UNIVERSIDADE NOVA
DE LISBOA

MScCBBi

MASTER IN
COMPUTATIONAL BIOLOGY
& BIOINFORMATICS

David Jeremy Lohmann
BSc in Biology

Multi- Omic approaches to unravel phenotypic responses of the tropical fish Lutjanus guttatus to ocean warming and acidification scenarios

Sep, 2024

NOVA SCHOOL OF
SCIENCE & TECHNOLOGY

DEPARTAMENT OF
NAME OF THE DEPARTMENT

Multi- Omic approaches to unravel phenotypic responses of the tropical fish *Lutjanus guttatus* to ocean warming and acidification scenarios

David Jeremy Lohmann

BSc in Biology

Adviser: Sara Carolina Gusmão Coito Madeira
Junior Researcher, NOVA University Lisbon

Co-advisers: Pedro M. Costa
Professor Auxiliar, NOVA University Lisbon

Examination Committee:

Chair: Paula Maria Theriaga Mendes Bernardo Gonçalves,
Associate Professor, NOVA School of Science and Technology,
NOVA University Lisbon

Rapporteurs: Bernardo Duarte,
Researcher, MARE – Universidade de Lisboa

Adviser: Sara Carolina Gusmão Coito Madeira,
Junior Researcher, NOVA University Lisbon

MASTER IN MSC BIOINFORMATICS AND COMPUTATIONAL BIOLOGY

NOVA University Lisbon
September, 2024

Multi- Omic approaches to unravel phenotypic responses of the tropical fish *Lutjanus guttatus* to ocean warming and acidification scenarios

Acknowledgments

Thank you to my supervisors Carolina Maderia and Pedro M. Costa for providing the opportunity and welcoming me to the SeaTox Lab.

To Carolina for all the patience, guidance and always helping when I was in need.

To Pedro for motivating me and always being able to help when I was stuck.

Thank you to the whole SeaTox team for the support and welcoming atmosphere.

To Inês for helping with my server problems, and to Ana and Marialena for the support and always being there when I needed help.

To João for always being there to talk and for the good times we had together.

A special thanks goes to Corinna for taking me in and always being there for me for the past two years.

.

ABSTRACT

Anthropogenic CO_2 emissions are contributing to rising seawater temperatures and ocean acidification, posing significant challenges for marine fish, which must either adapt physiologically or migrate to avoid deteriorating environmental conditions. This study aimed to investigate the molecular networks and biological pathways underlying the response of the spotted rose snapper (*Lutjanus guttatus*) to a mid-century scenario of ocean warming and acidification. To achieve this, quantitative metabolomics and proteomics were employed to analyze the livers of fish exposed for 28 days to two fully crossed experimental treatments (28 ºC – control, 30 ºC – warming) and pCO_2 (400 ppm – control, 700 ppm – acidification). Out of 45 identified primary metabolites, only glucose was significantly downregulated, while proteomics analysis identified 27 differentially regulated proteins (DRPs) across all contrasts. Pathway analysis revealed no significantly enriched pathways at the metabolite level, with retinol metabolism emerging as the only enriched pathway at the protein level. Ocean acidification during current temperatures did not affect protein regulation. However, when increased pCO_2 levels coincided with elevated temperatures, lipid metabolism was altered and retinol metabolism showed increased activity. In response to elevated temperature *L. guttatus* exhibited a clear molecular stress response, characterized by the upregulation of HSP90 regardless of pCO_2 treatment. Glucose concentrations decreased when increasing temperatures ambient pCO_2, which is the opposite reaction of what has been reported in literature. This suggests that *L. guttatus* may exhibit a unique metabolic response to warming. Under the most extreme conditions, glucose levels returned to control values. The combined stress of warming and acidification led to increased cellular repair investment, demonstrating the intricate interplay between temperature and pCO_2 on fish liver physiology. These findings suggest that *L. guttatus* is capable of responding to climate stressors through non-standard metabolic adaptations.

Keywords: : fish physiology, multi- omics, ocean warming, acidification, climate change

RESUMO

As emissões antropogénicas de CO_2 estão a contribuir para o aumento das temperaturas da água do mar e para a acidificação dos oceanos, colocando desafios significativos para os peixes marinhos, que devem adaptar-se fisiologicamente ou migrar para evitar o agravamento das condições ambientais. Este estudo teve como objetivo investigar as redes moleculares e as vias biológicas subjacentes à resposta do pargo manchado (*Lutjanus guttatus*) a um cenário de aquecimento e acidificação dos oceanos previsto para meados do século. Para tal, foram empregues abordagens de metabolómica e proteómica quantitativas para analisar os fígados de peixes expostos durante 28 dias a dois tratamentos experimentais cruzados (28 ºC – controlo, 30 ºC – aquecimento) e pCO2 (400 ppm – controlo, 700 ppm – acidificação). Dos 45 metabolitos primários identificados, apenas a glicose foi significativamente regulada em baixa, enquanto a análise proteómica identificou 27 proteínas diferencialmente reguladas (PDRs) em todos os contrastes. A análise das vias metabólicas não revelou vias significativamente enriquecidas ao nível dos metabolitos, com o metabolismo do retinol a surgir como a única via enriquecida ao nível das proteínas. A acidificação dos oceanos nas temperaturas atuais não afetou as condições de regulação das proteínas. No entanto, quando níveis elevados de pCO_2 coincidiram com temperaturas elevadas, o metabolismo lipídico foi alterado e o metabolismo do retinol apresentou atividade aumentada. Em resposta ao aumento da temperatura, o *L. guttatus* exibiu uma clara resposta molecular de stress, caracterizada pela regulação em alta da HSP90, independentemente do tratamento com pCO_2. As concentrações de glicose diminuíram com o aumento das temperaturas e pCO_2 ambiente, o que contraria o que tem sido relatado na literatura. Isto sugere que o *L. guttatus* pode apresentar uma resposta metabólica única ao aquecimento. Nas condições mais extremas, os níveis de glicose retornaram aos valores de controlo. O stress combinado de aquecimento e acidificação levou a um aumento do investimento na reparação celular, demonstrando a interação complexa entre temperatura e pCO_2 na fisiologia do fígado dos peixes. Estes resultados sugerem que o

L. guttatus é capaz de responder aos fatores de stress climáticos através de adaptações metabólicas não convencionais.

Palavras chave: fisiologia dos peixes, multi-ómicas, aquecimento oceânico, acidificação, mudanças climáticas

CONTENTS

1 INTRODUCTION 1

1.1 Global climate change 1

1.2 Ocean warming and acidification in the Eastern Tropical Pacific Ocean 2

1.3 Climate change impacts in tropical marine fauna and ecosystems 3

1.4 Stress physiology 5

1.4.1 Impacts of ocean warming 5

1.4.2 Impacts of ocean acidification 7

1.5 Multi- omics approaches to unravel response mechanisms to climate stressors at the molecular and cellular levels 8

1.5.1 The use of unconventional model species in omics studies 10

1.6 The fish model species 11

1.7 Thesis objectives and hypotheses 11

2 MATERIAL AND METHODS 13

2.1 Experimental design and fish sampling 13

2.2 Primary Metabolome profiling and quantification 15

2.3 Transcriptome assembly as a reference database for proteomics 16

2.4 Transcriptome assembly and annotation 17

2.5 Tandem mass tag (TMT) Proteomics 17

2.6 Bioinformatics 19

2.6.1 Metabolite analysis 19

2.6.2 Proteomic Analysis and Data integration 20

3 RESULTS 21
3.1 Fish metabolomic profiles across stress profiles 21
3.2 Proteomic analysis and data integration 26
4 DISCUSSION 37
4.1 Ocean acidification has only a marginal effect in fish metabolite and protein profiles 37
4.2 Temperature increase induces a cellular stress response 39
4.3 pCO_2 seems to affects the nature of the thermal response 41
4.4 Challenges working with non-model organisms 42
5 CONCLUSIONS 43

List of Figures

Figure 1.1. Currents and upwelling characterizing the Eastern Tropical Pacific 3
Figure 1.2. Illustration of the omics cascade 9
Figure 2.1. Schematic illustration of the experimental setup 15
Figure 3.1.1. Scatterplot of the first 2 components of the sPLS- DA analysis 24
Figure 3.1.2. Violin plot showing alpha – D – Glucose concentrations in all condition 26
Figure 3.1.3. Heatmap showing the relative abundances of each metabolite 27
Figure 3.2.1. Pie chart showing the number of identified proteins 28
Figure 3.2.2. Volcano plots of the differentially regulated Proteins found in all conditions ... 30
Figure 3.2.3. Heatmap illustrating expression levels of all unique differentially regulated proteins 32
Figure 3.2.4. PLS- DA analysis 33
Figure 3.2.5. Boxplots showing the log abundances of proteins 35
Figure 3.2.6. Dot plot showing the top 10 most enriched pathways for each of the four contrasts 37
Figure A.1. PCA scores plot of the first two components based on metabolite concentrations 65
Figure A.2. Scatterplot showcasing the differential abundance of all detected metabolites ... 66
Figure A.3. Bar plot containing the total number of up and down regulated genes per contrast 67
Figure A.4. Dot plot of top 25 enriched Metabolite sets in all contrasts 68

List of Tables

Table 3.1. One way- ANOVA results 25
Table 3.2.1. Number of up and downregulated proteins for each contrast 31
Table A1. List of all samples used in this work 69
Table A2. All identified differentially regulated Proteins found in all contrasts 70
Table A.3. Complete list of metabolites identified by LC- TOF- MS 71

ACRONYMS

ANOVA	Alcohol Dehydrogenase 1
BLAST	Analysis of Variance
DRP	Basic Local Alignment Search Tool
ETP	Differentially Regulated Proteins
FBP	Eastern Tropical Pacific
FC	Fructose-bisphosphatase 1
G6P	Fold Change
GSEA	Glucose-6-phosphate
GC-TOF MS	Gene Set Enrichment Analysis
ICD	Gas Chromatography Time-of-Flight Mass Spectrometry
HSP	Isocitrate Dehydrogenase
ICD	Heat Shock Protein
IPCC	Isocitrate Dehydrogenase
KEGG	Intergovernmental Panel on Climate Change
m/z	Kyoto Encyclopedia of Genes and Genomes
NADPH	Mass-to-charge ratio
NES	Nicotinamide Adenine Dinucleotide Phosphate
ORF	Normalized Enrichment Score
PCA	Open Reading Frame
PLS-DA	Principal Component Analysis
PPP	Partial Least Squares Discriminant Analysis
PC	Pentose Phosphate Pathway
pCO_2	Principal Component Partial Pressure of Carbon Dioxide
SDS	Sodium Dodecyl Sulfate
SSP	Shared Socioeconomic Pathway

SST	Sea Surface Temperature
Swiss-Prot	"UniProtKB/Swiss-Prot" (reviewed manually annotated)
TCA	Tricarboxylic Acid Cycle
TMT	Tandem Mass Tag
TEP	Tropical Eastern Pacific
UNIPROT	The Universal Protein Resource
VIP	Variable Importance in Projection

.

1

INTRODUCTION

1.1 Global climate change

Anthropogenic climate change is already impacting human and ecosystems globally and is likely to get more intense. Even under the most hopeful scenarios (SSP1, IPCC, 2023) predict an increase in mean global temperatures of 1.5°C (Rogelj *et al.*, 2018)and in sea surface temperature (SST) by 0.8-to 1 °C the end of the century (Leung *et al.*, 2022). Under the SSP5- 8.5 scenarios, which predict continued high CO_2 emissions, SST is projected to rise 1.5- 2 °C by 2050 and up to 4 °C by the end of the century (Escalante-Rojas *et al.*, 2018; Lucano-Ramírez *et al.*, 2023). Ecosystems around the world are already experiencing significant change due to anthropogenic climate change, as species all over the world are becoming threatened and are migrating polewards to escape rising temperatures (Loarie *et al.*, 2009; Gattuso *et al.*, 2015).

The anthropogenic impact is especially great in the oceans, since it has been a buffer for global climate change, absorbing up to 80% of the excess heat and dissolving excess CO_2, increasing its acidity (Halpern *et al.*, 2008; Venegas, Acevedo and Treml, 2023). While this might have protected us from the worst impacts of global warming, it also has drastically changed ocean physics, chemistry, ecology and the societal services it provides (Gattuso *et al.*, 2015). Increased CO_2 in the atmosphere has changed ocean chemistry by decreasing pH, consequently consuming carbonate ions (CO_3^{2-}) in the water to form carbonic acid, impeding calcification. Because of this, marine ecosystems are falling increasingly under stress. In the context of marine ecology, ocean warming, and acidification have been identified two of the main stressors affecting marine organisms around the globe (Wernberg, Smale and Thomsen, 2012; Bass *et al.*, 2021).

The effect this has on fishes is of particular interest, since fish are integral for the food supply of human populations around the world, especially coastal ones, and predicting how fishes will respond to climate change is essential for formulating effective ocean management, governance measures and harm reduction policies. Since the effects of climate change can no longer be fully mitigated the concept of adaptation to rising temperatures has increasingly come into the focus of researchers and policymakers (Dewulf, 2013). This approach focuses on increasing the resilience of communities and ecosystems to climate change. In the context of marine ecosystems increasing the role of local communities, adaptive management of bioresources and focusing on conservation of biodiversity are especially important (Ward *et al.*, 2022).

1.2 Ocean warming and acidification in the Eastern Tropical Pacific Ocean

The Eastern Tropical Pacific (ETP) is a geographic area that stretches from the gulf of California to the north of Peru (Lavín *et al.*, 2006). It contains the coastlines of 12 different counties, as well as several Islands archipelagos. It Is delineated in the north and south by different upwelling systems which lead to decreased water temperatures in these areas. To the west it borders the central American landmass and in the east by the East Pacific Barrier, a wide stretch of open ocean which acts as a biogeographic barrier(Briggs, 1961).

In the ETP there is an area of increased water temperature, known as the Eastern Pacific Warm Pool (EPWP) which is characterized by having a mean annual sea surface temperature of 28 °C or above (Fiedler *et al.*, 2013). This warm pool shifts seasonally north- or southward reflecting seasonal changes in the atmosphere (Ando and Hasegawa, 2009), leading to a high annual variability in SST throughout the region. Additionally, the ETP experiences significant interannual variation, of which the largest contributor is the El- Niño- Southern- Oscillation (Fiedler, 2002), a large-scale periodic shift in atmospheric and ocean water temperature in the Pacific which influences local and global climate (Alexander, 1992). This results in the Eastern Tropical Pacific being very heterogeneous, both on a spatial and a temporal scale.

The ETP contains three main distinct biogeographical zones, each with a distinct fauna; Baja California, the Panamic zone and the various Archipelagos and Islands off the coast (Robertson and Cramer, 2009). The ETP offers incredibly varied and distinct habitat which leads to one of the most diverse and productive marine ecosystems on the planet. It contains three UNESCO world heritage sites (the Galapagos Islands, Cocos Island and Coiba national park).

Climate change has been projected to especially impact the warm pool more than the neighboring areas with colder SST (Clarke *et al.*, 2021).

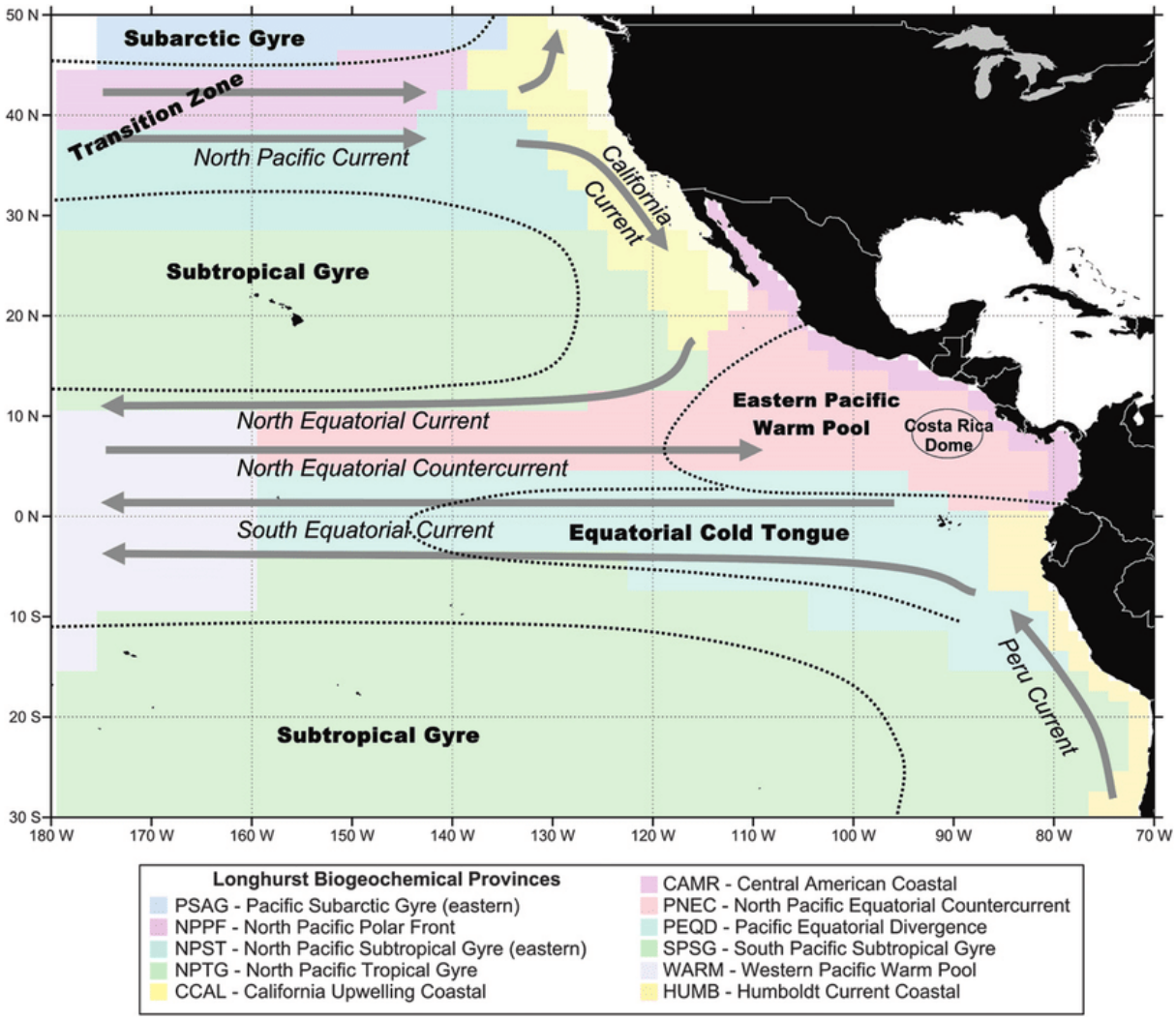

Figure 1.1. Currents and upwelling characterizing the Eastern Tropical Pacific, from Fiedler *et al.*, (2013).

1.3 Climate change impacts in tropical marine fauna and ecosystems

Assessing how species adapt to a changing climate involves the interplay of multiple factors and biological processes, such as the physiological responses of various species at different life stages, population dynamics, phenology (e.g. migration), and behavioral changes (Mora and Robertson, 2005). In the Eastern Tropical Pacific, which is home to numerous migratory fish species and endemic species with small ranges, these factors are particularly

significant (Mora and Robertson, 2005). Migratory species are crucial for ecosystem connectivity, yet this could be disrupted by range shifts or temperature-induced changes in ocean currents, as previously observed during El Niño events (Hearn *et al.*, 2010; Bessudo *et al.*, 2011; Baums *et al.*, 2012; Ganachaud *et al.*, 2013; Cortés *et al.*, 2017). Fish species with limited ranges are especially vulnerable to climate change, as they have less capacity to migrate and evade unfavorable conditions. Tropical marine species, in general, are more susceptible to climate change due to the low variance in yearly temperatures compared to temperate oceans, which leads to them to being warm adapted, but often having very narrow thermal ranges of only a few degrees (Tewksbury, Huey and Deutsch, 2008) which makes them less resilient to rising temperatures and increased environmental variability (Boyd *et al.*, 2016; Vinagre *et al.*, 2016). Several fish species from Papua New Guinea (Western Tropical Pacific) for example -were found to already live above their thermal optimum, showing significantly reduced performance to a temperature increase of +3 °C (Rummer *et al.*, 2014).

As environmental stress from climate change intensifies, organisms are pushed towards the edges of their ecological niches, leading them to either adapt or rely on phenotypic plasticity, which is the ability to adjust phenotypes without genetic changes (Reusch, 2014; Bernal *et al.*, 2022). Climate change is likely to increase oceanic variability through extreme weather events and marine heatwaves (Widlansky, Long and Schloesser, 2020), further challenging species to adapt. The restructuring of marine ecosystems is anticipated as organisms from different trophic levels respond differently to climate drivers. For instance, marine phytoplankton, the primary producers in the ocean, can quickly adapt to changing temperatures due to their short generation times and high thermal plasticity. Coral reefs, particularly in the tropics, face significant threats from acidification and warming, leading to widespread coral bleaching events that endanger entire ecosystems by removing essential habitats (Andini *et al.*, 2024). Beyond large-scale changes like sea temperature increases, small-scale processes, such as microclimates, behavioral adaptations, and physiological variability, also play a crucial role in species' responses to climate change (Dong *et al.*, 2017).

Species often respond to changing environments by migrating to conditions that better match their ecological niches, such as moving poleward or into deeper waters, a trend already observed in many marine species (Assan *et al.*, 2020). In the North and Baltic Seas, temperature has been identified as a key factor driving fish abundance, with further distributional changes likely in the future (Rutterford *et al.*, 2023). However, the situation is complex, as animals have various ways of responding to climate change—or their populations may risk local or regional extinction. Estuarine systems for instance are particularly at risk, with water temperature changes affecting species compositions and community assemblage (Anthony *et al.*, 2009; Hernández-Álvarez *et al.*, 2023). Ultimately, all higher-order effects stem from the physiological

responses of species or those they interact with, making the understanding of the underlying molecular mechanisms of responses critical for predicting species plasticity and population trajectories under climate-related stressors (Rijnsdorp *et al.*, 2009).

1.4 Stress physiology

In the research field of climate change biology, a robust understanding of physiological responses and the systems which facilitate them is indispensable (Huey *et al.*, 2012; Bass *et al.*, 2021), providing mechanistic explanations for higher-order changes like behavioral shifts or biogeographical movements (Somero, 2011). Physiological research is able to reveal how organisms adapt dynamically to changing environments (Reusch, 2014). Climate change is presenting species with multiple stressors at once, and recognition of the response mechanisms both alone and cumulatively is instrumental when trying to predict the transformations marine environments will go through and managing them accordingly (Wernberg, Smale and Thomsen, 2012; Boyd *et al.*, 2018).

The recent rapid technological advancements in the fields of Omics and Bioinformatics allow a unique opportunity for characterizing these mechanisms. One of the most powerful concepts hereby is the integration of information obtained from different omics data types, as it allows one to assess the biological pathways and mechanisms at multiple levels of cellular regulation (Manzoni *et al.*, 2018). With the number of options of analysis and the amount of data possibly generated ever increasing, the possibilities become almost endless, and it becomes necessary to carefully choose the methodologies best suited to answer the research question asked.

1.4.1 Impacts of ocean warming

Stress exerts a negative influence on body homeostasis in both vertebrates and invertebrates, (Sokolova *et al.*, 2012; Deck *et al.*, 2017; Gómez-Reyes *et al.*, 2023). In the former, abiotic stress triggers a range of physiological effects, especially in the liver, which is the organ that mainly regulates energy metabolism, as well as the immune response and detoxification processes (Bruslé and Gonzàlez, 2017). Among the various factors that can induce stress, temperature is particularly significant. Fish, being poikilothermic, have body temperatures that are directly influenced by the surrounding water, making them especially sensitive to fluctuations in ocean conditions. Changes in water temperature have been shown to affect a variety of performance parameters (Enes *et al.*, 2009). For instance, studies have found that with rising temperatures, food intake, feed efficiency, and the hepatosomatic index increases in species like

gilthead seabream (*Sparus aurata*) and European seabass *(Dicentrarchus labrax)* (Couto *et al.*, 2008; Zhang *et al.*, 2021).

As a general response, fishes have to boost their metabolic rate in order to keep up with the increased energy demand of maintaining homeostasis during stressful conditions leading to an increase in oxygen demand and a general decrease in aerobic scope (Sokolova, 2013) due to a greater energy expense in basal metabolism as well as cellular repair mechanisms (Claireaux and Lagardère, 1999; Cheng *et al.*, 2018). Aerobic scope is defined as the difference between resting and maximum metabolic rate. As a result, fish are either able to thermally compensate for their metabolism and acclimate to new conditions or they may have to upregulate anaerobic metabolic pathways to meet energy demands (Johansen *et al.*, 2021). For example, the coral reef snapper *Lutjanus carponotatus* showed increased oxygen consumption and lactate accumulation after marine heat waves (McMahon, Munday and Donelson, 2024), upregulation of glycolytic pathways in the American shad *Alosa sapidissima* (Luo *et al.*, 2024) and increased glucose levels in the freshwater carp *Labeo rohita* (Kumar *et al.*, 2015) in response to heat stress. Loss of aerobic scope was found in several fish species following rising temperatures. A mismatch between oxygen supply and demand has been suggested to be one of the main mechanisms setting thermal limits in marine organisms (Ern, Andreassen and Jutfelt, 2023). Still, there are other approaches to explain the physiological effect of above optimal temperatures on fishes, like the concept of "aerobic scope protection" (Jutfelt *et al.*, 2021).

Another significant impact of higher temperatures is the induction of oxidative stress. Thermal stress increases the production of reactive oxygen species (ROS) through increased metabolism with lower mitochondrial efficiency, potentially leading to oxidative damage within cells (Lushchak, 2011; Benedetti *et al.*, 2022), a process which is often linked to a higher oxygen consumption (Thoral *et al.*, 2022). Extreme temperatures have been found to elevate the expression of genes related to oxidative stress in estuarine fish (Madeira *et al.*, 2013), and increase oxidative biomarkers in *Hoplosternum littorale* (Rossi, Bacchetta and Cazenave, 2017). Additionally, heat stress can differentially regulate immune response genes, as the immune system often functions as a general stress response mechanism in many fish species (Tort, 2011; Mariana, Alfons and Badr, 2019). This complex interplay of metabolic and physiological responses underscores the profound impact that rising temperatures can have on marine organisms, particularly fish, which are highly sensitive to their thermal environment. Understanding these underlying molecular mechanisms is crucial for predicting how species will respond to ongoing climate change and its associated stressors. One model that attempts to integrate the responses organisms under different climate stressors exhibit is the Oxygen- and Capacity-Limited Thermal Tolerance (OCLTT) model (Pörtner, 2012). According to this model,

oxygen availability serves as the primary mechanistic link between several climate stressors such as temperature, hypoxia, hypocapnia and pollution

1.4.2 Impacts of ocean acidification

The reduced carbonate availability caused by increased absorption of atmospheric CO_2 into the seawater has had detrimental effects on many marine calcifying organisms, such as corals and some calcifying marine plankton like Coccolithophores, leading to habitat destruction and disruptions in biological systems like the carbon pump (Gattuso *et al.*, 1998; Wolf-Gladrow *et al.*, 1999). Teleost fishes also possess calcifying organs called otoliths, which are vital for hearing and balance (Schulz-Mirbach *et al.*, 2019). However, they can efficiently regulate their internal pH by increasing bicarbonate concentration in their cells or excreting hydrogen ions through their gills and intestines (Pörtner, Langenbuch and Reipschläger, 2004), making them much less vulnerable to ocean acidification. A 2019 review by Martin Grosell found that 11 studies reported an impact of CO_2 on otoliths, while six found no effect, with the differences primarily driven by species variations rather than the degree of pCO_2 increase or exposure time (Mirasole *et al.*, 2017; Grosell, 2019). Interestingly, when effects were observed, hypercapnia induced otolith growth rather than reducing it, because fishes react to a lower pH by increasing plasma carbonate concentration, which promotes calcification (Brauner *et al.*, 2019; Coll-Lladó *et al.*, 2021). Acidification has also been linked to behavioral changes in fish, such as disorientation and reduced prey-catching or foraging abilities (Watson, Fields and Munday, 2017; Jiahuan *et al.*, 2018).

