UNIVERSITY OF TURKU

Ecotype-specific transcriptomic differences in the telencephalon of Alaskan threespine stickleback, *Gasterosteus aculeatus*

Sanna Pausio

Ecology and Evolutionary Biology
Master's thesis
Credits: 40 ECTS

Supervisors:
Erica Leder
Heidi Viitaniemi
Irma Saloniemi

14.4.2022
Turku

Master's thesis

Subject: Ecology and Evolutionary Biology
Author: Sanna Pausio
Title: Ecotype-specific transcriptomic differences in the telencephalon of Alaskan threespine stickleback, *Gasterosteus aculeatus*
Supervisors: Erica Leder, Heidi Viitaniemi, Irma Saloniemi
Number of pages: 60 pages, 5 appendices
Date: 14.4.2022

Brain activity reacts to environmental complexity within an individual's reaction norm. Brain activity is highly energy demanding and a target for energetic trade-offs. Hence, cognition is an honest but complex signal of selection landscape. 'Clever foraging hypothesis' predicts increased fitness value for cognition in patchy, moderately stable environments.

In this thesis, I have studied telencephalon transcriptome of threespine stickleback, *Gasterosteus aculeatus* Linnaeus, 1758, females from four lakes with two levels of environmental complexity in subarctic Alaska. The threespine stickleback populations in these lakes represent benthic and limnetic ecotypes. Populations of these ecotypes have repeatedly and independently been established in North America after the last glaciation from the ancestral marine form. Despite their young evolutionary age, the ecotypes are phenotypically different. Because of the patchier nature of foraging in benthic environment, benthic populations should benefit more from investment in spatial cognition compared to limnetic populations. Therefore, I hypothesized increased expression of cognition-related genes in the telencephalons of benthic individuals.

Contrary to the expectations, more genes involved in the regulation of neurogenesis and neuron projection were upregulated in the limnetic ecotype. The most significant upregulated biological processes in the benthic ecotype were lipid transport and innate immunity. This suggests greater investment in cognition in the limnetic ecotype and immune defense in the benthic ecotype. These results highlight the impact of immunology instead of spatial complexity as the main factor in the telencephalon gene expression in the study populations. More research is needed to discover the environmental and pathogen community differences among the lake types and the impact of genetic differentiation on the reaction norm of immunity expression among the threespine stickleback ecotypes.

Key words: threespine stickleback, Gasterosteus aculeatus, transcriptome, habitat complexity, telencephalon, microarray

Pro gradu tutkielma

Oppiaine: Ekologia ja evoluutiobiologia
Tekijä: Sanna Pausio
Otsikko: Ecotype-specific transcriptomic differences in the telencephalon of Alaskan threespine stickleback, *Gasterosteus aculeatus*
Ohjaajat: Erica Leder, Heidi Viitaniemi, Irma Saloniemi
Sivumäärä: 60 sivua, 5 liitettä
Päivämäärä: 14.4.2022

Aivojen ja hermoston toiminta mukautuu nopeasti ympäristön vaatimuksiin yksilön reaktionormin asettamissa rajoissa. Niiden toiminta kuluttaa runsaasti energiaa. Jotta energian allokointi aivotoimintaan olisi kannattavaa, sen on tuotettava yksilölle valintahyötyä. Siksi älykkyyttä ja aivotoimintaa voidaan käyttää mittareina ympäristön valintapaineista. ”Clever foraging” -hypoteesi ennustaa resurssien allokoinnin aivotoimintaan ja oppimiseen hyödyttävän yksilöä eniten kohtuullisen stabiileissa ja laikuttaisissa elinympäristöissä.

Tässä pro gradu -työssä olen verrannut kolmipiikkinaaraiden *Gasterosteus aculeatus* Linne, 1758, aivojen nisäkkäiden hippokampusta vastaavan alueen, pääteaivojen, transkriptomeja neljässä subarktisen Alaskan järvipopulaatiossa. Tutkimuspopulaatiot edustavat benttistä (pohja) ja limneettistä (ulappa) ekotyyppiä ja niiden habitaatit kahta ympäristön laikuttaisuuden tasoa. Kolmipiikki on levinnyt Pohjois-Amerikan järviin viimeisimmän jäätiköitymisen jälkeen useiden erillisten leviämistapahtumien seurauksena. Populaatiot ovat toistuvasti ja toisistaan riippumattomasti kehittyneet alkuperäisestä mereisestä muodostaan kohti benttistä ja limneettistä ekotyyppiä. Huolimatta nuoresta evolutiivisesta iästään ekotyypit ovat fenotyypiltään erilaisia. Koska benttisissä elinympäristöissä resurssit ovat sijoittuneet limneettisiä ympäristöjä laikuttaisemmin, on todennäköistä, että benttisiä resursseja hyödyntävät kolmipiikit hyötyvät enemmän informaatiosta ja sen käsittelystä elinympäristössään. Tämän takia oletin tiedon käsittelyyn liittyvien geenien transkription olevan runsaampaa benttiseen ekotyyppiin kuuluvissa populaatioissa.

Vastoin oletusta hermosolujen kasvun ja synnyn säätelyyn liittyvien geeniryhmien transkriptio oli runsaampaa limneettisissä populaatioissa. Benttisissä populaatioissa parhaiten edustettuina olivat rasvan kuljetusta ja synnynnäisen immuniteetin ilmenemistä säätelevät geeniryhmät. Vaikuttaakin siltä, että limneettisten populaatioiden pääteaivojen geenieksprессio painottuu tiedon käsittelyyn ja benttisten populaatioiden synnynnäisen immuniteetin ilmenemiseen. Tuloksissa painottuu ympäristön laikuttaisuuden sijaan immuunipuolustuksen merkitys tutkimuspopulaatioiden geenieksprессiossa. Lisää tutkimusta tarvitaan benttisten ja limneettisten elinympäristöjen ja niiden patogeeniyhteisöjen eroista sekä siitä, onko tutkimusekotyyppien geneettinen erilaistuminen vaikuttanut niiden synnynnäisen immuniteetin ilmenemisen reaktionormeihin.

Avainsanat: kolmipiikki, Gasterosteus aculeatus, transkriptomi, ympäristön laikuttaisuus, pääteaivot, microarray

Table of contents

1 Introduction

1.1 Adaptation to environmental complexity

1.1.1 The significance of adaptation

Adaptation to fluctuating environmental conditions in space and time is essential for survival of both individuals and populations. The ability to adapt to altering conditions is also vital for dispersal (e.g., Sol et al., 2008) and a stepping stone for ecological speciation (Rundle, 2000; Savolainen et al., 2013). Understanding the processes, mechanisms, and prerequisites for adaptation is needed to assess the ecological and evolutionary consequences, that even subtle, small-scale changes in abiotic or biotic factors may induce (Urban et al., 2020).

Studying molecular level responses to environmental factors aims to identify the genetic basis and molecular interactions regulating the emergence of a phenotypic outcome. Together, the three levels, genes, interactions among gene products, and phenotype as the interface between individual and environment, comprise a complex and reciprocal network. Insight into the function of this network and its interaction with ecological and non-ecological factors that drive systemic and population-level change is essential; not only for ecological and evolutionary research but also for more applied fields such as conservation biology or climate change research (Savolainen et al., 2013).

1.1.2 Cognition as adaptation to complexity

Cognition is based on information and the ability to process it. From organismal perspective, it is a combination of three interacting aspects: perception, learning and memory (Arechavala-Lopez et al., 2020). It is a highly plastic (multiple phenotypes may arise from a single genotype) trait, and its physiology is regulated by several delicate mechanisms, as discussed later in this introduction.

While cognition itself is beneficial for fitness, the expenses coming from information gathering and processing may exceed the benefits (Dall et al., 2005). To fitness benefits outweigh the expenses, information has to be reliable and valuable enough to change an individual's behavior in a beneficial manner (Dall et al., 2005). This threshold depends on organismal characteristics (Leavell and Bernal, 2019) together with environmental predictability, patchiness, and productivity (Olsson and Brown, 2010). Patchiness can also be described as complexity, which consists of the inter- and intraspecies interactions as well as

spatial and temporal patchiness in the physical environment. Social environment, group size and parental care, (Gonzalez-Voyer et al., 2009; Kotrschal et al., 2012; Samuk et al., 2014; Tsuboi et al., 2015; Ashton et al., 2018), as well as ecology, above all diet, (Gonzalez-Voyer et al., 2009; Henke-von der Malsburg et al., 2020) has been shown to explain cognitional differences among populations. Often the aspects of environmental complexity act together (Pollen, 2007).

According to 'statistical decision theory' (Guitton et al., 1955; McNamara and Houston, 1980) and 'optimal foraging theory' (Emlen, 1966; MacArthur and Pianka, 1966; Charnov, 1976), the benefits from information-based strategies for an individual are highest in moderately predictable and complex environments. Hence, these environments should promote the evolution of cognition. 'Clever foraging hypothesis' (Parker and Gibson, 1977; Park and Bell, 2010), the theoretical basis of this thesis, is based on the presumptions of those theories. According to 'clever foraging hypothesis', cognitive evolution is mainly driven by the relative benefits of complex food searching strategy for an individual in an environment. Hence, it predicts increased information value and cognition in active predators with patchy food distribution (Park and Bell, 2010). This hypothesis has been tested in both intra- and interspecies cognition studies in e.g., threespine stickleback *Gasterosteus aculeatus* (Park and Bell, 2010; Ahmed et al., 2017), arctic char *Salvelinus alpinus* (Tamayo et al., 2020), sharks (Yopak et al., 2007) and carnivores (Gittleman, 1986).

Finding a robust correlation between cognition and fitness is complicated by the complexity of environmental relationships and their temporal fluctuation (Rowe and Healy, 2014). Moreover, biases arising from animal personality, life history, and the complexity of the concept of cognition may hinder distinguishing cognitional differences (Rowe and Healy, 2014). Not only the ability to learn as such, but also the cognitional strategy, e.g., trade-off between initial learning efficiency and learning flexibility (Bebus et al., 2016), may have different fitness value in different environments (Croston et al., 2017).

1.2 Physiological and molecular view to information processing and memory formation

1.2.1 The neural network for environmental information processing

Sensory organs are the first step in environmental information processing. Their neuron count and surface area is an indication of the type of cues an animal relies on in its environment

(Kotrschal et al., 1998; Keagy et al., 2018). Signals received by sensory organs are processed and the information is stored in specific sets of neural cells, neural circuits, that wire together to process specific information (Holtmaat and Svoboda, 2009; Chen et al., 2020).

In amniotes, the most important brain regions for processing environmental information are the hippocampus and the cortex (Squire et al., 2004; Frankland and Bontempi, 2005). In addition to spatial mapping, amniote hippocampus is associated with several other processes, such as emotions and stress-related behavior (Park et al., 2016; Herold et al., 2019).

The hippocampus is a crucial functional region in explicit memory that requires conscious recollection of facts or experiences. Hippocampus serves as initial memory storage for newly acquired memories (Tonegawa et al., 2018). According to well established systems consolidation theory, memories are gradually transferred from hippocampus to cortical network for final memory stabilization (Frankland and Bontempi, 2005). Still, hippocampus continues to contribute in remote memory management, recollection and cognitional flexibility (Abraham et al., 2002; Frankland and Bontempi, 2005; Herold et al., 2019), although hippocampus involvement in memory management after complete maturation of cortical storage is still debated (Tonegawa et al., 2018).

In teleost fish, brain organization differs from that of amniotes because of different developmental process during ontogeny (Salas et al., 2003). Teleost telencephalon dorsolateral region is generally considered to be functionally equivalent to amniote hippocampus, while the dorsomedial region is considered to be responsible for amniote amygdala-like functions, such as fear and social behavior (Demski, 2013). Teleost telencephalon has been demonstrated to mediate a wide variety of behaviors, e.g., spatial and conditional learning, aggressivity, processing of sensory information, and even reproduction (Demski, 2013).

1.2.2 Mechanisms for memory formation and neural plasticity

Cognition as a process consists of four phases: memory acquisition, consolidation, retrieval, and reconsolidation. These phases are enabled by neural plasticity: the ability of neural network to create novel connections and regulate the strength of existing ones to build and remodel memory engrams (sets of neurons that contribute to consolidation of a memory). Neural plasticity involves both functional (Oh and Smith, 2019) and structural (Holtmaat and

Svoboda, 2009; Zupanc and Sîrbulescu, 2011; Anggono and Huganir, 2012; Kol and Goshen, 2021) changes in synapses and neural networks.

Neural plasticity is mainly achieved by four mechanisms: (1) differentiation and co-operation of neurons into inhibitory neurons, whose facilitation decreases the possibility of action potential in postsynaptic neuron; and excitatory neurons, whose facilitation induces action potential in postsynaptic neuron (Oh and Smith, 2019), (2) plasticity in the number and location of synapses, which is regulated by the number and location of dendritic spines and axonal boutons as well as clustering of neurotransmitter receptors (Holtmaat and Svoboda, 2009), (3) change in synaptic efficacy (synaptic strengthening (potentiation) and weakening (depression)) defined mainly by the status and number of neurotransmitter receptors but also neurotransmitter release (Malenka and Bear, 2004), and (4) proliferation and integration of new neural cells in neural circuits (von Krogh et al., 2010; Steadman et al., 2020).

In addition to synaptic plasticity, memory formation is regulated by glial cells: oligodendrocytes, astrocytes, and microglia. Oligodendrocytes participate in the facilitation of memory consolidation (Steadman et al., 2020) and memory retrieval (Pan et al., 2020). Astrocytes have been shown to be capable of facilitating memory acquisition (Adamsky et al., 2018), long-term memory consolidation and memory allocation by secreting synaptic connection regulating factors (Kol and Goshen, 2021; Wang et al., 2021b), but see Agulhon et al. (2010) for opposing results. Microglia have immunological functions in brain, but they are also capable of promoting synaptogenesis, activity-dependent elimination of synapses and direct regulation of synaptic plasticity (Wang et al., 2021a).

Neural stimulation, both learning and memory retrieval, lower the threshold for synaptic wiring (increase intrinsic excitability) in stimulated synapses (Parsons, 2018; Chen et al., 2020). Learning also stimulates neural cell proliferation, and promotes the survival of relatively mature and the death of more immature neurons (Dupret et al., 2007). Ultimately, signaling intensity (Medina et al., 2019), pattern (Lee et al., 2017), and duration (Lee et al., 2017; Tyssowski et al., 2018) as well as previous experiences (Parsons, 2018) and systemic homeostasis (Johansen et al., 2012) determine the trajectory and the result of information processing.

1.2.3 Molecular regulation of information processing

Changes is synaptic efficacy initiate by voltage- and neurotransmitter-mediated opening of calcium-permeable AMPA-receptors (AMPARs) in glutamatergic excitatory neurons and chlorine-permeable GABA-receptors (GABARs) in GABAergic inhibitory neurons. The induction of long-term potentiation (LTP) usually further requires the opening of another glutamatergic calcium channel, NMDAR, although other forms of LTP exist (Frey and Morris, 1998). Besides calcium channel status, synaptic strength is regulated by their number: the LTP-related active transport of intracellular AMPARs to synaptic density and their long-term depression (LTD)-related endocytosis (Malenka and Bear, 2004). Early phase of LTP, as well as some forms of LTD can be maintained by local translation in dendrites, but later phases of LTP and memory consolidation requires somatic transcription and translation (Costa-Mattioli et al., 2009), which is accompanied by increased genome-wide chromatin accessibility and nuclear reorganization (Tyssowski et al., 2018; Marco et al., 2020).

Change in synaptic strength is mediated to gene expression by signaling pathways, which can be activated by intracellular calcium content change, but also regulated by a variety of other signaling molecules and regulatory pathways (Zheng et al., 2009; Miningou and Blackwell, 2020). The first wave of molecular response to a stimulus can be triggered by synaptic activity dependent activation of Ras-mitogen-associated protein kinase (MAPK or ERK), calcium/calmodulin-dependent protein kinases (CaMKs), and calcineurin-mediated signaling pathways (Yap and Greenberg, 2018). These pathways interact and co-act with each other and other signaling pathways in a neural stimulus-specific manner (Zheng et al., 2009).

The connectivity and reciprocity of signaling pathways enables the simultaneous sensitivity of transcriptional response to both internal (homeostasis) and external signals. MAPK signaling can be activated by several pathways (Miningou and Blackwell, 2020), including phosphatidylinositol 3 kinase/protein kinase B (pI3k/AKT) pathway, which in turn can be activated by variety of growth factors (Mendoza et al., 2011). Both MAPK and pI3k/AKT pathways participate in mammalian target of rapamycin (mTOR) pathway regulation (Mendoza et al., 2011), which in neural cells is involved with various brain functions from cell proliferation to synaptic plasticity and regulation of complex behaviors (Lipton and Sahin, 2014). It, as well as AKT signaling itself, has plethora of downstream targets (Brazil et al., 2004; Miningou and Blackwell, 2020). MAPK and AKT signaling have been demonstrated to both counter- and co-act, even compensate, each other (Mendoza et al.,

2011). The three pathways, MAPK, pI3k/AKT and mTOR, are central for memory formation (Chen et al., 2005; Costa-Mattioli et al., 2009), but are accompanied with vast number of other crucial pathways, e.g., p38 MAPK pathway, which is involved with the regulation of inflammatory responses, but also synaptic plasticity (Falcicchia et al., 2020).

The complexity of the neural regulatory mechanisms enables the differentiated response to environmental stimuli among cell types and brain regions (Malenka and Bear, 2004; Yap and Greenberg, 2018). Whilst the majority of the evidence for the mechanisms presented here comes from mammalian hippocampus, information processing is in many respects conserved across taxa (Salas et al., 2003; Costa-Mattioli et al., 2009).

1.3 Physical approaches for cognition assessment

1.3.1 Brain size and morphology

Brain allometry and morphometry are commonly used as physical measures of cognition. Because of its energy demands, the size of a neural tissue can be considered to signal the value of the information it processes in an environment-individual interaction (Dunbar, 1998; Niven and Laughlin, 2008; Tsuboi et al., 2015). Hence, both total brain size and the size of functional brain regions, but also brain region morphometry (Park and Bell, 2010) has been used to study the relationship between cognition and environment.

Positive correlation between ecological complexity and brain size has been demonstrated repeatedly in teleost fish (Pollen, 2007; Gonda et al., 2013; White and Brown, 2015; Keagy et al., 2018; Tamayo et al., 2020). This correlation has also been found in teleost telencephalon (Pollen, 2007; Gonda et al., 2009; White and Brown, 2015; Tamayo et al., 2020), the target tissue of this thesis. Both total brain (Gonzalez-Voyer et al., 2009; Kotrschal et al., 2012; Samuk et al., 2014; Tsuboi et al., 2015) and telencephalon (Pollen, 2007) sizes have also been shown to correlate with social complexity.

Although the relationship with environmental complexity and brain size and morphometry is well established, some studies (Ahmed et al., 2017; Keagy et al., 2018; Axelrod et al., 2021) have failed to demonstrate this, which highlights the complexity of cognitive evolution. Park & Bell (2010) have discovered negative correlation between habitat complexity and telencephalon size in threespine stickleback, although the telencephalon shape indicated

positive correlation between ecological complexity and telencephalon dorsolateral region size in the study.

The use of brain size as a proxy for cognition is supported by its experimentally established correlation with cognitive performance (Odling-Smee et al., 2008; Buechel et al., 2018; Keagy et al., 2018) and behavioral plasticity (Herczeg et al., 2019). Brain size also shows plastic response to environmental enrichment (Park et al., 2012; Fong et al., 2019), which has been demonstrated to have positive effects on learning ability (Salvanes et al., 2013; Carbia and Brown, 2019). However, morphometric measurement of brain functional regions instead of brain size may help to better distinguish the selective forces underlining cognitive evolution, as these regions wire together for specific cognitional tasks (Introduction 1.2.1).

1.3.2 Molecular measures for cognition

Molecular approach to cognition study, i.e., measurement of protein or RNA expression instead of morphometry has the advantage of enabling the research of ongoing physiological processes. Especially RNA expression approach, being less technically challenging than proteomics (Hill et al., 2020), has been widely applied. Several studies have measured the transcript abundance of individual genes known to contribute to cognition (e.g., Salvanes et al., 2013). However, studying the entire transcriptome gives a more comprehensive view of processes that generate phenotypic expression, by allowing functional grouping of expressed genes.

Transcriptomes can be studied by sequencing (RNA-seq) or by hybridizing mRNA to pre-selected gene oligonucleotides or open reading frames that are fixed on a slide or chip (i.e., microarray) (Malone and Oliver, 2011). RNA-seq studies have outnumbered microarrays in recent years (Mantione et al., 2014). Despite this, with appropriately selected, well-annotated candidate probes microarray studies can reveal ecologically and evolutionary meaningful molecular cascades by assaying mRNA abundances (e.g., Stanford et al., 2020).

Several studies have demonstrated the connection between environmental complexity and cognition-related gene expression. For example genes related to neuronal signaling and neuron growth (Rampon et al., 2000), growth and learning (Evans et al., 2015), and synaptic plasticity (Podgorniak et al., 2015; Liu et al., 2020) have been found to be upregulated after environmental enrichment or correlated with environmental complexity. An intraspecies study in black-capped chickadees *Poecile atricapillus* has discovered positive correlation between

temporally increased environmental patchiness, stemming from harsher climate, and heritable upregulation of learning and neurogenesis related genes, previously shown to correlate with increased level of spatial learning (Pravosudov et al., 2013).

1.4 Threespine stickleback as model species

Threespine stickleback is a small teleost fish, which is distributed around the northern hemisphere in the margins of the Atlantic and Pacific Oceans. It occupies marine environments and a wide range of freshwater habitats, forming marine, anadromous, and resident freshwater populations ranging from different fluvial environments to a variety of lentic habitats from ponds to large lakes (Bell and Foster, 1994).

Ecological factors, such as salinity, temperature, the level of spatial complexity and species interactions vary among different threespine stickleback habitats. Threespine stickleback has several, morphologically distinct, habitat-specific ecotypes that have evolved repeatedly from the ancestral marine form in response for environmental factors (Bell and Foster, 1994). The ability to quickly adapt to novel environments has made it a popular model for exploring genomic features that enable the rapid adaptive radiation (Hohenlohe et al., 2010; Jones et al., 2012; Morris et al., 2014; Härer et al., 2021). The threespine stickleback genome was sequenced for the first time in 2006 (Stickleback assembly and gene annotation, 2010).

The benthic-limnetic division is one of the best studied forms of adaptive radiation in threespine stickleback. Benthic (littoral) and limnetic (pelagic) lakes are structurally different (Foster et al., 2008; Park et al., 2012). The littoral zone of both benthic and limnetic lakes is complex, with landforms and rooted vegetation offering shelter from predation, contrary to the structurally monotonous, open pelagic area. Shallow, eutrophic benthic lakes have more littoral area than deep, oligotrophic limnetic lakes. In limnetic habitats, threespine sticklebacks feed mostly on open water, while in benthic habitats their diet consists of a mixture of benthic prey and plankton, although neither ecotype cannot be considered a specialist. Both ecotypes breed on benthos, and especially in allopatry, limnetic fish use littoral zone also for foraging (Park and Bell, 2010; Bolnick and Ballare, 2020).

Benthic and limnetic threespine sticklebacks are morphologically different. The benthic ecotype has a deeper body, reduced defensive structures, longer upper jaw, and fewer and shorter gill rakers than limnetic morph (Bell and Foster, 1994; McPhail, 1994). They also differ in pigmentation (Gygax et al., 2018). The morphological differences are more

prominent in sympatric populations (Vamosi and Schluter, 2004). There are also behavioral differences in courtship behavior (Shaw et al., 2007), although in both ecotypes, males show territoriality and paternal care (Bell and Foster, 1994). Wild-caught limnetic fish show increased sociality compared to sympatric benthic ecotypes (Kozak and Boughman, 2008).

Both limnetic and benthic ecotype have similar spatial learning strategies, but benthic individuals learn the tasks quicker and with fewer mistakes than limnetics (Odling-Smee et al., 2008; Park, 2013). Spatial performance has been discovered to have a heritable component (Martinez et al., 2016). Performance in reversal learning task (i.e., cognitional flexibility) has not been studied at the population level. Instead, at the individual level, personality type at the reactive/proactive axis has been shown to affect reversal learning performance (Bensky and Bell, 2020).

Benthic threespine sticklebacks have been reported to have bigger brains than limnetics (Keagy et al., 2018). While the brain size difference is congruent with the presumptions of 'clever foraging hypothesis' in this ecotype pair, brain size fluctuation has been found to be inconsistent with environmental complexity gradient in another ecotype pair that inhabits environments at two complexity levels, namely lake-stream ecotype pair (Ahmed et al., 2017).

Brain, including telencephalon, morphology is highly plastic in adult fish (Park et al., 2012; Buechel et al., 2019). The overall brain, but also telencephalon, size is bigger and more plastic in males than in females (Buechel et al., 2019). Over the breeding cycle, total brain and telencephalon volume grow in both sexes (Buechel et al., 2019). In Icelandic threespine stickleback populations, males, but not females, from structurally complex lava habitats has been reported to have bigger brain than those from more monotonous mud habitats (Kotrschal et al., 2012).

Noreikiene et al. (2015) have discovered a small heritable component in telencephalon morphology in a full- and half-sibling experiment on a stickleback population from the Baltic Sea. However, in Alaskan lake populations, the morphological distinction between the telencephalons of benthic and limnetic sticklebacks has been demonstrated to disappear in the next generation reared in common garden (Park et al., 2012).

1.5 Study populations

The samples for this thesis were collected from four Alaskan lakes in the Matanuska - Susitna and Kenai regions at the Gulf of Alaska (Figure 1, Table 1). The lakes have formed after the withdrawal of the Cordilleran ice sheet in the last glaciation ca 10 000 – 12 000 years ago (Bell and Foster, 1994), confirming the young evolutionary age of the study populations. All the study lakes locate in distinct drainages, suggesting the populations being established from distinct founder events. This is supported by genetic evidence from other freshwater populations in the area (Hohenlohe et al., 2010). The populations are from the same geographical area, which indicates phenotypic adaptive radiation to be based on shared ancestral genetic variation (Fang et al., 2020).

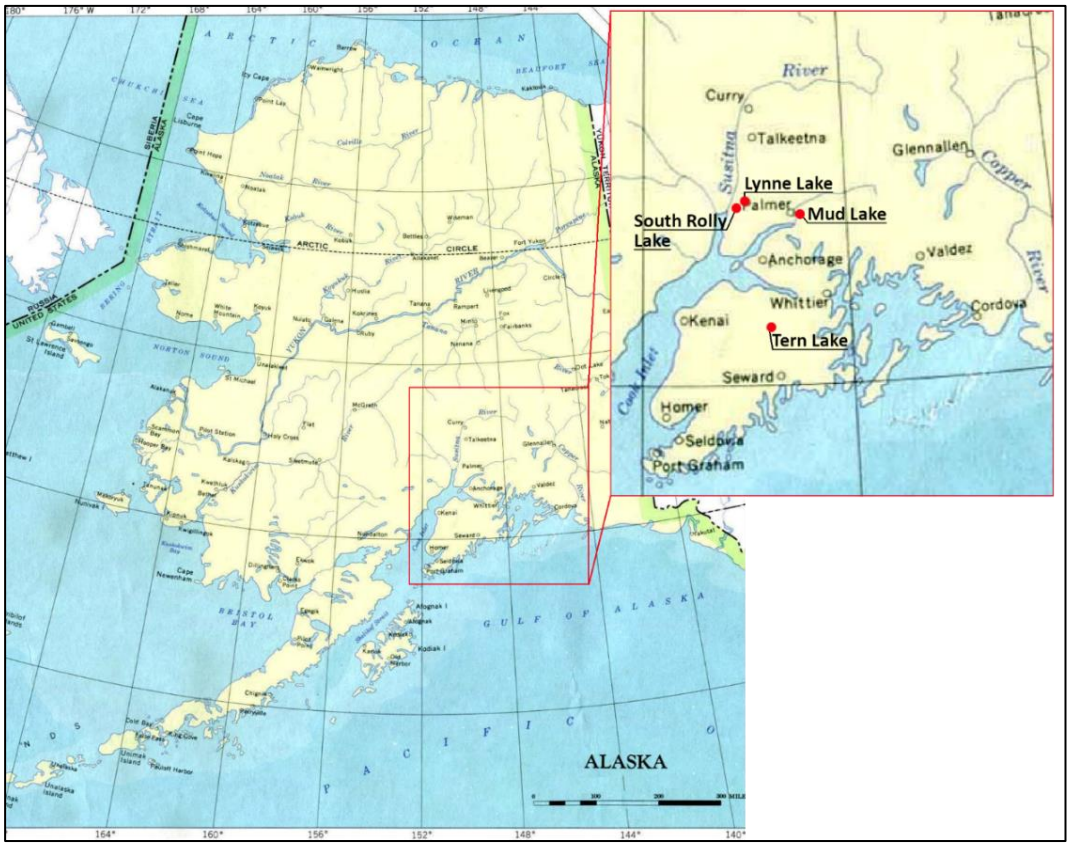


Figure 1. Study lake locations.

