Discovery of miRNAs and Their Corresponding miRNA Genes in Atlantic Cod (*Gadus morhua*): Use of Stable miRNAs as Reference Genes Reveals Subgroups of miRNAs That Are Highly Expressed in Particular Organs

Rune Andreassen[1]*, Fredrik Rangnes[1], Maria Sivertsen[1], Michelle Chiang[1], Michelle Tran[1], Merete Molton Worren[2]

1 Department of Pharmacy and Biomedical Laboratory Sciences, Faculty of Health Sciences, Oslo and Akershus University College of Applied Sciences, Oslo, Norway, **2** Bioinformatics Core Facility, Institute for Medical Informatics, Oslo University Hospital, Oslo, Norway

* Rune.Andreassen@hioa.no

Abstract

Background

Atlantic cod (*Gadus morhua*) is among the economically most important species in the northern Atlantic Ocean and a model species for studying development of the immune system in vertebrates. MicroRNAs (miRNAs) are an abundant class of small RNA molecules that regulate fundamental biological processes at the post-transcriptional level. Detailed knowledge about a species miRNA repertoire is necessary to study how the miRNA transcriptome modulate gene expression. We have therefore discovered and characterized mature miRNAs and their corresponding miRNA genes in Atlantic cod. We have also performed a validation study to identify suitable reference genes for RT-qPCR analysis of miRNA expression in Atlantic cod. Finally, we utilized the newly characterized miRNA repertoire and the dedicated RT-qPCR method to reveal miRNAs that are highly expressed in certain organs.

Results

The discovery analysis revealed 490 mature miRNAs (401 unique sequences) along with precursor sequences and genomic location of the miRNA genes. Twenty six of these were novel miRNA genes. Validation studies ranked *gmo*-miR-17-1—5p or the two-gene combination *gmo*-miR25-3p and *gmo*-miR210-5p as most suitable qPCR reference genes. Analysis by RT-qPCR revealed 45 miRNAs with significantly higher expression in tissues from one or a few organs. Comparisons to other vertebrates indicate that some of these miRNAs may regulate processes like growth, lipid metabolism, immune response to microbial

infections and scar damage repair. Three teleost-specific and three novel Atlantic cod miR-NAs were among the differentially expressed miRNAs.

Conclusions

The number of known mature miRNAs was considerably increased by our identification of miRNAs and miRNA genes in Atlantic cod. This will benefit further functional studies of miRNA expression using deep sequencing methods. The validation study showed that stable miRNAs are suitable reference genes for RT-qPCR analysis of miRNA expression. Applying RT-qPCR we have identified several miRNAs likely to have important regulatory functions in particular organs.

Introduction

Atlantic cod (*Gadus morhua*) is an economically important marine fish species for wild fisheries of the northern Atlantic with more than 1 million tonnes captured in commercial fisheries in 2014 [1]. Although there are challenges, the Atlantic cod is also considered an interesting species to the aquaculture industry. The commercial importance of Atlantic cod has led to the development of several genomic resources to better understand cod biology and how genetic variation may affect economically important traits [2–6]. Recently, a sequencing project using next generation sequencing and a whole genome shotgun approach was completed. The genome sequence assembled from these data represents a first Atlantic cod reference genome sequence publicly available and allows for miRNA discovery by use of deep sequencing methods [7].

MicroRNAs (miRNAs) are an abundant class of endogenous small RNA molecules that downregulate gene expression at the post-transcriptional level [8–10]. Most miRNAs are highly conserved among vertebrates with limited sequence variation across species. There are, on the other hand, also non-conserved miRNAs that are present in one species only (species specific miRNAs) [11, 12]. Evolutionary conserved miRNA genes are not only similar in their sequences among species. Some miRNA genes are closely located in the genome (clustered), and such miRNA gene clusters are also often conserved across species. Furthermore, if a miRNA gene cluster is located in an intron the clustered miRNA genes are often co-expressed with the host-gene [13, 14].

A prerequisite to study miRNAs and their functions in new species would be to discover and characterize the sequences of their biologically active mature miRNAs as well as the sequence and location of their corresponding miRNA genes. Recent advances in sequencing technology have led to increased sensitivity in sequencing analysis (deep sequencing) that allows even lowly abundant small RNAs to be detected [15]. Applying bioinformatics tools that utilize deep sequencing data together with current knowledge of the biosynthesis of miRNAs and a reference genome, both conserved miRNAs as well as novel miRNAs, may be discovered in a new species [15, 16]. Such miRNA discovery and characterization studies have been carried out in a large number of commercially important teleost species over the recent years (e.g. [17–20]). Recently the first miRNA discovery study was performed in Atlantic cod with a particular focus on characterization of miRNAs present in developmental tissues and their expression differences associated with temperature [21].

MiRNAs are major regulators of protein coding gene expression in vertebrates, and they act as key regulators of fundamental biological processes like growth, immune response, tissue development and maintenance of tissue specific functions [22–26]. They may also act as

causative mediators of disease. This has led to studies of how miRNAs affect health in humans as well as in other common model organisms [27, 28]. The motivation for miRNA research in commercially important teleost species would, however, not only be restricted to studies on how they affect health and disease, but would also include studies on how miRNAs may affect economically interesting traits like growth or food conversion. To fully understand the interaction between genes in networks that affect such interesting traits it is necessary to identify the miRNAs that participate as key regulators [29–32]. The miRNAs that show organ specific expression patterns are often key regulators of biological functions carried out by the organ in which they are uniquely expressed [13, 33]. Thus, to disclose which miRNAs that are particularly important in development and function of various tissues a first step could be to identify the miRNAs that are highly expressed in one or few particular organs. There have been relatively few studies aiming at identifying such miRNAs in non-model teleosts [34].

The aims of this study is to discover and characterize miRNAs in Atlantic cod as well as reveal those miRNAs that are expressed at significantly higher levels in particular organs. Most known mature miRNAs identified so far were discovered in tissues from early developmental stages [21]. Our discovery analysis was, therefore, carried out independently in eight organ samples from a 1 kg individual to enable identification of a large number of miRNAs with a high confidence. To facilitate analysis of differentially expressed miRNAs in Atlantic cod by RT-qPCR we carried out a validation study to identify miRNAs suitable as endogenous controls in this species. Finally, the expression profiles obtained from the deep sequencing datasets and the validated RT-qPCR method were utilized to identify a subset of miRNAs with higher expression in certain organs.

Material and Methods

Small RNA isolation from eight individuals

One wild-caught coastal Atlantic cod individual was used for deep sequencing, whereas samples from additional seven wild-caught coastal Atlantic cod individuals (a total of eight individuals) were used for qPCR. The individuals (0.6–1 kg) were from the Oslofjord, south-east Norway. No specific permissions were required for sampling fish at these locations and the field studies did not involve endangered or protected species. The fish were provided by local fishermen. The one individual that was used for small RNA deep sequencing (a female) was caught, kept alive in a saltwater tank, sacrificed in the harbour and immediately given to us. The remaining seven individuals (three males) were caught, sacrificed immediately upon capture and put on ice. Thus, the one individual that was used for deep sequencing (see next section of Material and Methods) was kept on ice less than 30 min before dissected, the other seven individuals were kept on ice less than 3 hours before dissected. Dissection of fish and sampling of materials was performed in agreement with the provisions enforced by the Norwegian Animal Research Authority. As fish were sacrificed prior to our study and delivered by local fishermen no permission was required from the Institutional Animal Care and Use Committee (IACUC). Immediately after dissection samples were collected from each of the individuals to RNA later (Ambion, Life technologies, Carlsbad, CA, USA) for isolation of total RNA. The samples were from nine different organs: head-kidney, heart, brain, gills, skeletal muscle, liver, spleen, skin and intestine. Thus, each sample consisted of tissues from a particular organ. Samples of all nine organs were taken from all eight individuals dissected. A sample of about 0.05 g was collected from each organ and stored in RNA later. The sample from the intestine was flushed with distilled water to remove exogenous material before stored in RNA later. The skin samples were rectangular pieces sampled close to the lateral line. Any muscle tissues were removed from the skin samples before they were stored in RNA later.

Total RNA was isolated from each of the samples by use of the mirVana miRNA isolation kit (Ambion, Life technologies, Carlsbad, CA, USA) following the manufacturer's protocol. The RNA concentration and purity were determined using the Bioanalyzer and/or Nanodrop following the manufacturer's protocol. The 72 samples (eight individuals, nine tissues) showed concentrations of total RNA ranging from 35–900 ng/μl (total volume 100 μl). All 72 samples were used for RT-qPCR analysis of miRNA expression.