A high degree of variability in the response to hypercapnia (an increased partial blood CO_2 pressure) across species is common, underscoring the need for a deeper understanding of the mechanisms of adaptation in fish (Sundin, 2023). Physiologically, hypercapnia can affect respiration, metabolism, and circulation in some species (Ishimatsu *et al.*, 2005). Increased pCO_2 has been shown to raise oxygen consumption in fish like rainbow trout and skate (Thomas *et al.*, 1983; Graham, Turner and Wood, 1990), though species like the carp *Cyprinus carpio* showed no response (Takeda, 1991). Although fish can generally cope with decreased pH levels, the additional energy demands for osmoregulation under elevated pCO_2 can lead to heightened oxidative stress and impact immune responses (De Souza *et al.*, 2014). Acidification may also amplify thermal stress on fish by acting on similar molecular mechanisms, with it and warming often showing synergistic effects (Servili *et al.*, 2022).

Furthermore, lower pH levels are often associated with hypoxic conditions in marine ecosystems, for example due to upwelling in coastal areas, potentially exacerbating stress for fish already dealing with increased oxygen demands. Some species in aquaculture have been found to tolerate pH levels exceeding those predicted by climate change models (Ishimatsu,

Hayashi and Kikkawa, 2008). The effects of acidification on general aerobic scope vary, with both positive and negative outcomes reported across different species (Munday, Crawley and Nilsson, 2009; Couturier *et al.*, 2013; Hannan and Rummer, 2018). While higher temperatures lead to increased growth rates, acidification tends to have the opposite effect (Moreira *et al.*, 2022). Much of the research on acidification has focused on freshwater fish, particularly in relation to acid rain, or in aquaculture environments where local pCO_2 levels can surpass climate change projections, making it challenging to assess the relevance of such findings in the context of global climate change impacts on marine environments without further research (Ishimatsu, Hayashi and Kikkawa, 2008).

1.5 Multi- omics approaches to unravel response mechanisms to climate stressors at the molecular and cellular levels

Proteomics and Metabolomics are both at the end of the omics cascade. That means they are able to represent a snapshot of the physiological response of an organism to stress, especially through the integration of different techniques. Metabolomics are the omics field most closely related to the phenotype, and therefore able to show the most accurate picture of the actual cellular processes affected by different stressors (Alfaro and Young, 2018). Previous metabolomic studies on fishes have been able to demonstrate the immediate consequences of heat stress in several organs, particularly highlighting an increased metabolism and protein catabolism and lipid metabolism (Li *et al.*, 2022; Chen *et al.*, 2023; Luo *et al.*, 2024). Due to the highly diverse nature of metabolites in an organism, for example ranging from small nucleotides to huge macromolecules like starch and lipids, getting a comprehensive picture of all metabolites in a sample can be technically challenging (Steinhauser and Kopka, 2007). Because of this, metabolomic works often must limit themselves to study a specific group of metabolites which may be most relevant to the research question at hand. In this work only primary metabolites, which are those involved in growth, development and reproduction, were analyzed. One major challenge when working with metabolomics is that, since metabolites abundances can be affected by multiple different pathways, real biological systems can be lost in the noise (Johnson and Gonzalez, 2012). One way to try to alleviate this problem is to integrate these analyses with different data types.

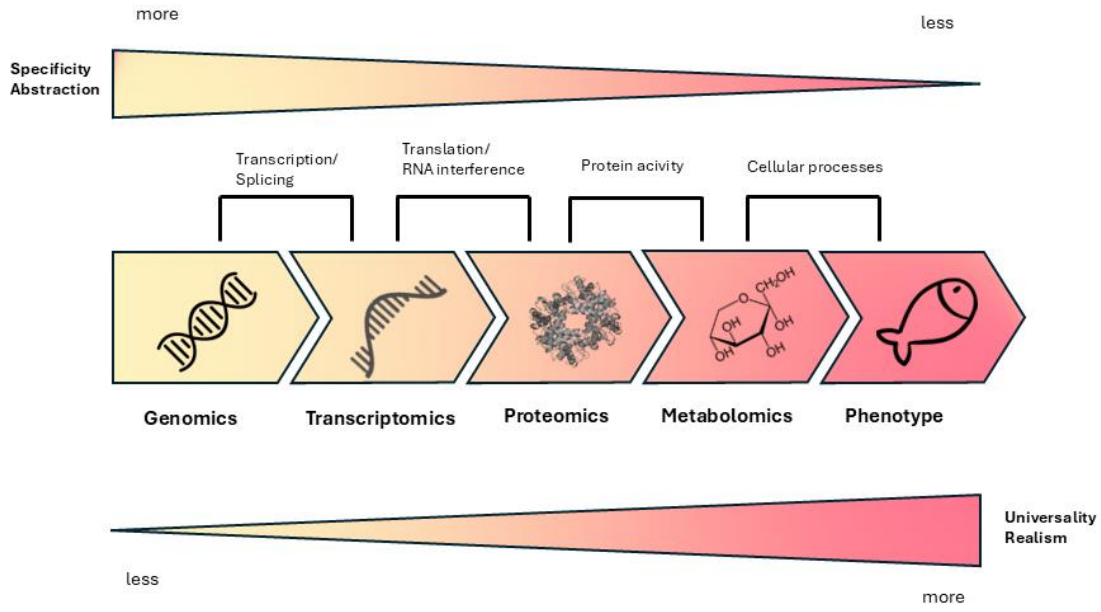

Figure 1.2. Illustration of the omics cascade and highlighting the tradeoffs between the different omics types

Proteomics is becoming more and more popular when researching climate change impacts, because modern high throughput techniques allow to simultaneously observe the potential change of thousands of molecules in response to a stressor (Sanchez, Ralston-Hooper and Sepúlveda, 2011). One of the most popular modern methodologies for high throughput proteomics is liquid chromatography tandem mass spectronomy (LC- MS/MS). The resulting spectral patterns contain information about the amino acid sequence of the peptides. These patterns are incredibly complex, so extracting the full sequence can be a challenging task. *De novo* sequencing is possible, but computationally very expensive and relies on high quality spectral data (Hoopmann and Moritz, 2013), so a more commonly used method is database searching. Here proteins with an already known sequence are digested *in silico* and their peptide spectral patterns are predicted. These are subsequently matched against the real spectral files. This method is much more inexpensive, but it relies on comprehensive and accurate sequence data of the proteome of the studied species (Marquioni, Nunes and Novo-Mansur, 2021). This is usually not a problem when working with well-studied model species but can pose a challenge when trying to investigate less well-studied organisms.

1.5.1 The use of unconventional model species in omics studies

Protein identification is not the only instance where bioinformatic analyses rely on large amounts of biological data (Joyce and Palsson, 2006). Detection of some difference in regulation of a given gene or protein means little without knowledge of the molecular systems it interacts with and is a part of. Since modern omics approaches can identify hundreds or even thousands of differentially regulated proteins or genes it is necessary for large amounts of prior knowledge to be available in order to apprehend biological meaning from such results (Joyce and Palsson, 2006). To this end there are a multitude of databases providing easy access to different types of information that might be needed for a given analysis. UNIPROT is a databank that provides protein sequences, as well as functional information in the form of gene ontology terms, providing wider context of the molecular function, biological process or cellular component of a given protein. The KEGG pathway database (https://www.genome.jp/kegg/pathway.html) assigns proteins to a collection of pathway maps, which provides information on how different proteins might interact, and which molecular processes or networks might be affected by them.

The amount of information that is available naturally is bigger the better studied a given species is. *Homo sapiens*, for which the most data exists, has 204.500 entries in the UNIPROT database (https://www.uniprot.org, as of 19.08.2024). The zebrafish (*Danio rerio*), which is the model organism for fishes, has 47.566 entries.

In recent years there has been a push towards studying a greater variety of species, in order to capture a larger part of the diversity in a given ecosystem (Heck and Neely, 2020). This presents a challenge when trying to perform multi- omics analyses, since here the needed biological information for these species is often spotty at best or entirely non- existent at worst. The *Lutjanus* family, to which *Lutjanus guttatus* belongs and is the subject of this study (Section 1.6), only has 1,643 entries in UNIPROT, and none for *L. guttatus* specifically. Here one of the main strategies to still be able to achieve biologically interpretable results is to match the identified proteins sequences and metabolites to already known ones in a database based on sequence similarity or identical mass-to-charge ratio within a specific range of molecular weights. While this is a powerful approach, it is not without problems. The main assumption these methods rely upon is that similarity implies homology, and that this homology can be used to infer function (Ponting, 2001). However, homologous proteins for example do not necessarily share significant sequence similarity (Marquioni, Nunes and Novo-Mansur, 2021), and often have differing functions, especially if the two organisms are only distantly related (Naveenkumar *et al.*, 2022).

1.6 The fish model species

The spotted rose snapper *Lutjanus guttatus* is a tropical teleost fish from the *Lutjanidae* family, which comprises 113 species, that mainly inhabits the western central American coast from Mexico to Peru (Hernández-Álvarez *et al.*, 2020). *Lutjanus guttatus* is a generalist feeder with a preference of small crustaceans and fish (Soto-Rojas, Hernández-Noguera and Vega-Alpízar, 2018). It is found in coral reefs with a preference for rocky substrate in proximity to the coast (Del Mar Palacios and Zapata, 2014). It is currently not endangered and is classified as "least concern" on the IUCN red list (IUCN, 2023). Due to its abundance, it plays a large role in local fisheries in several regions. In Mexico around 450t are caught annually (FAO, 2020). They are mainly exploited by local artisanal fisheries and as bycatch of shrimp trawling (Andrade, 2003). In the Gulf of Nicoya in Costa Rica, Lutjanus guttatus made up 20% of all bottom longline fishery in the area (Araya *et al.*, 2007). Additionally, it is well suited for aquaculture, since it accepts artificial food and is fully grown within 8 months (Castillo-Vargasmachuca *et al.*, 2018). Despite its economic importance, there has been relatively little research published on the biology of this species (Lucano-Ramírez *et al.*, 2023), and most of what has been published was aquaculture- related research.

1.7 Thesis objectives and hypotheses

The main goal of this study was to investigate the molecular networks and biological pathways of response of the tropical fish *Lutjanus guttatus* to mid- century scenarios of ocean warming and acidification. In order to achieve this, the fish were previously exposed to two fully crossed temperatures (28 ºC – control, and 30 ºC - warming) and pH (8.0 – control and 7.8 – acidification) for 28 days.

To this end, the specific objectives of this work were:

Objective 1:

Identify and quantify the effect of the proteins and metabolites that are affected by the climatic stressors studied.

H1. We hypothesize that the exposure to stressful conditions will up-regulate proteins associated with the minimal stress proteome (e.g. heat shock proteins, antioxidants) to counteract oxidative stress, enhance cellular repair and maintain homeostasis.

H2. We hypothesize that the exposure to stressful conditions will up-regulate specific metabolites such as amino acids and glucose to be used as major cellular fuels and osmolytes during unfavorable environmental conditions.

Objective 2:

Characterize pathways and physiological mechanisms that are involved in the adaptation to these stressors by integrating both datasets.

Ultimately, this work aims to uncover new information on the physiological mechanisms of how a sparsely investigated tropical fish species responds to future climate change, helping to understand how it will respond to a changing climate. Climate change is already disrupting ecosystems and natural environments worldwide, triggering biodiversity reorganizations, with mobile species such as fish on the move to escape worsening conditions. This understanding is critical as oceans provide the main source of protein for over a billion people and are a contributing factor to food security for billions more (Creighton *et al.*, 2016). The multi-omics approaches used in this study provide a powerful tool to predict organism responses and population trajectories. This knowledge is essential for accurately implementing effective management strategies ensuring that communities reliant on these resources can continue to thrive in the face of climate change. In doing so, this work supports the United Nations' Sustainable Development Goals 2 (Zero Hunger) and 14 (Life Below Water).

2

MATERIAL AND METHODS

The experimental tasks of this study (section 2.1) were led by Dr. Sonia Bejarano and conducted by the Reef Systems Work Group of the Leibniz Zentrum für Marine Tropenforschung at Parque Marino del Pacifico (PMP), Punteras, Costa Rica, in collaboration with researchers from Universidad Nacional de Costa Rica and NOVA University of Lisbon. The remaining tasks (sections 2.2 to 2.6) were conducted at NOVA University of Lisbon, within the scope of this thesis.

2.1 Experimental design and fish sampling

Lutjanus guttatus (n= 24) juveniles were obtained at approximately 3 months old from Parque Marino del Pacifico, Costa Rica. Prior to the experiment started, the fish were housed in a quarantine tank at PMP facilities for 7 days and fed twice a day ad libitum with BioMar (Cañas, Costa Rica) EFICO sigma 5012 commercial pallets until they reached saturation level, amounting to approximately 0.3-0.5% of their body mass per day. Prior to being transferred to the experimental tanks (ETs), individual fish were lightly anesthetized, tagged using colored elastomer tags, measured, and weighed.

The experimental design was made up of 4 treatment systems, each one consisting of a reservoir tank, a sump and 4 replicate 200L experimental tanks, which were filled with seawater from the Gulf of Nicoya. The water pumped from the ocean was initially passed through a chemical and a UV filter and was then transferred to each system and re-circulated circulated at a rate of 1.8 L/min. Each system had its own set of filters in place consisting of a biofilter, sand filter, UV filter and a de- gasifier. Target temperatures were 28°C (ambient) and 30°C (elevated) ± 0.5 °C were attained in each system using a high- capacity water- cooling system

(Jandy versatemp Heat/Chill pump, 130,000BTUS) for the ambient water. The 30° C water system was self-regulating, with fluctuations throughout the day, but in the end stabilizing around 30°C ± 0.5 °C. The pH was either kept at ambient levels (average 7.97, range 7.95 to 7.99) or adjusted to -0.2 pH (average 7.77, range 7.76 to 7.79) below average using a Sera Flore CO_2 active reactor 1000 to diffuse CO_2 in the water, which was automatically regulated using n Apex controller system (Neptune Systems) connected to a solenoid valve. The 16 experimental tanks were well aerated with air stones deriving from a common aerating system, with >95% oxygen saturation, and backwashes of approximately 280L per treatment were performed twice a day in order to clean the filters and partially exchange the re- circulated water. Per tank there were 11 fishes, which were acclimated for 2 days at ambient Temperature and pH of the water coming from the Gulf of Nicoya, before adjusting the water to experimental conditions over a period of another 2 days. Additionally, a 12-hour light cycle was installed to simulate normal photoperiods.

In order to monitor the water quality throughout the experiment, temperature and pH were measured daily using handheld probes. Additionally, Ammonia, nitrite and nitrate levels were measured biweekly in 2 tanks per treatment using a standard water testing kit.

The Fish were sampled after 28 days, and beforehand were starved for 24 hours. Per tank two fishes were randomly selected, anesthetized and euthanized by cervical dissection. After this the livers were dissected, frozen, and total of 48 liver samples were sent to NOVA-FCT, Portugal for further analysis.

Frozen liver samples (n=24, 6 from each experimental condition) were cut approximately in half, and one piece was used for metabolome profiling (64 mg to 478.9 mg of fresh weight), whereas the second one was used for proteome profiling (n = 18).

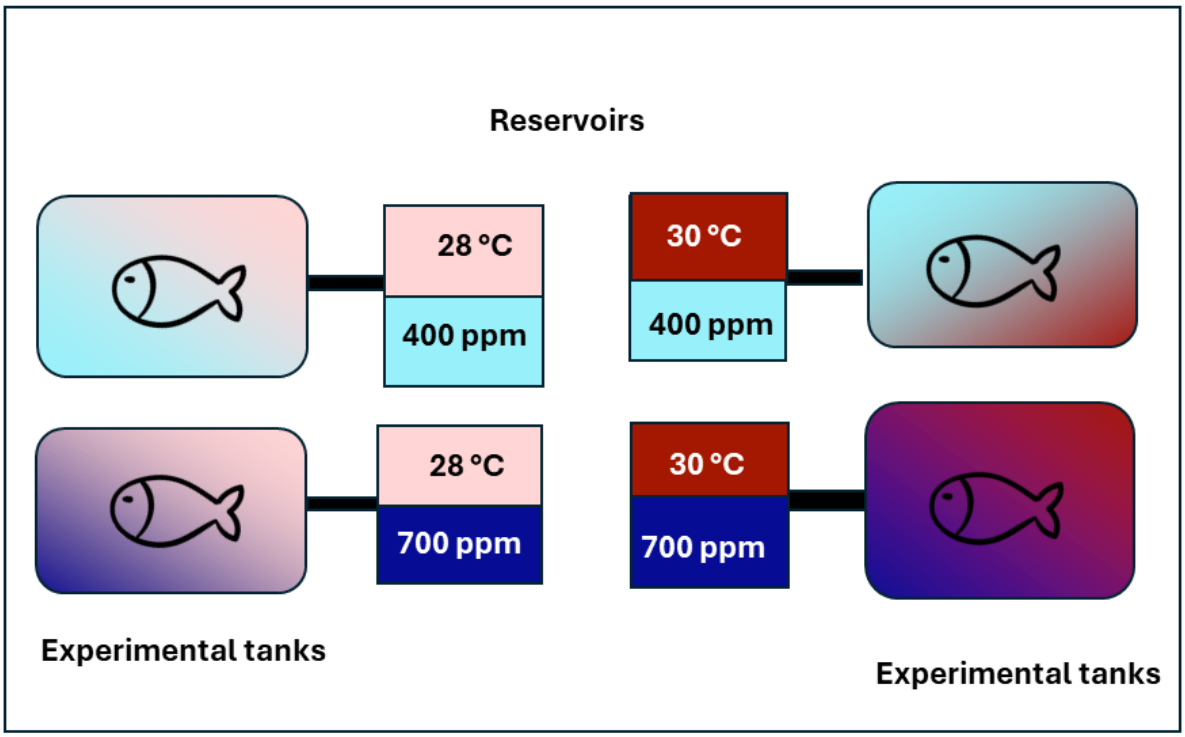

Figure 2.1. Schematic illustration of the experimental setup described in section 2.1. The water was treated in the reservoirs and then transferred into the experimental tanks where the fish were kept for the duration of the experiment. Control conditions are marked in black text and projected conditions in white text. The CO_2 levels were indicated in parts per million (ppm) of atmospheric CO_2 levels, resulting in average pH values of 7.97 (400 ppm CO_2) and 7.77 (700 ppm CO_2). The final Δ ppm CO_2 were 300 ppm, and the Δ pH were -0.2.

2.2 Primary Metabolome profiling and quantification

Frozen liver samples liver samples for metabolome profiling were subjected to further pre-processing were lyophilized and homogenized to a fine powder, then weighed again to determine and adjust their dry weight to the intended tissue mass. Each sample for analysis had a had a dry weight of 30 ± 1 mg. Samples were then shipped as frozen dry tissue in dry ice to the Plant Metabolomics Lab (PML) (https://www.plantmetabolomicslabpt.com), led by Dr. Carla António, based at CEF (Centro de Estudos Florestais) in Instituto Superior de Agronomia (Lisboa, Portugal).The methodology for primary metabolites extraction, detection and quantification were carried out in PML as a collaboration utilizing GC-TOF MS (Gas-Chromatography Time-Of-Flight Mass-Spectrometry), as follows.

Metabolite extraction was done in 700 µl methanol. As an internal standard, ribitol was used. The samples were incubated at 70 °C for 15 min while being vortexed in a shaker, then

centrifuged for 10 min at 12.000 × g. The supernatant was mixed with 370 Chloroform and 750 µl Water and again centrifuged for 15 min at 2.200 × g. 150 µl of the methanol phase was subsequently evaporated until dry at 30° C using a Vacufuge Plus centrifugal concentrator (Eppendorf, Hamburg, Germany).Trimethylsilyl (TMS) derivatization was carried out using 40 µL of 20 mg mL-1 methoxyamine hydrochloride in pyridine, followed by 70 µL of N-methyl-N-trimethylsilyltrifluoroacetamide and 20 µL mL-1 of a mixture of fatty acid methyl esters (FAMES). For the gas chromatography an Agilent 6890N gas chromatograph (Agilent Technologies, Böblingen, Germany) was used, in conjunction with a LECO Pegasus III TOF-Mass Spectrometer (LECO Instrumente, Mönchengladbach, Germany), which was running in electron ionization mode. The chromatographic separation was performed on a VF-5MS column (Varian Inc., 30 m length, 0.25 mm inner diameter, and 0.25 mm film thickness). The temperature program was as follows: isothermal for 2 min at 85 °C, followed by a 15 °C min-1 ramp to 360 °C, and hold at this temperature for 6 min. The injector and transfer line temperatures were set to 230 °C and 250 °C, respectively, and the helium carrier gas flow set to 2 mL min-1. After a solvent delay of 180 s, mass spectra were scanned from m/z 70–600 with acquisition rate of 20 spectra s-1 and a detector voltage between 1700 and 1850 V. The raw data was analyzed using AMIDS v. 2.71 (https://chemdata.nist.gov/dokuwiki/doku.php?id=chemdata:amdis). Annotation of the primary metabolites was done using TagFinder 4.0 (Luedemann *et al.*, 2008) using the Golm Metabolome Database as a reference library (Kopka *et al.*, 2005; Schauer *et al.*, 2005).

2.3 Transcriptome assembly as a reference database for proteomics

In order to improve the peptide detection rate and read depth, a reference library was constructed using RNA extracted from L. guttatus livers. RNA extraction was performed on not more than 30 mg of liver tissue using the Qiagen RNeasy Protect Mini Kit (Qiagen, Hilde, Germany) with an adapted protocol for animal tissues. Approximately 30 mg of *L. guttatus* liver tissue were put into an Eppendorf tube with a punctured cap and snap frozen using liquid nitrogen. The tissue was homogenized using a plastic pestle and 600 μl RLT Buffer plus with β-Mercaptoethanol (β-ME) (10 μl per 1ml Buffer) was added to further disrupt the tissue. The tissue was then further homogenized by passing the lysate 10- 15 times through a 23-gauge needle and syringe. After 4 minutes of centrifugation at 10.000 rpm the supernatant was transferred into a separate microcentrifuge tube.

One volume of 50% Ethanol was then added to the supernatants and the whole sample was transferred onto a supplied AllPred RNA spin column and centrifuged for 15 seconds at 10.000 rpm. The flow through was discarded. The spin column was washed by adding 350 µl RW1 buffer on the column and centrifuging again for 15 seconds at 10.000 rpm. On- column DNase digestion was accomplished by first mixing 10 µl DNase and 70 µl RDD buffer for each sample and pipetting it onto the spin column. Since the DNase is very sensitive mixing was done gently and without vortexing. Afterwards the samples were incubated for 15 minutes at room temperature on the workbench. Then 350 µl buffer RW1 was added and again centrifuged at 10.000 rpm for 15 seconds. The RNA spin column was cleaned again with 500 µl buffer RPE and again centrifuged. The final wash step was done with another 500 µl RPE buffer and centrifugation for 3 minutes. RNA elution was performed by placing the column in a 1.5 ml Eppendorf tube and adding 30 µl RNAse free water on the column with 4 minutes of centrifugation at 10.000 rpm. This step was done twice in order to increase RNA yield.

RNA quality was assessed via Nanodrop (Thermo Fisher, Scientific, Waltham, MA, USA). RNA -Sequencing was performed by Stabvida (https://www.stabvida.com).

2.4 Transcriptome assembly and annotation

Raw data quality was assessed using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/), then adapters and low-quality reads were removed with TrimGalore v. 0.6.10 8 (https://www.bioinformatics.babraham.ac.uk /projects/trim_galore/). De novo transcriptome assembly was accomplished using Trinity v2.14.0 (Grabherr *et al.*, 2011). Bowtie 2 (Langmead and Salzberg, 2012) and Trinity was used for quality control and calculating ExN50 statistics. Open reading frames were predicted using the TransDecoder v5.5.0 software (Haas, BJ. https://github.com/TransDecoder/TransDecoder) was used. Functional annotation was performed by BLAST against the reviewed Swiss- Prot Database from 19.02.2024 (https://ftp.uniprot.org/pub/databases/uniprot/current_release/knowledgebase/complete/). In order to maximize the search space in the proteomic database search, no e- value cutoff was used

2.5 Tandem mass tag (TMT) Proteomics

Eighteen samples (4 ambient temp & ambient pH, 5 ambient temp & acidified, 5 medium temp & acidified, 4 medium temp. & ambient pH) were selected for TMT mass spectronomy. About half of the whole liver was used per sample. Digestion, Separation, tagging and

spectronomy as well as protein identification and quantification was carried out by SciLifeLab (https://www.scilifelab.se/). The samples were first pulverized with a cryoPREP extraction system and lysed in 300 μl Lysis buffer (4% SDS, 50 mM HEPES pH 7,6, 1 mM DTT), afterwards heated to 95 °C and sonicated. The preparation of the proteins for mass spectronomy was done using a modified version of the SP3 protein clean up and digestion protocol by Moggridge et al. (2018) and Hughes et al. (2014). 200 μg of protein per sample was mixed with 4 mM Chloroacetamide. 30 μl of the Sera- Mag SP3 bead mix was then added in conjunction with 100% Acetonitrile up until a concentration of 70%. The samples were then incubated for 18 minutes at room temperature, and afterwards the tubes were placed on a magnetic rack and the supernatant was discarded. The remaining beads were washed with 70% ethanol and then again using 100% acetonitrile. Finally, they were incubated overnight in 100 μl Typrsin buffer (50 mM HEPES pH: 7.6, 10 mM $CaCl_2$, 10% Acetonitrile and 1:50 enzyme (Trypsin) to protein ratio). Afterwards, 20 µl of Sera-Mag SP3 bead mix (10 µg/µl) was added to the sample. Then, 100% acetonitrile was introduced to achieve a final concentration exceeding 95%. The samples were mixed by pipetting and left to incubate at room temperature for 18 minutes before being placed on a magnetic rack. After aspirating and discarding the supernatant, the beads were washed with 200 µl of acetonitrile. The samples were then removed from the magnetic rack, and the beads were reconstituted in 200 µl of MQ water, followed by peptide concentration assessment using the Bio-Rad DC Assay. After pH adjusting the samples with TEAB pH 8.5 (100 mM) they were then labelled with the 18 different isobaric TMT- tags following the manufacturer's protocol (Thermo Scientific). The labelling efficiency was assessed by LC-MS/MS. Sample clean-up was done by solid phase extraction (SPE strata-X-C, Phenomenex) and subsequent drying in a SpeedVac. 10 µg of the mixed samples were used in the Liquid chromatography mobile phase, of which 1 µg was inserted into the LC- MS/MS system.

