, using a Stable Isotope Analysis, we first showed that previously parasitized dace (since we used fin alteration as a proxy of past infection) had a higher hierarchical trophic level compared to healthy dace. Generally, a higher trophic position in the trophic network usually reflects an elevation in this network by eating prey from higher trophic levels (Post 2002). However, in a host-parasite context two (non-exclusive) hypotheses were proposed to explain such a higher trophic position of parasitized hosts compared to healthy individuals (Sánchez et al. 2013, Pulkkinen et al. 2016). First, Pulkkinen et al. (2016) suggested that parasites might have a direct effect on their hosts by altering their nutrients assimilation since they also deplete host tissues as resource for their own. For instance, Pulkkinen et al. (2016) found a direct effect of a microsporidian parasite on the isotopic signature of their hosts *Daphnia* in which the competitive performances of parasitized hosts were not impacted. Second, parasites might induce indirect effects on the trophic position of their hosts lowering their ability to deplete resources compared to healthy hosts (Pulkkinen et al. 2016). Indeed, when lacking of food supply

organism usually catalyze their own proteins, hence resulting in a higher signature for  $\delta^{15}\text{N}$ , and consequently a higher trophic position (Gaye-Siessegger et al. 2004, Haubert et al. 2005). In our case, indirect effects of *T. polycolpus* on dace trophic position seems to be a more likely hypothesis since –as shown here– highly parasitized hosts were less able to feed when competing with conspecifics. We found similar results regarding the isotopic signatures of carbon: The  $\delta^{13}\text{C}$  tends to vary according to the infestation level by *T. polycolpus*, hence suggesting that parasitized dace are likely using different preys. Although, once again such a result could be explained by direct parasite impact on host resource assimilation, the use of different prey items could similarly result from competitive exclusion by healthy conspecifics (Miura et al. 2006, Britton et al. 2011, Sánchez et al. 2013). Overall the Stable Isotope Analysis, interpreted congruently with results of the objective O3 (Fig. 1), tends to suggest that *T. polycolpus* could lead to different trophic niches between parasitized and healthy hosts such as use of less and/or different prey species. In the light of these results, it is thus likely that *T. polycolpus* induces modifications on its host that could impact the overall ecosystem functioning that would be worthy to investigate (Bernstein et al. 2002, Wood et al. 2007, Sato et al. 2012).

### **Hypotheses regarding *T. polycolpus* multi-level impacts on its host**

When interpreting our results as a whole, we hereafter propose an integrative hypothesis regarding causal links of *T. polycolpus* impact *L. burdigalensis*. *Tracheliastes polycolpus* infection probably triggers a proximate molecular response in its host notably the activation of the acute phase of host immune response. Host response likely requires a reallocation of energy from development toward immunity with a cost reinforced by parasite possible strategy to disrupt host immune response. We argue that this reallocation of energy normally invested in maintenance processes (as evidenced by lower growth rate of parasitized dace, Blanchet et al.

2009) could be translated in indirect parasite-mediated impact notably on host competitive performances (Hudson 1998). Finally indirect parasite-mediated effects on host competitive performances are likely to have cascading impacts on host trophic niche which could ultimately affect the ecosystem functioning. We consequently propose that even non-lethal parasites, such as *T. polycolpus*, could have (cryptic) profound impacts on the receiving ecosystems without displaying visible impact on their hosts at the individual level due to cascading effects.

Although our hypothesis of potential cascading effects linking parasite impact on its host through different levels of biological organization remains to be tested, we evidenced and argue that to fully investigate and measure parasite impact on its host, contextual interactions have to be taken into account (Sandland and Minchella 2003). Many parasites trigger an immune response in their host (except parasite relying on immune escape strategies, such as some malaria parasites, Gomes et al. 2016). Hence, a cost of such an immune response observed at broader biological levels of organization is likely a common associated cost of parasitism rather than a specificity of *T. polycolpus*. However, it is also worth noting that *T. polycolpus* is an emerging parasite that only arrived in French River Basins in 1920s through translocation of its original host the Common Ide (Rey et al. 2015a). *Tracheliastes polycolpus* is thus lacking of a long term co-evolution history with *L. burdigalensis* which could notably explain the strong and systemic immune response we observed in parasitized hosts (Lee and Klasing 2004). Hence, whether such indirect parasite impacts at broad organization levels triggered by a strong host immune response is a specificity of emerging parasites or a common characteristic of many parasites remains an open question. We consequently stand for more multi-levels studies on different host-parasite systems, including both emerging and native parasites, in order to better understand in which extent the contextual interactions such as competition or trophic interactions can exacerbate the impact of parasites on host populations, communities and ecosystem functioning.



## Conclusion

Overall, we conducted here one of the first multi-level study that considered parasite impact on host from the molecular level to the population level (but see Bankers et al. 2017). We progressively approached *in natura* conditions by successively evaluating parasite impact on host when considering no interaction, conspecific interaction and finally trophic interaction (Fig. 1). Doing so, we identified both proximate host response and indirect parasite-mediated impacts on host when conspecific and trophic interactions were included. We believe that this work, showing that parasite impacts goes from host gene expression to host performance in ecosystems, allowed (and hopefully will allow) formulating more realistic hypothesis to explain cascading impacts of parasites on hosts and ultimately ecosystem functioning.

## References

- Adamo, S. A. 2002. Modulating the Modulators: Parasites, Neuromodulators and Host Behavioral Change. *Brain, Behavior and Evolution* 60:370–377.
- Adamo, S. A. 2013. Parasites: evolution’s neurobiologists. *Journal of Experimental Biology* 216:3–10.
- Agnew, P., J. C. Koella, and Y. Michalakis. 2000. Host life history responses to parasitism. *Microbes and Infection* 2:891–896.
- Alvarez-Pellitero, P. 2008. Fish immunity and parasite infections: from innate immunity to immunoprophylactic prospects. *Veterinary Immunology and Immunopathology* 126:171–198.
- Bankers, L., P. Fields, K. E. McElroy, J. L. Boore, J. M. Logsdon, and M. Neiman. 2017. Genomic evidence for population-specific responses to co-evolving parasites in a New Zealand freshwater snail. *Molecular Ecology* 26:3663–3675.
- Bedhomme, S., P. Agnew, Y. Vital, C. Sidobre, and Y. Michalakis. 2005. Prevalence-Dependent Costs of Parasite Virulence. *PLoS Biology* 3:e262.
- Bernstein, C., A. Heizmann, and E. Desouhant. 2002. Intraspecific competition between healthy and parasitised hosts in a host-parasitoid system: consequences for life-history traits. *Ecological Entomology* 27:415–423.

- Blanchet, S., L. Méjean, J.-F. Bourque, S. Lek, F. Thomas, D. J. Marcogliese, J. J. Dodson, and G. Loot. 2009a. Why do parasitized hosts look different? Resolving the “chicken-egg” dilemma. *Oecologia* 160:37–47.
- Blanchet, S., O. Rey, P. Berthier, S. Lek, and G. Loot. 2009b. Evidence of parasite-mediated disruptive selection on genetic diversity in a wild fish population. *Molecular Ecology* 18:1112–1123.
- Boecklen, W. J., C. T. Yarnes, B. A. Cook, and A. C. James. 2011. On the Use of Stable Isotopes in Trophic Ecology. *Annual Review of Ecology, Evolution, and Systematics* 42:411–440.
- Bolker, B., H. Skaug, A. Magnusson, and A. Nielsen. 2012. Getting started with the glmmADMB package.
- Britton, J. R., J. Pegg, and C. F. Williams. 2011. Pathological and Ecological Host Consequences of Infection by an Introduced Fish Parasite. *PLoS ONE* 6:e26365.
- Camacho, C., T. Madden, G. Coulouris, V. Avagyan, N. Ma, T. Tao, and R. Agarwala. 2009. Manuel de l'utilisateur des applications BLAST en ligne de commande (BLAST Command Line Applications User Manual).
- Carroll, M. C. 2004. The complement system in regulation of adaptive immunity. *Nature Immunology* 5:981–986.
- Charbonnel, N., M. Galan, C. Tatard, A. Loiseau, C. A. Diagne, A. Dalecky, H. Parrinello, S. Rialle, D. Severac, and C. Brouat. 2019. Differential immune gene expression associated with contemporary range expansion of two invasive rodents in Senegal. *bioRxiv*.
- Chetouhi, C., J. Panek, L. Bonhomme, H. ElAlaoui, C. Texier, T. Langin, C. de Bekker, S. Urbach, E. Demetree, D. Missé, P. Holzmuller, D. P. Hughes, A. Zanzoni, C. Brun, and D. G. Biron. 2015. Cross-talk in host–parasite associations: What do past and recent proteomics approaches tell us? *Infection, Genetics and Evolution* 33:84–94.
- Deville, A.-S., D. Grémillet, M. Gauthier-Clerc, M. Guillemain, F. Von Houwald, B. Gardelli, and A. Béchet. 2013. Non-linear feeding functional responses in the Greater Flamingo ( *Phoenicopterus roseus* ) predict immediate negative impact of wetland degradation on this flagship species. *Ecology and Evolution* 3:1413–1425.
- Doucett, R. R., G. Power, D. R. Barton, R. J. Drimmie, and R. A. Cunjak. 1996. Stable isotope analysis of nutrient pathways leading to Atlantic salmon. *Canadian Journal of Fisheries and Aquatic Sciences* 53:2058–2066.
- Dray, S., A.-B. Dufour, and others. 2007. The ade4 package: implementing the duality diagram for ecologists. *Journal of statistical software* 22:1–20.
- Eldem, V., G. Zararsiz, T. Taşçi, I. P. Duru, Y. Bakir, and M. Erkan. 2017. Transcriptome Analysis for Non-Model Organism: Current Status and Best-Practices. Page in F. A. Marchi, P. D. R. Cirillo, and E. C. Mateo, editors. *Applications of RNA-Seq and Omics Strategies - From Microorganisms to Human Health*.

- Feldmeyer, B., J. Mazur, S. Beros, H. Lerp, H. Binder, and S. Foitzik. 2016. Gene expression patterns underlying parasite-induced alterations in host behaviour and life history. *Molecular Ecology* 25:648–660.
- Fry, B. 2006. *Stable isotope ecology*. Springer.
- Gaye-Siessegger, J., U. Focken, S. Muetzel, H. Abel, and K. Becker. 2004. Feeding level and individual metabolic rate affect  $\delta^{13}\text{C}$  and  $\delta^{15}\text{N}$  values in carp: implications for food web studies. *Oecologia* 138:175–183.
- Genomic Resources Development Consortium, S. Blanchet, O. Bouchez, C. A. Chapman, P. D. Etter, T. L. Goldberg, E. A. Johnson, J. H. Jones, G. Loot, P. A. Omeja, O. Rey, M. J. Ruiz-Lopez, W. M. Switzer, and N. Ting. 2015. *Molecular Ecology Resources* 15:684–684.
- Gomes, P. S., J. Bhardwaj, J. Rivera-Correa, C. G. Freire-De-Lima, and A. Morrot. 2016. Immune Escape Strategies of Malaria Parasites. *Frontiers in Microbiology* 7: 1617.
- Haas, B. J., A. Papanicolaou, M. Yassour, M. Grabherr, P. D. Blood, J. Bowden, M. B. Couger, D. Eccles, B. Li, M. Lieber, M. D. MacManes, M. Ott, J. Orvis, N. Pochet, F. Strozzi, N. Weeks, R. Westerman, T. William, C. N. Dewey, R. Henschel, R. D. LeDuc, N. Friedman, and A. Regev. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols* 8:1494–1512.
- Haddaway, N. R., R. H. Wilcox, R. E. A. Heptonstall, H. M. Griffiths, R. J. G. Mortimer, M. Christmas, and A. M. Dunn. 2012. Predatory Functional Response and Prey Choice Identify Predation Differences between Native/Invasive and Parasitised/Unparasitised Crayfish. *PLoS ONE* 7:e32229.
- Harper, D. G., and R. W. Blake. 1990. Fast-start performance of rainbow trout *Salmo gairdneri* and northern pike *Esox lucius*. *Journal of Experimental Biology* 150:321–342.
- Hatcher, M. J., J. T. Dick, and A. M. Dunn. 2012. Diverse effects of parasites in ecosystems: linking interdependent processes. *Frontiers in Ecology and the Environment* 10:186–194.
- Haubert, D., R. Langel, S. Scheu, and L. Ruess. 2005. Effects of food quality, starvation and life stage on stable isotope fractionation in *Collembola*. *Pedobiologia* 49:229–237.
- Higham, T. E. 2007. Feeding, fins and braking maneuvers: locomotion during prey capture in centrarchid fishes. *Journal of Experimental Biology* 210:107–117.
- Holling, C. S. 1959a. The Components of Predation as Revealed by a Study of Small-Mammal Predation of the European Pine Sawfly1. *The Canadian Entomologist* 91:293–320.
- Holling, C. S. 1959b. Some Characteristics of Simple Types of Predation and Parasitism. *The Canadian Entomologist* 91:385–398.
- Horký, P., K. Douda, M. Maciak, L. Závorka, and O. Slavík. 2014. Parasite-induced alterations of host behaviour in a riverine fish: the effects of glochidia on host dispersal. *Freshwater Biology* 59:1452–1461.

- Huang, Y., B. Niu, Y. Gao, L. Fu, and W. Li. 2010. CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics* 26:680–682.
- Hudson, P. 1998. Competition mediated by parasites: biological and theoretical progress. *Trends in Ecology & Evolution* 13:387–390.
- Jéru, I., and S. Amselem. 2011. Inflammasome et interleukine 1. *La Revue de Médecine Interne* 32:218–224.
- Jeschke, J. M., M. Kopp, and R. Tollrian. 2002. Predator functional responses: Discriminating between handling and digesting prey. *Ecological Monographs* 72:95–112.
- Jeschke, J. M., M. Kopp, and R. Tollrian. 2004. Consumer-food systems: why type I functional responses are exclusive to filter feeders. *Biological Reviews* 79:337–349.
- Keith, P., H. Persat, E. Feunteun, and J. Allardi. 2011. *Les Poissons d'eau douce de France*. Biotope Editions.
- Lafferty, K. D., and A. K. Morris. 1996. Altered Behavior of Parasitized Killifish Increases Susceptibility to Predation by Bird Final Hosts. *Ecology* 77:1390–1397.
- Langmead, B., and S. L. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9:357–359.
- Layman, C. A., M. S. Araujo, R. Boucek, C. M. Hammerschlag-Peyer, E. Harrison, Z. R. Jud, P. Matich, A. E. Rosenblatt, J. J. Vaudo, L. A. Yeager, D. M. Post, and S. Bearhop. 2012. Applying stable isotopes to examine food-web structure: an overview of analytical tools. *Biological Reviews* 87:545–562.
- Leaphart, J. C., and D. A. Zelmer. 2017. Wrecking the Curve: Altered Functional Response of *Tetragoneuria* (Odonata: Corduliidae) Naiads Infected with Metacercariae of *Haematoloechus floedae*. *Journal of Parasitology* 103:147–151.
- Lee, K. A., and K. C. Klasing. 2004. A role for immunology in invasion biology. *Trends in Ecology & Evolution* 19:523–529.
- Lehmann, T. 1993. Ectoparasites: Direct impact on host fitness. *Parasitology Today* 9:8–13.
- Li, B., and C. N. Dewey. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC bioinformatics* 12:323.
- Lochmiller, R. L., and C. Deerenberg. 2000. Trade-offs in evolutionary immunology: just what is the cost of immunity? *Oikos* 88:87–98.
- Loot, G., N. Poulet, Y. Reyjol, S. Blanchet, and S. Lek. 2004. The effects of the ectoparasite *Tracheliastes polycolpus* (Copepoda: Lernaeopodidae) on the fins of rostrum dace (*Leuciscus leuciscus burdigalensis*). *Parasitology Research* 94:16–23.
- Mathieu-Bégné, E., G. Loot, M. Chevalier, I. Paz-Vinas, and S. Blanchet. 2019. Demographic and genetic collapses in spatially structured populations: insights from a long-term survey in wild fish metapopulations. *Oikos* 128:196–207.