Pre-processing, deep sequencing and miRDeep analysis of deep sequencing data

Samples from one of the eight individuals (see above) were selected for deep sequencing and tissue from the following organs were included: head-kidney, brain, gills, skeletal muscle, liver, spleen, skin and intestine. The concentration of total RNA in these particular samples ranged from 75–865 ng/μl (total volume 100 μl) and with RIN values ≥ 7, see S1 Table. The Illumina TruSeq Small RNA Library Preparation Kit (Illumina, San Diego, USA) was used in the preparation of the libraries as described by the manufacturer with 1 ug total RNA input. After adapter ligation and cDNA synthesis the products were purified on a gel and the fraction between 145–160 bp was used for sequencing. The library construction was performed at the Norwegian Genomics Consortium's genomics core facility. The small RNA libraries constructed from the eight tissues were successfully subjected to next generation sequencing using Illumina Genome Analyzer IIx sequencing platform as described in Andreassen et al.[19]. The sequence reads (fastq-files) from each of the eight samples showed reads with Q-values ≥32. Adapter only sequences (5'TGGAATTCTCGGGTGCCAA GGAACTCCAGTCAC 3') were removed, reads were trimmed (removal of nucleotides from adapter sequences) and reads were size filtered to remove reads less than 18 nucleotides by use of the fastx-clipper tool (http://hannonlab.cshl.edu/fastx_toolkit/). The miRNA discovery analysis was independently performed in each of the eight samples to allow for detection of miRNAs highly expressed in particular tissues. The genome assembly from the *Gadus morhua* whole genome shotgun sequencing project (http://hgdownload.soe.ucsc.edu/goldenPath/gadMor1/bigZips/), Gen-Bank accession number: CAEA00000000.1, was used as reference genome. The high quality, adapter processed reads were used as the experimental data, and the discovery analysis was performed using the miRDeep2 software package (mapper and miRDeep2 analysis modules) [15, 35]. Default commands were used in the miRDeep2 analysis except that conservation scoring was omitted and the parameter g was set to -1 to allow all precursors to be analyzed during automatic excision gearing. We used the miRDeep2 score that yielded a signal-to-noise ratio of 30:1 as a cutoff threshold. All precursors with scores above this threshold and with reads that aligned perfectly, and in a discrete manner, to both 5' and 3' end of a precursor were regarded as putative miRNAs. These putative precursor sequences were further analyzed by BLAST searches against all known precursor sequences deposited in miRBase, release 21 (http://www.mirbase.org/search.shtml). We defined a significant hit as a match with an E-value ≤ 1E-06 to a stem-loop in miRBase. Any putative miRNA precursor sequence that provided a significant hit in the BLAST analyses was accepted as a true *Gadus morhua* miRNA precursor sequence, and each of these were annotated as the evolutionary conserved *Gadus morhua* ortholog of the miRNA gene in miRBase that retrieved the best hit. There are, at present, no miRNAs from *Gadus morhua* in the current version of miRBase, but Atlantic cod miRNAs have recently been characterized in materials from developmental stages [21] and the results submitted to miRBase. To facilitate comparison between our study and Bizuayehu et al [21] and to ensure that annotation are in agreement with the nomenclature guidelines [11, 36], our results from discovery and characterization of miRNAs were submitted to miRBase. The final annotation of all miRNAs and miRNA precursor sequences given in the results section was carried out by miRBase.

The precursors that were identified by miRDeep, but did not significantly match any miRNA precursor in miRBase were considered as putative novel miRNAs. All such precursors were used as queries in BLAST analysis that were performed against the nt/nr and refseqRNA databases in Genbank (http://blast.ncbi.nlm.nih.gov/Blast) and the small RNA family database in Rfam (http://rfam.xfam.org/search). Any putative precursor that showed a significant hit against these databases were considered to be other kinds of small RNA and excluded. Finally, all precursors were used as queries in BLAST analysis against the *Gadus morhua* genome sequence (http://www.ensembl.org/Gadus_morhua/Tools/Blast?db=core). Any putative precursor with a significant BLAST hit, defined as E-value ≤ 1E-06 against multiple loci (>5) in the genome reference sequence were considered to be part of interspersed repeats or tandem repeats and, consequently, excluded as novel miRNAs. The remaining putative novel miRNAs were validated in the following manner: they should be detected in at least two independent deep sequencing samples. More than five reads from the samples sequenced should match perfectly the expected mature products from both arms (5p and 3p), and the reads that aligned to the precursor should support that there was a consistent processing of the 5'end of the mature sequences. Passing all these criteria they were considered to be true novel *Gadus morhua* miRNAs.

The presence of clustered miRNA genes among the miRNA genes discovered in our study was investigated by comparing precursor locations within contigs. Any two miRNA precursors located within the same contig, less than 10 kb apart and with same direction of the transcription was considered part of a miRNA gene cluster. This definition (10 kb) is the same as the one used as default in miRBase (http://www.mirbase.org/search.shtml).

Sequencher software 5.3 (Gene Codes Corporation, Ann Arbor, USA) was used to align mature miRNA sequences (5p or 3p). By applying strict settings only identical mature sequences were allowed to align, thus, providing the total number of unique mature miRNAs in our materials.

cDNA synthesis and RT-qPCR

The miScript assays were used for cDNA synthesis and qPCR as described by the manufacturer (Qiagen, Hilden, Germany). A universal primer (reverse primer), provided with the miScript qPCR kit, was used in combination with our own custom designed forward primer in the qPCR assays amplifying each of the mature miRNAs. The sequence of the mature miRNAs characterized in this study was utilized to produce miRNA-specific forward primers. All primers were purchased from Sigma Aldridge, purified by desalt only and provided as liquid solution of 100 µM from the manufacturer. They were diluted to 10 µM for use in each of the qPCR assays. All forward primer sequences are given in S2 Table. The qPCR analysis was run on an Mx3000p (Stratagene). The qPCR reaction mixture consisted of 12.5 µL 2xQuantitec Syber Green Master Mix, 2.5 µL 10x miScript Universal Primer, 2.5 µL of 10 µM forward miRNA-specific primer, 5 µL Rnase free water, and 2.5 µL cDNA (template). The following program was used in qPCR: one thermal cycle at 95°C for 15 min followed by 40 cycles of 94°C for 15 sec, 55°C for 30 sec and 70°C for 30 sec. The Mx3000p software package was used for qPCR analysis. The cybergreen assay module includes a final melting point analysis that follows the 40 cycles of quantitative PCR. Plots from melting point analysis were manually inspected for all the miRNA assays tested to verify that forward primers were specific (data not shown).

Normalised read counts from deep sequencing datasets

The deep sequencing data from eight samples (S1 Table) was used to provide crude estimates of expression of individual miRNAs across different tissues. First, the mature miRNA sequences discovered and annotated in this study (from miRDeep) were used to generate an

index file for NovoAlign software (http://www.novocraft.com). Novoalign was then used to count the number of miRNA reads with a perfect match to any mature miRNA in each of the trimmed and size filtered deep sequencing samples. To obtain normalized read counts the total counts of a particular miRNA (reads in sizes >18 nt) was divided by the total number of all miRNAs in that sample and multiplied with 10^4. These normalized read counts of individual miRNAs could be compared across samples and were, thus, utilized to select candidate miRNAs that were likely to have higher expression in certain organs or candidate miRNAs that showed very similar expression level across all tissues (candidate reference genes). The read counts and normalized read counts from Novoalign analysis is given in S3 Table.

Stability testing of candidate reference genes

The average normalised read counts of each of the mature miRNAs across the eight samples and their standard deviation was estimated and utilized for comparison of individual miRNA stability across tissues. The relative standard deviation (RSD, coefficient of variation) was estimated by dividing the standard deviation of mean normalised read count from each mature miRNA by their mean. The RSD of each mature miRNA was used as a measurement of their relative stability across the tissues. Avoiding those miRNAs that were part of miRNA gene families (close sequence similarity) or belonging to same gene clusters [37] a set of candidate genes with low RSDs were selected for further stability testing by RT-qPCR. The commonly used endogenous control snRNA U6 was also tested in seven random samples (two brain, two intestine, head kidney, liver, heart and spleen samples) using the commercially available miScript assay MS00033740 (Qiagen, Hilden, Germany). The efficiency of the U6 assay was 57% and the Ct values showed rather low concentrations of U6 (n = 7, average Ct = 28,3). This indicated that U6 was not suitable as reference gene, and, consequently, it was not included as a candidate reference gene in NormFinder analysis (see below). The stability (standard deviation) from measurements of the seven samples using U6 was, however, compared to the stability of the best performing miRNAs.

The candidate reference miRNAs were analyzed by RT-qPCR in all samples from the nine different tissues (n = 72). All raw Ct values were transformed to linear scale and grouped by tissues before used as input in NormFinder. The NormFinder algorithm (http://moma.dk/normfinder-software) [38] calculates a "stability value" that is inversely correlated with the stability of gene expression (so a higher stability value indicates lower stability). This data normalization tool was used to rank the expressions of candidate genes by their stability and to identify the most suitable reference genes for normalization of RT-qPCR measurements of miRNA expression in Atlantic cod.