LC- MS was performed following (Branca *et al.*, 2014) utilizing a Dionex UltiMate™ 3000 RSLCnano System coupled to a Q-Exactive-HF mass spectrometer (Thermo Scientific). Each of the 72 plate wells was dissolved in 20ul solvent A and 10ul were injected. Samples were trapped on a C18 guard-desalting column (Acclaim PepMap 100, 75 μm x 2 cm, nanoViper, C18, 5 µm, 100Å), and separated on a 50 cm long C18 column (Easy spray PepMap RSLC, C18, 2 μm, 100Å, 75 μm x 50 cm). The nano capillary solvent A was 95% water, 5%DMSO, 0.1% formic acid; and solvent B was 5% water, 5% DMSO, 95% acetonitrile, 0.1% formic acid. At a constant flow of 0.25 μl min -1 , the curved gradient went from 6-8% B up to 40% B in each fraction in a dynamic range of gradient length, followed by a steep increase to 100% B in 5 min. FTMS master scans with 60,000 resolution (and mass range 300-1500 m/z) were followed by data-dependent MS/MS (30 000 resolution) on the top 5 ions using higher energy collision dissociation (HCD) at 30% normalized collision energy. The entire duty cycle lasted ~2.5 s.

Dynamic exclusion was used with 30 s duration. Precursors with unassigned charge state or charge state 1 were excluded with a used underfill ratio of 1%. The generated MS files were then analyzed by performing a database search against the predicted spectra of the Lutjanus_guttatus_annotated_sequences_no_threshhold database containing the previously sequenced and annotated *L. guttatus* sequences using Sequest-Percolator or Target Decoy PSM Validator under the software platform Proteome Discoverer 2.4 (Thermo Scientific). An FDR cutoff of 1% was used in conjunction with a precursor ion mass tolerance of 10 ppm and a product ion mass tolerance of 0.02 DA and 0.8 DA for HCD-FTMS and CID- ITMS respectively. Tryptic peptides with up to 2 missed cleavage sites were still considered. Fixed modifications TMT-18plex on lysines and peptide N- termini, and carbamidomethylation on cysteine residues, a variable modification was used for oxidation on methionine residues.

2.6 Bioinformatics

In order to fully unravel the effects of each stressor has independently and in conjunction with the other four contrasts were focused on: Δ Temperature (ambient pH), Δ Temperature (acidified pH), Δ pH (ambient Temperature) and Δ pH (elevated Temperature). The former term indicates the changed variable between the conditions, whereas the term in parenthesis stays constant, i.e. the Δ Temperature (ambient pH) contrasts fishes from the tanks with 28 °C and 400 ppm CO_2 and 30 °C and 400 ppm CO_2.

2.6.1 Metabolite analysis

Outliers were identified per PLS- DA with leave- one- out cross validation utilizing the "mix Omics" R package (Rohart *et al.*, 2017). The remaining statistical and functional analysis were done in the Metaboanalyst website (https://www.metaboanalyst.ca/). Before the analysis the data was normalized using base 10 log transformation and scaled using the auto scaling method (mean- centered then divided by standard deviation for each metabolite). Missing data imputation was performed by default by substituting the missing values with 1/5 of the lowest detected abundance value for that metabolite (response variable) Two- way ANOVA was performed False discovery rate as multiple testing correction (FDR cutoff < 0.05) testing for the effect of temperature, pH and their interaction followed by One- way ANOVA treating each condition as a separate variable, with Turkey's HSD correction. In order to visualize the metabolite data structure an unsupervised Principal Component Analysis (PCA) and supervised sparse partial least squares discriminant analysis (sPLS- DA) analyses were conducted.

Heatmaps and hierarchical clustering were performed using Euclidean distance as distance measure, Complete linkage as clustering method and feature auto scaling. Enrichment and pathway analysis was performed utilizing the Global test statistic (Goeman *et al.*, 2004). In both cases the supplied *Danio rerio* KEGG pathway database was used in this work.

2.6.2 Proteomic Analysis and Data integration

All proteomic and integrative bioinformatics analysis was performed in the R programming language utilizing Rstudio (R Core Team, 2023).Genewise likelihood tests and p- value adjustments were done using the edgeR and limma packages (Ritchie *et al.*, 2015; Chen *et al.*, 2024) respectively. P- values were adjusted by correcting for false discovery rate using a cutoff of 0.05.

Functional analysis was performed by Gene Set Enrichment Analysis (GSEA) using the MultiGSEA R package (Canzler and Hackermüller, 2020) using both proteomic and metabolomic datasets. Pathways utilized were the "drerio" KEGG databases. These were acquired using the and org.Dr.eg.db R packages (Carlson, 2023). Additionally, two methods of removing duplicate accessions were used; either taking the mean of all Ids with the same accession or only selecting the one with the highest abundance. GSEA analysis functions by first ranking all identified proteins by the direction of their fold change and their assigned p-value, then calculating an enrichment score based on this list for each pathway in the database. A significant score means that proteins belonging to this pathway are more likely to be on either extreme of the list than it would be by random chance. GSEA analysis was performed for each database separately and for each method of duplicate accession removal.

Additionally, the DRP list was searched for Proteins which are part of the most important metabolic pathways identified by the previous pathway analysis (section 2.6.1). Since very few pathways were significantly enriched, the selection was widened to the top three most significantly enriched pathways according to the metabolomic analysis. The pH (ambient temp) contrast was excluded, since here all adjusted p- values were very high.

3

RESULTS

3.1 Fish metabolomic profiles across stress profiles

Utilizing GC -TOF -Mass Spectrometry, a total of 45 primary metabolites were identified and quantified. After performing outlier analysis, one sample out of 24 was removed from the dataset (sample ID LT1F2 Table A1). Twenty (1.9%) missing values (below detection limit of the equipment) were detected in the dataset. After data normalization, transformation, and scaling PCA analysis was performed as an unsupervised method to explore the data structure. The PCA scores plot did not show any separation based on different metabolite abundance patterns between the experimental conditions. The first principal component explained 30.8% percent of total variance, the second 12.6% (cumulative 43.4%) (Figure A1). The Sparse PLS-DA (Figure 3.1.1) with 10 variables per component was able to show separation between samples of all experimental treatments except the condition at ambient temperature and ambient pCO_2, which had considerable overlap with all other conditions. The first two components only explained 24.2 % of the total observed variance (Component 1 =13.7%, Component 2 = 10.5%). The third component explained more variance than the first two combined, but all treatments overlapped almost completely when plotted against either component. The metabolites with the highest discriminating effect, indicated by larger absolute loading scores, in the first component were Glucose, followed by Urea. Both metabolites showed higher relative abundance at ambient temperatures and lower abundance during elevated temperatures. Asparagine and Succinic acid exhibited the largest loading scores in the second component. Tyrosine and trehalose being the most important metabolites in the third component.

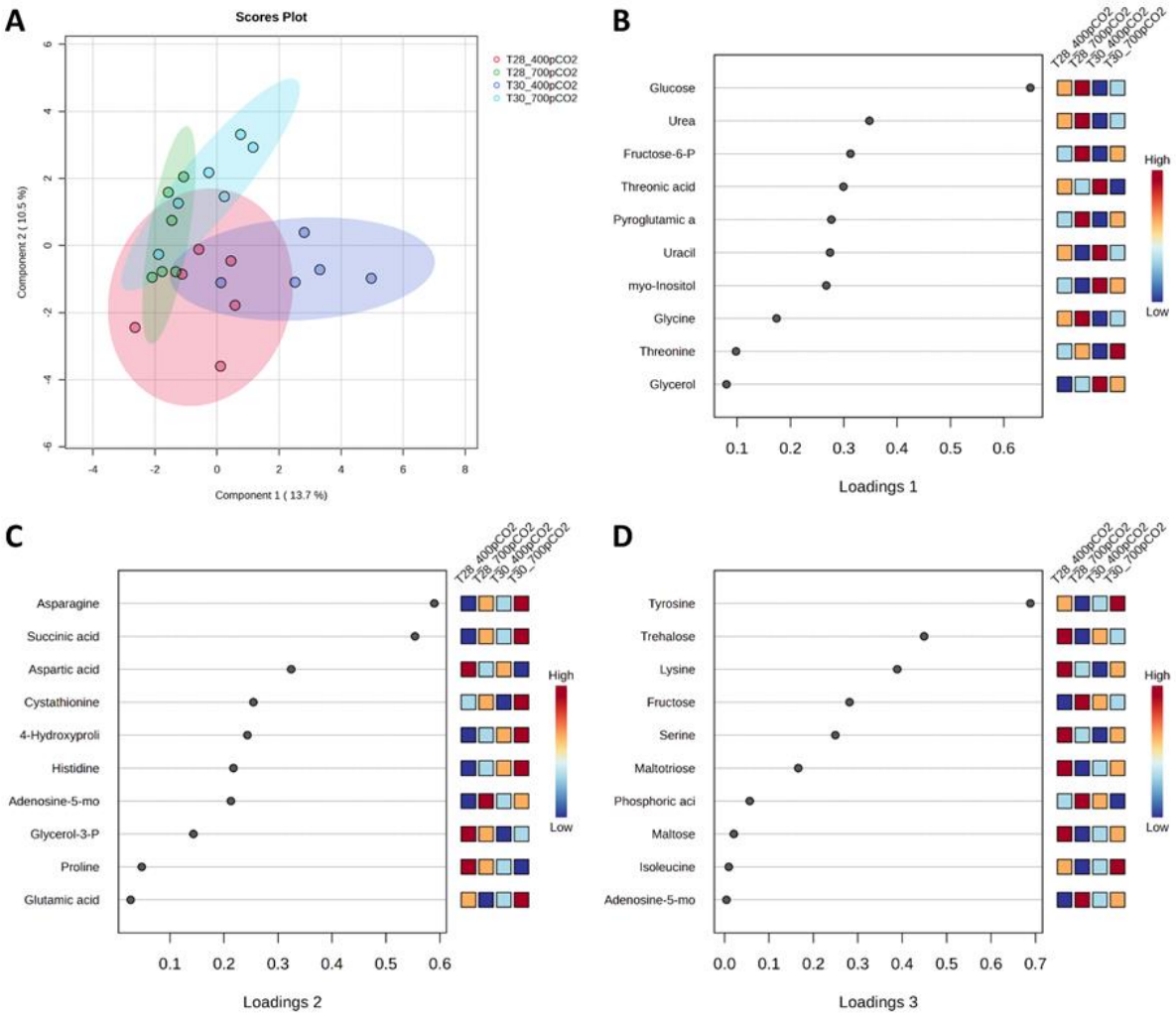

Figure 3.1.1. A) Scatterplot of the first 2 components of the sPLS- DA analysis. The different experimental conditions are color coded. **B) loading scores of the metabolites grouped into the first, second and third componen**t. Loading scores are indicated on the x- axis. The colored boxes on the right indicate the relative concentrations of the corresponding metabolite in each group under study. (note: the loadings plots depict the absolute loading scores)

Fold change analysis was performed (Figure A.2). The response was generally relatively muted, with all contrast having only a few metabolites that exhibit an absolute fold change of more than two. For this reason, the FC threshold was set down to 1.5. When comparing both temperatures at ambient pH, there were ten metabolites with lower abundance at higher temperatures and one with greater abundances. On the other hand, in the temperature contrast at acidified pH exhibited ten metabolites with greater abundances at elevated temperatures and only two with lower abundances showing a reversed trend compared to the other temperature contrast. When comparing both pH treatments, during ambient temperatures there were two more abundant and 9 less abundant metabolites, whereas there were ten more abundant and

four less abundant metabolites in the pH contrast during elevated temperatures. Two-way ANOVA analysis showed that Glucose levels were significantly changed by temperature increase (FDR = 0.034). This was the only metabolite that was significantly affected by any stressor. One- way ANOVA with Turkey HSD, treating all four experimental conditions as different factors in the same variable, showed Glucose to have a significantly different mean in the treatment at elevated temperatures and ambient pCO_2 levels (FDR 0.028056, Table 3.1, Figure 3.1.2).

Table 3.1. One way- ANOVA results, listing all metabolites found to have differing mean concentrations in at least one pair of experimental treatments. P values adjusted for false discovery rate are indicated in the FDR column. The post- hoc test column lists all pairwise comparisons in which the metabolite mean concentration is significantly different. All comparisons including the treatment at elevated temperatures and ambient pCO_2 were selected, and no pairings without this treatment.

Metabolite Name	f.value	p.value	FDR	Post-hoc tests
Glucose	9.0509	6.23E-04	0.028056	T30_700pCO_2-T30_400pCO_2
				T28_700pCO_2-T30_400pCO_2
				T28_400pCO_2-T30_400pCO_2

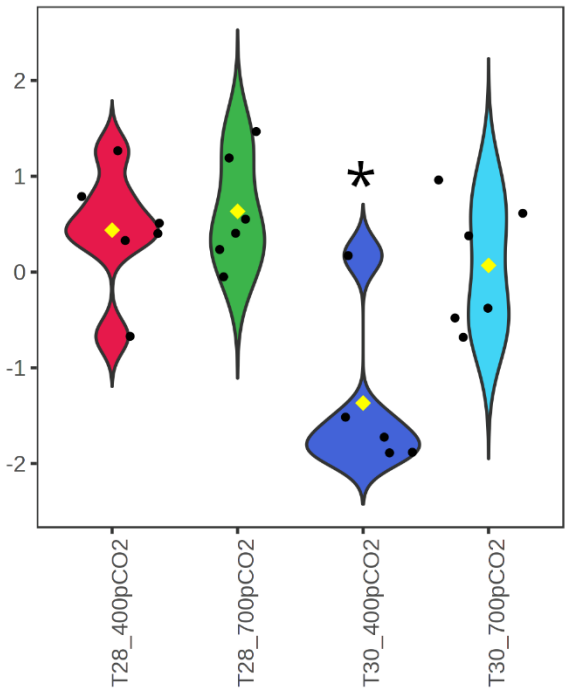

Figure 3.1.2. Violin plot showing alpha – D – Glucose concentrations in all conditions. The black dots represent individual sample concentrations, and the yellow dot the condition mean. Experimental conditions marked with a star have a significantly different mean (FDR > 0.05) from the other conditions

The sample clustering (horizontal dendrogram) did not group individuals from the same experimental conditions together (Figure 3.1.3). The dendrogram clearly clustered four samples, three of which belonged to the T28_ppm400 and one the T28_ppm700 condition. from the rest. These were further separated into two groups which showed a distinctive relative metabolite concentration pattern between each other and the remaining samples in the heatmap. These remaining samples were grouped into four groups, but each one contained sample from two to four different conditions. Visually there was for the most part no clear discernable pattern highlighting a high degree of variance between the samples. Metabolites were clustered into three main groups.

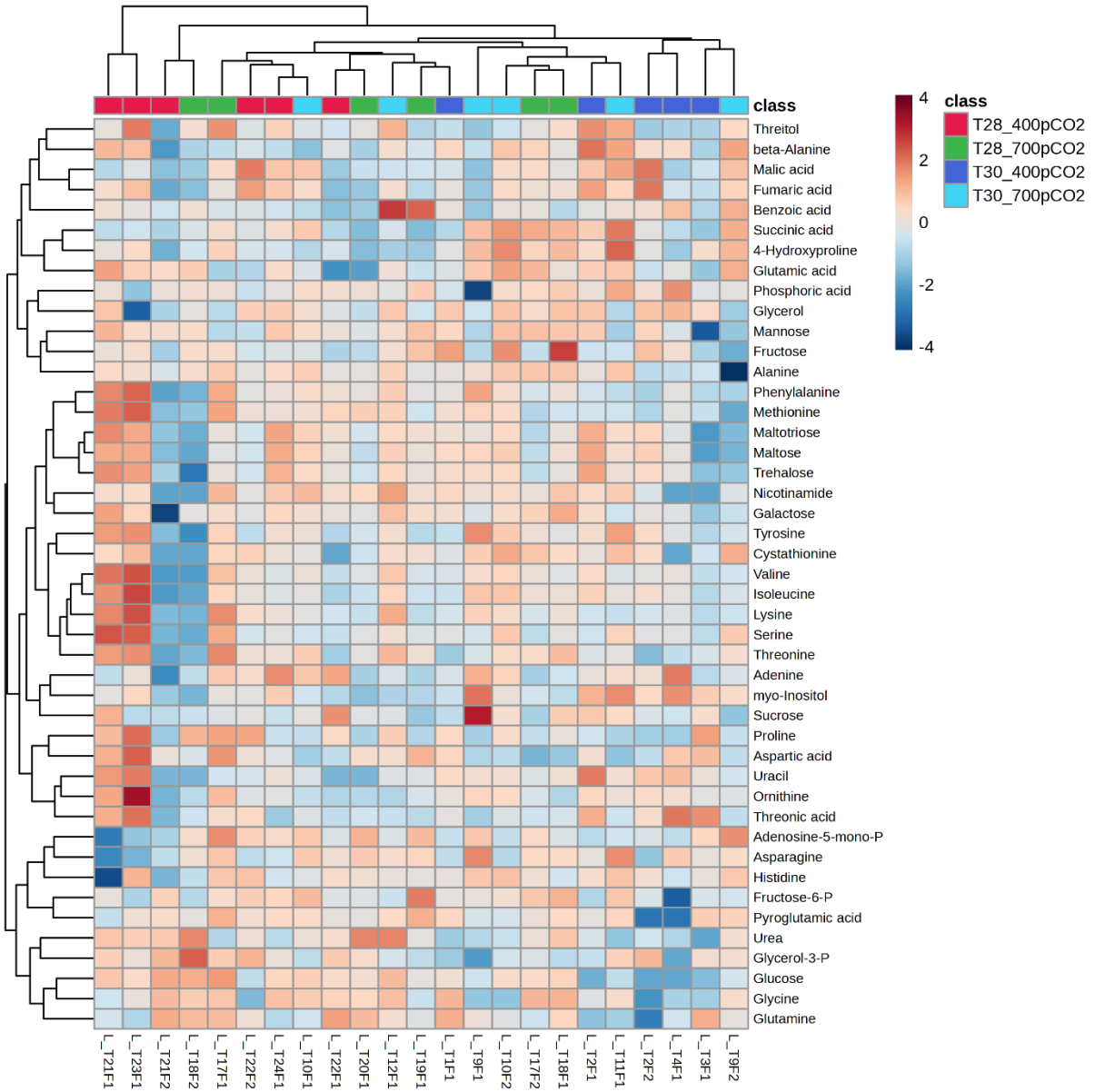

Figure 3.1.3. Heatmap showing the relative abundances of each metabolite. The dendrograms on the side indicate the hierarchical clustering results of the metabolites (columns, right dendrogram) and samples (rows, top dendrogram) respectively. The experimental conditions of the samples were additionally indicated by the colored squares on top of the figure. Cell color is determined by relative abundance, with negative values marked as blue and positive values as red.

Functional analysis using the global test did not show any significantly enriched pathways in any of the analyzed contrasts (Figure A.3)

Interestingly the results were also identical looking at the top 25 enriched metabolites when using the Homo sapiens or Danio rerio KEGG databases. In the different contrast of temperatures at ambient pCO_2 the most enriched Pathway was Glycolysis/ Gluconeogenesis, followed by fructose and mannose metabolism and amino sugar nucleotide metabolism,

although they were not significant with all three of these only having an FDR of 0.12 . In the temperature (acidified pH) and the pH (ambient temp) there were also no significant pathways found. In the pH (elevated pH) contrast the Glycine serine and threonine metabolism and the beta- Alanine metabolism pathways had low FDR values close to 0.05 (0.065 and 0.067 respectively) but were still non- significant.

3.2 Proteomic analysis and data integration

De novo transcriptome assembly resulted in a total of 211266 contigs and 145076 transcripts. Out of these 61992 open reading frames were predicted. A total of 61.442 predicted Lutjanus guttatus open reading frames were annotated by BLAST against a UNIPROT database.

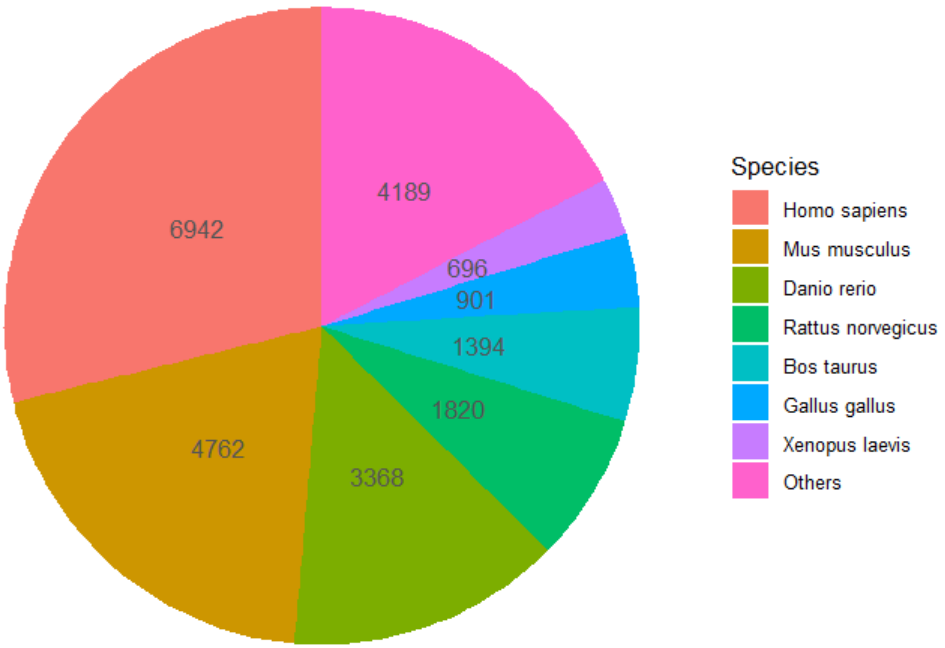

Figure 3.2.1. Pie chart showing the number of identified proteins attributed to each species. The top 7 most numerous species were names specifically. The "Others" category contains an additional 500 species, comprising 17.4% of total identified ORFs in the dataset.

In total 24.072 Proteins were identified using the Proteome Discoverer 2.4 (Thermo Scientific) platform by searching against a database of Lutjanus guttatus predicted open reading frames. In total the identified proteins had closest hits belonging to open reading frames

annotated from 507 different species. Of those, the most belonged to Homo sapiens (28.1%), Mus musculus (19.27%) and Danio rerio (13.6%)

Out of the 24.000 identified proteins only 37 were significantly differentially regulated in any of the chosen contrasts, with 27 unique Accessions found (Table A.2). A total of 35 Proteins were found to be upregulated between all contrasts, and 11 downregulated. Five Proteins were oppositely changed between different contrasts and as such were counted twice. All of these differences were found between the contrasts of temperature (at ambient pH), and pH (at high temperature).

The highest number of DRPs was found in the Temperature contrast at ambient pCO_2, with twenty-three (twelve unique Accessions), of which thirteen were upregulated , and ten (eight unique Accessions) were downregulated. The contrast between the pH treatments (at ambient temperature) exhibited no differentially regulated Proteins. The only two contrasts that showed an overlap in identified DRP accession codes are the two temperature contrasts at ambient and acidified pH, with four (three unique Accessions) upregulated DRPs.

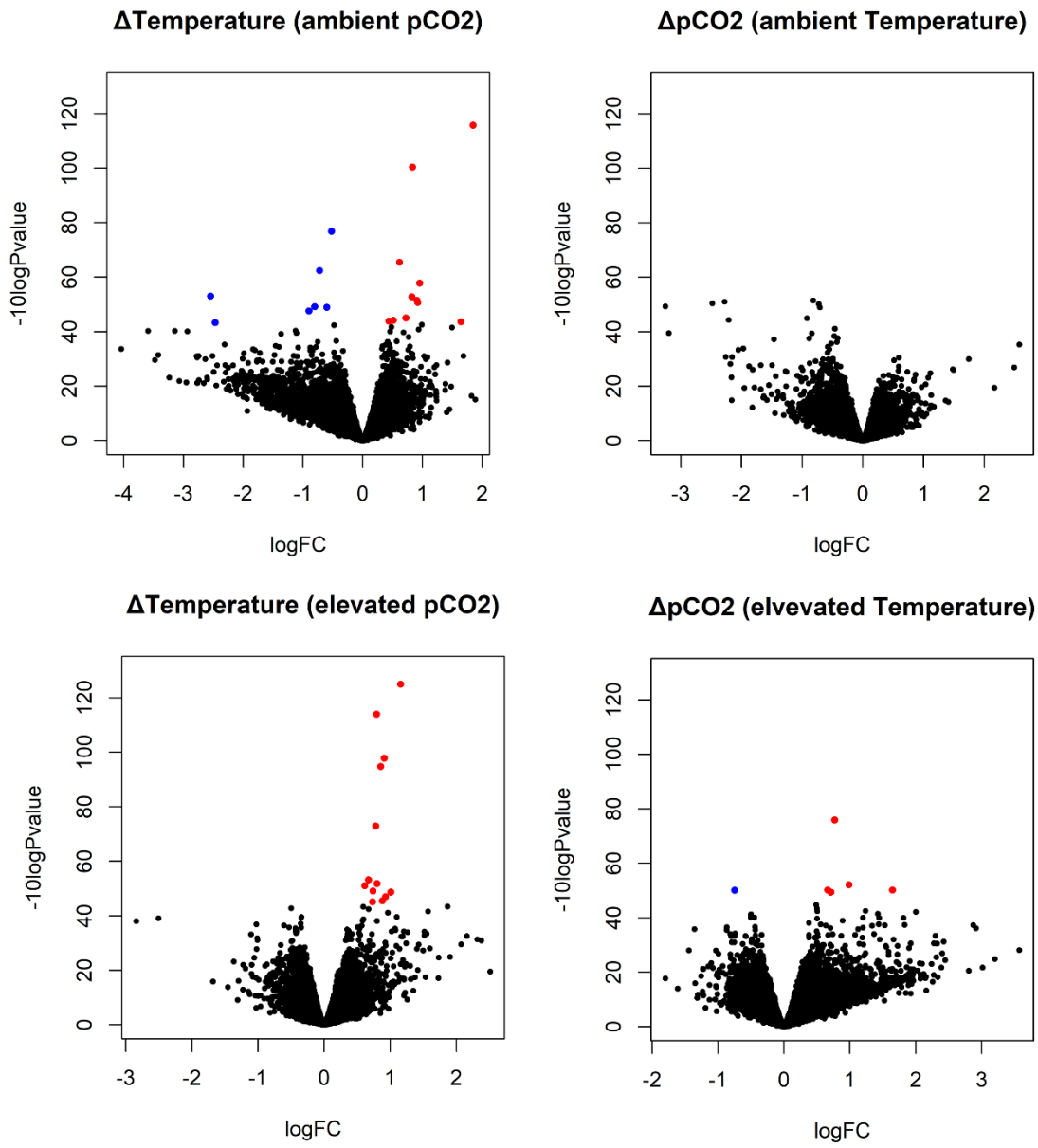

Figure 3.2.2. Volcano plots of the differentially regulated Proteins found in all conditions. Red dots indicate significantly upregulated and blue dots significantly downregulated genes (p. adjust < 0.05). The black dots are all non-differentially regulated proteins.

Table 3.2.1. Number of up and downregulated proteins for each contrast. The first number indicates the total number of identified variables, whereas the number in parenthesis is the number of unique accessions found in that set.

