

Adaptive phenotypic change over the lifetime of an invader - Transcriptomics of a *Cottus* lineage of hybrid origin

Dissertation

in fulfillment of the requirements for the degree
"Doctor rerum naturalium"
of the Faculty of Mathematics and Natural Sciences
at the Christian Albrechts University of Kiel

Submitted by

Elżbieta Anna Iwaszkiewicz-Eggebrecht

Department of Evolutionary Genetics

Max Planck Institute for Evolutionary Biology

Plön, July 2018

First referee: Prof. Dr. Diethard Tautz
Second referee: Prof. Dr. Thomas Roeder

Date of oral examination: 10 October 2018

Approved for publication: 10 October 2018

Signed: Kiel,
Prof. Dr. Frank Kempken, Dean

Table of Contents:

SUMMARY	7
ZUSAMMENFASSUNG	9
GENERAL INTRODUCTION	13
OVERVIEW OF THE DISSERTATION	25
CHAPTER 1	27
TEMPERATURE PREFERENCE AND GROWTH RATE IN THE INVASIVE <i>COTTUS</i> LINEAGE.....	27
INTRODUCTION.....	27
MATERIALS AND METHODS	31
RESULTS.....	36
DISCUSSION AND CONCLUSIONS	40
SUPPLEMENT	44
CHAPTER 2	47
ANNUAL CYCLE OF LOCAL ADAPTATION IN AN EMERGING HYBRID SPECIES.....	47
INTRODUCTION.....	47
MATERIALS AND METHODS	53
RESULTS.....	57
DISCUSSION	71
SUPPLEMENT	75
CHAPTER 3	80
TRANSGRESSIVE TRAIT EXPRESSION IN AN EMERGING HYBRID SPECIES	80
INTRODUCTION.....	80
MATERIALS AND METHODS	86
RESULTS.....	89
DISCUSSION	100
SUPPLEMENT	105
CONCLUDING REMARKS AND PERSPECTIVES:.....	113
REFERENCES:.....	116
DIGITAL SUPPLEMENT	128
DECLARATION OF CONTRIBUTIONS.....	129
ACKNOWLEDGEMENTS	131
CURRICULUM VITAE.....	135

Summary

A central idea about hybrid speciation is that recombination of parental traits can allow hybrids to colonize new ecological niches. However, the connections between ecological context and the genetic traits underlying adaptive evolutionary changes in hybrids remain elusive in most systems. I study a young hybrid lineage of *Cottus* fish that has invaded an environment where its parent species do not occur. While the parental species are limited to summer-cold headwaters, the hybrid *Cottus* colonize downstream areas of rivers that are more exposed to sunlight and significantly warmer in the summer. This suggests that invasive *Cottus* have adapted to a new ecological niche coarsely related to water temperature. I tested temperature preference of the three *Cottus* lineages in the laboratory conditions during summer and found that invasives, despite living in warmer waters, prefer same water temperature as their parental species when given a choice. Then I proceeded to transcriptome scans in natural habitats in order to understand the nature of adaptation to local environments. I used RNAseq to compare, in an unprecedented detail, transcriptome profiles of wild populations over the course of a complete year. I found that different clusters of genes contribute to the differentiation of hybrids from parental species in different seasons. Gene expression profiles that follow changes in temperature in nature were enriched for metabolism-related GO terms and suggest thermal adaptation in the hybrid lineage. However, the data suggest that much of the differentiation between hybrid *Cottus* and their parents is not related to temperature. The analysis is complemented by sampling of transcriptomes from fish raised under controlled laboratory conditions, which reveals heritable components of gene expression divergence. Some of these heritable differences distinguish hybrid *Cottus* from both parents and are likely to be involved in evolutionary novelty after hybridization. Prevailing theory about homoploid hybrid speciation posits that hybridization can give rise to new potentials for adaptation through

transgressive segregation. As a result, the phenotype of a hybrid lineage exceeds that of both parental species, which would equip hybrids with the ability to colonize new ecological niches and thus reinforce speciation. I found ecologically relevant transgressive traits in *Cottus* by comparing expression patterns identified as transgressive in the laboratory with the time-series data from natural habitats. Many of the transgressive genes were connected to metabolic functions suggesting that these genes have altered the metabolism of the invasive lineage relative to the parental species. Another interesting candidate trait shows a transgressive pattern likely due to an increase in copy number and may be involved in the extended red light vision. This possibly constitutes an adaptation to life in the murky waters of downstream habitats of the invasive *Cottus*. Overall, I explored *Cottus* responses to the environment in great depth and traced their genetic basis. The analysis inferred expression phenotypes that constitute worthy avenues for further investigation.

Zusammenfassung

Die Rekombination elterlicher Merkmale von unterschiedlichen Spezies an ihre Nachkommen ist ein wesentlicher Bestandteil der Artbildung durch Hybridisierung. Neu kombinierte Merkmale ermöglichen es den Nachkommen sich an neue ökologische Gegebenheiten anzupassen. Die Interaktionen zwischen ökologischen Faktoren und evolutionären Anpassungen der Hybride sind für die meisten Systeme allerdings kaum verstanden. In dieser Studie wird eine junge Hybridart von *Cottus* untersucht, die einen neuen Lebensraum besiedelt hat, in dem keine der beiden Elternarten vorkommt. Während die Elternarten bevorzugt beschattete, sommerkalte Oberläufe von Flüssen besiedeln, hat sich die Hybridart im Mittel- und Unterlauf von Flüssen angesiedelt. Dieser Bereich ist der Sonne stärker ausgesetzt, was höhere Wassertemperaturen während der Sommermonate verursacht. Dies legt nahe, dass eine Anpassung an eine neue ökologische Nische stattgefunden hat die im Zusammenhang mit der Wassertemperatur steht. Ich habe die Präferenzen der unterschiedlichen *Cottus* Arten für Wassertemperaturen im Sommer unter Laborbedingungen untersucht. Wenn unterschiedlich temperierte Bereiche zur Auswahl stehen, bevorzugt die invasive Hybridart die gleiche Wassertemperatur wie ihre Elternarten. Anschließend habe ich Transkriptome der unterschiedlichen Populationen verglichen um zu untersuchen, wie sie sich an unterschiedliche ökologische Gegebenheiten in natürlichen Lebensräumen anpassen. Mittels RNA-Sequenzierung konnte ich einen ausführlichen Vergleich der Transkriptome über den Zeitraum von einem Jahr durchführen. Es zeigte sich, dass sich Hybride von ihren Elternarten zu verschiedenen Jahreszeiten in mehr Gen-Clustern differenzieren. Genexpressionsprofile die mit Temperaturveränderungen assoziiert sind enthielten einen Überschuss an Genen die eine Rolle im Stoffwechsel spielen und Hinweise auf Temperaturanpassungen der Hybride liefert. Die Daten zeigen jedoch, dass ein Grossteil der Unterschiede zwischen invasiven *Cottus* und deren Elternarten

nicht durch unterschiedliche Temperaturen bewirkt werden. Dies wird durch Transkriptomanalysen unter standardisierten Laborbedingungen unterstützt, die darauf hinweisen, dass erbliche Faktoren Unterschiede in der Genexpression verursachen. Einige dieser vererbten Eigenschaften führen dazu, dass sich die Hybridlinie von beiden Elternarten unterscheidet und stehen vielleicht im Zusammenhang mit der Evolution neuer Merkmale nach der Hybridisierung. Theorien zur homoploiden hybriden Artbildung postulieren, dass Hybridisierung durch transgressive Segregation neue Potentiale für die Adaptation eröffnet, wobei Phänotypen der Hybride das Spektrum beider Elternarten überschreiten. Diese neuen Phänotypen ermöglichen dann die Besiedlung neuer ökologischer Nischen und fördern die Artbildung. Durch den Vergleich der Expressionsmuster von *Cottus*-Populationen aus natürlichen Lebensräumen und aus Laborhaltung konnte ich erstmalig transgressive Merkmale von *Cottus* in ihrem ökologischen Kontext charakterisieren. Mehrere der transgressiven Gene sind mit metabolischen Funktionen assoziiert, was auf Unterschiede beim Metabolismus zwischen den Hybriden und ihren Elternarten hinweist. Ein weiteres interessantes Merkmal der invasiven *Cottus*, welches transgressive Muster zeigt und durch Kopienzahlvariationen ausgelöst wird, hängt vielleicht mit einem verbesserten Sehen im roten Lichtspektrum zusammen. Hierbei könnte es sich um eine Adaptation an das Leben im trüben Wasser der Hauptströme handeln. Ich habe erstmals eine detaillierte Untersuchung der Interaktionen von *Cottus* und seiner Umwelt durchgeführt und deren genetische Grundlagen charakterisiert. Die in dieser Studie vorgestellten Expressionsmuster und Phänotypen tragen zum Verständnis der Umweltanpassungen von *Cottus* bei und ebnen den Weg für weitere Untersuchungen.

“In organic evolution, as in many other affairs, both natural and human, the decisive influence is wielded not by the orthodox majority, but by a significant, progressive minority.”

- G. LEDYARD STEBBINS, 1959

General Introduction

The diversity of life on Earth has always bewildered scientists. The basic unit of diversity is species – an abstract concept, which is itself a matter of wide debate (Coyne and Orr, 2004; De Queiroz, 2007; Mayr, 1996; Wiley, 1978). Despite disagreements on the precise definition, it is the sheer abundance of species that astounds people and many have tried to understand how new species form. A traditional view on the formation of new species – speciation – was that of very gradual slow changes happening in geographically separated populations. Step by step, little by little, those populations would drift apart until a “point of no return”, where they would be too different to recognize each other and reproduce (Mayr, 1963, 1942). More recently however, it has become clear that every so often speciation can happen very rapidly (Dieckmann et al., 2004; Mallet, 2008; Meyer, 1993; Pereyra et al., 2009; Rüber et al., 2003). One of the mechanisms facilitating such drastic evolutionary change is hybridization. This area of research is young and active. A few well-documented cases where hybridization led to the emergence of new species are accompanied by numerous putative cases. Our understanding of the evolutionary processes governing this mode of speciation is limited and demands more study. In my research I am contributing to this vibrant research area. As the introduction to my work I will start by presenting a brief history of research on hybridization.

Hybridization in evolution – brief historical perspective

Since Darwin’s groundbreaking work “On the origin of species” (1859), transitions from populations to species have been mostly viewed as gradual continuums of miniscule changes. In the early days of the integration of evolutionary theory and heredity - the Modern Synthesis – this opinion was still

dominant. Hybridization was considered to have close to no significance in evolutionary terms (Dobzhansky, 1937; Mayr, 1942) or to reinforce barriers between existing species (Bateson, 1909; Dobzhansky, 1933). This view was mostly created and maintained by the influential zoologists such as Theodosios Dobzhansky and Ernst Mayr, who rarely observed hybridization happening. Once the rules of heredity were uncovered, the new field of population genetics arose and gave exciting perspectives for reconstructing species histories and phylogenetic trees from individuals' genetic data. Of course, those reconstructions were possible only under the assumption that separated populations were transforming into species slowly through tiny neutral changes. If hybridization happened, the signal would be disturbed and would make reconstructions impossible. The spirit of that time is well reflected by comments by the "father" of population genetics Sir Ronald Fisher who wrote: "The grossest blunder in sexual preference, which we can conceive of an animal making, would be to mate with a species different from its own (...)" (Fisher, 1930).

Notwithstanding the prevailing opinion in the scientific community, botanists have frequently observed hybridization happening and have studied its role in evolutionary process. The famous botanist G. Ledyard Stebbins from the University of California Davis promoted the view that hybridization plays an important role in evolution throughout his career (Anderson and Stebbins, 1954; Stebbins, 1985, 1959, 1958). A thorough review of his ideas on hybridization is offered by Yakimowski and Rieseberg (2014) in their review paper and will be partially discussed in later parts. Despite work laid by G. Ledyard Stebbins's and other bright botanists, it took many decades to recognize the existence, prevalence and importance of hybridization in nature. Hybridization suddenly started to be frequently reported as researchers began to use genetic molecular techniques, which supplied tools to identify them.

Possible outcomes of hybridization

Nowadays hybridization is recognized as a relatively common phenomenon and a whole array of its evolutionary consequences has been described and documented. The possible outcomes of hybridization range from reinforcement of preexisting reproductive barriers – when interspecific hybrids have very low fitness, the cost of producing them is high and will be strongly selected against - to a complete merger of the species, where barriers between species are modest and the genetic pools of two species intermix. An example of the former is green toads from the Italian island of Sicily, where two closely related species inhabit the same region but do not produce hybrid offspring. As demonstrated by experimental crosses (Colliard et al., 2010), there is high cost of producing hybrids: the first generation hybrid green toads (F_1 hybrids) have reduced fitness and there is a complete and early mortality in the F_2 generation. In contrast, species mergers can be exemplified in the case of black ducks and mallards in North America. European colonization brought the two species into secondary contact where they hybridized, which led to shrinkage of the genetic distance between them over the following century (Mank et al., 2004).

In between those extremes we also distinguish three important outcomes. Firstly, introgression leading to transfer of adaptation after hybridization recombination “cuts out” fragments of the genome and rearranges them randomly. Certain gene alleles from one population, for example A, might turn out extremely useful in the other population, population B. In this case selection will favor the progeny of hybrids with particular allele(s) and the rest of the genome (or almost all of it) will be “cleaned”. An allele from population A might spread throughout population B. Sometimes this phenomenon can facilitate evolutionary rescue when the allele(s) from a foreign population A coming to rescue the genetic condition of a home population B

(Baskett and Gomulkiewicz, 2011). A second important possible outcome is formation of stable hybrid zones. This happens when populations meet, interbreed at a constant, stable rate but interbreeding does not drastically change the genetic composition of either population because admixture of genomes does not spread beyond the vicinity of the hybrid zone. For instance Baltic and Atlantic cod form a stable hybrid zone flanked by pure non-admixed populations in a transition area between the North Sea and the Baltic proper (Nielsen et al., 2003). Last but not the least, hybridization can lead to formation of new hybrid species (Mallet, 2007). After an initial hybridization two outcomes are possible: the progeny can be either polyploid – the whole set of chromosomes from both parental species were passed on and therefore the F_1 hybrid has doubled ploidy - or homoploid – without changes in chromosome number. The latter instance is the subject of our study and will be discussed further.

Homoploid hybrid speciation (HHS)

Hybridization has been primarily thought of as sort of gene flow between populations but in the 2000s this view was re-evaluated and hybridization got a new edge as a creative force in evolution (Abbott et al., 2013; Barton, 2001; Mallet, 2008; Nolte and Tautz, 2010). With this new perspective there was also a big riddle. How would homoploid hybrids gain reproductive isolation? And how would they compete with their parental species, if they are mostly genetically intermediate? Experimental and theoretical work points to the fact that mixing of genetic material of two species can give rise to a new adaptive potential – evolutionary novelty – which would equip hybrids to colonize new ecological niches and therefore promote speciation (Buerkle et al., 2000; Rieseberg et al., 1999). Further steps of the

speciation can be viewed as a special case of ecological speciation (Gross and Rieseberg, 2005; Seehausen, 2004).

Controversy about the limits of the HHS

HHS is a young area of research and the role of hybridization in speciation as well as the frequency at which it happens are still highly debated. Since the late 20th century the number of putative homoploid hybridization events reported has grown steadily, oftentimes instantly calling the lineages identified to have hybrid origins as new hybrid species. The vast majority of people in the field awaited clearer criteria for HHS identification to distill and many scientists contributed to that alongside their research on hybridization (Abbott et al., 2013; Mallet, 2008; Nolte and Tautz, 2010; Seehausen, 2013).

More recently, there has been a tumultuous discussion on the topic of where the limits of HHS are. Molly Schumer and collaborators (2014) proposed a set of stringent criteria to identify a new lineage as a homoploid hybrid species. The criteria included: 1) Strong reproductive isolation between hybrid and parental species, 2) Genetic evidence of hybridization and 3) Hybridization-derived reproductive isolation. The meta-analysis of putative HHS described to date in the literature showed that only one out of twenty-two animal models passed all three criteria. While all studied models presented genetic evidence of hybridization, criteria 1) and 3) proved impossible to meet by almost all studied systems, which led authors to conclude that hybridization might be much less important for speciation than previously thought.

A published response to Schumer et al. (2014) came from Nieto G. Feliner and collaborators (2017) and reflected the views of many other empiricists in the field. They challenged the importance of reproductive isolation in HHS, pointing out that it is not clear at this point whether reproductive isolation is a driver or merely a by-product of primary speciation. Feliner et al. (2017) also emphasized that concentrating on reproductive

isolation diverts attention away from the youngest, nascent hybrid species, which can be most interesting to observe.

The reply to this publication appeared promptly (Schumer et al., 2018) highlighting again the importance of strong reproductive isolation as a prerequisite for the HHS. Authors admitted that their conclusion about HHS being exceptionally rare might be a result of the lack of investigative effort into the reproductive isolation aspect of the HHS rather than the rarity of the phenomenon. However they encouraged studies in this direction, as they believed it to be of great importance to the development of the field. I believe that it is for the benefit of the community to put forward models and rules that stir the discussion. Reproductive isolation is definitely an aspect worth exploring in putative HHS cases. However, I believe that we should not limit ourselves with very restrictive criteria at these still very early days of hybrid speciation research. Assembling more examples and more evidence before we make rules might be the more open and appropriate approach in this case.

Schumer et al. (2018) also highlighted that most of studies done so far have demonstrated genetic evidence for hybridization but they oftentimes lack grounds to prove that traits/aspects that isolate hybrids from their parents are actually a result of hybridization. This opinion is not new and was already expressed multiple times before (Abbott et al., 2013; Barton, 2013; Gross and Rieseberg, 2005). The link between hybridization and isolation leading to speciation is undoubtedly the "Holy Grail" of research on putative HHS and I believe that everybody tries to prove it in his or her respective study system; nevertheless the task remains enormously difficult. Speciation acts on time scales often too wide for humans to directly see because conducting observations generation after generation over extended time and understanding genetic processes as well as ecological forces acting on population is arduous work and practically often not feasible.

However there are, as always, exceptions. A recent, impressive example of HHS comes from a classic study system – Darwin’s finches. Lamichhaney et al. (2018) presented 31 years of detailed history of hybrid speciation starting with one male finch immigrant to the Daphne Major Island in the Galapagos. This individual successfully bred with a local finch species and from the second generation on the progeny bred only within the lineage. The hybrids had big bills that gave them an advantage in competition for resources. The authors demonstrated that increased bill size was a result of hybridization via transgressive segregation. This spectacular example of hybrid speciation proved that sometimes evolution acts phenomenally rapidly and HHS might be in fact quite common.

Methods of investigation

The study by Lamichhaney et al. (2018) relied solely on samples collected in the finches’ natural habitat, but knowledge about the genes involved in bill morphology came from extensive whole-genome sequencing complemented by morphological data on all 13 finch species from Galápagos (Lamichhaney et al., 2015). Most studies demand lengthy observations, many samplings and artificial crosses. To this end, they usually combine observations in nature and experiments on lineages established and living in the laboratory. A frequently used method is common-garden experiments, where parental species and hybrids are reared in identical environments in order to determine how individuals from those groups differ in their responses to the same environment. Here it is assumed that if a hybrid lineage differs from its parental species under common-garden conditions it suggests that the difference is genetic. In studies on local adaptation this experiment is often extended to the common-garden with treatments (Kawecki and Ebert, 2004), where parental species and hybrids are placed in identical environments that differ only in one environmental aspect. Such a setting is designed to dissect plastic

and genetically fixed traits in their responses to the environmental stressors. Laboratory-based experiments allow near perfect uniformity of conditions and convenient manipulation of environmental factors, but to assure true biological relevance of results they are best complemented by observations in the wild and vice versa. To secure high quality data in our study we combined extensive sampling in the natural habitat with carefully designed common-garden experiment with two treatments.

Every experiment demands a phenotype or a set of phenotypes to be observed. A phenotype being increasingly used in comparative studies is the transcriptome. It is the whole set of RNA produced in a cell, population of cells or organism at a given time. Each gene in this set can have unique expression adding to the total gene expression profile. Profiles are therefore individual-specific, heritable and constitute a phenotype that can be subject to studies (Nachomy et al., 2007). The ability to express appropriate genes in appropriate amounts is crucial for organisms' success and survival. Differences between organisms' gene expression profiles are thought to play particularly important roles in early stages of evolutionary processes. It is known that it takes considerable evolutionary time to change protein sequences and that even species drastically morphologically different can have virtually no differences in almost all of their DNA coding sequences. It appears that regulation of gene expression - deciding where (which tissue, cell), when and how much of otherwise identical genes should be transcribed - accounts, in part, for important differences between species (King and Wilson, 1975; Wittkopp, 2013). Comparative transcriptomics focuses on studying variation in gene expression patterns among populations, among species or over time in order to shed light on evolutionary processes that shaped those patterns. The additional advantage of studying gene expression is that we can link it directly to organismal physiology, through genes' functions. For instance, (Derome et

al., 2006) compared transcription profiles of “normal” and “dwarf” lake whitefish ecotypes and found that differentially expressed genes were connected to energy metabolism, which led them to select candidate genes for functional studies on adaptive processes.

Study system

The present study concentrates on a recent homoploid hybridization event in European sculpins – *Cottus* species. (*Cottus*, Cottidae, Teleostei). *Cottus* are freshwater fish, and typically inhabit small and turbulent streams. *Cottus perifretum* is present in the British Isles as well as in the Netherlands and in Belgium, where it inhabits tributaries to river Scheldt. Concurrently *Cottus rhenanus* inhabits very similar small streams that are tributaries to the river Rhine in Germany (Fig. 1). The two species have been separated for over two million years (Englbrecht et al., 2000). Interestingly, in the late 1980s sculpins were reported in the Lower Rhine – a habitat that was previously free of this species (De Nie, 1997). Since then, the fish spread throughout their new habitat of the lower reaches of the river Rhine and its big tributaries such as the river Sieg (Nolte et al., 2005a). Due to this recent range expansion we refer to this lineage of sculpins as “invasives”. Genetic studies revealed that invasive fish are hybrids between *C. rhenanus* and *C. perifretum* (Nolte et al., 2005a, 2005b). Furthermore, it has been observed that despite the lack of geographic or reproductive barriers, invasive fish and their parental species stay separated with only very narrow hybrid zones (Nolte et al., 2006; Stemshorn et al., 2011). These findings suggest that hybridization gave rise to novel traits that allowed rapid adaptation to a new ecological niche and that we are in fact witnessing a homoploid hybrid speciation event in front of our eyes.

Criteria for HHS put forward by Schumer et al. (2014), include reproductive isolation, genetic evidence for hybridization and mechanisms connecting the two. The authors have applied those criteria to all published

putative HHS systems. The *Cottus* system was classified as fulfilling only criterion number 2 – presence of a hybrid genome – and failing in two others. However, Schumer and collaborators seem to consider solely the initial publication by Nolte et al. (2005a) as they did not cite any later works. However, there is a considerable amount of work suggesting that invasive *Cottus* could be considered reproductively isolated from the parental species. Shortly after the initial description of *Cottus* hybridization, Nolte and collaborators (Nolte et al., 2006) demonstrated steep genetic clines in multiple contact zones between *C. rhenanus* and the invasive lineage, suggesting strong selective forces in play. Following studies in the contact zones also observed that despite living in adjacent habitats and occasional mixing parental species and invasive lineage they stay genetically distinct suggesting a strong selection against mixed offspring (Nolte et al., 2009; Stemshorn et al., 2011). The invasive *Cottus* lineage is ecologically reproductively isolated from its parental species. The population genetic patterns are very well studied in the river Rhine system (Nolte et al., 2009; Nolte et al., 2006, 2005a) whereas the transition between populations of *C. perifretum* in their natural range and the invasive lineage in the Scheldt water system remains largely unexplored.

When it comes to the criterion number 3 – the evidence for link between the reproductive isolation and hybridization – it is notoriously difficult to demonstrate in many systems to this date, but the role of reproductive isolation in hybridization is the core question in HHS research. The observed patterns in the *Cottus* system suggest that the invasive lineage has an adaptive advantage compared to the parental species. Nonetheless, the source and nature of this adaptation are not yet fully explored, which is where our efforts concentrate. To date, Nolte and Sheets (2005) explored *Cottus* body shapes using geometric morphometric methods and concluded that the hybrid's lineage shape might be a result of transgressive segregation. Later, Czypionka

et al. (2012) used microarrays to compare transcriptomic phenotypes of the invasive lineage and the parentals in a laboratory common-garden setting. He found that the hybrid species is indeed characterized by a unique transcriptomic profile that is significantly different from the one of first-generation hybrids. The total profile is composed of a mixture of genes that are expressed similarly to either parent and also of genes that are expressed differently from both of them. The latter group, genes expressed differentially from both parental species, was enriched for genes connected to the metabolic pathway hinting at the possible basis of the invasives' adaptation to the new environment.

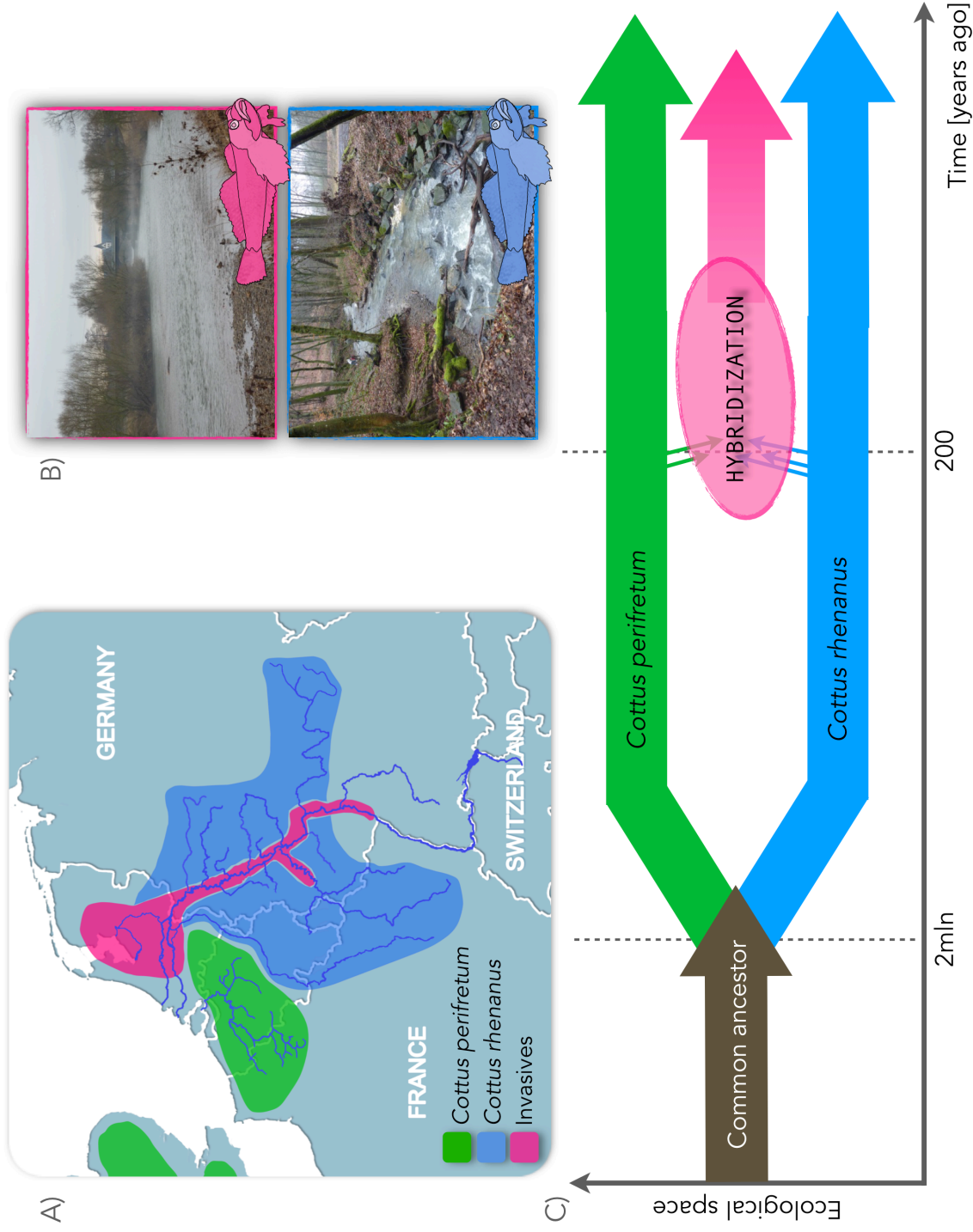


Figure 1. The Cottus system. Panel A presents a map of the Cottus species' distributions. Panel B shows pictures of typical *C. rhenanus*' habitat and invasives' habitat (magenta frame). In panel C we present a scheme of the Cottus system's history (modified from Nolte and Tautz (2010)).

Overview of the dissertation

In this dissertation I sought to uncover the adaptive phenotypic change in an invasive *Cottus* lineage. The lineage in question originated via hybridization and the work presented here contributes to the general understanding of homoploid hybrid speciation process – the formation of new species facilitated by homoploid hybridization.

In the first chapter I explore temperature preference and growth rate of the two parental species and the invasive *Cottus*. I demonstrate that temperature preference of *Cottus* is unaffected by thermal history (no acclimation effect) and remains unaltered in all three lineages, despite differences in summer temperatures in their respective natural habitats. The growth rate assessment under different temperature conditions turned out to be inconclusive due to an infection that affected the fish. Building on this experience I conclude that in order to understand adaptive change in *Cottus* we need to study more complex phenotypes and combine observations in the natural environment with laboratory-based experiments.

In the second chapter I perform an extensive comparative transcriptomic study combining samples from the natural and controlled environments. I describe in an unprecedented detail variation in the transcriptomic phenotype in *C. rhenanus* and the invasives in their natural habitats over the course of a complete year. I found that different clusters of genes contribute to the differentiation of hybrids from parental species in different seasons. Many genes follow changes in temperature in nature, however, the data suggest that much of the differentiation between invasive *Cottus* and their parent is not related to temperature. Further I proceed with comparing patterns from natural environment with the transcriptomes of fish raised under controlled laboratory

conditions in order to uncover the heritable component of the phenotype. I found clusters of genes with high heritability to be enriched for metabolism-related GO terms, which suggests that the hybrid lineage might have altered metabolism and therefore possibly thermal adaptation is taking place.

In the third chapter I explored transgressive transcriptomic traits in *Cottus* and I hypothesize on their possible role in the homoploid hybrid speciation process. I identify expression traits lying outside of the range of parental phenotypes in invasive *Cottus* - transgressive traits – at two time points during their yearly cycle. Further I verify if the expression profiles remain different in the natural contact zone of *C. rhenanus* and the invasives. I found that invasive-specific transgressive traits were enriched for up-regulated genes and their expression patterns were largely conserved in nature. Transgressive traits were representing mostly GO terms connected to primary metabolic processes, which suggest possible adaptation through altering the metabolic rate. Additionally I identified a transgressive gene possibly involved in widening vision in the red light, which might constitute an adaptation to murky waters. Those findings unravel new possible avenues for studies on adaptation in *Cottus* that will demand future functional studies.

Chapter 1

Temperature preference and growth rate in the invasive *Cottus* lineage

Introduction

Temperature affects virtually every living organism. Global scale, long-term observations suggest that temperature limits distribution of plants (Woodward, 1988), determines the precise set-off time for migratory birds (Jenni and Kéry, 2003; Tøttrup et al., 2010) and it is well known to regulate fish distribution ranges (Angilletta, 2009; Perry et al., 2005). Examples multiply in the recent surge of interest for the consequences of climate change (Kelly and Goulden, 2008; Walther et al., 2002). Despite the fact that it is not always clear if temperature is affecting those organisms directly or it is merely a correlate of other climatic factors, we still consider it an important indicator. However, undoubtedly temperature is of paramount importance for ectothermic organisms such as teleost fish or reptiles. Its influence on the distribution of aquatic ectotherms is explained by the concept of oxygen and capacity limited thermal tolerance (OCLTT) (Pörtner and Knust, 2007), which states that thermal range of a species depends on its capability to supply sufficient oxygen to its tissues at respective temperatures. Oxygen's solubility in water decreases with increasing temperature at a rate of around 0.4 mg/L per 1°C, therefore even a small change in temperature can potentially constitute an important challenge to fish physiology and fitness, hence drive adaptation.

Adaptation driven by an environmental factor can motivate a special case of speciation - the ecological speciation (Doebeli and Dieckmann, 2003). Ecological speciation is defined by Patrick Nosil (2012) as "the process by

which barriers to gene flow evolve between populations as a result of ecologically based divergent selection between environments". In this scenario speciation happens as sort of a "by-product" during the adaptive divergence. Naturally, ecological speciation theory was developed to describe primary speciation. However, known examples of species that originated from homoploid hybridization also point to the significance of adaptive divergence in the process. As predicted by simulation study (Buerkle et al., 2000), hybrid speciation happens accompanied with ecological selection and that ecological divergence is crucial to the maintenance of the new hybrid species. This might not be an absolute rule; nevertheless it gives a general direction and provides a framework to study early stages of homoploid hybrid speciation.

In the present study I focus on the freshwater fish system, which consists of two parental species – *Cottus rhenanus* and *Cottus perifretum* – and a lineage that resulted from hybridization between those two. Parental species inhabit small, summer-cold streams, tributaries to the river Rhine and to the river Scheldt respectively. The lineage of hybrid origin, on the other hand, has invaded new vast habitat of the main river Rhine. Due to this invasion we refer to them as "invasives". In some areas the distributions of *C. rhenanus* and the invasives come in direct contact nevertheless, gene flow between those populations remains limited as demonstrated by previous research (Nolte et al., 2006, 2005a; Stemshorn et al., 2011). Thorough analysis of genetic patterns in the contact zones done by Nolte and collaborators (2009) documented conspicuous deviations from neutral patterns of admixture, which can be explained with strong selection against immigrant genotypes. This confirmed that the differentiation into habitats is of adaptive nature.

