Study of morphogenesis and miRNA expression associated with craniofacial diversity in Arctic charr (*Salvelinus alpinus*) morphs

Kalina H. Kapralova

Faculty of Life and Environmental Sciences
University of Iceland
2014

Study of morphogenesis and miRNA expression associated with craniofacial diversity in Arctic charr (*Salvelinus alpinus*) morphs

Kalina H. Kapralova

Dissertation submitted in partial fulfillment of a
Philosophiae Doctor degree in Biology

Advisors
Professor Sigurður S. Snorrason
Professor Zophonías O. Jónsson
Dr. Arnar Pálsson

PhD Committee
Professor Sigurður S. Snorrason
Professor Zophonías O. Jónsson
Dr. Arnar Pálsson
Professor Ian A. Johnston

Opponents
Professor John H. Postlethwait
Professor Eiríkur Steingrímsson

Faculty of Life
School of Engineering and Natural Sciences
University of Iceland
Reykjavik, August 2014

Study of morphogenesis and miRNA expression associated with craniofacial diversity in Arctic charr (Salvelinus alpinus) morphs
Morphogenesis and miRNA expression in sympatric Arctic charr

Dissertation submitted in partial fulfillment of a *Philosophiae Doctor* degree in Biology

Faculty of Life and Environmental Sciences
School of Engineering and Natural Sciences
University of Iceland
Sæmundargötu 2
101, Reykjavik
Iceland

Telephone: 525 4000

Bibliographic information:
Kalina H. Kapralova, 2014, Study of morphogenesis and miRNA expression associated with craniofacial diversity in Arctic charr *(Salvelinus alpinus)* morphs, PhD dissertation, Faculty of Life and Environmental Sciences, University of Iceland, 189 pp.

ISBN 978-9935-9164-8-8

Printing: Háskólaprent
Reykjavik, Iceland, August 2014

Abstract

Trophic diversification of Arctic charr (*Salvelinus alpinus,* Linn. 1758) into four genetically distinct morphs, varying in life history characteristics, behavior, and trophic morphology, has occurred in lake Thingvallavatn following the last glaciation. The aim of this study was to investigate the genetic and developmental aspects of this diversification thereby gaining insights into the evolution and the maintenance of the Thingvallavatn morphs. In chapter I a population genetic screen of immunological candidate genes revealed differences among morphs for *Cath2* and *MHCII alpha* that far exceeded differentiation at neutral loci. This is consistent with a scenario were selection has led to divergence in parts of the immune system. In chapter II embryonic and early post-hatching craniofacial cartilage development is described. The ontogenetic trajectories of shape and size indicated developmental heterochrony as a possible mechanism of morph divergence. Chapter III describes subtle but significant differences in early post-hatching craniofacial morphology between the progeny of three morphs of Arctic charr. Moreover hybrid progeny of two contrasting morphs showed extreme (or transgressive) phenotypes well outside of the parental range, indicating that the ecological divergence within the lake might be enhanced by lowered fitness of hybrids. In chapter IV the level of integration and modularity in craniofacial traits in morphs and hybrids is analysed. Chapter V describes the annotation of Arctic charr miRNAs expressed during development and analyses of candidate miRNAs involved in Arctic charr morphogenesis and diversification.

Útdráttur

Frá lokum síðustu ísaldar hafa þróast fjögur afbrigði bleikju (*Salvelinus alpinus,* Linn. 1758) innan Þingvallavatns. Afbrigðin eru erfðafræðilega aðgreind og eru ólík hvað snertir lífsferla, atferli og útlit, og á það sérstaklega við um líkamshluta er tengjast fæðuöflun. Markmið þessarar rannsóknar var að kanna erfðafræðilegar og þroskunarfræðilegar orsakir þessa fjölbreytileika og öðlast þannig innsýn í þróun og varðveislu bleikjuafbrigðanna í Þingvallavatni. Stofnerfðafræðilegri leit að genum tengdum ónæmiskerfinu sem sýna mismun milli afbrigða er lýst í fyrsta kafla ritgerðarinnar. Þar á meðal eru *Cath2* og *MHCII alpha* sem sýna breytileika sem getur ekki talist hlutlaus og líklegast er að áhrif náttúrulegs vals á ónæmiskerfið hafi leitt til aðgreiningar á þessum erfðasetum. Annar kafli lýsir þroskun brjósks og beina í höfði fóstra og seiða stuttu eftir klak. Sá munur sem fram kemur milli afbrigða í þroskunarfræðilegum brautum útlits og stærðar þessara stoðeininga bendir til þess að orsakanna sé að leita í breytingum á tímasetningu þroskunaratburða. Í þriðja kafla segir frá litlum en marktækum mun í útliti höfuðbeina á fyrstu stigum eftir klak seiða þriggja afbrigða bleikju. Auk þess sýna blendingar tveggja ólíkra afbrigða svipgerð sem fellur að verulegu leyti fyrir utan útlitsmengi beggja foreldra-afbrigðanna. Það bendir til þess að aðskilnað afbrigðanna í vatninu megi rekja til minni hæfni blendinga. Fjórði kafli fjallar um þroskunarfræðileg tengsl valinna stoðeininga í höfði, þ.e. hversu sjálfstæðar eða samþættar þær eru, og hvernig þeim er háttað hjá kynblendingum ólíkra afbrigða Í fimmta kafla er miRNA sameindum bleikjunnar og tjáningu þeirrra í þroskun lýst í mismunandi afbrigðum. Sérstaklega athygli fengu miRNA-gen sem sýndu mismunandi tjáningarmynstur á afbrigðunum en slík gen kunna að leika mikilvægt hlutverk í formþroskun höfuðbeina og þannig verið undirstaða útlitsmunar milli afbrigða.

To my family,

Недка, Христо, Петя, Матчо и Фредрикчо

List of papers

This thesis is based on five papers, of which one has been published, one has been accepted for publication and three are manuscripts. In the text the papers are referred to with their respective numbers as follows:

- Paper I: Kalina H. Kapralova, Johannes Gudbrandsson, Sigrun Reynisdottir, Cristina B. Santos,Vanessa C. Baltana´ s, Valerie H. Maier, Sigurdur S. Snorrason, Arnar Palsson (2013). Differentiation at the MHCIIα and *Cath2* Loci in sympatric *Salvelinus Alpinus* Resource Morphs in Lake Thingvallavatn. *PLoS One* **8**: e69402

- Paper II: Kalina H. Kapralova, Zophonías O. Jónsson, Arnar Palsson, Sigrídur Rut Franzdóttir, Bjarni K. Kristjanson, Sigurður S. Snorrason. Bones in motion: ontogeny of craniofacial development in sympatric Arctic charr morphs (Manuscript)

- Paper III: Kalina H. Kapralova, Arnar Palsson, Bjarni K. Kristjanson, Zophonías O. Jónsson, Sigurður S. Snorrason. Evidence for reproductive isolation by phenotypic transgression in hybrids of Arctic charr (Manuscript).

- Paper IV: Kalina H. Kapralova, Arnar Palsson, Bjarni K. Kristjanson, Zophonías O. Jónsson, Sigurður S. Snorrason. Modularity and integration in craniofacial elements during development of sympatric Arctic charr morphs (Manuscript).

- Paper V: Kalina H. Kapralova, Sigrídur Rut Franzdóttir, Hákon Jónsson, Sigurður S. Snorrason, Zophonías O. Jónsson (2014). Patterns of miRNA expression in Arctic charr development (accepted for publication in *PLoS One*)

Other peer-reviewed papers not included in this thesis

- Ehsan Pashay Ahi, Kalina Hristova Kapralova, Arnar Pálsson, Valerie Helene Maier, Jóhannes Gudbrandsson, Sigurdur S. Snorrason, Zophonías O. Jónsson, Sigrídur Rut Franzdóttir. Transcriptional dynamics of a conserved gene expression network associated with benthic-limnetic craniofacial divergence in Arctic charr (Submitted to Evo-devo).

- Ehsan Pashay Ahi, Jóhannes Guðbrandsson, Kalina H. Kapralova, Sigríður R. Franzdóttir, Sigurður S. Snorrason, Valerie H. Maier, Zophonías O. Jónsson (2013). Validation of Reference Genes for Expression Studies during Craniofacial Development in Arctic Charr. PLoS ONE 8(6): e66389

Table of Contents

List of Figures

List of Tables

Acknowledgements

With great pleasure I would like to acknowledge the people who helped with the realisation of this thesis and supported me through this important period of my life.

First and foremost I would like to thank my main advisor Sigurður S. Snorrason for inviting me to join his team and giving me the opportunity to study Arctic charr. I am very grateful for his guidance through the years, his contagious interest in Evolution and in Science in general, his constant care and his unceasing optimism. Siggi, your open mind and positive attitude towards everything are very inspiring.

I thank my advisor Zophonías O. Jónsson for providing the opportunity to work on this project. His extensive molecular knowledge, brilliant ideas, willingness to take challenges, and implement novel methods to study this truly amazing and often not easy to work with species, made working on this project a great experience. Our chats in Bulgarian really made me feel like I was at home.

I am grateful to my advisor Arnar Pálsson for all his help. Arnar was the driver behind the MHC study. Also, his experience in geometric morphometrics and great sense of humour were of a great help during the dark times of multivariate analyses.

I would like to thank Professor Ian A. Johnston for agreeing to be in my committee, his critical reading of the thesis, and his comments.

I thank Professors John Postlethwait and Eiríkur Steingrímsson for allocating their time and expertise to review this thesis.

I thank Bjarni K. Kristjánson for the great sampling trips, although I realise now that having an untrained molecular biologist as an assistant in the field must not have been easy for him. I also enjoyed a lot our chats about morphometrics and evolution. The establishment of the fantastic beer club at Hólar is another thing that I and probably numerous other graduate students are very thankful for.

I was fortunate enough to work in a professional and yet friendly environment studying the development and evolution of a remarkable species. For that I would like to thank the members of the "charr group" (in no particular order): Valerie Maier, Sigrídur Rut Franzdóttir, Johannes Gudbrandsson and Ehsan Pashay Ahi.

I am very grateful to Soizic, Camille and Eva for their help with sampling. I would like to thank at Mjóanes, Jóhann Jónsson, for all his help with sampling from Thingvallavatn.

Special thanks to the bioinformatics and statistics gurus Hákon Jónsson and Johannes Guðbrandsson. Their vast knowledge and immense patience whilst sharing it made the work on this project even more enjoyable. You guys rock!

I thank Snæbjörn Pálsson for his help with statistics, Skúli Skúlason for our chats about charr ontogeny and his vivid interest in my findings. Special thanks to Guðmundur Hrafn Guðmundsson and Eva Benediktsdóttir for their advices and critical reading of the thesis and Ólafur S. Andrésson for his advices in the lab. I would like to thank Sigrún Reynisdóttir for her great help in the lab and her precious advice on everything from the best ways to order lab consumables to how to deal with Icelandic administration.

Thanks to all lab mates and crew from Askja: Borgný, Etienne, Ágústa, Rannveig, Lisa Anne (and her awesome cakes), Lovísa, Elísabet, Laurène, Ellen, Silke, Óli Patrick, Ágúst, Snorri Páll, Hlynur, Will, Ester, Ragnhildur, Sophie, Pamela, Derek, Jed, Fraser, Kai, Freydís, Martin, Oliver, Julian, Eduardo, Ubaldo, Ólöf Birna, Habba, Edda, Hrönn, Hildur, Ana, Andrey, Kata, Rakel and Gunnar. I am particularly grateful to my friend Kristen for her great support, especially in the last stages of writing. Her help with everything from proofreading and checking references to cooking lunch and baking tasty cakes has been instrumental for the successful completion of this thesis. I would like to thank Sindri for taking the most beautiful photo of Arctic charr I have seen and allowing me to use it, making it very easy to show people what truly remarkable creature I work with.

I would like to thank my first students Cristina, Javier, Vanessa, Xin, Jannika and Ástrós, and I wish them all the best for their future careers.

Last but not least I would like to thank my family Nedka, Hristo, Petya, Matthew and Fredrik for their unconditional love, support and understanding. Following their philosophy that when stressed one needs good food, a nice drink and if possible a visit to the beach or the mountains, has been extremely important especially in the last 6 months of writing. I feel privileged for having such an awesome family.

This project was supported by grants from The Icelandic Centre for Research (RANNIS#100204011, #110285-0061 and #130756-051) and the University of Iceland Research Fund.

1 Chapter I

1.1 General introduction

A central question in evolutionary biology is how the spectacular ecological and phenotypic diversity seen across the world's biota is generated and maintained. Adaptive divergence is regarded as the most important biological process leading to the evolution of ecological differences and ultimately to the emergence of new species (Schluter 2000; Gavrilets & Losos 2009). Adaptive radiations are examples of rapid divergence in particular groups, which generates many and distinct species. Like evolution in general, the process of adaptive radiation is dependent on historical contingencies and chance, and influenced by ecological, genetic and developmental factors (Gavrilets & Losos 2009). Adaptive radiation is characterized by four criteria: common ancestry, phenotype-environment correlation, trait utility and rapid speciation (Schluter 2000). Although an association between phenotype and environment can seemingly arise via non-adaptive processes (Gould & Lewontin 1979), replicated occurrences of certain phenotypes in multiple independent lineages are considered strong evidence for adaptive processes (Endler 1986; Schluter & Nagel 1995; Harvey & Pagel 1998).

The parallel evolution along a benthic-limnetic axis, seen in diverse lineages of bony fishes, is an excellent example of trophic adaptive divergence (Schluter & Rambaut 1996; Wainwright 1996; Wainwright & Shaw 1999; Bouton *et al.* 1999; Snorrason & Skúlason 2004). This ecological divergence is accompanied by changes in the morphology of the feeding apparatus: species that feed on mobile prey (limnetic species) have long, evenly protruding jaws adapted for suction feeding, whereas many benthic species that feed on slow moving benthic prey, have shorter and more robust lower jaws (see references in Willacker *et al.,* 2010). Phenotypic diversity arises through the combined effects of genes and the environment, and is built by the principles of development.

Development translates genotypes into phenotypes, which in turn will be available for natural selection to act upon. The understanding of how trophic morphologies evolve in response to environmental attributes will thus require knowledge of how these morphologies are produced during development (Atchley & Hall 1991). Hence, species with high within-species polymorphism and undergoing early stages of adaptive divergence may be well suited for studying the role of development in phenotypic variation.

The relationship between phenotypic divergence during evolution and individual development has, over the last three decades, been the focus of the field of evolutionary developmental biology (evo-devo) (Müller 2007). The integration of evolutionary theory with molecular and developmental genetics leads to the establishment of some general principles, namely, that developmental genes constitute a genetic toolkit and phenotypes most often evolve through changes in spatial and temporal regulation of functionally

conserved genes with *cis* regulatory elements being the focus of many studies (reviewed in Carroll, 2008). However, the importance of post-transciptional regulation in morphological evolution has recently recieved increasing attention. Micro-RNAs (miRNA's) are small (~ 22 nt) non-coding RNAs that post-transcriptionally regulate the expression of target genes, thus making for a specific and "fine-tunable" response (reviewed in Li & Zhang, 2012). It has been suggested that miRNAs may generally cover more restricted regulatory niches than transcription factors (Hobert 2008) and might be involved in enhancing species evolvability (Ebert & Sharp 2012).

A first step towards studying the role of development in producing variable trophic morphologies is to identify fundamental developmental units of the trophic apparatus. The next steps involve studies of the growth and shape changes during ontogeny as well as the integration of these developmental units into complex, functional assemblages (Atchley & Hall 1991). For example, the African cichlids, one of the most dramatic and best studied case of adaptive radiation, exhibit a limited number of distinct trophic morphologies that appear to have evolved repeatedly along the benthic/limnetic functional axis (Albertson & Kocher 2006). This interesting phenomenon might be the result of developmental processes constraining the variation in one or more axes of the morphospace (Brakefield 2011). Alternatively, genetic architecture or selection can shape the variance-covariance structure in populations, and hence evolution (Pitchers *et al.* 2014).

Three distinct approaches to identify underlying molecular mechanisms for morphological change can be employed: a candidate gene approach, mapping and genome wide screens. The first approach takes advantage of the wealth of information accumulated over decades of research in model organisms, whereas the second approach requires either genome wide quantitative trait loci or association mapping. The third approach encompasses other types of genome wide screens. Genome wide screens do not require any *a priori* information about the genetic and developmental nature of phenotypic variation under study. The candidate gene approach can take advantage of the extensive knowledge available from studies on model species to generate or narrow down lists of relevant candidates (Mallarino & Abzhanov 2012).

Because the majority of craniofacial mutants exhibit severe and often lethal defects, little is usually known about craniofacial morphogenesis in post embryonic stages (Albertson & Yelick 2004). Genome-wide screens represent a less biased way to find genes and pathways related to divergence, however, mapping requires considerable statistical power and large samples sizes. The vast amount of information generated in transcriptome analyses is often hard to interpret and validate. Thus, a combination of genome wide and candidate gene approaches can be employed to study morphological variation, especially in non-model species. Natural systems exhibiting high morphological variation can be used to assay the development and remodeling of traits beyond the embryonic stages (Abzhanov *et al.* 2004; Kimmel *et al.* 2005; Albertson & Kocher 2006).

In this study a highly variable fish species, Arctic charr (*Salvelinus alpinus*), is utilized to study the role of development in building variable trophic morphologies. Throughout its circumpolar distribution this species exhibits high levels of inter-population phenotypic variation with many populations showing trophic divergence along a benthic - limnetic habitat axis (Klemetsen *et al.* 2003). The young evolutionary age (post-glacial) of such

polymorphism in Arctic charr, readily observable phenotypic variation, and tractable ecological settings makes it an ideal species for the work presented in this thesis.

The Arctic charr in Lake Thingvallavatn, Iceland, constitute an extreme example of local phenotypic diversity. Four distinct morphs have been described in the lake, two benthic morphs (a large, LB, and a small, SB, benthivorous charr) and two limnetic morphs (a planktivorous, PL, and a piscivorous, PI, morph). The four morphs inhabit different ecological niches and, among other things, differ in parasite infection rate and prevalence (Frandsen *et al.* 1989). The benthic morphs inhabit the stony littoral zone and feed on slow moving benthic invertebrates, especially on the gastropod *Lymnaea peregrea* and littoral chironomid larvae. These fish are dark in color, have stocky bodies, long pectoral fins, blunt snouts, short lower jaws, and a low number of gill rakers. The planktivorous (PL) and piscivorous (PI) feed on zooplankton and threespine stickleback, respectively. These fish are more silvery, have fusiform bodies, smaller pectoral fins, pointed snouts, longer lower jaws, and a greater number of gill rakers compared to the two benthic morphs (Snorrason *et al.* 1989; Malmquist *et al.* 1992; Sandlund *et al.* 1992).

The morphs differ extensively in life-history characteristics: LB and PI take longer to mature sexually (3-11 and 5-10 years, respectively) and mature at a minimum 20.5 and 25.6 cm. fork length, respectively. The two smaller morphs (SB and PL) mature at a minimum of 7.2 and 15.2 cm. fork length, respectively, and have relatively short maturation time (2-4 and 3-4 years, respectively) (Jonsson *et al.* 1988).

While all Arctic charr morphs spawn in the stony littoral habitat the timing and the level of synchronization differ with the two smaller morphs (SB and PL) showing a partial spatial and temporal overlap (Skúlason *et al.* 1989).

Common garden experiments with the progeny of the four morphs have shown a clear genetic component to differences in growth, maturity age, development of trophic morphology, body color and foraging behavior (Skúlason *et al.* 1989, 1993; Smith & Skulason 1996; Eiriksson *et al.* 1999).Additionally, phenotypic plasticity is also a significant factor in molding the phenotypes (Parsons *et al.* 2010, 2011). Overall, these results emphasize the adaptive nature of morph formation of Arctic charr in Lake Thingvallavatn.

A variety of molecular and biochemical markers have been previously used to assess the population genetic structure of Arctic charr in Thingvallavatn. While some were insufficient to detect differentiation among morphs due to low amount of polymorphism between individuals (Magnusson & Ferguson 1987; Danzmann *et al.* 1991; Volpe & Ferguson 1996), others detected significant differences among morphs belonging to different morphotypes (limnetics vs benthics) but not within morphotypes (Volpe & Ferguson 1996). The most recent and rigorous study used 10 microsatellite markers and was based on samples of fully ripe spawners from five locations around the lakeshore. This study revealed subtle but significant genetic differentiation between the three most common morphs in Lake Thingvallavatn (PL, SB and LB). Further coalescent simulations indicated a scenario of early evolution of reproductive isolation followed by slow divergence by drift with restricted gene flow (Kapralova *et al.* 2011). The same study also suggested small benthic charr have evolved independently in several other springs and lakes in Iceland, providing the opportunity for comparative analyses of developmental

morphology and its genetic and molecular basis. Based on this previous research, the Arctic charr in Iceland and of Lake Thingvallavatn provide an ideal 'natural experimental' system for studying the role of development in the evolution in adaptive divergence.

This dissertation is divided into five chapters organized like manuscripts with each one addressing questions on i) population structure at putatively adaptive genes of the immune system, ii) characterizing major events of early Arctic charr craniofacial development, iii) how early in development shape changes arise between morphs and the effects of hybridization of contrasting morphologies on craniofacial shape and development, iv) patterns and level of modularity in Arctic charr pure morphs and hybrid crosses and v) patterns of miRNA expression during the development of contrasting Arctic charr morphologies.

In the first chapter I conducted a population genetics screen on four immunological candidate genes *Cathelicidin 2* (*Cath2*), *Hepcidin* (*Hamp*), *Liver expressed antimicrobial peptide 2a* (*Leap-2a*), and *Major Histocompatibility Complex IIa* (*MHCIIa*) and a mitochondrial marker (D-loop) among the three most common Lake Thingvallavatn charr morphs (LB, PL and SB). Two of the loci (*Cath2* and *MHCIIa*) showed significant differences in allele frequencies among morphs. In *Cath2*, SB deviated from the other two (F_{ST}=0.13). One of the substitutions detected was an amino acid replacement polymorphism in the antimicrobial peptide. This change is predicted to lead to an amino acid replacement (replacement of arginine by serine in position 115), altering the charge of the peptide and possibly its function.

A more striking difference was found in the *MHCIIa*. Two haplotypes were very common in the lake, and their frequency differed greatly between PL and SB (from 22% to 93.5%, F_{ST}=0.67). Next I surveyed the variation in *Cath2* and *MHCIIa* in nine Arctic charr populations from around Iceland. The populations varied greatly in terms of allele frequencies at *Cath2*. However the variation did not correlate with morphotype. The variation at the *MHCIIa* locus, was nearly identical to the variation in the two benthic morphs of Lake Thingvallavatn. The results are consistent with a scenario where parts of the immune systems have diverged substantially among Arctic charr populations in Iceland, possibly from standing genetic variation.

In the second chapter I used the recently evolved polymorphism in Icelandic Arctic charr to address questions about craniofacial development and evolution. First, dense developmental series' were established and the timings of major events in head cartilage and bone development were characterized. A total of eight (four pre- and four post-hatching) time points were selected to study the ontogenetic trajectories related to growth in three of the Thingvallavatn morphs, SB and LB (representing a benthic morphotype) and PL and an aquaculture strain AC (representing a limnetic morphotype). The charr morphs displayed segmental development of the pharyngeal arches, as is characteristic for all vertebrates, and the order of events accompanying the craniofacial development was the same as has been described for teleosts. The four Arctic charr varieties under study showed similar general patterns of head growth during this period of development. On a finer scale the growth rate differed among groups. The head starts out smaller in the benthic morphs but, due to a sharp increase in growth rate at hatching, the LB morph ends up with the largest heads at the post-hatching stages. The hatching period appears to be associated with significant allometric shape changes. The four varieties differed in the size, orientation

and/or shape of their ontogenetic trajectories of shape. LB had the largest trajectory of all the studied groups (i.e accumulated the most shape changes for the studied period) and differed significantly ($p < 0.05$) from the two limnetic groups (PL and AC). On the other hand, SB showed significant differences from the other three groups in the orientation of their ontogenetic trajectory of shape. SB also displayed marginally significant differences in the shape of ontogenetic trajectories compared to LB and AC. Interestingly, the two limnetic groups did not show significant differences in any of the three attributes of ontogenetic trajectories.

In the third chapter I used landmark based geometric morphometrics and multivariate analyses of shape to address questions on the evolution and development of elements of the Arctic charr feeding apparatus. I studied the progeny of pure morph crosses for SB, LB and PL. Compared to the stages studied in chapter II, here I focused on later stages that allowed the use of more landmarks. All studied groups displayed subtle differences in early development during cartilage formation and growth. Next I investigated the effect of hybridisation on the craniofacial morphology of Arctic charr by creating reciprocal crosses between PL and SB. Interestingly, the majority of hybrid embryos exhibited craniofacial phenotypes that were considerably displaced from the distributions seen in offspring of pure crosses. No significant differences in head shape were detected between the two reciprocal crosses, suggesting that the above transgressive genetic effects greatly outweighed any maternal effects.

In the fourth chapter, data from the crosses used in the third chapter were analyzed to address questions about integration and modularity in the developing trophic apparatus of Arctic charr. These data have not been fully analyzed, and the chapter summarizes the first descriptive and exploratory analyses. Preliminary results showed that during early post-hatching stages the craniofacial skeleton is modular and this modularity appears to reflect the developmental origins of the elements constituting it. The craniofacial integration was compared in AC, LB, PL and SB groups of Arctic charr. These groups did not appear to differ in the pattern, but rather in the level of their craniofacial modularity. Interestingly, hybrid crosses between two contrasting Arctic charr morphs may have different patterns of integration of their craniofacial skeleton compared to the pure crosses of the parental morphs.

In chapter five I studied the expression of small non-coding RNAs (miRNAs) during embryonic development of offspring from the two contrasting varieties; SB from Lake Thingvallavatn and AC from the Holar aquaculture stock. To this end, four time points (three embryonic and one just before first feeding) were selected for high-throughput small-RNA sequencing. A total of 326 conserved and 427 novel miRNA candidates were identified in Arctic charr of which 51 were conserved and six novel miRNA candidates were differentially expressed among developmental stages. Furthermore, 53 known and 19 novel miRNAs showed significantly different levels of expression in the two contrasting morphs. Some of these miRNAs are involved in regulating key developmental processes in other species such as development of brain and sensory epithelia, skeletogenesis, myogenesis and hematopoiesis. For example sal-miR-146, 183, 206 and 196a were highly expressed in the benthic embryos and sal-miR-130, 30, 451, 133, 26, and 199a were highly expressed in the limnetic embryos. The expression differences are confined to the embryonic stages and the two morphs exhibited similar miRNA expression profiles in the last stage. Interestingly, four of the 19 novel miRNA candidates were only detected in

either AC or SB.

As outlined above, ecological diversification of Arctic charr (Salvelinus alpinus) into four phenotypic variants within lake Thingvallavatn has occurred in just 10,000 years following the last glacial maximum. Ecological speciation may be progressing within this natural experimental system as evidenced by distinct variation in life history characteristics, behavior and trophic morphology. The research presented here contributes to the growing body of work on the underlying mechanisms of Arctic charr adaptive divergence.

In addition to the papers described above, during my PhD studies I took part in two additional studies, which have not been included in this dissertation. One, titled "Validation of reference genes for expression studies during craniofacial development of Arctic charr" was published in PloS One in 2013 (10.1371/journal.pone.0066389), and the other one titled "Transcriptional dynamics of a conserved gene expression network associated with benthic-limnetic craniofacial divergence in Arctic charr" has just been submitted to Evo-Devo.

1.2 References

Abzhanov A, Protas M, Grant BR, Grant PR, Tabin CJ (2004) Bmp4 and morphological variation of beaks in Darwin's finches. *Science (New York, N.Y.)*, **305**, 1462–1465.

Albertson RC, Kocher TD (2006) Genetic and developmental basis of cichlid trophic diversity. *Heredity*, **97**, 211–21.

Albertson RC, Yelick PC (2004) Morphogenesis of the jaw: development beyond the embryo. *Methods in cell biology*, **76**, 437–54.

Atchley WR, Hall BK (1991) A model for development and evolution of complex morphological structures. *Biological reviews of the Cambridge Philosophical Society*, **66**, 101–57.

Bouton N, Witte F, Alphen JJM v., Schenk A, Seehausen O (1999) Local adaptations in populations of rock-dwelling haplochromines (Pisces:Cichlidae) from southern Lake Victoria. *Proceedings of the Royal Society B: Biological Sciences*, **266**, 355–360.

Brakefield PM (2011) Evo-devo and accounting for Darwin's endless forms. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **366**, 2069–75.

Carroll SB (2008) Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell*, **134**, 25–36.

Danzmann RG, Ferguson MM, Skulason S, Snorrason SS, Noakes DLG (1991) Mitochondrial DNA diversity among four sympatric morphs of Arctic charr, Salvelinus alpinus L., from Thingvallavatn, Iceland. *Journal of Fish Biology*, **39**, 649–659.

Ebert MS, Sharp P a (2012) Roles for microRNAs in conferring robustness to biological processes. *Cell*, **149**, 515–24.

Eiriksson GM, Skulason S, Snorrason SS (1999) Heterochrony in skeletal development and body size in progeny of two morphs of Arctic charr from Thingvallavatn, Iceland. *Journal of Fish Biology*, **55**, 175–185.

Endler JA (1986) *Natural Selection in the Wild*.

Frandsen F, Malmquist HJ, Snorrason SS (1989) Ecological parasitology of polymorphic Arctic charr, Salvelinus alpinus (L.), in Thingvallavatn, Iceland. *Journal of Fish Biology*, **34**, 281–297.

Gavrilets S, Losos JB (2009) Adaptive radiation: contrasting theory with data. *Science (New York, N.Y.)*, **323**, 732–7.

Gould SJ, Lewontin RC (1979) The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme. *Proceedings of the Royal Society B: Biological Sciences*, **205**, 581–598.

Harvey PH, Pagel MD (1998) *The comparative method in evolutionary biology*. Oxford University Press.

Hobert O (2008) Gene regulation by transcription factors and microRNAs. *Science*, **319**, 1785–6.

Jonsson B, Skúlason S, Snorrason SS *et al.* (1988) Life History Variation of Polymorphic Arctic Charr (Salvelinus alpinus) in Thingvallavatn, Iceland. *Canadian Journal of Fisheries and Aquatic Sciences*, **45**, 1537–1547.

Kapralova KH, Morrissey MB, Kristjánsson BK *et al.* (2011) Evolution of adaptive diversity and genetic connectivity in Arctic charr (Salvelinus alpinus) in Iceland. *Heredity*, **106**, 472–487.

Kimmel CB, Ullmann B, Walker C *et al.* (2005) Evolution and development of facial bone morphology in threespine sticklebacks. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 5791–6.

Klemetsen A, Amundsen P-A, Dempson JB *et al.* (2003) Atlantic salmon Salmo salar L., brown trout Salmo trutta L. and Arctic charr Salvelinus alpinus (L.): a review of aspects of their life histories. *Ecology of Freshwater Fish*, **12**, 1–59.

Li J, Zhang Z (2012) miRNA regulatory variation in human evolution. *Trends in genetics : TIG*, 1–9.

Magnusson KP, Ferguson MM (1987) Genetic analysis of four sympatric morphs of Arctic charr, Salvelinus alpinus, from Thingvallavatn, Iceland. *Environmental Biology of Fishes*, **20**, 67–73.

Mallarino R, Abzhanov A (2012) Paths less traveled: evo-devo approaches to investigating animal morphological evolution. *Annual review of cell and developmental biology*, **28**, 743–63.

Malmquist HJ, Snorrason SS, Skulason S *et al.* (1992) Diet differentiation in polymorphic Arctic charr in Thingvallavatn , Iceland. *Journal of Animal Ecology*, **61**, 21–35.

Müller GB (2007) Evo-devo: extending the evolutionary synthesis. *Nature reviews. Genetics*, **8**, 943–9.

Parsons KJ, Sheets HD, Skúlason S, Ferguson MM (2011) Phenotypic plasticity, heterochrony and ontogenetic repatterning during juvenile development of divergent Arctic charr (Salvelinus alpinus). *Journal of evolutionary biology*, **24**, 1640–52.

Parsons KJ, Skúlason S, Ferguson M (2010) Morphological variation over ontogeny and environments in resource polymorphic arctic charr (Salvelinus alpinus). *Evolution & development*, **12**, 246–57.

Pitchers W, Wolf JB, Tregenza T, Hunt J, Dworkin I (2014) Evolutionary rates for multivariate traits: the role of selection and genetic variation. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **369**, 20130252–.

Sandlund OT, Gunnarson K, Jonasson PM *et al.* (1992) The Arctic charr Salvelinus alpinus in Thingvallavatn. *Oikos*, **64**, 305–351.

Schluter D (2000) *The Ecology of Adaptive Radiation*.

Schluter D, Nagel LM (1995) Parallel Speciation by Natural Selection. *The American Naturalist*, **146**, 292.

Schluter D, Rambaut A (1996) Ecological Speciation in Postglacial Fishes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **351**, 807–814.

Skúlason S, Snorrason SS, Noakes DLG, Ferguson MM, Malmquist HJ (1989) Segregation in spawning and early life history among polymorphic Arctic charr, Salvelinus alpinus, in Thingvallavatn, Iceland. *Journal of Fish Biology*, **35**, 225–232.

Skúlason S, Snorrason SS, Ota D, Noakes DLG (1993) Genetically based differences in foraging behaviour among sympatric morphs of arctic charr (Pisces: Salmonidae). *Animal Behaviour*, **45**, 1179–1192.

Smith TB, Skulason S (1996) Evolutionary Significance of Resource Polymorphisms in Fishes, Amphibians, and Birds. *Annual Review of Ecology and Systematics*, **27**, 111–133.

Snorrason SS, Skúlason S (2004) Adaptive speciation in northern fresh water fishes — patterns and processes. In: *Adaptive speciation* (eds Dieckmann U, Metz H, Doebeli M, Tautz D), pp. 210–228. Cambridge University Press, Cambridge,.

Snorrason SS, Skulason S, Sandlund OT *et al.* (1989) Shape polymorphism in sympatric Arctic charr, Salvelinus alpinus in Thingvallavath, Iceland. *Physiology and ecology Japan*, **1**, 393–404.

Volpe JP, Ferguson MM (1996) Molecular genetic examination of the polymorphic Arctic charr Salvelinus alpinus of Thingvallavatn, Iceland. *Molecular ecology*, **5**, 763–72.

Wainwright PC (1996) Ecological Explanation through Functional Morphology: The Feeding Biology of Sunfishes. *Ecology*, **77**, 1336.

Wainwright P, Shaw S (1999) Morphological basis of kinematic diversity in feeding sunfishes. *The Journal of experimental biology*, **202 Pt 22**.

Willacker JJ, von Hippel FA, Wilton PR, Walton KM (2010) Classification of threespine stickleback along the benthic-limnetic axis. *Biological journal of the Linnean Society. Linnean Society of London*, **101**, 595–608.

Paper 1

2 Paper I

Differentiation at the MHCIIα and Cath2 Loci in sympatric Salvelinus Alpinus Resource Morphs in Lake Thingvallavatn

Kalina H. Kapralova, Johannes Gudbrandsson, Sigrun Reynisdottir, Cristina B. Santos,Vanessa C. Baltana´ s, Valerie H. Maier, Sigurdur S. Snorrason, Arnar Palsson

Authors' contribution: Conceived and designed the experiments: AP KHK VHM SSS. Performed the experiments: KHK JG SR CBS VCB VHM SSS AP. Analysed the data: KHK JG VHM AP. Contributed reagents/materials/analysis tools: KHK SR VHM AP. Wrote the paper: KHK JG VHM SSS AP. Designed experiments: AP KHK VHM SSS. Molecular work: KHK SR. Parasite analyses: JG CBS KHK SR. Aged the specimens: VCB SSS.

OPEN ACCESS Freely available online

PLOS ONE

Differentiation at the *MHCIIα* and *Cath2* Loci in Sympatric *Salvelinus alpinus* Resource Morphs in Lake Thingvallavatn

Kalina H. Kapralova, Johannes Gudbrandsson, Sigrun Reynisdottir, Cristina B. Santos, Vanessa C. Baltanás, Valerie H. Maier, Sigurdur S. Snorrason, Arnar Palsson*

Institute of Life and Environmental Sciences, University of Iceland, Reykjavik, Iceland

Abstract

Northern freshwater fish may be suitable for the genetic dissection of ecological traits because they invaded new habitats after the last ice age (~10.000 years ago). Arctic charr (*Salvelinus alpinus*) colonizing streams and lakes in Iceland gave rise to multiple populations of small benthic morphotypes, often in sympatry with a pelagic morphotype. Earlier studies have revealed significant, but subtle, genetic differentiation between the three most common morphs in Lake Thingvallavatn. We conducted a population genetic screen on four immunological candidate genes *Cathelicidin 2* (*Cath2*), *Hepcidin* (*Hamp*), *Liver expressed antimicrobial peptide 2a* (*Leap-2a*), and *Major Histocompatibility Complex IIα* (*MHCIIα*) and a mitochondrial marker (D-loop) among the three most common Lake Thingvallavatn charr morphs. Significant differences in allele frequencies were found between morphs at the *Cath2* and *MHCIIα* loci. No such signal was detected in the D-loop nor in the other two immunological genes. In *Cath2* the small benthic morph deviated from the other two ($F_{ST}=0.13$), one of the substitutions detected constituting an amino acid replacement polymorphism in the antimicrobial peptide. A more striking difference was found in the *MHCIIα*. Two haplotypes were very common in the lake, and their frequency differed greatly between the morphotypes (from 22% to 93.5%, $F_{ST}=0.67$). We then expanded our study by surveying the variation in *Cath2* and *MHCIIα* in 9 Arctic charr populations from around Iceland. The populations varied greatly in terms of allele frequencies at *Cath2*, but the variation did not correlate with morphotype. At the *MHCIIα* locus, the variation was nearly identical to the variation in the two benthic morphs of Lake Thingvallavatn. The results are consistent with a scenario where parts of the immune systems have diverged substantially among Arctic charr populations in Iceland, after colonizing the island ~10.000 years ago.

Citation: Kapralova KH, Gudbrandsson J, Reynisdottir S, Santos CB, Baltanás VC, et al. (2013) Differentiation at the *MHCIIα* and *Cath2* Loci in Sympatric *Salvelinus alpinus* Resource Morphs in Lake Thingvallavatn. PLoS ONE 8(7): e69402. doi:10.1371/journal.pone.0069402

Editor: Gen Hua Yue, Temasek Life Sciences Laboratory, Singapore

Received January 3, 2013; **Accepted** June 9, 2013; **Published** July 24, 2013

Funding: The Palsson laboratory is supported by Icelandic Research foundation and the University of Iceland research fund. Icelandic research foundation (grant of excellence: nr 100204011) to S.S. Sigurdsson, A. Palsson, B.K. Kristjansson, Zophonias O. Jonsson and Ian A. Johnston paid for part of this work. Kalina H. Kapralova and Johannes Gudbrandsson were supported by the University of Iceland doctoral fund. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: apalsson@hi.is

Introduction

Processes of divergence and adaptation reflect evolutionary forces that alter the genetic make-up of populations over time [1]. While the bulk of these changes must be neutral, some are likely driven by natural selection. By identifying genes relating to adaptation we may be able to disentangle history, neutral forces and the contribution of positive and purifying selection on these evolutionary processes [2,3]. One approach to identify such loci is to dissect the molecular genetics of major adaptations in highly divergent species [4], another is to compare genetic architecture of adaptive traits between closely related species or populations [5]. One of the advantages in studying recent (or ongoing) divergence is that relatively few genetic changes differentiate populations or sibling species, compared to the vast number of changes separating major taxa. A potential downside to this approach is that, on short evolutionary time scale, divergence is mainly shaped by drift and fine tuning of preexisting adaptations. However, certain study systems have the advantage of rapid evolution, for instance when species respond to geographic catastrophes or when they colonize novel habitats [6,7].

Following the retreat of the last ice age cap (~10,000 years ago) anadromous and freshwater fishes in the northern hemisphere invaded and explored new habitats [8]. In some cases streams and lakes provided novel niches, which the colonizing populations may have adapted to. Multiple species (white fish, three spine sticklebacks, several salmonids) show signs of repeated adaptive changes in independent waterbodies [9–13], some of which have been dissected genetically [14–17].

Evolutionary Immunology of Fishes

The invasion into new habitats, changes from anadromous to "freshwater only" lifestyle, and sharing of habitat with other fishes provides novel challenges to the immune system of fishes [8]. The adaptive significance of immunological genes has been clearly illustrated. There are data supporting the role of frequency dependent selection, importance of local adaptation, the role of

generalist vs. specialist lifestyle and parasites, involvement in assortative/disassortative mating and even magic trait sympatric speciation as defined by [18], see [19] for review.

Fish possess both an adaptive and an innate immune system. The Major Histocompatibility complex (MHC) are cell surface molecules (class I on most cells and class II on specialized cells) that are involved in pathogen recognition and are central to adaptive immunity [20–22]. The MHCII is a heterodimer protein made of an α and a β chain, each with two domains (α1 and α2, β1 and β2 respectively). MHC genes have been identified in many teleost species and in general the β chain tends to be highly polymorphic [23]. The favoured explanation is that the multitude of infectious agents and environmental heterogeneity favours heterozygotes and rare alleles, which through balancing or frequency dependent selection result in high MHC diversity [19]. MHC allele diversity can be reduced in fish populations, as a consequence of local adaptation [24,25]. The distribution of *MHCIIα* alleles in Arctic charr is consistent with some degree of local adaptation [26], which will be studied further in this paper. Similarly data from brown trout (*Salmo trutta)* and Atlantic salmon (*S. salar*) how population differentiation in immunological genes, including TAP (Transporter associated with antigen processing) and interleukin-1 beta [27,28]. Curiously MHCII genes have been lost in Atlantic cod and related species [29], whereas in the Salmonidae they were duplicated along with the whole genome about 25–100 millon years ago [30]. There are two MHCII regions in Salmonids (observed in Atlantic salmon and rainbow trout (*Oncorhynchus mykiss*)), and evidence suggests at least four *MHCIIα* copies can be expressed [31].

The innate immunity system constitutes an evolutionarily old defense strategy, as the majority of gene families involved in it are present throughout the animal kingdom [32]. Innate immunity depends on a wide array of recognition, signal transduction and defence molecules, which are thought to evolve fast in response to pathogens. For instance, a comparison of 12 Drosophila species genomes revealed signs of positive selection on protein sequence and gene copy number in the sensory and effector genes of the innate immunity [33]. Innate immunity is considered to be of key importance in combating infections in fish [21,22]. Antimicrobial peptides (AMPs) play a major role in this system and in mammals these cationic peptides not only kill bacteria, but are multifunctional effectors of the innate immune system [34,35]. Many AMPs have been identified in fish including Cathelicidins (Cath), liver expressed antimicrobial peptides (LEAP) and hepcidins (HAMP) [36–40]. In salmonids two types of Cathelicidins have been identified; Cathelicidin 1 and 2 [39–41]. Cathelicidins are generally encoded by four exons with the exception of *Cathelicidin 2* (*Cath2*) in the *Salvelinus* genus, which have lost exon 3. In fish Cathelicidins expression increases due to bacterial infection and the mature antimicrobial peptide has been shown to have bactericidal activity [39,40,42–44]. Several studies have shown signs of positive selection on AMPs (reviewed by Tennessen [45]), specifically on the charged amino-acids. Population genetic studies of the AMPs and other innate immunity genes are needed to elucidate the distinct selection pressures that shape these ancient defense systems.

Arctic Charr Diversity and Resource Polymorphism

Arctic charr is a widespread circumpolar species. While it's distribution reaches south along the coastal areas of the N-Atlantic it is best described as an Arctic species and indisputably the most cold tolerant of the salmonids [46]. In the high north Arctic charr is often found in very cold waters and lakes with limited productivity and with few or no other fish species present. A body of ecological studies document high diversity among Arctic charr populations (e.g. refs. in [46–48]), and many instances of resource polymorphism within lakes (see refs. in [8,49,50]). The favored explanation is that diversity arises via ecological specialization in habitat use and diet, facilitated by relaxed inter-specific competition, leading to morphological divergence among and within lakes [8,51].

Icelandic Arctic charr descend from European charr [52] that colonized the island after the glacial retreat. Large parts of Iceland are constantly shaped by tectonic and volcanic activity which appear to have created special habitats for dwarf forms of Arctic charr that typically inhabit streams, ponds and lakes in the neo-volcanic zone that traverses Iceland from the south-west to the north-east. Kristjansson and coworkers have shown that in these habitats these small fish show similar phenotypes across locations, *e.g.* a typically benthic morphology, thus retaining a juvenile morphotype [53]. However, their evidence also shows that the morphological parallelism is incomplete [54,55]. In lakes with two or more distinct morphs they usually conform to two types in terms of morphology (i.e. morphotypes), a pelagic and a benthic type, that typically reflect their modes of habitat utilization. Multiple lines of evidence show that these differences stem both from environmental and genetic causes [56–58].

The best studied and most extreme example of sympatric charr morps are the four morphs in Lake Thingvallavatn [59]. Two large morphs are found, a large benthivorous (LB-charr) and a piscivorous morph (PI-charr), and two small forms (morphs), a small benthivorous (SB-charr) and planktivorous morph (PL-charr). PL- and PI-charr, display a pelagic morphotype and are more inclined to operate in open water and feed on free swimming prey, planktonic crustaceans and small fish, respectively. The two benthic morphs show a benthic morphotype and mainly reside on the bottom, feeding exclusively on benthic invertebrates. The very small size of the SB-charr also allows them to utilize interstitial spaces and crevices in the littoral zone typically consisting of submerged lava which offers a rich source of benthic invertebrate prey. As would be expected from the clear cut ecological diversification of the morphs their macroparasitic fauna differs distinctively [60].

Population genetic studies based on variation in mtDNA revealed a common ancestry of Arctic charr in the Nordic countries, Ireland and Iceland [52]. Within Iceland, allozyme, mtDNA and microsatellite data reveal significant genetic differences between localities and in some cases between sympatric morphs, like the four morphs in Lake Thingvallavatn [61–63]. The genetic differentiation among the Thingvallavatn morphs is rather weak however, the average F_{ST} over 10 microsatellites being 0.03, and a coalescence model suggests a scenario of early divergence with subsequent barriers to gene flow [63]. The strongest indication of genetic differentiation between sympatric charr morphs is a fixed difference in one microsatellite marker between two morphotypes in Lake Galtabol [64]. On a larger scale the available data suggest repeated evolution of dwarf forms (small fish with a benthic phenotype) in numerous Icelandic lakes and stream habitats in the neo-volcanic zone [53,63].

Molecular genetics have also been used to address the developmental basis of morphotype differences in Icelandic Arctic charr [65,66]. Macqueen and colleagues [66] conducted a study of the expression of 21 mTOR and growth regulation genes in 7 distinct Icelandic charr populations (thereof 5 with a small benthic morphotype), and revealed substantial divergence in gene expression of many pathway components. For instance mTOR is less and 4E-BP-1 more highly expressed in the populations of small benthic populations compared to other populations, a

finding consistent with the role of these genes in protein synthesis and growth regulation [55,66]. It is not clear whether those pathways are the foci of selection for changes in size and form, or realisitors of genes that promote dwarfism. Notably, considering our focus on immunological genes, the mTOR pathway is also involved in regulation of innate immunity [67,68].

We hypothesized that local differences in habitat use and diet between the morphs in Lake Thingvallavatn and among other Arctic charr populations and morphotypes in Iceland could impact variation in important immunological genes. Using samples from all major phylo-geographic groups of Arctic charr [52] Conejeros and colleagues [26] reported on rich allelic variation at the *MHCIIα* locus within and between charr populations. Their data showed considerable shared diversity within populations and across a broad geographic range, but are also consistent with differentiation among populations reflected in unique haplotypes and frequency differences. Here we present a study on a smaller geographic scale analyzing variation in *MHCIIα* and four other innate immunity genes in Icelandic Arctic charr. Our focus was on the three most common sympatric morphs from Lake Thingvallvatn and 9 populations of small benthic, anadromous and lake resident charr from the neo-volcanic zone (south, west and north) in Iceland – that we studied previously with 9 microsatellites [63]. Thus in this study we could interrogate local differences in gene frequencies and probe geographic patterns in these loci in small benthic charr in Iceland. The results indicate marked differentiation between sympatric morphotypes in Lake Thingvallavatn in two loci, *Cath2* and *MHCIIα* that we investigated further. Our findings have bearing on the understanding of those unique sympatric Arctic charr morphotypes, and immune system diversity in organisms with evolutionarily recent resource polymorphism.

Materials and Methods

Sampling

Specimens came from three collections of Arctic charr from Icelandic lakes and rivers. First, we utilized a sample of 30 large bentivorous charr (LB-charr, not sexed) caught on their spawning grounds at Olafsdrattur, and a total of 406 spawning small benthivorous charr (SB-charr, 102 females/83 males) and plantkivorous charr (PL-charr, 83 females/115 males) caught at Olafsdrattur and four other spawning locations in Lake Thingvallavatn in October 2005 (Table 1, Figure 1, inset) (for details see Kapralova *et al.* [63]). Second, we used another sample of 76 SB-charr (17 females/59 males), 102 PL-charr (51 females/males) and 17 LB-charr (1 female/16 males) collected in Olafsdrattur and Mjoanes, in September and October 2010 respectively. These two samples were pooled as our previous results [63] and the data from 2005, did not suggest genetic differentiation by location. The sampling in Lake Thingvallavatn focused on the SB and PL morphs, and the LB morph was mainly used for reference (hence the relatively lower sample size). For the 2010 sample, sex, fork length, weight, maturity and age were documented and parasite load (see below) assessed for every individual. DNA was extracted from a fin clip following a standard phenol-chloroform protocol. Third, we utilized samples from 9 populations of Arctic charr selected from a larger survey throughout Iceland collected in 2003–2006 (Table 1, Figure 1) previously described [63]. Those specimens were not sexed.

Fishing in Lake Thingvallavatn was with permissions obtained both from the owner of the land in Mjóanes and from the Thingvellir National Park commission. Ethics committee approval is not needed for regular or scientific fishing in Iceland (The Icelandic law on Animal protection, Law 15/1994, last updated with Law 157/2012). However, sampling was performed with University College Aquaculture Research Station (HUC-ARC) personnel. HUC-ARC has an operational license according to Icelandic law on aquaculture (Law 71/2008), that includes clauses of best practices for animal care and experiments.

Molecular Work and Data Processing

We screened for sequence variation in four immunological genes: *Cath2, Leap-2a, Hamp* and *MHCIIα* among the three Thingvallavatn morphs (SB-, PL- and LB-charr). Moreover we studied a 510 bp region of the D-loop (starting at base 25 in the *S. alpinus* mtDNA reference genome, accession number NC_000861.1) as a putative neutral marker or marker of maternal lineage sorting. Loci were amplified by PCR with TEQ polymerase (Prokaria-Matis). We used previously published primers for *MHCIIα* [26] and new primers for *D-loop*, *Leap-2a*, *Hamp* and *Cath2* (Table S1), designed with Primer3 (http://primer3.wi.mit.edu/[69]). The following PCR program was used for all primer pairs, except *MHCIIα*. Denaturation at 95°C for 5 min; 35 cycles of 95°C for 45 seconds; 45 seconds at a marker specific annealing temperature (Table S1); 1 min at 72°C, then a final step of 10 min at 72°C. For *MHCIIα* we used touchdown PCR, initial denaturation at 94°C for 5 min; 16 cycles of 94°C for 45 seconds, 62°C for 45 seconds (decreasing by 0.5°C every cycle), 1 min at 68°C; followed by 25 cycles of 94°C for 45 seconds, 53°C for 45 seconds, 1 min at 68°C; then a final step of 10 min at 68°C. PCR products were ExoSap purified, sequenced (BigDye) and run on an Applied Biosystems 3500xL Genetic Analyzer (Hitachi).

Raw sequencing data was base-called by Sequencing Analysis Software v5.4 with KBTMBasecaller v1.41 (Applied Biosystems), and run through Phred and Phrap [70], prior to trimming primer sequences, visual editing of ambiguous bases and putative polymorphisms in Consed [71]. Fasta files were exported and aligned with ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/, [72]) and manually inspected for alignment errors in Genedoc (www.psc.edu/biomed/genedoc) [73]. All sequences where deposited as Popsets in Genebank under the accession numbers KC590653-KC591103, KC591105-KC591218, KC591220-KC591303, KC591303-KC591626 and KC596075-KC596117.

Genotyping *MHCIIα*

Due to potential duplications or deletions of *MHC* genes and the ancestral genome duplications in salmonids [30] the presence of *MHC* paralogous genes has to be investigated in charr. Initially we used the SAALDAA primers from Conejeros *et al.* [26], (Table S1) that pick up part of exon 2 and intron 2 of *MHCIIα*, but obtained several satellite bands. To confirm the amplification of MHC, bands of various sizes (from a non-optimized PCR) were cloned into a TOPO vector (Invitrogen) and sequenced. Blastn was used to find related sequences in Genebank (NCBI – nucleotide collection – at latest in April 2013). We obtained bands from 4 size ranges. Most importantly, a ~400 bp fragment sequenced from 2 individuals (10 clones from each) yielded 3 different fragments of *MHCIIα* (Table S2). One of these fragments, represented by 5 clones from each individual, was 99% identical to Saal-DAA*0801 [26]. The other two versions, each restricted to one individual, had 98% and 99% identity to Saal-DAA*0305/0306/0307 and Saal-DAA*0305 [26], respectively (Table S2). The largest band (~720bp) was only present in ~1% of the samples and all ten clones from this band were identical to *MHCIIα* haplotype Saal-DAA*0104 (intron haplotype hap1 as defined by [26]. Two smaller fragments, ~250 bp and ~150 bp, contained mixed products of various origins unrelated to *MHCII*.

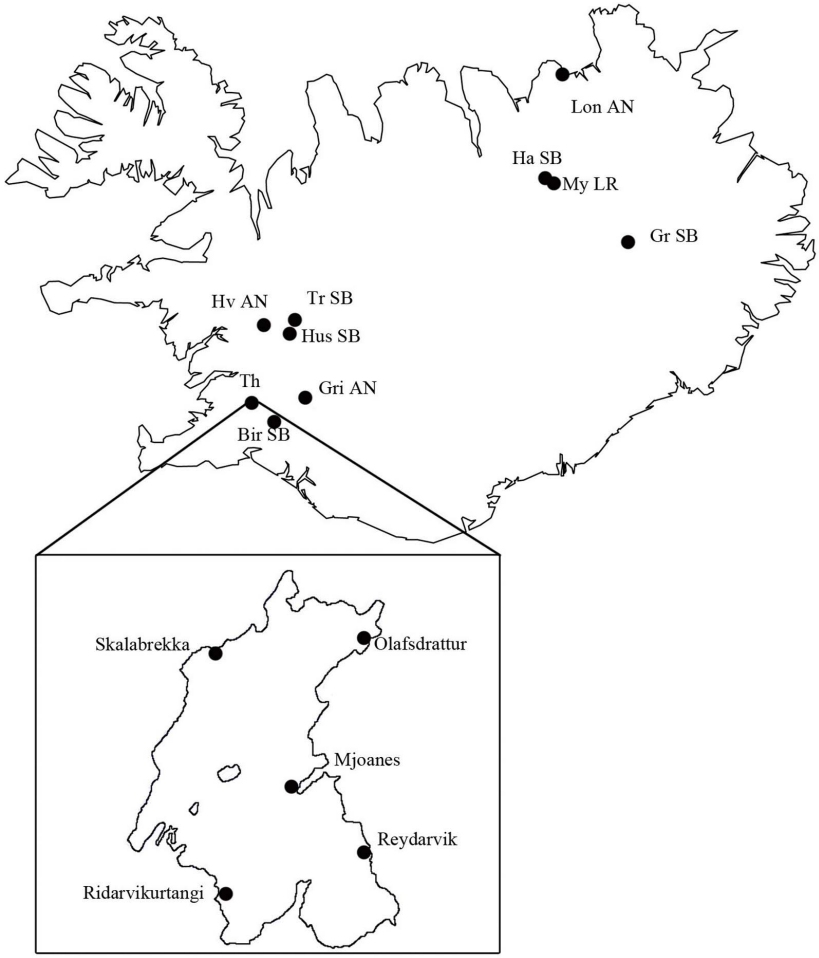

Figure 1. Sampling locations of Arctic charr in Lake Thingvallavatn and around Iceland. Fishes where collected in five locations within Lake Thingvallavatn (left), and from 9 other locations and populations around Iceland. In Lake Thingvallavatn, O: Olafsdrattur, M: Mjoanes, Re: Reydarvik, R: Ridvikurtangi and S: Skalabrekka. Around the island, either small benthic (SB) and lake resident (LR) or anadromous (AN) charr in Myvatn (My, LR), Haganes (Ha, SB), Lon (Lo, AN), Grafarlond (Gr, SB), Grimsnes (Gr, AN), Birkilundur (Bir, SB), Hvita (Hv, AN), Trussa (Tr, SB) and Husafell (Hus, SB).
doi:10.1371/journal.pone.0069402.g001

The PCR protocol was optimized to reduce unspecific small auxiliary bands (see above) and we proceeded with PCR and direct sequencing. The first 32 *MHCIIα* sequences from Lake Thingvallavatn (2005 sample) were amplified with the SAALDAA primers, and sequenced with both forward and reverse primers (error rate of Single nucleotide polymorphisms (SNP's) called was $<0.1\%$). Subsequently only the forward primer was used to sequence the PCR products. In total PCR and direct sequencing of 413 individuals from the 2005 sample gave sequences of three major types. Those corresponded to the large fragment (intron haplotype hap1) and the two versions (similar to Saal-DAA*0303 and Saal-DAA*0305), that we denote as second intron haplotypes 14 and 15. The fragment identical to Saal-DAA*0801 was never observed. PCR and direct sequencing clearly revealed individuals heterozygotic for a single base insertion/deletion polymorphism (indel) in the intron. To us the data suggest that two *MHCIIα* paralogous genes are present in Arctic charr, with hap14, hap15 and possibly hap1 being alleles of one paralog. The optimized PCR preferentially amplifies this paralog. This is supported by two observations. First, in the direct sequencing we never observe Saal-DAA*0801 (the two suspected paralogs are easy to distinguish) and second, the indel in the second intron conforms to Hardy Weinberg Equilibrium, within each morph (see below).

Because of low DNA availability and degradation in the 2005 Icelandic lake samples, we designed new primers (Table S1.) that gave a shorter amplicon and none of the satellite bands. With

Table 1. Details on sampling locations and the number of individuals collected in 2005 and 2010.

Location	Morphotype	Code	Latitude	Longitude	2005	2010
Thingvallavatn	Large benthic	TH_LB	64°11	21°08	30	17
Thingvallavatn	Small benthic	TH_SB	64°11	21°08	185	76
Thingvallavatn	Planktivorous	TH_PL	64°11	21°08	198	102
Grimsnes	Anadromous	Gri_AN	64°00	20°53	27	
Birkilundur	Small benthic	Bir_SB	64°01	20°57	30	
Hvita	Anadromous	Hv_AN	64°42	20°59	35	
Trussa	Small benthic	Tr_SB	64°43	20°46	29	
Husafell	Small benthic	Hus_SB	64°41	20°52	31	
Lon	Anadromous	Lon_AN	66°05	16°55	27	
Grafarlond	Small benthic	Gr_SB	65°15	16°09	31	
Myvatn	Lake resident	My_LR	65°37	17°03	34	
Myvatn-Haganes	Small benthic	Hag_SB	65°37	17°03	35	

doi:10.1371/journal.pone.0069402.t001

those primers fragments of *MHCIIα* from 6 individuals were amplified, cloned and sequenced (as before). We sequenced on average 8 clones per individual and in all cases the genotyping was in perfect concordance with the genotyping from PCR and direct sequencing. The suspected paralogous copy of *MHCIIα* (similar to Saal-DAA*0801) was found in a low proportion of the clones (5/45 sequences). The 2010 sample from Lake Thingvallavatn and the 9 Iceland wide populations were amplified and sequenced with these primers. Although there is a potential for ascertainment bias, as samples from two years (2005 and 2010) were genotyped with different primers, the results do not indicate a bias; the frequency of the indel variation was not statistically different between years (tested within morphs, see details below). Finally, we also did a restriction enzyme analysis, that could distinguish hap14 and hap15 on basis of a G/A polymorphism 13 bp down stream of the indel (TGAATGAATCAATAGGATTAATGTAGTAAA(A/−)TAGTCACCTCACT(G/A)TAACCTCTCACATGTTG-TATCATCTGTGGTATGG). These two polymorphisms were fully coupled in the sequencing data. This restriction digest of 28 individuals (equal number from 2005 and 2010) was in perfect concordance with the PCR and sequencing data.

Population Genetic Analyses

Tassel version 2.0.1 (www.maizegenetics.org) [74] and DNAsp 4 (www.ub.edu/dnasp/) [75] were used to calculate and analyze population genetic statistics. Tests of Hardy Weinberg proportions, allele and genotype frequencies between morphs, locations and were implemented in R (version 2.12, R Development Core Team, 2011). Arlequin v3.5.1.2 was also used to estimate F_{ST} [76–78]. We tested determinants of genetic differentiation between morphs within Lake Thingvallavatn with analyses of molecular variance (AMOVA) using Arlequin. We analyzed variation in 3 amplicons (*D-loop*, *Cath2* and *MHCIIα*), within Lake Thingvallavatn with a two level AMOVA with morph (LB, SB, SP) as a categorical variable, split by sex or sampling location.

The genetic relationships between and within morphs were estimated with an unrooted neighbor-joining tree. The tree was constructed using Cavalli–Sforza's genetic distances obtained from nine microsatellite loci [63] with the program NEIGHBOUR available in PHYLIP3.69 [79]. Confidence intervals were estimated by 1000 bootstrap replicates.

Parasite Analyses

The 2010 samples from Lake Thingvallavatn were used to assess infection rates and loads of the eye parasite *Diplostomum sp.*, the intestine parasite *Eubothrium salvelini*, Nematodes and *Diphyllobothrium sp.* Both eyes were extracted from each individual. The contents of each eye was poured on a flat slide, covered with a slip and processed under a Leica KL200 LED microscope at 2X magnification. The slide field was divided into 45 blocks, and the average number of metacercaria of *Diplostomum sp.* was estimated. We first screened all blocks, and in case of even distribution among them, counted the metacercaria in 5 randomly selected blocks, and then calculated average infection rate. In case of non-uniform distribution or low infection we counted the parasites in all 45 blocks. We recorded both counts and used an infection scale [60] ; 0 = total absence of parasites; 1 = 1 or fewer parasites per blocks; 2 = 1 to 3 individuals per block; 3 = 4 to 10 parasites per block and 4 represented more than 10 *Diplostomum sp.* individuals per block. The estimation was done by a single observer (S. Reynisdottir) on a single eye per specimen. The correlation of infection rate between eyes was high (Pearson $r = 0.75$, $p < 0.005$, for 25 pairs of eyes studied).

Infections by *Eubothrium salvelini* were assessed by carefully extracting the liver, stomach and intestine and documenting the presence or absence of the adult tapeworm. Infections of nematodes and plerocercoids of *Diphyllobothrium sp.* were estimated by counting individual nematodes and *Diphyllobothrium* cysts internal cavities and linings of flesh [60,80]. The *Diphyllobothrium sp.* infection rate was scored using the following infection scale: 0 = the total absence of parasites; 1 = 1 to 3 per individual; 2 = 4 to 7 per individual and 3 equaled more than 8 parasites per individual. For Nematodes the number per individual was recorded. All data on intestinal parasites were obtained by a single observer (C. B. Santos). Data of the 2010 and 2005 samples from Lake Thingvallavatn were deposited in the Dryad Repository: http://dx.doi.org/10.5061/dryad.81884.

Statistical Analyses of Parasite Infections

Statistical analyses were performed in R. The effects of morph, sex and weight on the load of individual parasite species was investigated with multivariate regression. Summary statistics were calculated for weight, age and parasite loads separately for each morph. Sex ratio was also calculated. For *Diphyllobothrium sp.* and

Table 2. Polymorphism in the mitochondrial D-loop and three immunological genes.

Gene/region	Morph	Size (bp)	N	S	Indel	π	θ	Haplotypes
D-loop	All	509	406	4	0	0.001	0.001	7
	PL	509	190	3	0	0.001	0.001	4
	SB	509	216	4	0	0.001	0.001	7
Hamp 5′ UTR	PL/SB*	454	12	0	0	0.000	0.000	1
Leap-2a 3′ UTR	All	559	15	4	1	0.001	0.004	3
	PL	559	8	2	0	0.001	0.003	2
	SB	559	7	3	1	0.002	0.003	3
Cath2 (intron 2)	PL/LB/SB*	219	258	0	0	0.000	0.000	1
Cath2 (peptide)	All	396	258	3	0	0.001	0.001	4
	PL	396	138	2	0	0.000	0.001	3
	LB	396	35	1	0	0.000	0.001	2
	SB	396	86	3	0	0.001	0.001	4
Cath2 (3′ UTR)	All	407	17	2	1	0.002	0.002	3
	PL	407	6	1	0	0.002	0.001	2
	SB	407	11	2	1	0.002	0.002	3

S: Segregating sites. Indel: Segregating insertion/deletion polymorphism. π: The average number of nucleotide differences per site. θ: Wattersons estimator of diversity per site. *The data from different morphs are summarized together as no differences in frequence were observed.

doi:10.1371/journal.pone.0069402.t002

Diplostomum sp. mean relative density (MRD) was calculated [60]. Statistical models for parasite load were applied to morph pairs to test for difference between the morphs. As parasite loads turned out to be different between morphs tests for other factors affecting the load were applied to the morphs separately. The models had the general structure:

$$Parasite\ load = Sex + Weight + Age + Error.$$

A term for genotype was also added to evaluate the impact of *MHCIIα* variation within morphotypes. The ANOVA function from the *car* package [81] was used to perform F-tests and log-likelihood tests. Raw counts of *Diphyllobothrium sp.* and *Diplostomum sp.* were analyzed by multivariate linear regression and variable effects tested with an F-test. The infections were also summarized with an infection scale [60] and analyzed using multinomial logit regression fitted with neural networks [82], with consistent results. Effects were tested with log-likelihood tests. Logistic regressions were applied to Nematodes and *Eubothrium salvelini* occurrence and effects were tested with log-likelihood test.