Contrast	Up	Down	total
all	35 (24)	11(9)	37
Temperature (at ambient pH)	23 (12)	10 (8)	23 (19)
Temperature (at acidified pH)	18 (10)	0 (0)	18 (10)
pH (at ambient Temperature)	0 (0)	0 (0)	0 (0)
pH (at high Temperature)	4 (2)	1 (1)	5 (3)

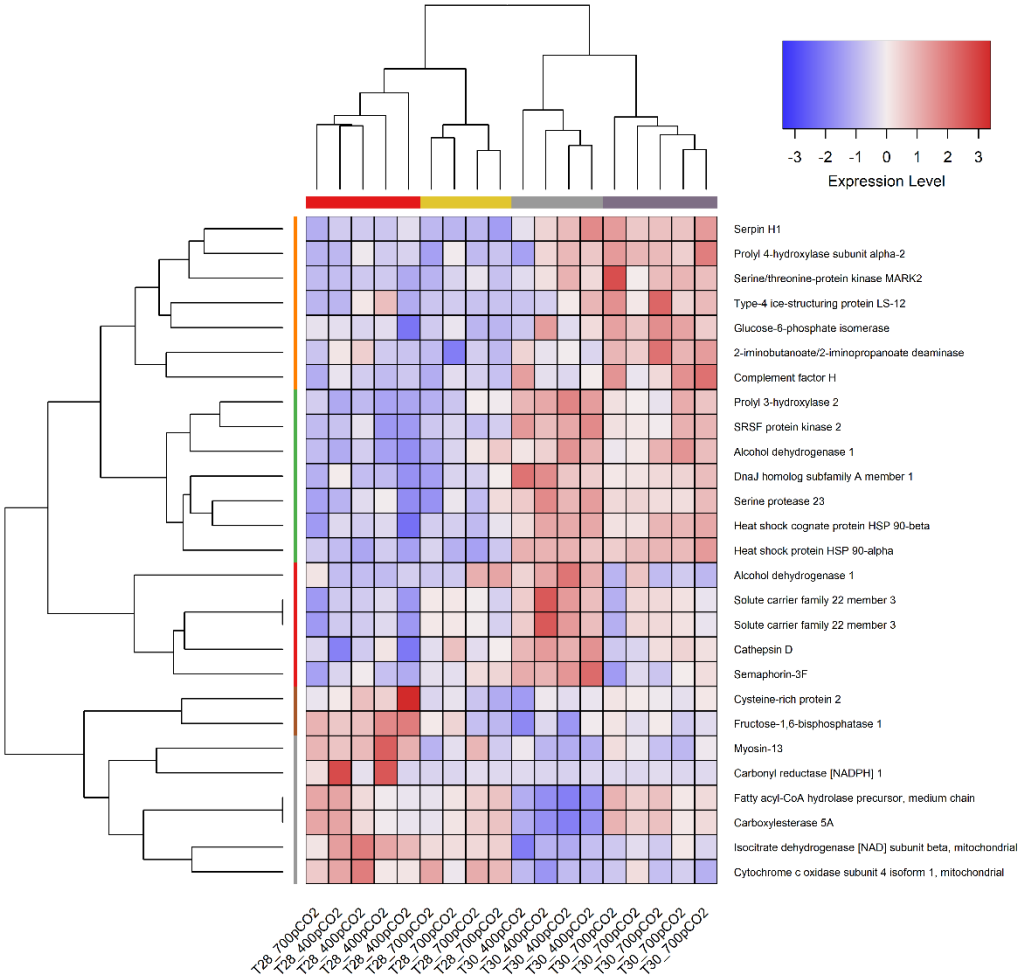

Figure 3.2.3. Heatmap illustrating expression levels of all unique differentially regulated proteins (p.adjust < 0.05). Columns represent samples, with only the experimental condition being indicated. Rows represent different Proteins. For duplicate accessions only one of them was included. Duplicate gene names for Solute carrier 22 family member 3 stem from these proteins being identified as stemming from different species, thus having different accessions. The clustering method as complete hierarchical clustering and the distance matrix was computed by Euclidean distance.

The heatmap showed some separation between the groups, with the clearest difference visible between the temperature treatments. The hierarchical clustering was able to group the individual samples based on their experimental condition quite well, with only one individual from the T28_700pCO_2 condition being grouped with all samples belonging to T28_ 400pCO_2. The dendrogram indicated that the proteins were roughly separated into 5 groups.

The abundance profiles of all 37 DRPs were used to perform Principal Component Analysis (PCA) and Partial Least Squares Principal Analysis (PLS-DA) for dimension reduction and pattern discovery. PCA analysis resulted in moderate separation between the groups, but still with considerable overlap. The percentages of explained variance were 45.9% for Principal component (PC) 1, 28.46% for PC 2 and 21.53% for PC 3 10 % (cumulative 49.99% for first 2 PCs). PLS – DA showed 23% explained variance in component 1 and 22% in component 2. Both the ambient and acidified pH conditions at 28°C showed almost no separation when using this method, whereas the two conditions at 30 °C were clearly separated between them and from

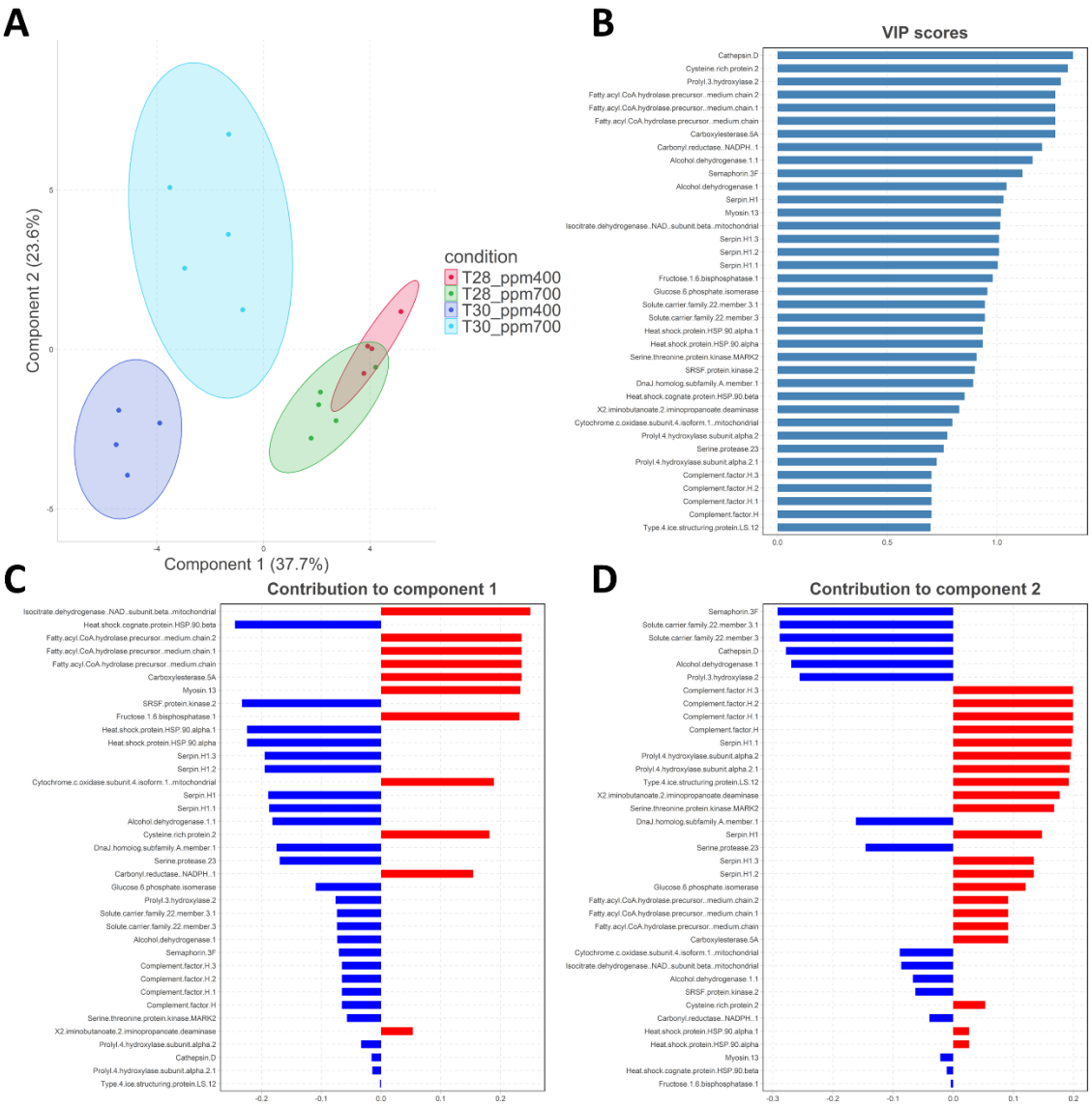

Figure 3.2.4. PLS- DA analysis investigating the regulation patterns of all DRPs. A) PLS- DA scores plot of the first two components. Samples belonging to different experimental conditions were color coded (red = control temperature and control ppm CO_2, green = control temperature and elevated ppm CO_2, dark blue = elevated temperature and control ppm CO_2 and light blue =elevated temperature and elevated ppm CO_2) VIP scores of all proteins used in the model. The VIP score summarizes the total contribution of the variable to the model. C) and D) loading scores of the first and second component, indicating the contribution of each to each individual component. The Proteins were ordered by absolute loading score, negative values are marked blue and positive ones red.

the other two. Using the VIP scores as well as the loadings for the first two components of the PLS- DA several Proteins for further investigation. The VIP score summarizes a variables importance in the PLSDA model, and higher loadings in a PC indicates a higher explained covariance. Six of these proteins' regulation patterns are illustrated in Figure 3.2.5.

Some Proteins, like the Heat shock protein HSP 90 show a clear upregulation in response to temperature increase,

Most of the proteins with high VIP values show a more differentiated response the stressors, which nevertheless follow a common trend observable for example in the Carboxylesterase 5A or Semaphorin 3F, where a clear change in abundance is visible in the T30 ppm400 condition before returning closer to control levels during the lower pH T30 ppm700 condition.

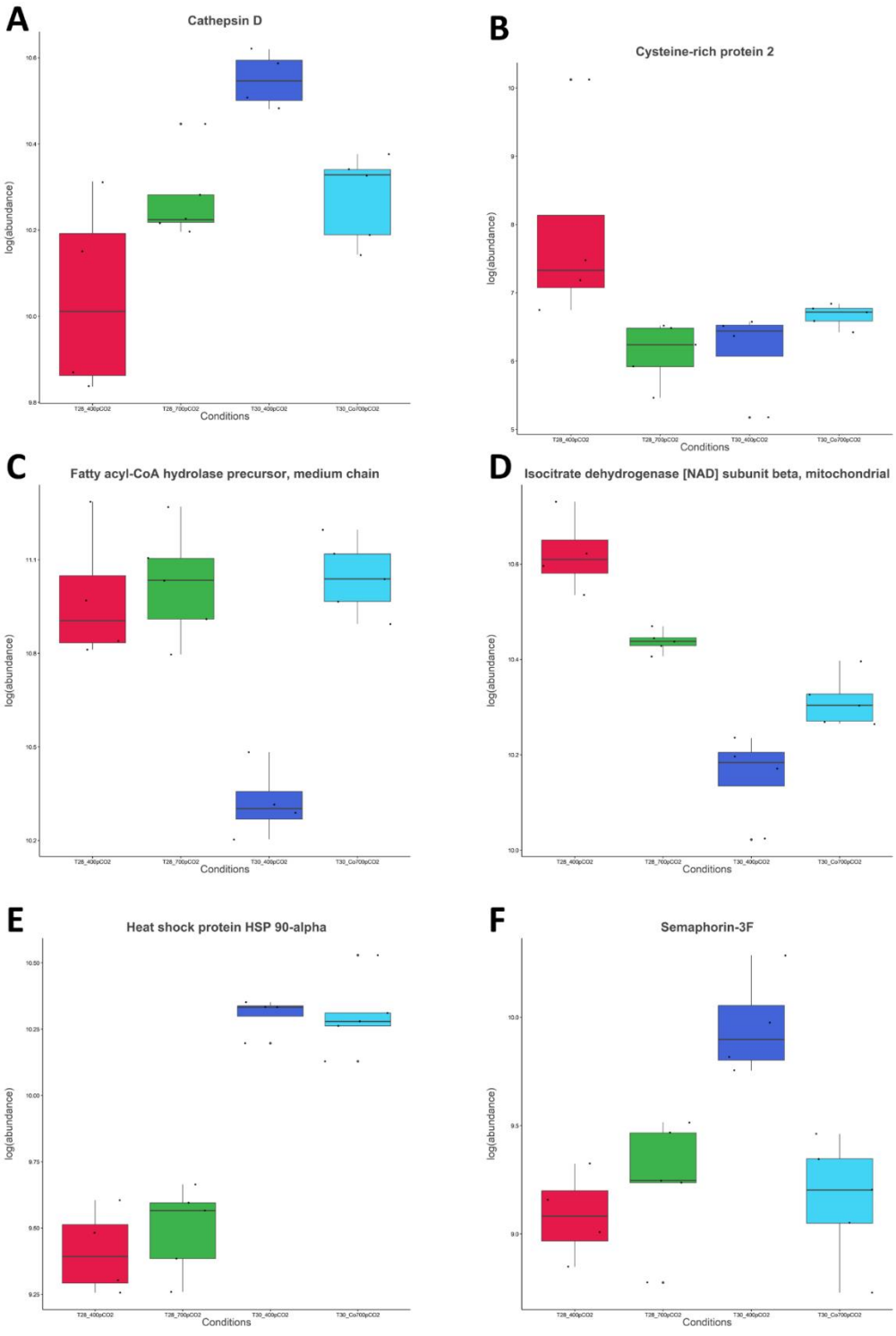

Figure 3.2.5. Boxplots showing the log abundances of proteins which PLS- DA analysis indicated had high correlation to the experimental conditions. A, B and C showed the highest PLS- DA VIP scores, D and E the highest component 1 loadings and E and F the highest component 2 loadings. The different treatments are indicated by color; red = ambient temperature and ambient pCO_2, green = ambient temperature and elevated pCO_2, dark blue = elevated temperature and ambient pCO_2, light blue = elevated temperature and elevated pCO_2.

Results from the multi-GSEA analysis conducted on *L. guttatus* proteomic datasets using *Danio rerio* KEGG pathway library are shown in Figure 3.2.6, specifically highlighting the top 10 most enriched pathways in each tested contrast. The only pathway found to be significantly enriched was the retinol pathway in the contrast of different pCO_2 treatments during elevated temperatures, where the pathway response was enriched (NES > 0). When comparing 400 ppm pCO_2 (control) with 700 ppm CO_2 (elevated) at each temperature, similar pathways were altered (9 of the top 10 most enriched) were the same, however, 7 of them (apoptosis, apelin signaling pathway, aminoacyl- tRNA biosynthesis, amino sugar and nucleotide sugar metabolism, adrenergic signaling in cardiomyocytes, adherens junctions and 2- Oxycarboxylic acid metabolism) showed contrary directions of responses depending on temperature. The contrasts comparing temperature treatments (28 vs 30 ºC) at ambient and acidified pH displayed only 2 pathways in common, both of which showed depletion (negative NES, red color), namely alanine, aspartate and glutamate metabolism as well as arginine biosynthesis. When comparing both temperatures specifically at the acidified pH, the 30 ºC treatment showed negative normalized enrichment scores (NES) for pathways such as carbon metabolism, and biosynthesis of biological molecules like sugars and amino acids (NES <0, blue color), whereas fatty acid elongation had positive scores (NES > 0, red color). These patterns were not discernible when comparing both control and elevated temperature at ambient pCO_2, which showed heterogenous responses at pathway level.

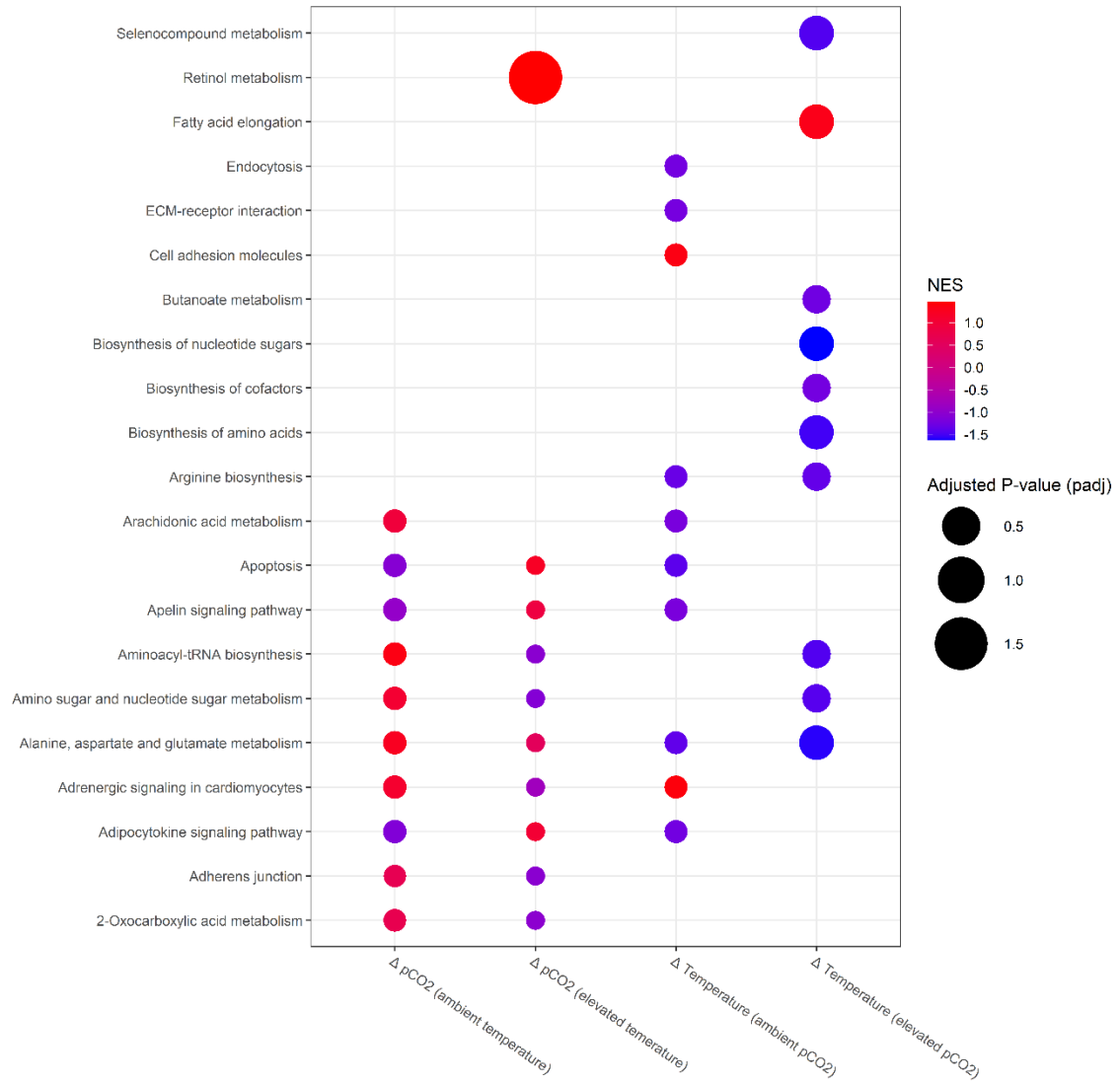

Figure 3.2.6. Dot plot showing the top 10 most enriched pathways for each of the four contrasts. The size of the dot represents the negative log10 of the FDR, the color the normalized enrichment score (NES). NES values greater than 0 (marked red) indicate that this pathway is upregulated, values smaller than 0 (marked blue) indicate that this pathway is depleted.

4

DISCUSSION

The discovery of the molecular mechanisms underpinning physiological responses of fish in ocean regions most vulnerable to climate change is crucial for the critical understanding of population viability and biodiversity loss or reorganization. In the present thesis, the commercial fish species *L. guttatus* was exposed to mid-century ocean warming and acidification conditions for 28 days, after which a multi-omics approach based on metabolomics coupled with proteomics was employed to study liver response mechanisms. Overall, our results showed that temperature had a greater effect in liver responses when compared to acidification, however pH chasnges seem to have a significant influence in the degree of thermal stress experienced by the fish. Our results also showed that the primary metabolome may be less affected by environmental stressors than the proteome. In both cases, only a small fraction of metabolites and proteins displayed significant concentration alterations (only 1 altered metabolite out of 45 identified ones, and 37 differentially regulated proteins out of 24,000 identified proteins). These results may be unsurprising since L. guttatus can tolerate temperatures up to 35 ºC (Larios-Soriano *et al.*, 2020), and aquaculture broodstocks in the ETP can be locally subjected to pCO_2 concentrations in seawater that exceed levels predicted for mid-century (Martínez-Cordero, Sanchez-Zazueta and Hernández, 2018). In the next sections, the main findings of this thesis will be discussed in detail.

4.1 Ocean acidification has only a marginal effect in fish metabolite and protein profiles

Results showed that at ambient temperatures, exposure to acidified water did not affect fish molecular profiles, neither metabolite nor protein levels. Both molecule types maintained

stable concentrations, similar to control pH conditions after 28 days of exposure. These results are corroborated by previous literature suggesting that fish are not affected by pH levels representing current or future ocean acidification conditions (Branch *et al.*, 2013). In fact, the present scientific consensus highlights the ability of fish to effectively regulate their blood pH (which is maintained within a narrow range) and the maintenance of basic performance indicators such as survival, swimming speed or aerobic capacity in acidified conditions (Cameron, 1978; Kinkead *et al.*, 1993; Perry *et al.*, 2015). For instance, Munday et al. (2011) found that exposure to a CO_2 increase of400 µatm had no effect on growth or survival in juvenile spiny damselfish (*Acanthochromis polyacanthus*). At the molecular level, another study performed by Shao et al. (2016), which exposed the orange-spotted grouper (*Epinephelus coioides)* to a similar pH decrease for 6 weeks showed that this treatment had no effect on the expression of small glycoproteins such as growth factor in the liver of the tested species. Nevertheless, some studies suggest that even if fish maintain homeostasis under acidification, they may need to perform more acid-base regulation under such conditions of environmental perturbation (Perry and Gilmour, 2006). However, the absence of concentration changes in metabolites or proteins in fish livers in our study points to a lack of evidence that the tested acidified condition (pH 7.8, equivalent to a pCO_2 of 700 ppm) induced a hypercapnia state in fish that would require any extra energy for osmoregulation. The lack of liver response - the central organ for metabolic regulation (Bruslé and Anadon, 2017), also suggests that no other energetically expensive biological processes or pathways were negatively affected, such as metabolic regulation, detoxification or the immune response (Fu *et al.*, 2021; Servili *et al.*, 2022). Albeit, since the main osmoregulatory organs in teleost fish are gills and kidney (Brauner *et al.*, 2019), multiomics analyses in these organs could provide further crucial information. Indeed, a more demanding osmoregulatory function during longer exposure times or in the presence of additional stressors, could potentially pose some energetic tradeoffs for these organisms (Huth and Place, 2016; Baag and Mandal, 2022).

In accordance, *L. guttatus* livers exposed to acidified conditions at an elevated temperature of 30°C increased glucose levels compared to those observed under elevated temperature with ambient pCO_2. However, glucose levels remained unchanged relative to those recorded under ambient temperature and ambient pCO_2 conditions. The increase in pCO_2 significantly affected the regulation of 3 proteins (Semaphorin-3F (downregulated), Carboxylesterase 5A (upregulated), and Fatty acyl-CoA hydrolase precursor (upregulated)) during elevated temperatures. Carboxylesterase 5A and Fatty acyl CoA hydrolase in are related to functions associated with breaking down lipids and fatty acids the cell for energy production (Ross *et al.*, 2010; Houten, Wanders and Ranea-Robles, 2020). Semaphorin 3F is a signaling protein that helps mediate, among others, angiogenesis and immune system regulation (Bougeret *et al.*,

1992; Carmeliet and Tessier-Lavigne, 2005; He, Crenshaw and Raper, 2019). This suggests that *L.guttatus* responds to acidification by increasing lipid metabolism. Retinol (Vitamin A) is an essential vitamin in fish that regulates lipid metabolism in the liver. These results suggest that acidification under elevated temperature led to increased lipid catabolism for energy production.

Despite the detected changes in protein regulation being small, the fact that the hierarchical clustering of protein data in the heatmap was able to cluster the majority of samples into the correct experimental pH groups still suggest that there is a specific pattern in protein abundances responding to the different pH conditions.

4.2 Temperature increase induces a cellular stress response

Increased Temperature during ambient pCO_2 lead to a significant decrease in glucose levels in the liver. Apart from that the metabolomic reaction was still limited and heterogenous, as no other metabolites were significantly changed or pathways significantly enriched.

On a protein regulation level, the fish showed a much more pronounced response to warming than to acidification. A total of twenty proteins were found to be differentially regulated in response to increasing temperatures during ambient pCO_2.

HSP 90 was upregulated in 30 °C water, independently of pCO_2 levels, indicating a molecular response to stressful thermal conditions (Hu *et al.*, 2022). HSPs have been used as a biomarker for thermal stress in various fish species such as crucian carp (*Carassius carassius*) or Pacific salmon (*Oncorhynchus spp.*) (An, Lei and Zheng, 2014; Akbarzadeh *et al.*, 2018).

In the temperature (ambient pH) contrast several proteins involved in key response mechanisms to heat stress were identified. The immune system plays an important role in the general stress response of fish (Elbahnaswy *et al.*, 2024). Serine protease 23 for example has anti- inflammatory properties (Qin *et al.*, 2022) and was upregulated in this contrast. Cysteine rich protein 2 may play a role in the antioxidant defense in zebrafish (Guru and Arockiaraj, 2023), but interestingly it was deregulated here.

Fructose-bisphosphatase 1 (FBP) and Alcohol dehydrogenase 1 (ADH) are part of the Glycolysis/ Gluconeogenesis pathway and were differentially regulated during increased temperatures under ambient pCO_2 conditions. Additionally, Isocitrate dehydrogenase (ICD), which is part of the citrate cycle, was downregulated in ambient pH conditions. FBP catalyzes the hydrolysis from Fructose 1,6- biphosphatase to Fructose 6- phosphate which is the rate-limiting step during gluconeogenesis (Timson, 2019). FBP and ICD expression was found to decrease during induced hypoglycemia in gilthead seabream (Rashidpour *et al.*, 2019).

The decreased in Glucose levels and the deregulation of Fructose -1,6- biphsosphatase a key rate limiting protein in Gluconeogenesis indicate that Glucose synthesis in the liver is negatively affected by temperature, even though this result was not reciprocated by either metabolomic or proteomic pathway analysis.

Previous studies have indicated that thermal stress increases gluconeogenesis in teleost fishes in order to mobilize energy for meeting higher demands for metabolic homeostasis (Rijnsdorp *et al.*, 2009). Temperature increase during ambient pH does not seem to follow this trend *in L. guttatus*. Fatty acyl-CoA hydrolase and Cathepsin D were differentially regulated in the temperature (ambient pH) contrast, which are participating in lipid metabolism (Houben et al., 2017) indicating a shift towards a utilization of lipids as an energy source.

Additionally, ICD was also deregulated, indicating a reduction in aerobic metabolism via the citrate cycle (Chao *et al.*, 1997). During heat stress fishes typically face low oxygen pressure because of increased anerobic metabolism increasing oxygen demand (He *et. al.* 2015).e

Moreover, the upregulation of alcohol dehydrogenase 1, which is involved in transforming pyruvate into ethanol, as an alternative to the normal reduction to lactate during anaerobic glycolysis (Dhillon *et al.*, 2018). This pathway has been shown to be activated in fishes to decrease lactate accumulation in the cells (Torres *et.al.* 2012) in conjunction with heat stress (Munday, Crawley and Nilsson, 2009). It was shown previously that several fish species respond to temperature cycling by increasing energy reserves when the stress levels were low enough (Spigarelli, Thommes and Prepejchal, 1982; Eldridge, Sweeney and Mac Law, 2015). Lutjanus guttatus has been shown to be able to tolerate temperatures of up to 32 °C, and sea surface temperature can exceed 28 °C in the TEP, especially during El Niño events (Amador *et al.*, 2016). It might therefore be possible that *L. guttatus* responds to heat stress by increasing energy reserves, potentially through lipogenesis

Glucose levels were only significantly affected during elevated temperatures in conjunction with ambient pC_2 levels and returned to control levels when fish were in warm acidified water, even though acidification was previously reported to increase energetic costs when occurring in conjunction with heat stress.