We know that parental species and the hybrid lineage inhabit different environments and that it is facilitated by adaptive processes. But what is the agent driving this divergent selection? *C. rhenanus* inhabits small streams of

the rhithral zone, which is characterized by fast water flow and high oxygen concentration. They are typically shaded by the forest surrounding them and maintain quite low water temperatures throughout the year. Invasives, on the other hand, live in potamal streams and rivers, which usually flow in between the fields and therefore are more exposed to the sunlight, where they get heated up in summer and cooled down in winter. They are deeper, have much slower water flow, higher temperature amplitude in a year, higher turbidity and low oxygen saturation on the bottom (Giller et al., 1998; Schönborn, 1992). As demonstrated, habitats of *C. rhenanus* and the invasives are fundamentally different: they differ in vast number of biotic and abiotic factors but what seems to be the structuring factor of those habitats is the availability of sunlight and, as a result, the temperature.

This raises the question if the invasive fish prefer different temperatures? Or if their preference stays the same but they differ in physiological response to thermal challenges?

In this study we address those questions in two experiments. In the first one we explore temperature preference of the parental species and hybrid lineage. Considering that invasive *Cottus* thrives in habitat characterized by different temperature regime, we want to know if their preferred, optimal temperature has changed as well? We hypothesize that hybrid fish are physiologically adapted to warmer summer temperatures and will prefer warmer water (H1). Importantly, when designing an experiment to test thermal preference it is necessary to account for acclimation. Acclimation allows individuals to adjust their functioning to changes in the environment in a short period of time. Temperature is an essential factor regulating fish physiology and adjusting to temperature changes is very costly (Schmidt-Nielsen, 1997). We expect that acclimation may affect our experiment. To correct for this we tested two groups of *Cottus* with different thermal histories.

In the second experiment we assess the physiological performance of sculpins in summer maximal temperatures by observing their growth rates. Based on observations and previous study (Nolte et al., 2005a) we know that invasive *Cottus* reaches maturity earlier than *C. rhenanus* and *C. perifretum*. Nolte et al. (2005a) have examined standard length, egg numbers and age, based on otoliths of female sculpins of all three lineages and demonstrated that all fish are one-time spawners. Invasive *Cottus* spawn in the first year of life while *C. rhenanus* grows slower and spawns in the second season. This difference in life history traits may constitute important adaptation to local environment and explain partially invasive *Cottus* success in the new environment. However, it might also be direct, plastic response to warmer temperatures and food availability. We want to explore whether this difference between *C. rhenanus* and invasives is genetically engrained or whether it is due to plasticity. We hypothesize that if this life history difference is genetic, the juvenile invasive fish will grow faster than *C. rhenanus* when put in the same conditions (H2). We predict that all fish will be growing faster in warmer water due to increased metabolic rates (H3) and that this effect will be particularly strong in the *Cottus* lineage of hybrid origin, as it is thriving in these summer-warm habitats in nature (H4).

Materials and Methods

Temperature preference - The preliminary experiment

In the summer of 2011 an intern high school student Jasmin Dehnen performed initial temperature preference experiment with all three lineages of *Cottus sp.* The experiment was done in a special aquarium system consisting of four chambers (27.6l volume each) connected by small passages (diameter ~24.5cm²) at the bottom. The system was set up in a climate-controlled chamber. The temperature of each aquarium was regulated and set to around 15, 18, 22 and 24°C. Fish were introduced to the set-up in groups of 15 individuals of the same lineage and were observed for 3 up to 4 days. Number of fish present in a given temperature compartment was noted and used as an approximation for the temperature preference of the group of fish.

Temperature preference - The main experiment

This experiment was performed in September 2015. We used two lineages - *Cottus perifretum* and invasive *Cottus*. Each lineage was divided into two groups. One was incubated in 19°C and the other in 16°C for 3 months prior to the experiment. In total we tested 30 invasives incubated in 16°C and 30 incubated in 19°C. For *C. perifretum* we used 30 individuals incubated in 16°C, 40 incubated in 19°C.

The experiment was performed in the same aquarium system as the preliminary study (Fig. 1). Water was passing through the aquaria at a flow rate of ~ 300 ml/min. The temperature of each aquarium was regulated and set to 16, 17, 18 and 19°C representing temperature range that fish encounter in nature between stream (16°C) and river habitat (19°C) during summer time. In order to maintain oxygen saturation and mixing of the water column all chambers were constantly aerated through air stones. Each aquarium was equipped with five small caves providing fish with hiding places.

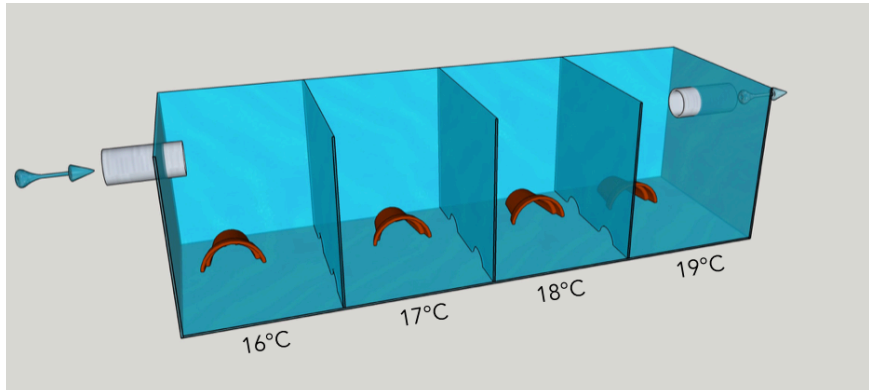


Figure 1. Experimental set-up for the temperature preference experiment. Schematic view of the section of the aquarium. Arrows indicate direction of the water flow.

Fish were put into the system in batches of 10 individuals. Each batch was kept in the experimental system for 2-3 days. During this time constant and equal food supply was granted for each chamber. Four times a day (always at the same hour) positions of fish were noted. Presence in a given temperature compartment was used as an approximation for the temperature preference of the group of fish and was subject to testing. Accounting for “settling in” in the new aquarium, we let the fish get used to new environment. We took counts from the last 24 hours of the experiment for each batch as the valid temperature choice and used them to test for most influential factors and differences between batches by means of generalized linear mixed model (GLMM) as reviewed in (Bolker et al., 2009). All analyses were performed in the R environment (R Core Team, 2017). First we tested for the influence of temperature (location), acclimation and lineage. Secondly we build a model with three-way interaction between temperature, acclimation and lineage. Both analyses were blocked by replicate nested in acclimation to avoid pseudo-replications and introduce repeated measures. Significant effects were analyzed by an *a posteriori* Tukey tests. For visual presentation of the temperature preference data were transformed from counts into percentages and plotted as a histogram.

Growth experiment – the experimental set-up

A temperature treatment was performed in a system similar to the one from the temperature choice experiment with few alterations (Fig. 2). Each tank had a plastic rack hanging at the surface with metal gauze on the bottom. Each rack was divided into 16 compartments (amounting to 64 compartments in total). Every such compartment was designed for one juvenile. Temperature in the first two tanks was adjusted to 17°C and in the other two to 21°C. Each tank was equipped with water pump and constantly aerated with air stones to assure thorough mixing and constant access of oxygen-saturated water to all cages. The total number of individuals tested in 17°C-incubation was 32 – 10 *C. rhenanus*, 11 *C. perifretum* and 11 invasives. In the in 21°C-incubation total number was 32 as well and it consisted of 11 *C. rhenanus*, 10 *C. perifretum* and 11 invasive *Cottus*. Details on individuals used are in the Table S1 in the Supplement. Experiment was performed between 28.07.2015 and 17.08.2015 in the facilities at the Max Planck Institute in Plön, Germany.

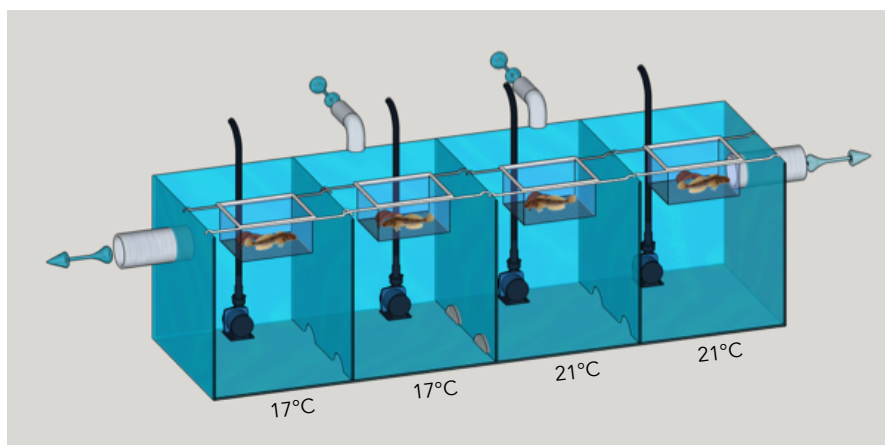


Figure 2. Experimental set-up for the growth rate experiment. Schematic view of a cross-section of the aquarium. Arrows indicate direction of the water flow.

Size assessment

To assess the growth rate, a picture of each individual was taken before and after the fish were incubated in the experimental set-up. Those pictures served as basis for a simple geometric morphometric analysis. A set of six anatomical landmarks was chosen to capture the shape of the specimen from a dorsal view (Fig. 4). Digitizing morphological landmarks was done with use of software `tpsDig ver. 2.10.` and afterwards data was imported to the `CoordGen8` software (Sheets, 2014) from the `Integrated Morphometrics Package` (Sheets, 2000). Size measurements between specimens were standardized based on two landmarks placed on the millimeter paper (Fig. 3). Differences between specimens arising from scaling and/or rotation were removed by Procrustes Superimposition (Zelditch et al., 2012). Afterwards, based on morphological landmarks fish centroids were calculated. Centroid size is a commonly used measure of size in geometric morphometrics. It is essentially calculated as the square root of the sum of the squared distances of a set of landmarks from their center of gravity or centroid (Zelditch et al., 2012). Those centroid sizes served as a proxy for fish size and were compared for each individual before and after the experiment.



Figure 3. Example of a picture of *Cottus* with landmarks placed on. The millimeter paper with a black square serves as a scale. Red dots are placed in the spots where landmarks were marked.

Statistical analyses

I used the centroid sizes to calculate percent growth for each individual. Then individuals' values were grouped according to the lineage the individuals belonged to and the temperature they were incubated in – all together six groups. Further data was processed in the R environment (R Core Team, 2017). I firstly examined normality of distribution by Shapiro-Wilk test (`shapiro.test` function) and homogeneity of variance between all groups by Levene's test (`leveneTest` function). Afterwards I proceeded to the ANOVA analysis by means of an `aoV` function and an *a posteriori* Tukey test (`TukeyHSD` function; confidence level set to 0.95) in order to identify groups with significantly different growth rates.

Study populations

For both experiments – the main temperature preference experiment and the growth rate experiment – I used fish from the same sampling batch. Fish eggs were collected during the breeding season (April 2015): *C. rhenanus* in the Bröl stream (50°48'03.7"N; 7°25'45.4"E) and the invasives in the Sieg river (50°46'23.0"N; 7°19'26.3"E). Eggs were brought to the aquarium facilities in the Max Planck Institute for Evolutionary Biology in Plön and after hatching larvae were fed *ad libitum* with brine shrimp. Due to legal restrictions on fishing in Belgium, we were not authorized to sample *C. perifretum*. However we had a number of families successfully breeding in the lab facilities. After *C. perifretum* laid and fertilized the eggs they were separated from their parent and further treated exactly like those of *C. rhenanus* and the invasives. All larvae and juveniles were kept in tanks with constant water flow through as well as identical photoperiod and temperature conditions imitating the ones that sculpins encounter in their natural habitat – central Europe, until the beginning of the experiment.

Results

Temperature preference - The preliminary experiment

Percentages of fish present in a given temperature tank are summarized in the histogram in the figure 4.

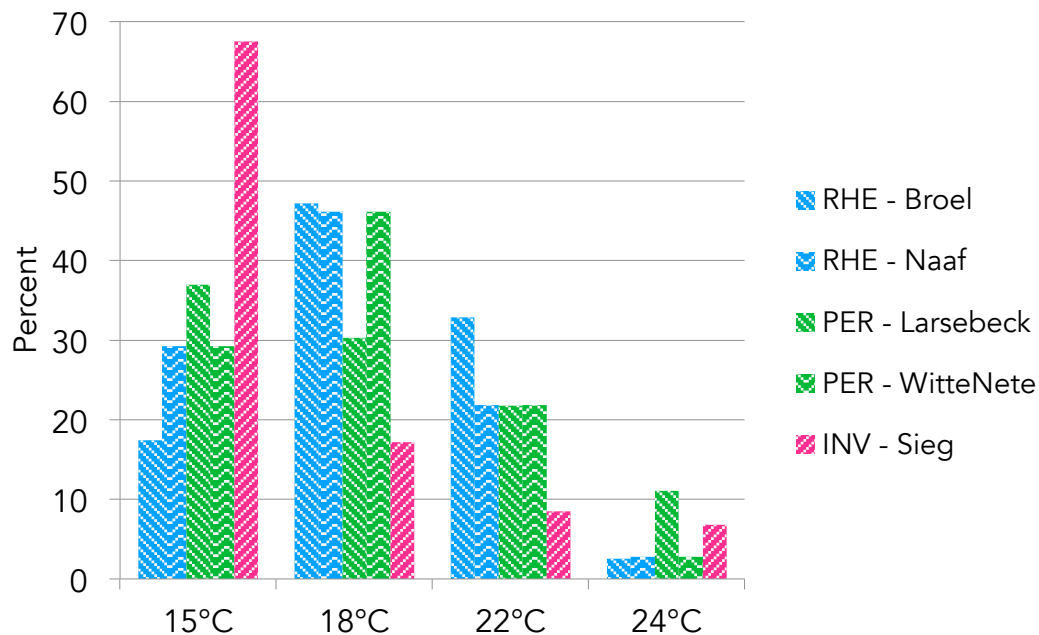


Figure 4. Histogram showing temperature preference of the *Cottus* fish. Bars represent percent of fish occupying a tank with a specified water temperature in the course of a two-day experiment.

Temperature preference - The main experiment

Acclimation was found to have no effect on the final temperature preference in neither of the lineages (GLMM: $\text{Chisq} = 0.7031$, $p = 0.4017$). Regardless of previous thermal experience as well as the genetic background (lineage) fish were showing a preference in terms of tanks (GLMM: $\text{Chisq} = 77.9048$; $p < 2.2e-16$). The Tukey test showed that the fish were choosing the 16°C tank. The percent of fish choosing the 16°C tanks were statistically different from all other tanks (p -values < 0.01 ; details in table S2). The rest of tanks was not different from one another in the acclimation 16°C group. For the 19°C acclimatized group the fish were also preferentially choosing the

16°C tank (p-values <0.01; details in table S3). Additionally, there was one result on the border of significance: the percentage of *C. perifretum* in the tank 17°C was statistically different from the tank 19°C (details in table S3).

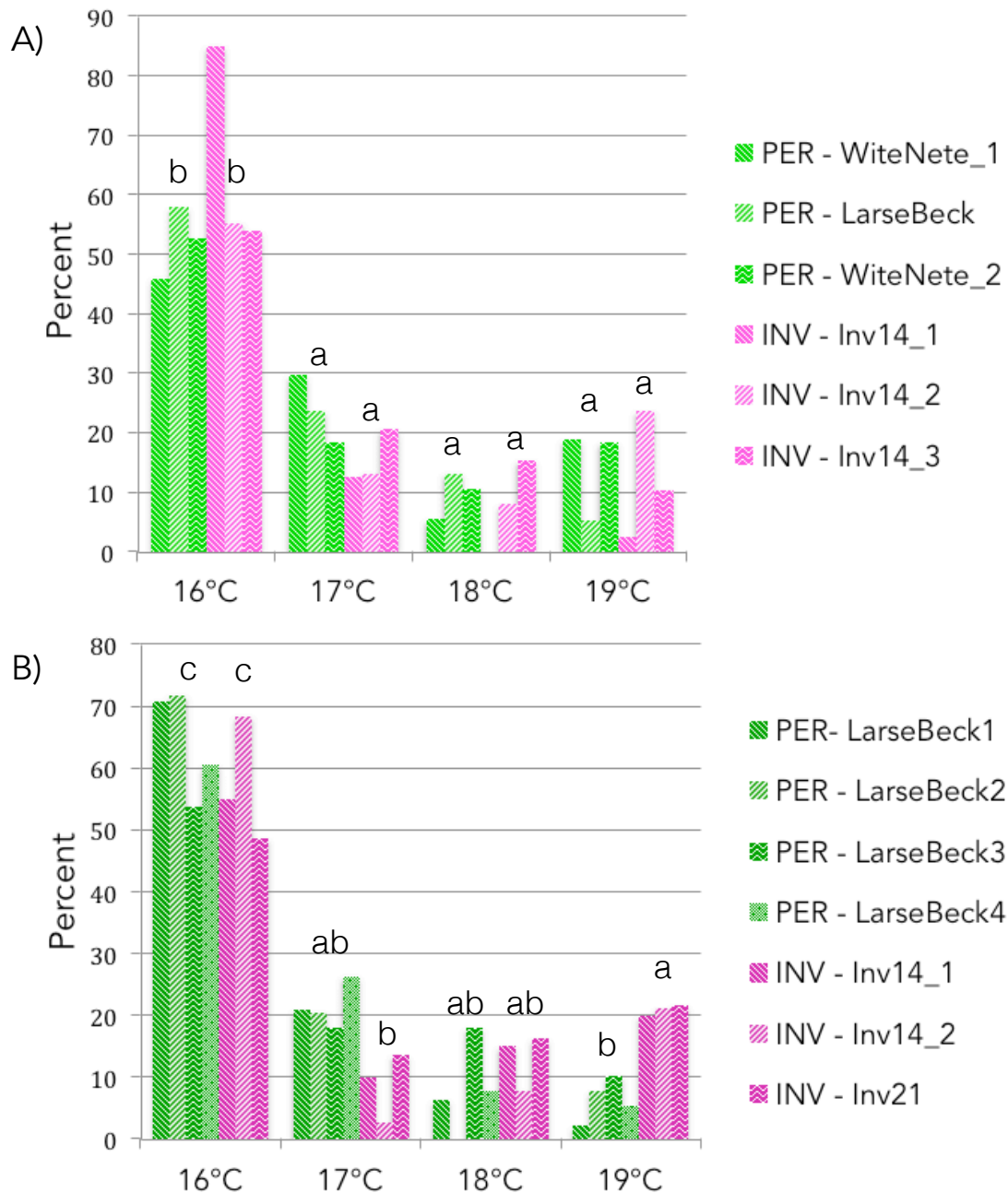


Figure 5. Temperature preference of *Cottus*. Histograms show fish temperature preference in terms of percentages of individuals present in a given tank during the experiment. Panel A shows results for fish acclimatized to 16°C and panel B for the 19°C-acclimatized fish. Water temperature in the tank on X-axis. Letters above bars represent results of multiple comparison test (Tukey test). Groups with same letters are not statistically different from one another. Those with different letter are significantly different (p-values in Table 1 and 2).

Growth experiment

During the experiment, I observed that fish were not feeding properly. Most of the food put into cages remained uneaten. While harvesting the experiment, I noticed that all fish were in poor physical condition. Under the binocular I could see that they had eroded lateral lines and small pits of receding skin around the fish's head. Nevertheless, the pictures were taken in the same fashion as at the beginning of the experiment and relative growth has been calculated. Despite their poor condition, the fish seemed to be growing slightly better in the 17°C incubation. Warmer water (21°C) was unfavorable to all lineages but particularly adverse to the invasives. According to the Shapiro-Wilk test, each group had normally distributed data ($W = 0.9822$, $p\text{-value} = 0.4838$). Levene's test showed that variances were homogenous across the experimental groups; therefore I proceeded with the parametric test. ANOVA analysis found invasives incubated in 21°C to be statistically different from *C. rhenanus* and *C. perifretum* from the 17°C incubation (Fig. 6). All other comparisons were statistically insignificant.

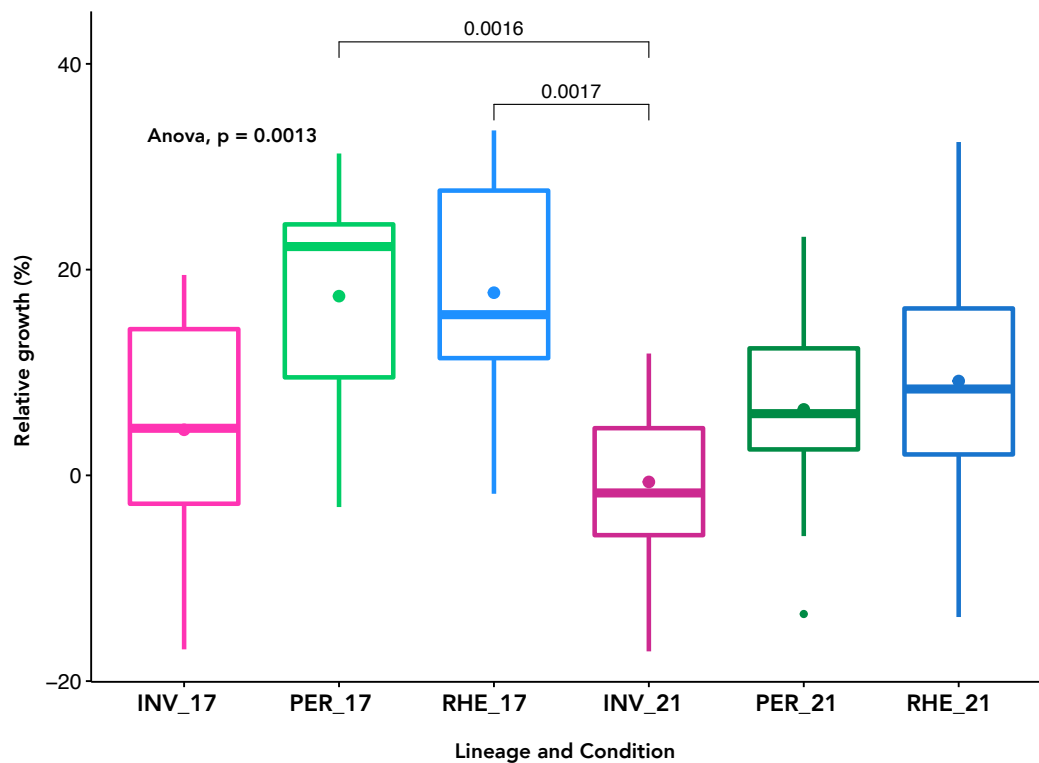


Figure 6. Relative growth of the Cottus fish in two incubation temperatures. Boxplot shows results of a growth experiment. Every box with whiskers represents the relative growth of an experimental group. Colors indicate lineages and the hue is the incubation (brighter: 17°C; darker: 21°C). Dots in boxes show the mean value of the group. Statistically significant ANOVA results are added above boxes.

Discussion and Conclusions

This study provides insights about temperature preference and growth rate of sculpins of hybrid origin under different temperatures. I showed that invasive *Cottus* do not have changed temperature preference when compared to their parental species. Nevertheless, in nature they persist in habitats that are warmer than their preferendum, which suggests that they react to those temperatures differently than the parental species. I attempted to assess growth rate of different *Cottus* lineages in the laboratory under different temperature scenarios. I arrive at the conclusion that in order to understand this system better we need to combine studies in the natural environment and laboratory-based experiments and we should set to study more complex phenotypes.

Thermal preference

Habitat preference has a potential to facilitate adaptive divergence. If two populations, which are at early stages of speciation, present differential habitat preference, they will be increasingly reproductively isolated and exposed to habitat-specific selective agents at the same time. Therefore habitat preference has the potential to act as a major reproductive barrier and therefore as a facilitator to the divergence (Berner and Thibert-Plante, 2015; Rice, 1987). In our study summer temperatures have been identified as the critical difference between the habitat of the hybrid lineage of *Cottus* and parental species (General Introduction). To verify whether preference influenced the adaptive divergence in this system I tested sculpins temperature choice as a proxy for their habitat preference. It turned out that preference, as tested in this study, did not differ between the parental species – *C. perifretum* and *C. rhenanus* – and the invasives. This result was unanticipated since we were expecting that invasive *Cottus*, which lives in summer-warm rivers, will be

adapted to warmer water and will show different preference. Therefore we needed to reject our initial hypothesis. Evidently invasive fish, if given a choice, prefer colder temperatures between 16-17°C. In nature there is a clear adaptive differentiation between *C. rhenanus* and invasive *Cottus*. The outcome of this experiment suggests that this differentiation is probably not driven by temperature choice. However, we still do not know what the nature of this adaptation is and it might still be connected to temperature.

Recognizing the absence of differential temperature preference in the *Cottus* parental species and hybrid lineage, I moved on to assess the performance trade-offs between different habitats. The invasive *Cottus* colonized habitats that have always been accessible to the *C. rhenanus*, nevertheless parental species have never extended their range there. It seems intuitive that there is an ecological barrier - which we believe is related to or at least coupled with temperature differences between habitats - which they could not cross. Perhaps the invasives do not prefer warmer water temperatures but they have special ability, which allows them to tolerate summer-warm waters – something that their parents could not do. It also implies that as far as temperature is concerned, the invasives would prefer to be in the parental's habitat. However, we do not observe them there, which lead us to the conclusion that they present lower fitness in the original habitat when compared with the parentals (Nolte et al., 2006, 2005a).

Acclimation's influence on the thermal preference

Before the actual assessment of the preferred temperature I tested if previous thermal experience (acclimation) was influencing the final temperature that fish choose. To that end, fish were divided into two groups, which were kept under different temperature regimes for 3 months. One group was experiencing temperatures as in the invasives' habitat, while the second one as

in the parentals' habitat. I observed that fish were choosing water temperature independent of their thermal history. This result corroborates previous results obtained in other fish species (Crawshaw, 1975; Schmidt-Nielsen, 1997). Repeating experiments that have demonstrated long time ago lack of acclimation's influence on the final thermal preferendum of fish might seem redundant however, it is worth mentioning that fish are highly diversified group with myriad of different systems of solving problems they encounter and adaptations to nature (Balcombe, 2016) and there are known examples of fast acclimation that not only persists throughout lifetime of the fish but also can be transferred to the next generation (Donelson et al., 2012).

In case of our particular study, we noted that among fish acclimatized to 19°C, invasive *Cottus* were encountered more often in the 19°C tank. Despite the fact that this difference between *C. perifretum* and the invasives was statistically insignificant, it presents a trend. At times a nearly significant result implies that an in depth analysis is warranted. It is perfectly conceivable that the sample size was too small to reach significance level given the effect size. Perhaps the hybrid lineage of *Cottus* has more flexible thermal preference.

Growth rate

In order to shed light on the differences in physiological performance between the hybrid lineage and the parental species I observed fish growth rates in the laboratory under simulated conditions of summer-maximal temperatures. My expectation was that the invasive *Cottus* would present more robust performance in warm water and grow faster. Unfortunately we could not draw any concrete conclusions from this experiment because the fish were in very poor physiological state suggesting that physiological performance including growth were impaired. Many of the fish suffered head lesions and lateral line erosion, which is known as a sign for disease in aquarium fish (Noga,

2011). Additionally, because of this disease, they were not feeding well and some of the fish even drastically lost weight during the experiment. Increased temperature seemed to further worsen the state of fish. As a result of those problems we observed average modest growth rate in the 17°C condition and visibly poor in the 21°C condition. Only two comparisons were statistically different. Although we believe that those differences reflected groups more or less severely touched by disease than the physiological performance that is relevant in the natural habitat. We therefore cannot treat those results as representative of the real situation. Work with live animals in the laboratory setting poses many challenges. It is very difficult to ensure the wellbeing of fish in aquaria sometimes even despite many years of experience.

Conclusions

Building on our experience gained in this project we acknowledge the need to study divergent adaptation in the *Cottus* system both in the natural habitat as well as in the controlled environment. Only by combining those approaches we will be able to get insight into real biological differences. Since the invasive *Cottus* range cannot be explained by simple differential habitat preference based on temperature, we should set to study wider range of temperatures and phenotypes. The natural environment in central Europe varies vastly in yearly cycles. I concentrated on summer temperatures, as it is the time when differences between habitats are most pronounced. It is however very possible that the crucial moment in the differentiation between invasives and parents is happening outside of summer temperatures. I note the necessity to study further the variation in the natural system, which could be achieved by sampling more time points. Finally I conclude that in order to understand the basis of this adaptation it is necessary to study more complex phenotypes.

Supplement

Table S1. List of individuals used in the growth experiment. Columns indicate lineage of origin (Lineage), population (Pop. of origin) and the temperature that the individual was exposed to during growth experiment (Incubation). "Bröl" means sampling station in the Bröl stream at 50°48'03.7"N; 7°25'45.4"E; "Pleis" is short for the sampling station located at 50°46'23.0"N; 7°19'26.3"E; "Witte Nete" is name of the stream where *C. perifretum* was sampled, exact location not known.

ID	Lineage	Pop. of Origin	Incubation	ID	Lineage	Pop. of Origin	Incubation
1	RHE	Bröl	17°C	33	RHE	Bröl	21°C
2	RHE	Bröl	17°C	34	RHE	Bröl	21°C
3	RHE	Bröl	17°C	35	RHE	Bröl	21°C
4	RHE	Bröl	17°C	36	RHE	Bröl	21°C
5	RHE	Bröl	17°C	37	RHE	Bröl	21°C
6	RHE	Bröl	17°C	38	RHE	Bröl	21°C
7	RHE	Bröl	17°C	39	RHE	Bröl	21°C
8	RHE	Bröl	17°C	40	RHE	Bröl	21°C
9	RHE	Bröl	17°C	41	RHE	Bröl	21°C
10	RHE	Bröl	17°C	42	RHE	Bröl	21°C
11	PER	WitteNete	17°C	43	RHE	Bröl	21°C
12	PER	WitteNete	17°C	44	PER	WitteNete	21°C
13	PER	WitteNete	17°C	45	PER	WitteNete	21°C
14	PER	WitteNete	17°C	46	PER	WitteNete	21°C
15	PER	WitteNete	17°C	47	PER	WitteNete	21°C
16	PER	WitteNete	17°C	48	PER	WitteNete	21°C
17	PER	WitteNete	17°C	49	PER	WitteNete	21°C
18	PER	WitteNete	17°C	50	PER	WitteNete	21°C
19	PER	WitteNete	17°C	51	PER	WitteNete	21°C
20	PER	WitteNete	17°C	52	PER	WitteNete	21°C
21	PER	WitteNete	17°C	53	PER	WitteNete	21°C
22	INV	Pleis	17°C	54	INV	Pleis	21°C
23	INV	Pleis	17°C	55	INV	Pleis	21°C
24	INV	Pleis	17°C	56	INV	Pleis	21°C
25	INV	Pleis	17°C	57	INV	Pleis	21°C
26	INV	Pleis	17°C	58	INV	Pleis	21°C
27	INV	Pleis	17°C	59	INV	Pleis	21°C
28	INV	Pleis	17°C	60	INV	Pleis	21°C
29	INV	Pleis	17°C	61	INV	Pleis	21°C
30	INV	Pleis	17°C	62	INV	Pleis	21°C
31	INV	Pleis	17°C	63	INV	Pleis	21°C
32	INV	Pleis	17°C	64	INV	Pleis	21°C

Table S2. Results of the Tukey HSD test for samples from the 16°C acclimation. Presents results of all comparisons between four experimental tanks and two experimental groups. PER – *C. perfretum*; RHE – *C. rhenanus*; INV – invasives; numbers signify tank of a specific temperature.

Contrast	Estimate	SE	z.ratio	p.value	signif.
INV.16 - PER.16	0.253	0.174	1.459	0.829	ns
INV.16 - INV.17	1.440	0.262	5.495	<.0001	***
INV.16 - PER.17	1.035	0.224	4.619	0.0001	***
INV.16 - INV.18	2.134	0.353	6.052	<.0001	***
INV.16 - PER.18	1.933	0.323	5.992	<.0001	***
INV.16 - INV.19	1.692	0.291	5.817	<.0001	***
INV.16 - PER.19	1.558	0.275	5.665	<.0001	***
PER.16 - INV.17	1.187	0.269	4.409	0.0003	***
PER.16 - PER.17	0.782	0.232	3.364	0.0175	*
PER.16 - INV.18	1.880	0.358	5.254	<.0001	***
PER.16 - PER.18	1.680	0.328	5.114	<.0001	***
PER.16 - INV.19	1.438	0.297	4.839	<.0001	***
PER.16 - PER.19	1.305	0.282	4.630	0.0001	***
INV.17 - PER.17	-0.405	0.304	-1.332	0.887	ns
INV.17 - INV.18	0.693	0.408	1.698	0.689	ns
INV.17 - PER.18	0.492	0.383	1.287	0.904	ns
INV.17 - INV.19	0.251	0.356	0.705	0.997	ns
INV.17 - PER.19	0.118	0.344	0.343	1.000	ns
PER.17 - INV.18	1.099	0.385	2.854	0.082	ns
PER.17 - PER.18	0.898	0.358	2.510	0.191	ns
PER.17 - INV.19	0.657	0.329	1.994	0.486	ns
PER.17 - PER.19	0.523	0.315	1.658	0.714	ns
INV.18 - PER.18	-0.201	0.449	-0.446	0.999	ns
INV.18 - INV.19	-0.442	0.427	-1.034	0.969	ns
INV.18 - PER.19	-0.575	0.417	-1.381	0.866	ns
PER.18 - INV.19	-0.241	0.403	-0.599	0.999	ns
PER.18 - PER.19	-0.375	0.392	-0.957	0.980	ns
INV.19 - PER.19	-0.134	0.366	-0.365	1.000	ns

Table S3. Results of the Tukey HSD test for samples from the 19°C acclimation. Presents results of all comparisons between four experimental tanks and two experimental groups. PER – *C. perfretum*; RHE – *C. rhenanus*; INV – invasives; numbers signify tank of a specific temperature.