Results

Nucleotide Polymorphism in Arctic Charr Morphs in Lake Thingvallavatn

Different molecular markers have revealed significant but weak genetic differentiation among the Lake Thingvallvatn charr morphs [61–63]. Here we make use of genetic material from individuals previously typed for 9 microsatellite markers [63] to explore variation in four immunological loci, and test for indications of population differentiation.

Four segregating sites were observed in the mitochondrial *D*-loop, but nucleotide diversity was rather low (Table 2). Of the four substitutions only one (m38A>G) had significant difference in frequency between PL and SB (χ^2 [1] = 9.36, p = 0.002). The F_{ST} = 0.001, which was lower than the F_{ST} for microsatellites between charr morphs in Lake Thingvallavatn [63]. A comparison with *S. alpinus D*-loop in genebank [52,83] shows that none of the four *D*-loop sites are restricted to Iceland. Analyzes of molecular variance (AMOVA) confirm that the observed variation in this part of the mtDNA of Lake Thingvallavatn charr is not affected by morph, sex or sampling location (Table 3).

We screened three innate immunity genes *Hamp*, *Leap-2a* and *Cath2* for nucleotide variation. The 454 bp *Hamp* amplicon, positioned in the untranslated 5′-region, proved invariant in a set of 12 specimens (6 PL- and 6 SB-charr). Four segregating sites and one insertion/deletion polymorphism (indel) were found in the 3′UTR of *Leap-2a*. These were at approximately equal frequency in SB- and PL-charr. The *Hamp* and *Leap-2a* genes where not studied further. Of the three regions surveyed in the *Cathelicidin* gene (spanning ~1 kb), only the peptide region showed frequency differences between morphs (Table 2) urging further investigation. The sequenced part of intron 2 was invariant in the sample, whereas the three mutations (one indel and two SNPs) in the 3′UTR were at about the same frequency in both morphs.

Sequencing of the antimicrobial peptide encoding region of *Cath2* in 264 individuals from Lake Thingvallavatn 2005 revealed three variant sites (including one singleton). One mutation (g558C>A) was found in intron 2. Another (g819C>A) was found in the exon encoding the mature antimicrobial peptide (in cathelicidins this region is on exon 4, but due to the lack of exon 3 in charr *Cath2* [40], it is encoded by the third exon in *S. alpinus*, Figure 2A). This mutation is predicted to lead to an amino acid replacement in the mature peptide (replacement of arginine by serine at position 115, Figure 2B). This alters the charge of the peptide, from +8 to +7.

We compared the frequency of the two mutations among morphs, sex and sampling locations in Lake Thingvallavatn. The g558C>A is largely restricted to the SB morph (11.3% frequency); it is not found in the LB-charr and only present in two of 134 PL-charr. The more common g819C>A variant shows significant frequency differences between morphs (χ^2 [2] = 43.91, p<0.0001). The A allele is at 27% frequency in SB-charr, but is rarer in LB-

Table 3. Analyses of molecular variance (AMOVA) of three loci by morphotypes (PL, LB and SB collected in 2005) and either location or sex.

Gene	Terms	d.f.	Sum of squares	Variance	Variation (%)	Fixation index	p-value
D-loop*	Among morphs	1	0.3	0	0.83	FSC : −0.01	ns.
	Among locations within morphs	7	0.39	0	−1.44	FST : −0.01	ns.
	Within locations	389	58.58	0.15	100.62	FCT : 0.01	****
	Total	397	59.27	0.15			
	Among morphs	1	0.17	0	0.4	FSC : −0.01	ns.
	Among sexes within morphs	2	0.11	0	−0.66	FST : 0	ns.
	Within sexes	393	58.17	0.15	100.26	FCT : 0	****
	Total	396	58.45	0.15			
Cath2	Among morphs	2	4.56	0.02	12.64	FSC : 0.03	****
	Among locations within morphs	7	2.2	0.01	3.02	FST : 0.16	**
	Within locations	253	42.18	0.17	84.34	FCT : 0.13	*
	Total	262	48.94	0.2			
	Among morphs	2	4.56	0.03	13.48	FSC : 0.01	****
	Among sexes within morphs	2	0.45	0	0.47	FST : 0.14	ns.
	Within sexes	258	43.94	0.17	86.05	FCT : 0.13	****
	Total	262	48.94	0.2			
MHCIIα	Among morphs	2	50.76	0.22	63.2	FSC : 0.03	****
	Among locations within morphs	8	0.88	0	−0.13	FST : 0.63	ns.
	Within locations	402	50.93	0.13	36.92	FCT : 0.63	***
	Total	417	102.57	0.34			
	Among morphs	2	51.44	0.22	64.06	FSC : −0.01	****
	Among sexes within morphs	2	0.04	0	−0.33	FST : 0.64	ns.
	Within sexes	408	50.91	0.12	36.28	FCT : 0.64	****
	Total	412	102.39	0.34			

*Only PL and SB were sequenced for the D-loop. d.f.: Degrees of freedom. Significance: ns. $p>0.05$,
*$p<0.05$,
**$p<0.01$,
***$p<0.001$,
****$p<0.001$.
doi:10.1371/journal.pone.0069402.t003

(5.7%) and PL-charr (6.4%). This translates into an F_{ST} of 0.17 ($p<0.0001$) between the SB- and PL morphs, and $F_{ST}=0.13$ ($p<0.0001$) between the LB and SB samples. No differences in allele frequency where found between PL- and LB-charr, sexes or sampling locations. Analyses of Molecular Variance (AMOVA) confirmed these patterns (Table 3).

MHCIIα Variation in Lake Thingvallavatn

Due to the structural richness of MHC regions and the fact that the common ancestor of salmonids underwent a whole genome duplication, studies of MHC variation in those species are rather complicated. We tackled this by genotyping with PCR and direct sequencing, and assessed the specificity and reproducibility of this genotyping method by cloning and restriction enzyme assays.

We concentrated on the highly variable intron 2 of *MHCIIα* [26], by DNA sequencing of 413 charr (LB, SB and PL) from Lake Thingvallavatn. There was high degree of polymorphism, with many segregating mutations (10 SNPs and 2 indels in ~300 bp). Two major and two minor versions of *MHCIIα* were identified. The two major haplotypes hap 14 and hap 15 are quite distinct, being separated by 6 segregating sites and 1 indel. These polymorphism were were described by Conejeros *et al.* [26], but the haplotypes involving them are unique and probably arose by recombination. In addition two rare versions were observed, hap16 (just one site diverged from hap14) and hap1 (Saal-DAA*0104) which contains a Hpa retrotransposon [26]. The hap1 and hap16 haplotype were extremely rare in all morphs, for instance hap1 was found in four SB-charr from 3 sampling locations (1.08%) and one LB-charr (1.67%). Our analyses focused on the two dominant haplotypes, hap14 and hap15.

As described in Materials and Methods, the cloning results suggest the presence of two distinct *MHCIIα* paralogs in Arctic charr in Iceland. One of these was never observed with the PCR and direct sequencing, but only detected in the cloning (prior to PCR optimization). The hap14 and hap15 haplotypes are readily distinguishable based on several markers, such as the indel in the intron. We are quite certain that these are allelic variations (true haplotypes, not paralogous genes) because Hardy Weinberg proportions are respected for the indel polymorphism in *MHCIIα* intron in all three morphs in Lake Thingvallavatn (LB: $\chi^2_{[1]}=0$, $p=1$, SB: $\chi^2_{[2]}=1.77$, $p=0.4$, PL: $\chi^2_{[2]}=6.2$, $p=0.05$). Furthermore restriction enzyme analysis of 28 individuals was in perfect concordance with the PCR and sequencing data.

As predicted [19,27] the nucleotide diversity was higher in *MHCIIα* than in the other sequences studied; π was an order of magnitude higher than for *Cath2* and the D-loop (Table 2 and

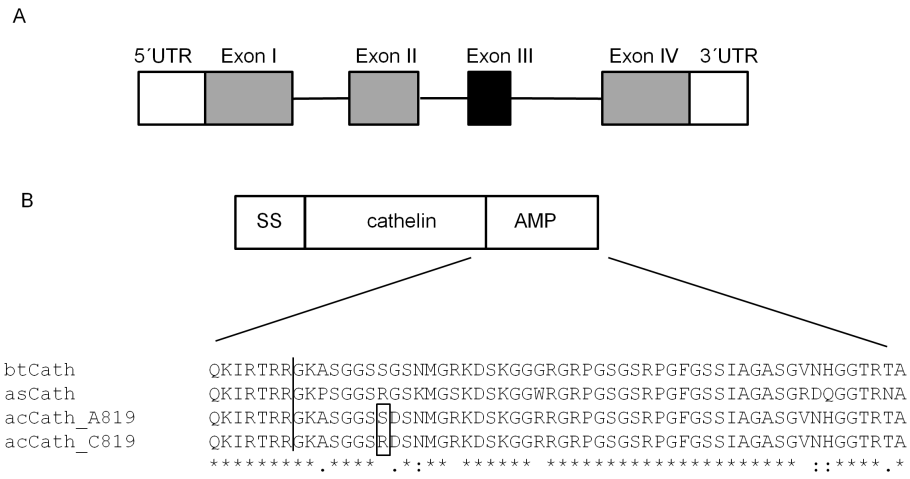

Figure 2. Polymorphism in the antimicrobial peptide Cathelicidin 2. *Cathelicidins* have a conserved 4 exon structure (A) with the exception of the *Salvelinus* Cathelicidins type 2 which have lost exon 3 (marked black). The peptides (B) are produced as pre-pro-peptides, where exon 1–3 encode the signal sequence (SS) and the conserved Cathelin region, while exon 4 encodes the processing site and the mature antimicrobial peptide (AMP). An amino acid alignment of this region for Cathelicidin 2 of Atlantic salmon (asCath), brook trout (btCath) and Arctic charr (acCath) shows the predicted processing site (vertical line) and the observed polymorphism (predicted peptide position 115) in Icelandic Arctic charr. Identical amino acids are marked with *, amino acids with a similar and somewhat similar function are marked with : and. respectively.
doi:10.1371/journal.pone.0069402.g002

below). We found large differences in *MHCIIα* frequencies among the three morphs studied from Lake Thingvallavatn (Table 4), with hap15 being dominant in both benthic morphs, 93.5% and 88.3% in SB- and LB-charr respectively. In contrast hap15 was at 22% frequency in the pelagic morph (PL). This translates into an F_{ST} of 0.56 ($p<0.0001$) between PL- and LB-charr, 0.67 between PL- and SB-charr ($p<0.0001$), and unsignificant F_{ST} between the two benthic morphs. This represents the strongest genetic differentiation reported to date between any of these three sympatric morphs. These findings were further supported by AMOVA, the effect of morphotype (benthic versus pelagic) dominating the explained variance (above 60%), while sex and sampling location did not have significant effects (Table 3).

This strong difference in *MHCIIα* frequency between morphotypes prompted several questions. Is the frequency difference consistent between years? What is the geographic distribution of variation in *MHCIIα* within Iceland? Do the haplotypes correlate with phenotypic attributes? We set out to answer these questions. Some hypotheses of MHC evolution involve temporal dynamics, for instance due to frequency dependent selection [19]. To evaluate this we used two approaches. We first compared the frequency of *MHCIIα* hap14 in the three morphs (PL-, LB- and SB-charr) in two cohorts sampled in 2005 and 2010 (Table 4). On all three morphs the haplotype frequencies were similar for the two years, χ^2 [1] = 0.0301, p = 0.9, χ^2 [1] = 0.08, p = 0.8, χ^2 [1] = 3.65, p = 0.06, for PL-, LB- and SB-charr, respectively. The age distribution was similar in the fishes collected, for instance the average age in PL sampled in 2005 and 2010 was 6.94 and 6.96 years respectively (weighted t-test, p = 0.96). No significant differences in haplotype frequency between years (χ^2 [2] = 0.59, p = 0.74) or age-classes were observed within morphs (Figure 3) (χ^2 [12] = 17.4, p = 0.13 for 2005 and χ^2 [12] = 10.5, p = 0.57 for 2010).

Cath2 and *MHCIIα* Polymorphism Across Morphotypes and Geographic Regions

As the frequency of variants both in *Cath2* and *MHCIIα* deviated significantly between morphs within Lake Thingvallavatn, we

Table 4. The frequency of the three most common *MHCIIα* haplotypes in the arctic charr morphotypes from Lake Thingvallavatn sampled in 2005 and 2010.

	LB		SB		PL	
Haplotypes	2005	2010	2005	2010	2005	2010
hap1	1 (1.7%)	0 (0.0%)	4 (1.1%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
hap14	6 (10.0%)	5 (14.7%)	20 (5.4%)	2 (1.3%)	309 (78.0%)	155 (76.0%)
hap15	53 (88.3%)	29 (85.3%)	346 (93.5%)	150 (98.7%)	87 (22.0%)	49 (24.0%)
N	30	17	185	76	198	102

N: the total number of fishes genotyped in each sample.
doi:10.1371/journal.pone.0069402.t004

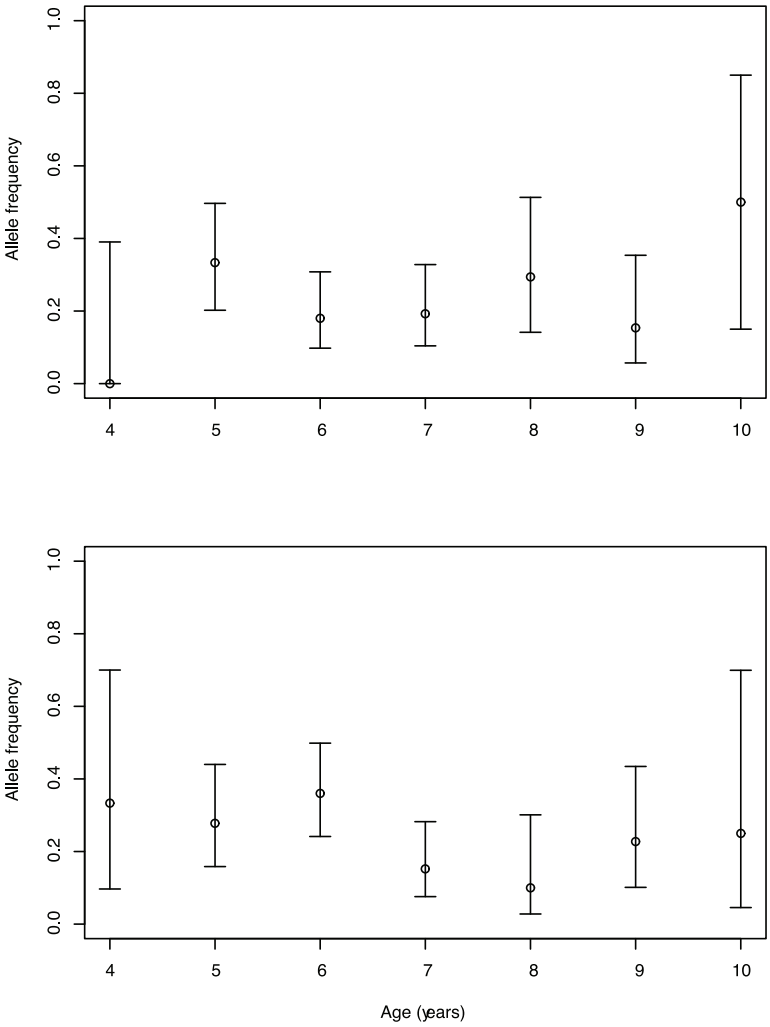

Figure 3. Frequency of *MHCIIα*variations in PL-charr from 2005 and 2010 by age classes. The frequency of *MHCIIα* hap14 (with 95% confidence intervals) by age of PL charr, collected in years 2005 (A) and 2010 (B) at the spawning grounds in Lake Thingvallavatn.
doi:10.1371/journal.pone.0069402.g003

wanted to know if the observations reflect a local or a broader geographic or ecological pattern. Our previous microsatellite study [63] enabled inference of relatedness among 9 Arctic charr populations from the north, west and south of Iceland (Figure 1 and 4A). We surveyed variations in both genes in those small benthic, anadromous and lake resident populations and superimposed on the microsatellite based tree.

There was very little polymorphism in *MHCIIα* in other populations and lakes, at maximum 3 haplotypes in each population (Table 5). The hap14 haplotype which dominated in the PL in Lake Thingvallavatn was only found in one other population (SB from Husafell), at 3% in 2 individuals (Figure 4B). The other haplotype (hap15), most common in the LB and SB morphs in Lake Thingvallavatn, dominated all other populations (average frequency 94%, lowest 81%). Several other haplotypes were observed, but all are one or few bases removed from hap15 and at very low frequency. The results show clearly reduced variation in this locus in Icelandic stocks of Arctic charr, except in the sympatric morphs in Lake Thingvallavatn. Summaries of nucleotide diversity reveal this pattern, as π (which responds to frequency and diversity of haplotypes) is larger in PL-charr from Lake Thingvallavatn than in the other charr populations surveyed (Table 5).

The *Cath2* g819C>A was genotyped in 7 populations (105 individuals total) and its frequency differed significantly between them (χ^2 [6] = 91.92, p<0.0001, Figure 4C). The g819C>A was dominant and even fixed in several small benthic charr populations (Birkilundur 100%, Haganes 86% and Grafarlond

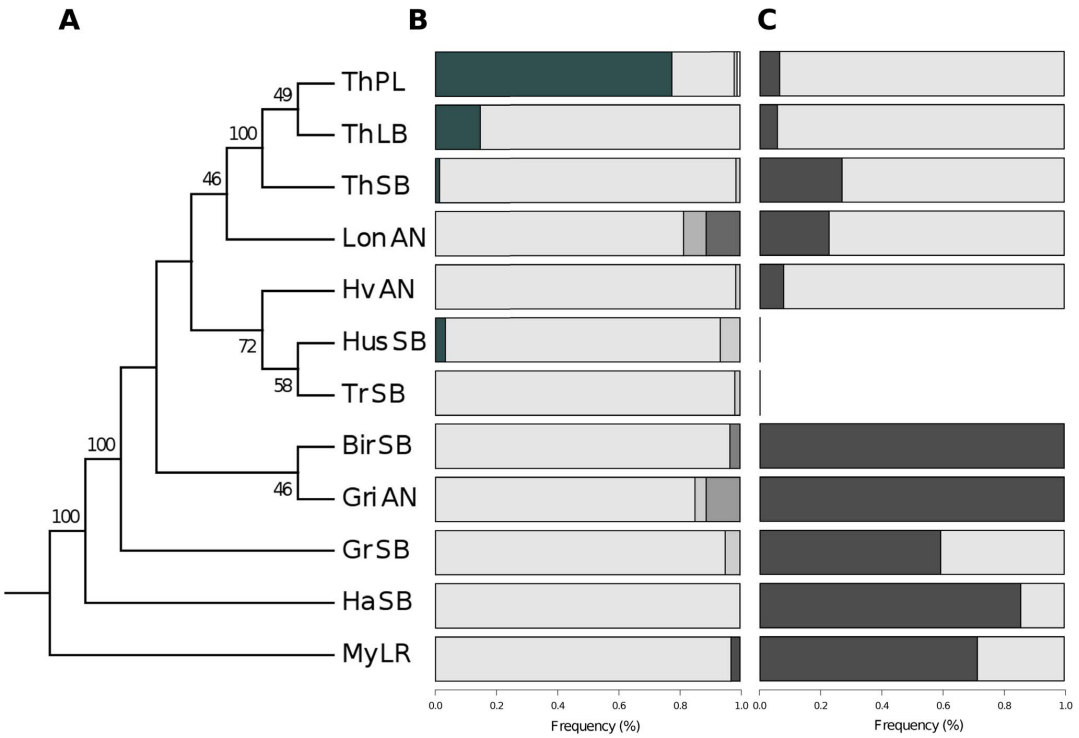

Figure 4. Arctic charr population history and variation in *Cath2* and *MHCIIα*. A) A genealogy of the sampled populations was built from 9 microsatellite markers and the confidence intervals were estimated by 1000 bootstrap replicates [63]. B) Frequencies of the *MHCIIα* intron haplotypes (hap 14 is dark, hap15 light gray, rare haplotypes are in intermediate shades of gray). C) The frequency of *Cath2* g819A (dark). Due to limited DNA available, the marker could not be typed in Husafell and Trussa. The same individuals where genotyped for all markers.
doi:10.1371/journal.pone.0069402.g004

59%). Recall, within Lake Thingvallavatn the variant was at highest frequency in the SB morph (27%), but lower in the other two. However g819C>A was also fixed in the anadromous Grimsnes population in the south of Iceland, and at high frequency in the lake resident population of large charr in Myvatn (71%) in the north. This translates into high interlocal F_{ST}, for instance 0.85 between the anadromous populations in Hvita and Grimsnes. The average F_{ST} for *Cath2* among all the populations was 0.29, while the average F_{ST} for microsatellites was 0.245 [63]. While the frequencies of the *Cath2* g819A certainly differ between the populations, the *Cath2* locus is not associated with morphotype, as for instance g819A is fixed in both anadromous and small benthic populations. However, the *Cath2* variation may correspond, to some extent, to the relatedness of populations (Figure 4). Note however that not all branches in the tree have strong bootstrap support. Finally, there is no concordance between the variation in the two loci (*MHCIIα* and *Cath2*), and no linkage disequilibrium was observed between *Cath2* and *MHCIIα* variations and the microsatellites (χ^2 [2] = 0.11, p = 0.94).

Tests of Association between *MHCIIα* Variation and Macroscopic Parasitic Infections

Frandsen and colleagues [60] reported a difference in parasite infection rate and prevalence between the four morphs in Lake Thingvallavatn. Can the differences in *MHCIIα* allele frequencies between the PL morph and the benthic morphs in Lake Thingvallvatn be driven by habitat-specific selection, caused by marked differences of infectious agents in habitat and diet? In immunity MHCII presents antigens of pathogens such as parasites [20], which may lead to evolutionary change [19]. We tested whether the *MHCIIα* variation is related to infection rate/prevelance of four classes of macroscopic parasites (*Diphyllobothrium sp., Diplostomum sp.*, parasitic nematodes and *Eubothrium salvelini*), in Lake Thingvallavatn charr. We sampled PL- (102), SB- (76) and LB charr (17) in the fall of 2010, screened for parasites and ascertained *MHCIIα* haplotypes. The pattern of parasite infection rate and prevalence (Table 6) is consistent with previous reports [60], with the *Diplostomum sp.* being most common in LB- and SB-charr, but the other three parasites infecting a very high fraction of PL-charr. This was confirmed by a generalized linear models analyses (Table 7), which also revealed the effects of age (*Eubothrium salvelini* in PL charr, *Diplostomum sp.* in SB- and LB charr), weight (*Diplostomum sp.* in SB- and LB charr and *Diphyllobothrium sp.* in PL charr) and sex (only significant for Nematodes in PL charr). We added a term for the genotype, to test the effects of *MHCIIα* on each of those parasite types. This was only done for the PL morph as there was almost no segregating variation in the benthic morphs. The genotype terms were not

Table 5. Nucleotide diversity in *MHCIIα* in Lake Thingvallavatn 2010 sample and 9 other populations around Iceland.

Location	Size (bp)	S	π	θ	*Haplotypes*
TH_LB	293	8	0.013	0.014	3
TH_SB	293	8	0.003	0.010	3
TH_PL	293	10	0.018	0.014	4
All Lake Thingvallavatn	293	10	0.024	0.012	4
Gri_AN	293	7	0.008	0.011	3
Bir_SB	293	2	0.001	0.003	2
Hv_AN	293	3	0.002	0.004	2
Tr_SB	293	3	0.002	0.005	2
Hus_SB	293	8	0.009	0.013	3
Lon_AN	293	3	0.003	0.005	3
Gr_SB	293	3	0.005	0.005	2
My_LR	293	2	0.001	0.003	2
Hag_SB	293	0	0.000	0.000	1
All Iceland w/o Lake Thingvallavatn	293	12	0.003	0.011	8

S: the total number of segregating sites. π: The average number of nucleotide differences per site. θ: Wattersons estimator of diversity per site. See Table 1 for population identification code.
doi:10.1371/journal.pone.0069402.t005

significant, neither as a class or quantitative variable (Table 7). The models were evaluated both on a parasite-scoring-scale and raw counts, with consistent results (Table 7). For exploration we also tested interaction of genotype with other terms, which yield borderline significance for Genotype by Sex interaction with nematodes ($p = 0.07$). Considering the number of tests preformed and the poor replicability of genetic interaction terms [84] this is almost certainly a spurious association. In summary, the data do not suggest that infection rate (or infection intensity) of those four parasite classes is affected by the frequency of *MHCIIα* alleles in Lake Thingvallavatn charr.

Table 6. Parasite infection rate in Lake Thingvallavatn Arctic charr in 2010.

		Morph		
Parasite	**Measure**	**SB**	**LB**	**PL**
Diphyllobothrium sp.	MRD	0.07	0.02	1.25
	Prevalence	15/113	5/19	125/131
	Count	0.22	0.4	10.13
	Score	0.15	0.37	2.32
Diplostomum sp.	MRD	46.84	8.69	9.93
	Prevalence	109/113	19/19	131/131
	Count	192.1	178.3	70.3
	Score	2.10	2.26	1.53
Nematodes	Prevalence	1/82	1/15	55/105
Eubothrium salvelini	Prevalence	6/82	2/15	68/105

MRD: mean relative density. Both count and score are summarized by arithmetic means.
doi:10.1371/journal.pone.0069402.t006

Discussion

The sharp distinction in form, size and ecology between the four sympatric Arctic charr morphs in Lake Thingvallavatn [8,59] calls for explanation. Earlier studies found evidence of subtle but significant genetic differentiation among the morphs within the lake [50,61–63,85]. Here we report substantial genetic differentiation among the morphs within the lake, in two of the four immunological genes investigated (*Cath2* and *MHCIIα*). The pattern of divergence is not the same for both loci. In *Cath2* the strongest differentiation is between SB charr and the other two morphs studied (LB- and PL charr). Whereas in the case of *MHCIIα* the PL charr deviates markedly from the two benthic morphs within the lake, which have very similar haplotype frequencies. No differentiation was detected in two other innate immunity genes (*Hamp* and *Leap-2a*) nor the D-loop. The lack of association between mtDNA haplotypes and morphotypes, is consistent with results on variation in Arctic charr (dwarf and large forms) in 56 Siberian lakes [83]. Allele frequency differences can be caused by neutral and selective forces, but several studies have documented the impact of selection on immunological genes, with most focus on MHC loci [19,33,45].

Which Evolutionary Forces Shaped the *MHCIIα* and *Cath2* Variation in Iceland?

We observe large frequency differences of the *MHCIIα* haplotypes in the three sympatric morphs in Lake Thingvallavatn. The highest F_{ST} was 0.67 between PL- and SB charr, while the F_{ST} was 0.03 on average for 10 microsatellites between these morphs [63]. This is in contrast to very little difference in *MHCIIα* variation among 9 Arctic charr populations from around Iceland (Figure 4). It is quite surprising to discover large differences at the *MHCIIα* among morphs within one lake, while the populations around Iceland were very similar. The pattern for *Cath2* was different. A modest F_{ST} of 0.23 among morphs in Lake Thingvallavatn is notably (~8X) higher than the F_{ST} for microsatellites [63]. On a larger geographic scale, we observe very large F_{ST} 's at *Cath2* among populations (highest 0.85). However there is no association of *Cath2* polymorphism with morphotype, while there may be a connection between relatedness and *Cath2* variation. The extent of differentiation in this locus is however stronger than seen in any individual microsatellite marker. In the absence of population genetic data spanning the relevant genomic regions, we cannot test for positive selection on those (or neighboring) genes.

Coalescence simulations [63] based on microsatellites (on the same fish studied here) support a model of very limited gene flow among the PL- and SB morphs in Lake Thingvallavatn, for the last 10.000 years. Also, the observed variation in microsatellites among arctic charr populations in Iceland and Lake Thingvallavatn, suggests substantial standing genetic variation in the anadromous stock(s) that colonized Icelandic waters. The reduced gene flow, due to isolation of populations or morphs, and local selective pressures could thus lead to differentiation in loci with fitness consequences. Thus the observed patterns in *MHCIIa* and *Cath2* within Lake Thingvallavatn and between Icelandic populations may reflect chance, history, and/or interplay of isolation and selection.

Table 7. Generalized linear model analyses of the contribution of morph, sex, weight, age and *MHCIIα* genotype on parasite infections in Lake Thingvallavatn charr in 2010.

Parasite	N	Morph	Weight	Age	Sex	*MHCIIα*
Diphyllobothrium sp.	263	PL vs. SB***; LB vs. PL***	PL**	ns.	ns.	ns.
Diplostomum sp.	263	PL vs. SB***; LB vs. SB***	SB***; LB***	SB*; LB***	ns.	ns.
Nematodes	202	PL vs. SB***; LB vs. PL**	ns.	ns.	PL*	ns.
Eubothrium salvelini	202	PL vs. SB***; LB. vs. PL***	ns.	PL*	ns.	ns.

Significance: ns. $p>0.05$,
*$p<0.05$,
**$p<0.01$,
***$p<0.001$.
doi:10.1371/journal.pone.0069402.t007

Reduced Variation in the *MHCIIα* in Iceland?

One feature in the data demands special attention. MHC loci often exhibit extreme polymorphism and signs of balancing selection in fish systems [19]. In Iceland *MHCIIα* variation is very much reduced in all populations, except for the PL morph in Lake Thingvallavatn (which has two common haplotypes). Conejeros and colleagues [26] studied *MHCIIα* variation in 6 populations of Arctic charr across Europe, Asia and North America, and found much higher diversity (7 or more haplotypes in 5 populations; at most 14 individuals sequenced in each). Only the population from Trinité (2 haplotypes at 50% frequency in 9 individuals) had comparable level of variation to that observed in Lake Thingvallavatn PL charr. Part of the explanation may be that, we are studying a slightly shorter fragment of the *MHCIIα* locus than Conejeros and associates [26]. Many studies have documented excessive variation in MHC genes within and between fish populations, but there are also examples of local differences, in part attributable to natural selection [19].

The low diversity in *MHCIIα* among Icelandic Arctic charr populations may reflect history, for instance low diversity within the colonizing stock or a bottleneck in recent history. Alternatively strong selection for certain *MHCIIα* alleles in specific populations may also have played a role. A putative case in point is is the observation that the PL-charr is clearly distinct from the two benthic morphs in Lake Thingvallavatn. MHC driven mate choice has been extensively studied, with documented examples of both assortative and disassortative mating [86–88]. Eizaguirre and Lenz [19] conclude that under parasite mediated selection, MHC mediated assortative mate-choice could promote local adaptation and divergence. Our data cannot be used to evaluate such scenarios, but it would be interesting to test whether *MHCII* variation correlates with mating preferences of Arctic charr.

F_{ST} mapping and Putative Functional Alleles

F_{ST} mapping can reveal both loci under positive selection and genes with relaxed purifying selection in certain populations, that stand out of the distribution of neutral variation. In this study a small fraction of the genome was interrogated and candidates were selected based on prior data and focus on particular pathways. This approach, although unlikely to find genes with the strongest signal of differentiation between groups, provided curious patterns for the sequenced candidates. In future genome wide single base polymorphism [89], microsatellite [90,91], Rad-tag screens [92,93] or even next generation sequencing of transcriptomes [94,95] from distinct populations/species are interesting strategies to study this system in more detail.

The *MHCII* genomic regions have been cloned and sequenced in *S. salar* [96], but not in *S. alpinus*. In light of the results, it would be most interesting to clone and sequence the MHCII regions from Arctic charr, possibly from distinct morphs, populations or continents. Also, in salmon the regions contain several immunological genes, so differentiation at *MHCIIα* could be caused by linked variants in other genes [31]. As we studied only a part of intron 2 in *MHCIIα* it is rather unlikely that functional polymorphism(s) were surveyed in the data. The situation is different with *Cath2* where the strongest signal was a segregating polymorphism that leads to an amino acid replacement, serine to an arginine (S115R), in the predicted antimicrobial peptide region. Cathelicidins are like most AMPs cationic and target specifically the negative charged bacterial membrane, which ultimately leads to the killing of the bacteria [34]. It has been suggested that Cod cathelicidins (codCath) kill bacteria through lysis [44], but so far little is known about the functional mechanisms of other fish cathelicidins, which are less charged than codCath. Therefore it is difficult to speculate on the effect an amino acid change in the mature cathelicidin antimicrobial peptide in Arctic charr. Phylogenetic comparisons show that positive selection operates on charged amino acids in AMPs [45]. Thus it is tempting to speculate that the Cath2 S115R replacement is functional. One way to test whether *Cath2* is under positive selection is to assess F_{ST} s along the locus and neighboring regions, to identify the marker with strongest signal of genetic differentiation between morphs and test formally for positive selection [97].

Tests of Association of Genes and Ecological Attributes

Several studies in *S. salar* and related species reveal strong differentiation in immunological genes among populations or morphotypes [27,98,99], which may be in part due to differences in parasite diversity in distinct habitats. Eizaguirre et al. [100] demonstrated with an experimental set up that parasitic nematode infections change *MHCIIβ* allele frequencies in a single generation. Here we tested for association of four classes of large and prevalent parasites (*Diplostomum sp.*, *Diphyllobothrium sp.*, *Eubothrium salvelini* and Nematodes) and the *MHCIIα* haplotypes, but found no significant associations. This does not formally exclude the possibility that those parasites were not involved in shaping *MHCIIα* diversity, for methodological and other reasons. On the methodology side, the sample size is relatively small, compared to association tests in human genetics [101,102] and the phenotypes

are not measured in controlled environment as in quantitative genetics [103,104]. Also, we only tested for association in a sample of 4–10 year old fish from 2010, but an association may have been between the genotype and parasites in the past (over many generations or during episodes of high infection) or only in juveniles. Reverse quantitative genetics can identify ecological variables of importance and shed light on the interplay of history, population genetic and ecological factors. However, failure of such phenotype hunts do not devalue the genetic signatures of differentiation among groups. QTL mapping within Arctic charr populations have identified chromosome regions that relate to ecologically important traits, *e.g.* spawning time and development [58,105,106].By combining population genetic and QTL mapping techniques, loci related to adaptation can be identified [107].

Freshwater Fishes to Study Adaptation

Following the last glaciation Nordic freshwater fishes expanded into new territories. Several features, like novel habitats, geographic isolation of stocks, in some cases small population sizes or bottlenecks, reduced gene flow and the relatively simpler ecosystem of arctic areas, could lead to rapid evolution via both drift and selection. Some Arctic charr populations show dedicated resource morphotypes while others retain ancestral phenotypes [8,108]. Similar to the stickleback and Mexican cavefish [9,109] the dozens of morphologically and ecologically distinct Arctic charr populations are *de facto* natural experiments in parallel evolution [53,63]. Genome-wide markers make it possible to elucidate the history of the distinct and even sympatric populations [93,107,110] and identify genes relating to adaptation [14,15,95,111]. Northern species like Arctic charr, which have invaded similar habitats multiple times and adapted to them in relatively short evolutionary time, provide an interesting system to dissect the genetics and ecology of parallel evolution, however complicated and challenging.

Supporting Information

Table S1 Specifics of primers and annealing temperatures.
(DOC)

Table S2 MHCIIα genotyping and polymorphism.
(DOC)

Acknowledgments

We thank Moira Ferguson and Xia Yue for sending charr DNA samples, Bjarni K. Kristjansson for help with collecting the 2010 sample and members of the Einar Arnason laboratory for help with the ABI sequencer. We thank Marta Gomez Munoz for preliminary work on *Cath2*, *Hamp* and *Leap-2a,* Jetty Ramadevi for DNA isolation and members of the University of Iceland Arctic charr group for discussion and advice.

Author Contributions

Conceived and designed the experiments: AP KHK VHM SSS. Performed the experiments: KHK JG SR CBS VCB VHM SSS AP. Analyzed the data: KHK JG VHM AP. Contributed reagents/materials/analysis tools: KHK SR VHM AP. Wrote the paper: KHK JG VHM SSS AP. Designed experiments: AP KHK VHM SSS. Molecular work: KHK SR. Parasite analyses: JG CBS KHK SR. Aged the specimens: VCB SSS.

References

1. Lewontin RC (1974) The Genetic Basis of Evolutionary Change. Columbia University Press. Available: http://www.abebooks.com/Genetic-Basis-Evolutionary-Change-LEWONTIN-Columbia/657524818/bd.
2. Rockman M V (2012) The QTN program and the alleles that matter for evolution: all that's gold does not glitter. Evolution; international journal of organic evolution 66: 1–17. Available: http://www.ncbi.nlm.nih.gov/pubmed/22220860. Accessed 25 October 2012.
3. Phillips PC (2005) Testing hypotheses regarding the genetics of adaptation. Genetica 123: 15–24. Available: http://www.ncbi.nlm.nih.gov/pubmed/15881677.
4. Raff R (1996) The Shape of Life: Genes, Development, and the Evolution of Animal Form. University Of Chicago Press. Available: http://www.citeulike.org/group/4585/article/2931409. Accessed 14 December 2012.
5. Stern DL (2000) Evolutionary biology. The problem of variation. Nature 408: 529, 531. Available: http://dx.doi.org/10.1038/35046183. Accessed 14 December 2012.
6. Lawton-Rauh A, Robichaux RH, Purugganan MD (2003) Patterns of nucleotide variation in homoeologous regulatory genes in the allotetraploid Hawaiian silversword alliance (Asteraceae). Molecular Ecology 12: 1301–1313. Available: http://doi.wiley.com/10.1046/j.1365-294X.2003.01814.x. Accessed 14 December 2012.
7. Elmer KR, Meyer A (2011) Adaptation in the age of ecological genomics: insights from parallelism and convergence. Trends in ecology & evolution 26: 298–306. Available: http://www.ncbi.nlm.nih.gov/pubmed/21459472. Accessed 4 November 2012.
8. Snorrason SS, Skúlason S (2004) Adaptive speciation in northern fresh water fishes – patterns and processes. In: Dieckmann U, Metz H, Doebeli M, Tautz D, editors. Adaptive speciation. Cambridge University Press, Cambridge, 210–228.
9. Rundle HD (2000) Natural Selection and Parallel Speciation in Sympatric Sticklebacks. Science 287: 306–308. Available: http://www.sciencemag.org/cgi/doi/10.1126/science.287.5451.306. Accessed 29 November 2012.
10. Reimchen TE, Nosil P (2006) Replicated ecological landscapes and the evolution of morphological diversity among Gasterosteus populations from an archipelago on the west coast of Canada. Canadian Journal of Zoology 84: 643–654. Available: http://www.nrcresearchpress.com/doi/abs/10.1139/z06-036#.UMp1HIqu_lc. Accessed 14 December 2012.
11. Landry L, Vincent WF, Bernatchez L (2007) Parallel evolution of lake whitefish dwarf ecotypes in association with limnological features of their adaptive landscape. Journal of evolutionary biology 20: 971–984. Available: http://www.ncbi.nlm.nih.gov/pubmed/17465908. Accessed 14 December 2012.
12. Noakes DLG (2008) Charr truth?: sympatric differentiation in Salvelinus species. Environmental Biology of Fishes: 7–15. doi:10.1007/s10641-008-9379-x.
13. Fraser DJ, Weir LK, Bernatchez L, Hansen MM, Taylor EB (2011) Extent and scale of local adaptation in salmonid fishes: review and meta-analysis. Heredity 106: 404–420. Available: http://dx.doi.org/10.1038/hdy.2010.167. Accessed 1 November 2012.
14. Peichel CL, Nereng KS, Ohgi K a, Cole BL, Colosimo PF, et al. (2001) The genetic architecture of divergence between threespine stickleback species. Nature 414: 901–905. Available: http://www.ncbi.nlm.nih.gov/pubmed/11780061.
15. Shapiro MD, Marks ME, Peichel CL, Blackman BK, Nereng KS, et al. (2004) Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. Nature 428: 717–723. Available: http://www.ncbi.nlm.nih.gov/pubmed/15085123.
16. Jones FC, Chan YF, Schmutz J, Grimwood J, Brady SD, et al. (2012) A genome-wide SNP genotyping array reveals patterns of global and repeated species-pair divergence in sticklebacks. Current biology?: CB 22: 83–90. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3319444&tool=pmcentrez&rendertype=abstract. Accessed 4 November 2012.
17. Jones FC, Grabherr MG, Chan YF, Russell P, Mauceli E, et al. (2012) The genomic basis of adaptive evolution in threespine sticklebacks. Nature 484: 55–61. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3322419&tool=pmcentrez&rendertype=abstract. Accessed 25 October 2012.
18. Gavrilets S (2004) Fitness Landscapes and the Origin of Species. Princeton University Press. Available: http://press.princeton.edu/titles/7799.html.
19. Eizaguirre C, Lenz TL (2010) Major histocompatibility complex polymorphism: dynamics and consequences of parasite-mediated local adaptation in fishes. Journal of fish biology 77: 2023–2047. Available: http://www.ncbi.nlm.nih.gov/pubmed/21133915. Accessed 13 August 2011.
20. Murphy K, Travers P, Walport M (2007) Janeway's Immunobiology. 7th ed. Garland Science.
21. Magnadóttir B (2006) Innate immunity of fish (overview). Fish & shellfish immunology 20: 137–151. Available: http://www.ncbi.nlm.nih.gov/pubmed/15950491. Accessed 29 October 2012.
22. Magnadottir B (2010) Immunological control of fish diseases. Marine biotechnology (New York, NY) 12: 361–379. Available: http://www.ncbi.nlm.nih.gov/pubmed/20352271. Accessed 29 October 2012.
23. Cuesta A, Angeles Esteban M, Meseguer J (2006) Cloning, distribution and up-regulation of the teleost fish MHC class II alpha suggests a role for granulocytes

as antigen-presenting cells. Molecular immunology 43: 1275–1285. Available: http://www.ncbi.nlm.nih.gov/pubmed/16168483. Accessed 29 November 2012.
24. Blais J, Rico C, Van Oosterhout C, Cable J, Turner GF, et al. (2007) MHC adaptive divergence between closely related and sympatric African cichlids. PloS one 2: e734. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid = 1939875&tool = pmcentrez&rendertype = abstract. Accessed 17 December 2012.
25. Matthews B, Harmon LJ, M'Gonigle L, Marchinko KB, Schaschl H (2010) Sympatric and allopatric divergence of MHC genes in threespine stickleback. PloS one 5: e10948. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid = 2886830&tool = pmcentrez&rendertype = abstract. Accessed 16 November 2012.
26. Conejeros P, Phan A, Power M, Alekseyev S, O'Connell M, et al. (2008) MH class IIalpha polymorphism in local and global adaptation of Arctic charr (Salvelinus alpinus L.). Immunogenetics 60: 325–337. Available: http://www.ncbi.nlm.nih.gov/pubmed/18488215. Accessed 8 December 2010.
27. Tonteri A, Vasemägi A, Lumme J, Primmer CR (2010) Beyond MHC: signals of elevated selection pressure on Atlantic salmon (Salmo salar) immune-relevant loci. Molecular ecology 19: 1273–1282. Available: http://www.ncbi.nlm.nih.gov/pubmed/20196809. Accessed 2 November 2012.
28. Keller I, Taverna a, Seehausen O (2011) Evidence of neutral and adaptive genetic divergence between European trout populations sampled along altitudinal gradients. Molecular ecology 20: 1888–1904. Available: http://www.ncbi.nlm.nih.gov/pubmed/21418113. Accessed 9 November 2012.
29. Star B, Nederbragt AJ, Jentoft S, Grimholt U, Malmstrøm M, et al. (2011) The genome sequence of Atlantic cod reveals a unique immune system. Nature 477: 207–210. Available: http://www.ncbi.nlm.nih.gov/pubmed/21832995. Accessed 1 November 2012.
30. Allendorf FW, Thorgaard GH (1984) Tetraploidy and the evolution of salmonid fishes. In: Turner BJ, editor. Evolutionary genetics of fishes. 1–53.
31. Harstad H, Lukacs MF, Bakke HG, Grimholt U (2008) Multiple expressed MHC class II loci in salmonids; details of one non-classical region in Atlantic salmon (Salmo salar). BMC genomics 9: 193. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid = 2386828&tool = pmcentrez&rendertype = abstract. Accessed 2 November 2012.
32. Flajnik MF, Du Pasquier L (2004) Evolution of innate and adaptive immunity: can we draw a line? Trends in immunology 25: 640–644. Available: http://www.ncbi.nlm.nih.gov/pubmed/15530832. Accessed 6 November 2012.
33. Sackton TB, Lazzaro BP, Schlenke TA, Evans JD, Hultmark D, et al. (2007) Dynamic evolution of the innate immune system in Drosophila. Nature genetics 39: 1461–1468. Available: http://www.ncbi.nlm.nih.gov/pubmed/17987029. Accessed 1 November 2012.
34. Zasloff M (2002) Antimicrobial peptides of multicellular organisms. Nature 415: 389–395. Available: http://www.ncbi.nlm.nih.gov/pubmed/11807545. Accessed 26 November 2012.
35. Lai Y, Gallo RL (2009) AMPed up immunity: how antimicrobial peptides have multiple roles in immune defense. Trends in immunology 30: 131–141. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid = 2765035&tool = pmcentrez&rendertype = abstract. Accessed 31 October 2012.
36. Shike H, Lauth X, Westerman ME, Ostland VE, Carlberg JM, et al. (2002) Bass hepcidin is a novel antimicrobial peptide induced by bacterial challenge. European Journal of Biochemistry 269: 2232–2237. Available: http://doi.wiley.com/10.1046/j.1432-1033.2002.02881.x. Accessed 12 November 2012.
37. Zhang Y-A, Zou J, Chang C-I, Secombes CJ (2004) Discovery and characterization of two types of liver-expressed antimicrobial peptide 2 (LEAP-2) genes in rainbow trout. Veterinary immunology and immunopathology 101: 259–269. Available: http://www.ncbi.nlm.nih.gov/pubmed/15350756. Accessed 29 November 2012.
38. Chang C-I, Pleguezuelos O, Zhang Y-A, Zou J, Secombes CJ (2005) Identification of a novel cathelicidin gene in the rainbow trout, Oncorhynchus mykiss. Infection and immunity 73: 5053–5064. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid = 1201231&tool = pmcentrez&rendertype = abstract. Accessed 19 December 2012.
39. Chang C-I, Zhang Y-A, Zou J, Nie P, Secombes CJ (2006) Two cathelicidin genes are present in both rainbow trout (Oncorhynchus mykiss) and atlantic salmon (Salmo salar). Antimicrobial agents and chemotherapy 50: 185–195. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid = 1346769&tool = pmcentrez&rendertype = abstract. Accessed 19 December 2012.
40. Maier VH, Dorn K V, Gudmundsdottir BK, Gudmundsson GH (2008) Characterisation of cathelicidin gene family members in divergent fish species. Molecular immunology 45: 3723–3730. Available: http://www.ncbi.nlm.nih.gov/pubmed/18614236. Accessed 24 October 2010.
41. Scocchi M, Pallavicini A, Salgaro R, Bocick K, Gennaro R (2009) The salmonid cathelicidins: a gene family with highly varied C-terminal antimicrobial domains. Comparative biochemistry and physiology Part B, Biochemistry & molecular biology 152: 376–381. Available: http://www.ncbi.nlm.nih.gov/pubmed/19168146. Accessed 24 October 2010.
42. Feng CY, Johnson SC, Hori TS, Rise M, Hall JR, et al. (2009) Identification and analysis of differentially expressed genes in immune tissues of Atlantic cod stimulated with formalin-killed, atypical Aeromonas salmonicida. Physiological genomics 37: 149–163. Available: http://www.ncbi.nlm.nih.gov/pubmed/19240301. Accessed 29 October 2012.
43. Broekman DC, Zenz A, Gudmundsdottir BK, Lohner K, Maier VH, et al. (2011) Functional characterization of codCath, the mature cathelicidin antimicrobial peptide from Atlantic cod (Gadus morhua). Peptides 32: 2044–2051. Available: http://www.ncbi.nlm.nih.gov/pubmed/21945422. Accessed 15 November 2012.
44. Broekman DC, Frei DM, Gylfason G a, Steinarsson A, Jörnvall H, et al. (2011) Cod cathelicidin: isolation of the mature peptide, cleavage site characterisation and developmental expression. Developmental and comparative immunology 35: 296–303. Available: http://www.ncbi.nlm.nih.gov/pubmed/20950641. Accessed 29 November 2012.
45. Tennessen J a (2005) Molecular evolution of animal antimicrobial peptides: widespread moderate positive selection. Journal of evolutionary biology 18: 1387–1394. Available: http://www.ncbi.nlm.nih.gov/pubmed/16313451. Accessed 6 November 2012.
46. Johnson L (1980) The Arctic charr, Salvelinus alpinus. In: Balon EK, editor. Charrs, Salmonid Fishes of the Genus Salvelinus. The Hague: Junk. 15–98. Available: http://www.cabdirect.org/abstracts/19801402600.html;jsessionid = 0C6C3ABB8538DE9A918A3009DE2B62EE?gitCommit = 4.13.8-6-g6c31ff9. Accessed 17 December 2012.
47. Behnke RJ (1972) Systematics of salmonid fishes of recently glaciated lakes | Aquaculture Association of Canada. Journal of the Fisheries Research Board of Canada 29: 639–671. Available: http://www.aquacultureassociation.ca/salmon/253.
48. Behnke RJ (1980) A systematic review of the genus Salvelinus. In: Balon E, editor. Charrs: Salmonid Fishes of the Genus Salvelinus. The Hague: Junk. 441–481. Available: http://www.powells.com/biblio?isbn = 9789061937012.
49. Skúlason S, Snorrason SS, Jonsson B (1999) Sympatric morphs, populations and speciation. In: Magurran A, May R, editors. Evolution of Biological Diversity. New York: Oxford University Press. 70–92.
50. Wilson AJ, Gíslason D, Skúlason S, Snorrason SS, Adams CE, et al. (2004) Population genetic structure of Arctic charr, Salvelinus alpinus from northwest Europe on large and small spatial scales. Molecular ecology 13: 1129–1142. Available: http://www.ncbi.nlm.nih.gov/pubmed/15078451.
51. Skulason S, Smith TB (1995) Resource polymorphisms in vertebrates. Trends in ecology & evolution 10: 366–370. Available: http://www.ncbi.nlm.nih.gov/pubmed/21237070. Accessed 16 December 2012.
52. Brunner PC, Douglas MR, Osinov a, Wilson CC, Bernatchez L (2001) Holarctic phylogeography of Arctic charr (Salvelinus alpinus L.) inferred from mitochondrial DNA sequences. Evolution; international journal of organic evolution 55: 573–586. Available: http://www.ncbi.nlm.nih.gov/pubmed/11327164. Accessed 26 November 2012.
53. Kristjansson BK, MALMQUIST HJ, INGIMARSSON F, ANTONSSON T, SNORRASON SS, et al. (2011) Relationships between lake ecology and morphological characters in Icelandic Arctic charr, Salvelinus alpinus. Biological Journal of the Linnean Society 103: 761–771. Available: http://doi.wiley.com/10.1111/j.1095-8312.2011.01670.x. Accessed 14 December 2012.
54. Sigursteinsdóttir RJ, Kristjánsson BK (2005) Parallel Evolution, not Always so Parallel: Comparison of Small Benthic Charr, Salvelinus alpinus, from Grímsnes and Thingvallavatn, Iceland. Environmental Biology of Fishes 74: 239–244. Available: http://www.springerlink.com/index/10.1007/s10641-005-0499-2. Accessed 16 December 2012.
55. Macqueen DJ, Kristjánsson BK, Johnston I a (2010) Salmonid genomes have a remarkably expanded akirin family, coexpressed with genes from conserved pathways governing skeletal muscle growth and catabolism. Physiological genomics 42: 134–148. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid = 2888561&tool = pmcentrez&rendertype = abstract. Accessed 29 November 2012.
56. Skúlason S, Noakes DLG, Snorrason SS (1989) Ontogeny of trophic morphology in four sympatric morphs of arctic charr Salvelinus alpinus in Thingvallavatn, Iceland*. Biological Journal of the Linnean Society 38: 281–301. Available: http://doi.wiley.com/10.1111/j.1095-8312.1989.tb01579.x. Accessed 16 December 2012.
57. Eiriksson GM, Skulason S, Snorrason SS (1999) Heterochrony in skeletal development and body size in progeny of two morphs of Arctic charr from Thingvallavatn, Iceland. Journal of Fish Biology 55: 175–185. Available: http://doi.wiley.com/10.1111/j.1095-8649.1999.tb01054.x. Accessed 16 December 2012.
58. Parsons KJ, Skúlason S, Ferguson M (n.d.) Morphological variation over ontogeny and environments in resource polymorphic arctic charr (Salvelinus alpinus). Evolution & development 12: 246–257. Available: http://www.ncbi.nlm.nih.gov/pubmed/20565535. Accessed 27 May 2013.
59. Snorrason SS, Skúlason S, Sandlund O, Malmquist H, Jonsson B, et al. (1989) Shape polymorphism in sympatric Arctic charr, Salvelinus alpinus, in Thingvallavatn, Iceland. In: Kawanabe H, Yamazaki F, Noakes D, editors. Biology of Charrs and Masu Salmon. Physiology and Ecology of Japan, Special Vol. 1. Kyoto: Kyoto University Press. 393–404.
60. Frandsen F, Malmquist HJ, Snorrason SS (1989) Ecological parasitology of polymorphic Arctic charr, Salvelinus alpinus (L.), in Thingvallavatn, Iceland. Journal of Fish Biology 34: 281–297. Available: http://doi.wiley.com/10.1111/j.1095-8649.1989.tb03309.x. Accessed 16 December 2012.
61. Magnusson KP, Ferguson MM (1987) Genetic analysis of four sympatric morphs of Arctic charr, Salvelinus alpinus, from Thingvallavatn, Iceland.

Environmental Biology of Fishes 20: 67–73. Available: http://www.springerlink.com/index/10.1007/BF00002026. Accessed 16 December 2012.
62. Volpe JP, Ferguson MM (1996) Molecular genetic examination of the polymorphic Arctic charr Salvelinus alpinus of Thingvallavatn, Iceland. Molecular ecology 5: 763–772. Available: http://www.ncbi.nlm.nih.gov/pubmed/8981767.
63. Kapralova KH, Morrissey MB, Kristjánsson BK, Olafsdóttir GÁ, Snorrason SS, et al. (2011) Evolution of adaptive diversity and genetic connectivity in Arctic charr (Salvelinus alpinus) in Iceland. Heredity 106: 472–487. Available: http://www.ncbi.nlm.nih.gov/pubmed/21224880.
64. Gíslason D, Ferguson MM, Skúlason S, Snorrason SS (1999) RAPID COMMUNICATIONS/COMMUNICATIONS RAPIDES Rapid and coupled phenotypic and genetic divergence in Icelandic Arctic char (Salvelinus alpinus). Arctic 2234: 2229–2234.
65. Sibthorpe D, Sturlaugsdóttir R, Kristjansson BK, Thorarensen H, Skúlason S, et al. (2006) Characterisation and expression of the paired box protein 7 (Pax7) gene in polymorphic Arctic charr (Salvelinus alpinus). Comparative biochemistry and physiology Part B, Biochemistry & molecular biology 145: 371–383. Available: http://www.ncbi.nlm.nih.gov/pubmed/17049897.
66. Macqueen DJ, Kristjánsson BK, Paxton CGM, Vieira VLA, Johnston IA (2011) The parallel evolution of dwarfism in Arctic charr is accompanied by adaptive divergence in mTOR-pathway gene expression. Molecular ecology 20: 3167–3184. Available: http://www.ncbi.nlm.nih.gov/pubmed/21714822. Accessed 27 May 2013.
67. Weichhart T, Costantino G, Poglitsch M, Rosner M, Zeyda M, et al. (2008) The TSC-mTOR signaling pathway regulates the innate inflammatory response. Immunity 29: 565–577. Available: http://www.ncbi.nlm.nih.gov/pubmed/18848473. Accessed 16 November 2012.
68. Smith TJ (2010) Insulin-like growth factor-I regulation of immune function: a potential therapeutic target in autoimmune diseases? Pharmacological reviews 62: 199–236. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid = 2879913&tool = pmcentrez&rendertype = abstract. Accessed 12 November 2012.
69. Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S, editors. Bioinformatics Methods and Protocols: Methods in Molecular Biology. Totowa, NJ: Humana Press. 365–386.
70. Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome research 8: 175–185. Available: http://www.ncbi.nlm.nih.gov/pubmed/9521921. Accessed 2 November 2012.
71. Gordon D (2003) Viewing and editing assembled sequences using Consed. Current protocols in bioinformatics/editoral board, Andreas D Baxevanis. [et al] Chapter 11: Unit11.2. Available: http://www.ncbi.nlm.nih.gov/pubmed/18428695. Accessed 26 November 2012.
72. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, et al. (2007) Clustal W and Clustal X version 2.0. Bioinformatics (Oxford, England) 23: 2947–2948. Available: http://www.ncbi.nlm.nih.gov/pubmed/17846036. Accessed 1 November 2012.
73. Nicholas KB (Bank of A, Nicholas Jr. HB (Pittsburg. SC, Deerfield II. DW (Pittsburgh SC (1997) GeneDoc: Analysis and Visualization of Genetic Variation. Available: http://www.nrbsc.org/gfx/genedoc/ebinet.htm.
74. Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, et al. (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics (Oxford, England) 23: 2633–2635. Available: http://www.ncbi.nlm.nih.gov/pubmed/17586829. Accessed 5 November 2012.
75. Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19: 2496–2497. Available: http://bioinformatics.oxfordjournals.org/content/19/18/2496.short. Accessed 30 October 2012.
76. Reynolds J, Weir BS, Cockerham CC (1983) Estimation of the coancestry coefficient: basis for a short-term genetic distance. Genetics 105: 767–779. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid = 1202185&tool = pmcentrez&rendertype = abstract. Accessed 26 November 2012.
77. Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. Genetics 139: 457–462. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid = 1206343&tool = pmcentrez&rendertype = abstract. Accessed 26 November 2012.
78. Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evolutionary bioinformatics online 1: 47–50. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid = 2658868&tool = pmcentrez&rendertype = abstract. Accessed 6 December 2012.
79. Felsenstein J (1989) PHYLIP–phylogeny inference package (version 3.2). Cladistics 5: 164–166.
80. Kristmundsson Á, Richter S (2009) Parasites of resident arctic charr, Salvelinus alpinus, and brown trout, Salmo trutta, in two lakes in Iceland. ICEL AGRIC SCI 22: 5–18.
81. Fox J, Weisberg S (2011) An R Companion to Applied Regression. Second Edi. Available: tinyurl.com/carbook.
82. Venables WN, Ripley BD (2002) Modern Applied Statistics with S, 4th ed. Springer. Available: http://www.stats.ox.ac.uk/pub/MASS4/.
83. Alekseyev SS, Bajno R, Gordeeva N V, Reist JD, Power M, et al. (2009) Phylogeography and sympatric differentiation of the Arctic charr Salvelinus alpinus (L.) complex in Siberia as revealed by mtDNA sequence analysis. Journal of fish biology 75: 368–392. Available: http://www.ncbi.nlm.nih.gov/pubmed/20738544. Accessed 17 October 2010.
84. Palsson A, Dodgson J, Dworkin I, Gibson G (2005) Tests for the replication of an association between Egfr and natural variation in Drosophila melanogaster wing morphology. BMC genetics 6: 44. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid = 1208880&tool = pmcentrez&rendertype = abstract. Accessed 26 November 2012.
85. Gislason D, Ferguson MM, Skulason S, Snorrason SS (1999) Rapid and coupled phenotypic and genetic divergence in Icelandic Arctic char … Library.
86. Thünken T, Meuthen D, Bakker TCM, Baldauf S a (2012) A sex-specific trade-off between mating preferences for genetic compatibility and body size in a cichlid fish with mutual mate choice. Proceedings Biological sciences/The Royal Society 279: 2959–2964. Available: http://www.ncbi.nlm.nih.gov/pubmed/22513859. Accessed 2 November 2012.
87. Lenz TL, Eizaguirre C, Scharsack JP, Kalbe M, Milinski M (2009) Disentangling the role of MHC-dependent "good genes" and "compatible genes" in mate-choice decisions of three-spined sticklebacks Gasterosteus aculeatus under semi-natural conditions. Journal of fish biology 75: 2122–2142. Available: http://www.ncbi.nlm.nih.gov/pubmed/20738677. Accessed 2 November 2012.
88. Eizaguirre C, Yeates SE, Lenz TL, Kalbe M, Milinski M (2009) MHC-based mate choice combines good genes and maintenance of MHC polymorphism. Molecular ecology 18: 3316–3329. Available: http://www.ncbi.nlm.nih.gov/pubmed/19523111. Accessed 2 November 2012.
89. Akey JM, Zhang G, Zhang K, Jin L, Shriver MD (2002) Interrogating a high-density SNP map for signatures of natural selection. Genome research 12: 1805–1814. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid = 187574&tool = pmcentrez&rendertype = abstract. Accessed 12 November 2012.
90. Lehmann T, Licht M, Elissa N, Maega BTA, Chimumbwa JM, et al. (n.d.) Population Structure of Anopheles gambiae in Africa. The Journal of heredity 94: 133–147. Available: http://www.ncbi.nlm.nih.gov/pubmed/12721225. Accessed 26 November 2012.
91. Storz JF, Payseur BA, Nachman MW (2004) Genome scans of DNA variability in humans reveal evidence for selective sweeps outside of Africa. Molecular biology and evolution 21: 1800–1811. Available: http://www.ncbi.nlm.nih.gov/pubmed/15201398. Accessed 12 November 2012.
92. Hohenlohe P a, Amish SJ, Catchen JM, Allendorf FW, Luikart G (2011) Next-generation RAD sequencing identifies thousands of SNPs for assessing hybridization between rainbow and westslope cutthroat trout. Molecular ecology resources 11 Suppl 1: 117–122. Available: http://www.ncbi.nlm.nih.gov/pubmed/21429168. Accessed 29 October 2012.
93. Emerson KJ, Merz CR, Catchen JM, Hohenlohe PA, Cresko WA, et al. (2010) Resolving postglacial phylogeography using high-throughput sequencing. Proceedings of the National Academy of Sciences of the United States of America 107: 16196–16200. Available: http://www.pnas.org/content/107/37/16196.long. Accessed 2 November 2012.
94. Hudson ME (2008) Sequencing breakthroughs for genomic ecology and evolutionary biology. Molecular ecology resources 8: 3–17. Available: http://www.ncbi.nlm.nih.gov/pubmed/21585713. Accessed 31 October 2012.
95. Stapley J, Reger J, Feulner PGD, Smadja C, Galindo J, et al. (2010) Adaptation genomics: the next generation. Trends in ecology & evolution 25: 705–712. Available: http://www.ncbi.nlm.nih.gov/pubmed/20952088. Accessed 26 October 2012.
96. Grimholt U, Larsen S, Nordmo R, Midtlyng P, Kjoeglum S, et al. (2003) MHC polymorphism and disease resistance in Atlantic salmon (Salmo salar); facing pathogens with single expressed major histocompatibility class I and class II loci. Immunogenetics 55: 210–219. Available: http://www.ncbi.nlm.nih.gov/pubmed/12811427. Accessed 18 November 2012.
97. Hudson RR, Bailey K, Skarecky D, Kwiatowski J, Ayala FJ (1994) Evidence for positive selection in the superoxide dismutase (Sod) region of Drosophila melanogaster. Genetics 136: 1329–1340. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid = 1205914&tool = pmcentrez&rendertype = abstract. Accessed 21 May 2013.
98. Evans ML, Neff BD, Heath DD (2010) MHC genetic structure and divergence across populations of Chinook salmon (Oncorhynchus tshawytscha). Heredity 104: 449–459. Available: http://www.ncbi.nlm.nih.gov/pubmed/19773808. Accessed 30 June 2010.
99. Consuegra S, De Eyto E, McGinnity P, Stet RJM, Jordan WC (2011) Contrasting responses to selection in class I and class IIα major histocompatibility-linked markers in salmon. Heredity 107: 143–154. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid = 3178404&tool = pmcentrez&rendertype = abstract. Accessed 10 November 2012.
100. Eizaguirre C, Lenz TL, Kalbe M, Milinski M (2012) Rapid and adaptive evolution of MHC genes under parasite selection in experimental vertebrate populations. Nature communications 3: 621. Available: http://www.nature.com/ncomms/journal/v3/n1/full/ncomms1632.html#/f1. Accessed 28 October 2012.
101. Amundadottir LT, Sulem P, Gudmundsson J, Helgason A, Baker A, et al. (2006) A common variant associated with prostate cancer in European and

African populations. Nature genetics 38: 652–658. Available: http://www.ncbi.nlm.nih.gov/pubmed/16682969. Accessed 4 November 2012.
102. Gudbjartsson DF, Arnar DO, Helgadottir A, Gretarsdottir S, Holm H, et al. (2007) Variants conferring risk of atrial fibrillation on chromosome 4q25. Nature 448: 353–357. Available: http://www.ncbi.nlm.nih.gov/pubmed/17603472. Accessed 2 November 2012.
103. Palsson A, Gibson G (2004) Association between nucleotide variation in Egfr and wing shape in Drosophila melanogaster. Genetics 167: 1187–1198. Available: http://www.genetics.org/content/167/3/1187.short. Accessed 6 November 2012.
104. Flint-Garcia SA, Thuillet A-C, Yu J, Pressoir G, Romero SM, et al. (2005) Maize association population: a high-resolution platform for quantitative trait locus dissection. The Plant journal?: for cell and molecular biology 44: 1054–1064. Available: http://www.ncbi.nlm.nih.gov/pubmed/16359397. Accessed 29 October 2012.
105. Leder EH, Danzmann RG, Ferguson MM (2006) The candidate gene, Clock, localizes to a strong spawning time quantitative trait locus region in rainbow trout. The Journal of heredity 97: 74–80. Available: http://jhered.oxfordjournals.org/content/97/1/74.short. Accessed 7 November 2012.
106. Parsons KJ, Sheets HD, Skúlason S, Ferguson MM (2011) Phenotypic plasticity, heterochrony and ontogenetic repatterning during juvenile development of divergent Arctic charr (Salvelinus alpinus). Journal of evolutionary biology 24: 1640–1652. Available: http://www.ncbi.nlm.nih.gov/pubmed/21599773. Accessed 27 May 2013.
107. Renaut S, Maillet N, Normandeau E, Sauvage C, Derome N, et al. (2012) Genome-wide patterns of divergence during speciation: the lake whitefish case study. Philosophical transactions of the Royal Society of London Series B, Biological sciences 367: 354–363. Available: http://www.ncbi.nlm.nih.gov/pubmed/22201165. Accessed 4 November 2012.
108. Noakes DLG (2008) Charr truth: sympatric differentiation in Salvelinus species. Environmental Biology of Fishes 83: 7–15. Available: http://apps.webofknowledge.com/full_record.do?product = WOS&search_mode = GeneralSearch&qid = 1&SID = P144M4CgPi63FGeh7lj&page = 1&doc = 1. Accessed 21 November 2012.
109. Jeffery WR (2001) Cavefish as a model system in evolutionary developmental biology. Developmental biology 231: 1–12. Available: http://www.ncbi.nlm.nih.gov/pubmed/11180948. Accessed 12 November 2012.
110. Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, et al. (2010) Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. PLoS genetics 6: e1000862. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid = 2829049&tool = pmcentrez&rendertype = abstract. Accessed 1 November 2012.
111. Linnen CR, Kingsley EP, Jensen JD, Hoekstra HE (2009) On the origin and spread of an adaptive allele in deer mice. Science (New York, NY) 325: 1095–1098. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid = 2736094&tool = pmcentrez&rendertype = abstract. Accessed 3 November 2012.