4.3 pCO_2 seems to affect the nature of the thermal response

There was a considerable difference in the physiologic response to thermal stress depending on the pH level. Out of the 18 and 9 DRPs in the temperature contrasts at ambient and acidified pH respectively only 3 were differentially regulated in both. These were for example the heat shock proteins HSP 90, which showed a clear elevation in increased temperature independently of pH levels. Most other DRPs showed a distinct regulatory reaction to a hike in temperature, depending on CO_2 levels. During ambient pCO_2 there was some evidence of decreased gluconeogenesis, but this was not the case during acidified pCO_2. The only differentially expressed protein in the latter contrast involved in Gluconeogenesis/Glycolysis was Glucose-6-phosphate isomerase (G6P), which was upregulated. G6P catalyzes the second step of glycolysis and is also a rate limiting enzyme in the pentose phosphate pathway (Tian *et al.*, 1998; Mojzikova *et al.*, 2018). No metabolites, including glucose were significantly altered. Since no evidence was found of a differential regulation of glycolysis, G6P overregulation might be evidence for a shift towards the pentose phosphate pathway, which is a major source of NHDAP, playing a critical role in the antioxidant defense (Patra and Hay, 2014). Additionally, all metabolites but one which had a fold change of less than 1.5 in the temperature (ambient pH) contrast were less abundant, whereas all but 2 were more abundant in the temperature (acidified pH) contrast. This, alongside the fact that no Proteins involved in lipid metabolism were differentially regulated during acidified pH levels, indicates a clear shift in metabolic activity depending on prevalent CO_2 levels.

Lutjanus guttatus showed DRPs related to thermal stress response by affecting the immune system (complement factor H (Sun *et al.*, 2010)) in temperature (acidified pH).

Serpin H1 and Prolyl 4-hydroxylase were both upregulated in temperature (acidified pH), which are both involved in collagen synthesis (Ishida and Nagata, 2011; Ishikawa, Holden and Bächinger, 2017). Collagen is involved in cell regeneration and wound healing and is a major component of connective tissue (Chung *et al.*, 2008). Serpin H1 was also identified as a potential biomarker for heat stress in salmonids (Akbarzadeh *et al.*, 2018)). The fact that these proteins were only upregulated in high temperature and acidified conditions indicates that both stressors combined can lead to higher tissue damage than just heat stress, leading to a larger investment into cellular repair mechanisms.

4.4 Challenges working with non-model organisms

One of the main challenges when working with non- model organisms is the lack of biological interpretability when faced with biological data. Since *L. guttatus* is not a model species and has not been well investigated by researchers in the past, there is not a lot of data available about this species. This has led to problems on multiple steps during this analysis.

The issue of needing a spectral library for the peptide and protein identification was solved by performing de novo transcriptome assembly with RNA extracted from a small amount of liver samples. Still, this method relies on the assumption that homologous genes will have the same functions and properties, which is not always the case, especially with more distantly related organisms.

For example, conserved sequences may have diverged in function due to different selective pressures in distinct environments (Nam *et al.*, 2021).

In the 24.000 identified proteins there were only 11.000 unique UNIPROT accessions. This can happen when multiple peptides are assigned to the same protein during spectral matching, leading to redundant protein identifications. This can happen for multiple reasons, for example with homologous or differently spliced peptides. Another reason for these false identifications is incomplete databases used for the peptide search. If the analyzed peptide belongs to a protein that is not included in the database, it will either not be identified at all or be assigned to the wrong peptide. There are statistical methodologies to minimize this problem, like controlling for false discovery rate and target- decoy searching during the database search, but the highly complex and heterogeneous nature of mass spectronomy data means that even peptides with high confidence scores and low FDRs may still be mismatched. Because of this one of the main challenges in proteomics analyses is the low analytical reproducibility.

5
CONCLUSIONS

This study aimed to provide new insight into the physiological reaction of *Lutjanus guttatus* when faced with ocean warming and acidification, and the molecular mechanisms involved therein, via a multi-omics approach using metabolomics and proteomics. To this end, the main findings of this thesis are as follows

(1) Lutjanus guttatus is well equipped to deal with future short term ocean acidification. There was no significant proteomic or metabolomic reaction to lower pH at current temperatures, although there was some evidence of altered amino acid metabolism during acidification in conjunction with higher temperatures.

(2) *Lutjanus guttatus* showed a clear molecular stress response to ocean warming, with areas such as metabolism as well as the anti-oxidative stress response being altered.

(3) The metabolic reaction to increased temperatures, but only during ambient pCO_2 levels, was unexpected, with Glucose being less abundant and Fructose -1,6- phosphatase. *L. guttatus* seems to exhibit the opposite reaction of what was reported in other species.

(4) The CO_2 concentration seemed to affect the nature of the temperature response in the rose spotted snapper. Glucose levels returned to control levels when the fish were faced with the most extreme conditions, and, apart from the Heat shock genes, there was no overlap between the DRPs of both temperature contrasts.

(5) Increased temperature and CO_2 combined leads to more investment into cellular repair in spotted rose snapper livers.

BIBLIOGRAPHY

Akbarzadeh, A. *et al.* (2018) 'Developing specific molecular biomarkers for thermal stress in salmonids', *BMC Genomics*, 19(1). Available at: https://doi.org/10.1186/s12864-018-5108-9.

Alexander, M.A. (1992) 'Midlatitude Atmosphere–Ocean Interaction during El Niño. Part I: The North Pacific Ocean', *Journal of Climate*, 5, pp. 944–958. Available at: https://doi.org/10.1175/1520-0442(1992)005<0944:MAIDEN>2.0.CO;2.

Alfaro, A.C. and Young, T. (2018) 'Showcasing metabolomic applications in aquaculture: a review', *Reviews in Aquaculture*, 10(1), pp. 135–152. Available at: https://doi.org/10.1111/raq.12152.

Amador, J.A. *et al.* (2016) 'The easternmost tropical Pacific. Part I: A climate review', *Revista de Biología Tropical*, 64(64), pp. 1–22. Available at: https://doi.org/10.15517/RBT.V64I1.23407.

An, L., Lei, K. and Zheng, B. (2014) 'Use of heat shock protein mRNA expressions as biomarkers in wild crucian carp for monitoring water quality', *Environmental Toxicology and Pharmacology*, 37(1), pp. 248–255. Available at: https://doi.org/10.1016/j.etap.2013.11.019.

Andini, N. *et al.* (2024) 'Global Warming Threats to Coral Bleaching Events and Their Impacts on Coastal Ecosystem Sustainability', *BIO Web of Conferences*, 134, p. 02009. Available at: https://doi.org/10.1051/bioconf/202413402009.

Ando, K. and Hasegawa, T. (2009) 'Annual Zonal Displacement of Pacific Warm Pool in Association with El Nino Onset', *SOLA*, 5, pp. 149–152. Available at: https://doi.org/10.2151/sola.2009-038.

Andrade, H. (2003) 'Age determination in the snapper Lutjanus guttatus (Pisces, Lutjanidae) and investigation of fishery management strategies in the Pacific Coast of Guatemala'. Available at: https://www.researchgate.net/publication/33416900.

Anthony, A. *et al.* (2009) 'Coastal lagoons and climate change: Ecological and social ramifications in U.S. Atlantic and Gulf coast ecosystems', *Ecology and Society*, 14(1). Available at: https://doi.org/10.5751/ES-02719-140108.

Araya, H. *et al.* (2007) 'Reporte del Manejo de los Recursos Pesqueros en el Golfo de Nicoya', *INCOPESCA— UNA—JICA Proyecto Manejo sostenible de la pesqueria para el Golfo de Nicoya, Costa Rica. Presentacion de conclusiones y reco- mandaciones. Comité de Evaluacion de Recursos Pesqueros. Puntarenas, Costa Rica,* pp. 5–67.

Assan, D. *et al.* (2020) 'Effects of Climate Change on Marine Organisms', *American Journal of Climate Change*, 09(03), pp. 204–216. Available at: https://doi.org/10.4236/ajcc.2020.93013.

Baag, S. and Mandal, S. (2022) 'Combined effects of ocean warming and acidification on marine fish and shellfish: A molecule to ecosystem perspective', *Science of The Total Environment*, 802, p. 149807. Available at: https://doi.org/10.1016/j.scitotenv.2021.149807.

Bass, A. *et al.* (2021) 'Another Decade of Marine Climate Change Experiments: Trends, Progress and Knowledge Gaps', *Frontiers in Marine Science*. Frontiers Media S.A. Available at: https://doi.org/10.3389/fmars.2021.714462.

Baums, I.B. *et al.* (2012) 'No gene flow across the Eastern Pacific Barrier in the reef-building coral Porites lobata', *Molecular Ecology*, 21(22), pp. 5418–5433. Available at: https://doi.org/10.1111/j.1365-294X.2012.05733.x.

Benedetti, M. *et al.* (2022) 'Emerging environmental stressors and oxidative pathways in marine organisms: Current knowledge on regulation mechanisms and functional effects', *Biocell*. Tech Science Press, pp. 37–49. Available at: https://doi.org/10.32604/biocell.2022.017507.

Bernal, M.A. *et al.* (2022) 'Plasticity to ocean warming is influenced by transgenerational, reproductive, and developmental exposure in a coral reef fish', *Evolutionary Applications*, 15(2), pp. 249–261. Available at: https://doi.org/10.1111/eva.13337.

Bessudo, S. *et al.* (2011) 'Residency of the scalloped hammerhead shark (Sphyrna lewini) at Malpelo Island and evidence of migration to other islands in the Eastern Tropical Pacific', *Environmental Biology of Fishes*, 91(2), pp. 165–176. Available at: https://doi.org/10.1007/s10641-011-9769-3.

Bougeret, C. *et al.* (1992) 'Increased surface expression of a newly identified 150-kDa dimer early after human T lymphocyte activation.', *The Journal of Immunology*, 148(2), pp. 318–323. Available at: https://doi.org/10.4049/jimmunol.148.2.318.

Boyd, P.W. *et al.* (2016) 'Biological responses to environmental heterogeneity under future ocean conditions', *Global change biology*. Blackwell Publishing Ltd, pp. 2633–2650. Available at: https://doi.org/10.1111/gcb.13287.

Boyd, P.W. *et al.* (2018) 'Experimental strategies to assess the biological ramifications of multiple drivers of global ocean change—A review', *Global Change Biology*. Blackwell Publishing Ltd, pp. 2239–2261. Available at: https://doi.org/10.1111/gcb.14102.

Branca, R.M.M. *et al.* (2014) 'HiRIEF LC-MS enables deep proteome coverage and unbiased proteogenomics', *Nature Methods*, 11(1), pp. 59–62. Available at: https://doi.org/10.1038/nmeth.2732.

Branch, T.A. *et al.* (2013) 'Impacts of ocean acidification on marine seafood', *Trends in Ecology & Evolution*, 28(3), pp. 178–186. Available at: https://doi.org/10.1016/j.tree.2012.10.001.

Brauner, C.J. *et al.* (2019) 'Acid-base physiology and CO2 homeostasis: Regulation and compensation in response to elevated environmental CO2', in *Fish Physiology*. Elsevier Inc., pp. 69–132. Available at: https://doi.org/10.1016/bs.fp.2019.08.003.

Briggs, J.C. (1961) *The East Pacific Barrier and the Distribution of Marine Shore Fishes*. Available at: https://doi.org/https://doi.org/10.2307/2406322.

Bruslé, J. and Gonzàlez I Anadon, G. (2017) 'The Structure and Function of Fish Liver', in *Fish Morphology*. Routledge, pp. 77–93. Available at: https://doi.org/10.1201/9780203755990-6.

Cameron, J.N. (1978) 'Regulation of blood pH in teleost fish', *Respiration Physiology*, 33(1), pp. 129–144. Available at: https://doi.org/10.1016/0034-5687(78)90092-0.

Canzler, S. and Hackermüller, J. (2020) 'multiGSEA: a GSEA-based pathway enrichment analysis for multi-omics data', *BMC Bioinformatics*, 21(1), p. 561. Available at: https://doi.org/10.1186/s12859-020-03910-x.

Carlson, M. (2023) 'org.Dr.eg.db: Genome wide annotation for Zebrafish. R package version 3.8.2. '. Available at: https://doi.org/10.18129/B9.bioc.org.Dr.eg.db.

Carmeliet, P. and Tessier-Lavigne, M. (2005) 'Common mechanisms of nerve and blood vessel wiring', *Nature*, 436(7048), pp. 193–200. Available at: https://doi.org/10.1038/nature03875.

Castillo-Vargasmachuca, S.G. *et al.* (2018) 'The spotted rose snapper (Lutjanus guttatus Steindachner 1869) farmed in marine cages: review of growth models', *Reviews in Aquaculture*. Wiley-Blackwell, pp. 376–384. Available at: https://doi.org/10.1111/raq.12166.

Chao, G. *et al.* (1997) 'Aerobic Regulation of Isocitrate Dehydrogenase Gene (icd) Expression in Escherichia coli by the arcA and fnr Gene Products', *JOURNAL OF BACTERIOLOGY*, 179(13), pp. 4299–4304. Available at: https://journals.asm.org/journal/jb.

Chen, Y. *et al.* (2023) 'Integrated biochemical, transcriptomic and metabolomic analyses provide insight into heat stress response in Yangtze sturgeon (Acipenser dabryanus)', *Ecotoxicology and Environmental Safety*, 249, p. 114366. Available at: https://doi.org/10.1016/j.ecoenv.2022.114366.

Chen, Y. *et al.* (2024) 'edgeR 4.0: powerful differential analysis of sequencing data with expanded functionality and improved support for small counts and larger datasets'. Available at: https://doi.org/10.1101/2024.01.21.576131.

Cheng, C.H. *et al.* (2018) 'Effects of high temperature on biochemical parameters, oxidative stress, DNA damage and apoptosis of pufferfish (Takifugu obscurus)', *Ecotoxicology and*

Environmental Safety, 150, pp. 190–198. Available at: https://doi.org/10.1016/J.ECOENV.2017.12.045.

Chung, H.J. *et al.* (2008) 'Collagen fibril formation: A new target to limit fibrosis', *Journal of Biological Chemistry*, 283(38), pp. 25879–25886. Available at: https://doi.org/10.1074/jbc.M804272200.

Claireaux, G. and Lagardère, J.P. (1999) 'Influence of temperature, oxygen and salinity on the metabolism of the European sea bass', *Journal of Sea Research*, 42(2), pp. 157–168. Available at: https://doi.org/10.1016/S1385-1101(99)00019-2.

Clarke, T.M. *et al.* (2021) 'Climate change impacts on living marine resources in the Eastern Tropical Pacific', *Diversity and Distributions*, 27(1), pp. 65–81. Available at: https://doi.org/10.1111/ddi.13181.

Coll-Lladó, C. *et al.* (2021) 'Pilot study to investigate the effect of long-term exposure to high pCO2 on adult cod (Gadus morhua) otolith morphology and calcium carbonate deposition', *Fish Physiology and Biochemistry*, 47(6), pp. 1879–1891. Available at: https://doi.org/10.1007/s10695-021-01016-6.

Cortés, J. *et al.* (2017) 'Marine Biodiversity of Eastern Tropical Pacific Coral Reefs', in. Springer, pp. 203–250. Available at: https://doi.org/10.1007/978-94-017-7499-4_7.

Couto, A. *et al.* (2008) 'Effect of water temperature and dietary starch on growth and metabolic utilization of diets in gilthead sea bream (Sparus aurata) juveniles', *Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology*, 151(1), pp. 45–50. Available at: https://doi.org/10.1016/j.cbpa.2008.05.013.

Couturier, C.S. *et al.* (2013) 'Species-specific effects of near-future CO2 on the respiratory performance of two tropical prey fish and their predator', *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, 166(3), pp. 482–489. Available at: https://doi.org/10.1016/j.cbpa.2013.07.025.

Creighton, C. *et al.* (2016) 'Adapting Management of Marine Environments to a Changing Climate: A Checklist to Guide Reform and Assess Progress', *Ecosystems*, 19(2), pp. 187–219. Available at: https://doi.org/10.1007/s10021-015-9925-2.

Deck, C.A. *et al.* (2017) 'Assessing the Functional Role of Leptin in Energy Homeostasis and the Stress Response in Vertebrates', *Frontiers in Endocrinology*, 8. Available at: https://doi.org/10.3389/fendo.2017.00063.

Dewulf, A. (2013) 'Contrasting frames in policy debates on climate change adaptation', *Wiley Interdisciplinary Reviews: Climate Change*, 4(4), pp. 321–330. Available at: https://doi.org/10.1002/wcc.227.

Dhillon, R.S. *et al.* (2018) 'Ethanol metabolism varies with hypoxia tolerance in ten cyprinid species', *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology*, 188(2), pp. 283–293. Available at: https://doi.org/10.1007/s00360-017-1131-4.

Dong, Y. wei *et al.* (2017) 'Untangling the roles of microclimate, behaviour and physiological polymorphism in governing vulnerability of intertidal snails to heat stress', *Proceedings of the Royal Society B: Biological Sciences*, 284(1854). Available at: https://doi.org/10.1098/rspb.2016.2367.

Elbahnaswy, S. *et al.* (2024) 'Immune protective, stress indicators, antioxidant, histopathological status, and heat shock protein gene expression impacts of dietary Bacillus spp. against heat shock in Nile tilapia, Oreochromis niloticus', *BMC Veterinary Research*, 20(1), p. 469. Available at: https://doi.org/10.1186/s12917-024-04303-5.

Eldridge, W.H., Sweeney, B.W. and Mac Law, J. (2015) 'Fish growth, physiological stress, and tissue condition in response to rate of temperature change during cool or warm diel thermal cycles', *Canadian Journal of Fisheries and Aquatic Sciences*, 72(10), pp. 1527–1537. Available at: https://doi.org/10.1139/cjfas-2014-0350.

Enes, P. *et al.* (2009) 'Nutritional regulation of hepatic glucose metabolism in fish', *Fish Physiol Biochem*, (35), pp. 519–539. Available at: https://doi.org/10.1007/s10695-008-9259-5.

Ern, R., Andreassen, A.H. and Jutfelt, F. (2023) 'Physiological mechanisms of acute upper thermal tolerance in fish', *Physiology*. American Physiological Society, pp. 141–158. Available at: https://doi.org/10.1152/physiol.00027.2022.

Escalante-Rojas, M. *et al.* (2018) 'De novo transcriptome assembly for the rose spotted snapper Lutjanus guttatus and expression analysis of growth/atrophy-related genes', *Aquaculture Research*, 49(4), pp. 1709–1722. Available at: https://doi.org/10.1111/are.13628.

Fiedler, P.C. (2002) 'Environmental change in the eastern tropical Pacific Ocean: review of ENSO and decadal variability', *Marine Ecology Progress Series*, 244, pp. 265–283.

Fiedler, P.C. *et al.* (2013) 'Pycnocline variations in the eastern tropical and north pacific, 1958-2008', *Journal of Climate*, 26(2), pp. 583–599. Available at: https://doi.org/10.1175/JCLI-D-11-00728.1.

Fu, Z. *et al.* (2021) 'Tissue comparison of transcriptional response to acute acidification stress of barramundi Lates calcarifer in coastal and estuarine areas', *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, 38, p. 100830. Available at: https://doi.org/10.1016/J.CBD.2021.100830.

Ganachaud, A. *et al.* (2013) 'Projected changes in the tropical Pacific Ocean of importance to tuna fisheries', *Climatic Change*, 119(1), pp. 163–179. Available at: https://doi.org/10.1007/s10584-012-0631-1.

Gattuso, J. *et al.* (1998) 'Effect of calcium carbonate saturation of seawater on coral calcification', *Global and Planetary Change*, 18(1–2), pp. 37–46.
Gattuso, J. *et al.* (2015) 'Contrasting futures for ocean and society from different anthropogenic CO2 emissions scenarios', *Science*. American Association for the Advancement of Science. Available at: https://doi.org/10.1126/science.aac4722.
Goeman, J.J. *et al.* (2004) 'A global test for groups of genes: Testing association with a clinical outcome', *Bioinformatics*, 20(1), pp. 93–99. Available at: https://doi.org/10.1093/bioinformatics/btg382.
Gómez-Reyes, R. *et al.* (2023) 'Individual Pattern Response to CO2-Induced Acidification Stress in Haliotis rufescens Suggests Stage-Specific Acclimatization during Its Early Life History', *Sustainability*, 15(18), p. 14010. Available at: https://doi.org/10.3390/su151814010.
Grabherr, M.G. *et al.* (2011) 'Full-length transcriptome assembly from RNA-Seq data without a reference genome', *Nature Biotechnology*, 29(7), pp. 644–652. Available at: https://doi.org/10.1038/nbt.1883.
Graham, M.S., Turner, D. and Wood, C.M. (1990) 'Control of ventilation in the hypercapnic skate Raja ocellata" I. Blood and extradural fluid', *Respiration Physiology*, 80(80), pp. 259–277.
Grosell, M. (2019) 'CO2 and calcification processes in fish', *Fish Physiology*, 37(37), pp. 133–159. Available at: https://doi.org/10.1016/bs.fp.2019.07.002.
Guru, A. and Arockiaraj, J. (2023) 'Exposure to environmental pollutant bisphenol A causes oxidative damage and lipid accumulation in Zebrafish larvae: Protective role of WL15 peptide derived from cysteine and glycine-rich protein 2', *Journal of Biochemical and Molecular Toxicology*, 37(1). Available at: https://doi.org/10.1002/jbt.23223.
Halpern, B.S. *et al.* (2008) 'A Global Map of Human Impact on Marine Ecosystems', *Science*, 319(5865), pp. 946–948. Available at: https://doi.org/10.1126/science.1151084.
Hannan, K.D. and Rummer, J.L. (2018) 'Aquatic acidification: a mechanism underpinning maintained oxygen transport and performance in fish experiencing elevated carbon dioxide conditions', *Journal of Experimental Biology*, 221(5). Available at: https://doi.org/10.1242/jeb.154559.
He, W., Cao, Z.D. and Fu, S.J. (2015) 'Effect of temperature on hypoxia tolerance and its underlying biochemical mechanism in two juvenile cyprinids exhibiting distinct hypoxia sensitivities', *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, 187, pp. 232–241. Available at: https://doi.org/10.1016/J.CBPA.2014.05.004.
He, Z., Crenshaw, E. and Raper, J.A. (2019) 'Semaphorin/neuropilin binding specificities are stable over 400 million years of evolution', *Biochemical and Biophysical Research Communications*, 517(1), pp. 23–28. Available at: https://doi.org/10.1016/j.bbrc.2019.06.133.

Hearn, A. *et al.* (2010) 'Hotspots within hotspots? Hammerhead shark movements around Wolf Island, Galapagos Marine Reserve', *Marine Biology*, 157(9), pp. 1899–1915. Available at: https://doi.org/10.1007/s00227-010-1460-2.

Heck, M. and Neely, B.A. (2020) 'Proteomics in Non-model Organisms: A New Analytical Frontier', *Journal of Proteome Research*. American Chemical Society, pp. 3595–3606. Available at: https://doi.org/10.1021/acs.jproteome.0c00448.

Hernández-Álvarez, C. *et al.* (2020) 'Phylogeography of the Pacific Red Snapper (Lutjanus Peru) and Spotted Rose Snapper (Lutjanus guttatus) in the Inshore Tropical Eastern Pacific', *Copeia*, 108(1), pp. 61–71. Available at: https://doi.org/10.1643/CG-18-157.

Hernández-Álvarez, Y. *et al.* (2023) 'Spatiotemporal Variability in Fish Assemblages in a Coastal and Estuarine System in the Tropical Eastern Pacific during the Anthropause', *Diversity*, 15(8). Available at: https://doi.org/10.3390/d15080934.

Hoopmann, M.R. and Moritz, R.L. (2013) 'Current algorithmic solutions for peptide-based proteomics data generation and identification', *Current Opinion in Biotechnology*, 24(1), pp. 31–38. Available at: https://doi.org/10.1016/j.copbio.2012.10.013.

Houten, S.M., Wanders, R.J.A. and Ranea-Robles, P. (2020) 'Metabolic interactions between peroxisomes and mitochondria with a special focus on acylcarnitine metabolism', *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1866(5), p. 165720. Available at: https://doi.org/10.1016/J.BBADIS.2020.165720.

Hu, C. *et al.* (2022) 'Heat shock proteins: Biological functions, pathological roles, and therapeutic opportunities', *MedComm*. John Wiley and Sons Inc. Available at: https://doi.org/10.1002/mco2.161.

Huey, R.B. *et al.* (2012) 'Predicting organismal vulnerability to climate warming: Roles of behaviour, physiology and adaptation', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1596), pp. 1665–1679. Available at: https://doi.org/10.1098/rstb.2012.0005.

Hughes, C.S. *et al.* (2014) 'Ultrasensitive proteome analysis using paramagnetic bead technology', *Molecular Systems Biology*, 10(10). Available at: https://doi.org/10.15252/msb.20145625.

Huth, T.J. and Place, S.P. (2016) 'RNA-seq reveals a diminished acclimation response to the combined effects of ocean acidification and elevated seawater temperature in Pagothenia borchgrevinki', *Marine Genomics*, 28, pp. 87–97. Available at: https://doi.org/10.1016/J.MARGEN.2016.02.004.

Ishida, Y. and Nagata, K. (2011) 'Hsp47 as a Collagen-Specific Molecular Chaperone', *Methods in Enzymology*, 499, pp. 167–182. Available at: https://doi.org/10.1016/B978-0-12-386471-0.00009-2.

Ishikawa, Y., Holden, P. and Bächinger, H.P. (2017) 'Heat shock protein 47 and 65-kDa FK506-binding protein weakly but synergistically interact during collagen folding in the endoplasmic reticulum', *Journal of Biological Chemistry*, 292(42), pp. 17216–17224. Available at: https://doi.org/10.1074/jbc.M117.802298.

Ishimatsu, A. *et al.* (2005) 'Physiological effects on fishes in a high-CO2 world', *Journal of Geophysical Research: Oceans*, pp. 1–8. Available at: https://doi.org/10.1029/2004JC002564.

Ishimatsu, A., Hayashi, M. and Kikkawa, T. (2008) 'Fishes in high-CO_2, acidified oceans', *Marine Ecology Progress Series*, 373, pp. 295–302. Available at: https://doi.org/10.2307/24872934.

IUCN (2023) 'The IUCN red list of threatened species', Version 2023-1.

Jiahuan, R. *et al.* (2018) 'Ocean acidification impairs foraging behavior by interfering with olfactory neural signal transduction in black sea bream, acanthopagrus schlegelii', *Frontiers in Physiology*, 9(NOV). Available at: https://doi.org/10.3389/fphys.2018.01592.

Johansen, J.L. *et al.* (2021) 'Thermal acclimation of tropical coral reef fishes to global heat waves', *eLife*, 10. Available at: https://doi.org/10.7554/eLife.59162.

Johnson, C.H. and Gonzalez, F.J. (2012) 'Challenges and opportunities of metabolomics', *Journal of Cellular Physiology*, 227(8), pp. 2975–2981. Available at: https://doi.org/10.1002/jcp.24002.

Joyce, A.R. and Palsson, B. (2006) 'The model organism as a system: Integrating "omics" data sets', *Nature Reviews Molecular Cell Biology*, pp. 198–210. Available at: https://doi.org/10.1038/nrm1857.