Contrast	Estimate	SE	z.ratio	p.value	signif.
INV.16 - PER.16	-0.118	0.159	-0.741	0.996	ns
INV.16 - INV.17	1.887	0.339	5.561	<.0001	***
INV.16 - PER.17	0.951	0.211	4.505	0.0002	***
INV.16 - INV.18	1.482	0.286	5.180	<.0001	***
INV.16 - PER.18	1.912	0.303	6.302	<.0001	***
INV.16 - INV.19	1.012	0.238	4.244	0.0006	***
INV.16 - PER.19	2.280	0.355	6.417	<.0001	***
PER.16 - INV.17	2.005	0.332	6.042	<.0001	***
PER.16 - PER.17	1.069	0.199	5.377	<.0001	***
PER.16 - INV.18	1.599	0.277	5.773	<.0001	***
PER.16 - PER.18	2.030	0.295	6.882	<.0001	***
PER.16 - INV.19	1.129	0.228	4.964	<.0001	***
PER.16 - PER.19	2.398	0.348	6.887	<.0001	***
INV.17 - PER.17	-0.936	0.360	-2.602	0.155	ns
INV.17 - INV.18	-0.405	0.408	-0.993	0.976	ns
INV.17 - PER.18	0.025	0.421	0.060	1.000	ns
INV.17 - INV.19	-0.875	0.376	-2.326	0.279	ns
INV.17 - PER.19	0.393	0.459	0.855	0.990	ns
PER.17 - INV.18	0.531	0.310	1.712	0.680	ns
PER.17 - PER.18	0.961	0.326	2.948	0.063	ns
PER.17 - INV.19	0.061	0.267	0.227	1.000	ns
PER.17 - PER.19	1.329	0.375	3.546	0.009	**
INV.18 - PER.18	0.431	0.379	1.137	0.949	ns
INV.18 - INV.19	-0.470	0.329	-1.428	0.845	ns
INV.18 - PER.19	0.799	0.422	1.894	0.555	ns
PER.18 - INV.19	-0.901	0.344	-2.616	0.150	ns
PER.18 - PER.19	0.368	0.434	0.848	0.990	ns
INV.19 - PER.19	1.269	0.391	3.245	0.026	*

Chapter 2

Annual cycle of local adaptation in an emerging hybrid species

Introduction

Homoploid hybridization is now accepted as a common phenomenon both in plants and animals (Abbott et al., 2016; Feliner et al., 2017; Yakimowski and Rieseberg, 2014). There is also growing evidence that hybrid lineages can give rise to new species (Abbott et al., 2013; Mallet, 2008; Payseur and Rieseberg, 2016; Wolf et al., 2010). The best-described example of such ecological hybrid speciation is the sunflower species complex, where three hybrid lineages adapted to extreme new environments (Rieseberg et al., 2003). Despite the fact that hybridization was long thought to take place predominantly in plants, examples from the animal kingdom have multiplied in recent years, including *Heliconius* sp. butterflies (Mavárez et al., 2006; Melo Maria et al., 2009; Salazar et al., 2010), Italian sparrow (Hermansen et al., 2011; Trier et al., 2014) as well as fish: swordtail fish (Cui et al., 2013; Kang et al., 2013), cichlid fish (Keller et al., 2012) and sculpins (Nolte et al., 2005a; Stemshorn et al., 2011). However our knowledge about the nature of hybrid speciation in animals is still limited.

According to theoretical modeling approaches (Buerkle et al., 2000), as well as the before mentioned empirical examples, hybrid speciation can happen when a hybrid lineage invades a new ecological niche, previously uninhabited by either of its parental species. Therefore, in order to understand the process of hybrid speciation, we need to uncover how hybrids cope with and adapt to new habitats (Mallet, 2008; Nolte and Tautz, 2010).

The success of any invasion depends on the ability of invaders to cope with a novel and heterogeneous environment. In fact, their survival and establishment success depends on the adaptive variation, which consists of two components: (i) the **plastic phenotype** – when a given genotype can produce several phenotypic states depending on the environment; we can consider this state “transient”, as in the absence of a given environmental cue the phenotype will not be realized; (ii) the **adaptive phenotype** – when natural selection has favored over time a specific genetic make-up of a population and that enhances fitness of individuals in a particular ecological setting (Price et al., 2003); it is heritable and is subject to adaptive evolution; we can call this state “constant” because even in the absence of the environmental factor to which the organism adapted the phenotype is genetically fixed and expressed.

In order to learn about the nature of evolutionary divergence during hybrid speciation, it is necessary to dissect plastic and genetically fixed adaptive phenotypic traits. This can be achieved by combining two complementary approaches: common-garden experiment and observation in the natural habitat as proposed by Czypionka et al. (2018).

Direct observations in the natural environment present us with a detailed picture of what is relevant for population’s survival and fitness. A natural environment is characterized by a multitude of variable factors that change constantly forcing organisms to adjust. This complexity is impossible to imitate in any laboratory setting and thus observations in nature provide us with an unmatched wealth of information. Those observations, however, encompass both heritable and plastic traits and to dissect those we need to use a complementary approach. A method widely used to expose the genetically determined, heritable component of adaptation are common-garden experiments (Kawecki and Ebert, 2004; Villemereuil et al., 2016). The guiding

principle here is that if you take two populations, which have adapted to different environments (by changes in allele frequencies, mutations or other) and put them in a consistent, identical, laboratory based environment, their plastic phenotypic responses will be similar, but genetically fixed differences will be exposed, as they are expressed no matter the environment.

Studying transcriptomics can provide insight into both plastic and adaptive components of a phenotype. Gene expression patterns are determined both by genetics – are heritable – as well as by the environment, can vary rapidly depending on external factors (Alvarez et al., 2015; Nachtomy et al., 2007; Todd et al., 2016). Expression profiles are multidimensional, provide instant information about thousands of genes and can suggest links between genetics and physiological traits involved. Since a little over a decade, thanks to decreasing costs, we have observed a rise in the number of transcriptomic studies done directly in the wild – the ecological transcriptomics (Alvarez et al., 2015). For instance (Czypionka et al., 2018) combined transcriptomic scans of two ecotypes of the European Fire Salamander in nature and in the laboratory setting. They were able to identify a fraction of genes that differed between ecotypes and for which the expression profile was determined by the underlying genetics – those constitute candidate genes for being involved in adaptive divergence between the two ecotypes.

Comparative transcriptomics shows a lot of promise but is also, just like any other method, imperfect and has its limitations. Almost all studies up to now describe transcriptomic profiles that capture one moment in time and therefore represent only a snapshot of patterns that most likely change over time. Moreover, an analysis of the temporal co-variation of gene expression patterns with environmental factors and throughout the lifetime of an organism may pinpoint the functional context of changes in gene expression patterns. Although our understanding of natural phenomena depends directly on

temporal and spatial scales that they were observed at, these types of analyses are scarce (Alvarez et al., 2015). Relevant ecological patterns reveal themselves only when observations are made over a reasonable temporal and spatial scale coverage (Estes et al., 2018). Gene expression is highly variable in response to the environment, can change instantly and is known to vary in circadian and yearly cycles (Dopico et al., 2015). It appears obvious that more gene expression scans are needed to fish out the relevant patterns.

I address this common shortcoming by using a time-series of gene expression scans. I compare, in an unprecedented detail, transcriptome profiles of wild populations of hybrid and its parental species in the course of a complete year. The analysis is complemented by sampling of transcriptomes from individuals raised under controlled laboratory conditions in two time points in the course of a year: in summer – when differences in temperature in the natural habitat are maximal – and in winter – when differences are small. This combined experimental design will reveal heritable components of gene expression divergence.

We study a recent homoploid hybridization event in European sculpins – *Cottus* sp. (*Cottus*, Cottidae, Teleostei) system. *Cottus* sp. are freshwater fish, which inhabit shallow, turbulent, summer-cold streams. *Cottus perifretum* is present on British Isles as well as in the Netherlands and Belgium, where it inhabits tributaries to river Scheldt. *Cottus rhenanus*, on the other hand, inhabits small streams that are tributaries to the river Rhine (Germany). The two species have been separated for over two million years. Interestingly, in late 1980' presence of sculpins has been reported in the Lower Rhine (De Nie, 1997) habitat previously free of this species. Since then, the fish have spread throughout new, summer-warm habitat - the lower reaches of the river Rhine and its big tributaries such as the river Sieg (Nolte et al., 2005a). Due to this recent range expansion we refer to this lineage of sculpins as “invasives”.

Genetic studies revealed that invasive fish are hybrids between *C. rhenanus* and *C. perifretum* (Nolte et al., 2005a). Furthermore, it has been observed that despite the lack of major geographic or reproductive barriers, invasive fish and their parental species stay separated with only very narrow contact hybrid zone (Nolte et al., 2009; Nolte et al., 2006; Stemshorn et al., 2011). Those findings are particularly exhilarating because they suggest that hybridization gave rise to new traits that allowed rapid adaptation in a new ecological niche.

Heretofore, Czypionka et al. (2012) have initially explored transcriptome divergence between *Cottus sp.* They studied gene expression variation in the two parental species, invasive lineage as well as lab-bred F2 hybrid crosses with the use of custom-designed oligonucleotide microarrays. The study revealed, among other results, that all three *Cottus* lineages are overall distinguished by specific gene expression patterns. *C. rhenanus* and *C. perifretum* were most different from each other and invasive *Cottus* showed an intermediate phenotype. However, it was also characterized by a set of uniquely regulated genes. This set was enriched for terms such as DNA metabolic processes, motor activity, transcription factor activity – suggesting that concerted changes in genes linked to metabolic rate regulation and other metabolic-related processes might have contributed to the adaptive advantage of the invasives (Czypionka et al., 2012). These results are encouraging and corroborate that transcriptomics is a revealing approach to study incipient hybrid speciation in *Cottus sp.*

In the *Cottus sp.* system ecological differences between upstream and downstream habitats involve rather different animal and plant communities, but are ultimately attributed to different light and temperature regimes. Temperature is of paramount importance to ectothermic organisms such as fish and it has been repeatedly identified as a limiting factor to species' distributions (Angilletta, 2009; Perry et al., 2005). It has been shown that

divergent selection between thermal environments can be strong enough to maintain a stable barrier between populations (Keller and Seehausen, 2012). As it is most likely the case that the thermal regimes between those two habitats differ, we aim to study changes in thermal differentiation across the seasons.

The goal was to use transcriptome differentiation as an indicator of different physiological performance of *Cottus* lineages. We did not have access to wild *C. perifretum* in this study, but we note that the precise patterns of distribution of *C. perifretum* and invasive *Cottus* in their shared range are not known. Hence this study focuses on the genetically and ecologically better studied contact zone of *C. rhenanus* and invasive *Cottus* (Nolte et al. 2005, 2006, 2009).

Under the assumption that temperature is the structuring factor in this system, one can hypothesize that (i) gene expression patterns in nature should be most different between lineages when temperature differs the most. Furthermore one can expect that there is (ii) a fixed genetic component that gives advantage in the rapid invasion of *Cottus* and one should observe important fixed adaptive gene expression differences between *C. rhenanus* and the invasive *Cottus* in the common garden experiment. Moreover one can anticipate as well finding (iii) patterns of gene expression divergence, which are temperature independent and signal other possible adaptive processes.

Materials and Methods

Sample collection – Time series from the natural environment

We collected individuals of one parental species (*C. rhenanus*) and invasive *Cottus* in the wild at nine time points throughout the year. *Cottus rhenanus* sampling station was located in the Bröl stream (50°48'03.7"N; 7°25'45.4"E) and invasives were caught in the Sieg river (50°46'23.0"N; 7°19'26.3"E). Sample sizes and exact dates of sampling are presented in the results section – Figure 2.

Directly after being captured every fish was anaesthetized with MS-222 (tricaine methanesulfonate, 200 mg/L) and killed by a blow on the head. Fish were subsequently dissected and put into RNAlater until transported to the laboratory. In the laboratory all fish were measured (standard length), sexed (if possible) and dissected. Pectoral fins and liver were removed and stored in RNAlater in -20°C until RNA extraction. The remaining fish bodies were also stored immersed in RNAlater in -20°C.

The two tissues were chosen for transcriptomic analysis because: 1) the liver is an organ that plays a major role in metabolism and thus particularly interesting in light of previous findings (GO enriched terms; Czypionka *et al.* 2012); 2) the fin is a good candidate peripheral tissue for non-lethal sampling potentially performed on same individual in the course of time, which could be useful for future studies.

Temperature measurements

Hourly measurements of the water temperature are made available by the Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen (LANUV). We were able to track temperature changes from 2012 until the end of 2017 and, for specific data, choose a station in the proximity to our sampling sites.

Sample collection – Controlled environment

Cottus sp. were reared and bred in our climate-controlled chambers in the laboratory. Light cycle and temperature were manipulated to imitate the conditions found in the wild in *Cottus rhenanus* habitat. We sampled 8 fish per lineage twice, once in January 2015 – winter condition – and once in September 2015 – summer condition. Every lineage was equally represented and balanced for sex whenever possible. The total number of individuals was 48 (details in table X and X in the supplement). Fish were anaesthetized with MS-222 (tricaine methanesulfonate, 250 mg/L) and killed by a blow on the head. Pectoral fins and livers were stored in RNAlater in -20°C until RNA extraction. The remaining fish bodies were also stored immersed in RNAlater in -20°C.

RNASeq and mapping

Total RNA was extracted from both tissues using a Trizol protocol (Invitrogen, Paisley, UK) and additionally cleaned with a lithium chloride precipitation. RNA quantity and integrity were assessed by spectrophotometer, gel electrophoresis and using an Agilent RNA 6000 Nano Chip (Agilent technologies, 2015). Samples that passed our quality assessment were shipped to Cologne Center for Genomics (CCG, Cologne, Germany) where RNA was poly-A enriched, reverse transcribed to cDNA with use of TrueSeq Stranded mRNA Kit and sequenced on Illumina HiSeq 4000 platform.

In the absence of a complete, annotated *Cottus sp.* genome I made a custom reference for mapping. I used a transcriptome assembly previously done for one male *Cottus rhenanus* individual with Trinity software (Grabherr et al., 2011). It was based on 6 libraries made for 6 tissues selected to represent different parts of the body for a complete inventory of genes – gills, muscle, liver, digestive tract, head, fin/skin. The complete assembly

counted ~108 thousand contigs. Considering that *Cottus* genome is ~20 thousand genes, there were probably multiple contigs representing same gene in many cases. I mapped them using the `bwa-mem` mapper (Houtgast et al., 2016) against all known stickleback genes from BioMart ENSEMBL database version 83 (Yates et al., 2016) and kept the best match for each contig. Those matches created a sort of “dictionary” that assigned *Cottus* Trinity contigs to known stickleback genes whenever unambiguous.

Paired-end RNASeq reads for each individual sample were mapped using `bwa-mem` to the complete Trinity assembly. In a second step, read counts were integrated for each annotated stickleback gene based on the matches that were found (illustrated in Fig. 1).

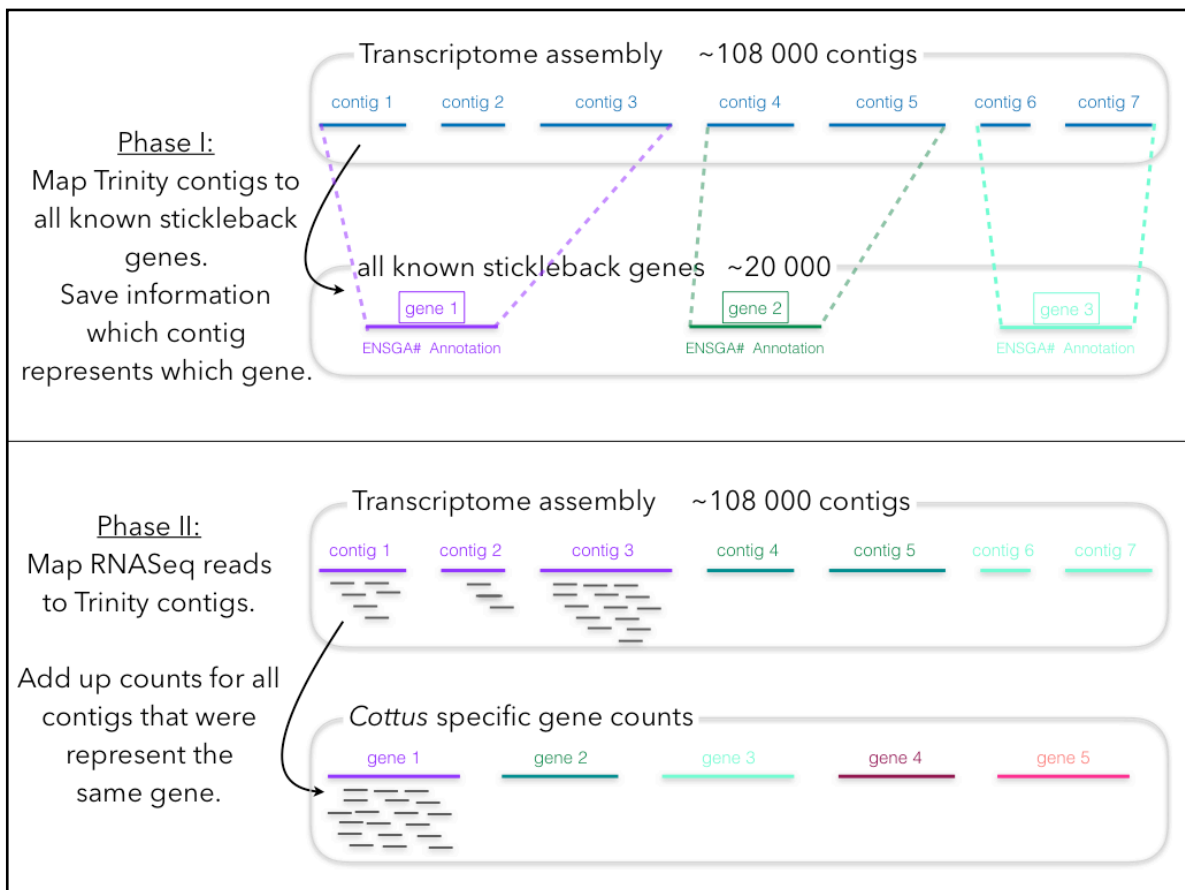


Figure 1. Strategy used to map *Cottus*-specific reads and assign them to annotated genes as identified in the stickleback genome.

Statistical analyses and Clustering – Natural environment samples

I performed a time series gene expression analysis using a dedicated R package: `maSigPro` (Conesa et al., 2006; Nueda et al., 2014). Its approach implements a two-step regression strategy to find differentially expressed genes in time course experiments. It finds genes with significant temporal expression changes and significant differences between experimental groups. Then it performs clustering to group genes showing similar expression profiles. Results are clusters of genes, which have similar expression pattern in time and differ significantly between the experimental groups (in our case: Invasive and *C. rhenanus*) in at least one time point. I conducted this analysis with a significance FDR-corrected p-value ≤ 0.05 and stringency $R^2 \leq 0.6$. I performed hierarchical clustering and visually checked cluster numbers ranging from 1 to 20 in order to choose the most informative number of clusters. I performed Gene Ontology (GO) term enrichment test for each cluster by Fisher's Exact Test in `Blast2GO` software (Conesa et al., 2005) with significant FDR-corrected p-value ≤ 0.05 .

Statistical analyses – Controlled environment samples

Mapped RNASeq read counts were imported into the R environment and normalized by trimmed mean of M-values method implemented by a `tmm` function in the `Noiseeq` package. I arranged the total normalized read counts into groups corresponding to gene clusters identified time series analysis. To get an overview of heritable and plastic component in each cluster I performed a Principal Component Analysis (PCA) for each group separately. The PCA was done by `rda` function in the `vegan` package . To find statistically significant differences between lineages and conditions ("summer" and "winter") I performed a differential expression analysis with `DESeq2` package were (Love et al., 2014) with significance level set at FDR-corrected p-value ≤ 0.01 .

Results

Sample collection – Time series from the natural environment

We collected fish from their natural environment nine times in the course of the year. For the purpose of transcriptomic study I used sample sizes varying from 3 to 8 individuals per sampling point per lineage (details in Fig. 2).



Date of sampling		06-04-16	03-07-16	23-07-16	23-08-16	13-09-16	02-10-16	26-10-16	12-12-16	28-01-17	
LIVER	RHE	8	6	6	5	4	5	5	6	7	Σ 52
	INV	3	6	7	6	5	6	6	3	6	Σ 48
FIN	RHE	6	7	6	6	4	6	5	6	5	Σ 51
	INV	5	4	7	6	6	6	6	6	6	Σ 52

Figure 2. Summary of mean water temperatures, dates of sampling and numbers of individuals analysed. Graph shows average day temperature on the day of sampling in two stations: Bröl – *C. rhenanus* habitat – and Eitorf (Sieg) – invasives' habitat. Specific dates of the 9 sampling points are given on the X-axis. The table summarizes numbers of individuals, per species, per tissue included in gene expression analyses. Sum of individuals analyzed in the time-series is given in the last column.

Temperature measurements

Temperature measurements from the field stations revealed a stable temperature pattern, showing that the habitat of hybrid *Cottus* is on average around 4°C warmer in summer and slightly colder in winter compared to the *C. rhenanus* habitat (Fig. 3). The exact difference in winter temperatures is, however, impossible to state, due to missing data. We were also able to capture this difference between environments in summer in our sampling from the wild (Fig. 2). Data from 6 past years shows as well that there is very little difference between habitats in spring and autumn.

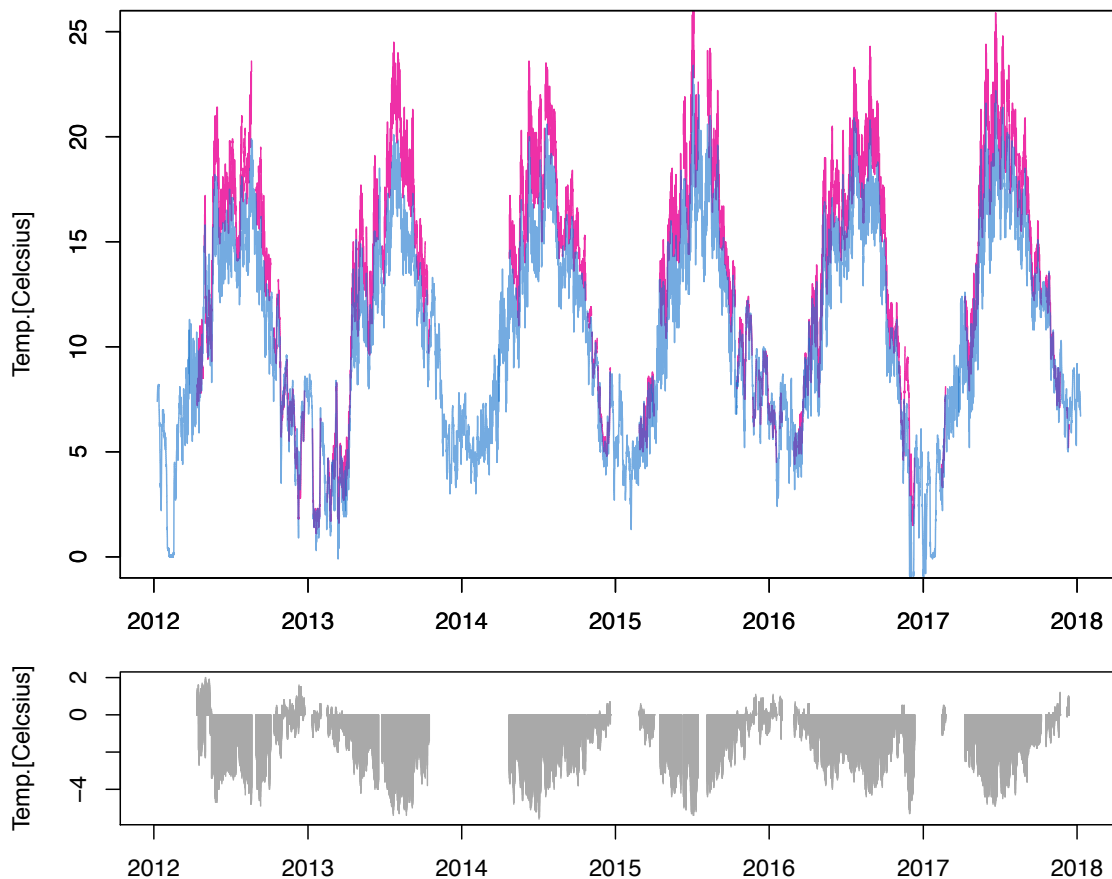


Figure 3. Water temperature in the natural habitat. Panel above shows a graph of hourly measures of water temperature over the past 6 years in two sampling stations: in magenta the Bröl (*C. rhenanus*' habitat) and in blue the Sieg (invasives' habitat). Panel below presents the difference in water temperature between the Bröl station and Sieg station. Periods lacking data are reported as missing values and are due to measuring equipment failure.

Growth assessed in the natural habitat

A proxy for fish growth – a standard length measurement – was generated for all the samples collected in the natural habitat (Fig 4).

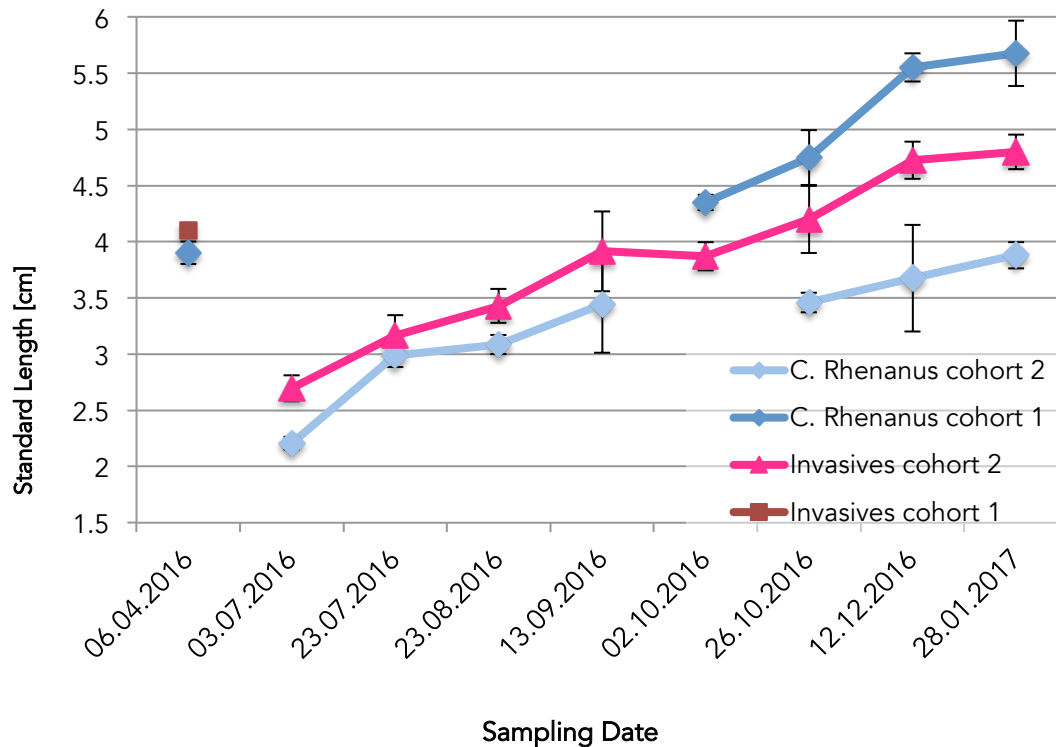


Figure 4. Standard length of fish sampled in the natural habitat. Lines connect points of medium standard length of fish samples at a given sampling date calculated per species and per cohort. Error bars show variance. The first sampling point (April 2016) represents adults from the previous breeding season. In the second sampling point (July 2016) we could sample only new fish and we follow their cohort till January 2017.

RNASeq and mapping

RNASeq performed on the natural environment samples yielded on average 25.9 million high quality paired and mapped reads per individual for fins and 25.0 million for livers. This summed up to 2.7 billion reads for fins and 2.6 billion for livers. In the controlled environment samples RNASeq gave on average 29.2 million properly paired and mapped reads in fin tissue and 34.5 million in liver. It amounted to 1.4 billion reads for fins and 1.6 billion for livers.

Liver samples - Statistical analyses and Clustering – Natural environment

Clustering analysis resulted in 812 genes grouped into eight clusters, which showed divergence in gene expression between lineages and followed similar expression profiles in the course of a year (Table 1). Among them, five clusters could be described as “winter clusters”, meaning that genes grouped in it were expressed at low levels during spring and summer and were “turned on” only in late autumn or winter (clusters 1, 2, 5, 6 and 7). Two clusters were “summer clusters”, because expression levels were higher in summer time and drop steeply in autumn to be “turned off” in winter (clusters 4 and 8). Finally one cluster (number 3) presented almost a constant expression pattern throughout the year. In this cluster invasives’ expression was consistently higher than *C. rhenanus*’.

Within the “winter clusters” I distinguish two types: 1) those where gene expression started to rise in October and peaked in January - clusters 1 and 2 belonged here. They show little, but still significant difference between the lineages; variance is high; 2) where genes increase expression gradually in December and become strongly up-regulated in January. Those were clusters 5, 6, 7 and invasives always showed higher expression than *C. rhenanus*.

Within the “summer clusters” gene expression patterns follow changes in water temperature – expression increases from April, peaks in late July and then lowers steadily until winter. In cluster 4 the invasives have higher expression from July to early October when river warms up and slightly lower expression in December when river is colder than the streams of *C. rhenanus*. Cluster 8 also roughly follows temperature changes, however the differences between lineages are less pronounced and appear to change during seasons. In the late July *C. rhenanus* seems to have a bit higher gene expression levels than invasives but in early July and in August the situation is reversed.

Liver samples - Statistical analyses – Controlled environment

The principal component analyses based on read counts are presented in Table 1. Testing for significant differential expression in the consistent environment showed overall that clusters classified as “winter” had low heritability and “summer” clusters higher.

Among the “winter-clusters” we observed strong differential expression between seasons, however, the number of genes differentially expressed between the lineages was modest. In the natural environment those genes were equally expressed in *C. rhenanus* and invasives in spring and summer but had significantly different expression levels in winter. We did not observe this pattern in the consistent environment.

Specifically, in cluster 1 the PCA ordination showed that the biggest difference was between the conditions: “summer” and “winter”. This was reflected in differential expression analysis where the majority of genes (77% for *C. rhenanus* and 70% for invasives) were expressed significantly higher in “winter”. The rest of the genes were not significantly differentially expressed, however, the trend was the same for all genes. The second PC explains 9% of variance and does not divide samples based on any factor known to us. Only a few genes were identified as differentially expressed between *Cottus* lineages – 4 in “summer” and 5 in “winter” - suggesting that plasticity is the dominating mechanism affecting this cluster. Genes from Cluster 2 present generally the same pattern as seen in Cluster 1. The first PC explains 62% of variance and divides samples according to the “summer” vs. “winter”. The majority of genes are expressed significantly higher in “winter” than in “summer” (70% for *C. rhenanus* and 66% for invasives). There is no clear difference in overall expression patterns between lineages. Only very few genes were differentially expressed between the lineages (3 in “summer” and 5 in “winter”) again suggesting plasticity as the dominating mechanism. When inspecting

expression in the consistent environment in “winter-clusters” 5, 6 and 7 a similar pattern emerges. In those three clusters, however, the most important differentiation was sex-specific gene expression within “winter” samples. Females from both *C. rhenanus* and the invasive *Cottus* are characterized by specific expression patterns differentiating them from “winter” males and all “summer” samples. In cluster 5, consisting of 72 genes, half differed between “summer” and “winter” (35 genes for *C. rhenanus* and 36 for invasives). There was virtually no difference between lineages (only 1 gene in “summer” and 2 in “winter”). In Cluster 6 the situation was very similar. Out of 86 genes, 37 in *C. rhenanus* and 23 in invasives differed between conditions and only few differed between the lineages (2 in “summer”) and (7 in “winter”). Cluster 7 presented most differences between “summer” and “winter”. Out of 124 making up the cluster, 92 in case of *C. rhenanus* and 76 invasives’ were expressed higher in “winter”. There was very little difference between lineages - 1 gene different in “summer” and 4 in “winter”. For those three clusters sex-specific expression is the ruling factor.