Identification of differentially expressed miRNAs by RT-qPCR

The normalized read counts of individual miRNAs from the deep sequencing samples was utilized to select miRNAs that were likely to have higher expression in one or more of the tissue samples. A total of 47 miRNAs were selected for RT-qPCR analysis from this manual comparison. In addition, a few evolutionary conserved miRNAs known to exhibit tissue enriched expression in other vertebrates (*gmo*-miR-736, *gmo*-miR-499 and *gmo*-miR-26a) were also selected for qPCR analysis. Thus, a total of 50 miRNAs were selected for RT-qPCR. All eight samples from the tissue(s) where a given miRNA was suspected to be expressed at higher levels (target tissue) were analysed together with four samples from each of the other tissues (reference tissues) in the same qPCR run. The Ct-value of a putative differentially expressed miRNA was normalized by subtracting the geometric mean of Ct-values generated from reference genes. The normalized measurements (ΔCts) were compared for all the tissues tested, an

approach similar to Liang et al. [13]. The four normalised Ct-values (ΔCt) from each different tissue in the reference group were inspected to make sure they showed low and non-overlapping expression levels compared to target tissue(s). This inspection revealed that in some cases the selected miRNAs seemed to have higher expression in more tissues than the initially selected target tissue. If so, these additional tissues were also defined as target tissues. The number of samples analysed in the reference group, thus, was from 20 to 32 samples depending on how many tissues that were defined as target and reference tissues for a given miRNA. The mean normalized Ct and the spread (standard deviation) was calculated for each of the groups, and Welchs t-test was used to test whether there was a significant difference in expression in target tissue vs reference tissue of a given target miRNA. A total of 76 tests were performed and the significant levels were adjusted according to number of tests to a "Bonferroni corrected" p-value of 6,6E-04. The relative increase in expression in target tissue of each miRNA was calculated using the principle of the comparative Ct method (ΔΔCt- method) [39]. Applying this method the average relative increase in target tissue could be calculated as the difference between mean normalized Ct from target tissue and mean normalized Ct from reference tissue. A correlation analysis between relative increases estimated from deep sequencing data and from RT-qPCR was performed. Those miRNAs that could be quantified in all eight deep sequencing samples (normalized read numbers larger than zero, a total of 34) were included in this comparison.

Results

Discovery of miRNAs and miRNA genes by deep sequencing and miRDeep analysis

The number of quality filtered (quality score above 32) and adapter trimmed reads from each sample ranged from 1.8 to 6.5 million. The concentration of total RNA and RIN values of the eight samples submitted for deep sequencing is given in S1 Table. This file also gives an overview of results from deep sequencing with total number of reads, number of adapter only reads, the total number of quality filtered, adapter trimmed and size filtered reads and total number of unique reads for each of the samples. All data from deep sequencing of the eight samples have been submitted to the NCBI Sequence Read Archive (SRA) database, and SRA accession numbers for each sample is given in S1 Table.

The quality filtered and adapter trimmed reads were used as input in miRDeep [39], and each sample was analyzed individually to facilitate discovery of miRNAs expressed in particular tissues only (see methods). Initially miRDeep identified a total of 544 putative mature miRNAs (5p and 3p) and their corresponding precursor sequences in the Atlantic cod genome assembly with score values above our threshold settings (see methods). All the precursor sequences were subsequently analyzed by BLAST homology searches against miRBase. Any precursor sequence that showed a significant match (E-value ≤ 1E-06) to a miRNA precursor in miRBase was assumed to be the *Gadus morhua* ortholog of this miRNA precursor and subsequently annotated as the *Gadus morhua* ortholog of this miRNA family. The homology analysis revealed that most of the putative miRNAs discovered by miRDeep, a total of 219 Atlantic cod miRNA genes with 438 corresponding mature miRNAs (5p and 3p), could be classified as orthologs to evolutionary conserved miRNAs. The alignment of perfectly matching reads to the precursor of each miRNA gene (by miRDeep) also revealed whether there was 5p or 3p arm domination among the mature miRNAs processed from a given precursor. An overview of all precursor sequences along with their corresponding 5p and 3p mature sequences from all evolutionary conserved miRNA genes is given in S4 Table. Information about arm dominance of mature sequences and genome location of each miRNA gene is also given in this file.

A total of 50 precursors identified by miRDeep did not show significant matches in the homology analyses against miRBase. These were analyzed by additional BLAST homology searches against smallRNA databases as well as the Atlantic cod genome sequence. These BLAST searches revealed that there were three precursors with significant matches to other small RNAs in the Rfam database. These were likely to be other small RNAs rather than miRNA precursors and were removed. Several of the putative precursors also showed significant matches to multiple loci (>50) in the cod genome. These were likely to be interspersed repeats rather than miRNAs and, consequently, were removed from our group of putative novel miRNAs. The remaining precursor and corresponding mature sequences all showed characteristics expected from true miRNAs. These 26 miRNA precursor sequences with their 5p and 3p mature sequences are likely to represent novel miRNAs. An overview of precursor sequences from all such novel miRNA genes along with their corresponding 5p and 3p mature sequences, the observed arm dominance of mature sequences and genome location of each miRNA gene is given in S5 Table. The miRDeep analysis and subsequent validations, thus, discovered 245 miRNA genes in the Atlantic cod genome sequence with 490 5p and 3p mature miRNAs. After aligning miRNA gene family members that shared identical mature miRNAs (Sequencher software), there were 401 unique 5p or 3p mature miRNAs discovered in our materials.

A miRNA gene cluster is defined by miRBase as two or more miRNA genes located less than 10 kb apart and transcribed in the same direction. Comparing the location of the discovered miRNA genes within the current version of the Atlantic cod genome sequence showed that there were 32 such miRNA gene clusters, and a total of 77 of the 245 miRNA genes identified in our study were located in a gene cluster of from two to six miRNA genes. An overview of all clustered miRNA genes is given in Fig 1 and in S6 Table. All the miRNA genes in the gene clusters were evolutionarily conserved.

Differentially expressed miRNAs validated by RT-qPCR

Five candidate miRNAs were tested as endogenous controls (see methods for selection criteria). One of the selected assays for a candidate reference gene, *gmo*-miR-455-5p, showed suboptimal performance with primer-dimer products appearing at Ct >36 and was not further tested. The remaining four candidate reference genes (*gmo*-miR-17-1—5p, *gmo*-miR-25-3p, *gmo*-miR-210-5p and *gmo*-miR-222-3p) were successfully analyzed by RT-qPCR in all samples (n = 72). The results were used as input in NormFinder [37, 38] to estimate most stable single gene and most stable two-gene combination. The stability values for each single gene and most stable two-gene combination as calculated by NormFinder is given in Fig 2 while Fig 3 shows the variation in stability across all nine tissues for each candidate reference gene. The best performing single reference gene was *gmo*-miR-17-1—5p while the best two-gene combination, revealing a slightly better stability score, was *gmo*-miR-25-3p and *gmo*-miR-210-5p. Initial analysis of U6 showed an efficiency of 57% indicating that it was not suitable as endogenous control. RT-qPCR measurements of U6, *gmo*-miR-25-3p and *gmo*-miR-210-5p across seven samples showed a standard deviation (S.D.) of 2.2 for U6 while it was 1.4 for *gmo*-miR-25-3p and 1.1 for *gmo*-miR-210-5p. This indicated that, not only poor efficiency made U6 unsuitable as a reference gene, it did also show less stability (larger S.D.) in direct comparison with the best performing miRNAs. The two miRNAs, *gmo*-miR-25-3p and *gmo*-miR-210-5p were, therefore, applied as reference genes in our RT-qPCR method for measuring miRNA expression across, and within, tissues in Atlantic cod.

A total of 50 putatively differentially expressed miRNAs (see methods for selection criteria) were analyzed with our dedicated RT-qPCR method. Specific forward primers (S2 Table) were