Jutfelt, F. *et al.* (2021) '"Aerobic scope protection" reduces ectotherm growth under warming', *Functional Ecology*, 35(7), pp. 1397–1407. Available at: https://doi.org/10.1111/1365-2435.13811.

Kinkead, R. *et al.* (1993) 'Propranolol Impairs the Hyperventilatory Response to Acute Hypercapnia in Rainbow Trout', *Journal of Experimental Biology*, 175(1), pp. 115–126. Available at: https://doi.org/10.1242/jeb.175.1.115.

Kopka, J. *et al.* (2005) 'GMD@CSB.DB: The Golm metabolome database', *Bioinformatics*, 21(8), pp. 1635–1638. Available at: https://doi.org/10.1093/bioinformatics/bti236.

Kumar, P. *et al.* (2015) 'Biochemical and Physiological Stress Responses to Heat Shock and Their Recovery in Labeo rohita Fingerlings', *Proceedings of the National Academy of Sciences India Section B - Biological Sciences*, 85(2), pp. 485–490. Available at: https://doi.org/10.1007/s40011-014-0357-0.

Langmead, B. and Salzberg, S.L. (2012) 'Fast gapped-read alignment with Bowtie 2', *Nature Methods*, 9(4), pp. 357–359. Available at: https://doi.org/10.1038/nmeth.1923.

Larios-Soriano, E. *et al.* (2020) 'Effect of acclimation temperature on thermoregulatory behaviour, thermal tolerance and respiratory metabolism of Lutjanus guttatus and the response of heat shock protein 70 (Hsp70) and lactate dehydrogenase (Ldh-a) genes', *Aquaculture Research*, 51(3), pp. 1089–1100. Available at: https://doi.org/10.1111/are.14455.

Lavín, M.F. *et al.* (2006) 'A review of eastern tropical Pacific oceanography: Summary', *Progress in Oceanography*, 69(2–4), pp. 391–398. Available at: https://doi.org/10.1016/j.pocean.2006.03.005.

Leung, J.C.H. *et al.* (2022) 'Differential expansion speeds of Indo-Pacific warm pool and deep convection favoring pool under greenhouse warming', *npj Climate and Atmospheric Science*, 5(1). Available at: https://doi.org/10.1038/s41612-022-00315-w.

Li, S. *et al.* (2022) 'Physiological responses to heat stress in the liver of rainbow trout (Oncorhynchus mykiss) revealed by UPLC-QTOF-MS metabolomics and biochemical assays', *Ecotoxicology and Environmental Safety*, 242, p. 113949. Available at: https://doi.org/10.1016/j.ecoenv.2022.113949.

Loarie, S.R. *et al.* (2009) 'The velocity of climate change', *Nature*, 462(7276), pp. 1052–1055. Available at: https://doi.org/10.1038/nature08649.

Lucano-Ramírez, G. *et al.* (2023) 'Reproduction of Lutjanus guttatus (Perciformes: Lutjanidae) captured in the Mexican Central Pacific', *Latin American Journal of Aquatic Research*, 51(4), pp. 503–520. Available at: https://doi.org/10.3856/vol51-issue4-fulltext-3008.

Luedemann, A. *et al.* (2008) 'TagFinder for the quantitative analysis of gas chromatography - Mass spectrometry (GC-MS)-based metabolite profiling experiments', *Bioinformatics*, 24(5), pp. 732–737. Available at: https://doi.org/10.1093/bioinformatics/btn023.

Luo, M. *et al.* (2024) 'Proteomics and metabolomics analysis of American shad (Alosa sapidissima) liver responses to heat stress', *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, 296, p. 111686. Available at: https://doi.org/10.1016/J.CBPA.2024.111686.

Lushchak, V.I. (2011) 'Environmentally induced oxidative stress in aquatic animals', *Aquatic Toxicology*, pp. 13–30. Available at: https://doi.org/10.1016/j.aquatox.2010.10.006.

Madeira, D. *et al.* (2013) 'Influence of temperature in thermal and oxidative stress responses in estuarine fish', *Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology*, 166(2), pp. 237–243. Available at: https://doi.org/10.1016/j.cbpa.2013.06.008.

Manzoni, C. *et al.* (2018) 'Genome, transcriptome and proteome: The rise of omics data and their integration in biomedical sciences', *Briefings in Bioinformatics*, 19(2), pp. 286–302. Available at: https://doi.org/10.1093/BIB/BBW114.

Del Mar Palacios, M. and Zapata, F.A. (2014) *Fish community structure on coral habitats with contrasting architecture in the Tropical Eastern Pacific*, *Rev. Biol. Trop. (Int. J. Trop. Biol. ISSN*.

Mariana, S., Alfons and Badr, G. (2019) 'Impact of heat stress on the immune response of fishes', *Journal of Survey in Fisheries Sciences*, 5(2), pp. 149–159. Available at: https://doi.org/10.18331/sfs2019.5.2.14.

Marquioni, V., Nunes, F.M.F. and Novo-Mansur, M.T.M. (2021) 'Protein Identification by Database Searching of Mass Spectrometry Data in the Teaching of Proteomics', *Journal of Chemical Education*, 98(3), pp. 812–823. Available at: https://doi.org/10.1021/acs.jchemed.0c00853.

Martínez-Cordero, F.J., Sanchez-Zazueta, E. and Hernández, C. (2018) 'Investment analysis of marine cage culture by applying bioeconomic reference points: A case study of the spotted rose snapper (*Lutjanus guttatus*) in Mexico', *Aquaculture Economics & Management*, 22(2), pp. 209–228. Available at: https://doi.org/10.1080/13657305.2017.1295489.

McMahon, S.J., Munday, P.L. and Donelson, J.M. (2024) 'The effects of marine heatwaves on a coral reef snapper: insights into aerobic and anaerobic physiology and recovery', *Conservation Physiology*. Edited by K. Ruthsatz, 12(1). Available at: https://doi.org/10.1093/conphys/coae060.

Mirasole, A. *et al.* (2017) 'The influence of high pCO2 on otolith shape, chemical and carbon isotope composition of six coastal fish species in a Mediterranean shallow CO2 vent', *Marine Biology*, 164(9). Available at: https://doi.org/10.1007/s00227-017-3221-y.

Moggridge, S. *et al.* (2018) 'Extending the Compatibility of the SP3 Paramagnetic Bead Processing Approach for Proteomics', *Journal of Proteome Research*, 17(4), pp. 1730–1740. Available at: https://doi.org/10.1021/acs.jproteome.7b00913.

Mojzikova, R. *et al.* (2018) 'Two novel mutations (p.(Ser160Pro) and p.(Arg472Cys)) causing glucose-6-phosphate isomerase deficiency are associated with erythroid dysplasia and inappropriately suppressed hepcidin', *Blood Cells, Molecules, and Diseases*, 69, pp. 23–29. Available at: https://doi.org/10.1016/J.BCMD.2017.04.003.

Mora, C. and Robertson, D.R. (2005) 'Factors shaping the range-size frequency distribution of the endemic fish fauna of the Tropical Eastern Pacific', *Journal of Biogeography*, 32(2), pp. 277–286. Available at: https://doi.org/10.1111/j.1365-2699.2004.01155.x.

Moreira, J.M. *et al.* (2022) 'Impacts of ocean warming and acidification on the energy budget of three commercially important fish species', *Conservation Physiology*, 10(1). Available at: https://doi.org/10.1093/conphys/coac048.

Munday, P. *et al.* (2011) 'Ocean acidification does not affect the early life history development of a tropical marine fish', *Marine Ecology Progress Series*, 423, pp. 211–221. Available at: https://doi.org/10.3354/meps08990.

Munday, P.L., Crawley, N.E. and Nilsson, G.E. (2009) 'Interacting effects of elevated temperature and ocean acidification on the aerobic performance of coral reef fishes', *Marine Ecology Progress Series*, 388, pp. 235–242. Available at: https://doi.org/10.3354/meps08137.

Naveenkumar, N. *et al.* (2022) 'Structures of distantly related interacting protein homologs are less divergent than non-interacting homologs', *FEBS Open Bio*, 12(12), pp. 2147–2153. Available at: https://doi.org/10.1002/2211-5463.13492.

Patra, K.C. and Hay, N. (2014) 'The pentose phosphate pathway and cancer', *Trends in Biochemical Sciences*, 39(8), pp. 347–354. Available at: https://doi.org/10.1016/J.TIBS.2014.06.005.

Perry, D.M. *et al.* (2015) 'Effect of ocean acidification on growth and otolith condition of juvenile scup, *Stenotomus chrysops*', *Ecology and Evolution*, 5(18), pp. 4187–4196. Available at: https://doi.org/10.1002/ece3.1678.

Perry, S.F. and Gilmour, K.M. (2006) 'Acid–base balance and CO2 excretion in fish: Unanswered questions and emerging models', *Respiratory Physiology & Neurobiology*, 154(1–2), pp. 199–215. Available at: https://doi.org/10.1016/j.resp.2006.04.010.

Ponting, C.P. (2001) 'Issues in predicting protein function from sequence', *Briefings in Bioinformatics*, 2(1), pp. 19–21. Available at: https://doi.org/10.1093/bib/2.1.19.

Pörtner, H.O. (2012) 'Integrating climate-related stressor effects on marine organisms: Unifying principles linking molecule to ecosystem-level changes', *Marine Ecology Progress Series*, 470, pp. 273–290. Available at: https://doi.org/10.3354/meps10123.

Pörtner, H.O., Langenbuch, M. and Reipschläger, A. (2004) *Biological Impact of Elevated Ocean CO 2 Concentrations: Lessons from Animal Physiology and Earth History*, *Journal of Oceanography*.

Qin, S. *et al.* (2022) 'Serine protease PRSS23 drives gastric cancer by enhancing tumor associated macrophage infiltration via FGF2', *Frontiers in Immunology*, 13. Available at: https://doi.org/10.3389/fimmu.2022.955841.

R Core Team (2023) 'R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.' Available at: https://www.R-project.org/.

Rashidpour, A. *et al.* (2019) 'Metformin counteracts glucose-dependent lipogenesis and impairs transdeamination in the liver of gilthead sea bream (Sparus aurata)', *American Journal of Physiology - Regulatory Integrative and Comparative Physiology*, 316(3), pp. R265–R273. Available at: https://doi.org/10.1152/ajpregu.00216.2018.

Reusch, T.B.H. (2014) 'Climate change in the oceans: Evolutionary versus phenotypically plastic responses of marine animals and plants', *Evolutionary Applications*, 7(1), pp. 104–122. Available at: https://doi.org/10.1111/eva.12109.

Rijnsdorp, A.D. *et al.* (2009) 'Resolving the effect of climate change on fish populations', *Journal of Marine Science*, 66, pp. 1570–1583. Available at: https://academic.oup.com/icesjms/article/66/7/1570/657377.

Ritchie, M.E. *et al.* (2015) 'limma powers differential expression analyses for RNA-sequencing and microarray studies', *Nucleic Acids Research*, 43(7), pp. e47–e47. Available at: https://doi.org/10.1093/nar/gkv007.

Robertson, D.R. and Cramer, K.L. (2009) 'Shore fishes and biogeographic subdivisions of the Tropical Eastern Pacific', *Marine Ecology Progress Series*, 380, pp. 1–17. Available at: https://doi.org/10.3354/meps07925.

Rogelj, J. *et al.* (2018) 'Mitigation Pathways Compatible with 1.5°C in the Context of Sustainable Development', *Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change* [Preprint].

Rohart, F. *et al.* (2017) 'mixOmics: An R package for 'omics feature selection and multiple data integration', *PLOS Computational Biology*, 13(11), p. e1005752. Available at: https://doi.org/10.1371/journal.pcbi.1005752.

Ross, M.K. *et al.* (2010) 'Carboxylesterases: dual roles in lipid and pesticide metabolism', *Journal of Pesticide Science*, 35(3), pp. 257–264. Available at: https://doi.org/10.1584/jpestics.R10-07.

Rossi, A., Bacchetta, C. and Cazenave, J. (2017) 'Effect of thermal stress on metabolic and oxidative stress biomarkers of Hoplosternum littorale (Teleostei, Callichthyidae)', *Ecological Indicators*, 79, pp. 361–370. Available at: https://doi.org/10.1016/j.ecolind.2017.04.042.

Rummer, J.L. *et al.* (2014) 'Life on the edge: Thermal optima for aerobic scope of equatorial reef fishes are close to current day temperatures', *Global Change Biology*, 20(4), pp. 1055–1066. Available at: https://doi.org/10.1111/gcb.12455.

Rutterford, L.A. *et al.* (2023) 'Sea temperature is the primary driver of recent and predicted fish community structure across Northeast Atlantic shelf seas', *Global Change Biology*, 29(9), pp. 2510–2521. Available at: https://doi.org/10.1111/gcb.16633.

Sanchez, B.C., Ralston-Hooper, K. and Sepúlveda, M.S. (2011) 'Review of recent proteomic applications in aquatic toxicology', *Environmental Toxicology and Chemistry*, pp. 274–282. Available at: https://doi.org/10.1002/etc.402.

Schauer, N. *et al.* (2005) 'GC-MS libraries for the rapid identification of metabolites in complex biological samples', *FEBS Letters*, 579(6), pp. 1332–1337. Available at: https://doi.org/10.1016/j.febslet.2005.01.029.

Schulz-Mirbach, T. *et al.* (2019) 'Enigmatic ear stones: what we know about the functional role and evolution of fish otoliths', *Biological Reviews*, 94(2), pp. 457–482. Available at: https://doi.org/10.1111/brv.12463.

Servili, A. *et al.* (2022) 'Ocean acidification alters the acute stress response of a marine fish', *Science of The Total Environment* [Preprint], (858).
Shao, Y.T. *et al.* (2016) 'Acidified seawater suppresses insulin-like growth factor I mRNA expression and reduces growth rate of juvenile orange-spotted groupers, Epinephelus coioides (Hamilton, 1822)', *Aquaculture Research*, 47(3), pp. 721–731. Available at: https://doi.org/10.1111/are.12533.
Sokolova, I.M. *et al.* (2012) 'Energy homeostasis as an integrative tool for assessing limits of environmental stress tolerance in aquatic invertebrates', *Marine Environmental Research*, 79, pp. 1–15. Available at: https://doi.org/10.1016/J.MARENVRES.2012.04.003.
Sokolova, I.M. (2013) 'Energy-limited tolerance to stress as a conceptual framework to integrate the effects of multiple stressors', *Integrative and Comparative Biology*, 53(4), pp. 597–608. Available at: https://doi.org/10.1093/icb/ict028.
Somero, G.N. (2011) 'Comparative physiology: a "crystal ball" for predicting consequences of global change', *Am J Physiol Regul Integr Comp Physiol*, 301, pp. 1–14. Available at: https://doi.org/10.1152/ajpregu.00719.2010.-Com.
Soto-Rojas, R.L., Hernández-Noguera, L.A. and Vega-Alpízar, J.L. (2018) 'Parámetros poblacionales y hábitos alimenticios del pargo mancha (Lutjanus guttatus) en el Área Marina de Pesca Responsable Paquera-Tambor, golfo de Nicoya, Costa Rica', *Uniciencia*, 32(2), p. 96. Available at: https://doi.org/10.15359/ru.32-2.7.
De Souza, K. *et al.* (2014) 'Effects of Increased CO 2 on Fish Gill and Plasma Proteome', *PLoS ONE*, 9(7), p. 102901. Available at: https://doi.org/10.1371/journal.pone.0102901.t001.
Spigarelli, S.A., Thommes, M.M. and Prepejchal, W. (1982) 'Feeding, Growth, and Fat Deposition by Brown Trout in Constant and Fluctuating Temperatures', *Transactions of the American Fisheries Society*, 111(2), pp. 199–209. Available at: https://doi.org/10.1577/1548-8659(1982)111<199:fgafdb>2.0.co;2.
Steinhauser, D. and Kopka, J. (2007) 'Methods, applications and concepts of metabolite profiling: Primary metabolism', in *Plant Systems Biology*. Basel: Birkhäuser Basel, pp. 171–194. Available at: https://doi.org/10.1007/978-3-7643-7439-6_8.
Sun, G. *et al.* (2010) 'Zebrafish complement factor H and its related genes: Identification, evolution, and expression', *Functional and Integrative Genomics*, 10(4), pp. 577–587. Available at: https://doi.org/10.1007/s10142-010-0182-3.
Sundin, J. (2023) 'The effects of ocean acidification on fishes – history and future outlook', *Journal of Fish Biology*. John Wiley and Sons Inc, pp. 765–772. Available at: https://doi.org/10.1111/jfb.15323.

Takeda, T. (1991) 'Regulation of blood oxygenation during short-term hypercapnia in the carp, Cyprinus carpio', *Comparative Biochemistry and Physiology*, 98(3–4), pp. 517–521. Available at: https://doi.org/https://doi.org/10.1016/0300-9629(91)90440-N.

Tewksbury, J.J., Huey, R.B. and Deutsch, C.A. (2008) 'Ecology: Putting the heat on tropical animals', *Science*, pp. 1296–1297. Available at: https://doi.org/10.1126/science.1159328.

Thomas, S. *et al.* (1983) *Comparison of the Effects of Exogenous and Endogenous Hypercapnia on Ventilation and Oxygen Uptake in the Rainbow Trout (SMmo gMrdneri R.)*, *Journal of Comparative Physiology. B*. Springer-Verlag.

Thoral, E. *et al.* (2022) 'Absence of mitochondrial responses in muscles of zebrafish exposed to several heat waves', *Comparative Biochemistry and Physiology -Part A : Molecular and Integrative Physiology*, 274. Available at: https://doi.org/10.1016/j.cbpa.2022.111299.

Tian, W.N. *et al.* (1998) 'Importance of glucose-6-phosphate dehydrogenase activity for cell growth', *Journal of Biological Chemistry*, 273(17), pp. 10609–10617. Available at: https://doi.org/10.1074/jbc.273.17.10609.

Timson, D.J. (2019) 'Fructose 1,6-bisphosphatase: getting the message across', *Bioscience Reports* [Preprint]. Available at: https://doi.org/10.1042/BSR20190124.

Torres, J.J., Grigsby, M.D. and Elizabeth Clarke, M. (2012) 'Aerobic and anaerobic metabolism in oxygen minimum layer fishes: The role of alcohol dehydrogenase', *Journal of Experimental Biology*, 215(11), pp. 1905–1914. Available at: https://doi.org/10.1242/jeb.060236.

Tort, L. (2011) 'Stress and immune modulation in fish', *Developmental & Comparative Immunology*, 35(12), pp. 1366–1375. Available at: https://doi.org/10.1016/J.DCI.2011.07.002.

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

Vinagre, C. *et al.* (2016) 'Vulnerability to climate warming and acclimation capacity of tropical and temperate coastal organisms', *Ecological Indicators*, 62, pp. 317–327. Available at: https://doi.org/10.1016/J.ECOLIND.2015.11.010.

Ward, D. *et al.* (2022) 'Safeguarding marine life: conservation of biodiversity and ecosystems', *Reviews in Fish Biology and Fisheries*, 32(1), pp. 65–100. Available at: https://doi.org/10.1007/s11160-022-09700-3.

Watson, S.A., Fields, J.B. and Munday, P.L. (2017) 'Ocean acidification alters predator behaviour and reduces predation rate', *Biology Letters*, 13(2). Available at: https://doi.org/10.1098/rsbl.2016.0797.

Wernberg, T., Smale, D.A. and Thomsen, M.S. (2012) 'A decade of climate change experiments on marine organisms: Procedures, patterns and problems', *Global Change Biology*, pp. 1491–1498. Available at: https://doi.org/10.1111/j.1365-2486.2012.02656.x.

Widlansky, M.J., Long, X. and Schloesser, F. (2020) 'Increase in sea level variability with ocean warming associated with the nonlinear thermal expansion of seawater', *Communications Earth & Environment*, 1(1), p. 9. Available at: https://doi.org/10.1038/s43247-020-0008-8.
Wolf-Gladrow, D.A. *et al.* (1999) 'Direct effects of CO2 concentration on growth and isotopic composition of marine plankton', *Tellus, Series B: Chemical and Physical Meteorology*, 51(2), pp. 461–476. Available at: https://doi.org/10.3402/tellusb.v51i2.16324.
Zhang, Z. *et al.* (2021) 'Metabolomics analysis of the effects of temperature on the growth and development of juvenile European seabass (Dicentrarchus labrax)', *Science of the Total Environment*, 769. Available at: https://doi.org/10.1016/j.scitotenv.2021.145155.

A

Appendix

A.1 Figures and Tables

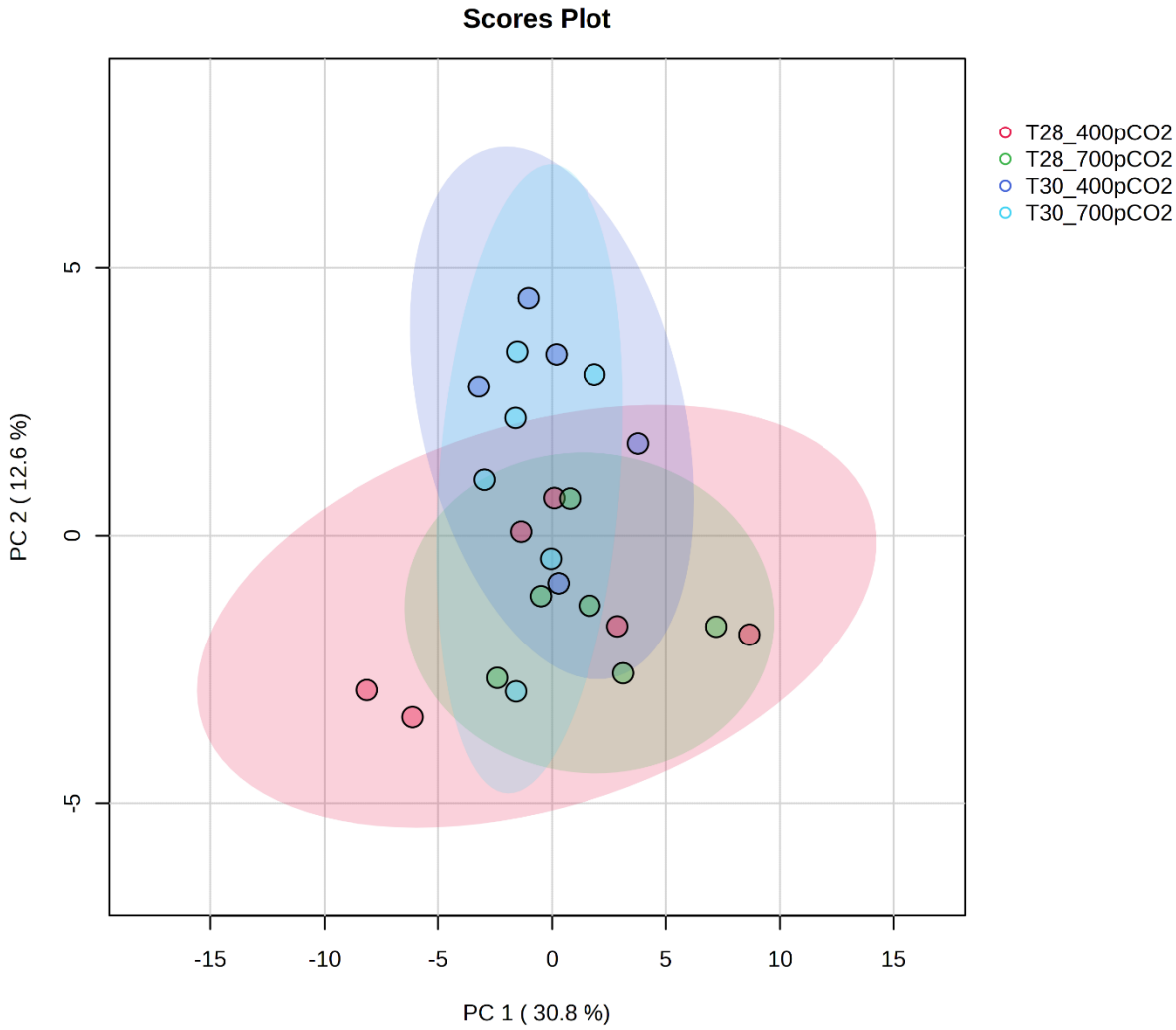

Figure A.1. PCA scores plot of the first two components based on metabolite concentrations

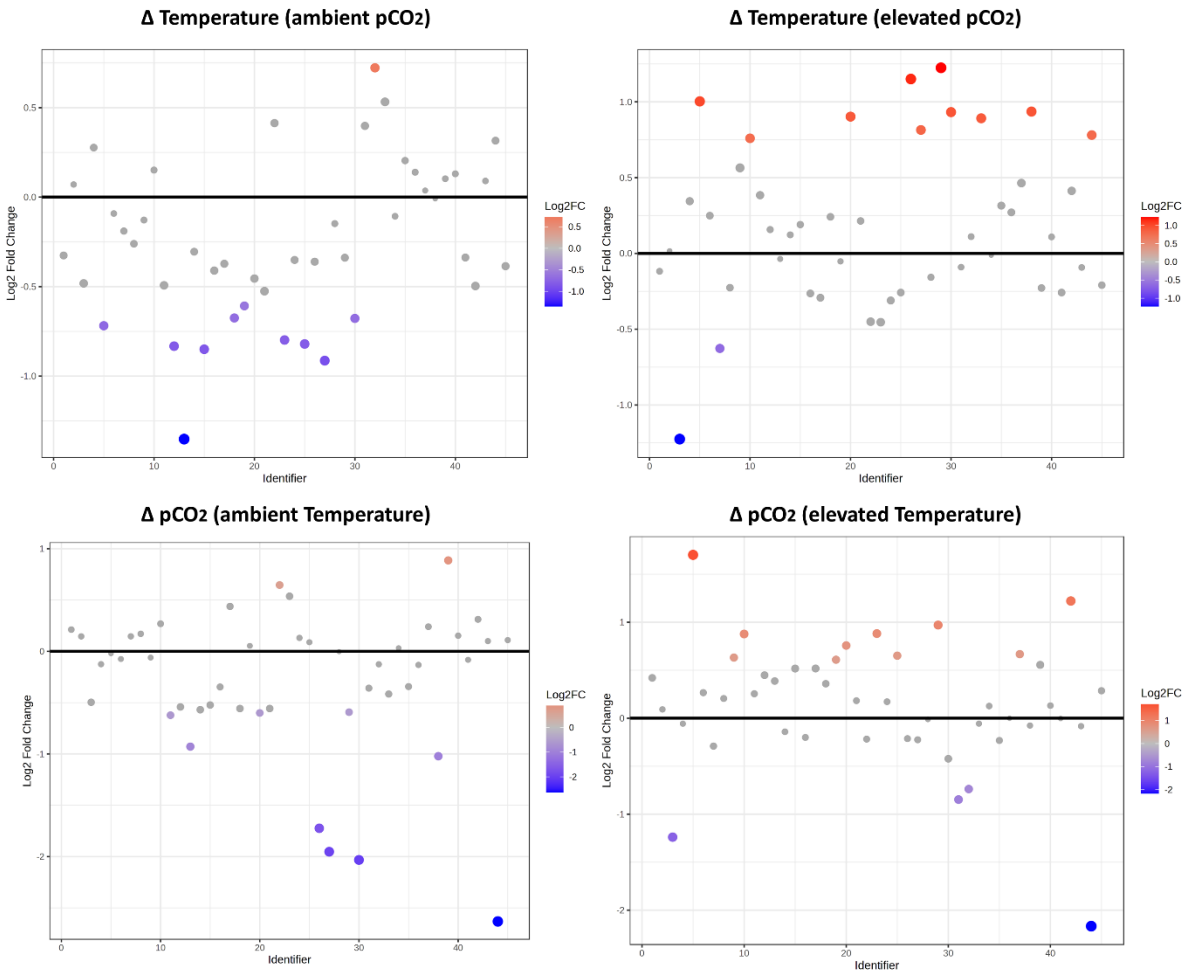

Figure A.2. Scatterplot showcasing the differential abundance of all detected metabolites in the four conditions. The y- axis shows the log2 fold change, and the vertical bar marks the zero position on the y- axis. The x- axis is an identifier with each metabolite being one value apart. Metabolites with a log2 fold change higher than 1.5 or lower than 0.66 were marked in red or blue respectively. A) comparing both temperature treatments during 400 ppm CO2 conditions B) comparing both temperature treatments during 700 ppm CO2 conditions C) comparing both CO2 treatments during ambient temperatures D) comparing both CO2 treatments during elevated temperatures