The “summer-clusters” – 4 and 8 – showed two very different patterns. Genes in both were following the general trend observed in the natural environment: the overall expression was higher in “summer”. Additionally there were many genes that were differentially expressed between *C. rhenanus* and the invasives – suggesting heritable genetically based phenotypes in cluster 4 but not in cluster 8. In more detail, genes from cluster 8 presented important difference between conditions along the first PC (65%). The second PC (10%) seems to separate males and females in “winter” condition. There were zero differentially expressed genes between lineages in “summer” and 7 in “winter”. Therefore heritability in this cluster was very low but it seemed to have strong sex-specific expression. Cluster 4, on the other hand, showed a strong heritable component. The first PC axis divided samples

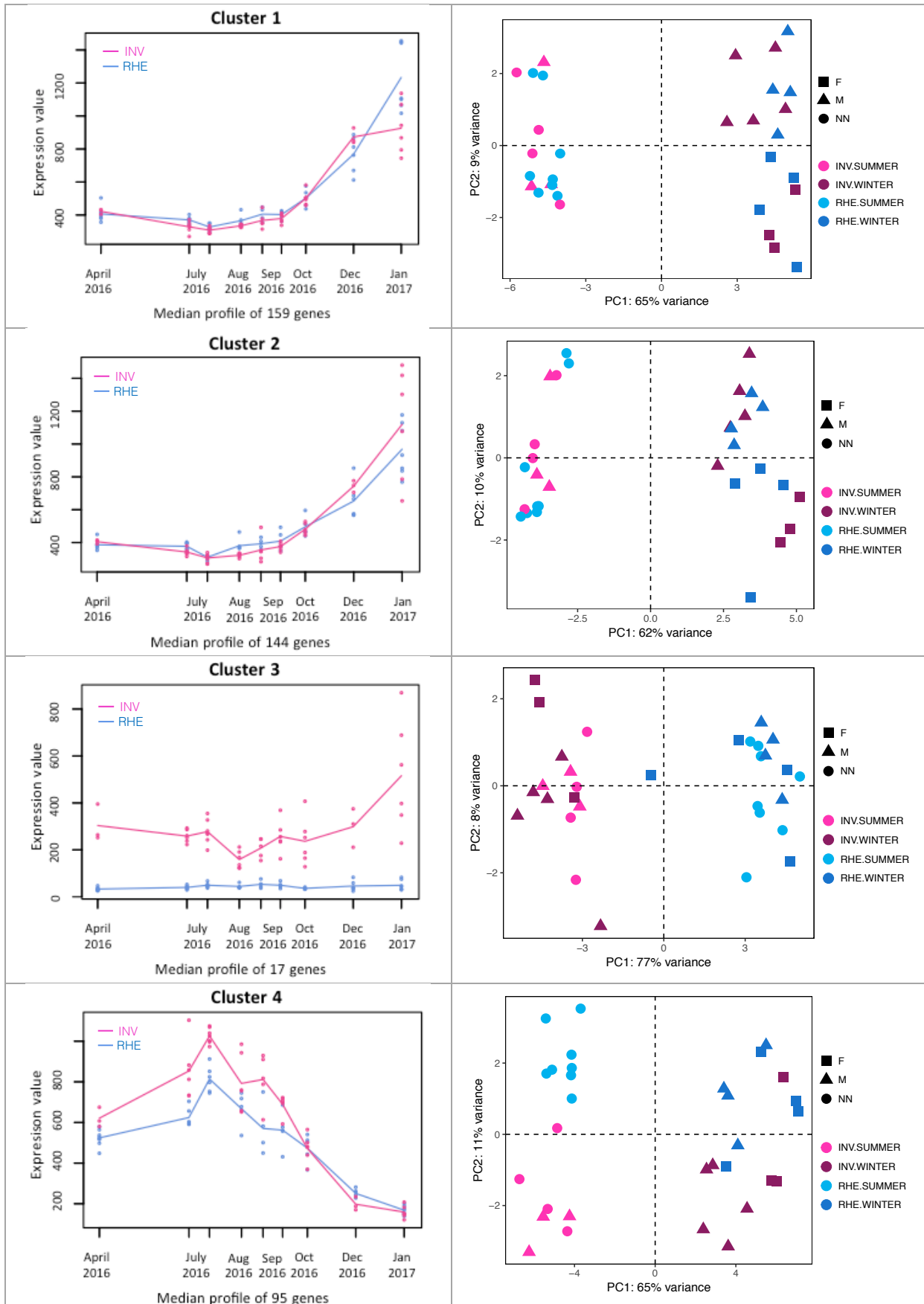
according to “summer” and “winter” differentiation. The second PC (11%), however, divided them according to the genotype in “summer” condition. Differences involved 6% of genes in “summer” and 4% in “winter”. Those percentages seem low, considering that the second PC explained 11% of total variance and was assigned to differences between lineages. It might indicate that the significance threshold chosen for the analysis (FDR-corrected $p < 0.01$) might be too strict to see meaningful results in this particular case. General patterns of expression are in accord with patterns observed in the natural habitat and suggests a genetic, heritable basis.

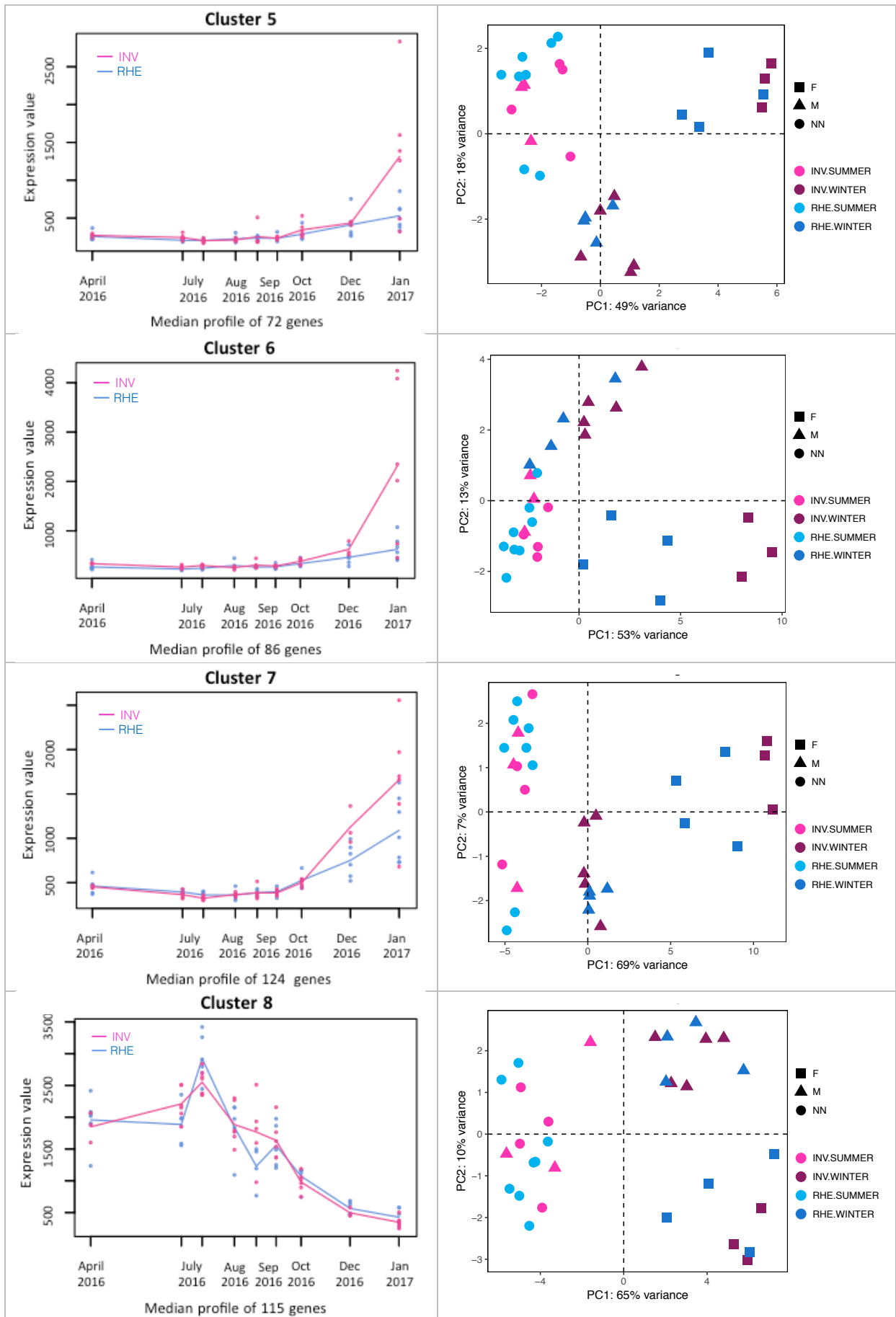
In the case of the unique cluster 3, the PCA exposed very strong grouping based on the lineage, independent of the condition. 77% of variance explained by the first PC separates *C. rhenanus* and invasive *Cottus*. The Second PC does not explain differences between conditions. Overall genes from Cluster 3 also show expression patterns in the consistent environment similar to the ones observed in nature. Tests for differential expression revealed only few genes that differed across conditions (3 for *C. rhenanus* and 0 for invasives). Almost all genes did not vary across conditions and were differentially expressed between the lineages (16 in “winter” and 15 in “summer”). The expression patterns of genes grouped in this cluster seem to be very strongly determined by genetic build-up.

Gene ontology enrichment analysis – Liver samples

Only one cluster – number 8 – showed significant GO term enrichment. Most prevalent enriched terms were connected to energy and metabolism: Oxidation-reduction ($p=0.0002$), oxidoreductase activity ($p=0.0012$), glycolytic processes ($p=0.0015$), mitochondrion ($p=0.03$). Additionally a glutathione transferase activity ($p=0.0017$), racemase and epimerase activity ($p=0.0017$), hydratase (0.0017) and hydrolase activity (0.0027). All p -values after FDR correction. Details available in Fig. S1 in the supplement.

Table 1. Eight gene clusters identified in the analysis of LIVER samples. First column shows expression in the natural environment. Second column is a PCA of the same group of genes but in the controlled environment.





Fin samples - Statistical analyses and Clustering – Natural environment

Clustering analysis of fin samples resulted in eight clusters that comprised 1342 genes and contigs (Table 2). Similar to liver samples we could distinguish between “summer-clusters” (clusters 2, 4, 6 and 7), “winter-clusters” (1 and 5) and the constant “all-year-round clusters” (3 and 8).

Among the “winter-clusters”, cluster 1 had very similar levels of expression for both lineages, except for December when invasives’ was higher. Cluster 5 had similar shape of gene expression pattern for both lineages but *C. rhenanus*’ levels were constantly higher than invasives’. Within the “summer-clusters” clusters 2, 4 and 7 had higher expression levels in invasives from July to the beginning of October. In cluster 6 it was the *C. rhenanus* that had higher expression and it persisted all year. Cluster 3 consisted of genes that were expressed higher in the invasives all year round. In cluster 8, on the other hand, all genes stayed expressed higher in *C. rhenanus* the whole year.

Fin samples - Statistical analyses – Controlled environment

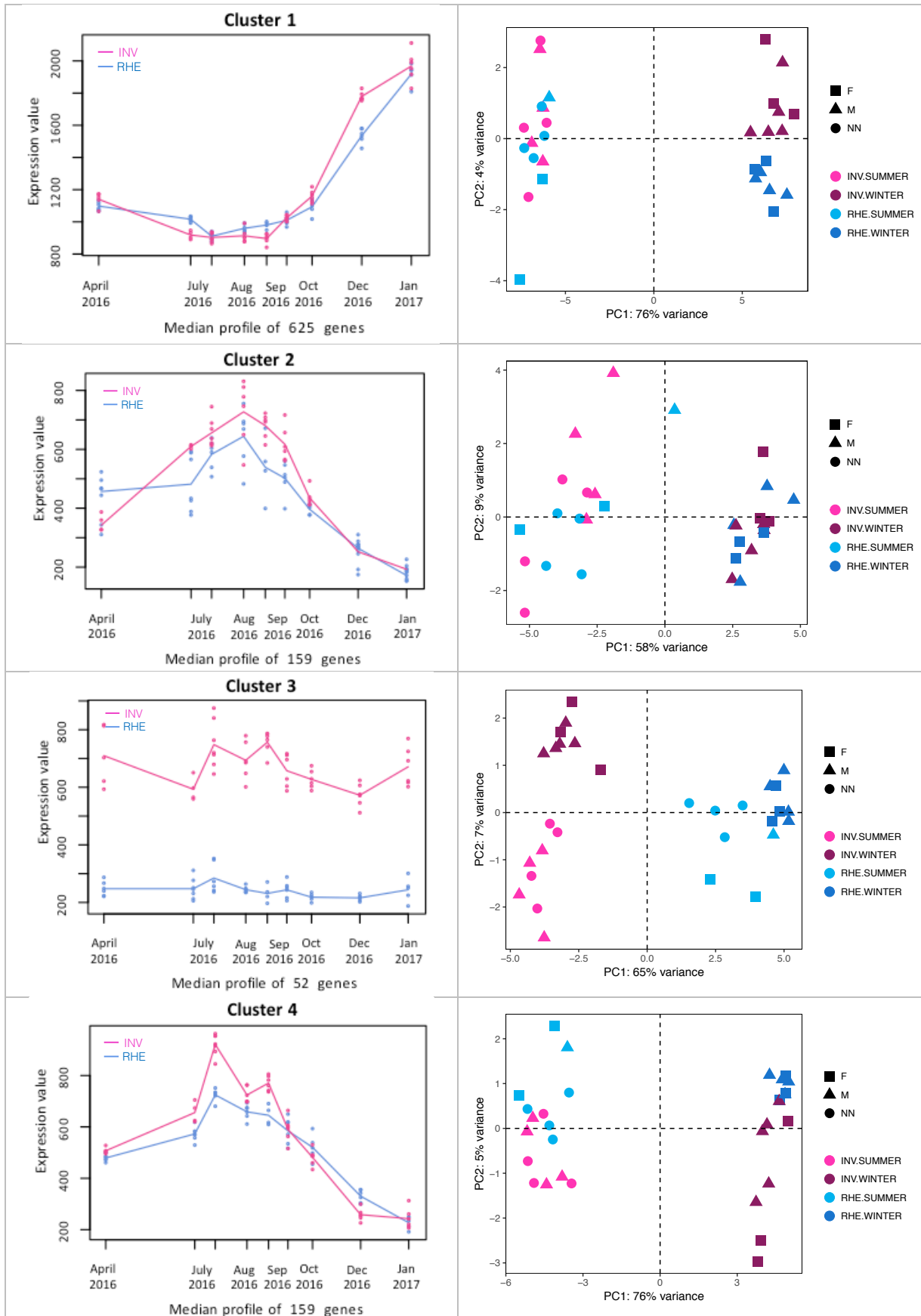
Graphs resulting from the PCA are presented in Table 2. Testing for significant differential expression in the consistent environment showed overall substantial heritability for genes in “winter-clusters”. “Summer-clusters” seemed to be mostly determined by plasticity, showing little heritable component. Genes from “all-year-round clusters” were strongly determined by genetics.

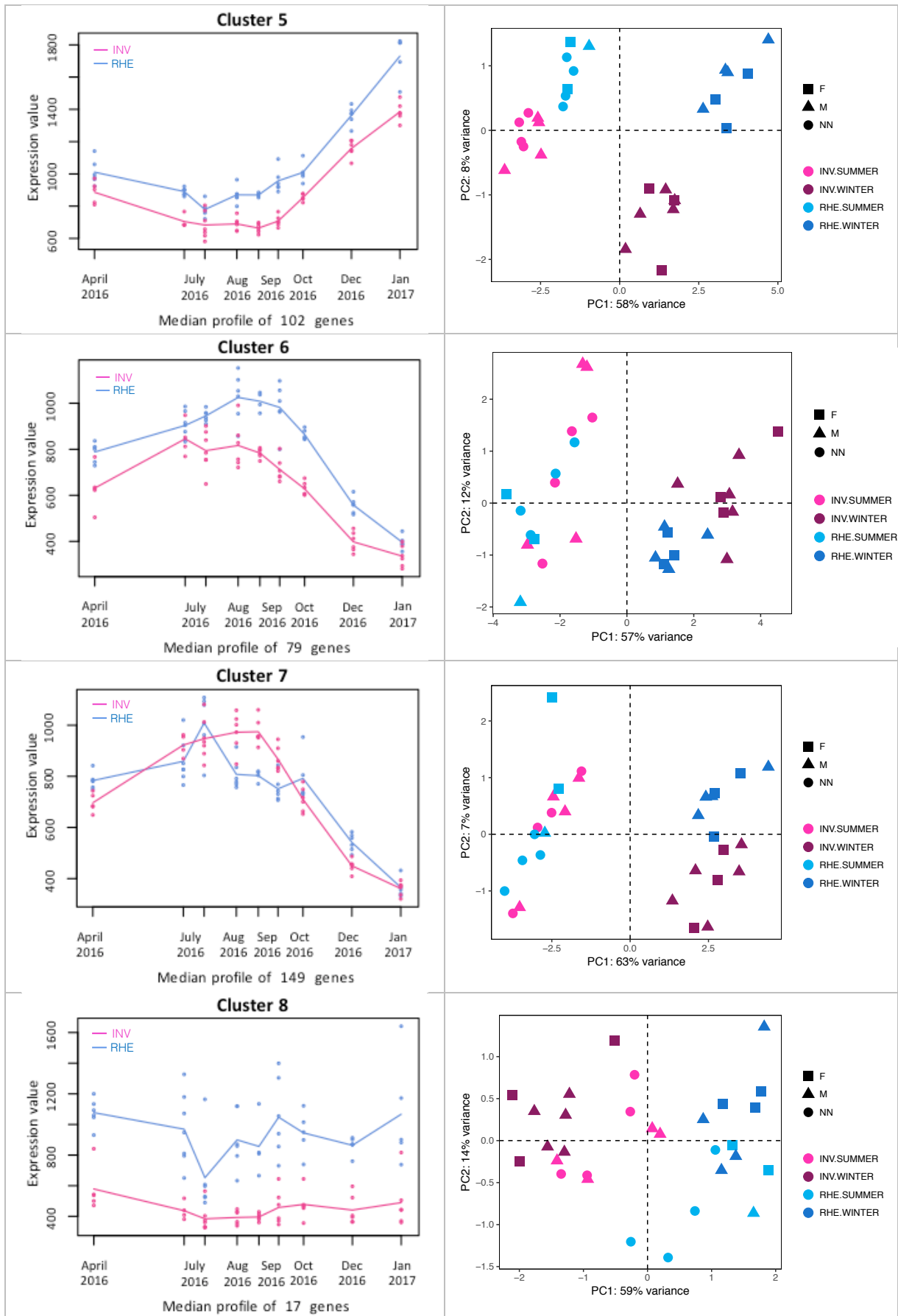
In more detail, among the “winter-clusters”, in the cluster 1 the first PC divided strongly “summer” and “winter” samples and the second PC indicated slight difference between lineages in “winter”. 82% of genes were expressed significantly higher in “winter” than in “summer” for both invasives and *C. rhenanus*. One gene was differentially expressed but in the opposite direction in both lineages. 17 (3%) genes were differentially expressed between lineages

in "summer" and 25 (4%) in "winter" exposing a heritable component of gene expression. Cluster 5 showed differentiation between conditions and between lineages, especially in "winter". All genes were expressed higher in "winter" in both lineages. Differential expression between lineages was detected for 10 genes (10%) in "summer" and 19 (19%) in "winter", hinting at fixed genetic effects.

Among "summer-clusters" there was very little heritable component detected. Genes from cluster 2 grouped on the PCA according to the condition. 62% of genes from invasives and 50% from *C. rhenanus* were differentially expressed between the conditions and always had higher expression in "summer" – a pattern similar to the general one observed in nature. There were no clear heritable differences between lineages. Only one gene (0.6%) was differentially expressed between lineages in "summer". Similarly in the PCA on genes from cluster 6 there was differentiation according to the condition along first axis. All genes were expressed higher in "summer" - 62% in *C. rhenanus* and 71% in invasives genes classified as differentially expressed. Differences between lineages on the other hand were modest: 3 genes (4%) in "summer" and 10 genes (13%) in "winter". Cluster 7 showed again strong difference based on the condition and some structuring between lineages in "winter". All of the genes showed higher expression in summer and 70% were statistically significant for *C. rhenanus* and 60% for invasives. In "winter" there were 5 differentially expressed genes (3%) between lineages and none in "summer". In cluster 4 the pattern was similar. Genes showed strong differentiation between conditions along the first PC and some very slight difference between lineages in "winter". Differences between lineages were: 25 genes (16%) in "summer" and 35 (22%) in "winter".

Table 2. Eight gene clusters identified in the analysis of FIN samples. First column shows expression in the natural environment. Second column is a PCA of the same group of genes but in the controlled environment.





The PCA of the “all-year-round” cluster 3 showed clustering according to lineage and, in case of invasives, some according to the condition along the second PC. Differences in gene expression patterns appeared heritable, as they depended strongly on the genotype (lineage) and persisted in the controlled environment. For cluster 8 the PCA also revealed differentiation between lineages but the pattern was less clear than in cluster 3. Almost all genes (15 out of 17) differed between lineages in “winter” (expression always higher in *C. rhenanus*). In “summer” there was more within-lineage variation and only 2 differed significantly between lineages. For *C. rhenanus* there was no differentially expressed genes between conditions. In case of invasives there were 2 differentially expressed genes between “summer” and “winter”.

Gene ontology enrichment analysis

GO term enrichment analysis returned significant results for two clusters among fin samples: clusters 1 and 2. Cluster 1 showed over-represented categories: ATP binding (p-value=0.0465), protein serine/threonine kinase activity (p-value=0.0324), protein modification by small protein conjugation or removal (p-value=0.0324), negative regulation of DNA binding transcription factor activity (p-value=0.0324). Cluster 2: homophilic cell adhesion via plasma membrane adhesion molecules (p-value=0.0409), cell-cell adhesion via plasma-membrane adhesion molecules (p-value=0.0158). All reported p-values were subject to FDR-correction. Details available in figures S2 and S3 in the supplement.

Discussion

In this study we focused on an incipient homoploid hybrid speciation event in sculpin fish (*Cottus sp.*). I did an extensive comparative transcriptomic study that was divided into two parts. First, I describe in an unprecedented detail variation in the transcriptomic phenotype in *C. rhenanus* and the *Cottus* lineage of hybrid origin in their natural habitats. Second, I compare those patterns with the transcriptomes of fish living under consistent environment in the laboratory. All analyses were done in two tissues: liver and fin. It was complemented by a collection of fish length measurement data from the field.

Ecological transcriptomics studies are potentially a great tool to connect phenotypic change with morphological function and underlying genetic architecture, but most of them represent only a 'snapshot' of the real situation, they were just sampled once (Alvarez et al., 2015). My results demonstrate how complex and ever-changing the gene expression landscape of an organism is in nature. By analyzing RNASeq time series I was able to observe relevant differences between populations or species without being pre-conditioned and biased by our starting hypothesis. Based on our understanding of the *Cottus* system one could have expected the biggest differences between lineages to appear in summer, when differences in water temperature between habitats are the largest. Surprisingly, however, most of the differences in gene expression within our defined clusters in liver tissue are seen in winter – December and January – and in fin tissue, differences were equally distributed across the year. I identified diverse gene expression profiles, some were peaking in winter, and some in summer, while others had equal expression level throughout sampling. This demonstrates how easy it is to bias ones results by selecting a preconceived time point for sampling. Careful and informed design is a crucial part of any experiment. Even more so in case of studies done in the wild populations of non-model species for which genomic

recourses are scarce. Nevertheless, even with the best of intentions and the maximum amount of work, we might find ourselves with biasing our own results by selective choice of sampling times and this way we miss the real important observation. Of course, it is not possible to study everything all the time. However, we should always keep in mind what we demonstrated here: biological responses are complex and very variable. One should not limit him/herself to the working hypothesis and stay open-minded.

After describing the patterns of gene expression observed in the natural habitat we were interested in comparing those with gene expression under constant environment.

In terms of results for liver samples, I identified eight clusters in total. Five clusters were “winter clusters” (expression was high in winter and very low in the rest of the year), two “summer clusters” and one where gene expression levels were similar across the year. When comparing patterns seen in the laboratory to those in nature I noted that most genes from each cluster were following the general same pattern of expression. However, in most of the clusters I could not observe fixed genetic differences between lineages (clusters 1, 2, 5, 6, 7, 8), either due to plasticity or because of the sampling time – in every cluster there are times when genes are expressed equally in both lineages. Interestingly, I saw also sex-specific gene expression in three of “winter clusters” (#5,6,7) and in one “summer cluster” (#8), where it was explaining all or part of the variation observed in nature. Thanks to the laboratory-based experiment, I was able to explain variation in natural habitat in those clusters of genes. Liver is well known as the organ where sex-specific gene expression takes place (Sharma et al., 2014). However, as the major powerhouse of the organism, it is also often analyzed in the context of adaptation to temperature. My results show how influential sex is on gene

expression in this organ. It should be always carefully taken into account when interpreting ecological transcriptomic data.

Two most interesting clusters in this analysis were clusters 3 and 4. Cluster number 3 consisted of only 17 features that appeared to be specific to the invasives - they were constantly expressed high in invasive, while very low in *C. rhenanus*. Among those features as many as 10 are unannotated contigs. They might be playing an important role in invasives' adaptation to the habitat, but we don't know what do they code for at this moment. The other 7 genes were representing biological processes such as signal transduction, lipid metabolic processes, DNA recombination, carbohydrate derivative metabolic processes. It is difficult to draw far-reaching conclusions based on only 7 genes, but even this small set suggests increased metabolic rate and faster growth (mitotic divisions).

Cluster 4 was characterized as a "summer cluster" in which gene expression seemed to follow exactly the rise and fall in temperature during the year. In the constant environment, however, it turned out that differences between invasives and parental species were not entirely plastic, but could be assigned, in large part, to fixed genetic differences. I therefore found it interesting to look at functions of genes in this cluster. Lo and behold, striking 70% of genes were connected to metabolic processes. Among those, 8 genes (accounting for the 16% of all genes in the cluster) were connected to oxidation-reduction processes. Other genes represented signaling processes and trans-membrane transport, however metabolic processes GO terms were undoubtedly the most significant result in this group.

Additionally I found gene ontology enrichment clearly pointing into the direction of metabolism and mitochondrion-related phenotype in cluster 8. These enriched terms make sense considering that genes in cluster 8 picked in summer. There was no direct response to the temperature though. Differences

in expression between lineages were small and varied substantially. Surprisingly, in the high of summer heat (end of July) *C. rhenanus* had higher expression levels. As mentioned before I did not detect any genetically fixed differences between lineages in the common-garden experiment, however, it might suggest a regulatory nature of differences between lineages. If genes in cluster 8 and 4 are connected functionally in a complex metabolic network, it is very well possible that while some are up-regulated by fixed genetic mechanisms, others will fluctuate accordingly. Maybe genes in this cluster are the “followers” of genes from the cluster 4. This salient result encourages me to conclude that what makes invasives different from their parents in their natural habitat is connected to the regulation of metabolic processes and has a strong genetic component.

Supplement

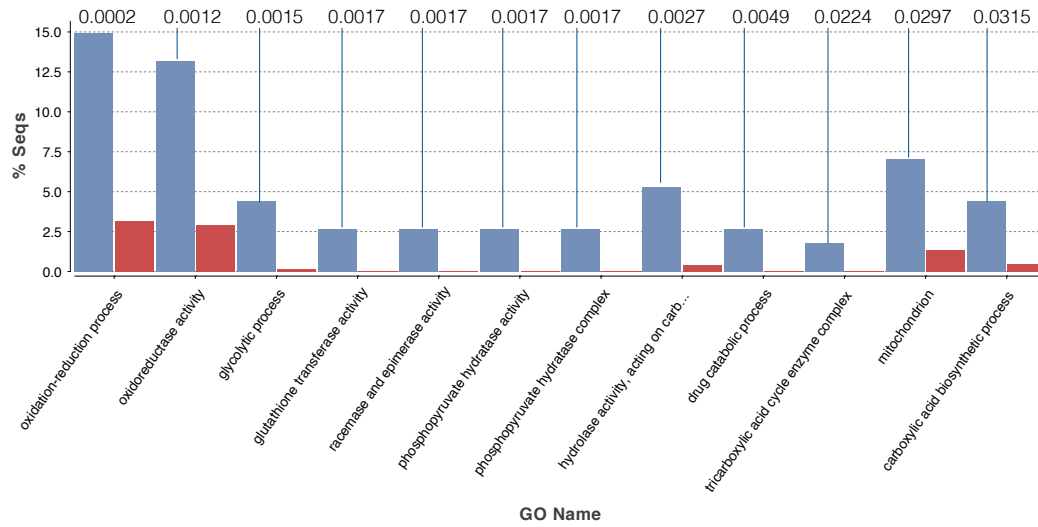


Figure S1. GO term enrichment in the gene cluster 8 in liver samples. Bar graph presenting statistically significant GO term enrichment. GO term names are given on the x-axis. Blue bars – test set; Red bars – reference set.

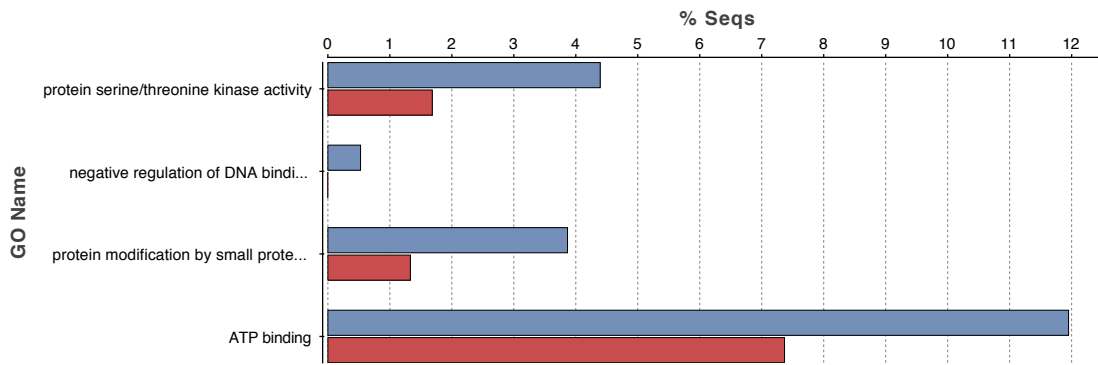


Figure S2. GO term enrichment in the gene cluster 1 in fin samples. Bar graph presenting statistically significant GO term enrichment. GO term names are given on the y-axis. Blue bars – test set; Red bars – reference set.

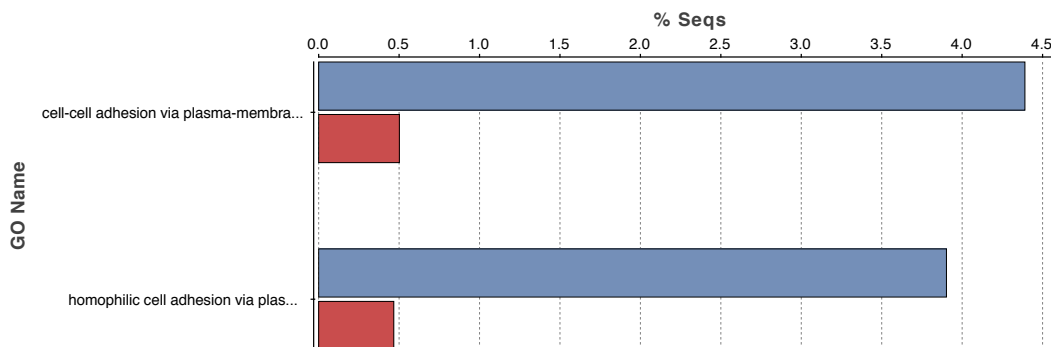


Figure S3. GO term enrichment in the gene cluster 2 in fin samples. Bar graph presenting statistically significant GO term enrichment. GO term names are given on the y-axis.