Paper 2

3 Paper II

Bones in motion: ontogeny of craniofacial development in sympatric Arctic charr morphs

Kalina H. Kapralova, Zophonías O. Jónsson, Arnar Palsson, Sigrídur Rut Franzdóttir, Bjarni K. Kristjanson, Sigurður S. Snorrason

Authors' contribution: Conceived and designed the experiments: KHK SSS ZOJ AP BKK. Performed the experiments: KHK AP BKK SRF ZOJ SSS. Analysed the data: KHK ZOJ AP BKK SSS. Contributed reagents/materials/analysis tools: KHK BKK SSS AP ZOJ SRF. Wrote the paper: KHK AP ZOJ SSS.

3.1 Abstract

This paper describes the embryonic and early post-hatching craniofacial cartilage development in Arctic charr. Four Arctic charr varieties are studied: three natural morphs (planctivorous (PL), large (LB) and small benthic (SB) from Thingvallavatn) and an aquaculture strain from the Holar Aquaculture stock (AC). Arctic charr display segmental development of the pharyngeal arches as is characteristic for all vertebrates, and the order of events accompanying the craniofacial development is the same as in other teleosts. The four Arctic charr varieties under study showed differences in their head growth rate. The head starts out smaller in the benthic morphs but, due to a sharp increase in growth rate at hatching, the LBs end up with the largest heads at the post-hatching stages. The hatching period appears to be associated with significant allometric shape changes. SB differed markedly from the others in terms of the orientation and/or shape of their ontogenetic trajectories and LB differed from AC and PL in the length of their ontogenetic trajectories. Together the data illustrate the strength of applying multivariate geometrics to analyses of recently evolved trophic polymorphism during early development

3.2 Introduction

Understanding the series of events through which complex phenotypes arise and evolve is a central aim in evolutionary-developmental biology. These processes (evolution and development) are highly intertwined: diverse phenotypes arise through development and changes in developmental processes will provide the necessary variation for natural selection to act upon (Gould, 1977; Hall, 1999). Ontogeny can be regarded as the progression of an organism through a multidimensional space defined by its size, shape and age (Klingenberg, 1998). Thus a common strategy in studying the relationship between evolution and development is to compare ontogenetic trajectories between closely related species (Klingenberg, 1998).

The vertebrate skull, one of the most complex anatomical units, is composed of three broad regions (spanchnocranium, neurocranium and dermatocranium) each of which is characterised by a unique developmental and evolutionary origin. The feeding apparatus (formed by the splanchnocranium and its dermal counterparts) is probably the most complex and evolutionary diverse mechanical unit (Westneat, 2005) and its development involves derivatives of all three germ layers (Szabo-Rogers *et al.*, 2010). The pharyngeal skeleton originates from neural crest cells, which migrate from the dorso-lateral margins of the neural folds (Basch and Bronner-Fraser, 2006) in a segmented manner following their rhombomeric origin (Lumsden *et al.*, 1991; Knight and Schilling, 2006). Each segment will give rise to a differentiated arch, which is subdivided, into individual dorsal and ventral structures (Schilling and Kimmel, 1997). Evolutionarily, the striking variation seen among vertebrates in the shape and function of pharyngeal cartilages is thought to be the result of subtle differences in the patterning of the neural crest cells (Knight and Schilling, 2006).

Many morphological characters can be measured by simply recording their presence or absence; other traits can be quantified based on simple linear measurements such as length and width of the beak in birds (Badyaev *et al.*, 2001), length of the jaws, or measures of mechanical levers in African cichlids (Albertson *et al.*, 2003, 2005) or measures of pelvic girdle etc. in sticklebacks (Shapiro *et al.*, 2004). Some structures however, call for more complex description of shape, that can summarise global and local changes in shape and proportion (Palsson and Gibson, 2004; Klingenberg, 2010). These methods have successively been used in studying morphological variation in development (Parsons *et al.*, 2008, 2014; Young *et al.*, 2010; Gonzalez *et al.*, 2011).

Using zebrafish as a model for studying the evolution and development of the craniofacial elements in fish undoubtedly has its advantages in terms of ease of experimental manipulation and controlled mutagenesis and a large body of work has generated valuable insights into these processes (Neuhauss *et al.*, 1996; Schilling, 1997; Yelick and Schilling, 2002; Dale *et al.*, 2009). On the other hand, the extraordinary variation in natural populations of freshwater fish provides an advantage model organisms do not have, namely a unique possibility to examine actual evolutionary processes (Klingenberg, 2010). Ongoing work on African cichlids (Albertson and Kocher, 2005; Albertson *et al.*, 2005; Roberts *et al.*, 2011; Parsons *et al.*, 2014) and threespine sticklebacks (Kimmel *et al.*, 2005, 2012; Jamniczky *et al.*, 2014) and Antarctic notothenioid fishes (Albertson *et al.*,

2010a) is generating important insights into the mechanisms involved in the developmental aspects of morphological evolution.

The four morphs of Arctic charr (*Salvelinus alpinus*) from Lake Thingvallavatn, Iceland offer an excellent opportunity to study craniofacial development in a context of a recent ecological diversification. These four morphs differ extensively in morphology, life history characteristics and ecology, as reflected in different habitat use, diet and endoparasite fauna (Jonsson *et al.*, 1988; Snorrason *et al.*, 1989; Frandsen *et al.*, 1989; Malmquist *et al.*, 1992; Sandlund *et al.*, 1992). The Arctic charr morphs show subtle but significant neutral genetic differentiation, and despite spatio-temporal overlap in spawning the level of gene flow between the two smallest, most abundant morphs in the lake, planktivorous (PL) and small benthivorous charr (SB), is restricted (Kapralova *et al.*, 2011). Even more pronounced genetic differentiation among the morphs in the lake was detected in a study on immune system genes (Kapralova *et al.*, 2013), suggesting that parts of the immune system had diverged among Arctic charr morphs in Thingvallavan since the colonization of Iceland 10 000 years ago.

Based on the morphology of their feeding apparatus and the shape of the snout, the four morphs can be classified into two morphotypes: a limnetic and a benthic (Snorrason *et al.*, 1989). The two morphs belonging to the limnetic morphotype, a planktivorous (PL) and piscivorous (PI) charr, have pointed snouts and evenly protruding jaws. The two morphs belonging to the benthic morphotype, a small (SB) and a large benthivorouscharr (LB), have blunt snouts, short lower jaws and relatively large pectoral fins (Snorrason *et al.*, 1989). Common garden experiments have indicated that the differences in trophic morphology of the Thingvallavatn morphs have a genetic basis (Skulason *et al.*, 1989). Moreover the characteristic short lower jaws and blunt snouts of LB and SB are thought to be embryonic characteristics retained in the adult through heterochronies, which are partially genetically determined (Skulason *et al.*, 1989) and to some extent the result of plastic responses to different environments (Parsons *et al.*, 2010, 2011). The role of heterochrony in the development of Arctic charr morphologies was further demonstrated in a study showing that some skeletal elements of the head start ossifying earlier and/or faster in small benthivorous embryos than in embryos derived from the planktivorous morph (Eiriksson *et al.*, 1999).

In this study we used the recently evolved variability in Icelandic Arctic charr to address questions about craniofacial development and evolution. We used dense developmental series to study early craniofacial development and to characterize the timing of major events in Arctic charr cartilage and bone development. We then selected 8 pre- and post-hatching time points during the period when the major craniofacial elements are laid down and the ossification of the trophic apparatus starts. We studied the ontogenetic trajectories related to growth in three of the Thingvallavatn morphs, SB and LB (representing a benthic morphotype) and PL and an aquaculture strain AC (representing a limnetic morphotype). Given that egg size and yolk quality differ between the studied varieties of Arctic charr (Skulason MSc thesis 1986, Leblanc PhD thesis 2012) and that maternal effects can influence embryo size considerably throughout embryonic development (see references in (Perry *et al.*, 2004), we expect that the Arctic charr varieties under study will differ in size during ontogeny. We also studied the role of allometry in the development of Arctic charr morphs. As the selected time-points covered stages of intense cartilage growth we hypothesised that the differences in shape associated with size will account for an important part of the variation. Finally, we selected four developmental time points to

study changes in head shape throughout ontogeny, before and after hatching. We defined the craniofacial shape changes throughout development for each Arctic charr variety as a multidimensional trajectory (ontogenetic shape trajectory) in morphological space (see (Adams and Collyer, 2009) for details). We then compared the three attributes (size, orientation and shape) of these trajectories between morphs.

3.3 Material and methods

3.3.1 Sampling of parent fish and rearing of offspring

For this study we used developmental time-series from pure crosses of four Arctic charr varieties, three morphs, LB-, SB- and PL-charr, from lake Thingvallavatn and an aquaculture strain from the Hólar breeding programme (AC) (Svavarsson, 2007). Sexually ripe fish from Thingvallavatn caught using gill-nets. Fishing permissions were obtained from the Thingvellir National Park Commission and the land-owner of Mjóanes farm. Fish were killed by a sharp blow to the head and for each group eggs from several females were pooled and fertilized using milt from several males from the same group. Eggs were reared under identical conditions in the same hatching tray (EWOS, Norway) with constant water flow (at approximately 5°C at all times) and in complete darkness at the Holar University College experimental facilities in Verið, Sauðárkrókur. The rearing and collection of the embryos was performed according to Icelandic regulations (license granted to Holar University College aquaculture and experimental facilities in Verið, Sauðárkrókur). Water temperature was recorded twice daily to estimate the relative age of the embryos using tau-somite (τs) units defined as the time it takes for one somite pair to form at a given temperature (Gorodilov, 1996).

3.3.2 Staining and photographing

To describe the events of craniofacial differentiation and growth of the head, samples were collected throughout development, fixed and stored in 4% PFA and stained for cartilage (alcian blue) and bone (alizarin red) using a modified protocol from (Walker and Kimmel, 2007).

Based on the information obtained from these developmental series four embryonic (200, 223, 246, 266 τs) and four post-hatching stages (293, 305, 315 and 336 τs), were selected for studying the variation in craniofacial development and growth of the four Arctic charr varieties. A total of 296 individuals (Table 3.1) were stained and photographed. The staining and photographing was performed in 3 staining batches: 1) embryonic stages 200, 223, 246 and 266 τs, 2) stages 293, 315 τs, and 3) stages 305 and 336 τs. In each batch, samples from the four morphs under study were stained simultaneously.

Stained individuals were placed in a petri dish containing 50 ml of 1% agarose gel and immobilized with dissecting needles to ensure the correct positioning of the embryo. The head of each individual was photographed ventrally facing left using a Leica (MZ10) stereomicroscope. The same magnification (2.0x) was used for each photo.

Table 3-1 Sampling scheme outlining the number of individuals per morph and time point.

		Morph			
Stage (τs)	Period	AC	LB	PL	SB
200	embryonic	8	9	9	8
223	embryonic	5	7	6	9
246	embryonic	5	8	7	8
266	embryonic	6	7	8	5
293	post hatching	12	10	12	12
305	post hatching	12	14	15	10
315	post hatching	15	12	8	9
336	post hatching	6	13	10	11
Total		69	80	75	72

3.3.3 Geometric morphometrics

We selected 15 landmarks (Figure 3.1) to describe the overall shape of the head as seen from the ventral side, focusing on elements such as the hyoid arch, the lower jaw and the ethmoid plate (Figure 3.2). Landmarks were digitized using tps.DIG2 (Rohlf, 2006). The landmarks do not occupy the same 2D plane, but since the embryos are small and the ventral aspect is rather flat at these stages of development, all of the landmarks are in comfortable focus for digitizing. The shape information (landmark co-ordinates) for each specimen was then extracted using a Generalized Procrustes analysis (GPA) in MorphoJ (Klingenberg, 2011), and after accounting for scale, position and orientation all specimens were superimposed on a common coordinate system (Rohlf and Slice, 1990; Goodall, 1991). Only the symmetric component of shape variation (Klingenberg *et al.*, 2002) was used for subsequent statistical analysis. The Centroid Size (defined as the square root of the sum of the squared distances of all landmarks from their centroid) for each specimen was retained after the Procrustes fit and was used as a measure of individual head size. Each individual was digitized twice and the results from the repeated measurements were averaged in the final data set. Measurement error was assessed by performing Principal Component Analysis (PCA) on a dataset containing the two landmarking sessions. The Principal component analysis did not show any separation between the landmarking sessions (Figure S3.1).

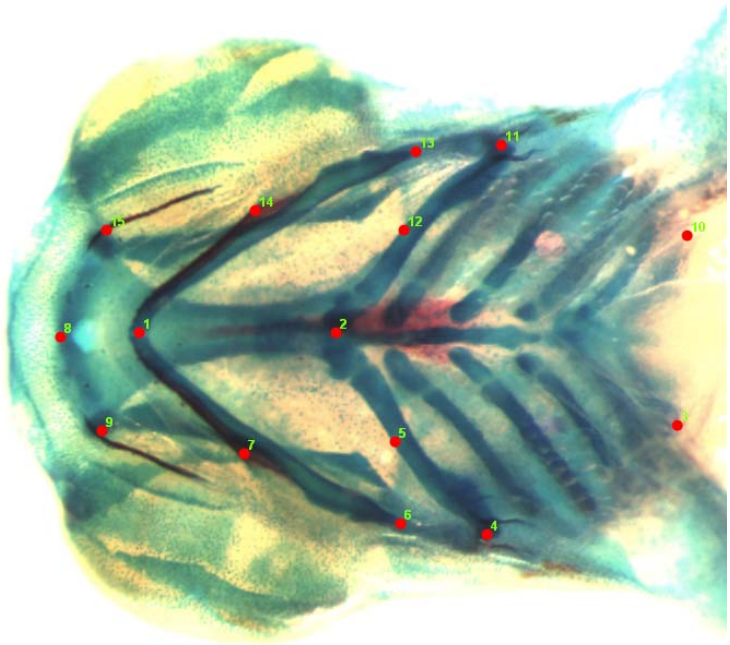

Figure 3.1 The head of an Arctic charr embryo/juvenile (stage 346 τs). The 15 landmarks (6 pairs of bilateral and 3 mid-line landmarks) used in this study are indicated by red dots. Landmarks were selected to describe the ventral aspect of the head shape of pre-hatching embryos and post-hatching juveniles of Arctic charr, focusing in the major craniofacial elements such as Meckel's cartilage, the hyoid and branchial arches and the ehtmoid plate.

3.3.4 Statistical analysis

First we studied the changes in head size (represented by the Log transformed Centroid size or LCS) throughout development for all four experimental groups. We compared the head size of morph offspring throughout ontogeny using a two-way ANOVA, followed by post hoc tests to evaluate which groups showed significant differences in head size.

To investigate whether morph offspring differed in shape we performed a Procrustes ANOVA (Klingenberg *et al.*, 2002). The shape variation was studied over eight developmental time points (200, 223, 246, 266, 293, 305, 315 and 336 τs) for the four varieties of Arctic charr AC, LB, PL and SB with a linear model (y~Morph+Stage, where y is a two dimensional array of shape data). The significance of the tests was assessed with 1000 permutations. These analyses were performed using the ProcDist.lm function in the geomorph package (Adams and Otárola-Castillo, 2013) in R.

We next investigated the nature of shape changes related to size (allometry) for the four Arctic charr varieties by performing Principal Component Analysis (PCA) on the whole dataset in MorphoJ. The shape changes associated with the three major Principle Components (PCs) were visualised using deformation grids.

Finally we used the *phenotypic trajectory analysi*s approach (Adams and Collyer, 2007, 2009; Collyer and Adams, 2007) to describe and compare shape craniofacial changes throughout ontogeny for the four studied groups, at four developmental stages. Briefly, ontogenetic shape means were computed for each combination of morph and time from linear models. The three attributes (size, orientation and shape) of the ontogenetic shape trajectories were then computed and compared statistically between morphs. Trajectory size is defined as the path-length distance along the ontogenetic trajectory and is calculated as the sum of the distances between adjacent developmental points. Trajectory orientation is described as the direction of its first PC. Trajectory shape is computed using Procrustes approaches and corresponds to the relative configuration of points expressed in the ontogenetic data space (for analytical details see (Adams and Collyer, 2009). The ontogenetic shape trajectory analyses were performed using the trajectory.analysis function in the geomorph package (Adams and Otárola-Castillo, 2013) in R.

3.4 Results

3.4.1 Development and growth of craniofacial elements

Although the trophic apparatus of salmonids can be described as primitive amongst teleosts, the key elements seen in development are the same and as in other fish species (Schilling and Kimmel, 1997; Albertson *et al.*, 2010b). Viewed from the ventral side the early, post-hatching (Stage 346 τs) craniofacial anatomy of Arctic charr is composed of seven pharyngeal arches (Figure 3.2, Table 3.2 for abbreviations). The first arch (the mandibular arch) is composed of a ventrally located Meckel's cartilage, forming the base for the dermal ossification of the lower jaw elements, and the adjoining palatoquadrates with dorsolateral projections that form the bases for the upper jaws. The dentary bone is taking shape and the articular-angular bone is starting to form. The ventral part of the second arch (the hyoid arch) is composed of a large, central basihyal cartilage, next to that, a pair of small cartilages (the hypohyals), and next to those, a pair of large cartilage bars (the ceratohyals) extending postero-laterally and dorsally around the pharynx. The remaining five pharyngeal arches are composed of four mid-ventral basibranchial cartilages, four paired hypobranchial cartilages and five large paired ceratohyals. The ventral aspect of the ethmoid plate and the first stage of ossification of upper jaw elements can be seen (maxillae, pre-maxillae and denticles of the pre-maxilla and the palatines) (Figure 3.2).

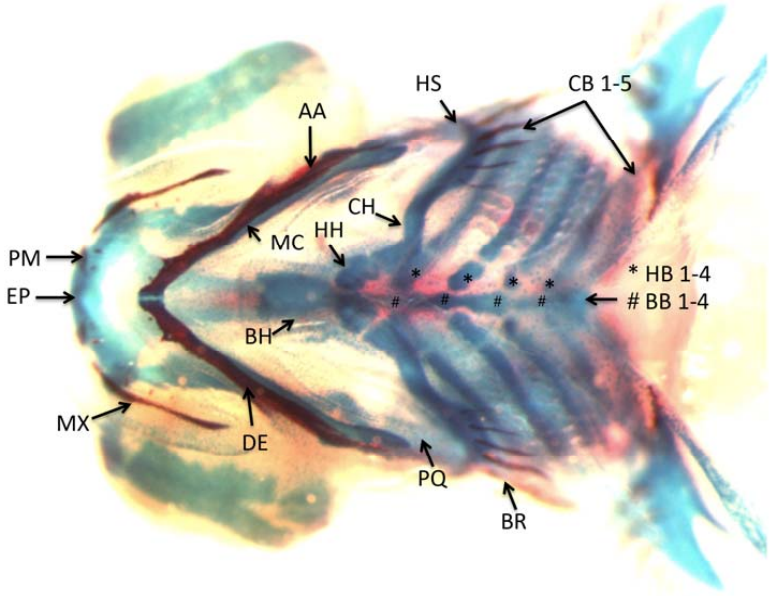

Figure 3.2 Arctic charr juvenile showing major elements of the splanchnocranium as well as the anterior projections of the neurochranium, e,g, the ethmoid plate (see Table 2 for abbreviations). Ossification of the lower jaw (dentary and Articular-angular bones) and upper jaw elements (premaxillas and maxillas) is advancing and teeth are being formed.

The first craniofacial elements to appear as clear units of cartilage formation in the ventral aspect are the two trabeculae (appearing at stage 138 τs) and the Meckel's cartilages and palatoquatrates (at stage140 τs), shortly followed by the major elements of the hyoid and branchial arches (stages 150-160 τs in Figure 3.3, Table 3.2). The minor elements (the hypo- and basi-branchials) of these arches appear later and over a more extended period (stages 187-266 τs in Figure 3.3, Table 3.2). The ethmoid plate starts forming around stage 180 τs and is fully fused centrally at stage 215 τs. Rudiments of the maxillae can be seen as early as stage 200 τs and ossification of the maxillas and the dentary has started at stage 246 τs. At the time of hatching (280-285 τs) all major craniofacial elements have formed.

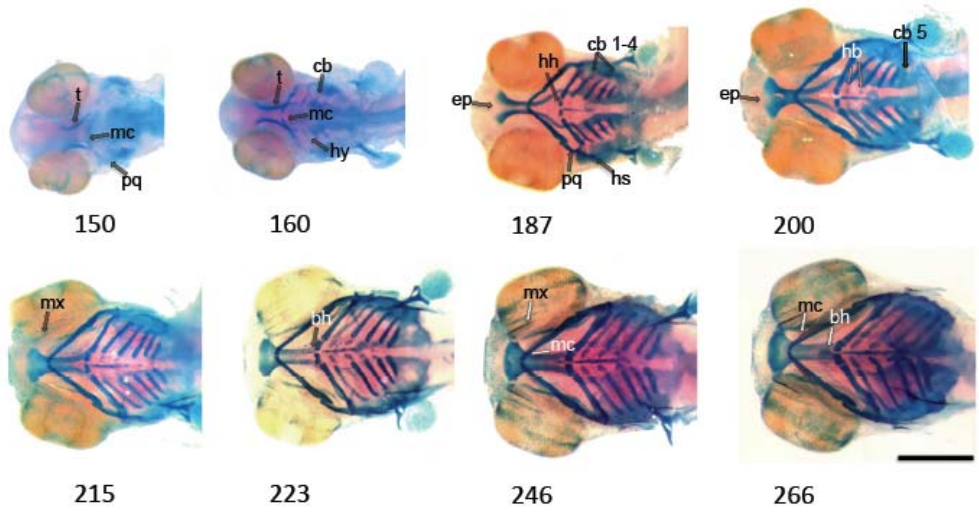

Figure 3.3 Development and growth of craniofacial cartilage elements in pre-hatching Arctic charr embryos. Ventral views of 8 stages: Stage 150 τs: the trabeculae (t), Meckel's cartilages (mc) and palatoquandrates (pq) can be seen clearly. Stage 160 τs: the hyoid arch (hy) and the ceratobranchials (cb) can be seen. Stage 187 τs: the ethmoid plate (ep) starts forming as well as the hyohyal (hh) and the basibranchial (bb) cartilages. Stage 200 τs: fusing of ep, hypobranchial cartilages (hb) 1-2 and cb 5 are visible. Stage 215 τs: rudements of the maxillae (mx) appear. Stage 223 τs: the basihyal cartilage (bh) starts forming and hb 3-4 appear. Stage 246 τs: mx and the dentary start ossifying. Stage 266 τs: pre-hatching, the majority of the craniofacial cartilage elements are in place and some have started ossifying. Scale bar: 1mm.

The post-hatching period is characterised by growth of all cartilage elements (Figure 3.4). The first craniofacial elements belonging to the dermatocranium to start ossifying are those bordering the Meckel's cartilages (Figure 3.4). Ossification can already be detected in late embryonic stages (246 τs) (Figure 3.4). A detailed description of the timing of major events during Arctic charr craniofacial development and the corresponding timing in zebrafish is given in Table 3.2.

Table 3-2 Sequence of appearance of craniofacial cartilages in Arctic charr.

Cartilage	Abreviation	Time of appearance in Arctic charr*	Time of appearance in zebrafish◆
Mandibular arch	ma		
Meckel's cartilage	mc	146	55
palatoquandrate	pq	146	53
Hyoid arch	ha		
basihyal	bh	223	54
ceratohyal	ch	150	54
hyosymplectic	hs	155	57
Branchial arches			
basibranchials	bb	187	68
hypobranchials	hb	200	74
ceratobranchial 1	cb 1	160	56
ceratobranchial 2	cb 2	160	60
ceratobranchial 3	cb 3	187	64
ceratobranchial 4	cb 4	187	68
ceratobranchial 5	cb 5	200	64
Neurocranium			
trabeculae	t	138	45
ethmoid plate	ep	187	52
Dermatocranium			
maxilae	mx	200	
pre-maxilae	pm	336	
anguloarticular	aa	293	
branchiostegal rays	br		

** Time is reported in τs (the time it takes for 1 somite to be formed, Gorodilov 1996).*

◆ Time is reported in hour post fertilisation at 28.5□C (from Schilling and Kimmel 1997)

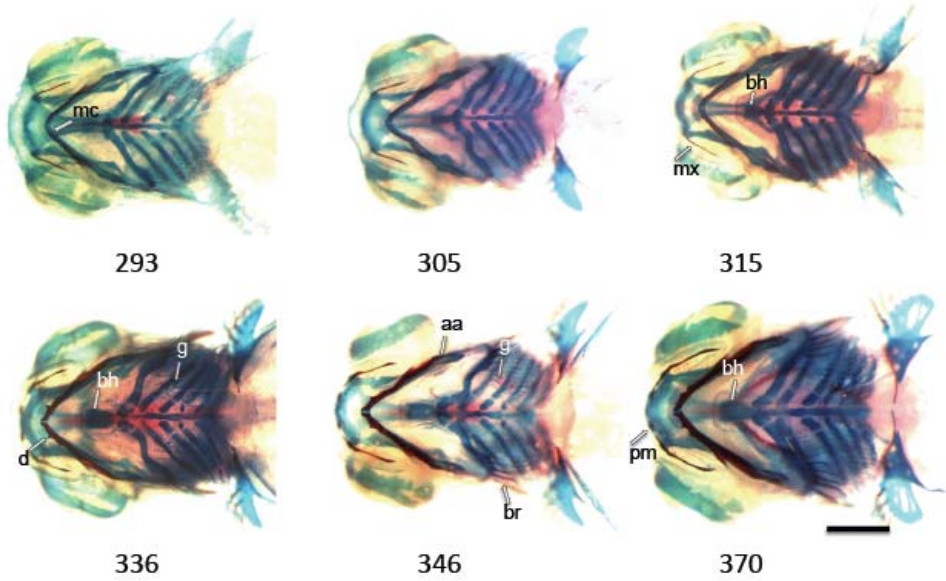

Figure 3.4 Development and growth of craniofacial cartilage elements in post-hatching Arctic charr juveniles. Ventral views of 6 stages (293, 305, 315, 336, 346, 370 τs). These stages are characterised by fast growth of the craniofacial elements, especially the basihyal cartilage (bh) and emergence and ossification of important dermal bones such as the dentary (d), articular-angular (aa), maxillae (mx), premaxillae (pm), and the branchiostegal rays (br). The emergence of the gills (g) can be seen clearly. Scale bar, 1mm

3.4.2 Ontogenic trajectories in four Arctic charr morphs

Overall, the growth rate of the head (represented by the slope of LCS) appears to be similar among Arctic charr varieties up to the first hatching stage (Figure 3.5).

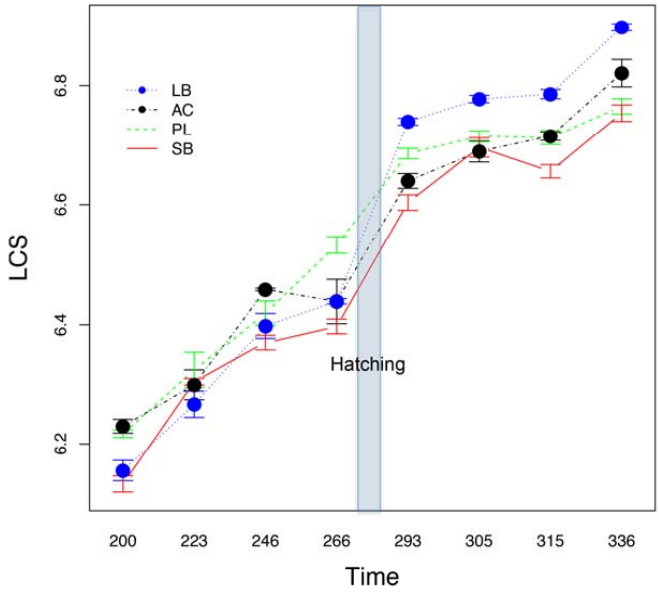

Figure 3.5 Pre- and post-hatching growth of the head of four Arctic charr varieties. Size is shown as the Log transformed centroid size (LCS) ± 95% c.l. for AC (black), LB (blue), PL (green) and SB (red) at eight developmental time points (200, 223, 246, 266, 293, 305, 315 and 336 τs). A light blue bar indicates the hatching period.

A more detailed analysis (ANOVA: LCS ~ Morph x Time) showed that the morphs differ significantly in head size ($F_{3,21} = 103.133$, $p < 0.0001$), and posthoc tests showed that the LCS is significantly different between all pairs of morphs except PL and AC (Table 3.3). The significant interaction of Morph and Time effects ($p < 0.0001$) also indicates differences in the slope of growth among morphs at this fine developmental time scale.

Table 3-3 Differences in head size (LCS) between four Arctic charr morphs (AC, LB, PL and SB) and eight developmental stages.

	Df	SS	MS	F	P
Morph	3	0.452	0.151	103.133	<2e-16
Time	7	10.417	1.488	1019.746	<2e-16
M x T	21	0.263	0.013	8.595	<2e-16
Total	226	0.330	0.002		

	diff	lwr	upr	p adj
LB-AC	0.053	0.035	0.070	0.000
PL-AC	0.016	-0.001	0.034	0.076
SB-AC	-0.063	-0.080	-0.045	0.000
PL-LB	-0.036	-0.053	-0.019	0.000
SB-LB	-0.115	-0.132	-0.098	0.000
SB-PL	-0.079	-0.097	-0.061	0.000

Noticeably, LB and SB start out having smaller heads than PL and AC, and while PL shows a steady and more or less constant rate of increase in head size during the first 5 intervals, the growth in AC, LB and SB appears to slow down prior to hatching but then may enjoy a spurt of growth around hatching. This was most prominent in LB which had the largest heads in all samples taken after hatching. Although the increase of head size was consistent throughout development for all studied groups, two deviations from this pattern can be seen (Figure 5): what appears to be a decrease in head size between stages 246-266 τs for AC and between stages 305-315 τs for SB. We think this likely reflects stochastic fluctuations as we had a relatively low sample size, especially for the pre-hatching stages.

The sharp increase in head size during hatching is also reflected in craniofacial shape changes. PCA analysis involving all the shape data (all stages and morphs) showed that the first three components captured most of the variation (Figure 6). The two periods (pre- and post-hatching) separate along PC1 (explains 34% of the variation) (Figure 6). Changes in shape along PC1 are usually strongly associated with changes in size, i.e. growth. This was confirmed by regressing PC1 scores on LCS (variation explained = 87%, $p < 0.0001$). So, as expected, the overall changes in shape along the PC1 axis are overwhelmingly allometric. Shape changes along PC2 (24%) mainly included mainly narrowing of the lower jaw (Figure 6A), whereas changes along PC3 (19%) consisted in a forward protrusion of both the hyoid arch and the lower jaw (Figure 6B). Interestingly the scores for PC2 and 3 appear to correlate with PC1 in the embryonic, but not in the post-hatching stages.

The role of allometry was further investigated with a multivariate regression of shape (represented by the Procrustes distance coordinates) on size (LCS). The results showed that size related shape changes during this period account for 29.9 % of the total shape variation and this result was highly significant ($p < 0.001$). Moreover this proportion was very similar for all groups, 35.6%, 33.07%, 25.25% and 31.95% for AC, LB, PL and SB, respectively. Interestingly, the PCA results clearly indicate that there is a shift in shape changes following hatching.

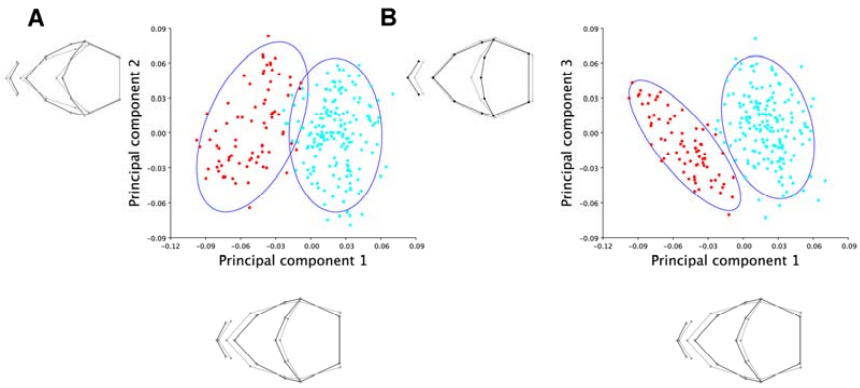

Figure 3.6 Scatter plots of the PCA scores (A) PC2 on PC1 and B) PC3 on PC1) for the entire dataset (all 8 developmental stages) of Procrustes distances. The embryonic stages are shown with red dots and post-hatching stages with blue dots. Confidence ellipses are set to 90%. The scale factor represents a change in Procrustes distance and it's set to 0.1 Wireframes depict shape changes associated with the two Principal Components shown in each graph. In the wire frames the extreme - shape is shown in black and the extreme + shape is shown in grey.

To investigate the shape changes associated with ontogeny in the different Arctic charr groups we did a Procrustes ANOVA (Table 4). Our results showed that, as expected, the largest amount of the total shape variation was attributed to developmental time ($F_{7, 257} = 37.9$, $p < 0.001$) and a significant but smaller amount of the shape variation was attributed to Morph ($F_{3,257} = 7.6$, $p < 0.001$) and the interaction of Morph by Time ($F_{3,257} = 3.09$, $p = 0.009$).

Table 3-4 Procrustes ANOVA comparing the shape variation over eight developmental time points for four varieties of Arctic charr (AC, LB, PL and SB).

	df	SS	MS	F	P	R^2
Morph	3	0.043	0.014	7.622	0.001	0.039
Time	7	0.503	0.072	37.900	0.001	0.458
M x T	21	0.123	0.006	3.094	0.009	0.112
Total	257	1.097	0.004			

To further analyze the developmental events related to hatching we studied the ontogenetic trajectories comprising four developmental stages, two before and two after hatching and compared them between AC, LB, PL and SB (Figure 3.7). Overall, the ontogenetic trajectories of shape change followed similar overall patterns for all four Arctic charr varieties and showed a positive association with PC1. Shape changes along PC1 mainly reflect extension of the lower jaw and a slight narrowing of the head. The trajectories showed substantial fluctuation along the PC2 axis, which captured a narrowing of the head in the pre-hatching period. Curiously, this is abruptly reversed upon hatching (Figure 3.7). All four morphs showed a somewhat comparable zig-zag pattern. This large fluctuation in PC2 may reflect changes in shape due to hatching (grey bar in Figure 3.7). The three natural morphs from Lake Thingvallavatn (LB, SB and PL) displaied similar starting shape for their trajectories, whereas the aquaculture variety (AC) had a sligtly different starting shape. Interestingly, this appeared to change just before hatching, where AC, LB and PL displaied similar shapes, whereas SB takes a dip along PC2. In the first post hatching stage LB appears to have a different shape from the other three groups and spurs along PC2. The two larger Arctic charr varieties (AC and LB) display similar end shapes of their ontogenetic tragectories. The end shape of the ontogenetic shape trajectory appears also to be relatively similar for the two small morphs from Lake Thingvallavatn (PL and SB).

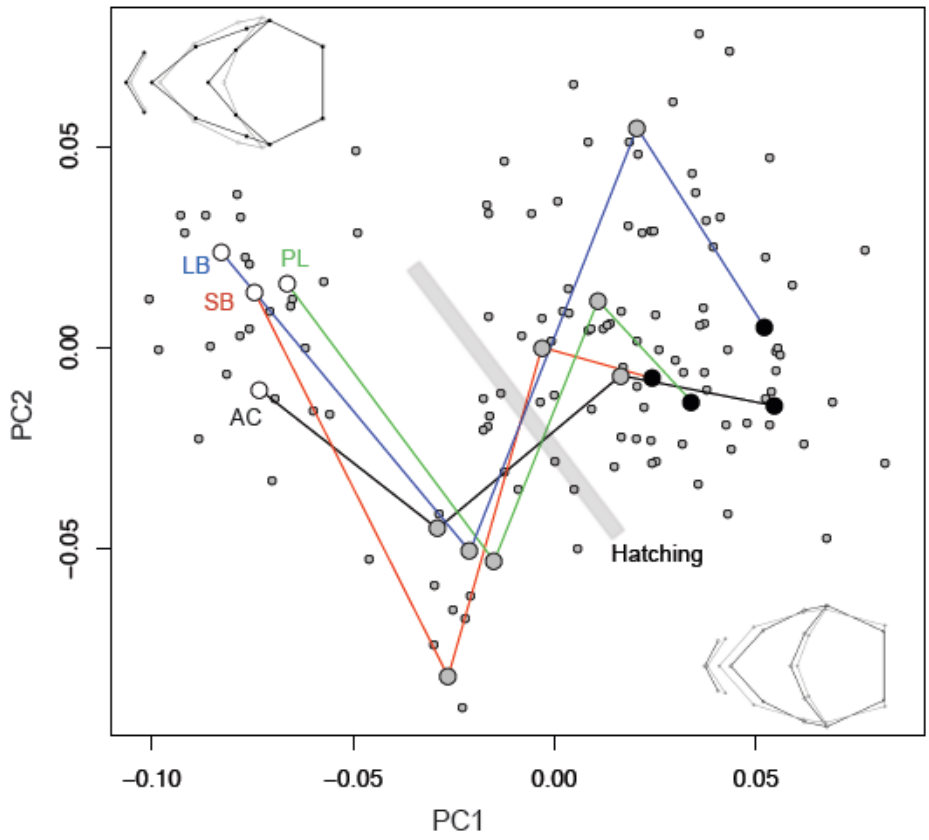

Figure 3.7 Scatter plot of the PCA scores for 4 of the developmental stages (200, 266, 293 and 336 τs). Overlaid are ontogenic trajectories of the four Arctic charr groups (AC = black, LB = blue, PL= green and SB=red) based on. For each trajectory, the first point is shown in white, the middle points are in gray and the last point is displayed in black. Wireframes depict shape changes associated with PC1 (upper left inset) and PC2 (lower right inset (see explanation in Fig. 3.6). Wireframes depict shape changes associated with the two Principal Components shown in each graph. In the wire frames the extreme - shape is shown in black and the extreme + shape is shown in grey.

We used the trajectory tool in geomorph to compare statistically the size, orientation and shape of the ontogenetic trajectories of the three studied groups. It showed (Table 3.5) that; i) the largest Arctic charr morph in this study, LB, differed significantly from AC and PL in the size or magnitude of the ontogenetic trajectory (Table 3.5A), ii) the dwarf form, SB, differed significantly from the three other Arctic charr varieties in the orientation of the trajectory (Table 3.5B), and iii) SB also differed significantly from LB and AC in the overall shape of the ontogenetic trajectory (Table 3.5C). The data indicate that the small benthic morph differs both in terms of the shape and orientation of the ontogenetic trajectory, at the stages around the hatching of the Arctic charr.

*Table 3-5 Size, orientation and shape defining the evolutionary-ontogenic trajectories of four varieties of Arctic charr (AC, LB, PL and SB). Each component of the evolutionary-ontogenic trajectories was compared between Arctic charr groups (varieties) using 1000 permutations. Significance codes are as follows: '**' 0.01 and '*' 0.05. Four developmental time points were used (200, 266, 293 and 336 τs).*

A) Size:

	AC	SB	PL	LB
AC	-	0.04	0.01	0.08*
SB		-	0.03	0.04
PL			-	0.07*
LB				-

B) Orientation:

	AC	SB	PL	LB
AC	-	90.67**	35.97	15.90
SB		-	69.03*	95.49**
PL			-	34.13
LB				-

C) Shape:

	AC	SB	PL	LB
AC	-	0.33*	0.22	0.23
SB		-	0.23	0.31*
PL			-	0.20
LB				-

3.5 Discussion

Here we studied the chronology of events accompanying the emergence of major craniofacial elements, their growth and shape changes, and the early ossification of accompanying dermal bones in Arctic charr. As expected, the elements taking part are the same and the sequence of events is very similar to what has previously been described for other teleosts (Schilling and Kimmel, 1997). This is not surprising as it is the rate and the timing of developmental processes that are thought to be most essential for evolution (Gould, 1977; Klingenberg, 1998). In Arctic charr the formation, growth and shape changes of the cartilage elements of the head skeleton occur mainly during late embryonic stages and by the time of hatching the majority of these elements are already in place (Figure 3.3). The late embryonic and post-hatching stages are characterised by rapid growth and by ossification of the mandibular elements that are essential for the start of efficient breathing movements and later for feeding (Figure 3.3 and 3.4).

The growth of the head follows a similar general pattern in all groups: a steady growth rate during the embryonic stages and a spurt of growth during hatching. The variation seen among the groups was mainly due to; (i) a temporary slowing of growth just before hatching followed by a spurt of growth during and just after hatching seen in SB, LB and AC, (ii) the spurt of growth being largest in LB, and (iii) the differences in head size at the start of measurements (SB, LB < PL, AC). The spurt of growth during hatching may result from the removal of the mechanical constrains of the chorion (Ninness *et al.*, 2006). In the early post hatching stages the growth of the head appears accelerated in LB compared to SB, the other benthic morph. To some extent the differences in the growth rates trajectories between the four groups under study may stem from maternal effects related to egg size and quality. It has been previously shown that SB has the smallest egg size of all Thingvallavatn morphs (Skulason MSc 1986) and the specific energy content of eggs is significantly higher in LB than in SB and PL (Leblanc, 2011 PhD thesis). This is relevant because the juveniles rely on maternally deposited energy after hatching, as feeding will not start for another 6 weeks. Irrespective of the underlying mechanisms, heterochronic shifts in the timing and rate of growth during development can affect craniofacial morphogenesis and lead to evolutionary change (Alberch *et al.*, 1979).

3.5.1 Changes in trophic apparatus shape related to hatching

As was previously noted the period when hatching occurs is accompanied by a marked increase in head size. This transition was also linked to significant allometric shape changes (Figure 3.6). These shape changes concerned mainly the Meckel's cartilage and the width of the head and to a lesser extent the hyoid arch. While Meckel's cartilage transition from embryonic to post-hatching stage is accompanied with an overall forward protrusion of the lower jaw, the hyoid arch appears to move slightly posteriorly. This phenomenon may have a mechanistic explanation, as upon hatching the lower jaw will be freed of all membranes attaching it to the remaining egg yolk, while the hyoid and gill arches remain attached to the egg yolk for some time after hatching. The allometric shape changes accompanying the hatching process appear to be relatively uniform for all the Arctic charr varieties studied here.

3.5.2 Morph specific changes in shape of the trophic apparatus during development

Studying allometric shape changes of the craniofacial elements during ontogeny in polymorphic groups is a step towards understanding the role of ontogeny in phenotypic divergence. During ontogeny, organisms will change shape and these shape changes will prepare them for vital functions such as respiration and feeding (Zelditch *et al.*, 2012). Here we studied the ontogenetic shape trajectories of four varieties of Arctic charr with different feeding regimes: three natural morphs from Thingvallavatn: two that feed on benthic invertebrates (LB and SB) and one that feeds on planktonic crustaceans in water column (PL), as well as an aquaculture variety (AC). Overall the four varieties displayed similar allometric shape changes with ontogeny as described above for the transition from embryonic to post hatching stages. Interestingly, the starting point of the shape trajectory was similar for the three natural morphs and different from AC, whereas the end point of the trajectory was more similar within the small and the large variaties. This may be due to the fact that during the first stages the cartilage is still being formed, while in the late embryonic stages are characterised with intense growth.

The four groups studied also exhibited significant differences in one or more of the attributes of the multidimensional ontogenetic trajectories, namely their size, their orientation and/or their shape. LB exhibited had the longest trajectory of all studied groups (i.e it accumulated the most shape changes throughput ontogeny) and differed significantly from PL and AC. On the other hand, the small benthic (SB) differed significantly from the other three groups, in terms of the orientation of their ontogenetic shape trajectory. While all the studied groups appear to undergo similar shape changes prior to hatching, the negative progression of SB along PC2 in the stage prior to hatching is much more pronounced and might represent a real SB specific phenomenon related to the pre-hatching stages. Furthermore, at the end of the trajectory, the average SB extends also shortest along PC1. Those two observations probably explain why this morph has significantly different axes of shape variation compared to AC, LB and PL.

Furthermore, SB displayed marginally significant differences in the shape of the ontogenetic trajectory compared to LB and AC (but not PL). Unlike the other two attributes of the ontogenetic trajectory (size and orientation) the differences in the shape of ontogenetic trajectories are more difficult to interpret (Collyer and Adams, 2013). These differences may imply that SB exhibits differences in the size and/or the orientation in specific portions of its ontogenetic trajectory (Collyer and Adams, 2013). Interestingly, AC and PL did not show significant differences in any of the three attributes of the ontogenetic trajectories. These two groups are characterised by limnetic type morphology, therefore the similarity of their ontogenetic shape trajectories might be related to the similarity in the processes of involved in the building of this morphology.

In summary, Arctic charr display segmental development of the pharyngeal arches as is characteristic for all vertebrates, and the order of events accompanying the craniofacial development is the same as in other teleosts. The four Arctic charr varieties under study showed similar general patterns of head growth during this period of development with a slowing down of growth rate around hatching. On a finer scale the growth rate differed among groups. The head starts out smaller in the benthic morphs but, due to a sharp increase in growth rate at hatching, the LBs end up with the largest heads at the post-

hatching stages. The hatching period appears to be associated with significant allometric shape changes. Finally, the SB morph differed markedly from the others in terms of the orientation and/or shape of their ontogenetic trajectories. Together the data illustrate the strength of applying multivariate geometrics to analyses of recently evolved trophic polymorphism during early development.

3.6 References

Adams DC, Collyer ML (2007). Analysis of character divergence along environmental gradients and other covariates. *Evolution* **61**: 510–5.

Adams DC, Collyer ML (2009). A general framework for the analysis of phenotypic trajectories in evolutionary studies. *Evolution* **63**: 1143–54.

Adams DC, Otárola-Castillo E (2013). geomorph: an r package for the collection and analysis of geometric morphometric shape data (E Paradis, Ed.). *Methods Ecol Evol* **4**: 393–399.

Alberch P, Gould SJ, Oster GF, Wake DB (1979). Size and shape in ontogeny and phylogeny. *Paleobiology* **5**: 296–317.

Albertson RC, Kocher TD (2005). Genetic architecture sets limits on transgressive segregation in hybrid cichlid fishes. *Evolution (N Y)* **59**: 686–690.

Albertson RC, Streelman JT, Kocher TD (2003). Directional selection has shaped the oral jaws of Lake Malawi cichlid fishes. *Proc Natl Acad Sci U S A* **100**: 5252–7.

Albertson RC, Streelman JT, Kocher TD, Yelick PC (2005). Integration and evolution of the cichlid mandible: the molecular basis of alternate feeding strategies. *Proc Natl Acad Sci U S A* **102**: 16287–92.

Albertson RC, Yan Y-L, Titus TA, Pisano E, Vacchi M, Yelick PC, *et al.* (2010a). Molecular pedomorphism underlies craniofacial skeletal evolution in Antarctic notothenioid fishes. *BMC Evol Biol* **10**: 4.

Albertson RC, Yan Y-L, Titus TA, Pisano E, Vacchi M, Yelick PC, *et al.* (2010b). Molecular pedomorphism underlies craniofacial skeletal evolution in Antarctic notothenioid fishes. *BMC Evol Biol* **10**: 4.

Badyaev A, Whittingham L, Hill G (2001). The evolution of sexual size dimorphism in the house finch. III. Developmental basis. *Evolution (N Y)* **55**: 176–189.

Basch ML, Bronner-Fraser M (2006). Neural crest inducing signals. *Adv Exp Med Biol* **589**: 24–31.

Collyer ML, Adams DC (2007). Analysis of two-state multivariate phenotypic change in ecological studies. *Ecology* **88**: 683–692.

Collyer ML, Adams DC (2013). Phenotypic trajectory analysis: comparison of shape change patterns in evolution and ecology. *Hystrix, Ital J Mammal* **24**: 75–83.

Dale RM, Sisson BE, Topczewski J (2009). The emerging role of Wnt/PCP signaling in organ formation. *Zebrafish* **6**: 9–14.

Eiriksson GM, Skulason S, Snorrason SS (1999). Heterochrony in skeletal development and body size in progeny of two morphs of Arctic charr from Thingvallavatn, Iceland. *J Fish Biol* **55**: 175–185.

Frandsen F, Malmquist HJ, Snorrason SS (1989). Ecological parasitology of polymorphic Arctic charr, Salvelinus alpinus (L.), in Thingvallavatn, Iceland. *J Fish Biol* **34**: 281–297.

Gonzalez PN, Hallgrímsson B, Oyhenart EE (2011). Developmental plasticity in covariance structure of the skull: effects of prenatal stress. *J Anat* **218**: 243–57.

Goodall C (1991). Procrustes methods in the statistical-analysis of shape. *J R Stat Soc Ser B* **53**: 285–339.

Gorodilov YN (1996). Description of the early ontogeny of the Atlantic salmon, Salmo salar, with a novel system of interval (state) identification. *Environ Biol Fishes* **47**: 109–127.

Gould SJ (1977). *Ontogeny and Phylogeny*.

Hall BK (1999). *Evolutionary Developmental Biology*. Springer Science & Business Media.

Jamniczky HA, Harper EE, Garner R, Cresko WA, Wainwright PC, Hallgrímsson B, *et al.* (2014). Association between integration structure and functional evolution in the opercular four-bar apparatus of the threespine stickleback, Gasterosteus aculeatus (Pisces: Gasterosteidae). *Biol J Linn Soc* **111**: 375–390.

Jonsson B, Skúlason S, Snorrason SS, Sandlund OT, Malmquist HJ, Jónasson PM, *et al.* (1988). Life History Variation of Polymorphic Arctic Charr (Salvelinus alpinus) in Thingvallavatn, Iceland. *Can J Fish Aquat Sci* **45**: 1537–1547.

Kapralova KH, Gudbrandsson J, Reynisdottir S, Santos CB, Baltanás VC, Maier VH, *et al.* (2013). Differentiation at the MHCIIα and Cath2 Loci in Sympatric Salvelinus alpinus Resource Morphs in Lake Thingvallavatn. (GH Yue, Ed.). *PLoS One* **8**: e69402.

Kapralova KH, Morrissey MB, Kristjánsson BK, Olafsdóttir GÁ, Snorrason SS, Ferguson MM (2011). Evolution of adaptive diversity and genetic connectivity in Arctic charr (Salvelinus alpinus) in Iceland. *Heredity (Edinb)* **106**: 472–487.

Kimmel CB, Hohenlohe PA, Ullmann B, Currey M, Cresko WA (2012). Developmental dissociation in morphological evolution of the stickleback opercle. *Evol Dev* **14**: 326–37.

Kimmel CB, Ullmann B, Walker C, Wilson C, Currey M, Phillips PC, *et al.* (2005). Evolution and development of facial bone morphology in threespine sticklebacks. *Proc Natl Acad Sci U S A* **102**: 5791–6.

Klingenberg CP (1998). Heterochrony and allometry: the analysis of evolutionary change in ontogeny. *Biol Rev Camb Philos Soc* **73**: 79–123.

Klingenberg CP (2010). Evolution and development of shape: integrating quantitative approaches. *Nat Rev Genet* **11**: 623–35.

Klingenberg CP (2011). MorphoJ: an integrated software package for geometric morphometrics. *Mol Ecol Resour* **11**: 353–7.

Klingenberg C, Barluenga M, Meyer A (2002). Shape analysis of symmetric structures: Quantifying variation among individuals and asymmetry. *Evolution (N Y)* **56**: 1909–1920.

Knight RD, Schilling TF (2006). Cranial neural crest and development of the head skeleton. *Adv Exp Med Biol* **589**: 120–33.

Leblanc C, PhD thesis 2012. The importance of egg size for the diversity of salmonids (http://hdl.handle.net/1946/10867)

Lumsden A, Sprawson N, Graham A (1991). Segmental origin and migration of neural crest cells in the hindbrain region of the chick embryo. *Development* **113**: 1281–1291.

Malmquist HJ, Snorrason SS, Skulason S, Jonsson B, Sandlund OT, Jonasson PM (1992). Diet differentiation in polymorphic Arctic charr in Thingvallavatn , Iceland. *J Anim Ecol* **61**: 21–35.

Neuhauss SCC, Solnica-Krezel L, Schier AFF, Zwartkruis F, Stemple DLL, Malicki J, *et al.* (1996). Mutations affecting craniofacial development in zebrafish. *Development* 123: 357–367.

Ninness MM, Stevens ED, Wright PA (2006). Removal of the chorion before hatching results in increased movement and accelerated growth in rainbow trout (Oncorhynchus mykiss) embryos. *J Exp Biol* **209**: 1874–82.

Palsson A, Gibson G (2004). Association between nucleotide variation in Egfr and wing shape in Drosophila melanogaster. *Genetics* **167**: 1187–98.

Parsons TE, Kristensen E, Hornung L, Diewert VM, Boyd SK, German RZ, *et al.* (2008). Phenotypic variability and craniofacial dysmorphology: increased shape variance in a mouse model for cleft lip. *J Anat* **212**: 135–43.

Parsons KJ, Sheets HD, Skúlason S, Ferguson MM (2011). Phenotypic plasticity, heterochrony and ontogenetic repatterning during juvenile development of divergent Arctic charr (Salvelinus alpinus). *J Evol Biol* **24**: 1640–52.

Parsons KJ, Skúlason S, Ferguson M (2010). Morphological variation over ontogeny and environments in resource polymorphic arctic charr (Salvelinus alpinus). *Evol Dev* **12**: 246–57.

Parsons KJ, Trent Taylor a., Powder KE, Albertson RC (2014). Wnt signalling underlies the evolution of new phenotypes and craniofacial variability in Lake Malawi cichlids. *Nat Commun* **5**: 1–11.

Perry G, Audet C, Laplatte B, Bernatchez L (2004). Shifting patterns in genetic control at the embryo-alevin boundary in brook charr. *Evolution (N Y)* **58**: 2002–2012.

Roberts RB, Hu Y, Albertson RC, Kocher TD (2011). Craniofacial divergence and ongoing adaptation via the hedgehog pathway. *Proc Natl Acad Sci U S A* **108**: 13194–9.

Rohlf FJ (2006). "tpsDig, version 2.10." *http://life.bio.sunysb.edu/morph/index.html*.

Rohlf FJ, Slice D (1990). Extensions of the Procrustes Method for the Optimal Superimposition of Landmarks. *Syst Zool* **39**: 40.

Sandlund OT, Gunnarson K, Jonasson PM, Jonsson B, Lindem T, Magnusson KP, *et al.* (1992). The Arctic charr Salvelinus alpinus in Thingvallavatn. *Oikos* **64**: 305–351.

Schilling TF (1997). Genetic analysis of craniofacial development in the vertebrate embryo. *Bioessays* **19**: 459–68.

Schilling T, Kimmel C (1997). Musculoskeletal patterning in the pharyngeal segments of the zebrafish embryo. *Development* **124**: 2945–2960.

Shapiro MD, Marks ME, Peichel CL, Blackman BK, Nereng KS, Jónsson B, *et al.* (2004). Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature* **428**: 717–723.

Skulason S, Noakes DL. G, Snorrason SS (1989). Ontogeny of trophic morphology in four sympatric morphs of arctic charr Salvelinus alpinus in Thingvallavatn, Iceland*. *Biol J Linn Soc* **38**: 281–301.

Snorrason SS, Skulason S, Sandlund OT, Malmquist HJ, Jonsson B, Jonasson PM (1989). Shape polymorphism in sympatric Arctic charr, Salvelinus alpinus in Thingvallavath, Iceland. *Physiol Ecol Japan* **1**: 393–404.

Svavarsson E (2007). Árangur í kynbótum á bleikju og næstu skref [reference in icelandic]. *Fræðaþing landbúnaðarins (conference proceedings)* **4**: 121–125.

Szabo-Rogers HL, Smithers LE, Yakob W, Liu KJ (2010). New directions in craniofacial morphogenesis. *Dev Biol* **341**: 84–94.

Walker MB, Kimmel CB (2007). A two-color acid-free cartilage and bone stain for zebrafish larvae. *Biotech Histochem* **82**: 23–8.

Westneat MW (2005). *Fish Biomechanics*. Elsevier.

Yelick P, Schilling T (2002). Molecular dissection of craniofacial development using zebrafish. *Crit Rev Oral Biol Med* **13**: 308–322.

Young NM, Chong HJ, Hu D, Hallgrímsson B, Marcucio RS (2010). Quantitative analyses link modulation of sonic hedgehog signaling to continuous variation in facial growth and shape. *Development* **137**: 3405–9.

Zelditch ML, Swiderski DL, Sheets HD (2012). *Geometric Morphometrics for Biologists: A Primer (Google eBook)*. Academic Press.

3.7 Appendix

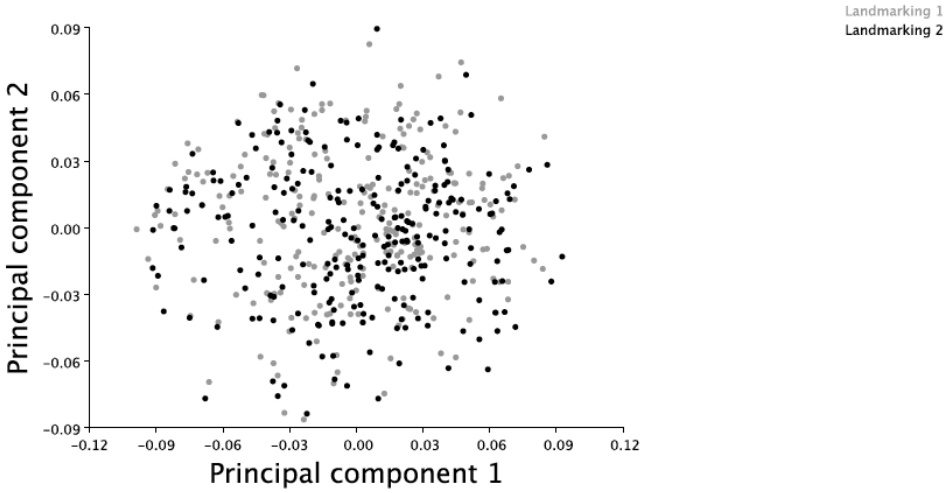

Figure S 3.1 - Scatter plot of the PCA scores for the entire dataset. The two separate landmarking sessions are shown in grey (Landmarking session 1) and black (Landmarking session 2). No separation between the two landmarking sessions could be detected.

Paper 3

4 Paper III

Evidence for reproductive isolation by phenotypic transgression in hybrids of Arctic charr

Kalina H. Kapralova, Arnar Palsson, Bjarni K. Kristjanson, Zophonías O. Jónsson, Sigurður S. Snorrason

Authors' contribution: Conceived and designed the experiments: KHK SSS ZOJ AP BKK. Performed the experiments: KHK AP BKK ZOJ SSS. Analysed the data: KHK AP BKK ZOJ SSS. Contributed reagents/materials/analysis tools: KHK BKK SSS AP ZOJ. Wrote the paper: KHK SSS AP ZOJ BKK.

4.1 Abstract

This paper describes subtle but significant differences in early post-hatching craniofacial morphology between the progeny of three sympatric morphs of Arctic charr from Lake Thingvallavatn, Iceland (LB, SB and PL). Furthermore the effect of of hybridisation on the craniofacial morphology of Arctic charr was investigated by creating reciprocal crosses between two contrasting Arctic charr morphologies PL and SB. Interestingly, the hybrid crosses exhibited extreme (or transgressive) craniofacial phenotypes compared to the pure morph breeds, indicating that the ecological divergence within the lake might be enhanced by lowered fitness of hybrids. While the shape of the feeding apparatus of the two reciprocal hybrid crosses differed significantly from both pure crosses, no significant differences in head shape were detected between the two reciprocal crosses, suggesting that genetic effects outweigh maternal effects.

4.2 Introduction

Understanding the developmental processes that translate genetic variation into phenotypes is a central aim in evolutionary and developmental biology. While a great deal of work has focused on genetics (Mousseau and Roff, 1987; Kruuk *et al.*, 2008) and selection (Kingsolver *et al.*, 2001), less is generally known about the process by which phenotypic variation arises (Travisano and Shaw, 2013). Evolution acts on a wide range of ecologically-important traits, but in general morphological traits of ecological relevance are easiest to study and the genotype to phenotype map best understood (Smith and Skulason, 1996; Grant, 1999; Marchinko, 2009). This is in part because we have considerable understanding of the principles of development and morphogenesis (Tickle, 2011) and many analytical methods are available for studies of variation in morphology (Bookstein, 1997; Slice, 2006; Zelditch *et al.*, 2012). Thus, morphological variation currently is an ideal tool for disentangling the genotype-phenotype map for ecologically improtant traits in an evolutionary context.

Morphological traits related to habitat use and especially to feeding have been widely studied, not only on a macroevolutionary scale (Sidor, 2001; Darrin Hulsey, 2006; Bhullar *et al.*, 2012), but also within closely related species (Schluter, 2000; Abzhanov *et al.*, 2004; Roberts *et al.*, 2011a) and even within populations (Smith and Skulason, 1996; Schluter and Rambaut, 1996; Kimmel *et al.*, 2005; Landry and Bernatchez, 2010). Classic examples of adaptive radiations resulting in trophic morphology diversification include Darwin's finches (Lack, 1945; Bowman, 1961; Grant, 1999), African cichlids (Fryer, 1972; Cooper *et al.*, 2010) and honeycreepers (Amadon, 1950; Freed *et al.*, 1987). Adaptive radiation and divergence can be envisaged as localized ecological selection acting on an ancestral phenotype where variations in the natural range of ancestral phenotypes tend towards maximizing fitness at “adaptive peaks”. As fitness is maximized at adaptive peaks, intermediate phenotypes in valleys between the adaptive peaks are expected to have lower fitness (Simpson, 1965). Differences in fitness between “peak” and “valley” phenotypes is a mechanism that eventually leads to the emergence of reproductive isolation (Gavrilets and Losos, 2009). The tempo and mode of reproductive isolation is related to the relative fitness differences between peak and valley phenotypes but gene flow among incompletely isolated populations may still occur (Seehausen *et al.*, 2014). Although gene flow is often considered to break down adaptive divergence by homogenizing genetic variation among populations, occasionally it can also facilitate adaptive divergence by maintaining genetic diversity (Swindell and Bouzat, 2006). Whether incompletely isolated populations progress towards speciation or divergence breaks down will depend on the underlying genetic architecture (Nosil *et al.*, 2009) and the antagonism between selection and recombination (Felsenstein, 1981).

Crossbreeding between closely related species and/or divergent populations can sometimes have a positive effect on fitness and ultimately lead to diversification. This diversification by hybridisation can be triggered by a fluctuating environment, where even a small proportion of the hybrids are fitter than the parental populations, leading to establishment of new alleles by introgression (Dowling and DeMarais, 1993; Salzburger *et al.*, 2002) or even to new evolutionary lineages (Barton, 2001). However, outcrossing does not usually enhance fitness but most often leads to outbreeding depression, where hybrids between closely related species and even divergent populations of the same species exhibit lower

fitness than the parental groups (Burke and Arnold, 2001). Outbreeding depression in hybrids reflects accumulation of incompatible genes and is usually attributed to the breaking up of co-adapted loci or favourable epistatic relationships (Lynch, 1991; Matute *et al.*, 2010).