Mud Lake (Mud) and Tern Lake (Tern) are benthic lakes, shallow with high relative littoral area. Lynne Lake (Lynne) and South Rolly Lake (SRolley) are deeper and have more pelagic area (Park and Bell, 2010). Previously, the fish from these lakes has been morphologically determined as benthic and limnetic, respectively (Table 1) (Park and Bell, 2010). Also, the telencephalon shape of the limnetic and benthic fish has been shown to correspond to both

fish morphology and lake topography. Limnetic fish in the study lakes have laterally concave telencephalons in comparison to benthic fish convex ones, with fish from Tern and South Rolly lakes having the most extreme telencephalon shapes (Figure 2) (Park and Bell, 2010). This indicates smaller dorsolateral telencephalic regions in limnetic fish (Park and Bell, 2010). Cognitive performance in a spatial task has been shown to vary inconsistently in relation to ecotype in these populations, with the Mud population having the highest performance, followed by Lynne, Tern, and South Rolly populations (Park, 2013).

Table 1. Properties of the study lakes and study populations

Mat-Sus refers to Matanuska - Susitna region. *No information available. [1]Benthic foraging reported (Park, Chase and Bell, 2012). Modified from (Park and Bell, 2010).

	Lynne	**South Rolly**	**Mud**	**Tern**
Topography	Limnetic	*	Benthic	Benthic
Location Lat (N)	61.712	61.401	61.563	60.533
Location Long (W)	150.039	150.073	148.949	149.55
Drainage	Little Susitna River	Susitna River	Knik River	Kenai River
Region	Mat-Sus	Mat-Sus	Mat-Sus	Kenai
Fish body shape	*	Limnetic	Benthic	Benthic
Foraging behavior	Limnetic[1]	Limnetic	*	*
Telencephalon shape	Limnetic	Limnetic	Benthic	Benthic
Number of egg-carrying individuals sampled	1	3	*	2
Number of parasitized individuals sampled	1	-	*	1

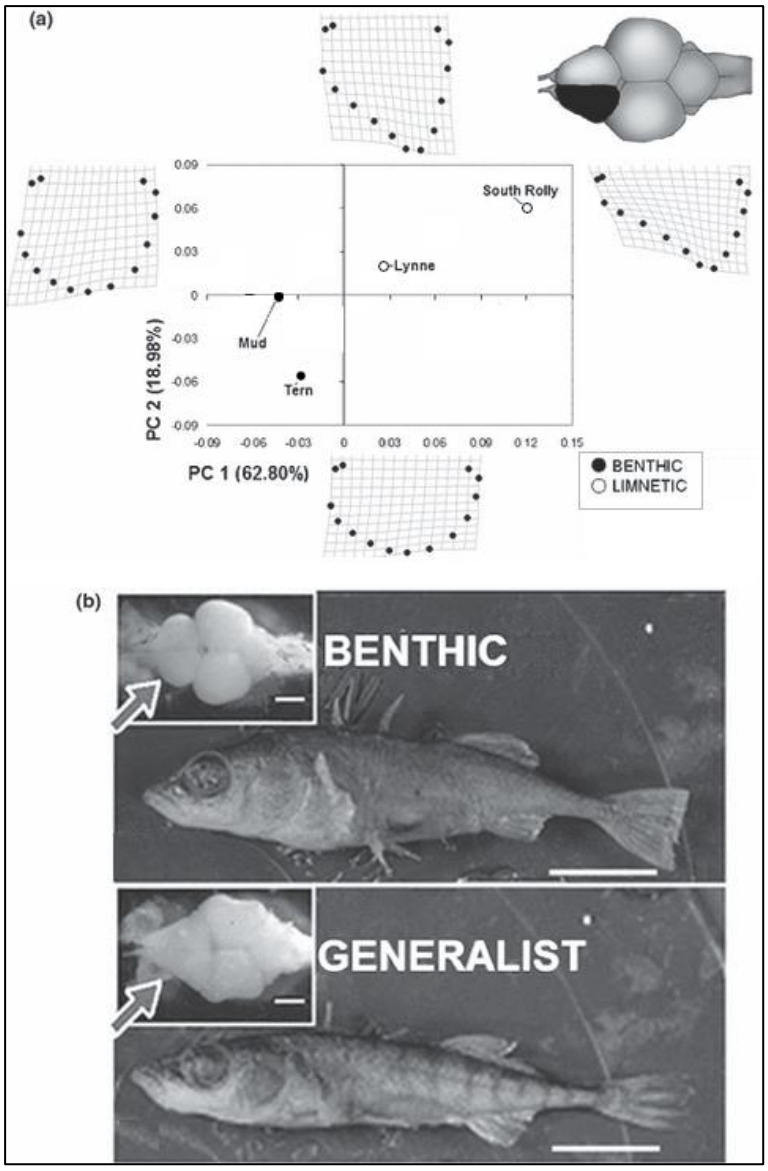


Figure 2. Telencephalon morphology. 1(a) A principal component analysis plot of the morphology of the telencephalons of the threespine sticklebacks sampled in (Park & Bell, 2010). The plot represents the sample means of the study populations. Only the populations shared with this study are included for clarity. 1(b) The telencephalon and body shape difference between benthic and limnetic ecotype. Generalist refers to limnetic ecotype. Modified from (Park and Bell, 2010).

All the study lakes have native predatory fish (Park and Bell, 2010). The benthic stickleback in Mud lake is sympatric with, but reproductively isolated from, anadromous threespine stickleback (Karve et al., 2008). The fish used in this study were confirmed to be resident benthic ecotype. All the telencephalons were collected from females to avoid the confounding factors of sex differences. The number of confirmed reproductively active sampled individuals varied among the populations (Table 1). Two samples in this study, one individual from Tern Lake and one from Lynne Lake were infected with *Glugea anomala* (microsporidia) (Table 1).

1.6 Aim of the thesis

In this thesis, I have examined transcriptome profile differences between benthic and limnetic threespine stickleback telencephalon and discussed the differences in ongoing functional processes in the study populations. I have also discussed the possible ecological reasons for the detected processes as well as their compatibility with 'clever foraging hypothesis'. The aim of the thesis was not to test the impact of any specific selective force on the adaptive radiation of threespine stickleback. Instead, I aimed to examine how gene expression differs in the differentiated ecotypes from distinct levels of habitat complexity, using the assumptions of past studies as a background to explain the detected gene expression differences.

Considering the well-established connection between environmental complexity and spatial cognition, I hypothesized up-regulated expression of neuron growth and connectivity related genes and functional pathways in the benthic ecotype of threespine stickleback. This expectation was supported by the earlier detected bigger telencephalon dorsolateral region (Park and Bell, 2010) and better overall performance in a spatial learning experiment (Park, 2013) in the benthic populations.

2 Material and methods

2.1 Sample collection and RNA extraction

The samples for the thesis were collected prior the start of the thesis. Telencephalons from female threespine sticklebacks from benthic (Mud and Tern) and limnetic (Lynne and South Rolly) populations in Alaska were collected during breeding season mid-June 2010 between 1-7 pm, using non-baited minnow traps. The fish were euthanized by decapitation and the telencephalons were immediately dissected and placed in 1,5 µl RNase free Eppendorf tubes. The tubes were instantly placed in dry ice. The samples were kept at -70°C until extraction to avoid RNA degradation. All the equipment used in sample collection was treated with 70 % ethanol and RNase Away (ThermoFisher) prior use.

The RNA was extracted in October 2010 - May 2011 with a phenol-based phase separation, following the RNA isolation protocol from the Tri-reagent manufacturer (Ambion, Applied Biosystems) (TRI Reagent® Solution Protocol (PN 9738M Rev D), 2010). The protocol is explained in detail in appendix 1.

After extraction, the samples were treated with RNase-free DNase, RQ1 1U/µl (Promega), following the manufacture's protocol (Appendix 1) and the isolation protocol was conducted again immediately after DNase treatment, with downscaling of reagent amounts for 400 µl Tri-reagent, to remove any traces of DNase. Finally, the RNA was dissolved in 15 µl NF water. After dissolving, the samples were stored at -70°C until processing for microarray.

All reagents used for RNA isolation were molecular grade. The glassware used for storing the working solutions was treated with 0.1 % diethylpyrocarbonate (DEPC), incubated at room temperature for 12 hours, and autoclaved, to inactivate RNase enzymes. The pipette tips and Eppendorf tubes were RNase free. All the surfaces and equipment were wiped with 70 % ethanol and RNase Away before extraction to avoid contamination.

RNA purity and yield were checked with the NanoDrop ND-1000 spectrophotometer (Thermofisher). RNA quality was further assessed with automated gel-based electrophoresis with Experion (Bio-Rad) in Turku Centre for Biotechnology. Only samples with spectrophotometer purity value 260/280 ~2.0 and RNA integrity value ≥7 were sent for microarray analysis. For sending, an aliquot of 600 – 800 ng RNA/sample was precipitated

with 3 M sodium acetate and stored in 100 µl of 75 % ethanol, to preserve RNA integrity during transport. For the microarray hybridization, 100 ng RNA/sample were used.

2.2 Microarray analysis

Four samples from each of the study populations were used for the analysis. The samples were processed in the UHN microarray center, Toronto, Canada, and analyzed on a custom two-color, oligonucleotide, 8x60k Agilent array. The platform consists of 60 bp, custom-designed probes, which are designed based on previously sequenced DNA sequences.

The microarray design was a modified version of a design described in (Leder et al., 2009). The study array is publicly available in the ArrayExpress repository, accession A-MTAB-573 (Agilent ID 041818). Briefly, the microarray consisted of 1319 controls and 61597 probes, which were present in 3 or 10 replicates. Instead of splice variant specific probes, consensus probes were chosen.

To measure the gene expression level differences between the ecotypes, 100 ng RNA per sample was labeled with Cy3 or Cy5 and hybridized against a sample from another ecotype in a microarray. After that, the fluorescence intensity of the probe spots in green and red channels was measured and used as a proxy for the probe expression level in the hybridized sample (Two-Color Microarray-Based Gene Expression Analysis (Low Input Quick Amp Labeling) Protocol, 2015). The array was designed with dye swap between the ecotypes to control dye-specific bias (Appendix 2).

2.3 Differential expression analysis

Before differential expression analysis, the technical quality of microarray experiments was evaluated. Agilent microarrays have a set of negative control probes for evaluating background noise levels due to unspecific cRNA binding and chemical emission. They also use positive control probes with the addition of spike-ins, the RNA sequences known to hybridize with the probes. These control features together with the signal intensity distribution of non-control probes are used for testing the microarrays' technical performance (Agilent technologies Inc., 2015). The microarrays used in this study were labeled as 'good' technical quality in all evaluation metrics, except for array 25418180002 1-2, where the slope of the expected/observed log ratio of Spike In signal intensity was slightly below the 'good' technical quality threshold.

Limma -package (Ritchie et al., 2015), designed for statistical program R (R Core Team, 2014) especially for linear modelling of microarray data, was used to analyze the gene expression differences among the populations. Prior to model construction, the data was normalized between arrays and filtered based on signal intensity across probe replicates. All the saturated probes and control probes were removed from the data and the signals were normalized across arrays using the method **Aquantile**. There was no apparent bias in the MA-plots of log2FC x log2 average intensity in any of the arrays, showing no need for within array normalization (Appendix 2).

Only probes with at least two replicates flagged as being well above the background signal and median log2 raw expression value being >5.4 were included in the final analysis to ensure data reliability. Population clustering was examined by creating a heatmap and a dendrogram from raw expression values using R-package Heatmaply (Galili et al., 2018). Subsequently, a principal component analysis (PCA) using raw expression values was performed with Limma-package, method **plotMDS** (gene.selection = ”common”), with all the expressed and with differentially expressed genes (DEG).

A linear regression model was fitted to median signal intensities between the probes from each population comparison, using habitat as the factor. The model fit was evaluated by the empirical Bayes method. Finally, the p-values were corrected for multiple comparisons using Benjamini-Yekutieli correction under dependency (Yekutieli and Benjamini, 2001) to control false discovery rate. All the p-values reported in this thesis are adjusted, unless stated otherwise.

2.4 Gene set enrichment analysis

2.4.1 Probe sequence annotation

The assignment of the probes to the *Gasterosteus aculeatus* genome was based on Ensembl release version 103 (Yates et al., 2020) and version 104 (Howe et al., 2021), stickleback genes (BROAD S1) reference genome using BLASTn (Basic Local Alignment Search Tool for nucleotides). From the blast, a full-length (60 base), 100 % identity, single alignment was present for 11 536 probe sequences and similar alignment with over 50 base length for additional 17 sequences of the total of 11 698 expressed probes. For 44 sequences, there were multiple alignments present. Most of these had the same gene product. For hits with different gene products, the ones with tissue-specific expression in the subject tissue were chosen. For

101 sequences, no alignment could be found or assigned reliably in the reference genome. From the 11 597 successfully aligned probes 253 assigned to duplicate genes.

The human orthologs for threespine stickleback genes were annotated using BioMart, Ensembl. One to one orthologs were present for 5 866 stickleback genes, one to many for 2 933 genes, and many to many for 405 genes. In case of multiple matches, the best human ortholog match was chosen based on GOC-score and whole-genome alignment score. For the stickleback genes for which no ortholog were returned from Ensembl BioMart, the annotation was supplemented with the Basic Local Alignment Search tool for proteins, (BLASTp) from the stickleback gene encoded protein to human protein in the Uniprot database (Bateman et al., 2021). Only the matches with e-score below 1e-50 and reviewed protein function were accepted. This way, 9 341 human orthologs could be assigned to expressed stickleback genes.

2.4.2 Functional enrichment analysis

Gene set enrichment analysis (GSEA) and the subsequent semantic similarity clustering was performed using ViSEAGO package (Brionne et al., 2019) from Bioconductor, R, following the pipeline from ViSEAGO vignette. The enrichment was performed with **fgseamultilevel** algorithm (Korotkevich et al., 2021), with minimum Gene Ontology (GO) term size 10 for biological process and 5 for molecular function, due to GO term count difference between the categories. The algorithm returns enriched GO terms with adjusted p-values and IC-value.

The gene set for GSEA was organized to a ranked list by multiplying the -log10 of the adjusted p-value of each probe by the sign of log fold change of the probe between the ecotypes from the differential gene expression model. All the expressed genes were included in the analysis. The method has no background gene set. In the analysis, the gene sets in the positive and negative extremes of the ranked list are weighted in the enrichment score calculation to avoid high enrichment scores in the middle of the list (Subramanian et al., 2005). GO annotations for the analysis were retrieved from Ensembl version 104 database (Howe et al., 2021). Human orthologues were used for superior annotation (20 358 biological process GO terms in Ensembl) compared to stickleback genes (BROAD S1) database (12 055 biological process GO terms) (Howe et al., 2021).

Hierarchical clustering based on semantic similarity of enriched GO terms was performed using Ward D2 aggregation method (Ward, 1963) with Resnik's distance (Resnik, 1995). The distance between GO clusters was computed using Best-Match Average (BMA) method.

3 Results

3.1 Differential gene expression

When exploring between-array normalized median raw expression values, most of the probes showed similar expression patterns in all the populations (Figure 4). Based on raw expression values, the three populations from Matanuska – Susitna valley (South Rolly, Mud, and Lynne) clustered together, apart from the Kenai peninsula population (Tern).

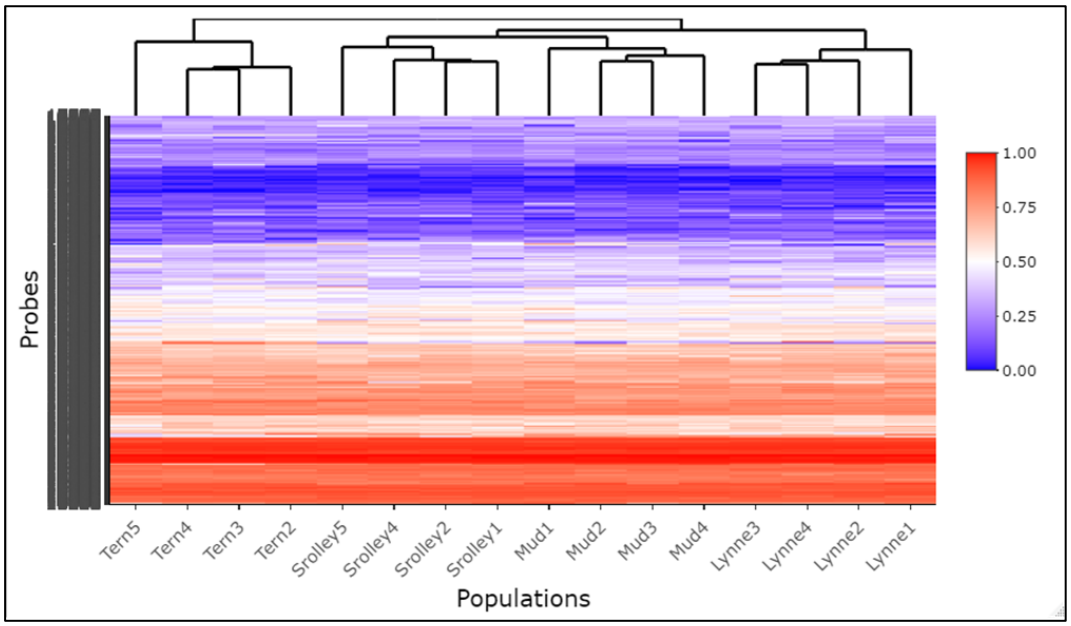


Figure 3. Heatmap based on between-array normalized median raw expression values of all the expressed probes using method percentize and eucledian distance. The data shows little ecotype-specific difference in gene expression. The populations from Matanuska - Susitna (South Rolly, Mud, and Lynne) cluster together.

Principal component analysis (PCA) with all the expressed probes showed ecotype-specific division in PC 2 (Figure 5a), apart from one outlier in the limnetic group. There was not much population-level clustering in the first two PCs. The method used here does not return the percentage of variation explained by the principal components.

In the PCA based on DEGs, the populations from each ecotype clustered together on the PC 1 axis (Figure 5b). Lynne and Mud, and South Rolly and Tern, respectively, clustered in PC 2, apart from one outlier in both limnetic populations (Figure 5b). Overall, the limnetic populations showed tighter clustering than the benthic populations.

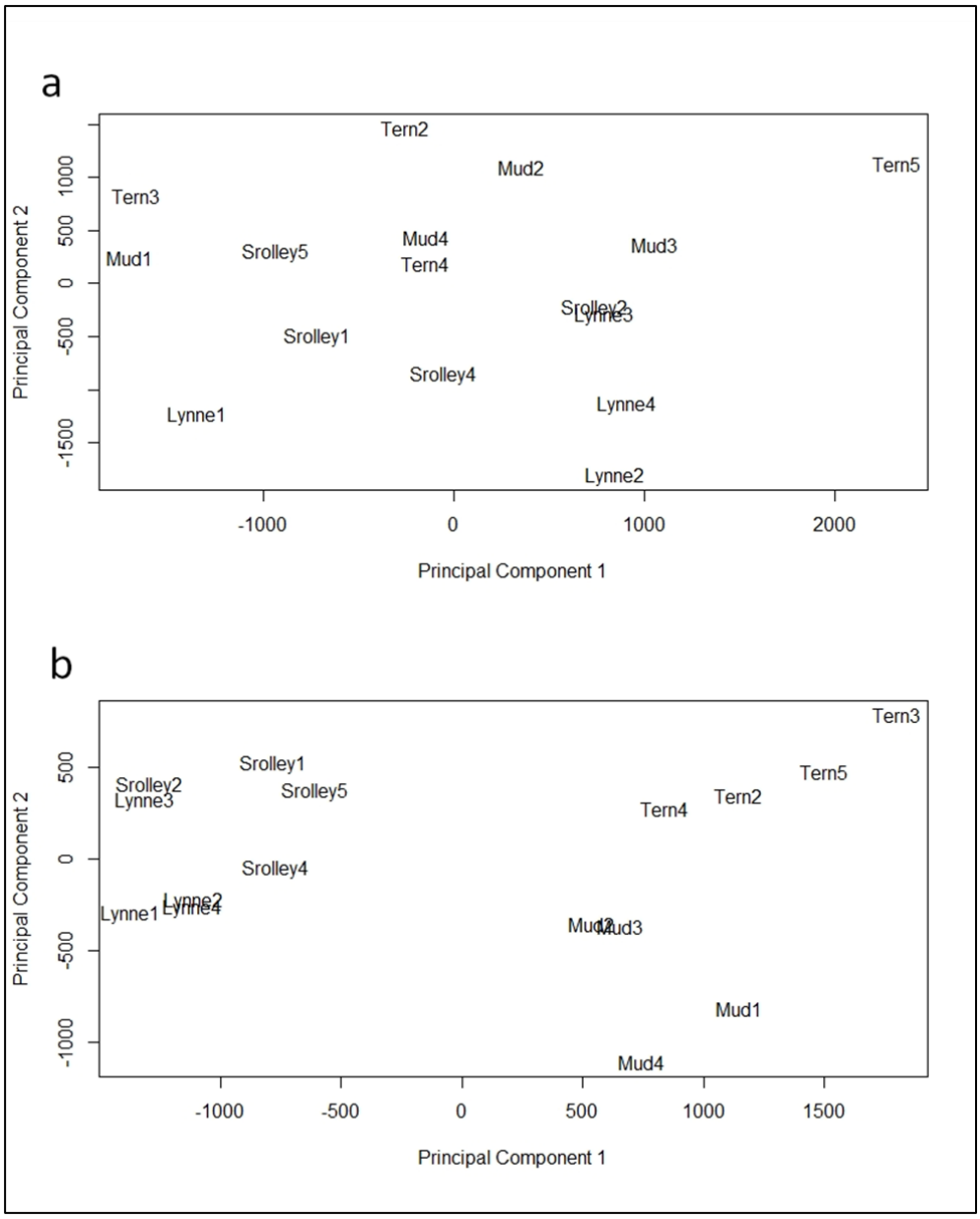


Figure 4. MDS-plots from between-array normalized median raw expression values. Figure 5a represents principal component analysis from all the expressed genes and 5b from differentially expressed genes adjusted p<0.05. Figure 5a shows weak separation between the ecotypes in PC 2. In figure 5b, the two ecotypes separate in PC 1. Limnetic populations cluster closer together, with one outlier in the Lynne population, whereas benthic populations form two clusters in PC 2.

In total 11 698 probes (57 % of the total number hybridized in microarray) were included in the differential expression model. Of those, 434 (9.4 %) were differentially expressed ($p<0.05$). Of differentially expressed probes, 42 % were upregulated in the benthic populations and 58 % in the limnetic populations (Figure 6).

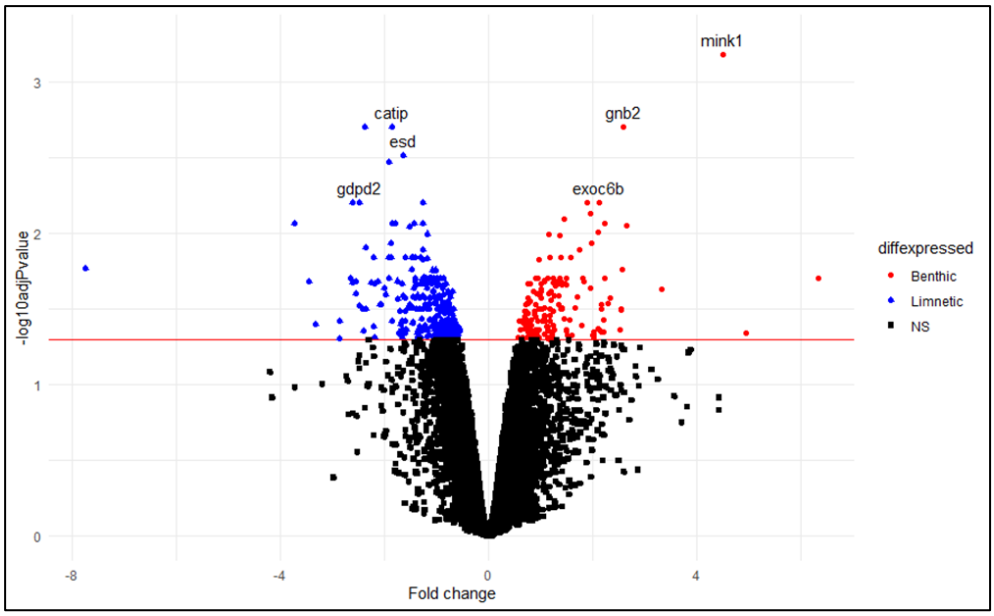


Figure 5. Volcano plot displaying differentially expressed genes between ecotypes, based on probe expression model p-values and fold change between ecotypes. The red line is the significance threshold p=0.05.

Of the 22 top DEGs with p<0.01, nine were upregulated in benthic (Table 3) and 13 in limnetic ecotype (Table 4). Of those genes, three genes upregulated in the benthic group (ENSGACG00000011017, ENSGACG00000014279, and ENSGACG00000018099) and one gene in the limnetic group (ENSGACG00000007572), lacked human ortholog and GO annotation.

The most significant DEG (logFC=4.51, p<0.0007), ENSGACG00000020394, MINK1 (MAP4K4, Mitogen-Activated Protein Kinase 6), was upregulated in the benthic ecotype (Table 3). It is a protein kinase, which gene product is an upstream regulator of JNK in neural cells (Larhammar et al., 2017). It has been associated with variety of other regulatory functions on other cell types as well (Yue et al., 2016). Two other well-characterized benthic upregulated genes, protein phosphatase PTEN and guanine nucleotide binding protein GNB2, are upstream regulators in protein kinase B (AKT) cascade (Lasarge and Danzer, 2014). Hence, all three are involved in MAPK signaling cascades that interact with p38 signaling pathway and participate in cell signaling.

In addition to cell signaling, another function observed in the benthic top genes was compound transport. It is associated with gene products of EXOC6B, NAPB, and WASHC3. All the characterized benthic top genes (p<0.01), except for WASHC3, engaged in signal transduction or synaptic transmission.

Table 2. Top genes (adjusted p<0.01) upregulated in the benthic ecotype

Go terms from Ensembl release 104 human ortholog annotations. *GO term from stickleback broad s1 annotation.