- McTaggart, S. J., T. Cézard, J. S. Garbutt, P. J. Wilson, and T. J. Little. 2015. Transcriptome profiling during a natural host-parasite interaction. *BMC Genomics* 16: 643–652.
- Miller, J. L., B. K. Sack, M. Baldwin, A. M. Vaughan, and S. H. I. Kappe. 2014. Interferon-Mediated Innate Immune Responses against Malaria Parasite Liver Stages. *Cell Reports* 7:436–447.
- Miura, O., A. M. Kuris, M. E. Torchin, R. F. Hechinger, and S. Chiba. 2006. Parasites alter host phenotype and may create a new ecological niche for snail hosts. *Proceedings of the Royal Society B: Biological Sciences* 273:1323–1328.
- Mouritsen, K. N., and R. Poulin. 2005. Parasites boosts biodiversity and changes animal community structure by trait-mediated indirect effects. *Oikos* 108:344–350.
- Murray, G. P. D., R. A. Stillman, R. E. Gozlan, and J. R. Britton. 2013. Experimental Predictions of The Functional Response of A Freshwater Fish. *Ethology* 119:751–761.
- Newsome, S. D., M. T. Tinker, D. H. Monson, O. T. Oftedal, K. Ralls, M. M. Staedler, M. L. Fogel, and J. A. Estes. 2009. Using stable isotopes to investigate individual diet specialization in California sea otters (*Enhydra lutris nereis*). *Ecology* 90:961–974.
- Ontoria, E., Y. E. Hernández-Santana, A. C. González-García, M. C. López, B. Valladares, and E. Carmelo. 2018. Transcriptional Profiling of Immune-Related Genes in *Leishmania infantum*-Infected Mice: Identification of Potential Biomarkers of Infection and Progression of Disease. *Frontiers in Cellular and Infection Microbiology* 8: 197.
- Post, D. M. 2002. Using stable isotopes to estimate trophic position: models, methods, and assumptions. *Ecology* 83:703–718.
- Pulkkinen, K., S. L. Aalto, and H. Nykänen. 2016. Parasite infection alters host stable-isotope composition under controlled feeding. *Freshwater Biology* 61:1981–1990.
- R Core Team. 2017. R: A language and environment for statistical computing. R foundation for statistical Computing, Vienna, Austria.
- Rey, O., L. Fournet, I. Paz-Vinas, G. Loot, C. Veyssi re, B. Roche, and S. Blanchet. 2015a. Elucidating the spatio-temporal dynamics of an emerging wildlife pathogen using approximate Bayesian computation. *Molecular Ecology* 24:5348–5363.
- Rey, O., G. Loot, O. Bouchez, and S. Blanchet. 2015b. Data from: De novo assembly transcriptome for the rostrum dace (*Leuciscus burdigalensis*, Cyprinidae: fish) naturally infected by a copepod ectoparasite. Dryad Digital Repository.
- Richman, A. M. 1997. Plasmodium activates the innate immune response of *Anopheles gambiae* mosquitoes. *The EMBO Journal* 16:6114–6119.
- Robinson, M. D., D. J. McCarthy, and G. K. Smyth. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26:139–140.

Robinson, M. D., and A. Oshlack. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome biology* 11:R25.

Robledo, D., P. Ronza, P. W. Harrison, A. P. Losada, R. Bermúdez, B. G. Pardo, M. J. Redondo, A. Sitjà-Bobadilla, M. I. Quiroga, and P. Martínez. 2014. RNA-seq analysis reveals significant transcriptome changes in turbot (*Scophthalmus maximus*) suffering severe enteromyxosis. *Bmc Genomics* 15:1149.

Ronza, P., D. Robledo, R. Bermúdez, A. P. Losada, B. G. Pardo, A. Sitjà-Bobadilla, M. I. Quiroga, and P. Martínez. 2016. RNA-seq analysis of early enteromyxosis in turbot (*Scophthalmus maximus*): new insights into parasite invasion and immune evasion strategies. *International Journal for Parasitology* 46:507–517.

Sánchez, M. I., N. Varo, C. Matesanz, C. Ramo, J. A. Amat, and A. J. Green. 2013. Cestodes change the isotopic signature of brine shrimp, *Artemia*, hosts: Implications for aquatic food webs. *International Journal for Parasitology* 43:73–80.

Sandland, G. J., and D. J. Minchella. 2003. Costs of immune defense: an enigma wrapped in an environmental cloak? *Trends in Parasitology* 19:571–574.

Sato, T., T. Egusa, K. Fukushima, T. Oda, N. Ohte, N. Tokuchi, K. Watanabe, M. Kanaiwa, I. Murakami, and K. D. Lafferty. 2012. Nematomorph parasites indirectly alter the food web and ecosystem function of streams through behavioural manipulation of their cricket hosts. *Ecology Letters* 15:786–793.

Schmid Hempel, P. 2011. Evolutionary parasitology: The integrated study of infections, immunology, ecology, and genetics.

Soler, J. J., L. de Neve, T. Pérez-Contreras, M. Soler, and G. Sorci. 2003. Trade-off between immunocompetence and growth in magpies: an experimental study. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270:241–248.

Storey, J. D. 2011. False discovery rate. Pages 504–508 *International encyclopedia of statistical science*. Springer.

The Gene Ontology Consortium. 2015. Gene Ontology Consortium: going forward. *Nucleic Acids Research* 43:D1049–D1056.

Thompson, S. N., and M. Kavaliers. 1994. Physiological bases for parasite-induced alterations of host behaviour. *Parasitology* 109:S119–S138.

Tornøe, C. W., H. Agersø, E. N. Jonsson, H. Madsen, and H. A. Nielsen. 2004. Non-linear mixed-effects pharmacokinetic/pharmacodynamic modelling in NLME using differential equations. *Computer Methods and Programs in Biomedicine* 76:31–40.

Toscano, B. J., B. Newsome, and B. D. Griffen. 2014. Parasite modification of predator functional response. *Oecologia* 175:345–352.

Verma, I. M., J. K. Stevenson, E. M. Schwarz, and D. Van Antwerp. 1995.  $\text{R}^{\sim\sim}$ /NF-KB/IKB family: intimate tales of association and dissociation. *Genes Dev*:2723–35.

Vivas Muñoz, J. C., G. Staaks, and K. Knopf. 2017. The eye fluke *Tylodelphys clavata* affects prey detection and intraspecific competition of European perch (*Perca fluviatilis*). *Parasitology Research* 116:2561–2567.

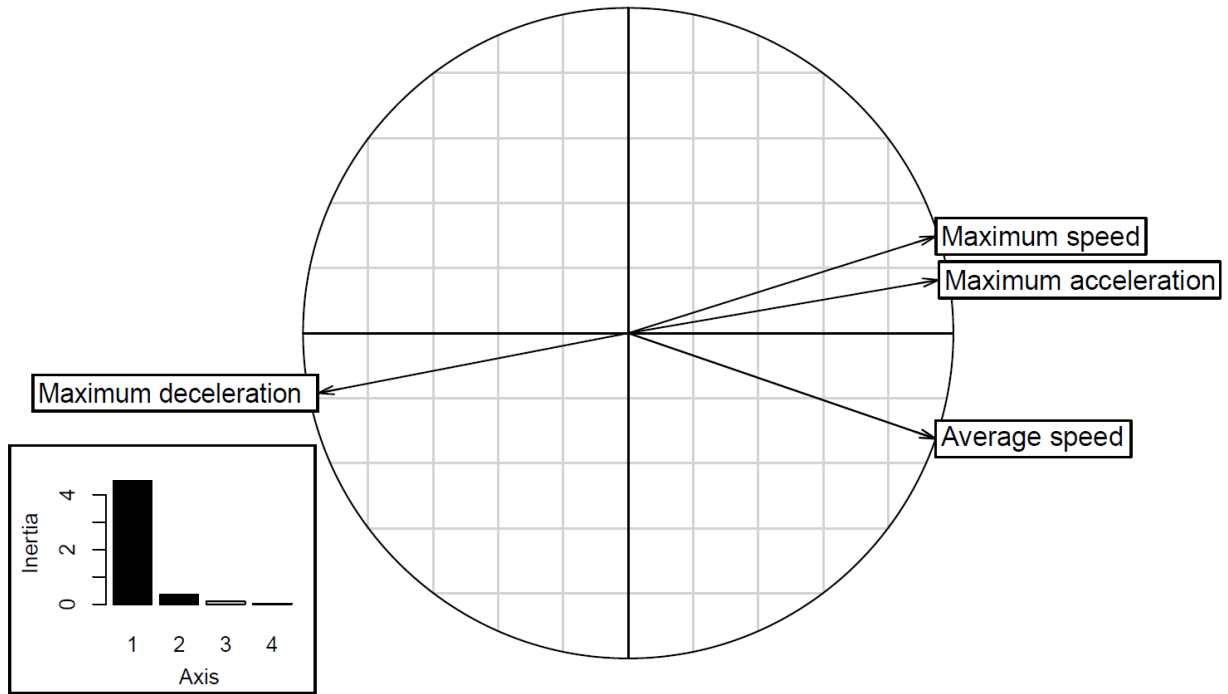
Voolstra, C. R., S. Sunagawa, M. V. Matz, T. Bayer, M. Aranda, E. Buschiazzi, M. K. DeSalvo, E. Lindquist, A. M. Szmant, M. A. Coffroth, and M. Medina. 2011. Rapid Evolution of Coral Proteins Responsible for Interaction with the Environment. *PLoS ONE* 6:e20392.

Wolinska, J., B. Keller, K. Bittner, S. Lass, and P. Spaak. 2004. Do parasites lower *Daphnia* hybrid fitness? *Limnology and Oceanography* 49:1401–1407.

Wood, C. L., J. E. Byers, K. L. Cottingham, I. Altman, M. J. Donahue, and A. M. H. Blakeslee. 2007. Parasites alter community structure. *Proceedings of the National Academy of Sciences* 104:9335–9339.

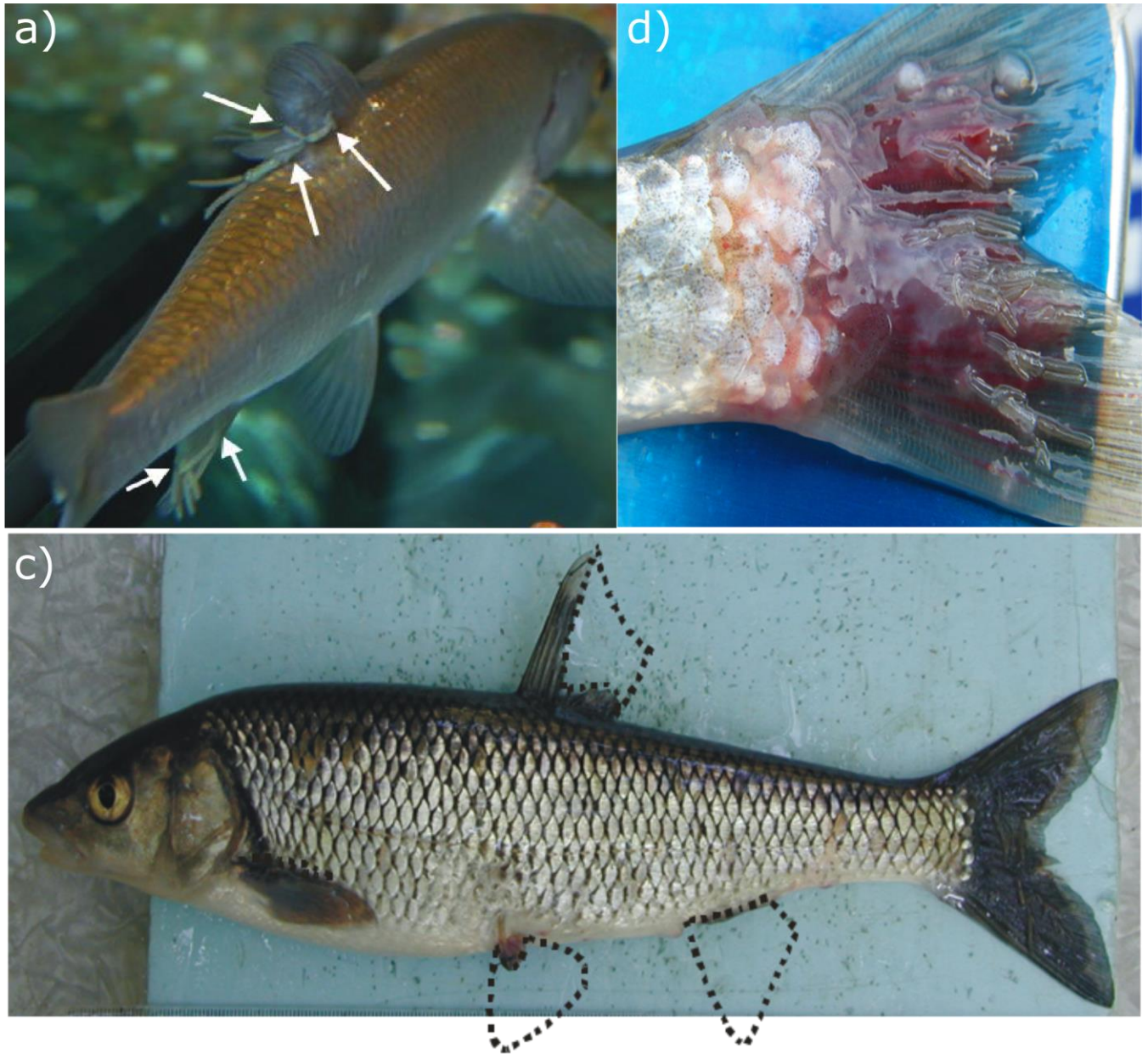
Wright, R. M., G. V. Aglyamova, E. Meyer, and M. V. Matz. 2015. Gene expression associated with white syndromes in a reef building coral, *Acropora hyacinthus*. *BMC Genomics* 16:371–383.

## Supplementary

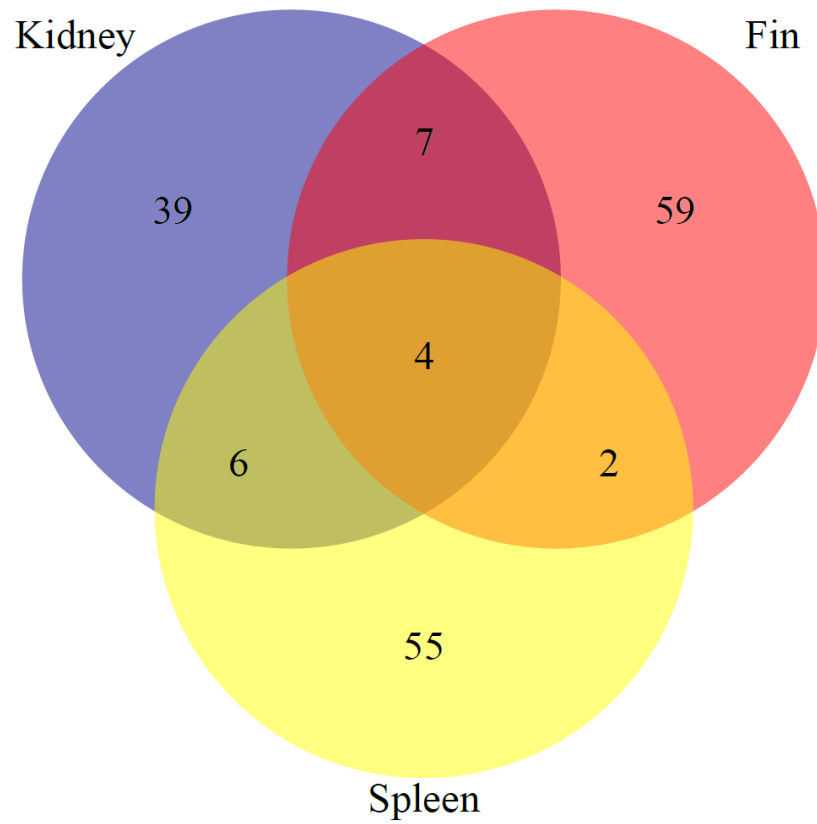


**Figure S1:** Scatter plot of the Principal Component Analysis realized on the raw variables describing locomotion of dace in the experiment related the objective 03. Bar plots indicate how much of the variance is explained by each PCA axis.