The vertebrate skull is one of the most complex anatomical units and displays an astonishing diversity in shapes among taxa, involving interactions between derivatives of all three germ layers (neural crest, mesodermal mesenchyme and surrounding epithelia). Bones determine how adjacent soft tissue and musculoskeletal elements connect and function together. Studying the growth and shaping of cartilage and bone elements of the feeding apparatus during morphogenesis and its subsequent remodelling can help us understand how trophic morphologies evolve in response to environmental cues. Although natural populations do not have the advantages of model organisms in terms of ease of experimental and genetic manipulation, they offer a unique possibility to examine actual evolutionary processes (Klingenberg, 2010). Ongoing work on various evolutionary models such as Galapagos finches (Abzhanov *et al.*, 2004, 2006; Mallarino *et al.*, 2011, 2012), African cichlids (Albertson and Kocher, 2005; Albertson *et al.*, 2005; Roberts *et al.*, 2011b; Parsons *et al.*, 2014) and threespine sticklebacks (Kimmel *et al.*, 2005, 2012; Jamniczky *et al.*, 2014) is generating important insights into the mechanisms involved in the developmental aspects of morphological evolution.

Many morphological traits can be quantified on the basis of univariate measurements of key structural and/or functional elements, for example the length and width of the beak in birds (Badyaev *et al.*, 2001), length of the jaws, or measures of mechanical levers in fish (Albertson *et al.*, 2003, 2005). Some structures call for more complex description of shape that can summarise global and local changes in shape and proportions (Palsson and Gibson, 2004; Klingenberg, 2010). Such methods have successively been applied in studying morphological variation in trophic development (Parsons *et al.*, 2008, 2014; Young *et al.*, 2010; Gonzalez *et al.*, 2011).

North-temperate lacustrine ecosystems are ideal for studying the development and evolution of ecologically important morphological traits. These systems are evolutionarily youngas much of their diversity arose after the last glacial period (about 10,000 years ago), and often they exhibit high levels of phenotypic polymorphism. Such polymorphisms can be subtle, but often populations are composed of distinct phenotypes (morphs) that show differences in morphological, ecological and life history traits. Looking at these traits across populations (and even species) also reveals parallelisms as seen in the repeated occurrence of benthic and limnetic morphs which are thought to have arisen via common adaptations to the different challenges in foraging on benthic versus open water (pelagic) prey (see references in (Skulason and Smith, 1995; Schluter and Rambaut, 1996)).

Icelandic Arctic charr (*Salivelunus alpinus*) originates from a single Atlantic lineage (Brunner *et al.* 2001). This species shows an extremely high level of variation in phenotype between populations and many examples of polymorphism (i.e. sympatric morphs) have been documented (Gíslason *et al.*, 1999; Snorrason and Skúlason, 2004; Woods *et al.*, 2012). The four morphs of Arctic charr in Lake Thingvallavatn represent an extreme case of intralacustrine diversity. These morphs, grouped into two morphotypes, differ greatly in morphology of the trophic apparatus (Snorrason *et al.*, 1989). The two morphs belonging to the limnetic morphotype, a planktivorous (PL) and a piscivorous (PI) charr have pointed snouts and evenly protruding jaws, while the two benthic morphs, a

small (SB) and a large benthic (LB) have blunt snouts, short lower jaws and relatively large pectoral fins (Snorrason *et al.*, 1989). The four Arctic charr morphs also exhibit strikingly clear differentiation in life history characteristics and ecology as reflected in different habitat use, diet and endoparasite fauna (Jonsson *et al.*, 1988; Frandsen *et al.*, 1989; Malmquist *et al.*, 1992; Sandlund *et al.*, 1992). Common-garden experiments have indicated that the development of Arctic charr trophic morphologies and behaviour are the result of both genetic differences (Skulason *et al*., 1989a; Skúlason *et al.*, 1993) and putatively adaptive plastic responses to different environments (Parsons *et al.*, 2010; Parsons, Sheets, *et al.*, 2011).

The charr morphs all spawn in the stony littoral habitat but the timing of spawning and the level of synchronization differs among the morphs (Skúlason *et al*., 1989b). Opportunities for interbreeding among morphs do exist, and in the case of the two smallest and most abundant morphs (PL and SB) interbreeding opportunities seem wide open. Yet, a recent study, using neutral microsatellite markers revealed subtle but significant genetic differentiation between the three most common morphs in Lake Thingvallavatn (LB, PL, and SB). Further, coalescent simulations indicated a scenario of early evolution of reproductive isolation followed by slow divergence by drift with restricted gene flow (Kapralova *et al.*, 2011). Notably, a study of immune system genes revealed more pronounced genetic differentiation, consistent with a scenario where parts of the immune systems have diverged substantially among Arctic charr morphs in lake Thingvallavatn (Kapralova *et al.*, 2013). Previous studies of the ontogeny of the Thingvallavatn morphs indicate a clear genetic basis for their morphological differences, likely rooted in developmental heterochrony with significant maternal effects (Skulason *et al*., 1989a). The role of developmental heterochrony in the evolution of the Thingvallavatn Arctic charr morphs was further demonstrated in a study showing that some skeletal elements of the head start ossifying earlier and/or faster in small benthivorous embryos than in embryos derived from the planktivorous morph (Eiriksson *et al.*, 1999).

Here we use the recently evolved Arctic charr morphs of Lake Thingvallavatn and landmark based geometric morphometrics to address questions about the evolution and the development of diverging craniofacial morphologies. To this end we generated pure morph crosses between three natural morphs of the Thingvallavatn morphs (SB, LB and PL), as well as an Aquaculture strain (AC). When juveniles emerge and start active foraging and feeding, the environment can induce plastic responses, e.g. different types of prey can induce changes in craniofacial morphology (Parsons *et al.*, 2010). To minimise the environmental effects on morphological variation, we used a common garden set-up and studied early post-hatching stages spanning from immediately after hatching to shortly before the start of exogenous feeding. We predicted that craniofacial variation between morphs, although subtle, would already be detectable in the early post-hatching stages. We used a similar experimental set up to address questions regarding the effect of hybridisation between two contrasting Arctic charr morphologies (SB and PL) on the craniofacial morphology of the hybrid progeny. To account for maternal effects on craniofacial morphology we established reciprocal hybrid crosses between PL and SB. Using a similar experimental setup, although at later developmental stages, (Skulason *et al*., 1989a) showed that some hybrid crosses between Arctic charr morphs have intermediate morphologies and other crosses strongly resemble the maternal parental group. We expect to detect these effects of hybridisation on craniofacial morphology in early post hatching stages before the onset of exogenous feeding.

4.3 Material and methods

4.3.1 Sampling

For this study we established developmental series for six crosses, four pure morph and two hybrid crosses. The pure morph crosses included three morphs (LB, SB and PL) from Lake Thingvallavatn and one of a strain from the Hólar aquaculture station (AC). The two reciprocal hybrid crosses were made between PL and SB. For that mature fish from the planktivorous (PL), small and large benthic morphs (SB and LB) were caught in Lake Thingvallavatn using gill-nets. Fishing permissions were obtained from the Thingvellir National Park Commission and the landowner of Mjóanes farm. The AC crosses were made with parents from the Hólar breeding program (Svavarsson, 2007). A total of six experimental groups (Table 4.1) were set up for this study (4 pure morph crosses and 2 reciprocal hybrid crosses). For each pure morph cross, eggs from several females were pooled and fertilized using milt from several males from the same group. In the case of the reciprocal hybrid crosses between PL and SB, eggs from several PL females were pooled and fertilized with the milt of several SB males and vice versa. After stripping of gametes the fish were killed by a sharp blow to the head. The eggs were reared at approximately 5°C in a hatching tray (EWOS, Norway) under constant water flow and in complete darkness at the Holar University College experimental facilities in Verið, Sauðárkrókur. The rearing and collection of the embryos was performed according to Icelandic regulations (licence granted to Holar University College aquaculture and experimental facilities in Verið, Sauðárkrókur). Exact water temperature was recorded twice daily to estimate the relative age of the embryos using tau-somite τ_s units defined as the time it takes for one somite pair to form at a given temperature (Gorodilov, 1996). Embryos were collected throughout development (Table 4.1) and fixed in 4% PFA.

Table 4-1 Sampling scheme: shown are the developmental stages (in tau units), the number of individuals per morph for each time point and staining batches. AC=Aquaculture charr from the Hólar breeding stock, LB and SB =large and small benthic charr, respectively, PL=Planctivotouscharr, PLx= hybrid cross between PL female and SB male, SBx=hybrid cross between SB female and PL male.

Stage	Batch	AC	LB	PL	SB	PLx	SBx
293	1	12	10	12	12	8	10
305	2	12	14	15	10	-	-
315	1	15	12	8	9	9	9
336	2	6	13	10	11	-	-
346	1	10	9	11	10	10	10
370	2	10	8	9	10	-	-
Total		65	66	65	62	27	29

4.3.2 Staining and photographing

A total of 314 individuals (Table 4.1) were stained for cartilage (alcian blue) and bone (alizarin red) using a modified protocol from (Walker and Kimmel, 2007). Individuals were placed in a petri dish containing 50 ml of 1% agarose gel and immobilised with dissecting needles to insure the correct positioning of the embryo. The head of each individual was photographed ventrally facing left using a Leica (MZ10) stereomicroscope at the same magnification (2.0x) for each photo. Samples were stained and photographed in two batches (Table 4.1). To test for the effect of staining batch on shape, we performed Principal components analysis (PCA) of the morph dataset (excluding the hybrids, as they were all part of the same staining batch). PCA showed no separation between the two staining batches (Figure S 4.1).

4.3.3 Geometric morphometrics

We selected 46 landmarks to describe the craniofacial shape including the developing lower jaw, the hyoid arch, pharyngeal arches and the tip of the ethmoid plate and maxillas (Figure 4.1) and digitized them using tps.DIG2 (Rohlf, 2006). Although the mandible and the ethmoid plate do not occupy the same 2D space, at the stages of development under study the head is rather flat and the landmarks describing the front of the ethmoid plate were used as an approximation representing the tip of the mouth. The shape information for each specimen was extracted using a Generalized Procrustes analysis (GPA) in MorphoJ (Klingenberg, 2011) where, after accounting for scale, position and orientation, all specimens are superimposed to a common coordinate system (Rohlf and Slice, 1990; Goodall, 1991). Only the symmetric component of shape variation (Klingenberg *et al.*, 2002) was used for subsequent statistical analyses. The Centroid Size (defined as the square root of the sum of the squared distances of all landmarks from their centroid) of each specimen was retained after the Procrustes fit and used as a measure of individual size.

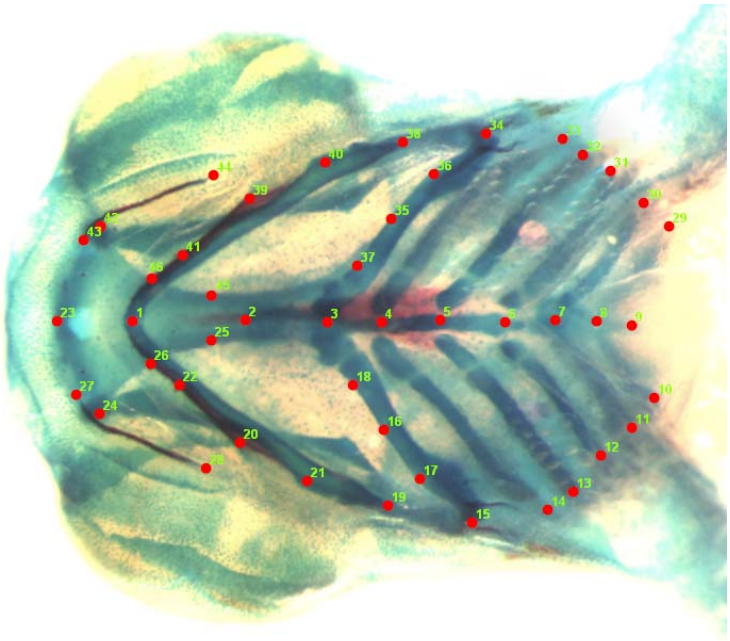

Figure 4.1 The 46 landmarks (red dots) used in this study. The landmarks were selected to describe major craniofacial elements such as the ethmoid plate, the maxillas, Meckel's cartilages, the hyoid arches, the gill arches, the basihyal and basibranchial cartilages.

The data was screened for outliers using the function "find outliers" in MorpoJ. Samples showing large deviations from the average landmark position were examined carefully and either re-scored, when the deviation was present in only one of the two symmetric sides, excluded if the specimen appeared to be damaged (11 specimen) or left in the data set if they appeared to represent valid extremes in the natural sample variation (9 specimen).

To quantify measurement error we used Procrustes ANOVA (Klingenberg *et al.*, 2002) in MorphoJ. Digitizing error was about 7x smaller than the smallest biological effect (Individual/Side). Although the digitizing error was rather small, we digitized each individual 3 times and the results from the repeated measurements were averaged in the final dataset.

To characterise allometry (variation in shape caused by variation in size), we regressed shape on size (represented by the log transformed Centroid size or LCS) in MorphoJ. The null hypothesis of this test is that shape does not change as a function of size (i.e that growth is isometric). We used 10.000 permutations to estimate the statistical significance of the test.

In order to investigate whether the different morphs undergo the same allometric change in shape i.e whether the observed differences in shapes can be due to the impact of size on shape, we conducted a permutational MANCOVA (multivariate analysis of covariance) with LCS as a covariate using the "vegan" package in R (Dixon, 2003) and following the methodology described in (Zelditch *et al.*, 2012). Briefly, we tested whether the difference in shape between experimental groups depends on the size at which they are compared. If

the interaction between LCS and Morph is not significant and the null hypothesis cannot be rejected, the effect of shape on size can be removed by using the regression residuals obtained from the regression of shape on LCS. If, the null hypothesis is rejected, regression can still be used to account for allometry; but must be done for each group separately (Zelditch *et al.*, 2012).

4.3.4 Quantifying shape differences

To quantify shape differences among morphs we used a combination of multivariate analysis in MorphoJ. After accounting for allometry, the data containing all the morphs and developmental time-points was subjected to Principal component analysis (PCA). Next we studied the effect of Morph, Time and (Morph x Time) of each of the 6 major Principal Components (PCs) with a generalised linear model in R (3.0.2 R Core Team 2012).

Canonical Variate analysis (CVA) was used to visualize the differences among groups. CVA is an ordination method, which explores shape features distinguishing among *a priori* defined groups. Differences between extremes are used to illustrate shape differences for both PCA and CVA. Both Mahalanobis distances (which measures the distances of separation between two groups scaled by the standard deviation in the respective directions) and Procrustes distances (which measures the absolute amount of shape variation) were generated and their statistical significance was assessed with 10.000 permutations. The same procedure as described above was used to quantify shape differences between hybrids and the pure parental crosses. All analyses were performed with the reciprocal hybrid crosses as separate groups and as a single combined "hybrid" group.

4.4 Results

We set out to characterize the variation in trophic apparatus in developing Arctic charr, using geometric morphometrics of bone and cartilege structures on the ventral side of the developing head. First we tested for a potential effects of growth on shape. Allometry was found to play a significant role ($p < 0.0001$) in post-hatching craniofacial development, both in pure Morph (Figure S4.2) and hybrid progeny (Figure S 4.3) where it accounted for 8% and 11.8% of the total shape variation, respectively. The permutational MANCOVA showed highly significant ($p < 0.001$) Morph and Size effects, with the Size effect being 4.5 times larger than the Morph effect (Table S4.1). No significant Morph x Size interaction was detected, indicating that that the differences in shape between morphs do not depend on the size at which they are compared. A similar pattern was observed for the data including the hybrid and parental progeny groups, although the Morph effect on shape variation was larger and comparable to the LCS (Table S4.2).

4.4.1 Quantifying craniofacial shape differences among 4 morphs of Arctic charr (AC, SB, LB and PL)

The first 6 Principal Components (PC) accounted for 82.3% of the variation in craniofacial shape, with the first 3 alone accounting for 69.5% of the variation (Table S3). All of the 3

major PCs had a significant Morph effect. PC1 (accounting for 30.5% of the variation) and PC2 (22.7%) showed a significant time effect and PC3 (16.3%) exhibited a significant (Morph x Time) effect (Table S4.3). The 3 major PCs showed shape variation in all craniofacial elements under study: the pharyngeal arches, maxillas the lower jaw, hyoid arch and ethmoid plate (Figure S4.3). The craniofacial shape differences among morphs were further characterized with Canonical Variate Analysis (CVA). Shape changes associated with CV1 (50.27%) include the pharyngeal arches, maxillas and the hyoid arch (Figure 4.2). CV2 (28.32%) showed more subtle changes in the pharyngeal arches, maxilla length and more pronounced differences in the shape of the lower jaw and hyoid arch (Figure 4.2). Finally, CV3 (21.41%) showed pronounced overall shape changes, including the pharyngeal arches, maxillas the shape of the lower jaw and hyoid arch and ethmoid plate. AC separated from the three morphs from Thingvallavatn (LB, SB and PL) along CV1, PL separated from the two benthic morphs LB and SB along CV2 whereas SB separated from the rest along CV3 (Figure 4.2).

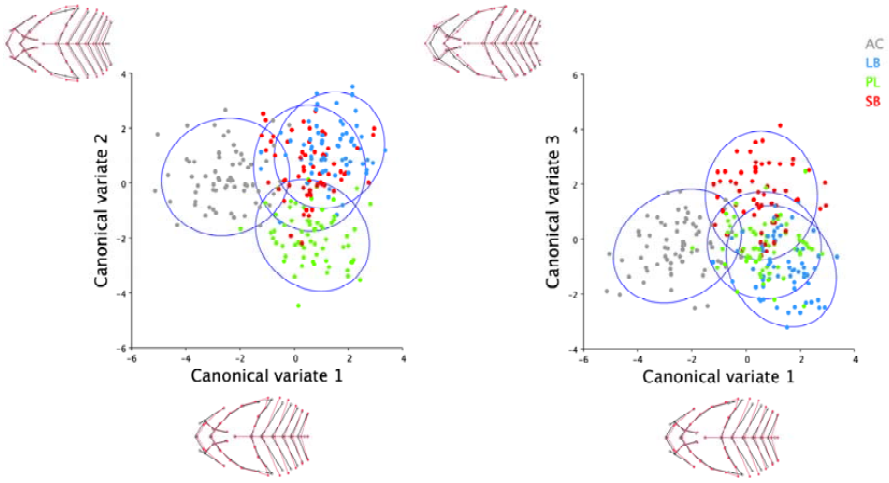

Figure 4.2 Scatter plot of the CVA scores for four morphs of Arctic charr (AC=grey, LB=blue, PL=green and SB=red). Wireframes depict shape changes associated with the two Canonical Variates shown in each graph. In the wire frames the extreme negative value is shown in black and the extreme positive values in red. The scale factor is in units of Mahalanobis distance and it's set to 5. Confidence ellipses are set to 90%.

Pairwise comparisons using both Mahalanobis and Procrustes distances showed significant differences between morphs (Table 4.2). Although significant for all comparisons, the shape changes between all pairs of morphs measured by Procrustes distances are fairly subtle (Table 4.2). The shape changes between pairs of morphs measured by Mahalanobis distance are all highly significant. The largest Mahalanobis distances are between AC and the three natural morphs from Lake Thingvallavatn (Table 4.2). Within Lake Thingvallavatn the lowest Mahalanobis distance is between the two benthic morphs LB and SB (2.76; equivalent to 2.76 times the standard deviation for the discriminant function), while the Mahalanobis distances of the two benthic morphs (LB and SB) to the pelagic morph (PL) are larger (3.25 and 3.05) respectively. In sum, the data show clearly

that the four Arctic charr morphs studied here have separable morphologies at this early stage of development, reflecting variation in multiple aspects of the cartilege and bones in the head (visualized from the ventral side).

Table 4-2 Pairwise Mahalanobis (upper panel) and Procrustes distances (lower panel) between morphs and their significance obtained with 10.000 permutations. AC = Aquaculture charr from the Hólar breeding stock, LB and SB = large and small benthivorous charr, respectively, PL = Planktivotous charr.

	AC	LB	PL	SB
AC	-	4.02***	3.85***	3.61***
LB	0.032***	-	3.25***	2.76***
PL	0.028***	0.017*	-	3.05***
SB	0.015*	0.026**	0.021**	-

******p<0.001,**p<0.001, * p<0.05***

4.4.2 Hybrids show significant craniofacial shape differences from the pure parental crosses

The data show that F1 hybrids of SB and PL are significantly different from both parental strains in the shape of craniofacial morphology and these differences are so extreme that they can be qualified as "transgressive" (i.e extreme or novel phenotypes relative to the parental populations).

The first 6 Principal Components (PC) comparing SB, PL and the hybrids, accounted for 84% of the variation, with the first 2 alone accounting for 63.8% of the variation (Table S4.4). The first 2 Principal Components PC1 (44.3%) and PC2 (19.5%) showed a highly significant Morph effect ($p < 0.001$), a significant Time effect and PC1 also exhibited a significant ($p < 0.05$) Morph x Time effect. PC3 (8.7%) however didn't show any Morph, Time or Morph x Time effect (Table S4.4). The hybrid crosses separate from the pure parental crosses along PC1 and to a lesser extend along PC2 (Figure 4.3). Shape changes along PC1 include overall narrowing of the head, seen mainly in the shape of the lower jaw, the hyoid arch and the pharyngeal arches (Figure 4.3). Shape changes along PC2 involve retraction of the ethmoid plate and broadening of the mandibular arches (Figure 4.3).

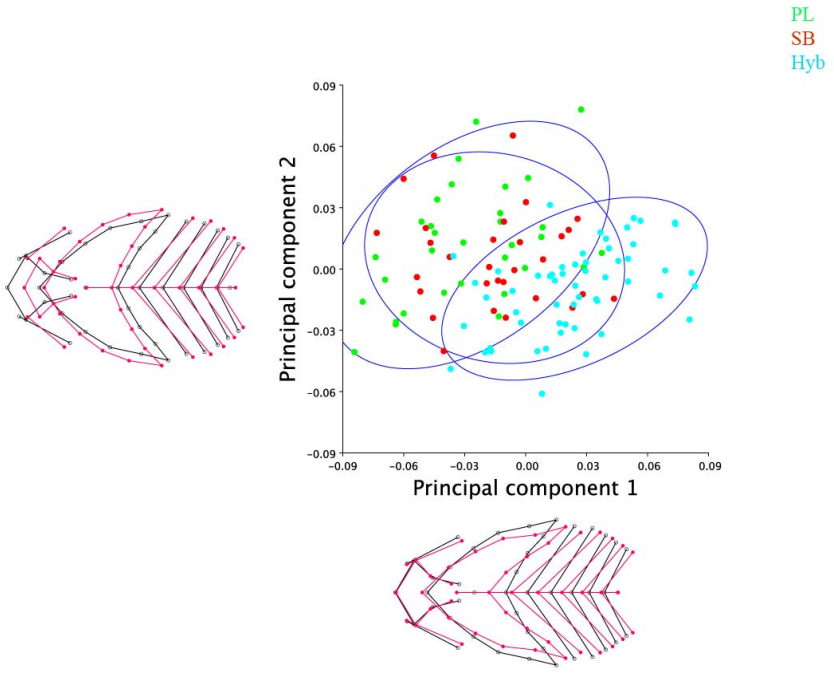

Figure 4.3 Scatter plot of the PCA scores for the pure crosses SB (red) and PL (green) and the hybrid crosses (PL♀ x SB♂and SB♀ x PL♂ combined in blue). Wireframes depict shape changes associated with the two major Principal Components PC 1 (30.5%) and PC2 (22.7%). Wireframes depict shape changes associated PC1 and PC2. In the wire frames the extreme negative value is shown in black and the extreme positive values in red.The scale factor represents a change in Procrustes distance and is set to 0.05. Confidence ellipses are set to 90%.

The differences between the hybrids and among the hybrids and the parental pure crosses were further explored with Canonical Variate Analysis (CVA) (Figure 4.4). Hybrid crosses separate from the SB pure cross along CV1 (52.8%) and from the PL pure cross along CV2 (31.4%). The shape changes along CV1 and CV2 to an extent resemble the ones seen along PC1. i.e an overall narrowing of the head. The reciprocal hybrid crosses PLx and SBx show some separation along CV3 (15.8%) but this is not significant.

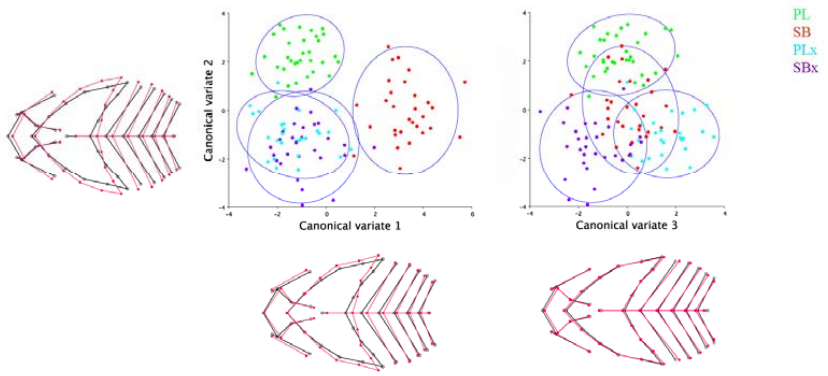

Figure 4.4 Scatter plot of the CVA scores for the pure parental crosses (PL and SB) and the resiprocal hybrid crosses (PLx = PL eggs and SBx = SB eggs). Wireframes depict shape changes associated with each Canonical Variate (CV). The scale factor is in units of Mahalanobis distance and it's set to 5. Wireframes depict shape changes associated with the two Canonical Variates shown in each graph. In the wire frames the extreme negative value is shown in black and the extreme positive values in red. Confidence ellipses are set to 90%.

Pairwise comparisons using Mahalanobis distances showed highly significant differences ($p < 0.001$) between the hybrids and the pure parental strains and between the progeny of the two pure parental strains (Table 4.3). Procrustes distances between the two reciprocal crosses and the hybrids and the pure parental strains are considerably larger than the ones seen between the pure morph crosses (Table 4.3) and highly significant ($p < 0.001$). The Procrustes distances between the reciprocal crosses (SBx and PLx) were very low and not significant ($p = 0.06$) (Table 4.3).

Table 4-3 Pairwise Mahalanobis(upper panel) and Procrustes distances (lower panel) between hybrid and pure parental crosses and their significance obtained with 10.000 permutations.

	PL	SB	PLx	SBx
PL	-	4.91***	3.76***	3.95***
SB	0.020*	-	4.96***	4.72***
PLx	0.058***	0.045***	-	3.08***
SBx	0.057***	0.044***	0.018	-

***** p<0.001, ** p<0.001, * p<0.05**

Because of the strong maternal effects influencing embryonic size in early development (Perry *et al.*, 2004), we compared the head size (represented by LCS) of the hybrids to the pure maternal morph (Figure S4..4). In both comparisons PL vs. PLx and SB vs. SBx and for the three stages tested, the hybrid progeny exhibited lower head size compared to the progeny of the maternal morph. A more detailed analysis (ANOVA: LCS ~ Morph x Time) showed that the morphs differ significantly in head size ($F_{3,106} = 25.448$, $p < 0.0001$), and *post hoc* tests showed that the LCS was significantly different between all pairs of morphs except PLx and SBx (Table S4.5). The significant interaction of Morph and Time ($p < 0.05$) also indicates differences in size among morphs at finer time scales during development.

4.5 Discussion

Teleost fish vary greatly in head morphology and especially in morphology of craniofacial elements related to feeding, such as the lower and the upper jaw elements, operculum bones etc. (Albertson *et al.*, 2003; Kimmel *et al.*, 2005; Jamniczky *et al.*, 2014). These elements emerge early and undergo extensive morphological changes, not only during embryonic development, but also in larval stages and even during adulthood. In this study we focused on a short time frame including only stages from hatching to just before the first feeding starts. During these stages the splanchnocranium (the upper and lower jaws, the ethmoid plate, the hyoid and the pharyngeal arches) is still a relatively simple structure although all of its anatomical elements have already been formed and, as in the case for the Meckel's cartilage, have started to be surrounded by dermal ossification. Significant inter-morph differences in the shape of the major trophic elements observed well before the start of feeding suggest a role of genetic architecture and/or development in the formation of divergent Arctic charr trophic morphologies. The emergence of transgressive phenotypes in the developing splanchnocranium in hybrids, made by crossing the benthic (SB) and limnetic (PL) sympatric morphs, further emphasize a strong genetic component in the trophic polymorphism in the Thingvallavatn charr. This phenomenon and the smaller embryo head sizes seen in the hybrids both indicate the existence of developmental incompatibilities with potentially detrimental effects on hybrid fitness.

4.5.1 Craniofacial shape development and evolution

An important first question in studyies of the developmental origins of morphological differences is how early differences in shape can be detected. For example, the black-bellied African seedcracker *(Pyrenestes ostrinus*) exhibits morphological distinct polymorphism in bill size, which is evident by the time they start feeding on their respective adult diets (Smith, 1987). Differences in mandibular morphology between the two African cichlid species *Labeotropheus fuelleborni* (a biter feeder) and *Metriaclima zebra* (a generalist feeder) can be detected as early as 7 days post-fertilisation (Albertson and Kocher, 2005). Our study indicates that in Arctic charr, shape differences in trophic elements can be detected during cartilage formation and well before first feeding (about 5 weeks before feeding if developing at 5°C). Although subtle, these differences were significant among all the studied progeny groups. The largest shape difference (in Mahalanobis distance) were seen between AC and the three morphs from Lake Thingvallavatn, while within the lake the distance between PL and the two benthic morphs

was larger than the distance between SB and LB. This morphological pattern is similar to what has previously been described for the adult fish in the lake (Snorrason *et al.*, 1994), where it was demonstrated that the four morphs of Arctic charr can be divided into two morphological variants: a benthic and a limnetic morphotype. In our analysis the two benthic morphs separate from the limnetic morph (PL) along the second canonical variate (CV2). These shape differences are subtle, reflecting narrowing of the lower jaw in the two benthic morphs, longer maxilla bones in PL and slightly shorter lower jaws in SB and LB. While the length of the lower jaws is a major predictor of diet type (and hence morph) in the adults (Snorrason *et al.*, 1994), juveniles of all morphs studied here have subterminal mouths. Clearly, the lengthening of the lower jaw is not completed at the developmental stages we are looking at. Paedomorphosis, the retention of juvenile features in adults, has been hypothesised as the evolutionary mechanism behind the evolution of the derived benthic morphs in Lake Thingvallavatn (Skulason *et al*., 1989a). The differences between morphs, especially in these trophic traits, become clearer as development progresses with the benthic morphs retaining their embryonic characteristics including the subterminal mouth in their adult phenotype, while the lower jaws of the limnetic morphs become more protruded (Skulason *et al*., 1989a). These differences can be further enhanced through plastic responses once the fish start feeding on their adult diets (Parsons *et al.*, 2010).

4.5.2 Transgressive phenotypes new phenotypes or post-zygotic reproductive barriers?

Despite the recent colonization of the lake, the two smallest and most abundant morphs PL and SB are phenotypically (Snorrason *et al.*, 1994) and genetically divergent (Volpe and Ferguson, 1996; Kapralova *et al.*, 2011) Yet the possibilities for interbreeding seem wide open as both spawn in the stony littoral zone of the lake and although the breeding period of PL is much more synchronized it completely overlaps with the breeding period of active SB spawners. Importantly, intermediate, adult sized phenotypes of pelagic and benthic morphs are rarely observed in the lake (Sigurður S. Snorrason personal observations). This fits with the hypothesis that the transgressive craniofacial morphology seen in F1 hybrid crosses between SB and PL is suboptimal. A number of explanations have been offered as to the existence of transgressive phenotypes in hybridizing populations (Rieseberg *et al.*, 1999), yet the consequences of this phenomenon will depend on the heritability, the genetic architecture of traits and their effects on fitness (Burke & Arnold 2001). For example, transgressive segregation can generate phenotypic-genotypic diversity necessary for the successful establishment of hybrid lineages in novel unexploited niches (Lewontin and Birch, 1966; Lexer *et al.*, 2003; Seehausen, 2004; Albertson and Kocher, 2005; Bell and Travis, 2005). However, if transgressive traits in hybrid populations are maladaptive compared to the parental strains this may contribute to the formation of a post-zygotic barrier (Rogers and Bernatchez, 2006).

As discussed above the estimated levels of gene flow between small benthic and pelagic charr are very low and although pre-zygotic mechanisms such as mating behaviour and small-scale temporal isolation can partly explain these observations, the indiscriminate nature of male salmonid mating behaviour (Foote *et al.*, 1997) suggests the presence of post-zygotic barriers (Kapralova *et al.*, 2011). Given the importance of the feeding apparatus for fitness, we can speculate that strong natural selection against extreme hybrid phenotypes could be an important barrier to gene flow between SB and PL charr at present. Here we observed deviant, transgressive phenotypes in head shape for the hybrid crosses

compared to the pure parental crosses, affecting the majority of the craniofacial elements under study. As these elements will give rise to important parts of the trophic apparatus this might have important consequences later in life, especially once feeding starts. Furthermore, the head of the hybrid offspring was significantly smaller indicating slower growth rate in the hybrids. This could affect important functional dimensions such as gape-width of the alevins as they start feeding. Although not conclusive these findings suggest that hybrids might have decreased fitness compared to offspringintra-morph, similar to what was observed in whitefish hybrids (Rogers and Bernatchez, 2006).

As populations evolving via genetic drift or stabilizing selection are likely to possess alleles with antagonistic effects, they are more likely to exhibit transgressive phenotypes than populations diverging via constant directional selection (Rieseberg *et al.*, 1999). Evidence of such effects was found in a study by (Albertson and Kocher, 2005), where the hybrid progeny of two closely related but morphologically divergent cichlid species exhibited intermediate phenotypes in the shape of the lower mandible, whereas the shape of the neurocranium exceeded the parental phenotypes. These two morphological traits appear to have different selection histories (Albertson *et al.*, 2003) and with the lower jaw evolved in response to directional selection, and selective forces such as stabilizing selection involved in the divergence of the cranium (Albertson and Kocher, 2005).

A previous study concluded that laboratory generated hybrid crosses between different morphs of Arctic charr from Lake Thingvallavatn strongly resembled the maternal pure breed crosses (Skulason *et al*., 1989a). Interestingly, in our study the two reciprocal hybrid crosses differed largely from both parental strains in craniofacial shape, but were rather similar to one another. These findings do not necessarily contradict each other for couple of reasons. Our study concentrated on earlier developmental stages, i.e. immediately after hatching and before the onset of feeding as opposed to after the feeding had started in the (Skulason *et al*., 1989a). Our data show that not all hybrids exhibit transgressive phenotypes, a proportion of hybrid offspring falls into the parental range of craniofacial shape variation (Figure 3). As hybrid mortality was not monitored here nor in the (Skulason *et al*., 1989a) study, it is possible that higher mortally of transgressive phenotypes preceeding their analysis could have led to biased phenotypic distributions. Also, (Skulason *et al*., 1989a) used length measurements to estimate shape differences and measurements were done using a lateral view of fish heads, while in our study we concentrated on the ventral view. The levels of transgressive segregation can vary between lateral and ventral aspects of the head as a consequence of the different genetic architecture and/or selective history of the underlying traits (Parsons, Son, *et al.*, 2011). While no maternal effects on craniofacial elements were observed here, it is possible that they will be revealed once the hybrid larvae start taking food.

4.5.3 A look ahead

The natural variation in trophic morphology seen in Icelandic Arctic charr both among and within systems presents an excellent model for studying the early steps of divergence and its developmental and genetic basis. At the developmental level, the nature of the molecular mechanisms underlying the diverse morphologies of Arctic charr need to be explored. Such studies are now underway in a number of vertebrate species. Recent studies have indicated that BMP4 and CAM1 are important for the shaping of the beaks of the Galapagos finches and the evolution of craniofacial diversity in vertebrates (Abzhanov *et*

al., 2004, 2006; Albertson *et al.*, 2005; Parsons and Albertson, 2009). Moreover expanded Wnt signalling through ontogeny has been associated with early ossification and retention of rounded craniofacial profile (paedomorphosis) in cichlids and zebrafish (Parsons *et al.*, 2014). These molecular pathways are likely to have parallels in Arctic charr, and it is possible that segregating variation in them is used during trophic evolution. Given the short evolutionary history of freshwater Arctic charr in Iceland and other Arctic areas it is likely that the molecular basis of this morphological differences arose mostly by differences in gene regulation as opposed to changes in protein coding sequences. In our recent studies (Ahi *et al.*, 2013) (Ahi *et al*. in submitted) several genes involved in matrix remodeling in bone formation showed consistent differences in expression during the development of embryos derived from benthic and limnetic morphs of Arctic charr, suggesting that these genes might be involved in the development of distinct Arctic charr morphologies. Moreover a recent study found interesting candidate miRNAs (Kapralova *et al*. in revision) for studying the involvement of non coding RNAs in developmental regulation.

In our view, the transgressive phenotypes in head morphology seen during the development of the hybrid offspring of SB and PL, begs further study of the genetics of trophic differences among sympatric morphs and of other small benthic charr populations. Given the recent evolutionary history of the Lake Thingvallavatn Arctic charr morphs, their co-existence in sympatry, and the possibility of interbreeding, the barriers to gene flow between SB and PL appear to be strong. Although the data suggest some post-zygotic barriers to gene flow might be in place, tests for developmental mortality of hybrid crosses, their fitness and fertility are needed. Importantly, these traits should be measured also in the F2 and in backcrosses to both parental strains thus providing the opportunity to map the fitness related loci and also study the segregation of the transgressive morphological traits. These studies are not trivial because of the slow development of Arctic charr and the difficulty of obtaining sufficient numbers of fertile offspring. Yet this is definitely an important future research goal.

4.6 Conclusions

We use the recently evolved and highly polymorphic Arctic charr species to study the evolution and development of elements of the feeding apparatus by using landmark based geometric morphometrics and multivariate analyses of shape. Subtle differences among three sympatric morphs of Arctic charr from Lake Thingvallavatn, Iceland were detected early in development during cartilage formation and growth. Furthermore we investigated the effect of hybridisation on the craniofacial morphology of Arctic charr by creating reciprocal crosses between PL and SB. Interestingly, the hybrid crosses exhibit transgressive craniofacial phenotypes compared to the pure morph breeds. While the shape of the feeding apparatus of the two reciprocal hybrid crosses differed significantly from both pure crosses, no significant differences in head shape were detected between the two reciprocal crosses, suggesting genetic effects greatly outweigh maternal effects.

4.7 References

Abzhanov A, Kuo WP, Hartmann C, Grant BR, Grant PR, Tabin CJ (2006). The calmodulin pathway and evolution of elongated beak morphology in Darwin's finches. *Nature* **442**: 563–567.

Abzhanov A, Protas M, Grant BR, Grant PR, Tabin CJ (2004). Bmp4 and morphological variation of beaks in Darwin's finches. *Science* **305**: 1462–1465.

Ahi EP, Guðbrandsson J, Kapralova KH, Franzdóttir SR, Snorrason SS, Maier VH, *et al.* (2013). Validation of reference genes for expression studies during craniofacial development in arctic charr. *PLoS One* **8**: e66389.

Albertson RC, Kocher TD (2005). Genetic architectute sets limits on transgressive segregation in hybrid cichlid fishes. *Evolution (N Y)* **59**: 686–690.

Albertson RC, Streelman JT, Kocher TD (2003). Directional selection has shaped the oral jaws of Lake Malawi cichlid fishes. *Proc Natl Acad Sci U S A* **100**: 5252–7.

Albertson RC, Streelman JT, Kocher TD, Yelick PC (2005). Integration and evolution of the cichlid mandible: the molecular basis of alternate feeding strategies. *Proc Natl Acad Sci U S A* **102**: 16287–92.

Amadon D (1950). The Hawaiian honeycreepers (Aves, Drepaniidae). *Bull AMNH* **95**.

Badyaev A, Whittingham L, Hill G (2001). The evolution of sexual size dimorphism in the house finch. III. Developmental basis. *Evolution (N Y)* **55**: 176–189.

Barton NH (2001). The role of hybridization in evolution. *Mol Ecol* **10**: 551–568.

Bell MA, Travis MP (2005). Hybridization, transgressive segregation, genetic covariation, and adaptive radiation. *Trends Ecol Evol* **20**: 358–61.

Bhullar B-AS, Marugán-Lobón J, Racimo F, Bever GS, Rowe TB, Norell MA, *et al.* (2012). Birds have paedomorphic dinosaur skulls. *Nature* **487**: 223–6.

Bookstein FL (1997). *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge University Press.

Bowman R (1961). *Morphological differentiation and adaptation in the Galápagos finches.* University of California Press: Berkeley.

Burke JM, Arnold ML (2001). Genetics and the fitness of hybrids. *Annu Rev Genet* **35**: 31–52.

Cooper WJ, Parsons K, McIntyre A, Kern B, McGee-Moore A, Albertson RC (2010). Bentho-pelagic divergence of cichlid feeding architecture was prodigious and consistent during multiple adaptive radiations within African rift-lakes. (S Humphries, Ed.). *PLoS One* **5**: e9551.

Darrin Hulsey C (2006). Function of a key morphological innovation: fusion of the cichlid pharyngeal jaw. *Proc Biol Sci* **273**: 669–75.

Dixon P (2003). VEGAN, a package of R functions for community ecology. *J Veg Sci* **14**: 927–930.

Dowling TE, DeMarais BD (1993). Evolutionary significance of introgressive hybridization in cyprinid fishes. *Nature* **362**: 444–446.

Eiriksson GM, Skulason S, Snorrason SS (1999). Heterochrony in skeletal development and body size in progeny of two morphs of Arctic charr from Thingvallavatn, Iceland. *J Fish Biol* **55**: 175–185.

Felsenstein J (1981). Skepticism towards Santa Rosalia, or why are there so few kinds of animals? *Evolution (N Y)* **35**: 124–138.

Foote CJ, Brown GS, Wood CC (1997). Spawning success of males using alternative mating tactics in sockeye salmon, (*Oncorhynchus nerka*). *Can J Fish Aquat Sci* **54**: 1785–1795.

Frandsen F, Malmquist HJ, Snorrason SS (1989). Ecological parasitology of polymorphic Arctic charr, Salvelinus alpinus (L.), in Thingvallavatn, Iceland. *J Fish Biol* **34**: 281–297.

Freed LA, Conant S, Fleischer RC (1987). Evolutionary ecology and radiation of Hawaiian passerine birds. *Trends Ecol Evol* **2**: 196–203.

Fryer G (1972). *The cichlid fishes of the great lakes of Africa: their biology and evolution,*. Oliver and Boyd: Edinburgh.

Gavrilets S, Losos JB (2009). Adaptive radiation: contrasting theory with data. *Science* **323**: 732–7.

Gíslason D, M Ferguson M, Skúlason S, S Snorrason S (1999). Rapid and coupled phenotypic and genetic divergence in Icelandic Arctic char (Salvelinus alpinus). *Can J Fish Aquat Sci* **56**: 2229–2234.

Gonzalez PN, Hallgrímsson B, Oyhenart EE (2011). Developmental plasticity in covariance structure of the skull: effects of prenatal stress. *J Anat* **218**: 243–57.

Goodall C (1991). Procrustes methods in the statistical-analysis of shape. *J R Stat Soc Ser B* **53**: 285–339.

Gorodilov YN (1996). Description of the early ontogeny of the Atlantic salmon, Salmo salar, with a novel system of interval (state) identification. *Environ Biol Fishes* **47**: 109–127.

Grant PR (1999). *Ecology and evolution of Darwin's finches*. Princeton University Press: Princeton (N.J.).

Jamniczky HA, Harper EE, Garner R, Cresko WA, Wainwright PC, Hallgrímsson B, *et al.* (2014). Association between integration structure and functional evolution in the opercular four-bar apparatus of the threespine stickleback, Gasterosteus aculeatus (Pisces: Gasterosteidae). *Biol J Linn Soc* **111**: 375–390.

Jonsson B, Skúlason S, Snorrason SS, Sandlund OT, Malmquist HJ, Jónasson PM, *et al.* (1988). Life History Variation of Polymorphic Arctic Charr (Salvelinus alpinus) in Thingvallavatn, Iceland. *Can J Fish Aquat Sci* **45**: 1537–1547.

Kapralova KH, Gudbrandsson J, Reynisdottir S, Santos CB, Baltanás VC, Maier VH, *et al.* (2013). Differentiation at the MHCIIα and Cath2 Loci in Sympatric Salvelinus

alpinus Resource Morphs in Lake Thingvallavatn. (GH Yue, Ed.). *PLoS One* **8**: e69402.

Kapralova KH, Morrissey MB, Kristjánsson BK, Olafsdóttir GÁ, Snorrason SS, Ferguson MM (2011). Evolution of adaptive diversity and genetic connectivity in Arctic charr (Salvelinus alpinus) in Iceland. *Heredity (Edinb)* **106**: 472–487.

Kimmel CB, Hohenlohe PA, Ullmann B, Currey M, Cresko WA (2012). Developmental dissociation in morphological evolution of the stickleback opercle. *Evol Dev* **14**: 326–37.

Kimmel CB, Ullmann B, Walker C, Wilson C, Currey M, Phillips PC, *et al.* (2005). Evolution and development of facial bone morphology in threespine sticklebacks. *Proc Natl Acad Sci U S A* **102**: 5791–6.

Kingsolver JG, Hoekstra HE, Hoekstra JM, Berrigan D, Vignieri SN, Hill CE, *et al.* (2001). The strength of phenotypic selection in natural populations. *Am Nat* **157**: 245–61.

Klingenberg CP (2010). Evolution and development of shape: integrating quantitative approaches. *Nat Rev Genet* **11**: 623–35.

Klingenberg CP (2011). MorphoJ: an integrated software package for geometric morphometrics. *Mol Ecol Resour* **11**: 353–7.

Klingenberg C, Barluenga M, Meyer A (2002). Shape analysis of symmetric structures: Quantifying variation among individuals and asymmetry. *Evolution (N Y)* **56**: 1909–1920.

Kruuk LEB, Slate J, Wilson AJ (2008). New Answers for Old Questions: The Evolutionary Quantitative Genetics of Wild Animal Populations. *Annu Rev Ecol Evol Syst* **39**: 525–548.

Lack D (1945). *The Galapagos finches (Geospizinae) a study in variation, by David Lack.* California academy of sciences: San Francisco.

Landry L, Bernatchez L (2010). Role of epibenthic resource opportunities in the parallel evolution of lake whitefish species pairs (Coregonus sp.). *J Evol Biol* **23**: 2602–13.

Lewontin RC, Birch LC (1966). Hybridization as a Source of Variation for Adaptation to New Environments. *Evolution (N Y)* **20**: 315.

Lexer C, Welch ME, Durphy JL, Rieseberg LH (2003). Natural selection for salt tolerance quantitative trait loci (QTLs) in wild sunflower hybrids: Implications for the origin of Helianthus paradoxus, a diploid hybrid species. *Mol Ecol* **12**: 1225–1235.

Lynch M (1991). The Genetic Interpretation of Inbreeding Depression and Outbreeding Depression. *Evolution (N Y)* **45**: 622–629.

Mallarino R, Campàs O, Fritz JA, Burns KJ, Weeks OG, Brenner MP, *et al.* (2012). Closely related bird species demonstrate flexibility between beak morphology and underlying developmental programs. *Proc Natl Acad Sci U S A* **109**: 16222–7.

Mallarino R, Grant PR, Grant BR, Herrel A, Kuo WP, Abzhanov A (2011). Two developmental modules establish 3D beak-shape variation in Darwin's finches. *Proc Natl Acad Sci U S A* **108**: 4057–62.

Malmquist HJ, Snorrason SS, Skulason S, Jonsson B, Sandlund OT, Jonasson PM (1992). Diet differentiation in polymorphic Arctic charr in Thingvallavatn , Iceland. *J Anim Ecol* **61**: 21–35.

Marchinko KB (2009). Predation's role in repeated phenotypic and genetic divergence of armor in threespine stickleback. *Evolution* **63**: 127–38.

Matute DR, Butler IA, Turissini DA, Coyne JA (2010). A test of the snowball theory for the rate of evolution of hybrid incompatibilities. *Science* **329**: 1518–21.

Mousseau TA, Roff DA (1987). Natural selection and the heritability of fitness components. *Heredity (Edinb)* **59 (Pt 2)**: 181–97.

Nosil P, Harmon LJ, Seehausen O (2009). Ecological explanations for (incomplete) speciation. *Trends Ecol Evol* **24**: 145–56.

Palsson A, Gibson G (2004). Association between nucleotide variation in Egfr and wing shape in Drosophila melanogaster. *Genetics* **167**: 1187–98.

Parsons KJ, Albertson RC (2009). Roles for Bmp4 and CaM1 in shaping the jaw: evo-devo and beyond. *Annu Rev Genet* **43**: 369–88.

Parsons TE, Kristensen E, Hornung L, Diewert VM, Boyd SK, German RZ, *et al.* (2008). Phenotypic variability and craniofacial dysmorphology: increased shape variance in a mouse model for cleft lip. *J Anat* **212**: 135–43.

Parsons KJ, Sheets HD, Skúlason S, Ferguson MM (2011). Phenotypic plasticity, heterochrony and ontogenetic repatterning during juvenile development of divergent Arctic charr (Salvelinus alpinus). *J Evol Biol* **24**: 1640–52.

Parsons KJ, Skúlason S, Ferguson M (2010). Morphological variation over ontogeny and environments in resource polymorphic arctic charr (Salvelinus alpinus). *Evol Dev* **12**: 246–57.

Parsons KJ, Son YH, Craig Albertson R (2011). Hybridization Promotes Evolvability in African Cichlids: Connections Between Transgressive Segregation and Phenotypic Integration. *Evol Biol* **38**: 306–315.

Parsons KJ, Trent Taylor a., Powder KE, Albertson RC (2014). Wnt signalling underlies the evolution of new phenotypes and craniofacial variability in Lake Malawi cichlids. *Nat Commun* **5**: 1–11.

Perry G, Audet C, Laplatte B, Bernatchez L (2004). Shifting patterns in genetic control at the embryo-alevin boundary in brook charr. *Evolution (N Y)* **58**: 2002–2012.

Rieseberg LH, Archer MA, Wayne RK (1999). Transgressive segregation, adaptation and speciation. *Heredity (Edinb)* **83**: 363–372.

Roberts RB, Hu Y, Albertson RC, Kocher TD (2011a). Craniofacial divergence and ongoing adaptation via the hedgehog pathway. *Proc Natl Acad Sci U S A* **108**: 13194–9.

Roberts RB, Hu Y, Albertson RC, Kocher TD (2011b). Craniofacial divergence and ongoing adaptation via the hedgehog pathway. *Proc Natl Acad Sci U S A* **108**: 13194–9.

Rogers SM, Bernatchez L (2006). The genetic basis of intrinsic and extrinsic post-zygotic reproductive isolation jointly promoting speciation in the lake whitefish species complex (Coregonus clupeaformis). *J Evol Biol* **19**: 1979–94.

Rohlf FJ (2006). "tpsDig, version 2.10." *http://life.bio.sunysb.edu/morph/index.html*.

Rohlf FJ, Slice D (1990). Extensions of the Procrustes Method for the Optimal Superimposition of Landmarks. *Syst Zool* **39**: 40.

Salzburger W, Baric S, Sturmbauer C (2002). Speciation via introgressive hybridization in East African cichlids? *Mol Ecol* **11**: 619–625.

Sandlund OT, Gunnarson K, Jonasson PM, Jonsson B, Lindem T, Magnusson KP, *et al.* (1992). The Arctic charr Salvelinus alpinus in Thingvallavatn. *Oikos* **64**: 305–351.

Schluter D (2000). *The Ecology of Adaptive Radiation*.

Schluter D, Rambaut A (1996). Ecological Speciation in Postglacial Fishes. *Philos Trans R Soc B Biol Sci* **351**: 807–814.

Seehausen O (2004). Hybridization and adaptive radiation. *Trends Ecol Evol* **19**: 198–207.

Seehausen O, Butlin RK, Keller I, Wagner CE, Boughman JW, Hohenlohe PA, *et al.* (2014). Genomics and the origin of species. *Nat Rev Genet* **15**: 176–92.

Sidor C (2001). Simplification as a trend in synapsid cranial evolution. *Evolution (N Y)* **55**: 1419–1442.

Simpson GG (1965). *The Major Features of Evolution*. Columbia University Press.

Skulason S, Noakes DL. G, Snorrason SS (1989a). Ontogeny of trophic morphology in four sympatric morphs of arctic charr Salvelinus alpinus in Thingvallavatn, Iceland*. *Biol J Linn Soc* **38**: 281–301.

Skulason S, Smith TB (1995). Resource polymorphisms in vertebrates. *Trends Ecol Evol* **10**: 366–370.

Skúlason S, Snorrason SS, Noakes DLG, Ferguson MM, Malmquist HJ (1989b). Segregation in spawning and early life history among polymorphic Arctic charr, Salvelinus alpinus, in Thingvallavatn, Iceland. *J Fish Biol* **35**: 225–232.

Skúlason S, Snorrason SS, Ota D, Noakes DLG (1993). Genetically based differences in foraging behaviour among sympatric morphs of arctic charr (Pisces: Salmonidae). *Anim Behav* **45**: 1179–1192.

Slice DE (2006). *Modern Morphometrics in Physical Anthropology (Google eBook)*. Springer.

Smith TB (1987). Bill size polymorphism and intraspecific niche utilization in an African finch. *Nature* **329**: 717–719.

Smith TB, Skulason S (1996). Evolutionary Significance of Resource Polymorphisms in Fishes, Amphibians, and Birds. *Annu Rev Ecol Syst* **27**: 111–133.

Snorrason SS, Skúlason S (2004). Adaptive speciation in northern fresh water fishes — patterns and processes. In: Dieckmann U, Metz H, Doebeli M, Tautz D (eds) *Adaptive speciation*, Cambridge University Press, Cambridge,, pp 210–228.

Snorrason SS, Skulason S, Jonsson B, Malmquist HJ, Jonasson PM, Sandlund ODD., *et al.* (1994). Trophic specialization in Arctic charr Salvelinus alpinus (Pisces:Salmonidae): Morphological divergence and ontogenetic niche shifts. *Biol J Linn Soc* **52**: 1–18.

Snorrason SS, Skulason S, Sandlund OT, Malmquist HJ, Jonsson B, Jonasson PM (1989). Shape polymorphism in sympatric Arctic charr, Salvelinus alpinus in Thingvallavath, Iceland. *Physiol Ecol Japan* **1**: 393–404.

Svavarsson E (2007). Árangur í kynbótum á bleikju og næstu skref [reference in icelandic]. *Fræðaþing landbúnaðarins (conference proceedings)* **4**: 121–125.

Swindell WR, Bouzat JL (2006). Gene Flow and Adaptive Potential in Drosophila melanogaster. *Conserv Genet* **7**: 79–89.

Tickle C (2011). *Principles of Development*. Oxford University Press.

Travisano M, Shaw RG (2013). Lost in the map. *Evolution* **67**: 305–14.

Volpe JP, Ferguson MM (1996). Molecular genetic examination of the polymorphic Arctic charr Salvelinus alpinus of Thingvallavatn, Iceland. *Mol Ecol* **5**: 763–72.

Walker MB, Kimmel CB (2007). A two-color acid-free cartilage and bone stain for zebrafish larvae. *Biotech Histochem* **82**: 23–8.

Woods PJ, Skulason S, Snorrason SS, Kristjansson BK, Malmquist HJ, Quinn TP (2012). Intraspecific diversity in Arctic charr, Salvelinus alpinus, in Iceland: I. Detection using mixture models. *Evol Ecol Res* **14**: 973–992.

Young NM, Chong HJ, Hu D, Hallgrímsson B, Marcucio RS (2010). Quantitative analyses link modulation of sonic hedgehog signaling to continuous variation in facial growth and shape. *Development* **137**: 3405–9.

Zelditch ML, Swiderski DL, Sheets HD (2012). *Geometric Morphometrics for Biologists: A Primer (Google eBook)*. Academic Press.

4.8 Appendix

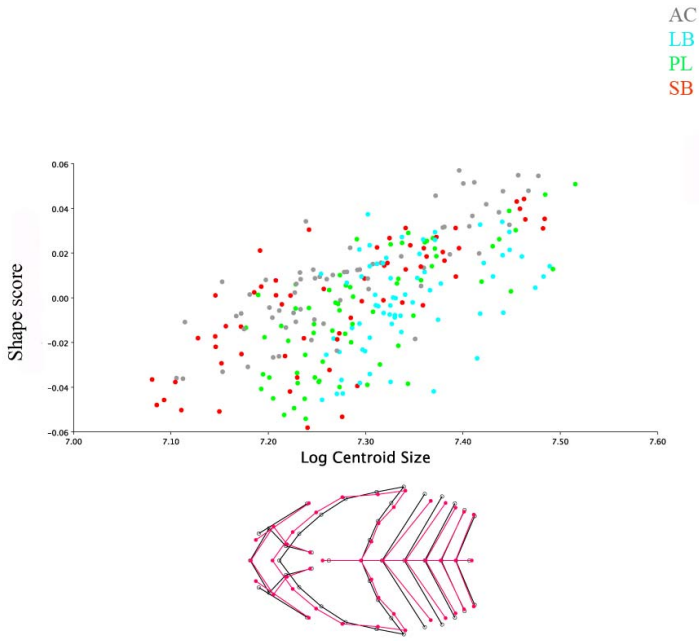

Figure S 4.1 - Multivariate regression of shape (symmetric component) on size (LCS) for the offspring of the four pure crosses. Allometry accounts for 8% of the total shape variance. Wireframe depict shape changes associated with LCS, the shape associated with low LCS values are shown in black and the shape associated with high LCS values in red.

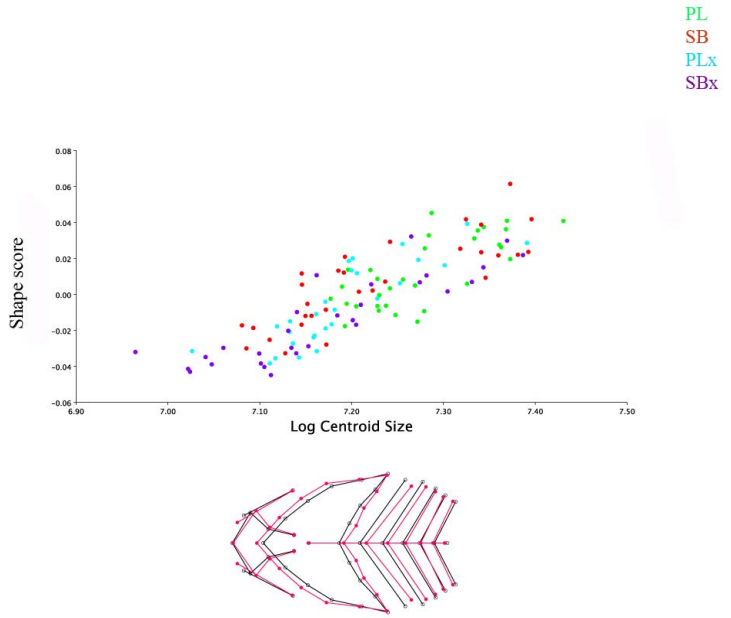

Figure S 4.2 – Multivariate regression of shape (symmetric component) on size (LCS) for the offspring of PL and SB pure crosses and their reciprocal crosses PLx and SBx.

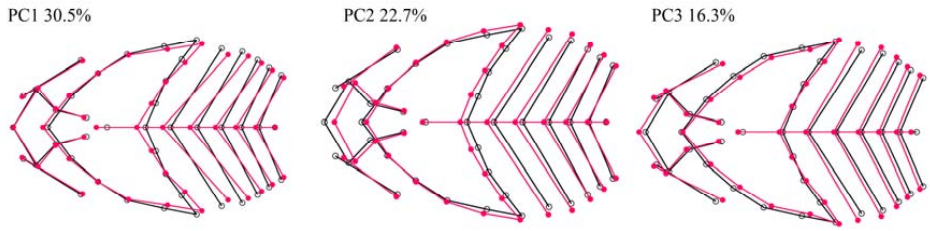

Figure S 4.3 - Shape changes associated with the three major PCs from a PCA of shape data (size removed) of pure morph crosses: wireframes depict shape changes associated with the three major Principal Components (PCs). The scale factor is in units of Procrustes distance and it's set to 0.05. Wireframes depict shape changes associated with the two Principal Components shown in each graph. In the wire frames the extreme negative value is shown in black and the extreme positive values in red.

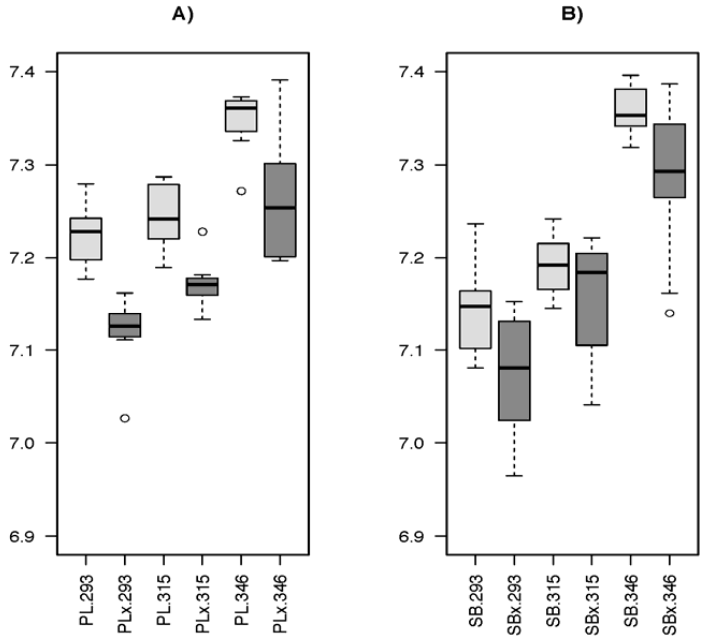

Figure S 4.4 - Boxplots of pure and hybrid cross head size (measured by the Log transformed Centroid Size) at three developmental time points. A) Head size of PL (Planctivorous charr) and PLx (PL female x SB male), B) Head size of SB (Small benthic charr) and SBx (SB female x PL male.)

Table 4-4 - MANCOVA of Shape (represented by the Procrustes coordinates) with morph (AC, LB, PL and SB) as a variate and LCS (Log transformed centroid size) as a covariate.

	Df	SS	MS	F.Model	R2	Pr(>F)
LCS	1	0.078	0.077998	19.4278	0.06794	0.001
Morph	3	0.05133	0.017111	4.2621	0.04472	0.001
LCS:Morph	3	0.01494	0.004979	1.2402	0.01301	0.216
Residuals	250	1.00369	0.004015	0.87433		
Total	257	1.14796	1			

Table 4-5 - MANCOVA of Shape (represented by the Procrustes coordinates) with morph (PL, SB, PLx and SBx) as a variate and LCS (Log transformed centroid size) as a covariate.

	Df	SS	MS	F.Model	R2	Pr(>F)
LCS	1	0.0543	0.054299	14.21	0.094	0.001
Morph	2	0.08419	0.04209	11.01	0.146	0.001
LCS:Morph	2	0.01048	0.005241	1.37	0.018	0.146
Residuals	112	0.01048	0.003822	0.74		
Total	117	0.57704	1			

*Table 4-6 - Effect of Time, Morph and Time x Morph on each of the 6 major Principal Components (PC) for LB, PL, SB and AC tested with a generalized linear model. Significance codes '***' 0.001, '**' 0.01, '*' 0.05*

	%	Time	Morph	Time x Morph
PC1	30.5	*	***	NS
PC2	22.7	***	***	NS
PC3	16.3	NS	*	*
PC4	6.4	**	***	NS
PC5	3.5	*	***	*
PC6	3	NS	NS	NS

*Table 4-7 - Effect of Time, Morph and Time x Morph on each of the 6 major Principal components two pure breeds (PL and SB) and two hybrid crosses PLx (PL female x SB male) and SBx (SB female and PL male) tested with a generalized linear model. Significance codes '***' 0.001, '**' 0.01, '*' 0.05*

	%	Time	Morph	Time x Morph
PC1	44.3	*	***	*
PC2	19.5	***	***	NS
PC3	8.7	NS	NS	NS
PC4	4.6	NS	NS	*
PC5	4.2	NS	NS	NS
PC6	2.8	NS	NS	NS

Table 4-8- Differences in head size (LCS) between hybrid and pure morph crosses. ANOVA of the Log transformed centroid size (LCS) for four varieties of Arctic charr AC, LB, PL and SB and all stages followed by post hoc tests for Morph.