Nr	Clustered miRNA genes	Accession # to contig in Genbank with postion and strand (+/-) for the clustered genes
1	gmo-mir-27d, gmo-mir-23a-2	CAEA01040759:456..519:-, :805..865:-
2	gmo-mir-144, gmo-mir-451	CAEA01183658:343..401:+, 488..531:+
3	gmo-mir-222-1, gmo-mir-221-1	CAEA01202201:317..379:+, 767..830:+
4	gmo-mir-29b-2, gmo-mir-29a-2	CAEA01475371:1119..1181:+, 1317..1377:+
5	gmo-let-7a-4, gmo-let-7e-2	CAEA01553329:496..574:-, 703..775:-
6	gmo-mir-30e-1, gmo-mir-30e-2	HE566387:147043..147105:-, 148854..148917:-
7	gmo-mir-449a, gmo-mir-449b	HE566499:61062..61121:-, :61372..61428:-
8	gmo-mir-133a-1, gmo-mir-1-2	HE566612:109175..109235:-, 112878..112939:-
9	gmo-mir-19b-2, gmo-mir-20a-1, gmo-mir-19a-1, gmo-mir-18a-2, gmo-mir-17-1	HE566687:992644..992704:-, :992762..992821:-, 992926..992983:-, 993064..993126:-, 993185..993245:-
10	gmo-mir-125a-2, gmo-let-7d-2	HE566792:123225..123283:-, 128930..129001:-
11	gmo-let-7a-2, gmo-mir-125b-1	HE566819:120032..120099:+, 131490..131551:+
12	gmo-mir-148, gmo-mir-10a-2	HE566826:222662..222721:-, 350427..350489:-
13	gmo-mir-99, gmo-let-7c, gmo-mir-125b-2	HE566869:184077..184135:+, 185169..185236:+, 194779..194842:+
14	gmo-mir-183, gmo-mir-96, gmo-mir-182	HE567039:12669..12729:+, 12976..13040:+, 13679..13743:+
15	gmo-let-7d-1, gmo-mir-125a-1	HE567108:269036..269106:+, 272623..272680:+
16	gmo-let-7h-1, gmo-let-7g-2	HE567345:1052383..1052465:+, 1052708..1052780:-,
17	gmo-mir-212-1, gmo-mir-132-2	HE568217:132512..132577:+, 132868..132930:+
18	gmo-mir-15b, gmo-mir-16a	HE568600:341901..341962:+, 342080..342145:+
19	gmo-mir-133b, gmo-mir-206	HE568667:96738..96797:-, 99341..99401:-
20	gmo-mir-7e-1, gmo-let-7a-5	HE568845:697528..697601:+, 697772..697849:+
21	gmo-mir-125b-3, gmo-let-7a-1, gmo-mir-100a	HE569363:2583112..2583174:-, 2590773..2590841:-, 2591157..2591214:-
22	gmo-mir-29d, gmo-mir-29a-1	HE569614:926750..926809:+, 927151..927210:+
23	gmo-mir-17-2, gmo-mir-18a-1, gmo-mir-19a-2, gmo-mir-20a-2, gmo-mir-19b-1, gmo-mir-92-1	HE569758:354880..354939:+, 355007..355070:+, 355142..355198:+, 355323..355382:+, 355452..355513:+, 355560..355618:+
24	gmo-mir-30b, gmo-mir-30e-1	HE570053:81943..82003:-, 82256..82321:-
25	gmo-mir-132-1, gmo-mir-212-2	HE570340:947712..947775:-, 947995..948061:-
26	gmo-mir-20a-3, gmo-mir-93, gmo-mir-19d, gmo-mir-25	HE570406:50465..50526:+, 50691..50752:+, 50960..51025:+, 51119..51179:+
27	gmo-mir-193a, gmo-mir-365	HE570758:401123..401181:+, 403785..403847:+
28	gmo-let-7h-2, gmo-let-7g-1	HE570778:22986..23070:-, 23472..23544:-
29	gmo-mir-16b, gmo-mir-15a	HE571102:46215..46281:-, 46493..46553:-
30	gmo-mir-731, gmo-mir-462	HE571228:149014..149078:-, 149805..149862:-
31	gmo-mir-200b, gmo-mir-200a, gmo-mir-429	HE571528:73084..73143:+, 73296..73357:+, 74829..74887:+
32	gmo-mir-199-1, gmo-mir-214-1	HE572020:79409..79469:+, 81842..81903:+

Fig 1. miRNA gene clusters in Atlantic cod.

doi:10.1371/journal.pone.0153324.g001

successfully applied in RT-qPCR assays for each mature miRNA except *gmo*-miR-7552. Forward primers failed to produce a single specific mature PCR-product in this assay, and *gmo*-miR-7552 was therefore not analyzed. Among the remaining miRNAs that were successfully analyzed by RT-qPCR, 45 out of 49 miRNAs showed a significantly higher expression level in one or more tissue. An overview of all miRNAs analyzed by qPCR, their target tissue(s), the relative increase in expression compared to reference tissues and p-values from t-tests is given in Fig 4. A correlation analysis between the relative increase of miRNAs from comparison of normalized reads and the relative increase of these miRNAs as measured by RT-qPCR showed a correlation coefficient of 0.63 S7 Table.

The magnitude of the increase in expression in target tissue(s) *versus* reference tissues were from rather small (1.5 fold) to very large differences with increases of several thousand-fold in some target tissues (e.g. *gmo*-miR-122 in liver). *Gmo*-miR-1 and *gmo*-miR-133a showed higher expression in four tissues (skeletal muscle, heart, skin and intestine). *Gmo*-miR-206 showed

miRNA gene	Stability value	Best combination of two genes	Stability value for best two-combination
gmo-miR-17-5p[1]	0.331	gmo-miR-25 and gmo-miR-210	0.318
gmo-miR-25-3p	0.475		
gmo-miR-210-5p	0.509		
gmo-miR-222-3p	0.565		

[1] Best single miRNA gene

Fig 2. Estimated stability values by NormFinder.
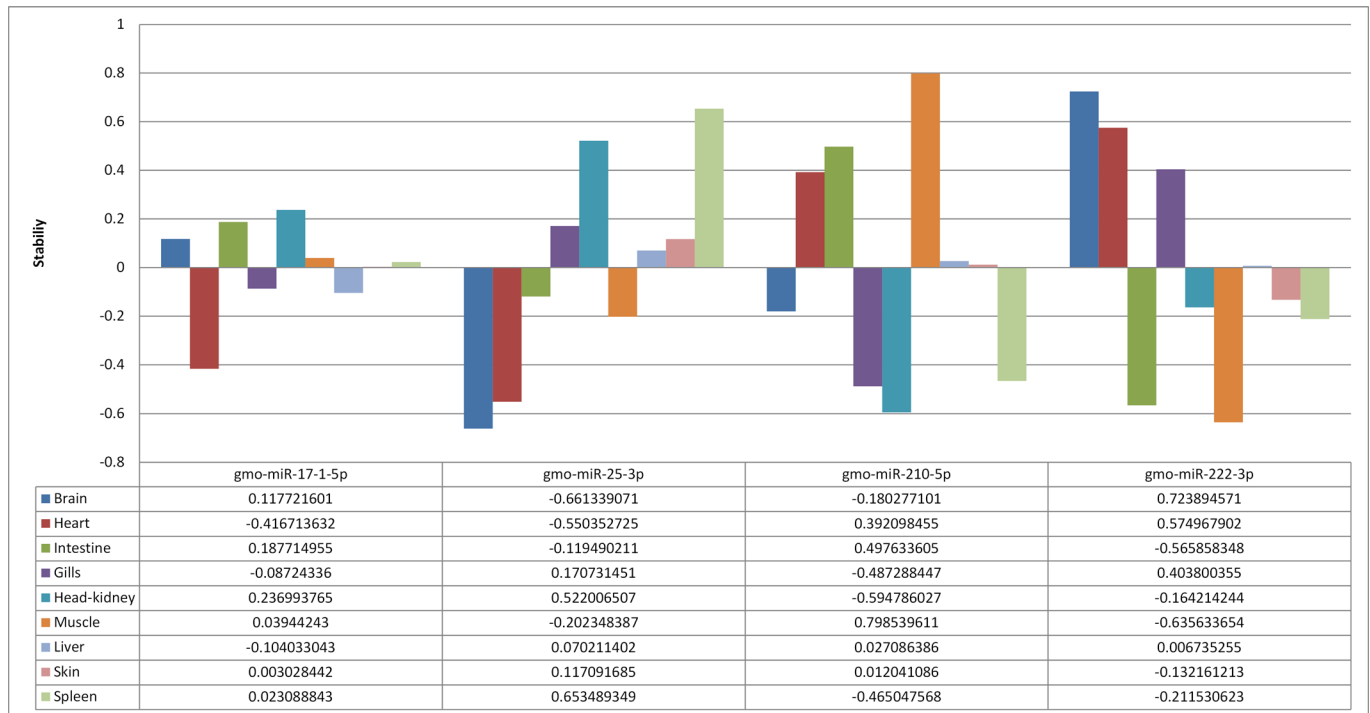
doi:10.1371/journal.pone.0153324.g002

	gmo-miR-17-1-5p	gmo-miR-25-3p	gmo-miR-210-5p	gmo-miR-222-3p
Brain	0.117721601	-0.661339071	-0.180277101	0.723894571
Heart	-0.416713632	-0.550352725	0.392098455	0.574967902
Intestine	0.187714955	-0.119490211	0.497633605	-0.565858348
Gills	-0.08724336	0.170731451	-0.487288447	0.403800355
Head-kidney	0.236993765	0.522006507	-0.594786027	-0.164214244
Muscle	0.03944243	-0.202348387	0.798539611	-0.635633654
Liver	-0.104033043	0.070211402	0.027086386	0.006735255
Skin	0.003028442	0.117091685	0.012041086	-0.132161213
Spleen	0.023088843	0.653489349	-0.465047568	-0.211530623

Fig 3. Variation across nine tissues for the four candidate reference genes. Fig3 shows the variation across tissues sampled from nine different organs for the four candidate reference genes.

doi:10.1371/journal.pone.0153324.g003

higher expression in tissues from skeletal muscle, heart and skin while *gmo*-miR-725 showed higher expression in skeletal muscle, skin and gills. Fifteen miRNAs showed higher expression in two out of the nine tissues tested, while the remaining 25 were highly expressed in tissue from one organ, only.