Stickleback gene ID	**Human ortholog gene symbol**	**Uniprot ID**	**GO terms**
ENSGACG00000020394	MINK1	MINK	protein kinase activity; transferase activity; phosphorylation; chemical synaptic transmission; neuron projection morphogenesis; actin cytoskeleton reorganization; positive regulation of JNK cascade; positive regulation of p38MAPK; regulation of AMPA receptor activity cascade
ENSGACG00000020395	GNB2	GNB2	GTPase activity; G protein-coupled receptor signaling pathway; calcium channel regulator activity; signal transduction
ENSGACG00000018808	EXOC6B	EXOC6B	exocytosis; Golgi to plasma membrane transport
ENSGACG00000011818	NAPB	NAPB	SNARE complex disassembly; syntaxin binding; vesicle-mediated transport; synaptic transmission, glutamatergic
ENSGACG00000009998	PTEN	PTEN	negative regulation of protein phosphorylation; positive/negative regulation of apoptotic process; neuron-neuron synaptic transmission; synapse assembly; regulation of neuron projection development; central nervous system myelin maintenance; protein kinase B signaling
ENSGACG00000019865	WASHC3	CCDC53	WASH complex; protein transport; actin filament polymerization
ENSGACG00000018099			integral component of membrane*

The most upregulated gene in the limnetic group, ENSGACG00000012847, CATIP (Ciliogenesis Associated TTC17 Interacting Protein) (Table 4), is involved in ciliogenesis and actin filament organization (Bontems et al., 2014). Three other top genes upregulated in limnetic ecotype, COL5A3, GDPD2, and NTNG2 are also involved with neuron projection and cell morphological organization. XRCC2 and SPIN1 are annotated with cell cycle. Important GO term groups in DEGs of limnetic ecotype are also histone

methylation/demethylation and transcription regulation. AASDHPPT, ESD, SAMM50, and CRYM participate in reversible histone methylation process. SPIN1 and CRYM function as transcription activity regulators.

Table 3. Top genes (adjusted p<0.01) upregulated in the limnetic ecotype

Go terms from Ensembl release 104 human ortholog annotations. *GO term from stickleback broad s1 annotation.

Stickleback gene ID	**Human ortholog gene symbol**	**Uniprot ID**	**GO terms**
ENSGACG00000012847	CATIP	C2ORF62	actin filament polymerization; cilium organization; cell projection organization
ENSGACG00000016535	COL5A3	COL5A3	collagen fibril organization; cell-matrix adhesion
ENSGACG00000008096	ESD	ESD	formaldehyde catabolic process; methylumbelliferyl-acetate deacetylase activity; glutathione derivative biosynthetic process
ENSGACG00000019330	SAMM50	SAMM50	SAM complex; MIB complex; protein insertion into mitochondrial outer membrane; integral component of membrane
ENSGACG00000017135	GDPD2	GDPD2	lipid metabolic process; glycerophosphoinositol inositolphosphodiesterase activity; actin filament reorganization
ENSGACG00000020819	NOP16	NOP16	RNA binding; ribosomal large subunit biogenesis
ENSGACG00000007572			integral component of membrane*
ENSGACG00000016021	NTNG2	NTNG2	synaptic membrane adhesion; intercellular bridge; regulation of neuron projection arborization; regulation of neuron migration; modulation of chemical synaptic transmission
ENSGACG00000016430	SEC11A	SEC11A	proteolysis; signal peptide processing; serine-type endopeptidase activity
ENSGACG00000008154	AASDHPPT	AASDHPPT	lysine biosynthetic process via aminoadipic acid; magnesium ion binding; transferase activity

ENSGACG00000001170	XRCC2	XRCC2	positive regulation of neurogenesis; DNA-dependent ATPase activity; mitotic cell cycle; DNA repair; replication fork
ENSGACG00000009554	CRYM	CRYM	negative regulation of transcription by RNA polymerase II; lysine catabolic process; NADP binding; oxidoreductase activity; thyroid hormone transport
ENSGACG00000018280	SPIN1	SPIN1	positive regulation of Wnt signaling pathway; positive regulation of transcription, DNA-templated; methylated histone binding

3.2 Functional enrichment analysis

Gene set enrichment analysis on all the expressed genes returned 59 enriched GO terms (Figure 7); 27 GO terms were enriched in the limnetic and 32 in the benthic ecotype. They clustered into 13 semantic similarity groups (Figure 8). After multiple testing adjustment, the six strongest supported GO terms remained significantly enriched (Table 5). These were lipid transport, GO:0006869, neutrophil degranulation, GO:0043312, innate immune response, GO:0045087, and negative regulation of protein kinase B signaling, GO:0051898 in the benthic ecotype and non-motile cilium assembly, GO:1905515 and regulation of smoothened signaling pathway, GO:0008589, in limnetic ecotype.

Most of the semantic similarity clusters were shared among the ecotypes, indicating similar functions being active in both groups (Figure 8). However, two clusters in the benthic ecotype, ‘lipid transport’ and ‘endocytosis’ (clusters 7 and 8) were distinctive for benthic ecotype, while ‘cell cycle’, ‘cell projection organization’ and ‘multicellular organism development’ (clusters 4, 5 and 9) were specific for limnetic ecotype (Figure 8). In shared clusters, most limnetic enriched GO terms were cell cycle related. In total 16 of 27 enriched GO terms in the limnetic ecotype were directly associated with cell cycle or morphology and projection, and additional five terms with transcription (Figure 7). The benthic ecotype was enriched for more diverse functions, including immune reactions, electron transfer, and regulation of signaling cascades (Figure 7). Both ecotypes shared similar functions in cluster 12, ‘programmed cell death’.

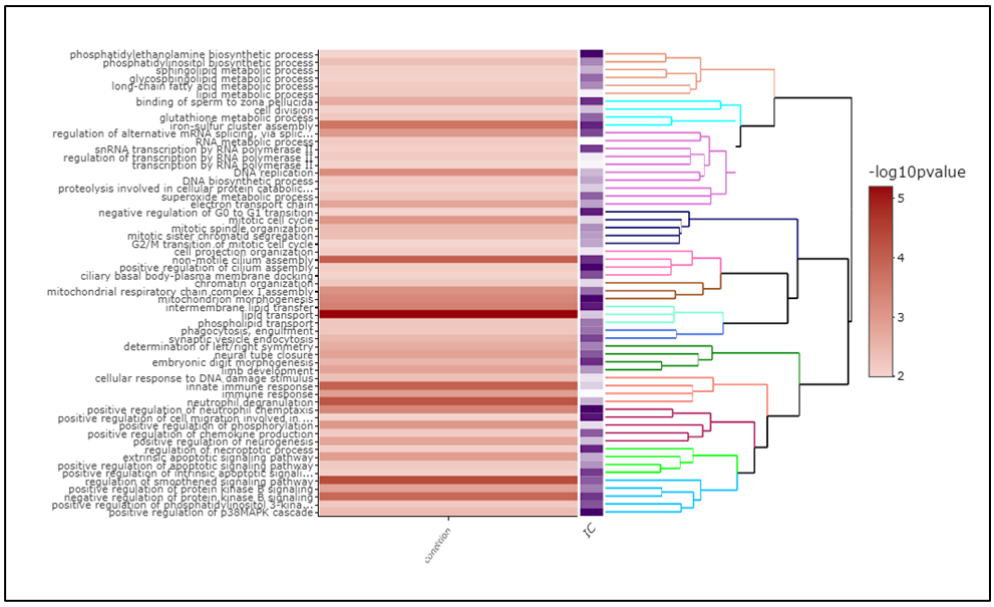


Figure 6. Gene set enrichment heatmap of enriched biological process GO terms with Resnik method semantic similarity clustering.

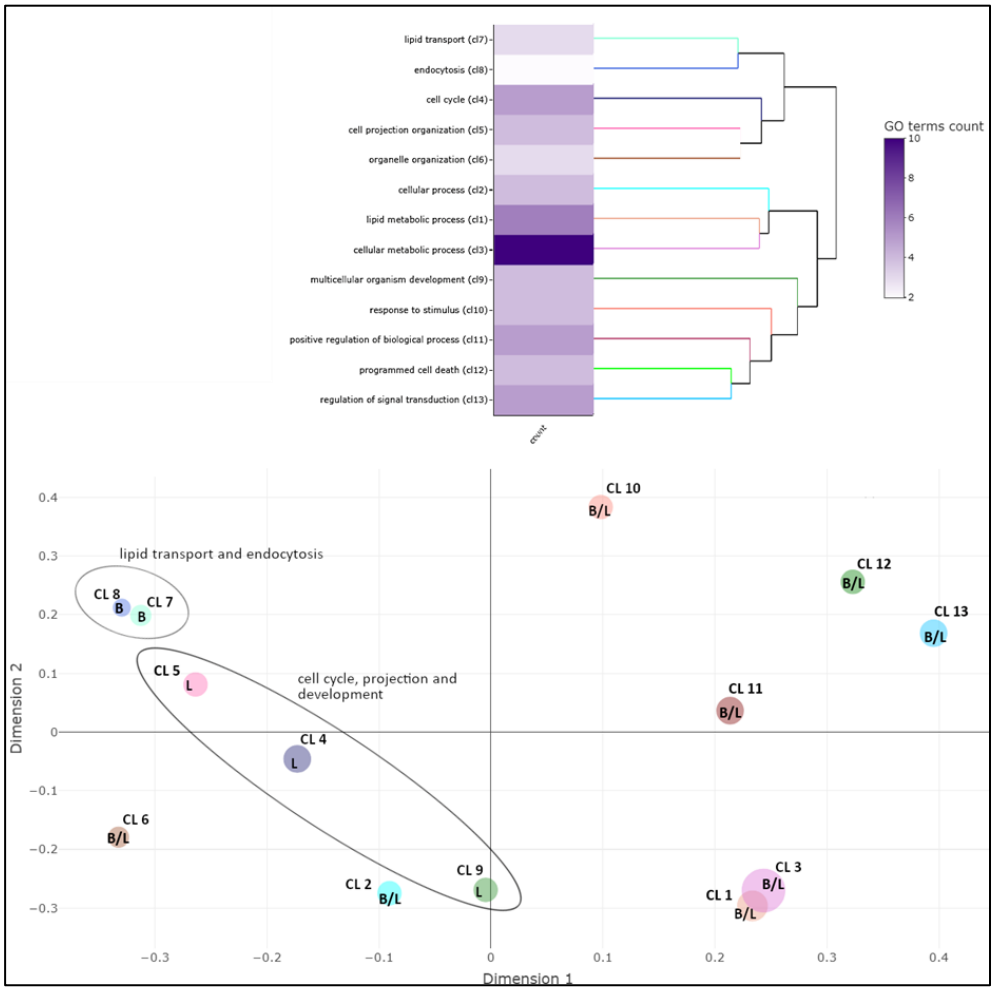


Figure 7. Multidimensional scaling plot from the semantic similarity of the enriched ‘biological process’ GO term clusters using Resnik clustering method and BMA distance. The clusters are labeled by the ecotype the GO term cluster is enriched in. B=benthic, L=limnetic.

GSEA for molecular functions returned 19 enriched terms (Figure 9). Three terms, DNA binding, GO:0003677, gamma-tubulin binding, GO:0043015, and tumor necrosis factor receptor binding, GO:0005164 were significantly enriched after multiple testing adjustment (Table 5). They were all enriched in limnetic ecotype. The most significantly enriched term in benthic ecotype was 4 iron, 4 sulfur cluster binding, GO:0051539 (Table 5).

The molecular function enriched terms clustered in 5 functional groups, three of which were shared among ecotypes (Figure 10). The ecotype-specific clusters were cluster 2, 'metal cluster binding' for benthic ecotype and clusters 3 and 5, 'signaling receptor binding' and 'DNA binding' for the limnetic ecotype, respectively. In the shared clusters, the molecular functions were mainly DNA and chromatin binding in the limnetic ecotype, and protein and lipid binding and phosphatase activity regulation in the benthic ecotype (Figures 9 and 10).

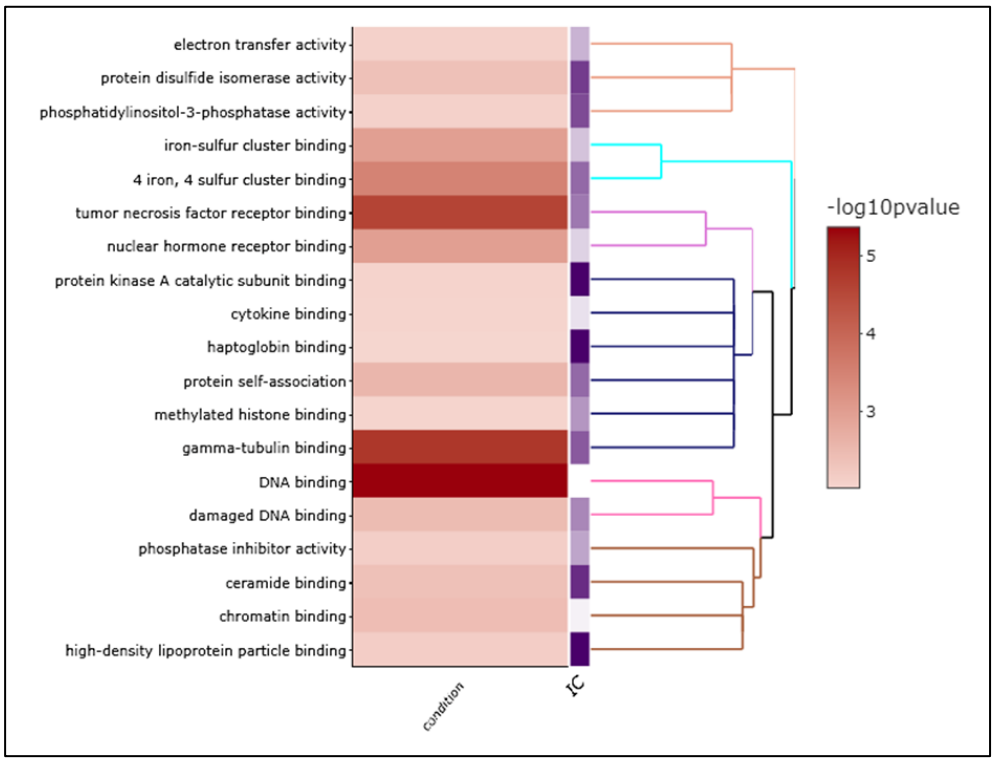


Figure 8. Gene set enrichment of molecular functions GO terms. The enriched terms are clustered by semantic similarity using Resnik method.

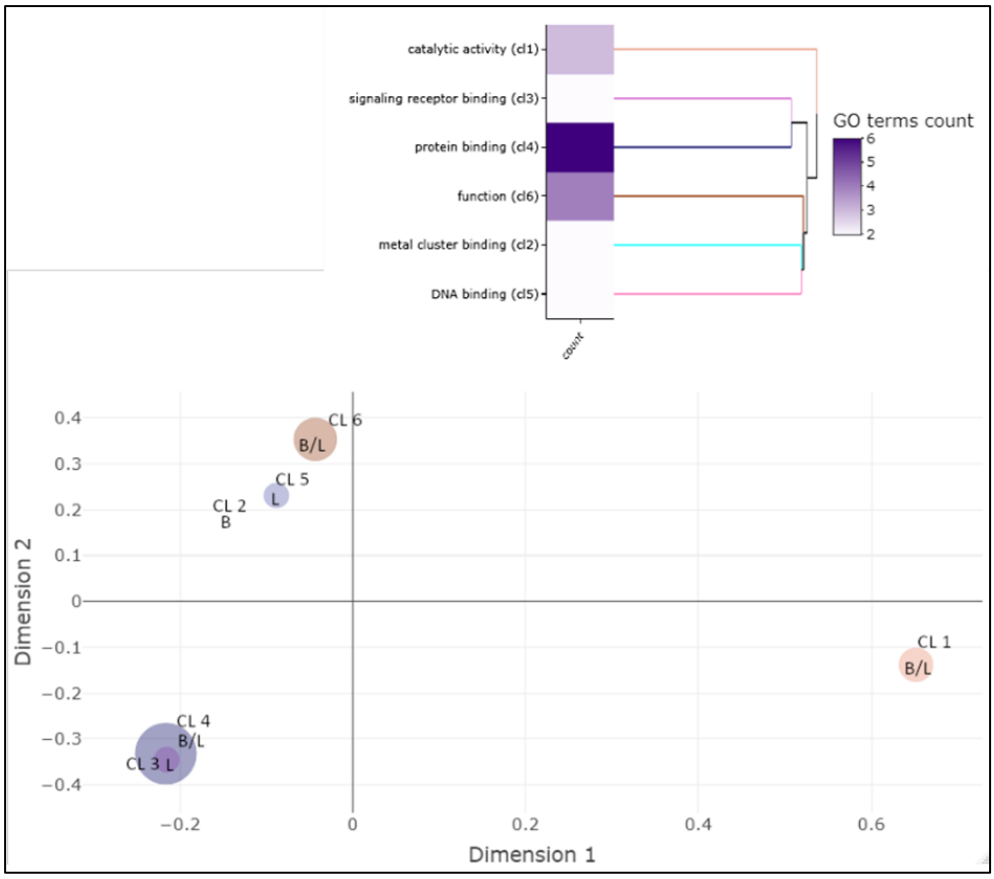


Figure 9. MDS plot of molecular function GO terms semantic similarity clusters using Resnik method and BMA distance.

Table 4. FGSEA enriched GO terms at significance level p<0.1.

GO class: BP=Biological process, MF=Molecular function. Ecotype: B=benthic, L=limnetic.

GO cluster	GO class	Ecotype	GO ID	GO term	Condition padj
7	BP	B	GO:0006869	lipid transport	0.01
10	BP	B	GO:0043312	neutrophil degranulation	0.03
10	BP	B	GO:0045087	innate immune response	0.03
13	BP	B	GO:0051898	negative regulation of protein kinase B signaling	0.04
2	BP	B	GO:0016226	iron-sulfur cluster assembly	0.05
7	BP	B	GO:0120009	intermembrane lipid transfer	0.07
6	BP	B	GO:0070584	mitochondrion morphogenesis	0.08
11	BP	B	GO:0090023	positive regulation of neutrophil chemotaxis	0.08
6	BP	B	GO:0032981	mitochondrial respiratory chain complex I assembly	0.1
5	BP	L	GO:1905515	non-motile cilium assembly	0.03

GO cluster	GO class	Ecotype	GO ID	GO term	Condition padj
13	BP	L	GO:0008589	regulation of smoothened signaling pathway	0.03
3	BP	L	GO:0006260	DNA replication	0.09
2	MF	B	GO:0051539	4 iron, 4 sulfur cluster binding	0.07
5	MF	L	GO:0003677	DNA binding	0
4	MF	L	GO:0043015	gamma-tubulin binding	0.01
3	MF	L	GO:0005164	tumor necrosis factor receptor binding	0.01

4 Discussion

4.1 Population differentiation

When examining the eucledian distance of raw expression values, there was weak clustering by populations and geographical regions, but not by geographical proximity within regions (Figure 1, Figure 4). There is no direct contact between the Matanuska – Susitna area study drainages, which excludes the possibility of gene flow between the study lakes. The separation of the Tern Lake population from the other study populations can be explained by its location at higher elevation and further inlands, as well as by its narrower genetic diversity (Park, 2013). Genetic clustering by watersheds and geographical proximity within watersheds has previously been detected in threespine stickleback populations from Vancouver Island, British Columbia (Bolnick and Ballare, 2020). Although the distinction between the gene expression profiles of the Matanuska - Susitna area populations and the Tern Lake population was not strongly supported (Figure 4 and 5a), it may have diminished the statistical power in detecting transcription level differences between the ecotypes. Despite the geographical factor, PCA with DEGs supported the presence of ecotype-specific gene expression pattern (Figure 5b). This is concordant with the differences in telencephalon shape and size (Figure 2), which has been previously reported in these populations (Park and Bell, 2010).

The greater distribution between the benthic populations compared to the limnetic populations in the PCA with DEGs may relate to above discussed factors concerning geographical location of the benthic populations or differences in the reaction norms of the populations. Sympatric Mud population (Introduction 1.5) may experience more extreme habitat specialization compared to allopatric populations. The anadromous ecotype sympatric with the Mud population resembles limnetic ecotype (Karve et al., 2008). Non-hybridizing, sympatric populations are often more specialized than allopatric populations (Svanbäck and Schluter, 2012). The values in PC 2 were more extreme in the Mud Lake population than the in the Tern Lake population (Figure 5b). On the other hand, the lack of genetic diversity in the Tern Lake population may show here as a separating factor. These two populations have previously been found to differ in spatial performance (Park, 2013) (Introduction 1.5). Interestingly, the populations with poor spatial performance and the populations with good spatial performance in Park (2013) clustered in PC 2 based on DEGs in this study (Figure 5b).

4.2 Ecotype-specific patterns in neural regulation

4.2.1 Upregulated neurogenesis and transcription in the limnetic ecotype

The enrichment of cell cycle and projection associated gene sets (Figure 8) indicates ongoing active engram formation in the limnetic ecotype. Engram formation is accompanied by neural cell proliferation and structural changes in neurons (Introduction 1.2.2).

Gamma tubulin, the target for a significantly enriched molecular function in the limnetic ecotype, 'gamma tubulin binding', is a primary compound in the formation of mitotic spindle (Alvarado-Kristensson, 2018). It also forms nucleating centers that initiate microtubule formation (Alvarado-Kristensson, 2018), which in turn contribute to cytoskeleton formation and direct neuron growth and morphology (Kapitein and Hoogenraad, 2015). Non-motile cilia, the target for 'non-motile cilium assembly', which was significantly enriched in the limnetic ecotype (Table 5), act as cell antennas. They are essential for various biological processes ranging from non-synaptic neuronal signaling in mature neurons to cell differentiation and projection and axon guidance (Guemez-Gamboa et al., 2014). Microtubules are central structural elements in cilia (Kapitein and Hoogenraad, 2015). The most significant DEG in the limnetic ecotype, CATIP (Table 4), is involved in actin polymerization modulation and cilia initiation (Bontems et al., 2014). The importance of cell cycle regulation in the limnetic ecotype is further consolidated by the enrichment of the biological process 'DNA replication' and several non-significantly enriched biological processes (Table 5, Figures 7-8).

The only limnetic enriched signaling pathway, 'smoothened signaling' (Table 5), is signal transduction in the 'sonic hedgehog' pathway, which is associated with neuron differentiation, proliferation, and axonal elongation (Delmotte et al., 2020). 'Smoothened signaling' activation itself promotes GABAergic neuron transition from excitatory to inhibitory stage during maturation (Delmotte et al., 2020), which is an important regulatory mechanism for synaptic signal transduction and is needed for systemic homeostasis. The enrichment of processes that are associated with neural circuit formation is further supported by the upregulation of two genes from the WNT-signaling pathway in the limnetic ecotype, SPIN1 (Table 4) and REG1, which has the largest fold change of the DEGs (Appendix 3). WNT signaling regulates a variety of neural circuit formation processes (Rosso and Inestrosa, 2013), and is an upstream regulator in mTOR pathway (Lipton and Sahin, 2014).

Transcription-related gene sets were non-significantly enriched in the limnetic ecotype (Figure 7). The enrichment of transcription-related processes solely in the limnetic ecotype indicates increased biological significance for transcription regulation in the limnetic group. Gene sets involved in transcription regulation have been demonstrated to be upregulated in memory consolidation and throughout memory formation (Marco et al., 2020).

Adult neurogenesis is substantially more frequent in teleost fish than in mammalian brain (Zupanc and Sîrbulescu, 2011), which emphasizes its significance for fish behavior. To stress the connection between neurogenesis and learning, environmental enrichment has been shown to increase proliferation in teleost fish (Lema et al., 2005; von Krogh et al., 2010). Together the evidence from DEGs and enriched functional gene sets indicate the individuals in the limnetic ecotype experiencing conditions that promote learning and memory formation.

4.2.2 Metabolic functions and immune defense enriched in the benthic ecotype

Functional groups associated with increased metabolism and immune defense were enriched in the benthic ecotype. Biological processes ‘Lipid transport’, ‘iron-sulfur cluster assembly’, ‘mitochondrion morphogenesis’ and ‘mitochondrial respiratory chain complex I assembly’ are associated with cellular energy production. The strongest supported biological process, ‘lipid transport’ (Table 5), is needed for both intake of fatty acids and localization of synthesized lipids. Twenty percent of the energy needed in the brain has been estimated to be produced by fatty acid oxidation (Tracey et al., 2018). Cellular energy production mainly takes place in mitochondrial respiratory complexes by electron transport in iron-sulfur complexes, which are also assembled in mitochondria (Ciofi-Baffoni et al., 2018; Khacho and Slack, 2018). The function of these complexes is regulated by morphological modifications in mitochondria (Khacho and Slack, 2018). Mitochondrial dynamics is important in the regulation of various energy-demanding neuronal processes because it directly regulates cellular metabolism (Khacho and Slack, 2018).

Apart from energy production, lipids in brain have major roles as structural molecules in cellular membranes and in neural signaling (Tracey et al., 2018). Phosphatidylinositol (PPI) signaling was a major compounding factor in the benthic ecotype’s enrichment profile in this study. ‘Phosphatidylinositol biosynthetic process’ was non-significantly enriched in the benthic ecotype. PPIs are lipid species, that have minor roles as membrane structural proteins, but also act as signaling molecules in different signaling cascades and cellular types (Balla,

2013). Since PPI signaling is cellular-location dependent, lipid transport and transfer are important in PPI signaling.

Phosphorylation state, together with cellular location, determines the regulatory impacts of PPIs (Balla, 2013). 'Positive regulation of phosphorylation' was enriched in the benthic ecotype. Phosphorylation of PPIs is regulated by a variety of kinases and phosphatases. The best-studied PPI pathway is pI3k/PTEN regulated pI3k/AKT pathway (Introduction 1.2.3). In this study, both positive and negative regulation of AKT signaling ('regulation of protein kinase B signaling') were enriched in the benthic ecotype, but only the enrichment of negative regulation was significant (Figure 7, Table 5). The enrichment of negative regulation of AKT was driven by PTEN (Table 3), a protein phosphatase which inhibits PPI-induced activation of AKT on plasma membrane (Balla, 2013) and downregulates the activation of downstream mTOR pathway (Introduction 1.2.3). Besides pI3k/PTEN regulation, AKT activity is regulated by variety of other signaling cascades (Balla, 2013) and in turn participates in downstream regulation of several processes, including neurodegeneration, apoptosis and synaptic strength (Brazil et al., 2004).

AKT signaling in a MAPK-family protein kinase c-Jun N-terminal kinase (JNK) pathway promotes cell survival and prevents apoptosis (Brazil et al., 2004; Larhammar et al., 2017). Interestingly, the strongest upregulated DEG in the benthic ecotype, MINK, participates in activation of neurodegenerative signaling cascade in JNK pathway (Larhammar et al., 2017). Within this pathway, MINK is also involved in induction of depotentiation after LTP by mediating AMPA-receptor removal from synapses (Zhu et al., 2005) and suppressing dendritic spines and excitatory synapses (Han et al., 2010). JNK signaling was not enriched in this study, but it strongly interacts with p38MAPK pathway, a stress induced pathway involved in e.g., neurodegeneration and regulation of synaptic plasticity, notably the astrocyte-mediated induction of LTD (Navarrete et al., 2019; Falcicchia et al., 2020). 'Positive regulation of p38MAPK cascade' was non-significantly enriched in the benthic ecotype (Figure 7).

PI3k/AKT pathway is also involved in the regulation of 'neutrophil chemotaxis' and 'neutrophil degranulation' (Table 5) enriched in the benthic ecotype (Heit et al., 2002; Lodge et al., 2020). Neutrophils, which migrate to brain via bloodstream, attend to innate immune defense together with astrocytes and microglia (Kim et al., 2020). They move to targets by chemotaxis and can destroy the target by releasing proteolytic enzymes, reactive oxidative

species (ROS), and stimulating inflammatory cytokines (Rane et al., 2005). Together with pI3k/AKT signaling, p38MAPK pathway participates in regulating neutrophil chemotaxis (Table 5) (Heit et al., 2002). Neutrophil chemotaxis is regulated by the mitochondrial electron transport chain and redox status (Zhou et al., 2018). Thus, the enrichment of neutrophil chemotaxis, mitochondrial dynamics, and the assembly of electron transfer machinery have a complementary function. The enrichment of immunological functions is also supported by the gene with the greatest fold change in the benthic group, GIMAP7 (logFC=5.0, p=0.46), a member of immune-associated subfamily of nucleotide-binding proteins.

The pathways that were enriched in the benthic ecotype, together with mitochondrial functions, are essential for maintaining systemic homeostasis, but also for various organismal processes, including neuronal dynamics in engram formation (Chen et al., 2005; Khacho and Slack, 2018; Tracey et al., 2018; Falcicchia et al., 2020) and immune defense. Immune defense and memory formation share elements (Falcicchia et al., 2020; Cornell et al., 2022) and organismal stress has been demonstrated to negatively influence learning (Wood et al., 2011) and brain cell proliferation (Øverli and Sørensen, 2016). Hence, the environment-organism interaction does not appear to promote memory formation through active gene expression in the benthic ecotype during the sampling time.