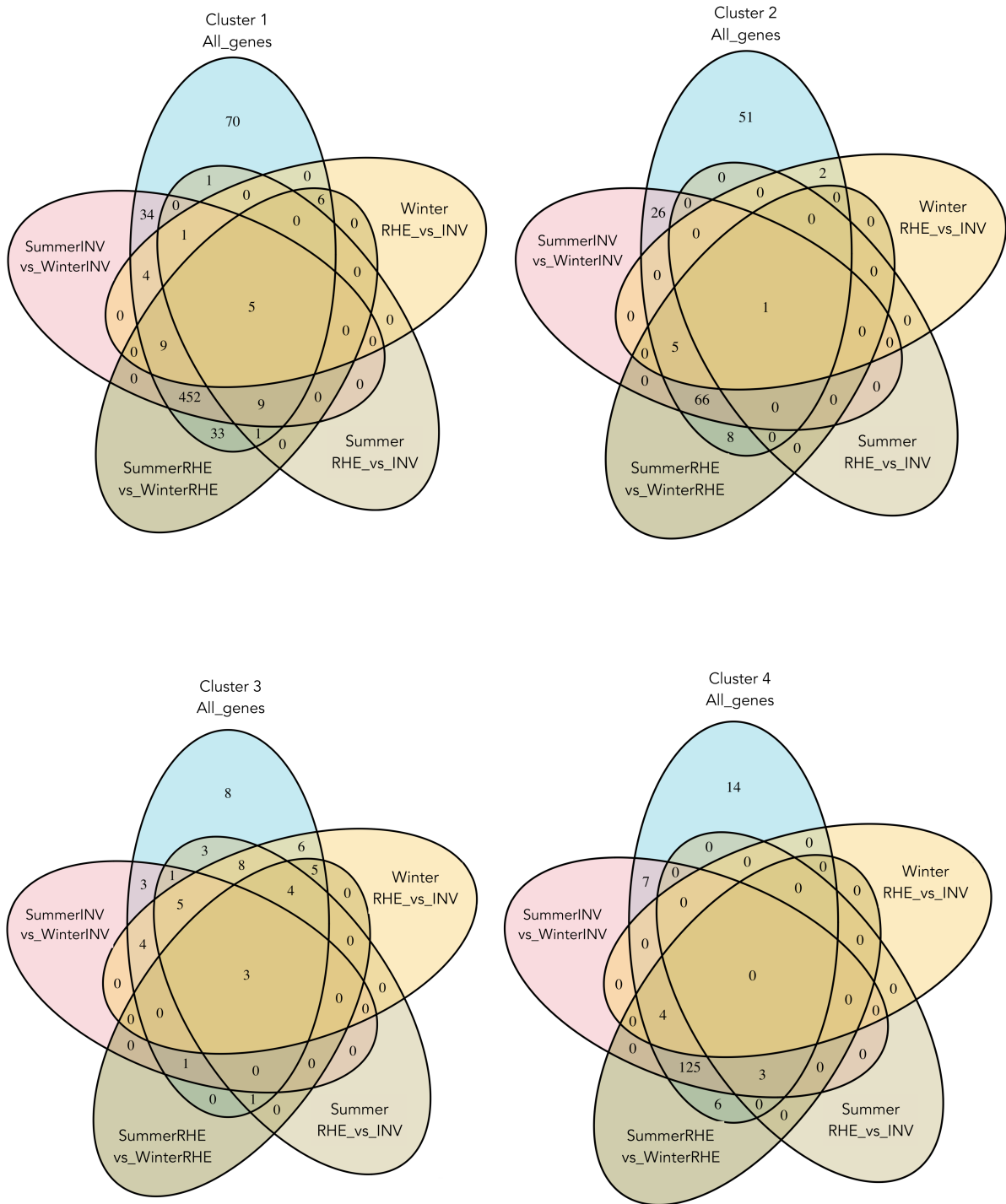


Fig. S4. Venn diagrams presenting numbers of differentially expressed genes for each gene cluster identified in the natural experiment in fin tissue. Every Venn diagram shows number of genes in a given cluster and four contrasts tested for differential expression: 1) Invasives in "summer" vs. Invasives in "winter"; 2) *C. rhenanus* in "summer" vs. *C. rhenanus* in "winter"; 3) *C. rhenanus* in "summer" vs. Invasives in "summer"; 4) *C. rhenanus* in "summer" vs. Invasives in "winter"

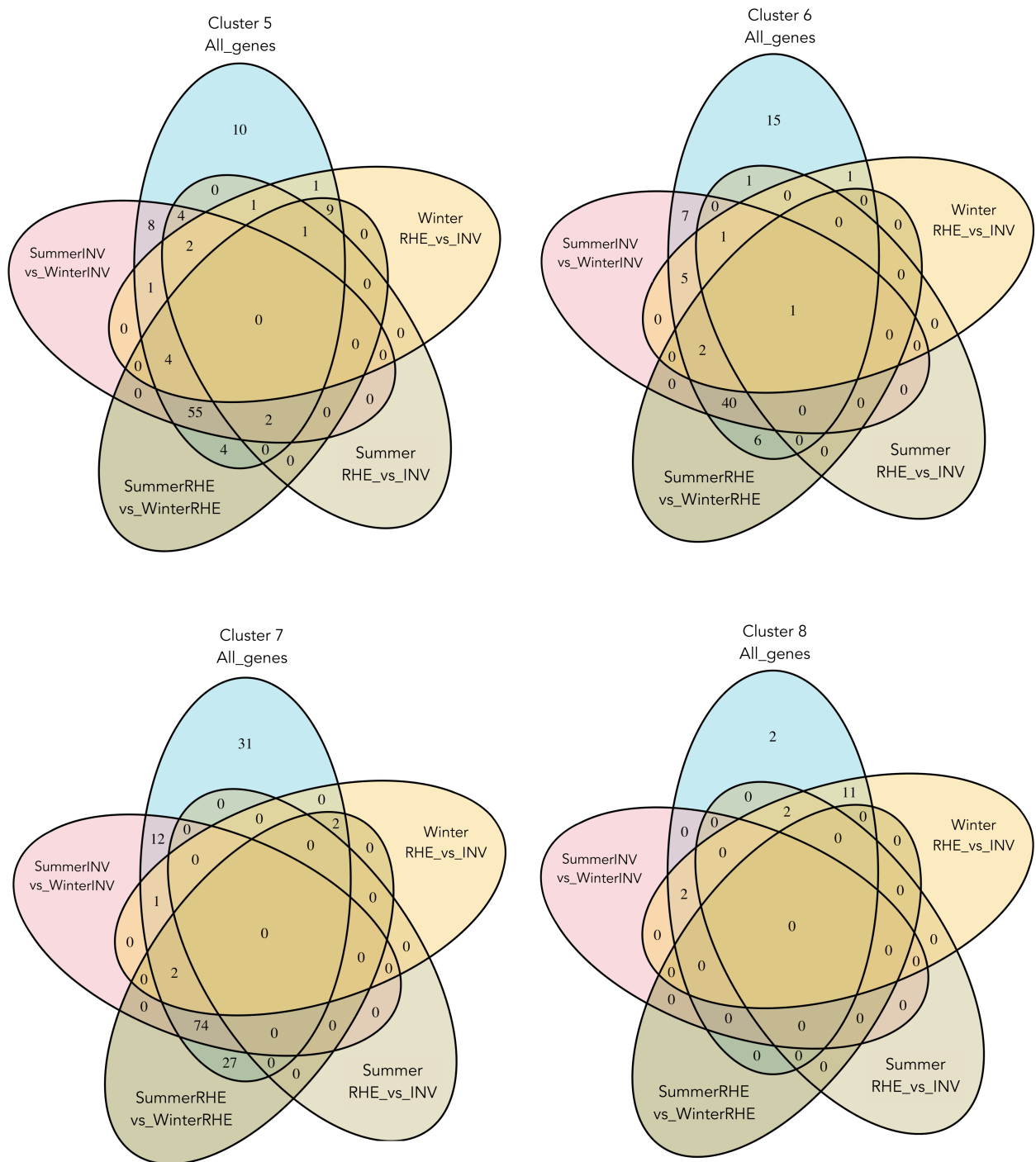


Fig. S5. Venn diagrams presenting numbers of differentially expressed genes for each gene cluster identified in the natural experiment in fin tissue. Every Venn diagram shows number of genes in a given cluster and four contrasts tested for differential expression: 1) Invasives in “summer” vs. Invasives in “winter”; 2) *C. rhenanus* in “summer” vs. *C. rhenanus* in “winter”; 3) *C. rhenanus* in “summer” vs. Invasives in “summer”; 4) *C. rhenanus* in “summer” vs. Invasives in “winter”

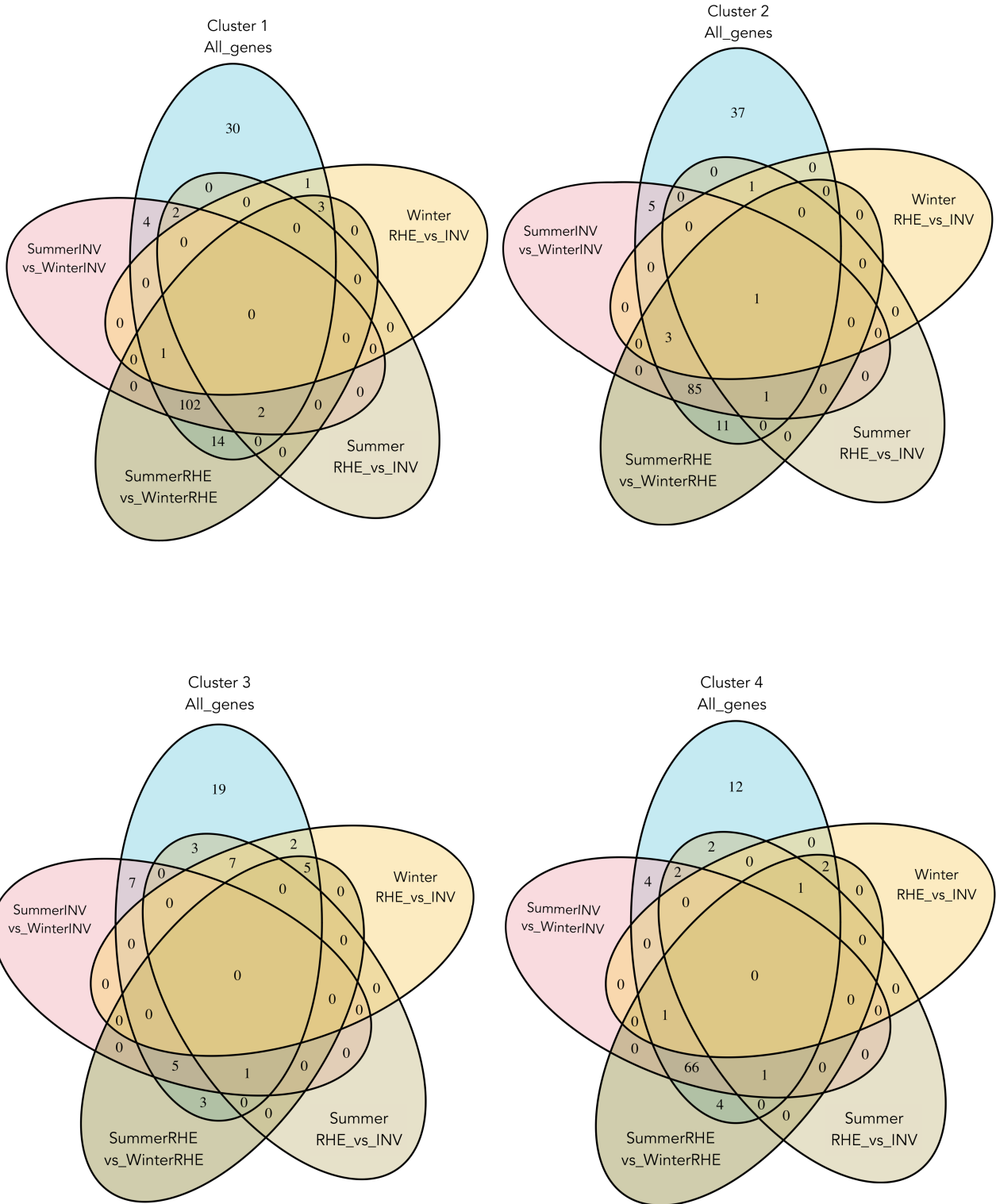


Fig. S6. Venn diagrams presenting numbers of differentially expressed genes for each gene cluster identified in the natural experiment in liver tissue. Every Venn diagram shows number of genes in a given cluster and four contrasts tested for differential expression: 1) Invasives in "summer" vs. Invasives in "winter"; 2) *C. rhenanus* in "summer" vs. *C. rhenanus* in "winter"; 3) *C. rhenanus* in "summer" vs. Invasives in "summer"; 4) *C. rhenanus* in "summer" vs. Invasives in "winter"

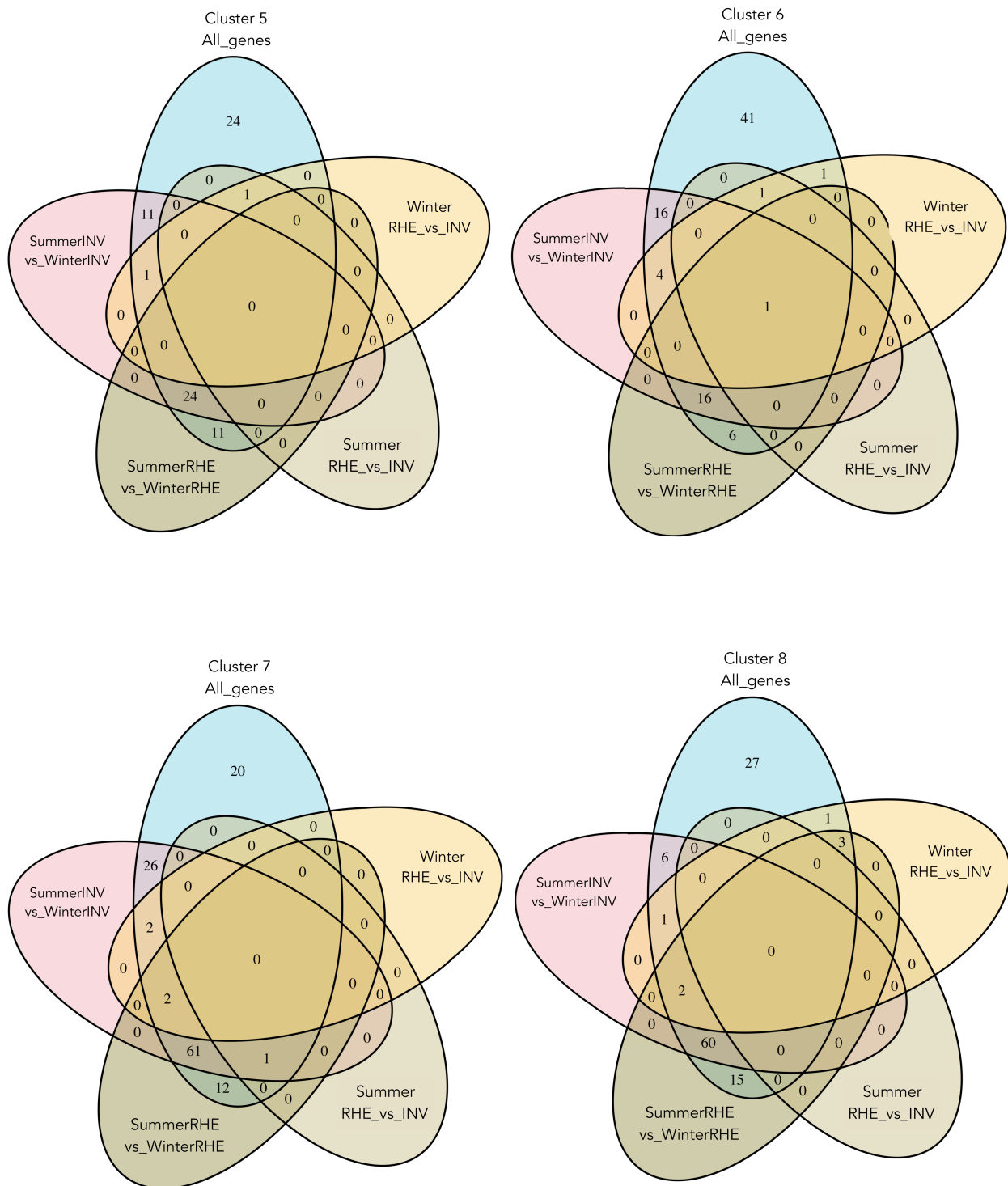


Fig. S7. Venn diagrams presenting numbers of differentially expressed genes for each gene cluster identified in the natural experiment in liver tissue. Every Venn diagram shows number of genes in a given cluster and four contrasts tested for differential expression: 1) Invasives in "summer" vs. Invasives in "winter"; 2) *C. rhenanus* in "summer" vs. *C. rhenanus* in "winter"; 3) *C. rhenanus* in "summer" vs. Invasives in "summer"; 4) *C. rhenanus* in "summer" vs. Invasives in "winter"

Chapter 3

Transgressive trait expression in an emerging hybrid species

Introduction

Evidence that hybridization even without a change of ploidy can lead to the formation of new species has accumulated over the past two decades. The phenomenon is named homoploid hybrid speciation (HHS) and a number of possible examples have been described in plant and animal systems (Abbott et al., 2013; Feliner et al., 2017; Mavárez et al., 2006; Payseur and Rieseberg, 2016; Rieseberg et al., 2003). In the case of HHS, the establishment of a new hybrid lineage is usually connected with the expansion to a novel, previously unoccupied habitat. Such ecological transition can be best explained if hybrids develop novel traits that confer a new adaptive potential. James Mallet in his review on hybrid speciation (2007) called the hybrids “hopeful monsters” (a term originally coined by Richard (Goldschmidt, 1940)) referring to the transgressive traits they expressed, which were “strangely” outside of the parental’s phenotypic range but at the same time facilitated adaptations to a novel ecological niche. A poster child for the research in this domain is the wild sunflower of genus *Helianthus*. Gross and Rieseberg (2005) and Rieseberg et al. (2003) demonstrated that the hybrid lineage of sunflowers expressed transgressive phenotypes as a direct result of hybridization that made it more resistant to desiccation and therefore better adapted to the extreme desert habitat. Another illustrative example is given by Parsons et al. (2011) who showed how hybridization in cichlid fish from the Lake Malawi could possibly promote trophic specialization and adaptive radiation by broadening phenotypic diversity of the mandible.

It appears that hybridization can create the genetic and phenotypic diversity that generates such novel traits, however, our knowledge about specific mechanisms causing it is limited. The best described and widely recognized genetic mechanism is transgressive segregation (Rieseberg et al., 1999, 1996), which causes the phenotypes of segregating hybrid offspring to exceed the phenotypic values observed in parental genotypes. It is thought that extreme expression of traits in the hybrids results from the merger of quantitative trait locus (QTL) alleles of opposite effects that originate from alternative parents. In the F_2 and later generations after the initial hybridization, recombination and segregation generates inter-species genotypes that could not have occurred in either of the parental species alone and that enhance traits beyond the phenotypes expressed by the parents. Subsequently selection on such new phenotypes can facilitate colonization of a new environment (Arnold, 1997; Arnold et al., 2012; Bell and Travis, 2005; Dittrich-Reed and Fitzpatrick, 2013; Kagawa and Takimoto, 2018; Stelkens and Seehausen, 2009). In the specific case of transcriptomic studies we can observe transgressive gene expression patterns due to the combination of distant and close regulatory regions from both parental species. Recombination can break up established interactions of close and distant regulatory elements and fix new combinations that will result in increased or decreased patterns of gene expression (Wittkopp, 2013). In yeast the divergence in distant regulatory regions has been observed to happen fast and was associated with a differential response to the environment (Tirosh et al., 2009). It suggests that mixing of rapidly diverging regulatory regions in hybrids may result in widening gene expression phenotypes, which would in turn increase the adaptive potential in a new environment. The other, much less explored, mechanism underlying transgressive traits in hybrid lineage is *de novo* structural mutation. Hybridization can be described as a “genetic shock” that can possibly

accelerate genetic changes such as gene duplications and increased copy number variation (CNV) has been observed in animal hybrids in the laboratory (Scavetta and Tautz, 2010). It has been speculated that they may contribute to the novel phenotypic diversity by gene dosage effect (Machado et al., 2014). (Dennenmoser et al., 2017) however found limited evidence for increase in number of structural variants in the natural fish hybrid lineage. Phenotypic changes resulting from structural variation should be considered as a possible source of transgressive traits but remain largely unexplored to date.

The focus of our study is on a young homoploid hybridization event in the Genus *Cottus*. The two parental species (*C. perifretum* and *C. rhenanus*) hybridized possibly less than 200 years ago and gave rise to a hybrid lineage that invaded new ecological niche previously uninhabited by its' parental species (Nolte et al., 2005a). Due to this invasion we refer to them as the "invasives". Despite the lack of geographic barriers between the parentals' and invasives' habitats the gene flow is limited and fish of hybrid origin remain a separate lineage (Nolte et al., 2006; Stemshorn et al., 2011). This suggests that the young hybrid lineage has an adaptive advantage in the new habitat, which is likely a result of genetic admixture following hybridization. This very young system provides exceptional opportunity to observe and study homoploid hybrid speciation in action, however, details about the nature of the traits that confer an adaptive advantage in alternative habitats are still not known.

Heretofore transgressive traits in invasive *Cottus* have been studied by Czypionka et al. (2012), who concentrated on gene expression changes from the initial hybridization event (F_2 crosses) towards the established hybrid lineage. They have found important differences between transgressive traits expressed in the two. Namely the amount of up- and down-regulated genes identified in the F_2 hybrids were balanced while in the invasive lineage there were four times more up-regulated genes than down-regulated ones. This

imbalance suggests that the expression phenotype of invasives has changed after the first generations of admixture. Czypionka et al. (2012) have also explored gene ontology terms of the invasive lineage-specific transgressive traits. They found that the up-regulated genes were enriched for gene ontology terms related to the mitochondrial phenotype and metabolism pointing to the putative adaptations in the metabolic pathway. Moreover, the most recent study by Dennenmoser et al. (2017) has explored the extent of genetic structural variants (SV) in *Cottus*. They detected no significant increase in the number of SV in invasive *Cottus* compared to its parental species but they identified a set of invasive-specific gene copy number variants.

The study by Czypionka et al. (2012) provided valuable insight into transgressive transcriptomic traits in *Cottus* but was limited in that the analyses of gene expression were conducted exclusively on fish kept in a constant laboratory environment. Controlled conditions are useful in comparative studies, however the connection with the complex ecological context, which is crucial for the formation of hybrid species in nature, is lost (Alvarez et al., 2015). An integration of laboratory studies with field studies can potentially highlight plastic and genetically fixed changes in gene expression in relation to the environmental factors (Czypionka et al. 2018). Another limitation of the study by Czypionka et al. (2012) as well as most other studies on transgressive phenotypes is that the comparisons between parentals and the hybrids are done at only one time point. Many phenotypic traits are highly variable and depend on the context (i.e. environment), which is also reflected in the results presented in chapter 2 of this thesis. For instance gene expression may vary greatly depending on the time of the day, season (chapter 2), developmental stage (Nolte et al., 2009), habitat and temperature (Czypionka et al., 2018) etc.. Hence, measuring a phenotype once in the laboratory setting and drawing conclusions for the whole natural system might be misleading.

In the present study I address those limitations by studying transgressive phenotypes in the controlled environment and by comparing them with data collected in nature across the whole yearly cycle. By means of this comparative approach I am able to cross the boundaries of the controlled laboratory setting and test the relevance of transgressive gene expression patterns in a natural habitat. In the laboratory I mimicked the natural conditions of the parentals' streams in terms of temperature and photoperiod throughout the year. We sampled laboratory-bred fish in summer and in winter to test whether transgressive trait expression changes across the seasons. My sampling from the natural environment concentrates on the natural contact area between *C. rhenanus* and the invasives situated in the river Sieg, which is a perfect study site to verify our lab-based findings in nature. This area has been studied since many years and all evidence points to the fact that despite the lack of geographic barrier and a narrow stretch of overlap between distributions of the two *Cottus* lineages', they stay genetically separated (Nolte et al., 2006; Stemshorn et al., 2011).

Additionally I evaluate whether proportions of up- and down-regulated genes corroborate previous findings by Czypionka et al. (2012). They pointed out that they extracted RNA from the whole animal and therefore their result might be an artifact due to oversized livers. If a conspicuous enrichment for up-regulated genes holds true in our study, I hypothesize that this might be a result of selection. We will therefore predict that the majority of transgressive patterns should be playing a role in adaptation and will persist in the natural environment. Further, profiting from the study by Dennenmoser et al. (2017) on CNV in the invasive *Cottus*, I can distinguish between possible mechanism causing transgressive expression in hybrids.

Czypionka et al. (2012) used the whole fish and an 8K microarray to study gene expression. I extend their experimental design by selecting two

tissues for a more specific comparison. The tissues are: liver – the powerhouse of the organism (Gracey et al., 2001; Orczewska et al., 2010) – and fin – a peripheral tissue that could be re-sampled multiple times and was found to nonetheless represent a broad diversity of transcriptomic patterns in a study on salamanders (Czypionka et al., 2015). In terms of gene expression measurement we use state-of-the-art technology: paired-end RNA sequencing (RNASeq). RNASeq provides us with an instantaneous insight into tens of thousands of individual gene expression profiles painting a complex, multidimensional picture of the phenotype (Todd et al., 2016). The transcriptome is increasingly being studied by evolutionary biologists and ecologists as a transient phenotype that can give us great insight into the ecology and habitat context of the organism's physiological response when studied in the natural setting – the field called “ecological transcriptomics” (Alvarez et al., 2015; Czypionka et al., 2018; Richards et al., 2012; Romero et al., 2012). In high-throughput transcriptomic studies, however, the great challenge is to connect the resulting lists of genes with the higher order phenotype and possible physiological functions. The strategy commonly employed to solve that issue is the analysis of gene ontology terms (GO terms) (Gene Ontology Consortium, 2004). The GO terms offer comprehensive information on attributes of genes based on current published literature and experimental evidence in model organisms, which is extended to less explored, non-model organisms based on sequence similarity (Camon et al., 2003; Pouliot et al., 2001; Xie et al., 2002). The GO terms method is limited by the extent of sequencing and functional studies available to date. It remains, however, a useful tool for the exploration of molecular functions and biological processes in both model and non-model organisms. I use GO term analysis to investigate functions of genes identified as consistently up-regulated in the natural and controlled environment.

Materials and Methods

Sample collection – Controlled environment

Samples were collected as described in the Chapter 2. In short, our sampling was done in two groups: in the controlled environment and natural environment. For the former, we sampled 8 fish from each lineage of *Cottus*: *C. perifretum*, *C. rhenanus* and the invasives. All fish were reared and raised under same controlled laboratory conditions, which coarsely reflected photoperiod and temperature changes in the River Sieg across the year. Sampling was performed two times in the course of a year – once in summer and once in winter – summing up to 48 individuals.

Sample collection – Natural environment

In case of sampling in the natural environment we sampled *C. rhenanus* and the invasive *Cottus* from neighboring habitats – Bröl stream and Sieg river - at nine time points spread throughout the year. In total we sampled 48 invasives and 52 *C. rhenanus* individuals.

Sample treatment

In both groups we anesthetized the fish with MS-222 (tricaine methanesulfonate, 250 mg/L) and killed by a blow on the head. We harvested pectoral fins and livers and stored them in RNAlater in -20°C until RNA extraction. The remaining fish bodies were also stored immersed in RNAlater in -20°C.

RNASeq and mapping

Total RNA was extracted from both tissues using a Trizol protocol (Invitrogen, Paisley, UK) and additionally cleaned with a lithium chloride precipitation. RNA quantity and integrity were assessed by spectrophotometer, gel electrophoresis and using an Agilent RNA 6000 Nano Chip (Agilent

technologies, 2015). Samples that passed our quality assessment were shipped to Cologne Center for Genomics (CCG, Cologne, Germany) where RNA was poly-A enriched, reverse transcribed to cDNA with use of TrueSeq Stranded mRNA Kit and sequenced on Illumina HiSeq 4000 platform. Resulting RNASeq reads were then mapped to a custom-made transcriptome reference (details in Chapter 2) with the `bwa-mem` mapper (Houtgast et al., 2016).

Statistical analyses – Controlled environment samples

After mapping raw RNASeq read counts were imported into the R environment (R Core Team, 2017) and normalized by trimmed mean of the M-values method implemented by a `tmm` function in `Noiseq` package. Normalized reads were basis for a basic explorative analysis - Principal Component Analysis (PCA) - done by the `rda` function in `vegan` package (Dixon, 2009). Because the PCA revealed two outliers and some inconsistencies in the fin samples I decided to look into the genome data in those samples in detail. I recovered whole genome single nucleotide polymorphism data (SNP) from the RNASeq by using `samtools mpileup` function. Later I used a custom `perl` script to build a neighbor-joining phylogenetic tree based on the `pileup` file for all fins samples.

After getting the final dataset we proceeded to the differential expression analyses. For that we used `DESeq2` package (Love et al., 2014). I performed analysis separately for each tissue (fin and liver) and each condition ("summer" and "winter"). Firstly I identified genes that were differentially expressed between invasives and both of their parents (FDR-corrected-p-value < 0.01). Those were classified into two categories: **up-regulated** – where gene expression in invasives was significantly higher than in both parents and **down-regulated** – where gene expression in invasives was significantly lower than in both parents.

Statistical analyses – Natural environment samples

The field data and analysis was already described in Chapter two. Briefly I performed a time series gene expression analysis using `maSigPro`, a `Bioconductor` package dedicated to this type of analyses (Conesa et al., 2006; Nueda et al., 2014). Its approach implements a two-step regression strategy to find genes differentially expressed in time course experiments. It finds genes that display significant temporal expression changes and significant differences between experimental groups (in our case condition is species: Invasive and *C. rhenanus*). I conducted this analysis with a significance FDR-corrected value of $p \leq 0.05$ and stringency $R^2 \leq 0.6$.

Comparing natural and controlled environment samples

For every gene classified as an up and down-regulated in the laboratory I verified manually results obtained in the natural environment. If the data was available I checked the p-values for significant differentiation in summer and/or winter (depending on the result from the laboratory). When differential gene expression was detected both in the controlled and natural environment a gene was classified as “consistently up-regulated genes” and “consistently down-regulated genes”.

Gene Ontology terms analysis

All consistently down- and up-regulated genes identified in this study were assembled into lists, which were further studied with respect to gene ontology terms present. I performed Fischer’s Exact test implemented in `Blast2GO` software (Conesa et al., 2005) to test for enrichment of GO terms. I also looked specifically at all GO terms present in those list of genes to get insight into possible functions of transgressively expressed genes. This was done by means of hierarchical GO terms graphs produced in `Blast2GO` software.

Results

RNASeq and mapping

Sequencing results are same as in chapter 2. In short, RNASeq performed on the natural environment samples yielded on average 25.9 million high quality paired and mapped reads per individual for fins and 25.0 million for livers. This summed up to 2.7 billion reads for fins and 2.6 billion for livers. In the controlled environment samples RNASeq gave on average 29.2 million properly paired and mapped reads in fin tissue and 34.5 million in liver. It amounted to 1.4 billion reads for fins and 1.6 billion for livers.

Statistical analyses – Controlled environment samples

PCA of liver samples showed substantial inter-species variation in gene expression as well as some lineage-specific clustering. The first PC axis divided samples according to lineage and the second divided the “summer” samples from “winter” and the “winter” samples according to sex. Along this second PC females were strongly differentiated from males and males were closer to “summer” samples from their own lineage (Fig. 1).

Fin samples analyzed by PCA revealed two strong outliers – both from *C. perfretum* (Fig. 2A). Such extreme within-samples variation may be a result of technical problems during extraction or library preparation and can skew all further analysis. Therefore I eliminated the two samples and re-run the PCA. In the re-run we observed strong differentiation between “summer” and “winter” samples along the first PC axis (Fig. 2B). The second PC was dividing lineages: *C. rhenanus* and *C. peifretum* most different from each other and the invasives intermediate in both “summer” and “winter” conditions. Here, however, two *C. rhenanus* samples were grouping with the invasives. To verify the source of this odd grouping, I examined SNP variation in the raw sequence data. The genetic distance tree revealed that those individuals were genetically

intermediate between *C. rhenanus* and the invasives, closer to the invasives (Fig. 3). It might be that those are natural hybrids between *C. rhenanus* and invasives, however such hybrids were not found in nature making this possibility less probable. Another and more likely explanation is that samples were contaminated during extraction. As it is not possible to solve this issue *post factum* the samples were eliminated from further analysis.

Details of differential gene expression analysis results are reported in tables S2.1 and S2.2 in the digital supplement.

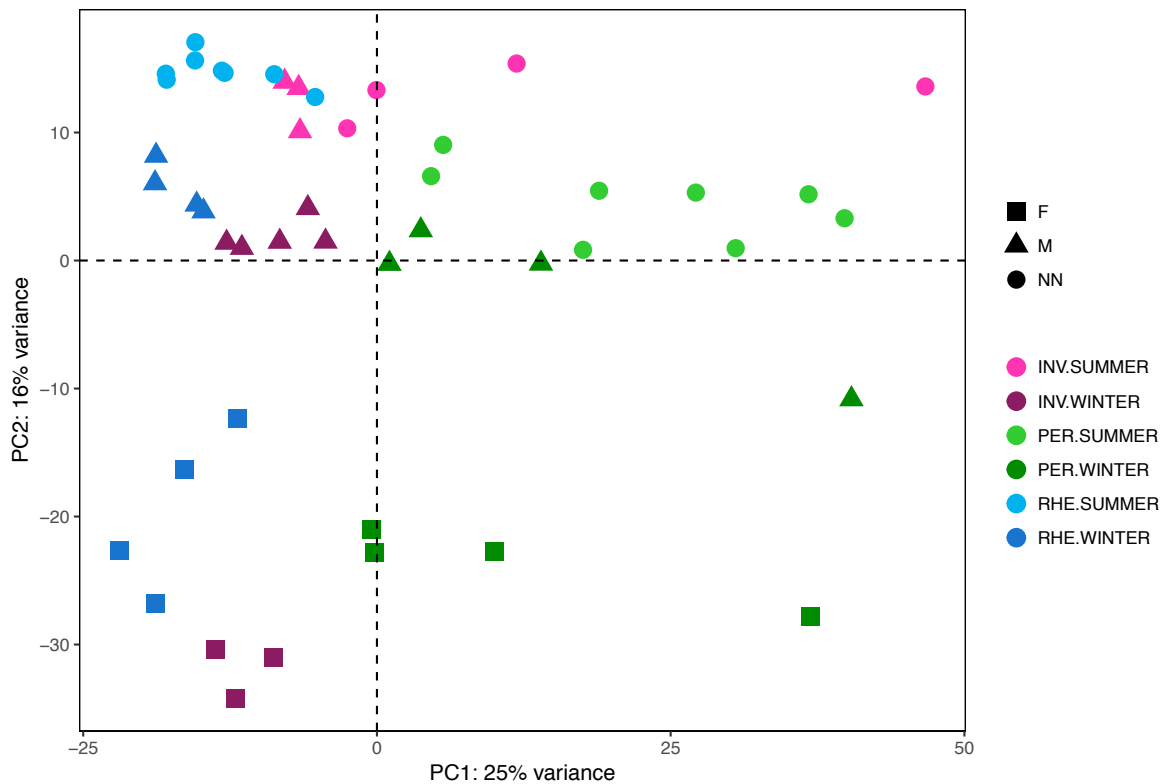


Figure 1. Principal Component Analysis of gene expression profiles from liver tissue of fish raised in the consistent environment. PCAs are based on whole-genome gene expression (17549 genes and 7389 unannotated contigs) for each individual. Colors distinguish lineages (INV - invasive *Cottus*, PER - *C.perifretum*, RHE - *C.rhenanus*) and conditions (SUMMER, WINTER). Shapes distinguish the sex (Female, Male, Not Known/Juvenile).

