Panmixia under thermal stress: the case of the worm pipefish

Lurdes Patrícia Neto da Silva Gomes
Master in Biodiversity, Genetics and Evolution
Department of Biology
Faculty of Sciences of the University of Porto
2024

Panmixia under thermal stress: the case of the worm pipefish

Lurdes Patrícia Neto da Silva Gomes
Dissertation carried out as part of the Master in Biodiversity, Genetics and Evolution
Department of Biology
2024

Supervisor
Nuno Monteiro, Researcher at BIOPOLIS, CIBIO-InBIO

Co-supervisor
Olivia Roth, Professor at the Zoological Institute, Kiel University

Dedicated to all the supporters.

Acknowledgements

I would like to thank all the supporters, my family, my boyfriend, my friends, my psychologist, and my supervisors. This whole year, this thesis and this experience, have been very fulfilling and I'm proud of myself to be able to get through it, even despite all stressful times, hard picky analysis, (sort of) coding and self-learning. In the end, it was an experience, and it helped me develop different levels of my persona.

I wanted to start my million thank you with my base, that is my family. I am thankful for all the support of them on my journey and to allow me reaching this higher degree of education. Despite no one understands what I'm doing on my thesis or what I will be doing in the future, my family proudly brags me as the scientist of the family and listens (not understanding anything) all little details about the research I am performing. Without their support and love, I would not be where I am today. I am thankful to my boyfriend who has always been my number one fan, listening to my presentations, giving critical aesthetic advice so I could have pretty graphs and dealing with all my thesis related anxiety, always patiently saying "you can do this". He is one of a kind, a wonderful person and the best I wanted by my side. I am thankful to my friends, the ones with I laugh the most and have the nicest times to get away from reality. Despite we are all in different work areas, they were always ready to support and help me from computer problems to emotional ones. Since I cannot mention all of my incredible friends, I will refer some that really impacted me through the thesis: Raquel, the partner on the long travels, our conversations on the end of a tiring day at CIBIO, were the joy of my many dark days; Inês, the house stealer, she would come on holiday to my home as it was an hotel and wake up at 6 A.M., in order to spend time with me at CIBIO while I was working; Claudia, the supportive, she was always a good listener to my many complains and also a good advisor with her incredible comebacks; Sofia, the care giver, she was always prepared to hug me during difficult times, give me a kiss and pet my head as a mother; Sandiares, my love, she would always make me laugh so much with her jokes and porky laugh; Leonor, the red haired twin, she is an incredible person and in who I got a lot of expiration to become a better me; and lastly, Alexandre, the bestie, he would always answer my panic messages about the million computer problems, especially with the Mac and its uncooperativeness, and then spend one hour gossiping so we could catch each other's life. To the unmentioned ones, a big thank you for being my friends and also my second family. I am thankful to my psychologist, Dr. Susana, for being so warm and patient with me. She was a crucial person in this thesis, as science is a challenging path and gets more difficult when you are under a strong burnout. She would always advise me on how

I could overcome the difficult times and remember me about my capacities. Lastly, I am thankful to my supervisors which are incredible researchers and persons. I would like to start with Olivia and her team, whose received me so well in their group and house (in the case of Olivia). It was such a nice experience working on Germany and learning so much. I would like to individually thank Arseny, who thought me a lot about transcriptomics and patiently answered my many questions. Lastly, I wanted to thank my supervisor, Nuno, who was one of the few who recognized my effort and potential. Through all my thesis, he was always supportive, but at the same time, he gave me freedom to do my work and proudly call it a "*desmame*" from bachelor. On all our meetings he would always tell me an important lesson/advice from his experience through a story time, so I could overcome my work struggles. Despite the cold face he has, deep in, he is a caring and sweet hearted person, who emphasized I was now a scientist colleague of him and not his student. Mine and other close friends' thesis experiences, lead me to conclude that supervisors really make a difference on this journey, and sometimes, don't even understand the huge impact they have on their students and how much they have helped them in so many levels besides the academic work. In the end, I am glad I have had the opportunity to work with humanized researchers, that understood that I was a student, a colleague and a human which has a personality, feelings, dreams, struggles and insecurities.

As a final shot, I want to dedicate a paragraph to my cats, that accept me in their home, living rent free. These little creatures (Tonico, Banico and Emi) are the joy of my days, and usually, my main supporters during hard times, seeing me crying, with angry rages and complaining about everything. These cats have been with me since nineth grade (Tonico and Banico) and bachelor (Emi) and they are experts in biology, if we account the many classes they attended with me during lockdown, when they carelessly were my audience during presentations training and how I engaged in brainstorms with them. For that, they deserve a diploma. But despite being my four-leg co-workers, they are my emotional supporters and have done so much for me, besides any human could have done, only being there, with a cute little face, demanding for food and attention.

Resumo

No passado foi assumido que as populações marinhas não eram geneticamente estruturadas. Contudo, esta suposição foi refutada, e a bibliografia tem demonstrado que fatores como a ecologia, demografia e biogeografia, têm um importante papel no processo de diferenciação. As alterações climáticas têm influenciado estas dinâmicas e os seus efeitos estão a agravar-se impactando negativamente os ecossistemas marinhos, e os singnatídeos (cavalos-marinhos, marinhas e dragões do mar) não são exceção. Embora vários estudos laboratoriais já demonstrem o impacto negativo das altas temperaturas nestes peixes, poucos foram realizados em populações naturais. Para conduzir este tipo de estudos, *Nerophis lumbriciformis* é um ótimo candidato, já que experiencia diferentes regimes de temperatura ao longo da sua distribuição e existe um basto reportório de estudos relacionados com a sua reprodução. Além disso, Monteiro et al. (2017a) sugeriram uma maior pressão da seleção sexual nos limites da distribuição, onde as fêmeas, com ornamentações faciais expandidas, produzem mais e maiores ovos. Tendo em conta a variação no clima e a pressão diferencial da seleção sexual ao longo da distribuição de *N. lumbriciformis*, era esperado pelo menos algum nível de diferenciação entre populações, o que foi rejeitado por Mendes et al. (2020). Contudo, sinais de diferenciação podem estar presentes, mas restritos a zonas específicas do genoma, só detetadas utilizando métodos de alta resolução, como a análise genómica. Assim, nesta tese, tivemos como primeiro objetivo investigar os níveis de estruturação populacional de *N. lumbriciformis* utilizando a genómica, tentando depois avaliar, ao nível da expressão génica, o impacto de regimes contrastantes de temperatura junto aos limites da distribuição, utilizando a transcriptómica. Os resultados obtidos demonstraram um cenário de panmixia, onde aparentemente nenhum gene é positivamente selecionado em associação com o clima. A análise transcriptómica, contudo, identificou diferenças entre os limites da distribuição, onde se detetaram genes associados a choques de frio, no Norte, e de calor, no Sul. Diferentes padrões espaciais de expressão génica sobre um genoma homogéneo fazem-nos suspeitar que a variação fenotípica observada ao longo da latitude (e.g., maior investimento das fêmeas na reprodução nos limites da distribuição ou desenvolvimento de carateres sexuais secundários) poderá não resultar de um processo de seleção sexual, mas antes traduza a amplitude da plasticidade fenotípica induzida pelas condições stressantes (temperaturas baixas e altas) junto aos limites da distribuição. Os nossos resultados também sugerem que as cada vez mais altas temperaturas, decorrentes das alterações climáticas, já estão a afetar as populações do Sul da Europa, onde mecanismos de

resposta a choques de calor e genes apoptóticos estão a ser altamente expressos, insinuando a gradual aproximação aos limites fisiológicos da espécie. Mais estudos serão necessários para confirmar a ausência de estruturação ao longo da distribuição do *N. lumbriciformis*, nomeadamente através da inclusão de populações do centro da distribuição, o que permitirá também perceber se, junto ao hipotético ótimo térmico, existem sinais de condições stressantes.

Palavras-chave: Syngnathidae, Genómica, Transcriptómica, Panmixia, Alterações Climáticas, Stress Térmico.

Abstract

Population structure has been assumed in the past to be absent on marine environments. However, this assumption was debunked, and literature demonstrated that ecology, demography and biogeographic factors have an important role on the differentiation process. Climate change impacts these marine dynamics, and its effects are worsening, jeopardizing ecosystems. Syngnathids (seahorses, pipefishes and seadragons) are no exception, while several laboratory studies have shown the negative effects of high temperatures on these fishes, few studies were conducted on wild populations. To conduct these kind of studies, the worm pipefish, *Nerophis lumbriciformis*, can be viewed as an interesting model, since it experiences distinct climatic regimes in its distribution range, and a vast repertoire of reproduction-related studies already exists. Also, Monteiro et al. (2017a) hypothesized higher sexual selection near the range limits, where females, with expanded facial ornamentations, produce more and larger eggs. Taking the opposite climatic regimes and differential sexual selection along pipefish's range into account, some degree of population differentiation was hypothesized and quickly rejected by Mendes et al. (2020). Differentiation can be present, however reduced to specific areas of the genome whose detection requires high-resolution methodologies, such as genomics. Accordingly, in this thesis, we aimed to assess the existence of structure in worm pipefish populations, using a genomics approach, and then investigate how the temperature regimes at the range limits are impacting gene expression, using a transcriptomics approach. Genomic results showcased a panmictic scenario with no apparent positive selected genes associated with climate. However, transcriptomics indicated differences between range limits, where expression changes in genes related to cold shock, in the North, and heat shock, on the South, were identified. Different spatial patterns of gene expression over a homogeneous genome make us suspect that the phenotypic variation observed across latitude (e.g., greater female investment in reproduction, or development of secondary sexual characters at the limits of distribution) may not result from a process of sexual selection, but rather translate the amplitude of phenotypic plasticity induced by stressful conditions (low and high temperatures) along the limits of the distribution. Our results also suggest that increasingly high temperatures, resulting from climate change, are already affecting populations in Southern Europe, where heat shock response mechanisms and apoptotic genes are being highly expressed, hinting at the gradual approach to the physiological limits of the species. More studies will be necessary to confirm the lack of structure throughout the distribution of *N. lumbriciformis*, namely through the inclusion of

populations from the center of the distribution, which will allow us to understand if signs of stressful conditions are also apparent close to the species' hypothetical thermal optimum.

Keywords: Syngnathidae, Genomics, Transcriptomics, Panmixia, Climate Change, Thermal Stress.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

M	MALE
F	FEMALE
NO	NORWAY
SC	SCOTLAND
SP	SPAIN
PT	PORTUGAL
DNA	DEOXYRIBONUCLEIC ACID
RNA	RIBONUCLEIC ACID
SNP	SINGLE NUCLEOTIDE POLYMORPHISM
MAF	MINOR ALLELE FREQUENCY
LD	LINKAGE DISEQUILIBRIUM
ROH	RUNS OF HOMOZYGOSITY
HO	OBSERVED HETEROZYGOSITY
UHE	UNBIASED EXPECTED HETEROZYGOSITY
FIS	INBREEDING COEFFICIENT
PA	PRIVATE ALLELES
Π	NUCLEOTIDE DIVERISTY
FST	FIXATION INDEX
PCA	PRINCIPAL COMPONENT ANALYSIS
SFS	SITE FREQUENCY SPECTRUM
MDS	MULTIDIMENSIONAL SCALLING
ANOSIM	ANALYSIS OF SIMILARITY
DEG	DIFFERENTIALLY EXPRESSED GENE
ID	IDENTIFICATION
NAD	NORTH ATLANTIC DRIFT
NC	NORWEGIAN CURRENT
CC	CANARY CURRENT

1. Introduction

Literature relating to differentiation in marine environments is an extensive list of assumptions and exceptions. Ocean presents as a wide-open space, where species are free to connect with farther conspecifics. Thus, marine organisms are assumed to have low population structure due to high levels of gene flow, enhanced by factors such as pelagic larvae phase (Braga-Goncalves et al., 2017) or passive drafting (Bertola et al., 2020). However, the generalized assumption has been proven to have a lot of exceptions. Several barriers to gene flow have been identified in marine environments, which isolate populations, leading to differentiation and culminating on speciation (Palumbi, 1994). Some of the identified barriers are larvae behavior, demographic history, different selection pressures, habitat patchiness, distance between populations, and biogeographic factors such as currents and temperature (Cowen et al., 2006; Palumbi, 1994; Schiebelhut & Dawson, 2018). With current climate change, population dynamics will be affected (Kendall et al., 2016) threatening marine organisms (Venegas et al., 2023). Syngnathids (seahorses, pipefishes and seadragons) are no exception, with simulations based on current climatic trends suggesting that, within this century, all autochthonous species might vanish from the South of Europe as a result of episodes of local extinctions and poleward migrations (Monteiro et al., 2023) (Figure 1).

For syngnathids, temperature is known to be a key abiotic factor governing their ecology by impacting growth (Monteiro et al., 2017a; Monteiro & Lyons, 2012), pregnancy length (Monteiro et al., 2003), timing and extent of the breeding season (Monteiro et al., 2017a), and modulating mating behaviours (Silva et al., 2007). Laboratory studies showcased the perverse effects of high temperatures on gonad development, survival and gene expression (Lin et al., 2006; Mascaró et al., 2019; Miranda et al., 2024; Roth & Landis, 2017), implying that climate change may pose an inescapable challenge to the survival of syngnathids. Surprisingly, to date, little research has been conducted on wild syngnathid populations that allows us to understand the true impact of present-day temperatures along a species' range. The worm pipefish, *Nerophis lumbriciformis*, is an ideal candidate for such a study given the wealth of information related to the species' reproductive biology (*e.g.*, Monteiro et al., 2001; Monteiro et al., 2005a; Monteiro et al., 2017b) coupled with a latitudinally wide-ranging distribution that spreads from the warmer waters of Southern Europe to the colder North of the continent.

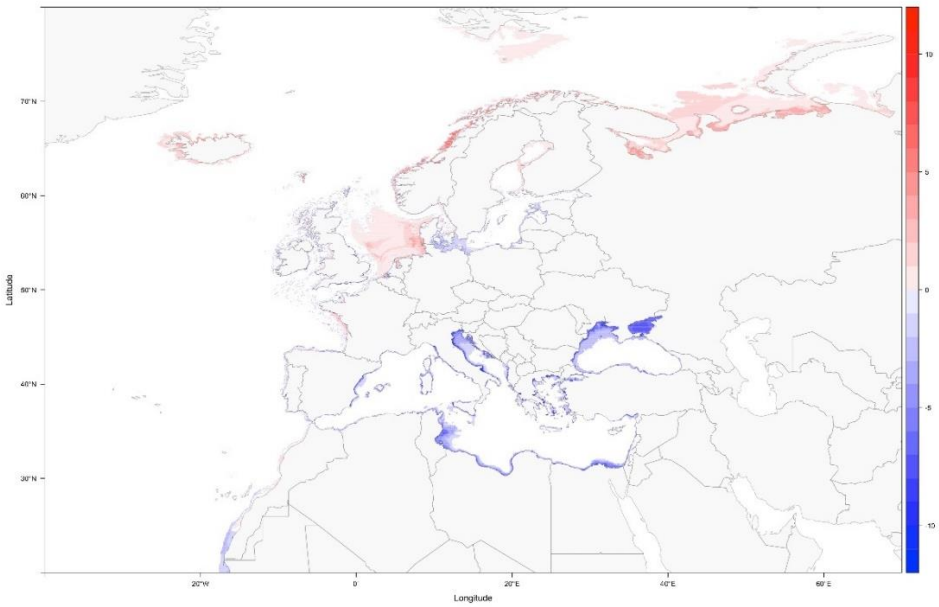

Figure 1 – Syngnathid species richness forecast for the 2090-2100 time period, under a "business-as-usual" RCP 8.5 scenario (blue shows areas with loss of species while red displays areas of increased diversity; from Monteiro et al., 2023).