4.3 Potential sources for the ecotype-specific gene expression profiles

4.3.1 Limnetic populations

The upregulation of cognitive processes in the limnetic ecotype can be explained by both environmental and social complexity. Benthic populations forage in complex littoral environments throughout their lifespan, while limnetic populations have been reported to use mostly pelagic prey outside the breeding season, although the distinction is not stark (Introduction 1.4). It is possible, that limnetic individuals undergo habitat transition in the breeding season and as a result experience increased engram formation and restructuring compared to benthic individuals in familiar environments. Habitat generalists have indeed been suggested to outperform specialists in cognitional performance (Henke-von der Malsburg et al., 2020). Outside the breeding season, benthic populations live in patchier habitats, but given the temporal aspect, limnetic populations may experience more variable environment. However, the potential of habitat transition as a factor in the transcriptomic difference is diminished by the discovery of the recent diet change from pelagic feeding to

feeding on benthos in the Lynne population (Park et al., 2012), although the topography of Lynne Lake and the telencephalon shape of the Lynne population are classified as limnetic (Park and Bell, 2010). On the other hand, the enrichment of cognition-related functions can be explained by more variable habitat use throughout the year in the limnetic populations, as suggested by Park & Bell (2010).

Besides feeding behavior, increased level of sociality in the limnetic populations in adulthood compared to the benthic populations, which are more solitary, may be a key factor in promoting learning and memory. Interestingly, in the comparison between male and female telencephalon, Park and Bell (2010) found the telencephalons of males, besides being bigger than in females morphologically resembled that of limnetics. Since threespine stickleback males show paternal care, they experience more socially complex environment than females. Park and Bell (2010) have suggested social factors as possible explanation for the telencephalon size fluctuation among populations and sexes. The triangular shape is concurrent with the idea of neuron growth in the dorsomedial region of the telencephalon, which is primarily considered to be responsible for social and emotional processes (Demski, 2013). Biological processes that are associated with neuron growth and projection have been found to be upregulated in socially reared threespine stickleback females (Greenwood and Peichel, 2015). As discussed in the introduction (1.1.3, 1.3.1), social interactions have repeatedly been shown to have major impact on cognitive processes. Together, the concave shape of limnetic telencephalon and enrichment of learning- and memory-related functions in this study emphasize the possibility of the impact of social interactions on cognition of limnetic females.

4.3.2 Benthic populations

Several lake and organismal features may contribute to the detected enrichment of the immunological and metabolic functions in the benthic ecotype. First, parasite abundance, parasite community, and parasites' ability to infect and exploit their hosts may vary among the lakes and the habitats. Stutz et al. (2015) have discovered significant lake- but also habitat-specific, plastic differences in immune gene expression between threespine stickleback populations from small and large lakes. The discovery underlines the impact of parasite community structure on host's gene expression. Benthic and limnetic threespine sticklebacks have been suggested to differ in their exposure to limnetic and benthic parasites due to different proportions of consumed intermediate hosts (MacColl, 2009; Stutz et al.,

2014), There is a complex relationship concerning exposure rates and parasite infections, though, perhaps because of avoidance behavior or resistance to infection (Stutz et al., 2014; Arostegui et al., 2018).

Second, sympatric limnetic/benthic populations have been hypothesized to differ in their investment in immune defense (MacColl, 2009). Exposure to different parasites, together with host genetics and exposure history have been shown to impact the activation, strength, and mode of immune defense in threespine stickleback populations (Kalbe and Kurtz, 2006; Lenz et al., 2013; Piecyk et al., 2021). Functional groups that have previously been enriched in threespine stickleback upon parasite exposure include immunological and metabolic processes (Lenz et al., 2013). Infection with a parasite has been demonstrated to alter the expression of inositol pathway in threespine stickleback brain, simultaneously with a behavioral change (Grecias et al., 2020). While the inositol pathway was not enriched in this study, PPIs enriched in the benthic ecotype (Discussion 4.2.2) are an important part of inositol signaling.

One individual from both ecotypes in this study was observed to be infected with a parasite. However, the early stages of infection are difficult to detect, and mere exposure to a parasite has been shown to be sufficient to alter gene expression in the brain (Grecias et al., 2020) and the head kidney (Piecyk et al., 2021). The head kidney has been suggested to be involved in the redistribution of energy substrates during immune response in teleost fish (Geven and Klaren, 2017). It must be noted that previous studies describing transcriptional response to parasite exposure in threespine stickleback are mainly transplant studies between novel parasite/host combinations and lab-bred, previously unexposed individuals. Thus, they are not readily comparable to wild populations in their native habitats where threespine stickleback populations are adapted to their local parasite communities.

The other potential source for the need of immune defense stem from behavior and physiological stress associated with the breeding season. Stressful courtship behaviors, such as biting and chasing females, exist in both ecotypes, but are more common in the benthic ecotype (Rowland, 1994; Shaw et al., 2007). Social stress has been demonstrated to increase neutrophil mobilization in mice (Ishikawa et al., 2021).

The third possible explanation for the enrichment of immunological functions is the consumption of algal toxins. The high ratio of surface area to water volume makes shallow lakes more susceptible to warming than deeper lakes (Moran et al., 2010). Warming, especially together with thermal stratification and eutrophication, exposes them to toxic algal

(including cyanobacteria) blooms (Vilas et al., 2018). While the lakes in the study area are mostly oligo-or mesotrophic (Jones et al., 2003), toxic algae are frequently encountered in Alaskan lakes. The impact of lake morphology and rooted vegetation on algal blooms is complex. Lakes with widespread rooted vegetation, as both benthic lakes in this study, are often less phytoplankton dominated (Bayley and Prather, 2003; Vilas et al., 2018). Shallow lakes are also less prone to thermal stratification than deeper lakes. On the other hand, infrequent thorough mixing promotes the release of nutrients from lake sediment for algae use (Vilas et al., 2018). Panmictic conditions also promote the mixing of toxic algae throughout the water column, thus making it difficult to avoid their consumption. Algae and cyanobacteria toxins have direct immuno- and neurotoxic effects on teleost fish (Banerjee et al., 2021). Exposure to cyanobacteria toxin microcystin has been shown to have multifaceted impacts on the fish immune system, depending on dosage and exposure route, including up- and downregulation of innate immune response, induction of inflammatory reaction, and apoptosis (Lin et al., 2021).

Finally, a possible cause for the enrichment of lipid transport and metabolism is physiological stress that threespine stickleback females experience because of energy-demanding, repeated spawning and egg production (Wootton, 1994; Foster et al., 2008). Benthic fish have been suggested to produce greater clutch size per spawning than limnetics (Bell and Foster, 1994). The difference in egg production is predicted to even out during the spawning season (Bell and Foster, 1994), but the one-time energy investment and thus energy demand remains higher in benthic than in limnetic females. To meet the energy demand, females increase their energy intake, but also consume systemic lipid and glycogen reservoirs and decrease their growth rate (Wootton, 1994). While there is no evidence for food availability limitation in the benthic populations, oxygen depletion has been suggested to limit feeding ability in male threespine stickleback during the breeding season (Guderley, 1994). A similar mechanism may shift cellular energy production in brain from external sources to the use of systemic lipid reservoirs in benthic females and contribute to the detected enrichment of lipid metabolism; however, no directly hypoxia associated processes were detected in this study, although benthic functional enrichment shared similarities with hypoxia associated decrease in DNA replication and mitosis and upregulation of oxidase and oxygenase activity (Leveelahti et al., 2011).

The enrichment of stress-related functions in geographically distant populations of Mud and Tern lakes is interesting. It is an intriguing question whether the ecotype-specific expression

of immune defense is triggered by environmental factors, which are either correlated with lake topography or pathogen community, or by genetic differences between the ecotypes in the immune response reaction norm. Genomic markers near genes linked to both immunology and cognition has been found to vary in frequency in threespine stickleback lake populations in relation to lake size (Bolnick and Ballare, 2020).

4.4 Elements of uncertainty

4.4.1 Populations and tissue collection

Sampled populations in this study represent benthic and limnetic threespine stickleback populations. The ecological specialization (Park and Bell, 2010) in allopatric study populations is not as extreme as in sympatric populations of the same ecotypes and one of the limnetic populations has atypical feeding habits. Also, cognitional performance among the populations (Park, 2013) does not strictly follow the expectations of benthic/limnetic division. This adds uncertainty to the results. However, the telencephalon and body morphology of the study populations, as well as lake morphology, are typical for benthic and limnetic ecotypes, which justifies their use as sample populations in this study.

The samples were collected using minnow traps. Although the samples were collected within an hour per site, the time the fish spent in the trap may have varied. This could potentially have had an impact on the stress level of individual fish. This impact could not be controlled. However, it is unlikely that the results are severely affected by the variation in trapping time because there is no evidence for systematic bias in trapping time between the lake types.

Fish age may change its transcriptome. While the age of the sampled fish was not determined, a proportion of samples from every population except for Mud was confirmed sexually mature (with eggs). Hence, it is unlikely that these results are driven by age-dependent changes in RNA abundance.

4.4.2 Data analysis

Functional enrichment is sensitive to gene set selection and the annotation coverage of selected genes (Primmer et al., 2013; Stanford et al., 2020). The use of annotations from distant phylogenic group is complicated by evolutionary differences among the groups (Primmer et al., 2013). While GO annotations are not species-specific, some uncertainty to their interpretation comes from the fact that many of the regulatory pathways have been

studied in mammals and invertebrates, and it is not known how conserved they are between the study organisms and teleost fish (Pravosudov et al., 2013; Sun and Lin, 2016). The use of human annotations in the functional enrichment instead of threespine stickleback or other teleost species increased annotation coverage but reduced the number of genes included in this study. To diminish the ambiguity from human ortholog use, the orthologs were chosen with strict selection criteria (Material and methods 2.4.1). Cognitive processes are relatively conserved in vertebrates. Still, it must be remembered, that the knowledge of regulatory pathways is incomplete.

The gene set enrichment method applied in this thesis allows potentially biologically meaningful genes with subtle, non-significant changes in the expression levels to be included. The problem of the enrichment of gene sets with non-significant up- or downregulation that lack phenotypic correlation is resolved by weighting the positive and the negative values in the ranked list of genes (Subramanian et al., 2005). The gene set for the analysis can be ranked by fold change or significance. While the fold change is biologically meaningful, the significance value, which has been used here, better considers the consistency among samples. Because of the weighting of the extreme values, the ranking method may have influenced the enrichment.

Since transcriptome is not equivalent to protein expression, it does not accurately describe the phenotypic outcome. Measuring mRNA contents also leaves out non-coding functional RNA (Stanford et al., 2020). Finally, the biologically most meaningful transcripts may not be among the most expressed ones (Stanford et al., 2020). Despite this, RNA profiling has been proven to reveal biologically meaningful differences between populations (Introduction 1.3.2). The use of FGSEA in this study is intended to minimize the effects of the difference between statistical and biological significance.

5 Conclusions

In this thesis, I have discovered divergent gene expression in the telencephalons of benthic and limnetic threespine sticklebacks from allopatric populations with non-extreme foraging behavior. Contrary to my hypothesis, the limnetic populations in this study were enriched with more neural growth-associated functions compared to benthic populations. This contradicts the presumptions of 'clever foraging hypothesis', as benthic foraging habitat has complex elements compared to limnetic habitat. However, it is not clear whether the results in this study have been driven by the enhancement of cognitional processes in the limnetic ecotype or rather by environmentally or socially induced organismal stress in the benthic ecotype, severe enough to trigger energetic trade-off and counteract the impact of environmental complexity on gene expression.

Considering the earlier research on the foraging habits and cognition of the study populations, it is plausible that the distinction in the foraging strategy among the ecotypes in these populations is too subtle and inconsistent to be the main driving force of the ecotype-specific transcriptome profiles. Hence, these results do not unambiguously neither support nor oppose 'clever foraging hypothesis' in explaining cognitional differences between benthic and limnetic ecotypes in general. Instead, they support the presence of trade-off between energy investment on immunology and cognition.

Because of the transient nature of transcriptome and plasticity of both cognitional and immunological traits, these results reflect only a limited time span and a unique combination of environmental factors. Further research is needed to discover whether they can be applied to other populations in the benthic - limnetic axis. The presence of lake size correlated variation on the genetic markers of both traits indicates heritable differences between the ecotypes and potential importance of those traits for adaptive radiation of threespine stickleback in benthic and limnetic lakes (Discussion 4.3.2). Especially interesting question is the extent of environmental versus genetic impact on the expression of innate immune defense among the ecotypes. What is the role of immunology in the adaptive radiation of threespine stickleback? What are the unifying environmental factors that trigger the expression of innate immunology in lakes that are structurally similar, but differ in their geographical location?

Acknowledgments

I would like to thank my supervisors Erica Leder, Heidi Viitaniemi, and Irma Saloniemi for their endless patience, support, and insightful comments. I would also like to thank Erica for an interesting topic and material for the thesis. A special thanks to Professor Michael A. Bell for sharing his expertise on Alaskan threespine stickleback populations during sample collection. Finally, many thanks to Heidi for everything, especially for teaching me the art of RNA handling, and Tiina Sävilammi for invaluable help with statistical analysis and R.

References

Abraham, W. C., Logan, B., Greenwood, J. M., and Dragunow, M. (2002). Induction and experience-dependent consolidation of stable long-term potentiation lasting months in the Hippocampus. *J. Neurosci.* 22, 9626–9634. doi:10.1523/jneurosci.22-21-09626.2002.

Adamsky, A., Kol, A., Kreisel, T., Doron, A., Ozeri-Engelhard, N., Melcer, T., et al. (2018). Astrocytic Activation Generates De Novo Neuronal Potentiation and Memory Enhancement. *Cell* 174, 59–71. doi:10.1016/j.cell.2018.05.002.

Agilent technologies Inc. (2015). Agilent Feature Extraction 11.5 Reference Guide. 328. Available at: https://www.agilent.com/cs/library/usermanuals/Public/G4460-90045_FE_RefGuide.pdf [Accessed March 16, 2021].

Agulhon, C., Fiacco, T. A., and McCarthy, K. D. (2010). Hippocampal short- and long-term plasticity are not modulated by astrocyte Ca2+ signaling. *Science* 327, 1250–1254. doi:10.1126/science.1184821.

Ahmed, N. I., Thompson, C., Bolnick, D. I., and Stuart, Y. E. (2017). Brain morphology of the threespine stickleback (*Gasterosteus aculeatus*) varies inconsistently with respect to habitat complexity: A test of the Clever Foraging Hypothesis. *Ecol. Evol.* 7, 3372–3380. doi:10.1002/ece3.2918.

Alvarado-Kristensson, M. (2018). γ-tubulin as a signal-transducing molecule and meshwork with therapeutic potential. *Signal Transduct. Target. Ther.* 3, 1–6. doi:10.1038/s41392-018-0021-x.

Anggono, V., and Huganir, R. L. (2012). Regulation of AMPA receptor trafficking and synaptic plasticity. *Curr. Opin. Neurobiol.* 22, 461–469. doi:10.1016/j.conb.2011.12.006.

Arechavala-Lopez, P., Caballero-Froilán, J. C., Jiménez-García, M., Capó, X., Tejada, S., Saraiva, J. L., et al. (2020). Enriched environments enhance cognition, exploratory behaviour and brain physiological functions of *Sparus aurata*. *Sci. Rep.* 10. doi:10.1038/s41598-020-68306-6.

Arostegui, M. C., Hovel, R. A., and Quinn, T. P. (2018). *Schistocephalus solidus* parasite prevalence and biomass intensity in threespine stickleback vary by habitat and diet in boreal lakes. *Environ. Biol. Fishes* 101, 501–514. doi:10.1007/s10641-018-0719-1.

Ashton, B. J., Ridley, A. R., Edwards, E. K., and Thornton, A. (2018). Cognitive performance is linked to group size and affects fitness in Australian magpies. *Nature* 554, 364–367. doi:10.1038/nature25503.

Axelrod, C. J., Laberge, F., and Robinson, B. W. (2021). Interspecific and intraspecific comparisons reveal the importance of evolutionary context in sunfish brain form divergence. *J. Evol. Biol.* 34, 639–652. doi:10.1111/jeb.13763.

Balla, T. (2013). Phosphoinositides: Tiny lipids with giant impact on cell regulation. *Physiol. Rev.* 93, 1019–1137. doi:10.1152/physrev.00028.2012.

Banerjee, S., Maity, S., Guchhait, R., Chatterjee, A., Biswas, C., Adhikari, M., et al. (2021). Toxic effects of cyanotoxins in teleost fish: A comprehensive review. *Aquat. Toxicol.* 240, 105971. doi:10.1016/j.aquatox.2021.105971.

Bateman, A., Martin, M. J., Orchard, S., Magrane, M., Agivetova, R., Ahmad, S., et al. (2021). UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Res.* 49, D480–D489. doi:10.1093/nar/gkaa1100.

Bayley, S. E., and Prather, C. M. (2003). Do wetland lakes exhibit alternative stable states? Submersed aquatic vegetation and chlorophyll in western boreal shallow lakes. *Limnol. Oceanogr.* 48, 2335–2345. doi:10.4319/lo.2003.48.6.2335.

Bebus, S. E., Small, T. W., Jones, B. C., Elderbrock, E. K., and Schoech, S. J. (2016). Associative learning is inversely related to reversal learning and varies with nestling corticosterone exposure. *Anim. Behav.* 111, 251–260. doi:10.1016/j.anbehav.2015.10.027.

Bell, M. A., and Foster, S. A. (1994). "Introduction To the Evolutionary Biology of the Threespine Stickleback," in *The Evolutionary Biology of the Threespine Stickleback*, eds. M. A. Bell and S. A. Foster (Oxford University Press), 1–27.

Bensky, M. K., and Bell, A. M. (2020). Predictors of individual variation in reversal learning performance in three-spined sticklebacks. *Anim. Cogn.* 23, 925–938. doi:10.1007/s10071-020-01399-8.

Bolnick, D. I., and Ballare, K. M. (2020). Resource diversity promotes among-individual diet variation, but not genomic diversity, in lake stickleback. *Ecol. Lett.* 23, 495–505. doi:10.1111/ele.13448.

Bontems, F., Fish, R. J., Borlat, I., Lembo, F. F., Chocu, S., Chalmel, F. F., et al. (2014). C2orf62 and TTC17 are involved in actin organization and ciliogenesis in zebrafish and human. *PLoS One* 9, 86476. doi:10.1371/journal.pone.0086476.

Brazil, D. P., Yang, Z. Z., and Hemmings, B. A. (2004). Advances in protein kinase B signalling: AKTion on multiple fronts. *Trends Biochem. Sci.* 29, 233–242. doi:10.1016/J.TIBS.2004.03.006.

Brionne, A., Juanchich, A., and Hennequet-Antier, C. (2019). ViSEAGO: A Bioconductor package for clustering biological functions using Gene Ontology and semantic similarity. *BioData Min.* 12. doi:10.1186/s13040-019-0204-1.

Buechel, S. D., Boussard, A., Kotrschal, A., van Der Bijl, W., and Kolm, N. (2018). Brain size affects performance in a reversal-learning test. *Proc. R. Soc. B Biol. Sci.* 285. doi:10.1098/rspb.2017.2031.

Buechel, S. D., Noreikiene, K., DeFaveri, J., Toli, E., Kolm, N., and Merilä, J. (2019). Variation in sexual brain size dimorphism over the breeding cycle in the three-spined stickleback. *J. Exp. Biol.* 222. doi:10.1242/JEB.194464.

Carbia, P. S., and Brown, C. (2019). Environmental enrichment influences spatial learning ability in captive-reared intertidal gobies (*Bathygobius cocosensis*). *Anim. Cogn.* 22, 89–98. doi:10.1007/s10071-018-1225-8.

Charnov, E. L. (1976). Optimal Foraging, the Marginal Value Theorem. *Theor. Popul. Biol.* 9, 129–136.

Chen, L., Cummings, K. A., Mau, W., Zaki, Y., Dong, Z., Rabinowitz, S., et al. (2020). The role of intrinsic excitability in the evolution of memory: Significance in memory allocation, consolidation, and updating. *Neurobiol. Learn. Mem.* 173, 107266. doi:10.1016/j.nlm.2020.107266.

Chen, X., Garelick, M. G., Wang, H., Li, V., Athos, J., and Storm, D. R. (2005). PI3 kinase signaling is required for retrieval and extinction of contextual memory. *Nat. Neurosci.* 8, 925–931. doi:10.1038/nn1482.

Ciofi-Baffoni, S., Nasta, V., and Banci, L. (2018). Protein networks in the maturation of human iron–sulfur proteins. *Metallomics* 10, 49–72. doi:10.1039/C7MT00269F.

Cornell, J., Salinas, S., Huang, H. Y., and Zhou, M. (2022). Microglia regulation of synaptic plasticity and learning and memory. *Neural Regen. Res.* 17, 705–716. doi:10.4103/1673-5374.322423.

Costa-Mattioli, M., Sossin, W. S., Klann, E., and Sonenberg, N. (2009). Translational Control of Long-Lasting Synaptic Plasticity and Memory. *Neuron* 61, 10–26. doi:10.1016/j.neuron.2008.10.055.

Croston, R., Branch, C. L., Pitera, A. M., Kozlovsky, D. Y., Bridge, E. S., Parchman, T. L., et al. (2017). Predictably harsh environment is associated with reduced cognitive

flexibility in wild food-caching mountain chickadees. *Anim. Behav.* 123, 139–149. doi:10.1016/j.anbehav.2016.10.004.

Dall, S. R. X., Giraldeau, L.-A., Olsson, O., McNamara, J. M., and Stephens, D. W. (2005). Information and its use by animals in evolutionary ecology. *Trends Ecol. Evol.* 20, 187–93. doi:10.1016/j.tree.2005.01.010.

Delmotte, Q., Hamze, M., Medina, I., Buhler, E., Zhang, J., Belgacem, Y. H., et al. (2020). Smoothened receptor signaling regulates the developmental shift of GABA polarity in rat somatosensory cortex. *J. Cell Sci.* 133. doi:10.1242/jcs.247700.

Demski, L. S. (2013). The pallium and mind/behavior relationships in teleost fishes. in *Brain, Behavior and Evolution* (Karger Publishers), 31–44. doi:10.1159/000351994.

Dunbar, R. I. M. (1998). The social brain hypothesis. *Evol. Anthropol. Issues, News, Rev.* 6, 178–190. doi:10.1002/(SICI)1520-6505(1998)6:5<178::AID-EVAN5>3.3.CO;2-P.

Dupret, D., Fabre, A., Döbrössy, M. D., Panatier, A., Rodríguez, J. J., Lamarque, S., et al. (2007). Spatial learning depends on both the addition and removal of new hippocampal neurons. *PLoS Biol.* 5, 1683–1694. doi:10.1371/journal.pbio.0050214.

Emlen, J. M. (1966). The Role of Time and Energy in Food Preference. *Am. Nat.* 100, 611–617. doi:10.1086/282455.

Evans, M. L., Hori, T. S., Rise, M. L., and Fleming, I. A. (2015). Transcriptomic Responses of Atlantic Salmon (*Salmo salar*) to Environmental Enrichment during Juvenile Rearing. *PLoS One* 10, e0118378. doi:10.1371/journal.pone.0118378.

Falcicchia, C., Tozzi, F., Arancio, O., Watterson, D. M., and Origlia, N. (2020). Involvement of p38 MAPK in Synaptic Function and Dysfunction. *Int. J. Mol. Sci.* 21, 5624. doi:10.3390/ijms21165624.

Fang, B., Kemppainen, P., Momigliano, P., Feng, X., and Merilä, J. (2020). On the causes of geographically heterogeneous parallel evolution in sticklebacks. *Nat. Ecol. Evol.* 4, 1105–1115. doi:10.1038/s41559-020-1222-6.

Fong, S., Buechel, S. D., Boussard, A., Kotrschal, A., and Kolm, N. (2019). Plastic changes in brain morphology in relation to learning and environmental enrichment in the guppy (*Poecilia reticulata*). *J. Exp. Biol.* 222. doi:10.1242/JEB.200402.

Foster, S. A., Shaw, K. A., Robert, K. L., and Baker, J. A. (2008). Benthic, limnetic and oceanic threespine stickleback: Profiles of reproductive behaviour. *Behaviour* 145, 485–508. doi:10.1163/156853908792451421.

Frankland, P. W., and Bontempi, B. (2005). The organization of recent and remote memories. *Nat. Rev. Neurosci.* 6, 119–130. doi:10.1038/nrn1607.

Frey, U., and Morris, R. G. M. (1998). Synaptic tagging: Implications for late maintenance of hippocampal long- term potentiation. *Trends Neurosci.* 21, 181–188. doi:10.1016/S0166-2236(97)01189-2.

Galili, T., O'Callaghan, A., Sidi, J., and Sievert, C. (2018). heatmaply: an R package for creating interactive cluster heatmaps for online publishing. *Bioinformatics* 34, 1600–1602. doi:10.1093/BIOINFORMATICS/BTX657.

Geven, E. J. W., and Klaren, P. H. M. (2017). The teleost head kidney: Integrating thyroid and immune signalling. *Dev. Comp. Immunol.* 66, 73–83. doi:10.1016/j.dci.2016.06.025.

Gittleman, J. L. (1986). Carnivore Brain Size, Behavioral Ecology, and Phylogeny. *J. Mammal.* 67, 23–36. doi:10.2307/1380998.

Gonda, A., Herczeg, G., and Merilä, J. (2009). Adaptive brain size divergence in nine-spined sticklebacks (*Pungitius pungitius*)? *J. Evol. Biol.* 22, 1721–1726. doi:10.1111/j.1420-9101.2009.01782.x.

Gonda, A., Herczeg, G., and Merilä, J. (2013). Evolutionary ecology of intraspecific brain size variation: a review. *Ecol. Evol.* 3, 2751–2764. doi:10.1002/ece3.627.

Gonzalez-Voyer, A., Winberg, S., and Kolm, N. (2009). Social fishes and single mothers: brain evolution in African cichlids. *Proc. R. Soc. B Biol. Sci.* 276, 161–167. doi:10.1098/rspb.2008.0979.

Grecias, L., Hebert, F. O., Alves, V. A., Barber, I., and Aubin-Horth, N. (2020). Host behaviour alteration by its parasite: from brain gene expression to functional test. *Proc. R. Soc. B Biol. Sci.* 287. doi:10.1098/rspb.2020.2252.

Greenwood, A. K., and Peichel, C. L. (2015). Social regulation of gene expression in threespine sticklebacks. *PLoS One* 10. doi:10.1371/JOURNAL.PONE.0137726.

Guderley, H. E. (1994). "Physiological ecology and evolution in the threespine stickleback," in *The Evolutionary Biology of the threespine stickleback*, eds. M. A. Bell and S. A. Foster (Oxford University Press), 85–113.

Guemez-Gamboa, A., Coufal, N. G., and Gleeson, J. G. (2014). Primary cilia in the developing and mature brain. *Neuron* 82, 511. doi:10.1016/J.NEURON.2014.04.024.

Guitton, H., Blackwell, D., and Girshick, M. A. (1955). Theory of Games and Statistical Decisions. *Rev. économique* 6, 508. doi:10.2307/3497265.

Gygax, M., Rentsch, A. K., Rudman, S. M., and Rennison, D. J. (2018). Differential predation alters pigmentation in threespine stickleback (*Gasterosteus aculeatus*). *J. Evol. Biol.* 31, 1589–1598. doi:10.1111/jeb.13354.

Han, S., Nam, J., Li, Y., Kim, S., Cho, S.-H., Cho, Y. S., et al. (2010). Regulation of Dendritic Spines, Spatial Memory, and Embryonic Development by the TANC Family of PSD-95-Interacting Proteins. *J. Neurosci.* 30, 15102–15112. doi:10.1523/JNEUROSCI.3128-10.2010.

Härer, A., Bolnick, D. I., and Rennison, D. J. (2021). The genomic signature of ecological divergence along the benthic-limnetic axis in allopatric and sympatric threespine stickleback. *Mol. Ecol.* 30, 451–463. doi:10.1111/mec.15746.