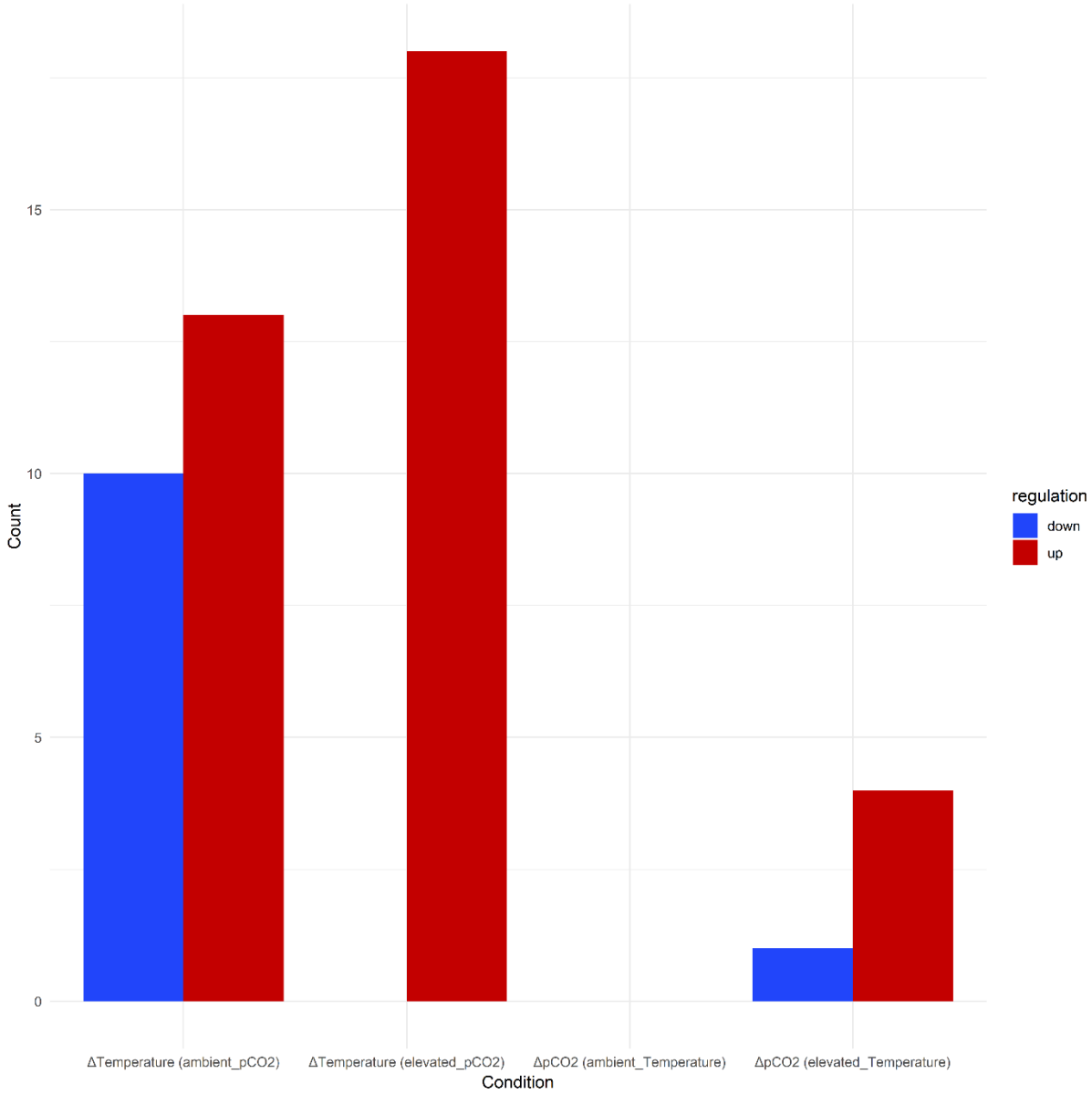

Figure A.3. Bar plot containing the total number of up and down regulated genes per contrast. Upregulated proteins are marked red and downregulated ones are marked blue.

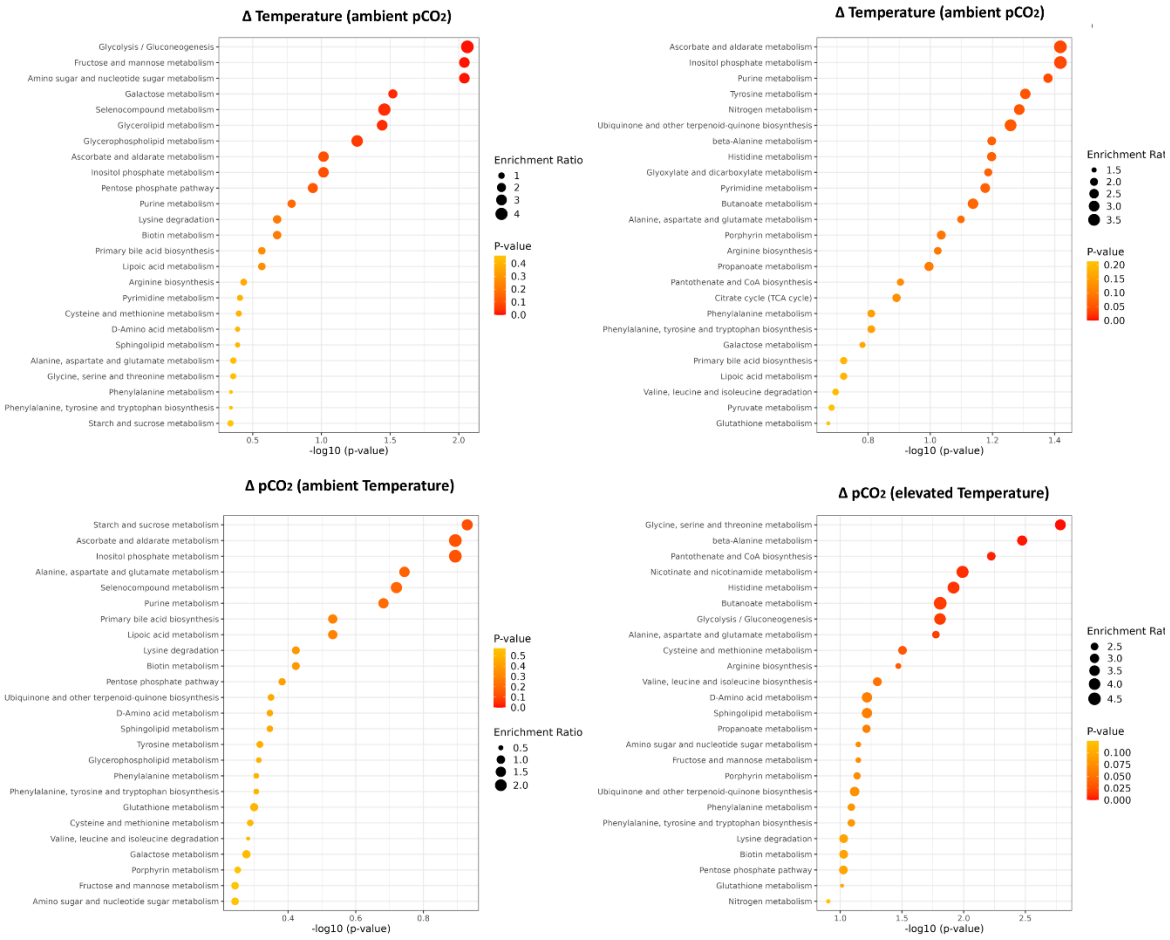

Figure A.4. Dot plot of top 25 enriched Metabolite sets in all contrasts. Dot size indicates the enrichment ratio, and dot color the p- value. The x- axis is the negative logarithm of the raw p- value. (note: the enrichment ratio is calculated by dividing observed metabolite hits per pathway by the expected hits) . A) describe the contrast; b) describe the contrast; C) describe the contrast; D) describe the contrast.

Table A1. List of all samples used in this work. The X indicate for which omics analysis the sample was used. Sampling day indicates how long the individuals were subjected to the experimental treatment. (note: sample LT1F2 was used in metabolomic analysis but was discarded as an outlier before bioinformatic analysis).

Sample ID	Metabolomics	Proteomics	Transcriptomics	temp treat (°C)	CO2 treat (ppm)	sampling day	sampling date	sample tissue
L_T21F1	x	x		ambient 28	ambient 400	exp_day_27	06.01.2023 00:00	Liver
L_T21F2	x		x	ambient 28	ambient 400	exp_day_27	06.01.2023 00:00	Liver
L_T22F1	x	x		ambient 28	ambient 400	exp_day_27	06.01.2023 00:00	Liver
L_T22F2	x		x	ambient 28	ambient 400	exp_day_27	06.01.2023 00:00	Liver
L_T23F1	x	x		ambient 28	ambient 400	exp_day_27	06.01.2023 00:00	Liver
L_T24F1	x	x		ambient 28	ambient 400	exp_day_27	06.01.2023 00:00	Liver
L_24_F2			x	ambient_28	ambient 400	exp_day_27	06.01.2023 00:00	Liver
L_T17F1	x	x		ambient 28	acidified 700	exp_day_27	06.01.2023 00:00	Liver
L_T17F2	x	x		ambient 28	acidified 700	exp_day_27	06.01.2023 00:00	Liver
L_T18F1	x	x		ambient 28	acidified 700	exp_day_27	06.01.2023 00:00	Liver
L_T18F2	x			ambient 28	acidified 700	exp_day_27	06.01.2023 00:00	Liver
L_T19F1	x	x		ambient 28	acidified 700	exp_day_27	06.01.2023 00:00	Liver
L_T20F1	x	x		ambient 28	acidified 700	exp_day_27	06.01.2023 00:00	Liver
L_T9F1	x	x		elevated 30	acidified 700	exp_day_26	06.01.2023 00:00	Liver
L_T9F2	x	x		elevated 30	acidified 700	exp_day_26	06.01.2023 00:00	Liver
L_T10F1	x	x		elevated 30	acidified 700	exp_day_26	06.01.2023 00:00	Liver
L_T10F2	x			elevated 30	acidified 700	exp_day_26	06.01.2023 00:00	Liver
L_T11F1	x	x		elevated 30	acidified 700	exp_day_26	06.01.2023 00:00	Liver
L_T12F1	x	x		elevated 30	acidified 700	exp_day_26	06.01.2023 00:00	Liver
L_T1F1	x	x		elevated 30	ambient 400	exp_day_26	06.01.2023 00:00	Liver
L_T1F2	x			elevated 30	ambient 400	exp_day_26	06.01.2023 00:00	Liver
L_T2F1	x	x		elevated 30	ambient 400	exp_day_26	06.01.2023 00:00	Liver
L_T2F2	x			elevated 30	ambient 400	exp_day_26	06.01.2023 00:00	Liver
L_T3F1	x	x		elevated 30	ambient 400	exp_day_26	06.01.2023 00:00	Liver
L_T4F1	x	x		elevated 30	ambient 400	exp_day_26	06.01.2023 00:00	Liver

Table A2. All identified differentially regulated Proteins found in all contrasts. Uniprot ID and Genename were identified using homology matching. The Species coloumn indicates the original species of the matched sequence. The regulations columns indicate if a protein was significantly up- or downregulated (FDR < 0.05).

UNIPROT ID	Genename	Species	logFC temp (ambient pH)	regulation	logFC pH (ambient temp)	regulation	logFC temp (acidified pH)	regulation	logFC pH (elevated temp)	regulation
P52758	2-iminobutanoate/2-iminopropanoate deaminase	Homo sapiens	-0.0185		-0.2975		0.7802	upregulated	0.5013	
P49645	Alcohol dehydrogenase 1	Apteryx australis	1.6435	upregulated	0.9689		-0.6835		-1.3581	
P26325	Alcohol dehydrogenase 1	Gadus morhua subsp. callarias	0.9074	upregulated	0.4289		0.4718		-0.0067	
Q8MI29	Carbonyl reductase [NADPH] 1	Macaca fascicularis	-2.5456	downregulated	-2.2787		-0.1553		0.1116	
Q6AW47	Carboxylesterase 5A	Canis lupus familiaris	-0.7206	downregulated	0.0425		0.0057		0.7688	upregulated
Q9DEX3	Cathepsin D	Clupea harengus	0.4390	upregulated	0.2210		-0.0064		-0.2244	
Q28085	Complement factor H	Bos taurus	0.2606		-0.0793		0.7392	upregulated	0.3993	
Q28085	Complement factor H	Bos taurus	0.2606		-0.0793		0.7392	upregulated	0.3993	
Q28085	Complement factor H	Bos taurus	0.2606		-0.0793		0.7392	upregulated	0.3993	
Q28085	Complement factor H	Bos taurus	0.2606		-0.0793		0.7392	upregulated	0.3993	
Q9DCT8	Cysteine-rich protein 2	Mus musculus	-2.4688	downregulated	-2.4826		0.4873		0.4736	
Q9I8U0	Cytochrome c oxidase subunit 4 isoform 1, mitochondrial	Thunnus obesus	-0.6016	downregulated	-0.0355		-0.3742		0.1918	
P31689	DnaJ homolog subfamily A member 1	Homo sapiens	0.5126	upregulated	0.0333		0.2758		-0.2036	
Q04791	Fatty acyl-CoA hydrolase precursor, medium chain	Anas platyrhynchos	-0.7206	downregulated	0.0425		0.0057		0.7688	upregulated
Q04791	Fatty acyl-CoA hydrolase precursor, medium chain	Anas platyrhynchos	-0.7206	downregulated	0.0425		0.0057		0.7688	upregulated
Q04791	Fatty acyl-CoA hydrolase precursor, medium chain	Anas platyrhynchos	-0.7206	downregulated	0.0425		0.0057		0.7688	upregulated
P00636	Fructose-1,6-bisphosphatase 1	Sus scrofa	-0.8038	downregulated	-0.4444		-0.0783		0.2811	
Q3ZBD7	Glucose-6-phosphate isomerase	Bos taurus	0.4622		0.0712		0.8827	upregulated	0.4917	
Q6AZV1	Heat shock cognate protein HSP 90-beta	Xenopus laevis	0.9546	upregulated	0.0507		0.8007	upregulated	-0.1032	
P07901	Heat shock protein HSP 90-alpha	Mus musculus	0.8305	upregulated	0.0853		0.7937	upregulated	0.0485	
P07901	Heat shock protein HSP 90-alpha	Mus musculus	0.8305	upregulated	0.0853		0.7937	upregulated	0.0485	
Q5RBT4	Isocitrate dehydrogenase [NAD] subunit beta, mitochondrial	Pongo abelii	-0.5238	downregulated	-0.1870		-0.1327		0.2041	
Q076A3	Myosin-13	Canis lupus familiaris	-0.8987	downregulated	-0.4840		-0.1736		0.2411	
Q6JHU7	Prolyl 3-hydroxylase 2	Gallus gallus	0.6140	upregulated	0.1868		0.2209		-0.2063	
Q5ZLK5	Prolyl 4-hydroxylase subunit alpha-2	Gallus gallus	0.1996		-0.1034		0.6719	upregulated	0.3689	
Q5ZLK5	Prolyl 4-hydroxylase subunit alpha-2	Gallus gallus	0.1661		-0.1033		0.6159	upregulated	0.3465	
Q13275	Semaphorin-3F	Homo sapiens	0.8225	upregulated	0.1725		-0.0987		-0.7488	downregulated
Q1LZE9	Serine protease 23	Bos taurus	0.9224	upregulated	0.0320		0.5575		-0.3329	
O08679	Serine/threonine-protein kinase MARK2	Rattus norvegicus	0.4115		0.0705		0.7349	upregulated	0.3939	
P13731	Serpin H1	Gallus gallus	0.5111		-0.2432		0.9108	upregulated	0.1566	
P13731	Serpin H1	Gallus gallus	0.6840		-0.1532		1.1585	upregulated	0.3212	
P13731	Serpin H1	Gallus gallus	0.5031		-0.2213		0.8557	upregulated	0.1313	
P13731	Serpin H1	Gallus gallus	0.5031		-0.2213		0.8557	upregulated	0.1313	
O75751	Solute carrier family 22 member 3	Homo sapiens	0.7222	upregulated	0.1515		0.0996		-0.4711	
O88446	Solute carrier family 22 member 3	Rattus norvegicus	0.7222	upregulated	0.1515		0.0996		-0.4711	
P78362	SRSF protein kinase 2	Homo sapiens	1.8459	upregulated	0.1919		1.0089	upregulated	-0.6451	
P80961	Type-4 ice-structuring protein LS-12	Myoxocephalus octodecemspinosus	0.0531		-0.2580		0.9287	upregulated	0.6176	

Table A.3. Complete list of metabolites identified by LC- TOF- MS,

Metabolite name	Stereoisomer	Formula	Class	CAS number	Kegg ID	HMDBID
Alanine	L	C3H7NO2	Amino acids, peptides, and analogues	302-72-7	C00041	HMDB0000161
Asparagine	L	C4H8N2O3	Amino acids, peptides, and analogues	2058-58-4	C00152	HMDB0000168
Aspartic acid		C4H7NO4	Amino acids, peptides, and analogues	56-84-8	C00049	HMDB0000191
beta-Alanine	L	C3H7NO2	Amino acids, peptides, and analogues	107-95-9	C00099	HMDB0000056
Cystathionine	L	C7H14N2O4S	Amino acids, peptides, and analogues	56-88-2	C02291	HMDB0000099
Glutamic acid	L	C5H9NO4	Amino acids, peptides, and analogues	56-86-0	C00025	HMDB0000148
Glutamine	L	C5H10N2O3	Amino acids, peptides, and analogues	56-85-9	C00064	HMDB0000641
Glycine		C2H5NO2	Amino acids, peptides, and analogues	56-40-6	C00037	HMDB0000123
Histidine	L	C6H9N3O2	Amino acids, peptides, and analogues	71-00-1	C00135	HMDB0000177
4-Hydroxyproline	L	C5H9NO3	Amino acids, peptides, and analogues	51-35-4	C01157	HMDB0000725
Isoleucine	L	C6H13NO2	Amino acids, peptides, and analogues	73-32-5	C00407	HMDB0000172
Lysine	L	C6H14N2O2	Amino acids, peptides, and analogues	56-87-1	C00047	HMDB0000182
Methionine	L	C5H11NO2S	Amino acids, peptides, and analogues	59-51-8	C00073	HMDB0000696
Ornithine	L	C5H12N2O2	Amino acids, peptides, and analogues	616-07-9	C00077	HMDB0003374
Phenylalanine	L	C9H11NO2	Amino acids, peptides, and analogues	63-91-2	C00079	HMDB0000159
Proline		C5H9NO2	Amino acids, peptides, and analogues	147-85-3	C00148	HMDB0000162
Pyroglutamic acid	L	C5H7NO3	Amino acids, peptides, and analogues	98-79-3	C01879	HMDB0000267
Serine	L	C3H7NO3	Amino acids, peptides, and analogues	56-45-1	C00716	HMDB0000187
Threonine	L	C4H9NO3	Amino acids, peptides, and analogues	72-19-5	C00188	HMDB0000167
Tyrosine	L	C9H11NO3	Amino acids, peptides, and analogues	60-18-4	C00082	HMDB0000158
Valine	L	C5H11NO2	Amino acids, peptides, and analogues	72-18-4	C00183	HMDB0000883
Fructose	D	C6H12O6	Carbohydrates and carbohydrate conjugate	57-48-7	C00095	HMDB0000660
Fructose-6-Posphate	D	C6H13O9P	Carbohydrates and carbohydrate conjugate	643-13-0	C00085	HMDB0000124
Galactose	D	C6H12O6	Carbohydrates and carbohydrate conjugate	59-23-4	C00124	HMDB0000143
Glucose	D-,alpha	C6H12O6	Carbohydrates and carbohydrate conjugate	50-99-7	C00031	HMDB0000122
Maltose		C12H22O11	Carbohydrates and carbohydrate conjugate	69-79-4	C00208	HMDB0000163
Maltotriose	D	C18H32O16	Carbohydrates and carbohydrate conjugate	1109-28-0	C01835	HMDB0001262
Mannose	D	C6H12O6	Carbohydrates and carbohydrate conjugate	3458-28-4	C00159	HMDB0000169
Sucrose	D	C12H22O11	Carbohydrates and carbohydrate conjugate	57-50-1	C00089	HMDB0000258
Trehalose	D-,alpha	C12H22O11	Carbohydrates and carbohydrate conjugate	99-20-7	C01083	HMDB0000975
Threonic acid	L	C4H8O5	Carbohydrates and carbohydrate conjugate	7306-96-9	C01620	HMDB0000943
Glycerol		C3H8O3	Carbohydrates and carbohydrate conjugate	56-81-5	C00116	HMDB0000131
myo -Inositol		C6H12O6	Alcohols and polyols	87-89-8	C00137	HMDB0000211
Threitol	D	C4H10O4	Carbohydrates and carbohydrate conjugate	2319-57-5	C16884	HMDB0004136
Fumaric acid		C4H4O4	Dicarboxylic acids and derivatives	110-17-8	C00122	HMDB0000134
Malic acid	L	C4H6O5	Beta hydroxy acids and derivatives	6915-15-7	C00149	HMDB0000156
Succinic acid		C4H6O4	Carboxylic acids and derivatives	110-15-6	C00042	HMDB0000254
Adenine		C4H6O5	Imidazopyrimidines	73-24-5	C00147	HMDB0000034
Inosine-5-mono-Phosphate		C10H14N5O7P	Purine nucleotides	61-19-8	C00020	HMDB0000045
Benzoic acid		C7H6O2	Benzene and substituted derivatives	65-85-0	C00180	HMDB0001870
Glycerol-3-Posphate	D	C3H9O6P	Glycerophospholipids	57-03-4	C03189	HMDB0000126
Nicotinamide		C6H6N2O	Pyridines and derivatives	98-92-0	C00153	HMDB0001406
Phosphoric acid		H3O4P	Non-metal oxoanionic compounds	7664-38-2	C00009	HMDB0001429
Uracil		C4H4N2O2	Diazines	66-22-8	C00106	HMDB0000300
Urea		CH4N2O	Organic carbonic acids and derivatives	57-13-6	C00086	HMDB0000294

A.2 Command line code

Trimming

trim_galore --paired --illumina --phred33 --output_dir [path]/Trimming --length 20 --stringency 1 -e 0.1 --gzip [path]/L_T21F2_N3143_R1_001.fastq.gz [path]/L_T21F2_N3143_R2_001.fastq.gz
trim_galore --paired --illumina --phred33 --output_dir [path]/Trimming --length 20 --stringency 1 -e 0.1 --gzip [path]/L_T23F2_N3144_R1_001.fastq.gz [path]/L_T23F2_N3144_R2_001.fastq.gz
trim_galore --paired --illumina --phred33 --output_dir [path]/Trimming --length 20 --stringency 1 -e 0.1 --gzip [path]/L_T24F2_N3145_R1_001.fastq.gz [path]/L_T24F2_N3145_R2_001.fastq.gz

FastQC

fastqc [path]/Trimming/L_T21F2_N3143_R1_001_val_1.fq.gz --outdir [path]/FastQC
fastqc [path]Trimming/L_T21F2_N3143_R2_001_val_2.fq.gz --outdir [path]/FastQC
fastqc [path]/Trimming/L_T23F2_N3144_R1_001_val_1.fq.gz --outdir [path]/FastQC
fastqc [path]/Trimming/L_T23F2_N3144_R2_001_val_2.fq.gz --outdir [path]/FastQC
fastqc [path]/Trimming/L_T24F2_N3145_R1_001_val_1.fq.gz --outdir [path]/FastQC
fastqc [path]/Trimming/L_T24F2_N3145_R2_001_val_2.fq.gz --outdir [path]/FastQC

Trinity

Trinity --seqType fq --left [path]/Trimming/L_T21F2_N3143_R1_001_val_1.fq.gz --right [path]/L_T21F2_N3143_R2_001_val_2.fq.gz --max_memory 200G --CPU 30 --no_salmon --full_cleanup --output [path]/Trinity
Trinity --seqType fq --left [path]/Trimming/L_T23F2_N3144_R1_001_val_1.fq.gz --right [path]/Trimming/L_T23F2_N3144_R2_001_val_2.fq.gz --max_memory 200G --CPU 30 --no_salmon --full_cleanup --output [path]/Trinity
Trinity --seqType fq --left [path]/Trimming/L_T24F2_N3145_R1_001_val_1.fq.gz --right [path]/Trimming/L_T24F2_N3145_R2_001_val_2.fq.gz --max_memory 200G --CPU 30 --no_salmon --full_cleanup --output [path]/Trinity

Nx statistics

TrinityStats.pl [path]/Trinity.Trinity.fasta

ExN50 statistics

align_and_estimate_abundance.pl --transcripts [path]/Trinity.Trinity.fasta --seqType fq --left [path]/Trimming/L_T21F2_N3143_R1_001_val_1.fq.gz --right [path]/Trimming/L_T21F2_N3143_R2_001_val_2.fq.gz --est_method kallisto --output_dir [path]/Trinity_statistics/ExN50/L_21_F2 --prep_reference

align_and_estimate_abundance.pl --transcripts [path]/Trinity.Trinity.fasta --seqType fq --left [path]/Trimming/L_T23F2_N3144_R1_001_val_1.fq.gz --right [path]/Trimming/L_T23F2_N3144_R2_001_val_2.fq.gz --est_method kallisto --output_dir [path]/Trinity_statistics/ExN50/L_23F2 --prep_reference

align_and_estimate_abundance.pl --transcripts [path]/Trinity.Trinity.fasta --seqType fq --left [path]/Trimming/L_T24F2_N3145_R1_001_val_1.fq.gz --right [path]/Trimming/L_T24F2_N3145_R2_001_val_2.fq.gz --est_method kallisto --output_dir [path]/Trinity_statistics/ExN50/L_24F2 --prep_reference

abundance_estimates_to_matrix.pl --est_method kallisto --gene_trans_map none --out_prefix kallisto --name_sample_by_basedir [path]/Trinity_statistics/ExN50/L_21_F2/abundance.tsv [path]/Trinity_statistics/ExN50/L_23F2/abundance.tsv [path]/Trinity_statistics/ExN50/L_24F2/abundance.tsv

echo "Kallisto done"

[path]/trinity-2.14.0.FULL-[path]/contig_ExN50_statistic.pl kallisto.isoform.TMM.EXPR.matrix [path]/Trinity.Trinity.fasta > ExN50.stats
[path]/trinity-2.14.0.FULL-[path]/plot_ExN50_statistic.Rscript ExN50.statsx

Transdecoder

TransDecoder.LongOrfs -t [path]/Trinity.Trinity.fasta
TransDecoder.Predict -t [path]/Trinity.Trinity.fasta