**Figure S2:** Pictures illustrating damages caused by *Tracheliastes polycolpus* on a rostrum dace (*Leuciscus burdigalensis*). a) Dace on which several *T. polycolpus* have attached on dorsal and caudal fin (indicated by white arrows, from (Blanchet et al. 2009b)). b) Severe inflammations caused by *T. polycolpus* on the caudal fin of a dace. c) Fin degradation caused by *T. polycolpus* on a dace, missing fin area are indicated by dashes.



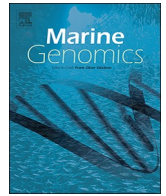
**Figure S3:** Venn diagram displaying the number of DEGs between parasitized and healthy hosts in their fin, their cephalic kidney and their spleen.

**Table S1:** Anova table for the model linking host food intake to host size and health status (Objective O3). Degree of freedom (Df), Chi-squared (Chisq) statistics and p-values are displayed for each fixed effect.

	Df	Chisq	P-value
Body size	1,16	1.565	0.211
Health status	1,16	9.135	0.003

**Table S2:** Anova table for the model linking host locomotion synthetic variable to host size and health status (Objective O3). Degree of freedom (Df), Fisher (F) statistics and p-values are displayed for each fixed effect.

	Df	F	P-value
Body size	1,15	1.119	0.307
Health status	1,15	5.326	0.036



## De novo transcriptome assembly for *Tracheliastes polycolpus*, an invasive ectoparasite of freshwater fish in western Europe

Eglantine Mathieu-Bégné<sup>a,b,c,\*</sup>, Géraldine Loot<sup>a</sup>, Simon Blanchet<sup>a,b</sup>, Eve Toulza<sup>c</sup>, Clémence Genthon<sup>d</sup>, Olivier Rey<sup>c</sup>

<sup>a</sup> Université Paul Sabatier, Centre National pour la Recherche Scientifique, Ecole Nationale de Formation Agronomique, Laboratoire Evolution et Diversité Biologique (UMR5174), 118 route de Narbonne, F-31062 Toulouse, France

<sup>b</sup> Université Paul Sabatier, Centre National pour la Recherche Scientifique, Station d'Ecologie Théorique et Expérimentale (UMR5321), 2 route du CNRS, F-09200 Moulis, France

<sup>c</sup> Université Perpignan Via Domitia, Laboratoire Interactions Hôte-Parasite-Environnement, UMR5244, 58 Avenue Paul Alduy, F-66860 Perpignan, France

<sup>d</sup> INRA, US 1426, GeT-PlaGe, Genotoul, Castanet-Tolosan, France

### ARTICLE INFO

#### Keywords:

Copepods  
Monoxenous parasite  
Parasite specificity  
Curation

### ABSTRACT

*Tracheliastes polycolpus* is an ectoparasitic copepod that recently emerged in western Europe and that infects several freshwater fish species. Its recent successful spread might be due to its fascinating ability to shift to new host populations and/or species. Here, we present the first non-redundant and protein-coding de novo transcriptome assembly for *T. polycolpus* along with a quality assessment and reliable transcript annotations. This assembly was built from fifteen adult female parasites sampled from three different host species from a single river in southwestern France. Overall, 17,157 non-redundant contigs likely corresponding to protein-coding transcripts were identified, of which 13,093 (i.e., 76%) were successfully annotated. This assembly displayed good representativeness since 65.8% of the raw reads properly aligned back to the assembly. Similarly, 90.5% of the single copies of orthologues conserved across arthropods were retrieved in this assembly, which reflects a very good completeness. Finally, this transcriptome assembly gathered 7979 *T. polycolpus* specific transcripts when compared with the two closest referenced species (*Lepeophtheirus salmonis* and *Caligus rogercresseyi*), thus constituting an original genomic resource. This high-quality transcriptome is an important genomic resource for deciphering the molecular bases underlying host shifts in generalist parasites, and for studying the evolutionary biology of copepods that are major biological components of both freshwater and marine ecosystems.

### 1. Introduction

*Tracheliastes polycolpus* (Copepoda, Neocopepoda, Podoplea, Siphonostomatoida, Lerneapodidae) is a monoxene ectoparasitic copepod of freshwater fish (Fryer, 1982). Only adult females have a parasitic lifestyle, whereas males are microscopic and free-living organisms (Kabata, 1986). Adult females anchor on fish fins and feed on their epidermal cells. This causes partial to total fin degradation, which generally favours secondary infections by bacteria or/and viruses (Fryer, 1982), and ultimately reduces the fitness of their hosts (Loot et al., 2004; Blanchet et al., 2009).

In the 1920s, *T. polycolpus* was introduced from eastern Eurasia to western Europe through fish trades, and has since spread over several watersheds in England, France and Spain (Rey et al., 2015). In France,

*T. polycolpus* is primarily associated with the common dace and the rostrum dace (*Leuciscus leuciscus* and *L. burdigalensis*), but is also commonly observed on several alternative cyprinid species (Lootvoet et al., 2013). The recent invasion history of *T. polycolpus*, together with its switch to new alternative host species, constitutes an unprecedented opportunity to study the mechanisms underlying parasite specificity and the rapid adaptive processes associated with host shifts (Lootvoet et al., 2013; Dybdahl et al., 2014; De Fine Licht, 2018). Developing “omics” resources is a key step towards a better understanding of the molecular bases of such processes, notably the role of plasticity and/or selection associated with host shift based on gene expression profiles (De Fine Licht, 2018; Hébert et al., 2017).

Here, we built a curated de novo transcriptome assembly for this species. More specifically, we present the first non-redundant and

\* Corresponding author at: UMR5174 EDB, 118 Route de Narbonne, 31062 Toulouse, France.

E-mail addresses: [eglantine.mb@gmail.com](mailto:eglantine.mb@gmail.com) (E. Mathieu-Bégné), [geraldine.loot@univ-tlse3.fr](mailto:geraldine.loot@univ-tlse3.fr) (G. Loot), [simon.blanchet@sete.cnrs.fr](mailto:simon.blanchet@sete.cnrs.fr) (S. Blanchet), [eve.toulza@univ-perp.fr](mailto:eve.toulza@univ-perp.fr) (E. Toulza), [clemence.genthon@inra.fr](mailto:clemence.genthon@inra.fr) (C. Genthon).

<https://doi.org/10.1016/j.margen.2018.12.001>

Received 13 July 2018; Received in revised form 7 December 2018; Accepted 7 December 2018

Available online 07 January 2019

1874-7787/ © 2019 Elsevier B.V. All rights reserved.

**Table 1**  
 MixS specifications, Assembly description and Annotation statistics for *T. polycolpus* de novo transcriptome assembly.

MixS descriptors	
Investigation type	Eukaryote
Project name	<i>Tracheliastes polycolpus</i> Transcriptome
Organism	<i>Tracheliastes polycolpus</i>
Geographical location name	The Salat River
Geographical location	43°04'43.0"N; 0°57'29.0"E
Collection date	11/07/2013
Environment (biome)	Aquatic biome
Environment (feature)	River
Environment (material)	Freshwater
Host-associated	<i>Leuciscus burdigalensis</i> , <i>Gobio occitaniae</i> and <i>Phoxinus phoxinus</i>
Tissue type	Trunk
Developmental stage	Adult
Sex	Female
Assembly description	
Sequencing method	Illumina HiSeq 2000
Assembly	Trinity (version 2014-07-17)
Coding-protein transcript selection	TransDecoder
Assembly redundancy reduction	CD-HIT-EST (version 4.6)
Transcript number	17,157
Total assembled bases	20,355,465
N25	2676
N50	1599
N75	954
Average contig length	1186
Longest contig length	23,889
% GC	38.55
Annotation (total)	
Predicted ORF	17,157
Complete protein (including start and stop codons)	10,112
Swiss-Prot top BLASTP hit	12,343
Gene ontology in blastp	12,098
Pfam	11,751
KEGG	10,924
eggNOG	10,400
Gene ontology in Pfam	7641

protein-coding transcriptome assembly together with functional annotations for *T. polycolpus* obtained from fifteen adult female parasites sampled on three different host species. Only a few assemblies exist for copepods and notably for parasitic copepods, although some have led to novel discoveries (including mechanisms for toxin resistance, gene expression patterns associated with molting or development) (Lenz et al., 2014; Nuñez-Acuña et al., 2014; Carmona-Antoñanzas et al., 2015). Consequently, this database constitutes an important genomic resource for an emerging parasite in wild populations and for the diverse group of Crustacean copepods.

## 2. Data description

### 2.1. Sampling

Fifteen adult females of *T. polycolpus* were sampled on three different fish host species at a single location in the Salat River in southwestern France on a single day to limit confounding environmental effects (Table 1). The three host species were the rostrum dace (*L. burdigalensis*) and two alternative hosts; the Occitan gudgeon (*Gobio occitaniae*) and the European minnow (*Phoxinus phoxinus*). Fish were caught by electric fishing using a DEKA® 7000, anaesthetized according to a standardized protocol, and inspected for the presence of parasites. Parasites were collected using sterile forceps as follows: five parasites were sampled on five dace, five parasites were sampled on four gudgeons, and five parasites were sampled on four minnows. The fifteen parasites were immediately stored in RNAlater for 24 h and then conditioned at  $-80^{\circ}\text{C}$  (to stabilize and protect cellular RNAs) until RNA

extraction.

### 2.2. Sequencing

Total RNA from each parasite was extracted using the RNeasy Plus Mini Kit (Qiagen) following the manufacturer's instructions from the parasite trunk only—to minimize possible contamination with host RNA—with a final elution volume of 40  $\mu\text{L}$  RNase-free water. The quantity and quality of RNA extractions were assessed using a nanodrop ND-8000 (Thermo Scientific) and a BioAnalyser (Agilent Technologies), respectively. The individual RNA-seq libraries were prepared on a Tecan EVO2000 using the IlluminaTruSeq Stranded mRNA protocol. The individual libraries obtained were multiplexed then pooled and sequenced on two lanes of an Illumina HiSeq 2000 (High Throughput mode) using a paired-end read length of  $2 \times 100$  bp. The reads were demultiplexed, those that did not pass the chastity filter (i.e., internal filtering procedure from Illumina sequencers) were filtered out (about 7% of the total reads), and the adapters were trimmed automatically at this stage. The sequencing resulted in approximately 420 million  $2 \times 100$  bp paired-end reads, with an average of 28 million paired-end reads per sample. The library preparation, sequencing and pre-processing automatic filtering steps were performed at the GeT-PlaGe core facility (Toulouse, France).

### 2.3. De novo transcriptome assembly and curation

Based on the reads obtained from all individuals, we first assembled a raw de novo transcriptome using the Trinity pipeline (version 2014-07-17) (Haas et al., 2013). Reads were trimmed using Trimmomatic (Bolger et al., 2014) with default parameters to discard low-quality and/or poorly informative reads. Then, each sample was parsed separately as Trinity input, and the option “-SS\_lib\_type” was selected to keep strand specificity information. Other parameters were set to the default values. At this step the assembly contained a total of 101,636 contigs.

To curate the raw transcriptome assembly from non-protein-coding contigs, we discarded contigs lacking an open reading frame (ORF) or having an ORF shorter than 100 amino acids using the Trinity plugin TransDecoder (Haas et al., 2013). This filtering step resulted in 33,984 contigs enriched with only messenger RNAs (putative protein-coding genes). Finally, to reduce redundancy, we ran CD-HIT-EST (Version 4.6, (Huang et al., 2010)) to cluster the resulting transcripts according to their similarity with a minimum sequence identity set to 95%, (options -c 0.95 -n 8, where c and n are “similarity threshold” and “word size” parameters respectively). These filtering steps resulted in a final protein-coding and non-redundant assembly containing 17,157 transcripts corresponding to as many putative unique protein-coding genes.

### 2.4. Annotation

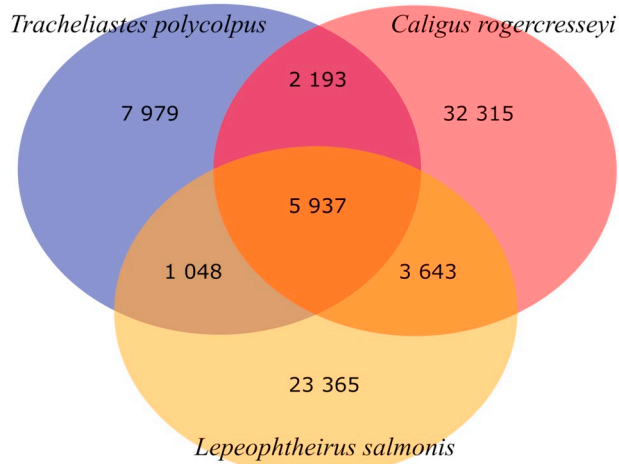
Transcriptome functional annotation was performed using the software suite Trinotate (version 3.1.0, (Haas et al., 2013)). Translated sequences from the final assembly were used to perform a blastp search on the Swiss-Prot database and a protein domain search on the Pfam database (both uploaded from Trinotate version 3.1.10 on 07/03/2016). Results were then integrated in Trinotate to retrieve functional annotations leveraging eggNOG, GO and Kegg databases (Huerta-Cepas et al., 2016; Gene Ontology Consortium, 2004; Kanehisa and Kegg, n.d.). Only highly significant matches were reported in the annotation file (i.e., those with an e-value  $< 10^{-5}$  for the blast hits and the domain noise cut-off for the Pfam domains). Overall, 13,093 contigs were annotated from at least one database. Details about the number of contigs annotated on each database are given in Table 1. Transcript nucleotide and amino acid sequences are also provided in the annotation file (Supplementary S1).

**Table 2**  
Quality scores for the raw assembly and the final assembly of *Tracheliastes polycopus*.

		Raw assembly <sup>a</sup>	Final assembly <sup>b</sup>
Representativeness	Re-mapped reads (1 time)	71.00%	65.77%
	Re-mapped reads (> 1 times)	16.40%	6.89%
Completeness	Re-mapped reads (0 time)	12.60%	27.34%
	Single-copy BUSCOs	59.10%	90.50%
	Duplicated BUSCOs	37.70%	5.70%
	Missing BUSCOs	2.20%	2.60%
	Fragmented BUSCOs	1.00%	1.20%
	Total BUSCOs groups searched	1066	1066

<sup>a</sup> *Tracheliastes polycopus* assembly generated from Trinity only.

<sup>b</sup> *Tracheliastes polycopus* reference assembly, resulting from curation steps.



**Fig. 1.** Venn plot showing the number of specific and shared genes between *T. polycopus*, *C. rogercresseyi* and *L. salmonis*, resulting from the RBH analysis.