Heit, B., Tavener, S., Raharjo, E., and Kubes, P. (2002). An intracellular signaling hierarchy determines direction of migration in opposing chemotactic gradients. *J. Cell Biol.* 159, 91–102. doi:10.1083/jcb.200202114.

Henke-von der Malsburg, J., Kappeler, P. M., and Fichtel, C. (2020). Linking ecology and cognition: does ecological specialisation predict cognitive test performance? *Behav. Ecol. Sociobiol.* 74. doi:10.1007/s00265-020-02923-z.

Herczeg, G., Urszán, T. J., Orf, S., Nagy, G., Kotrschal, A., and Kolm, N. (2019). Brain size predicts behavioural plasticity in guppies (*Poecilia reticulata*): An experiment. *J. Evol. Biol.* 32, 218–226. doi:10.1111/jeb.13405.

Herold, C., Schlömer, P., Mafoppa-Fomat, I., Mehlhorn, J., Amunts, K., and Axer, M. (2019). The hippocampus of birds in a view of evolutionary connectomics. *Cortex* 118, 165–187. doi:10.1016/j.cortex.2018.09.025.

Hill, M. S., Vande Zande, P., and Wittkopp, P. J. (2020). Molecular and evolutionary processes generating variation in gene expression. *Nat. Rev. Genet.*, 1–13. doi:10.1038/s41576-020-00304-w.

Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A., and Cresko, W. A. (2010). Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genet.* 6, e1000862. doi:10.1371/journal.pgen.1000862.

Holtmaat, A., and Svoboda, K. (2009). Experience-dependent structural synaptic plasticity in the mammalian brain. *Nat. Rev. Neurosci.* 10, 647–658. doi:10.1038/nrn2699.

Howe, K. L., Achuthan, P., Allen, J., Allen, J., Alvarez-Jarreta, J., Ridwan Amode, M., et al. (2021). Ensembl 2021. *Nucleic Acids Res.* 49, D884–D891. doi:10.1093/nar/gkaa942.

Ishikawa, Y., Kitaoka, S., Kawano, Y., Ishii, S., Suzuki, T., Wakahashi, K., et al. (2021). Repeated social defeat stress induces neutrophil mobilization in mice: maintenance after cessation of stress and strain-dependent difference in response. *Br. J. Pharmacol.* 178, 827–844. doi:10.1111/BPH.15203.

Johansen, I. B., Sørensen, C., Sandvik, G. K., Nilsson, G. E., Höglund, E., Bakken, M., et al. (2012). Neural plasticity is affected by stress and heritable variation in stress coping style. *Comp. Biochem. Physiol. Part D Genomics Proteomics* 7, 161–171. doi:10.1016/J.CBD.2012.01.002.

Jones, F. C., Grabherr, M. G., Chan, Y. F., Russell, P., Mauceli, E., Johnson, J., et al. (2012). The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* 484, 55–61. doi:10.1038/nature10944.

Jones, J. R., Bell, M. A., Baker, J. A., and Koenings, J. P. (2003). General limnology of lakes near cook inlet, southcentral alaska. *Lake Reserv. Manag.* 19, 141–149. doi:10.1080/07438140309354080.

Kalbe, M., and Kurtz, J. (2006). Local differences in immunocompetence reflect resistance of sticklebacks against the eye fluke Diplostomum pseudospathaceum. *Parasitology* 132, 105–116. doi:10.1017/S0031182005008681.

Kapitein, L. C., and Hoogenraad, C. C. (2015). Building the Neuronal Microtubule Cytoskeleton. *Neuron* 87, 492–506. doi:10.1016/J.NEURON.2015.05.046.

Karve, A. D., Von Hippel, F. A., and Bell, M. A. (2008). Isolation between sympatric anadromous and resident threespine stickleback species in Mud Lake, Alaska. *Environ. Biol. Fishes* 81, 287–296. doi:10.1007/s10641-007-9200-2.

Keagy, J., Braithwaite, V. A., and Boughman, J. W. (2018). Brain differences in ecologically differentiated sticklebacks. *Curr. Zool.* 64, 243–250. doi:10.1093/cz/zox074.

Khacho, M., and Slack, R. S. (2018). Mitochondrial dynamics in the regulation of neurogenesis: From development to the adult brain. *Dev. Dyn.* 247, 47–53. doi:10.1002/DVDY.24538.

Kim, Y. R., Kim, Y. M., Lee, J., Park, J., Lee, J. E., and Hyun, Y.-M. (2020). Neutrophils Return to Bloodstream Through the Brain Blood Vessel After Crosstalk With Microglia During LPS-Induced Neuroinflammation. *Front. Cell Dev. Biol.* 8. doi:10.3389/FCELL.2020.613733.

Kol, A., and Goshen, I. (2021). The memory orchestra: the role of astrocytes and oligodendrocytes in parallel to neurons. *Curr. Opin. Neurobiol.* 67, 131–137. doi:10.1016/j.conb.2020.10.022.

Korotkevich, G., Sukhov, V., Budin, N., Shpak, B., Artyomov, M., and Sergushichev, A. (2021). Fast gene set enrichment analysis. *bioRxiv*, 060012. doi:10.1101/060012.

Kotrschal, A., Räsänen, K., Kristjánsson, B. K., Senn, M., and Kolm, N. (2012). Extreme sexual brain size dimorphism in sticklebacks: A consequence of the cognitive challenges of sex and parenting? *PLoS One* 7. doi:10.1371/journal.pone.0030055.

Kotrschal, K., Staaden, M. J. Van, and Huber, R. (1998). Fish brains: evolution and environmental relationships. *Rev. Fish Biol. Fish.* 8, 373–408. doi:10.1023/A:1008839605380.

Kozak, G. M., and Boughman, J. W. (2008). Experience influences shoal member preference in a species pair of sticklebacks. *Behav. Ecol.* 19, 667–676. doi:10.1093/beheco/arn022.

Larhammar, M., Huntwork-Rodriguez, S., Rudhard, Y., Sengupta-Ghosh, A., and Lewcock, J. W. (2017). The Ste20 family kinases MAP4K4, MINK1, and TNIK converge to regulate stress-induced JNK signaling in neurons. *J. Neurosci.* 37, 11074–11084. doi:10.1523/JNEUROSCI.0905-17.2017.

Lasarge, C. L., and Danzer, S. C. (2014). Mechanisms regulating neuronal excitability and seizure development following mTOR pathway hyperactivation. *Front. Mol. Neurosci.* 7. doi:10.3389/fnmol.2014.00018.

Leavell, B. C., and Bernal, X. E. (2019). The Cognitive Ecology of Stimulus Ambiguity: A Predator–Prey Perspective. *Trends Ecol. Evol.* 34, 1048–1060. doi:10.1016/j.tree.2019.07.004.

Leder, E. H., Merilä, J., and Primmer, C. R. (2009). A flexible whole-genome microarray for transcriptomics in three-spine stickleback (*Gasterosteus aculeatus*). *BMC Genomics* 10, 426. doi:10.1186/1471-2164-10-426.

Lee, P. R., Cohen, J. E., Iacobas, D. A., Iacobas, S., and Fields, R. D. (2017). Gene networks activated by specific patterns of action potentials in dorsal root ganglia neurons. *Sci. Rep.* 7. doi:10.1038/srep43765.

Lema, S. C., Hodges, M. J., Marchetti, M. P., and Nevitt, G. A. (2005). Proliferation zones in the salmon telencephalon and evidence for environmental influence on proliferation rate. *Comp. Biochem. Physiol. Part A Mol. Integr. Physiol.* 141, 327–335. doi:10.1016/j.cbpb.2005.06.003.

Lenz, T. L., Eizaguirre, C., Rotter, B., Kalbe, M., and Milinski, M. (2013). Exploring local immunological adaptation of two stickleback ecotypes by experimental infection and transcriptome-wide digital gene expression analysis. *Mol. Ecol.* 22, 774–786. doi:10.1111/j.1365-294X.2012.05756.x.

Leveelahti, L., Leskinen, P., Leder, E. H., Waser, W., and Nikinmaa, M. (2011). Responses of threespine stickleback (*Gasterosteus aculeatus*, L) transcriptome to hypoxia. *Comp. Biochem. Physiol. - Part D Genomics Proteomics* 6, 370–381. doi:10.1016/j.cbd.2011.08.001.

Lin, W., Hung, T.-C., Kurobe, T., Wang, Y., and Yang, P. (2021). Microcystin-Induced Immunotoxicity in Fishes: A Scoping Review. *Toxins (Basel).* 13, 765. doi:10.3390/toxins13110765.

Lipton, J. O., and Sahin, M. (2014). The Neurology of mTOR. *Neuron* 84, 275–291. doi:10.1016/J.NEURON.2014.09.034.

Liu, Y., Jones, C. D., Day, L. B., Summers, K., and Burmeister, S. S. (2020). Cognitive Phenotype and Differential Gene Expression in a Hippocampal Homologue in Two Species of Frog. *Integr. Comp. Biol.* 60, 1007–1023. doi:10.1093/icb/icaa032.

Lodge, K. M., Cowburn, A. S., Li, W., and Condliffe, A. M. (2020). The impact of hypoxia on neutrophil degranulation and consequences for the host. *Int. J. Mol. Sci.* 21, 1183. doi:10.3390/ijms21041183.

MacArthur, R. H., and Pianka, E. R. (1966). On Optimal Use of a Patchy Environment. *Am. Nat.* 100, 603–609. doi:10.1086/282454.

MacColl, A. D. C. (2009). Parasite burdens differ between sympatric three-spined stickleback species. *Ecography (Cop.).* 32, 153–160. doi:10.1111/j.1600-0587.2008.05486.x.

Malenka, R. C., and Bear, M. F. (2004). LTP and LTD: An embarrassment of riches. *Neuron* 44, 5–21. doi:10.1016/j.neuron.2004.09.012.

Malone, J. H., and Oliver, B. (2011). Microarrays, deep sequencing and the true measure of the transcriptome. *BMC Biol.* 9. doi:10.1186/1741-7007-9-34.

Mantione, K. J., Kream, R. M., Kuzelova, H., Ptacek, R., Raboch, J., Samuel, J. M., et al. (2014). Comparing bioinformatic gene expression profiling methods: microarray and RNA-Seq. *Med. Sci. Monit. Basic Res.* 20, 138–142. doi:10.12659/MSMBR.892101.

Marco, A., Meharena, H. S., Dileep, V., Raju, R. M., Davila-Velderrain, J., Zhang, A. L., et al. (2020). Mapping the epigenomic and transcriptomic interplay during memory formation and recall in the hippocampal engram ensemble. *Nat. Neurosci.* 23, 1606–1617. doi:10.1038/s41593-020-00717-0.

Martinez, J., Keagy, J., Wurst, B., Fetzner, W., and Boughman, J. W. (2016). The relative roles of genes and rearing environment on the spatial cognitive ability of two sympatric species of threespine stickleback. *Evol. Ecol. Res.* 17, 565–581.

McNamara, J., and Houston, A. (1980). The application of statistical decision theory to animal behaviour. *J. Theor. Biol.* 85, 673–690. doi:10.1016/0022-5193(80)90265-9.

McPhail, J. D. (1994). "Speciation and the evolution of reproductive isolation in the sticklebacks (Gasterosteus) of south-western British Columbia," in *The evolutionary biology of the threespine stickleback*, eds. M. A. Bell and S. A. Foster (Oxford University Press), 399–437.

Medina, A. C., Torres-García, M. E., Rodríguez-Serrano, L. M., Bello-Medina, P. C., Quirarte, G. L., McGaugh, J. L., et al. (2019). Inhibition of transcription and translation in dorsal hippocampus does not interfere with consolidation of memory of intense training. *Neurobiol. Learn. Mem.* 166, 107092. doi:10.1016/j.nlm.2019.107092.

Mendoza, M. C., Er, E. E., and Blenis, J. (2011). The Ras-ERK and PI3K-mTOR pathways: cross-talk and compensation. *Trends Biochem. Sci.* 36, 320–328. doi:10.1016/j.tibs.2011.03.006.

Miningou, N., and Blackwell, K. T. (2020). The road to ERK activation: Do neurons take alternate routes? *Cell. Signal.* 68, 109541. doi:10.1016/j.cellsig.2020.109541.

Moran, R., Harvey, I., Moss, B., Feuchtmayr, H., Hatton, K., Heyes, T., et al. (2010). Influence of simulated climate change and eutrophication on three-spined stickleback populations: a large scale mesocosm experiment. *Freshw. Biol.* 55, 315–325. doi:10.1111/J.1365-2427.2009.02276.X.

Morris, M. R. J., Richard, R., Leder, E. H., Barrett, R. D. H., Aubin-Horth, N., and Rogers, S. M. (2014). Gene expression plasticity evolves in response to colonization of freshwater lakes in threespine stickleback. *Mol. Ecol.* 23, 3226–3240. doi:10.1111/mec.12820.

Navarrete, M., Cuartero, M. I., Palenzuela, R., Draffin, J. E., Konomi, A., Serra, I., et al. (2019). Astrocytic p38α MAPK drives NMDA receptor-dependent long-term depression and modulates long-term memory. *Nat. Commun.* 10. doi:10.1038/s41467-019-10830-9.

Niven, J. E., and Laughlin, S. B. (2008). Energy limitation as a selective pressure on the evolution of sensory systems. *J. Exp. Biol.* 211, 1792–1804. doi:10.1242/jeb.017574.

Noreikiene, K., Herczeg, G., Gonda, A., Balázs, G., Husby, A., and Merilä, J. (2015). Quantitative genetic analysis of brain size variation in sticklebacks: Support for the mosaic model of brain evolution. *Proc. R. Soc. B Biol. Sci.* 282. doi:10.1098/RSPB.2015.1008.

Odling-Smee, L. C., Boughman, J. W., and Braithwaite, V. A. (2008). Sympatric species of threespine stickleback differ in their performance in a spatial learning task. *Behav. Ecol. Sociobiol. 2008 6212* 62, 1935–1945. doi:10.1007/S00265-008-0625-1.

Oh, W. C., and Smith, K. R. (2019). Activity-dependent development of GABAergic synapses. *Brain Res.* 1707, 18–26. doi:10.1016/J.BRAINRES.2018.11.014.

Olsson, O., and Brown, J. S. (2010). Smart, smarter, smartest: foraging information states and coexistence. *Oikos* 119, 292–303. doi:10.1111/j.1600-0706.2009.17784.x.

Øverli, Ø., and Sørensen, C. (2016). On the Role of Neurogenesis and Neural Plasticity in the Evolution of Animal Personalities and Stress Coping Styles. *Brain. Behav. Evol.* 87, 167–174. doi:10.1159/000447085.

Pan, S., Mayoral, S. R., Choi, H. S., Chan, J. R., and Kheirbek, M. A. (2020). Preservation of a remote fear memory requires new myelin formation. *Nat. Neurosci.* 23, 487–499. doi:10.1038/s41593-019-0582-1.

Park, P. J. (2013). Spatial learning ability of the threespine stickleback (Gasterosteus aculeatus) in relation to inferred ecology and ancestry. *Evol. Ecol. Res.* 15, 213–239.

Park, P. J., and Bell, M. A. (2010). Variation of telencephalon morphology of the threespine stickleback (Gasterosteus aculeatus) in relation to inferred ecology. *J. Evol. Biol.* 23, 1261–1277. doi:10.1111/j.1420-9101.2010.01987.x.

Park, P. J., Chase, I., and Bell, M. A. (2012). Phenotypic plasticity of the threespine stickleback Gasterosteus aculeatus telencephalon in response to experience in captivity. *Curr. Zool.* 58, 189–210. doi:10.1093/czoolo/58.1.189.

Park, S., Kramer, E. E., Mercaldo, V., Rashid, A. J., Insel, N., Frankland, P. W., et al. (2016). Neuronal Allocation to a Hippocampal Engram. *Neuropsychopharmacology* 41, 2987–2993. doi:10.1038/npp.2016.73.

Parker, S. T., and Gibson, K. R. (1977). Object manipulation, tool use and sensorimotor intelligence as feeding adaptations in cebus monkeys and great apes. *J. Hum. Evol.* 6, 623–641. doi:10.1016/S0047-2484(77)80135-8.

Parsons, R. G. (2018). Behavioral and neural mechanisms by which prior experience impacts subsequent learning. *Neurobiol. Learn. Mem.* 154, 22–29. doi:https://doi.org/10.1016/j.nlm.2017.11.008.

Piecyk, A., Hahn, M. A., Roth, O., Dheilly, N. M., Heins, D. C., Bell, M. A., et al. (2021). Cross-continental experimental infections reveal distinct defence mechanisms in populations of the three-spined stickleback Gasterosteus aculeatus. *Proc. R. Soc. B* 288. doi:10.1098/RSPB.2021.1758.

Podgorniak, T., Milan, M., Pujolar, J. M., Maes, G. E., Bargelloni, L., De Oliveira, E., et al. (2015). Differences in brain gene transcription profiles advocate for an important role of cognitive function in upstream migration and water obstacles crossing in European eel. *BMC Genomics* 16. doi:10.1186/s12864-015-1589-y.

Pollen, A. A. (2007). Environmental complexity and social organization sculpt the brain in Lake Tanganyikan cichlid fish. *Brain. Behav. Evol.* 70, 21–39. doi:DOI: 10.1159/000101067.

Pravosudov, V. V., Roth, T. C., Forister, M. L., Ladage, L. D., Kramer, R., Schilkey, F., et al. (2013). Differential hippocampal gene expression is associated with climate-related natural variation in memory and the hippocampus in food-caching chickadees. *Mol. Ecol.* 22, 397–408. doi:10.1111/mec.12146.

R Core Team (2014). R: A Language and Environment for Statistical Computing.

Rampon, C., Jiang, C. H., Dong, H., Tang, Y. P., Lockhart, D. J., Schultz, P. G., et al. (2000). Effects of environmental enrichment on gene expression in the brain. *Proc. Natl. Acad. Sci. U. S. A.* 97, 12880–12884. doi:10.1073/pnas.97.23.12880.

Rane, M. J., Gozal, D., Butt, W., Gozal, E., Pierce, W. M., Guo, S. Z., et al. (2005). γ-Amino Butyric Acid Type B Receptors Stimulate Neutrophil Chemotaxis during Ischemia-Reperfusion. *J. Immunol.* 174, 7242–7249. doi:10.4049/JIMMUNOL.174.11.7242.

Resnik, P. (1995). Using Information Content to Evaluate Semantic Similarity in a Taxonomy. in *Proceedings of the 14th international joint conference on Artificial intelligence - Volume 1 (IJCAI'95)*, ed. C. S. Mellish (San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.), 448–453.

Rosso, S. B., and Inestrosa, N. C. (2013). WNT signalling in neuronal maturation and synaptogenesis. *Front. Cell. Neurosci.* 7. doi:10.3389/fncel.2013.00103.

Rowe, C., and Healy, S. D. (2014). Measuring variation in cognition. *Behav. Ecol.* 25, 1287–1292. doi:10.1093/beheco/aru090.

Rowland, W. J. (1994). "Proximate determinants of stickleback behavior: an evolutionary perspective," in *The Evolutionary Biology of the threespine stickleback*, eds. M. A. Bell and S. A. Foster (Oxford University Press), 297–344.

Rundle, H. D. (2000). Natural Selection and Parallel Speciation in Sympatric Sticklebacks. *Science (80-.).* 287, 306–308. doi:10.1126/science.287.5451.306.

Salas, C., Broglio, C., and Rodríguez, F. (2003). Evolution of forebrain and spatial cognition in vertebrates: Conservation across diversity. *Brain. Behav. Evol.* 62, 72–82. doi:10.1159/000072438.

Salvanes, A. G. V., Moberg, O., Ebbesson, L. O. E., Nilsen, T. O., Jensen, K. H., and Braithwaite, V. A. (2013). Environmental enrichment promotes neural plasticity and cognitive ability in fish. *Proc. R. Soc. B Biol. Sci.* 280, 20131331. doi:10.1098/rspb.2013.1331.

Samuk, K., Iritani, D., and Schluter, D. (2014). Reversed brain size sexual dimorphism accompanies loss of parental care in white sticklebacks. *Ecol. Evol.* 4, 3236–3243. doi:10.1002/ece3.1175.

Savolainen, O., Lascoux, M., and Merilä, J. (2013). Ecological genomics of local adaptation. *Nat. Rev. Genet.* 14, 807–820. doi:10.1038/nrg3522.

Shaw, K. A., Scotti, M. L., and Foster, S. A. (2007). Ancestral plasticity and the evolutionary diversification of courtship behaviour in threespine sticklebacks. *Anim. Behav.* 73, 415–422. doi:10.1016/j.anbehav.2006.09.002.

Sol, D., Bacher, S., Reader, S. M., and Lefebvre, L. (2008). Brain size predicts the success of mammal species introduced into novel environments. *Am. Nat.* 172, s63–s71. doi:10.1086/588304.

Squire, L. R., Stark, C. E. L., and Clark, R. E. (2004). The medial temporal lobe. *Annu. Rev. Neurosci.* 27, 279–306. doi:10.1146/annurev.neuro.27.070203.144130.

Stanford, B. C. M., Clake, D. J., Morris, M. R. J., and Rogers, S. M. (2020). The power and limitations of gene expression pathway analyses toward predicting population response to environmental stressors. *Evol. Appl.* 13, 1166–1182. doi:10.1111/eva.12935.

Steadman, P. E., Xia, F., Ahmed, M., Mocle, A. J., Penning, A. R. A., Geraghty, A. C., et al. (2020). Disruption of Oligodendrogenesis Impairs Memory Consolidation in Adult Mice. *Neuron* 105, 150–164. doi:10.1016/j.neuron.2019.10.013.

Stickleback assembly and gene annotation (2010). *Gasterosteus_aculeatus - Ensembl genome Brows. 103*. Available at: http://www.ensembl.org/Gasterosteus_aculeatus/Info/Annotation [Accessed March 9, 2021].

Stutz, W. E., Lau, O. L., and Bolnick, D. I. (2014). Contrasting patterns of phenotype-dependent parasitism within and among populations of threespine stickleback. *Am. Nat.* 183, 810–825. doi:10.1086/676005.

Stutz, W. E., Schmerer, M., Coates, J. L., and Bolnick, D. I. (2015). Among-lake reciprocal transplants induce convergent expression of immune genes in threespine stickleback. *Mol. Ecol.* 24, 4629–4646. doi:10.1111/mec.13295.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 102, 15545–15550. doi:10.1073/pnas.0506580102.

Svanbäck, R., and Schluter, D. (2012). Niche Specialization Influences Adaptive Phenotypic Plasticity in the Threespine Stickleback. *Am. Nat.* 180, 50–59. doi:10.1086/666000.

Tamayo, A.-M. P., Devineau, O., Præbel, K., Kahilainen, K. K., and Østbye, K. (2020). A brain and a head for a different habitat: Size variation in four morphs of Arctic charr (*Salvelinus alpinus* (L.)) in a deep oligotrophic lake. *Ecol. Evol.* 10, 11335–11351. doi:10.1002/ECE3.6771.

Tonegawa, S., Morrissey, M. D., and Kitamura, T. (2018). The role of engram cells in the systems consolidation of memory. *Nat. Rev. Neurosci.* 19, 485–498. doi:10.1038/s41583-018-0031-2.

Tracey, T. J., Steyn, F. J., Wolvetang, E. J., and Ngo, S. T. (2018). Neuronal lipid metabolism: Multiple pathways driving functional outcomes in health and disease. *Front. Mol. Neurosci.* 11, 10. doi:10.3389/fnmol.2018.00010.

TRI Reagent® Solution Protocol (PN 9738M Rev D) (2010). Available at: https://assets.thermofisher.com/TFS-Assets/LSG/manuals/9738M_D.pdf [Accessed March 12, 2021].

Tsuboi, M., Husby, A., Kotrschal, A., Hayward, A., Buechel, S. D., Zidar, J., et al. (2015). Comparative support for the expensive tissue hypothesis: Big brains are correlated with smaller gut and greater parental investment in Lake Tanganyika cichlids. *Evolution (N. Y).* 69, 190–200. doi:10.1111/evo.12556.

Two-Color Microarray-Based Gene Expression Analysis (Low Input Quick Amp Labeling) Protocol (2015). Available at: https://www.agilent.com/cs/library/usermanuals/Public/G4140-90050_GeneExpression_TwoColor_6.9.pdf [Accessed March 16, 2021].

Tyssowski, K. M., DeStefino, N. R., Cho, J. H., Dunn, C. J., Poston, R. G., Carty, C. E., et al. (2018). Different Neuronal Activity Patterns Induce Different Gene Expression Programs. *Neuron* 98, 530–546. doi:10.1016/j.neuron.2018.04.001.

Urban, M. C., Strauss, S. Y., Pelletier, F., Palkovacs, E. P., Leibold, M. A., Hendry, A. P., et al. (2020). Evolutionary origins for ecological patterns in space. *Proc. Natl. Acad. Sci. U. S. A.* 117, 17482–17490. doi:10.1073/pnas.1918960117.

Vamosi, S. M., and Schluter, D. (2004). Character shifts in the defensive armor of sympatric sticklebacks. *Evolution (N. Y).* 58, 376–385. doi:10.1111/j.0014-3820.2004.tb01653.x.

Vilas, M. P., Marti, C. L., Oldham, C. E., and Hipsey, M. R. (2018). Macrophyte-induced thermal stratification in a shallow urban lake promotes conditions suitable for nitrogen-fixing cyanobacteria. *Hydrobiologia* 806, 411–426. doi:10.1007/s10750-017-3376-z.

von Krogh, K., Sørensen, C., Nilsson, G. E., and Øverli, Ø. (2010). Forebrain cell proliferation, behavior, and physiology of zebrafish, *Danio rerio*, kept in enriched or barren environments. *Physiol. Behav.* 101, 32–39. doi:10.1016/j.physbeh.2010.04.003.

Wang, C., Wang, L., and Gu, Y. (2021a). Microglia, synaptic dynamics and forgetting. *Brain Res. Bull.* 174, 173–183. doi:10.1016/J.BRAINRESBULL.2021.06.005.

Wang, Y., Fu, W. Y., Cheung, K., Hung, K. W., Chen, C., Geng, H., et al. (2021b). Astrocyte-secreted IL-33 mediates homeostatic synaptic plasticity in the adult hippocampus. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2020810118. doi:10.1073/pnas.2020810118.

Ward, J. H. (1963). Hierarchical Grouping to Optimize an Objective Function. *J. Am. Stat. Assoc.* 58, 236. doi:10.2307/2282967.

White, G. E., and Brown, C. (2015). Microhabitat Use Affects Brain Size and Structure in Intertidal Gobies. *Brain. Behav. Evol.* 85, 107–116. doi:http://dx.doi.org/10.1159/000380875.

Wood, L. S., Desjardins, J. K., and Fernald, R. D. (2011). Effects of stress and motivation on performing a spatial task. *Neurobiol. Learn. Mem.* 95, 277–285. doi:10.1016/J.NLM.2010.12.002.

Wootton, R. J. (1994). “Energy allocation in the threespine stickleback,” in *The Evolutionary Biology of the Threespine Stickleback*, eds. M. A. Bell and S. Forster (Oxford University Press), 116–140.

Yap, E. L., and Greenberg, M. E. (2018). Activity-Regulated Transcription: Bridging the Gap between Neural Activity and Behavior. *Neuron* 100, 330–348. doi:10.1016/j.neuron.2018.10.013.

Yates, A. D., Achuthan, P., Akanni, W., Allen, J., Allen, J., Alvarez-Jarreta, J., et al. (2020). Ensembl 2020. *Nucleic Acids Res.* 48, D682–D688. doi:10.1093/nar/gkz966.

Yekutieli, D., and Benjamini, Y. (2001). The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* 29, 1165–1188. doi:10.1214/aos/1013699998.

Yopak, K. E., Lisney, T. J., Collin, S. P., and Montgomery, J. C. (2007). Variation in Brain Organization and Cerebellar Foliation in Chondrichthyans: Sharks and Holocephalans. *Brain. Behav. Evol.* 69, 280–300. doi:10.1159/000100037.