The worm pipefish is distributed along the Western and Northern European coasts, from Portugal to Norway (Monteiro et al., 2017a). Females are usually the competing sex (Monteiro et al., 2017b; Monteiro et al., 2002), and exhibit a patchy facial coloration (part of their secondary sexual characters' repertoire), especially during the reproductive season (Monteiro et al, 2002), that signals their sexual aptitude and intrinsic quality (Monteiro et al., 2017b). Previous research, conducted by Monteiro et al. (2017a), showed that worm pipefish females from populations at range limits, when compared to those near the centre of distribution, display similarly expanded facial coloration (Figure 2) and production of more and larger eggs. While the expression of reproduction-related variables shows clear similarities at both range extremes, growth patterns clearly differ with latitude (poleward increase in body size accompanied by contracting sexual size dimorphism) (Monteiro et al., 2017a). Considering that sexual selection is known to often lead to reproductive isolation and speciation (Ritchie, 2007), it was suggested that higher sexual selection pressure, induced by the contrasting temperature regimes experienced at the range limits which enforce shorter breeding seasons or gamete production intervals (Monteiro et al., 2017a), could hypothetically explain the observed parabolic variation in the expression of secondary sexual characters and overall reproduction investment. Given the observed sexual selection intensity variation along the species' range, Mendes et al. (2020) initially hypothesised some degree of population differentiation. Surprisingly, results obtained with (mainly) mitochondrial markers showed

an apparent lack of population structure, even when contrasting populations separated by more than 20º of latitude and experiencing very distinct temperature regimes (Mendes et al., 2020). These results, however, do not fully dismiss the possibility that, while harder to detect, differentiation exists, restricted to a few islands scattered in a genome otherwise homogenised by gene flow (Hemmer-Hansen et al., 2014; Pujolar et al., 2014).

To our knowledge, no study was ever conducted on syngnathid natural populations using a two-omics approach, focused on disentangling the long- *vs.* short-term effects of current climate change-modulated temperatures. To fulfil this gap, using a genomic approach, which provides a resolution several orders of magnitude greater than that of the previous study (Mendes et al., 2020), we aim to investigate how the different temperature regimes are impacting the genome of the worm pipefish in order to confirm the absence of structure along the species' range and detect if populations show signs of genetic adaptation to distinct environments over evolutionary time scales. Also, we aim to investigate how the distinct temperature regimes at the North and South range limits are influencing gene expression on the short run, using a transcriptomic approach. We hypothesize that, as current data suggests (*see* Mendes et al., 2020), no demonstrable evidence of genome differentiation between populations at the range extremes will be detected, as the most likely 'culprit' behind the observed morphological differences should reside downstream of the genome to phenotype path. Within a panmixia scenario, since transcriptomic modulation is a faster response mechanism to the environmental conditions, we can still expect differences at gene expression level between North and South range limits modulated by the cold and warm temperature regimes. Among the transcriptomic differences between range limits, we expect to detect thermal stress responses, especially noticeable on the South due to the contemporary climate change heightened sea temperatures that fast approach the species' physiological tolerance limits.

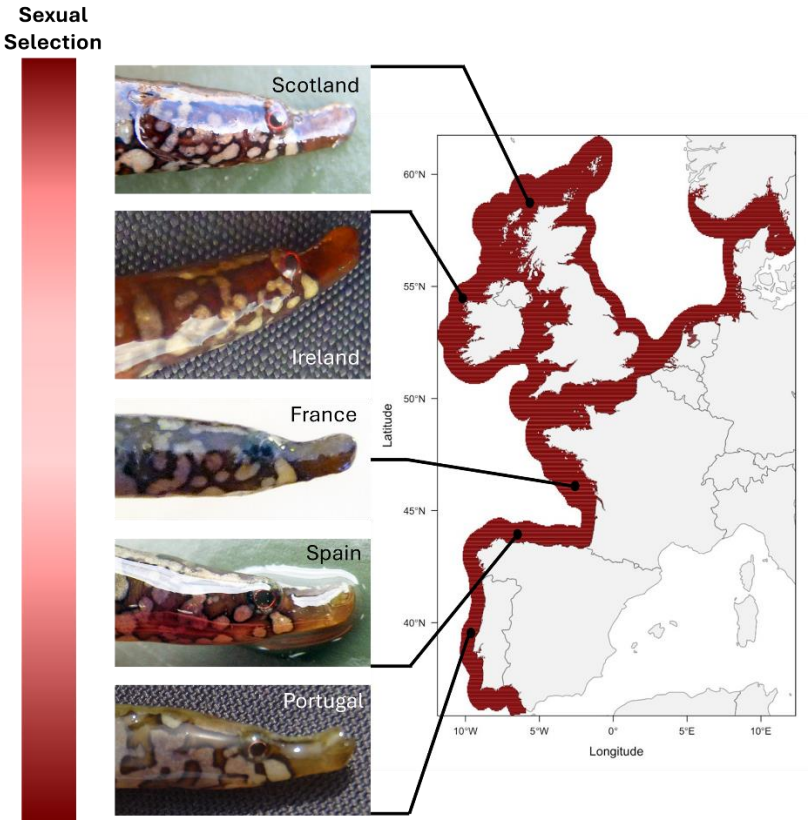

Figure 2 - Distribution of *Nerophis lumbriciformis*, according to IUCN, and examples of female facial coloration from Scotland, Ireland, France, Spain and Portugal. On the left, there is a visual representation of the hypothesised parabolic variation in sexual selection intensity along the species' distribution.

2. Material and Methods

2.1. Sampling

A total of 43 worm pipefish individuals [22 males (M) and 21 females (F)] were sampled during the breeding season, from the intertidal or near subtidal, in Norway (NO; N=9, 4F and 5M; on 08/08/2019), Scotland (SC; N=14, 7F and 7M; on 08/07/2019), Spain (SP; N=10, 5F and 5M; on 06/06/2023), and Portugal (PT; N=10, 5F and 5M; on 24/01/2023) (Figure 3). Sampling locations were selected so as to gather representatives from the North (NO and SC) and South (SP and PT) limits of the species' range (*see* Monteiro et al., 2017a), and this grouping will be termed as *latitude groups*. Sampled individuals were humanely euthanised and immediately preserved in RNAlater (Thermo Fisher Scientific) until DNA and RNA extraction.

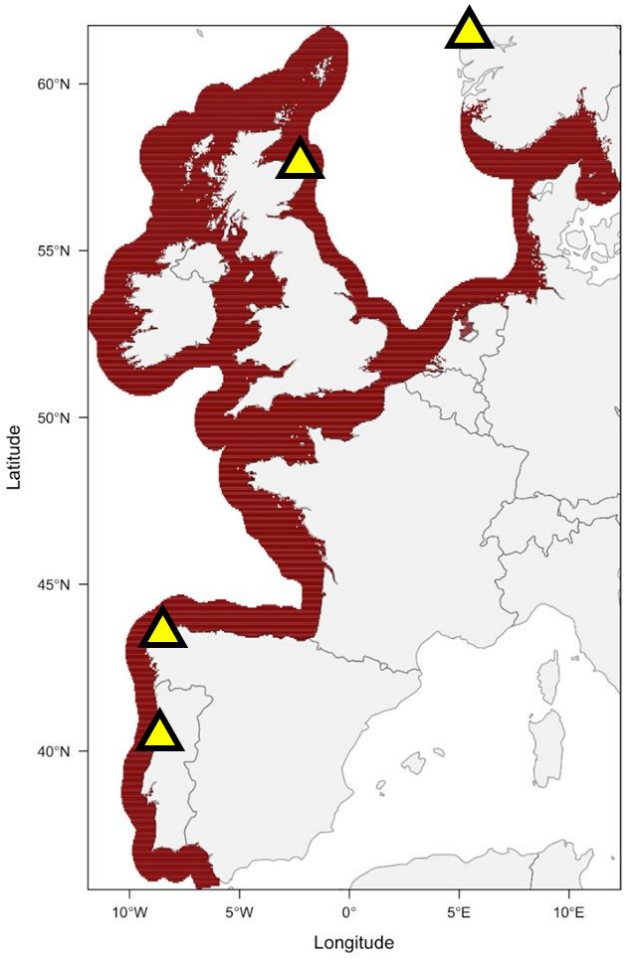

Figure 3 - Distribution of *Nerophis lumbriciformis*, according to IUCN, with yellow triangles indicating the selected sampling sites.

DNA was extracted from muscle tissue, using EZ-10 Spin Column Animal Genomic DNA Miniprep Kit (Bio Basic), and RNA was extracted from liver and ovaries using RNeasy Micro Kit (Qiagen), following the kits manufacturer's instructions. Obtained DNA and RNA were quantified using Qubit (Thermo Fisher Scientific), and whole genomes and transcriptomes were sequenced by Novogene (Beijing, China) using Illumina NovaSeq 6000.

2.2. Genomics

Raw genome reads were mapped to the reference genome of *N. lumbriciformis* (assession number - GCF_033978685.1, downloaded on 18/06/2024), using BWA MEM v.0.7.17 (Li & Durbin, 2009), and the obtained files were sorted and indexed using SAMtools v.1.15.1 (Danecek et al., 2021). Variant calling was then performed with ANGSD v.0.935 (Korneliussen et al., 2014), using SAMtools model (*-GL 1*). Major and minor alleles were inferred from genotype likelihoods (*-doMajorMinor 1*) and allele frequencies were calculated based on known major and unknown minor allele (*-doMaf 2*). Filtering was implemented to dump poor base (*-minQ 20*) and mapping *(-minMapQ 20*) quality, maintaining polymorphic sites *(-SNP_pval 1e-6*), discarding sites with missing data in 20% of the samples *(-minInd 35*), while defining a minimum depth (*-setMinDepthInd 7*). ANGSD output was filtered, using PLINK v.1.9 (Chang et al., 2015), in order to exclude variants present on unplaced scaffolds, maintaining those from the 39 assembled chromosomes. A final pruned SNP dataset was obtained after filtering, in PLINK, for *Minor Allele Frequency* (MAF<0.05) and *Linkage Disequilibrium* (LD, R^2>0.2 within a 50 SNP window and sliding 10 SNPs for each evaluation). Analysis used the pruned SNP dataset, except detection of *Runs of Homozygosity* (ROH) (chapter 2.2.1.) and selection analysis (chapter 2.2.2.), where an unpruned dataset (i.e., without LD and MAF filtering) was used.

2.2.1. Population Genomics

For each population, several genomic diversity measures were computed. Observed Heterozygosity (Ho), Unbiased Expected Heterozygosity corrected for sample size (uHe), Inbreeding Coefficient (FIS), and Private Alleles (PA) were obtained with the R package *dartR* v.2.9.7 (Gruber et al., 2018; Mijangos et al., 2022). Nucleotide diversity (π) per site was calculated with VCFtools v.0.1.16 (Danecek et al., 2011) and averaged

using R (R Core Team, 2023). The presence of large homozygous genomic regions (ROHs) was investigated using PLINK, following Meyermans et al. (2020) recommendations *(--homozyg-density 75; --homozyg-gap 800; --homozyg-window-snp 36; --homozyg-window-threshold 0.139; --homozyg-window-missing 1; --homozyg-window-het 0; --homozyg-snp 36*), with ROH size as PLINK's default. To understand worm pipefish's population structure, pairwise weighted average F_{ST} was calculated with VCFtools, among all population pairs (NOxSC, NOxSP, NOxPT, SCxSP, SCxPT, SPxPT). Also, *Principal Component Analysis* (PCA) was computed on PLINK, and an introgression analysis was performed on ADMIXTURE v.1.3.0 (Alexander et al., 2009), testing K values between 1 and 5. The best K was chosen based on the minor cross-validation error. All plots were obtained using *ggplot2* v.3.5.1 (Wickham, 2016).

2.2.2. Detection of Selection Signatures

Since *latitude groups* are under different environmental conditions, it can be hypothesized that positive selection will act differently on each, selecting different genes that can contribute to local adaptation. Thus, to understand which genes are being selected on each population, RAiSD v.2.9 (Alachiotis & Pavlidis, 2018) was used to search for signatures of positive selection. This software calculates a composite measure, *μ statistic*, based on three signatures of positive selection - high LD, changes on *Site Frequency Spectrum* (SFS), and local reduction in polymorphisms. An independent RAiSD analysis was conducted on each population, and results were concatenated to form a unique dataset. The 0.05% highest μ values were considered candidate sites under positive selection (Alachiotis & Pavlidis, 2018), and were annotated using the reference genome annotation file, a process that was automated with R (R Core Team, 2023).

2.3. Transcriptomics

Raw reads from liver and ovary were quality checked before and after trimming, by FastQC v.0.12.1 (Andrews, 2010) and summarized by MultiQC v.1.0 (Ewels et al., 2016). Trimming was performed with Fastp v.0.23.2 to remove adapters *(--detect_adapter_for_pe*) and low-quality bases (*-q 20*). Reads were then aligned against the reference genome of worm pipefish (assession number - GCF_033978685.1, downloaded on 09/04/2024) using STAR v.2.7.10b (Dobin et al., 2013), employing the

quantification mode of gene counting (*--quantMode GeneCounts*) and only considering reads aligned to unique regions (*--outFilterMultimapNmax 1*). The alignment quality was accessed by summarizing STAR reports with MultiQC.

2.3.1. Differential Gene Expression Analysis

Gene counts from liver and ovary were analysed independently on a differential gene expression analysis, but first they were filtered to exclude genes present on unplaced scaffolds, leaving only those associated to the 39 assembled chromosomes.

To preliminarily detect relationships between countries, *Multidimensional Scaling* (MDS) plots were built for each organ, using *limma* v.3.50.3 (Ritchie et al., 2015), which based sample distances on log fold changes. Inferred relationships were tested for their significance with an *Analysis of Similarity* (ANOSIM), performed in the R package *vegan* v.2.6-6.1 (Oksanen et al., 2015), using the distance method “Bray-Curtis” with significance accessed by 9,999 permutations. For both organs, ANOSIM was used to perform comparisons between countries and *latitude groups* (North and South). For liver gene counts, ANOSIM was also performed to compare sexes. These results were then used to build the model matrix for each organ analysis and to define contrasts.

On each organ gene counts, *edgeR* v.3.36.0 (Chen et al., 2024) was employed to compute normalization factors and filter low or unexpressed genes [*filterByExpr* command (Chen et al., 2016)]. Then, differential gene expression analysis was performed by *limma* R package v.3.50.3 (Ritchie et al., 2015) implementing *voom* transformation (Law et al., 2014). Contrasts were defined between *latitude groups* and sexes (only on liver). Standard errors of log fold changes were smoothed using Empirical Bayes method and results were corrected for multiple testing with the Benjamini & Hochberg (BH) correction, both implemented in *limma*. Genes with adjusted p-values equal or lower than 0.05 were considered as differentially expressed genes and were included in the enrichment analysis (*see* chapter 2.4). To obtain a visual representation of results, Volcano and MA plots were built using *limma*.

2.4. Orthology and Enrichment Analysis

The worm pipefish is not a model organism, a fact that poses difficulties when using tools such as g:Profiler (Kolberg et al., 2023). Thus, to overcome this difficulty, one approach

is to discover orthologs from a close model organism (*e.g. Danio rerio*). To obtain orthologs, OrthoFinder v.2.9 (Emms & Kelly, 2019) was run using proteomes from *Nerophis lumbriciformis* (assession number: GCF_033978685.1), *Danio rerio* (GCA_000002035.4), *Hippocampus comes* (GCA_001891065.1), *Oryzias latipes* (GCA_002234675.1), *Gasterosteus aculeatus* (GCA_016920845.1), *Salmo salar* (GCA_905237065.2), *Gadus morhua* (GCA_902167405.1), *Nothobranchius furzeri* (GCA_001465895.2), *Poecilia reticulata* (GCA_000633615.2), *Cyprinus carpio carpio* (GCA_905221575.1) and *Lepisosteus oculatus* (GCA_000242695.1) as an outgroup. All proteomes were obtained from ensembl (Harrison et al., 2024), except for worm pipefish which was obtained from NCBI (O'Leary et al., 2024). Gene products and protein IDs were assigned to the differentially expressed genes, and pipefish's protein IDs were converted into *D. rerio*'s gene ID, using OrthoFinder's results. To overcome the issue of multiple orthologs for each pipefish protein, biasing the results of g:Profiler, only the first *D. rerio's* ortholog was considered. The described procedure was automated in R (R Core Team, 2023). To understand which pathways were being differentially expressed, an enrichment analysis was performed on g:GOST from g:Profiler (accessed on 14/08/2024) for each organ's differentially expressed genes. Organism was set as *Danio rerio*, the significance was defined by Benjamini-Hochberg FDR set as 0.05 and numeric IDs were treated as ENTREZGENE_ACC. Only GO biological processes and Reactome pathways were used as data sources. To increase accuracy of the results, no electronic GO annotations were considered.