Those differentially expressed miRNAs that were from closely located miRNA genes (gene clusters) showed, in general, similar high expression in the same tissues. *Gmo*-miR-1 and *gmo*-133a were e.g. highly expressed in the same four tissues (skeletal muscle, heart, skin and intestine). *Gmo*-miR-144 and *gmo*-miR-451 were highly expressed in head-kidney and spleen. *Gmo*-miR-212 and *gmo*-miR-132 were preferentially expressed in brain while *gmo*-miR-200a and *gmo*-miR-429 were highly expressed in gills. The proportion of miRNAs that were preferentially expressed in particular tissues is shown in Fig 5. This comparison showed that brain was the one tissue with largest proportion (28%) of differentially expressed miRNAs. The number of individuals in our study was small. Although we did not observe any differences in miRNA expression between subgroups of individuals (data not shown), we cannot rule out that the general miRNA expression, or expression of certain miRNAs may be affected by sex. The one individual that was deep sequenced showed Ct-values that grouped together with the remaining seven individuals indicating that a different sampling procedure did not affect the RT-qPCR measurements (data not shown).

A comparison to findings in other species could reveal whether the difference in expression of the particular miRNAs from our study are evolutionary conserved among vertebrates. Our results were, thus, compared to studies in other teleost as well as higher vertebrates (see references [34, 40–57]) and the miRNAs listed in databases TsmiR (http://bioeng.swjtu.edu.cn/TSmiR/) and Genecards (http://www.genecards.org/). A summary of this comparison showing the miRNAs preferentially expressed in each of the nine tissues and those miRNAs with similar

Assay No.	miRNA	Target tissue[1]	target tissue m_ΔC_T[2]	Reference tissues m_ΔCT[3]	ΔΔC_T[4]	Increase in target-tissue[5]	p-value[6]
1.	1-3p	MU	-6.08	5.98	-12.1	4289	1.10E-14
		HA	-1.97	5.98	-7.96	249	1.30E-09
		SK	-2.99	5.98	-8.98	504	2.60E-11
		I	-3.07	5.98	-9.06	533	1.90E-11
2.	let-7i-5p	B	-3.19	2.22	-5.4	43	2.00E-06
		I	0.39	2.22	-1.8	3.5	5.40E-11
3.	9-5-3p	B	-5.336	6.28	-11.6	3144	1.20E-26
4.	10d-5p	SP	-0.69	0.32	-1.01	2	4.10E-04
		SK	-2.92	0.32	-3.24	9.4	1.10E-13
5.	26a-5p	MU	-2.71	-2.13	-0.6	1.5	2.00E-06
		B	-3.94	-2.13	-1.8	3.5	5.50E-10
6.	101b-3p	L	-1.6	1.18	-2.78	6.9	2.50E-08
7.	103-3p	B	-0.99	3.03	-4.02	16.3	1.10E-08
8.	122-5p	L	-6.23	5.66	-11.9	3799	4.40E-21
9.	206-3p	MU	-5.064	7.85	-12.9	7721	1.20E-18
		HA	-0.97	7.85	-8.8	452	1.30E-14
		SK	-2.91	7.85	-10.8	1740	3.70E-17
10.	128-3p	B	-5.53	0.46	-6	64	1.50E-09
11.	129-3p	B	-2.41	6.7	-9.32	637	1.60E-20
12.	132-3p	B	-0.72	8.87	-9.6	774	1.00E-20
13.	133a-3p	MU	-5.16	6.29	-11.5	2819	1.30E-14
		HA	-1.48	6.29	-7.8	218	8.30E-11
		SK	-1.9	6.29	-8.2	293	1.90E-11
		I	-1.16	6.29	-7.5	176	9.20E-11
14.	135b-5p	B	-0.39	7.84	-8.2	301	4.10E-27
15.	135d-5p	B	-1.08	9.25	-10.3	1292	3.10E-08
16.	140-3p	G	-0.03	3.93	-4	15.6	5.80E-20
		HA	1.83	3.93	-2.1	4.3	3.40E-11
17.	143-3p	I	-4.61	1.38	-6	63	3.40E-16
18.	144-5p	HK	2.37	7.65	-5.28	39	8.50E-17
		SP	3.53	7.65	-4.11	17	1.50E-08
19.	148-5p	L	-0.53	1.8	-2.34	5.1	4.40E-14
		SK	-0.55	1.8	-2.35	5.1	3.80E-08
20.	150-5p	G	-4.64	0.87	-5.5	46	6.10E-24
21.	153b-3p	B	0.87	10.93	-10.1	1072	2.00E-15
22.	182-5p	G	4.54	10.04	-5.5	45	3.30E-19
23.	192-5p	I	-1.82	6.35	-8.2	289	2.20E-14
		L	2.61	6.35	-3.7	13.2	5.80E-10
24.	194b-5p	I	0.77	9.83	-9.1	533	1.90E-16
		HK	4.38	9.83	-5.4	43	1.50E-06
25.	194a-5p	I	-2.12	5.61	-7.7	213	3.40E-16
26.	200a-3p	G	-2.35	4.02	-6.4	83	2.10E-12
		HK	2.18	4.02	-1.84	3.6	*1.50E-03*
27.	204-5p	SK	-2.98	5.67	-8.7	402	2.40E-20
		B	-1.66	5.67	-7.3	161	7.10E-19
28.	205-5p	G	-2.94	4.46	-7.4	169	9.30E-15
		SK	-0.77	4.46	-5.22	37	2.30E-05
29.	212-3p	B	0.321	9.608	-9.3	625	1.20E-16
30.	219-3p	B	-0.459	12.654	-13.1	8858	5.50E-13
31	223a-5p	HK	2.79	7.24	-4.46	22	2.00E-18
		SP	3.52	7.24	-3.73	13	1.60E-11
32.	338-3p	B	-2.64	2.84	-5.5	45	4.40E-09
		MU	-0.66	2.84	-3.5	11.2	6.40E-09
33.	375-3p	HK	3.31	4.73	-1.42	2.7	*2.40E-02*
		G	1.36	4.79	-3.43	10.8	2.00E-09
34.	429-3p	G	-1.21	5.19	-6.4	85	6.60E-14
		HK	3.03	5.2	-2.17	4.5	*2.80E-03*
35.	449b-5p	B	3.94	11.25	-7.3	159	2.60E-05
36.	451-5p	SP	-4.11	-0.35	-3.76	13.5	2.60E-10
		HK	-3.73	-0.35	-3.38	10.4	1.10E-15
		G	-2.24	-0.35	-1.89	3.7	*1.20E-03*
37.	455-5p	G	1.01	4.07	-3.06	8.4	6.00E-12
38.	458-3p	B	1.51	8.04	-6.53	92	5.30E-12
39.	499-5p	HA	-3.329	8.259	-11.6	3080	5.10E-19
		SK	-1.673	7.967	-9.6	798	6.20E-17
40.	723-3p	B	3.689	13.905	-10.2	1190	5.70E-16
41.	724-5p	B	0.799	10.203	-9.4	678	1.70E-06
42.	725-3p	G	-0.35	7.45	-7.8	223	1.30E-14
		SK	2.15	7.45	-5.3	39	1.60E-04
		MU	2.69	7.45	-4.76	27	2.60E-05
43.	736-3p	HA	-0.119	11.977	-12.1	4380	1.40E-18
44.	8160a-5p	B	3.69	5.93	-2.24	4.7	*3.10E-03*
		I	3.78	5.93	-2.15	4.4	*2.20E-02*
45.	11223-5p	B	5.38	7.84	-2.46	5.5	*7.50E-04*
46.	11272-3p	MU	6.508	11.054	-4.5	23	7.10E-19
		B	9.094	11.054	-2	3.9	2.80E-08
47.	11935-3p	MU	6.07	6.48	-0.41	1.3	*3.60E-02*
		I	4.69	6.48	-1.79	3.5	*1.80E-03*
48.	11183-5p	MU	6.83	6.98	-0.14	1.1	*2.60E-01*
		SP	5.72	6.98	-1.26	2.4	*1.30E-03*
		I	5.23	6.98	-1.74	3.3	*6.40E-03*
49.	11260-3p	B	9.105	13.69	-4.6	24	3.00E-04

[1] Target tissues are those tissues with expected enriched expression of miRNAs tested. B: brain, G: gills, HA: heart, HK: head-kidney, I: intestine, L: liver, MU: skeletal muscle, SK: skin, SP: spleen. [2] Target tissue m_ΔCT is the mean of normalised Ct-values from measurement of target tissue samples. [3] Reference tissue m_ΔCT is the mean of normalised Ct-values from measurement of reference tissue samples. [4] ΔΔCT is the difference between target tissue m_ΔCT and reference tissue m_ΔCT. [5] Increase in target-tissue is the relative fold change in expression in target tissue for this miRNA (linearised values from ΔΔCT). [6] p-values from testing for significant difference in distribution of normalized Ct-values in the target tissue group *versus* reference tissue group by Welchs t-test.