Yue, M., Luo, D., Liu, S. Y. P., Zhou, Q., Hu, M., Liu, Y., et al. (2016). Misshapen/NIK-related kinase (mink1) is involved in platelet function, hemostasis, and thrombus formation. *Blood* 127, 927–937. doi:10.1182/blood-2015-07-659185.

Zheng, F., Luo, Y., and Wang, H. (2009). Regulation of BDNF-mediated transcription of immediate early gene Arc by intracellular calcium and calmodulin. *J. Neurosci. Res.* 87, 380. doi:10.1002/JNR.21863.

Zhou, W., Cao, L., Jeffries, J., Zhu, X., Staiger, C. J., and Deng, Q. (2018). Neutrophil-specific knockout demonstrates a role for mitochondria in regulating neutrophil motility in zebrafish. *Dis. Model. Mech.* 11. doi:10.1242/dmm.033027.

Zhu, Y., Pak, D., Qin, Y., Mccormack, S. G., Kim, M. J., Baumgart, J. P., et al. (2005). Rap2-JNK removes synaptic AMPA receptors during depotentiation. *Neuron* 46, 905–916. doi:10.1016/j.neuron.2005.04.037.

Zupanc, G. K. H., and Sîrbulescu, R. F. (2011). Adult neurogenesis and neuronal regeneration in the central nervous system of teleost fish. *Eur. J. Neurosci.* 34, 917–929. doi:10.1111/j.1460-9568.2011.07854.x.

Appendices

Appendix 1. RNA isolation protocol

1000 µl Tri-reageant (Ambion, Applied Biosystems) were added to a sample immediately after removing it from dry ice. A carbide bead was added to the tube, and the tissue was homogenized by shaking at 30 shakes/second in Tissue Lyser (Qiagen) for 60 seconds, in 30 second cycles, until completely homogenized. The bead was removed by magnet and the lysis was incubated at room temperature for 10 minutes to ensure homogenization. Because of the small amount and soluble nature of the tissue, the additional centrifugation phase was not necessary.

100 µl Bromo-3-chloropropane (Sigma) was added to lysis and mixed by vigorous shaking by hand for 15 seconds followed by a 10-minute incubation in room temperature. The mixture was centrifuged in +4°C at 12,000 x g for 15 minutes.

At the centrifugation, the mixture was separated into three layers, with the middle, aqueous, layer containing RNA and lower, red, phase DNA and proteins. The aqueous layer was carefully transferred to a fresh tube by pipetting, avoiding phase mixture. 500 µl isopropanol was added to aqueous solution and mixed by shaking, followed by a 10 - minute incubation in room temperature. The samples were centrifuged for 15 minutes at 12,000 x g in +4°C, to precipitate and collect the RNA.

After centrifugation, the supernatant was discarded, and the pellet was washed twice with 1 ml of 75 % ethanol. After the second wash, the samples were incubated in 75 % ethanol in +4°C over night. The ethanol was removed from the samples by pouring after centrifugation in +4°C at 7,500 x g for 5 minutes. The residual ethanol was removed by pipetting with fine tip and brief drying on a paper towel, until no traces of ethanol were visible, avoiding the complete drying of the pellet to ensure RNA solubility.

DNAse treatment

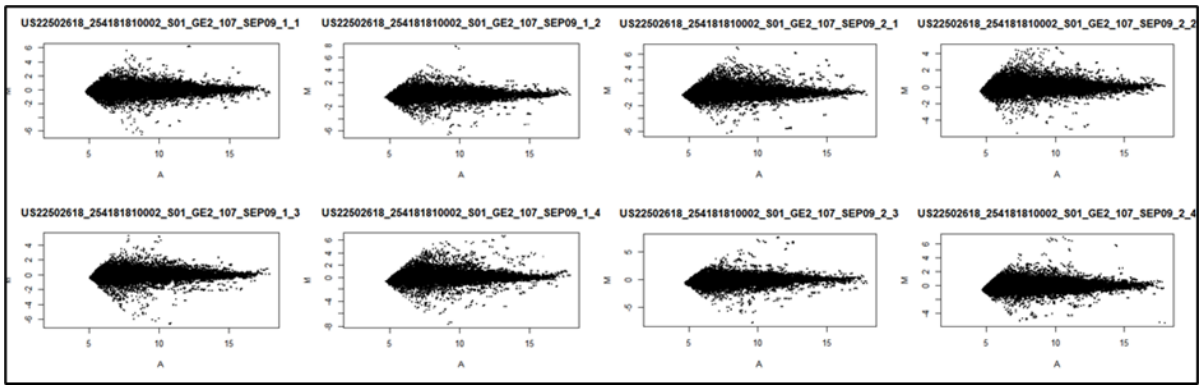
The master mix of 2.5 µl enzyme and 3 µl 10 x reaction buffer were prepared. The precipitated RNA was dissolved in 25 µl nuclease free H_2O on ice. 5.5 µl master mix was added and mixed by careful pipetting.

The samples were incubated in a 37°C water for 45 minutes and cooled on ice.

Appendix 2. Microarray design and MA-plots

Microarray design

Array: GacV4_2012, ID:041818		
Cy5	**Cy3**	**Barcode**
Mud 3	Lynne 1	25418180002 1-1
Tern 4	South Rolly 1	25418180002 1-2
Lynne 4	Mud 4	25418180002 1-3
South Rolly 4	Tern 5	25418180002 1-4
Tern 2	Lynne 2	25418180002 2-1
Mud 2	South Rolly 2	25418180002 2-2
Lynne 3	Tern 3	25418180002 2-3
South Rolly 5	Mud 1	25418180002 2-4



Microarray MA-plots.

Appendix 3. Differentially expressed genes and their human orthologs.

ProbeName	Log FC	padj	Gasterosteus aculeatus Ensembl ID	Human ortholog Ensembl ID	Human ortholog gene symbol
Gac_32274_PI222095114	4.51	0.0007	ENSGACG00000020394	ENSG00000141503	MINK1
Gac_53227_PI225017757	-2.37	0.0020	ENSGACG00000016535	ENSG00000080573	COL5A3
Gac_25502_PI225017757	-1.86	0.0020	ENSGACG00000012847	ENSG00000158428	CATIP
Gac_40437_PI225017757	2.59	0.0020	ENSGACG00000020395	ENSG00000172354	GNB2
Gac_15926_PI225017757	-1.63	0.0030	ENSGACG00000008096	ENSG00000139684	ESD
Gac_16650_PI225017757	-1.91	0.0034	ENSGACG00000019330	ENSG00000100347	SAMM50
Gac_46919_PI225017757	-2.62	0.0062	ENSGACG00000020819	ENSG00000048162	NOP16
Gac_11884_PI225017757	-2.48	0.0062	ENSGACG00000017135	ENSG00000130055	GDPD2
Gac_6186_PI225017757	2.12	0.0062	ENSGACG00000018808	ENSG00000144036	EXOC6B
Gac_15447_PI225017757	1.96	0.0074	ENSGACG00000011818	ENSG00000125814	NAPB
Gac_2808_PI225017757	1.45	0.0079	ENSGACG00000009998	ENSG00000171862	PTEN
Gac_28462_PI225017757	2.24	0.0085	ENSGACG00000019865	ENSG00000120860	WASHC3
Gac_44710_PI225017757	-1.42	0.0085	ENSGACG00000016430	ENSG00000140612	SEC11A
Gac_6455_PI225017757	-1.26	0.0085	ENSGACG00000008154	ENSG00000149313	AASDHPPT
Gac_27832_PI225017757	-1.85	0.0085	ENSGACG00000016021	ENSG00000196358	NTNG2
Gac_52854_PI222095114	-3.74	0.0086	ENSGACG00000009554	ENSG00000103316	CRYM
Gac_39320_PI225017757	-1.79	0.0086	ENSGACG00000001170	ENSG00000196584	XRCC2
Gac_38489_PI225017757	-1.52	0.0090	ENSGACG00000018280	ENSG00000106723	SPIN1
Gac_47680_PI225017757	1.17	0.0100	ENSGACG00000016607	ENSG00000106246	PTCD1
Gac_39550_PI225017757	-1.17	0.0100	ENSGACG00000018088	ENSG00000242852	ZNF709

ProbeName	Log FC	padj	Gasterosteus aculeatus Ensembl ID	Human ortholog Ensembl ID	Human ortholog gene symbol
Gac_24528_PI222095114	1.37	0.0103	ENSGACG00000007021	ENSG00000157851	DPYSL5
Gac_51475_PI225017757	-1.88	0.0115	ENSGACG00000020288	ENSG00000035687	ADSS2
Gac_43044_PI225017757	1.99	0.0115	ENSGACG00000007602	ENSG00000119686	FLVCR2
Gac_23428_PI225017757	-1.26	0.0128	ENSGACG00000013296	ENSG00000137692	DCUN1D5
Gac_54100_PI225017757	1.18	0.0144	ENSGACG00000014142	ENSG00000053900	ANAPC4
Gac_3404_PI225017757	-1.59	0.0144	ENSGACG00000012321	ENSG00000104824	HNRNPL
Gac_52175_PI225017757	-1.31	0.0144	ENSGACG00000015229	ENSG00000114107	CEP70
Gac_49223_PI225017757	-1.42	0.0144	ENSGACG00000004936	ENSG00000115282	TTC31
Gac_6546_PI225017757	-1.47	0.0144	ENSGACG00000011844	ENSG00000151135	TMEM263
Gac_T09453xG07120_753	1.40	0.0144	ENSGACG00000007120	ENSG00000164182	NDUFAF2
Gac_3967_PI225017757	-1.46	0.0144	ENSGACG00000010417	ENSG00000166128	RAB8B
Gac_29870_PI225017757	-1.92	0.0144	ENSGACG00000004903	ENSG00000169032	MAP2K1
Gac_40893_PI225017757	1.58	0.0144	ENSGACG00000018932	ENSG00000184672	RALYL
Gac_26794_PI222095114	-1.18	0.0147	ENSGACG00000016450	ENSG00000198586	TLK1
Gac_43969_PI225017757	0.98	0.0149	ENSGACG00000007296	ENSG00000198408	OGA
Gac_27306_PI222095114	-7.74	0.0169	ENSGACG00000015818	ENSG00000115386	REG1A
Gac_34558_PI222095114	-1.48	0.0171	ENSGACG00000020284	ENSG00000162188	GNG3
Gac_40410_PI225017757	2.58	0.0172	ENSGACG00000011359	ENSG00000172366	MCRIP2
Gac_6999_PI225017757	-1.02	0.0174	ENSGACG00000004234	ENSG00000169727	GPS1
Gac_5956_PI225017757	-1.25	0.0194	ENSGACG00000020505	ENSG00000189143	CLDN4
Gac_21488_PI225017757	1.25	0.0196	ENSGACG00000003218	ENSG00000064601	CTSA

ProbeName	Log FC	padj	Gasterosteus aculeatus Ensembl ID	Human ortholog Ensembl ID	Human ortholog gene symbol
Gac_Gene_T23700_2215	-1.06	0.0196	ENSGACG00000017901	ENSG00000242419	PCDHGC4
Gac_12079_PI225017757	2.23	0.0197	ENSGACG00000015943	ENSG00000051523	CYBA
Gac_22618_PI225017757	1.82	0.0197	ENSGACG00000010118	ENSG00000071553	ATP6AP1
Gac_9438_PI225017757	-1.12	0.0197	ENSGACG00000019329	ENSG00000072778	ACADVL
Gac_27698_PI225017757	1.49	0.0197	ENSGACG00000007011	ENSG00000084764	MAPRE3
Gac_46173_PI225017757	-1.41	0.0197	ENSGACG00000006249	ENSG00000087460	GNAS
Gac_T10167xG07609_4805	-1.13	0.0197	ENSGACG00000007609	ENSG00000116852	KIF21B
Gac_5808_PI225017757	-1.09	0.0197	ENSGACG00000007747	ENSG00000117632	STMN1
Gac_463_PI222095114	-1.10	0.0197	ENSGACG00000012623	ENSG00000117877	POLR1G
Gac_137_PI222095114	-1.91	0.0197	ENSGACG00000005981	ENSG00000125952	MAX
Gac_33481_PI225017757	-2.64	0.0197	ENSGACG00000010775	ENSG00000128886	ELL3
Gac_36919_PI225017757	0.96	0.0197	ENSGACG00000017521	ENSG00000135632	SMYD5
Gac_54235_PI225017757	-0.93	0.0197	ENSGACG00000005068	ENSG00000138162	TACC2
Gac_42912_PI225017757	0.90	0.0197	ENSGACG00000010813	ENSG00000151413	NUBPL
Gac_26283_PI225017757	-1.26	0.0197	ENSGACG00000008117	ENSG00000162813	BPNT1
Gac_T20445xG15460_607	-1.21	0.0197	ENSGACG00000015460	ENSG00000163864	NMNAT3
Gac_T27086xG20444_5636	1.14	0.0197	ENSGACG00000020444	ENSG00000173442	EHBP1L1
Gac_42837_PI225017757	1.36	0.0197	ENSGACG00000010841	ENSG00000181704	YIPF6
Gac_16985_PI225017757	-0.94	0.0197	ENSGACG00000001149	ENSG00000187416	LHFPL3
Gac_23096_PI222095114	-0.98	0.0197	ENSGACG00000000467	ENSG00000257727	CNPY2
Gac_34391_PI225017757	-0.87	0.0197	ENSGACG00000011044	ENSG00000274386	TMEM269

ProbeName	Log FC	padj	Gasterosteus aculeatus Ensembl ID	Human ortholog Ensembl ID	Human ortholog gene symbol
Gac_34965_PI225017757	-1.28	0.0205	ENSGACG00000013509	ENSG00000108602	ALDH3A1
Gac_23507_PI225017757	1.35	0.0205	ENSGACG00000013295	ENSG00000109089	CDR2L
Gac_41226_PI225017757	1.07	0.0205	ENSGACG00000017747	ENSG00000156170	NDUFAF6
Gac_26599_PI225017757	1.27	0.0205	ENSGACG00000014987	ENSG00000182173	TSEN54
Gac_54605_PI225017757	-0.98	0.0205	ENSGACG00000001974	ENSG00000253729	PRKDC
Gac_40496_PI225017757	-1.75	0.0205	ENSGACG00000020388	ENSG00000106384	MOGAT3
Gac_2992_PI222095114	1.83	0.0205	ENSGACG00000013619	ENSG00000167536	DHRS13
Gac_54661_PI222095114	-2.13	0.0208	ENSGACG00000012888	ENSG00000080824	HSP90AA1
Gac_19543_PI225017757	-1.02	0.0208	ENSGACG00000000750	ENSG00000113851	CRBN
Gac_54911_PI225017757	-1.40	0.0208	ENSGACG00000003928	ENSG00000128159	TUBGCP6
Gac_33929_PI222095114	1.40	0.0208	ENSGACG00000016157	ENSG00000137094	DNAJB5
Gac_2306_PI225017757	-1.06	0.0208	ENSGACG00000012991	ENSG00000137275	RIPK1
Gac_22805_PI225017757	-1.08	0.0208	ENSGACG00000006602	ENSG00000155111	CDK19
Gac_1446_PI225017757	1.08	0.0208	ENSGACG00000018583	ENSG00000165029	ABCA1
Gac_52822_PI225017757	-3.45	0.0208	ENSGACG00000007077	ENSG00000241563	CORT
Gac_10332_PI225017757	-1.15	0.0210	ENSGACG00000018256	ENSG00000147234	FRMPD3
Gac_38119_PI225017757	-1.04	0.0210	ENSGACG00000016401	ENSG00000107897	ACBD5
Gac_15067_PI222095114	-1.22	0.0210	ENSGACG00000010415	ENSG00000156795	NTAQ1
Gac_50463_PI222095114	1.24	0.0210	ENSGACG00000000950	ENSG00000164794	KCNV1
Gac_11438_PI225017757	-1.17	0.0210	ENSGACG00000009527	ENSG00000176171	BNIP3
Gac_6190_PI225017757	-2.20	0.0214	ENSGACG00000018805	ENSG00000149476	TKFC

ProbeName	Log FC	padj	Gasterosteus aculeatus Ensembl ID	Human ortholog Ensembl ID	Human ortholog gene symbol
Gac_27217_PI222095114	0.81	0.0215	ENSGACG00000006384	ENSG00000001084	GCLC
Gac_34613_PI225017757	1.01	0.0215	ENSGACG00000010619	ENSG00000080823	MOK
Gac_45842_PI225017757	1.35	0.0215	ENSGACG00000013081	ENSG00000100749	VRK1
Gac_37433_PI225017757	-1.66	0.0215	ENSGACG00000011396	ENSG00000105058	FAM32A
Gac_28890_PI225017757	-1.25	0.0215	ENSGACG00000000064	ENSG00000132182	NUP210
Gac_6815_PI225017757	1.27	0.0215	ENSGACG00000000623	ENSG00000151490	PTPRO
Gac_30511_PI225017757	1.47	0.0215	ENSGACG00000011930	ENSG00000157077	ZFYVE9
Gac_15402_PI225017757	1.26	0.0215	ENSGACG00000011815	ENSG00000163421	PROK2
Gac_34511_PI225017757	-1.10	0.0215	ENSGACG00000010635	ENSG00000186687	LYRM7
Gac_23035_PI225017757	-1.02	0.0215	ENSGACG00000000492	ENSG00000204463	BAG6
Gac_698_PI225017757	-1.40	0.0215	ENSGACG00000017643	ENSG00000014138	POLA2
Gac_16589_PI222095114	0.75	0.0215	ENSGACG00000004853	ENSG00000131844	MCCC2
Gac_44606_PI222095114	-0.82	0.0215	ENSGACG00000017580	ENSG00000185085	INTS5
Gac_43143_PI225017757	-1.39	0.0216	ENSGACG00000005197	ENSG00000006625	GGCT
Gac_T00532xG00416_431	1.51	0.0216	ENSGACG00000000416	ENSG00000080371	RAB21
Gac_6498_PI225017757	-0.98	0.0216	ENSGACG00000020900	ENSG00000137500	CCDC90B
Gac_Gene_07735	-0.99	0.0221	ENSGACG00000007735	ENSG00000177302	TOP3A
Gac_52894_PI225017757	1.29	0.0221	ENSGACG00000013288	ENSG00000186111	PIP5K1C
Gac_45396_PI225017757	-0.88	0.0221	ENSGACG00000013213	ENSG00000231824	AKAIN1
Gac_18920_PI222095114	-2.01	0.0227	ENSGACG00000020758	ENSG00000011143	MKS1
Gac_2855_PI225017757	1.95	0.0227	ENSGACG00000009990	ENSG00000124818	OPN5

ProbeName	Log FC	padj	Gasterosteus aculeatus Ensembl ID	Human ortholog Ensembl ID	Human ortholog gene symbol
Gac_5555_PI225017757	-0.88	0.0229	ENSGACG00000010339	ENSG00000119801	YPEL5
Gac_32436_PI225017757	-1.44	0.0229	ENSGACG00000005331	ENSG00000177830	CHID1
Gac_47830_PI225017757	1.01	0.0231	ENSGACG00000005761	ENSG00000131467	PSME3
Gac_28991_PI225017757	3.32	0.0234	ENSGACG00000018138	ENSG00000171587	DSCAM
Gac_24883_PI225017757	-1.05	0.0242	ENSGACG00000001958	ENSG00000050393	MCUR1
Gac_48659_PI225017757	-0.77	0.0242	ENSGACG00000001042	ENSG00000112081	SRSF3
Gac_29298_PI225017757	-1.07	0.0242	ENSGACG00000016291	ENSG00000125821	DTD1
Gac_49604_PI225017757	0.87	0.0242	ENSGACG00000005124	ENSG00000159363	ATP13A2
Gac_47557_PI225017757	-0.69	0.0242	ENSGACG00000010222	ENSG00000162755	KLHDC9
Gac_10503_PI225017757	1.01	0.0242	ENSGACG00000007738	ENSG00000257103	LSM14A
Gac_27791_PI225017757	-0.74	0.0248	ENSGACG00000013536	ENSG00000119707	RBM25
Gac_T06996xG05264_904	-2.56	0.0248	ENSGACG00000005264	ENSG00000135437	RDH5
Gac_40364_PI222095114	0.92	0.0248	ENSGACG00000009122	ENSG00000138231	DBR1
Gac_31286_PI225017757	1.16	0.0248	ENSGACG00000013420	ENSG00000167889	MGAT5B
Gac_30294_PI225017757	-1.98	0.0251	ENSGACG00000018447	ENSG00000155858	LSM11
Gac_22139_PI222095114	-0.87	0.0251	ENSGACG00000018458	ENSG00000171169	NAIF1
Gac_47121_PI225017757	1.00	0.0251	ENSGACG00000010251	ENSG00000198829	SUCNR1
Gac_46690_PI225017757	1.35	0.0256	ENSGACG00000008392	ENSG00000164252	AGGF1
Gac_1621_PI222095114	-0.94	0.0256	ENSGACG00000018561	ENSG00000115163	CENPA
Gac_11429_PI225017757	-1.65	0.0256	ENSGACG00000009536	ENSG00000162702	ZNF281
Gac_42136_PI225017757	1.00	0.0256	ENSGACG00000014382	ENSG00000182108	DEXI

ProbeName	Log FC	padj	Gasterosteus aculeatus Ensembl ID	Human ortholog Ensembl ID	Human ortholog gene symbol
Gac_3314_PI222095114	-1.26	0.0259	ENSGACG00000016703	ENSG00000204438	GPANK1
Gac_39130_PI225017757	1.71	0.0260	ENSGACG00000019975	ENSG00000006459	KDM7A
Gac_T15334xG11552_1371	-1.06	0.0260	ENSGACG00000011552	ENSG00000167553	TUBA1C
Gac_49294_PI222095114	-0.95	0.0260	ENSGACG00000019396	ENSG00000170296	GABARAP
Gac_21164_PI225017757	-0.98	0.0260	ENSGACG00000008942	ENSG00000177030	DEAF1
Gac_10352_PI222095114	-1.45	0.0261	ENSGACG00000018175	ENSG00000113328	CCNG1
Gac_8716_PI225017757	-1.14	0.0265	ENSGACG00000011734	ENSG00000185418	TARS3
Gac_32127_PI225017757	-1.12	0.0267	ENSGACG00000000858	ENSG00000003756	RBM5
Gac_38942_PI222095114	-1.21	0.0267	ENSGACG00000018248	ENSG00000091157	WDR7
Gac_11576_PI222095114	1.01	0.0267	ENSGACG00000009858	ENSG00000105655	ISYNA1
Gac_29543_PI225017757	-0.93	0.0267	ENSGACG00000013187	ENSG00000112282	MED23
Gac_5048_PI222095114	-1.41	0.0267	ENSGACG00000010310	ENSG00000125534	PPDPF
Gac_40946_PI225017757	0.75	0.0267	ENSGACG00000007387	ENSG00000144840	RABL3
Gac_53571_PI225017757	2.35	0.0267	ENSGACG00000000116	ENSG00000153029	MR1
Gac_44249_PI225017757	-1.32	0.0267	ENSGACG00000001248	ENSG00000168237	GLYCTK
Gac_19947_PI225017757	-1.60	0.0267	ENSGACG00000000341	ENSG00000198113	TOR4A
Gac_11182_PI225017757	-1.75	0.0271	ENSGACG00000006959	ENSG00000115828	QPCT
Gac_4970_PI225017757	1.19	0.0273	ENSGACG00000011641	ENSG00000143627	PKLR
Gac_20173_PI225017757	-1.02	0.0275	ENSGACG00000006958	ENSG00000147419	CCDC25
Gac_6095_PI225017757	-0.77	0.0277	ENSGACG00000008438	ENSG00000104883	PEX11G
Gac_53988_PI225017757	0.81	0.0277	ENSGACG00000005632	ENSG00000110756	HPS5

ProbeName	Log FC	padj	Gasterosteus aculeatus Ensembl ID	Human ortholog Ensembl ID	Human ortholog gene symbol
Gac_45855_PI225017757	-0.87	0.0283	ENSGACG00000015461	ENSG00000134987	WDR36
Gac_40870_PI225017757	-0.79	0.0283	ENSGACG00000018936	ENSG00000171311	EXOSC1
Gac_99_PI225017757	0.76	0.0287	ENSGACG00000012089	ENSG00000160323	ADAMTS13
Gac_37554_PI225017757	-1.02	0.0287	ENSGACG00000013783	ENSG00000170571	EMB
Gac_T22298xG16852_2347	-0.85	0.0291	ENSGACG00000016852	ENSG00000099889	ARVCF
Gac_46029_PI225017757	0.73	0.0291	ENSGACG00000004619	ENSG00000118363	SPCS2
Gac_37472_PI225017757	-1.14	0.0291	ENSGACG00000011385	ENSG00000160633	SAFB
Gac_Gene_T04434_242	-1.01	0.0291	ENSGACG00000003355	ENSG00000165169	DYNLT3
Gac_Gene_10285_a_20	1.35	0.0293	ENSGACG00000010285	ENSG00000148688	RPP30
Gac_52026_PI222095114	0.73	0.0293	ENSGACG00000012733	ENSG00000198822	GRM3
Gac_2810_PI225017757	-2.09	0.0294	ENSGACG00000009978	ENSG00000076924	XAB2
Gac_Gene_20432	-1.14	0.0294	ENSGACG00000020432	ENSG00000022267	FHL1
Gac_55099_PI225017757	-1.52	0.0294	ENSGACG00000017478	ENSG00000102393	GLA
Gac_35187_PI225017757	-0.89	0.0294	ENSGACG00000010020	ENSG00000122008	POLK
Gac_29012_PI225017757	-2.06	0.0294	ENSGACG00000000047	ENSG00000133466	C1QTNF6
Gac_43841_PI225017757	-0.79	0.0294	ENSGACG00000007353	ENSG00000144559	TAMM41
Gac_5553_PI225017757	-1.21	0.0294	ENSGACG00000010340	ENSG00000159840	ZYX
Gac_26564_PI225017757	0.92	0.0294	ENSGACG00000017430	ENSG00000189134	NKAPL
Gac_5387_PI225017757	-2.48	0.0298	ENSGACG00000018220	ENSG00000141527	CARD14
Gac_T06417xG04826_4268	-0.94	0.0307	ENSGACG00000004826	ENSG00000077044	DGKD
Gac_34197_PI225017757	1.08	0.0307	ENSGACG00000014965	ENSG00000127824	TUBA4A

ProbeName	Log FC	padj	Gasterosteus aculeatus Ensembl ID	Human ortholog Ensembl ID	Human ortholog gene symbol
Gac_307_PI222095114	-1.46	0.0307	ENSGACG00000003637	ENSG00000132330	SCLY
Gac_9259_PI225017757	1.41	0.0307	ENSGACG00000015399	ENSG00000152213	ARL11
Gac_43622_PI225017757	-1.54	0.0307	ENSGACG00000016526	ENSG00000166851	PLK1
Gac_23544_PI225017757	-1.49	0.0307	ENSGACG00000004091	ENSG00000164645	TEX47
Gac_36324_PI225017757	1.67	0.0310	ENSGACG00000001337	ENSG00000135164	DMTF1
Gac_40029_PI225017757	-1.15	0.0312	ENSGACG00000018162	ENSG00000104765	BNIP3L
Gac_4395_PI225017757	0.97	0.0312	ENSGACG00000013812	ENSG00000111241	FGF6
Gac_50637_PI225017757	-1.32	0.0312	ENSGACG00000006084	ENSG00000124193	SRSF6
Gac_9324_PI225017757	-1.14	0.0312	ENSGACG00000015389	ENSG00000169683	LRRC45
Gac_33359_PI225017757	-1.39	0.0312	ENSGACG00000020663	ENSG00000196218	RYR1
Gac_43036_PI225017757	-1.26	0.0313	ENSGACG00000007591	ENSG00000105507	CABP5
Gac_49062_PI225017757	2.16	0.0313	ENSGACG00000013996	ENSG00000134716	CYP2J2
Gac_46574_PI225017757	-0.71	0.0313	ENSGACG00000017548	ENSG00000138777	PPA2
Gac_23643_PI225017757	-2.35	0.0313	ENSGACG00000013064	ENSG00000139304	PTPRQ
Gac_41174_PI225017757	2.56	0.0313	ENSGACG00000003287	ENSG00000159208	CIART
Gac_35988_PI225017757	-0.76	0.0313	ENSGACG00000018402	ENSG00000173611	SCAI
Gac_38936_PI222095114	-1.17	0.0313	ENSGACG00000006743	ENSG00000184924	PTRHD1
Gac_35132_PI225017757	-0.98	0.0313	ENSGACG00000010024	ENSG00000139154	AEBP2
Gac_13762_PI222095114	1.04	0.0315	ENSGACG00000004145	ENSG00000087470	DNM1L
Gac_14343_PI225017757	-0.77	0.0315	ENSGACG00000017794	ENSG00000129484	PARP2
Gac_29685_PI225017757	1.17	0.0315	ENSGACG00000004128	ENSG00000132199	ENOSF1