3. Results

3.1. Genomics

All genomes were successfully sequenced, and ANGSD detected 242,513 SNPs. After scaffold filtering, the unpruned dataset dropped to 241,659 SNPs and, after MAF and LD filtering, the pruned dataset retained 78,906 SNPs.

3.1.1. Population Genomics

We computed several measures of genomic diversity that revealed similarly high levels of diversity for the sampled populations (Table 1). Ho is higher than uHe, accompanied by negative values of FIS, low numbers of PA, and high nucleotide diversity for each population ($\pi \approx 0.30$). The high levels of diversity are corroborated by the absence or very low number of ROH on each population. In total, only 5 ROH were detected, with lengths varying from 1,045 kb to 2,667 kb. The majority of ROH (80%) were present on SC individuals, and 2 were from the same individual (F4). The remaining ROH was present on a SP individual (Table 2).

Table 1 - Diversity measures for each country (Ho = Observed Heterozygosity, uHe = Unbiased Expected Heterozygosity, FIS = Inbreeding Coefficient, PA = Private Alleles, π = Nucleotide Diversity).

Country	Ho	uHe	FIS	PA	π
NO	0.412	0.301	-0.369	15	0.301
SC	0.415	0.300	-0.386	89	0.300
SP	0.411	0.302	-0.361	4	0.302
PT	0.412	0.302	-0.365	13	0.303

Table 2 - Runs of Homozygosity, as detected by PLINK, specifying the individual where it was detected, the chromosome and its length in kb (F = female, M =male, SC = Scotland, SP = Spain).

Country	Individual	Chromosome	Length (kb)
SC	F3	37	1,636
SC	F4	16	1,177
SC	F4	27	2,667
SC	M2	3	1,045
SP	F3	11	1,175

Structure analysis highlighted an apparent absence of differentiation across populations, with very low pairwise F_{ST} values (NOxSC = 0.0007, NOxSP = 0.0016, NOxPT = 0.0011, SCxSP = 0.0008, SCxPT = 0.0007, SPxPT = 0.0001). PCA first two *Principal Components* (PC) explain 10.35% of variation and depict a lack of structuring (Figure 4). These results were also corroborated by ADMIXTURE (best K=1), implying a single population within our dataset.

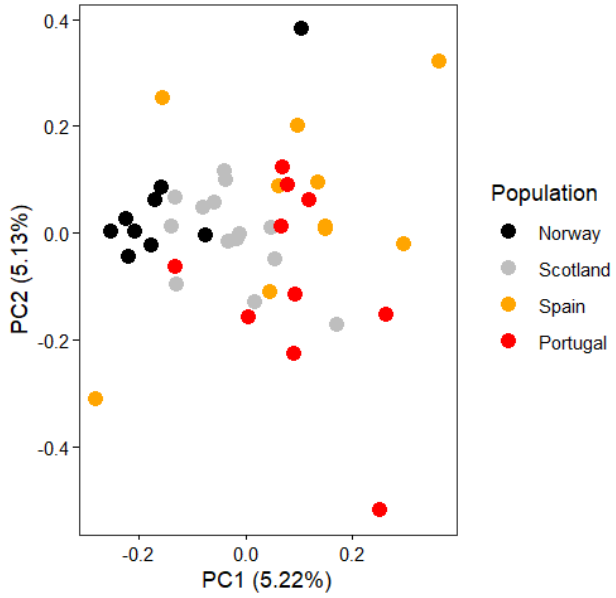

Figure 4 - PCA plotted with the two first *Principal Components* (PC) that explain 10.35% of variation.

3.1.2. Selection

We implemented a RAiSD analysis on each population, to investigate if the two climate regimes, experienced on the North and South range limits, induced positive selection on different genes. RAiSD highest 0.05% μ statistic values returned 350 candidate sites under positive selection, where the lowest μ statistic value was 101.8. Among the total number of selected sites, 60 were identified on NO individuals, 47 on SC, 78 on SP and 164 on PT. The highest μ values were present on chromosome 1 and 23 on all populations (Figure 5). Annotation of selected sites revealed that 12 genes were under selection on NO, 6 on SC, 13 on SP and 30 on PT. Within all populations, only 2 genes were common to all. They codify a cadherin-23 and a protocadherin-10, that are involved in cell adhesion. Also, a codifying gene for sulfhydryl oxidase 1 protein, involved on protein folding, was under selection in NO, SP and PT. When comparing selected genes between populations, excluding those common to all, 3 genes were common between NO and SP, 3 between NO and PT, 2 between SC and PT and 1 between SP and PT.

Between NO and SC, and SC and SP, no more common genes under selection besides those common to all populations, were detected. Most of the selected genes seem to be associated with cell adhesion, synaptic transmission, protein metabolism and transcription processes (*see* Table 3 *for more details*).

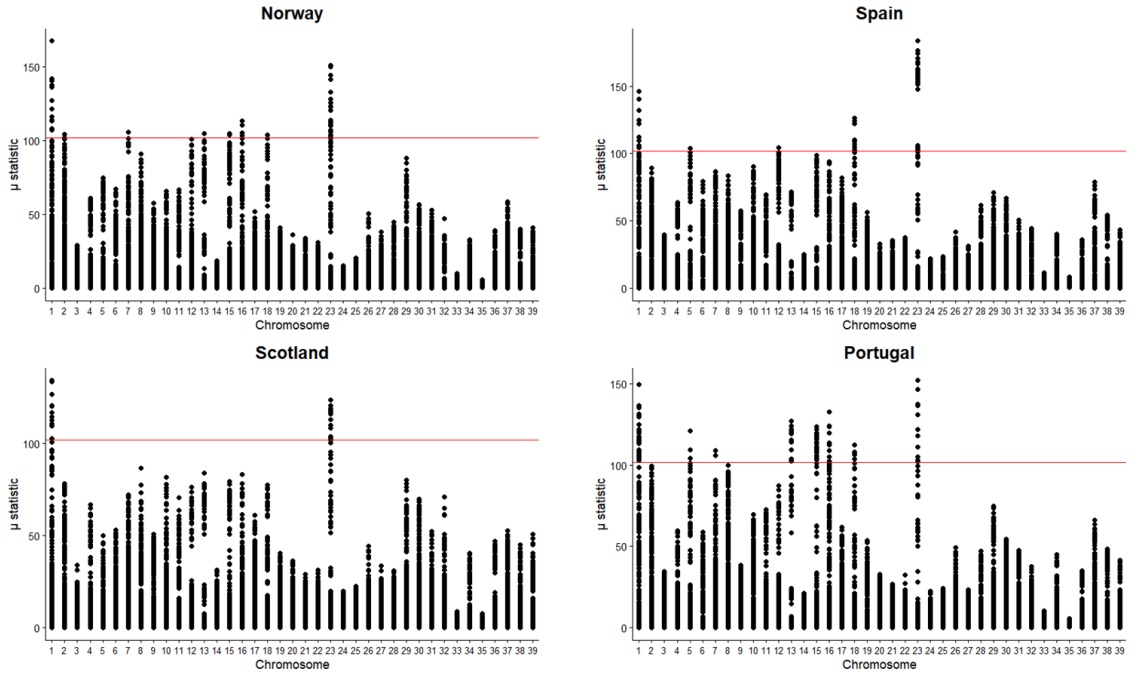

Figure 5 - RAiSD results plotted as Manhattan plots with μ values on the y axis and chromosome on the x axis. A threshold was set at μ=101.8 according to the lowest μ statistic from RAiSD highest 0.05% values.

Table 3 – Genes detected as under selection. The 'Identifier' column shows the gene name used on reference genome annotation. Information regarding 'Gene description' and 'Process' were obtained from NCBI (O'Leary et al., 2024) and UniProt (UniProt Consortium, 2023) websites. The column 'Country' shows where the gene was detected to be under selection.

Identifier	Gene Description	Country	Process
LOC133606963	cadherin-23-like	NO, SC, SP, PT	Responsible for cell-cell adhesion.
LOC133621667	protocadherin-10-like	NO, SC, SP, PT	Responsible for cell-cell adhesion.
LOC133617669	sulfhydryl oxidase 1-like	NO, SP, PT	Protein folding.
LOC133607347	rab11 family-interacting protein 2-like	NO, SP	Involved in regulated exocytosis.
LOC133621659	hydroperoxide isomerase ALOXE3-like	NO, SP	Linoleic acid metabolism and lipid oxidation.
LOC133621662	dynein axonemal heavy chain 6-like	NO, SP	Microtubule-based movement of cellular components.

Identifier	Gene Description	Country	Process
LOC133606899	arginyl-tRNA--protein transferase 1-like	NO, PT	Argination of proteins for further degradation via ubiquitin pathway.
LOC133615997	amyloid-beta A4 precursor protein-binding family A member 2-like	NO, PT	Chemical synapse transmission.
LOC133616457	adenylate kinase 4, mitochondrial-like	NO, PT	Metabolic processes involving nucleus components.
LOC133606837	inactive tyrosine-protein kinase 7-like	SC, PT	Cell-cell adhesion, protein phosphorylation and synapse organization
LOC133606914	serum response factor-like	SC, PT	Regulation of RNA polymerase II associated processes.
LOC133622130	hydroperoxide isomerase ALOXE3-like	SP, PT	Linoleic acid metabolism and lipid oxidation.
LOC133606971	eukaryotic translation initiation factor 3 subunit L-like	NO	Translation initiation.
LOC133622873	calcium-binding and coiled-coil domain-containing protein 1-like	NO	Transcription activation of several genes.
LOC133614701	CLOCK-interacting pacemaker-like	NO	Negative regulator of DNA transcription and circadian cycle.
LOC133621663	sodium channel and clathrin linker 1-like	SC	Cilium assembly and on voltage-gated sodium channels grouping on the membrane.
LOC133621665	UPF0462 protein C4orf33 homolog	SC	No identified function or related process.
srd5a2b	steroid-5-alpha-reductase, alpha polypeptide 2b	SP	Biosynthesis of steroids.
LOC133606489	vitrin-like	SP	Organization of extracellular matrix and regulation (positive) of cell-substrate adhesion.
LOC133607535	regulator of microtubule dynamics protein 2-like	SP	Microtubule binding.
LOC133613193	uncharacterized LOC133613193	SP	No identified function or related process.
LOC133617664	tyrosine-protein kinase ABL2-like	SP	Protein phosphorylation and apoptosis regulation.
LOC133617674	follistatin-related protein 3-like	SP	Cell differentiation and development. Regulator in activin and BMP receptor pathways.
LOC133607830	inositol polyphosphate-5-phosphatase A-like	PT	Dephosphorylates the phosphatidylinositol.
LOC133607838	vertebrate ancient opsin-like	PT	Regulation of G protein binding receptor signalling pathway and participation in visual perception.
LOC133606899	arginyl-tRNA--protein transferase 1-like	PT	Argination of proteins for further degradation via ubiquitin pathway.
LOC133606560	AT-rich interactive domain-containing protein 4B-like	PT	Regulation of RNA polymerase II associated processes.
LOC133607431	sphingosine-1-phosphate lyase 1-like	PT	Metabolism of carboxylic acid and catabolism of sphingolipid.
LOC133606602	lysosomal-trafficking regulator-like	PT	Mediation of protein localization.

Identifier	Gene Description	Country	Process
LOC133606397	neurexin-2-like	PT	Probable involvement on cell recognition, adhesion and intracellular signalling.
LOC133605717	mitochondrial fission 1 protein-like	PT	Involvement on several processes relating to mitochondria (autophagy, fission, fragmentation during apoptosis) and on the division of peroxisome.
LOC133606381	short transient receptor potential channel 5-like	PT	Transmembrane calcium transport, regulating its cytosolic concentrations. Involvement on manganese ion transport.
LOC133614703	cyclin-P-like	PT	Cyclin-dependent protein serine/threonine kinase activity and mitotic phase transition.
LOC133614684	serine/Arginine-related protein 53-like	PT	Alternative mRNA splicing.
LOC133614243	P2X purinoceptor 5-like	PT	Transmembrane transport of monoatomic cation and mediates responses of cell/organism to ATP.
LOC133615997	amyloid-beta A4 precursor protein-binding family A member 2-like	PT	Chemical synapses transmission.
LOC133615995	tight junction protein ZO-1-like	PT	Functions related with cell membrane adhesion and permeability.
LOC133616020	12S rRNA N4-methylcytidine methyltransferase-like	PT	Involvement in the addition of a methyl group to rRNA.
LOC133616003	protein ENTREP2-like	PT	No identified function or related process.
LOC133616006	cartilage intermediate layer protein 1-like	PT	Assumed to be related to cartilage scaffolding.
LOC133616449	dolichyl pyrophosphate Man9GlcNAc2 alpha-1,3-glucosyltransferase-like	PT	N-linked glycosylation.
LOC133616457	adenylate kinase 4, mitochondrial-like	PT	Metabolic processes involving nucleus components.
LOC133616442	KN motif and ankyrin repeat domain-containing protein 4-like	PT	Negative regulation of actin filament polymerization.
LOC133616445	ubiquitin carboxyl-terminal hydrolase 1-like	PT	Deubiquitination of proteins.
LOC133617669	sulfhydryl oxidase 1-like	PT	Protein folding.
LOC133618094	acyl-CoA-binding domain-containing protein 6-like	PT	Protein and fatty-acyl-CoA binding.
LOC133607736	cilia- and flagella-associated protein 46-like	PT	Axoneme organization and cilium movement.
LOC133606914	serum response factor-like	PT	Regulation of RNA polymerase II associated processes.
LOC133606837	inactive tyrosine-protein kinase 7-like	PT	Cell adhesion, protein phosphorylation and synapse organization.
LOC133622132	neuronal acetylcholine receptor subunit alpha-7-like	PT	Membrane potential and synaptic processes.
LOC133622130	hydroperoxide isomerase ALOXE3-like	PT	Linoleic acid metabolism and lipid oxidation.

3.3. Transcriptomics

Most samples were successfully sequenced, with the exception of 4 liver NO males, 1 liver NO female and 1 ovary SC female. The remaining 39 liver samples and 19 ovary samples were successfully sequenced, and the obtained reads were trimmed and aligned against the reference genome. On average, 34M ± 3.37 reads per sample were sequenced from liver, and 37M ± 5.46 from ovary. Samples exhibited elevated levels of read duplication, higher on liver ($\bar{x}_{Liver}$= 84.14% ± 4.33%) when compared to the ovary ($\bar{x}_{Ovary}$ = 61.82% ± 1.60%). Despite this, alignment performed by STAR seemed unaffected since, on average, 83.70% ± 5.14% liver reads, and 88.50% ± 1.32% ovary reads aligned against the reference genome. Reads were initially aligned against 53,119 genes, obtaining gene counts information for those genes. After unplaced scaffold filtering, gene counts information associated to 28,850 genes were retained and used on the analysis.

Preliminary relationships between samples were inferred by MDS on each organ using gene counts information. Plots showed a separation according to the *latitude group* (North and South) and sex (on liver, where both were present), resulting in 4 clusters using liver, and 2 using ovary gene counts (Figure 6). On liver MDS, the first two dimensions explained 30% of variation, and 4 outliers were detected: 2 SP females were in the South males' group, and 1 PT male and 1 SC female were in the North males' group. On ovary MDS, the first two dimensions represented 33% of variation and two females (1 SP and 1 NO) showed to be distant from their respective latitude groups. To associate a statistical significance to the MDS clustering, ANOSIM was employed using gene counts information, comparing countries and *latitude groups* on each organ (and sexes on liver). ANOSIM confirmed the suggested division between *latitude groups* on both organs (Table 4). Significant differences were detected on North (NO and SC) against South countries (SP and PT), and non-significant results were obtained within *latitude groups* (NOxSC and SPxPT). The single exception was the comparison between NO and SP (P=0.058 on liver, P=0.142 on ovary). Significant differences between sexes were detected on liver. Based on these results, contrasts used on the differential gene expression analysis were set between *latitude groups* and sexes (only on liver).