## 2.5. Data validation and quality control

Quality of our final transcriptome assembly was evaluated based on three main criteria (Table 2 and Fig. 1, (Haas et al., 2013; Hara et al., 2015)). 1. The representativeness criterion refers to raw read content represented by the assembly. 2. The completeness criterion is based on evolutionary expectations of near-universal gene content that the assembly includes. 3. Specificity refers to the number of genes that are specific to *T. polycopus* when compared to closely related species.

### 2.5.1. Representativeness

Representativeness was evaluated by measuring the percentage of quality filtered reads that were properly re-aligned against the assembly. To do so, the software Bowtie 2 (version 2.1.0, (Langmead and Salzberg, 2012)) was used with default “end-to-end” alignment parameters. We estimated representativeness as the percentage of paired reads that aligned back to the assembly exactly once and with respect to the forward and reverse directions (Langmead and Salzberg, 2012). We also reported the percentage of reads that mapped back more than once (as an indication of redundancy or repeated sequences) and the percentage of reads that did not map back to the final assembly (i.e., the proportion of non-represented reads within the assembly).

### 2.5.2. Completeness

Completeness was assessed by counting the percentage of orthologues conserved across arthropods that are present in the assembly using the software BUSCO (Simão et al., 2015). However, if more than one transcript was found to be aligned with the same orthologous gene, then the assembly was therefore considered partially redundant. Thus, to account for both the number of orthologues and redundancy, we

quantified the number of orthologues conserved across arthropods that were of a single copy within the assembly (i.e., BUSCOs: “Benchmarking Universal Single-Copy Orthologs”, (Simão et al., 2015)).

### 2.5.3. Specificity

Specificity was assessed by quantifying the number of non-homologous transcripts obtained for *T. polycopus* when compared to *Lepeophtheirus salmonis* and *Caligus rogercresseyi*, two parasitic copepod species sharing the same order with *T. polycopus* and for which transcriptomic resources exist (available on Genbank under Bioproject numbers PRJEB1804 and PRJNA234316 respectively). These transcriptomes were obtained with comparable technologies to those we used in this study and from different development stages including adult females (Nuñez-Acuña et al., 2014; Carmona-Antoñanzas et al., 2015), which make the comparison relevant. We expect to find common orthologs between the three species, but also a large amount of development stage specific genes in the two sea lice species. We searched for reciprocal blast hit (RBH) between the assembly of *T. polycopus* (after in silico RNA translation) and that of each of the two copepod species using “tblastx”. The e-value threshold was set to 0.001 in order to select the most significant matches (Camacho et al., 2009), and transcripts not having a reciprocal match according to this criteria were considered specific to *T. polycopus*.

### 2.5.4. Quality control results & discussion

Regarding representativeness, we obtained a mapping rate of 65.8%, which is lower than expected for a Trinity assemblage (i.e., approximately 90% (Clarke et al., 2013)). Importantly, a large proportion of raw reads (27.3%) could not be mapped back to the non-redundant and protein-coding assembly (Table 2) likely because some unique contigs were also lost during the filtering steps. Given that the raw assembly size was reduced by almost a factor of six in the non-redundant and protein-coding assembly, this assembly still displays a good compromise between its informative content (representativeness) and its complexity reduction. Indeed, when limiting redundancy and removing chimaera contigs, usually up to 47% of reference genes are retrieved in de novo assemblies conducted with several assemblers (Yang and Smith, 2013). Furthermore, high completeness was achieved with overall 90.5% of BUSCOs retrieved in the assembly and only 2.6% of conserved orthologues across arthropods having been missed (Table 2). Interestingly, the BUSCO analysis also confirmed that we significantly limited redundancy during the curation process, as 37.7% of BUSCOs were duplicated in the raw assembly and only 5.7% of duplicated BUSCOs were found in the final assembly (Table 2). Finally, we found 6910 potentially specific transcripts in our *T. polycopus* assembly compared with the two most closely related species currently available (Fig. 1). This high number of specific transcripts might be partly explained by the fact that the three species are from the same order but different families, except for *L. salmonis* and *C. rogercresseyi* that share the same family, which may explain their higher number of shared genes (see also Supplementary Fig. S2 for similarity histograms of the

shared genes between the three species). Overall, this highlights that this assembly constitutes an original specific resource of interest. Our quality assessment thus clearly revealed that the resulting non-redundant and protein-coding assembly fulfilled the three criteria we evaluated (i.e., representativeness, completeness and specificity). As such, we propose this non-redundant and protein-coding transcriptome assembly as a reference transcriptome assembly for *T. polycolpus*.

### 3. Conclusion

The transcriptome assembly reconstructed in this study is the most complete genomic/transcriptomic resource available for *T. polycolpus*, and more generally for freshwater copepods (Bron et al., 2009). As such, this paves the way for future studies on this species, notably about the molecular bases of host shift by this parasite. Beyond the studies on this targeted species, this transcriptome will also open new avenues for studying freshwater copepods that constitute cornerstone elements for the biological functioning of freshwater ecosystems.

### Availability of supporting data

Raw reads are deposited in Sequence Read Archive (GenBank) under project number PRJNA476682. Accession numbers are SRR7411022, SRR7411021, SRR7411024, SRR7411023, SRR7411026, SRR7411025, SRR7411028, SRR7411027, SRR7411030, SRR7411029, SRR7411019, SRR7411018, SRR7411017, SRR7411016, and SRR7411020. This Transcriptome Shotgun Assembly project has been deposited at DDBJ/EMBL/GenBank under the accession GGQW00000000. The version described in this paper is the first version, GGQW01000000.

### Ethics statement

Fish carrying the parasites used in this study were caught under prefectural order and in accordance to current ethical laws in France.

### Competing interests

None.

### Funding

This work was funded by the Agence Nationale de la Recherche (project INCLIMPAR, grant no. ANR-11-JSV7-0010), BiodivERsA (project PROBIS) and the Institut Universitaire de France. EMB was supported by a PhD grant from the French Ministry for Education and Sciences. This work was performed in collaboration with the GeT core facility, Toulouse, France (<http://get.genotoul.fr>), and was supported by France Génomique National infrastructure, funded as part of “Investissement d’avenir” program managed by Agence Nationale pour la Recherche (contract ANR-10-INBS-09). This work was undertaken in laboratories being part of the “Laboratoire d’Excellence” (LABEX) entitled TULIP (ANR-10-LABX-41).

### Authors' contributions

GL, SB and OR designed the experiment and sampled the specimens. OR and CG performed molecular analyses and sequencing. EMB, OR and ET assembled and evaluated the transcriptome assembly. EMB, GL, ET, OR and SB wrote the paper. All authors read and approved the final version of the manuscript.

### Acknowledgements

We warmly thank all the students involved in field sampling.

### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.margen.2018.12.001>.

### References

- Blanchet, S., Méjean, L., Bourque, J.-F., Lek, S., Thomas, F., Marcogliese, D.J., et al., 2009. Why do parasitized hosts look different? Resolving the “chicken-egg” dilemma. *Oecologia* 160, 37–47.
- Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120.
- Bron, J., Boore, J., Boxshall, G., Bricknell, I., Frisch, D., Goetze, E., et al., 2009. Copepod Genome Initiative.
- Camacho, C., Madden, T., Coulouris, G., Avagyan, V., Ma, N., Tao, T., et al., 2009. Manuel de l'utilisateur des applications BLAST en ligne de commande (BLAST Command Line Applications User Manual).
- Carmona-Antoñanzas, G., Carmichael, S.N., Heumann, J., Taggart, J.B., Gharbi, K., Bron, J.E., et al., 2015. A survey of the ATP-binding cassette (ABC) gene superfamily in the salmon louse (*Lepeophtheirus salmonis*). Corsi I, editor. *PLOS ONE* 10, e0137394.
- Clarke, K., Yang, Y., Marsh, R., Xie, L., ZK, K., 2013. Comparative analysis of de novo transcriptome assembly. *Sci. China Life Sci.* 56, 156–162.
- De Fine Licht, H.H., 2018. Does pathogen plasticity facilitate host shifts? Round JL, editor. *PLoS Pathog.* 14, e1006961.
- Dybdahl, M.F., Jenkins, C.E., Nuismer, S.L., 2014. Identifying the molecular basis of host-parasite coevolution: merging models and mechanisms. *Am. Nat.* 184, 1–13.
- Fryer, G., 1982. The parasitic copepoda and branchiura of British freshwater fishes. A handbook and key. Freshwater Biological Association, Ambleside.
- Gene Ontology Consortium, 2004. The gene ontology (GO) database and informatics resource. *Nucleic Acids Res.* 32, 258D–261.
- Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., et al., 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* 8, 1494–1512.
- Hara, Y., Tatsumi, K., Yoshida, M., Kajikawa, E., Kiyonari, H., Kuraku, S., 2015. Optimizing and benchmarking de novo transcriptome sequencing: from library preparation to assembly evaluation. *BMC Genomics* 16.
- Hébert, F.O., Grambauer, S., Barber, I., Landry, C.R., Aubin-Horth, N., 2017. Major host transitions are modulated through transcriptome-wide reprogramming events in *Schistocephalus solidus*, a threespine stickleback parasite. *Mol. Ecol.* 26, 1118–1130.
- Huang, Y., Niu, B., Gao, Y., Fu, L., Li, W., 2010. CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics* 26, 680–682.
- Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M.C., et al., 2016. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res.* 44, D286–D293.
- Kabata, Z., 1986. Redescrptions of and comments on four little-known Lernaepodidae (Crustacea: Copepoda). *Can. J. Zool.* 64, 1852–1859.
- Kanehisa, M., Goto, S., 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30.
- Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359.
- Lenz, P.H., Roncalli, V., Hassett, R.P., Wu, L.-S., Cieslak, M.C., Hartline, D.K., et al., 2014. De Novo Assembly of a Transcriptome for Calanus finmarchicus (Crustacea, Copepoda) – the Dominant Zooplankton of the North Atlantic Ocean. *Ianora A*, editor. *PLoS One* 9, e88589.
- Loot, G., Poulet, N., Reyjol, Y., Blanchet, S., Lek, S., 2004. The effects of the ectoparasite *Tracheliastes polycolpus* (Copepoda: Lernaepodidae) on the fins of rostrum dace (*Leuciscus leuciscus burdigalensis*). *Parasitol. Res.* 94, 16–23.
- Lootvoet, A., Blanchet, S., Gevrey, M., Buisson, L., Tudesque, L., Loot, G., 2013. Patterns and processes of alternative host use in a generalist parasite: insights from a natural host-parasite interaction. Mayhew P, editor. *Funct. Ecol.* 27, 1403–1414.
- Núñez-Acuña, G., Valenzuela-Muñoz, V., Gallardo-Escárate, C., 2014. High-throughput SNP discovery and transcriptome expression profiles from the salmon louse *Caligus rogercresseyi* (Copepoda: Caligidae). *Comp. Biochem. Physiol. Part D Genomics Proteomics* 10, 9–21.
- Rey, O., Fournet, L., Paz-Vinas, I., Loot, G., Veyssiére, C., Roche, B., et al., 2015. Elucidating the spatio-temporal dynamics of an emerging wildlife pathogen using approximate Bayesian computation. *Mol. Ecol.* 24, 5348–5363.
- Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., Zdobnov, E.M., 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212.
- Yang, Y., Smith, S.A., 2013. Optimizing de novo assembly of short-read RNA-seq data for phylogenomics. *BMC Genomics* 14, 328.

## Research

### Demographic and genetic collapses in spatially structured populations: insights from a long-term survey in wild fish metapopulations

Eglantine Mathieu-Bégné, Géraldine Loot, Mathieu Chevalier, Ivan Paz-Vinas and Simon Blanchet

*E. Mathieu-Bégné (https://orcid.org/0000-0002-6846-5491) (eglantine.mb@gmail.com), G. Loot, M. Chevalier and S. Blanchet, Centre National de la Recherche Scientifique (CNRS), Univ. Paul Sabatier (UPS), Inst. de Recherche pour le Développement (IRD), Ecole Nationale Supérieure de Formation de l'Enseignement Agricole (ENSFEA); UMR5174, Evolution et Diversité Biologique, 118 route de Narbonne, FR-31062 Toulouse, France. LG also at: Inst. Universitaire de France, Paris, France. BS also at: CNRS, UPS; UMR5321, Station d'Ecologie Théorique et Expérimentale, Moulis, France. – I. Paz-Vinas, Univ. de Lyon, Ecole Nationale des Travaux Publics de l'Etat (ENTPE), CNRS; UMR5023, Laboratoire d'Ecologie des Hydrosystèmes Naturels et Anthropisés, Villeurbanne, France, and: UPS, INP, CNRS, Univ. de Toulouse, UMR 5245 Laboratoire Écologie Fonctionnelle et Environnement, Ecolab, Toulouse, France.*

#### Oikos

128: 196–207, 2019

doi: 10.1111/oik.05511

Subject Editor: Florian Altermatt

Editor-in-Chief: Dries Bonte

Accepted 17 August 2018

Unraveling the relationship between demographic declines and genetic changes over time is of critical importance to predict the persistence of at-risk populations and to propose efficient conservation plans. This is particularly relevant in spatially structured populations (i.e. metapopulations) in which the spatial arrangement of local populations can modulate both demographic and genetic changes. We used ten-year demo-genetic monitoring to test 1) whether demographic declines were associated with genetic diversity declines and 2) whether the spatial structure of a metapopulation can weaken or reinforce these demographic and genetic temporal trends. We continuously surveyed, over time and across their entire range, two metapopulations of an endemic freshwater fish species *Leuciscus burdigalensis*: one metapopulation that had experienced a recent demographic decline and a second metapopulation that was stable over time. In the declining metapopulation, the number of alleles rapidly decreased, the inbreeding coefficient increased, and a genetic bottleneck emerged over time. In contrast, genetic indices were constant over time in the stable metapopulation. We further show that, in the declining metapopulation, demographic and genetic declines were not homogeneously distributed across the metapopulation. We notably identify one local population situated downstream as a 'reservoir' of individuals and genetic variability that dampens both the demographic and genetic declines measured at the metapopulation level. We demonstrate the usefulness of long-term monitoring that combines both genetic and demographic parameters to understand and predict temporal population fluctuations of at-risk species living in a metapopulation context.

Keywords: conservation genetics, rescue effect, fragmentation



## Introduction

Understanding how demographic and genetic changes occur over time is critical to resolve evolutionary and conservation issues. Joint changes in demographic and genetic parameters over time have often been investigated through snap-shot approaches that investigate genetic changes between two (or more) periods during which demographic parameters have changed in the surveyed population (Schwartz et al. 2007, Habel et al. 2014). Nonetheless, using continuous monitoring (i.e. monitoring characterized by a regular temporal resolution) to conjointly study demographic and genetic changes over time should provide complementary information at a higher temporal resolution (Schwartz et al. 2007). For instance, demographic monitoring can provide information about the mechanisms responsible for temporal demographic changes (such as the intrinsic growth rate or density dependence, Hostetler and Chandler 2015), whereas genetic monitoring provides information regarding the evolutionary sustainability of populations (Schwartz et al. 2007). Thus, continuous real-time demo-genetic monitoring is a promising approach that should provide fundamental insights for the long-term conservation of populations (Hoban et al. 2014).

However, continuous real-time demo-genetic monitoring is still rarely used (but see Devillard et al. 2011, Chen et al. 2016). Most continuous real-time demo-genetic monitoring studies have focused on stable or increasing populations to assess their long-term genetic viability (Hansson et al. 2000, Kaeuffer et al. 2007, Devillard et al. 2011). Consequently, association between decreases in abundance (demographic collapses) and genetic diversity losses (genetic collapses) in declining populations has rarely been investigated in the wild. This is surprising given that association of demographic collapses and genetic diversity losses may sustain extinction vortices in at-risk populations. Particularly, Spielman et al. (2004) have shown that a loss in genetic diversity is expected to become critical before populations go extinct and this has consequences for the mechanisms that underline demographic changes (Cappuccino and Price 1995, Hildner et al. 2003). For instance, a loss in genetic diversity such as higher inbreeding depression in populations should decrease the intrinsic growth rate of these populations, hence favoring demographic declines (Cappuccino and Price 1995, Hostetler et al. 2013).