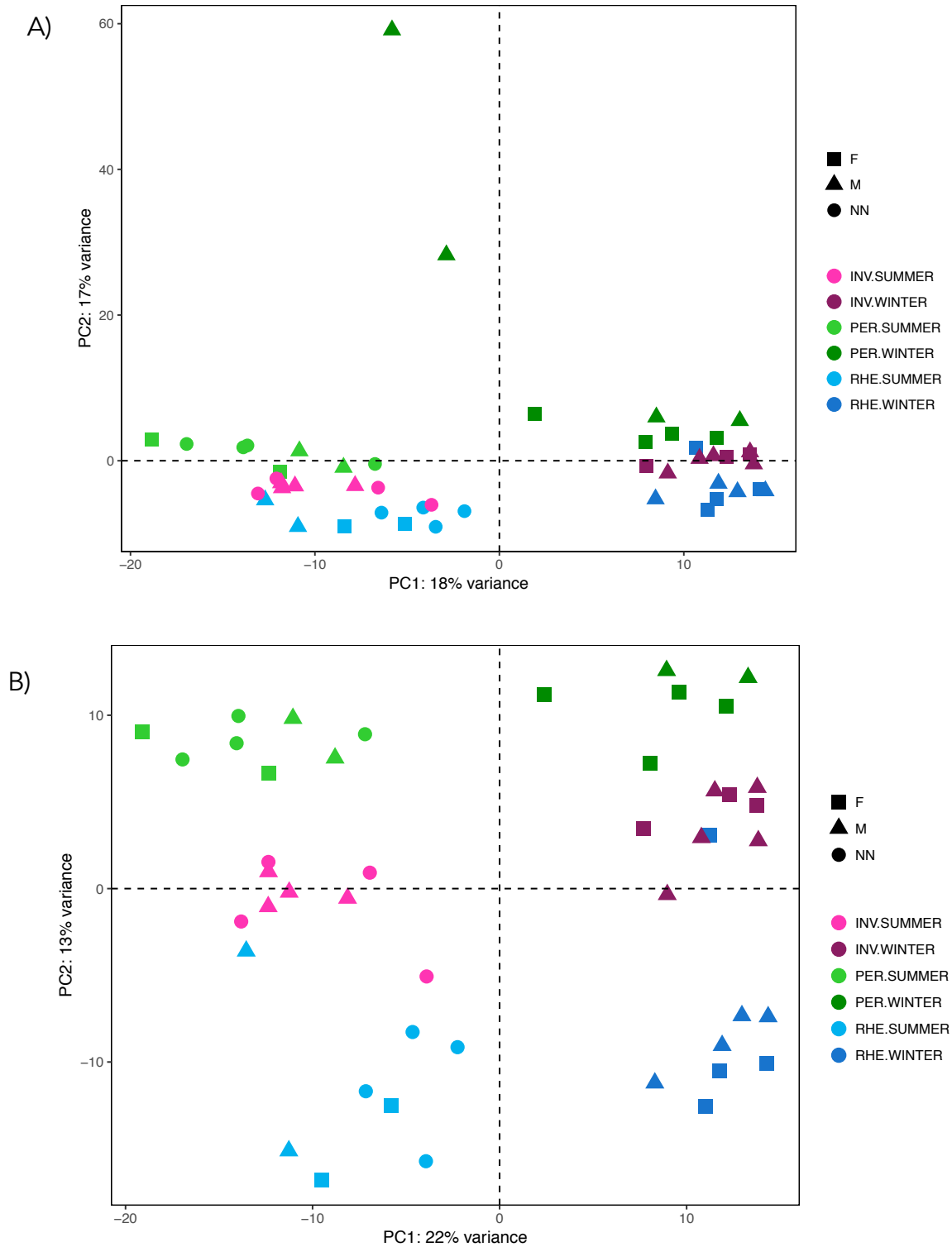


Figure 2. Principal Component Analysis of gene expression profiles from fin tissue of fish raised in the consistent environment. PCAs are based on whole-genome gene expression (17549 genes and 7389 unannotated contigs) for each sample. Colors distinguish lineages (INV-invasive *Cottus*, PER - *C.perifretum*, RHE - *C.rhenanus*) and conditions (SUMMER, WINTER). Shapes distinguish the sex (Female, Male, Not Known/Juvenile). Panel A: all fin samples; Panel B: after excluding two *C. perifretum* outliers.

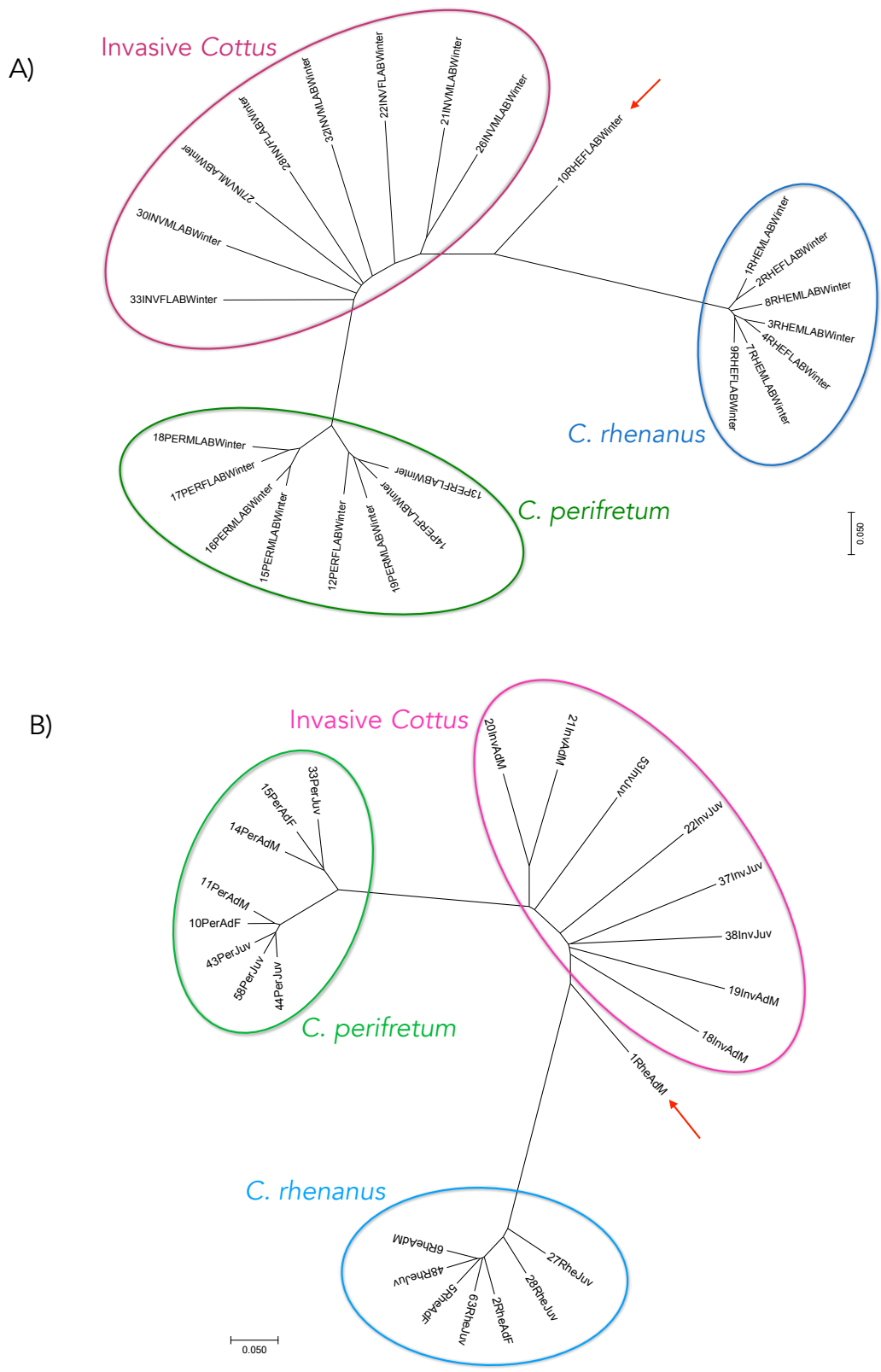


Figure 3. Genetic distance trees of fin tissue samples from the fish bred in consistent environment. Neighbor-joining trees based on the SNPs of all individuals in the consistent environment experiment in “winter” – panel A – and “summer” – panel B. Sample names in the graph specify the sample number, abbreviation of the lineage (RHE, PER, INV) and sex (AdM, AdF, Juv). Color circles and descriptions indicate lineages. Red arrows point out the individuals that did not group according to expectation.

Transgressive gene identification

Differential gene expression analysis revealed that both categories of genes - up- and down-regulated - could be found in the hybrid lineage (Fig. 4). The majority of the transgressively expressed genes in the invasives were the up-regulated ones (liver: N = 73 genes; fin: N = 98 genes), while down-regulated gene numbers were much smaller (liver: N = 10 genes; fin: N = 24 genes).

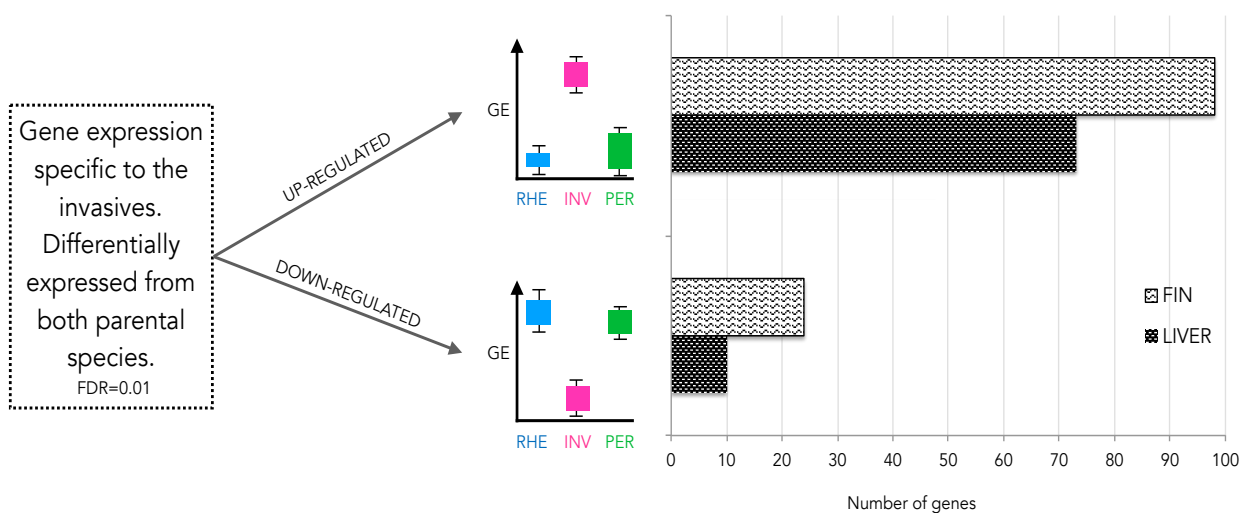


Figure 4. Numbers of differentially expressed genes in the constant environment. Expression patterns specific to the invasive *Cottus*. Genes were tested for differential gene expression at the p -value=0.01 (corrected by False Discovery Rate control according to Benjamini-Hochberg). Genes that were differentially expressed in invasives when compared with both parental species were categorized into up-regulated and down-regulated groups. Histogram shows numbers of genes in each category for both fin and liver tissue.

Statistical analyses – Natural environment samples

Time series gene expression data served as basis for building yearly expression profiles for every gene in each of the two lineages. Differential expression analysis between the two lineages in a time series was performed by *maSigPro* program. The analysis demands strict filtering of lowly expressed genes – at least 20 reads per sample in all samples for a given population are demanded. Genes that did not pass this filtering threshold or were not differentially expressed (neither in time nor between the lineages) were excluded from the analysis. Resulting expression profiles and p-values associated with differential expression analysis are reported in Table S2.1. and Table S2.2 in the digital supplement.

Comparing natural and controlled environment samples

Among the genes identified as up-regulated in the controlled environment in liver samples only 28 out of 73 (21.4% of all differentially expressed genes) could be compared to the time series data collected from the natural environment. 20 of those 28 (71.4%) were differential expressed both in the natural and controlled environment. In fin samples, 37 out of 98 up-regulated genes (23.7%) could be compared to the dataset from natural environment. Differential expression was consistent in lab and nature for 32 of those (86.5%).

Among the down-regulated genes in the controlled environment in liver samples 8 out of 17 (47% of all transgressive down-regulated genes) could be compared to the time series data collected from the natural environment. 6 out of those 8 (75%) were differential expressed both in the natural and controlled environment. In fin samples, 15 out of 24 down-regulated genes (63%) could be compared to the dataset from natural environment. Differential expression was consistent in lab and nature for 10 of those (66.7%).

Gene Ontology terms analysis

Gene ontology (GO) term enrichment analysis did not return any significant enrichment in the consistently up- nor down-regulated genes. We further inspected all GO terms represented to get a detailed picture of the transgressive up-regulated gene set (Biological processes terms in Fig. 5 and 6; all terms in the supplementary materials Tab. S3 and S4) and down-regulated sets (Biological processes terms in Fig. 8; and supplementary materials Tab. S5 and S6).

Among consistently up-regulated genes identified in liver tissue (20 genes) three main biological processes GO terms were represented: 1) Circulatory system development (sprouting angiogenesis); 2) Primary metabolic processes (DNA, protein and lipid metabolic processes); and most of all 3) Transport processes: exocytosis, lipid transport and secretion (Fig. 5).

In fin tissue analysis (32 genes) we could also summarize the GO terms by three main processes: 1) Primary and cell metabolic processes (Oxydation-reduction processes, Protein catabolic processes, Proteolysis, Glyoxylane cycle), 2) Response to light (Melanocyte differentiation, Pigmentation), 3) Protein import into peroxisome (Fig. 6). I selected two examples out of those consistently up-regulated genes in fins, which appeared particularly interesting with respect to possible phenotypic advantage of the invasives. Those are dihydrolipoamide dehydrogenase gene (DLD) (Fig. 9A) and the cytochrome P450 family 27 subfamily C member 1 gene (Cyp27C1) (Fig. 9B).

Consistently down-regulated genes gave relatively little information about GO terms due to small gene sets. In fin tissue (10 genes) we had only five genes that could supply information about biological processes GO terms. Among those were: "regulation of transcription", "signal transduction by phosphorylation", "oxidation-reduction" and "cell redox homeostasis" (Fig. 8A). Liver tissue (6 genes) had only one gene with attached information about

the biological processes GO terms and those were: 1) Carbohydrate metabolic processes; and 2) Primary metabolic processes (Fig. 8B).

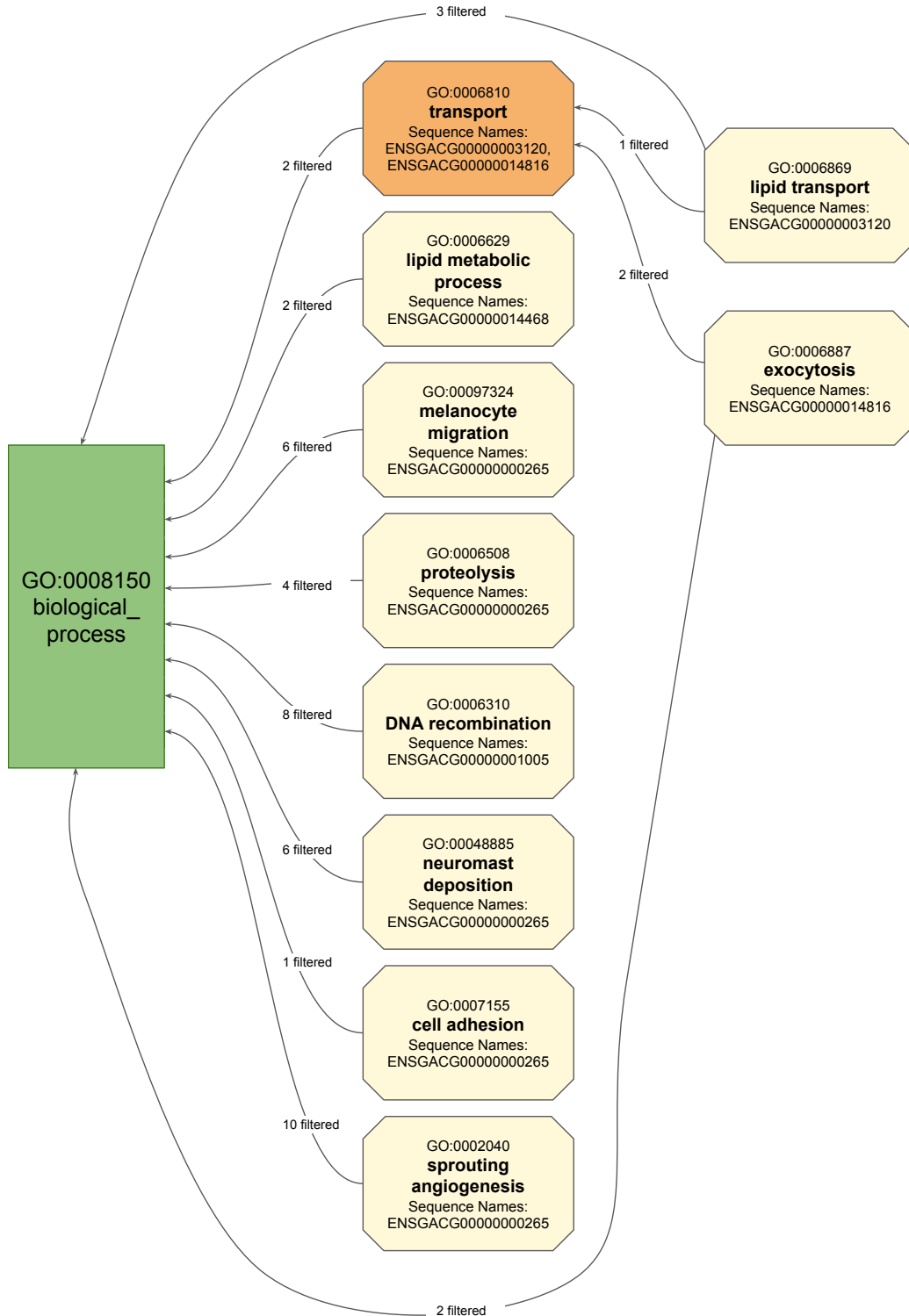


Figure 5. GO terms found in the up-regulated gene list from liver samples. Biological processes GO terms represented in the gene set of genes transgressively overexpressed in the invasives' liver tissue both in natural and controlled habitat.

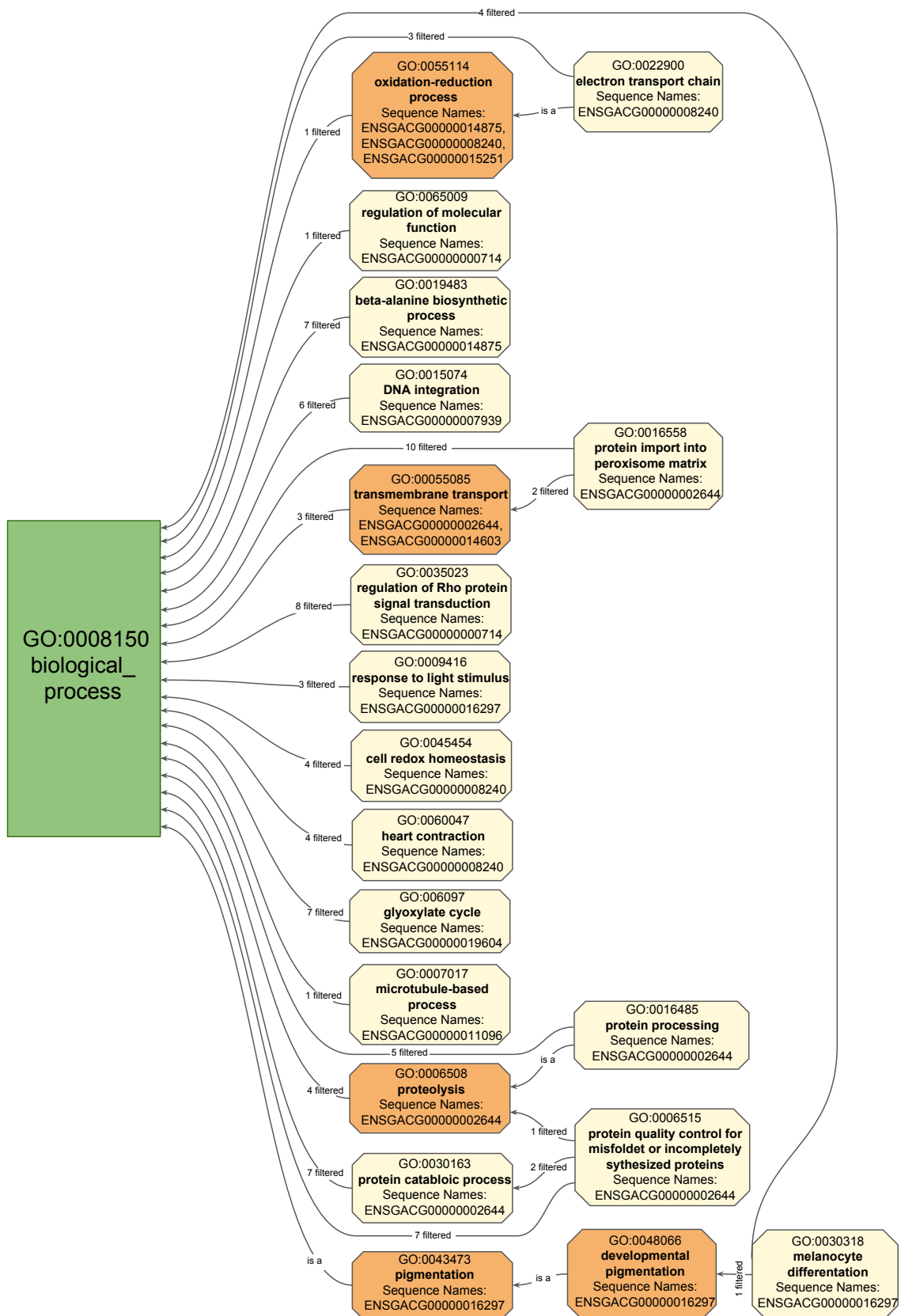


Figure 6. GO terms found in the up-regulated gene list from fin samples. Biological processes GO terms represented in the gene set of genes transgressively overexpressed in the invasives' fin tissue both in natural and controlled habitat.

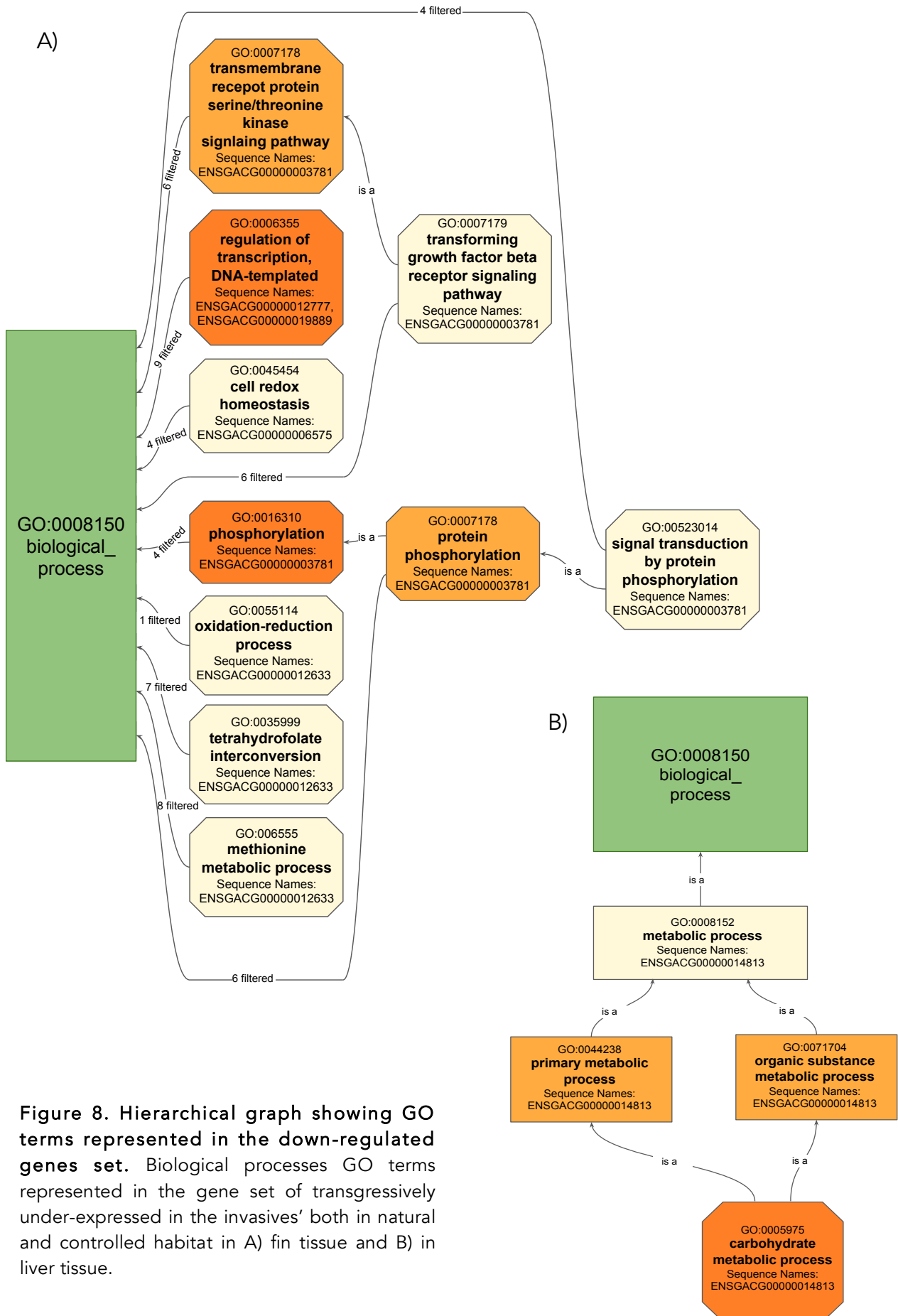


Figure 8. Hierarchical graph showing GO terms represented in the down-regulated genes set. Biological processes GO terms represented in the gene set of transgressively under-expressed in the invasives' both in natural and controlled habitat in A) fin tissue and B) in liver tissue.

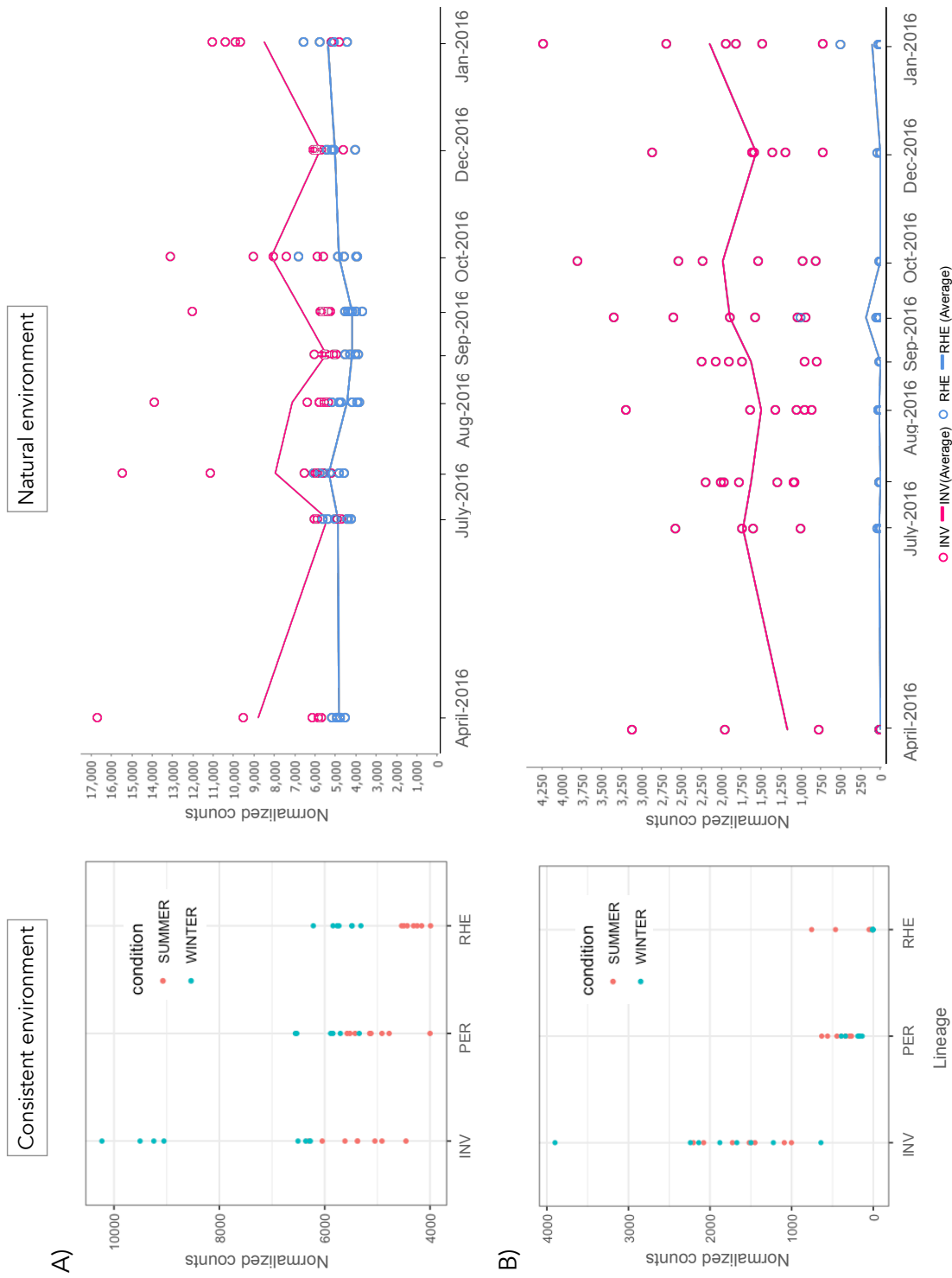


Figure 9. Gene expression in the controlled and natural environment for two genes A) dihydroipoamide dehydrogenase (DLD; ENSGACG00000008240) and B) cytochrome P450 family 27 subfamily C member 1 (Cyp27c1; ENSGACG00000015251). Ge

Discussion

I assessed transgressive gene expression in the invasive *Cottus* lineage of hybrid origin compared to their two parental species *C. rhenanus* and *C. perifretum* at two time points in a controlled, common-garden experiment. These time points represented summer and winter seasons of their life cycle in nature. The results were compared with gene expression data from specimens collected in the natural environment in a well-studied contact zone between *C. rhenanus* and the invasives in the river Sieg.

I found that each of the three lineages studied was characterized by a distinctive genome-wide gene expression profile specific to their genotype. Among transgressively expressed genes in the lineage of hybrid origin I found substantially more up-regulated than down-regulated genes. This result is in accordance with previous work on the *Cottus* system, which indicated enrichment for the gene up-regulation during the short, possibly less than 200 generations, hybrid lineage history (Nolte et al. 2005, Czypionka et al., 2012). Czypionka and collaborators (2012) were careful to accept such result and credited it to possible over-representation of some organs or tissues in the analyzed samples, as they were analyzing RNA extracted from the whole fish. However, in the present work we analyze gene expression in two separate tissues and we corroborate the enrichment for up-regulated gene expression in the hybrid lineage.

We expect that the overexpressed traits might be connected with adaptation to new environment, as in the case of desert sunflowers (Rieseberg et al., 2003, 1999). As the nature of this adaptation is at the heart of our study, I further analyzed up- and down-regulated gene sets. I evaluated to what extent the patterns measured in the laboratory were also detectable in the natural setting, where a range of ecological forces besides temperature is affecting the fish. The analysis pipeline for time series gene expression applies

strict selection criteria on expression levels and expression differences and all genes that do not pass the criteria are excluded from the analysis. It appears that around 60% of the differentially expressed genes in our lab experiment were excluded from the analysis and we could not verify their expression profiles in the natural environment. Among the genes for which profiles in the natural environment were available, the vast majority (~73% in liver and ~76% in fin) showed consistent up- or down-regulation. Precisely, those genes were identified as transgressive when compared with both parental species in controlled conditions and also expressed significantly higher (or lower) in the invasives than in *C. rhenanus* in natural conditions. This suggests that the expression of transgressive phenotypes is genetically fixed and remains largely unmodified by environmental influences. Those genes are promising candidates for further studies as they present an invasive-specific gene expression pattern that persists in the natural environment and is therefore likely to be ecologically relevant.

Careful examination of gene ontology terms and gene functions from this candidate gene list revealed common themes and possible functions. Most strikingly, despite the fact that transgressive gene lists from liver and fin tissues did not overlap at all, they shared common gene ontology terms. These were connected to "Primary Metabolic Processes", which encompass lipid metabolism, protein metabolism, sugar metabolism and oxidation-reduction processes. This category of GO terms was the most important common theme among the candidates in both tissues studied. This result strongly suggests that invasive *Cottus* has an increased metabolic rate, more mitotic divisions and higher metabolic turnover. Another important function represented in the liver transgressive gene list was "Angiogenesis", which suggests that there might be differences in the cardiovascular system development between lineages. Our results provide additional evidence for the initial *Cottus*

transcriptome exploration by Czypionka et al. (2012), which also found increased metabolism related functions in invasives. Accumulation of those GO terms among the transgressive genes might suggest that the invasives have indeed increased their metabolic rate and extended cardiovascular system to deal with summer heat. Oxygen availability in water decreases as the temperature rises (Pörtner and Knust, 2007) and the invasive *Cottus* do experience higher temperatures in summer when compared with *C. rhenanus* (Chapter 2). Our results warrant further studies on possible adaptive metabolism and blood transportation system modifications in the invasives.