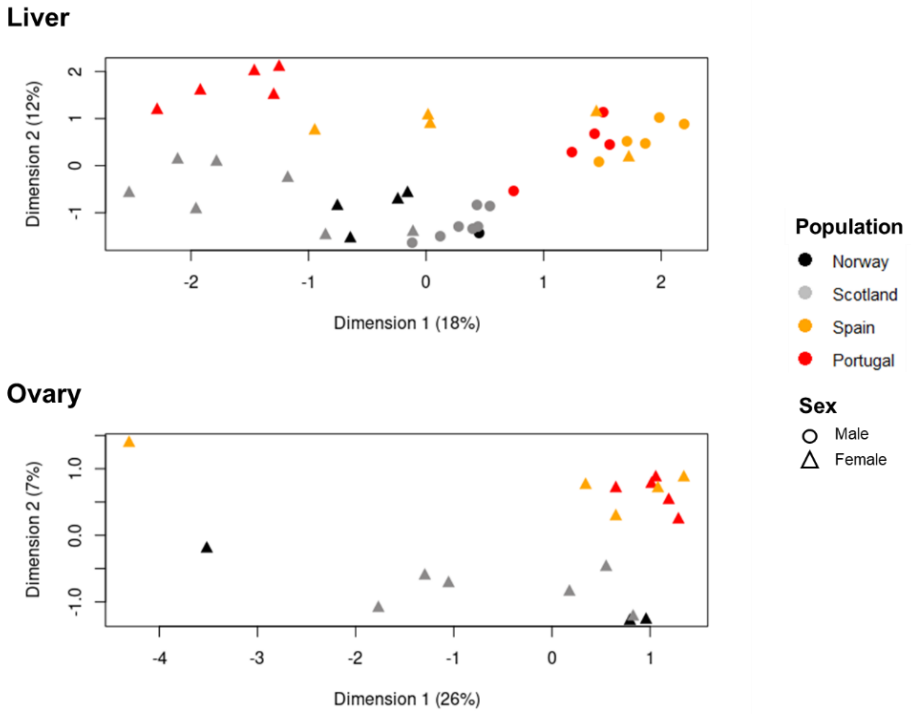

Figure 6 – MDS plots of STAR gene counts obtained from liver and ovary tissues. On each plot sex and country are differentiated by shape and colour, respectively. On the liver plot the first two dimension explain 30% of variation and on the ovary plot 33%.

Table 4 - ANOSIM R values for each independently computed comparison between pairs of countries and sexes, using STAR gene counts (bold* R values indicate statistically significant results (P<0.05). "---" indicates a comparison that was not performed).

Comparison	Liver	Ovary
NO *vs* SC	0.04554	-0.0123
NO *vs* SP	0.2204	0.2513
NO *vs* PT	**0.3527***	**0.7333***
SC *vs* SP	**0.2397***	**0.288***
SC *vs* PT	**0.2109***	**0.547***
SP *vs* PT	0.1273	0
North *vs* South	**0.2395***	**0.431***
Male *vs* Female	**0.3843***	---

3.3.1. Differential gene expression analysis

On liver, filtering performed in *edgeR*, retained gene counts information for 13,256 genes, which were further used on a differential gene expression analysis performed by *limma:voom*, to identify differential expressed genes (DEG) between latitude groups and

sexes. The analysis identified 6,015 DEG between latitude groups, where 2,879 genes were upregulated on the North and 3,136 on the South. Also, 742 DEG were identified between the sexes, with 455 genes upregulated on males and 287 on females (Figure 7). To understand if specific pathways were being enriched within DEG results, an enrichment analysis was performed in g:Gost, identifying 8 and 12 enriched pathways between *latitude groups* and sexes, respectively (Table 5).

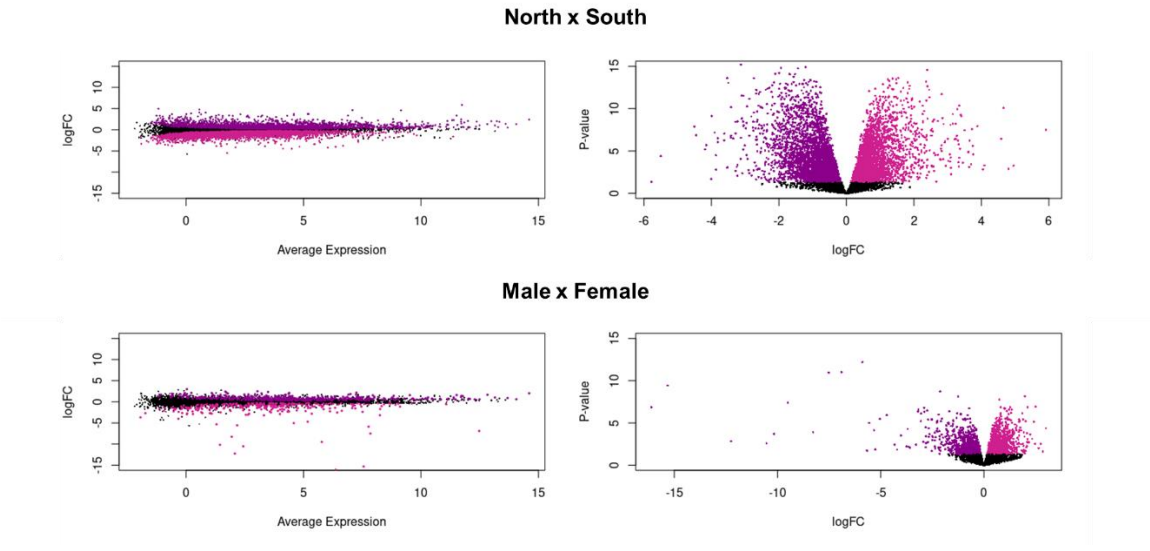

Figure 7 - MA (left) and Volcano plots (right) built using differentially expressed genes in the liver, between *latitude groups* and sexes. MA plot x-axis represents the average expression among all samples on log counts per million, and y-axis represents the log fold changes. Volcano plot x-axis represents the log2 fold changes and y-axis is the *P*-value obtained from the moderated t-test. Purple and pink dots show differentially expressed genes (*P*>0.05).

Enriched pathways between *latitude groups* could be divided into three categories: “Metabolism and Biosynthesis”, “Transcription”, and “Stress Response”. Generally, genes upregulated both on the North and South were present on each enriched pathway. However, water-soluble vitamin metabolic process (GO:0006767), oxoacid metabolic process (GO:0043436) and ribosome biogenesis (GO:0042254) pathways had a higher proportion of North upregulated DEG (71.43%, 67.92% and 61.54%, respectively).

Between sexes, enriched pathways could be divided into four categories: “Metabolism and Biosynthesis”, “Stress Response”, “Reproduction and Development”, and “Homeostasis”. Two enriched pathways related to stress response [cellular response to oxidative stress (GO:0034599) and removal of superoxide radicals (GO:0019430)] were uniquely constituted by male DEG. Also, two additional pathways related to metabolism were mainly associated to male DEGs [hemoglobin biosynthetic process (GO:0042541) and small molecule metabolic process (GO:0044281)]. Furthermore, four pathways were constituted by mainly female DEG [lipid metabolic process (GO:0006629), response to

estrogen (GO:0043627), biomineral tissue development (GO:0031214) and myeloid cell homeostasis (GO:0002262)]. Lastly, the remaining stress response pathways [response to stress (GO:0006950) and intrinsic apoptotic signalling pathway (GO:0097193)], and homeostasis pathways [homeostatic process (GO:0042592) and endoplasmic reticulum plasma membrane tethering (GO:0061817)] showed equal number of DEG for each group.

On the ovary, 17,857 genes were retained after filtering and used on a differential gene expression analysis to infer DEG between *latitude groups*. Between North and South, 4,585 genes were retrieved as DEG where 2,710 were upregulated on North and 1,875 on South (Figure 8). Enrichment analysis performed on g:Gost detected three enriched pathways, composed by proportional parts of upregulated genes from both *latitude groups*: GO:0071704 - organic substance metabolic process, GO:0048477 – oogenesis, GO:0035845 - photoreceptor cell outer segment organization (Table 5).

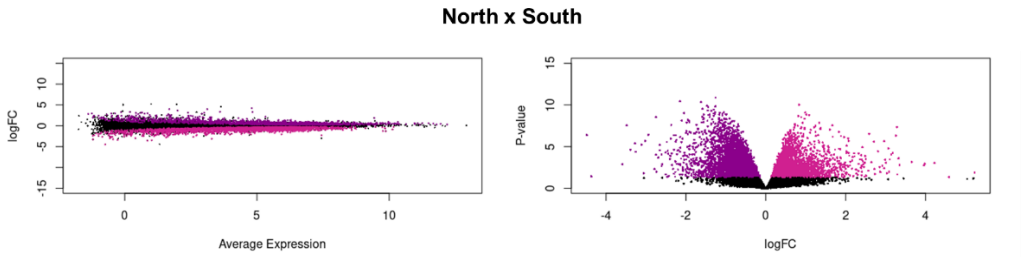

Figure 8 - MA (left) and Volcano plots (right) built using differentially expressed genes in the ovary, between *latitude groups*. MA plot x-axis represents the average expression among all samples on log counts per million, and y-axis represents the log fold changes. Volcano plot x-axis represents the log2 fold changes and y-axis is the *P*-value obtained from the moderated t-test. Purple and pink dots show differentially expressed genes (*P*>0.05).

Table 5 – Enriched pathways present on differentially expressed genes between *latitude groups* (liver and ovary) and sexes (liver) (when a pathway is composed by more than 60% of a specific's group genes, a symbol is added, when it's more than 90%, an asterisk is added to the symbol).

North x South		Male x Female (liver)	GO term	Pathway	Category
Liver	Ovary				
■			GO:0008152	Metabolic process	Metabolism and Biosynthesis
❄			GO:0006767	Water-soluble vitamin metabolic process	
	■		GO:0071704	Organic substance metabolic process	
		▲	GO:0006629	Lipid metabolic process	
		●	GO:0044281	Small molecule metabolic process	
		●	GO:0042541	Hemoglobin biosynthetic process	
■			GO:0006396	RNA processing	Transcription
■			GO:0016071	mRNA metabolic process	
■			GO:0034660	ncRNA metabolic process	
❄			GO:0042254	Ribossome biogenesis	
■			GO:0033554	Cellular response to stress	Stress Response
		■	GO:0006950	Response to stress	
		●*	GO:0034599	Cellular response to oxidative stress	
		●*	GO:0019430	Removal of superoxide radicals	
❄			GO:0043436	Oxoacid metabolic process	
		■	GO:0097193	Instrinsic apoptotic signaling pathway	
		▲	GO:0043627	Response to estrogen	Reproduction and Development
	■		GO:0048477	Oogenesis	
	■		GO:0035845	Photoreceptor cell outer segment organisation	
		▲	GO:0031214	Biomineral tissue development	
		▲	GO:0002262	Myeloid cell homeostasis	Homoeostasis
		■	GO:0042592	Homeostatic process	
		■	GO:0061917	Endoplasmatic reticulum-plasma membrane tethering	

Legend:

❄ North latitude group (Norway and Scotland) ● Male ▲ Female

■ Equal proportions (0.4 ≤ proportion ≤ 0.6) * Proportion higher than 0.9

3.3.2. Possible genes associated to climate differences

To understand how temperature may be affecting transcriptomic profiles in worm pipefish, genes associated to cold and warm clime adaptation were searched within differential expressed genes between *latitude groups* (North and South), for each organ. A summary of differentially expressed genes that may be influenced by the cold from the North and warm from the South are represented on Table 6 and 7, respectively (*for the complete list of differential expressed genes see* Attachment 1). In total, possible climate-influenced genes can be divided into 6 categories on the North and 4 on the South *latitude group*. In the North, genes related to heat shock proteins, apoptosis, transcription, oxoacid metabolism, fatty acid and energy were upregulated. On the South, genes related to heat shock proteins, apoptosis, hypoxia and circadian rhythm were upregulated.

Several members of the heat shock proteins (HSP) family were detected, mainly HSP70, HSP40/DNAJ and *hspb1* on the South, and small HSP, HSP40/DNAJ, HSP70 and HSP90 on the North. HSP are a family of proteins that work as molecular chaperones involved in protein folding processes and regulation, transduction of cell signals, cell cycle and apoptosis (Hu et al., 2022). The HSP can be divided into four categories according to their molecular size and function: heat shock protein factors, small HSP, HSP40 or DNAJ, HSP70 and HSP90. Heat shock factors control the production of HSP, and the detected *hspb1* is a negative regulator of the heat shock factor 1 (*hsf1*), reducing the binding capacity of HSF1 to the promoters of HSP (UniProt Consortium, 2023). Small HSP work as the first attack to the accumulation of unfolded proteins, binding to them and preventing their aggregation (Hu et al., 2022). HSP70 are dependent on ATPase activity to change their conformation and bind to unfolded proteins, to properly fold them again. HSP40 works as a chaperone similar to HSP70, but is also a cochaperone of the last, enhancing the last ATPase activity and consequently its folding function (Hu et al., 2022; Mayer & Bukau, 2005). HSP90 has a chaperone activity, whose main functions relate to protein folding processes and protein conformational maturation (Hoter et al., 2018; Hu et al., 2022).

Apoptosis related genes were upregulated both on the North and South, with the majority (78%) being upregulated on the latter. In the South, apoptotic genes were related to apoptosis inducers (*aifm1*, *aifm2*, *naif1* and *bax*), regulators (*triap1* and *bcl2l13*) and a caspase (*casp3)*. Apoptosis-inducing factors (AIF) trigger DNA condensation and degradation in a caspase-independent apoptosis mechanism (Sevrioukova, 2011), while *bax* mediates the permeabilization of the mitochondria membrane, allowing cytochrome

C release and apoptosis (Peña-Blanco & García-Sáez, 2018). Relating to the apoptotic regulators, *triap1* is an inhibitor (Adams et al., 2015) while *bcl2l13* can be an inhibitor or an inducer of apoptosis, according to the conditions (Kim et al., 2023). Genes related to P53 and BCL2 regulation, caspases and apoptotic inducers were identified in the South. Tumour protein p53 (TP53, in the study identified *tp53inp1*) is important to prevent tumor formation by regulating apoptosis, cell cycle and DNA repair. It is regulated by apoptosis-stimulating protein of p53 family (ASPP), where the inducers *aspp1* and *aspp2* were identified as upregulated in this study (Yin et al., 2018). BCL2 family members were identified among upregulated genes, they work as regulators of apoptosis where they can promote or repress it, depending on the conditions. Lastly, caspases *casp3* (identified on North), *casp6* and *casp7* are crucial during apoptosis where they act as executers. Executers caspases, during apoptosis scenarios, are activated by a group of initiation caspases, that cleave the executers, allowing them to acquire their protease capacity and participate on the apoptosis process (McIlwain et al., 2013).

Oxoacid metabolism related genes, upregulated in the North, were associated to glutathione peroxidase, glutathione S-transferase, peroxiredoxin and superoxide dismutase. Glutathione peroxidase reduces peroxide and lipid hydro-peroxydes using glutathione and an electron donor (Shah et al., 2013). Glutathione S-transferase is involved in the conjunction of toxic exogenous and endogenous compounds with reduced glutathione, so that they can be excreted (Sainas et al., 2018). Peroxiredoxin is part of a family of peroxidases involved on the reduction of peroxide and alkyl hydroperoxides (Echtay, 2007).

Among the transcription related genes upregulated on the North, genes related to transcription process initiation and regulation of alternative splicing were identified. The family of eukaryotic translation initiation factor (eIF) is involved on the initiation of the translation process where it composes the ribosomal machinery (Hao et al., 2020). The other identified genes are associated to the alternative splicing process where they incorporate the spliceosome (UniProt Consortium, 2023) or as members of the SR family (serine/aginine-rich proteins), being involved in alternative pre-mRNA splicing and post-splicing processes (Shepard & Hertel, 2009). Lastly, *cold-inducible RNA-binding protein A-like* (*cirp*) is usually expressed under cold-shock temperatures, where it regulates certain mRNAs by binding to their 3'-UTR (UniProt Consortium, 2023).