Association between demographic decline and loss of genetic diversity has been initially tested in populations homogeneously distributed in space (Spielman et al. 2004), but populations are often spatially structured across continuous or discontinuous habitat patches (Hanski 1998). This is the case for many natural ecosystems forming mosaics of habitats, such as rivers or oceans, but also for many ecosystems that are increasingly fragmented by human activities (e.g. through deforestation or the building of infrastructures). In these spatially structured landscapes, 'local' populations are connected to one another to form metapopulations, and some local populations may act as sources of individuals

and genes through emigration, whereas others may act as sinks, receiving individuals and genes through immigration (Gaggiotti 1996).

Predicting demo-genetic declines in metapopulations is not straightforward since both demographic changes and the distribution of genetic diversity are not necessarily spatially homogeneous (Naranjo and Bodmer 2007, Morrissey and de Kerckhove 2009). For instance, within a demographically declining metapopulation, if local source populations failed to provide enough individuals and genes to local sink populations, we can predict strong effects of genetic drift in local sink populations, but not necessarily in local source populations (Dunning et al. 1992, Spielman et al. 2004). As the population size decreases, the intensity of genetic drift increases in local sink populations, which may ultimately lead to local decreases in genetic diversity due to the loss or the fixation of particular alleles (Spielman et al. 2004, Potvin et al. 2017). In contrast, genetic rescue effects may be expected in local sink populations if they receive genetically distinct migrants from source populations (Dunning et al. 1992, Carlson et al. 2014). By dampening declines in genetic diversity through dispersal, local source populations may hence counteract the pervasive genetic consequences of demographic declines that occur at the metapopulation scale (Barson et al. 2009, Jangjoo et al. 2016). Testing whether demographic and genetic temporal changes are concomitant in declining metapopulations is consequently an important prerequisite to predict the sustainability of metapopulations in the long term (Vuilleumier et al. 2010, Palstra and Ruzzante 2011).

Here, we aimed to 1) test whether demographic declines over time are associated with changes in genetic diversity in the same timeframe in metapopulations and 2) identify to which extent the spatial structure of metapopulations may modulate these temporal trends. We focused on an endemic freshwater fish species from southwestern France (the rostrum dace, *Leuciscus burdigalensis*). We considered two metapopulations that belong to two sub-river basins that were surveyed demographically (by measuring local fish abundance) and genetically (by assessing genetic variation using neutral markers) for ten years on a year-to-year basis. One metapopulation underwent a sharp demographic decline during that period, whereas the other was stable (or slightly increased) demographically over time. We first quantified these demographic trends for both metapopulations and tested whether demographic changes were homogeneously distributed along the upstream–downstream gradients of the riverscapes. Second, we tested whether processes such as density-dependence and/or the effect of the intrinsic growth rates underlay these demographic changes. Then, we tested whether the general demographic decline observed in one of the two metapopulations was associated with a loss of genetic diversity over time. We further tested whether genetic diversity was homogeneously distributed along the upstream–downstream gradients of the riverscapes to identify the putative sources of genetic diversity. Finally, we evaluated the marginal impact of the spatial structure of the metapopulation on temporal genetic trends, notably by testing whether some local populations were

dampening genetic diversity losses occurring in other local populations by acting as reservoirs of individuals and genes.

According to Spielman et al. (2004), we predict a sharp and general decline in genetic diversity in the demographically collapsing metapopulation. Notably we expected an intrinsic negative growth rate underlying the demographic collapse and a departure from the Hardy–Weinberg equilibrium (which represents a genetic equilibrium expected in a stable population). However, riverscapes are strongly structured spatially and environmentally along the upstream–downstream gradient, which influences the demo-genetic dynamics of riverine metapopulations (Labonne et al. 2008, Altermatt 2013, Paz-Vinas and Blanchet 2015). Specifically, genetic diversity is generally higher in downstream local populations compared to upstream local populations in dendritic river networks (Paz-Vinas et al. 2015), and theory predicts that local extinction rates should be higher in local populations situated in isolated upstream localities (Gotelli and Taylor 1999, Mari et al. 2014). We can therefore expect that downstream localities will be 1) more demographically stable than upstream localities, and 2) potentially acting as reservoirs of genetic diversity, hence dampening (and potentially rescuing) the overall genetic consequences of the demographic decline (Fronhofer and Altermatt 2017).

## Material and methods

### Biological model and sampling design

Rostrum dace *Leuciscus burdigalensis* is a cyprinid species living in running waters with intermediate water temperature (from 14°C to 24°C in warmer months). It is a gregarious species living in small shoals of 3–5 individuals, and it generally feeds on benthic or drifting invertebrates. During the active season (June–October) dace forage in groups in habitats characterized by moderate to high water velocity (10–70 cm s<sup>-1</sup>) and moderate water depth (20–60 cm depth), two characteristics that are generally not uniformly distributed across a whole stream. We can therefore consider that dace form metapopulations along the upstream–downstream gradient composed of discrete sub-populations (i.e. in habitats suitable for foraging) connected by corridors (i.e. unsuitable habitats for foraging). Dace can measure up to 30 cm in length, and can reach the age of 10–12 years. They are generally mature at ~2–3 years and they reproduce (several times along its lifetime) during spring on gravel beds of rivers, when the water temperature reaches ~9°C (Keith et al. 2011).

We sampled dace in the Viaur and Célé Rivers, two similar mid-size rivers (~150 km long) in the Garonne River Basin (southwestern France, Fig. 1). The Viaur and Célé Rivers are situated in two different sub-river basins (the Aveyron and Lot sub-river basins, respectively) separated from each other by more than 390 km of topological distance and by several impassable dams. Contemporary dispersal between the Viaur and Célé Rivers is hence highly unlikely (if not

impossible). Hence, we will hereafter consider that the local dace populations we sampled in these two rivers constitute two metapopulations whose contemporary dynamics are independent. Sampling was conducted yearly during the periods 2005–2014 and 2006–2014 for the Viaur and Célé Rivers, respectively, at eleven localities (i.e. river stretches of ~200 m long covering favourable foraging habitats for dace, Fig. 1). Sampled localities, which sustain local dace populations for each river, were separated by topological distance of at least 3.5 km that were often characterized by less favourable habitats (e.g. lentic river sections with low velocity and high water depth) that constitute dispersal corridors between two adjacent local dace populations. Sampling localities were distributed along the upstream–downstream gradient of each river to sample the whole spatial structure of each metapopulation. For logistical reasons, some localities were not sampled each year, but an average of 8–9 localities was sampled each year per metapopulation. Sampling was done according to a standardized electrofishing protocol covering a 500–2000 m<sup>2</sup> surface section depending on the sampled locality. To ensure year-to-year comparisons, we did not vary the sampling surface at the locality level over time. All dace ( $N_{\text{tot}} = 2264$ ) were anaesthetized, counted (to quantify dace abundance for each locality) and measured, and a piece of a pelvic fin was removed and stored in ethanol for each individual. All individuals were returned alive to their sampling locality. Permit numbers from regional authorities (i.e. the ‘Directions Départementales des Territoires’) are provided in Supplementary material Appendix 1 Table A1.

### Laboratory analyses and genetic diversity assessment

We extracted the total DNA from fins following Aljanabi and Martinez (1997). Individual genotypes were obtained at twelve microsatellite loci (Supplementary material Appendix 1 Table A2). We estimated the frequencies of null alleles using the R package ‘PopGenReport’ (Adamack and Gruber 2014). We found a relatively high frequency of null alleles in one marker (Rhca20, Supplementary material Appendix 1 Table A3). Analyses were therefore run with and without this specific marker to ensure the reliability of results. Since there were no significant differences between the results obtained with and without Rhca20, we only present those with the twelve markers.

Genetic diversity was quantified using seven genetic indices. Five of them were calculated using the R package ‘adegenet’ (Jombart 2008) at the marker level: the number of alleles (NA), the allelic richness (AR, i.e. the mean number of alleles across loci standardized for the lowest sampling size using a rarefaction procedure), the inbreeding coefficient (Fis), the expected heterozygosity (He) and the observed heterozygosity (Ho). We also estimated the M-ratio (Garza and Williamson 2001) because this index efficiently captures the signatures of recent genetic bottlenecks in spatially structured populations potentially experiencing downstream-biased asymmetric gene flow, such as riverine populations (Paz-Vinas et al. 2013). The M-ratio detects

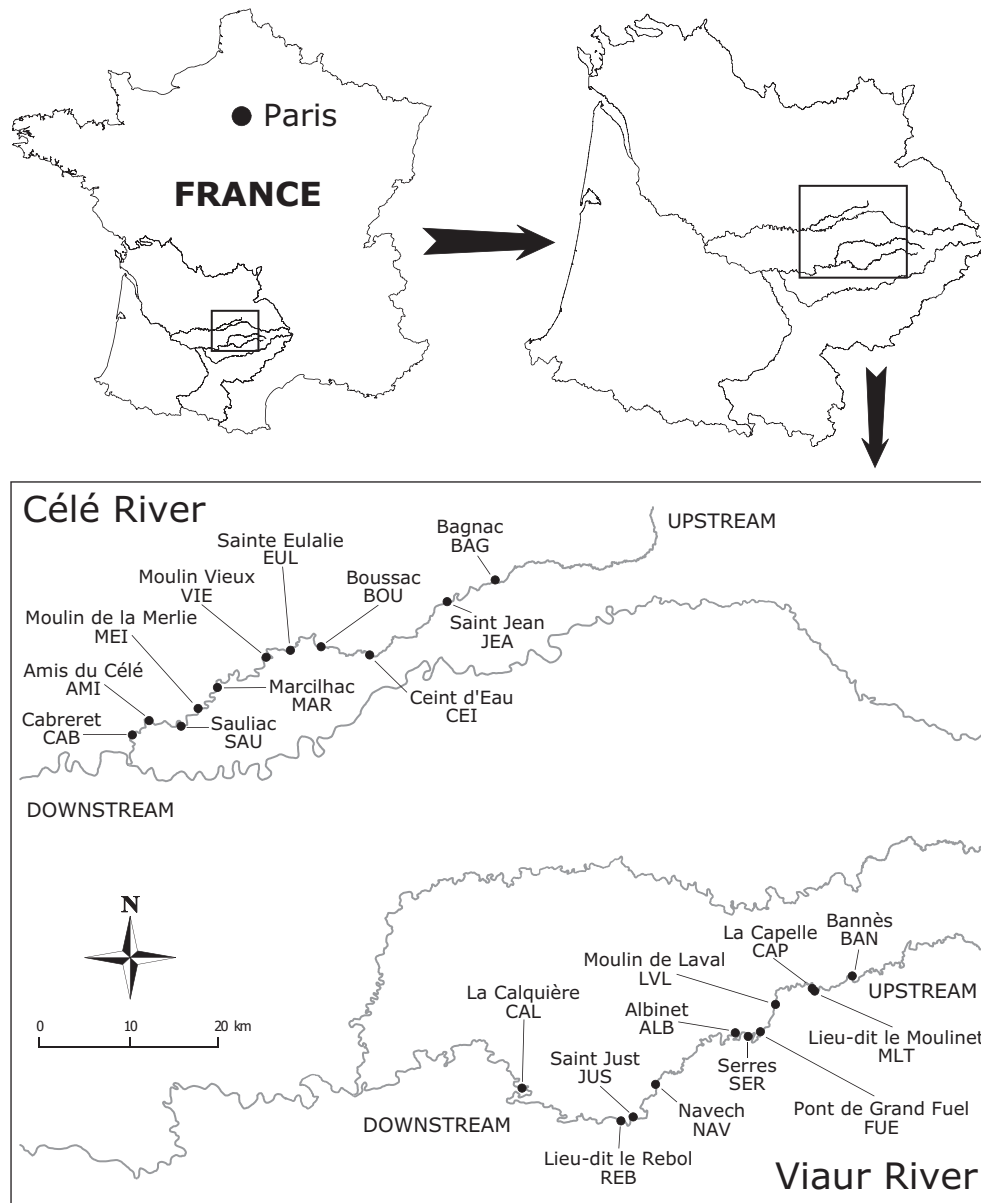


Figure 1. Maps of the Viaur and Célé Rivers, together with the different sampling localities (11 per river) and their distribution along the upstream–downstream gradient of each river.

genetic bottlenecks via a loss of alleles relative to the allelic range, the latter of which is supposed to decrease slower than the number of alleles during a demographic collapse (Garza and Williamson 2001). Empirical  $M$ -ratio values that were lower than a critical value of 0.68 indicate significant bottlenecks (Garza and Williamson 2001). The  $M$ -ratio was calculated at the marker level using a customized R script (Mathieu-Bégné et al. 2018). Finally, we also estimated the effective metapopulation size ( $N_e$ ) for each year and each metapopulation using the linkage disequilibrium method implemented in the software NeEstimator V2.01 (Do et al. 2014) and a sliding window of three years corresponding to the generation time of dace (Skrbinšek et al. 2012).

## Statistical analyses

### *Temporal demographic trends and impact of the spatial structure of metapopulation*

We computed a generalized linear mixed model (GLMM) to test whether there were significant temporal trends in dace abundance and whether these trends varied between the two metapopulations. Since the response variable was the number of dace caught per year and per locality, a Poisson error term distribution was fitted. Explanatory variables (fixed effects) were metapopulation identity (i.e. the Viaur or Célé River), sampling year and the resulting two-term interaction. Locality identity was included as an intercept-only random variable.

We further investigated whether temporal demographic changes were homogeneously distributed across each metapopulation by testing the relationship between temporal demographic changes measured at the locality level and the spatial position of each locality along the upstream–downstream gradient of each metapopulation. We specifically used an additive linear model that assumed a Gaussian error term distribution, the distance from the river source of each locality and metapopulation identity as explanatory variables, and the strength of the temporal trend at the locality level (i.e. the slope of the relationship between dace abundance and year of sampling) as the response variable.

#### **Population dynamics models and processes underlying temporal demographic trends**

We aimed at evaluating, for each metapopulation, whether demographic changes observed over time resulted from changes in the intrinsic growth rate (IGR) and/or in the strength of density-dependence processes. To do so, we fitted hierarchical N-mixture models (Hostetler and Chandler 2015) to the abundances recorded across localities in each metapopulation, in order to estimate the two parameters cited above for each metapopulation and for each sampling year. These models are a class of state–space models integrating 1) a sampling process relating the observed abundances (denoted  $N$ , i.e. the number of individuals detected during a sampling event) to the true abundances (denoted  $X$ , i.e. the actual number of individuals present during a sampling event) and 2) a population dynamics process that describes the evolution of abundances through time (De Valpine and Hastings 2002). The observation process was described by a binomial distribution:

$$X_{i,t} \sim \text{Bin}(N_{i,t}, p_i)$$

where  $p_i$  is the probability of detection at locality  $i$ . To model dace abundances at locality  $i$  and time  $t$ , we used a Poisson distribution:

$$N_{i,t} \sim \text{Pois}(E(N_{i,t}))$$

where  $E(N_{i,t})$  is the expected dace abundance at locality  $i$  and time  $t$  under a stock–recruitment Ricker model. On the log-scale the model is represented by:

$$\log(E(N_{i,t})) = \log(N_{i,t-1}) + \rho - \eta \times N_{i,t-1}$$

where  $\rho$  and  $\eta$  are fixed effects representing the intrinsic metapopulation growth rate (i.e. the rate of metapopulation increase when  $N_{i,t-1}=0$ ) and the strength of density-dependence, respectively. Both parameters were assumed to follow normal distributions centred on 0 with large standard deviations. To allow for stochastic differences between localities in the IGR and for stochastic year effects common to all localities, we introduced the parameters  $S_i$  and  $T_t$  to the

model described above, representing the random locality and random year effects, respectively.  $S_i$  and  $T_t$  were assumed to follow normal distributions centred on 0, with standard deviations  $\sigma_S$  and  $\sigma_T$  representing the magnitude of the variability among localities and years, respectively. Half-Cauchy distributions were used as priors for  $\sigma_S$  and  $\sigma_T$  (Gelman 2006). Finally, the priors for the initial values  $N_1$  of each time series were Poisson distributions with means equal to the average number of individuals captured during the survey, increased by one (Kéry and Schaub 2012). To obtain posterior distributions of parameters (Clark and Bjørnstad 2004), we adopted a Bayesian approach using JAGS 3.3.0 (Plummer 2003) run through the R package ‘R2jags’ (Su and Yajima 2012). Convergence was visually assessed and confirmed using the Gelman and Rubin statistic for all parameters (Gelman and Rubin 1992), and highest posterior density (HPD) intervals were used as 95% credible intervals.