In the transgressive gene list from fin tissue we found the GO terms: "Response to light", "Melanocyte differentiation", "Pigmentation" and "Protein import into peroxisome". This reflects the general character of skin, an organ in constant contact with the external world, which protects the organism from adversities such as UV light and pollution (Slominski et al., 2005). It is not clear at this point whether invasive *Cottus* are more exposed to the UV light.

Among the consistently up-regulated gene list I found two particularly interesting candidates: the dihydrolipoamide dehydrogenase gene (DLD) and cytochrome P450 family 27 subfamily C member 1 gene (*Cyp27c1*). Dihydrolipoamide dehydrogenase is a mitochondrial enzyme that is known to play crucial role in energy metabolism. DLD is indispensable for the complete reaction of at least five different multi-enzyme complexes and has a wide range of effects on the metabolism of an organism (Brautigam et al., 2006; Ciszak et al., 2006). Functions of this gene are very much in line with the hypothesis proposed by Czypionka et al. (2012) however this particular gene was not identified as transgressive in their study. This is not surprising because Czypionka et al. (2012) studied gene expression solely in summer. This study demonstrated that DLD's expression in the consistent environment stays similar and stable across seasons in two parental species and the invasives in summer

however in winter it becomes transgressive in invasives. In the natural habitat its expression stayed stable in *C. rhenanus* across the year but fluctuated significantly in the invasives. We observed relatively low gene expression, similar to *C. rhenanus*, during summer but much higher levels in spring (April), autumn (August-October) and winter (January) when the most important growing phase takes place, which suggests that the expression of this gene might be connected to metabolism and growth and that it is fine-tuned to fluctuations of the environmental factors.

Cyp27c1, on the other hand, was expressed higher than in both parental species across seasons in the consistent environment. In natural environment it was also up-regulated in the invasives compared to *C. rhenanus* all year long. Interestingly, the study by Dennenmoser et al. (2017) suggests that this gene is duplicated in the hybrid lineage and that may explain higher gene expression levels by a gene dosage effect. Czypionka et al. (2012) have also classified Cyp27c1 as an up-regulated gene, however its function was not known at the time. Interestingly, the Cyp27c1-coded protein has recently been described to convert vitamin A₁ to vitamin A₂, which increases vision in the infrared in zebrafish (Enright et al., 2015). As the authors explain, improved infrared vision may constitute a relevant fitness advantage for freshwater fish. When waters are murky, either seasonally or all year round, visibility drops drastically. Increased capabilities to see in red are useful and can help spot prey or possible predators. This might directly refer to the *Cottus* system. The invasive lineage colonized rivers that were never inhabited by the parental species. While *C. rhenanus* and *C. perifretum* live in small, turbulent, shallow streams, the invasives moved to bigger, deeper rivers that are oftentimes murky. Supposing that this gene is indeed performing the function as described by Enright and collaborators, it is possible that the additional copy of Cyp27c1 gene and its increased gene expression might have contributed to the survival,

establishment and eventual success of the invasives in the new environment. This remains speculative, however, until proven by functional studies.

Overall the study improved the understanding the ecological context of gene expression response in *Cottus*. Sampling multiple times during the year allowed to significantly improve the ability to detect possible functions involved in adaptation and narrow down the list of candidate genes involved. Combining experiments based in the laboratory with transcriptomics in the natural environment assured that the inferred functions are relevant in nature as well. The study could confirm some of the initially proposed functions by Czypionka et al. (2012), giving confidence that those results are robust. However the study also unravels new possible avenues that will demand more work. Still, it represents a major progress in describing how the invasive lineage differs from their parental species.

Supplement

Table S1. List of individuals with origin and sex specifications for three lineages of *Cottus* sp. sampled in WINTER conditions (January 2015).

#	Sample ID	Species	Origin		Sex	FIN	LIVER
1	1	<i>C. rhenanus</i>	Wahnbach	Wild 2014	M	Yes	Yes
2	2	<i>C. rhenanus</i>	Wahnbach	Wild 2014	F	Yes	Yes
3	3	<i>C. rhenanus</i>	Krabach	Wild 2014	M	Yes	Yes
4	4	<i>C. rhenanus</i>	Krabach	Wild 2014	F	Yes	Yes
5	7	<i>C. rhenanus</i>	Naafbach	ex ovo	M	Yes	Yes
6	8	<i>C. rhenanus</i>	Fokkenbach	Wildfang 2013	M	Yes	Yes
7	9	<i>C. rhenanus</i>	Bröl	ex ovo	F	Yes	Yes
8	10	<i>C. rhenanus</i>	Bröl	ex ovo	F	Yes	Yes
9	12	<i>C. perifretum</i>	LB-2-2014	ex ovo	F	Yes	Yes
10	13	<i>C. perifretum</i>	LB-3-2014	ex ovo	F	Yes	Yes
11	14	<i>C. perifretum</i>	LB-3-2014	ex ovo	F	Yes	Yes
12	15	<i>C. perifretum</i>	WN-2014	ex ovo	M	Yes	Yes
13	16	<i>C. perifretum</i>	WN-2014	ex ovo	M	Yes	Yes
14	17	<i>C. perifretum</i>	WN	ex ovo	F	Yes	Yes
15	18	<i>C. perifretum</i>	WN	ex ovo	M	Yes	Yes
16	19	<i>C. perifretum</i>	WN	ex ovo	M	Yes	Yes
17	21	<i>C. invasive</i>	Lower Wahnbach	wild	M	Yes	Yes
18	22	<i>C. invasive</i>	Sieg Bustorf	wild	F	Yes	Yes
19	26	<i>C. invasive</i>	Invasive O- 2014	ex ovo	M	Yes	Yes
20	27	<i>C. invasive</i>	Sieg Bustorf	wild	M	Yes	Yes
21	28	<i>C. Invasive</i>	Sieg Bustorf	wild	F	Yes	Yes
22	30	<i>C. Invasive</i>	Sieg Bustorf	wild	M	Yes	Yes
23	32	<i>C. Invasive</i>	Lower Wahnbach	wild	M	Yes	Yes
24	33	<i>C. Invasive</i>	Sieg Bustorf	wild	F	Yes	Yes

Table S2. Origin, developmental stage and sex of *Cottus* samples for which transcriptome was analyzed in SUMMER conditions (Septembr 2015).

#	Sample ID	Species	Pop. of Origin	Sex	Age	FIN	LIVE R
1	1	<i>C. rhenanus</i>	Broel pure	M	adult	Yes	No
2	2	<i>C. rhenanus</i>	Broel pure	F	adult	Yes	No
3	5	<i>C. rhenanus</i>	Wahnbach oben	F	adult	Yes	No
4	6	<i>C. rhenanus</i>	Krabach	M	adult	Yes	No
5	27	<i>C. rhenanus</i>	Broel 2015		juvenile	Yes	Yes
6	28	<i>C. rhenanus</i>	Broel 2015		juvenile	Yes	Yes
7	48	<i>C. rhenanus</i>	Broel 2015 (belege x WF)		juvenile	Yes	Yes
8	63	<i>C. rhenanus</i>	Wanne (Ruhr)/Rhen9_2015		juvenile	Yes	Yes
9	10	<i>C. perifretum</i>	WitteNete	F	adult	Yes	No
10	11	<i>C. perifretum</i>	WitteNete	M	adult	Yes	No
11	14	<i>C. perifretum</i>	LaarseBeck	M	adult	Yes	No
12	15	<i>C. perifretum</i>	LaarseBeck	F	adult	Yes	No
13	33	<i>C. perifretum</i>	LaarseBeck 2015		juvenile	Yes	Yes
14	43	<i>C. perifretum</i>	WitteNete 2015 pure		juvenile	Yes	Yes
15	44	<i>C. perifretum</i>	WitteNete 2015 pure		juvenile	Yes	Yes
16	58	<i>C. perifretum</i>	WitteNete 2015		juvenile	Yes	Yes
17	18	Invasive	Pleisbach 2015	M	adult	Yes	Yes
18	19	Invasive	Pleisbach 2015	M	adult	Yes	Yes
19	20	Invasive	Inv8_2014	M	adult	Yes	Yes
20	21	Invasive	Inv8_2014	M	adult	Yes	Yes
21	22	Invasive	Pleis 2015 (F1 5 belege pool)		juvenile	Yes	Yes
22	37	Invasive	Inv7_2015		juvenile	Yes	Yes
23	38	Invasive	Pleis 2015		juvenile	Yes	Yes
24	53	Invasive	Inv10_2015/Inv11_2015	M	adult	Yes	Yes
25	29	<i>C. rhenanus</i>	Broel 2015		juvenile	No	Yes
26	32	<i>C. perifretum</i>	LaarseBeck 2015		juvenile	No	Yes
27	45	<i>C. perifretum</i>	WitteNete 2015 pure		juvenile	No	Yes
28	49	<i>C. rhenanus</i>	Broel 2015 (belege x WF)		juvenile	No	Yes
29	50	<i>C. rhenanus</i>	Broel 2015 (belege x WF)		juvenile	No	Yes
30	59	<i>C. perifretum</i>	WitteNete 2015		juvenile	No	Yes
31	60	<i>C. perifretum</i>	WitteNete 2015		juvenile	No	Yes
32	64	<i>C. rhenanus</i>	Wanne (Ruhr)/Rhen9_2015		juvenile	No	Yes

Table S3. List of up-regulated genes in invasives' liver tissue, both in natural and controlled environment. Gene names and GO terms associated with them. Last column specifies in a given gene was differentially expressed in winter, summer or both.

Sequence ID	Description	GO Names	DE
ENSGACG00000000265	leishmanolysin-like (metallopeptidase M8 family)	P:sprouting angiogenesis; F:metalloendopeptidase activity; P:proteolysis; P:cell adhesion; C:membrane; C:integral component of membrane; P:neuromast deposition; P:melanocyte migration	WINTER
ENSGACG00000001005	RecQ helicase-like	F:nucleotide binding; F:nucleic acid binding; F:helicase activity; F:ATP binding; C:nucleus; P:DNA recombination; F:ATP-dependent helicase activity; F:hydrolase activity	WINTER& SUMMER
ENSGACG00000003120	oxysterol binding protein like 10	P:transport; P:lipid transport	WINTER
ENSGACG00000003701	dopey family member 2	-	WINTER
ENSGACG00000005448	Mov10 RISC complex RNA helicase a	-	WINTER
ENSGACG00000006453	-	F:GTPase activity; F:GTP binding	WINTER
ENSGACG00000008769	WSC domain containing 2	C:membrane; C:integral component of membrane	WINTER& SUMMER
ENSGACG00000010247	si:ch211-248a14.8	C:membrane; C:integral component of membrane	WINTER
ENSGACG00000012548		F:GTP binding	SUMMER
ENSGACG00000013166	dual adaptor of phosphotyrosine and 3-phosphoinositides	-	WINTER
ENSGACG00000014468	-	P:lipid metabolic process; C:membrane; C:integral component of membrane	WINTER
ENSGACG00000014816	exocyst complex component 3	C:exocyst; P:exocytosis	WINTER
ENSGACG00000015277	NTPase, KAP family P-loop domain containing 1	C:membrane; C:integral component of membrane	WINTER
ENSGACG00000017217	-	-	WINTER
ENSGACG00000018359	NEDD4 binding protein 3	-	WINTER
ENSGACG00000018624	solute carrier family 38, member 9	C:membrane; C:integral component of membrane	SUMMER
comp53119_c1_seq1	-	-	WINTER
comp58544_c9_seq1	-	-	WINTER
comp59166_c13_seq1	-	-	WINTER& SUMMER
comp55697_c0_seq1	-	-	WINTER

Table S4. List of up-regulated genes in invasives' fin tissue both in natural and controlled environment. Gene names and GO terms associated with them. Last column specifies in a given gene was differentially expressed in winter, summer or both.

Sequence ID	Description	GO Names	DE
ENSGACG00000000419		F:nucleic acid binding	SUMMER
ENSGACG00000000714	Rho guanine nucleotide exchange factor (GEF) 25b	F:Rho guanyl-nucleotide exchange factor activity; P:regulation of Rho protein signal transduction; P:regulation of molecular function	WINTER
ENSGACG00000002644	Ion peptidase 2, peroxisomal	F:nucleotide binding; F:ATP-dependent peptidase activity; F:serine-type endopeptidase activity; F:ATP binding; C:peroxisome; C:peroxisomal matrix; P:proteolysis; P:protein quality control for misfolded or incompletely synthesized proteins; F:peptidase activity; F:serine-type peptidase activity; P:protein processing; P:protein import into peroxisome matrix; F:hydrolase activity; F:ATPase activity; P:protein catabolic process	SUMMER
ENSGACG00000006992	excision repair cross-complementation group 6-like 2	F:ATP binding	SUMMER
ENSGACG00000007063	envoplakin	C:cytoskeleton	WINTER
ENSGACG00000007616	RMI1, RecQ mediated genome instability 1, homolog (S. cerevisiae)	F:nucleotide binding	WINTER
ENSGACG00000007939		F:nucleic acid binding; P:DNA integration	WINTER&SUMMER
ENSGACG00000008240	dihydrolipoamide dehydrogenase	F:dihydrolipoyl dehydrogenase activity; C:cell; F:electron transfer activity; F:oxidoreductase activity; F:oxidoreductase activity, acting on a sulfur group of donors, NAD(P) as acceptor; P:electron transport chain; P:cell redox homeostasis; F:flavin adenine dinucleotide binding; P:oxidation-reduction process; P:heart contraction	WINTER
ENSGACG00000008379	si:ch73-103b11.2	-	SUMMER
ENSGACG00000011096	dynein light chain LC8-type 2	C:microtubule associated complex; P:microtubule-based process; C:dynein complex	WINTER
ENSGACG00000013166	dual adaptor of phosphotyrosine and 3-phosphoinositides	-	WINTER&SUMMER
ENSGACG00000014572	LSM14A mRNA processing body assembly factor b	-	WINTER
ENSGACG00000014577	ribosomal protein L37	-	SUMMER
ENSGACG00000014603	solute carrier family 25, member 39	P:transport; C:membrane; C:integral component of membrane; P:transmembrane transport	WINTER

ENSGACG00000014875	dihydropyrimidine dehydrogenase b	F:catalytic activity; C:cytoplasm; F:oxidoreductase activity; F:oxidoreductase activity, acting on the CH-CH group of donors; F:dihydropyrimidine dehydrogenase (NADP+) activity; P:beta-alanine biosynthetic process; F:iron-sulfur cluster binding; F:4 iron, 4 sulfur cluster binding; P:oxidation-reduction process	WINTER
ENSGACG00000015028	glycine amidinotransferase (L-arginine:glycine amidinotransferase)	F:amidinotransferase activity	WINTER
ENSGACG00000015251	cytochrome P450 family 27 subfamily C member 1	F:monooxygenase activity; F:iron ion binding; F:oxidoreductase activity; F:oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen; F:heme binding; F:metal ion binding; P:oxidation-reduction process	WINTER& SUMMER
ENSGACG00000015815	EFR3 homolog A (S. cerevisiae)		SUMMER
ENSGACG00000016297	solute carrier family 45, member 2	P:response to light stimulus; C:membrane; C:integral component of membrane; P:melanocyte differentiation; P:pigmentation; P:developmental pigmentation	SUMMER
ENSGACG00000016305		C:lysosome; C:membrane; C:integral component of membrane; F:signaling receptor activity	SUMMER
ENSGACG00000016452	kelch-like family member 3		WINTER
ENSGACG00000017852	EF-hand calcium binding domain 14	C:membrane; C:integral component of membrane	WINTER
ENSGACG00000018624	solute carrier family 38, member 9	C:membrane; C:integral component of membrane	SUMMER
ENSGACG00000019604		F:catalytic activity; F:malate synthase activity; P:glyoxylate cycle	SUMMER
ENSGACG00000020236		F:benzodiazepine receptor binding	WINTER
ENSGACG00000020603	listerin E3 ubiquitin protein ligase 1	F:zinc ion binding; F:metal ion binding	SUMMER
comp38619_c0_seq2	-	-	SUMMER
comp58293_c4_seq6	-	-	SUMMER
comp58611_c1_seq1	-	-	WINTER& SUMMER
comp58685_c6_seq21	-	-	SUMMER
comp59238_c0_seq8	-	-	SUMMER
comp68504_c0_seq1	-	-	WINTER& SUMMER

Table S5. List of down-regulated genes in invasives' liver tissue, both in natural and controlled environment. Gene names and GO terms associated with them. Last column specifies in a given gene was differentially expressed in winter, summer or both.

Sequence ID	Description	GO Names	DE
ENSGACG00000003632	reactive intermediate imine deaminase A homolog	-	SUMMER
ENSGACG00000008282	fibroblast growth factor receptor like 1	C:membrane; C:integral component of membrane	WINTER
ENSGACG00000009554	crystallin, mu	-	SUMMER
ENSGACG00000014813	-	P:carbohydrate metabolic process; C:membrane; C:integral component of membrane; F:transferase activity, transferring hexosyl groups	WINTER& SUMMER
ENSGACG00000015532	zinc finger protein 513	F:nucleic acid binding	WINTER
ENSGACG00000020401	si:ch211-137i24.10	C:membrane; C:integral component of membrane	WINTER

Table S6. List of down-regulated genes in invasives' fin tissue, both in natural and controlled environment. Gene names and GO terms associated with them. Last column specifies in a given gene was differentially expressed in winter, summer or both.

Sequence ID	Description	GO Names	DE
ENSGACG00000003781	-	F:nucleotide binding; F:protein kinase activity; F:protein serine/threonine kinase activity; F:obsolete signal transducer, downstream of receptor, with serine/threonine kinase activity; F:transforming growth factor beta receptor activity, type II; F:ATP binding; P:protein phosphorylation; P:transmembrane receptor protein serine/threonine kinase signaling pathway; P:transforming growth factor beta receptor signaling pathway; C:membrane; C:integral component of membrane; F:kinase activity; P:phosphorylation; F:transferase activity; P:signal transduction by protein phosphorylation; F:signaling receptor activity; C:receptor complex; F:metal ion binding	WINTER
ENSGACG00000006575	fibroblast growth factor receptor like 1	C:cell; P:cell redox homeostasis	WINTER
ENSGACG00000006706	tumor necrosis factor receptor superfamily, member 11b	-	SUMMER
ENSGACG00000012633	methylenetetrahydrofolate reductase (NAD(P)H)	F:methylenetetrahydrofolate reductase (NAD(P)H) activity; P:methionine metabolic process; F:oxidoreductase activity; P:tetrahydrofolate interconversion; P:oxidation-reduction process	WINTER
ENSGACG00000012777	forkhead box P3b	F:DNA binding; F:DNA-binding transcription factor activity; C:nucleus; P:regulation of transcription, DNA-templated; F:sequence-specific DNA binding	WINTER
ENSGACG00000015516	-	F:catalytic activity	WINTER
ENSGACG00000019889	Spi-C transcription factor	F:DNA binding; F:DNA-binding transcription factor activity; C:nucleus; P:regulation of transcription, DNA-templated; F:sequence-specific DNA binding	SUMMER
comp46419_c0_seq1	-	-	WINTER
comp49059_c0_seq12	-	-	WINTER
comp59062_c2_seq7	-	-	WINTER

Concluding remarks and Perspectives:

In this thesis I have assessed phenotypic differentiation of the emerging homoploid hybrid species – the invasive *Cottus* lineage – from its parental species and with respect to the environment.

I showed that the invasive lineage, despite living in summer-warm environment, does not display an altered temperature preference - all three *Cottus* lineages were choosing the same temperature (Chapter 1). The laboratory-based experimental study on growth rate turned out to be inconclusive, because of technical problems and persuaded me to study a more complex phenotype – gene expression - both in the laboratory and natural environment (Chapters 2 and 3).

Since ecology is thought to play a crucial role in the emergence of homoploid hybrid speciation, sampling phenotypic variation in the right ecological context is the key to understand evolutionary processes shaping the system. I observed in an unprecedented detail gene expression changes in nature by sampling the invasive lineage and *C. rhenanus* nine times in the course of a complete year. Contrary to initial expectations, the largest differentiation in expression patterns between the lineages was detected during autumn and winter and was mostly independent of the water temperature. Furthermore, the majority of the divergently expressed genes was up-regulated in the invasives and these were connected to metabolism related gene ontology terms. Complementary common-garden experiments revealed that most of the genes had genetically determined, heritable expression patterns (Chapter 2).

Transgressive trait expression in the hybrids is thought to play an important role in the homoploid hybrid speciation process. Recombination of parental alleles can result in phenotypes lying outside of the normal parental

range, which might in turn increase hybrid's fitness in the new ecological niche. I assessed transgressive trait expression in the invasive lineage of *Cottus* and concluded that there is clear enrichment for the up-regulated transgressive traits (Chapter 3). Among those I identified a gene with increased copy number, which is possibly involved in enhancing vision in murky waters (*Cyp27c1*) and therefore might contribute to the adaptation in the invasives. It is an interesting candidate for follow-up functional studies.

Interestingly, however, the majority of genes with the up-regulated expression in invasives were connected, again, to metabolic functions. This is a recurring result in my as well as previous studies (Czypionka et al., 2012) and therefore encourages serious consideration of possible mechanisms behind it. Metabolic rate is well established as an important phenotypic trait, which is tightly linked with ecology. "The Metabolic Theory of Ecology" conceptualized by Brown et al. (2004) links metabolic processes with the ecological processes happening in populations. However, its connection with speciation research has not been fully explored to date. McFarlane et al. (2016) initiated this direction of research by studying resting metabolic rate in the hybrid zone of flycatchers. The secondary contact zone in flycatchers is the specific case of hybridization that led to reinforcement of reproductive barriers between the species. Hybrid offspring has decreased fitness compared to both parental species and is therefore selected against. McFarlane et al. (2016) have demonstrated that hybridization was causing incompatibilities in the oxidation-reduction pathway genes, which was resulting in decreased metabolic dysfunction and lowered fitness.

However, as we are nowadays well aware that hybridization can also have "positive" effects and lead to the formation of new species, we might consider that hybridization could also result in improvement of metabolic processes and therefore allowing hybrids to establish in a new ecological

niche. In my study I observed that genes related to basic metabolic functions are up-regulated in the invasives. This might indicate that they have higher metabolic rate and maybe it improves their performance. Of course it remains purely speculative at this point and will require functional studies. For that purpose I propose to study resting metabolic rate – a proxy for metabolic performance – in both parental species and the invasive lineage, which will determine whether invasives indeed have enhanced metabolism. Furthermore, we should investigate divergence between lineages in the genes coding proteins of the oxidation-reduction chain and test if there are certain parental allele combinations fixed in the invasives.

Overall my study laid ground for further study on homoploid hybrid speciation by functional studies on the candidate traits possibly involved in adaptation in the *Cottus* system and testing whether or not those traits are a result of hybridization.

References:

- Abbott, R., Albach, D., Ansell, S., Arntzen, J.W., Baird, S.J.E., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C.A., Buggs, R., Butlin, R.K., Dieckmann, U., Eroukhmanoff, F., Grill, A., Cahan, S.H., Hermansen, J.S., Hewitt, G., Hudson, A.G., Jiggins, C., Jones, J., Keller, B., Marczewski, T., Mallet, J., Martinez-Rodriguez, P., Möst, M., Mullen, S., Nichols, R., Nolte, A.W., Parisod, C., Pfennig, K., Rice, A.M., Ritchie, M.G., Seifert, B., Smadja, C.M., Stelkens, R., Szymura, J.M., Väinölä, R., Wolf, J.B.W., Zinner, D., 2013. Hybridization and speciation. *J. Evol. Biol.* 26, 229–246. <https://doi.org/10.1111/j.1420-9101.2012.02599.x>
- Abbott Richard J., Barton Nicholas H., Good Jeffrey M., 2016. Genomics of hybridization and its evolutionary consequences. *Mol. Ecol.* 25, 2325–2332. <https://doi.org/10.1111/mec.13685>
- Alvarez, M., Schrey, A.W., Richards, C.L., 2015. Ten years of transcriptomics in wild populations: what have we learned about their ecology and evolution? *Mol. Ecol.* 24, 710–725. <https://doi.org/10.1111/mec.13055>
- Anderson, E., Stebbins, G.L., 1954. Hybridization as an Evolutionary Stimulus. *Evolution* 8, 378–388. <https://doi.org/10.1111/j.1558-5646.1954.tb01504.x>
- Angilletta, M.J., 2009. *Thermal Adaptation: A Theoretical and Empirical Synthesis*. OUP Oxford.
- Arnold, M.L., 1997. *Natural Hybridization and Evolution*. Oxford University Press, USA.
- Arnold, M.L., Ballerini, E.S., Brothers, A.N., 2012. Hybrid fitness, adaptation and evolutionary diversification: lessons learned from Louisiana Irises. *Heredity* 108, 159–166. <https://doi.org/10.1038/hdy.2011.65>
- Balcombe, J., 2016. What a Fish Knows: The Inner Lives of Our Underwater Cousins. *Scientific American*.
- Barton, N. H., 2013. Does hybridization influence speciation? *J. Evol. Biol.* 26, 267–269. <https://doi.org/10.1111/jeb.12015>
- Barton, N.H., 2001. The role of hybridization in evolution. *Mol. Ecol.* 10, 551–568. <https://doi.org/10.1046/j.1365-294x.2001.01216.x>
- Baskett, M.L., Gomulkiewicz, R., 2011. Introgressive hybridization as a mechanism for species rescue. *Theor. Ecol.* 4, 223–239. <https://doi.org/10.1007/s12080-011-0118-0>
- Bateson, W., 1909. *Heredity and variation in modern lights*, in: *Darwin and Modern Science*,. Cambridge University Press, pp. 85–101.
- Bell, M.A., Travis, M.P., 2005. Hybridization, transgressive segregation, genetic covariation, and adaptive radiation. *Trends Ecol. Evol.* 20, 358–361. <https://doi.org/10.1016/j.tree.2005.04.021>

- Berner, D., Thibert-Plante, X., 2015. How mechanisms of habitat preference evolve and promote divergence with gene flow. *J. Evol. Biol.* 28, 1641–1655. <https://doi.org/10.1111/jeb.12683>
- Bolker, B.M., Brooks, M.E., Clark, C.J., Geange, S.W., Poulsen, J.R., Stevens, M.H.H., White, J.-S.S., 2009. Generalized linear mixed models: a practical guide for ecology and evolution. *Trends Ecol. Evol.* 24, 127–135. <https://doi.org/10.1016/j.tree.2008.10.008>
- Brautigam, C.A., Wynn, R.M., Chuang, J.L., Machius, M., Tomchick, D.R., Chuang, D.T., 2006. Structural insight into interactions between dihydrolipoamide dehydrogenase (E3) and E3 binding protein of human pyruvate dehydrogenase complex. *Struct. Lond. Engl.* 1993 14, 611–621. <https://doi.org/10.1016/j.str.2006.01.001>
- Brown, J.H., Gillooly, J.F., Allen, A.P., Savage, V.M., West, G.B., 2004. Toward a Metabolic Theory of Ecology. *Ecology* 85, 1771–1789. <https://doi.org/10.1890/03-9000>
- Buerkle, C.A., Morris, R.J., Asmussen, M.A., Rieseberg, L.H., 2000. The likelihood of homoploid hybrid speciation. *Heredity* 84, 441–451. <https://doi.org/10.1046/j.1365-2540.2000.00680.x>
- Camon, E., Magrane, M., Barrell, D., Binns, D., Fleischmann, W., Kersey, P., Mulder, N., Oinn, T., Maslen, J., Cox, A., Apweiler, R., 2003. The Gene Ontology Annotation (GOA) Project: Implementation of GO in SWISS-PROT, TrEMBL, and InterPro. *Genome Res.* 13, 662–672. <https://doi.org/10.1101/gr.461403>
- Ciszak, E.M., Makal, A., Hong, Y.S., Vettaikorumakankau, A.K., Korotchkina, L.G., Patel, M.S., 2006. How dihydrolipoamide dehydrogenase-binding protein binds dihydrolipoamide dehydrogenase in the human pyruvate dehydrogenase complex. *J. Biol. Chem.* 281, 648–655. <https://doi.org/10.1074/jbc.M507850200>
- Colliard, C., Sicilia, A., Turrisi, G.F., Arculeo, M., Perrin, N., Stöck, M., 2010. Strong reproductive barriers in a narrow hybrid zone of West-Mediterranean green toads (*Bufo viridissubgroup*) with Plio-Pleistocene divergence. *BMC Evol. Biol.* 10, 232. <https://doi.org/10.1186/1471-2148-10-232>
- Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M., Robles, M., 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674–3676. <https://doi.org/10.1093/bioinformatics/bti610>
- Conesa, A., Nueda, M.J., Ferrer, A., Talón, M., 2006. maSigPro: a method to identify significantly differential expression profiles in time-course microarray experiments. *Bioinformatics* 22, 1096–1102. <https://doi.org/10.1093/bioinformatics/btl056>
- Coyne, J.A., Orr, H.A., 2004. *Speciation*. Sunderland, MA.

- Crawshaw, L.I., 1975. Attainment of the final thermal preferendum in brown bullheads acclimated to different temperatures. *Comp. Biochem. Physiol. A Physiol.* 52, 171–173. [https://doi.org/10.1016/S0300-9629\(75\)80148-4](https://doi.org/10.1016/S0300-9629(75)80148-4)
- Cui, R., Schumer, M., Kruesi, K., Walter, R., Andolfatto, P., Rosenthal, G.G., 2013. Phylogenomics Reveals Extensive Reticulate Evolution in Xiphophorus Fishes. *Evolution* 67, 2166–2179. <https://doi.org/10.1111/evo.12099>
- Czypionka, T., Cheng, J., Pozhitkov, A., Nolte, A.W., 2012. Transcriptome changes after genome-wide admixture in invasive sculpins (*Cottus*). *Mol. Ecol.* 21, 4797–4810. <https://doi.org/10.1111/j.1365-294X.2012.05645.x>
- Czypionka, T., Goedbloed, D.J., Steinfartz, S., Nolte, A.W., 2018. Plasticity and evolutionary divergence in gene expression associated with alternative habitat use in larvae of the European Fire Salamander. *Mol. Ecol.* 0. <https://doi.org/10.1111/mec.14713>
- Czypionka, T., Krugman, T., Altmüller, J., Blaustein, L., Steinfartz, S., Templeton, A.R., Nolte, A.W., 2015. Ecological transcriptomics – a non-lethal sampling approach for endangered fire salamanders. *Methods Ecol. Evol.* n/a-n/a. <https://doi.org/10.1111/2041-210X.12431>
- Darwin, C., 1859. *On the Origin of the Species by Natural Selection*. Murray.
- De Nie, H.W., 1997. *Bedreigde en kwetsbare zoetwatervissen in Nederland*. Stichting Atlas Verspreiding Nederlandse Zoetwatervissen, Nieuwegein.
- De Queiroz, K., 2007. Species concepts and species delimitation. *Syst. Biol.* 56, 879–886. <https://doi.org/10.1080/10635150701701083>
- Dennenmoser Stefan, Sedlazeck Fritz J., Iwaszkiewicz Elzbieta, Li Xiang-Yi, Altmüller Janine, Nolte Arne W., 2017. Copy number increases of transposable elements and protein-coding genes in an invasive fish of hybrid origin. *Mol. Ecol.* 26, 4712–4724. <https://doi.org/10.1111/mec.14134>
- Derome, N., Duchesne, P., Bernatchez, L., 2006. Parallelism in gene transcription among sympatric lake whitefish (*Coregonus clupeaformis* Mitchill) ecotypes. *Mol. Ecol.* 15, 1239–1249. <https://doi.org/10.1111/j.1365-294X.2005.02968.x>
- Dieckmann, U., Doebeli, M., Metz, J.A.J., Tautz, D., 2004. *Adaptive Speciation*. Cambridge University Press.
- Dittrich-Reed, D.R., Fitzpatrick, B.M., 2013. Transgressive Hybrids as Hopeful Monsters. *Evol. Biol.* 40, 310–315. <https://doi.org/10.1007/s11692-012-9209-0>
- Dixon, P., 2009. VEGAN, a package of R functions for community ecology [WWW Document]. [Httpdxdoiorg1016581100-923320030140927VAPORF20CO2](http://dx.doi.org/10.1658/1100-923320030140927VAPORF20CO2). URL <http://www.bioone.org/doi/abs/10.1658/1100->

9233%282003%29014%5B0927%3AVAPORF%5D2.0.CO%3B2

(accessed 4.5.17).