ProbeName	Log FC	padj	Gasterosteus aculeatus Ensembl ID	Human ortholog Ensembl ID	Human ortholog gene symbol
Gac_9657_PI225017757	1.40	0.0315	ENSGACG00000008302	ENSG00000149043	SYT8
Gac_3229_PI225017757	-0.98	0.0315	ENSGACG00000016720	ENSG00000181004	BBS12
Gac_40850_PI225017757	-2.42	0.0316	ENSGACG00000018939	ENSG00000167191	GPRC5B
Gac_8545_PI225017757	-0.67	0.0316	ENSGACG00000011804	ENSG00000156256	USP16
Gac_T14264xG10741_1968	-1.17	0.0316	ENSGACG00000010741	ENSG00000150764	DIXDC1
Gac_7006_PI225017757	-0.84	0.0316	ENSGACG00000010949	ENSG00000162873	KLHDC8A
Gac_23381_PI225017757	-0.75	0.0319	ENSGACG00000013313	ENSG00000013375	PGM3
Gac_47144_PI225017757	2.55	0.0320	ENSGACG00000019282	ENSG00000166710	B2M
Gac_43411_PI225017757	-1.70	0.0326	ENSGACG00000001380	ENSG00000032742	IFT88
Gac_19686_PI222095114	0.71	0.0326	ENSGACG00000018791	ENSG00000068654	POLR1A
Gac_4094_PI222095114	-0.79	0.0326	ENSGACG00000010353	ENSG00000090661	CERS4
Gac_42082_PI225017757	0.88	0.0326	ENSGACG00000014386	ENSG00000103274	NUBP1
Gac_5170_PI222095114	-1.60	0.0326	ENSGACG00000002628	ENSG00000148848	ADAM12
Gac_4143_PI225017757	-0.90	0.0326	ENSGACG00000008609	ENSG00000156521	TYSND1
Gac_25688_PI225017757	0.72	0.0326	ENSGACG00000007842	ENSG00000183431	SF3A3
Gac_4285_PI225017757	1.24	0.0327	ENSGACG00000017630	ENSG00000110104	CCDC86
Gac_39547_PI225017757	0.72	0.0329	ENSGACG00000018089	ENSG00000171202	TMEM126A
Gac_54703_PI225017757	0.75	0.0336	ENSGACG00000018960	ENSG00000130638	ATXN10
Gac_T23532xG17768_2549	-0.71	0.0336	ENSGACG00000017768	ENSG00000132692	BCAN
Gac_459_PI225017757	-0.69	0.0339	ENSGACG00000018061	ENSG00000176101	SSNA1
Gac_12062_PI225017757	0.88	0.0340	ENSGACG00000004393	ENSG00000143595	AQP10

ProbeName	Log FC	padj	Gasterosteus aculeatus Ensembl ID	Human ortholog Ensembl ID	Human ortholog gene symbol
Gac_38384_PI225017757	-0.98	0.0348	ENSGACG00000017285	ENSG00000039560	RAI14
Gac_49541_PI225017757	-0.96	0.0348	ENSGACG00000016673	ENSG00000138686	BBS7
Gac_50689_PI222095114	1.07	0.0348	ENSGACG00000020577	ENSG00000166979	EVA1C
Gac_27055_PI225017757	1.18	0.0348	ENSGACG00000015835	ENSG00000151575	TEX9
Gac_3321_PI225017757	1.08	0.0351	ENSGACG00000012340	ENSG00000151702	FLI1
Gac_31529_PI225017757	0.69	0.0351	ENSGACG00000017579	ENSG00000169230	PRELID1
Gac_17609_PI225017757	-1.01	0.0354	ENSGACG00000017363	ENSG00000107020	PLGRKT
Gac_55179_PI225017757	-0.79	0.0368	ENSGACG00000008409	ENSG00000100418	DESI1
Gac_49042_PI225017757	-1.29	0.0368	ENSGACG00000011079	ENSG00000104442	ARMC1
Gac_49878_PI225017757	1.45	0.0368	ENSGACG00000009224	ENSG00000124257	NEURL2
Gac_25316_PI225017757	-1.06	0.0368	ENSGACG00000012971	ENSG00000139044	B4GALNT3
Gac_48643_PI225017757	0.83	0.0369	ENSGACG00000019099	ENSG00000070610	GBA2
Gac_9621_PI222095114	0.83	0.0369	ENSGACG00000002229	ENSG00000132600	PRMT7
Gac_41235_PI222095114	-1.13	0.0371	ENSGACG00000000829	ENSG00000163528	CHCHD4
Gac_30404_PI222095114	1.56	0.0371	ENSGACG00000017996	ENSG00000164270	HTR4
Gac_16216_PI225017757	-1.12	0.0372	ENSGACG00000018710	ENSG00000109158	GABRA4
Gac_23425_PI225017757	1.05	0.0372	ENSGACG00000013304	ENSG00000116251	RPL22
Gac_14320_PI225017757	-0.66	0.0372	ENSGACG00000008765	ENSG00000134779	TPGS2
Gac_33570_PI225017757	0.75	0.0372	ENSGACG00000019825	ENSG00000136021	SCYL2
Gac_8178_PI225017757	-0.69	0.0372	ENSGACG00000000917	ENSG00000138698	RAP1GDS1
Gac_35113_PI225017757	-0.78	0.0372	ENSGACG00000004434	ENSG00000143575	HAX1

ProbeName	Log FC	padj	Gasterosteus aculeatus Ensembl ID	Human ortholog Ensembl ID	Human ortholog gene symbol
Gac_33164_PI222095114	0.83	0.0372	ENSGACG00000013146	ENSG00000187664	HAPLN4
Gac_40871_PI225017757	-0.91	0.0375	ENSGACG00000007400	ENSG00000119929	CUTC
Gac_19093_PI225017757	0.94	0.0376	ENSGACG00000002312	ENSG00000175305	CCNE2
Gac_52378_PI225017757	-0.71	0.0379	ENSGACG00000011903	ENSG00000072756	TRNT1
Gac_13023_PI225017757	1.18	0.0379	ENSGACG00000012149	ENSG00000100731	PCNX1
Gac_39375_PI225017757	-1.04	0.0379	ENSGACG00000013899	ENSG00000102931	ARL2BP
Gac_27030_PI225017757	0.63	0.0379	ENSGACG00000014081	ENSG00000103047	TANGO6
Gac_6574_PI225017757	0.60	0.0379	ENSGACG00000011836	ENSG00000125375	DMAC2L
Gac_25997_PI225017757	-0.75	0.0379	ENSGACG00000004659	ENSG00000125871	MGME1
Gac_20427_PI225017757	1.04	0.0379	ENSGACG00000020325	ENSG00000132510	KDM6B
Gac_28561_PI225017757	-1.08	0.0379	ENSGACG00000016628	ENSG00000136933	RABEPK
Gac_10390_PI225017757	0.95	0.0379	ENSGACG00000002371	ENSG00000138303	ASCC1
Gac_41358_PI225017757	-1.85	0.0379	ENSGACG00000006904	ENSG00000151778	SERP2
Gac_47734_PI225017757	-1.66	0.0379	ENSGACG00000014845	ENSG00000162745	OLFML2B
Gac_19249_PI225017757	-0.83	0.0379	ENSGACG00000007706	ENSG00000164190	NIPBL
Gac_37943_PI225017757	-0.80	0.0379	ENSGACG00000015632	ENSG00000166689	PLEKHA7
Gac_34034_PI222095114	-0.84	0.0379	ENSGACG00000014400	ENSG00000168040	FADD
Gac_16284_PI225017757	1.11	0.0379	ENSGACG00000009678	ENSG00000169641	LUZP1
Gac_33774_PI225017757	-0.89	0.0379	ENSGACG00000014633	ENSG00000188316	ENO4
Gac_16256_PI225017757	-1.60	0.0379	ENSGACG00000009661	ENSG00000189043	NDUFA4
Gac_32981_PI225017757	-1.45	0.0386	ENSGACG00000011096	ENSG00000264364	DYNLL2

ProbeName	Log FC	padj	Gasterosteus aculeatus Ensembl ID	Human ortholog Ensembl ID	Human ortholog gene symbol
Gac_14656_PI222095114	0.67	0.0388	ENSGACG00000016519	ENSG00000128989	ARPP19
Gac_51467_PI225017757	-1.46	0.0389	ENSGACG00000020292	ENSG00000011677	GABRA3
Gac_6276_PI222095114	-3.33	0.0394	ENSGACG00000018765	ENSG00000164305	CASP3
Gac_4299_PI225017757	0.79	0.0394	ENSGACG00000006823	ENSG00000168569	TMEM223
Gac_16041_PI225017757	0.71	0.0395	ENSGACG00000006347	ENSG00000179083	FAM133A
Gac_28048_PI225017757	-0.74	0.0396	ENSGACG00000011536	ENSG00000105251	SHD
Gac_9487_PI225017757	0.82	0.0401	ENSGACG00000019308	ENSG00000179094	PER1
Gac_T16900xG12768_793	-0.97	0.0405	ENSGACG00000012768	ENSG00000104805	NUCB1
Gac_20919_PI225017757	1.80	0.0405	ENSGACG00000002649	ENSG00000152527	PLEKHH2
Gac_12444_PI222095114	-1.22	0.0406	ENSGACG00000019755	ENSG00000111224	PARP11
Gac_53908_PI225017757	1.25	0.0406	ENSGACG00000016466	ENSG00000137273	FOXF2
Gac_5604_PI225017757	-1.10	0.0411	ENSGACG00000010332	ENSG00000166321	NUDT13
Gac_24938_PI225017757	-0.82	0.0411	ENSGACG00000010954	ENSG00000186889	TMEM17
Gac_45693_PI225017757	-0.85	0.0414	ENSGACG00000009458	ENSG00000051620	HEBP2
Gac_T03511xG02673_2093	-1.32	0.0414	ENSGACG00000002673	ENSG00000115415	STAT1
Gac_32145_PI225017757	-0.97	0.0414	ENSGACG00000000859	ENSG00000133878	DUSP26
Gac_27640_PI222095114	-1.03	0.0414	ENSGACG00000011710	ENSG00000158014	SLC30A2
Gac_41905_PI225017757	-2.20	0.0414	ENSGACG00000001667	ENSG00000166848	TERF2IP
Gac_47608_PI225017757	1.13	0.0414	ENSGACG00000010199	ENSG00000171552	BCL2L1
Gac_15353_PI225017757	1.11	0.0414	ENSGACG00000020869	ENSG00000175175	PPM1E
Gac_32857_PI225017757	-0.81	0.0414	ENSGACG00000020179	ENSG00000188375	H3-5

ProbeName	Log FC	padj	Gasterosteus aculeatus Ensembl ID	Human ortholog Ensembl ID	Human ortholog gene symbol
Gac_4066_PI225017757	-0.61	0.0417	ENSGACG00000010389	ENSG00000108510	MED13
Gac_27607_PI225017757	-0.56	0.0417	ENSGACG00000017826	ENSG00000109606	DHX15
Gac_24256_PI225017757	-0.82	0.0417	ENSGACG00000015260	ENSG00000184226	PCDH9
Gac_40257_PI225017757	-0.84	0.0418	ENSGACG00000003886	ENSG00000143179	UCK2
Gac_38321_PI222095114	-0.97	0.0421	ENSGACG00000013826	ENSG00000138764	CCNG2
Gac_50213_PI225017757	0.70	0.0424	ENSGACG00000018871	ENSG00000262660	AC139530.2
Gac_6544_PI225017757	2.11	0.0424	ENSGACG00000011841	ENSG00000265491	RNF115
Gac_35624_PI225017757	1.26	0.0428	ENSGACG00000004345	ENSG00000055332	EIF2AK2
Gac_43517_PI225017757	0.92	0.0428	ENSGACG00000005738	ENSG00000088808	PPP1R13B
Gac_16025_PI225017757	-1.01	0.0428	ENSGACG00000006361	ENSG00000106733	NMRK1
Gac_31435_PI225017757	-0.96	0.0428	ENSGACG00000012151	ENSG00000110318	CEP126
Gac_22750_PI222095114	-0.99	0.0428	ENSGACG00000010073	ENSG00000151640	DPYSL4
Gac_16469_PI222095114	0.90	0.0428	ENSGACG00000001224	ENSG00000157911	PEX10
Gac_19580_PI225017757	0.78	0.0428	ENSGACG00000000309	ENSG00000163938	GNL3
Gac_Gene_T07924_111	1.15	0.0428	ENSGACG00000005975	ENSG00000167996	FTH1
Gac_29761_PI225017757	-0.89	0.0428	ENSGACG00000014037	ENSG00000175643	RMI2
Gac_44767_PI222095114	-1.35	0.0428	ENSGACG00000001975	ENSG00000182162	P2RY8
Gac_24299_PI225017757	-1.65	0.0428	ENSGACG00000003741	ENSG00000198959	TGM2
Gac_47880_PI225017757	1.50	0.0429	ENSGACG00000005778	ENSG00000112977	DAP
Gac_17567_PI225017757	2.53	0.0429	ENSGACG00000002886	ENSG00000140853	NLRC5
Gac_12296_PI225017757	-1.32	0.0429	ENSGACG00000019226	ENSG00000183307	TMEM121B

ProbeName	Log FC	padj	Gasterosteus aculeatus Ensembl ID	Human ortholog Ensembl ID	Human ortholog gene symbol
Gac_31152_PI225017757	-0.64	0.0429	ENSGACG00000019996	ENSG00000197620	EOLA1
Gac_53729_PI222095114	0.98	0.0432	ENSGACG00000010986	ENSG00000169599	NFU1
Gac_1281_PI225017757	-0.82	0.0433	ENSGACG00000004523	ENSG00000100216	TOMM22
Gac_17280_PI222095114	-0.72	0.0433	ENSGACG00000004572	ENSG00000125848	FLRT3
Gac_50175_PI225017757	-1.58	0.0434	ENSGACG00000000821	ENSG00000173914	RBM4B
Gac_37899_PI225017757	-1.59	0.0434	ENSGACG00000004803	ENSG00000115128	SF3B6
Gac_20867_PI225017757	-1.21	0.0434	ENSGACG00000011667	ENSG00000163029	SMC6
Gac_21765_PI225017757	-1.57	0.0435	ENSGACG00000016240	ENSG00000152229	PSTPIP2
Gac_9311_PI225017757	-1.00	0.0438	ENSGACG00000013606	ENSG00000121152	NCAPH
Gac_900_PI225017757	-0.74	0.0438	ENSGACG00000019571	ENSG00000124587	PEX6
Gac_36490_PI225017757	-0.73	0.0438	ENSGACG00000014042	ENSG00000131778	CHD1L
Gac_47439_PI222095114	-0.83	0.0438	ENSGACG00000011573	ENSG00000140521	POLG
Gac_14064_PI225017757	1.51	0.0438	ENSGACG00000018660	ENSG00000145428	RNF175
Gac_42245_PI225017757	-0.65	0.0438	ENSGACG00000005536	ENSG00000151247	EIF4E
Gac_23352_PI225017757	1.16	0.0438	ENSGACG00000013318	ENSG00000160818	GPATCH4
Gac_41239_PI225017757	0.83	0.0438	ENSGACG00000017741	ENSG00000173120	KDM2A
Gac_50548_PI225017757	0.94	0.0438	ENSGACG00000007860	ENSG00000182768	NGRN
Gac_47383_PI225017757	-1.28	0.0438	ENSGACG00000002579	ENSG00000187475	H1-6
Gac_49618_PI225017757	0.89	0.0438	ENSGACG00000005108	ENSG00000197746	PSAP
Gac_45862_PI225017757	-0.78	0.0439	ENSGACG00000015463	ENSG00000136169	SETDB2
Gac_39336_PI225017757	-2.41	0.0442	ENSGACG00000001165	ENSG00000106327	TFR2

ProbeName	Log FC	padj	Gasterosteus aculeatus Ensembl ID	Human ortholog Ensembl ID	Human ortholog gene symbol
Gac_42763_PI225017757	2.22	0.0445	ENSGACG00000018734	ENSG00000110063	DCPS
Gac_23541_PI222095114	0.75	0.0448	ENSGACG00000009598	ENSG00000108312	UBTF
Gac_8079_PI222095114	2.17	0.0448	ENSGACG00000009911	ENSG00000163947	ARHGEF3
Gac_7842_PI225017757	0.84	0.0449	ENSGACG00000008196	ENSG00000159685	CHCHD6
Gac_41744_PI225017757	-0.62	0.0450	ENSGACG00000016968	ENSG00000109472	CPE
Gac_T21148xG15996_1209	-1.72	0.0450	ENSGACG00000015996	ENSG00000183569	SERHL2
Gac_45113_PI225017757	1.23	0.0453	ENSGACG00000007869	ENSG00000053770	AP5M1
Gac_13549_PI225017757	-1.61	0.0454	ENSGACG00000010655	ENSG00000160563	MED27
Gac_16991_PI225017757	0.93	0.0455	ENSGACG00000001145	ENSG00000102003	SYP
Gac_Gene_T01558_571	4.96	0.0457	ENSGACG00000001198	ENSG00000179144	GIMAP7
Gac_12016_PI225017757	0.91	0.0460	ENSGACG00000015955	ENSG00000103248	MTHFSD
Gac_Gene_03206_633	0.86	0.0460	ENSGACG00000003206	ENSG00000112531	QKI
Gac_24372_PI225017757	1.07	0.0460	ENSGACG00000008909	ENSG00000136270	TBRG4
Gac_2076_PI225017757	-1.03	0.0462	ENSGACG00000012291	ENSG00000188612	SUMO2
Gac_21902_PI225017757	-0.80	0.0463	ENSGACG00000018509	ENSG00000101843	PSMD10
Gac_20362_PI225017757	-0.80	0.0465	ENSGACG00000011692	ENSG00000154240	CEP112
Gac_2690_PI225017757	-0.63	0.0467	ENSGACG00000008279	ENSG00000118432	CNR1
Gac_29495_PI225017757	1.51	0.0467	ENSGACG00000001439	ENSG00000151693	ASAP2
Gac_26051_PI225017757	-0.70	0.0467	ENSGACG00000004663	ENSG00000170779	CDCA4
Gac_286_PI225017757	-0.87	0.0467	ENSGACG00000012734	ENSG00000197077	KIAA1671
Gac_49072_PI225017757	2.03	0.0467	ENSGACG00000014009	ENSG00000204305	AGER

ProbeName	Log FC	padj	Gasterosteus aculeatus Ensembl ID	Human ortholog Ensembl ID	Human ortholog gene symbol
Gac_25443_PI225017757	-0.98	0.0467	ENSGACG00000003827	ENSG00000204392	LSM2
Gac_6857_PI225017757	-0.63	0.0473	ENSGACG00000013358	ENSG00000131943	C19orf12
Gac_T17650xG13329_1515	-0.60	0.0476	ENSGACG00000013329	ENSG00000066044	ELAVL1
Gac_20515_PI225017757	-0.69	0.0476	ENSGACG00000011255	ENSG00000100154	TTC28
Gac_41448_PI225017757	-0.83	0.0476	ENSGACG00000013037	ENSG00000204256	BRD2
Gac_42727_PI225017757	1.23	0.0478	ENSGACG00000009724	ENSG00000141337	ARSG
Gac_21679_PI222095114	-0.92	0.0482	ENSGACG00000008584	ENSG00000109072	VTN
Gac_Gene_04780b	-2.18	0.0482	ENSGACG00000004780	ENSG00000114554	PLXNA1
Gac_274_PI225017757	-0.94	0.0482	ENSGACG00000012738	ENSG00000139977	NAA30
Gac_11678_PI225017757	1.54	0.0482	ENSGACG00000015312	ENSG00000143153	ATP1B1
Gac_26818_PI225017757	-1.66	0.0482	ENSGACG00000017237	ENSG00000147274	RBMX
Gac_10155_PI225017757	-0.91	0.0482	ENSGACG00000002950	ENSG00000198198	SZT2
Gac_33631_PI225017757	0.56	0.0482	ENSGACG00000019812	ENSG00000213465	ARL2
Gac_21867_PI222095114	0.73	0.0483	ENSGACG00000010760	ENSG00000163754	GYG1
Gac_15553_PI225017757	-0.79	0.0483	ENSGACG00000002186	ENSG00000082258	CCNT2
Gac_49207_PI225017757	0.66	0.0483	ENSGACG00000015744	ENSG00000113163	CERT1
Gac_7030_PI225017757	0.83	0.0483	ENSGACG00000010947	ENSG00000173141	MRPL57
Gac_44179_PI225017757	-1.36	0.0487	ENSGACG00000002399	ENSG00000145491	ROPN1L
Gac_20950_PI222095114	0.97	0.0487	ENSGACG00000003016	ENSG00000144834	TAGLN3
Gac_14062_PI225017757	-2.87	0.0491	ENSGACG00000018668	ENSG00000166801	FAM111A
Gac_28203_PI225017757	-0.81	0.0496	ENSGACG00000001725	ENSG00000172007	RAB33B

ProbeName	Log FC	padj	Gasterosteus aculeatus Ensembl ID	Human ortholog Ensembl ID	Human ortholog gene symbol
Gac_51945_PI225017757	-0.73	0.0496	ENSGACG00000008449	ENSG00000196876	SCN8A
Gac_18132_PI225017757	1.17	0.0499	ENSGACG00000002021	ENSG00000111834	RSPH4A

Appendix 4. FGSEA enriched biological processes and molecular functions

Negative NES indicates enrichment in limnetic ecotype, positive NES in benthic ecotype. Gene frequency is the ratio of genes expressed in this study in an enriched GO term versus all the genes annotated to the GO term in the Ensembl data base.

GO class	GO cluster	GO ID	GOterm	padj	NES	genes frequency
BP	1	GO:0006629	lipid metabolic process	0.25	1.32	27.052% (89/329)
BP	1	GO:0006661	phosphatidylinositol biosynthetic process	0.22	1.64	28.571% (14/49)
BP	1	GO:0006665	sphingolipid metabolic process	0.26	1.65	28.571% (10/35)
BP	1	GO:0006687	glycosphingolipid metabolic process	0.26	1.69	32.143% (9/28)
BP	1	GO:0006646	phosphatidylethanolamine biosynthetic process	0.27	1.74	50% (6/12)
BP	1	GO:0001676	long-chain fatty acid metabolic process	0.25	1.78	66.667% (8/12)
BP	2	GO:0051301	cell division	0.26	-1.39	26.57% (55/207)
BP	2	GO:0006749	glutathione metabolic process	0.25	1.73	21.429% (6/28)
BP	2	GO:0016226	iron-sulfur cluster assembly	0.05	2.04	36.842% (7/19)
BP	2	GO:0007339	binding of sperm to zona pellucida	0.14	1.90	61.538% (8/13)
BP	3	GO:0071897	DNA biosynthetic process	0.25	-1.69	40.909% (9/22)
BP	3	GO:0000381	regulation of alternative mRNA splicing. via spliceosome	0.10	-1.74	39.474% (15/38)
BP	3	GO:0016070	RNA metabolic process	0.25	-1.64	45.161% (14/31)
BP	3	GO:0042795	snRNA transcription by RNA polymerase II	0.26	-1.57	35.185% (19/54)
BP	3	GO:0006260	DNA replication	0.09	-1.70	29.07% (25/86)
BP	3	GO:0006366	transcription by RNA polymerase II	0.26	-1.42	33.036% (37/112)
BP	3	GO:0006357	regulation of transcription by RNA polymerase II	0.25	-1.28	24.696% (122/494)
BP	3	GO:0022900	electron transport chain	0.11	1.66	51.515% (34/66)
BP	3	GO:0051603	proteolysis involved in cellular protein catabolic process	0.26	1.63	52.632% (20/38)

GO class	GO cluster	GO ID	GOterm	padj	NES	genes frequency
BP	3	GO:0006801	superoxide metabolic process	0.25	1.83	42.857% (6/14)
BP	4	GO:0000070	mitotic sister chromatid segregation	0.22	-1.71	40% (6/15)
BP	4	GO:0070317	negative regulation of G0 to G1 transition	0.26	-1.62	34.783% (8/23)
BP	4	GO:0000278	mitotic cell cycle	0.10	-1.64	26.829% (22/82)
BP	4	GO:0007052	mitotic spindle organization	0.21	-1.57	45.946% (34/74)
BP	4	GO:0000086	G2/M transition of mitotic cell cycle	0.27	-1.48	27.397% (20/73)
BP	5	GO:0045724	positive regulation of cilium assembly	0.26	-1.67	36.364% (4/11)
BP	5	GO:1905515	non-motile cilium assembly	0.03	-1.95	31.034% (9/29)
BP	5	GO:0097711	ciliary basal body-plasma membrane docking	0.26	-1.56	27.778% (15/54)
BP	5	GO:0030030	cell projection organization	0.26	-1.46	20.952% (22/105)
BP	6	GO:0006325	chromatin organization	0.25	-1.41	25.389% (49/193)
BP	6	GO:0032981	mitochondrial respiratory chain complex I assembly	0.10	1.80	28.846% (15/52)
BP	6	GO:0070584	mitochondrion morphogenesis	0.08	2.02	42.857% (6/14)
BP	7	GO:0006869	lipid transport	0.01	2.00	50.649% (39/77)
BP	7	GO:0015914	phospholipid transport	0.25	1.82	41.379% (12/29)
BP	7	GO:0120009	intermembrane lipid transfer	0.07	1.99	46.154% (12/26)
BP	8	GO:0048488	synaptic vesicle endocytosis	0.20	1.81	53.571% (15/28)
BP	8	GO:0006911	phagocytosis. engulfment	0.25	1.82	57.143% (8/14)
BP	9	GO:0060173	limb development	0.11	-1.80	47.368% (9/19)
BP	9	GO:0042733	embryonic digit morphogenesis	0.18	-1.78	33.333% (6/18)
BP	9	GO:0007368	determination of left/right symmetry	0.15	-1.75	42.857% (12/28)
BP	9	GO:0001843	neural tube closure	0.11	-1.67	32.653% (16/49)

GO class	GO cluster	GO ID	GOterm	padj	NES	genes frequency
BP	10	GO:0006974	cellular response to DNA damage stimulus	0.22	-1.36	32.119% (97/302)
BP	10	GO:0043312	neutrophil degranulation	0.03	1.55	32.4% (81/250)
BP	10	GO:0045087	innate immune response	0.03	1.66	31.915% (45/141)
BP	10	GO:0006955	immune response	0.11	1.67	31.507% (23/73)
BP	11	GO:0050769	positive regulation of neurogenesis	0.11	-1.81	57.143% (8/14)
BP	11	GO:0032722	positive regulation of chemokine production	0.25	1.78	66.667% (8/12)
BP	11	GO:0090050	positive regulation of cell migration involved in sprouting angiogenesis	0.26	1.74	40% (4/10)
BP	11	GO:0042327	positive regulation of phosphorylation	0.10	1.97	35.714% (5/14)
BP	11	GO:0090023	positive regulation of neutrophil chemotaxis	0.08	1.99	69.231% (9/13)
BP	12	GO:0060544	regulation of necroptotic process	0.26	-1.67	54.545% (6/11)
BP	12	GO:0097191	extrinsic apoptotic signaling pathway	0.11	-1.82	36.842% (7/19)
BP	12	GO:2001244	positive regulation of intrinsic apoptotic signaling pathway	0.26	1.77	40% (6/15)
BP	12	GO:2001235	positive regulation of apoptotic signaling pathway	0.25	1.78	75% (9/12)
BP	13	GO:0008589	regulation of smoothened signaling pathway	0.03	-1.96	63.636% (7/11)
BP	13	GO:0051897	positive regulation of protein kinase B signaling	0.11	1.66	34.286% (24/70)
BP	13	GO:0014068	positive regulation of phosphatidylinositol 3-kinase signaling	0.25	1.59	27.907% (12/43)
BP	13	GO:0051898	negative regulation of protein kinase B signaling	0.04	2.05	44% (11/25)
BP	13	GO:1900745	positive regulation of p38MAPK cascade	0.21	1.85	53.333% (8/15)
MF	1	GO:0003756	protein disulfide isomerase activity	0.33	-1.70	54.545% (6/11)
MF	1	GO:0009055	electron transfer activity	0.44	1.55	52.83% (28/53)
MF	1	GO:0004438	phosphatidylinositol-3-phosphatase activity	0.44	1.69	25% (3/12)