#### **Temporal genetic diversity trends**

We computed independent GLMMs or GLMs to test the effect of temporal trends on each genetic index (i.e. NA, AR, Fis, Ho, He, M-ratio and Ne used as response variables). Gaussian error term distributions were assumed for all indices but NA and Ne, for which a Poisson error term distribution and a Quasi-Poisson error term distribution (to deal with overdispersion) were assumed, respectively. Metapopulation identity, sampling year and the resulting two-term interaction were set as fixed effects. The marker identity was set as a random intercept for each genetic index except for Ne, for which we used a simple GLM (because Ne was not calculated at the marker level). Additionally, we tested whether rare alleles were more likely to be lost (rather than common alleles) in the Viaur metapopulation, as expected theoretically in declining populations (Nei et al. 1975). To do so, we used a Wilcoxon rank-sum test comparing for the Viaur metapopulation 1) the allele frequencies measured across the first four years of the study (2005–2008, i.e. before the population collapse) for alleles that are still present in the four last years of the survey (2011–2014, i.e. after the population collapse) to 2) allele frequencies measured across 2005–2008 for alleles that have been lost in the last four years of the survey. Finally, we plotted direct relationships between temporal demographic and genetic diversity variations for each genetic diversity index, and tested their relationships using Mantel correlation tests conducted for each metapopulation independently (i.e. resulting in twelve Mantel tests). The significance of each Mantel test was evaluated using Spearman correlation coefficients and 1000 permutations. Mantel tests were performed using the R package ‘vegan’ (Oksanen et al. 2011).

#### **The impact of metapopulation spatial structure on temporal trends of genetic diversity**

To investigate the impact of the metapopulation spatial structure on genetic diversity trends, we first tested the spatial distribution of the ‘initial’ genetic diversity (i.e. the

genetic diversity observed for the first year of the survey) using GLMs that assumed Gaussian error term distributions. Initial genetic diversity was computed for NA, Fis, He, Ho and the M-ratio at each locality for the first monitoring year (2005 for the Viaur River metapopulation and 2006 for the Célé River metapopulation) and used as response variables in separated models. The explanatory variables were metapopulation identity and the distance from the source of each locality.

Second, we evaluated the marginal impact of each local population on changes in temporal genetic diversity in the declining metapopulation. To do so, we re-ran the analyses in order to test for temporal trends in the genetic diversity presented above, but by iteratively removing data from one locality from the whole dataset (from the most upstream locality to the most downstream locality, i.e. localities were not randomly removed). For each combination (all localities sampled in each year minus one being excluded) we calculated genetic indices for each year at the marker level and used GLMMs with the marker as a random effect as described above to estimate the slopes of the relationships between genetic diversity indices and the year of sampling, together with its standard error. We excluded Ne from this analysis to allow robust models with information at the marker level. We finally compared these slope estimates to the estimate calculated when all localities are included in the analyses. If the removal of one specific locality leads to a slope estimate steeper than the slope estimated from the whole dataset, one can consider that the population in this specific locality may serve as a genetic reservoir (and potentially as a source) in the metapopulation, since this local population actually provides some marginal genetic diversity to the metapopulation, hence dampening the expected genetic diversity loss.

All statistical analyses were done using the R software (ver. 3.4.2 <[www.r-project.org](http://www.r-project.org)>), and all GLMMs were fitted using the package 'glmmADMB' (Bolker et al. 2012).

## Data deposition

Raw data and scripts required for statistical analyses are available at <<https://doi.org/10.6084/m9.figshare.5999171>>.

## Results

### Temporal demographic trends and impact of the spatial structure of metapopulations

As expected, we found that demographic trends over time significantly varied between the two metapopulations (Table 1a). Dace displayed a slight but significant increase in metapopulation size over time in the Célé River, whereas dace abundance severely declined in the Viaur River (Fig. 2a). We further detected a positive relationship between the magnitude of the demographic changes measured at the local population level and their distance from the river source in both metapopulations ( $F = 5.916$ ,  $df = 1, 15$ ,  $p\text{-value} = 0.029$ ,

Table 1. Output of GLMMs (a–f) and GLM (h) testing the effect of metapopulation identity, year, and the resulting interaction (included only when significant) on (a) dace abundance, (b) allelic richness, (c) expected heterozygosity, (d) observed heterozygosity, (e) number of alleles, (f) Fis, (g) M-ratio and (h) effective population size.

	df <sup>a</sup>	Statistic <sup>b</sup>	p-value
a) Number of dace <sup>c</sup>			
Year	1,156	56.333	< 0.001
Metapopulation	1,156	177.34	< 0.001
Year: metapopulation	1,156	171.875	< 0.001
b) Allelic richness <sup>b</sup>			
Year	1,223	0.098	0.754
Metapopulation	1,223	256.637	< 0.001
c) Expected heterozygosity <sup>b</sup>			
Year	1,223	0.786	0.376
Metapopulation	1,223	65.885	< 0.001
d) Observed heterozygosity <sup>b</sup>			
Year	1,223	0.837	0.361
Metapopulation	1,223	35.224	< 0.001
e) Number of alleles <sup>c</sup>			
Year	1,223	3.185	0.074
Metapopulation	1,223	16.694	< 0.001
Year: metapopulation	1,223	14.693	< 0.001
f) Fis <sup>b</sup>			
Year	1,222	4.187	0.042
Metapopulation	1,222	6.471	0.012
Year: metapopulation	1,222	6.473	0.012
g) M-ratio <sup>b</sup>			
Year	1,222	1.157	0.283
Metapopulation	1,222	7.921	0.005
Year: metapopulation	1,222	8.014	0.005
h) Effective population size <sup>b</sup>			
Year	1,15	141.270	< 0.001
Metapopulation	1,15	263.639	< 0.001
Year: Metapopulation	1,15	15.817	< 0.001

<sup>a</sup>Degrees of freedom at the numerator and denominator, respectively.

<sup>b</sup>Statistics used in these models are Fisher statistics.

<sup>c</sup>Statistics used in these models are  $\chi^2$ -statistics.

Fig. 2b). This correlation indicated that demographic declines in the Viaur metapopulation were stronger in upstream sites than in downstream sites (where the most downstream local population was actually demographically stable, Fig. 2b) and that demographic increases in the Célé metapopulation were stronger in downstream sites than in upstream sites.

### Population dynamics models and processes underlying temporal demographic trends

In support of previous demographic temporal trend analyses, we found a positive IGR with HPD intervals that did not overlap zero for dace from the Célé River, whereas a negative IGR was measured for dace from the Viaur River, although the HPD intervals overlapped zero (Fig. 3a). Our metapopulation dynamics models further revealed significant regulation through density-dependence processes for the Célé River metapopulation (positive value of  $\eta$ ), whereas a slight negative value of  $\eta$  was found for the Viaur River metapopulation (Fig. 3b). In both metapopulations, we found strong year-to-year variations in the IGR. More specifically,

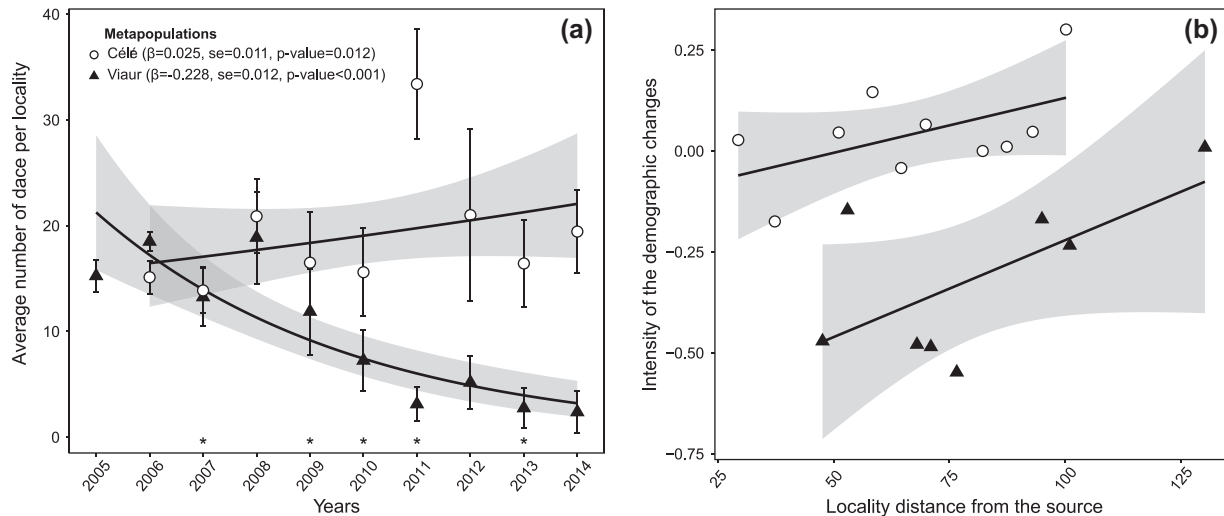


Figure 2. (a) Temporal changes in dace *Leuciscus burdigalensis* abundances in the Viaur and Célé metapopulations. Error bars are standard deviation. Grey areas are confidence intervals of the prediction (lines) from a linear model. Stars represent years for which intrinsic growth rates are significantly negative. (b) Relationship between the strength of the demographic changes (i.e. the slope of the relationship between dace abundance and years) for each local population and the distance from the source of their respective locality. Grey areas are confidence intervals of the prediction (lines) from models.

IGR values were always positive for the Célé River metapopulation, with HPD intervals that never overlapped zero (Fig. 3c). In contrast, IGR values were negative for most years for the Viaur River metapopulation, with five years specifically showing significant negative IGR values (2007, 2009–2011, 2013, Fig. 3c).

### Temporal genetic diversity trends

He, Ho and AR were, on average, 7–26% lower for the Viaur River metapopulation than from the Célé River metapopulation (Table 1b–d, Supplementary material Appendix 1 Fig. A1), although we did not detect any significant temporal trends for these genetic indices (Table 1b–d). In contrast, significant interaction terms between ‘metapopulation identity’ and ‘sampling year’ were detected for NA, Fis and the M-ratio (Table 1e–g), indicating that temporal trends vary between metapopulations for these three genetic parameters. For both NA and M-ratio, we detected a strong and significant decrease over time in the Viaur River metapopulation, whereas these two genetic parameters were stable over time in the Célé River metapopulation (Fig. 4a–b). Interestingly, the values observed for the M-ratio for the Viaur River metapopulation were always lower than 0.68 (Fig. 4b). Fis significantly increased over time in the Viaur River metapopulation, whereas it significantly decreased in the Célé River metapopulation (Fig. 4c). Temporal trends in Ne estimated from genetic markers to lesser extent mirrored the temporal trends in dace abundance highlighted from demographic data: Ne increased sharply and significantly over time in the Célé River metapopulation, whereas Ne remained very low and stable over time in the Viaur River metapopulation (Table 1h, Fig. 4d). Ne values estimated for the Viaur River metapopulation never exceeded ~50 individuals and

were always smaller than values estimated in the Célé River metapopulation (Table 1h, Fig. 4d).

Furthermore, alleles that were lost in the last four years had lower initial frequencies than alleles that were still present in the last four years, indicating that, as expected, rare alleles were first lost during the decline ( $W = 5237$ ,  $p\text{-value} < 0.001$ , Supplementary material Appendix 1 Fig. A2). Finally, the observed relationship between temporal demographic and genetic variations for NA and the M-ratio also supports that the demographic decline observed in the Viaur metapopulation is associated with genetic diversity loss (Supplementary material Appendix 1 Fig. A3). Notably, we identified significant relationships between temporal demographic and genetic variations, implying both NA and the M-ratio in the Viaur metapopulation, and a significant relationship between temporal demographic and genetic variations for NA in the Célé metapopulation (Supplementary material Appendix 1 Table A4).

### The impact of metapopulation spatial structure on temporal trends of genetic diversity

We found that initial genetic diversity was spatially structured for two indices (NA and AR). We indeed highlighted a positive trend between these two indices and the distance from the source of each locality ( $F_{NA} = 6.023$ ,  $df_{NA} = 1, 12$ ,  $p\text{-value}_{NA} = 0.030$ ,  $F_{AR} = 5.246$ ,  $df_{AR} = 1, 12$ ,  $p\text{-value}_{AR} = 0.041$ , Fig. 5), indicating that genetic diversity tended, as expected, to be higher downstream.

Regarding the impact of metapopulation spatial structure on temporal trends of genetic diversity, we found that the dace population from one locality of the Viaur River metapopulation (‘La Calquière’, localized in the most downstream part of the river, Fig. 1) had a marginal impact on the loss

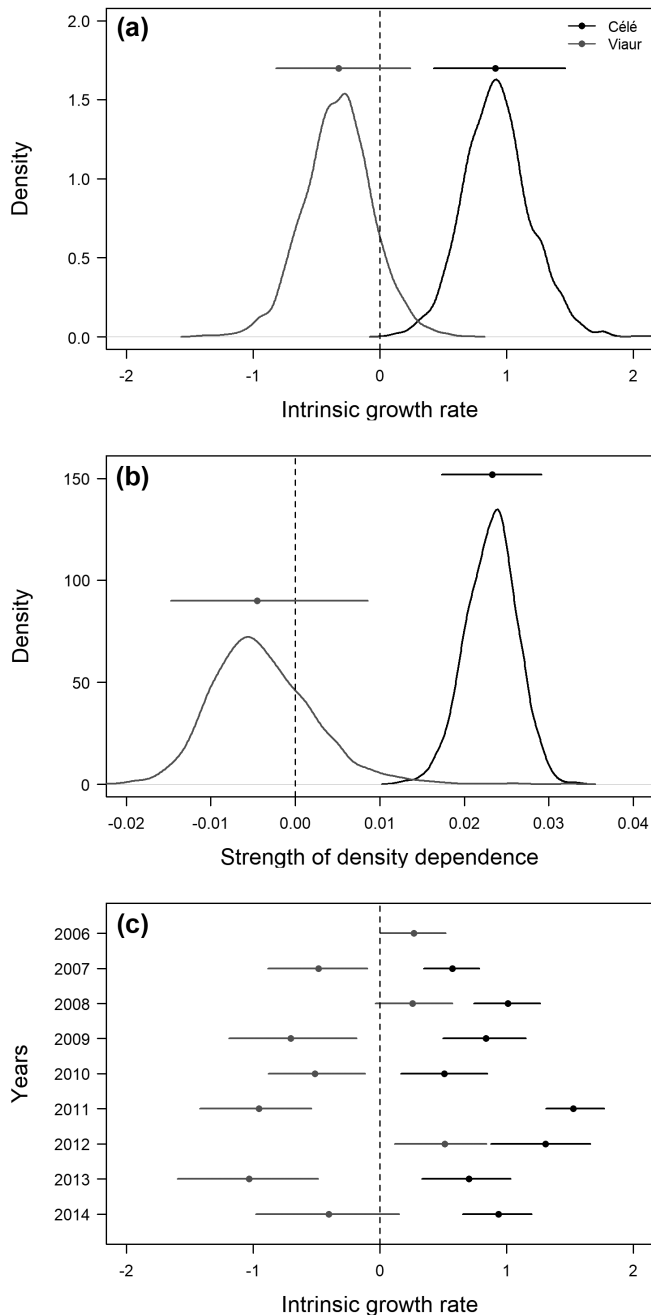


Figure 3. Posterior distributions of (a) the intrinsic growth rate (IGR), (b) the strength of density dependence and (c) temporal variations in IGR. The latter was obtained by adding random sampling sites and years effects to the fixed value of IGR. In (a) and (b), the curves represent the posterior density distribution of the parameters for both metapopulations. In the three panels, the horizontal lines represent the 95% of the highest posterior density (HPD) intervals whereas the points represent the median of the posterior distribution and the vertical dashed line is indicative of the zero value. HPD intervals that do not overlap this line represent significant effects.

of genetic diversity measured over time for several indices. More precisely, when this local population was removed from the whole dataset, steeper slope estimates (compared to the

estimates based on the whole dataset) were detected for the relationship between NA and year of sampling (Fig. 6a) for the relationship between Fis and year of sampling (Fig. 6e) and for the relationship between He and year of sampling (Fig. 6f). These results suggest that this most downstream local population likely acts as a genetic diversity reservoir in the Viaur River metapopulation.