Fig 4. Summary of RT-qPCR results and tests to reveal differentially expressed miRNAs.
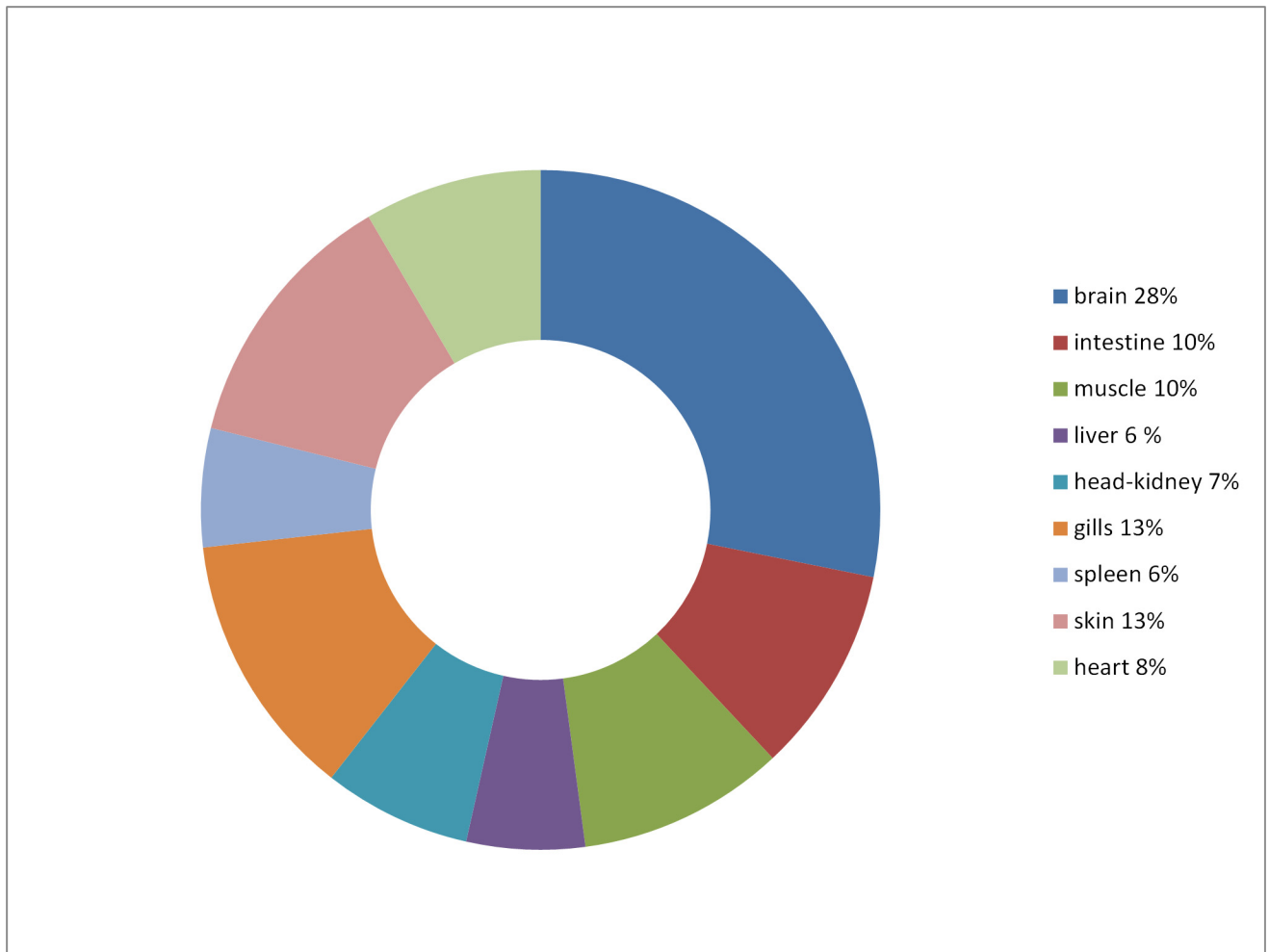
Fig 5. Distribution of differentially expressed miRNAs. The figure shows the distribution of differentially expressed miRNAs in nine different organ tissues from Atlantic cod.

finding in other species is given in Fig 6. This comparison revealed that most of the evolutionarily conserved miRNAs have been reported with similar expression patterns in other vertebrates.

Sixteen miRNAs (*Gmo*-miR-101b-3p, 140-3p, 148-5p, 150-5p, 182-5p, 205-5p, 338-3p, 375-3p, 455-5p, 458-3p, 723-3p, 724-5p, 725-3p, 11272-3p, 11223-5p, 11260-3p) have, to our knowledge, not been reported as with similar expression patterns in other vertebrates (all miRNAs shown in bold in Fig 6). Three of these (miR-723, 724 and 725) are teleost-specific, while three are novel miRNAs discovered in Atlantic cod (*gmo*-miR-11272-3p, 11223-5p and 11260-3p). Although their target transcripts are unknown, it is likely that they are regulators of genes/gene networks associated with the biological function of the tissue where they are preferentially expressed.

Discussion

A recent study by Bizuayehu et al. [21] is, to our knowledge, the only miRNA discovery study in Atlantic cod. A total of 281 mature miRNA sequences (245 unique sequences) discovered in materials mainly from early developmental tissues were characterized in this study. We have

Tissue	miRNAs[1]	Other teleosts/mammals[2]	References[3]
Brain	let-7i-5p, 9-5-3p, 26a-5p, 103-3p, 128-3p, 129-3p, 132-3p, 135d-5p, 135b-5p, 153b-3p, 204-5p, 212-3p, 219-3p, 338-3p, 449b-5p, **458-3, 723-3p, 724-5p, 11223-5p, 11260-3p**	miR-9, 128, 129, 212, 132,135 (D, O), 204, 338 (O),26, 107, 153, 219 (D), let-7i (D, AA, O), 449 (H)	[34, 41-42] TSmiR and Genecards
Skeletal Muscle	1-3p, 26a-5p, 206-3p, 133a-3p, **338-3p, 725-3p, 11272-3p**	miR-1, 133 (D, NT, CC, BSB), 26a (CC), 126a(D)	[34, 50] TSmiR
Heart	1-3p, 206-3p, 133a-3p, **140-3p**, 499-5p, 736-3p	miR-1, 133, 499 (D), 126a (AA), 736 (D)	[34, 44, 51] TSmiR
Intestine	1-3p, let-7i-5p, 133a-3p, 143-3p, 192-5p, 194a-5p, 194b-5p	miR-192, miR-194a, 194b (D, M), miR-1, 133 (Genecards), let-7i, miR-143(H)	[43, 45, 51-52] Genecards
Skin	1-3p, 10d-5p, 206-3p, 133a-3p, **148-5p**, 204-5p, 205-5p, **499a-5p, 725-3p**	miR-204, miR_205, 126 (H), miR-1, 133 (Genecards), miR-10a (H)	[46-47, 51, 53-55] Genecards
Gills	140-3p, **150-5p, 182-5p**, 200a-3p, **205a-5p, 375-3p**, 429-3p, **455-5p, 725-3p**	miR-140 (D), miR-200, 429 (NT)	[34, 43, 56]
Liver	**101b-3p**, 122-5p, **148-5p**, 192-5p	miR-122 (O), miR-192 (H, M)	[48-49] TSmiR
Head-kidney	144-5p, 194b-5p, 223-5p, 451-5p	miR-429 (NT), 194, 223, 144,451 (H)	[34, 45, 57, 59]
Spleen	10d-5p, 144-5p, 223a-5p, 451-5p	miR-10a (C), 223, 144, 451 (H)	[45, 57-59] TSmiR

[1]The miRNAs in bold have to our knowledge not been reported as with similar differential expression in other species.

[2]D: Zebrafish, NT: Nile tilapia, CC: Common carp BSB: Blunt snout bream, AA: Asian seabass, O: Rainbow trout, H: Human, M: Mouse, C: Chicken

[3]TSmiR is the tissue specific miRNA database, Genecards is the database Genercards.org

Fig 6. Differentially expressed miRNAs in Atlantic cod versus other teleosts or higher vertebrates.

doi:10.1371/journal.pone.0153324.g006

characterized the sequences of 490 mature miRNAs (401 unique sequences). This represents a considerable increase in number of known mature miRNAs in Atlantic cod. A comparison to other fish species showed that all of the miRNA genes that were clustered in Atlantic cod were also clustered in a similar manner in other teleosts [14, 19]. The discovery that these evolutionarily conserved Atlantic cod miRNA genes are clustered as expected from other teleosts provides additional support that they are true miRNA genes in *Gadus morhua*. Comparing findings from both studies show that there is, presently, a total of 450 unique mature miRNAs (5p and 3p) characterized and annotated in Atlantic cod. Detailed and reliable knowledge of the mature miRNA sequences in a species is a prerequisite for studying miRNA expression by deep sequencing methods. The discovery and characterization of the miRNA repertoire in this study will therefore greatly benefit further functional studies of the miRNA transcriptome in Atlantic cod.