	Df	SS	MS	F	Pr (>F)
Morph	3	0.1827	0.0609	25.448	1.77E-12
Time	2	0.6551	0.3275	136.877	2.00E-16
Morph x Time	6	0.0393	0.0066	2.738	0.0164
Residuals	106	0.2537	0.0024		

	diff	lwr	upr	p-adj
PLx-PL	-0.0918	-0.1266	-0.0569	0.0000
SB-PL	-0.0447	-0.0786	-0.0108	0.0045
SBx-PL	-0.1020	-0.1362	-0.0678	0.0000
SB-PLx	0.0470	0.0117	0.0824	0.0041
SBx-PLx	-0.0103	-0.0459	0.0254	0.8762
SBx-SB	-0.0573	-0.0920	-0.0226	0.0002

Paper 4

5 Paper IV

Modularity and integration in craniofacial elements during development of sympatric Arctic charr morphs

Kalina H. Kapralova, Arnar Palsson, Bjarni K. Kristjanson, Zophonías O. Jónsson, Sigurður S. Snorrason

Authors' contribution: Conceived and designed the experiments: KHK SSS ZOJ AP BKK. Performed the experiments: KHK AP BKK ZOJ SSS. Analysed the data: KHK AP BKK ZOJ SSS. Contributed reagents/materials/analysis tools: KHK BKK SSS AP ZOJ. Wrote the paper: KHK SSS AP ZOJ BKK

5.1 Abstract

This paper analyzes the level of integration and modularity in craniofacial traits in morphs and hybrids. The results indicated that during early post-hatching stages the craniofacial skeleton is modular and that this modularity appears to reflect the developmental origins of the elements constituting it. We compared the craniofacial integration in four groups of Arctic charr (LB, SB, PL from Thingvallavatn and AC) and saw indications that these groups may differ in the level of their craniofacial modularity. The hybrid progeny of two contrasting morphs appeared to have different patterns of integration of their craniofacial skeleton compared to the pure crosses of the parental morphs. The hybrid crosses also exhibited different patterns of allometry compared to the pure morphs. These results taken together with the transgressive phenotypes in head morphology the hybrids exhibit may indicate developmental instabilities during craniofacial morphogenesis in the hybrids.

5.2 Introduction

A central question in evolutionary biology is how the spectacular ecological and phenotypic biological diversity is generated and maintained. Adaptive radiation, a key process in the evolution of ecological diversity, is regarded as the most important biological process leading to the evolution of ecological differences and ultimately to the emergence of new species (Schluter, 2000; Gavrilets and Losos, 2009). This process is dependent on historical contingencies and influenced by ecological, genetic and developmental factors (Gavrilets and Losos, 2009). Complex phenotypes arise through development and their evolvability is influenced by the level of integration and modularity of the morphological units constituting them (Klingenberg, 2005). Morphological integration has a developmental, functional and often adaptive basis (Klingenberg, 2005) and refers to the developmental and functional coordinated variation of traits (Hallgrimsson *et al.*, 2009). Modularity on the other hand refers to the organization of organisms into distinct units or modules and is characterized by a strong integration within each module and a relative independence among modules (Klingenberg, 2005). Strong integration among developmental or anatomical units will have a constraining effect on morphological evolution, while modularity will enhance evolutionary flexibility (Drake and Klingenberg, 2010). Modularity will manifest itself as the relative covariation of traits within an integrated functional unit and varies in strength rather than being an all-or-nothing phenomenon (Klingenberg, 2003).

The vertebrate head is one of the most complex anatomical units and its morphological integration has been the focus of many studies (see references in (Jamniczky and Hallgrimsson, 2011)). A small proportion of these studies have addressed questions on the relationships between variation in the shapes, proportions and placement of bones and other tissues (e.g. (Richtsmeier *et al.*, 2006; Jamniczky and Hallgrimsson, 2011; Zelditch *et al.*, 2012; Tsuboi *et al.*, 2014). However the skull and its different bone elements have received most attention when it comes to studying morphological integration of the head (Drake and Klingenberg, 2010; Parsons *et al.*, 2012a; Jamniczky *et al.*, 2014). The feeding apparatus and especially the mandible has been a major inspiration for the development and implementation of various statistical methods for studies of modularity (Klingenberg *et al.*, 2003; Márquez, 2008; Parsons *et al.*, 2012b).

In vertebrates, the majority of the head skeleton originates from cranial neural crest cells, which appear to migrate in a segmented manner according to their rhombomeric origin (Lumsden *et al.*, 1991). Each segment will give rise to a differentiated arch, which is subdivided, into individual dorsal and ventral structures (Schilling and Kimmel, 1997). In zebrafish the different segments show similar patterns of pre-cartilage condensation and chondrification (cartilage differentiation), however a certain degree of variation in size and shape can be detected between the anterior (Meckel's cartilage and the hyoid arch) and the posterior arches. Differences in the developmental timing of their formation are considered to be the underlying cause (Schilling and Kimmel, 1997).

The four morphs of Arctic charr (*Salvelinus alpinus*) from Lake Thingvallavatn, Iceland offer an excellent opportunity to study morphological integration in the context of a recent ecological diversification. These morphs (belonging to two morphotypes: limnetic and benthic) exhibit striking differentiation in morphology of the trophic apparatus (Snorrason *et al.*, 1989), life history characteristics and ecology, as reflected in different habitat use,

diet and endoparasite fauna (Jonsson *et al.*, 1988; Frandsen *et al.*, 1989; Malmquist *et al.*, 1992; Sandlund *et al.*, 1992) The two limnetic morphs, a planktivorous (PL) and a piscivorous (PI) morph, have pointed snouts and evenly protruding jaws, while the two benthic morphs, a small (SB) and a large benthivorous (LB) morph have blunt snouts, short lower jaws and relatively large pectoral fins (Snorrason *et al.*, 1989). This phenotypic diversity covaries with ecological features of each morph's niche thus providing suggestive evidence of the adaptive nature of this variation. Common-garden experiments have demonstrated that the development of Arctic charr trophic morphologies and behaviors are the result of both genetic differences (Skulason *et al.*, 1989; Skúlason *et al.*, 1993; Eiriksson *et al.*, 1999) and plastic responses to different environments (Parsons *et al.*, 2010, 2011). Although the Thingvallavatn morphs show significant but subtle neutral genetic differentiation, the levels of gene flow between morphs are restricted (Kapralova *et al.*, 2011). A more pronounced genetic differentiation among the morphs in the lake was detected in a study on immune system genes (Kapralova *et al.*, 2013). Subtle but significant differences are detected during the development of Arctic charr head morphologies, mainly in shape of the craniofacial elements during cartilage formation and early growth, long before ossification of the craniofacial elements is completed. Moreover, hybrids between PL and SB showed highly significant differences in craniofacial morphology when compared to the pure parental crosses. These differences are so distinct that hybrids can be qualified as “transgressive” in craniofacial morphology (i.e appearance of extreme or novel phenotypes relative to the parental populations) (Paper III in this thesis).

Here we use landmark based geometric morphometrics to address developmental and evolutionary questions regarding Arctic charr craniofacial modularity. We base our hypotheses on the segmented developmental origin of craniofacial bones (Lumsden *et al.*, 1991) and we predict that during early post-hatching stages the craniofacial elements will exhibit low levels of integration and different modules will be defined by the developmental origins of each bone group. We use the progeny of four Arctic charr varieties that have presumably undergone different selection regimes: three natural morphs from Lake Thingvallavatn (PL, LB and SB) and an aquaculture strain (AC) to address evolutionarily questions on modularity. Because of the genetic divergence between these varieties we predict that patterns and/or levels of integration of craniofacial elements will vary among morphs. Finally, we investigated the level of post-hatching craniofacial integration in hybrid crosses between SB and PL and compared it to the pure parental crosses. As noted above these hybrids have been shown to have extreme (or transgressive) phenotypes in craniofacial morphology (Paper III in this thesis). In light of this we expect that the underlying mechanisms for the observed transgressive phenotypes might have also influenced the degree or even the patterns of integration of the craniofacial elements in hybrids.

5.3 Material and methods

5.3.1 Sampling

For this study we used post hatching embryos from pure crosses of four Arctic charr varieties, three morphs (LB, SB and PL) from lake Thingvallavatn and an aquaculture

strain from the Hólar aquaculture station (AC), as well as two reciprocal hybrid crosses of contrasting Thingvallavatn morphs (SB and PL). Fish from the planktivorous (PL), small and large benthic morphs (SB and LB) were caught in lake Thingvallavatn using gill-nets. Fishing permissions were obtained from the Thingvellir National Park Commission and the land-owner of Mjóanes farm. The AC crosses were made with parents from the Hólar breeding program (Svavarsson, 2007). Fish were killed by a sharp blow to the head and for each experimental group eggs from several females were pooled and fertilized using milt from several males from the same group. Hybrid crosses were also made between SB and PL. Eggs were reared at approximately 5°C in a hatching tray (EWOS, Norway) under constant water flow and in complete darkness at the Hólar University College experimental facilities in Verið, Sauðárkrókur. The rearing and collection of the embryos was performed according to Icelandic regulations (licence granted to Hólar University College aquaculture and experimental facilities in Verið, Sauðárkrókur). Exact water temperature was recorded twice daily to estimate the relative age of the embryos using tau-somite τ_s units defined as the time it takes for one somite pair to form at a given temperature (Gorodilov, 1996). Embryos were collected throughout development (Table 5.1) and fixed in 4% PFA.

Table 5-1 Numbers of embryos sampled for the study of developmental modularity in Arctic charr. Embryos were sampled at six developmental stages,just after hatching (293, 305, 315, 336, 346, 370 τs)). AC=Aquaculture charr from the Hólar breeding stock, PL=planktivorouscharr, LB and SB =large and small benthic charr, respectively, and HY = hybrid crosses between SB and PL.

Stage	Batch	AC	LB	PL	SB	HY
293	1	12	10	12	12	18
305	2	12	14	15	10	-
315	1	15	12	8	9	18
336	2	6	13	10	11	-
346	1	10	9	11	10	20
370	2	10	8	9	10	-
Total		65	66	65	62	56

5.3.2 Staining and photographing

A total of 314 individuals (258 pure morph and 56 hybrid) were stained for cartilage (alcian blue) and bone (alizarin red) using a modified protocol from Walker and Kimmel (2007). Individuals were placed in a petri dish containing 50 ml of 1% agarose gel and immobilised with dissecting needles to insure the correct positioning of the embryo. The head of each individual was photographed ventrally facing left using Leica (MZ10) stereomicroscope and the same magnification (2.0x) was used for each photo. Samples all studied groups (the pure morphs and the hybrid crosses) were stained and photographed simultaneously.

5.3.3 Geometric morphometrics

We selected 35 landmarks (13 pairs of bilateral and 9 mid-line) to describe the craniofacial shape of developing Arctic charr (Figure 5.1) and digitized them using tps.DIG2 (Rohlf, 2006). The landmarks were positioned in the lower jaw, the hyoid arch and the pharyngeal arches (Figure 1). The shape information for each specimen was extracted using a Generalized Procrustes analysis (GPA) in MorphoJ (Klingenberg, 2011), where after accounting for scale, position and orientation, all specimen are superimposed on a common coordinate system (Rohlf and Slice, 1990; Goodall, 1991). Only the symmetric component of shape variation (Klingenberg *et al.*, 2002) was used for subsequent statistical analysis. The Centroid Size (defined as the square root of the sum of the squared distances of all landmarks from their centroid) for each specimen was retained after the Procrustes fit and the Log transformed Centroid size (LCS) was used as a measure of individual size in subsequent analysis.

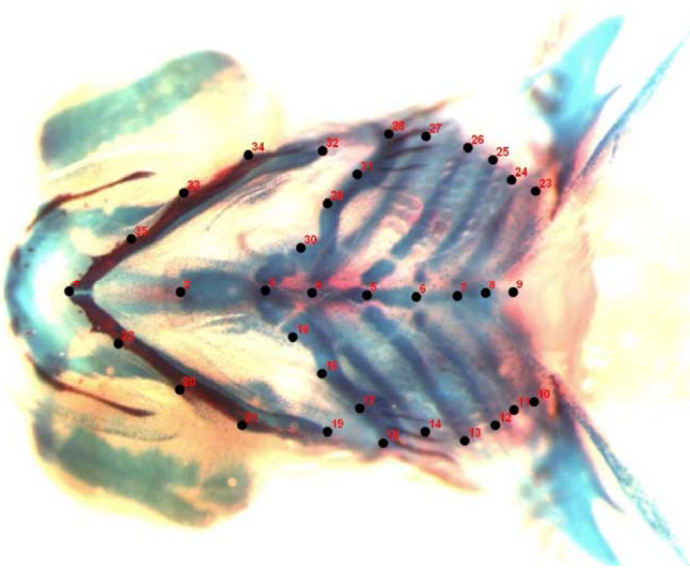

Figure 5.1 The 35 landmarks (13 pairs of bilateral and 9 mid-line landmarks) used in this study. Landmarks were selected to describe major craniofacial elements such as Meckel's cartilage (MC), the hyoid arch (HA), the ceratobranchial arches (CB 1-5), the basihyal (BH) and basibranchial cartilage (BB). The specimen on the figure is at 346 τ_s.

Each landmark for every individual was digitised twice by the same observer and the results from the repeated measurements were averaged in the final data-set. Measurement error was accessed by performing Principal Component Analysis (PCA) on the two landmarking sessions followed by discriminant function analysis (DFA). The Principal component analysis did not show any separation between landmarking sessions (Figure S 5.1) nor could observations be classified as belonging to landmarking session 1 or 2 by DFA (p=0.593).

5.3.4 Evaluating modularity hypothesis for the developing Arctic charr

We tested four hypotheses of the developmental modularity of Arctic charr by comparing patterns of covariation among landmark positions following (Klingenberg, 2009a). As allometry has an overall integrative effect and can conceal modularity (Klingenberg, 2009a), all analyses were performed both with and without accounting for allometry. To account for allometry we regressed shape on size (measured by LCS) and used the regression residuals for subsequent analyses. We used the Modulatiry function in MorphoJ to test four hypotheses for landmark partitions (Figure 5.2) H1: 2 partitions (anterior vs posterior), H2: 3a partitions: 1) Meckel's cartilage, 2) the hyoid arch and the basihyal and basibranchial cartilage and 3) the branchial arches, H3: 3b partitions: 1) Meckel's cartilage, 2) the hyoid arch and the basihyal cartilage 3) the basibranchial cartilage and the branchial arches and H4: 4 partitions: 1) Meckel's cartilage, 2) the hyoid arch, 3) the basibranchial cartilage and 4) the branchial arches. The covariation between subsets of landmarks for each hypothesis was measured by an RV coefficient (Escoufier, 1973) and the modularity for each hypothesis was assessed by comparing its RV coefficient to all the RV coefficients of spatially contiguous subset of landmarks (Figure 2) with the same number of landmarks as the hypothesized partitions (Klingenberg, 2009a). If an *a-priori* hypothesis of modularity is supported we would expect the hypothesized landmark partitions to show a weaker correlation between modules than would be seen for other random partitions containing the same number of landmarks. In other words the RV coefficient between the hypothesized partitions is expected to be among the lowest of the RV coefficients for all partitions.

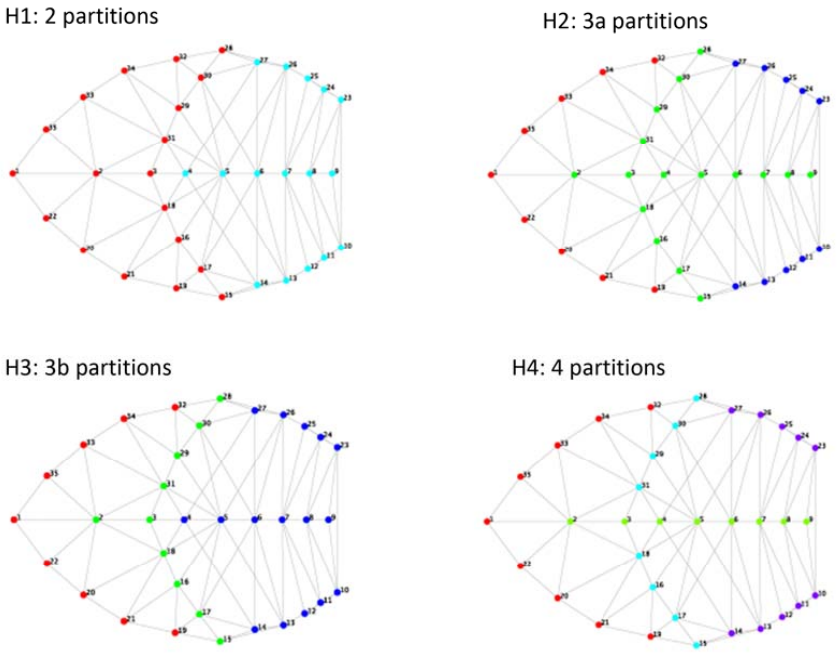

Figure 5.2 Four hypotheses on developmental modularity shown as partitioning of landmarks. Hypothetical modules are shown as different coloured landmark dots: A) H1: two partitions (anterior (red) vs posterior (bright blue) elements); B) H2: three partitions (3a) 1) Meckel's cartilage (red) 2) the hyoid arch, the basihyal and basibranchial cartilage (green) 3) branchial arches (dark blue) C) H3: three partitions (3b): 1) Meckel's cartilage (red), 2) the hyoid arch and the basihyal cartilage (green) and 3) the branchial arches and basibranchial cartilage (dark blue), and D) H4: four partitions: 1) Meckel's cartilage (red), 2) the hyoid arch (bright blue), 3) the branchial arches and basibranchial cartilage (green) and 4) the branchial arch cage (purple).

Note the analyses presented here just represent analyses of modules within each of the groups. We have not compared modules or integration between groups or stages. Thus here we just present preliminary data and analyses bearing on a comparison of integration in hybrids and parental strains. A more thorough analyses will be conducted with custom made scripts and permutation analyses, using methods developed by (Mitteroecker and Bookstein, 2009; Mitteroecker *et al.*, 2012).

5.4 Results

5.4.1 Developmental modularity in craniofacial elements of Arctic charr

To analyse size-related shape changes (allometry) in craniofacial elements in the period between hatching and first feeding we regressed shape (represented by Procrustes shape coordinates) on size (LCS). The results show that allometry accounts for a significant ($p < 0.0001$) part (10.5%) of the total shape changes during this period. Within each morph allometry accounted for 2.4%, 12.9%, 14.5%, 15.9%, and 18.6% of the total shape variation in LB, PL, SB, AC and the hybrids, respectively. These results were highly significant ($p < 0.0001$) for all morphs except LB. Thus the subsequent analyses were performed both without and with accounting for allometry.

Although removing the effect of allometry resulted in slightly lower RV coefficients and lower p-values for the majority of the tested partitions with the symmetric component, the overall results did not change between the two tests and in both cases all modularity hypotheses of landmark partitions were supported (Table S 5.1). The RV coefficient obtained for H1 was 0.492 and only 7 out of 1036 partitions ($p = 0.007$) showed lower RV coefficient than the a-priori hypothesis (Figure 5.3 H1, Table S1). The multi-set RV-coefficients for the 3a, 3b and 4 partitions (0.31, 0.44 and 0.39), respectively, were also among the lowest values calculated for all partitions (Figure 3 H2, H3, H4, Table S1). Landmark partition 3a (H2) had the lowest RV-coefficient and was the best supported partition among the four tested partitions, while the RV-coefficients for 2 (H1) and 3b (H3) partitions were among the highest (Figure 3 H1 and H3, Table S1).

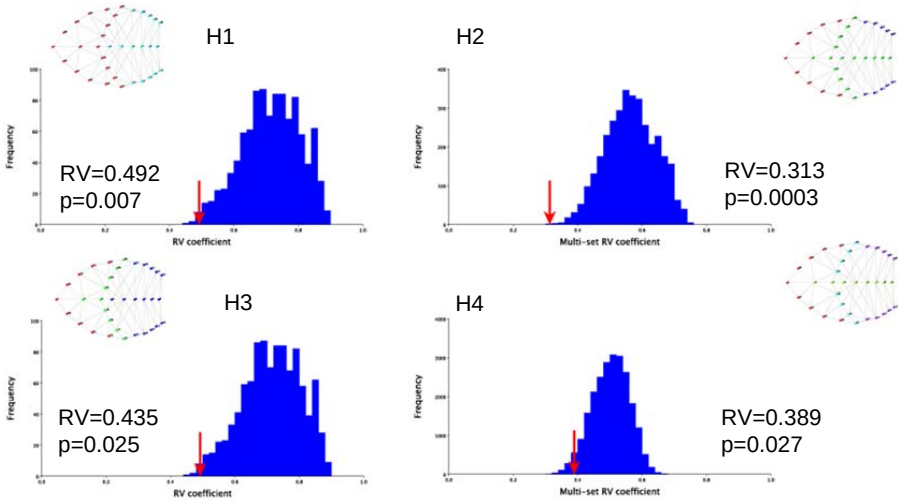

Figure 5.3 Modularity of craniofacial elements in post hatching Arctic charr: The histogram shows the permutational distribution of RV-coefficients for contiguous partitions of the four modularity hypothesis: H1 (two partitions), H2 (3a partitions, H3 (3b partitions), H4 (4 partitions). RV-coefficient and the associated P-value for each hypothesis are shown. The red arrows show the position of the relevant RV-coefficients with respect to the corresponding permutational distributions. The analyses were performed using the residuals from the regression of shape on size (LCS).

5.4.2 Craniofacial modularity and integration in different morphs of Arctic charr

Accounting for allometry did not have an effect on the RV coefficients and p-values for AC and LB, but it decreased the RV-coefficients and the p-values for the two smaller morphs, PL and SB (Table 2). Out of all landmark partitions tested in this study, 3a (H2) had the lowest RV coefficients and p-values for all studied morphs, while the anterior-posterior partitioning (H1) had the highest RV-coefficients (Table 5.2). With a few exceptions (the 3b partition (H3) for the two benthic morphs, SB and LB, and the 4 partition (H4) for SB) the RV coefficient between the hypothesized partitions were among the lowest of the RV coefficients for all contiguous partitions (Table 5.2). While removing the effect of allometry decreased both the RV coefficients and the p-values for SB, it had no effect for LB. This is not surprising as size accounts only for 2.4% of the shape changes and these results were not significant, which indicates that at the stages under study LB appears to be growing isometrically.

Table 5-2 Testing four hypotheses (H1-H4) on craniofacial modularity in post hatching Arctic charr: Shown are RV-coefficients and p-values without (RV and P-val) and with accounting for allometry (RV residual and P-val residual) for four pure crosses; three morphs from Lake Thingvallavatn (LB, SB and PL) and an aquaculture strain from Hólar Aquaculture station. The hypotheses of landmark partitions tested: H1) 2 partitions, H2) 3a partitions, H3) 3b partitions, H4) 4 partitions (See Figure 5.2).

Morph	RV	P-val	RV residuals	P-val residuals	Hypothesis
AC	0.515	0.0738	0.504	0.0430	**H1**
LB	0.593	0.0092	0.601	0.0055	
PL	0.463	0.0022	0.426	0.0022	
SB	0.517	0.0319	0.467	0.0135	
AC	0.312	0.0000	0.292	0.0000	**H2**
LB	0.409	0.0011	0.422	0.0007	
PL	0.351	0.0003	0.314	0.0003	
SB	0.373	0.0084	0.314	0.0003	
AC	0.402	0.0377	0.406	0.0326	**H3**
LB	0.532	0.0583	0.546	0.0570	
PL	0.448	0.0207	0.416	0.0184	
SB	0.455	0.0892	0.406	0.0262	
AC	0.374	0.0874	0.374	0.0148	**H4**
LB	0.480	0.0392	0.498	0.0380	
PL	0.423	0.0236	0.385	0.0093	
SB	0.404	0.0576	0.381	0.0267	

5.4.3 Craniofacial modularity in hybrid progeny of PL and SB

Note, here I present descriptive analyses of the four modularity hypotheses for two morphs and their hybrids. Those are not direct comparisons of integration in those groups. For hypotheses H1 and H2 the analyses of RV-coefficients showed similar results for the progeny of the hybrid crosses as seen in the pure crosses. Despite being somewhat higher in the hybrids, the RV-coefficients departed from the permutational distributions with a high degree of significance (Figure 5.4, Table 5.3). Contrary to what was observed for the progeny of the pure crosses (Table 5.2), the third hypothesis of landmark partition (H3)

was not supported in the hybrid crosses (Figure 5.4 H3, Table 3). Similar results were obtained with and without accounting for allometry (Table 5.3).

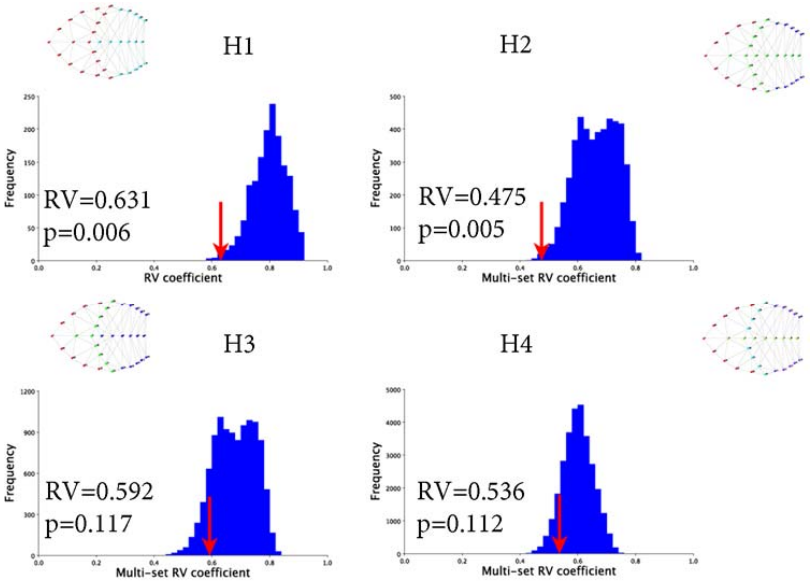

Figure 5.4 Craniofacial modularity in post hatching offspring of hybrid crosses between SB and PL: See Figure 3 for explanations.

Table 5-3 Testing four hypotheses (H1-H4) on craniofacial modularity in post hatching offspring of hybrid crosses of two Arctic charr morphs (SB and PL) from Lake Thingvallavatn. See explanations in Table 3.

	RV	**P-val**	**RV residuals**	**P-val residuals**
H1	0.580	0.0083	0.631	0.0064
H2	0.450	0.0008	0.475	0.0049
H3	0.549	0.1192	0.592	0.1166
H4	0.489	0.0793	0.536	0.1120

Interestingly, accounting for allometry lead to higher RV-coefficients for all tested partitions (Table 5.3). Given the overall integrative effect of allometry and its tendency to mask existing modularity (Klingenberg, 2009b) obtaining higher RV-coefficients after accounting for allometry was somewhat surprising. As changes of shape associated with changes in size (allometry) usually affect the entire organism and are major factors of morphological variation (Klingenberg, 2010), we decided to look further into how shape changes with size for the pure morph and hybrid crosses. For that did Principal components (PCA) of shape from the procrustes adjusted coordinates of the hybrid and each of the pure morph crosses separately. The first three Principal Components (PCs) with size (LCS) for each of the pure morph and hybrid groups were then regressed on size

(LCS). PC1 and PC2 showed significant correlation with size for all studied groups, while PC3 did not correlate with size for any of the studied groups (Table 5.4). Size explained similar amount of the shape variation for PC1 and PC2 in the two pure morph crosses, but it explained more than two times the variance in PC2 in the hybrid crosses (Table 5.3). These results indicate that hybrids exhibit differences in their allometric growth compared to the two pure parental crosses, that manifests most strongly in shape changes captured by PC2 (Figure S 5.2). To further disentangle the allometric relationships between pure and hybrid crosses, we next examined the correlation of two contrasting, simple size measurements, i.e. the length (described by the log transformed euclidean distance between LM 1 and 9) and the width (described by the log transformed euclidean distance between LM 15 and 28) with the PC1 and PC2 scores for the pure and hybrid crosses (Figure S 5.3). While the correlation with the PC2 scores for the width differed among the three groups, the correlation of the width with the PC1 scores was similar for all three groups (Figure S 5.3A). Similarly the correlation of the length with the PC2 differed among the groups, while the correlation with PC2 scores differed among the groups, being more similar for the hybrid and the pure SB cross (Figure S 5.3A). Finally we studied the correlation of the log transformed ratio length/width (Figure S 5.3B) with PC1 and PC2 for the pure and hybrid crosses. While all three groups show similar correlation of the ratio with PC1 scores, the ratio in hybrids did not correlate with the PC2 scores (Figure S 5.3B).

*Table 5-4 Shape changes explained by size (allometry) for the three major Principal components (PCs) of shape. The PCs were computed separately for each group. The individual scores from each PC were then regressed against LCS for PL, SB and Hyb separately. % depict the variation of shape explained by size and the significance was obtained with 10.000 permutations (*** $p<0.001$, ** $p<0.001$, * $p<0.05$).*

	PL	SB	Hyb
PC1	17%*	29%**	10%*
PC2	21%**	29%**	62%***
PC3	ns	ns	ns

5.5 Discussion

The craniofacial development in all vertebrates is a segmental process where each pharyngeal arch develops according to their rhomomeric origin (Lumsden *et al.*, 1991). Here we investigated the modularity of craniofacial elements in early post hatching stages of Arctic charr. We predicted that during this developmental period the craniofacial elements will exhibit low levels of integration and modules will be defined by the developmental origins of the different cartilage groups. In accordance with our prediction the Arctic charr face is highly modular and all four *a priori* hypothesis of landmark partitions we tested were supported at the developmental stages studied here, just after hatching and prior to feeding. The best-supported hypothesis, however (H2) has the basihyal, basibranchial cartilages and the hyoid arch in the same module. Interestingly, these are three separate cartilages with different developmental origins. This finding

indicates that with the onset of breathing movements during late embryonic and early post-hatching stages, the Arctic charr craniofacial elements become more integrated and this integration is likely related to the onset of vital functional demands such as breathing. As expected, allometry had an overall integrative effect (Klingenberg, 2009) and removing the effect of allometry lead to lower RV coefficients (and in the majority of the cases to lower p-values) for the four hypotheses of landmark partitions.

Natural selection can act on development to produce more functionally integrated structures and thus can influence the direction of evolutionary change (Cheverud, 1996). Two alternative hypotheses exist as to the relationship between modularity and evolution: a hypothesis of modular stasis and the alternative hypothesis of modular reorganization (Rogers and Jamniczky, 2014). (Jamniczky *et al.*, 2014) found that patterns of modularity of the trophic apparatus are conserved across oceanic and fresh-water stickleback populations despite the differences of feeding behavior between populations, which is in support of the hypothesis of modular stasis. However, in agreement of the contrasting hypothesis of modular reorganization (Parsons *et al.*, 2012b) showed that patterns of modularity differed between cichlids adapted to two contrasting trophic niches (suction feeders and biting species).

Here we investigated the modularity during early post-hatching of two distinct phenotypes of Arctic charr (benthic and limnetic). Although the data do not indicate difference in the pattern of modularity between morphotypes, the two benthic morphs (LB and SB) had slightly higher RV coefficients and P-values for all a priory hypotheses of landmark partitions tested in the study (Table 5.2). This indicates that the benthic and limnetic morphologies of Arctic charr might differ in their level of craniofacial integration rather than in patterns of integration. Another possibility is that the functional integration of the craniofacial elements starts earlier in the benthic morphs. However more work is needed to unravel the nature of phenotypic modularity in this system.

Finally, we studied the craniofacial modularity in developing hybrid crosses between SB and PL. These hybrids exhibit extreme craniofacial phenotypes when compared to the pure morph crosses (Paper III in this thesis). Interestingly, the RV coefficients and associated p-values were higher in the hybrid than either of the two pure morph crosses. Although it is hard to compare directly between datasets it is nonetheless an interesting pattern that deserves further study. The changes of shape, associated with changes in size (allometry) are major factors of morphological variation (Klingenberg, 2010), especially during ontogeny when growth is the most intense. As growth affects the entire organism, allometry usually has a strong integrative effect (Klingenberg, 2009a). Thus the fact that, after accounting for allometry, the RV coefficients (and some of the associated P-values), increased for all tested landmark partitions in the hybrids was intriguing. As mentioned above allometry is a major factor in morphological variation and an evolutionary level it reflects changes in growth patterns (Klingenberg, 2010) among closely related species (Cardini and Elton, 2008; Gidaszewski *et al.*, 2009; Wilson and Sánchez-Villagra, 2010) or even between populations of the same species (Kimmel *et al.*, 2007; Aguirre *et al.*, 2008). Our analysis indicated that size had a similar effect on major axes of shape, seen in both PC1 and PC2 in the pure parental crosses, but had a stronger effect on PC2 in the hybrids and a relatively weak effect on PC1. The data show that allometric growth differs substantially between the hybrid and the pure morph. This phenomenon is unlikely to reflect morphological divergence, but might rather point towards developmental

instabilities resulting from the combination of two diverged genomes (PL and SB). This phenomenon is very interesting and deserves further study, from both developmental (covering longer developmental period) and molecular perspectives (studying the underlying molecular mechanisms).

In summary we showed that during early post-hatching stages the craniofacial skeleton is modular and that this modularity appears to reflect the developmental origins of the elements constituting it. We compared the craniofacial integration in four groups of Arctic charr and saw indications that these groups may differ in the level of their craniofacial modularity. The hybrid progeny of two contrasting morphs appeared to have different patterns of integration of their craniofacial skeleton compared to the pure crosses of the parental morphs. The hybrid crosses also exhibited different patterns of allometry compared to the pure morphs. These results taken together with the transgressive phenotypes in head morphology (see Paper III in this thesis) the hybrids exhibit may indicate developmental instabilities during craniofacial morphogenesis in the hybrids. As mentioned in the methods, however these results need to be treated with caution and more thorough analyses will be conducted with the data.

5.6 References

Aguirre W, Ullmann B, Currey M, Cresko W, Kimmel C (2008). Allometric change accompanies opercular shape evolution in Alaskan threespine sticklebacks. *Behaviour* **145**: 669–691.

Cardini A, Elton S (2008). Variation in guenon skulls (II): sexual dimorphism. *J Hum Evol* **54**: 638–47.

Cheverud JM (1996). Developmental Integration and the Evolution of Pleiotropy. *Integr Comp Biol* **36**: 44–50.

Drake AG, Klingenberg CP (2010). Large-scale diversification of skull shape in domestic dogs: disparity and modularity. *Am Nat* **175**: 289–301.

Eiriksson GM, Skulason S, Snorrason SS (1999). Heterochrony in skeletal development and body size in progeny of two morphs of Arctic charr from Thingvallavatn, Iceland. *J Fish Biol* **55**: 175–185.

Escoufier Y (1973). Le Traitement des Variables Vectorielles. *Biometrics* **29**: 751.

Frandsen F, Malmquist HJ, Snorrason SS (1989). Ecological parasitology of polymorphic Arctic charr, Salvelinus alpinus (L.), in Thingvallavatn, Iceland. *J Fish Biol* **34**: 281–297.

Gavrilets S, Losos JB (2009). Adaptive radiation: contrasting theory with data. *Science* **323**: 732–7.

Gidaszewski NA, Baylac M, Klingenberg CP (2009). Evolution of sexual dimorphism of wing shape in the Drosophila melanogaster subgroup. *BMC Evol Biol* **9**: 110.

Goodall C (1991). Procrustes methods in the statistical-analysis of shape. *J R Stat Soc Ser B* **53**: 285–339.

Gorodilov YN (1996). Description of the early ontogeny of the Atlantic salmon, Salmo salar, with a novel system of interval (state) identification. *Environ Biol Fishes* **47**: 109–127.

Hallgrimsson B, Jamniczky H, Young NM, Rolian C, Parsons TE, Boughner JC, *et al.* (2009). Deciphering the palimpsest: Studying the relationship between morphological integration and phenotypic covariation. *Evol Biol* **36**: 355–376.

Jamniczky HA, Hallgrimsson B (2011). Modularity in the skull and cranial vasculature of laboratory mice: Implications for the evolution of complex phenotypes. *Evol Dev* **13**: 28–37.

Jamniczky HA, Harper EE, Garner R, Cresko WA, Wainwright PC, Hallgrímsson B, *et al.* (2014). Association between integration structure and functional evolution in the opercular four-bar apparatus of the threespine stickleback, Gasterosteus aculeatus (Pisces: Gasterosteidae). *Biol J Linn Soc* **111**: 375–390.

Jonsson B, Skúlason S, Snorrason SS, Sandlund OT, Malmquist HJ, Jónasson PM, *et al.* (1988). Life History Variation of Polymorphic Arctic Charr (Salvelinus alpinus) in Thingvallavatn, Iceland. *Can J Fish Aquat Sci* **45**: 1537–1547.

Kapralova KH, Gudbrandsson J, Reynisdottir S, Santos CB, Baltanás VC, Maier VH, *et al.* (2013). Differentiation at the MHCIIα and Cath2 Loci in Sympatric Salvelinus alpinus Resource Morphs in Lake Thingvallavatn. (GH Yue, Ed.). *PLoS One* **8**: e69402.

Kapralova KH, Morrissey MB, Kristjánsson BK, Olafsdóttir GÁ, Snorrason SS, Ferguson MM (2011). Evolution of adaptive diversity and genetic connectivity in Arctic charr (Salvelinus alpinus) in Iceland. *Heredity (Edinb)* **106**: 472–487.

Kimmel CB, Aguirre WE, Ullmann B, Currey M, Cresko WA (2007). Allometric change accompanies opercular shape evolution in Alaskan threespine sticklebacks. *Behaviour*: 669–691.

Klingenberg CP Morphometric integration and modularity in configurations of landmarks: tools for evaluating a priori hypotheses. *Evol Dev* **11**: 405–21.

Klingenberg C (2003). Developmental instability as a research tool: using patterns of fluctuating asymmetry to infer the developmental origins of morphological integration. *Dev Instab causes …*: 1–30.

Klingenberg CP (2005). Developmental constraints, modules, and evolvability. In: *Variation*, Elsevier Inc., pp 219–247.

Klingenberg CP (2009a). Morphometric integration and modularity in configurations of landmarks: tools for evaluating a priori hypotheses. *Evol Dev* **11**: 405–21.

Klingenberg CP (2009b). Morphometric integration and modularity in configurations of landmarks: tools for evaluating a priori hypotheses. *Evol Dev* **11**: 405–21.

Klingenberg CP (2010). Evolution and development of shape: integrating quantitative approaches. *Nat Rev Genet* **11**: 623–35.

Klingenberg CP (2011). MorphoJ: an integrated software package for geometric morphometrics. *Mol Ecol Resour* **11**: 353–7.

Klingenberg C, Barluenga M, Meyer A (2002). Shape analysis of symmetric structures: Quantifying variation among individuals and asymmetry. *Evolution (N Y)* **56**: 1909–1920.

Klingenberg CP, Mebus K, Auffray J-C (2003). Developmental integration in a complex morphological structure: how distinct are the modules in the mouse mandible? *Evol Dev* **5**: 522–531.

Lumsden A, Sprawson N, Graham A (1991). Segmental origin and migration of neural crest cells in the hindbrain region of the chick embryo. *Development* **113**: 1281–1291.

Malmquist HJ, Snorrason SS, Skulason S, Jonsson B, Sandlund OT, Jonasson PM (1992). Diet differentiation in polymorphic Arctic charr in Thingvallavatn , Iceland. *J Anim Ecol* **61**: 21–35.

Márquez EJ (2008). A statistical framework for testing modularity in multidimensional data. *Evolution* **62**: 2688–708.

Mitteroecker P, Bookstein F (2009). The ontogenetic trajectory of the phenotypic covariance matrix, with examples from craniofacial shape in rats and humans. *Evolution* **63**: 727–37.

Mitteroecker P, Gunz P, Neubauer S, Müller G (2012). How to Explore Morphological Integration in Human Evolution and Development? *Evol Biol* **39**: 536–553.

Parsons KJ, Márquez E, Albertson RC (2012a). Constraint and opportunity: the genetic basis and evolution of modularity in the cichlid mandible. *Am Nat* **179**: 64–78.

Parsons KJ, Márquez E, Albertson RC (2012b). Constraint and opportunity: the genetic basis and evolution of modularity in the cichlid mandible. *Am Nat* **179**: 64–78.

Parsons KJ, Sheets HD, Skúlason S, Ferguson MM (2011). Phenotypic plasticity, heterochrony and ontogenetic repatterning during juvenile development of divergent Arctic charr (Salvelinus alpinus). *J Evol Biol* **24**: 1640–52.

Parsons KJ, Skúlason S, Ferguson M (2010). Morphological variation over ontogeny and environments in resource polymorphic arctic charr (Salvelinus alpinus). *Evol Dev* **12**: 246–57.

Richtsmeier JT, Aldridge K, DeLeon VB, Panchal J, Kane AA, Marsh JL, *et al.* (2006). Phenotypic integration of neurocranium and brain. *J Exp Zool Part B Mol Dev Evol* **306**: 360–378.

Rogers SM, Jamniczky HA (2014). The shape of things to come in the study of the origin of species? *Mol Ecol* **23**: 1650–2.

Rohlf FJ (2006). "tpsDig, version 2.10." *http://life.bio.sunysb.edu/morph/index.html*.

Rohlf FJ, Slice D (1990). Extensions of the Procrustes Method for the Optimal Superimposition of Landmarks. *Syst Zool* **39**: 40.

Sandlund OT, Gunnarson K, Jonasson PM, Jonsson B, Lindem T, Magnusson KP, *et al.* (1992). The Arctic charr Salvelinus alpinus in Thingvallavatn. *Oikos* **64**: 305–351.

Schilling T, Kimmel C (1997). Musculoskeletal patterning in the pharyngeal segments of the zebrafish embryo. *Development* **124**: 2945–2960.

Schluter D (2000). *The Ecology of Adaptive Radiation*.

Skulason S, Noakes DL. G, Snorrason SS (1989). Ontogeny of trophic morphology in four sympatric morphs of arctic charr Salvelinus alpinus in Thingvallavatn, Iceland*. *Biol J Linn Soc* **38**: 281–301.

Skúlason S, Snorrason SS, Ota D, Noakes DLG (1993). Genetically based differences in foraging behaviour among sympatric morphs of arctic charr (Pisces: Salmonidae). *Anim Behav* **45**: 1179–1192.

Snorrason SS, Skulason S, Sandlund OT, Malmquist HJ, Jonsson B, Jonasson PM (1989). Shape polymorphism in sympatric Arctic charr, Salvelinus alpinus in Thingvallavath, Iceland. *Physiol Ecol Japan* **1**: 393–404.

Svavarsson E (2007). Árangur í kynbótum á bleikju og næstu skref [reference in icelandic]. *Fræðaþing landbúnaðarins (conference proceedings)* **4**: 121–125.

Tsuboi M, Gonzalez-Voyer A, Kolm N (2014). Phenotypic integration of brain size and head morphology in Lake Tanganyika Cichlids. *BMC Evol Biol* **14**: 39.

Walker MB, Kimmel CB (2007). A two-color acid-free cartilage and bone stain for zebrafish larvae. *Biotech Histochem* **82**: 23–8.

Wilson LAB, Sánchez-Villagra MR (2010). Diversity trends and their ontogenetic basis: an exploration of allometric disparity in rodents. *Proc Biol Sci* **277**: 1227–34.

Zelditch ML, Swiderski DL, Sheets HD (2012). *Geometric Morphometrics for Biologists: A Primer (Google eBook)*. Academic Press.

5.7 Appendix

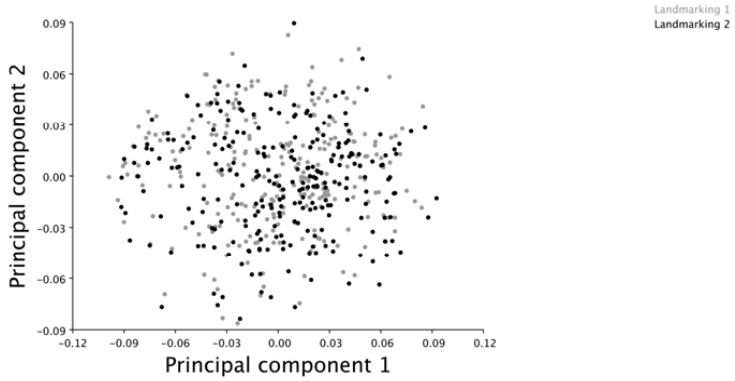

Figure S 5.1 - Scatter plot of the Principal Component Analysis (PCA) scores for the two separated landmarking sessions. The first landmarking session is shown in grey and the second landmarking session is shown in black. Principal Component Analysis (PCA) showed no separation between the two landmarking sessions.

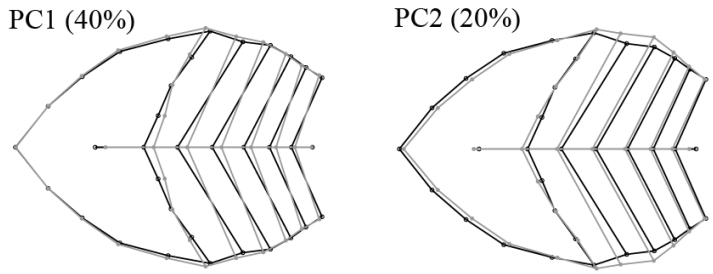

Figure S 5.2: Shape changes associated with PC1 and PC2 from a PCA of pure and hybrid crosses. The scale factor is in units of Procrustes distance and it's set to 0.05. Wireframes depict shape changes associated with the two Principal Components shown in each graph. In the wire frames the extreme negative value is shown in black and the extreme positive values in grey.

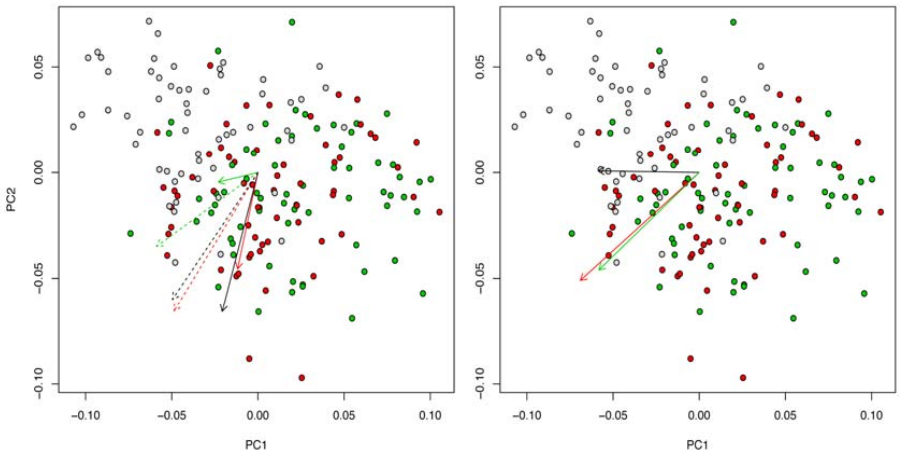

Figure S 5.3 - Scatter plot of the PCA scores for the two major Principal Components (PC1-39% and PC2-20%) of shape variation for two pure parental crosses, SB (red) and PL (green), and the two reciprocal hybrid crosses (in grey) combined. A) Arrows indicate correlation vectors for the length of the head (solid arrows) and the width of the head (dashed arrows) for each cross (SB=red arrows, PL=green arrows, Hyb=black arrows). B) The scatter plot is the same as in A. Arrows indicate the correlation vectors of the ratio of length/width of the head for each cross (SB=red arrow, PL=green arrow, Hyb=black arrow).

Table S1 - Testing four hypotheses (H1-H4) on craniofacial modularity in post hatching Arctic charr: Shown are RV-coefficients and p-values without (RV and P-val) and with accounting for allometry (RV residual and P-val residual) for 258 Arctic charr.The hypotheses of landmark partitions tested: H1) 2 partitions, H2) 3a partitions, H3) 3b partitions, H4) 4 partitions.

	RV	P-val	RV residuals	P-val residuals
H1	0.512	0.0174	0.492	0.0068
H2	0.338	0.0003	0.313	0.0003
H3	0.441	0.0471	0.435	0.0254
H4	0.400	0.0444	0.389	0.0272

Paper 5

6 Paper V

Patterns of miRNA expression in Arctic charr development

Kalina H. Kapralova, Sigrídur Rut Franzdóttir, Hákon Jónsson, Sigurður S. Snorrason, Zophonías O. Jónsson

Authors' contributions: Conceived and designed the experiments: SRF ZOJ KHK SSS. Performed the experiments: SRF KHK ZOJ SSS. Analysed the data: HJ KHK. Contributed reagents/materials/analysis tools: ZOJ KHK SRF SSS. qPCR and alaysis: KHK HJ. Wrote the paper: KHK ZOJ SRF HJ SSS .

6.1 Abstract

Micro-RNAs (miRNAs) are now recognized as a major class of developmental regulators. Sequences of many miRNAs are highly conserved, yet they often exhibit temporal and spatial heterogeneity in expression among species and have been proposed as an important reservoir for adaptive evolution and divergence. With this in mind we studied miRNA expression during embryonic development of offspring from two contrasting morphs of the highly polymorphic salmonid Arctic charr (*Salvelinus alpinus*), a small benthic morph from Lake Thingvallavatn (SB) and an aquaculture stock (AC). These morphs differ extensively in morphology and adult body size. We established offspring groups of the two morphs and sampled at several time points during development. Four time points (3 embryonic and one just before first feeding) were selected for high-throughput small-RNA sequencing. We identified a total of 326 conserved and 427 novel miRNA candidates in Arctic charr, of which 51 conserved and 6 novel miRNA candidates were differentially expressed among developmental stages. Furthermore, 53 known and 19 novel miRNAs showed significantly different levels of expression in the two contrasting morphs. Hierarchical clustering of the 53 conserved miRNAs revealed that the expression differences are confined to the embryonic stages, where miRNAs such as sal-miR-130, 30, 451, 133, 26 and 199a were highly expressed in AC, whereas sal-miR-146, 183, 206 and 196a were highly expressed in SB embryos. The majority of these miRNAs have previously been found to be involved in key developmental processes in other species such as development of brain and sensory epithelia, skeletogenesis and myogenesis. Four of the novel miRNA candidates were only detected in either AC or SB. miRNA candidates identified in this study will be combined with available mRNA expression data to identify potential targets and involvement in developmental regulation.

6.2 Introduction

Since the initial discoveries of lin-4 and let-7 miRNAs have emerged as key regulators of animal development (reviewed in [1,2]). These small (~ 22 nt) non coding RNAs, regulate gene expression by inducing mRNA degradation or translational repression, making for a specific and "fine-tunable" response (reviewed in [3]). miRNAs originate from different parts of the genome (intergenic regions, exons or intronic sequences [4] and are transcribed either as independent transcriptional units or as clusters of several miRNAs (reviewed in [5]). A common feature of all miRNA genes, regardless of their genome location, is the folding of their primary transcript into a stem-loop structure. This hairpin structure is recognized and converted into a miRNA-miRNA* duplex by the miRNA processing machinery (see [5]). One of the strands, dubbed the "mature" miRNA, is then loaded into the miRISC complex while the complementary "star" sequence is often degraded [6]. In most cases of miRNA mediated gene regulation the target repertoire is determined by the "seed" region (nt 2-8 located at the 5' end of the mature miRNA) of the miRNA [7].

In general miRNAs are highly conserved among taxa [5]. Comparative studies show how new miRNAs have continuously been emerging during the evolution of metazoan genomes [8,9] through various mechanisms including gene duplications of preexisting miRNAs followed by changes in their sequences, or *de novo* appearance from random hairpins [10]. However, once they become integrated into the regulatory network, their primary sequence and particularly their seed region, becomes subject to strict selective constraints [5,11]. Variation in timing and expression patterns among species suggests that these molecules may play an important role in shaping physiological differences. For example comparison of two fish species (medaka and zebrafish) showed that heterochrony in miRNAs expression is associated with neuromast and craniofacial development [12]. This was suggested to reflect the differences in formation of the head and sensory epithelia observed between medaka and zebrafish. Morphological differences arising in development can potentially drive evolutionary change, adaptive divergence and speciation (discussed in [1]). More specifically, it has been suggested that miRNAs may generally cover more restricted regulatory niches than transcription factors and thus frequently be more important in terminal differentiation programs [13]. It has also been proposed that miRNAs are involved in enhancing species evolvability by stabilizing gene expression and signaling cascades leading to the increased distinctness of developmental phenotypes, thereby increasing heritability of traits and facilitating natural selection ([14] and discussion in [15]).

6.2.1 Arctic charr as a model species to study adaptive divergence.

The high level of phenotypic polymorphism present in Northern freshwater systems offers an excellent opportunity to study adaptive divergence [16]. These watersheds, with their rivers and lakes, were formed after the last glacial epoch 10 000 - 16 000 years ago. The short evolutionary history characterized by physical variability and topographic dynamics sets a stage where the early steps of divergence may be playing out in multiple locations and species. Studies of whitefish (*Coregonus clupeaformis*), threespine stickleback (*Gasterostreus aculeatus*) and Arctic charr (*Salivelunus alpinus*) have shown that fish inhabiting these systems exhibit an extremely high level of inter-population variation in

phenotype with many populations diversifying along a benthic to limnetic habitat axis [17–20]. Although Arctic charr in Iceland originates from a single Atlantic lineage [21], this species shows an extremely high level of variation in phenotype between populations and many examples of polymorphism (i.e. sympatric morphs) have been documented [17,22–24]. The Arctic charr morphs of Lake Thingvallavatn constitute an extreme example of local phenotypic diversity. Four morphs grouped into two morphotypes have been described in the lake: a limnetic morphotype represented by planktivorous (PL) and piscivorous (PI) charr, with pointed snout and evenly protruding jaws, and a derived, benthic morphotype represented by small (SB) and large benthivorous (LB) charr, blunt snout, short lower jaw and relatively large pectoral fins [25]. These morphs also differ extensively in life history characteristics (size and age at maturity) and embryology [26–29]. The morphs also exhibit strikingly clear differentiation in ecology as reflected in different habitat use, diet and endoparasite fauna [27,28,30]. Several common garden experiments have shown that some key morph specific traits have a definite genetic basis [31,32]. A recent study, using neutral microsatellite markers, revealed significant but subtle genetic differentiation between the three most common morphs in Lake Thingvallavatn, which is consistent with a scenario of early evolution of reproductive isolation, followed by slow divergence by drift with restrictive gene flow [33]. Notably, a study of immune system genes revealed more pronounced genetic differentiation among the morphs in the lake, consistent with a scenario where parts of the immune systems have diverged substantially among Arctic charr morphs from Lake Thingvallavatn [34]. The adaptive nature of the trophic morphology and feeding behavior of the Thingvallavatn morphs has been demonstrated in a series of laboratory rearing experiments [29,31,35]. Moreover the role of developmental heterochrony in the evolution of the Thingvallavatn Arctic charr morphs was demonstrated in a study showing that some skeletal elements of the head start ossifying earlier and/or faster in small benthivorous embryos than in embryos derived from the planktivorous morph [35].

Some of the key differences in functional traits that define the charr morphs are without doubt rooted in differences in the expression of developmental genes. We hypothesize that miRNAs may, through their potentially stabilizing effect of phenotypes [14], play a fundamental role in the divergence of developmental processes that induce differential cranial morphologies in Arctic charr morphs. As a first step of addressing this hypothesis we utilized high-throughput sequencing techniques to identify and annotate Arctic charr miRNAs and to study their expression during the development of two contrasting Arctic charr morphologies. To this end we used a common garden set up to generate embryonic series of two contrasting Arctic charr morphotypes, a benthic morphotype, represented by the SB-charr from Thingvallavatn and a limnetic morphotype represented by fish from the Hólar aquaculture stock (AC). These two morphs differ greatly in adult size, color and head morphology (Figure 1): SB are small, dark and have a sub-terminal mouth and rounded snout whereas AC are large, silvery and have a pointed snout and a longer lower jaw. We sampled AC and SB embryos at four developmental time-points reflecting important events in Arctic charr craniofacial development and used high-throughput sequencing to quantify differences in miRNA expression between the morphs. More specifically we identified and annotated Arctic charr miRNAs using homology to known miRNAs in other species. Furthermore, we identified a large set of novel miRNA candidates by aligning reads to genomic sequences from the closely related Atlantic salmon, *Salmo salar*. Expression levels for both known and novel miRNAs were compared between AC and SB.

6.3 Material and Methods

6.3.1 Sampling and Methodology

All sampling from the wild and rearing in aquaculture was performed according to Icelandic law and with proper permissions. Fish from Lake Thingvallavatn were caught by the authors for the purpose of this study with fishing permissions obtained from the Thingvellir National Park Commission and the owner of the Mjóanes farm. SSS and ZOJ hold special permits for sampling fish from nature for scientific purposes according to Icelandic law (clause 26 of law 61/2006 on salmonid fishing). Control fish from Hólar aquaculture stock were obtained from a national breeding programme, and were not specifically bred for the purpose of this project. These fish are held at the arctic charr breeding station, a quarantined rearing and holding facility, at Hólar University College. After stripping for gametes, parent fish were killed by a sharp blow to the head and checked for absence of breathing when placed in water. Setting up crosses and the subsequent killing of parents was performed by the authors. Ethics committee approval is not needed for regular or scientific fishing in Iceland (The Icelandic law on animal protection, Law 15/1994, last updated with Law 157/2012). The rearing of embryos was performed according to Icelandic regulations (licence granted to Hólar University College aquaculture and experimental facilities) in Verið, Sauðárkrókur, Iceland. Sampling of embryos was performed by University College Aquaculture Research Station (HUC-ARC) personnel. HUC-ARC has an operational license according to Icelandic law on aquaculture (Law 71/2008), that includes clauses of best practices for animal care and experiments. For this study the last gestation age at which embryos were sacrificed was 434 (τ_S) units. For RNA extraction, samples were flash frozen in RNA later. Prior to freezing eggs were permeabilized by puncture with a needle. Samples for staining (not described in this study) were treated with an overdose of phenoxyethoanol before fixing.

For this study we used developmental time-series from pure crosses of two Arctic charr morphs, Hólar aquaculture charr (AC) and small benthic charr (SB) from Lake Thingvallavatn. These strains were selected mainly for their pronounced differences in body size, coloration and head morphology (Figure 6.1). As stated above, the AC crosses were made with parents from the Hólar breeding programme [36]. Fish from the small benthic morph (SB) were caught in Lake Thingvallavatn using gill-nets. Eggs from several females were pooled and fertilized using milt from several males from the same group. Eggs were reared at approximately 5°C in a hatching tray (EWOS, Norway) under constant water flow and in complete darkness at the Holar University College experimental facilities in Verið, Sauðárkrókur. Exact water temperature was recorded twice daily to estimate the relative age of the embryos using tau-somite (τ_S) units defined as the time it takes for one somite pair to form at a given temperature [37]. Embryos were collected throughout development and either fixed in 4% PFA or stored in RNAlater (Ambion) at -80°C. Based on embryos sampled at different developmental stages and stained with alcian blue (cartilage) and alizarin red (bone), four time-points (141, 161, 200 and 434 τ_S) were selected to represent important stages of bone and cartilage development. Stages 141, 161 and 200 are embryonic whereas stage 434 is a fry stage and for simplicity these stages will be referred to as stages 1, 2, 3 and 4, respectively. Two independent samplings were performed: one was used for high-throughput small-RNA-sequencing (miRNA-seq) and the other one for qRT-PCR.

Figure 6.1 Two contrasting Arctic charr morphs differing in size, coloration and head morphology: Top: Arctic charr from aquaculture stock (AC) is large, silvery and has a pointed snout and long lower jaw, Bottom: Small benthic charr from Thingvallavatn (SB) is small, dark and has a sub-terminal mouth and rounded snout.

6.3.2 Small RNA sequencing

Total RNA from each stage of each morph was isolated from a pool of 6 whole embryos and enriched for small RNAs using the mirVana kit (Ambion). The purity and amount of small RNA was verified on a BioAnalyzer (Agilent Technologies). The samples were prepared for sequencing following the small RNA v1.5 sample preparation protocol from Illumina. Briefly, 3‘ and 5‘ RNA adapters were ligated to small RNAs, which were subsequently, reverse transcribed into DNA and PCR amplified. The samples were then run on polyacrylamide gels and the DNA eluted from bands corresponding to 20-30 nucleotide RNA fragments. miRNA and transcriptome sequencing (mRNA-seq) was performed at deCODE Genetics (Reykjavik, Iceland) using the TruSeq smallRNA (v1.5) kit (Illumina) on an Illumina $GAII_X$ instrument. Raw reads were submitted to NCBI Sequence Read Archive (SRA) under accession number SRP039492.

6.3.3 miRNA-seq data processing

Raw reads were processed with cutadapt [38] as follows: First, adaptor sequences were removed and only reads with adaptors were kept. Next, we used the FastX toolkit (http://hannonlab.cshl.edu/fastx_toolkit/index.html) (script available on request) and the quality scores associated with the reads to remove bases with a Phred based quality score [39] of less than 20 from read ends. Sequences retaining less than 15 nucleotides after filtering were discarded. Reads where 10% or more of the bases had a Phred quality score lower than 20 were also discarded. Finally, identical reads were reduced to one copy with the redundancy noted in the read name. The sequence filtering and collapsing was repeated for each sample.

6.3.4 Annotation of ncRNAs

To annotate sequences using known RNAs we used Rfam version 10.1[40] and miRBase [41–44] version 20 databases. The Rfam database was searched with HMMER (version 3.0; http://hmmer.janelia.org/) with an e-value cutoff of 0.01. For the miRBase the ssearch command from the fasta package version 36.3.6d [45] was used to detect homology between the mature miRNAs and the collapsed sequences (e-value cutoff 0.01).

6.3.5 On a quest for novel miRNAs

To identify novel miRNAs we used a probabilistic model of miRNA biosynthesis implemented in miRDeep 2 [46]. As a sequenced Arctic charr genome is not currently available, we used the genome sequences from the closely related Atlantic salmon [47] for reference. Collapsed reads were aligned to the Atlantic salmon genome with the mirDeep 2 mapping program (10 minimum reads per miRNA) for both morphs with the time-points combined. To facilitate mapping, collapsed sequences with strictly lower read count than 4 were omitted from the detection. The miRDeep2 algorithm then uses the reference regions bracketing the aligned reads to compute a hairpin structure and estimates the probability of each sequence being a true miRNA precursor based on the position of the reads, their frequency, the energetic stability of their secondary structure and conservation of the 5' ends.

Sequences with log score greater or equal to 2 were considered as potential miRNAs, the predicted hairpins were searched against hairpins from mirBase (version 20) with blastall (version 2.2.26, -W 7)[48]. Candidates were annotated as known miRNAs if the alignment length was greater or equal to 60 nucleotides and expected value for the match was lower than 0.01 (-e 0.01) otherwise the hairpins were classified as novel.

6.3.6 PCR amplification and sequencing of miRNA clusters

To assess the degree of sequence conservation for genomic clusters containing known and novel miRNAs between Arctic charr and Atlantic salmon we selected 4 clusters containing known miRNAs (miR-19c, 18b* and 20b; miR-133a and miR-133b and miR-143-3p and miR-143-5p; miR-219-3p and miR-219-5p) and 3 clusters containing novel miRNA candidates (sal-nov-235, sal-nov-242 and sal-nov-343) and PCR amplified their genomic regions from the Arctic charr genome. Primers were designed with Primer3 (http://primer3.wi.mit.edu/) (Table S1). The same PCR program was used for all primer pairs: an initial denaturation at 95 °C for 5 min; 35 cycles of 95 °C for 45 seconds; 45 seconds at a 53 °C; 1 min at 72 °C, then a final step of 10 min at 72 °C. PCR products were treated with ExoSap and sequenced on an Applied Biosystems 3500xL Genetic Analyzer using BigDye chemistry. Raw sequencing data was base-called by Sequencing Analysis Software v5.4 with KBTMBasecaller v1.41 (Applied Biosystems), and run through Phred and Phrap, prior to trimming primer sequences, visual editing of ambiguous bases and putative polymorphisms in Consed [49]. Fasta files were exported and aligned with ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/, and manually inspected for alignment errors in Genedoc (www.psc.edu/biomed/genedoc). All sequences were deposited in Genebank under the accession numbers [KJ573796-KJ573802]. These sequences were then searched using blast against the salmon database. The conservation

between Arctic charr and salmon ranged between 91-94% for the known miRNAs and 92-98% for the novel miRNA candidates. Mismatches were always located outside of the miRNA mature-star sequence.

6.3.7 Differential expression analysis

The R package edgeR [50] was used to study the differential expression of conserved and novel miRNA candidates between morphs and among developmental time-points in a generalized linear model, where the additive covariates (no-interaction) corresponded to developmental time-point and different morphs. The normalization factors were calculated for each sample using the function calcNormFactors. As there are no replicates in any of the experimental conditions, the options method="deviance", robust=TRUE and subset=NULL were used for estimating the common dispersion (function estimateGLMCommonDisp) parameters as recommended by the edgeR user manual. The trended and tagwise dispersion were also estimated (function estimateGLMTrendedDisp and estimateGLMTagwiseDisp) with default options. The statistical significance of the terms was assessed by comparing likelihood difference to a reduced model without the time or the morph terms, with the function glmLRT in edgeR. The first 20 bases of each annotated sequence (novel and previously described miRNAs) were used as an identifier and the counts were aggregated for sequences that share the first 20 bases. This allowed us to work at the sequence level without lumping together isoforms (isomiRs). An entry (first 20 bases) was only considered for the statistical testing if the counts per million reads were strictly greater than 3 in at least two experimental points resulting in 1862 tags. We adjusted for multiple testing using the Benjamini-Hochberg false discovery rate [51]. The R script used for this analysis is available in supplementary File S 6.1.

6.3.8 Descriptive analysis

Cluster analysis was performed using the heatmap function and plotted using the gplots package in R (http://www.r-project.org/). Prior to the clustering analysis expression levels for each miRNA were normalized across samples using a Variance Stabilizing Transformation.