## Discussion

We confirmed the incidence of a severe demographic decline during the monitoring period in the dace metapopulation from the Viaur River, corresponding to an ~80% abundance decrease that occurred in less than ten years. Meanwhile, dace abundance increased over time in the Célé River metapopulation. We further show that the intrinsic growth rates of dace metapopulations were mostly negative in the Viaur River, whereas they were always positive in the Célé River, thus confirming the contrasting demographic trends observed for these two and supporting an expected decline of genetic diversity in the Viaur metapopulation (Reed et al. 2007, Hostetler et al. 2013). Second, our study highlights that the demographic decline observed in the Viaur River was not homogeneous among local populations. We specifically demonstrate that the most downstream local dace population sampled in this riverscape was actually stable over time (a tendency also observed in the Célé River Fig. 2b). This finding is in line with the general idea that upstream populations are more prone to local extinctions than downstream populations in riverine networks (Gotelli and Taylor 1999, Fagan 2002, Mari et al. 2014).

At the genetic level, our longitudinal survey revealed that, contrary to the stable Célé River metapopulation, the collapsing metapopulation was suffering a genetic bottleneck, as it was experiencing a loss of genetic diversity and an increase of its inbreeding coefficient (generally associated with higher extinction risk, Frankham 1995). This global loss of genetic diversity in the Viaur River metapopulation seems to be intrinsically associated with the demographic collapse we identified as documented by Spielman et al. (2004). For instance, the decrease in the number of observed alleles is directly linked to the loss of individuals observed in many localities of the Viaur River. We indeed found that rare alleles were more prone to disappear from the Viaur metapopulation compared to common alleles, as it is expected in collapsing populations (Nei et al. 1975), and we can therefore assume that these alleles will only be recovered if individuals from peripheral populations immigrate (Chen et al. 2016). We further found that the inbreeding coefficient decreased over time in the Célé River metapopulation, which indicates that this metapopulation became closer to the Hardy–Weinberg equilibrium. This may be attributed to the significant increase in dace abundance detected in this metapopulation, and shows that a slight increase in metapopulation abundance may help reduce inbreeding in wild

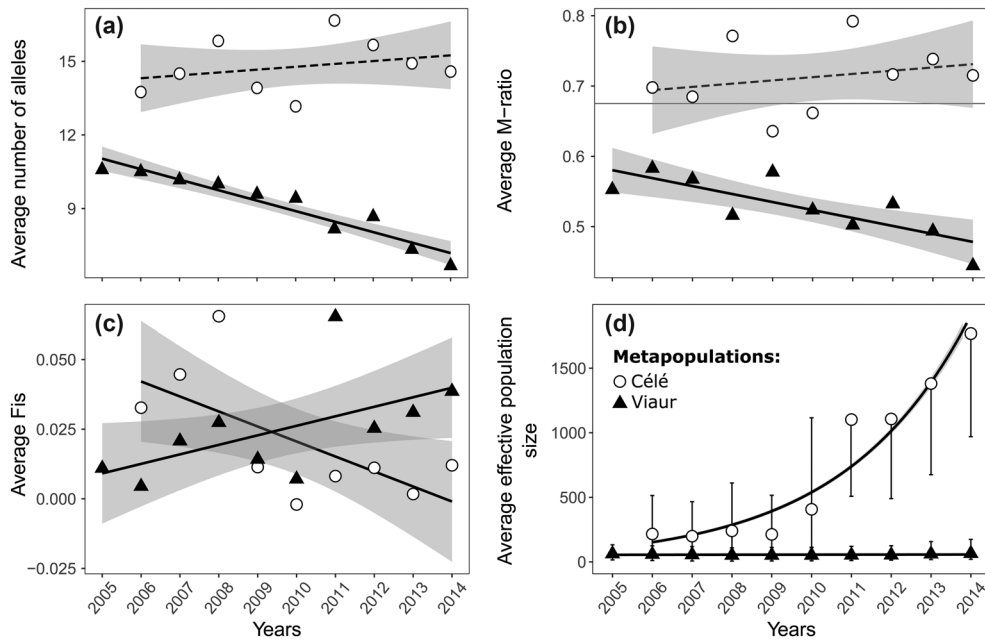


Figure 4. Temporal changes in (a) mean number of alleles, (b) mean M-ratios, with the grey horizontal line that displays the significant limit of 0.68 (Garza and Williamson 2001), (c) mean Fis, and (d) Ne for dace from the Viaur and Célé metapopulations. Note that for Ne estimates, when no upper error bars are displayed this indicates infinite estimates of the upper limit. Grey areas are confidence intervals of the predictions (lines) from a linear model. Solid lines are significant, whereas dashed lines are non-significant.

metapopulations. In contrast, departure from the Hardy–Weinberg equilibrium was stronger over time in the declining metapopulation, which probably reflects the higher risk of inbreeding due to the extremely low and declining number of effective breeders measured in this metapopulation

and likely resulting in the negative intrinsic growth rate we identified (Waples 2002, Reed et al. 2007).

Interestingly, not all genetic diversity indices co-varied with this demographic decline, as we failed to detect temporal changes for heterozygosity and allelic richness. This

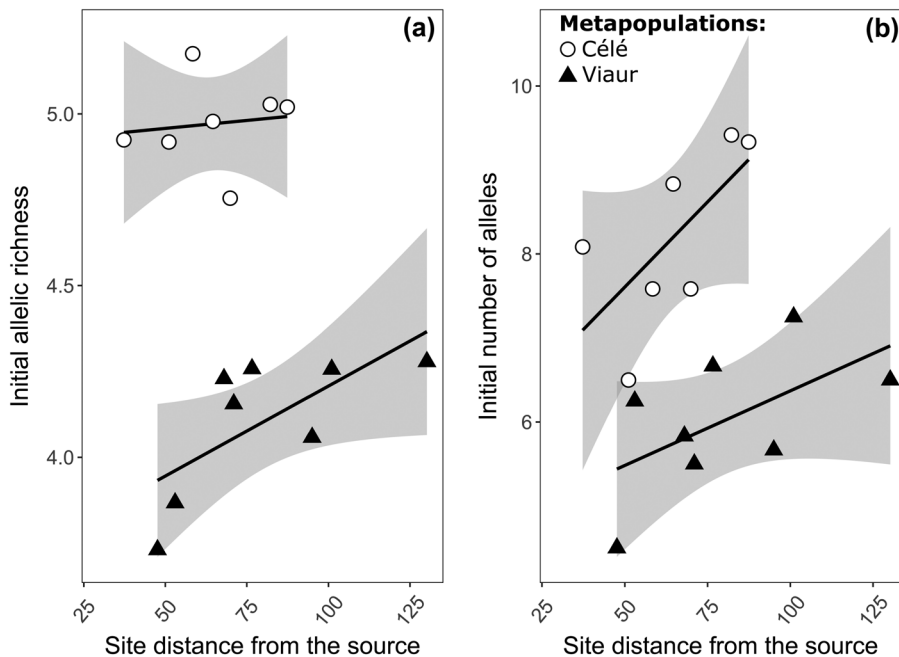


Figure 5. Relationship between the initial genetic diversity measured as the allelic richness (a) and the number of alleles (b) in each local population and the distance from the source of each locality. Grey areas are confidence intervals of the predictions (lines) from linear models.



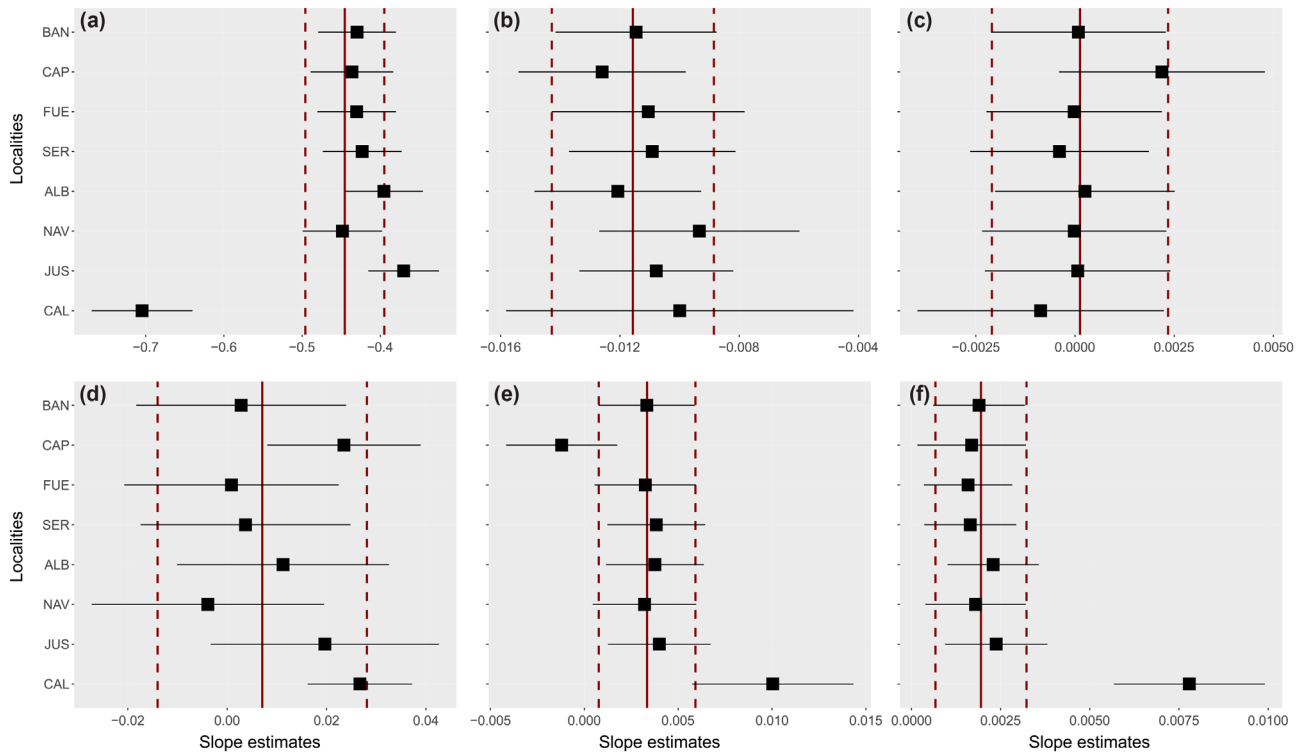


Figure 6. Forest plots showing for the number of alleles (a), the M-ratio (b), the observed heterozygosity (c), the allelic richness (d), the Fis (e) and the expected heterozygosity (f) slope estimates for the relationship between each of these indices and the year of sampling, in the Vïaur metapopulation. Each local population names (BAN, CAP, FUE, SER, ALB, NAV, JUS and CAL, ranked from upstream to downstream) refers to the dataset used to estimate slopes (i.e. the whole dataset minus the named local population). Note that only the sites that were sampled in all years were considered in this analysis. Red vertical line refers to the observed estimate for the whole dataset and dashed lines refer to the confident interval of this estimate.

probably reflects the differential sensitivity of genetic indices to demographic declines, which was recently demonstrated theoretically (Hoban et al. 2014). Some indices could be characterized by non-linear temporal changes that may be detectable only once a threshold is reached (Scheffer et al. 2001), which may require a longer time series to be attained (Antao et al. 2011). In any cases, genetic diversity was globally lower in the Vïaur than in the C  l   metapopulation and both effective metapopulation sizes and the M-ratio were extremely low in the Vïaur metapopulation, even before the beginning of the demographic decline.

Finally, we exemplified how metapopulation structure could dampen the loss of genetic diversity at the metapopulation scale, notably through source local populations, such as the one sampled at the most downstream locality ('La Calqui  re') in the Vïaur River. We notably found that if this local population would not have sustained large dace abundances over time (i.e. would have critically declined instead of being stable, as we actually observed) the decrease in the number of alleles at the metapopulation scale would have been much steeper over this 10-year period. Similarly, this local population also contributed to limit deviation from the Hardy-Weinberg equilibrium and reduced the bottleneck signal observed at the metapopulation scale. Indeed, without this local population, we would have recorded increasing

expected heterozygosity over time, together with a constant number of heterozygotes (measured by Hobs, Fig. 6c, f), which is generally interpreted as a signal of bottleneck (Luikart and Cornuet 1998). This downstream local population was both the most stable demographically and the most diverse genetically, hence contributing to the explanation of its marginal influence at the metapopulation scale. This demonstrates that the spatial variation between demographic parameters and genetic parameters observed in certain metapopulations (such as those living in riverscapes) may dampen the temporal variation observed at the metapopulation scale ('portfolio effect', Doak et al. 1998, Anderson et al. 2013). However, the high level of fragmentation by dams and weirs in the Vïaur River might limit potential genetic and demographic rescue effects provided by emigration from downstream localities to upstream localities (Chen et al. 2016). Thus, any genetic and demographic rescue effect in the Vïaur River would be only possible through human-assisted migration.

It is worth noting that in this study the two metapopulations were analyzed from a linear perspective, i.e. along the upstream-downstream gradient of the main stem. In stream metapopulations, however, we can expect specific demographic and genetic dynamics arising from the dendritic structure of the network (Paz-Vinas et al. 2015, Morrissey and de Kerckhove 2009). For instance, tributaries can in some

cases serve as genetic reservoirs that sustain the entire genetic diversity observed downstream through ‘genetic mass effects’ (Kunin 1998, Campbell Grant et al. 2007, also known as ‘downstream-biased gene flow’, Paz-Vinas et al. 2015). This is nonetheless unlikely in our case study since dace (in the two sub-river basins considered in this study) mainly live the main stem and rarely use tributaries as main habitats (because they are generally too cold and too steep for dace). In both the Célé and Viaur metapopulations, higher levels of genetic diversity observed downstream are hence more likely to be the produce of a past upstream-directed colonization process or of higher effective population sizes downstream rather than from mass effects from upstream to downstream (as observed in many river systems for many organisms, Paz-Vinas et al. 2015). We hence reasonably assume that the demographic and genetic dynamics observed in the most downstream locality of the Viaur River likely reflect a ‘reservoir’ in this area, rather than a sink resulting from a ‘mass effect’. Nonetheless, we argue that further long-term demo-genetic monitoring should be conducted in dendritic river networks to better understand the role of spatial structures in shaping the demographic and genetic stability of the metapopulations inhabiting these specific ecosystems (Anderson et al. 2013).