Energy related genes upregulated in the North are associated to ATP synthase and NADH dehydrogenase ubiquinone. ATP synthase genes codify for several components of the ATP synthase enzyme, which is involved on ATP production using an

electrochemical gradient of protons (Whitehouse & Moore, 2013). NADH dehydrogenase ubiquinone genes codify parts of the complex NADH-ubiquinone oxidoreductase present on the mitochondrial membrane and participates on the electron transport chain, during respiration and consequent production of ATP, where it catalyses the transfer of NADH electrons to the ubiquinone (Paradies et al., 2014). On fatty acid related genes, several acyl-CoA enzymes and *elovl6* were upregulated in the North and are mainly associated to lipid and fatty acid metabolism (Benjamins et al., 2012).

Hypoxia related genes were upregulated in the South. *Hypoxia upregulated protein 1-like* belongs to HSP70 and it may play an important role on protein folding and secretion on endoplasmic reticulum and protect the cell on hypoxic conditions (Rao et al., 2021). Hypoxia-inducible factors (*hif1a* and *hif3a*) are transcription factors that activate genes that increase oxygen delivery in order to support hypoxic conditions (Averett et al., 2014; UniProt Consortium, 2023).

Circadian rhythm genes as *period*, *CLOCK-interacting pacemaker-like* and *cry* were also identified as upregulated on the South. PERIOD genes modulate the molecular circadian clock rhythm, by negative feedback on circadian clock loop, to regulate species' physiology and cellular rhythms. *CLOCK-interacting pacemaker-like* is a negative regulator of CLOCK-BMAL1 activity. CRY genes, *cry1* and *cry2*, are involved on light perception and, consequently, on the modulation of the circadian rhythm (UniProt Consortium, 2023).

Also, to investigate whether differentially expressed genes between sexes using liver could be indicative of stress, we searched for stress response associated genes, where we summarized them on Table 8. The identified genes could be divided into 5 categories: "Energy", "Oxoacid Metabolism", "Heat Shock Proteins" and "Apoptosis". The first two categories were only composed by male upregulated genes. The majority of differentially expressed genes between sexes were associated to reproduction (*see* Attachment 1 *for more information*) and anchored the highest values of log fold changes. For example, female upregulated *vitellogenin-1-like* and *2-like* log fold changes were higher than 10. These genes code for vitellogenin protein, which is the main source of protein on the yolk sack of fishes (Sullivan & Yilmaz, 2018).

Table 6 – Summary of upregulated genes in the North, possibly associated to survival in colder climate (asterisk indicates log fold changes between 1 and 5).

	Gene ID	Description/Abreviation	North x South	
			Liver	Ovary
Heat Shock Protein	hsbp1b	heat shock factor binding protein 1b / *hsbp1b*		
	dnajc24	DnaJ (Hsp40) homolog, subfamily C, member 24 / *dnajc24*		
	LOC133570215	dnaJ homolog subfamily B member 14-like / *dnajb14*		
	LOC133614269	dnaJ homolog subfamily C member 30, mitochondrial-like / *dnajc30*		
	LOC133576211	dnaJ homolog subfamily C member 5-like / *dnajc5*	*	
	LOC133611571	dnaJ homolog subfamily C member 9-like / *dnajc9*		
	LOC133623995	heat shock 70 kDa protein / *hsp70*		*
	LOC133606861	heat shock 70 kDa protein 14-like / *hspa14*		
Apoptosis	LOC133607465	apoptosis-inducing factor 1, mitochondrial-like / *aifm1*		
	aifm2	apoptosis inducing factor mitochondria associated 2 / *aifm2*		*
	LOC133607343	nuclear apoptosis-inducing factor 1-like / *naif1*		
	LOC133614162	TP53-regulated inhibitor of apoptosis 1-like / *triap1*		
	LOC133605668	bcl-2-like protein 13 / *bcl2l13*	*	
	LOC133619458	apoptosis regulator BAX-like / *bax*		
	LOC133618827	caspase-3-like / *casp3*		
Oxoacid Metabolism	LOC133578302	glutathione peroxidase 1-like / *gpx1*	*	
	LOC133611264	glutathione S-transferase 3, mitochondrial-like / *mgst3*		
	LOC133602560	glutathione S-transferase A4-like / *gsta4*		
	LOC133620825	glutathione S-transferase A-like / *gsta*	*	
	LOC133620824			*
	LOC133595603	glutathione S-transferase kappa 1-like / gstk1		
	LOC133624276	glutathione S-transferase Mu 3-like / *gstm3*	*	
	LOC133611491	glutathione S-transferase omega-1-like / *gsto1*		
	LOC133619423	glutathione S-transferase theta-3-like / *gstt3*	*	
	LOC133605748	microsomal glutathione S-transferase 1-like / *mgst1*	*	
	LOC133605747		*	
	mgst3a	microsomal glutathione S-transferase 3a / *mgst3*		
	prdx6	peroxiredoxin 6 / *prdx6*	*	
	LOC133577236	peroxiredoxin-1-like / *prdx1*		
	LOC133624542			
	LOC133608143	peroxiredoxin-5, mitochondrial-like / *prdx5*		
	LOC133577746	peroxiredoxin-like 2A / *prxl2a*	*	*
	LOC133607847	superoxide dismutase [Cu-Zn]-like / *sod1*		
	LOC133623842	superoxide dismutase [Mn], mitochondrial-like / *mnsod*	*	
Transcription	LOC133609405	eukaryotic translation initiation factor 1A, X-chromosomal / *eif1ax*		
	LOC133622702	eukaryotic translation initiation factor 2 subunit 2-like / *eif2s2*		
	LOC133609426	eukaryotic translation initiation factor 2 subunit 3-like / *eif2s3*		
	eif2a	eukaryotic translation initiation factor 2A / *eif2a*		
	LOC133620235	eukaryotic translation initiation factor 2D-like / eif2d		
	LOC133620574	eukaryotic translation initiation factor 3 subunit D-like / *eif3d*		
	LOC133577589	eukaryotic translation initiation factor 3 subunit E-B-like / *eif3eb*		
	LOC133618544	eukaryotic translation initiation factor 3 subunit F-like / *eif3f*		
	LOC133594133	eukaryotic translation initiation factor 3 subunit H-like / *eif3h*		
	LOC133614395	eukaryotic translation initiation factor 3 subunit K-like / *eif3k*		
	LOC133618912	eukaryotic translation initiation factor 4E type 2-like / *eif4e2*	*	
	LOC133622362	eukaryotic translation initiation factor 4E type 3-like / *eif4e3*		
	LOC133623214	eukaryotic translation initiation factor 4E-binding protein 1-like / *eif4ebp1*		
	LOC133614462	eukaryotic translation initiation factor 4H-like / *eif4h*		
	LOC133609418	eukaryotic translation initiation factor 5A-1-like / *eif5a1*		
	LOC133570933	eukaryotic translation initiation factor 6 / *eif6*		
	prpf31	PRP31 pre-mRNA processing factor 31 homolog (yeast) / *prp31*		
	LOC133576755	RNA-binding motif protein, X chromosome-like / *rbmx*		
	LOC133578264	pre-mRNA-splicing factor ISY1 homolog / *isy1*		
	LOC133623579	splicing factor 3B subunit 6-like / *sf3b6*		
	LOC133616654	splicing factor 3B subunit 5 / *sf3b5*		
	LOC133574653	serine/arginine-rich splicing factor 9-like / *srsf9*		
	LOC133620220	serine/arginine-rich splicing factor 2-like / *srsf2*		
	LOC133618932	cold-inducible RNA-binding protein A-like / *cirp*		

Fatty Acid	LOC133622947	acyl-CoA dehydrogenase family member 11-like / *acad11*		
	acad11	acyl-CoA dehydrogenase family, member 11 / *acad11*		
	LOC133612278	acyl-CoA Delta-6 desaturase-like / *fads2*	*	
	LOC133610363	acyl-CoA desaturase-like / *fads*	*	
	LOC133584877	acyl-CoA desaturase-like / fads	*	
	LOC133621081	acyl-coenzyme A diphosphatase FITM2-like / *fitm2*		
	LOC133610561	acyl-coenzyme A thioesterase 13-like / *acot13*		
	LOC133618565	acyl-coenzyme A thioesterase 3-like / *acot3*	*	
	LOC133619489	acyl-coenzyme A thioesterase MBLAC2-like / *mblac2*	*	
	LOC133606065	fatty acyl-CoA reductase 1-like / *far1*		
	fads6	fatty acid desaturase 6 / *fads6*	*	
	LOC133605380	fatty acid-binding protein 10-A, liver basic-like / *fabp10a*	*	
	LOC133576337	fatty acid-binding protein, brain / *fabp7*	*	
	LOC133577245	fatty acid-binding protein, heart-like / *fabp3*	*	
	LOC133610771	elongation of very long chain fatty acids protein 6-like / *elovl6*	*	
	LOC133616978	elongation of very long chain fatty acids protein 6-like / elovl6	*	
Energy	LOC133606092	catalase-like / *cat*	*	
	LOC133609592	5'-AMP-activated protein kinase subunit gamma-2-like / *prkag2*		
	LOC133608618	AMP deaminase 3-like / *ampd3*		
	crebl2	cAMP responsive element binding protein-like 2 / *crebl2*		
	pkig	protein kinase (cAMP-dependent, catalytic) inhibitor gamma / pkig		
	LOC133576442	cAMP-specific 3',5'-cyclic phosphodiesterase 7B-like / *pde7b*	*	
	LOC133613275	cyclic AMP receptor 4-like / *car4*		
	LOC133624262	ATP synthase F(0) complex subunit B1, mitochondrial-like / *atp5f1b*		
	LOC133620298	ATP synthase F(0) complex subunit C2, mitochondrial-like / *atp5mc2*		
	LOC133613794			
	LOC133615885	ATP synthase F(0) complex subunit C3, mitochondrial-like / *atp5mc3*		
	LOC133624595	ATP synthase membrane subunit K, mitochondrial-like / *atp5mk*		
	LOC133610799	ATP synthase mitochondrial F1 complex assembly factor 2-like / *atpaf2*		
	LOC133574935	ATP synthase subunit alpha, mitochondrial-like / *atp5f1a*		
	LOC133575653			
	LOC133571974	ATP synthase subunit beta, mitochondrial / *atp5f1b*		
	LOC133613838			
	LOC133582341	ATP synthase subunit C lysine N-methyltransferase-like / *atp5mcl*		
	LOC133570098	ATP synthase subunit d, mitochondrial-like / *atp5f1d*		
	LOC133617676	ATP synthase subunit delta, mitochondrial-like / *atp5f1d*		
	LOC133622905	ATP synthase subunit e, mitochondrial-like / *atp5me*		
	LOC133615437	ATP synthase subunit f, mitochondrial-like / *atp5mf*		
	LOC133607991	ATP synthase subunit g, mitochondrial-like / *atp5mg*		
	LOC133606895	ATP synthase subunit gamma, mitochondrial-like / *atp5f1g*		
	LOC133576863	ATP synthase subunit O, mitochondrial-like / *atp5o*		
	LOC133576009	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex assembly factor 2-like / *ndufaf2*		
	LOC133576192	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex assembly factor 4-like / *ndufaf4*		
	LOC133613386	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 10, mitochondrial-like / *ndufa10*		
	LOC133618013	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 11-like / *ndufa11*		
	LOC133621555	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 12-like / *ndufa12*		
	LOC133618931	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 13-like / *ndufa13*		
	LOC133617023	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 1-like / *ndufa1*		
	LOC133607821	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 2-like / *ndufa2*		
	LOC133606562	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 5-like / ndufa5		
	LOC133616791	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 7-like / *ndufa7*		
	LOC133610251	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 8-like / *ndufa8*		
	LOC133612607	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 9, mitochondrial-like / *ndufa9*		
	LOC133615622	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 10-like / *ndufb10*		
	LOC133612626	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 11, mitochondrial-like / *ndufb11*		
	LOC133623721	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 1-like / *ndufb1*		
	LOC133581304	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 3-like / *ndufb3*		
	LOC133609723	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 4-like / *ndufb4*		
	LOC133616633	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 5, mitochondrial-like / *ndufb5*		
	LOC133622005	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 6-like / *ndufb6*		
	LOC133624311	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 7-like / *ndufb7*		
	LOC133610361	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 8, mitochondrial-like / *ndufb8*		
	LOC133577306	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 9-like / *ndufb9*		
	LOC133614152	NADH dehydrogenase [ubiquinone] 1 subunit C2-like / *ndufc2*		
	LOC133618491	NADH dehydrogenase [ubiquinone] flavoprotein 2, mitochondrial-like / *ndufv2*		
	LOC133606410	NADH dehydrogenase [ubiquinone] iron-sulfur protein 2, mitochondrial-like / *ndufs2*		
	LOC133616042	NADH dehydrogenase [ubiquinone] iron-sulfur protein 3, mitochondrial-like / *ndufs3*		
	LOC133623353	NADH dehydrogenase [ubiquinone] iron-sulfur protein 4, mitochondrial-like / *ndufs4*		
	LOC133577321	NADH dehydrogenase [ubiquinone] iron-sulfur protein 5-like / *ndufs5*		
	LOC133621308	NADH dehydrogenase [ubiquinone] iron-sulfur protein 6, mitochondrial-like / *ndufs6*		

Table 7 - Summary of upregulated genes in the South, possibly associated to survival in warmer climates (asterisk indicates log fold changes between 1 and 5).

	Gene ID	Description/Abreviation	North x South	
			Liver	Ovary
Heat Shock Protein	LOC133578417	heat shock protein beta-8-like / *hspb8*	*	
	dnajc1	DnaJ (Hsp40) homolog, subfamily C, member 1 / *dnajc1*		
	dnajc14	DnaJ (Hsp40) homolog, subfamily C, member 14 / *dnajc14*		
	LOC133613341	dnaJ homolog subfamily B member 11-like / *dnajb11*	*	
	LOC133607493	dnaJ homolog subfamily B member 13-like / dnajb13		
	LOC133620513	dnaJ homolog subfamily B member 1-like / *dnajb1*		
	LOC133609565	dnaJ homolog subfamily C member 13-like / *dnajc13*		
	LOC133619603	dnaJ homolog subfamily C member 25-like / *dnajc25*		
	LOC133572863		*	
	LOC133571774	dnaJ homolog subfamily C member 5-like / *dnajc5*		
	LOC133608568	heat shock 70 kDa protein 12A-like / *hspa12a*	*	
	LOC133607685			
	LOC133577103	*heat shock 70 kDa protein 4-like* / *hspa4*		
	LOC133618846	stress-70 protein, mitochondrial-like	*	
	LOC133576405	heat shock protein HSP 90-alpha-like / *hsp90*		
Apoptosis	LOC133623667			*
	LOC133576568	apoptosis-stimulating of p53 protein 1-like / *aspp1*		
	LOC133588048			
	LOC133624434	apoptosis-stimulating of p53 protein 2-like / *aspp2*		
	LOC133603505	tumor protein p53-inducible nuclear protein 1-like / *tp53inp1*	*	
	LOC133621549	apoptosis facilitator Bcl-2-like protein 14 / *bcl2l14*	*	
	LOC133618785	apoptosis regulator Bcl-2-like / *bcl2*	*	
	bag6	BCL2 associated athanogene 6 / *bag6*		
	bcl2l10	BCL2 like 10 / *bcl2l10*	*	
	LOC133572153	bcl-2-like protein 1 / *bcl2l1*		
	LOC133611297	bcl-2-modifying factor-like / *bmf1*	*	
	LOC133617509	bcl-2-related ovarian killer protein homolog A-like / *bok*		
	LOC133595613	caspase-2-like / *casp2*		
	LOC133570226	caspase-6-like / *casp6*	*	
	LOC133610439	caspase-7-like / *casp7*		
	acin1a	apoptotic chromatin condensation inducer 1a / *acin1*	*	
	LOC133570957	death-inducer obliterator 1-like / *dido1*		*
	ddias	DNA damage-induced apoptosis suppressor / *ddias*		
Hypoxia	LOC133607366	hypoxia up-regulated protein 1-like / *hyou1*	*	
	LOC133582015	hypoxia-inducible factor 1-alpha inhibitor-like	*	
	LOC133611750	hypoxia-inducible factor 1-alpha-like / *hif1a*		
	LOC133614396	hypoxia-inducible factor 3-alpha-like / *hif3a*		
Circadian Rythm	LOC133611722		*	
	LOC133576125	CLOCK-interacting pacemaker-like / *cirp*		
	LOC133606783	period circadian protein homolog 1-like / *per1*	*	
	LOC133572556	period circadian protein homolog 2-like / per2	*	
	per2	period circadian clock 2 / *per2*		
	LOC133572293			
	LOC133612565	cryptochrome-1-like / *cry1*		
	LOC133612398	cryptochrome-2-like / *cry2*		

Table 8 - Summary of stress associated genes, in the liver, between the sexes (one asterisk indicates a log fold change between 1 and 5. Blue indicates genes upregulated on males and pink on females).