6.3.9 Real-time qualtitative PCR analysis

In order to verify the observed differential expression between morphs in our miRNA-seq data, we selected 9 miRNAs (sal-miR-17, sal-miR-26a, sal-miR-30b, sal-miR-122, sal-miR-140, sal-miR-181a*, sal-miR-196a, sal-miR-199a and sal-miR-206) for qPCR analysis. We concentrated on the 3 embryonic stages, as in both our analyses (for morph or developmental effect) the expression profiles between the samples of the last stage appeared to be very similar (see results). For the qPCR analysis two separate RNA extractions (biological replicates) were used for each data point. RNA was extracted from pools of 6 whole embryos/fry using a standard TRI Reagent (Sigma) protocol and treated with DNaseI (New England Biolabs) in order to limit genomic DNA contamination. All samples were from the same sampling effort and were extracted and processed simultaneously. cDNA was synthesized using the Exiqon universal cDNA Synthesis Kit II. The consistency of the cDNA synthesis among samples was verified using a spike in template along with a Control primer set (Exiqon). For the qPCRs we used SYBR Green

master mix (Exiqon) and LNA primers (Exiqon). All qPCRs were done in duplicates (technical replicates) in a 10 µl reaction volume in 96 well-PCR plates on an ABI 7500 real-time PCR System (Applied Biosystems) following manufacturer instructions (Exiqon). The same PCR program was used for all miRNA primer pairs: starting with a 2 min hold at 50°C followed by a 10 min initial denaturation at 95°C and 45 cycles of 10 sec denaturation at 95°C and 1 min annealing/extension at 60°C. A melting curve analysis was performed at the end of each PCR to verify the specificity of the amplification. U2 spliceosomal snRNA (Primer sequence: GGTACTGCAATACCGGGG) was initially selected as a reference gene. However, the use of non-miRNA genes as reference has been shown to be problematic and the use of mean expression is often more appropriate [52]. We therefore opted to use the geometric mean for the expression values of the miRNAs under study as a reference. Relative expression (fold change) for each miRNA compared to stage 1 in AC was calculated in R using a script provided in supplementary File S 6.2.

6.4 Results

6.4.1 Small RNA descriptive statistics

In order to identify miRNAs involved in Arctic charr development and morph differences, we made 8 small RNA libraries from four developmental time-points of two contrasting morphs of Arctic charr. The sequencing depth ranged from 29.1 to 33.9 million reads with a mean depth of 32.4 million reads per sample. After removing the adapters using cutadapt [38] and filtering out low quality reads using the FastX toolkit (http://hannonlab.cshl.edu/fastx_toolkit/), we obtained on average 24.8 million reads per sample (Table 6.1).

Table 6-1 Summary of read numbers from small RNA sequencing. Number of reads (NR, in millions of reads) in high-throughput data from small RNA libraries of four developmental points and two morphs of Arctic charr. AC = Aquaculture charr, SB = Small Benthic charr.

Sample	Number of reads (NR)	NR after adapter trimming	NR after collapsing
SB 1	33.4M	23.1M	2.4M
SB 2	32.8M	28.2M	0.9M
SB 3	33.9M	30.5M	0.7M
SB 4	29.1M	21.6M	0.7M
AC 1	32.8M	24.2M	1.7M
AC 2	32.5M	21.1M	2.0M
AC 3	31.3M	23.2M	0.7M
AC 4	33.6M	26.2M	1.0M

The size distribution of the collapsed reads of all 8 libraries accounting for redundancy is shown in Figure S1. The majority of the reads were 21-23 nucleotides, corresponding to the typical miRNA size range. Details of the size distribution for both unique and collapsed reads for all 8 libraries are shown in Figure S 6.2. All 8 libraries showed similar distribution with a peak at 21-23 nt. Furthermore, annotation of the collapsed reads using the Rfam database confirmed that our small RNA libraries were highly enriched with miRNAs (Table 6.2).

Table 6-2 High-throughput reads annotated using the Rfam database.

snc RNA	Number of reads
miRNA	50841311
rRNA	49033
mRNA	5755
tRNA	24451
SNORD	151004
U	208259
sno	25857

6.4.2 A total of 326 conserved and 427 novel miRNA candidates were found in the data

All collapsed reads were compared to the mature miRNA sequences available in miRBase (release 19) using ssearch [45] and 326 candidates (Table S 6.2) were identified with high confidence (e-value < 0.001). The 10 most abundant miRNAs account for 65% of the total conserved miRNAs (Figure 6.2) with sal-miR-206 and sal-miR-1 alone accounting for 36% of the total miRNAs. We identified 427 novel miRNA candidates (Table S 6.3) of which 37% were represented by the 10 highest expressed putative miRNAs. We sequenced the genomic regions of three novel miRNAs (sal-nov-235, sal-nov-242 and sal-nov-334). They were all highly conserved between Arctic charr and Salmon (Table S 6.1). Furthermore their mature and star sequences are located in highly conserved blocks in medaka, fugu, tetraodon and stickleback. Several of the conserved miRNAs were present in two or three isoforms (isomiRs). For example sal-miR-451 exists in 3 isoforms (Table S 6.2). Two of these (sal-miR-451_1 and sal-miR-451_3) are highly conserved among vertebrates, whereas the third (sal-miR-451_2) has not previously been described in other species. This derived isoform differs in one base (G->U substitution) located at the 3' end of the mature sequence and is the predominant isoform of sal-miR-451 in our data (Table S2). Another interesting example is sal-miR-152, where 4 isoforms are found in our data (Table S 6.2) with the most abundant being the ancestor sequence. The three other isoforms are one mutation away from the ancestral form. Interestingly, these mutations (T->A, C, or G) are located at the same site (position 5) for all 3 derived isoforms (Table S 6.2).

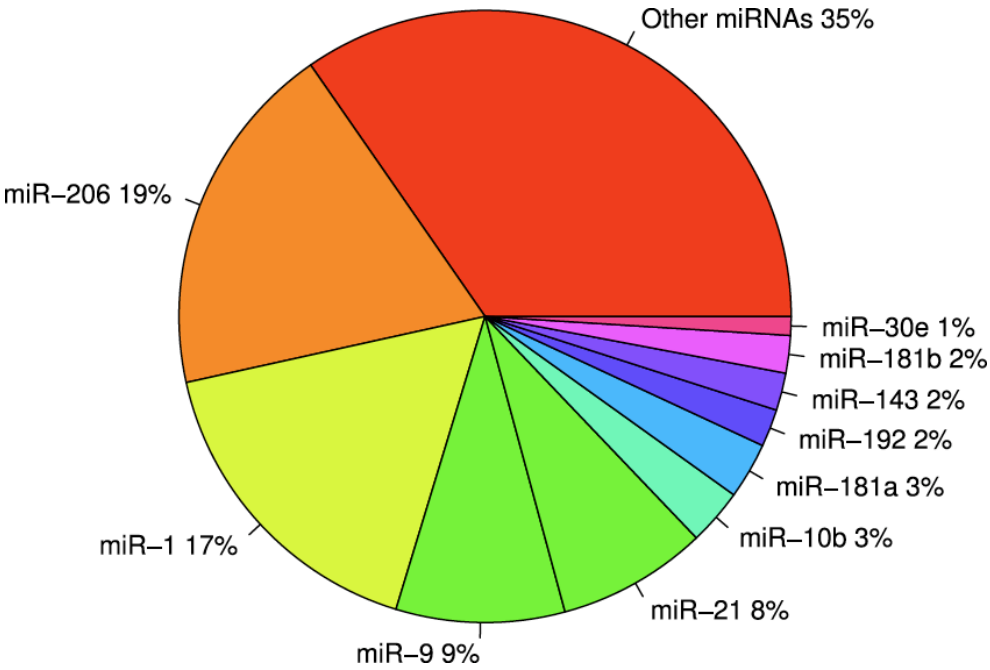

Figure 6.2 Relative abundance of known miRNAs in all samples combined. Together the 10 most abundant known miRNAs constitute 65% of all known miRNAs.

6.4.3 51 known miRNAs and 6 novel miRNA candidates are differentially expressed among developmental stages

We found 51 known miRNAs to be differentially expressed among developmental time-points. Hierarchical clustering (Figure 6.3, see also Table S4 for background data) of miRNA expression showed that the 8 samples grouped into four clusters according to developmental time: one major division separates the three embryonic stages (1, 2 and 3) from the last post-hatching stage (4) and three divisions for each of the three embryonic stages (Figure 6.3). This major division between embryonic and post hatching stages indicates a clear shift in miRNAs expression between these developmental phases. The second division separates stage 3 from stages 1 and 2 and the third division separates stages 1 and 2 (Figure 6.3). Interestingly, there are two major divisions in the miRNA expression pattern clustering: node one depicts miRNAs that are highly expressed during the embryonic stages and their expression decreases in the last stage while the second node includes miRNAs with high expression in stage 4 and low expression in the embryonic stages. For example members of the 430 family (miR-430 a, b, c and d) are highly expressed in the embryonic stages and their expression decreases markedly in late development. In addition other miRNAs, such as sal-miR-219 a and b and miR-181c, show higher expression in the embryonic stages. On the other hand, miRNAs such as sal-miR-22a, 140, 182, 183, 192, 215 and different members of the let-7 family show increasing expression over time. Of the novel candidates, 6 putative miRNAs were found to be differentially expressed among developmental points (Table 6.3). Three of them sal-nov-1, sal-nov-5 and sal-nov-18 are also differentially expressed between morphs.

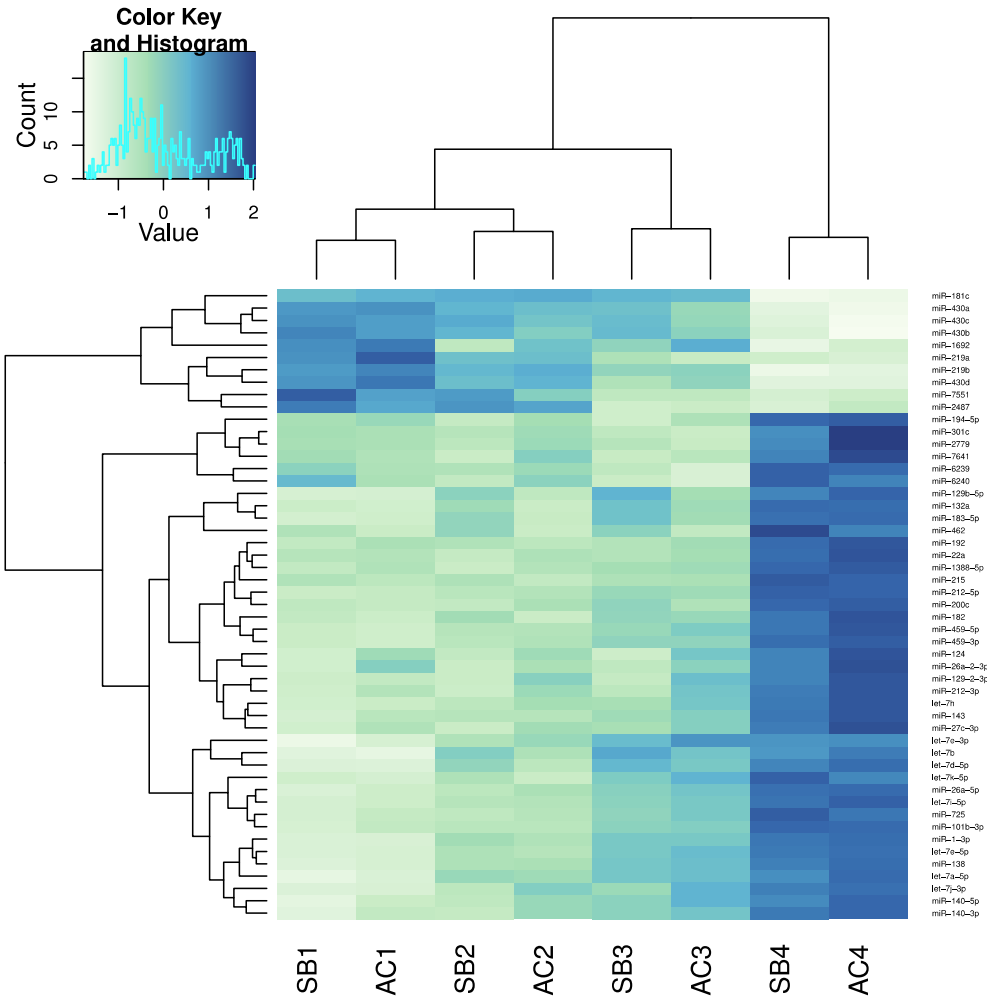

Figure 6.3 Heat-map showing relative expression of the 51 miRNAs significantly differentially expressed among developmental stages. Expression levels for each miRNA were normalized across samples using variance stabilizing transformation. Blue denotes high and white low relative expression. AC stands for Aquaculture charr and SB stands for Small benthic charr. Numbers 1, 2, 3 and 4 depict the four developmental time-points.

Table 6-3 Novel miRNA candidates differentially represented between morphs and/or developmental stages in the small-RNA-seq data. Included are miRNA names, differential expression and number of raw reads per stage per morph.

Name	Expressed	SB1	SB2	SB3	SB4	AC1	AC2	AC3	AC4
sal-nov_1	Morph/Time	7	4	1	1	17	31	25	10
sal-nov_2	Morph	8	13	16	9	104	89	151	88
sal-nov_3	Morph	34	78	63	17	0	0	0	0
sal-nov_4	Morph	0	0	0	0	13076	5397	5755	560
sal-nov_5	Morph/Time	0	0	0	0	1800	563	343	0
sal-nov_6	Morph	47	9	14	18	73	58	23	28
sal-nov_7	Morph	25	5	1	9	48	45	14	5
sal-nov_8	Morph	207	168	20	7	256	117	47	71
sal-nov_9	Morph	130	59	13	54	268	361	173	70
sal-nov_10	Morph	16	6	12	10	45	53	52	25
sal-nov_11	Morph	8	13	16	9	104	89	151	88
sal-nov_12	Morph	8	12	13	6	14	17	15	6
sal-nov_13	Morph	21	44	112	133	113	110	236	308
sal-nov_14	Morph	2	32	31	53	27	25	96	103
sal-nov_15	Morph	3	22	32	31	6	7	8	10
sal-nov_16	Morph	644	1472	1725	983	966	648	944	1011
sal-nov_17	Morph	24	30	27	2	65	41	69	50
sal-nov_18	Morph/Time	13	89	121	372	3	11	17	219
sal-nov_19	Morph	8	4	9	0	207	63	107	15
sal-nov_20	Time	70	274	967	5358	128	213	988	6078
sal-nov_21	Time	3472	7090	12625	26369	5005	4127	10188	20766
sal-nov_22	Time	121	107	56	1	206	71	26	0

6.4.4 53 known miRNAs and 19 novel miRNA candidates are differentially expressed between AC and SB embryos

We tested for differential expression between morphs using a Generalised Linear Model and adjusted for multiple testing using the Benjamin-Hochberg false discovery rate as decribed in methods. We found 53 miRNAs to be differentially expressed between AC and SB. These miRNAs cluster by morph during the embryonic stages (stages 1-3) (Figure 6.4, see also Table S 6.4 for background data). During these 3 stages miRNAs such as sal-miR-130, 133, 153, 17, 30, 451, 219, 26, 199a and 145 are highly expressed in AC, whereas sal-miR- 206, 133, 122, 181a, 192, 196a and 223 are highly expressed in SB. The expression of some of these "morph specific" miRNAs for example sal-miR-130, 153, 17, 30b and 30c in AC and sal-miR-196a, 206, 192 and 122 in SB observed in the embryonic stages decreases markedly in the last stage. During the last stage the observed miRNA expression differences between the two morphs disappear (Figure 6.4). Of the novel miRNA candidates, 19 putative miRNAs were found to be differentially expressed between AC and SB (Table 6.3). Two of them, sal-nov-4 and 5 were only expressed in AC and at most/all stages whereas expression of another putative novel miRNA, sal-nov-3, was only detected in SB offspring and at all four developmental points. With three exceptions (sal-nov-4, 5 and 16) all of the differentially expressed putative miRNAs showed very low expression levels (Table 6.3).

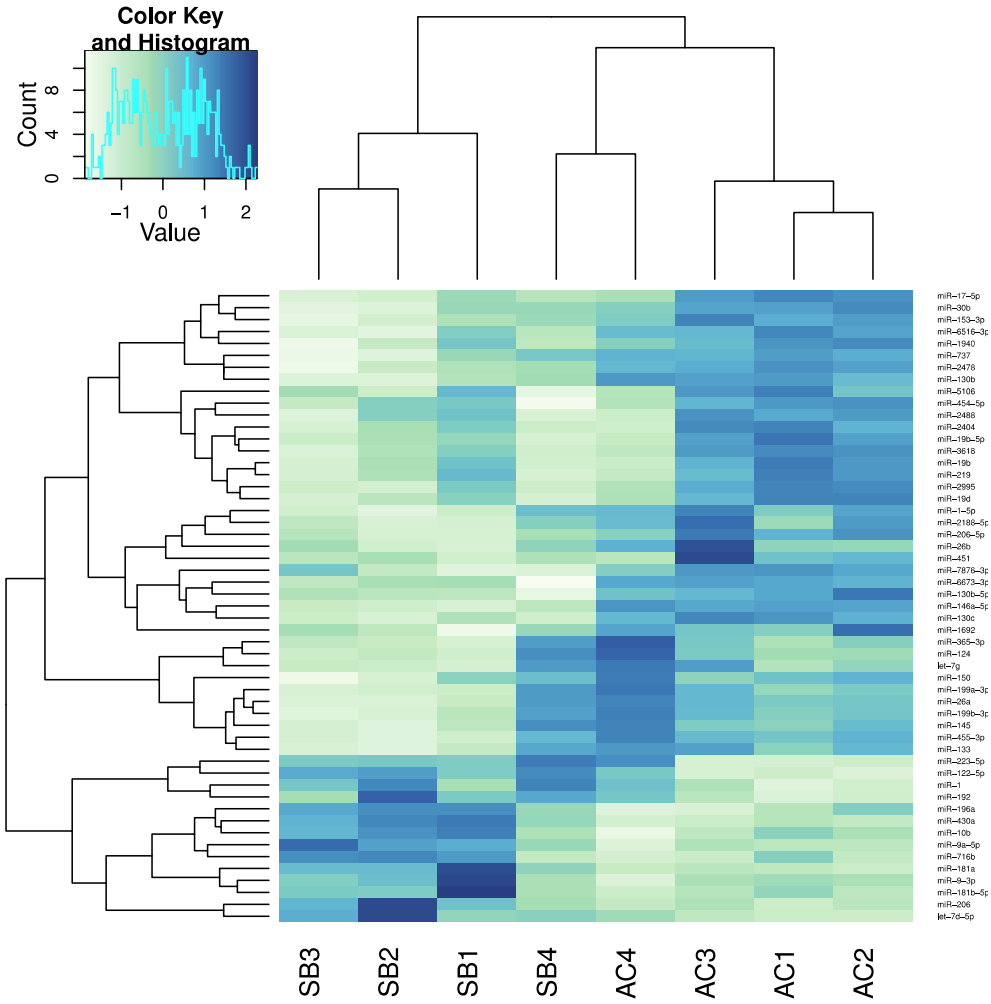

Figure 6.4 Heat-map showing relative expression of the 53 miRNAs significantly differentially expressed between AC and SB morphs. Expression levels for each miRNA were normalized across samples using variance stabilizing transformation. Blue denotes high and white low relative expression. AC stands for Aquaculture charr and SB stands for Small benthic charr. Numbers 1, 2, 3 and 4 depict the four developmental points.

We selected 9 miRNAs and further studied their expression by qPCR using independent biological replicates. The selection was based on the dynamics and degree of differential expression between morphs and/or among developmental points seen in the sequencing data. We concentrated on the three embryonic stages, as in both our analyses (for morph or developmental effect), the expression profiles between the samples of stage as the expression profiles between the samples of stage 4 appeared to be very similar. Eight of these miRNAs (miR-17, miR-26a, miR-30b, miR-140, miR-181a*, miR-196a, miR-199a and miR-206) amplified well (Figure 5), whereas miR-122 showed double peaks in melting curve analysis and was discarded from further analysis. Five (miR 17, 26a, 30b,

140 and 206) out of eight miRNAs tested with qPCR showed similar expression patterns to what was expected from the high-throughput sequencing (Figure 6.5, A-E). Three miRNAs (miR-196a and miR-199a and miR-181a) exhibited similar expression patterns in one or two of the three stages under study, (Figure 6.5, F-H).

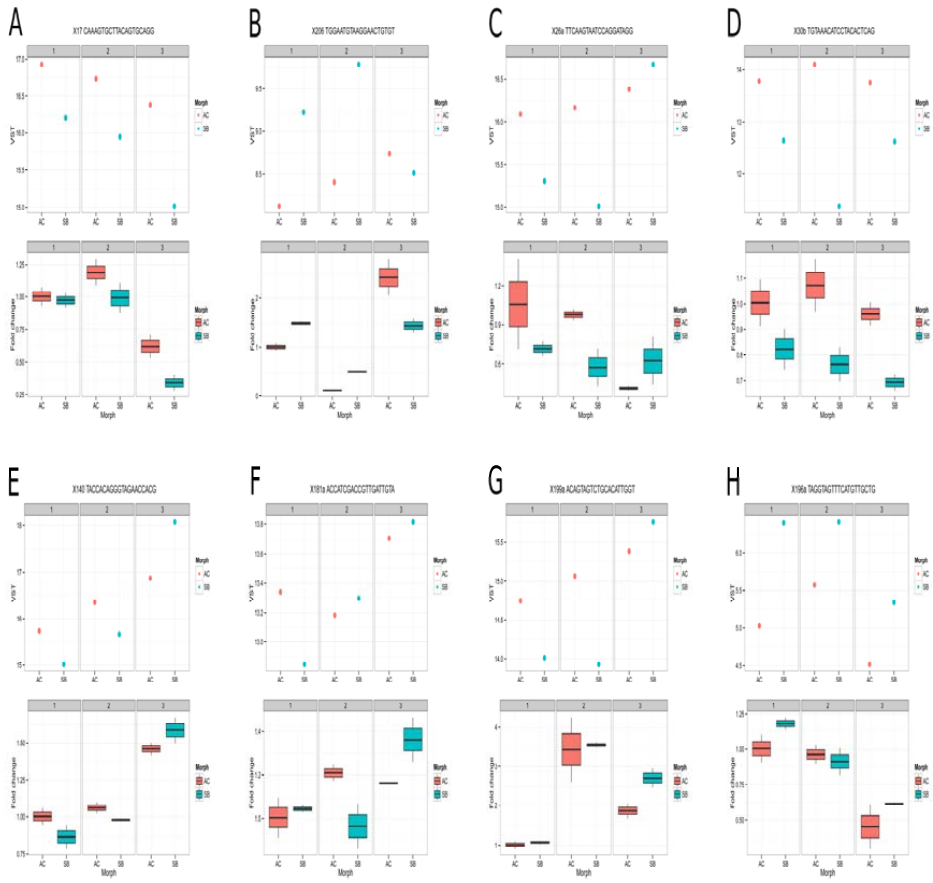

Figure 6.5 Comparison of expression of 8 selected miRNAs (miR-17, 26a, 30b, 140, 181a, 196a and 199a) at three developmental time-points for two contrasting Arctic charr morphs (AC and SB) quantified by small RNA-seq (upper panel) or qPCR (lower panel).

6.5 Discussion

The molecular mechanisms underlying the development of Arctic charr morphologies are likely to have parallels in other vertebrate species and studying them is of interest in both developmental and evolutionary contexts. Given the recent divergence in northern populations of Arctic charr, it is likely that the observed phenotypic polymorphism rooted in development in this species arose mostly by differences in gene regulation as opposed to changes in protein coding sequences. In a recent study [53] two genes involved in matrix remodeling in bone formation (sparc and mmp2) showed consistent differences in expression during the development of embryos derived from benthic and limnetic morphs of Arctic charr, suggesting that these genes might be involved in the development of these distinct Arctic charr morphologies. Little is known about what controls such differences in expression. While numerous studies demonstrate the involvement of transcription factors and other regulatory elements in the phenotypic evolution of birds [54] and fish [55–60], there are few known examples of miRNAs playing a role in morphological variation. In a recent study Arif *et al.* [61] experimentally demonstrated that differences in the "naked valley" phenotype observed among natural populations of *D. melanogaster* were caused by variation of miR-92a expression. Our study is the first phase of assessing the involvement of miRNAs in the development of key morphological traits and their potential role in the morphological evolution of the highly polymorphic Arctic charr. In so doing we also hope to shed light on some of the developmental circuitry operating at these levels of development.

Using small RNA-seq we found 326 known and 427 novel miRNA candidates in Arctic charr. A few of the candidates, termed novel (Tables 3 and S2) are absent from miRbase but have been previously identified in other salmonid species [62–64]. When only the 326 known miRNAs are considered, the 10 most abundant ones account for 65% of reads (Figure 2). These miRNA are highly conserved among taxa and have been shown to have important functions during development. The two most abundant miRNAs in our data, miR-206 and miR-1, together account for 36% of the total known miRNAs (Figure 2). Their role in skeletogenesis and myogenesis has been studied in some detail, for example miR-206 has been found to induce myogenic differentiation [65–67] while inhibiting osteoblast differentiation [68], and miR-1 has been found to regulate skeletal muscle and cardiac development [69,70]. These miRNAs are highly conserved in animal evolution [8] with miR-1 retaining its muscle-specific expression from *C. elegans* to human [71]. Other highly expressed miRNAs are involved in cardiogenesis (miR-21), neurogenesis (miR-9), gut and gall bladder development (miR-143 and miR-192) [72]. One of the oldest miRNAs in the animal kingdom (miR-10) [8] is also among the 10 most highly expressed miRNAs in our data. Encoded in the intron of hoxB4, this miRNA is suggested to play a role in anterior posterior patterning.

A few miRNAs exist in multiple forms in our data. Among these the most interesting examples are miR-152 and miR-451. In the case of miR-152 the polymorphism is located in the seed region, which suggests functional divergence. In the case of miR-451 the ancestral and derived forms differ in one base (G->U substitution) located in the 3' end of the mature sequence. Interestingly the derived form of miR-451 is also the most abundant one. Although not as essential as the seed region, 3' miRNA-target pairing has a role in defining target specificity within miRNA families [73]. The derived form of miR-451 might have evolved following the whole genome duplication Salmonids underwent 25-100 million years ago or as a result of gene duplication. Other possibilities for the presence of

this miRNA in our data include post-transcriptional editing of the ancestral form. However, without a sequenced Arctic charr genome, distinguishing between evolutionary scenarios represents a challenge. The derived form of mir-451 might be specific to Arctic charr as it is not present in the salmon genome and has not been reported in rainbow trout. As miRNAs often co-evolve with their targets [5] further phylogenetic analysis will help shedding light on the evolution of miR-152 and miR-451 and their targets in Arctic charr.

51 previously annotated miRNAs were found to be differentially expressed among developmental points. The cluster analysis of these miRNAs showed a clear shift in miRNA expression between the embryonic stages and the post-hatching stage, visible from both dendrograms. Although samples grouped by developmental stage, the major division was between stages 1, 2, 3 and stage 4. These findings were further confirmed by the existence of two major clades of miRNA expression: one containing miRNAs highly expressed during the embryonic stages and one the last stage. Among the miRNAs showing high expression in early development are the members of the 430 family (miR-430 a, b, c and d). In zebrafish these miRNAs are involved in the maternal to zygotic transition by deadenylation and clearing of maternal transcripts [74]. The majority of the miRNAs showed higher expression in the last developmental stage. Examples include miR-1, members of the let-7 family, miR-22a, miR-140, miR-182, miR-183, miR-192 and miR-215. The evolutionarily ancient and highly conserved let-7 family is involved in the regulation of the timing of developmental events in *C. elegans,* in particular the transition from larval stage 4 (L4) to adult [75]. In vertebrate development let-7 is temporally regulated and it is thought to play a role in late temporal transitions during development [75]. Other miRNAs found to be highly expressed in the last developmental stage are also involved in major developmental processes such as muscle differentiation (miR-1), endochondrial bone development (miR-140) and neuromast differentiation (miR-182 and 183) [12,82]. Overall 72 miRNAs (19 novel miRNA candidates and 53 conserved miRNAs) were found to be differentially expressed between AC and SB at the developmental points under study. Of those sal-miR-196a, sal-miR-206, 122, 192, 196a, 223 and 181a were more highly expressed in SB whereas sal-miR-26a, 30b, 17-5p, 153-3p, 130b and c, 199a were more highly expressed in AC (Figure 4). All of the conserved miRNAs showing variation in expression between the Arctic charr morphs have been found to play an important role in development. For example, miR-196a, which is encoded in a Hox cluster, has been found to be involved in axial and appendicular patterning in chicken [76] and zebrafish [77]. Another muscle specific miRNA, miR-206, shows large expression differences between SB and AC especially at stage 3, where there is a 2.5 fold difference between the two morphs. This miRNA is involved in muscle differentiation and its expression is up-regulated by MyoD in differentiating muscle fibers. Loss of function of MyoD leads to down-regulation of miR-206 and severe deformities in the craniofacial elements [78]. miR-206 has also been shown to directly affect osteoblast differentiation and its overexpression in the osteoblasts of transgenic mice leads to bone abnormalities [68].

Other conserved miRNAs, miR-130b and c, miR-133, miR-199a-3p, miR-26a and miR-451 were highly expressed in AC throughout early development compared to SB. Some of these miRNAs are involved in myogenesis and skeletogenesis, for example, miR-199a-3p is important for normal skeletal development. In mouse a knockdown of Dnm3os (the primary precursor of a miR-214-miR-199a cluster) leads to skeletal abnormalities such as craniofacial hypoplasia [79]. miR-26 contributes to neurogenesis and myogenesis [80] and is involved in rainbow trout embryonic development [62] whereas miR-451 has been found

to be involved in erythroid maturation in zebrafish [81]. Nineteen novel miRNA candidates were found to be differentially expressed between the two Arctic charr morphs in this study. Of those, two were only expressed in AC while one putative miRNA was only expressed in SB. These novel miRNA candidates were detected in most/all stages in one morph and not detectable in any of the stages in the other morph (Table 3), therefore it is unlikely that they represent a sequencing or technical error. As we used the Salmon genome to detect novel miRNAs, none of the novel miRNAs is likely to be morph or even Arctic charr specific, although expression differences can be expected. Several scenarios exist as to why these miRNAs are not expressed in both morphs, for example they might have been lost, their sequence might have been modified leading to the instability of the miRNA secondary structure or their expression repressed.

miRNAs are understudied in fishes and at present represent just a fraction of the miRNAs described in mammals. In the latest release of mirBase (version 20) from June 2013 there are only 255 mature miRNA sequences available for zebrafish, whereas 2578 mature sequences have been described in humans. Here we find 427 novel miRNA candidates, which are not Arctic charr specific. Some of these miRNAs are in highly conserved blocks (sal-nov235, 242, 334) among fishes, indicating that these miRNAs might have arisen early in the evolutionary history of fishes.

6.6 Concluding remarks and future directions

The theoretical underpinnings of our study are based on the general proposition that differences in the level, timing and pattern of miRNA expression or acquisition of new miRNAs can influence variation in developmental circuits, so as to generate diverse and possibly discrete morphological phenotypes, thereby creating substrate for natural selection to act upon. We use a system of two contrasting morphs of Arctic charr and as a first step we surveyed miRNA expression at four developmental stages thereby homing in on the miRNA genes that may have a bearing on the morphological and functional differences between the morphs. Differences in expression levels were detected in 72 miRNAs. Interestingly, the majority of these miRNAs (53/72) are evolutionarily stable and have been previously described as part of important developmental processes such as neurogenesis, erythropoiesis, skeleto- and myogenesis, specifically in craniofacial elements. Some miRNAs (e.g. the let-7 and miR430 families) show indications of differences in timing of expression. Other miRNAs (sal-miR-152 and sal-miR-451) exhibit sequence divergence. We are currently working on follow up experiments e.g. looking for the putative targets of the interesting miRNA candidates found in this study and defining their expression pattern using *in situ* hybridization in embryos derived from additional morphs and populations.

6.7 Acknowledgements

We acknowledge Bjarni K. Kristjánsson, Einar Svavarsson and Soizic Le Deuff for assisting with the sampling of parents, generation and maintenance of and sampling from embryo groups. We thank Jóhannes Guðbrandsson and Arnar Pálsson for discussions and advice on statistics. We also thank Valerie H. Maier and Ehsan Pashay Ahi for discussing the project at various stages. We thank Guðbjörg Þ. Örlygsdóttir, Steinunn Snorradóttir and Ólafur Þ. Magnússon at deCODE Genetics for technical support.

6.8 References

1. Plasterk RHA (2006) Micro RNAs in animal development. Cell 124: 877–881.

2. Mishima Y (2012) Widespread roles of microRNAs during zebrafish development and beyond. Dev Growth Differ 54: 55–65. doi:10.1111/j.1440-169X.2011.01306.x.

3. Li J, Zhang Z (2012) miRNA regulatory variation in human evolution. Trends Genet: 1–9. doi:10.1016/j.tig.2012.10.008.

4. Zhao Y, Srivastava D (2007) A developmental view of microRNA function. Trends Biochem Sci 32: 189–197. doi:10.1016/j.tibs.2007.02.006.

5. Berezikov E (2011) Evolution of microRNA diversity and regulation in animals. Nat Rev Genet 12: 846–860. doi:10.1038/nrg3079.

6. Bartel DP, Lee R, Feinbaum R (2004) MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. Genomics: The miRNA Genes. 116: 281–297.

7. Lewis BP, Shih I, Jones-Rhoades MW, Bartel DP, Burge CB (2003) Prediction of Mammalian MicroRNA Targets. Cell 115: 787–798.

8. Heimberg AM, Sempere LF, Moy VN, Donoghue PCJ, Peterson KJ (2008) MicroRNAs and the advent of vertebrate morphological complexity. Proc Natl Acad Sci U S A 105: 2946–2950. doi:10.1073/pnas.0712259105.

9. Tarver JE, Sperling EA, Nailor A, Heimberg AM, Robinson JM, et al. (2013) miRNAs: Small Genes with Big Potential in Metazoan Phylogenetics. Mol Biol Evol 30: 2369–2382.

10. Liu N, Okamura K, Tyler DM, Phillips MD, Chung W-J, et al. (2008) The evolution and functional diversification of animal microRNA genes. Cell Res 18: 985–996. doi:10.1038/cr.2008.278.

11. Saunders MA, Liang H, Li W-H (2007) Human polymorphism at microRNAs and microRNA target sites. Proc Natl Acad Sci U S A 104: 3300–3305.

12. Ason B, Darnell DK, Wittbrodt B, Berezikov E, Kloosterman WP, et al. (2006) Differences in vertebrate microRNA expression. Proc Natl Acad Sci U S A 103: 14385–14389. doi:10.1073/pnas.0603529103.

13. Hobert O (2008) Gene regulation by transcription factors and microRNAs. Science (80-) 319: 1785–1786. doi:10.1126/science.1151651.

14. Peterson KJ, Dietrich MR, McPeek MA (2009) MicroRNAs and metazoan macroevolution: insights into canalization, complexity, and the Cambrian explosion. Bioessays 31: 736–747.

15. Ebert MS, Sharp P a (2012) Roles for microRNAs in conferring robustness to biological processes. Cell 149: 515–524. doi:10.1016/j.cell.2012.04.005.

16. Wilson AJ, Gíslason D, Skúlason S, Snorrason SS, Adams CE, et al. (2004) Population genetic structure of Arctic charr, Salvelinus alpinus from northwest Europe on large and small spatial scales. Mol Ecol 13: 1129–1142. doi:10.1111/j.1365-294X.2004.02149.x.

17. Snorrason SS, Skúlason S (2004) Adaptive speciation in northern fresh water fishes —

patterns and processes. In: Dieckmann U, Metz H, Doebeli M, Tautz D, editors. Adaptive speciation. Cambridge University Press, Cambridge. pp. 210–228.

18. Robinson, B. W. and DS (2000) Natural selection and the evolution of adaptive genetic variation in northern freshwater fishes. Adaptive genetic variation in the wild. New York: Oxford University Press. pp. 65–94.

19. Schluter D (1993) Adaptive Radiation in Sticklebacks: Size, Shape, and Habitat Use Efficiency. Ecology 3: 699–709.

20. Schluter D, McPhail JD (1993) Character displacement and replicate adaptive radiation. Trends Ecol Evol 8: 197–200.

21. Brunner PC, Douglas MR, Osinov a, Wilson CC, Bernatchez L (2001) Holarctic phylogeography of Arctic charr (Salvelinus alpinus L.) inferred from mitochondrial DNA sequences. Evolution 55: 573–586.

22. Gíslason D, M Ferguson M, Skúlason S, S Snorrason S, Ferguson MM, et al. (1999) Rapid and coupled phenotypic and genetic divergence in Icelandic Arctic char (Salvelinus alpinus). Can J Fish Aquat Sci 56: 2229–2234.

23. Cresko WA, Amores A, Wilson C, Murphy J, Currey M, et al. (2004) Parallel genetic basis for repeated evolution of armor loss in Alaskan threespine stickleback populations. Proc Natl Acad Sci U S A 101: 6050–6055. doi:10.1073/pnas.0308479101.

24. Woods PJ, Skulason S, Snorrason SS, Kristjansson BK, Malmquist HJ, et al. (2012) Intraspecific diversity in Arctic charr, Salvelinus alpinus, in Iceland: I. Detection using mixture models.

25. Snorrason S., Skúlason S, Sandlund O., Malmquist H., Jonsson B, et al. (1989) Shape polymorphism in Arctic charr, Salvelinus alpinus, in Thingvallavatn, Iceland. Pysiological Ecol Japan 1: 393–404.

26. Jonsson B, Skúlason S, Snorrason SS, Sandlund OT, Malmquist HJ, et al. (1988) Life History Variation of Polymorphic Arctic Charr (Salvelinus alpinus) in Thingvallavatn, Iceland. Can J Fish Aquat Sci 45: 1537–1547.

27. Malmquist HJ, Snorrason SS, Skulason S, Jonsson B, Sandlund OT, et al. (1992) Diet differentiation in polymorphic Arctic charr in Thingvallavatn , Iceland. J Anim Ecol 61: 21–35. doi:10.2307/5505.

28. Sandlund, O.T., Gunnarson, K., Jonasson, P.M., Jonsson, B., Lindem, T., Magnusson, K.P., Malmquist, H.J., Sigurjonsdottir, H., Skulason, S. & Snorrason SS (1992) The Arctic charr Salvelinus alpinus in Thingvallavatn. Oikos 64: 305–351.

29. Skulason S, Noakes DL. G, Snorrason SS (1989) Ontogeny of trophic morphology in four sympatric morphs of arctic charr Salvelinus alpinus in Thingvallavatn, Iceland. Biol J Linn Soc 38: 281–301. doi:10.1111/j.1095-8312.1989.tb01579.x.

30. Frandsen F, Malmquist HJ, Snorrason SS (1989) Ecological parasitology of polymorphic Arctic charr, Salvelinus alpinus (L.), in Thingvallavatn, Iceland. J Fish Biol 34: 281–297.

31. Skúlason S, Snorrason SS, Ota D, Noakes DLG (1993) Genetically based differences in foraging behaviour among sympatric morphs of arctic charr (Pisces: Salmonidae). Anim Behav 45: 1179–1192.

32. Skúlason S, Snorrason SS, Noakes DLG, Ferguson MM (1996) Genetic basis of life history variations among sympatric morphs of Arctic char , Salvelinus alpinus. Can J Fish Aquat Sci 53: 1807–1813.

33. Kapralova KH, Morrissey MB, Kristjánsson BK, Olafsdóttir GÁ, Snorrason SS, et al. (2011) Evolution of adaptive diversity and genetic connectivity in Arctic charr (Salvelinus alpinus) in Iceland. Heredity (Edinb) 106: 472–487.

34. Kapralova KH, Gudbrandsson J, Reynisdottir S, Santos CB, Baltanás VC, et al. (2013) Differentiation at the MHCIIα and Cath2 Loci in Sympatric Salvelinus alpinus Resource Morphs in Lake Thingvallavatn. PLoS One 8: e69402.

35. Eiriksson GM, Skulason S, Snorrason SS (1999) Heterochrony in skeletal development and body size in progeny of two morphs of Arctic charr from Thingvallavatn, Iceland. J Fish Biol 55: 175–185.

36. Svavarsson E (2007) Árangur í kynbótum á bleikju og næstu skref [reference in icelandic]. Fræðaþing landbúnaðarins (conference proceedings) 4: 121–125.

37. Gorodilov YN (1996) Description of the early ontogeny of the Atlantic salmon, Salmo salar, with a novel system of interval (state) identification. Environ Biol Fishes 47: 109–127. doi:10.1007/BF00005034.

38. Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.jounal 17: 10–12.

39. Ewing B, Green P (1998) Base-Calling of Automated Sequencer Traces Using Phred. II. Error Probabilities. Genome Res 8: 186–194.

40. Burge SW, Daub J, Eberhardt R, Tate J, Barquist L, et al. (2013) Rfam 11.0: 10 years of RNA families. Nucleic Acids Res 41: D226–32.

41. Griffiths-Jones S (2004) The microRNA Registry. Nucleic Acids Res 32: D109–11. doi:10.1093/nar/gkh023.

42. Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ (2006) miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res 34: D140–4.

43. Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ (2008) miRBase: tools for microRNA genomics. Nucleic Acids Res 36: D154–8.

44. Kozomara A, Griffiths-Jones S (2011) miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res 39: D152–7.

45. Pearson WR, Lipman DJ (1988) Improved tools for biological sequence comparison. Proc Natl Acad Sci 85: 2444–2448.

46. Friedländer MR, Chen W, Adamidi C, Maaskola J, Einspanier R, et al. (2008) Discovering microRNAs from deep sequencing data using miRDeep. Nat Biotechnol 26: 407–415. doi:10.1038/nbt1394.

47. Di Génova A, Aravena A, Zapata L, González M, Maass A, et al. (2011) SalmonDB: a bioinformatics resource for Salmo salar and Oncorhynchus mykiss. Database (Oxford) 2011: bar050. doi:10.1093/database/bar050.

48. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403–410.

49. Gordon D, Abajian C, Green P (1998) Consed: A Graphical Tool for Sequence Finishing. Genome Res 8: 195–202. doi:10.1101/gr.8.3.195.

50. Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26: 139–140.

51. Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B 57: 289–300.

52. Mestdagh P, Van Vlierberghe P, De Weer A, Muth D, Westermann F, et al. (2009) A novel and universal method for microRNA RT-qPCR data normalization. Genome Biol 10: R64.

53. Ahi EP, Guðbrandsson J, Kapralova KH, Franzdóttir SR, Snorrason SS, et al. (2013) Validation of reference genes for expression studies during craniofacial development in arctic charr. PLoS One 8: e66389.

54. Mallarino R, Campàs O, Fritz JA, Burns KJ, Weeks OG, et al. (2012) Closely related bird species demonstrate flexibility between beak morphology and underlying developmental programs. Proc Natl Acad Sci U S A 109: 16222–16227.

55. Sylvester JB, Rich CA, Loh Y-HE, van Staaden MJ, Fraser GJ, et al. (2010) Brain diversity evolves via differences in patterning. Proc Natl Acad Sci U S A 107: 9718–9723.

56. Roberts RB, Hu Y, Albertson RC, Kocher TD (2011) Craniofacial divergence and ongoing adaptation via the hedgehog pathway. Proc Natl Acad Sci U S A 108: 13194–13199.

57. Manousaki T, Hull PM, Kusche H, Machado-Schiaffino G, Franchini P, et al. (2013) Parsing parallel evolution: ecological divergence and differential gene expression in the adaptive radiations of thick-lipped Midas cichlid fishes from Nicaragua. Mol Ecol 22: 650–669. doi:10.1111/mec.12034.

58. Henning F, Jones JC, Franchini P, Meyer A (2013) Transcriptomics of morphological color change in polychromatic Midas cichlids. BMC Genomics 14: 171.

59. Filteau M, Pavey SA, St-Cyr J, Bernatchez L (2013) Gene coexpression networks reveal key drivers of phenotypic divergence in lake whitefish. Mol Biol Evol 30: 1384–1396.

60. Yamamoto Y, Byerly MS, Jackman WR, Jeffery WR (2009) Pleiotropic functions of embryonic sonic hedgehog expression link jaw and taste bud amplification with eye loss during cavefish evolution. Dev Biol 330: 200–211.

61. Arif S, Murat S, Almudi I, Nunes MDS, Bortolamiol-Becet D, et al. (2013) Evolution of mir-92a Underlies Natural Morphological Variation in Drosophila melanogaster.

62. Ramachandra RK, Salem M, Gahr S, Rexroad CE, Yao J (2008) Cloning and characterization of microRNAs from rainbow trout (Oncorhynchus mykiss): their expression during early embryonic development. BMC Dev Biol 8: 41. doi:10.1186/1471-213X-8-41.

63. Ma H, Hostuttler M, Wei H, Rexroad CE, Yao J (2012) Characterization of the rainbow trout egg microRNA transcriptome. PLoS One 7: e39649.

64. Bekaert M, Lowe NR, Bishop SC, Bron JE, Taggart JB, et al. (2013) Sequencing and Characterisation of an Extensive Atlantic Salmon (Salmo salar L.) MicroRNA Repertoire. PLoS One 8: e70136.

65. Kim HK, Lee YS, Sivaprasad U, Malhotra A, Dutta A (2006) Muscle-specific microRNA miR-206 promotes muscle differentiation. J Cell Biol 174: 677–687.

66. Sweetman D, Goljanek K, Rathjen T, Oustanina S, Braun T, et al. (2008) Specific requirements of MRFs for the expression of muscle specific microRNAs, miR-1, miR-206 and miR-133. Dev Biol 321: 491–499.

67. Dey BK, Gagan J, Dutta A (2011) miR-206 and -486 induce myoblast differentiation by downregulating Pax7. Mol Cell Biol 31: 203–214. doi:10.1128/MCB.01009-10.

68. Inose H, Ochi H, Kimura A, Fujita K, Xu R, et al. (2009) A microRNA regulatory mechanism of osteoblast differentiation. Proc Natl Acad Sci U S A 106: 20794–20799. doi:10.1073/pnas.0909311106.

69. Chen J-F, Mandel EM, Thomson JM, Wu Q, Callis TE, et al. (2006) The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. Nat Genet 38: 228–233.

70. Liu N, Williams AH, Kim Y, McAnally J, Bezprozvannaya S, et al. (2007) An intragenic MEF2-dependent enhancer directs muscle-specific expression of microRNAs 1 and 133. Proc Natl Acad Sci U S A 104: 20844–20849.

71. Niwa R, Slack FJ (2007) The evolution of animal microRNA function. Curr Opin Genet Dev 17: 145–150. doi:10.1016/j.gde.2007.02.004.

72. Wienholds E, Kloosterman WP, Miska E, Alvarez-Saavedra E, Berezikov E, et al. (2005) MicroRNA expression in zebrafish embryonic development. Science 309: 310–311.

73. Brennecke J, Stark A, Russell RB, Cohen SM (2005) Principles of microRNA-target recognition. PLoS Biol 3: e85.

74. Giraldez AJ, Mishima Y, Rihel J, Grocock RJ, Van Dongen S, et al. (2006) Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. Science 312: 75–79. doi:10.1126/science.1122689.

75. Pasquinelli AE, Reinhart BJ, Slack F, Martindale MQ, Kuroda MI, et al. (2000) Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. Nature 408: 86–89.

76. McGlinn E, Yekta S, Mansfield JH, Soutschek J, Bartel DP, et al. (2009) In ovo application of antagomiRs indicates a role for miR-196 in patterning the chick axial skeleton through Hox gene regulation. Proc Natl Acad Sci U S A 106: 18610–18615.

77. He X, Yan Y-L, Eberhart JK, Herpin A, Wagner TU, et al. (2011) miR-196 regulates axial patterning and pectoral appendage initiation. Dev Biol 357: 463–477. doi:10.1016/j.ydbio.2011.07.014.

78. Hinits Y, Williams VC, Sweetman D, Donn TM, Ma TP, et al. (2011) Defective cranial skeletal development, larval lethality and haploinsufficiency in Myod mutant zebrafish. Dev Biol 358: 102–112.

79. Watanabe T, Sato T, Amano T, Kawamura Y, Kawamura N, et al. (2008) Dnm3os, a

non-coding RNA, is required for normal growth and skeletal development in mice. Dev Dyn 237: 3738–3748. doi:10.1002/dvdy.21787.

80. Wong CF, Tellam RL (2008) MicroRNA-26a targets the histone methyltransferase Enhancer of Zeste homolog 2 during myogenesis. J Biol Chem 283: 9836–9843.

81. Pase L, Layton JE, Kloosterman WP, Carradice D, Waterhouse PM, et al. (2008) miR-451 regulates zebrafish erythroid maturation in vivo via its target gata2. Blood 113: 1794–1804. doi:10.1182/blood-2008-05-155812.

82. Papaioannou G, Inloes JB, Nakamura Y, Paltrinieri E, Kobayashi T (2013) let-7 and miR-140 microRNAs coordinately regulate skeletal development. Proc Natl Acad Sci U S A 110: E3291–300.

6.9 Appendix

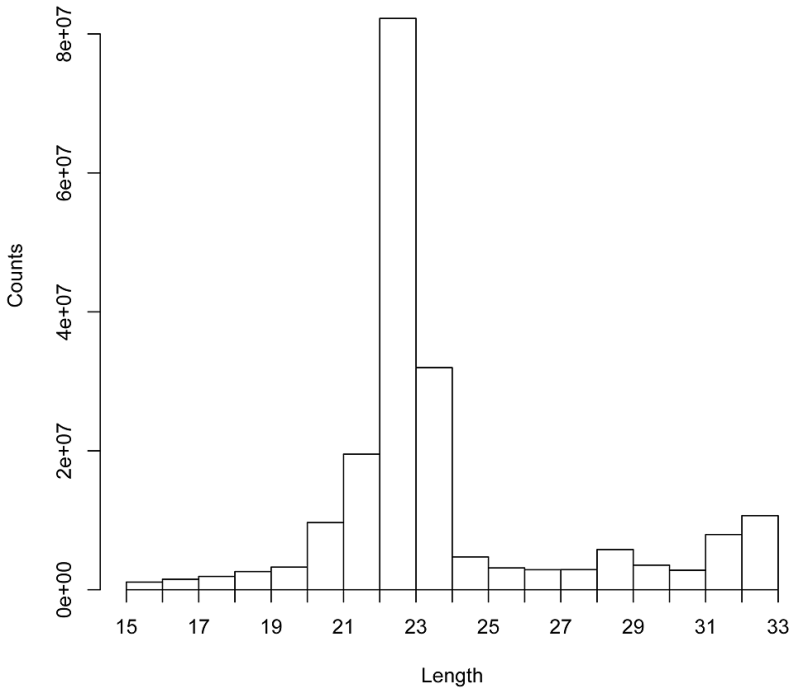

Figure S 6.1 - Length distribution of reads in the small-RNA-seq data for all samples combined. A major peak is observed at 22 nt, corresponding to the typical miRNA

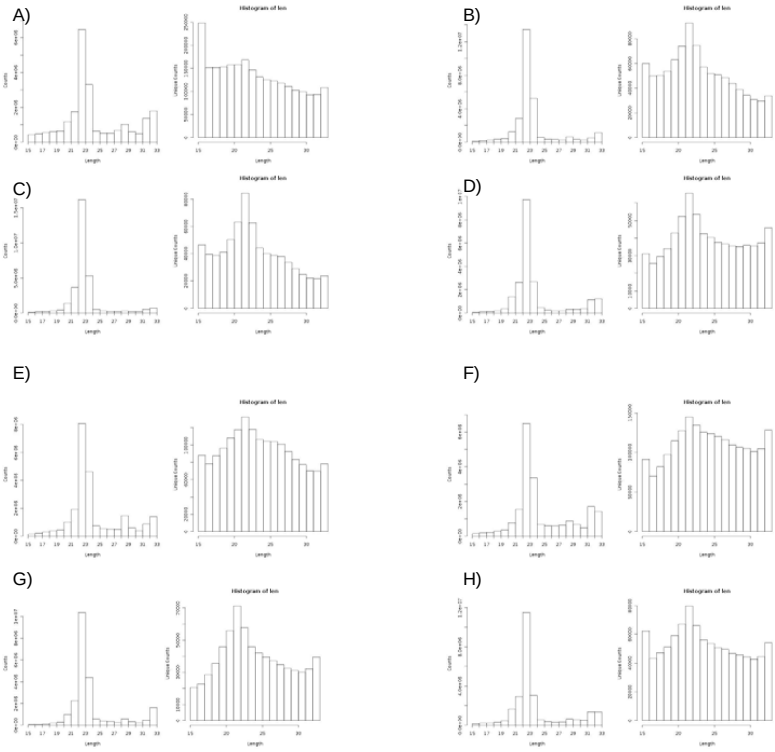

Figure S 6.2 Length distribution of reads in individual miRNA-seq samples. A-D: Small benthic (SB) stages 1-4; E-H: Aquaculture (AC) stages 1-4. Left panel: redundant reads, Right panel: unique reads.

Table S 6.1 Background information and primers used for amplification of selected miRNA clusters. The names of S. salar contigs and miRNAs in clusters used for primer design, sequence identity (%) between Arctic charr and Salmon, and forward and reverse primer sequences are shown.

Contig in *Salmo Salar*	miRNA cluster	Primer sequence (5'-3')	Identity %
AGKD01198198.1	20b-3p, 18b*, 19c	TGAGGATATTGCAGTTTTCTGAAC TAATCCGCCAGATGTTGTGA	94
AGKD01020894.1	133b, 133a	TTTTTCTTTCTCTCTTTTCTAAACAGG TGAATTTGACGTGATTCAGACAC	91
AGKD01079956.1	143-3p, 143-5p	CATTCCAAAACACCCCAAGT GGCCAGGGTAATGCAGTAAA	92
AGKD01039727.1	219-5p, 219-3p	GAGACATACTTTGAGCCCTTGC AGCACTAAGAGCCGCAAAAA	91
AGKD01040644.1	sal-nov-242	CCCTTTACACAAAGCACTCG TGTTTTGCACTGGTGTGAGA	96
AGKD01042615.1	sal-nov-334	AAATGACTTGGGTTTATTTTGTAGA TGCAAGCAGTATATTAAGAGGATTTG	92
AGKD01070756.1	sal-nov-235	AAGCTCATTCTCCATATCCAACA TCACCCACTGGACCAAAACT	98

Table S 6.2 - Conserved Arctic charr miRNAs. 326 known miRNAs were identified in the small-RNA-seq data. Included are miRNA names, miRNA sequences, number of raw reads per sample and names of miRNA orthologs.

ID	Sequence	SB1	SB2	SB3	SB4	AC1	AC2	AC3	AC4	Ortholog
let-7a-5p	TGAGGTAGTAGGTTGTATAG	43968	199133	334022	356101	61663	98284	246955	528993	has-let-7a-5p
let-7a-3p	CTGTACAACCTCCTAGCTTC	153	239	404	423	243	142	237	381	gga-let-7a-2-3p
let-7b	TGAGGTAGTAGGTTGTGTGG	148	2444	8220	6084	132	520	2445	10107	dre-let-7b
let-7c-5p	TGAGGTAGTAGGTTGTATGG	28482	98174	131230	43246	34956	43222	66206	65927	rno-let-7c-5p
let-7c-3p	CTGTACAACCTTCTAGCTTT	253	533	801	458	555	387	781	439	hsa-let-7c-3p
let-7d-5p	TGAGGTAGTTGGTTGTATGG	114842	462518	682690	261278	144404	158167	320932	367455	tni-let-7d
let-7d-3p	CGGTACAACCTTCTAGCTTT	39	84	93	36	111	103	153	58	pol-let-7d-3p
let-7e-5p	TGAGGTAGTAGATTGAATAG	3080	16244	59218	219783	3839	6877	55331	271976	aca-let-7e-5p
let-7e-3p	TAACTATACAATCTACTGTC	0	28	180	223	3	29	272	238	aca-let-7e-3p
let-7f-2-3p	CTATACAGTCTACTGTCTTT	23	25	44	45	64	22	65	59	mdo-let-7f-2-3p
let-7g	TGAGGTAGTAGTTTGTATAG	8377	15225	17507	39759	14953	15309	48190	57785	dre-let-7g
let-7g-3p	CTGTACAAGCCACTGCCTTG	26	47	57	222	42	42	109	274	mml-let-7g-3p
let-7h	TGAGGTAGTAAGTTGTGTTG	454	1118	1810	8992	650	890	2609	16017	ipu-let-7h
let-7i-5p	TGAGGTAGTAGTTTGTGCTG	169	611	1722	11066	250	342	1744	15766	aca-let-7i-5p
let-7j	TGAGGTAGTTGTTTGTACAG	184	456	636	1936	325	373	1104	2513	ccr-let-7j
let-7j-3p	CTATACAGTCTATTGCCTTC	5	26	54	177	7	40	100	218	gga-let-7j-3p
let-7k-5p	TGAGGGAGTAGATTGAATAG	7	41	155	1039	5	8	217	439	gga-let-7k-5p
miR-1-3p	TGGAATGTAAAGAAGTATGT	808598	2312903	3723254	4607787	924705	1058002	2436555	5046138	pol-miR-1-3p
miR-1-5p	TGGAATGTAAAGGAGTATGT	522	363	574	691	154	168	328	778	bbe-miR-1-5p
miR-100-5p	AACCCGTAGATCCGAACTTG	53588	106498	123200	97688	78350	51913	73667	91579	asu-miR-100a-5p
miR-100-3p	ACAAGCTCGTGTCTATAGGT	22	62	80	99	51	27	37	113	mmu-miR-100-3p
miR-101-3p	TACAGTACTGTGATAACTGA	36594	77892	131303	213593	53439	40370	92306	207826	cgr-miR-101b-3p
miR-101b	ACAGTACTATGATAACTGAA	90	105	97	92	78	54	70	107	tni-miR-101b
miR-101b-5p	CAGTTATCATGGTACCGGTG	235	356	361	153	259	164	132	145	ola-miR-101b-5p
miR-103	AGCAGCATTGTACAGGGCTA	11702	20444	27648	24667	15661	11389	16773	23276	eca-miR-103

miR-103-3p	CAGCATTGTACAGGGCTATG	130	178	276	148	165	120	176	150	gga-miR-103-3p
miR-106	AAAGTGCTTATAGTGCAGGT	1192	1456	1199	354	979	654	521	136	xtr-miR-106
miR-10a	TACCCTGTAGATCCGAATTT	62	92	96	24	130	37	59	22	bta-miR-10a
miR-10b	TACCCTGTAGAACCGAATTT	499652	812880	779435	224327	732238	416190	418152	215077	tni-miR-10b
miR-10b-3p	ACAGATTCGATTCTAGGGGA	397	583	486	132	657	387	533	135	pol-miR-10b-3p
miR-10c	TACCCTGTAGATCCGGATTT	217602	291504	286306	76349	370209	216238	184573	89525	ccr-miR-10c
miR-10c-5p	TACCCTGTAGAATCGAATTT	954	125	121	37	112	65	57	25	aca-miR-10c-5p
miR-10d	ACCCTGCAGAACCGAATTTG	348	127	115	38	106	67	72	12	fru-miR-10d
miR-122-5p	TGGAGTGTGACAATGGTGTT	45163	86851	95296	38243	27321	15547	15583	53689	mdo-miR-122-5p
miR-122-3p	AACGCCATTATCACACTAAA	40	53	92	56	91	26	82	66	gga-miR-122-3p
miR-124-3p	TAAGGCACGCGGTGAATGCC	783	1525	1367	7246	1931	1371	3135	13813	bmo-miR-124
miR-124b	CGTGTTCACGGCGGACCTTG	416	785	1264	282	729	282	408	284	ccr-miR-124b
miR-124-5p	CGTGTTCACAGCGGACCTTG	634	1042	1503	617	1710	735	1246	746	pol-miR-124-5p
miR-125a	TCCCTGAGACCCTTAACCTG	6133	8555	24464	5872	8593	9484	8611	9334	dre-miR-125a
miR-125a-3p	ACAGGTGAGGTCCTCGGGAA	1497	2442	2788	551	2374	990	1137	490	ola-miR-125a-3p
miR-125b-3p	ACGGGTTAGGCTCTTGGGAC	47903	99776	108454	31348	79829	39124	59095	34454	hsa-miR-125b-1-3p
miR-125c	TCCCTGAGACCCTAACTTGT	14299	17632	50319	16172	20354	17199	19163	28961	ola-miR-125c
miR-126-3p	CGTACCGTGAGTAATAATGC	1311	1420	1710	1134	1698	891	859	1354	gga-miR-126-3p
miR-126-5p	CATTATTACTTTTGGTACGC	9256	8947	7481	5148	12747	6825	6595	6766	bta-miR-126-5p
miR-126-5p	ATTATTACTTTTGGTACGCG	534	556	412	420	766	476	599	437	dre-miR-126a-5p
miR-1260	ATCCCACCGCTGCCACCA	696	782	562	267	1530	430	463	300	cgr-miR-1260
miR-128	TCACAGTGAACCGGTCTCTT	14628	16001	17064	10608	18059	13777	15993	9475	sla-miR-128
miR-128-5p	CGGGGCCGGGGCACTGTCTG	38	84	128	18	21	28	45	54	hsa-miR-128-1-5p
miR-129-5p	CTTTTTGCGGTCTGGGCTTG	861	4290	9195	9028	1005	1092	2137	14442	hhi-miR-129b-5p
miR-129-3p	AAGCCCTTACCCCAAAAAGC	54	92	113	775	81	150	337	1465	ppy-miR-129-2-3p
miR-130a-3p	CAGTGCAATGTTAAAAGGGC	2809	3136	4107	3465	5334	3573	5456	3995	mmu-miR-130a-3p
miR-130a-5p	GCCCTTTTACTGCTGTACTA	165	328	479	56	201	114	95	55	gga-miR-130a-5p
miR-130b	CAGTGCAACAATGAAAGGGC	44803	49103	54941	41954	89702	51880	81628	68908	dre-miR-130b
miR-130b-5p	ACTCTTTCCCTGTTGCACT	38	40	41	33	65	51	49	30	mmu-miR-130b-5p

miR-130c	CAGTGCAATATTAAAAGGGC	34788	40211	49261	25207	65759	40351	65735	42031	dre-miR-130c
miR-132-3p	TAACAGTCTACAGCCATGGT	153	409	653	765	140	186	292	819	cfa-miR-132
miR-132-5p	ACCGTGGCTTTAGATTGTTA	67	213	467	918	64	54	150	887	ipu-miR-132a
miR-133-3p	TTTGGTCCCCTTCAACCAGC	3566	3157	4075	12774	8196	9187	16925	15285	spu-miR-133
miR-133-5p	AGCTGGTAAAATGGAACCAA	2470	6056	8826	7000	4222	4259	6806	7491	gga-miR-133a-5p
miR-135a-3p	ATATAGGGATGGAAGCCATG	414	653	607	392	879	472	681	485	mdo-miR-135a-3p
miR-135a-5p	TATGGCTTTTTATTCCTATG	12	68	148	152	27	33	132	197	gga-miR-135a-5p
miR-135b-3p	TATGGCTTTTTATTCCTATC	57470	80529	82182	33641	94876	45365	59058	27394	tni-miR-135b
miR-135b-5p	TATGGCTTTTTATTCCTATA	194	247	327	270	221	128	174	241	oan-miR-135b-5p
miR-135c	TATGGCTTTCTATTCCTATG	2139	4228	5673	4057	3756	3899	6947	4052	ipu-miR-135c
miR-137-3p	TTATTGCTTAAGAATACGCG	13293	29192	36348	14655	16293	16842	29154	11678	cgr-miR-137-3p
miR-137-5p	ACGGGTATTCTTGGGTGTAT	878	2877	1952	389	1074	845	1242	391	gga-miR-137-5p
miR-138	AGCTGGTGTTGTGAATCAGG	2026	5483	10170	10986	2530	2959	7325	13040	ppy-miR-138
miR-138-2-3p	GCTATTTCACAACACCAGGG	53	57	53	50	51	39	67	62	hsa-miR-138-2-3p
miR-1388-3p	ATCTCAGGTTCGTCAGCCCA	1774	3498	4313	6443	1816	1091	1559	5811	ipu-miR-1388
miR-1388-5p	AGGACTGTCCAACTTGAGAA	1015	1421	1379	1492	1079	758	743	1553	ola-miR-1388-5p
miR-139-3p	TGGAGATGCTGCTCTGTTGG	164	322	454	197	222	148	213	205	tgu-miR-139-3p
miR-139-5p	TCTACAGTGCATGTGTCTCC	432	405	529	556	625	711	820	629	aca-miR-139-5p
miR-140-3p	TACCACAGGGTAGAACCACG	30365	70246	142747	216951	56094	59644	116669	255118	dre-miR-140-3p
miR-140-5p	CAGTGGTTTTACCCTATGGT	10811	26302	45983	51596	18063	19430	44955	69429	cgr-miR-140-5p
miR-142-3p	TGTAGTGTTTCCTACTTTAT	4015	4062	3831	8170	3216	2308	2996	7344	eca-miR-142-3p
miR-142-5p	CATAAAGTAGAAAGCACTAC	145	142	211	616	153	115	206	580	tni-miR-142b
miR-143-3p	TGAGATGAAGCACTGTAGCT	106262	219013	254815	431009	177107	132284	247463	560637	bta-miR-143
miR-143-5p	GGTGCAGTGCTGCATCTCTG	312	645	737	539	517	298	335	638	oan-miR-143-5p
miR-144-3p	CTACAGTATAGATGATGTAC	714	1038	3665	749	662	975	2626	1430	ptr-miR-144
miR-144-5p	AGGATATCATCATATACTGT	7420	9692	14085	3069	12004	10142	12066	2837	rno-miR-144-5p
miR-145-5p	GTCCAGTTTTCCCAGGAATC	846	545	803	5675	1986	2435	2348	7001	ppy-miR-145
miR-145-3p	GGATTCCTGGAAATACTGTT	13153	35244	28255	13076	15872	10518	11406	9812	mml-miR-145-3p
miR-146a	TGAGAACTGAATTCCATAGA	38758	80403	100751	97315	39182	29504	35845	64425	ola-miR-146a-5p