To conclude, our study demonstrates, in a continuous real-time demo-genetic design, that a demographic decline of a wild metapopulation can be associated with a significant loss in genetic diversity over a period as short as ten years. This shows the usefulness of continuous real-time monitoring to quantify, at a fine time-scale, both the genetic and demographic changes occurring in declining populations, which has strong implications for the prediction and prevention of extinction vortices in the wild. Continuous real-time monitoring combining genetic and demographic indicators is still rarely used (Habel et al. 2014), but we demonstrate that these indicators can jointly vary over short time scales. Moreover, we demonstrate how the spatial arrangement and the specificity of local populations within a landscape can modulate both local extinction rates and local genetic diversity distributions, which may dampen the pervasive genetic effects of demographic declines at the metapopulation scale. We call for more integrated monitoring programmes, because they can realistically inform both the ecological and evolutionary viability of endangered metapopulations.

*Acknowledgements* – We thank all the colleagues that helped with sampling. This work was performed using HPC resources from CALMIP (allocation P1003). This work has been done in two research units (SETE & EDB) that are part of the ‘Laboratoire d’Excellence’ (LABEX) entitled TULIP (ANR-10-LABX-41). This work is part of the project INCLIMPAR (ANR-11-JSV7-0010) supported by a grant from the ‘Agence Nationale de la Recherche’. *Author contributions* – GL and SB coordinated the study, organized the long-term surveys and performed the molecular analyses. EMB, SB, IPV and MC ran statistical analyses. EMB, SB, GL, IPV and MC wrote the manuscript. All authors agreed to be held accountable for the content therein and approved the final version of the manuscript.

*Ethical statement* – Permit numbers required to capture and manipulate fish are provided in Supplementary material Appendix 1 Table A1. These permits are in accordance with current ethical laws in France.

*Conflict of interest* – We have no competing interest.

## References

- Adamack, A. T. and Gruber, B. 2014. Popgenreport: simplifying basic population genetic analyses in R. – *Methods Ecol. Evol.* 5: 384–387.
- Aljanabi, S. and Martinez, I. 1997. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. – *Nucleic Acids Res.* 25: 4692–4693.
- Altermatt, F. 2013. Diversity in riverine metacommunities: a network perspective. – *Aquat. Ecol.* 47: 365–377.
- Anderson, S. C. et al. 2013. Ecological prophets: quantifying metapopulation portfolio effects. – *Methods Ecol. Evol.* 4: 971–981.
- Antao, T. et al. 2011. Early detection of population declines: high power of genetic monitoring using effective population size estimators. – *Evol Appl.* 4: 144–154.
- Barson, N. J. et al. 2009. Population genetic analysis of microsatellite variation of guppies (*Poecilia reticulata*) in Trinidad and Tobago: evidence for a dynamic source–sink metapopulation structure, founder events and population bottlenecks. – *J. Evol. Biol.* 22: 485–497.
- Bolker, B. et al. 2012. Getting started with the glmmADMB. – R package ver. 3.4.3 <<http://glmmadmb.r-forge.r-project.org>>.
- Campbell Grant, E. H. et al. 2007. Living in the branches: population dynamics and ecological processes in dendritic networks. – *Ecol. Lett.* 10: 165–175.
- Cappuccino, N. and Price, P. W. 1995. Population dynamics: new approaches and synthesis. – Elsevier.
- Carlson, S. M. et al. 2014. Evolutionary rescue in a changing world. – *Trends Ecol. Evol.* 29: 521–530.
- Chen, N. et al. 2016. Genomic consequences of population decline in the endangered Florida scrub-jay. – *Curr. Biol.* 26: 2974–2979.
- Clark, J. and Bjørnstad, O. 2004. Population time series: process variability, observation errors, missing values, lags, and hidden states. – *Ecology* 85: 3140–3150.
- De Valpine, P. and Hastings, A. 2002. Fitting population models incorporating process noise and observation error. – *Ecol. Monogr.* 72: 57–76.
- Devillard, S. et al. 2011. Linking genetic diversity and temporal fluctuations in population abundance of the introduced feral cat (*Felis silvestris catus*) on the Kerguelen archipelago. – *Mol. Ecol.* 20: 5141–5153.
- Do, C. et al. 2014. NeEstimator: re-implementation of software for the estimation of contemporary effective population size from genetic data. – *Mol. Ecol. Resour.* 14: 209–214.
- Doak, D. F. et al. 1998. The statistical inevitability of stability–diversity relationships in community ecology. – *Am. Nat.* 151: 264–276.
- Dunning, J. B. et al. 1992. Ecological processes that affect populations in complex landscapes. – *Oikos* 65: 169–175.
- Fagan, W. F. 2002. Connectivity, fragmentation, and extinction risk in dendritic metapopulations. – *Ecology* 83: 3243–3249.
- Frankham, R. 1995. Inbreeding and extinction: a threshold effect. – *Conserv. Biol.* 9: 792–799.

- Fronhofer, E. A. and Altermatt, F. 2017. Classical metapopulation dynamics and eco-evolutionary feedbacks in dendritic networks. – *Ecography* 40: 1455–1466.
- Gaggiotti, O. E. 1996. Population genetic models of source–sink metapopulations. – *Theor. Popul. Biol.* 50: 178–208.
- Garza, J. C. and Williamson, E. G. 2001. Detection of reduction in population size using data from microsatellite loci. – *Mol. Ecol.* 10: 305–318.
- Gelman, A. 2006. Prior distributions for variance parameters in hierarchical models. – *Bayesian Anal.* 1: 515–534.
- Gelman, A. and Rubin, D. 1992. Inference from iterative simulation using multiple sequences. – *Stat. Sci.* 7: 457–511.
- Gotelli, N. J. and Taylor, C. M. 1999. Testing macroecology models with stream-fish assemblages. – *Evol. Ecol. Res.* 1: 847–858.
- Habel, J. C. et al. 2014. The relevance of time series in molecular ecology and conservation biology: temporal comparison of genetic data. – *Biol. Rev.* 89: 484–492.
- Hanski, I. 1998. Metapopulation dynamics. – *Nature* 396: 41–49.
- Hansson, B. et al. 2000. Increase of genetic variation over time in a recently founded population of great reed warblers (*Acrocephalus arundinaceus*) revealed by microsatellites and DNA fingerprinting. – *Mol. Ecol.* 9: 1529–1538.
- Hildner, K. K. et al. 2003. The relationship between genetic variability and growth rate among populations of the pocket gopher, *Thomomys bottae*. – *Conserv. Genet.* 4: 233–240.
- Hoban, S. et al. 2014. Comparative evaluation of potential indicators and temporal sampling protocols for monitoring genetic erosion. – *Evol. Appl.* 7: 984–998.
- Hostetler, J. A. and Chandler, R. B. 2015. Improved state–space models for inference about spatial and temporal variation in abundance from count data. – *Ecology* 96: 1713–1723.
- Hostetler, J. A. et al. 2013. A cat’s tale: the impact of genetic restoration on Florida panther population dynamics and persistence. – *J. Anim. Ecol.* 82: 608–620.
- Jangjoo, M. et al. 2016. Connectivity rescues genetic diversity after a demographic bottleneck in a butterfly population network. – *Proc. Natl Acad. Sci. USA* 113: 10914–10919.
- Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. – *Bioinformatics* 24: 1403–1405.
- Kaeuffer, R. et al. 2007. Unexpected heterozygosity in an island mouflon population founded by a single pair of individuals. – *Proc. R. Soc. B* 274: 527–533.
- Keith, P. et al. 2011. Les poissons d’eau douce de France. – Biotope.
- Kéry, M. and Schaub, M. 2012. Bayesian population analysis using WinBUGS: a hierarchical perspective. – Academic Press.
- Kunin, W. E. 1998. Biodiversity at the edge: a test of the importance of spatial ‘mass effects’ in the Rothamsted Park grass experiments. – *Proc. Natl Acad. Sci. USA* 95: 207–212.
- Labonne, J. et al. 2008. Linking dendritic network structures to population demogenetics: the downside of connectivity. – *Oikos* 117: 1479–1490.
- Luikart, G. and Cornuet, J.-M. 1998. Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. – *Conserv. Biol.* 12: 228–237.
- Mari, L. et al. 2014. Metapopulation persistence and species spread in river networks. – *Ecol. Lett.* 17: 426–434.
- Mathieu-Bégné, E. et al. 2018. Resources for ‘Demographic and genetic collapses in spatially-structured populations: insights from a long-term survey in wild fish metapopulations.’. – Figshare fileset <<https://doi.org/10.6084/m9.figshare.5999171>>.
- Morrissey, M. B. and de Kerckhove, D. T. 2009. The maintenance of genetic variation due to asymmetric gene flow in dendritic metapopulations. – *Am. Nat.* 174: 875–889.
- Naranjo, E. J. and Bodmer, R. E. 2007. Source–sink systems and conservation of hunted ungulates in the Lacandon Forest, Mexico. – *Biol. Conserv.* 138: 412–420.
- Nei, M. et al. 1975. The bottleneck effect and genetic variability in populations. – *Evolution* 29: 1–10.
- Oksanen, J. et al. 2011. vegan: Community Ecology Package. – R package ver. 1.17-2 <<http://CRAN.R-project.org/package=vegan>>.
- Palstra, F. P. and Ruzzante, D. E. 2011. Demographic and genetic factors shaping contemporary metapopulation effective size and its empirical estimation in salmonid fish. – *Heredity* 107: 444–455.
- Paz-Vinas, I. and Blanchet, S. 2015. Dendritic connectivity shapes spatial patterns of genetic diversity: a simulation-based study. – *J. Evol. Biol.* 28: 986–994.
- Paz-Vinas, I. et al. 2013. The demographic history of populations experiencing asymmetric gene flow: combining simulated and empirical data. – *Mol. Ecol.* 22: 3279–3291.
- Paz-Vinas, I. et al. 2015. Evolutionary processes driving spatial patterns of intraspecific genetic diversity in river ecosystems. – *Mol. Ecol.* 24: 4586–4604.
- Plummer, M. 2003. JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. – *Proc. 3rd Int. Workshop Distrib. Stat. Comput.* 124: 125.
- Potvin, D. A. et al. 2017. Genetic erosion and escalating extinction risk in frogs with increasing wildfire frequency. – *J. Appl. Ecol.* 54: 945–954.
- Reed, D. H. et al. 2007. Genetic quality of individuals impacts population dynamics. – *Anim. Conserv.* 10: 275–283.
- Scheffer, M. et al. 2001. Catastrophic shifts in ecosystems. – *Nature* 413: 591–596.
- Schwartz, M. et al. 2007. Genetic monitoring as a promising tool for conservation and management. – *Trends Ecol. Evol.* 22: 25–33.
- Skrbinšek, T. et al. 2012. Monitoring the effective population size of a brown bear (*Ursus arctos*) population using new single-sample approaches. – *Mol. Ecol.* 21: 862–875.
- Spielman, D. et al. 2004. Most species are not driven to extinction before genetic factors impact them. – *Proc. Natl Acad. Sci. USA* 101: 15261–15264.
- Su, Y.-S. and Yajima, M. 2012. R2jags: a package for running jags from R. – R package ver. 0.03-08 <<http://CRAN.R-project.org/package=R2jags>>.
- Vuilleumier, S. et al. 2010. Effects of colonization asymmetries on metapopulation persistence. – *Theor. Popul. Biol.* 78: 225–238.
- Waples, R. 2002. Population viability analysis. – In: Beissinger S. R. and McCullough, D. R. (eds), *Definition and estimation of effective population size in the conservation of endangered species*. Univ. of Chicago Press, pp. 147–168.

Supplementary material (available online as Appendix oik-05511 at <[www.oikosjournal.org/appendix/oik-05511](http://www.oikosjournal.org/appendix/oik-05511)>). Appendix 1.

**AUTHOR:**

Eglantine MATHIEU-BÉGNÉ

**TITLE:**

Infection mechanisms of the ectoparasite of freshwater fish *Tracheliastes polycolpus*

**THESIS SUPERVISORS:**

Géraldine LOOT, Simon BLANCHET and Olivier REY

**DATE AND LOCATION OF THE DEFENCE:**

January, 20<sup>th</sup> 2020 at the Laboratory Evolution and Biological Diversity, Toulouse, France

---

**SUMMARY**

What are the ecological evolutionary mechanisms governing species interactions? I addressed this issue considering the interaction between the parasite *Tracheliastes polycolpus* and several of its cyprinid hosts. First, I highlighted the role of the environment at a very fine spatial scale (few hundred meters) on the encountering rate between this parasite and its hosts. At larger spatial scale (France) I illustrated the limiting role of resistance of host populations on the distribution of *T. polycolpus*. I further showed that *T. polycolpus* manages to infect different host species by relying on a plastic expression of its genes. Finally, I showed that host bacterial communities change after *T. polycolpus* infection in a co-infection dynamics with the parasite. Through the deployment of various and complementary approaches, it appears that both environmental factors, factors related to the host, to the parasite and to their associated bacterial communities contribute to the success of an intimate interaction such as parasitism.

---

**KEY-WORDS:**

Host-Parasite Interactions, Encountering, Compatibility, Genomics, Freshwater Ecosystems

**DOCTORAL SCHOOL AND DISCIPLINE:**

ED SEVAB - Ecology and Evolutionary Biology

**RESEARCH UNIT:**

UMR 5174 – Laboratory Evolution and Biological Diversity, 118 Route de Narbonne, 31077 Toulouse, France

**AUTEUR :**  
Eglantine MATHIEU-BÉGNÉ

**TITRE :**  
Mécanismes d'infection de l'ectoparasite de poissons d'eau douce *Tracheliastes polycolpus*

**DIRECTEURS DE THÈSE :**  
Géraldine LOOT, Simon BLANCHET et Olivier REY

**DATE ET LIEU DE SOUTENANCE :**  
Le 20 janvier 2020 au Laboratoire Evolution et Diversité Biologique à Toulouse, France

---

**RÉSUMÉ**

Quels mécanismes écologiques et évolutifs rendent les interactions entre espèces possibles ? J'ai abordé cette question en considérant l'interaction entre un parasite de poisson d'eau douce (*Tracheliastes polycolpus*) et plusieurs de ses hôtes cyprinidés. J'ai d'abord mis en avant le rôle de l'environnement à une échelle spatiale très fine (quelques centaines de mètres) sur le taux de rencontre entre ce parasite et ses hôtes. A plus large échelle spatiale (la France) j'ai illustré le rôle limitant de la résistance des populations hôtes sur la distribution de *T. polycolpus*. Par ailleurs j'ai montré que *T. polycolpus* parvient à parasiter différentes espèces hôtes en s'appuyant sur une expression plastique de ses gènes. Enfin, j'ai montré que les communautés bactériennes associées à l'hôte changent consécutivement à l'infection parasitaire dans une dynamique de co-infection avec le parasite. Par le déploiement d'approches variées et complémentaires, il apparaît que des facteurs environnementaux, liés à l'hôte, au parasite ou encore à leur cortège de bactéries contribuent à la réalisation d'une interaction étroite et durable tel que le parasitisme.

---

**MOTS-CLEFS :**  
Interaction hôte-parasite, Filtre de rencontre, Filtre de compatibilité, Génomique, Ecosystèmes aquatiques

**ÉCOLE DOCTORALE ET DISCIPLINE :**  
ED SEVAB - Écologie et Évolution

**UNITÉ DE RECHERCHE :**  
UMR 5174 – Laboratoire Evolution et Diversité Biologique, 118 Route de Narbonne, 31077  
Toulouse, France