	Gene ID	Description/Abreviation	Male x Female (liver)
Energy	LOC133609723	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 4-like / *ndufb4*	
	LOC133589138	cAMP-specific 3',5'-cyclic phosphodiesterase 4D-like / *pde4d*	*
Oxoacid Metabolism	LOC133611264	glutathione S-transferase 3, mitochondrial-like / *gst4*	*
	LOC133620825	glutathione S-transferase A-like / *gsta*	
	LOC133619423	glutathione S-transferase theta-3-like / *gstt3*	*
	mgst3a	microsomal glutathione S-transferase 3a / *mgst3a*	
	LOC133607847	superoxide dismutase [Cu-Zn]-like / *sod*	
	LOC133623842	superoxide dismutase [Mn], mitochondrial-like / *mnsod*	
	LOC133605748	microsomal glutathione S-transferase 1-like / *mgst1*	*
	LOC133577236	peroxiredoxin-1-like / *pex1*	*
Heat Shock Protein	LOC133614482	stress-associated endoplasmic reticulum protein 1 / *serp1*	
	LOC133607493	dnaJ homolog subfamily B member 13-like / *dnajb13*	*
	LOC133618846	stress-70 protein, mitochondrial-like	
Apoptosis	LOC133619458	apoptosis regulator BAX-like / *bax*	
	LOC133618785	apoptosis regulator Bcl-2-like / *bcl2*	

4. Discussion

4.1. Panmixia in bad swimmers

Population genomics results suggest a panmictic scenario for the worm pipefish. The high genetic diversity levels (*e.g.*, Ho>uHe; FIS<0 and π≈0.30) combined with the portrayed absence of population structure (*e.g.*, minute FST values, PCA and ADMIXTURE single population clustering), indicate that even the range limits are genomically homogeneous. This trend probably extends to the entire distribution, as suggested by Mendes et al. (2020) results. Thus, the worm pipefish genomic results contrasts with the common intermediate/high levels of structure among syngnathids, as identified by lower-resolution methods (*e.g.*, *Hippocampus guttulatus* - Woodall et al., 2015; *Hippocampus hippocampus* - Woodall et al., 2011; *Syngnathus typhle* - Wilson & Veraguth, 2010; and *Syngnathus abaster* - Sanna et al., 2013). Although all above-mentioned syngnathids share similar life-history traits, namely adult poor swimming abilities and strong site fidelity (Monteiro et al., 2005b, 2006), the worm pipefish displays a pelagic larvae phase, a trait that is (uncoincidentally) shared with another panmictic syngnathid, the snake pipefish, *Entelurus aequoreus* (Braga Gonçalves et al., 2017; Monteiro et al., 2003). A pelagic larvae life stage enhances the dispersal potential of these pipefishes, allowing gene flow with geographically distant populations (Cowen et al., 2006). While there are no ocean currents directly connecting the two latitudinally distant worm pipefish range limits, the observed genome homogeneity should derive from continuous pelagic larvae migrations stemming from the range centre, possibly the British Isles (Figure 14), where the species is most abundant. The North Atlantic Drift (NAD) morphing into the Norwegian Current (NC) may continuously transport larvae from the British Isles (*e.g.*, Ireland and Scotland) towards the North range limit, while the Canary Current (CC) could support the constant transport of larvae (*e.g.*, from Ireland) towards the South, down to the Iberian Peninsula. This “British Isles leakage” seems to be the most plausible scenario able to explain the Atlantic panmictic population currently observed in the worm pipefish. To test this hypothesis, it will be important to include samples from the range core in order to further validate the lack of structure along the whole distribution, thus further expanding our knowledge on the connectivity patterns of the worm pipefish - a ‘bad swimmer’ that started as a highly travelled toddler.

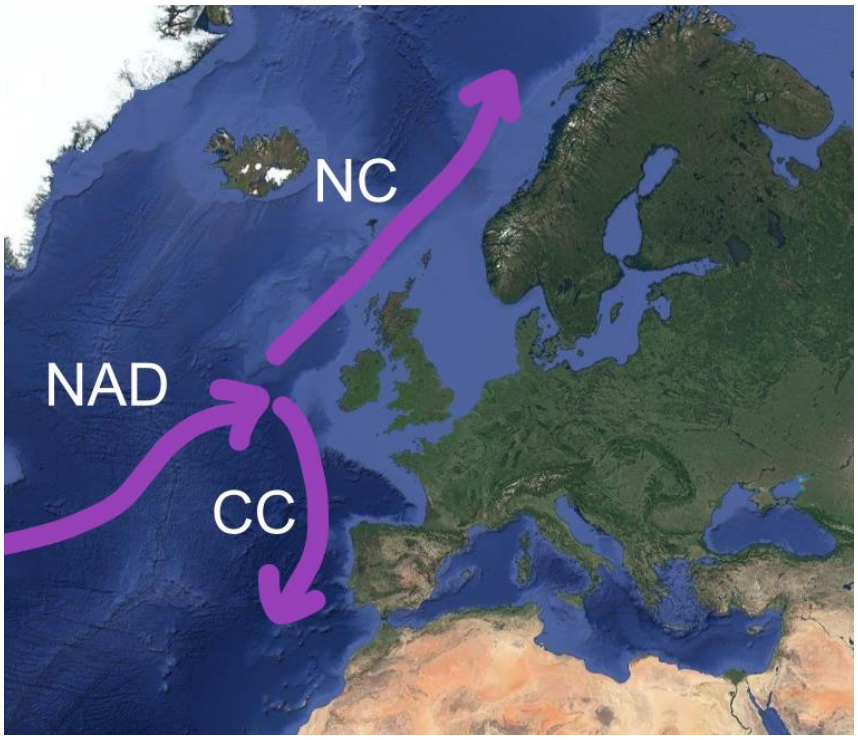

Figure 9 – Sea current patterns on the Northeast Atlantic Ocean (NAD – North Atlantic Drift; NC – Norwegian Current; CC – Canary Current).

4.2. Temperature has not markedly impacted the genome… (yet?)

It could be hypothesised that, under contrasting climate regimes (cold North *vs.* warm South), selection would tend to inexorably work towards promoting adaptation, differentially impacting each of the two range limits by targeting different genes, as observed in other species (e.g. *Maccullochella peelii*, Harrisson et al., 2017; *Symphodus tinca* and *Symphodus ocellatus*, Torrado et al., 2020; and *Gadus morhua*, Nielsen et al., 2009). Our RAiSD results, reconfirming the general absence of genomic differentiation between range limits, seem to discard this hypothesis, since no apparent adaptation to the different climates was detected, with none (NOxSC) or only one selected gene (SPxPT) being shared by populations within the same *latitude group*. Also, the detected selected genes were mainly associated to cellular/molecular processes as cell adhesion, synaptic transmission, protein metabolism and transcription, which aren't obvious candidates to support temperature adaptation and are mostly associated to housekeeping chores. However, we found a gene possibly associated to adaptation to higher latitudes in Norway. *CLOCK-interacting pacemaker-like* or *cipc*, is suggested to be a negative-feedback regulator on circadian cycle and its downregulation is associated

with the shortening of the circadian period length (Zhao et al., 2007). Since, on higher latitudes, photoperiod range is wider and temperatures lower, worm pipefish individuals may need to adapt to the longer days that overlap the breeding season, where CIPC may be crucial to help adjust their physiology and behaviour (Hut et al., 2013). A similar result was obtained on the European eel (*Anguilla anguilla*), where gene *period* was detected to be under selection and in higher frequencies at higher latitudes (Pujolar et al., 2014). Thus, latitude, even if highly correlated with temperature, can also be a modulator of selection in the worm pipefish.

The absence of clear adaptation signals to warm and cold climates could be due to at least three factors. First, technical problems, as sequencing errors and depth could have ignored important SNPs associated to selected genes. Second, selection can be in an early stage, where its signals are not yet detectable or happening on standing variation which create a soft selective sweep. RAiSD only detects hard selective sweeps, which are associated to the hitchhike effect of new acquired mutations that are beneficial on the surrounding environment (Alachiotis & Pavlidis, 2018). However, research has been showing that selection on standing variation due to local adaptation is common, leading to soft selective sweeps, where several haplotypes are associated to the selected locus (Messer & Petrov, 2013). Consequently, soft sweeps will be missed, since they are usually associated to some degree of diversity, and RAiSD only searches for sites with low levels of polymorphisms (Alachiotis & Pavlidis, 2018; Vitti et al., 2013). Lastly, gene flow may prevent or continuously dilute the effect of selection on the genome. This scenario might be the most unlikely (Tigano & Friesen, 2016) as several studies have been demonstrating the possibility of local adaptation even in highly connected populations. For instance, along the range of *Gadus morhua* and *Maccullochella peelii*, researchers detected genes/loci under positive selection as adaptative to different environmental features such as temperature, precipitation, salinity and latitude (Nielsen et al., 2009; Harrisson et al., 2017). Also, a study conducted on two sympatric species, where one has some degree of population structure (*Symphodus ocellatus*) and the other lacks it (*Symphodus tinca*), both showed positive selection and adaptation to the environment (Torrado et al., 2020). Since temperature is far from a negligible factor for syngnathids, ruling their reproduction and survival (Lin et al., 2006; Monteiro et al., 2017a, 2023), temperature-dependent selection pressure is expected to be strong and thus able to create detectable signals, even in a highly connected population as that of the worm pipefish. Nonetheless, if the “British Isles leakage” phenomenon is a reality, its constant rhythm and/or the magnitude of migrating pipefish from the core of the distribution towards the range limits might be enough to dilute any incipient temperature-

related modification of the genome. On future work, alternative selection detection methods and a landscape genomics approach could help verify if the apparent absence of adaptation-related genomic fingerprints is indeed true. A closer look at chromosomes 1 and 23, for instance, could be an appropriate starting point.

4.3. Harsh places require stress responses

Before advancing further into the discussion of our transcriptomics results, I have to acknowledge that our sampling scheme was not ideal. Sampling both range limits was an adequate strategy for the detection of population structure as, if there was differentiation, it would most certainly occur between the two geographically distant extremes of distribution. The same rationale applies also to the detection of positively selected genes. But, once we established that no obvious differences existed between the range extremes, despite the clear variation in morphological (e.g., increasing size and decreasing sexual size dimorphism as we climb in latitude) and reproductive variables (e.g., heightened colouration and egg investment at the range extremes), we were forced to take the next obvious step leading from genome to phenotype, and look into the transcriptome. Here, the previously defined strategy of contrasting range extremes partially lacks the required power to fully grasp the dynamics of gene expression throughout a temperature cline.

In essence, our sampling design has two main shortcomings of distinct dimensions: temporal and geographical. In terms of time, we are aware that we would profit by sampling distinct time points throughout the year (e.g., at least the colder and warmer months, within each location). Although we made sure that our sampling was standardised, in the sense that it was performed, at each location, during the breeding season, where seawater temperatures during pipefish collection did not obviously translate the ‘hot to cold’ latitudinal cline. As the worm pipefish breeds during a relatively narrow temperature interval (7-14ºC), the timing of the breeding season latitudinally varies, winter-centred in the South and summer-centred in the North (Monteiro et al., 2017a). As such, if gene expression is merely an ephemeral response to the surrounding environment, genes associated to cold temperatures would not be expected to be upregulated in the North, at least during the warmer months of the breeding season. Conversely, genes associated to hot temperatures would be less expressed in the South, especially at the beginning of the breeding season, when seawaters are colder. In terms

of geography, we also believe that we would gain from incorporating samples from more range-central locations, where we know that reproduction investment is diminished and seawater temperatures are less extreme, when compared to the latitudinal edges of distribution. Here, we employed a differential gene expression analysis, contrasting natural populations that seem to be under thermal stress. Given that, when considering reproduction-related variables, both hot and cold range limits display similar phenotypes (initially interpreted by Monteiro et al. (2017a) as the end-result of heightened sexual selection), it could be hypothesised that similar sets of upregulated genes could be present in both distributional extremes. If they do exist, these similarities could not have been picked up by our analysis without a control population (i.e., geographically central, thermally less stressed). Furthermore, if a set of genes is upregulated, albeit differentially, on both range extremes when compared with a control population, our analysis would pick up the signal of only one group. Despite the fact that our sampling scheme was adequate for the analysis of the worm pipefish population structure pattern, the results from our transcriptomic analysis reveal only part of the story. Even considering the limitations listed above, our results are far from uninteresting.

Unlike the inferred genomic homogeneity across the worm pipefish range, transcriptomic data showed a clear differentiation between North (NO and SC) and South (SP and PT). Since both groups experience overall different climates, cold on the North and warm on the South, pipefish will necessarily have to adjust their physiology to cope with the environment. Here, near the limits of distribution, with the ‘same’ genome producing different transcriptomic profiles, we are probably witnessing the approach to the thermal boundaries of phenotypic plasticity, which may either reflect a genetic adaptation-free acclimatization process or the first mechanism of adaptation in a path ultimately leading to genomic rearrangement (Sommer, 2020).

Between North and South, the main detected differences were related to metabolism, transcription, reproduction and stress responses. As observed in other species, several genes upregulated during cold shock and heat shock experiments were also upregulated on the worm pipefish North and South *latitude group*, respectively (Cold shock: Gracey et al., 2004; Hu et al., 2015; Li et al., 2020; Liu et al., 2023; Long et al., 2013; Heat Shock: Cheng et al., 2015; Del Vecchio et al., 2022; Hori et al., 2010; Liu et al., 2022). In the colder North, the worm pipefish: i) activated transcription associated genes such as *cirp*, typically involved on mRNA regulation and activated during cold-shock treatments (UniProt Consortium, 2023), as well as alternative splicing genes involved on the spliceosome machinery or members of the SR family (serine/arginine-rich proteins),

which will produce different isoforms that may be better adapted to cold temperatures (Li et al., 2020; Singh & Ahi, 2022); ii) activated oxoacid metabolism, upregulating glutathione peroxidase, glutathione S-transferase, peroxiredoxin and superoxide dismutase genes, to cope with increased reactive oxygen species (ROS) due to higher production of energy through O_2 consumption and anabolism of lipids (Chowdhury & Saikia, 2020; Menon et al., 2023; Reid et al., 2022); iii) probable increased energy production due to the upregulation of genes associated with ATP synthase and NADH dehydrogenase ubiquinone machinery; iv) activated upregulation of acyl-CoA enzymes and *elovl6*, hinting at a fatty acid metabolism increase, further used for fat accumulation to regulate body temperature or for energy production (Li et al., 2020); and v) activated heat shock proteins, especially HSP70 but also HSP40 which is a cochaperone of the former and may be consequently upregulated in order to enhance its activity (Gracey et al., 2004; Hu et al., 2015; Li et al., 2020; Liu et al., 2023; Long et al., 2013; Reid et al., 2022). In the warmer South, worm pipefish increased the expression of heat shock proteins (small HSP, HSP70, HSP40/DNAJ and HSP90) that usually intervene as a first response to heat stress and prevent the accumulation of unfolded proteins caused by high temperatures (Yin et al., 2018). Also, there was an activation of hypoxia related genes (*hyou1*, *hif1a* and *hif3a*), since higher temperatures tend to decrease the dissolved oxygen on the water, forcing organisms to maximize O_2 transport and better cope with the depleted oxygen environment.