miR-146b	TGAGAACTGAATTCCAAGGG	1871	2906	3188	2322	1834	1156	1259	1883	ipu-miR-146b
miR-148	TCAGTGCATTACAGAACGTT	6884	12431	19269	16457	8036	5821	10744	16861	fru-miR-148
miR-148-5p	AAGTTCTGTGATACACTTAG	683	1854	3437	1995	658	599	784	1753	mml-miR-148b-5p
miR-150	TCTCCCAATCCTTGTACCAG	187	79	48	246	297	284	186	616	dre-miR-150
miR-152	TCAGTGCATAACAGAACTTT	39201	89006	162195	164218	62535	46885	98817	188809	ola-miR-152
miR-152-5p	AAGTTCTGTGATACACTTCG	1798	4175	6127	3921	2270	1881	3448	3536	hsa-miR-152-5p
miR-153-3p	TTGCATAGTCACAAAAGTGA	3169	3133	2634	3199	7787	6061	10578	4050	pma-miR-153-3p
miR-153-5p	ATTTTTGTGATGTTGCAGCT	2167	4320	3660	875	2942	1773	1859	857	oan-miR-153-2-5p
miR-155	TTAATGCTAATCGTGATAGG	29	34	64	130	38	25	72	172	ipu-miR-155
miR-1595-3p	AGCCTCCTGGTCCTACTGCT	103	118	121	62	126	86	73	58	gga-miR-1595-3p
miR-15a-3p	CAGGCCGTACTGTGCTGCCG	57	111	163	77	75	44	58	93	dre-miR-15a-3p
miR-15a-5p	TAGCAGCACAGAATGGTTTG	1574	1630	2153	2863	2490	2287	3310	3069	dre-miR-15a-5p
miR-15b	TAGCAGCGCATCATGGTTTG	3511	3017	2656	2100	4852	2880	3072	1887	fru-miR-15b
miR-15c	TAGCAGCACGTCATGGTTTG	9101	9713	10720	3538	14252	9820	11547	4811	meu-miR-15c
miR-15c-5p	GTAGCAGCACGTCATGGTTT	84	74	67	18	95	72	84	24	gga-miR-15c-5p
miR-16-5p	TAGCAGCACGTAAATATTGG	57859	70896	68503	48619	70958	51848	59651	52813	pma-miR-16-5p
miR-16c-5p	TAGCAGCACGTAAATACTGG	12943	12743	12485	4591	18158	11898	13562	5325	xtr-miR-16c
miR-1692	CCTTCGATAGCTCAGTTGGT	1875	5184	7286	4237	6418	11713	6844	8062	gga-miR-1692
miR-17-5p	CAAAGTGCTTACAGTGCAGG	69302	85782	76525	26091	128386	78017	83363	27176	mne-miR-17-5p
miR-17-3p	ACTGCAGTGGAGGCACTTCT	57	53	46	13	94	54	87	15	ccr-miR-17-3p
miR-1788-3p	CAGGCAGCTAAAGCAAGTCT	461	1592	2730	1430	447	778	1698	1383	aca-miR-1788-3p
miR-18b-5p	TAAGGTGCATCTAGTGTAGT	23515	35195	30493	7412	32256	19230	19939	7278	hsa-miR-18b-5p
miR-18a-5p	TAAGGTGCATCTAGTGCAGA	19936	24050	16874	4444	29616	19434	17962	5198	rno-miR-18a-5p
miR-18a-3p	ACTGCCCTAAGTGCTCCTTC	41	58	28	8	67	60	60	14	mdo-miR-18a-3p
miR-18b-3p	ACTGCCCTAGTTGCTCCTTC	50	43	49	14	140	76	81	7	mdo-miR-18b-3p
miR-181a-5p	AACATTCAACGCTGTCGGTG	220537	493750	749490	434815	296178	211783	499454	355550	bta-miR-181a
miR-181a-3p	ACCATCGACCGTTGACTGTG	25485	58650	99819	26256	50651	19074	31146	17849	mdo-miR-181a-2-3p
miR-181b-5p	AACATTCATTGCTGTCGGTG	159462	418164	586975	165505	194864	141272	292426	137632	ola-miR-181b-5p
miR-181b-3p	ACTCGCTGAACAATGAATGC	331	606	853	508	445	303	486	342	rno-miR-181b-1-3p

miR-181c	CACATTCATCGCTGTCGGTG	10067	18952	18672	1567	13124	10292	11271	1669	ipu-miR-181c
miR-181c-3p	AACCATCGACCGTTGAGTGT	272	557	873	322	345	284	284	235	hsa-miR-181c-3p
miR-182-5p	TTTGGCAATGGTAGAACTCA	58966	159293	166847	258268	80093	53698	132902	248806	rno-miR-182
miR-182-3p	TGGTTCTAGACTTGCCAACT	46	58	84	155	56	42	62	166	ccr-miR-182-3p
miR-183-5p	TATGGCACTGGTAGAATTCA	30539	90131	135716	166939	36513	27771	55737	174639	bbe-miR-183-5p
miR-183-3p	TGAATTACCATAGGGCCATA	910	1307	1677	1365	1077	900	1006	1163	mdo-miR-183-3p
miR-184-3p	TGGACGGAGAACTGATAAGG	213794	247395	231272	172449	306525	179373	235340	205753	xtr-miR-184
miR-184-5p	TTATCACTTTTCCAGCCCAG	325	345	451	362	693	469	647	369	gga-miR-184-5p
miR-187-3p	TCGTGTCTTGTGTTGCAGCC	181	348	874	684	226	177	238	888	ccr-miR-187
miR-187-5p	GGCTGCAACACAGGACATGG	131	273	442	478	146	104	151	687	gga-miR-187-5p
miR-1889	ACATTACAGACTGGTATCTC	19	46	166	138	25	32	112	140	aga-miR-1889
miR-190a-5p	TGATATGTTTGATATATTAG	7003	8864	7777	6468	10225	6706	8712	5828	aca-miR-190a-5p
miR-190b	TGATATGTTTGATATTCGGT	874	2356	4404	4304	1247	1754	4739	4282	ola-miR-190b
miR-192	ATGACCTATGAATTGACAGC	124737	223978	247767	741275	181521	103806	189877	898410	bta-miR-192
miR-192-3p	CCTGTCAGTTATGTAGGCCA	105	301	459	1080	188	153	328	1959	ola-miR-192-3p
miR-193a-3p	AACTGGCCTACAAAGTCCCA	30	32	50	147	68	53	93	174	bta-miR-193a-3p
miR-193a-5p	TGGGTCTTTGCGGGCAAGGT	117	285	386	168	141	100	167	234	mmu-miR-193a-5p
miR-194-3p	CCAGTGGAGCTGCTGTTATG	229	872	1298	1040	197	130	259	1324	oan-miR-194-3p
miR-194-5p	TGTAACAGCAACTCCATGTG	15071	17501	17407	43899	19183	12170	15442	47000	pma-miR-194-5p
miR-1940	TTGGAGGATTGAGAAGGTGT	433	199	77	124	1093	836	550	293	mmu-miR-1940
miR-1957a	GGGGGTATAGCTCAGTGGTA	3549	1782	537	3156	4144	2988	5904	3354	mmu-miR-1957a
miR-1957b	GGGGATGTAACTCAGTGGTA	72	26	12	21	51	67	54	57	mmu-miR-1957b
miR-196a	TAGGTAGTTTCATGTTGTTG	95036	178935	157972	28487	86687	53121	50144	34575	ipu-miR-196a
miR-196b	TAGGTAGTTTCAAGTTGTTG	47537	94261	83663	23821	60226	46840	65588	26232	dre-miR-196b
miR-196c	TAGGTAGTTTCACGTTGTTG	532	197	155	39	114	53	67	31	aca-miR-196c-5p
miR-1973	ATTTTGACCGTGCGAAGGTA	422	13	25	52	104	562	35	82	cgr-miR-1973
miR-1989	TGAATAGAGCTGACATGATT	113	111	63	6	68	37	15	4	cte-miR-1989
miR-199a-5p	CCCAGTGTTCAGACTACCTG	28224	50342	68537	42142	44942	40762	60831	58713	ppy-miR-199a
miR-199a-3p	ACAGTAGTCTGCACATTGGT	15129	21046	20689	43599	28153	24568	41576	60234	bta-miR-199a-3p

miR-199b-5p	CAGTGTTCAGACTACCTGTT	142	262	314	227	247	198	260	282	mdo-miR-199b-2-5p
miR-199b-3p	CAGTAGTCTGCACATTGGTT	3522	3454	3344	8547	6376	5294	8329	11738	pma-miR-199b-3p
miR-19a	TTGTGCAAATCTATGCAAAA	3817	3888	3258	1783	6939	3672	4154	1414	mne-miR-19a
miR-19a-5p	AGTTTTGCATAGTTGCACTA	37	43	34	7	70	42	25	13	hsa-miR-19a-5p
miR-19b	TGTGCAAACCCAAGCAAAAC	30981	28499	20287	10588	55142	33595	30626	9485	sha-miR-19b
miR-19b-5p	AGTTTTGCAGGTTTGCTTTC	91	104	77	31	368	161	217	42	pma-miR-19b-5p
miR-19d	TGTGCAAACCCATGCAAAAC	501	432	315	163	1581	1118	835	287	ccr-miR-19d
miR-20a-5p	TAAAGTGCTTATAGTGCAGG	196902	239532	193121	57944	301635	180676	163958	58266	xtr-miR-20a-5p
miR-20b	CAAAGTGCTCACAGTGCAGG	17522	23612	19138	4124	29063	18461	17919	4768	ssc-miR-20b
miR-200a-3p	TAACACTGTCTGGTAACGAT	40381	60716	71904	67058	47590	37438	48793	51893	oan-miR-200a-3p
miR-200a-5p	CATCTTACCTGACAGTGCTG	1586	3374	4056	1585	1651	840	1297	1272	mml-miR-200a-5p
miR-200b	TAATACTGCCTGGTAATGAT	28836	60757	103573	107162	40180	33278	51003	82140	cgr-miR-200b
miR-200c	TAACACTGTCTGGTAATGAT	10695	23589	39110	39859	12291	11786	17118	33012	bfl-miR-200c
miR-202-5p	TTCCTATGCATATACCGCTT	20	20	53	61	51	53	60	81	dre-miR-202-5p
miR-203a-3p	GTGAAATGTTTAGGACCACT	53343	77847	112786	88273	84603	60800	85675	74605	pma-miR-203a-3p
miR-203b-5p	TGAAATGTTTAGGACCACT	1063	1751	2810	1491	958	722	1244	1046	rno-miR-203b-5p
miR-204	TTCCCTTTGTCATCCTATGC	5151	3964	5263	7060	7535	8496	8069	8578	gga-miR-204
miR-205-5p	TCCTTCATTCCACCGGAGTC	3389	3116	6519	3160	3311	2483	1694	4370	pma-miR-205a-5p
miR-205-3p	ATTCCCAGTGGGATGAAGCT	76	81	94	99	54	33	43	30	mmu-miR-205-3p
miR-206-3p	TGGAATGTAAGGAAGTGTGT	1906021	4854337	5536080	1468874	1899574	1833563	3402216	2260072	ssc-miR-206
miR-206-5p	ACATGCTTCCTTATATCCCC	591	864	1377	811	1962	1808	2970	1025	pol-miR-206-5p
miR-21	TAGCTTATCAGACTGGTGTT	1243080	1862527	1322089	845043	1533429	977044	891452	1053130	ccr-miR-21
miR-21-3p	CGACAACAGTCTGTAGGCTG	1381	2164	1301	766	1583	709	742	670	pol-miR-21-3p
miR-21c	TAGCTTATCAGACTGG	146	139	66	88	165	85	99	67	mmu-miR-21c
miR-210	CTGTGCGTGTGACAGCGGCT	781	1613	3601	2652	2019	1409	3880	6346	xtr-miR-210
miR-210-5p	AGCCACTGACTAACGTACAT	106	172	106	203	253	168	334	274	dre-miR-210-5p
miR-212-3p	TAACAGTCTACAGTCATGGC	8	14	25	247	21	26	77	433	aca-miR-212-3p
miR-212-5p	ACCTTGGCTTTAGACTGCTT	20	42	90	437	26	26	51	407	rno-miR-212-5p
miR-214-3p	ACAGCAGGCACAGACAGGCA	628	927	1289	1098	1095	1103	2190	1697	tgu-miR-214-3p

miR-214-5p	GCCTGTCTACACTTGCTGTG	988	1955	3348	1680	1562	1303	1616	2307	hsa-miR-214-5p
miR-215	ATGACCTATGAATTGACAG	1475	2256	2556	9100	1417	935	1671	8300	mdo-miR-215
miR-216	TAATCTCAGCTGGCAACTGT	14820	40500	59438	13828	26697	28225	48391	16421	ssc-miR-216
miR-216b	TAATCTCTGCAGGCAACTGT	11233	25905	40831	12155	19925	15760	25222	11640	fru-miR-216b
miR-217	TACTGCATCAGGAACTGATT	108944	216338	271702	83819	178954	134670	208027	77735	cfa-miR-217
miR-2187	TTAATTAGTATAGCCTGTAT	36	89	151	463	69	49	137	458	ipu-miR-2187
miR-2188-3p	GCTGTGTGAGGTCGGACCTA	1511	10617	20949	1703	1590	1644	3033	1962	dre-miR-2188-3p
miR-2188-5p	AAGGTCCAACCTCACATGTC	750	1027	1642	1492	1615	2612	5422	1963	tgu-miR-2188-5p
miR-2189	TTGATTATTTGAATCGGCTG	1127	857	820	312	1104	816	542	252	dre-miR-2189
miR-218b	TTGTGCTTGATCTAACCATG	6844	10921	13265	6303	10633	9890	10608	6993	pma-miR-218b
miR-219a	AGAATTGTGCATGGACATCT	28732	23183	11987	3907	59344	12735	5527	3152	ipu-miR-219a
miR-219-1-3p	AGAATTGTATCTGGACATCT	6293	5597	3199	786	16670	3157	2088	682	aca-miR-219-1-3p
miR-219b	GGAGTTGTGGATGGACATCA	3450	2729	1164	48	5982	1898	834	72	ipu-miR-219b
miR-22a	AAGCTGCCAGCTGAAGAACT	18586	21996	31884	162354	21988	16716	25331	227843	ccr-miR-22a
miR-22-5p	AGTTCTTCACTGGCAAGCTT	852	860	1024	2829	1463	1021	1668	3275	pol-miR-22-5p
miR-221-3p	AGCTACATTGTCTGCTGGGT	9154	8452	8229	7276	21110	12597	13839	6411	tgu-miR-221-3p
miR-221-5p	ACCTAGCATACAATGTAGAT	3358	9701	10122	6223	5733	1764	4147	4129	tgu-miR-221-5p
miR-222-3p	AGCTACATCTGGCTACTGGG	29736	31788	22712	7420	37644	19791	16403	6111	cgr-miR-222-3p
miR-222-5p	TGCTCAGTAGGCAGTGTAGA	4953	6358	9373	4548	8359	4494	4787	4711	aca-miR-222-5p
miR-222a	TGCTCAGTAGTCAGTGTAGA	89674	130146	116544	16081	156325	55638	37830	19191	ipu-miR-222a
miR-222a-5p	CTCAGTAGTCAGTGTAGAT	144	193	159	26	271	79	113	20	oan-miR-222a-5p
miR-222b-3p	AGCTACATCTGATTACTGGG	286	475	486	656	145	92	154	529	gga-miR-222b-3p
miR-223	TGTCAGTTTGTCAAATACCC	813	532	600	1371	493	505	751	1325	bta-miR-223
miR-223-5p	AGTGTATTTGACAAGCTGAG	619	1012	1066	1992	175	151	162	1533	mdo-miR-223-5p
miR-2281-3p	ATCTGTTATTGCTGTATTGC	70	153	209	95	47	42	52	53	dme-miR-2281-3p
miR-23a-3p	ATCACATTGCCAGGGATTTC	2596	3070	3705	11641	3558	4004	4411	13477	oan-miR-23a-3p
miR-23b-3p	ATCACATTGCCAGGGATTAC	3945	5149	6504	9492	5382	5315	6617	8790	bta-miR-23b-3p
miR-24-3p	TGGCTCAGTTCAGCAGGAAC	13494	14523	15270	16268	19639	16390	18710	19919	cfa-miR-24
miR-24-5p	TGCCTACTGAACTGGTATCA	46	44	53	89	60	55	70	97	fru-miR-24-5p

miR-24b-5p	TGCCTACTGAGCTGATAACA	223	248	398	287	426	382	427	372	ola-miR-24b-5p
miR-2404	AAGCACTGCATGGTATCTGC	31	21	5	5	127	42	105	4	bta-miR-2404
miR-2478	ATCCCACTTCTGACACCA	27442	42978	22209	6041	42033	15079	17960	4639	bta-miR-2478
miR-2487	CTAAGGGCTGGGTCGGTCGG	218	219	76	32	150	180	22	76	bta-miR-2487
miR-2488	GAGGGAATATGATCAGAGTG	9282	9135	1434	894	18652	16815	27137	1348	bta-miR-2488
miR-25-3p	CATTGCACTTGTCTCGGTCT	21980	37470	44933	12199	39952	23322	29928	13995	ipu-miR-25
miR-25-5p	AGGCGGAGACTTGGGCAATT	105	109	92	31	129	48	41	20	cgr-miR-25-5p
miR-26a	TTCAAGTAATCCAGGATAGG	37081	44686	48124	82286	71578	52647	82838	94506	ipu-miR-26a
miR-26a-3p	CCTATTCTTGATTACTTGTT	10	19	26	109	45	18	39	189	mml-miR-26a-2-3p
miR-26b	TTCAAGCAATCCAGGATAGG	127	36	26	52	52	46	46	49	ipu-miR-26b
miR-2779	TTAGGTCGCTGGTTCATCCG	88799	21129	10380	74410	96386	80895	24732	27523	bmo-miR-2779
miR-27a-3p	TTCACAGTGGCTAAGTTCCG	2610	4359	6275	7809	4842	4096	6154	11975	ipu-miR-27a
miR-27b-3p	TTCACAGTGGCTAAGTTCTG	22996	40394	71638	53040	35802	27507	48649	55107	tni-miR-27b
miR-27b-5p	AGAGCTTAGCTGATTGGTGA	619	1341	1703	704	874	529	917	590	rno-miR-27b-5p
miR-27c-3p	TTCACAGTGGTTAAGTTCTG	2008	3435	5414	10604	3389	2723	4907	16323	ccr-miR-27c-3p
miR-27c-5p	CAGGACTTAGCACACATGTG	545	1366	1453	1899	959	384	756	1883	ola-miR-27c-5p
miR-27d-3p	TTCACAGTGGCTAAGTTC	35	90	205	202	61	44	82	244	ola-miR-27d-3p
miR-27e	TTCACAGTGGCTAAGTTCAG	1151	1719	2875	4820	1833	1532	2298	6111	tni-miR-27e
miR-2898	TCCCCGGCATCTCCACCA	632	365	591	222	617	328	172	346	bta-miR-2898
miR-2995	TGGCACTGTTCGTAACCTGG	143	39	37	38	364	217	118	27	tgu-miR-2995
miR-29a-5p	ACTGATTTCTTCTGGTGTTT	156	470	828	469	148	86	158	344	cgr-miR-29a-5p
miR-29a-3p	TAGCACCATTTTAAATCAGT	80	87	112	45	166	179	229	63	pma-miR-29a-3p
miR-29b	TAGCACCATTTGAAATCAGT	76	86	101	585	166	168	237	888	ssc-miR-29b
miR-29c	CTGGTTTCACATGGTGGTTT	33	139	216	81	49	24	39	76	ipu-miR-29c
miR-29c-3p	TAGCACCATTTGAAATCGGT	79	79	108	271	137	105	149	452	mmu-miR-29c-3p
miR-301a	CAGTGCAATAGTATTGTCAA	1611	2060	2198	1573	2687	2000	2984	2145	bta-miR-301a
miR-301b-3p	CAGTGCAATAGTATTGTC	66	44	36	21	80	45	51	19	ola-miR-301b-3p
miR-301b-5p	GCTTTGACAATGTTGCACTA	1027	969	719	224	2022	1113	945	254	ola-miR-301b-5p
miR-301c	CAGTGCAATAGTATTGTCAT	19626	25644	27072	10262	32391	22999	35770	16037	ipu-miR-301c

miR-30a-3p	CTTTCAGTCGGATGTTTGCA	51355	96104	126957	62780	82396	44440	58348	71393	ptr-miR-30a-3p
miR-30b-5p	TGTAAACATCCTACACTCAG	5857	3611	3626	4969	14491	12608	13445	5992	ppy-miR-30b
miR-30c-5p	TGTAAACATCCTACACTCTC	27457	28008	25963	32818	51579	44341	48268	32133	tgu-miR-30c-5p
miR-30c-3p	CTGGGAGAAGGCTGTTTACT	317	355	356	81	313	131	154	79	tgu-miR-30c-3p
miR-30d-5p	TGTAAACATCCCCGACTGGA	105440	107644	109878	116181	198318	106241	131489	109785	cfa-miR-30d
miR-30d-3p	CTTTCAGTAGGGTGTTTGCT	219	494	692	182	460	197	317	253	ipu-miR-30d
miR-30e-5p	TGTAAACATCCTTGACTGGA	189102	243103	220526	225923	305867	194160	257746	216754	sha-miR-30e
miR-3120-5p	TACAGCAGGCACAGACAG	42	69	88	167	40	16	33	161	hsa-miR-3120-5p
miR-3126	TCAGTAACTGGAATCTGTCC	1267	3314	3986	5621	1217	694	1019	2924	ggo-miR-3126
miR-3294	TTGCTGGTGATACTGTCTGT	98	93	78	12	64	32	22	13	bmo-miR-3294
miR-33-3p	CAATGTTCCTGCAGTGCAAG	322	670	1330	1225	842	540	968	1334	ola-miR-33
miR-33-5p	GTGCATTGTAGTTGCATTGC	656	914	942	359	1041	662	728	345	tca-miR-33-5p
miR-3306	CTCAAAGGATGACAGATTTT	377	286	250	175	429	361	294	146	bmo-miR-3306
miR-338-3p	TCCAGCATCAGTGATTTTGT	544	780	1149	3123	988	934	1564	4089	mdo-miR-338
miR-338-5p	AACAATATCCTGGTGCTGTA	181	404	453	129	265	184	246	177	ola-miR-338-5p
miR-344b-5p	AGAATAATGCCAGCAGTCGG	52	103	133	190	115	98	150	213	rno-miR-344b-5p
miR-34a-5p	TGGCAGTGTCTTAGCTGGTT	821	347	443	694	423	346	329	841	gga-miR-34a-5p
miR-3618	TGATTTCCAATAATTGAGAC	1975	2363	1741	805	2207	1647	1387	588	ipu-miR-3618
miR-365-5p	AGGGACTTTTAGGGGCAGCT	892	1752	4065	2358	1167	751	1573	2031	mmu-miR-365-2-5p
miR-365-3p	TAATGCCCCTAAAAATCCTT	173	355	461	1321	396	482	806	3158	tgu-miR-365-3p
miR-375-3p	TTTGTTCGTTCGGCTCGCGT	14301	31532	49042	12278	21952	15814	24661	10264	bbe-miR-375-3p
miR-3963	ATCCCACTTCTGACAC	105	68	53	6	91	71	25	2	mmu-miR-3963
miR-4015-3p	GGATAGGTTGATTTCATTGC	165	147	154	13	267	163	150	38	cin-miR-4015-3p
miR-4185-3p	ATGCTTGTTGTATTCATACT	115	245	338	194	326	323	233	91	cin-miR-4185-3p
miR-4286	ACCCCACTCCTGGTACCA	46	13	13	22	56	46	28	37	bta-miR-4286
miR-429-3p	TAATACTGTCTGGTAATGCC	12337	22078	26182	15219	12290	11788	9887	12486	gga-miR-429-3p
miR-429-5p	GTCTTACCAGACATGGTTAG	114	178	121	33	85	75	67	28	mmu-miR-429-5p
miR-430a	TAAGTGCTACTTGTTGGGGT	29239	23513	19137	540	38030	9620	5045	359	ola-miR-430a
miR-430b	AAAGTGCTATTAAGTTGGAG	3315	3031	2298	104	5843	2200	1236	76	dre-miR-430b

miR-430c	TAAGTGCTTCTCTTTAGGGT	93116	86723	64634	2454	83965	24542	17168	1262	dre-miR-430c
miR-430d	TAAGTGCTTCTCTTTGGGG	95	54	33	0	227	71	39	3	ola-miR-430d
miR-4448	GGCTCGTTGGTCTAGGGGTA	17008	6981	2883	2405	19696	20787	7578	2344	hsa-miR-4448
miR-4454	CGGTTCGTATCCGGGTCACG	226	94	76	108	245	246	118	440	hsa-miR-4454
miR-449	AGGCAGTGTCTTGTTAGCTG	553	456	363	551	356	309	319	396	hhi-miR-449
miR-451_1	AAACCGTTACCATTACTGAT	49014	130851	116760	69922	187059	159995	805577	60315	bta-miR-451
miR-451_2	AAACCGTTACCATTACTGAG	3198	5998	10344	7424	15251	15320	122360	5868	gga-miR-451
miR-451_3	TAACCGTTACCATTACTGAG	1129	2272	5978	3159	3342	3929	22710	3120	gga-miR-451
miR-454-3p	TAGTGCAATATTGCTTATAG	3539	4058	4657	1892	4807	3561	3509	2155	ggo-miR-454
miR-454-5p	ACCCTATCAATATTGCCTCT	83	96	72	24	110	94	74	29	tgu-miR-454-5p
miR-455a	TATGTGCCCTTGGACTACAT	1779	2680	5763	7696	1826	1894	1380	10307	fru-miR-455
miR-455-3p	TGCAGTCCATGGGCATATAC	144	158	215	433	457	422	508	751	gga-miR-455-3p
miR-455b	GTATGTGCCCTTGGACTAC	636	1169	2130	1124	574	407	271	1165	dre-miR-455b
miR-456-3p	CAGGCTGGTTAGATGGTTGT	11001	13678	14379	8020	11432	6006	8311	6093	dre-miR-456
miR-458	ATAGCTCTTTGAATGGTACT	10646	20674	28661	46191	18459	11743	23730	45944	ipu-miR-458
miR-459-5p	TCAGTAACAAGGATTCATCC	40	97	216	1325	36	57	230	2377	ipu-miR-459
miR-460a-5p	CCTGCATTGTACACACTGTG	1523	3603	4816	5388	2484	1932	3428	6662	aca-miR-460a-5p
miR-460a-3p	CACAGCGCATACAATGTGGA	122	171	195	296	137	125	172	327	gga-miR-460a-3p
miR-462	TAACGGAACCCATAATGCAG	753	1682	2069	8116	600	416	635	3887	ipu-miR-462
miR-4735-3p	ACTTCAACTTGAGCACCTTT	441	151	110	22	467	186	137	15	hsa-miR-4735-3p
miR-4856b-3p	TAAGTCTCACACCAGTGCAA	62	106	71	68	123	71	67	101	bfl-miR-4856b-3p
miR-488-3p	TTACAATTAAAGGATATTTC	49	117	194	429	66	107	324	441	hsa-miR-488-3p
miR-489	TGACATCATATGTACGGCTG	428	643	753	1186	408	452	579	974	ipu-miR-489
miR-499-5p	TTAAGACTTGCAGTGATGTT	9219	11218	10589	9871	8260	5710	3992	6353	xtr-miR-499
miR-499-3p	AACATCACTTTAAGTCTGTG	34	56	41	62	85	49	84	46	tgu-miR-499-3p
miR-4991	ATCGTATCAGTTTGGATTCC	499	1158	1437	809	789	516	782	826	egr-miR-4991
miR-5106	AGGGGTGTAGCTCAATTGGC	184	30	92	4	585	104	351	31	mmu-miR-5106
miR-551	GCGACCCATCCTTGGTTTCT	185	194	238	126	344	219	231	173	ccr-miR-551

miR-611	CACAGTAGAGCTAGCAGACC	6178	11494	13256	5347	8835	6777	7386	5060	hsa-miR-611
miR-6239	AACTCTTAGCGGTGGATCAC	5505	3560	2792	105131	2939	3121	737	86015	mmu-miR-6239
miR-6240	CAAAGCATCGCGAAGGCCCA	176	42	29	5	54	38	8	4	mmu-miR-6240
miR-6412	TCGAAACCATCCTCTGCTAC	449	221	307	119	359	178	196	113	mmu-miR-6412
miR-643	TTTCTGTCTTGTATTCTACC	1868	1645	1564	238	872	516	200	140	mml-miR-643
miR-6516-3p	ATGTATGATACTGCGGACAA	46	9	6	9	99	77	63	18	gga-miR-6516-3p
miR-6673-3p	GTAGAATCAGATGTGCTTGT	20	29	23	2	63	38	64	47	gga-miR-6673-3p
miR-7-5p	TGGAAGACTAGTGATTTTGT	5507	4688	4187	5211	5468	3669	4930	3895	tca-miR-7-5p
miR-7a-1-3p	CAACAAATCACAGTCTGCCA	47	61	60	27	80	81	64	24	mmu-miR-7a-1-3p
miR-7a-2-3p	CAACAAGTCCCAGTCTGCCT	31	37	65	12	67	49	56	13	rno-miR-7a-2-3p
miR-7132	TGAGGCGTTTAGAACAAGTT	1466	1767	1906	4047	1062	551	984	3746	ccr-miR-7132
miR-716b	GCAGATCTTGGTGGTAGTAG	509	477	140	42	103	70	30	45	sha-miR-716b
miR-722	TTTTGCAGAAACGTTTCAGA	31	73	84	191	66	26	75	251	dre-miR-722
miR-723-3p	AGACATCAGATAAATCTGTG	39	62	105	186	47	50	173	214	dre-miR-723-3p
miR-724	TTAAAGGGAATTTGCGACTG	3470	6622	11100	11913	3780	3049	6093	10819	ccr-miR-724
miR-725	TTCAGTCATTGTTTCTGGTA	3472	7090	12625	26369	5005	4127	10188	20796	dre-miR-725
miR-727-3p	GTTGAGGCGAGTTGAAGACT	53	74	61	41	37	48	54	83	ccr-miR-727-3p
miR-727-5p	TCAGTCTTCAATTCCTCCCA	180	223	305	131	329	332	435	200	ccr-miR-727-5p
miR-728	ATACTAAGTACACTACGTTT	38	17	25	70	77	63	127	49	ipu-miR-728
miR-730	TCCTCATTGTGCATGCTGTG	345	468	1014	349	394	204	208	572	dre-miR-730
miR-731	AATGACACGTTTTCTCCCGG	227	583	896	2405	216	127	183	801	dre-miR-731
miR-733	GCGCTGGTGTAGCTCAGTGG	72	29	20	37	68	97	28	50	dre-miR-733
miR-736	GTAAGACGAACAAAAAGTTT	513	559	330	155	595	355	433	149	dre-miR-736
miR-737	GTTTTTTTACGTTTAGATTT	81	40	33	98	225	132	164	136	ipu-miR-737
miR-749	GCCCGGATGAGCCTCGGTGG	118	99	113	38	155	76	131	53	sme-miR-749
miR-7551	TGTCGATCGGGGGCCTGAGT	403	220	21	4	153	36	13	6	ipu-miR-7551
miR-7552	AAATGTCCCTTAATTGTTTG	8	40	87	144	21	18	80	129	ipu-miR-7552
miR-7641	TCGTCTGATCTCGGAAGCTA	951	451	422	200	778	931	364	295	hsa-miR-7641
miR-7876-3p	AATACCACTTAGAATGCATG	98	149	168	61	69	75	43	33	prd-miR-7876-3p

miR-7977	TCGTTTCCCGGCCAACGCAC	216	128	288	216	192	127	166	227	hsa-miR-7977
miR-9-3p	TAAAGCTAGAGAACCGAATG	178730	416411	400453	116108	263613	160061	228959	78661	gga-miR-9-3p
miR-9b-3p	TAAAGCTAGAGAACCGAAAG	25688	52113	47114	15053	34163	22741	30708	10572	pol-miR-9b-3p
miR-9-5p	TCTTTGGTTATCTAGCTGTA	630628	2284664	2879902	636745	1077974	1038213	1787142	726463	aae-miR-9a
miR-92-3p	TATTGCACTTGTCCCGGCCT	77894	84513	84205	23555	141416	84198	74788	22606	mdo-miR-92a-3p
miR-92-5p	AGGTTGGGCTAGGTAGCAAT	1703	2505	1954	332	2793	1338	1269	348	hsa-miR-92a-1-5p
miR-93	AAAAGTGCTGTTTGTGCAGG	84730	141915	123156	29019	120828	73778	74631	27733	ccr-miR-93
miR-96-5p	TTTGGCACTAGCACATTTTT	2491	4486	5433	11622	2962	2474	5111	12060	rno-miR-96-5p
miR-99-5p	AACCCGTAGATCCGATCTTG	5587	10095	11376	7242	10537	10438	12313	7943	ipu-miR-99b
miR-99-3p	CAAGCTCGCCTCTGTGGGTC	75	101	163	41	140	68	104	52	ola-miR-99
miR166g-3p	ACAAACTGATCTTGGACCAG	666	1176	1032	318	833	534	690	249	bdi-miR166g-3p
miR1859	TTCCCTATGTAGTGCATTAC	89	78	63	36	113	100	52	38	osa-miR1859
miR444d.3	GATCTGTTTGTGCTTTCTTG	99	219	371	27	65	54	45	46	osa-miR444d.3
miR821a	AACGTTCCTGTTGATGACTG	76	133	133	40	74	46	51	21	sbi-miR821a
miR-1001-5p	ATCCTGATCTCGCGAGTTTA	367	494	542	64	353	252	176	24	dsi-miR-1001-5p

Table S 6.3 - Putative novel arctic charr miRNAs. 427 novel miRNA candidates. Included are novel miRNA candidate IDs, miRDeep2 score (the log-odds score assigned to the hairpin by miRDeep2), significant ($p<0.005$) randfold p-value of the putative miRNA to form a hairpin structure, consensus mature, consensus star, consensus precursor sequences and genomic coordinates in Salmon.

ID	miRDeep2 score	significant randfold p-value	consensus mature sequence	consensus star sequence	precursor coordinate in the Salmon genome
sal-nov-1	6.60E+02	no	acuguaugucuguagucaug	uuggccaguguaauacuguaacu	gi|354430582|gb|AGKD01028454.1|:1548..1657:+
sal-nov-2	6.50E+02	yes	aauacuguaacuauuggcca	uuggccaguauaauacuguaaau	gi|354399331|gb|AGKD01059705.1|:6064..6174:+
sal-nov-3	3.70E+02	yes	cagagaagauugaacgacugcu	aaucguucaauauucucugug	gi|354362801|gb|AGKD01096235.1|:2935..3045:
sal-nov-4	2.5	yes	uaagcucgcgauaucaggaugg	aucuugaacgcugcgcucucg	gi|354435043|gb|AGKD01023993.1|:3981..4088:+
sal-nov-5	1869.2	yes	gacaugacuggauaggugaaagc	uugaccuauccaguaguguuag	gi|354430159|gb|AGKD01028877.1|:7405..7517:+
sal-nov-6	7.20E+01	no	auucuaacucuacucugaggcu	ccucguaugagagcugguaac	gi|354381996|gb|AGKD01077040.1|:5017..5126:+
sal-nov-7	239	no	uaccuguaucgaaaguggcuguggc	caggucauugauaaugcaggaaga	gi|354270665|gb|AGKD01187862.1|:8452..8559:+
sal-nov-8	744.6	no	aggacuuaguacccgcugc	gucggacuguccucagu	gi|354451648|gb|AGKD01007397.1|:16197..16304:+
sal-nov-9	498.1	no	aggcauggaugaauucugug	cagacaccaugcaaaagcau	gi|354036921|gb|AGKD01421549.1|:4918..5027:-
sal-nov-10	101.3	yes	agugcacuauaaagggaauagg	ccuauucucuaugaagugcac	gi|354369005|gb|AGKD01090031.1|:4188..4296:+
sal-nov-11	54.4	yes	auacuguaacuauuggccagu	uuggccaguauaauacuguaacu	gi|354222094|gb|AGKD01236392.1|:3874..3986:+
sal-nov-12	999.2	yes	uugcugguaacacugucuguga	aaggcacacuuuaaccagcaugg	gi|354227323|gb|AGKD01231204.1|:1505..1616:+
sal-nov-13	174.3	yes	uugguuaucugaaucagcagug	cagcugauugaaauaaucaaag	gi|354112797|gb|AGKD01345673.1|:583..693:-
sal-nov-14	2.4	yes	auggacuauguagggaguagga	cccuauuccuuaugaagugca	gi|354360416|gb|AGKD01098620.1|:4839..4948:+
sal-nov-15	106.7	yes	gaauggaaugguaucaaacaca	uuugaugucauuccauuuacuc	gi|354418205|gb|AGKD01040831.1|:11162..11271:-
sal-nov-16	3913.5	yes	aucguaucaguuuggauucca	aaaccaaacugauaugcucugu	gi|354333716|gb|AGKD01124812.1|:2099..2206:+
sal-nov-17	2.4	yes	guagaaucagaugugcuugucc	aaaaacaccugauucaacuuga	gi|354406799|gb|AGKD01052237.1|:15415..15524:+
sal-nov-18	559.4	yes	gugucgacaccgauaaggcagc	ugccaucucuguaucgacaccg	gi|354025355|gb|AGKD01433115.1|:1166..1277:-
sal-nov-19	1483.3	yes	auccugaucucgcgaguuuaca	uaagcucgcgauaucaggaugg	gi|354458648|gb|AGKD01000401.1|:21731..21839:-
sal-nov-20	3676.2	yes	uucaaguaaucuaggauaggcu	ccuauucgggaugacuugguuc	gi|354329862|gb|AGKD01128666.1|:8691..8799:+
sal-nov-21	25941.6	yes	uucagucauuguuucugguagu	agcuagaaaccuugccugagau	gi|354209511|gb|AGKD01248975.1|:753..860:+
sal-nov-22	4611	no	uaagagcuuuuuguugaggua	acccuaacaaaagcauugacu	gi|354456619|gb|AGKD01002427.1|:19058..19168:+

sal-nov-23	153	yes	aagccgcguguaucgcccuggac	cagggcauuacgugaggcugg	gi|354440892|gb|AGKD01018150.1|:2406..2515:+
sal-nov-24	2749.6	yes	aaugacacguuuucucccggauugc	caccgggaauuucgugucagc	gi|354263513|gb|AGKD01195014.1|:4033..4147:-
sal-nov-25	76.4	yes	acucuccugaccacagcagu	ugcugcggucgacagagugaga	gi|354006162|gb|AGKD01452308.1|:73..181:-
sal-nov-26	3.1	yes	agugcacuaaauagggaauagg	uauuucccuauuuagugcucuac	gi|354433811|gb|AGKD01025225.1|:2340..2450:-
sal-nov-27	3.00E+02	no	aucaaaugauguuaacugaugugg	caacaaaucacagcacaua	gi|354191381|gb|AGKD01267105.1|:1631..1744:-
sal-nov-28	1634.9	yes	auccugaucucgcgaguuuaca	uaagcucgcgagaucaggaugg	gi|354313422|gb|AGKD01145106.1|:2669..2776:+
sal-nov-29	2.60E+01	yes	caaauucgguucuacagggugc	ucccguagaggcgaauuugcgg	gi|354400493|gb|AGKD01058543.1|:13387..13497:-
sal-nov-30	1.30E+03	yes	ccauuaguuccuguugccaagg	aggcagcagcuacucuuccugggg	gi|354171609|gb|AGKD01286877.1|:544..655:+
sal-nov-31	340.4	no	ugagaaguuuggaacaaguaaa	gacuuggucaaagcuccucagu	gi|354205253|gb|AGKD01253233.1|:5094..5205:-
sal-nov-32	727.2	yes	gcgucgacacugauaaggcagc	ugccgucucugcgucgacacug	gi|354275311|gb|AGKD01183217.1|:1499..1610:-
sal-nov-33	1236	yes	gcuuugacgauguugcacuacu	cagugcaacaguauugucauggc	gi|354322613|gb|AGKD01135915.1|:3449..3559:+
sal-nov-34	471.9	no	ggauagguugauuucauugcau	cugaugaaaucacucuauggucug	gi|354443253|gb|AGKD01015789.1|:1577..1688:-
sal-nov-35	3.60E+01	yes	gguguuguguauuaaacugua	uaguuugauacacagcauaaa	gi|354444988|gb|AGKD01014054.1|:4064..4174:-
sal-nov-36	2766.5	yes	guuggguggagaugcugaau	uccaguaccaccaccccaaca	gi|354352628|gb|AGKD01106408.1|:3188..3297:-
sal-nov-37	2.80E+02	yes	uauucccuauauagugcacuac	aauguacuauauaggggauagg	gi|354300798|gb|AGKD01157730.1|:271..381:+
sal-nov-38	24207.9	no	ucuguacacaguucauggaagc	uuccauggccugcauacuc	gi|354448672|gb|AGKD01010371.1|:12909..13020:-
sal-nov-39	1.30E+02	yes	uggcgcuuguucacugggacga	cucccucguaagcaagcuacaca	gi|354447037|gb|AGKD01012005.1|:15515..15626:+
sal-nov-40	2.6	yes	ugggcccuggauuugaucgcu	cgaacaaaccccagggcuccgg	gi|354457572|gb|AGKD01001477.1|:21470..21580:-
sal-nov-41	4.20E+03	yes	uguauucuaccugauagagaga	uuucugucuuguauucua	gi|354453021|gb|AGKD01006024.1|:11742..11849:-
sal-nov-42	1743.3	yes	uuaaauuaggaguagaggcaga	aguuuucuacugcugaauuuaag	gi|354332432|gb|AGKD01126096.1|:792..903:-
sal-nov-43	2.20E+01	yes	uucagaacuagcggaauuccaga	ccgaguucugcuagucgugaac	gi|354454944|gb|AGKD01004102.1|:6462..6574:+
sal-nov-44	2.70E+03	no	uuuucuucugaacuggggc	caagacggacgaaagcg	gi|354299001|gb|AGKD01159527.1|:1795..1906:+
sal-nov-45	181.2	yes	aaaccagacguaacacuacgc	uguaguauucugucugguuuaa	gi|354418693|gb|AGKD01040343.1|:12112..12222:-
sal-nov-46	0.4	yes	aaagugcuauuaaauugg	ucaacuugagcaccuuucu	gi|354456619|gb|AGKD01002427.1|:26509..26616:+
sal-nov-47	241.6	yes	uuaauuaguauagccuguuuaag	uuacaggcuaugcuaaucugg	gi|354427047|gb|AGKD01031989.1|:2210..2322:+
sal-nov-48	2.5	yes	aaaguuguguacuauaugggga	uagggcacuacuuuuggcc	gi|354330087|gb|AGKD01128441.1|:5183..5291:+
sal-nov-49	937.7	yes	aaaugacaaagaacaaugacc	gauuuguuucuugccguuuga	gi|354040241|gb|AGKD01418229.1|:276..386:-
sal-nov-50	2.6	yes	aaauggaauuguauugaacaca	uuugauaccauuccauuuacuc	gi|354359244|gb|AGKD01099792.1|:1..112:-
sal-nov-51	2.6	yes	aaauguccgacuugcuaacugg	uauuaacgagucggacauuauc	gi|354454517|gb|AGKD01004529.1|:25185..25296:-

sal-nov-52	1621.6	yes	aacaauauccuggugcuguaugagu	uccagcaucagugauuuuguuac	gi|354424344|gb|AGKD01034692.1|:1821..1935:-
sal-nov-53	2.6	yes	aacacugguggaaccagc	ugguuccaccaagcuacc	gi|354446452|gb|AGKD01012590.1|:4538..4648:-
sal-nov-54	26654.5	no	aacauucauugcugucgcuggguu	cucacugaacaaugagugcaa	gi|354184920|gb|AGKD01273566.1|:294..407:-
sal-nov-55	389.7	no	aacguuccuguugaugacugua	cuguugauggcuguaacguucc	gi|354349350|gb|AGKD01109418.1|:6344..6455:-
sal-nov-56	332.4	yes	aagaagaucuuagugcuguac	cguagcacugaaauccu	gi|354253451|gb|AGKD01205076.1|:640..748:-
sal-nov-57	223.6	yes	aagcagcacagaaugguuugug	caggcuguacugugcuguugc	gi|354134594|gb|AGKD01323892.1|:43..154:-
sal-nov-58	79.1	yes	aagcgcgugagaucagaauggu	cauccugaucucgugcgcuuac	gi|354458268|gb|AGKD01000781.1|:7132..7243:-
sal-nov-59	1.8	yes	aaggcacacaaccagcauggcu	uugcuggugccacggucuguga	gi|354415304|gb|AGKD01043732.1|:11167..11276:+
sal-nov-60	1519.7	yes	aaggcacacuuaaccagcaugg	uuguuggugacauugucuguga	gi|354403808|gb|AGKD01055228.1|:4931..5042:+
sal-nov-61	1505.8	yes	aaggcacacuuaaccagcaugg	uugcuggugacauugccuguga	gi|354420722|gb|AGKD01038314.1|:10397..10508:+
sal-nov-62	1334	no	aaggcacacuuaaccagcaugg	uugcuguuaacacugucuguga	gi|354025093|gb|AGKD01433377.1|:316..424:+
sal-nov-63	1261	yes	aaggcacacuuaaccagcaugg	uuguuggugacaaugucuguga	gi|354267671|gb|AGKD01190856.1|:554..664:+
sal-nov-64	217.3	yes	aaggcacacuuuaccagcaugg	uuguuggugacauugucugug	gi|354451876|gb|AGKD01007169.1|:20788..20898:-
sal-nov-65	53.1	no	aagguacacuuaaccagcaugg	uugcugguggaacugucugug	gi|354388443|gb|AGKD01070593.1|:4173..4283:-
sal-nov-66	0.8	yes	aaguaaaacucaacuccacg	aggaguugaguuuaggcuac	gi|354436821|gb|AGKD01022215.1|:18521..18631:+
sal-nov-67	6372	yes	aagugacuguauaaguggaagca	cuuucaccaauccagccuuga	gi|354349138|gb|AGKD01109603.1|:10845..10956:-
sal-nov-68	395.9	yes	aauaccacuuagaaugcaugcc	uucauucuaaguaguauacug	gi|354398861|gb|AGKD01060175.1|:4684..4794:+
sal-nov-69	2	yes	aaucagacugaacacugcacu	cacagcauucaaucuggguua	gi|354378978|gb|AGKD01080058.1|:3040..3147:+
sal-nov-70	368.4	yes	aauggaauggcaucaaaca	uuugauaccauuccacugauu	gi|354195753|gb|AGKD01262733.1|:3946..4057:+
sal-nov-71	0.9	yes	aauggauaaggcacugu	augccaaccauauc	gi|354451382|gb|AGKD01007663.1|:10954..11062:-
sal-nov-72	2.2	yes	acaaacgccuauagcagaggu	cucugcuacagguuguugugg	gi|354311500|gb|AGKD01147028.1|:972..1081:-
sal-nov-73	2.1	yes	acaaggauugucaaggagugg	acuucuugucacuccuuguca	gi|354074706|gb|AGKD01383764.1|:395..505:-
sal-nov-74	72.4	yes	acaccggucuggaugauugcagaga	ucggcaagaacagacuggccc	gi|354438627|gb|AGKD01020411.1|:24235..24346:+
sal-nov-75	1.1	yes	acaccggucuggaugauugcagaga	ucagcaagaacagacuggccc	gi|354383173|gb|AGKD01075863.1|:3161..3275:-
sal-nov-76	0	yes	acacugguaguaacacuggu	uaguaacacugguaguaguaa	gi|354449096|gb|AGKD01009947.1|:6947..7057:+
sal-nov-77	2	yes	acagcaggcacagacagg	ggucauguggaggcugcuac	gi|354437134|gb|AGKD01021902.1|:999..1107:-
sal-nov-78	363.5	yes	acagggcgucauucggaugcgc	cacagucgaacaacgcccugga	gi|354444619|gb|AGKD01014423.1|:5106..5217:+
sal-nov-79	245.1	yes	acagucucuggaaagcaauagc	auugcugugcuuggaucucaga	gi|354254685|gb|AGKD01203842.1|:1456..1567:+
sal-nov-80	1.4	yes	accaggugcuguaagca	ugacacaagaugcacaguag	gi|354436330|gb|AGKD01022706.1|:3840..3947:+

sal-nov-81	1.6	yes	acccuguaucgaaagugacuauggc	cauagucauugauguuacgggaguug	gi|354380520|gb|AGKD01078516.1|:12456..12570:-
sal-nov-82	2.2	yes	acggaauggcaucaaauacau	uuugauaccauuccacuccagc	gi|354401603|gb|AGKD01057433.1|:5156..5267:-
sal-nov-83	1.7	yes	acggagcuaauggaauggcauc	uaucauuccagugauuccacuc	gi|354333061|gb|AGKD01125467.1|:1614..1723:+
sal-nov-84	57.1	yes	acuaacacccuggaccagagcu	ucuuguccaacggcguuagugc	gi|354397195|gb|AGKD01061841.1|:6241..6352:-
sal-nov-85	101.7	yes	acuaucugugcucacuggcaga	uucagugaacacagguacuag	gi|354293713|gb|AGKD01164815.1|:1490..1601:-
sal-nov-86	1.8	yes	acugauuuccucugguguucaga	uagcaccauuugaagucaauug	gi|354256244|gb|AGKD01202283.1|:2250..2362:-
sal-nov-87	2.6	yes	acuggaacagaauaagugga	uccauuuauucaguuccagcca	gi|354455250|gb|AGKD01003796.1|:26438..26547:-
sal-nov-88	2.1	yes	acugggauuguacaggggu	ucccugggcucucccagag	:..:
sal-nov-89	64.7	yes	acugguauacuggucuggacu	gucuggacugguauacuggucu	gi|354345399|gb|AGKD01113342.1|:2576..2687:-
sal-nov-90	55.6	yes	acugguauacuggucuggacu	gucuggacugguauacuggucu	gi|354345399|gb|AGKD01113342.1|:2626..2737:-
sal-nov-91	2.6	yes	acuuugguaacagaaugacggu	cauucuguuacuaaacuuugca	gi|354325726|gb|AGKD01132802.1|:7573..7684:-
sal-nov-92	1041.4	yes	agaagacgaauagaacucuguu	uagaacucuguuguguucuga	gi|354232742|gb|AGKD01225785.1|:509..620:-
sal-nov-93	2.2	yes	agaagacgaauagaacucuguu	cucuguuguguucugagaaga	gi|354232742|gb|AGKD01225785.1|:717..828:-
sal-nov-94	1.7	yes	agaagacgauuagaacucuguug	uagaacucuguuguguucuga	gi|354232742|gb|AGKD01225785.1|:836..948:-
sal-nov-95	1.4	yes	agaagacgauuagaacucuguug	auuagaacucuguuguguucuga	gi|354232742|gb|AGKD01225785.1|:786..898:-
sal-nov-96	483.6	yes	agaauaaugccagcagucggcc	ccagcgcugguguuauuggga	gi|354379705|gb|AGKD01079331.1|:7706..7815:+
sal-nov-97	450.5	yes	agaauaaugccagcagucggcc	ccagcacugguguuauuggga	gi|354393131|gb|AGKD01065905.1|:6703..6812:+
sal-nov-98	395.7	no	agaauaaugccagcagucggcc	ccagcacugguguuauugga	gi|354366030|gb|AGKD01093006.1|:1624..1732:+
sal-nov-99	73.3	yes	agacagucugacacuacugcug	agcaguaguguaagacuggc	gi|354396972|gb|AGKD01062064.1|:14481..14591:-
sal-nov-100	257.5	yes	agacaucagauaaaucugugcu	ggcagcuuugaaugauguuauuuc	gi|354315732|gb|AGKD01142796.1|:2092..2200:+
sal-nov-101	6058	no	agacaugaacuacuggguu	ccauagauacagaucaaa	gi|354450397|gb|AGKD01008648.1|:395..502:+
sal-nov-102	98.5	yes	agacugagguacucacaaagg	uuugugagugccucagaagga	gi|354441278|gb|AGKD01017764.1|:10213..10323:-
sal-nov-103	171.6	yes	agagcaucuggucagcggggugg	uccucgucugcccacauguucu	gi|354272263|gb|AGKD01186264.1|:1000..1112:+
sal-nov-104	84.3	yes	agaggagcaagcgaugaguggcu	ccgacugucuugcuucccgu	gi|354378557|gb|AGKD01080479.1|:7218..7325:+
sal-nov-105	156.3	yes	agcgagaucaauggaagguuu	aaccuuccguugaucucgcugu	gi|354416816|gb|AGKD01042220.1|:6803..6913:+
sal-nov-106	1.1	yes	agcgcauuggauuucuau	agaccggugcgccau	gi|354426089|gb|AGKD01032947.1|:1831..1938:-
sal-nov-107	324.8	yes	agcugaccaacaaaaacucugg	ggaguuuuuccuggucagguca	gi|354224672|gb|AGKD01233855.1|:213..324:-
sal-nov-108	65	no	aggcaguguaauguuaacuga	cagcucacgcuacagcugc	gi|354272907|gb|AGKD01185620.1|:1864..1974:+

sal-nov-109	972	no	aggcagugucuuguuagcuga	aguuuaccuguugcugcccuc	gi|354272907|gb|AGKD01185620.1|:3209..3318:+
sal-nov-110	2.3	yes	aggcagugucuuguuagcuga	agcuuaccugguacugcccuc	gi|354360418|gb|AGKD01098618.1|:506..616:-
sal-nov-111	431.3	yes	aggcuagcuucugcaagagacg	ucucugguguuacuauccauuu	gi|354242112|gb|AGKD01216415.1|:415..526:-
sal-nov-112	72.5	yes	agucugaagcagcagcaagga	cuugucuucugagauagacugg	gi|354423729|gb|AGKD01035307.1|:10087..10197:-
sal-nov-113	644.2	yes	agugcacuauauaggaaauagg	auucccuauguagugcacuacu	gi|354186051|gb|AGKD01272435.1|:396..505:-
sal-nov-114	2.6	yes	agugcacuauauagggaauagg	ccuauucucaauguagugcac	gi|354387935|gb|AGKD01071101.1|:3888..3997:-
sal-nov-115	2.1	yes	agugcacuauauagggaauagg	cuaaaguagugcacuauauagg	gi|354245745|gb|AGKD01212782.1|:1569..1680:+
sal-nov-116	2	yes	agugcacuauauagggaauagg	uauuagaccagaccccuaugugccuug	gi|354246166|gb|AGKD01212361.1|:925..1034:+
sal-nov-117	400	yes	agugcacuauauagguacuagg	uguucccuauguagagcacuac	gi|354321127|gb|AGKD01137401.1|:2339..2450:-
sal-nov-118	432.6	yes	agugcacuauaucgguaauagg	auucccuauuuagugcaccacu	gi|354005939|gb|AGKD01452531.1|:48..159:-
sal-nov-119	2.3	yes	agugcacuauaucgguaauagg	uauucccuacauaguggacuac	gi|354322831|gb|AGKD01135697.1|:331..442:+
sal-nov-120	1.6	yes	agugcacuguaaagggaauaag	uauucccuuuauagggcccauag	gi|354382539|gb|AGKD01076497.1|:251..361:-
sal-nov-121	214	yes	agugcacuguguagggaguagg	auucccuacauagugcacuaca	gi|354389922|gb|AGKD01069114.1|:1128..1239:-
sal-nov-122	2.9	yes	agugcacuguguagggaguagg	uauuaccaagugcacuac	gi|354384351|gb|AGKD01074685.1|:457..567:+
sal-nov-123	2.4	yes	aguggaaucaguggaauggua	ccguuccauuugcucu	gi|354447051|gb|AGKD01011991.1|:16740..16851:+
sal-nov-124	1.7	yes	agugggagagcauuugg	agagcucuccuggugcuga	gi|354445658|gb|AGKD01013384.1|:17440..17547:+
sal-nov-125	2.7	yes	aguguuacuugcgguuugcuga	aacaagccacaaguaacgcccu	gi|354295074|gb|AGKD01163454.1|:1804..1913:+
sal-nov-126	91.6	yes	auacggucucuguagaagac	cgucuacaguggcuguacagca	gi|354449921|gb|AGKD01009124.1|:7346..7453:+
sal-nov-127	102.7	yes	auacuagcaucccacuuagaau	uucuaaguggugcgcuagucuag	gi|354406901|gb|AGKD01052135.1|:9254..9366:+
sal-nov-128	1.9	yes	auagaaauguguuguauagagcu	uucuacagaauauauuucugucu	gi|354419545|gb|AGKD01039491.1|:9462..9572:+
sal-nov-129	1.8	yes	auagggaauagggugccauuug	uauucccuuuguggugc	gi|354409667|gb|AGKD01049369.1|:6748..6856:-
sal-nov-130	2.8	yes	auagugcacuacuuuugaccag	acuuuuaaguaguguacuauca	gi|354421134|gb|AGKD01037902.1|:10118..10229:-
sal-nov-131	2	yes	auagugcacuacuuuugaccag	cuauuccuaauguagugcacu	gi|354368753|gb|AGKD01090283.1|:1633..1743:+
sal-nov-132	158.6	yes	aucacauugccagggauaccac	gugucucuggcaugcugauu	gi|354318549|gb|AGKD01139979.1|:6738..6847:+
sal-nov-133	1.4	yes	aucacgucggggucacca	ucugacuuguuuuaacu	gi|354407399|gb|AGKD01051637.1|:4089..4197:-
sal-nov-134	641.3	yes	auccugaucucgcgaguuuac	uaagcucgcgagaucaggaugg	gi|354458604|gb|AGKD01000445.1|:3403..3514:+
sal-nov-135	2.7	yes	auccugaucucgcgaguuuac	guacgcucgcgagaucaggauggu	gi|354417862|gb|AGKD01041174.1|:4989..5096:-
sal-nov-136	147.4	yes	aucggacauucauccaacagagu	ucuguugaaugaauguccgacc	gi|354247748|gb|AGKD01210779.1|:1716..1827:-

sal-nov-137	1.1	yes	aucguaucaguuuggauucca	gaguccuaacuugagaucu	gi|354333716|gb|AGKD01124812.1|:2149..2256:+
sal-nov-138	2.5	yes	aucuguuugugcuguugagcca	gcuuaacaguacaaacuaaucu	gi|354429959|gb|AGKD01029077.1|:18571..18682:+
sal-nov-139	140.7	yes	auggcuggaauggaaucaaugg	auugauuccaauccagccauu	gi|354404037|gb|AGKD01054999.1|:5991..6102:+
sal-nov-140	75.9	yes	augucaagacagcacaaacaga	uguuugugccgucuugccaacuc	gi|354369313|gb|AGKD01089723.1|:909..1020:-
sal-nov-141	2.5	yes	auuacucuguaggcuauugacc	uaaguagccuacagagugguau	gi|354453938|gb|AGKD01005108.1|:11958..12068:-
sal-nov-142	2.6	yes	auuaguggauucaggugugguag	cuacacaccugauuccacua	gi|354426768|gb|AGKD01032268.1|:11217..11328:+
sal-nov-143	3	yes	auuaguuugugcugucuugcca	gcaagacagaacaaauaaaucu	gi|354382724|gb|AGKD01076312.1|:7408..7516:+
sal-nov-144	2.7	yes	auuaguuugugcugucuugcca	cgcacaaacagaucuggga	gi|354413602|gb|AGKD01045434.1|:2530..2640:-
sal-nov-145	86.7	yes	auuauaccacgucuaagggcugu	agccguucgccgugguacguugg	gi|354327489|gb|AGKD01131039.1|:777..888:-
sal-nov-146	2.7	yes	auuauaccacgucuaagggcugu	acagcccuuagccguggua	gi|354403326|gb|AGKD01055710.1|:129..240:-
sal-nov-147	307.5	yes	auucccuauguagugcacuacu	guagugcacuauauaggga	gi|354369847|gb|AGKD01089189.1|:2084..2195:+
sal-nov-148	1.3	yes	auucccuauguagugcacuacu	gagcccuguguagggaauaga	gi|354405463|gb|AGKD01053573.1|:2911..3021:+
sal-nov-149	2.2	yes	auugcacuauguagagaauagg	acuauucccuauauagug	gi|354390619|gb|AGKD01068417.1|:9770..9878:+
sal-nov-150	2.3	yes	auugcacuauguagggaauagg	uauucccuaaguagugcacuuu	gi|354309103|gb|AGKD01149425.1|:3658..3765:+
sal-nov-151	1.9	yes	auugcacuauguagggaauagg	uauucccuaaguagugcacuuu	gi|354309103|gb|AGKD01149425.1|:3608..3715:+
sal-nov-152	148	yes	auuuuauuccagauuuuguaag	uuuaaaauguggaauauuggac	gi|354405565|gb|AGKD01053471.1|:8175..8286:-
sal-nov-153	2631.6	yes	auuuuguuaggcguagcuugc	caaacuacacugaacaaaaaua	gi|354301542|gb|AGKD01156986.1|:16657..16767:+
sal-nov-154	2.1	yes	caaaagaucugaauugggcugc	gcagcucaauucugauauuuuu	gi|354416503|gb|AGKD01042533.1|:8627..8735:-
sal-nov-155	17245.2	yes	caaagugcuguuugugcaaguag	acugcaaaaccagcacuuccuga	gi|354409868|gb|AGKD01049168.1|:15638..15750:-
sal-nov-156	2.7	yes	caagccacauguaacgc	guuaggugguggcuuacu	gi|354362178|gb|AGKD01096858.1|:6701..6811:+
sal-nov-157	2.4	yes	caagccacauguaacgc	guuacguucaguuugcu	gi|354179699|gb|AGKD01278787.1|:163..273:-
sal-nov-158	1.8	yes	caauaauggcuggaacg	cuccagccauuaccaca	gi|354449782|gb|AGKD01009263.1|:5843..5954:+
sal-nov-159	120.6	no	cacaagcugauugucacucugg	ggguugcaaacagcacugugga	gi|354381484|gb|AGKD01077552.1|:1957..2067:+
sal-nov-160	2	yes	cacaggcagcguaauucagauu	uccaauugagcugccuguguu	gi|354061421|gb|AGKD01397049.1|:357..464:-
sal-nov-161	32325.6	no	cacaguagagcuagcagacccc	agccuccugguccuacugcugcu	gi|354428317|gb|AGKD01030719.1|:7599..7710:-
sal-nov-162	2.2	yes	cacccuauucccuugauuguga	ccauaaggggauaggguacc	gi|354415329|gb|AGKD01043707.1|:13489..13600:-
sal-nov-163	479.3	yes	cagaacugaucugauugucaa	ugaccaaucacaucagaucuuuu	gi|354448241|gb|AGKD01010802.1|:6374..6483:+
sal-nov-164	2.6	yes	cagagaagauugaacgacugcu	caaucguucaauauucucugug	gi|354362801|gb|AGKD01096235.1|:2935..3045:-
sal-nov-165	2.3	yes	cagagagcagaacugguugga	caaccaguuuugcucacuggg	gi|354168500|gb|AGKD01289986.1|:856..966:-