BLAST

makeblastdb -in [path]/Blast/Database/uniprot_sprot.fasta -out [path]/Blast/Database/uniprot_sprot -dbtype prot

blastp -query [path]/Transdecoder/Trinity.Trinity.fasta.transdecoder.pep -db [path]/Blast/Database/uniprot_sprot -max_target_seqs 1 -max_hsps 1 -outfmt '10 delim=; qseqid sseqid pident length mismatch gapopen qstart qend sstart send evalue bitscore sacc qcovs stitle' -evalue 1 -num_threads 20 > [path]/Blast/Lutjanus_Blastp_SwissProt.csv

A.3 R code

```
graphics.off()
windowsFonts(Calibri = windowsFont("Calibri"))

###packages
library(edgeR)
library(tidyr)
library(dplyr)
library(tidyverse)
library(ggthemr)
library(readr)
library(ggplot2)
library(plyr)


#### function to save plots
savePlot<-function(fileName, xDimension, yDimension){
 print(getwd())
 print(fileName)
 dev.print(
  tiff,
  fileName,
  height = 25,
  width = 25,
  units = 'cm',
  type="windows",
  res=600
 )
}
```

```
####### READ METADATA FILE ##############
setwd("WOKING DIR NAME")
folder<-"DIR PATH" ### path to data folder
filename <- "FILENAME"

file <- c(paste(sep="",folder,"/",filename))
metadata <- read.csv(file = file, sep = ";")

names<- metadata$Tube.ID
temp <- metadata$Temp.treat
pH <- metadata$CO2.treat

condition <- paste(temp, pH, sep = "_")


####### READ PROTEIN DATA #########################

filename_p
file_data <- (paste(sep="",folder,"/",filename_p))

data <- as.data.frame(read.csv(file = file_data, sep = ";"))
data <- na.omit(data)


abundances <- as.data.frame(data[,c(19:36)]) ###
Ids <- data$Proteins.Unique.Sequence.ID
rownames(abundances) <- Ids
Annotation <- data[c(5,6)]
colnames(Annotation) <- c("Accession", "Gene")
rownames(Annotation) <- Ids


Data_full <- as.data.frame(cbind(Annotation, abundances))

##### function to generate excel files
```

```
write.table(Data_full, "Data_full.csv", sep=";", col.names=NA)



dataMatrix<-as.data.frame(
 cbind(
  names,
  condition
 )
)


#############################################################################
########## PIECHART OF ANNOTATED PROTEINS SPECIES COMPOSITION #############
#############################################################################


tnames <-  sapply(strsplit(Data_full$Gene, split = "OX="),`[`,1)
prot_names <- sapply(strsplit(tnames, split = "OS="),`[`,1)
species_names <- sapply(strsplit(tnames, split = "OS="),`[`,2)



species_names <- data.frame(species_names = factor(species_names))
species_count <- table(species_names$species_names)
species_count <- sort(species_count, decreasing = TRUE)


top_species <- names(species_count)[1:7]

species_grouped <- ifelse(species_names$species_names %in% top_species, as.character(species_names$species_names), "Others")

species_grouped <- factor(species_grouped)
```

```
species_grouped_count <- table(species_grouped)
species_grouped_df <- as.data.frame(species_grouped_count)

colnames(species_grouped_df) <- c("species_grouped", "n")
species_grouped_df<- species_grouped_df[order(species_grouped_df$n),]
#ggthemr('solarized')
ggthemr("pale")
#swatch()
ggplot(species_grouped_df, aes(x = "", y = n, fill = species_grouped)) +
 geom_bar(stat = "identity", width = 1) +
 coord_polar(theta = "y") +
 geom_text(aes(label = n),
        position = position_stack(vjust = 0.5))+
 theme_minimal() +
 labs(title = "Species Identity of found proteins", fill = "Species")+
 theme_minimal() +
 theme(
  axis.title.x = element_blank(),
  axis.title.y = element_blank(),
  axis.text = element_blank(),
  axis.ticks = element_blank(),
  panel.grid = element_blank(),
  axis.line = element_blank()
 )



################################################################
################################################################
###################### DRP ANALYSIS ############################
################################################################
################################################################


####################BUILD DATA STRUCTURE#####################
write.table(dataMatrix, "dataMatrix.csv", sep=";", col.names=NA)
```

```
LgutData<-DGEList(counts=abundances, group=condition, genes=Ids, remove.zeros = TRUE)



#my Data is already Normalized
#LgutData<-calcNormFactors(LgutData)
LgutDesign<-model.matrix(~0+condition,data=LgutData$samples)
colnames(LgutDesign)<-levels(LgutData$samples$group)

########### FIT GLM MODEL ######################
LgutData<-estimateDisp(LgutData,LgutDesign)
Lgutfit<-glmFit(LgutData,LgutDesign)


############# #SET CONTRASTS BETWEEN CONDITIONS ########################
T28ppm400vsT30ppm400 <-makeContrasts(T30_ppm400-T28_ppm400, levels=LgutDesign)
LgutT28ppm400vsT30ppm400<-glmLRT(Lgutfit,contrast=T28ppm400vsT30ppm400)

topTags(LgutT28ppm400vsT30ppm400)

T28_ppm400vsT28_ppm700<-makeContrasts(T28_ppm700-T28_ppm400, levels=LgutDesign)

LgutT28_ppm400vsT28_ppm700<-glmLRT(Lgutfit,contrast=T28_ppm400vsT28_ppm700)
topTags(LgutT28_ppm400vsT28_ppm700)


T28_ppm700vsT30_ppm700<-makeContrasts(T30_ppm700-T28_ppm700, levels=LgutDesign)
LgutT28_ppm700vsT30_ppm700<-glmLRT(Lgutfit,contrast=T28_ppm700vsT30_ppm700)
topTags(LgutT28_ppm700vsT30_ppm700)

T30_ppm400vsT30_ppm700<-makeContrasts(T30_ppm700-T30_ppm400, levels=LgutDesign)
LgutT30_ppm400vsT30_ppm700<-glmLRT(Lgutfit,contrast=T30_ppm400vsT30_ppm700)
topTags(LgutT30_ppm400vsT30_ppm700)
```

```
T28_ppm400vsT30_ppm700<-makeContrasts(T30_ppm700-T28_ppm400, levels=LgutDesign)
LgutT28_ppm400vsT30_ppm700<-glmLRT(Lgutfit,contrast=T28_ppm400vsT30_ppm700)


topTags(LgutT28_ppm400vsT30_ppm700)



LgutResults<-cbind(Data_full[,c(1,2)],                    LgutT28ppm400vsT30ppm400$table,
LgutT28_ppm400vsT28_ppm700$table,                    LgutT28_ppm700vsT30_ppm700$table,
LgutT30_ppm400vsT30_ppm700$table )


colnames(LgutResults)<-c(
  "ID",
  "Gene",
  "logFC-T28ppm400vsT30ppm400",
  "logCPM-T28ppm400vsT30ppm400",
  "LR-T28ppm400vsT30ppm400",
  "FDRp-T28ppm400vsT30ppm400",

  "logFC-T28_ppm400vsT28_ppm700",
  "logCPM-T28_ppm400vsT28_ppm700",
  "LR-T28_ppm400vsT28_ppm700",
  "FDRp-T28_ppm400vsT28_ppm700",

  "logFC-T28_ppm700vsT30_ppm700",
  "logCPM-T28_ppm700vsT30_ppm700",
  "LR-T28_ppm700vsT30_ppm700",
  "FDRp-T28_ppm700vsT30_ppm700",

  "logFC-T30_ppm400vsT30_ppm700",
  "logCPM-T30_ppm400vsT30_ppm700",
  "LR-T30_ppm400vsT30_ppm700",
  "FDRp-T30_ppm400vsT30_ppm700")
```

```
write.table(LgutResults,"LgutResults.csv",sep=";",col.names=NA)


expressionTable<-decideTests(LgutResults[,grepl("FDRp", colnames(LgutResults))], coefficients=LgutResults[,grepl("logFC", colnames(LgutResults))], lfc=0.2, adjust.method="fdr")


head(expressionTable)
colnames(expressionTable)<-c("T28ppm400vsT30ppm400", "T28_ppm400vsT28_ppm700", "T28_ppm700vsT30_ppm700", "T30_ppm400vsT30_ppm700")


LgutDEG<-cbind(LgutResults,expressionTable)
degTable<-LgutDEG[which(abs(LgutDEG$T28ppm400vsT30ppm400) == 1 | abs(LgutDEG$T28_ppm400vsT28_ppm700) == 1 | abs(LgutDEG$T28_ppm700vsT30_ppm700) | abs(LgutDEG$T28_ppm700vsT30_ppm700) == 1 ),]

DEG <- rownames(degTable)


write.table(degTable,"degTable.csv",sep=";",col.names=NA)

########################################
############## VOLCANO PLOTS ############
########################################


degT28ppm400vsT30ppm400up <- LgutDEG[which(LgutDEG$T28ppm400vsT30ppm400 == 1),]
```

```
degT28ppm400vsT30ppm400down <- LgutDEG[which(LgutDEG$T28ppm400vsT30ppm400 == -1),]

ttup <- rownames(LgutT28ppm400vsT30ppm400$table) %in% rownames(degT28ppm400vsT30ppm400up)
ttdown <- rownames(LgutT28ppm400vsT30ppm400$table) %in% rownames(degT28ppm400vsT30ppm400down)
par(pin = c(3,3))
plot(x = LgutT28ppm400vsT30ppm400$table$logFC, y = -10*log10(LgutT28ppm400vsT30ppm400$table$PValue), main = "Δ Temperature (ambient pCO2)", xlab = "logFC", ylab = "-10logPvalue",cex = 0.5, pch = 19, ylim = c(0,130))
points(LgutT28ppm400vsT30ppm400$table$logFC[ttup], -10*log10(LgutT28ppm400vsT30ppm400$table$PValue[ttup]), col="red",cex = 0.65, pch = 19)
points(LgutT28ppm400vsT30ppm400$table$logFC[ttdown], -10*log10(LgutT28ppm400vsT30ppm400$table$PValue[ttdown]), col="blue",cex = 0.65, pch = 19)

savePlot("Volcano1.tif", 10,10)



pT28_ppm400vsT28_ppm700up <- rownames(LgutT28_ppm400vsT28_ppm700$table) %in% rownames(LgutDEG[which(LgutDEG$T28_ppm400vsT28_ppm700 == 1),])
pT28_ppm400vsT28_ppm700down <- rownames(LgutT28_ppm400vsT28_ppm700$table) %in% rownames(LgutDEG[which(LgutDEG$T28_ppm400vsT28_ppm700 == -1),])



plot(x = LgutT28_ppm400vsT28_ppm700$table$logFC, y = -10*log10(LgutT28_ppm400vsT28_ppm700$table$PValue), main = "Δ pCO2 (ambient Temperature)", xlab = "logFC", ylab = "-10logPvalue",cex = 0.5, pch = 19, ylim = c(0,130))

points(LgutT28_ppm400vsT28_ppm700$table$logFC[pT28_ppm400vsT28_ppm700up], -10*log10(LgutT28_ppm400vsT28_ppm700$table$PValue[pT28_ppm400vsT28_ppm700up]), col="red",cex = 0.65, pch = 19)
```

```
points(LgutT28_ppm400vsT28_ppm700$table$logFC[pT28_ppm400vsT28_ppm700down], -10*log10(LgutT28_ppm400vsT28_ppm700$table$PValue[pT28_ppm400vsT28_ppm700down]), col="blue",cex = 0.65, pch = 19)


savePlot("Volcano2.tif", 10,10)

pT28_ppm700vsT30_ppm700up <- rownames(LgutT28_ppm700vsT30_ppm700$table) %in% rownames(LgutDEG[which(LgutDEG$T28_ppm700vsT30_ppm700 == 1),])
pT28_ppm700vsT30_ppm700down <- rownames(LgutT28_ppm700vsT30_ppm700$table) %in% rownames(LgutDEG[which(LgutDEG$T28_ppm700vsT30_ppm700 == -1),])


plot(x = LgutT28_ppm700vsT30_ppm700$table$logFC, y = -10*log10(LgutT28_ppm700vsT30_ppm700$table$PValue), main = "Δ Temperature (elevated pCO2)", xlab = "logFC", ylab = "-10logPvalue",cex = 0.5, pch = 19, ylim = c(0,130))

points(LgutT28_ppm700vsT30_ppm700$table$logFC[pT28_ppm700vsT30_ppm700up], -10*log10(LgutT28_ppm700vsT30_ppm700$table$PValue[pT28_ppm700vsT30_ppm700up]), col="red",cex = 0.65, pch = 19)
points(LgutT28_ppm700vsT30_ppm700$table$logFC[pT28_ppm700vsT30_ppm700down], -10*log10(LgutT28_ppm700vsT30_ppm700$table$PValue[pT28_ppm700vsT30_ppm700down]), col="blue",cex = 0.65, pch = 19)


savePlot("Volcano3.tif", 10,10)


pT30_ppm400vsT30_ppm700up <- rownames(LgutT30_ppm400vsT30_ppm700$table) %in% rownames(LgutDEG[which(LgutDEG$T30_ppm400vsT30_ppm700 == 1),])
pT30_ppm400vsT30_ppm700down <- rownames(LgutT30_ppm400vsT30_ppm700$table) %in% rownames(LgutDEG[which(LgutDEG$T30_ppm400vsT30_ppm700 == -1),])
```

```
plot(x = LgutT30_ppm400vsT30_ppm700$table$logFC, y = -10*log10(LgutT30_ppm400vsT30_ppm700$table$PValue), main = " Δ pCO2 (elvevated Temperature)", xlab = "logFC", ylab = "-10logPvalue",cex = 0.5, pch = 19, ylim = c(0,130))

points(LgutT30_ppm400vsT30_ppm700$table$logFC[pT30_ppm400vsT30_ppm700up], -10*log10(LgutT30_ppm400vsT30_ppm700$table$PValue[pT30_ppm400vsT30_ppm700up]), col="red",cex = 0.65, pch = 19)
points(LgutT30_ppm400vsT30_ppm700$table$logFC[pT30_ppm400vsT30_ppm700down], -10*log10(LgutT30_ppm400vsT30_ppm700$table$PValue[pT30_ppm400vsT30_ppm700down]), col="blue",cex = 0.65, pch = 19)


savePlot("Volcano4.tif", 10,10)



################
### Bar Plot####
#non-uique-hits#
################



par(mfrow = c(1,1))

con <- c(" Δ Temperature (ambient_pCO2)"," Δ Temperature (ambient_pCO2)",
        " Δ pCO2 (ambient_Temperature)"," Δ pCO2 (ambient_Temperature)",
        " Δ Temperature (elevated_pCO2)"," Δ Temperature (elevated_pCO2)",
        " Δ pCO2 (elevated_Temperature)"," Δ pCO2 (elevated_Temperature)")

regulation <- c("up", "down", "up", "down", "up", "down", "up", "down")
Value <- c(
  length(which(degTable$T28ppm400vsT30ppm400 == 1)),
  length(which(degTable$T28ppm400vsT30ppm400 == -1)),
  length(which(degTable$T28_ppm400vsT28_ppm700 == 1)),
  length(which(degTable$T28_ppm400vsT28_ppm700 == -1)),
```

```
  length(which(degTable$T28_ppm700vsT30_ppm700 == 1)),
  length(which(degTable$T28_ppm700vsT30_ppm700 == -1)),
  length(which(degTable$T30_ppm400vsT30_ppm700 == 1)),
  length(which(degTable$T30_ppm400vsT30_ppm700 == -1))
)
xpos = c(1,1,3,3,2,2,4,4)
bar_data<- data.frame(con,regulation,Value)

bar_data <- bar_data[c(1,2,5,6,3,4,7,8),]
bar_data$con <- factor(bar_data$con, levels = c("Δ Temperature (ambient_pCO2)","Δ Temperature (elevated_pCO2)","Δ pCO2 (ambient_Temperature)", "Δ pCO2 (elevated_Temperature)"))
barplot2 <- ggplot(bar_data, aes(fill=regulation, x =con, y=Value)) +
  geom_bar(position = "dodge", stat = "identity", width = 0.9) +
  xlab("Condition") +
  ylab("Count") +
  scale_fill_manual(values = c("up" = "#C30000", "down" = "#2145FA")) +
  scale_x_discrete(drop = FALSE) +
  theme_minimal()




print(barplot2)
savePlot(fileName = "BAR.tif", xDimension = 10, yDimension = 10)



############################################################
############################################################
########## PROTEIN SPECIFIC ABUNDANCE BAR PLOTS ###############
############################################################
############################################################
```

```
##### takes a Unique sequence ID and cosntructs a Barplot of the abundances of the corresponding Proteins
#####

build_box <- function(ID){

 gene <- Data_full[which(rownames(Data_full) == ID),]

 v1 <- gene[c(3,5,7,9)]
 v2 <- gene[c(11,13,15,17,19)]



 v3 <- gene[c(14,16,18,20)]
 v4 <- gene[c(4,6,8,10,12)]

 v1 <- laply(v1, function(x) x)

 v2 <- laply(v2, function(x) x)
 v3 <- laply(v3, function(x) x)
 v4 <- laply(v4, function(x) x)

 #box.data <- data.frame(name = names(gene)[-c(1,2)], value = gene[-c(1,2)])
 #names(gene)[-c(1,2)]

 data <- data.frame(
  name=c( rep("T28_400pCO2",4), rep("T28_700pCO2",5), rep("T30_400pCO2",4), rep("T30_Co700pCO2",5)),
  value=log2(c( v1, v2, v3,v4))
 )

 data %>%
  ggplot( aes(x=name, y=value, fill=name)) +
  geom_boxplot() +
  #scale_fill_viridis(discrete = TRUE, alpha=0.6) +
```

```
  scale_fill_manual(values = c("#e6194B", "#3cb44b", "#4363d8", "#42d4f4")) +
  geom_jitter(color = "black", size = 0.9) +
  theme(
   legend.position="none",
   text = element_text(family="Serif"),
   axis.title.x = element_text(margin=margin(r=20), size = 18),
   axis.title.y = element_text(margin=margin(r=20), size = 18),
   plot.title = element_text(size=22, hjust = 0.5),
   panel.background = element_rect(fill = "white", color = NA),
   plot.background = element_rect(fill = "white", color = NA),
   panel.grid.major = element_blank(),
   panel.grid.minor = element_blank(),
   axis.line = element_line(color = "black"),
   axis.ticks = element_line(color = "black"),
   axis.text = element_text(color = "black"),
   plot.margin = margin(5, 5, 5, 5)
  ) +
  ggtitle(strsplit(gene$Gene,split = " OS=")[[1]][1]) +
  xlab("Conditions") +
  ylab("log(abundance)")


}




build_box(-9077811903498609664)
savePlot("P1.tif",10,10)
build_box(-8170166660701259776)
savePlot("P2.tif",10,10)
build_box(-3459209202849469952)
savePlot("P3.tif",10,10)
build_box(929444232977523968)
savePlot("P4.tif",10,10)
build_box(-83651836573142896664)
```

```
savePlot("P5.tif",10,10)
build_box(-7244779768486219776)
savePlot("P6.tif",10,10)


##################################
############ HEATMAP ##############
##################################

### all individuals unique accession

un <- unique(degTable$Gene)
index <- c()
for (i in c(1:length(un))){
  ind <- which(degTable$Gene %in% un[i])[1]
  index <- c(index,ind)

}
a <- rownames(degTable[index,])

heat_samples <- Data_full[(which(rownames(Data_full) %in% a)),]


######
heatData <- as.matrix(heat_samples[-c(1,2)])
rownames(heatData) <- sapply(strsplit(heat_samples$Gene," OS="),`[`, 1)

condition_label <- c("T28_400pCO2", "T30_700pCO2", "T28_400pCO2" ,"T30_700pCO2",
"T28_400pCO2",
                     "T30_700pCO2" ,"T28_400pCO2", "T30_700pCO2", "T28_700pCO2",
"T30_700pCO2",
                     "T28_700pCO2", "T30_400pCO2", "T28_700pCO2", "T30_400pCO2",
"T28_700pCO2",
                     "T30_400pCO2", "T28_700pCO2", "T30_400pCO2")
```

```
colnames(heatData) <- condition_label
head(heatData)


heatData <- log10(heatData)
heatData <- t(scale(t(heatData)))
clustFunction <- function(x) hclust(x, method="complete")
distFunction <- function(x) dist(x,method="euclidean")

#clusterint (sidebars)
colCutoff = 8
cCol<-colorRampPalette(brewer.pal(9,"Set1"))
colFit<-clustFunction(distFunction(t(heatData)))
cClusters<-cutree(colFit,h=colCutoff)
cHeight<-length(unique(as.vector(cClusters)));
colHeight = cCol(cHeight)
rowCutoff = 5
cRow<-colorRampPalette(brewer.pal(9,"Set1"))
rowFit<-clustFunction(distFunction(heatData))
rClusters<-cutree(rowFit,h=rowCutoff)
rHeight<-length(unique(as.vector(rClusters)));
rowHeight = cRow(rHeight)

#Heatmap
heatColour<-colorRampPalette(c("#3832FA", "#F3EDED", "#D22B2B"))(n = 500)
fileName<-"Heatmap2.tif"
xDimension<-10
yDimension<-10
#windows(width = 10, height = 20)
#windows(width = 10, height = 10)
grDevices::windows(xDimension, yDimension)
par(family="Calibri")
heatmap.2(
 heatData,
 hclust=clustFunction,
```

```
  distfun=distFunction,
  ColSideColors=colHeight[cClusters],
  RowSideColors=rowHeight[rClusters],
  density.info="none",
  col=heatColour,
  colsep=0:ncol(heatData),
  rowsep=0:nrow(heatData),
  sepcolor="black",
  sepwidth=c(0.001,0.001),
  trace="none",
  #scale="row",
  srtCol = 45,
  cexRow=0.8,
  cexCol=1,
  lhei = c(1,10),
  lwid = c(1.5,15),
  keysize=1,
  margins = c(20,30),

)
savePlot(fileName, xDimension, yDimension)



grDevices::windows(xDimension, yDimension)
heatmap.2(
  heatData,
  hclust = clustFunction,
  distfun = distFunction,
  ColSideColors = colHeight[cClusters],
  RowSideColors = rowHeight[rClusters],
  density.info = "none",
  col = heatColour,
  colsep = 0:ncol(heatData),
  rowsep = 0:nrow(heatData),
  sepcolor = "black",
```

```
  sepwidth = c(0.001, 0.001),
  trace = "none",
  # scale = "row",
  srtCol = 45,
  key = TRUE,
  key.title = NA,
  key.xlab = "Expression Level",
  cexRow = 0.8,
  cexCol = 1,
  lhei = c(2,10),
  lwid = c(5, 15),
  keysize = 1.5,
  margins = c(20, 30)


)
savePlot("__just_the_legend.tif",10,10)

####################
####PCA & PSLDA#####
####################


Dex <- Data_full[rownames(Data_full) %in% rownames(degTable),]

ext <- Dex[-c(1,2)]
ext <- log10(ext)
ext <- scale(ext)


new_names <- strsplit2(Dex$Gene, split = " OS")[,1]
new_names<- make.names(new_names, unique = TRUE)
Fac2 <- factor(condition)
colnames(ext) <- names
rownames(ext) <- new_names
ext<-as.matrix(ext)
```

```
pca.data <- PCA(t(ext),scale.unit = FALSE)
nb.cols <- 4
mycolors <- c("#e6194B","#3cb44b","#4363d8","#42d4f4")
pca <- fviz_pca_ind(pca.data,col.ind = Fac2,
                    palette = mycolors, addEllipses = T)

PC1 <- round(pca.data$eig[1,2],digits = 2)
PC2 <- round(pca.data$eig[2,2], digits = 2)

ggpar(pca,
    title = "PCA",
    xlab = paste("PC1 (",PC1, "%)",sep = ""), ylab = paste("PC2 (",PC2, "%)",sep = ""),
    legend.title = "condition", legend.position = "top",
    ggtheme = theme_minimal()
)




plsda <- opls(x = t(ext),y = Fac2,  predI = NA)




################################################################################
#####
###############          Nicer          PLS-          DA          plot
#################################################
################################################################################
#####

pls_data <- data.frame(plsda@scoreMN)
```

```
pls_data$condition <- condition
PC1 <- round(plsda@modelDF$R2X[1], digits = 4) * 100
PC2 <- round(plsda@modelDF$R2X[2], digits = 4) * 100
PC1



ggplot(pls_data, aes(x = p1, y = p2, color = condition, group = condition)) +
  geom_point(size = 2.4) +
  xlab(paste("Component 1 (",PC1, "%)", sep = "")) +
  ylab(paste("Component 2 (",PC2, "%)", sep = "")) +
  #geom_text(label = rownames(pls_data), nudge_x = 0.5, nudge_y = 0.5, check_overlap = T) +
  scale_color_manual(values = c("#e6194B","#3cb44b","#4363d8","#42d4f4")) +
  theme(panel.background = element_rect(fill = "white"),
        axis.title.x = element_text(size = 22),
        axis.title.y = element_text(size = 22),

        legend.text = element_text(size = 22),
        legend.title = element_text(size = 24)) +
  stat_ellipse(geom = "polygon", level = 0.90, alpha = 0.25, aes(fill = condition)) +
  scale_fill_manual(values = c("#e6194B","#3cb44b","#4363d8","#42d4f4"))

savePlot("plsda.tif", xDimension = 10,yDimension = 20)

###################################
###PLSDA LOADINGS AND VIP PLOTS####
###################################

VIP <- plsda@vipVn
VIP <- VIP[order(VIP, decreasing = TRUE)]
VIP <- data.frame(VIP)
VIP$Protein <- rownames(VIP)



ggplot(VIP, aes(x = reorder(Protein, -VIP, decreasing = T), y = VIP)) +
  geom_bar(position = "stack", stat = "identity", width = 0.6, fill = "steelblue") +
```

```
  coord_flip() +
  labs(x = "", y = "", title = "VIP scores") +
  theme(panel.border = element_rect(colour = "black", fill = NA, size = 1),
        plot.title = element_text(hjust = 0.5, size = 22, face = "bold"))
savePlot("VIPSCORES.tif", 10,7)
```

```
pls_loading <- data.frame(plsda@loadingMN)
pls_loading$Protein <- rownames(pls_loading)
ggplot(pls_loading, aes(x = reorder(Protein, abs(p1), decreasing = F), y = p1, fill = ifelse(p1 >
0, "red", "blue"))) +
  geom_bar(position = "stack", stat = "identity", width = 0.6) +
  coord_flip() +
  labs(x = "", y = " ", title = "Contribution to component 1") +
  theme(panel.border = element_rect(colour = "black", fill = NA, size = 1),
        plot.title = element_text(hjust = 0.5, size = 22, face = "bold")) +
  scale_y_continuous(position = "right") +
  scale_fill_identity()

savePlot("1_loadings.tif",10, 7)
```

```
ggplot(pls_loading, aes(x = reorder(Protein, abs(p2), decreasing = F), y = p2, fill = ifelse(p2 >
0, "red", "blue"))) +
  geom_bar(position = "stack", stat = "identity", width = 0.6) +
  coord_flip() +
  labs(x = "", y = "  ", title = "Contribution to component 2") +
  theme(panel.border = element_rect(colour = "black", fill = NA, size = 1),
        plot.title = element_text(hjust = 0.5, size = 22, face = "bold")) +
  scale_fill_identity()
savePlot("2_loadings.tif", 10,7)
```

UNIVERSIDADE NOVA
DE LISBOA

2024

Multi- Omic approaches to unravel phenotypic responses of the tropical fish Lutjanus guttatus to ocean warming and acidification scenarios

David Jeremy Lohmann

MASTER IN **COMPUTATIONAL BIOLOGY & BIOINFORMATICS**

SPECIALIZATION Multi-Omics for Life and Health Sciences