The results presented above suggest that both groups may be under thermal stress associated to the two distinct year-round temperature regimes, and gene expression remodulation eases the adaptation process to each particular environment. Also, in both groups (but mainly in the South), several genes associated to the apoptotic process were upregulated (e.g., *aspp1*, *aspp2*, *casp6* and *casp7*). This upregulation can be indicative of an approach to the “exhaustion phase” where the damages induced by thermal stress are accumulating at the cellular level, leading to apoptotic processes and, in the most extreme scenario, to the death of the organism (Reid et al., 2022; Roychowdhury et al., 2024). This unveils a clearer picture of how contemporary high temperatures, such as those imposed by climate change, are already impacting syngnathids, especially on the southern part of their distribution. High temperatures influence syngnathid population health, increasing their mortality and impacting their reproduction (Del Vecchio et al., 2022; Lin et al., 2006; Miranda et al., 2024). Since future climate trends are not optimist for the resilience of southern syngnathid populations (Monteiro et al., 2023), conservation actions must be taken to preserve these species.

When considering stress responses, the lack of substantial sexual dimorphism on gene expression, except for the upregulation of oxoacid metabolism on males, could highlight that the response to stressful conditions is largely sex independent. Accordingly, although Monteiro et al. (2017a) only suggested that females were coerced to increase investment in reproduction in thermally sub-optimal areas, our results hint at the fact that males could be equally constrained by stressful environments. Since gamete production (Guyomarc'h, 1972; Lahaye, 1971), mating (Monteiro et al., 2001) and pregnancies (Monteiro et al., 2003) are highly dependent on water temperature, reproduction will be constrained on the range limits. In the North, although pipefish might have a larger period to produce eggs or sperm, the breeding season is shorter and pregnancies longer. In the South, while the breeding season is longer and pregnancies shorter, the appropriate period for gamete production is briefer. These restrictions will surely constrain females, the competing sex (Monteiro et al., 2002) in this sex-role reversed pipefish, but also males.

Despite the similarity between the many enriched pathways highlighted for other teleosts and those from the worm pipefish, a few discrepancies were observed such as the circadian rhythm genes associated to cold stress (Li et al., 2020), that we found upregulated in the South (warmer climate) *latitude group*. Also, oxoacid metabolism associated to heat stress (Hori et al., 2010; Liu et al., 2022), was absent and only upregulated in the cold-stressed northern populations. These discrepancies can obviously derive from interspecific differences in the physiology of the fish. Alternatively, we could have unintentionally uncovered a subgroup of genes/pathways that translate proximate responses to thermal stress impacts. If expression patterns are much more than an immediate response to the environmental conditions, we expect to see transcriptomic fingerprints of both immediate responses to thermal stress (*e.g.*, translating mechanisms to cope with a heat or cold wave, for instance) and more perennial responses (*e.g.*, translating mechanisms that allow pipefish to be better adapted to an environment closer to its physiological thermal limits, even if momentarily experiencing more pleasant temperatures).

Reusing the metaphor of climbing the stairway leading from genome to phenotype, it seems that the obvious next step will be to look into the epigenome – a trait that allows for the remodulation of gene expression in order to enhance adaptation, and that can be transmitted to the future generations (Baduel et al., 2024), without changing the genome. Latitudinal differences in the epigenome, possibly modulated by temperature, would help

justify how an apparently homogenous genome, when differentially translated, supports the observed latitudinal linear clines in morphology or parabolic clines in reproduction.

To further secure our steps, it will be crucial to include populations near the worm pipefish distribution core. We already did this, by sampling northern France, but, due to time constraints, it was not possible to include this data on our final report.

4.4. Future Prospects

Along the discussion section, the need for central range populations became especially apparent, to further validate the panmixia scenario and properly contrast the thermal stress transcriptomic profiles. To this extent, we sampled a population in France. However, due to time constraints, it was not possible to include this data on the final report. This thesis can be viewed as an exploratory study, where we obtained preliminary results that should guide future work. On the next step, with more sampled populations, we will more robustly confirm the lack of structure along species' range and lack of genes strongly associated to the North and South climate regimes. Also, we will possibly extend our transcriptomic investigation, i) confirming the temperature driven stress responses on *latitude groups*, ii) understanding if individuals at the range limits are differentially investing on reproduction, when compared to the centre, iii) possibly compare populations using only female/male related information, and finally, iv) deepen our understanding on gene expression sexual dimorphism, comparing with the available literature.

5. Conclusion

The work developed in this thesis, using a two-omics approach, is innovative within syngnathids. Our results unveil a panmictic scenario where range limits seem to be inhabited by thermally stressed individuals. Importantly, our results help clarify how current climate change-driven seawater temperatures impact individuals, especially those inhabiting the warmer limits of their thermal tolerance, an outcome whose relevance extends well beyond *Nerophis lumbriciformis* or syngnathids. As climate change gives no signs of slowing down on the upcoming years, the already discernible consequences, entailing range shifts and local extinctions, might possibly be just the tip of the iceberg. We urgently need to expand our understanding on the mechanisms supporting adaptation, in order to take meaningful actions that can at least attenuate the current scenario of biodiversity loss, thus protecting our marine ecosystems from a mass extinction.

References

Adams, C., Cazzanelli, G., Rasul, S., Hitchinson, B., Hu, Y., Coombes, R. C., Raguz, S., & Yagüe, E. (2015). Apoptosis inhibitor TRIAP1 is a novel effector of drug resistance. *Oncology Reports*, *34*(1), 415–422. https://doi.org/10.3892/or.2015.3988

Alachiotis, N., & Pavlidis, P. (2018). RAiSD detects positive selection based on multiple signatures of a selective sweep and SNP vectors. *Communications Biology*, *1*, 79. https://doi.org/10.1038/s42003-018-0085-8

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*(9), 1655–1664. https://doi.org/https://doi.org/10.1101%2Fgr.094052.109

Andrews, S. (2010). *FastQC: A quality control tool for high throughput sequence data*. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Averett, C., Arora, S., Zubair, H., Singh, S., Bhardwaj, A., & Singh, A. P. (2014). Molecular Targets of Honokiol: A Promising Phytochemical for Effective Cancer Management. *The Enzymes*, *36*, 175–193. https://doi.org/10.1016/B978-0-12-802215-3.00009-4

Baduel, P., Sammarco, I., Barrett, R., Coronado-Zamora, M., Crespel, A., Díez-Rodríguez, B., Fox, J., Galanti, D., González, J., Jueterbock, A., Wootton, E., & Harney, E. (2024). The evolutionary consequences of interactions between the epigenome, the genome and the environment. *Evolutionary Applications*, *17*(7), e13730. https://doi.org/10.1111/eva.13730

Benjamins, J. A., Murphy, E. J., & Seyfried, T. N. (2012). Chapter 5 - Lipids. In S. T. Brady, G. J. Siegel, R. W. Albers, & D. L. Price (Eds.), *Basic Neurochemistry* (8th ed., pp. 81–100). Academic Press. https://doi.org/10.1016/B978-0-12-374947-5.00005-5

Bertola, L. D., Boehm, J. T., Putman, N. F., Xue, A. T., Robinson, J. D., Harris, S., Baldwin, C. C., Overcast, I., & Hickerson, M. J. (2020). Asymmetrical gene flow in five co-distributed syngnathids explained by ocean currents and rafting propensity. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1926). https://doi.org/10.1098/rspb.2020.0657

Braga-Gonçalves, I., Cornetti, L., Couperus, A. S., van Damme, C. J. G., & Mobley, K. B. (2017). Phylogeography of the snake pipefish, *Entelurus aequoreus* (Family: Syngnathidae) in the northeastern Atlantic Ocean. *Biological Journal of the Linnean Society*, *122*(4), 787–800. https://doi.org/10.1093/biolinnean/blx112

Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, *4*(1), 7. https://doi.org/10.1186/s13742-015-0047-8

Chen, Y., Chen, L., Lun, A. T. L., Baldoni, P. L., & Smyth, G. K. (2024). edgeR 4.0: powerful differential analysis of sequencing data with expanded functionality and improved support for small counts and larger datasets. *BioRxiv*. https://doi.org/10.1101/2024.01.21.576131

Chen, Y., Lun, A. T. L., & Smyth, G. K. (2016). From reads to genes to pathways: differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. *F1000Research*, *5*, 1438. https://doi.org/10.12688/f1000research.8987.2

Cheng, C.-H., Yang, F. F., Liao, S. A., Miao, Y. T., Ye, C. X., Wang, A. L., Tan, J. W., & Chen, X. Y. (2015). High temperature induces apoptosis and oxidative stress in pufferfish (*Takifugu obscurus*) blood cells. *Journal of Thermal Biology*, *53*, 172–179. https://doi.org/10.1016/j.jtherbio.2015.08.002

Chowdhury, S., & Saikia, S. K. (2020). Oxidative Stress in Fish: A Review. *Journal of Scientific Research*, *12*(1), 145–160. https://doi.org/10.3329/jsr.v12i1.41716

Cowen, R. K., Paris, C. B., & Srinivasan, A. (2006). Scaling of Connectivity in Marine Populations. *Science*, *311*(5760), 522–527. https://doi.org/10.1126/science.1122039

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021).

Twelve years of SAMtools and BCFtools. *GigaScience*, *10*(2), giab008. https://doi.org/10.1093/gigascience/giab008

Del Vecchio, G., Galindo-Sánchez, C. E., Tripp-Valdez, M. A., López-Landavery, E. A., Rosas, C., & Mascaró, M. (2022). Transcriptomic response in thermally challenged seahorses Hippocampus erectus: The effect of magnitude and rate of temperature change. *Comparative Biochemistry and Physiology, Part B*, *262*, 110771. https://doi.org/10.1016/J.CBPB.2022.110771

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15–21. https://doi.org/10.1093/bioinformatics/bts635

Echtay, K. S. (2007). Mitochondrial uncoupling proteins - What is their physiological role? *Free Radical Biology and Medicine*, *43*(10), 1351–1371. https://doi.org/10.1016/j.freeradbiomed.2007.08.011

Emms, D. M., & Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology*, *20*(1), 238. https://doi.org/10.1186/s13059-019-1832-y

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, *32*(19), 3047–3048. https://doi.org/10.1093/bioinformatics/btw354

Gracey, A. Y., Fraser, E. J., Li, W., Fang, Y., Taylor, R. R., Rogers, J., Brass, A., & Cossins, A. R. (2004). Coping with cold: An integrative, multitissue analysis of the transcriptome of a poikilothermic vertebrate. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(48), 16970–16975. https://doi.org/10.1073/pnas.0403627101

Gruber, B., Unmack, P. J., Berry, O. F., & Georges, A. (2018). dartr : An r package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Molecular Ecology Resources*, *18*(3), 691–699. https://doi.org/10.1111/1755-0998.12745

Guyomarc'h, J. (1972). Spermatogenic cycle in *Nerophis lumbriciformis* (Teleostean, Syngnathidae). *Cahiers De Biologie Marine*, *13*, 21–35.

Hao, P., Yu, J., Ward, R., Liu, Y., Hao, Q., An, S., & Xu, T. (2020). Eukaryotic translation initiation factors as promising targets in cancer therapy. *Cell Communication and Signaling*, *18*(1), 175. https://doi.org/10.1186/s12964-020-00607-9

Harrison, P. W., Amode, M. R., Austine-Orimoloye, O., Azov, A. G., Barba, M., Barnes, I., Becker, A., Bennett, R., Berry, A., Bhai, J., Bhurji, S. K., Boddu, S., Branco Lins, P. R., Brooks, L., Ramaraju, S. B., Campbell, L. I., Martinez, M. C., Charkhchi, M., Chougule, K., … Yates, A. D. (2024). Ensembl 2024. *Nucleic Acids Research*, *52*(D1), D891–D899. https://doi.org/10.1093/nar/gkad1049

Harrisson, K. A., Amish, S. J., Pavlova, A., Narum, S. R., Telonis-Scott, M., Rourke, M. L., Lyon, J., Tonkin, Z., Gilligan, D. M., Ingram, B. A., Lintermans, M., Gan, H. M., Austin, C. M., Luikart, G., & Sunnucks, P. (2017). Signatures of polygenic adaptation associated with climate across the range of a threatened fish species with high genetic connectivity. *Molecular Ecology*, *26*(22), 6253–6269. https://doi.org/10.1111/mec.14368

Hemmer-Hansen, J., Therkildsen, N. O., & Pujolar, J. M. (2014). Population Genomics of Marine Fishes: Next-Generation Prospects and Challenges. *The Biological Bulletin*, *227*(2), 117–132. https://doi.org/10.1086/BBLv227n2p117

Hori, T. S., Gamperl, A. K., Afonso, L. O., Johnson, S. C., Hubert, S., Kimball, J., Bowman, S., & Rise, M. L. (2010). Heat-shock responsive genes identified and validated in Atlantic cod (*Gadus morhua*) liver, head kidney and skeletal muscle using genomic techniques. *BMC Genomics*, *11*(1), 72. https://doi.org/10.1186/1471-2164-11-72

Hoter, A., El-Sabban, M. E., & Naim, H. Y. (2018). The HSP90 Family: Structure, Regulation, Function, and Implications in Health and Disease. *International Journal of Molecular Sciences*, *19*(9), 2560. https://doi.org/10.3390/ijms19092560

Hu, C., Yang, J., Qi, Z., Wu, H., Wang, B., Zou, F., Mei, H., Liu, J., Wang, W., & Liu, Q. (2022). Heat shock proteins: Biological functions, pathological roles, and therapeutic opportunities. *MedComm*, *3*(3). https://doi.org/10.1002/mco2.161

Hu, P., Liu, M., Zhang, D., Wang, J., Niu, H., Liu, Y., Wu, Z., Han, B., Zhai, W., Shen, Y., & Chen, L. (2015). Global identification of the genetic networks and

cis-regulatory elements of the cold response in zebrafish. *Nucleic Acids Research*, *43*(19), 9198–9213. https://doi.org/10.1093/nar/gkv780

Hut, R. A., Paolucci, S., Dor, R., Kyriacou, C. P., & Daan, S. (2013). Latitudinal clines: an evolutionary view on biological rhythms. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1765), 20130433. https://doi.org/10.1098/rspb.2013.0433

Kendall, M. S., Poti, M., & Karnauskas, K. B. (2016). Climate change and larval transport in the ocean: fractional effects from physical and physiological factors. *Global Change Biology*, *22*(4), 1532–1547. https://doi.org/10.1111/gcb.13159

Kim, J. H., Kim, H. S., & Lee, S. (2023). BCL2L13 protein prevents apoptosis in acute myeloid leukemia cells. *Oncologie*, *25*(4), 345–354. https://doi.org/10.1515/oncologie-2023-0112

Kolberg, L., Raudvere, U., Kuzmin, I., Adler, P., Vilo, J., & Peterson, H. (2023). g:Profiler—interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). *Nucleic Acids Research*, *51*(W1), W207–W212. https://doi.org/10.1093/nar/gkad347

Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics*, *15*(1), 356. https://doi.org/10.1186/1471-2105-12-398

Lahaye, J. (1971). L'Ovogenèse chez *Nerophis lumbriciformis* (Pennant, 1776). Le cycle sexuel. *Cahiers De Biologie Marine*, *13*, 239–254.