Quantitative studies of the miRNA transcriptome may be carried out by deep sequencing followed by confirmation of any miRNA expression differences by RT-qPCR. Measurements of miRNA expression by RT-qPCR require proper reference genes. The ideal reference genes are the ones with similar biological properties as the target genes except they are stably expressed in the species (or across materials) studied [37, 40]. The validation of reference genes in our study indicated that three miRNAs, either *gmo*-17-1—5p alone, or the combination of *gmo*-miR-25-3p and *gmo*-miR-210-5p, were particularly suitable as reference genes in Atlantic cod. In agreement with recent finding in similar studies [58] U6 performed less well compared to the miRNAs as an endogenous control. MiR-25-3p seems to perform well as a miRNA reference gene in several teleosts [37, 40].

Most of the miRNAs revealed as differentially expressed have been reported as to have similar expression patterns in other teleosts or higher vertebrates (Fig 6) [34, 40–57]. It is likely that these miRNAs, with both conserved mature miRNA sequences and enriched expression in same tissues across species, regulate the same gene pathways and, thus, have same functions in Atlantic cod as those revealed in other vertebrates. Among the miRNAs that may be of particular interest to the aquaculture industry are those regulating growth (miR-1 and miR-133) and

lipid metabolism (miR-122) [34, 47–49]. Others of the differentially expressed miRNAs are known from other species to be important to health e.g. immune response to microbial infection in epithelial cells (let-7i) and scar damage repair (miR-10d) [44, 45, 51].

The differentially expressed miRNAs revealed in this study may be utilized to discover orthologous host genes in Atlantic cod. A host gene is a protein coding gene with one or more miRNA gene located within their introns. These host genes are usually expressed in a similar manner as the miRNA gene(s) that are located within their introns [13, 14]. Some miRNAs discovered in our study are located in well described host genes in other vertebrates. MiR-338 with preferential expression in brain in Atlantic cod is e.g. located in the gene brain apoptosis-associated tyrosine kinase (AATK) in other vertebrates [59]. Another miRNA showing significantly higher expression in skin (miRNA-204) is known from other species to be located in the host gene TRPM3 [13]. As the location of the Atlantic cod miRNA genes in our study are known, this may facilitate characterization and annotation of such orthologous host genes in the Atlantic cod genome.

Conclusions

The number of known mature miRNAs was considerably increased by our discovery and characterization of miRNAs and miRNA genes in Atlantic cod. This will benefit further functional studies of miRNA expression using deep sequencing methods. The validation study identified suitable endogenous controls and showed that stable miRNAs are suitable reference genes for RT-qPCR analysis of miRNA expression. Applying the RT-qPCR method we identified several miRNAs likely to have important regulatory functions in particular organs.

Ethical considerations

Dissection of fish and sampling of materials was performed in agreement with the provisions enforced by the Norwegian Animal Research Authority.

Supporting Information

S1 Table. The table shows sequence Read Archive (SRA) database accession numbers to all samples sequenced using the Illumina small RNA sequencing platform.
(XLSX)

S2 Table. The table shows all forward primer sequences used in the miRNA-qPCR assays.
(XLSX)

S3 Table. The table shows read counts of mature miRNAs from analysis with Novoalign.
(XLSX)

S4 Table. The table shows all evolutionary conserved miRNAs identified and characterized in Atlantic cod.
(XLSX)

S5 Table. The table shows all novel miRNAs identified and characterized in Atlantic cod.
(XLSX)

S6 Table. The table shows all genes located in gene clusters.
(XLSX)

S7 Table. Correlation between relative increase of differentially expressed miRNAs in RT-qPCR measurements vs estimates of the relative increase in deep sequencing samples.
(XLSX)

Acknowledgments

Author Contributions

Conceived and designed the experiments: RA. Performed the experiments: RA FR MS MC MT. Analyzed the data: RA MMW FR. Contributed reagents/materials/analysis tools: RA. Wrote the paper: RA.

References

1. Food and Agriculture Organization of the United Nations. Globale capture production, FishStat. Available: http://www.fao.org/fishery/species/2218/en

2. Booman M, Borza T, Feng CY, Hori TS, Higgins B, Culf A, et al. Development and experimental validation of a 20K Atlantic cod (Gadus morhua) oligonucleotide microarray based on a collection of over 150,000 ESTs. MarBiotechnol(NY) 2011, 13(4):733–750.

3. Bowman S, Hubert S, Higgins B, Stone C, Kimball J, Borza T, et al. An integrated approach to gene discovery and marker development in Atlantic cod (Gadus morhua). MarBiotechnol(NY) 2011, 13(2):242–255.

4. Moen T, Hayes B, Nilsen F, Delghandi M, Fjalestad KT, Fevolden SE, et al. Identification and characterisation of novel SNP markers in Atlantic cod: evidence for directional selection. BMCGenet 2008, 9:18.

5. Higgins B, Hubert S, Simpson G, Stone C, Bowman S: Characterization of 155 EST-derived microsatellites from Atlantic cod (Gadus morhua) and validation for linkage mapping. MolEcolResour 2009, 9 (3):733–737.

6. Hubert S, Higgins B, Borza T, Bowman S: Development of a SNP resource and a genetic linkage map for Atlantic cod (Gadus morhua). BMCGenomics 2010, 11:191.

7. Star B, Nederbragt AJ, Jentoft S, Grimholt U, Malmstrom M, Gregers TF, et al. The genome sequence of Atlantic cod reveals a unique immune system. Nature 2011, 477(7363):207–210. doi: 10.1038/nature10342 PMID: 21832995

8. Bartel DP: MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 2004, 116(2):281–297. PMID: 14744438

9. Chekulaeva M, Filipowicz W: Mechanisms of miRNA-mediated post-transcriptional regulation in animal cells. CurrOpinCell Biol 2009, 21(3):452–460. doi: 10.1016/j.ceb.2009.04.009 PMID: 19450959

10. Winter J, Jung S, Keller S, Gregory RI, Diederichs S: Many roads to maturity: microRNA biogenesis pathways and their regulation. NatCell Biol 2009, 11(3):228–234. doi: 10.1038/ncb0309-228 PMID: 19255566

11. Griffiths-Jones S, Grocock RJ, van DS, Bateman A, Enright AJ: miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res 2006, 34(Database issue):D140–D144. PMID: 16381832

12. Pasquinelli AE, Reinhart BJ, Slack F, Martindale MQ, Kuroda MI, Maller B, et al. Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. Nature 2000, 408 (6808):86–89. PMID: 11081512

13. Liang Y, Ridzon D, Wong L, Chen C: Characterization of microRNA expression profiles in normal human tissues. BMCG¿enomics 2007, 8:166. PMID: 17565689

14. Thatcher EJ, Bond J, Paydar I, Patton JG: Genomic organization of zebrafish microRNAs. BMCGenomics 2008, 9:253. doi: 10.1186/1471-2164-9-253 PMID: 18510755

15. Friedlander MR, Chen W, Adamidi C, Maaskola J, Einspanier R, Knespel S, Rajewsky N: Discovering microRNAs from deep sequencing data using miRDeep. NatBiotechnol 2008, 26(4):407–415. doi: 10.1038/nbt1394 PMID: 18392026

16. Li Y, Zhang Z, Liu F, Vongsangnak W, Jing Q, Shen B: Performance comparison and evaluation of software tools for microRNA deep-sequencing data analysis. Nucleic Acids Res 2012, 40(10):4298–4305. doi: 10.1093/nar/gks043 PMID: 22287634

17. Xu Z, Chen J, Li X, Ge J, Pan J, Xu X: Identification and characterization of microRNAs in channel catfish (Ictalurus punctatus) by using Solexa sequencing technology. PLoSOne 2013, 8(1):e54174.

18. Bizuayehu TT, Fernandes JM, Johansen SD, Babiak I: Characterization of novel precursor miRNAs using next generation sequencing and prediction of miRNA targets in Atlantic halibut. PLoSOne 2013, 8(4):e61378. doi: 10.1371/journal.pone.0061378 PMID: 23626677

19. Andreassen R, Worren MM, Hoyheim B: Discovery and characterization of miRNA genes in Atlantic salmon (Salmo salar) by use of a deep sequencing approach. BMCGenomics 2013, 14:482. doi: 10.1186/1471-2164-14-482 PMID: 23865519

20. Ma H, Hostuttler M, Wei H, Rexroad CE III, Yao J: Characterization of the rainbow trout egg microRNA transcriptome. PLoSOne 2012, 7(6):e39649. doi: 10.1371/journal.pone.0039649 PMID: 22761856

21. Bizuayehu TT, Johansen SD, Puvanendran V, Toften H, Babiak I: Temperature during early development has long-term effects on microRNA expression in Atlantic cod. BMCGenomics 2015, 16(1):305.