sal-nov-166	1196.1	yes	cagcgaaugacucugacuagca	cuggcucagacuugcugcgcug	gi|354121201|gb|AGKD01337285.1|:80..190:-
sal-nov-167	453.3	yes	cagguaguuucguguuguuagg	ccaccacacgaaacugc	gi|354071994|gb|AGKD01386476.1|:5470..5581:-
sal-nov-168	12073.5	no	cagugcaauaguauugucaugg	uuugaugacguugcacuacu	gi|354187706|gb|AGKD01270780.1|:1648..1761:+
sal-nov-169	1.6	yes	caguguugccuguuacaucaua	uggucggggaacaaccuga	gi|354440526|gb|AGKD01018516.1|:6085..6192:+
sal-nov-170	1.7	yes	caucccacuacugacacca	uaguuacuggugggguuuc	gi|354380604|gb|AGKD01078432.1|:2401..2508:+
sal-nov-171	2.3	yes	caugacaaagaguggaaugaua	acauucgacgccuugucaagcc	gi|354280977|gb|AGKD01177551.1|:1822..1931:-
sal-nov-172	2.1	yes	caugcauuguagaagauccaag	uggaccuucuccaauguauuuugag	gi|354455244|gb|AGKD01003802.1|:9532..9643:-
sal-nov-173	0	yes	cauggcaguaguuauuucaca	ugacuuaauuauaugccaggaa	gi|354199881|gb|AGKD01258605.1|:1319..1428:-
sal-nov-174	2.2	yes	ccaaguuugguaacagaaugac	gucauucuguuaccuaacuuug	gi|354339756|gb|AGKD01118985.1|:2048..2158:-
sal-nov-175	2.8	yes	ccagacagcaucagauaca	uaucugaugcggucuggcca	gi|354438408|gb|AGKD01020630.1|:9112..9223:-
sal-nov-176	2.5	yes	ccagacagcaucagauaca	ucuggugcugucuggccaaa	gi|354444273|gb|AGKD01014769.1|:14744..14854:+
sal-nov-177	118.7	yes	ccagacgacacaaauagaucu	aucuguuugugcagucuugc	gi|354427402|gb|AGKD01031634.1|:11425..11535:-
sal-nov-178	1918.7	yes	ccaguggagcugcuguuauggg	uguaacagcaacuccauauggaau	gi|354452024|gb|AGKD01007021.1|:17436..17547:-
sal-nov-179	1872.7	yes	ccaguggagcugcuguuauggg	uguaacagcaucuccauaug	gi|354426828|gb|AGKD01032208.1|:4509..4620:-
sal-nov-180	2578.1	yes	ccauuaguuccuguugccaag	aggcagcagcuacucuuccugggg	gi|354171609|gb|AGKD01286877.1|:544..655:+
sal-nov-181	545.6	yes	cccauuaguuccuauugccaag	aggcagcagcuacucuuccugggg	gi|353898887|gb|AGKD01552432.1|:12..123:+
sal-nov-182	2.4	yes	cccuauucccuauacagugca	cacuauauagggaaugggucc	gi|354444067|gb|AGKD01014975.1|:5498..5609:+
sal-nov-183	2.4	yes	ccggggaaaaaaaaaaa	uuuuuuuuuuuaaaggg	gi|354000193|gb|AGKD01458277.1|:57..164:-
sal-nov-184	1.1	yes	ccggggaaaacaaaaaa	uuuuuuuuuuuugggg	gi|354271845|gb|AGKD01186682.1|:1864..1971:+
sal-nov-185	1.7	yes	ccuaguccuggacaaugaagcau	aagcuuuaguccaggauuaga	gi|354433050|gb|AGKD01025986.1|:9869..9978:+
sal-nov-186	2.5	yes	ccugaucugaugcugcuguauug	cacaaacagaucugggaccagg	gi|354070248|gb|AGKD01388222.1|:1020..1131:-
sal-nov-187	2.7	yes	cuaaaaauagugcacuauauagg	auagugcacuauuuuugaacag	gi|354256041|gb|AGKD01202486.1|:751..863:-
sal-nov-188	157.5	yes	cuagacacgggccuuaggauugg	gauccuagaucugugucugc	gi|354251766|gb|AGKD01206761.1|:2223..2335:+
sal-nov-189	882.2	no	cuagaggcucuguccagcggcu	ucuguccagcggcucuguc	gi|354297047|gb|AGKD01161481.1|:1638..1745:+
sal-nov-190	880.1	no	cuagaggcucuguccagcggcu	cucuguccagcggcucugucc	gi|354297047|gb|AGKD01161481.1|:1688..1795:+
sal-nov-191	2.3	yes	cucccuauguaguacacuacuu	agcacuguauagggaauagggu	gi|354216047|gb|AGKD01242439.1|:2097..2208:-
sal-nov-192	2.1	yes	cuccguuccagacauuauuaug	ugguaauaauggggaguggaguc	gi|354439164|gb|AGKD01019876.1|:360..469:-
sal-nov-193	1.5	yes	cucuggucaaaaguagggcac	gccguuuggggcagaagca	gi|354181248|gb|AGKD01277238.1|:334..444:-
sal-nov-194	375.1	no	cugaaucugucugacugccauu	agguagacaacauguagu	gi|354416607|gb|AGKD01042429.1|:8294..8401:+

sal-nov-195	221.4	yes	cugguuuccuaugguggguuuaga	uagcaccauaugaaaucagugu	gi|354446031|gb|AGKD01013011.1|:9817..9929:-
sal-nov-196	2.6	yes	cuucucgcggaucuccc	caggugagcaagaaggg	gi|353993913|gb|AGKD01464557.1|:91..198:-
sal-nov-197	138.3	yes	cuucuuugguuucuuuaugauu	aaucauauaguaaccaaaaaagu	gi|354243208|gb|AGKD01215319.1|:720..831:-
sal-nov-198	2.2	yes	cuugauuguuugacuaaggugg	aacugauucaacuaaucgugg	gi|354413406|gb|AGKD01045630.1|:376..487:+
sal-nov-199	2.8	yes	gacaugacuggauaggugaaagc	uugccuauccaguaguguuaga	gi|354428406|gb|AGKD01030630.1|:8177..8289:-
sal-nov-200	1.1	yes	gacaugacuggauaggugaaagc	ucauaccuagccuuguguuag	gi|354366226|gb|AGKD01092810.1|:2620..2732:-
sal-nov-201	67.6	no	gagauaauagguguaguauaacg	ugaacuaucccauuauucucac	gi|354440820|gb|AGKD01018222.1|:9951..10064:+
sal-nov-202	163.2	yes	gaugcaggauaaguuacagugc	ucugucguucccccugcaca	gi|354351987|gb|AGKD01107049.1|:2069..2180:+
sal-nov-203	1.7	yes	gauuaguggauucaggugugu	acacccaauuccacuaaucaa	gi|354440404|gb|AGKD01018638.1|:10380..10490:+
sal-nov-204	55.9	yes	gcacuaacagaucugggacca	acucagaucuguuuaugcugc	gi|354444027|gb|AGKD01015015.1|:206..316:-
sal-nov-205	2.7	yes	gcacuaacagaucugggacca	gucccagaucuguucaugaug	gi|354425411|gb|AGKD01033625.1|:11696..11806:+
sal-nov-206	767.4	yes	gcgacccacucaugauuucaga	uggagucagguguggguguag	gi|354207349|gb|AGKD01251137.1|:59..170:-
sal-nov-207	406.8	yes	gcgacccauccuugguuucuga	gaaaccaagaguggguguggg	gi|354208640|gb|AGKD01249846.1|:350..460:-
sal-nov-208	2.3	yes	gcgacccauccuugguuucuga	agaaaccaaguguggguguggccu	gi|354442731|gb|AGKD01016311.1|:12292..12402:+
sal-nov-209	2.3	yes	gcggaaugaguggaaugguauc	aaugaugugaguuaugcgcuguuc	gi|354381460|gb|AGKD01077576.1|:2022..2135:-
sal-nov-210	0.7	yes	gcgugguacucuguaaugugua	caugguacagccaccacuagcuu	gi|354430042|gb|AGKD01028994.1|:14750..14858:+
sal-nov-211	1.7	yes	ggaaggagguuguuggucaaga	auggccccauccauccucc	gi|354371042|gb|AGKD01087994.1|:8294..8401:+
sal-nov-212	2	yes	ggaaggauguuguuggccaaga	auggccccauccauccucc	gi|354361105|gb|AGKD01097931.1|:8463..8574:-
sal-nov-213	1.8	yes	ggaggggaauguaaaacuggc	caguuuaccauccacaaacccua	gi|354320809|gb|AGKD01137719.1|:4153..4262:+
sal-nov-214	659.3	yes	ggcagcuuugaaugauguuacuuc	agacaucagauaaaucugugcu	gi|354329653|gb|AGKD01128875.1|:8380..8493:+
sal-nov-215	0.4	yes	ggcgguagagcaucaga	ugaaaugugcaucucuguccc	gi|354426465|gb|AGKD01032571.1|:13773..13880:+
sal-nov-216	1.9	yes	ggcucguuggucuaggg	caaaucccgaugagcccu	gi|353961935|gb|AGKD01496535.1|:1210..1324:-
sal-nov-217	1.9	yes	gggggauuagcucaaaugguagagc	uuguacaggagcuagug	gi|354329571|gb|AGKD01128957.1|:4206..4319:-
sal-nov-218	57.8	no	gguguugugaaucaggccg	gcuacuucccaacaccagggu	:..:
sal-nov-219	889.9	yes	guaagacgaacaaaaaguuugu	aggcuuucuguuuguauuaugu	gi|354395220|gb|AGKD01063816.1|:1104..1211:+
sal-nov-220	2.2	yes	guagaacauacaugauuc	aucgugccaguucuguac	gi|354341013|gb|AGKD01117728.1|:168..275:+
sal-nov-221	2.3	yes	guagaaucagaugugcuugucc	aaaaacaccugauucaacuuga	gi|354427133|gb|AGKD01031903.1|:8226..8335:+
sal-nov-222	1.9	yes	guagaaucagaugugcuugucc	acaaaaacaacugauucaacuu	gi|354430822|gb|AGKD01028214.1|:18045..18156:-
sal-nov-223	2.5	yes	guagugcacuauauagggagua	uucccaauaucgugcacuacuu	gi|354208283|gb|AGKD01250203.1|:2309..2421:+

sal-nov-224	146.4	yes	guagugcacuauguagggaaua	uucccuauaaaauguauuacuu	gi|354373158|gb|AGKD01085878.1|:2700..2811:+
sal-nov-225	1.2	yes	guaugguucccaaucagagaca	ccugauugagaacccuacccggc	gi|354244341|gb|AGKD01214186.1|:1354..1465:+
sal-nov-226	79.8	yes	guauugggcugcuuguguuaac	uacacaggcagcccaauu	gi|354348668|gb|AGKD01110073.1|:12520..12631:+
sal-nov-227	0.9	yes	gucuaguggguuaggauu	uucuagcuuuguguga	gi|354446567|gb|AGKD01012475.1|:9506..9614:-
sal-nov-228	55.9	yes	gugaacguagcgauuaggcugg	aacuugaucgcuguguucacca	gi|354433629|gb|AGKD01025407.1|:10090..10200:-
sal-nov-229	1.3	yes	gugagaaaucuguugaugu	aucaggguuucucaaac	gi|354399934|gb|AGKD01059102.1|:12797..12904:+
sal-nov-230	194.8	yes	gugguaauggcuggaguggaau	cuccguuccagccauuauuau	gi|354188140|gb|AGKD01270346.1|:1592..1702:+
sal-nov-231	148.8	no	guugacguauucugugcgaga	aagcacuagccuaugucaaucu	gi|354434539|gb|AGKD01024497.1|:6678..6788:-
sal-nov-232	1.8	yes	guuuuuuuagguuuagacuuuug	aaaucuaaaccuaaagaaaaua	gi|354448264|gb|AGKD01010779.1|:22913..23020:+
sal-nov-233	1.6	yes	guuuuuuuagguuuagauuuuu	aaaucucaaccuaaagaaaaua	gi|354425235|gb|AGKD01033801.1|:2892..2999:+
sal-nov-234	157.3	yes	uaaaaggcacuuucaguggaga	accauuggaagugauuuuuagu	gi|354435951|gb|AGKD01023085.1|:3747..3858:-
sal-nov-235	528.6	yes	uaaaugcugcagaauugugcuu	gaacuauucugcaacauuuguu	gi|354388280|gb|AGKD01070756.1|:12051..12161:+
sal-nov-236	1.2	yes	uaaaaguaguucacuauguagg	ugcacuauauaggggauagg	gi|353980992|gb|AGKD01477478.1|:953..1064:-
sal-nov-237	5861.3	yes	uaaagugcuuacagugcaggua	ccgcaaugugagcacuucuguc	gi|354449516|gb|AGKD01009527.1|:5161..5270:+
sal-nov-238	1.4	yes	uaaaaguaguucacuauguagg	uagugcacuauguaggggaua	gi|353980992|gb|AGKD01477478.1|:815..926:-
sal-nov-239	0.5	yes	uaaaugcuuuuuguuggggua	cccuaacaaaagcguugacc	gi|354352934|gb|AGKD01106102.1|:1904..2014:+
sal-nov-240	146.1	yes	uaagaacugaacaggucacaggc	ccaccuuucaagucuuuucca	gi|354453903|gb|AGKD01005143.1|:16305..16414:+
sal-nov-241	181.7	yes	uaagacggcacaaacugaacug	aucuauuugugucgucuugcca	gi|354412133|gb|AGKD01046903.1|:7855..7965:+
sal-nov-242	282.2	yes	uaagucucacaccagugcaaaac	cugcucaggugugggaca	gi|354418392|gb|AGKD01040644.1|:2277..2384:+
sal-nov-243	288	yes	uaagucucacaccagugcaaaac	cugcuccgguguggacaaua	gi|354457173|gb|AGKD01001876.1|:19101..19213:-
Sal-nov-244	54967.9	yes	uaaggugcaucuaguguaguuag	acugcccuaguugcuccuucuggc	gi|354305223|gb|AGKD01153305.1|:8766..8878:-
sal-nov-245	0.2	yes	uaagugauuuuuguuggggua	cccuaacaaaagcguugac	gi|354352934|gb|AGKD01106102.1|:10365..10472:+
sal-nov-246	0.1	yes	uaagugcuuaucuuuggguag	accuuaaaaggaagcauacauu	gi|354456619|gb|AGKD01002427.1|:23588..23699:+
sal-nov-247	0	yes	uaagugcuucucuuuaggguag	accugaaauggaagcauacauu	gi|354456619|gb|AGKD01002427.1|:22580..22687:+
sal-nov-248	20630.3	yes	uaagugcuuuuuguuugggua	acccuaacaaaagcauugacu	gi|353942433|gb|AGKD01516037.1|:1294..1404:-
sal-nov-249	2.4	yes	uaaucucagaucuaggaucag	gacccuguaucuguggauagg	gi|354266910|gb|AGKD01191617.1|:1574..1685:-
sal-nov-250	245	yes	uaaugcgcuauauaggaaauag	uauucccuuuauagugcacuac	gi|354422235|gb|AGKD01036801.1|:453..563:-
sal-nov-251	1.4	yes	uacaaggauuggaauguuagc	uaucauuccacuccuugccau	gi|354368281|gb|AGKD01090755.1|:721..830:+
sal-nov-252	0.8	yes	uacacuguauagggaauaggg	cuuuucccuauuuagugcagca	gi|354382891|gb|AGKD01076145.1|:0..111:-

sal-nov-253	2.1	yes	uacacuguauagggaauagggu	uauucccuacauagugcacua	gi|354386422|gb|AGKD01072614.1|:5471..5581:+
sal-nov-254	1.9	yes	uacacuguauagggaauagggu	ccuauucccuucccugcacuuccug	gi|354320551|gb|AGKD01137977.1|:490..601:-
sal-nov-255	2.3	yes	uacaucaucuauacuguagug	uuacaguauaugaugauauccu	gi|354419876|gb|AGKD01039160.1|:12776..12886:+
sal-nov-256	1032.9	yes	uacuacauggaauaaggggcc	cacccuauuccccauguauugc	gi|354271646|gb|AGKD01186881.1|:1121..1231:-
sal-nov-257	2	yes	uagagcggacuuugagacacggg	ugguuucacaaacaccucucuugc	gi|354449569|gb|AGKD01009474.1|:1459..1570:+
sal-nov-258	269.2	yes	uagcaccauuuuaaaucaguua	acugauuuccucugguguucaga	gi|354282807|gb|AGKD01175721.1|:2216..2327:-
sal-nov-259	2309.6	yes	uagcagcacaucauguuuugca	caaaucauuaugugcugucacu	gi|354424302|gb|AGKD01034734.1|:1325..1436:-
sal-nov-260	1456.2	yes	uagcagcacaucauguuuugca	caaaucauuaugugcugccgcu	gi|354378776|gb|AGKD01080260.1|:11816..11924:+
sal-nov-261	20735.6	yes	uagcagcacgucaugguuugua	cagaccaugaagagcugcugcu	gi|354413093|gb|AGKD01045943.1|:4133..4244:-
sal-nov-262	6292.4	yes	uagcagcgcaucaugguuuga	cgaaccaucauuugcug	gi|354452782|gb|AGKD01006263.1|:9144..9253:+
sal-nov-263	0.2	yes	uagcagcgcaucaugguuuga	gaaucaucauuugcugcuuua	gi|354441185|gb|AGKD01017857.1|:342..451:+
sal-nov-264	2	yes	uagcguagcacuuagaaugaag	gcacucuaagugguauguuagu	gi|354382063|gb|AGKD01076973.1|:192..301:+
sal-nov-265	0.8	yes	uagcguagcacuuagaaugaag	gcauuauaagugguauugcuagu	gi|354403896|gb|AGKD01055140.1|:8599..8708:+
sal-nov-266	2.2	yes	uagcuuaucagacugguguug	cgacaacggucuguaagcuguc	gi|354172448|gb|AGKD01286038.1|:2740..2851:-
sal-nov-267	2.6	yes	uaggaucaguuuugccuuuuaga	agggggaacugauccuagauc	gi|354345670|gb|AGKD01113071.1|:4246..4355:+
sal-nov-268	161.4	yes	uaggcaugauagcacaaacagacug	aucuguuugugccaucuccaug	gi|354446119|gb|AGKD01012923.1|:10126..10237:+
sal-nov-269	1.7	yes	uaggcaugauagcacaaacagacug	gaucaguuuauacucucgccuacu	gi|354339935|gb|AGKD01118806.1|:498..609:+
sal-nov-270	198.8	yes	uagugcacuacuuuugaccaga	caggucaaaaggagugc	gi|354265980|gb|AGKD01192547.1|:2525..2634:+
sal-nov-271	2.4	yes	uagugcacuacuuuugaccaga	aaguaaggcacuaaaua	gi|354315070|gb|AGKD01143458.1|:4668..4778:-
sal-nov-272	2.5	yes	uagugcacuauguaggggaua	uauuccauauauacugcacuau	gi|354233686|gb|AGKD01224841.1|:541..649:-
sal-nov-273	164.2	no	uaguugagucaggugugcaagu	aaacgcaccugaugcaacucau	gi|354408739|gb|AGKD01050297.1|:5000..5107:-
sal-nov-274	64.9	yes	uauaaugggaagauuacugugg	acaguaaucuucccauuauag	gi|354395578|gb|AGKD01063458.1|:6773..6882:-
sal-nov-275	2.5	yes	uauaggacaauacuuuugacc	ucaauaguagugaacuauaaagg	gi|354437073|gb|AGKD01021963.1|:186..296:-
sal-nov-276	2.2	yes	uaugacaguguuaugacagcgu	gcauuauggcaguguuaugac	gi|354236632|gb|AGKD01221895.1|:265..375:+
sal-nov-277	1.2	yes	uaugacaguguuaugacagcgu	gcauuaugacaguguuaugac	gi|354236632|gb|AGKD01221895.1|:958..1068:+
sal-nov-278	1.2	yes	uaugacaguguuaugacagcgu	gcauuaugacaguguuaugac	gi|354236632|gb|AGKD01221895.1|:1186..1296:+
sal-nov-279	0.7	yes	uaugagaacucacacaggagaga	ucuuugugucaguucccuuaug	gi|354311442|gb|AGKD01147086.1|:451..563:-
sal-nov-280	1259	yes	uaugauuggucaaaagaacaau	uggucuuuugaccaaucagauc	gi|354405038|gb|AGKD01053998.1|:3410..3521:+
sal-nov-281	743.4	yes	uauggcuuuuuauuccuauaug	auguagggauagaagccacu	gi|354448902|gb|AGKD01010141.1|:3190..3299:+

sal-nov-282	210.2	yes	uauggguaugauacgaccuggg	caggguuguaucauaaccaagc	gi|354047645|gb|AGKD01410825.1|:951..1062:-
sal-nov-283	2.1	yes	uauggguaugauacgaccuggg	cucaggguuguaucauaaccaac	gi|354358154|gb|AGKD01100882.1|:8460..8570:+
sal-nov-284	2.6	yes	uauguagggaauagggugcca	cccuauuccuaauguagugca	gi|354392844|gb|AGKD01066192.1|:4643..4753:-
sal-nov-285	2	yes	ucaagaguagugcacuuuauagg	uauauagugcacuauucuaugggac	gi|354443392|gb|AGKD01015650.1|:6398..6507:+
sal-nov-286	1.2	yes	ucaagauggacucaaacacagg	ugggguuuggguucaucuuguuc	gi|354299792|gb|AGKD01158736.1|:2377..2485:+
sal-nov-287	64	yes	ucaaggaaugggauucuauuuu	aauggcacccuauucccuguau	gi|354244817|gb|AGKD01213710.1|:249..359:+
sal-nov-288	23350	yes	ucaguaacuggaaucugucccugc	agggccggcugguuacugcgc	gi|354344435|gb|AGKD01114306.1|:3891..4004:-
sal-nov-289	150.6	yes	ucagucuucaguuccucccagc	ugagaggagcugaagacuua	gi|354110537|gb|AGKD01347933.1|:691..802:-
sal-nov-290	161.3	yes	ucagugaacacagguaacagag	ugauaccugugcucacuguagg	gi|354447769|gb|AGKD01011274.1|:14870..14981:-
sal-nov-291	239654	yes	ucagugcauaacagaacuuug	aaguucugugauacacuuaggcu	gi|354361920|gb|AGKD01097116.1|:7327..7434:+
sal-nov-292	6645.9	no	ucagugcauuacagaacuuua	aaguucggugauacacuucgacu	gi|354393765|gb|AGKD01065271.1|:5690..5799:+
sal-nov-293	1.4	yes	ucagugcuacuuguugggguagu	uaccugaacacuagcauugacu	gi|354456619|gb|AGKD01002427.1|:19701..19811:+
sal-nov-294	2.3	yes	ucaucccugauugugaccccc	gggucccaggaaagucaugaga	gi|354092083|gb|AGKD01366387.1|:2317..2426:-
sal-nov-295	0.8	yes	uccaacaucuguuugugcugu	ucaucacaaacagaucugggauc	gi|354450556|gb|AGKD01008489.1|:13698..13808:+
sal-nov-296	1.3	yes	uccaguccggucugucugugga	caucaaagccagauuggacugauc	gi|354378937|gb|AGKD01080099.1|:2551..2658:+
sal-nov-297	0.9	yes	uccaguccggucugucugugga	aacuccaguguggacugagcc	gi|354301016|gb|AGKD01157512.1|:1465..1572:+
sal-nov-298	2.1	yes	uccaguucugucguuaugucagc	gguguaagacgacagaacuggugu	gi|354440113|gb|AGKD01018927.1|:6221..6334:+
sal-nov-299	70	yes	uccaucagucaugugaccuga	gaggucacacgaguggugggcag	gi|354418923|gb|AGKD01040113.1|:7292..7401:+
sal-nov-300	259	yes	uccaugucugaaaagagccagg	uaacucuuuuagaccuuguauu	gi|354440376|gb|AGKD01018666.1|:3645..3755:-
sal-nov-301	1.1	yes	ucccacuucugacacca	augucagagguggagu	gi|354175378|gb|AGKD01283108.1|:752..859:-
sal-nov-302	2.2	yes	ucccagaucuguuuguagcaaug	cagcacaaacagaucugg	gi|354373488|gb|AGKD01085548.1|:1180..1292:-
sal-nov-303	1.8	yes	ucccagaucuguuuguagcaaug	cagcacaagaagaucuggaa	gi|354348661|gb|AGKD01110080.1|:3821..3932:+
sal-nov-304	417.5	yes	ucccagaucuguuugugcuguc	caacacaaacugaucuagga	gi|354452147|gb|AGKD01006898.1|:4577..4686:+
sal-nov-305	2.5	yes	ucccagaucuguuugugcuguc	aacagaucugggaucaggcu	gi|354346773|gb|AGKD01111968.1|:2864..2974:+
sal-nov-306	2.5	yes	ucccagaucuguuugugcuguc	cagcacaacugaucugggaca	gi|354420385|gb|AGKD01038651.1|:1771..1882:+
sal-nov-307	2	yes	ucccagaucuguuugugcuguc	uuuggcacaaacagaucugu	gi|354336349|gb|AGKD01122392.1|:891..999:+
sal-nov-308	645	no	ucccucauucuguacuaaaugg	auucuguucuaaaugggaguuu	gi|354367930|gb|AGKD01091106.1|:1429..1539:+
sal-nov-309	2	yes	uccucgucugcccacauguucu	agcaucuggccagcggggugg	gi|354385901|gb|AGKD01073135.1|:8230..8342:+

sal-nov-310	82	yes	ucggaaucgagucggcucuaca	uagaacagacaucguuuucgg	gi|354435056|gb|AGKD01023980.1|:7882..7993:-
sal-nov-311	2.2	yes	ucucauccaaugggguuuugagc	guacccauugguggagaaggucag	gi|354337952|gb|AGKD01120789.1|:5470..5582:+
sal-nov-312	118.1	yes	ucucugaaaucuacccacugca	cuguggguggcagcucggg	gi|354442832|gb|AGKD01016210.1|:1984..2095:-
sal-nov-313	214.1	yes	ucucugauugggagccauauu	uaugguucucaaucagggaca	gi|354452832|gb|AGKD01006213.1|:6176..6286:+
sal-nov-314	2.2	yes	ucucugauugggagccauauu	ugguuuccaaucagaggcagcu	gi|354216758|gb|AGKD01241728.1|:443..553:-
sal-nov-315	1127.9	yes	ucugauuucagcuaaaugccau	agguaauuaguuguuuuuagga	gi|354190890|gb|AGKD01267596.1|:293..402:+
sal-nov-316	1.7	yes	ucuggucaaaaguagugc	acucauuucgaucagggc	gi|354423732|gb|AGKD01035304.1|:49..159:+
sal-nov-317	0.8	yes	ucuguauucagguucugaa	cugaaccuagaaccacagccc	gi|354256332|gb|AGKD01202195.1|:52..160:+
sal-nov-318	87.1	yes	ucugugcuggcugucuuucaca	uggagagacuggucuugcacaca	gi|354384114|gb|AGKD01074922.1|:4629..4740:-
sal-nov-319	189.6	yes	ucuguguuccucagauuugcau	uuuaucugaggacucucaucuga	gi|354406357|gb|AGKD01052679.1|:2300..2409:+
sal-nov-320	1.8	yes	ucuguguuccucagauuugcau	guuuaucagaggacuucaucugacu	gi|354409850|gb|AGKD01049186.1|:3883..3992:+
sal-nov-321	1.6	yes	ucuguuauuuccuguuga	aacugccugguagagagagc	gi|354321385|gb|AGKD01137143.1|:1537..1644:-
sal-nov-322	79.7	yes	ucuuauuuguagaucagucagu	ugacugaucuacaaauca	gi|354328882|gb|AGKD01129646.1|:432..542:+
sal-nov-323	75.1	yes	ucuuauuuguagaucagucagu	ugacugaucuacaaauca	gi|354442491|gb|AGKD01016551.1|:3290..3401:-
sal-nov-324	96.7	yes	ucuuuugguuaauggauugugg	ucauuccaucuacacaaaaagau	gi|354170757|gb|AGKD01287729.1|:813..924:-
sal-nov-325	135.7	yes	ugaaacaauagugaaauguagugu	acuucuguucauguuuguuuc	gi|354345027|gb|AGKD01113714.1|:7852..7963:+
sal-nov-326	0	yes	ugaacaugcuuuuugguccagg	uggaauaaaaugcacacucaau	gi|354257971|gb|AGKD01200556.1|:1508..1615:+
sal-nov-327	1.2	yes	ugaacuguuucuagucugcuaga	uaguccgcuagacggaacaaaacgaa	gi|354417413|gb|AGKD01041623.1|:6181..6292:+
sal-nov-328	1	yes	ugaaggaucuggagaaguggcaga	ugucacuucuaaucuuucuuu	gi|354437736|gb|AGKD01021300.1|:5899..6012:-
sal-nov-329	2	yes	ugaauaauggaaucagaugugug	caccugauuacacuaaucaau	gi|354419986|gb|AGKD01039050.1|:352..463:+
sal-nov-330	2.3	yes	ugaaugcagcaaucaggcugg	agcacaaucaauguguucacc	gi|354398234|gb|AGKD01060802.1|:1894..2004:-
sal-nov-331	1.4	yes	ugaaugcagcaaucaggcugg	aggcuaguguugcucaag	gi|354426410|gb|AGKD01032626.1|:6637..6747:-
sal-nov-332	136	yes	ugaauggaacagagucauacaug	uguaugauacuguuccauuu	gi|354358027|gb|AGKD01101009.1|:2060..2170:-
sal-nov-333	2.1	yes	ugagaacugaauuccaaggguu	cccaugggcucaguucuuaaa	gi|354375728|gb|AGKD01083308.1|:7962..8069:+
sal-nov-334	272.5	yes	ugagccuaguccuggacuaag	ucaguccaugacuaggcuu	gi|354455877|gb|AGKD01003169.1|:34441..34551:+
sal-nov-335	1	yes	ugaggaaguagguuguaugu	uccaaaccugcauaccuuagu	gi|354354342|gb|AGKD01104694.1|:4237..4345:-
sal-nov-336	6073.6	yes	ugaggcguuuagaacaaguuca	gacuuggucaaagcuccucagc	gi|354324807|gb|AGKD01133721.1|:5133..5242:+
sal-nov-337	0.4	yes	ugaggcguuuagaacaaguuca	gacucuuuuugacgacuacaau	gi|354403045|gb|AGKD01055991.1|:7176..7285:+
sal-nov-338	19767.5	yes	ugagguaguagguuggauaguu	cuguacaaccuacugucuuucu	gi|354412458|gb|AGKD01046578.1|:1442..1550:+

sal-nov-339	69.5	yes	ugaguggaugaaggugugguag	uaacgcacccgaguccacuaau	gi|354351526|gb|AGKD01107510.1|:3149..3259:+
sal-nov-340	364.1	yes	ugaguuggucagcuguguagag	uacacaucugauucaaaucauc	gi|354339001|gb|AGKD01119740.1|:2408..2518:+
sal-nov-341	2.2	yes	ugaugaguuauacucugcaggg	cugcaggggugaugaguuuuc	:..:
sal-nov-342	1.9	yes	ugauguugaaccgggcaguca	gcugcuccaauacagu	gi|354435290|gb|AGKD01023746.1|:2608..2715:+
sal-nov-343	116.9	yes	ugcagcguagucuucaauauggu	ucagacugaacacugggcucacc	gi|354367290|gb|AGKD01091746.1|:1577..1687:+
sal-nov-344	2	yes	ugcacuauauagagguuagg	cccuacacagugcacuacuu	gi|354291647|gb|AGKD01166881.1|:753..862:-
sal-nov-345	0.1	yes	ugauuuccaauaauugagacag	uacauuaaugaaaagaacaaug	gi|354415760|gb|AGKD01043276.1|:16950..17061:+
sal-nov-346	2.8	yes	ugcgagaucaggaugguucgua	aaaaccauccugaucucguaag	gi|354432630|gb|AGKD01026406.1|:7113..7224:+
sal-nov-347	158.9	no	ugcgcacggggccacgcucugc	uaggcgugucacugcgugucaca	gi|354391891|gb|AGKD01067145.1|:7526..7637:-
sal-nov-348	1.3	yes	ugcgccguccguuuuuauugucg	acagugaccguaaacagacgacguuu	gi|354365729|gb|AGKD01093307.1|:4283..4395:-
sal-nov-349	67.1	yes	ugcgcucugaauuuacaaaug	uugugaauuuacgagcgcacuc	gi|354408096|gb|AGKD01050940.1|:7433..7542:+
sal-nov-350	168.1	yes	ugcugauggacuaccugcaga	ugugaggugaugaccaugcagcagg	gi|354402305|gb|AGKD01056731.1|:8290..8402:+
sal-nov-351	357.1	yes	ugcuucacugugaccucauagc	ugugacaucacaaugaauccacagaaug	gi|354302152|gb|AGKD01156376.1|:5094..5204:+
sal-nov-352	1.9	yes	ugcuucacugugaccucauagc	ugugacaucacaaggaauccaca	gi|354401056|gb|AGKD01057980.1|:11434..11545:-
sal-nov-353	474.8	yes	uggacagcaggugguggcgguggu	cgcggaucaucuguuucccagc	gi|354385003|gb|AGKD01074033.1|:6169..6276:+
sal-nov-354	148.8	yes	uggacuaagaagcguuuucaau	ugaacaugcuuuuuagucu	gi|354392180|gb|AGKD01066856.1|:1605..1712:+
sal-nov-355	1.7	yes	uggacuaagaagcguuuucaau	uuaauaugcuuuucggucuagc	gi|354294803|gb|AGKD01163725.1|:3990..4097:-
sal-nov-356	101.8	yes	uggaggggaaugaguggaaugg	uuccauucgcuccguucccg	gi|354439101|gb|AGKD01019939.1|:8462..8573:-
sal-nov-357	356.2	yes	uggccaagcacaaacagaucu	gucuguuucugccaaggcuuga	gi|354392275|gb|AGKD01066761.1|:1353..1462:+
sal-nov-358	2.4	yes	ugggaucuguugauuugucugcgu	acagauguuggcagaccacc	gi|354371690|gb|AGKD01087346.1|:2943..3053:-
sal-nov-359	118.5	yes	uggucaaaaguaguacacuaugg	cuccuguacuucuuuugaccagg	gi|354415042|gb|AGKD01043994.1|:408..519:+
sal-nov-360	96.6	no	uggucaaauaucuguugaagug	cuacaaugaucugggaccgg	gi|354391298|gb|AGKD01067738.1|:1098..1209:-
sal-nov-361	1.3	yes	uggucuuuugacugaucagauc	ucugaucauuugccacu	gi|354425710|gb|AGKD01033326.1|:8787..8898:-
sal-nov-362	21260.4	yes	uguaaacaucccgacuggaauc	cuuucagucaggugguugcug	gi|354423667|gb|AGKD01035369.1|:4659..4769:+
sal-nov-363	265819.3	yes	uguaaacauccuugacugaaagc	uuucagucggauguuugcagc	gi|354362733|gb|AGKD01096303.1|:488..599:+
sal-nov-364	54.8	yes	uguaccaugcugguagccagu	ugggaaccaguguugugcuac	gi|354327145|gb|AGKD01131383.1|:3423..3533:-
sal-nov-365	175.2	yes	uguaccuauguggguguuugugg	uccuacacccauacaggacacu	gi|354224874|gb|AGKD01233653.1|:859..971:-
sal-nov-366	2.3	yes	uguaccuauguggguguuugugg	cuacacccauacaggaca	gi|354103671|gb|AGKD01354799.1|:157..268:+

sal-nov-367	2.6	yes	uguauccccugagacugagacu	gcucagccuggggggagaugcagg	gi|354424201|gb|AGKD01034835.1|:6474..6586:-
sal-nov-368	1.4	yes	uguauuacaguauauagggaau	ucccuacauauagugcacu	gi|354421046|gb|AGKD01037990.1|:8624..8733:+
sal-nov-369	2.2	yes	ugucacugacuguugaaaucag	uucaacagucagugacucaggc	gi|354417874|gb|AGKD01041162.1|:4176..4286:-
sal-nov-370	240	yes	ugucucagaccuguuugugcug	gucacaacggaucugggaucagu	gi|354348584|gb|AGKD01110157.1|:1726..1837:-
sal-nov-371	498	yes	ugucuuucaaagcauagcuugg	caacuguuuuugagggccauga	gi|354152828|gb|AGKD01305658.1|:1669..1779:+
sal-nov-372	441.6	yes	ugucuuucaaagcauagcuugg	caacuauuuuugagggccacg	gi|354143605|gb|AGKD01314881.1|:1655..1766:+
sal-nov-373	47608.9	no	ugugcaaacccaagcaaaacuga	agcuuugcggggugggcagucagc	gi|354452722|gb|AGKD01006323.1|:1295..1405:+
sal-nov-374	1575.8	yes	ugugcaaauccaugcaaaacucu	aguuuugcaggguugcuuucagc	gi|354260329|gb|AGKD01198198.1|:4711..4820:-
sal-nov-375	65.4	yes	ugugcagauagaccgacugaag	gucaguuagucuaucuuauacagc	gi|354329384|gb|AGKD01129144.1|:7186..7295:-
sal-nov-376	217.2	no	ugugcccaugucugaugugaaac	aucacagucaaacaggcuccuacc	gi|354423573|gb|AGKD01035463.1|:14332..14443:-
sal-nov-377	1.3	yes	ugugugugugugugu	auauauagucuuacu	gi|354269892|gb|AGKD01188635.1|:2023..2137:-
sal-nov-378	2.3	yes	uguguuaggccuccgagucug	gcgucucagaggucaaacacagu	gi|354449383|gb|AGKD01009660.1|:24121..24231:-
sal-nov-379	0	yes	uguuaaaugugauauuguagaau	ucuacaccagauucuauugu	gi|354454466|gb|AGKD01004580.1|:1819..1931:-
sal-nov-380	1.5	yes	uguuauuuguaaaucagucagu	ugacagaucuacaaauaauagu	gi|354456510|gb|AGKD01002536.1|:5812..5923:-
sal-nov-381	2	yes	uuaaacugcagucugacu	gcauacugcaguuagacu	gi|354421218|gb|AGKD01037818.1|:4461..4568:-
sal-nov-382	0.2	yes	uuaaauuaggaguagaggcaga	guuuucuacugcuaaauuuauuu	gi|354420292|gb|AGKD01038744.1|:16800..16911:-
sal-nov-383	0.6	yes	uuacaauuaaaggauauuucuug	ggaaaucucucuuaauuguuugg	gi|354364716|gb|AGKD01094320.1|:2588..2700:-
sal-nov-384	0.3	yes	uuacaauuaaaggauauuucuug	ggaaauguccccuuaauuguuugg	gi|354370269|gb|AGKD01088767.1|:4077..4189:-
sal-nov-385	0.4	yes	uuaccguaguaacggacuguuacu	caacagacuauuacugcaacu	gi|354026448|gb|AGKD01432022.1|:292..405:-
sal-nov-386	2.2	yes	uuaggccaggacuaggcuuaug	aaagcuuaguccuggauuaaaa	gi|354405890|gb|AGKD01053146.1|:8591..8701:-
sal-nov-387	176.4	no	uuaggguuaggguuagg	uuaggguuaagguuaggg	gi|354261929|gb|AGKD01196598.1|:2605..2712:+
sal-nov-388	643.4	yes	uucacuauaucgggaauagggu	cccuauucccuguaaagugca	gi|354192538|gb|AGKD01265948.1|:850..958:+
sal-nov-389	1	yes	uucuaaguggaaugcuagugugg	auagcauacuacucagaguaua	gi|354401484|gb|AGKD01057552.1|:4763..4875:-
sal-nov-390	0.7	yes	uucuguccagugaaaugcuguagc	ugcaguuuuauagaccaaugc	gi|354449168|gb|AGKD01009875.1|:11122..11234:+
sal-nov-391	79.8	yes	uucugucccacugagccuggca	gccgguagugaggagcagaa	gi|354412458|gb|AGKD01046578.1|:12837..12946:+
sal-nov-392	66	no	uugaagaaacaggaaguauaga	uuuuuguuguugcuucu	gi|354413112|gb|AGKD01045924.1|:1139..1249:-
sal-nov-393	124.6	no	uugacuaggacaaucaggcga	ccugauucaacuaaucaucaag	gi|354443736|gb|AGKD01015306.1|:13263..13372:+
sal-nov-394	2393.8	yes	uugauuauuuaaaucggcugug	cagcugauucaaauaaccaacu	gi|354427434|gb|AGKD01031602.1|:14155..14266:-
sal-nov-395	2.9	yes	uugauuauuugaaucagcug	gcugauucaaauacccaacu	gi|354434162|gb|AGKD01024874.1|:1783..1892:-

sal-nov-396	2515.3	yes	uugauuauuugaauuggcugug	cagcugauucaaauaaccaacu	gi|354389165|gb|AGKD01069871.1|:8249..8360:-
sal-nov-397	2.9	yes	uugauuauuugaauuggcugug	cacagcugauucaaauaa	gi|354293853|gb|AGKD01164675.1|:50..161:+
sal-nov-398	2.1	yes	uugcacaccguuagaacccagc	ggguguaggaugugugcuacagaaacgca	gi|354339445|gb|AGKD01119296.1|:6769..6879:-
sal-nov-399	2.1	yes	uugcccacauucuccacca	guucaagggugugggguuuc	gi|354367655|gb|AGKD01091381.1|:10289..10398:-
sal-nov-400	971.3	yes	uugcugauuacacugucuguga	aaggcacaguuaaccagcaugg	gi|354269531|gb|AGKD01188996.1|:1482..1593:+
sal-nov-401	1.7	yes	uugcuggugaaauugucuguga	aagggauauuuaaccagcaugu	gi|354091020|gb|AGKD01367450.1|:156..266:+
sal-nov-402	340.8	no	uugcuggugacauugccuguga	cagaguaggugaacggauuau	gi|354420722|gb|AGKD01038314.1|:10347..10458:+
sal-nov-403	397.1	yes	uugcuggugagacugucuguga	aagaaacacuuaaccagcaugg	gi|354306704|gb|AGKD01151824.1|:2042..2153:+
sal-nov-404	2091.4	yes	uugcugguuaaacugucuguga	aaggcacacuuaaccagcaugg	gi|354404616|gb|AGKD01054420.1|:6542..6652:-
sal-nov-405	1822.1	yes	uugcugguuacauggucuguga	aaggcacacuuaaccagcaugg	gi|354023037|gb|AGKD01435433.1|:42..153:-
sal-nov-406	129.6	no	uuggagguaaauuagcagaaugg	aaccugccagucccccuucauaca	gi|354449514|gb|AGKD01009529.1|:2203..2313:+
sal-nov-407	2.7	yes	uugggcagggcuguguuuuuga	aaaagcaagcccuggccgaca	gi|354410116|gb|AGKD01048920.1|:1385..1495:+
sal-nov-408	1.9	yes	uuggucuuguggaugggagaua	uguccccacuugcaagagcuaau	gi|354351556|gb|AGKD01107480.1|:4927..5035:+
sal-nov-409	2.4	yes	uugguuauuugaaucggcugug	cagcggauuaaaauaaucaacu	gi|354444825|gb|AGKD01014217.1|:2366..2478:-
sal-nov-410	406.1	yes	uugucucugauuggggagcaua	ugguucccaauaagaggcagcu	gi|354383168|gb|AGKD01075868.1|:5365..5472:+
sal-nov-411	80.2	yes	uuguuggugacaaugucuguga	caagacacacuuaaccagcauggcu	gi|354443844|gb|AGKD01015198.1|:12095..12205:+
sal-nov-412	337	yes	uuguuggugacauugucuguga	caagacacacuuaaccagcauggcu	gi|354419031|gb|AGKD01040005.1|:2910..3021:-
sal-nov-413	53	yes	uuuaaaaugugaacuauaccu	ggauaauucacuuuuugaagu	gi|354452752|gb|AGKD01006293.1|:5076..5186:-
sal-nov-414	600.4	yes	uuuaacaugcaccuguggaacgg	uguuucccaugcuuguucaau	gi|354087696|gb|AGKD01370774.1|:12..124:+
sal-nov-415	1.9	yes	uuuaacgucaguggucaugaag	acauaacccugauacuuaaaau	gi|354452389|gb|AGKD01006656.1|:16699..16810:+
sal-nov-416	2754.2	yes	uuuauuauuugaaucggcugug	acagcugauuaaaauaauc	gi|354422157|gb|AGKD01036879.1|:345..456:-
sal-nov-417	2	yes	uuuauuauuugaaucggcugug	cagcugauuucauaucaaagcu	gi|354426489|gb|AGKD01032547.1|:830..941:-
sal-nov-418	2	yes	uuuauuauuugaaucggcugug	cacaacugauuuaaauaa	gi|354457889|gb|AGKD01001160.1|:33720..33829:+
sal-nov-419	483.6	yes	uuucugaccaugugaccuggggg	ucaggucacauguucaggaua	gi|354452966|gb|AGKD01006079.1|:5407..5518:-
sal-nov-420	156.1	yes	uuugaaagggcaacaauacuc	cuauuguugucauaccauaca	gi|354352111|gb|AGKD01106925.1|:10638..10748:-
sal-nov-421	2.5	yes	uuugaaccagcuguguagugc	acuacacagcugauucuaauc	gi|354402419|gb|AGKD01056617.1|:216..326:+
sal-nov-422	0.7	yes	uuugaauguuaacagacacuuu	ggugucuuuugauauuucauauu	gi|354451551|gb|AGKD01007494.1|:624..731:+
sal-nov-423	2.3	yes	uuugacauugugcugggagaca	ucucccagcacaacaacgucaacag	gi|354348345|gb|AGKD01110396.1|:1773..1882:+

sal-nov-424	0.2	yes	uuugagucagcuguguagugc	acuacacaucuuauucaaagc	gi|354306015|gb|AGKD01152513.1|:6830..6942:-
sal-nov-425	2.3	yes	uuugguucucaauuagaggca	cucuaauuggggaucaa	gi|354338378|gb|AGKD01120363.1|:4357..4468:-
sal-nov-426	66.1	yes	uuuugaaugguguuauugugugu	accagucaccccaucaggacc	gi|353957833|gb|AGKD01500637.1|:166..278:-
sal-nov-427	290.9	yes	uuuugaauguagccuagugcu	cacuaggcugcauucaaaa	gi|354383266|gb|AGKD01075770.1|:620..730:-

Table S 6.4 - Background data for figures 3 and 4. Normalized number of reads in all 8 samples for miRNAs that showed differential expression between morphs and/or developmental time-points.

ID	SB1	SB2	SB3	SB4	AC1	AC2	AC3	AC4
let-7a-5p	15.55162557	17.16718146	17.75712637	18.78763299	15.87325519	17.06750347	17.95481125	19.37568089
let-7b	7.451093021	10.82952683	12.4159762	12.91900645	7.147321398	9.531741259	11.30424829	13.66734152
let-7d-5p	13.81152187	15.57314152	16.50489514	17.53831875	13.7731879	14.74660243	16.03425346	18.02164096
let-7e-3p	2.768323414	5.040373538	7.051450943	8.213559793	3.696639123	5.735471991	8.196129368	8.319771158
let-7e-5p	11.72159152	13.5529306	15.26170262	18.09145325	11.87257248	13.23229858	15.7969909	18.41593891
let-7g	13.16141442	13.45957531	13.50479003	15.62504546	13.83030749	14.38570234	15.59768786	16.18143483
let-7h	8.992455945	9.713884013	10.24546783	13.48180782	9.335692572	10.29603189	11.3974233	14.33103323
let-7i-5p	7.628989894	8.861287072	10.17438589	13.78090578	8.004197141	8.940827091	10.81992513	14.30826113
let-7j-3p	4.020216046	4.97087762	5.599320452	7.896947271	4.156736453	6.1053007	6.859200804	8.198648786
let-7k-5p	4.232803926	5.42096856	6.858202366	10.38141816	3.953912382	4.508818085	7.886681927	9.175050522
miR-1	8.76866889	10.70220339	9.586686581	10.79347173	7.551376753	7.987475886	8.677961884	9.728110431
miR-1-3p	19.75213532	20.70495992	21.23559459	22.48130854	19.77946571	20.49561561	21.25726441	22.62951414
miR-1-5p	6.999206207	6.286603805	6.833159295	8.744996226	8.378064297	9.253180573	9.819244353	8.836541123
miR-101b-3p	15.28686125	15.81320408	16.41022011	18.05023985	15.66679352	15.78403578	16.53518903	18.02785307
miR-10b	7.980729848	7.7542285	7.332937526	6.374294489	6.800943192	6.445480798	6.182949377	5.395716756
miR-122-5p	5.902283369	6.526654444	6.415627892	6.744534662	5.028633279	4.770649166	4.884757455	5.978184715
miR-124	9.76263877	10.15555588	9.845696052	13.17076432	10.88634709	10.91391143	11.66117136	14.1176081
miR-129-2-3p	6.168969455	6.340613889	6.46004518	9.963472715	6.52279016	7.801031736	8.492553587	10.88982716
miR-129b-5p	9.897603307	11.63661785	12.57730635	13.48756535	9.953682612	10.58824634	11.11111749	14.18180381
miR-130b	15.57875937	15.14774023	15.15358897	15.70255201	16.4138879	16.14585338	16.35785578	16.43536858
miR-130b-5p	6.124076226	6.157002656	6.248249609	5.512496463	7.127328618	7.565651556	6.970234587	6.826758351
miR-130c	15.21386939	14.85963899	14.99621389	14.96781293	15.96601807	15.78335678	16.04550998	15.72229271
miR-132a	6.430554057	7.415865348	8.33621106	10.20471933	6.234476411	6.468566154	7.389367995	10.17263638
miR-133	11.93217581	11.19640303	11.40728545	13.98779563	12.96397593	13.64959963	14.08884185	14.26359149

miR-138	11.12031285	11.98926422	12.72239646	13.77044831	11.2736831	12.01832814	12.88206358	14.03459081
miR-1388-5p	7.035787211	6.838772904	7.598197075	10.63648809	7.401425351	7.32244955	7.72757713	10.82483251
miR-140-3p	15.01776435	15.66418215	16.53076256	18.0727436	15.73672909	16.34701662	16.87307074	18.32362797
miR-140-5p	13.52892597	14.24757081	14.89691144	16.00093398	14.10267776	14.72941345	15.49746417	16.44623382
miR-143	9.44314268	10.3230552	10.87277037	13.8631746	10.23805347	10.31830845	11.4735072	14.56125361
miR-145	9.872610498	8.701402981	9.092936972	12.81878945	10.92657792	11.73815779	11.24616646	13.13828938
miR-146a-5p	5.538583609	5.758995234	5.968619157	6.209340939	8.217449568	8.159603171	8.134053145	8.442210557
miR-150	7.76571991	6.157002656	5.47292272	8.349102664	8.240856501	8.680746423	7.677655027	9.654053371
miR-153-3p	11.76252567	11.18545753	10.78168249	11.99389172	12.89025268	13.05034987	13.41142934	12.3502365
miR-1692	7.477893696	4.274819484	5.381715751	2.768323414	7.965371661	6.134567547	6.76647011	3.64131647
miR-1692	11.00927802	11.90864284	12.24242526	12.39813432	12.61186053	13.99971475	12.78425742	13.34157939
miR-17-5p	11.55145954	10.68157803	10.51695786	11.14470701	13.04229295	12.91229355	12.78614969	11.37299949
miR-181a	8.992455945	6.721966346	6.737395446	5.934402919	4.935468415	4.689335451	5.081440964	4.77401557
miR-181b-5p	8.896519052	6.17200163	6.222644069	5.383397931	5.822736564	5.052583552	5.169815671	4.77401557
miR-181c	13.42618265	13.77514015	13.59762914	10.96927431	13.64226801	13.81329309	13.50286711	11.07664778
miR-182	7.698977592	8.342557092	8.644358594	10.89612466	7.434319276	7.522387638	8.517076081	11.50779349
miR-183-5p	15.02600442	16.02370596	16.45790371	17.69471676	15.11748586	15.24448935	15.80753574	17.77686667
miR-192	17.05572419	17.33679135	17.32619818	19.84533279	17.43069841	17.14635721	17.57565149	20.13979436
miR-194-5p	14.00773117	13.6603436	13.49653544	15.76791538	14.18940338	14.05488824	13.95665394	15.88346124
miR-1940	8.925976119	7.325538038	5.999206879	7.417226468	10.07319401	10.20673355	9.178266896	8.608336174
miR-196a	6.393311143	6.405364654	6.142931118	5.337513556	5.028633279	5.572419971	4.518322935	4.445088471
miR-199a-3p	14.01326808	13.92619665	13.74546206	15.75802478	14.74252848	15.06775665	15.38476808	16.24130711
miR-199b-3p	11.91432675	11.32539834	11.12363721	13.40867332	12.60240903	12.85548765	13.06706443	13.88296343
miR-19b	14.24991639	13.51080362	12.82392666	12.94108758	15.24610405	14.90644749	14.53321443	13.09080937
miR-19b-5p	6.81495211	6.491338395	5.999206879	5.704003461	8.53776004	7.897268495	7.886681927	6.064798546
miR-19d	9.130994571	8.378256639	7.796471572	7.784932882	10.6001245	10.62189797	9.769096063	8.579520692
miR-200c	5.271807049	5.229414251	6.058461845	8.250165972	5.157134422	5.615011052	5.541556656	8.382290458
miR-206	9.92054561	11.85698049	10.21841796	9.151918851	8.49946375	8.816642251	9.004198135	8.608336174
miR-206-5p	9.364052245	9.348869495	9.85605734	10.02810981	10.90916044	11.31062032	11.58350531	10.37897185

miR-212-3p	4.325336144	4.449420655	4.83701006	8.35471628	5.028633279	5.615011052	6.52880172	9.155662255
miR-212-5p	5.091973752	5.446265565	6.183347835	9.151918851	5.236424268	5.615011052	6.030796594	9.06849752
miR-215	10.66567684	10.71494608	10.73865439	13.49901195	10.44358518	10.36644413	10.75871001	13.38349867
miR-2188-5p	9.701504573	9.593455341	10.10657851	10.89900778	10.6305558	11.83900166	12.44897267	11.30936407
miR-219	10.86552636	9.097635173	7.912299365	7.889228832	12.38576184	11.78289642	10.8338154	8.51503835
miR-219a	14.93804707	14.06560149	12.95910683	12.28145963	15.81796508	14.1203068	12.47657746	11.98965277
miR-219b	11.8846343	10.98753031	9.617465452	6.209340939	12.51059327	11.38033955	9.767394054	6.724803743
miR-223-5p	9.429496952	9.572602258	9.492860148	11.31339955	7.521080857	7.81005117	7.492017432	10.95482615
miR-22a	14.3099467	13.98983726	14.36892831	17.65454127	14.38615721	14.51247468	14.67021574	18.1605105
miR-2404	5.538583609	4.778990818	3.755420608	4.112384304	7.096807585	6.163241635	6.921944664	3.985428795
miR-2478	5.928722034	5.633488579	4.801019421	6.105689937	7.474406168	7.244674843	7.095215402	6.952306559
miR-2487	7.414563813	7.030939716	4.260169007	3.978631984	6.654689225	6.738957653	4.44509393	4.6204733
miR-2488	13.30920737	12.7237717	9.913736436	10.16694454	14.1489353	14.52098884	14.76952277	10.77064261
miR-26a	15.30592782	15.01180548	14.96253875	16.67420995	16.08832098	16.16702227	16.37908098	16.89104294
miR-26a-2-3p	4.490710775	4.693404219	4.872021897	7.246319152	5.822736564	5.230438377	5.724329273	8.002690165
miR-26a-5p	6.484654562	7.7542285	9.355003745	12.73602445	7.107053525	8.281291342	10.00836276	12.93465586
miR-26b	4.939142338	5.106402544	5.80483058	6.050871504	6.032765911	5.949118392	7.942088984	6.742308493
miR-2779	5.980154508	5.471105197	5.619308385	8.867832093	5.919971907	6.34920595	5.211910485	10.837054
miR-27c-3p	11.10751686	11.31748262	11.81539299	13.71944135	11.69339618	11.89882538	12.30535731	14.35831763
miR-2995	6.728102997	5.006083399	5.178669292	5.136686117	8.179189286	8.072366036	7.424404845	5.752579485
miR-301c	5.791250741	5.286950938	5.064388114	8.670444049	5.634403905	5.949118392	4.714834405	10.78753427
miR-30b	12.64606472	11.38919643	11.23975792	12.62751074	13.78507269	14.1058584	13.75704763	12.91413281
miR-3618	6.412054323	5.817756958	5.122750021	5.337513556	7.474406168	7.384194312	7.278823433	5.643521845
miR-365-3p	7.660516367	8.107521811	8.318330295	10.72471771	8.639874897	9.424335332	9.718903723	11.99238724
miR-430a	14.96327177	14.08597676	13.63307885	9.450709348	15.17619153	13.71597388	12.34526421	8.892274864
miR-430b	11.49417138	10.04282776	9.855024554	5.872344995	10.80156831	8.903418743	8.648669312	4.445088471
miR-430c	16.63398471	15.96810882	15.38791821	11.61286567	16.31855121	15.06622966	14.10939204	10.67629759
miR-430d	6.502241311	5.738846213	4.560480872	3.386637256	7.012095415	5.981768652	5.126358453	3.390256856
miR-451	15.70832342	16.56144378	16.24089446	16.43934264	17.47405204	17.77044661	19.6605319	16.24324553

miR-454-5p	6.275351271	6.09536571	5.151001258	3.978631984	7.137359803	7.231293095	6.696352032	5.395716756
miR-455-3p	7.414563813	7.022741405	7.284573983	9.138968011	8.840079322	9.236564917	9.066432672	9.93557671
miR-459-3p	4.635992375	5.286950938	6.653977356	10.54496529	4.327152748	5.48309015	6.592025894	10.97524338
miR-459-5p	5.819843907	6.405364654	7.290709173	10.72904467	5.575559619	6.535661227	7.966046411	11.5840993
miR-462	9.707170535	10.29532378	10.43639377	13.33412	9.222721065	9.216370118	9.381092051	12.29112936
miR-5106	7.743814454	5.106402544	6.209665923	3.978631984	9.187028988	7.309775865	8.549137084	5.717166371
miR-6239	12.55684127	11.36877998	10.86509636	17.02763834	11.48876848	12.0949812	9.591952567	16.75524125
miR-6240	6.448816665	4.141353128	3.927735096	8.740715897	4.726326214	5.48309015	3.325448819	7.913102283
miR-6516-3p	5.819843907	4.141353128	4.369074255	5.023750031	7.033746857	6.620480398	6.320340209	6.248385873
miR-6673-3p	5.091973752	5.073798002	4.763987433	3.636308541	6.215549646	6.044884947	6.301508063	6.198313684
miR-716b	6.974290152	7.178420645	7.10817014	5.2900305	6.075473973	5.284856188	5.169815671	4.974441236
miR-725	11.893772	12.35898962	13.03379492	15.03280433	12.25399207	12.49697198	13.35730235	14.70751618
miR-737	6.667129676	5.395196176	5.093888633	7.106656992	7.860446846	7.628202187	7.508435429	7.556584661
miR-7551	8.825351571	7.458986732	4.686514533	3.978631984	7.34198578	5.981768652	4.652730127	4.239049627
miR-7641	6.635639498	5.342147494	5.47292272	10.28332163	6.196366952	7.489063962	5.960213241	11.90704512
miR-7876-3p	4.410934372	4.897494783	5.65844139	4.526631728	6.215549646	6.044884947	6.182949377	5.525172089
miR-9-3p	8.980027146	6.691211881	6.209665923	5.470796268	5.690856839	5.436156249	5.474849871	4.239049627
miR-9a-5p	7.512868462	7.646489416	8.24450286	6.778462132	6.394345712	6.398156864	6.52880172	5.853737235

File S 6.1: R code used for differential expression analysis of miRNA-seq data

```
#! /usr/bin/Rscript
library(edgeR) #the package
USEDB <- FALSE
library(reshape)

if (USEDB){
    require('RPostgreSQL')
    drv <- dbDriver("PostgreSQL")
    con <- dbConnect(drv, dbname="smallrna")

    counts <- dbGetQuery(con,"
                        SELECT
                           substr(id,1,20) AS first20bases,
                           sum(count1) AS count1,
                           sum(count2) AS count2,
                           sum(count3) AS count3,
                           sum(count4) AS count4,
                           sum(count5) AS count5,
                           sum(count6) AS count6,
                           sum(count7) AS count7,
                           sum(count8) AS count8
                        FROM
                           seqhomomircount
                        GROUP BY
                           first20bases
                        ORDER BY
                           first20bases;")

   # max_sw_score  <- dbGetQuery(con,"
   #                         SELECT
   #                            substr(id,1,20) as first20bases,
   #                            max(bitscore)
```

```
  #                       FROM
  #                          mirbasesw
  #                       GROUP BY
  #                          first20bases;
  #                       ")

   writeToDataBase <- function(x,name,conn){
         if(dbExistsTable(conn, name)){
            dbRemoveTable(conn, name)
            dbWriteTable(conn,name,x)
         }else{
               dbWriteTable(conn,name,x)
         }
   }

   #READ in data
   #counts <- read.table("homology_counts_grouped_on_20_firstbases.txt",header=T,row.names=1)

   write.table(counts,"counts.txt")
}else{
   counts <- read.table("counts.txt")
}

#The entries here are sequence that have occurences over 20
rownames(counts) <- counts$first20bases
counts_long_format <- melt(counts,id="first20bases")
counts <- counts[,paste("count",1:8,sep="")]
colnames(counts) <- paste("s",1:8,sep='')



#Filter out contigs with fewer than 20 reads for all samples
more_than_80 <- DGEList(counts[rowSums(counts)>80,])
```

```
pdf("MDS_more_than_80.pdf")
   plotMDS(more_than_80)
dev.off()

counts_cpm3 <- counts[rowSums(cpm(counts)>3)>2,]
cpm3_exp2 <- DGEList(counts_cpm3)
pdf("cpm3_exp2.pdf")#This looks better based on the multidimensional scaling
   plotMDS(cpm3_exp2)
dev.off()




da <- DGEList(cpm3_exp2)

cat("Number of tags used",nrow(da),"\n")
write.table(counts_cpm3,"counts_cpm3.txt",quote=FALSE,col.names=FALSE)


da <- calcNormFactors(da)
design <- model.matrix(~factor(rep(1:4,2))+factor(rep(0:1,each=4)))#design matrix
colnames(design) <- c("mu",paste("T",2:4,sep=""),"morph")#Name the colums

#Estimate dispersion,
da <- estimateGLMCommonDisp(da,design,method="deviance", robust=TRUE,subset=NULL)
da <- estimateGLMTrendedDisp(da,design)
da <- estimateGLMTagwiseDisp(da,design)


#Fitting
fit <- glmFit(da,design)

#Testing
lrtM <- glmLRT(fit)             #Drop the morph term
```

```
lrtT <- glmLRT(fit,coef=2:4)      #Drop the time term

pdf("smear_plot_morph.pdf")
    plotSmear(lrtM)
dev.off()

write.csv(topTags(lrtM,100),file="first20bases_morphDEC.csv")
write.csv(topTags(lrtT,100),file="first20bases_timeDEC.csv")

if (USEDB){
    writeToDataBase(as.data.frame(topTags(lrtM,100)),"morphdec",con)
    writeToDataBase(as.data.frame(topTags(lrtT,100)),"timedec",con)
}
```

File S2: R code used to calculate relative expression (fold change) for each miRNA compared to stage 1 in AC.

```
library(ggplot2)
library(gridExtra)
library(reshape)

deltaDeltaTransform <- function(dat,geneList,all_gene=FALSE){
   for (gene in geneList){
      diff_gene <- paste(gene,"_diff",sep="")
      norm_gene <- paste(gene,"_norm",sep="")
      if  (!all_gene){
         dat[[diff_gene]] <- dat[[gene]]-dat[["U2"]]
      }else{
         dat[[diff_gene]] <- dat[[gene]]-
apply(dat[,c("X196a","X199a","X181a","X140","X30b","X26a","X206","X17")],1,function(x){exp(mean(log(x)))})
      }
      diff_gene_mean <- mean(subset(dat,Time.point=="150" & Morph=="AC" )[[diff_gene]])
      dat[[norm_gene]] <- 2^(-(dat[[diff_gene]]-diff_gene_mean))
   }
   return(dat)
}

qpcr_dat <- read.table("QPCR_ct.csv",sep=",",header=TRUE)
geneList <- c("X196a","X199a","X181a","X140","X30b","X26a","X206","X17")
qpcr_dat <- deltaDeltaTransform(qpcr_dat,geneList,all_gene=TRUE)
#qpcr_dat <- deltaDeltaTransform(qpcr_dat,geneList,all_gene=FALSE)

# Now the hts-part
sequences <-
c(X196a="TAGGTAGTTTCATGTTGCTG",X199a="ACAGTAGTCTGCACATTGGT",X181a="ACCATCGACCGTTGATTGTA",X140="TACCA
CAGGGTAGAACCACG",X30b="TGTAAACATCCTACACTCAG",X26a="TTCAAGTAATCCAGGATAGG",X206="TGGAATGTAAGGAACTGT
GT",X17="CAAAGTGCTTACAGTGCAGG")

seq2names  <- names(sequences)
```

```
names(seq2names) <- sequences
if (TRUE){
   hts_dat <- read.table("counts_cpm3_vsd.txt")
}else {
   hts_dat <- read.table("counts_cpm3.txt")
   rownames(hts_dat) <- hts_dat[,1]
   hts_dat <- hts_dat[,-1]
   colnames(hts_dat) <- c("FISH1.t1","FISH1.t2","FISH1.t3","FISH1.t4","FISH2.t1","FISH2.t2","FISH2.t3","FISH2.t4")
}

#Transform everything to have the same names
sel_hts_dat <- hts_dat[rownames(hts_dat) %in% sequences,]
sel_hts_dat$seqs <- rownames(sel_hts_dat)
sel_hts_dat <- transform(sel_hts_dat,names=seq2names[seqs])
me_sel_hts_dat <- melt(sel_hts_dat)
me_sel_hts_dat <- transform(me_sel_hts_dat,Time.point=sub("FISH[12]\\.t","",variable,perl=TRUE))
me_sel_hts_dat <- transform(me_sel_hts_dat,Morph=sub("\\.t[1234]","",variable,perl=TRUE))
me_sel_hts_dat <- transform(me_sel_hts_dat,Morph=ifelse(Morph=="FISH1","SB","AC"))


plot_miRNAs <- function(x="X181a",xseq="AAA"){
   plot_hts_dat <- subset(me_sel_hts_dat,names==x)
   p_hts <- ggplot(data=plot_hts_dat)+
         geom_point(aes(Morph,value,col=Morph))+
         facet_wrap(~Time.point,ncol=4)+
         theme_bw()+
         xlab(NULL)+
         ylab("VST")
   qpcr_dat_cp <- qpcr_dat
   qpcr_dat_cp[["plot_val"]] <- qpcr_dat[[paste(x,"_norm",sep="")]]
   p <- ggplot(data=qpcr_dat_cp)+geom_boxplot(aes(Morph,plot_val,fill=Morph)) +facet_wrap(~Time.point)+theme_bw()+ylab("Fold
change")
   grid.arrange(p_hts,p,main=paste(x,xseq))
}
```

```
pdf("comparison.pdf")
    for (name in names(sequences)){
        print(name)
        plot_miRNAs(name,sequences[name])
    }
dev.off()


u2 <- ggplot(data=qpcr_dat)+geom_boxplot(aes(Morph,U2,fill=Morph)) +facet_wrap(~Time.point)+theme_bw()
pdf("U2.pdf")
    plot(u2)
dev.off()
```