Law, C. W., Chen, Y., Shi, W., & Smyth, G. K. (2014). voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biology*, *15*(2), R29. https://doi.org/10.1186/gb-2014-15-2-r29

Li, B. J., Zhu, Z. X., Qin, H., Meng, Z. N., Lin, H. R., & Xia, J. H. (2020). Genome-Wide Characterization of Alternative Splicing Events and Their Responses to Cold Stress in Tilapia. *Frontiers in Genetics*, *11*, 244. https://doi.org/10.3389/fgene.2020.00244

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Lin, Q., Lu, J., Gao, Y., Shen, L., Cai, J., & Luo, J. (2006). The effect of temperature on gonad, embryonic development and survival rate of juvenile seahorses, *Hippocampus kuda* Bleeker. *Aquaculture*, *254*(1–4), 701–713. https://doi.org/10.1016/j.aquaculture.2005.11.005

Liu, E., Zhao, X., Li, C., Wang, Y., Li, L., Zhu, H., & Ling, Q. (2022). Effects of acute heat stress on liver damage, apoptosis and inflammation of pikeperch (*Sander lucioperca*). *Journal of Thermal Biology*, *106*, 103251. https://doi.org/10.1016/j.jtherbio.2022.103251

Liu, S., Tian, F., Qi, D., Qi, H., Wang, Y., Xu, S., & Zhao, K. (2023). Physiological, metabolomic, and transcriptomic reveal metabolic pathway alterations in *Gymnocypris przewalskii* due to cold exposure. *BMC Genomics*, *24*(1), 545. https://doi.org/10.1186/s12864-023-09587-9

Long, Y., Song, G., Yan, J., He, X., Li, Q., & Cui, Z. (2013). Transcriptomic characterization of cold acclimation in larval zebrafish. *BMC Genomics*, *14*, 612. https://doi.org/10.1186/1471-2164-14-612

Mascaró, M., Horta, J. L., Diaz, F., Paschke, K., Rosas, C., & Simões, N. (2019). Effect of a gradually increasing temperature on the behavioural and physiological response of juvenile *Hippocampus erectus*: Thermal preference, tolerance, energy balance and growth. *Journal of Thermal Biology*, *85*, 102406. https://doi.org/10.1016/J.JTHERBIO.2019.102406

Mayer, M. P., & Bukau, B. (2005). Hsp70 chaperones: Cellular functions and molecular mechanism. *Cellular and Molecular Life Sciences*, *62*(6), 670. https://doi.org/10.1007/s00018-004-4464-6

McIlwain, D. R., Berger, T., & Mak, T. W. (2013). Caspase Functions in Cell Death and Disease. *Cold Spring Harbor Perspectives in Biology*, *5*(4), a008656–a008656. https://doi.org/10.1101/cshperspect.a008656

Mendes, T., Gomes, C., Monteiro, N., & Antunes, A. (2020). Strong Sexual Selection Does Not Induce Population Differentiation in a Fish Species with High Dispersal Potential: The Curious Case of the Worm Pipefish *Nerophis lumbriciformis* (Teleostei: Syngnathidae). *Journal of Heredity*, *111*(7), 585–592. https://doi.org/10.1093/jhered/esaa052

Menon, S. V., Kumar, A., Middha, S. K., Paital, B., Mathur, S., Johnson, R., Kademan, A., Usha, T., Hemavathi, K. N., Dayal, S., Ramalingam, N.,

Subaramaniyam, U., Sahoo, D. K., & Asthana, M. (2023). Water physicochemical factors and oxidative stress physiology in fish, a review. *Frontiers in Environmental Science*, *11*. https://doi.org/10.3389/fenvs.2023.1240813

Messer, P. W., & Petrov, D. A. (2013). Population genomics of rapid adaptation by soft selective sweeps. *Trends in Ecology & Evolution*, *28*(11), 659–669. https://doi.org/10.1016/J.TREE.2013.08.003

Meyermans, R., Gorssen, W., Buys, N., & Janssens, S. (2020). How to study runs of homozygosity using plink? A guide for analyzing medium density SNPdata in livestock and pet species. *BMC Genomics*, *21*(1), 94. https://doi.org/10.1186/s12864-020-6463-x

Mijangos, J. L., Gruber, B., Berry, O., Pacioni, C., & Georges, A. (2022). dartR v2: An accessible genetic analysis platform for conservation, ecology and agriculture. *Methods in Ecology and Evolution*, *13*(10), 2150–2158. https://doi.org/10.1111/2041-210X.13918

Miranda, L. M., Correia, M., Barraca, C., Baylina, N., & Faria, A. M. (2024). Prolonged exposure to elevated temperature leads to severe thermal stress in the broad-nosed pipefish, *Syngnathus typhle*. *Marine and Freshwater Behaviour and Physiology*, *57*(1–3), 1–16. https://doi.org/10.1080/10236244.2024.2378752

Monteiro, N., Almada, V. C., Santos, A. M., & Vieira, M. N. (2001). The breeding ecology of the pipefish *Nerophis lumbriciformis* and its relation to latitude and water temperature. *Journal of the Marine Biological Association of the United Kingdom*, *81*(6), 1031–1033. https://doi.org/10.1017/S002531540100501X

Monteiro, N., Almada, V. C., & Vieira, M. N. (2003). Early life history of the pipefish *Nerophis lumbriciformis* (Pisces: Syngnathidae). *Journal of the Marine Biological Association of the United Kingdom*, *83*(5), 1179–1182. https://doi.org/10.1017/S0025315403008452h

Monteiro, N., Almada, V. C., & Vieira, M. N. (2005a). Temporal patterns of breeding and recruitment in *Nerophis lumbriciformis* (Pisces; Syngnathidae) related to seawater temperatures. *Journal of Fish Biology*, *67*(5), 1475–1480. https://doi.org/10.1111/j.0022-1112.2005.00818.x

Monteiro, N., Cunha, M., Ferreira, L., Vieira, N., Antunes, A., Lyons, D., & Jones, A. G. (2017a). Parabolic variation in sexual selection intensity across the range of a cold-water pipefish: implications for susceptibility to climate change. *Global Change Biology*, *23*(9), 3600–3609. https://doi.org/10.1111/gcb.13630

Monteiro, N., Berglund, A., Vieira, M. N., & Almada, V. C. (2006). Reproductive migrations of the sex role reversed pipefish *Nerophis lumbriciformis* (Pisces; Syngnathidae). *Journal of Fish Biology*, *69*(1), 66–74. https://doi.org/10.1111/j.0022-1112.2006.01064.x

Monteiro, N., Carneiro, D., Antunes, A., Queiroz, N., Vieira, M. N., & Jones, A. G. (2017b). The lek mating system of the worm pipefish (*Nerophis lumbriciformis*): a molecular maternity analysis and test of the phenotype-linked fertility hypothesis. *Molecular Ecology*, *26*(5), 1371–1385. https://doi.org/10.1111/mec.13931

Monteiro, N., & Lyons, D. O. (2012). Stronger Sexual Selection in Warmer Waters: The Case of a Sex Role Reversed Pipefish. *PloS One*, *7*(8), e44251. https://doi.org/10.1371/journal.pone.0044251

Monteiro, N., Vieira, M. da N., & Almada, V. (2005b). Homing behaviour and individual identification of the pipefish Nerophis lumbriciformis (Pisces; Syngnathidae): a true intertidal resident? *Estuarine, Coastal and Shelf Science*, *63*, 93–99. https://doi.org/10.1016/j.ecss.2004.10.012

Monteiro, N., Pinheiro, S., Magalhães, S., Tarroso, P., & Vincent, A. (2023). Predicting the impacts of climate change on the distribution of European syngnathids over the next century. *Frontiers in Marine Science*, *10*. https://doi.org/10.3389/fmars.2023.1138657

Monteiro, N., Vieira, N. R., & Almada, V. (2002). The courtship behaviour of the pipefish *Nerophis lumbriciformis*: reflections of an adaptation to intertidal life. *Acta Ethologica*, *4*(2), 109–111. https://doi.org/10.1007/s102110100048

Nielsen, E. E., Hemmer-Hansen, J., Poulsen, N. A., Loeschcke, V., Moen, T., Johansen, T., Mittelholzer, C., Taranger, G.-L., Ogden, R., & Carvalho, G. R. (2009). Genomic signatures of local directional selection in a high gene flow marine organism; the Atlantic cod (*Gadus morhua*). *BMC Evolutionary Biology*, *9*, 276. https://doi.org/10.1186/1471-2148-9-276

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P., O'Hara, B., Simpson, G., Solymos, P., Stevens, H., & Wagner, H. (2015). Vegan: Community Ecology Package. *R Package Version 2.2-1*, *2*, 1–2.

O'Leary, N. A., Cox, E., Holmes, J. B., Anderson, W. R., Falk, R., Hem, V., Tsuchiya, M. T. N., Schuler, G. D., Zhang, X., Torcivia, J., Ketter, A., Breen, L., Cothran, J., Bajwa, H., Tinne, J., Meric, P. A., Hlavina, W., & Schneider, V. A. (2024). Exploring and retrieving sequence and metadata for species across the tree of life with NCBI Datasets. *Scientific Data*, *11*(1), 732. https://doi.org/10.1038/s41597-024-03571-y

Palumbi, S. R. (1994). Genetic divergence, reproductive isolation and marine speciation. *Annual Review of Ecology and Systematics*, *25*(1), 547–572. https://doi.org/10.1146/annurev.es.25.110194.002555

Paradies, G., Paradies, V., De Benedictis, V., Ruggiero, F. M., & Petrosillo, G. (2014). Functional role of cardiolipin in mitochondrial bioenergetics. *Biochimica et Biophysica Acta - Bioenergetics*, *1837*(4), 408–417. https://doi.org/10.1016/j.bbabio.2013.10.006

Peña-Blanco, A., & García-Sáez, A. J. (2018). Bax, Bak and beyond — mitochondrial performance in apoptosis. *The FEBS Journal*, *285*(3), 416–431. https://doi.org/10.1111/febs.14186

Pujolar, J. M., Jacobsen, M. W., Als, T. D., Frydenberg, J., Munch, K., Jónsson, B., Jian, J. B., Cheng, L., Maes, G. E., Bernatchez, L., & Hansen, M. M. (2014). Genome-wide single-generation signatures of local selection in the panmictic European eel. *Molecular Ecology*, *23*(10), 2514–2528. https://doi.org/10.1111/mec.12753

R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rao, S., Oyang, L., Liang, J., Yi, P., Han, Y., Luo, X., Xia, L., Lin, J., Tan, S., Hu, J., Wang, H., Tang, L., Pan, Q., Tang, Y., Zhou, Y., & Liao, Q. (2021). Biological Function of HYOU1 in Tumors and Other Diseases. *OncoTargets and Therapy*, *14*, 1727–1735. https://doi.org/10.2147/OTT.S297332

Reid, C. H., Patrick, P. H., Rytwinski, T., Taylor, J. J., Willmore, W. G., Reesor, B., & Cooke, S. J. (2022). An updated review of cold shock and cold stress in fish. *Journal of Fish Biology*, *100*(5), 1102–1137. https://doi.org/10.1111/jfb.15037

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, *43*(7), e47. https://doi.org/10.1093/nar/gkv007

Ritchie, M. G. (2007). Sexual Selection and Speciation. *Annual Review of Ecology, Evolution, and Systematics*, *38*, 79–102. https://doi.org/10.1146/annurev.ecolsys.38.091206.095733

Roth, O., & Landis, S. H. (2017). Trans-generational plasticity in response to immune challenge is constrained by heat stress. *Evolutionary Applications*, *10*(5), 514–528. https://doi.org/10.1111/eva.12473

Roychowdhury, P., Aftabuddin, M., & Pati, M. (2024). A Review on The Impact of Thermal Stress on Fish Biochemistry. *Aquatic Sciences and Engineering*, *39*(2), 121–129. https://doi.org/10.26650/ASE20231341460

Sainas, S., Dosio, F., Boschi, D., & Lolli, M. L. (2018). Targeting Human Onchocerciasis: Recent Advances Beyond Ivermectin. *Annual Reports in Medicinal Chemistry*, *51*, 1–38. https://doi.org/10.1016/BS.ARMC.2018.08.001

Sanna, D., Biagi, F., Alaya, H. B., Maltagliati, F., Addis, A., Romero, A., De Juan, J., Quignard, J. P., Castelli, A., Franzoi, P., Torricelli, P., Casu, M., Carcupino, M., & Francalacci, P. (2013). Mitochondrial DNA variability of the pipefish Syngnathus abaster. *Journal of Fish Biology*, *82*(3), 856–876. https://doi.org/10.1111/jfb.12027

Schiebelhut, L. M., & Dawson, M. N. (2018). Correlates of population genetic differentiation in marine and terrestrial environments. *Journal of Biogeography*, *45*(11), 2427–2441. https://doi.org/10.1111/jbi.13437

Sevrioukova, I. F. (2011). Apoptosis-Inducing Factor: Structure, Function, and Redox Regulation. *Antioxidants & Redox Signaling*, *14*(12), 2545–2579. https://doi.org/10.1089/ars.2010.3445

Shah, D., Sah, S., & Nath, S. K. (2013). Interaction between glutathione and apoptosis in systemic lupus erythematosus. *Autoimmunity Reviews*, *12*(7), 741–751. https://doi.org/10.1016/J.AUTREV.2012.12.007

Shepard, P. J., & Hertel, K. J. (2009). The SR protein family. *Genome Biology*, *10*(10), 242. https://doi.org/10.1186/gb-2009-10-10-242

Silva, K., Vieira, M. N., Almada, V. C., & Monteiro, N. M. (2007). The effect of temperature on mate preferences and female–female interactions in *Syngnathus abaster*. *Animal Behaviour*, *74*(5), 1525–1533. https://doi.org/10.1016/J.ANBEHAV.2007.03.008

Singh, P., & Ahi, E. P. (2022). The importance of alternative splicing in adaptive evolution. *Molecular Ecology*, *31*(7), 1928–1938. https://doi.org/10.1111/mec.16377

Sommer, R. J. (2020). Phenotypic Plasticity: From Theory and Genetics to Current and Future Challenges. *Genetics*, *215*(1), 1–13. https://doi.org/10.1534/genetics.120.303163

Sullivan, C. V., & Yilmaz, O. (2018). Vitellogenesis and Yolk Proteins, Fish. *Encyclopedia of Reproduction*, 266–277. https://doi.org/10.1016/B978-0-12-809633-8.20567-0

Tigano, A., & Friesen, V. L. (2016). Genomics of local adaptation with gene flow. *Molecular Ecology*, *25*(10), 2144–2164. https://doi.org/10.1111/mec.13606

Torrado, H., Carreras, C., Raventos, N., Macpherson, E., & Pascual, M. (2020). Individual-based population genomics reveal different drivers of adaptation in sympatric fish. *Scientific Reports*, *10*, 12683. https://doi.org/10.1038/s41598-020-69160-2

UniProt Consortium. (2023). UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Research*, *51*(D1), D523–D531. https://doi.org/10.1093/nar/gkac1052

Venegas, R. M., Acevedo, J., & Treml, E. A. (2023). Three decades of ocean warming impacts on marine ecosystems: A review and perspective. *Deep Sea Research Part II*, *212*, 105318. https://doi.org/10.1016/J.DSR2.2023.105318

Vitti, J. J., Grossman, S. R., & Sabeti, P. C. (2013). Detecting Natural Selection in Genomic Data. *Annual Review of Genetics*, *47*, 97–120. https://doi.org/10.1146/annurev-genet-111212-133526

Whitehouse, D. G., & Moore, A. L. (2013). Respiratory Chain and ATP Synthase. *Encyclopedia of Biological Chemistry*, 83–86. https://doi.org/10.1016/B978-0-12-378630-2.00214-0

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer International Publishing. https://doi.org/10.1007/978-3-319-24277-4

Wilson, A. B., & Veraguth, I. E. (2010). The impact of Pleistocene glaciation across the range of a widespread European coastal species. *Molecular Ecology*, *19*(20), 4535–4553. https://doi.org/10.1111/j.1365-294X.2010.04811.x

Woodall, L. C., Koldewey, H. J., Boehm, J. T., & Shaw, P. W. (2015). Past and present drivers of population structure in a small coastal fish, the European long snouted seahorse Hippocampus guttulatus. *Conservation Genetics*, *16*, 1139–1153. https://doi.org/10.1007/s10592-015-0728-y

Woodall, L. C., Koldewey, H. J., & Shaw, P. W. (2011). Historical and contemporary population genetic connectivity of the European short-snouted seahorse *Hippocampus hippocampus* and implications for management. *Journal of Fish Biology*, *78*(6), 1738–1756. https://doi.org/10.1111/j.1095-8649.2011.02974.x

Yin, L., Lin, Y., Wang, X., Su, Y., Hu, H., Li, C., Wang, L., & Jiang, Y. (2018). The family of apoptosis-stimulating proteins of p53 is dysregulated in colorectal cancer patients. *Oncology Letters*, *15*(5), 6409–6417. https://doi.org/10.3892/ol.2018.8151

Zhao, W.-N., Malinin, N., Yang, F.-C., Staknis, D., Gekakis, N., Maier, B., Reischl, S., Kramer, A., & Weitz, C. J. (2007). CIPC is a mammalian circadian clock protein without invertebrate homologues. *Nature Cell Biology*, *9*(3), 268–275. https://doi.org/10.1038/ncb1539

Attachments

Attachment 1: Web link to access the whole differential gene expression analysis results.

https://uporto-my.sharepoint.com/:x:/g/personal/up201904997_up_pt/EcSQXF8jGjpBvQaJrpT1eZIBMqsrbHgdvlgPnV_GzqLlXA?e=iYBccK