22. Bushati N, Cohen SM: microRNA functions. AnnuRevCell DevBiol 2007, 23:175–205. PMID: 17506695

23. Lagos-Quintana M, Rauhut R, Yalcin A, Meyer J, Lendeckel W, Tuschl T: Identification of tissue-specific microRNAs from mouse. CurrBiol 2002, 12(9):735–739. PMID: 12007417

24. Stefani G, Slack FJ: Small non-coding RNAs in animal development. NatRevMolCell Biol 2008, 9(3):219–230. doi: 10.1038/nrm2347 PMID: 18270516

25. Friedman RC, Farh KK, Burge CB, Bartel DP: Most mammalian mRNAs are conserved targets of microRNAs. Genome Res 2009, 19(1):92–105. doi: 10.1101/gr.082701.108 PMID: 18955434

26. Sonkoly E, Stahle M, Pivarcsi A: MicroRNAs and immunity: novel players in the regulation of normal immune function and inflammation. SeminCancer Biol 2008, 18(2):131–140. doi: 10.1016/j.semcancer.2008.01.005 PMID: 18291670

27. Jiang Q, Wang Y, Hao Y, Juan L, Teng M, Zhang X, et al. miR2Disease: a manually curated database for microRNA deregulation in human disease. Nucleic Acids Res 2009, 37(Database issue):D98–104. doi: 10.1093/nar/gkn714 PMID: 18927107

28. Esquela-Kerscher A, Slack FJ: Oncomirs—microRNAs with a role in cancer. NatRevCancer 2006, 6(4):259–269. PMID: 16557279

29. Malachowicz M, Kijewska A, Wenne R: Transcriptome analysis of gill tissue of Atlantic cod Gadus morhua L. from the Baltic Sea. MarGenomics 2015.

30. Yadetie F, Karlsen OA, Eide M, Hogstrand C, Goksoyr A: Liver transcriptome analysis of Atlantic cod (Gadus morhua) exposed to PCB 153 indicates effects on cell cycle regulation and lipid metabolism. BMCGenomics 2014, 15:481.

31. Barendse W, Reverter A, Bunch RJ, Harrison BE, Barris W, Thomas MB: A validated whole-genome association study of efficient food conversion in cattle. Genetics 2007, 176(3):1893–1905. PMID: 17507676

32. Zhang D, Lu K, Dong Z, Jiang G, Xu W, Liu W: The effect of exposure to a high-fat diet on microRNA expression in the liver of blunt snout bream (Megalobrama amblycephala). PLoSOne 2014, 9(5):e96132.

33. Guo Z, Maki M, Ding R, Yang Y, Zhang B, Xiong L: Genome-wide survey of tissue-specific microRNA and transcription factor regulatory networks in 12 tissues. SciRep 2014, 4:5150. doi: 10.1038/srep05150 PMID: 24889152

34. Bizuayehu TT, Babiak I: MicroRNA in teleost fish. Genome BiolEvol 2014, 6(8):1911–1937.

35. Friedlander MR, Mackowiak SD, Li N, Chen W, Rajewsky N: miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. Nucleic Acids Res 2012, 40(1):37–52. doi: 10.1093/nar/gkr688 PMID: 21911355

36. Ambros V, Bartel B, Bartel DP, Burge CB, Carrington JC, Chen X, et al. A uniform system for microRNA annotation. RNA 2003, 9(3):277–279. PMID: 12592000

37. Johansen I, Andreassen R: Validation of miRNA genes suitable as reference genes in qPCR analyses of miRNA gene expression in Atlantic salmon (Salmo salar). BMCResNotes 2014, 8:945. doi: 10.1186/1756-0500-7-945 PMID: 25533334

38. Andersen CL, Jensen JL, Orntoft TF: Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Res 2004, 64(15):5245–5250. PMID: 15289330

39. Schmittgen TD, Livak KJ: Analyzing real-time PCR data by the comparative C(T) method. NatProtoc 2008, 3(6):1101–1108. PMID: 18546601

40. Farlora R, Valenzuela-Miranda D, Alarcon-Matus P, Gallardo-Escarate C: Identification of microRNAs associated with sexual maturity in rainbow trout brain and testis through small RNA deep sequencing. MolReprodDev 2015. doi: 10.1002/mrd.22499 PMID: 25983107

41. Dill H, Linder B, Fehr A, Fischer U: Intronic miR-26b controls neuronal differentiation by repressing its host transcript, ctdsp2. Genes Dev 2012, 26(1):25–30. doi: 10.1101/gad.177774.111 PMID: 22215807

42. Wienholds E, Kloosterman WP, Miska E, Alvarez-Saavedra E, Berezikov E, de BE, et al. MicroRNA expression in zebrafish embryonic development. Science 2005, 309(5732):310–311. PMID: 15919954

43. Deep sequencing of small RNA facilitates tissue and sex specific microRNA discovery in zebrafish. Available: http://dx.doi.org/10.1101/019760

44. Nahid MA, Satoh M, Chan EK: MicroRNA in TLR signaling and endotoxin tolerance. Cell MolImmunol 2011, 8(5):388–403.

45. Li C, Zhu HY, Bai WD, Su LL, Liu JQ, Cai WX, et al. MiR-10a and miR-181c regulate collagen type I generation in hypertrophic scars by targeting PAI-1 and uPA. FEBS Lett 2015, 589(3):380–389. doi: 10.1016/j.febslet.2014.12.024 PMID: 25554417

46. Yi R, Fuchs E: MicroRNA-mediated control in the skin. Cell DeathDiffer 2010, 17(2):229–235.

47. Menningen J, Plagnes-Juan E, Figueredo-Silva CA, Seiliez I, Panserat S, Skiba-Cassy S: Acute endocrine and nutritional co-regulation of the hepatic omy-miRNA-122b and the lipogenic gene fas in rainbow trout, Onchorhybchus mykiss. Comparative Biochemistry and Physiology 2014, 169:16–24. doi: 10.1016/j.cbpb.2013.12.002 PMID: 24333236

48. Wang K, Zhang S, Marzolf B, Troisch P, Brightman A, Hu Z, et al. Circulating microRNAs, potential biomarkers for drug-induced liver injury. ProcNatlAcadSciUSA 2009, 106(11):4402–4407.

49. Yan X, Ding L, Li Y, Zhang X, Liang Y, Sun X, et al.: Identification and profiling of microRNAs from skeletal muscle of the common carp. PLoSOne 2012, 7(1):e30925.

50. Kent OA, McCall MN, Cornish TC, Halushka MK: Lessons from miR-143/145: the importance of cell-type localization of miRNAs. Nucleic Acids Res 2014, 42(12):7528–7538. doi: 10.1093/nar/gku461 PMID: 24875473

51. Chen XM, Splinter PL, O'Hara SP, LaRusso NF: A cellular micro-RNA, let-7i, regulates Toll-like receptor 4 expression and contributes to cholangiocyte immune responses against Cryptosporidium parvum infection. JBiolChem 2007, 2 82(39:):28929–28938.

52. Banerjee J, Chan YC, Sen CK: MicroRNAs in skin and wound healing. Physiol Genomics 2011, 43 (10):543–556. doi: 10.1152/physiolgenomics.00157.2010 PMID: 20959495

53. Yan B, Liu B, Zhu CD, Li KL, Yue LJ, Zhao JL, et al.: microRNA regulation of skin pigmentation in fish. JCell Sci 2013, 126(Pt 15):3401–3408.

54. Haflidadottir BS, Bergsteinsdottir K, Praetorius C, Steingrimsson E: miR-148 regulates Mitf in melanoma cells. PLoSOne 2010, 5(7):e11574. doi: 10.1371/journal.pone.0011574 PMID: 20644734

55. Miyaki S, Sato T, Inoue A, Otsuki S, Ito Y, Yokoyama S, et al.: MicroRNA-140 plays dual roles in both cartilage development and homeostasis. Genes Dev 2010, 24(11):1173–1185. doi: 10.1101/gad.1915510 PMID: 20466812

56. Saal S, Harvey SJ: MicroRNAs and the kidney: coming of age. CurrOpinNephrolHypertens 2009, 18 (4):317–323. doi: 10.1097/MNH.0b013e32832c9da2 PMID: 19424061

57. Trakooljul N, Hicks JA, Liu HC: Characterization of miR-10a mediated gene regulation in avian splenocytes. Gene 2012, 500(1):107–114. doi: 10.1016/j.gene.2012.03.028 PMID: 22446044

58. Zhu X, Li YL, Chen DX, Wu P, Yi T, Chen T, et al.: Selection of reference genes for microRNA quantitative expression analysis in Chinese perch, Siniperca chuatsi. IntJMolSci 2015, 16(4):8310–8323.

59. Baskerville S, Bartel DP: Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. RNA 2005, 11(3):241–247. PMID: 15701730