GO class	GO cluster	GO ID	GOterm	padj	NES	genes frequency
MF	2	GO:0051536	iron-sulfur cluster binding	0.15	1.77	56.25% (27/48)
MF	2	GO:0051539	4 iron. 4 sulfur cluster binding	0.07	1.92	61.29% (19/31)
MF	3	GO:0005164	tumor necrosis factor receptor binding	0.01	-1.91	37.5% (3/8)
MF	3	GO:0035257	nuclear hormone receptor binding	0.15	-1.81	46.154% (6/13)
MF	4	GO:0043015	gamma-tubulin binding	0.01	-2.04	25% (4/16)
MF	4	GO:0035064	methylated histone binding	0.44	-1.62	38.71% (12/31)
MF	4	GO:0043621	protein self-association	0.33	1.70	51.724% (15/29)
MF	4	GO:0019955	cytokine binding	0.44	1.66	46.154% (6/13)
MF	4	GO:0034236	protein kinase A catalytic subunit binding	0.44	1.70	37.5% (3/8)
MF	4	GO:0031720	haptoglobin binding	0.44	1.62	100% (5/5)
MF	5	GO:0003684	damaged DNA binding	0.33	-1.70	36.585% (15/41)
MF	5	GO:0003677	DNA binding	0.00	-1.47	26.011% (193/742)
MF	6	GO:0003682	chromatin binding	0.33	-1.41	34.222% (77/225)
MF	6	GO:0008035	high-density lipoprotein particle binding	0.44	1.72	42.857% (3/7)
MF	6	GO:0019212	phosphatase inhibitor activity	0.44	1.65	80% (4/5)
MF	6	GO:0097001	ceramide binding	0.33	1.77	50% (3/6)

Appendix 5. Expressed genes annotated to FGSEA enriched GO terms

GO class	GO cluster	GO term ID	Human ortholog gene symbols
BP	1	GO:0006629	ABCB4;ST3GAL1;ABHD5;AGPS;PNPLA6;IDI1;PLPP1;GBA2;HACD3;HADHA;HSD17B14;CETP;SCD;TECR;MTMR3;DDHD1;PLTP;RUBCNL;GDPD3;ASAH1;EPHX3;ISYNA1;CPT1A;CHKA;PLCH1;HDLBP;ELOVL4;GPAM;EPHX2;AKR1D1;ACSL3;MBOAT7;INSIG2;ECHS1;PLD2;APOE;SLC27A1;PNPLA7;PRKAB2;ERG28;CEPT1;CYP2J2;FADS2;PNPLA8;HADHB;ABHD2;ETNK2;OSBPL10;PLA1A;ERLIN2;IPMK;ACSL1;LPCAT1;PLCL2;SLC16A1;PTDSS1;LRP8;AASDH;G6PD;AGPAT3;FDPS;ACOX1;TPRA1;LIPH;ENPP6;ABCA1;FAAH2;ECI1;AGPAT2;FASN;FABP6;PTEN;FADS6;PTDSS2;ACER2;ACOT4;PSAPL1;PLCXD1;PLCB1;ACSM2A;HACD4;GM2A;FAR1;SLC22A24;PSAP;HACD2;CERS1;PISD;HTD2
BP	1	GO:0006661	MTMR7;SH3YL1;PITPNM2;MTMR3;VAC14;PIK3R2;PIP5K1B;SLC27A1;INPP5K;PIK3R5;PIP5K1A;PTPN13;PTEN;PIP5K1C
BP	1	GO:0006665	AGK;GBA2;CERK;ASAH1;ACER2;PSAPL1;TH;GM2A;PSAP;CERS1
BP	1	GO:0006687	CTSA;GBA2;CERK;ASAH1;ESYT1;GLB1;GM2A;PSAP;HEXA
BP	1	GO:0006646	CHKA;SLC27A1;CEPT1;ETNK2;PHOSPHO1;PISD
BP	1	GO:0001676	CPT1A;ACSL3;SLC27A1;ACSL1;CPT2;SLC27A4;ACOT4;ACAD9
BP	2	GO:0051301	CCDC124;KIF2A;CDC42;NDE1;CCNT2;KLHL42;BIRC5;TRIOBP;TTC28;KATNAL1;UBE2I;FSD1;CDC37;TIMELESS;RNF8;CCNG1;ACTR8;GNAI2;NEK4;CENPA;BIRC6;ARHGEF2;NEK6;NCAPH;CDK7;BORA;SETDB2;DCTN3;NEK1;NUSAP1;CCNG2;POGZ;JTB;DYNC1LI1;TACC1;EML3;INCENP;USP16;PMF1;DYNLT3;PLK1;TUBA1C;CDCA4;PRKCE;TPPP;CKS1B;LIG4;ANKLE2;CETN1;FIGN;KNTC1;SAPCD2;TUBB;DCTN1;PPP1CB
BP	2	GO:0006749	GCLC;GGT5;GSTZ1;ARL6IP5;G6PD;GSTT4
BP	2	GO:0016226	NUBP1;ISCU;NUBPL;XDH;CIAO2B;NFU1;NFS1
BP	2	GO:0007339	CCT4;TCP1;ALDOA;CCT5;CLGN;CCT8;CCT2;ZP3
BP	3	GO:0071897	POLB;POLE4;POLK;REV1;POLG;POLE3;POLL;TK2;LIG4
BP	3	GO:0000381	RBM5;YTHDC1;CELF4;FMR1;HNRNPL;RBM25;SRSF6;ZC3H10;TRA2B;WTAP;RBMX;CELF3;RBM15;RBM10;RNPS1
BP	3	GO:0016070	HNRNPM;POLR2E;HNRNPL;POLR2I;GTF2F1;HNRNPR;NCBP1;HNRNPD;DUSP11;POLR2D;WTAP;RBMX;POLR2H;HNRNPK
BP	3	GO:0042795	CCNT2;SRRT;POLR2E;INTS9;POLR2I;RPAP2;GTF2F1;ELL3;CDK7;NCBP1;INTS3;POLR2D;TAF5;RPRD2;POLR2H;PHAX;GTF2A1;POLR2G;INTS5

GO class	GO cluster	GO term ID	Human ortholog gene symbols
BP	3	GO:0006260	POLA2;RFC1;POLB;SUPT16H;GINS1;TIMELESS;ORC4;RPA2;POLK;DTD1;GINS2;REV1;BARD1;POLG;POLE3;KIN;POLL;MCM7;FAM111A;LIG4;RMI2;TOP3A;PRIM1;TOP1;NT5M
BP	3	GO:0006366	ETV1;GTF2IRD1;CCNT2;SUPT16H;POLR2E;TRIP11;NELFCD;RPAP1;POLR2I;SUPT6H;GTF2H3;RNGTT;TAF12;HIF3A;GTF2F1;ELL3;CDK7;NCBP1;WDR61;NFIC;POLR2D;RBMX;TAF5;PKNOX1;PMF1;IWS1;POLR2H;GTF2A1;DDX21;POLR2G;STAT3;DRAP1;DEAF1;JUN;MAF;TCEA1;GTF2H4
BP	3	GO:0006357	ZFX;ETV1;TARBP1;MED29;RORA;TCF3;SMARCE1;RFX3;MEF2C;PGR;CCNT2;NFE2L1;ATXN7L3;KDM2B;IKZF5;HIF1A;MYEF2;TLE5;LHX2;MED13;KAT2A;DLX4;DBX1;CHD4;PPARD;MED23;FHL2;EPAS1;CHD5;TFAP2E;ARID1A;ATF6;BCL11A;MXI1;ENY2;ZNF644;INHBA;TWIST1;TTC21B;BAZ2B;RBPJL;HIF3A;MAX;STAT5A;UXT;BCL11B;KLF2;FEZF1;MED18;ZNF236;RFXAP;TFCP2;EMX1;VEZF1;TBR1;DMRT1;BBS7;AEBP2;TCF12;ZFHX3;MED9;SMAD4;NFIC;DMRTA2;HDGF;CIAO1;OGT;ST18;THRB;ZIC1;TGIF2LX;BATF;USP16;ZNF362;PKNOX1;MED27;SAFB;GMEB1;LHX8;NIPBL;UNCX;GBX1;HNRNPK;PRKCB;CTNNB1;IRF2BP2;STAT3;ZEB2;HIC2;NFXL1;IRF2BP1;ZBTB26;BPTF;CAMTA1;KLF17;PHF8;TADA2B;ESRRA;STOX2;NR2F1;MAMSTR;DEAF1;JUN;MAF;MED14;BHLHA15;BHLHE22;NPM1;ZBTB20;RFX7;TSHZ2;POU3F1;ZFP91;ZNF292;MYT1;HES5;FOXO6;BRD2;EHMT2;SP9;ZNF709;SEBOX
BP	3	GO:0022900	CIAPIN1;CYBA;NDUFB4;ME1;P4HA2;ME2;PHGDH;DHRS2;ETFB;COX7A2;NDUFS7;COX7A2L;NCF2;SDHB;AKR1A1;COX7C;COX4I1;AOC2;FDX1;ETFA;SDHC;SURF1;XDH;NCF1;DHRS3;NDUFS4;CYBB;ETFDH;GLRX;COX5A;NDUFA4L2;ADH5;MT-ND3;MT-ND1
BP	3	GO:0051603	PSMB1;PSMA4;PSMA3;PSMB5;CTSZ;CTSH;FZR1;CTSC;CTSL;MDM2;CTSV;CTSK;PSMA8;PSMB4;CTSS;RNF123;CTSB;CLPX;PSMB10;PSMB9
BP	3	GO:0006801	CYBA;SOD2;NCF1;CYBB;SH3PXD2B;NOXA1
BP	4	GO:0000070	WRAP73;TUBG1;ESPL1;NUSAP1;INCENP;PLK1
BP	4	GO:0070317	YAF2;L3MBTL2;RNF2;MAX;UXT;SUZ12;EHMT1;EHMT2
BP	4	GO:0000278	HGF;PPP1R12A;TUBB1;NEK4;DNMT3A;RNF2;TUBGCP6;TUBG1;ESPL1;SETDB2;DCTN3;JTB;USP16;NOLC1;PLK1;TUBA1C;TUBB6;CETN1;TUBB;XRCC2;DCTN1;TUBA4B
BP	4	GO:0007052	MAD1L1;NUP160;INTS13;KIF2A;NDE1;SEH1L;BIRC5;NUDC;KIF4A;RANGAP1;DCTN6;CEP126;NUP107;CENPA;STMN1;NEK6;CENPL;B9D2;NUP85;TUBG1;DYNC1LI2;TACC2;DYNC1LI1;TACC1;INCENP;BUB3;PMF1;CLASP2;MAD2L1;PLK1;DCTN2;CENPS;KNTC1;DYNLL2
BP	4	GO:0000086	PPP1R12A;PRKACA;HSP90AA1;CCP110;FBXL15;SKP1;CEP70;OPTN;TUBG1;BORA;DCTN3;PLK1;YWHAG;DCTN2;SSNA1;TUBB;CEP290;DCTN1;CEP43;PPP1CB
BP	5	GO:0045724	FUZ;CROCC;CCP110;DZIP1

GO class	GO cluster	GO term ID	Human ortholog gene symbols
BP	5	GO:1905515	FUZ;MKS1;IFT88;CEP126;IFT172;BBS7;TMEM17;DYNC2H1;DCTN1
BP	5	GO:0097711	MKS1;PRKACA;HSP90AA1;CCP110;CEP70;B9D2;TUBG1;PLK1;YWHAG;DCTN2;SSNA1;TUBB;CEP290;DCTN1;CEP43
BP	5	GO:0030030	FUZ;MKS1;IFT88;CROCC;CCP110;CEP126;WRAP73;B9D2;CEP20;NEK1;BBS7;DCDC2;TMEM237;CATIP;DNAAF3;GPR22;UNC119B;TMEM17;DYNC2H1;DNAJB13;CEP290;CEP43
BP	6	GO:0006325	HLTF;SMARCE1;BAZ2A;TDRD3;ATXN7L3;KDM2B;BRPF3;L3MBTL2;RNF40;SPIN1;EZH1;KMT5B;CHD4;RNF8;CHD5;RBBP5;ARID1A;DNMT3A;HAT1;RRP8;DPF2;SETDB2;AEBP2;SETD1B;CBX4;BRD4;KDM6A;OGT;RNF20;N6AMT1;USP16;SAFB;DPY30;SMARCAD1;MEAF6;BAP1;PRKCB;RCCD1;BPTF;PHF8;SUZ12;EHMT1;SMYD3;KDM4D;TLK1;BRD2;EHMT2;BAG6;C17orf49
BP	6	GO:0032981	NDUFB7;TAZ;FOXRED1;NDUFS7;NDUFAF4;NDUFB5;NUBPL;NDUFAF6;NDUFAF2;TMEM126A;ACAD9;NDUFV2;NDUFA6;MT-ND3;MT-ND4
BP	6	GO:0070584	POLDIP2;DNM1L;CERT1;NUBPL;MFF;BCL2L1
BP	7	GO:0006869	ABCB4;CETP;GRAMD1A;OSBPL8;PLTP;ATP11C;OSBP;CHKA;TMEM30A;CERT1;HDLBP;ESYT2;LBP;APOE;SLC27A1;OSBPL2;ATP8A2;ANO3;ESYT1;PRELID3A;OSBPL1A;ABCG8;OSBPL10;ATP10D;ANO7;VLDLR;PEX19;ABCA1;SLC27A4;MFSD2A;PRELID1;SPNS1;FABP6;APOLD1;APOO;GM2A;MFSD2B;APOC2;YJEFN3
BP	7	GO:0015914	CETP;PITPNM2;OSBPL8;PLTP;ATP11C;OSBPL2;ATP8A2;PRELID3A;ABCG8;ATP10D;ABCA1;PRELID1
BP	7	GO:0120009	CETP;PITPNM2;OSBPL8;PLTP;OSBP;CERT1;APOE;OSBPL2;PRELID3A;ABCG8;ABCA1;PRELID1
BP	8	GO:0048488	ACTB;DNM1L;SYP;DNM1;PACSIN1;SYT2;SYT8;AP2M1;CDK5;ARF6;NLGN1;PIP5K1C;NLGN3;STON1;NLGN2
BP	8	GO:0006911	MARCO;MYH9;BIN2;MFGE8;ELMO1;ABCA1;ARHGAP12;IGLL5
BP	9	GO:0060173	WNT3;KAT2A;KAT2B;IFT172;BBS7;RNF165;SLC39A3;CTNNB1;BMPR2
BP	9	GO:0042733	CREBBP;MKS1;B9D1;TWIST1;SMAD4;CTNNB1
BP	9	GO:0007368	MKS1;RFX3;IFT74;SUFU;NPHP3;ACVR2A;IFT172;DYNC2LI1;BBS7;DPCD;DYNC2H1;DNAAF4
BP	9	GO:0001843	FUZ;MKS1;PRKACA;LHX2;SUFU;KAT2A;TWIST1;IFT172;PRICKLE1;KDM6A;SPINT1;RPS7;TRAF6;DEAF1;RGMA;BRD2
BP	10	GO:0006974	BAZ1B;ERCC1;NUCKS1;POLB;MLH1;UNG;XAB2;UBE2A;XRCC5;SIRT4;SUPT16H;ASCC2;SUSD6;APEX1;XIAP;FMR1;UBR5;FBXW7;GTF2H3;TIMELESS;TDP2;RNF8;PHF1;ACTR8;WDR48;NEK4;GADD45A;RBBP5;RPA2;MLH3;NRDE2;POLK;OARD1;RNF113A;MGME1;PARP2;TOP2A;CHD1L;CDK7;FBH1;DDB2;REV1;TTC5;CDK9;XPA;APTX;DCUN1D5;BARD1;NEIL1;BRD4;INTS3;FANCD2;SETD7;SHPRH;TLK2;POLE3;KIN;DCLRE1C;BATF;USP16;PAXIP1;RNF111;UBQLN4;MAPKAPK2;SMC6;SMARCAD1;DTX3L;NIPBL;CASP3;SPIDR;POLL;

GO class	GO cluster	GO term ID	Human ortholog gene symbols
			MCM7;FAM111A;INO80E;USP47;NFRKB;ZDHHC16;ZMAT3;FBXO45;LIG4;TRAF6;CENPS;ERCC4;PARP10;SLX1B;MORF4L1;KDM4D;NHEJ1;MCRS1;XRCC2;TOPORS;TLK1;GTF2H4;PARG;SHLD3;PRKDC;EID3
BP	10	GO:0043312	MGST1;ALOX5;DERA;ATP6V0A1;AGA;CYBA;SLC2A3;CTSA;PKM;PTPRC;CMTM6;UNC13D;ADA2;TOM1;CTSZ;MAGT1;PSMD7;HMOX2;CTSH;ASAH1;RAB3D;ALDOC;CTSC;TCIRG1;BIN2;PTPN6;FAF2;LAMTOR2;PADI2;CTSD;VAMP8;PLAU;HVCN1;NCKAP1L;RAP2C;FGL2;OSTF1;TMBIM1;ADAM10;RHOF;SRP14;IQGAP1;CD53;CREG1;ILF2;IQGAP2;SURF4;ALDOA;LPCAT1;CCT8;EEF1A1;NBEAL2;PSMC2;CTSS;PTX3;GYG1;CTSB;CYBB;CCT2;B2M;EEF2;FTH1;PGM2;AGPAT2;GLB1;RAB37;HPSE;PSMD1;RHOG;ARHGAP45;CXCR2;RAP2B;ATP6V0C;TUBB4B;HMGB1;GM2A;ATG7;PSAP;PLEKHO2;HBB;PECAM1
BP	10	GO:0045087	MARCO;CYBA;EIF2AK2;IFI35;DDX1;HCK;USP14;TLR8;CORO1A;C1QBP;DHX58;C7;TRIM23;CAPZA1;TRIM32;NMI;LBP;CRP;APCS;RSAD2;PARP9;NLRC5;APP;LYAR;MFHAS1;MR1;SLA;NCF1;C1QC;TNFAIP8L2;SLC15A2;PTX3;CYBB;IFI27;CLEC4E;B2M;PARP14;CSF1R;POLR3C;HMGB1;ILRUN;SRC;AGER;TTC4;IGLL5
BP	10	GO:0006955	CD74;LCP2;TLR8;C1QBP;CTSC;C7;CERT1;EXOSC9;S1PR4;CTSL;CTSV;IFI44;IRF8;MR1;C1QC;CTSS;B2M;SP2;FTH1;PRELID1;CMKLR1;CXCR2;TAPBP
BP	11	GO:0050769	SPEN;XRCC5;SRRT;OPRM1;ELL3;LIG4;XRCC2;TGM2
BP	11	GO:0032722	CD74;EIF2AK2;HMOX1;SELENOK;LBP;APP;CSF1R;AGER
BP	11	GO:0090050	HDAC9;HMOX1;VEGFA;AKT3
BP	11	GO:0042327	VEGFA;NCKAP1L;APP;AR;DSCAM
BP	11	GO:0090023	CD74;DAPK2;DNM1L;C1QBP;RIPOR2;NCKAP1L;RAC2;LBP;CXCR2
BP	12	GO:0060544	HSP90AA1;CDC37;BIRC2;RIPK1;OGT;FADD
BP	12	GO:0097191	KRT18;INHBA;PARP2;ACVR1B;RIPK1;FADD;KRT8
BP	12	GO:2001244	IL20RA;DNM1L;NKX3-1;BCL2L1;BOK;BCAP31
BP	12	GO:2001235	CTSH;TGFBR1;CTSC;TPD52L1;TRAF7;DAB2IP;NKX3-1;ING5;PTEN
BP	13	GO:0008589	CREBBP;FUZ;MKS1;RORA;TTC21B;IFT172;ZIC1
BP	13	GO:0051897	MYOC;FGFR1;MAZ;AKT2;PIK3R2;C1QBP;FGF6;VEGFA;HBEGF;RTN4;RAC2;LIN28A;KL;ARRB2;PIK3R5;MFHAS1;MYORG;P2RY12;HPSE;VEGFB;RHOG;HCLS1;IRS2;SRC
BP	13	GO:0014068	MYOC;FGFR1;MAZ;PTPN6;VEGFA;SOX9;PRR5L;PIK3R5;NCF1;HCLS1;IRS2;SRC

GO class	GO cluster	GO term ID	Human ortholog gene symbols
BP	13	GO:0051898	PHLPP2;HYAL2;MUL1;NOP53;INPP5K;ARRB2;FLCN;XDH;OTUD3;PTEN;PDCD6
BP	13	GO:1900745	SPI1;GADD45B;VEGFA;MINK1;MFHAS1;XDH;NCF1;AGER
MF	1	GO:0003756	PDIA5;PDIA6;PDIA4;PDIA3;CRELD2;PDIA2
MF	1	GO:0009055	CIAPIN1;NDUFS1;CYBA;ME1;P4HA2;ME2;PHGDH;ETFB;COX7A2;COX7A2L;NCF2;SDHB;AKR1A1;AOC2;FDX1;ETFA;SDHC;XDH;NCF1;NDUFS2;DHRS3;NDUFAF2;CYBB;ETFDH;GLRX;NDUFV2;COX5A;ADH5
MF	1	GO:0004438	MTMR7;MTMR3;PTEN
MF	2	GO:0051536	CIAPIN1;NDUFS1;POLD1;NTHL1;FECH;ACO2;NUBP1;NDUFS7;SDHB;LIAS;ACO1;MOCS1;ELP3;RSAD2;ISCU;FDX1;NUBPL;XDH;NDUFS2;UQCRFS1;NFU1;ETFDH;RFESD;POLE;NDUFV2;BOLA2;NFS1
MF	2	GO:0051539	CIAPIN1;NDUFS1;POLD1;NTHL1;ACO2;NUBP1;NDUFS7;SDHB;LIAS;ACO1;MOCS1;ELP3;RSAD2;ISCU;NUBPL;NDUFS2;NFU1;ETFDH;POLE
MF	3	GO:0005164	STAT1;FADD;TRAF6
MF	3	GO:0035257	CRY1;HIF1A;STAT1;TACC2;CTNNB1;JUP
MF	4	GO:0043015	CEP70;B9D2;TUBGCP6;DIXDC1
MF	4	GO:0035064	TDRD3;L3MBTL2;FMR1;SPIN1;RBBP5;RRP8;CBX4;CDYL;BPTF;PHF8;SUZ12;JMJD7
MF	4	GO:0043621	AGA;SYP;SGTA;CTSC;TRIM32;GPSM2;ATXN1;RSAD2;TDG;MDH2;NKX3-1;CRK;TMEM43;MYD88;LDB1
MF	4	GO:0019955	CD74;IFNGR1;CSF3R;TNFRSF11A;FZD4;CSF1R
MF	4	GO:0034236	CSK;SOX9;PRKAR1B
MF	4	GO:0031720	HBA1;HBE1;HBG1;HBD;HBB
MF	5	GO:0003684	CREBBP;ERCC1;POLB;UNG;XRCC5;APEX1;RPA2;POLK;DDB2;REV1;XPA;APTX;DCLRE1C;SMC6;KDM4D
MF	5	GO:0003677	RBM5;RECQL;ZFX;ETV1;GTF2IRD1;CRY1;ERCC1;POLR3B;POLA2;YAF2;SARS1;RFC1;SPEN;RORA;POLB;TCF3;MBD3;HLTF;SMARCE1;BAZ2A;TOP2B;THOC1;XRCC5;RFX3;MEF2C;PGR;NFE2L1;APLP2;SRRT;KDM2B;CERS4;KIF4A;IKZF5;POLR2E;HIF1A;APEX1;DNTTIP1;RPAP1;MYEF2;NUCB1;PDCD5;LHX2;DLX4;SUPT6H;DBX1;TIMELESS;CHD4;PPARD;CENPA;POLE4;STAT1;ORC4;EPAS1;CHD5;ASH1L;TFAP2E;CR2;ARID1A;RPA2;ATF6;DNMT3A;MXI1;TAF12;MTERF2;TARDBP;POLK;ZNF644;TWIST1;RNF11;BAZ2B;RBPJL;HIF3A;SOX4;GTF3C4;GTF2F1;DTD1;MAX;STAT5A;KLF2;FEZF1;PARP2;ZNF236;TOP2A;MYBBP1A;TERF2;RFXAP;FBH1;DDB2;TFCP2;EMX1;REV1;SETDB2;TTC5;VEZF1;TBR1;XPA;APTX;DMRT1;NUSAP1;HNRN

GO class	GO cluster	GO term ID	Human ortholog gene symbols
			PD;AEBP2;TCF12;POLG;ZFHX3;SMAD4;NFIC;DMRTA2;HDGF;CERS2;POGZ;ASXL2;SHPRH;CCDC25;ST18;POLE3;GTF3C5;THRB;KIN;ZIC1;FEZF2;TGIF2LX;RABGEF1;BATF;SUPV3L1;BRPF1;ZNF618;ZNF362;PKNOX1;SAFB;GMEB1;FUBP1;LHX8;ZNF281;H3-3A;SMARCAD1;POLR2H;RAD54L2;UNCX;GBX1;HNRNPK;GTF2A1;POLL;MCM7;FAM111A;TERF2IP;KMT2D;STAT3;ZEB2;HIC2;ZBTB26;CAMTA1;KLF17;ESRRA;THAP2;AGFG1;ENDOV;LIG4;NR1D2;BANF1;DRAP1;ERCC4;RMI2;NR2F1;SOX11;DEAF1;TOP3A;JUN;MAF;PUF60;BHLHA15;ZBTB20;RFX7;TSHZ2;POU3F1;POLR1D;H1-6;TCEA1;NHEJ1;H3-5;ZNF292;MYT1;LONP1;XRCC2;HES5;QRICH1;ZNF536;TOP1;FOXO6;SP9;ZNF709;PRKDC;C17orf49;SEBOX
MF	6	GO:0003682	CREBBP;SCMH1;NCAPH2;NUCKS1;MBD3;SMARCE1;MLH1;TOP2B;SMARCA2;CCNT2;TDRD3;FUS;IFT74;L3MBTL2;STAG2;FMR1;LHX2;KAT2A;EZH1;KMT5B;RNF8;PHF1;KAT2B;CENPA;ASH1L;MLH3;DNMT3A;ENY2;NCAPH;RNF2;TTC21B;RBPJL;KLHDC3;UXT;PARP2;TOP2A;TTC5;CDK9;APTX;DMRT1;HNRNPD;POLG;CBX4;SMAD4;BRD4;ASXL2;DLX1;SETD7;SCML4;RBMX;THRB;CDYL;FEZF2;RNF20;AUTS2;PKNOX1;MCM3AP;SAFB;SMARCAD1;BAP1;WDR82;NIPBL;PRKCB;CTNNB1;SIN3A;PHF8;TADA2B;CENPS;NPM1;HES5;TOP1;SAMD13;BRD2;CORT;BAHCC1;PCGF2;
MF	6	GO:0008035	PLTP;LRP8;ABCA1
MF	6	GO:0019212	TESC;ARPP19;ANP32E;ENSA
MF	6	GO:0097001	PLTP;CERT1;VDAC1