- Dobzhansky, T., 1937. *Genetics and the Origin of Species*, Columbia University Biological Series. Columbia University Press.
- Dobzhansky, T., 1933. On the Sterility of the Interracial Hybrids in *Drosophila Pseudoobscura*. *Proc. Natl. Acad. Sci.* 19, 397–403.
- Doebeli, M., Dieckmann, U., 2003. Speciation along environmental gradients. *Nature* 421, 259–264. <https://doi.org/10.1038/nature01274>
- Donelson, J.M., Munday, P.L., McCormick, M.I., Pitcher, C.R., 2012. Rapid transgenerational acclimation of a tropical reef fish to climate change. *Nat. Clim. Change* 2, 30–32. <https://doi.org/10.1038/nclimate1323>
- Dopico, X.C., Evangelou, M., Ferreira, R.C., Guo, H., Pekalski, M.L., Smyth, D.J., Cooper, N., Burren, O.S., Fulford, A.J., Hennig, B.J., Prentice, A.M., Ziegler, A.-G., Bonifacio, E., Wallace, C., Todd, J.A., 2015. Widespread seasonal gene expression reveals annual differences in human immunity and physiology. *Nat. Commun.* 6, 7000. <https://doi.org/10.1038/ncomms8000>
- Englbrecht, C.C., Freyhof, J., Nolte, A., Rassmann, K., Schliewen, U., Tautz, D., 2000. Phylogeography of the bullhead *Cottus gobio* (Pisces: Teleostei: Cottidae) suggests a pre-Pleistocene origin of the major central European populations. *Mol. Ecol.* 9, 709–722. <https://doi.org/10.1046/j.1365-294x.2000.00912.x>
- Enright, J.M., Toomey, M.B., Sato, S., Temple, S.E., Allen, J.R., Fujiwara, R., Kramlinger, V.M., Nagy, L.D., Johnson, K.M., Xiao, Y., How, M.J., Johnson, S.L., Roberts, N.W., Kefalov, V.J., Guengerich, F.P., Corbo, J.C., 2015. Cyp27c1 Red-Shifts the Spectral Sensitivity of Photoreceptors by Converting Vitamin A1 into A2. *Curr. Biol.* 25, 3048–3057. <https://doi.org/10.1016/j.cub.2015.10.018>
- Estes, L., Elsen, P.R., Treuer, T., Ahmed, L., Caylor, K., Chang, J., Choi, J.J., Ellis, E.C., 2018. The spatial and temporal domains of modern ecology. *Nat. Ecol. Evol.* 1. <https://doi.org/10.1038/s41559-018-0524-4>
- Feliner, G.N., Álvarez, I., Fuertes-Aguilar, J., Heuertz, M., Marques, I., Moharrek, F., Piñeiro, R., Riina, R., Rosselló, J.A., Soltis, P.S., Villa-Machío, I., 2017. Is homoploid hybrid speciation that rare? An empiricist's view [WWW Document]. *Heredity*. <https://doi.org/10.1038/hdy.2017.7>
- Fisher, R., 1930. *The genetical theory of natural selection*. Oxford at the Clarendon Press.
- Gene Ontology Consortium, 2004. The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res.* 32, D258–D261. <https://doi.org/10.1093/nar/gkh036>

- Giller, P.S., Giller, P., Malmqvist, B., 1998. *The Biology of Streams and Rivers*. OUP Oxford.
- Goldschmidt, R., 1940. *The material basis of evolution*. New Haven: Yale Univ. Press.
- Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nat. Biotechnol.* 29, 644–652. <https://doi.org/10.1038/nbt.1883>
- Gracey, A.Y., Troll, J.V., Somero, G.N., 2001. Hypoxia-induced gene expression profiling in the euryoxic fish *Gillichthys mirabilis*. *Proc. Natl. Acad. Sci.* 98, 1993–1998. <https://doi.org/10.1073/pnas.98.4.1993>
- Gross, B.L., Rieseberg, L.H., 2005. The Ecological Genetics of Homoploid Hybrid Speciation. *J. Hered.* 96, 241–252. <https://doi.org/10.1093/jhered/esi026>
- Hermansen, J.S., Sæther, S.A., Elgvin, T.O., Borge, T., Hjelle, E., Sætre, G.-P., 2011. Hybrid speciation in sparrows I: phenotypic intermediacy, genetic admixture and barriers to gene flow. *Mol. Ecol.* 20, 3812–3822. <https://doi.org/10.1111/j.1365-294X.2011.05183.x>
- Houtgast, E.J., Sima, V.-M., Bertels, K., Al-Ars, Z., 2016. GPU-Accelerated BWA-MEM Genomic Mapping Algorithm Using Adaptive Load Balancing, in: *Architecture of Computing Systems – ARCS 2016, Lecture Notes in Computer Science*. Presented at the International Conference on Architecture of Computing Systems, Springer, Cham, pp. 130–142. https://doi.org/10.1007/978-3-319-30695-7_10
- Jenni, L., Kéry, M., 2003. Timing of autumn bird migration under climate change: advances in long-distance migrants, delays in short-distance migrants. *Proc. R. Soc. Lond. B Biol. Sci.* 270, 1467–1471. <https://doi.org/10.1098/rspb.2003.2394>
- Kagawa, K., Takimoto, G., 2018. Hybridization can promote adaptive radiation by means of transgressive segregation. *Ecol. Lett.* 21, 264–274. <https://doi.org/10.1111/ele.12891>
- Kang, J.H., Schartl, M., Walter, R.B., Meyer, A., 2013. Comprehensive phylogenetic analysis of all species of swordtails and platies (Pisces: Genus *Xiphophorus*) uncovers a hybrid origin of a swordtail fish, *Xiphophorus monticolus*, and demonstrates that the sexually selected sword originated in the ancestral lineage of the genus, but was lost again secondarily. *BMC Evol. Biol.* 13, 25. <https://doi.org/10.1186/1471-2148-13-25>

- Kawecki, T.J., Ebert, D., 2004. Conceptual issues in local adaptation. *Ecol. Lett.* 7, 1225–1241. <https://doi.org/10.1111/j.1461-0248.2004.00684.x>
- Keller, I., Seehausen, O., 2012. Thermal adaptation and ecological speciation. *Mol. Ecol.* 21, 782–799. <https://doi.org/10.1111/j.1365-294X.2011.05397.x>
- Keller, I., Wagner, C.E., Greuter, L., Mwaiko, S., Selz, O.M., Sivasundar, A., Wittwer, S., Seehausen, O., 2012. Population genomic signatures of divergent adaptation, gene flow and hybrid speciation in the rapid radiation of Lake Victoria cichlid fishes. *Mol. Ecol.* 22, 2848–2863. <https://doi.org/10.1111/mec.12083>
- Kelly, A.E., Goulden, M.L., 2008. Rapid shifts in plant distribution with recent climate change. *Proc. Natl. Acad. Sci.* 105, 11823–11826. <https://doi.org/10.1073/pnas.0802891105>
- King, M.C., Wilson, A.C., 1975. Evolution at two levels in humans and chimpanzees. *Science* 188, 107–116. <https://doi.org/10.1126/science.1090005>
- Lamichhaney, S., Berglund, J., Almén, M.S., Maqbool, K., Grabherr, M., Martinez-Barrio, A., Promerová, M., Rubin, C.-J., Wang, C., Zamani, N., Grant, B.R., Grant, P.R., Webster, M.T., Andersson, L., 2015. Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature* 518, 371–375. <https://doi.org/10.1038/nature14181>
- Lamichhaney, S., Han, F., Webster, M.T., Andersson, L., Grant, B.R., Grant, P.R., 2018. Rapid hybrid speciation in Darwin's finches. *Science* 359, 224–228. <https://doi.org/10.1126/science.aao4593>
- Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 1–21. <https://doi.org/10.1186/s13059-014-0550-8>
- Machado, H.E., Jui, G., Joyce, D.A., Reilly, C.R., Lunt, D.H., Renn, S.C., 2014. Gene duplication in an African cichlid adaptive radiation. *BMC Genomics* 15, 161. <https://doi.org/10.1186/1471-2164-15-161>
- Mallet, J., 2008. Hybridization, ecological races and the nature of species: empirical evidence for the ease of speciation. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 2971–2986. <https://doi.org/10.1098/rstb.2008.0081>
- Mallet, J., 2007. Hybrid speciation. *Nature* 446, 279–283. <https://doi.org/10.1038/nature05706>
- Mank, J.E., Carlson, J.E., Brittingham, M.C., 2004. A Century of Hybridization: Decreasing Genetic Distance Between American Black Ducks and Mallards. *Conserv. Genet.* 5, 395–403. <https://doi.org/10.1023/B:COGE.0000031139.55389.b1>

- Mavárez, J., Salazar, C.A., Bermingham, E., Salcedo, C., Jiggins, C.D., Linares, M., 2006. Speciation by hybridization in *Heliconius* butterflies. *Nature* 441, 868–871. <https://doi.org/10.1038/nature04738>
- Mayr, E., 1996. What Is a Species, and What Is Not? *Philos. Sci.* 63, 262–277. <https://doi.org/10.1086/289912>
- Mayr, E., 1963. *Animal Species and Evolution*. Harvard University Press.
- Mayr, E., 1942. *Systematics and the origin of species*. Columbia Univ. Press.
- McFarlane, S.E., Sirkiä, P.M., Ålund, M., Qvarnström, A., 2016. Hybrid Dysfunction Expressed as Elevated Metabolic Rate in Male *Ficedula* Flycatchers. *PLOS ONE* 11, e0161547. <https://doi.org/10.1371/journal.pone.0161547>
- Melo Maria C., Salazar Camilo, Jiggins Chris D., Linares Mauricio, 2009. Assortative mating preferences among hybrids offers a route to hybrid speciation. *Evolution* 63, 1660–1665. <https://doi.org/10.1111/j.1558-5646.2009.00633.x>
- Meyer, A., 1993. Phylogenetic relationships and evolutionary processes in East African cichlid fishes. *Trends Ecol. Evol.* 8, 279–284. [https://doi.org/10.1016/0169-5347\(93\)90255-N](https://doi.org/10.1016/0169-5347(93)90255-N)
- Nachtomy, O., Shavit, A., Yakhini, Z., 2007. Gene expression and the concept of the phenotype. *Stud. Hist. Philos. Sci. Part C Stud. Hist. Philos. Biol. Biomed. Sci.* 38, 238–254. <https://doi.org/10.1016/j.shpsc.2006.12.014>
- Nielsen, E.E., Hansen, M.M., Ruzzante, D.E., Meldrup, D., Grønkjær, P., 2003. Evidence of a hybrid-zone in Atlantic cod (*Gadus morhua*) in the Baltic and the Danish Belt Sea revealed by individual admixture analysis. *Mol. Ecol.* 12, 1497–1508. <https://doi.org/10.1046/j.1365-294X.2003.01819.x>
- Noga, E.J., 2011. *Fish Disease: Diagnosis and Treatment*. John Wiley & Sons.
- Nolte, A.W., Freyhof, J., Stemshorn, K.C., Tautz, D., 2005a. An invasive lineage of sculpins, *Cottus* sp. (Pisces, Teleostei) in the Rhine with new habitat adaptations has originated from hybridization between old phylogeographic groups. *Proc. R. Soc. Lond. B Biol. Sci.* 272, 2379–2387. <https://doi.org/10.1098/rspb.2005.3231>
- Nolte, A.W., Freyhof, J., Tautz, D., 2006. When invaders meet locally adapted types: rapid moulding of hybrid zones between sculpins (*Cottus*, Pisces) in the Rhine system. *Mol. Ecol.* 15, 1983–1993. <https://doi.org/10.1111/j.1365-294X.2006.02906.x>
- Nolte, A. W., Gompert, Z., Buerkle, C.A., 2009. Variable patterns of introgression in two sculpin hybrid zones suggest that genomic isolation differs among populations. *Mol. Ecol.* 18, 2615–2627. <https://doi.org/10.1111/j.1365-294X.2009.04208.x>
- Nolte, Arne W., Renaut, S., Bernatchez, L., 2009. Divergence in gene regulation at young life history stages of whitefish (*Coregonus* sp.) and

- the emergence of genomic isolation. *BMC Evol. Biol.* 9, 59. <https://doi.org/10.1186/1471-2148-9-59>
- Nolte, A.W., Sheets, H.D., 2005. Shape based assignment tests suggest transgressive phenotypes in natural sculpin hybrids (Teleostei, Scorpaeniformes, Cottidae). *Front. Zool.* 2, 11. <https://doi.org/10.1186/1742-9994-2-11>
- Nolte, A.W., Stemshorn, K.C., Tautz, D., 2005b. Direct cloning of microsatellite loci from *Cottus gobio* through a simplified enrichment procedure. *Mol. Ecol. Notes* 5, 628–636. <https://doi.org/10.1111/j.1471-8286.2005.01026.x>
- Nolte, A.W., Tautz, D., 2010. Understanding the onset of hybrid speciation. *Trends Genet.* 26, 54–58. <https://doi.org/10.1016/j.tig.2009.12.001>
- Nosil, P., 2012. *Ecological Speciation*. Oxford University Press.
- Nueda, M.J., Tarazona, S., Conesa, A., 2014. Next maSigPro: updating maSigPro Bioconductor package for RNA-seq time series. *Bioinformatics* btu333. <https://doi.org/10.1093/bioinformatics/btu333>
- Orczewska, J.I., Hartleben, G., O'Brien, K.M., 2010. The molecular basis of aerobic metabolic remodeling differs between oxidative muscle and liver of threespine sticklebacks in response to cold acclimation. *Am. J. Physiol.-Regul. Integr. Comp. Physiol.* 299, R352–R364. <https://doi.org/10.1152/ajpregu.00189.2010>
- Parsons, K.J., Son, Y.H., Albertson, R.C., 2011. Hybridization Promotes Evolvability in African Cichlids: Connections Between Transgressive Segregation and Phenotypic Integration. *Evol. Biol.* 38, 306–315. <https://doi.org/10.1007/s11692-011-9126-7>
- Payseur, B.A., Rieseberg, L.H., 2016. A genomic perspective on hybridization and speciation. *Mol. Ecol.* 25, 2337–2360. <https://doi.org/10.1111/mec.13557>
- Pereyra, R.T., Bergström, L., Kautsky, L., Johannesson, K., 2009. Rapid speciation in a newly opened postglacial marine environment, the Baltic Sea. *BMC Evol. Biol.* 9, 70. <https://doi.org/10.1186/1471-2148-9-70>
- Perry, A.L., Low, P.J., Ellis, J.R., Reynolds, J.D., 2005. Climate Change and Distribution Shifts in Marine Fishes. *Science* 308, 1912–1915. <https://doi.org/10.1126/science.1111322>
- Pörtner, H.O., Knust, R., 2007. Climate Change Affects Marine Fishes Through the Oxygen Limitation of Thermal Tolerance. *Science* 315, 95–97. <https://doi.org/10.1126/science.1135471>
- Pouliot, Y., Gao, J., Su, Q.J., Liu, G.G., Ling, X.B., 2001. DIAN: A Novel Algorithm for Genome Ontological Classification. *Genome Res.* 11, 1766–1779. <https://doi.org/10.1101/gr.183301>

- Price, T.D., Qvarnström, A., Irwin, D.E., 2003. The role of phenotypic plasticity in driving genetic evolution. *Proc. R. Soc. B Biol. Sci.* 270, 1433–1440. <https://doi.org/10.1098/rspb.2003.2372>
- R Core Team, 2017. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rice, W.R., 1987. Speciation via habitat specialization: the evolution of reproductive isolation as a correlated character. *Evol. Ecol.* 1, 301–314. <https://doi.org/10.1007/BF02071555>
- Richards, C.L., Rosas, U., Banta, J., Bhambhra, N., Purugganan, M.D., 2012. Genome-Wide Patterns of Arabidopsis Gene Expression in Nature. *PLOS Genet.* 8, e1002662. <https://doi.org/10.1371/journal.pgen.1002662>
- Rieseberg, L.H., Archer, M.A., Wayne, R.K., 1999. Transgressive segregation, adaptation and speciation. *Heredity* 83, 363–372. <https://doi.org/10.1038/sj.hdy.6886170>
- Rieseberg, L.H., Raymond, O., Rosenthal, D.M., Lai, Z., Livingstone, K., Nakazato, T., Durphy, J.L., Schwarzbach, A.E., Donovan, L.A., Lexer, C., 2003. Major Ecological Transitions in Wild Sunflowers Facilitated by Hybridization. *Science* 301, 1211–1216. <https://doi.org/10.1126/science.1086949>
- Rieseberg, L.H., Sinervo, B., Linder, C.R., Ungerer, M.C., Arias, D.M., 1996. Role of gene interactions in hybrid speciation: Evidence from ancient and experimental hybrids. *Science* 272, 741.
- Romero, I.G., Ruvinsky, I., Gilad, Y., 2012. Comparative studies of gene expression and the evolution of gene regulation. *Nat. Rev. Genet.* 13, 505–516. <https://doi.org/10.1038/nrg3229>
- Rüber, L., Van Tassell, J.L., Zardoya, R., Karl, S., 2003. Rapid speciation and ecological divergence in the american seven-spined gobies (gobiidae, gobiosomatini) inferred from a molecular phylogeny. *Evolution* 57, 1584–1598. <https://doi.org/10.1554/02-668>
- Salazar, C., Baxter, S.W., Pardo-Diaz, C., Wu, G., SurrIDGE, A., Linares, M., Bermingham, E., Jiggins, C.D., 2010. Genetic Evidence for Hybrid Trait Speciation in *Heliconius* Butterflies. *PLOS Genet.* 6, e1000930. <https://doi.org/10.1371/journal.pgen.1000930>
- Scavetta, R.J., Tautz, D., 2010. Copy Number Changes of CNV Regions in Interspecific Crosses of the House Mouse. *Mol. Biol. Evol.* 27, 1845–1856. <https://doi.org/10.1093/molbev/msq064>
- Schmidt-Nielsen, K., 1997. *Animal Physiology: Adaptation and Environment*. Cambridge University Press.
- Schönborn, W., 1992. *Fließgewasserbiologie*. Gustav Fischer Verlag Jena, Stuttgart.

- Schumer, M., Rosenthal, G.G., Andolfatto, P., 2018. What do we mean when we talk about hybrid speciation? *Heredity* 120, 379–382. <https://doi.org/10.1038/s41437-017-0036-z>
- Schumer Molly, Rosenthal Gil G., Andolfatto Peter, 2014. How common is homoploid hybrid speciation? *Evolution* 68, 1553–1560. <https://doi.org/10.1111/evo.12399>
- Seehausen, O., 2013. Conditions when hybridization might predispose populations for adaptive radiation. *J. Evol. Biol.* 26, 279–281. <https://doi.org/10.1111/jeb.12026>
- Seehausen, O., 2004. Hybridization and adaptive radiation. *Trends Ecol. Evol.* 19, 198–207. <https://doi.org/10.1016/j.tree.2004.01.003>
- Sharma, E., Künstner, A., Fraser, B.A., Zipprich, G., Kottler, V.A., Henz, S.R., Weigel, D., Dreyer, C., 2014. Transcriptome assemblies for studying sex-biased gene expression in the guppy, *Poecilia reticulata*. *BMC Genomics* 15, 400. <https://doi.org/10.1186/1471-2164-15-400>
- Sheets, H.D., 2014. CoordGen8. Dept. of Physics, Canisius College, Buffalo, NY 14208.
- Sheets, H.D., 2000. Integrated Morphometrics Package (IMP).
- Slominski, A., Fischer, T.W., Zmijewski, M.A., Wortsman, J., Semak, I., Zbytek, B., Slominski, R.M., Tobin, D.J., 2005. On the role of melatonin in skin physiology and pathology. *Endocrine* 27, 137–147. <https://doi.org/10.1385/ENDO:27:2:137>
- Stebbins, G.L., 1985. Polyploidy, Hybridization, and the Invasion of New Habitats. *Ann. Mo. Bot. Gard.* 72, 824–832. <https://doi.org/10.2307/2399224>
- Stebbins, G.L., 1959. The Role of Hybridization in Evolution. *Proc. Am. Philos. Soc.* 103, 231–251.
- Stebbins, G.L., 1958. The Inviability, Weakness, and Sterility of Interspecific Hybrids, in: Demerec, M. (Ed.), *Advances in Genetics*. Academic Press, pp. 147–215. [https://doi.org/10.1016/S0065-2660\(08\)60162-5](https://doi.org/10.1016/S0065-2660(08)60162-5)
- Stelkens, R., Seehausen, O., 2009. Genetic Distance Between Species Predicts Novel Trait Expression in Their Hybrids. *Evolution* 63, 884–897. <https://doi.org/10.1111/j.1558-5646.2008.00599.x>
- Stemshorn, K.C., Reed, F.A., Nolte, A.W., Tautz, D., 2011. Rapid formation of distinct hybrid lineages after secondary contact of two fish species (*Cottus* sp.). *Mol. Ecol.* 20, 1475–1491. <https://doi.org/10.1111/j.1365-294X.2010.04997.x>
- Tirosh, I., Reikhav, S., Levy, A.A., Barkai, N., 2009. A Yeast Hybrid Provides Insight into the Evolution of Gene Expression Regulation. *Science* 324, 659–662. <https://doi.org/10.1126/science.1169766>

- Todd, E.V., Black, M.A., Gemmell, N.J., 2016. The power and promise of RNA-seq in ecology and evolution. *Mol. Ecol.* 25, 1224–1241. <https://doi.org/10.1111/mec.13526>
- Tøttrup, A.P., Rainio, K., Coppack, T., Lehikoinen, E., Rahbek, C., Thorup, K., 2010. Local Temperature Fine-Tunes the Timing of Spring Migration in Birds. *Integr. Comp. Biol.* 50, 293–304. <https://doi.org/10.1093/icb/icq028>
- Trier, C.N., Hermansen, J.S., Sætre, G.-P., Bailey, R.I., 2014. Evidence for Mitochondrial and Sex-Linked Reproductive Barriers between the Hybrid Italian Sparrow and Its Parent Species. *PLOS Genet* 10, e1004075. <https://doi.org/10.1371/journal.pgen.1004075>
- Villemereuil, P. de, Gaggiotti, O.E., Mouterde, M., Till-Bottraud, I., 2016. Common garden experiments in the genomic era: new perspectives and opportunities. *Heredity* 116, 249–254. <https://doi.org/10.1038/hdy.2015.93>
- Walther, G.-R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T.J.C., Fromentin, J.-M., Hoegh-Guldberg, O., Bairlein, F., 2002. Ecological responses to recent climate change. *Nature* 416, 389–395. <https://doi.org/10.1038/416389a>
- Wiley, E.O., 1978. The Evolutionary Species Concept Reconsidered. *Syst. Biol.* 27, 17–26. <https://doi.org/10.2307/2412809>
- Wittkopp, P.J., 2013. Evolution of Gene Expression. pp. 413–419.
- Wolf, J.B.W., Lindell, J., Backström, N., 2010. Speciation genetics: current status and evolving approaches. *Philos. Trans. R. Soc. B Biol. Sci.* 365, 1717–1733. <https://doi.org/10.1098/rstb.2010.0023>
- Woodward, F.I., 1988. Temperature and the distribution of plant species. *Symp. Soc. Exp. Biol.* 42, 59–75.
- Xie, H., Wasserman, A., Levine, Z., Novik, A., Grebinskiy, V., Shoshan, A., Mintz, L., 2002. Large-Scale Protein Annotation through Gene Ontology. *Genome Res.* 12, 785–794. <https://doi.org/10.1101/gr.86902>
- Yakimowski, S.B., Rieseberg, L.H., 2014. The role of homoploid hybridization in evolution: A century of studies synthesizing genetics and ecology. *Am. J. Bot.* 101, 1247–1258. <https://doi.org/10.3732/ajb.1400201>
- Yates, A., Akanni, W., Amode, M.R., Barrell, D., Billis, K., Carvalho-Silva, D., Cummins, C., Clapham, P., Fitzgerald, S., Gil, L., Girón, C.G., Gordon, L., Hourlier, T., Hunt, S.E., Janacek, S.H., Johnson, N., Juettemann, T., Keenan, S., Lavidas, I., Martin, F.J., Maurel, T., McLaren, W., Murphy, D.N., Nag, R., Nuhn, M., Parker, A., Patricio, M., Pignatelli, M., Rahtz, M., Riat, H.S., Sheppard, D., Taylor, K., Thormann, A., Vullo, A., Wilder, S.P., Zadissa, A., Birney, E., Harrow, J., Muffato, M., Perry, E., Ruffier, M., Spudich, G., Trevanion, S.J., Cunningham, F., Aken, B.L., Zerbino, D.R.,

Flicek, P., 2016. Ensembl 2016. *Nucleic Acids Res.* 44, D710–D716.
<https://doi.org/10.1093/nar/gkv1157>

Zelditch, M.L., Swiderski, D.L., Sheets, H.D., 2012. *Geometric Morphometrics for Biologists: A Primer*. Academic Press.

Digital Supplement

Description of data:

Chapter 1

- File S1.1. Table of counts for the temperature choice experiment.
(DOI: <http://doi.org/10.5281/zenodo.2535621>)
- File S1.2. Morphometric study; Collection of photos of fish before the incubation. (DOI: <http://doi.org/10.5281/zenodo.2535536>)
- File S1.3. Morphometric study; Collection of photos of fish after the incubation. (DOI: <http://doi.org/10.5281/zenodo.2535536>)

Chapter 2

File S2.1. Table with the results of differential gene expression in the controlled environment for liver tissues. (DOI: <https://doi.org/10.5281/zenodo.2535631>)

File S2.2. Table with the results of differential gene expression in the controlled environment for fin tissues. (DOI: <https://doi.org/10.5281/zenodo.2535640>)

File S2.3. Table with the results of the time series analysis for livers.
(DOI: <https://doi.org/10.5281/zenodo.2535642>)

File S2.4. Table with the results of the time series analysis for fins.
(DOI: <https://doi.org/10.5281/zenodo.2535646>)

Chapter 3

File S3.1. Collection of graphs presenting gene expression for transgressively up-regulated genes in fins. (DOI: <https://doi.org/10.5281/zenodo.2535648>)

File S3.2. Collection of graphs presenting gene expression for transgressively down-regulated genes in fins. (DOI: <https://doi.org/10.5281/zenodo.2535650>)

File S3.3. Collection of graphs presenting gene expression for transgressively up-regulated genes in livers. (DOI: <https://doi.org/10.5281/zenodo.2535652>)

File S3.4. Collection of graphs presenting gene expression for transgressively down-regulated genes in livers.
(DOI: <https://doi.org/10.5281/zenodo.2535654>)

Declaration of Contributions

The design of the studies comprised in this dissertation was developed together with Prof. Arne W. Nolte. Specific contributions to this work are as follows:

Chapter 1

Fish used in the temperature preference experiment and the growth experiment were bred in laboratory facilities by Arne W. Nolte and Ralf Schmuck in spring 2014, before I joined the group. The preliminary temperature choice experiment was set up and performed by summer inter Jasmin Dehnen under supervision of Arne W. Nolte. In the main experiment I set up the experimental chamber and controlled the conditions. I performed all the observations of the fish. For the growth experiment I set up the chamber and controlled the conditions. During the experiment I received help from Ezgi Ozkurt in performing temperature measurements and to feeding the fish. Pictures of fish were taken by Ezgi Ozkurt and myself. I developed and installed the set-up for photography.

Chapter 2 and Chapter 3:

Fish from the controlled environment were bred and maintained in the laboratory by Arne W. Nolte and Ralf Schmuck. Elke Bustorf and I scarified the fish. I sampled tissues and sexed all samples. I extracted RNA and performed quality checks for samples from the "summer" batch while Simone Lenci did extracted RNA for the "winter" batch. Fish from the natural environment were caught and sacrificed by Arne W. Nolte, Fabian Herder (Zoologisches Forschungsmuseum Alexander Koenig, Bonn) and myself. I carried out

dissections, length measurements, RNA extractions and quality checks. Library preparations and RNA sequencing were done by the Cologne Center for Genomics (Cologne). I performed sequence processing and data analyses. Arne W. Nolte supplied a perl script for calculating the genetic distance tree. I wrote both manuscripts with contributions from Arne W. Nolte.

Acknowledgements

First and foremost I want to thank Prof. Arne Nolte for believing in me and letting me work on his precious study system. He transferred all his knowledge about *Cottus* fish (as well as many other fish) during very entertaining breakfasts, dinners, beer-hours and field trips. It was a pleasure to be part of this group. Despite my lack of background in bioinformatics Arne trusted that I can make the transition to new transcriptomics technology in his lab happen and I am grateful for that. His passion for fishing (*Cottus* and else) is impossible to overestimate and he has a special talent to talk about it with so much passion that it becomes contagious.

I am deeply grateful to Prof. Diethard Tautz for taking over supervision of my thesis and enabling me to remain in the IMPRS after Arne was appointed professorship at the University of Oldenburg. I am very thankful for the time he took to discuss my doubts and issues. His support and advice were invaluable help for me especially in the last phase of my PhD.

I would also like to thank for helpful discussions and time to the other members of my thesis committee: Dr. Oscar Puebla and Prof. Lutz Becks. Lutz Becks was particularly solicitous and always equipped with helpful advice about statistics, even when approached out of the blue on the train home late after work. I appreciate that a lot.

This work would not have been possible without the support from colleagues, authorities and fisheries societies that enables fieldwork at the River Sieg. Wilhelm Kreuzmann representing the Sieg-Fischerei Genossenschaft supported our fieldwork and I am grateful for the necessary permits to conduct the research given by K. Lessenich and A. Mellin (Bezirksregierung Köln). Fabian Herder (Museum König Bonn) has volunteered on several occasions to help me to collect necessary samples. *Cottus* were kept

and bred in the laboratory of the MPI in Plön with permission from S. Hauschildt (Veterinäeramt in Plön).

I would like to thank a lot the Cottus group: Sunna Ellendt, Stefan Dennenmoser, Jie Cheng, Till Czypionka, Guénolé Le Pennec, Thijs Janzen, Kristin Tietje. You provided me with good discussions and a really good, collegial atmosphere. I definitely felt I could always count on you and that was very reassuring. I also thank a lot the invaluable technicians Malte Dittmann, Elke Bustorf and Elke Blohm-Sievers for keeping lab-stuff in perfect order. And of course huge thanks to the fish caretaker Ralf Schmuck for taking good care of my study subjects.

I am also very grateful to the Behavioral Genomics (Miriam Liedvogel) and Biological Clocks (Tobias Kaiser) groups for adopting me to their bunch once my group moved to Oldenburg. I profited greatly from listening and participating in their joint lab seminars. Thanks for letting me talk about my research and learn about yours – it was a great experience.

I would also place big thank you to the Meiotic Recombination group who embraced me in their office: Alina Jeschke, Elena Kovalchuk, Khawla Abualia. Thanks for good times, countless nice conversations and laughs. I am however specially indebted to the group leader Linda Odenthal-Hesse who made my transition from one-of-the-Cottus-group-members to the-last-Cottus-person-in-Plön easy and painless.

During a PhD time in a small city like Plön many of your colleagues will become friends and they are crucial to survival. Huge thanks for all the scientific and non-scientific discussions and just great company to: Luisa Pallares, Rafik Neme, Jatin Arora, Chen Xie, Çemali Bekpen, Joana Bernardes, Neva Skrabar and many many more. Thanks for great company in the office and countless lunches to Alina Jeschke. To Maryam Keshavarz and Zahra Khomarbaghi for adopting me to their Persian circle. Noémie Erin, Stefan

Dennenmoser and Primrose Boynton helped me a lot with improving this thesis and I am very grateful for that.

Additional huge thanks goes to Noémie Erin who was an amazing companion in the cultural excursions throughout my whole PhD time and contributed greatly to my wellbeing.

I would like to thank as well to my Polish base-camp support team: Paulina Włostowska, Natalia Kobryń, Ania Siekierska and Marta Prus. They are always there for me and when I was in doubt they reminded me why I started this PhD in the first place.

Huge thanks to my family! First of all my parents, Barbara and Maciej Iwaszkiewicz, thanks for always believing in me and encouraging me to go forward. Without your support none of that could have happened. My brother Jan and his fantastic family were always a safe (but never quiet) bay where I could rest. Thank you for that support.

Last but not the least I would like to thank my husband Wolfram Eggebrecht for immense love, support and what seems to be an absolute and bottomless patience. Not only did he endure me throughout the ups and downs of PhD time but even decided to marry me and went through with this plan! Truly fearless. Thank you for all.

Affidavit

Hereby I declare that

- i) apart from my supervisor's guidance, the content and design of this dissertation is the product of my own work. The co-author's contributions are listed in the dedicated section;
- ii) this thesis has not already been submitted either partially or wholly as part of a doctoral degree to another examination body, and no other materials are published or submitted for publication than indicated in the thesis;
- iii) the preparation of the thesis has been subjected to the Rules of Good Scientific Practice of the German Research Foundation.

Plön, 17 July 2018

Elżbieta Anna Iwaszkiewicz-Eggebrecht

Curriculum vitae

ELŻBIETA ANNA IWASZKIEWICZ-EGGEBRECHT

Date and place of birth: 25/01/1987, Warsaw, Poland

Nationality: Polish

E-mail: elzbieta.iwaszkeiwicz@yahoo.fr

Address: Lübecker Str. 8, 24306 Plön

SCHOOL EDUCATION

1994-2000	Primary School, Warsaw, Poland
2000-2003	Stefania Sempołowska Middle School, Warsaw, Poland
2003-2006	Mikołaj Kopernik High School, Warsaw, Poland – Abitur (Matura)

HIGHER EDUCATION

2006-2007	Interdisciplinary Studies in Mathematics and Natural Sciences, University of Warsaw, Poland - main topic Psychology, secondary Biology
2007-2008	Oceanography with French, University of Southampton, United Kingdom
2008-2011	Interdisciplinary Studies in Mathematics and Natural Sciences, University of Warsaw, Poland – B.Sc. degree in Biology
2011-2012	Independent work on a research project in population genetics, Polish Academy of Sciences, Warsaw, Poland
2012-2014	Oceanography and Marine Environment master programme, Pierre and Marie Curie University (Paris VI, Sorbonne), Paris, France M. Sc. degree in Oceanography
2015-2018	International Max Planck School for Evolutionary Biology, Germany PhD candidate at the Christian-Albrecht-University of Kiel, Germany

