Comparative Biochemistry and Physiology - Part D 31 (2019) 100597

Contents lists available at ScienceDirect

Comparative Biochemistry and Physiology - Part D

journal homepage: www.elsevier.com/locate/cbpd

This article is a part of the Special Issue on Aquaculture

Whole transcriptome analysis of the Atlantic cod vaccine response reveals subtle changes in adaptive immunity☆

Monica Hongrø Solbakken[a], Sissel Jentoft[a,*], Trond Reitan[a], Helene Mikkelsen[b], Kjetill S. Jakobsen[a], Marit Seppola[c,*]

[a] Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biosciences, University of Oslo, Oslo, Norway
[b] The Northern Norway Regional Health Authority, Tromsø, Norway
[c] Department of Medical Biology, The Arctic University of Norway, Tromsø, Norway

A B S T R A C T

Atlantic cod has lost the Major Histocompatibility complex class II pathway – central to pathogen presentation, humoral response and immunity. Here, we investigate the immunological response of Atlantic cod subsequent to dip vaccination with *Vibrio anguillarum* bacterin using transcriptome sequencing. The experiment was conducted on siblings from an Atlantic cod family found to be highly susceptible towards vibriosis where vaccination has demonstrated improved pathogen resistance. Gene expression analyses at 2, 4, 21 and 42 days post vaccination revealed GO-term enrichment for muscle, neuron and metabolism-related pathways. In-depth characterization of immune-related GO terms demonstrated down-regulation of MHCI antigen presentation, C-type lectin receptor signaling and granulocyte activation over time. Phagocytosis, interferon-gamma signaling and negative regulation of innate immunity were increasingly up-regulated over time. Individual differentially expressed immune genes implies weak initiation of acute phase proteins with little or no inflammation. Furthermore, gene expression indicates presence of T-cells, NK-like cells, B-cells and monocytes/macrophages. Three *MHCI* transcripts were up-regulated with *B2M* and *SEC61*. Overall, we find no clear immune-related transcriptomic response which could be attributed to Atlantic cod's alternative immune system. However, we cannot rule out that this response is related to vaccination protocol/sampling strategy. Earlier functional studies demonstrate significant memory in Atlantic cod post dip vaccination and combined with our results indicate the presence of other adaptive immunity mechanisms. In particular, we suggest that further investigations should look into CD8+ memory T-cells, γδ T-cells, T-cell independent memory or memory induced through NK-like/other lymphoid cells locally in the mucosal lining for this particular vaccination strategy.

1. Introduction

Teleosts are among the oldest, most diverse and numerous vertebrate infraclasses (Eschmeyer and Fricke, 2015). Recently, it became evident that the genetic and phenotypic diversification of this lineage also has manifested through a diverse repertoire of immune-related genes (Zhu et al., 2013; Malmstrøm et al., 2016; Solbakken et al., 2016a). In teleost species classified with a conventional set of genes related to adaptive immunity such as Major Histocompatibility Complex class II (*MHCII*), *CD4*, T-cell and B-cell receptors, functional and transcriptomic studies demonstrate more or less a classic immune response (Gao et al., 2014; Jiang et al., 2014; Makesh et al., 2015; Haase et al., 2016; Jiang et al., 2016; Hoare et al., 2017; Maekawa et al., 2017; Tang et al., 2018). The expression profiles of specific immune genes varies some according to tissue examined, type of pathogen and route of pathogen transmission, but tend to demonstrate up-regulation of pro-inflammatory genes, pattern recognition, phagocytosis, complement activation, antigen processing and B- and T-cell activation to mention some (Jiang et al., 2014; Haase et al., 2016; Jiang et al., 2016; Maekawa et al., 2017). In terms of vaccination experiments, a large degree of variation in response and immunity are observed – mainly linked to vaccination method and tissues samples. However, responses could in general be attributed to conventional adaptive immunity (Jiang et al., 2014; Liu et al., 2015; Makesh et al., 2015; Hwang et al., 2017; Zhang et al., 2017). This can be exemplified by zebrafish, which upon vaccination elicit both innate and adaptive responses. More specifically, there is an initial acute phase response and likely detection performed by a range of recognition receptors (PRRs). Subsequently, there is an onset of inflammation in parallel with antimicrobial defenses. Antigen processing and presentation ensues together with up-regulation of *MHCI* and *MHCII*, B- and T-cell activation and in most cases expression of antibodies (Liu et al., 2015). Intriguingly, for some teleost species, such as Atlantic cod (*Gadus morhua*) and haddock (*Melanogrammus aeglefinus*), functional studies collectively describe inconsistent antibody response patterns and somewhat variable antibody specificity (Corripio-Miyar et al., 2007; Lund et al., 2007; Caipang et al., 2008; Caipang et al., 2009; Gudmundsdottir et al., 2009; Magnadottir et al., 2009; Schrøder et al., 2009; Ellingsen et al., 2011; Mikkelsen

☆ This article is part of a special issue entitled: Aquaculture- edited by Dr. Matt Rise, Dr. Muyan Chen and Dr. Chris Martyniuk.

* Corresponding authors.
E-mail addresses: sissel.jentoft@ibv.uio.no (S. Jentoft), marit.seppola@uit.no (M. Seppola).

https://doi.org/10.1016/j.cbd.2019.100597
Received 19 October 2018; Received in revised form 21 May 2019; Accepted 24 May 2019
Available online 28 May 2019

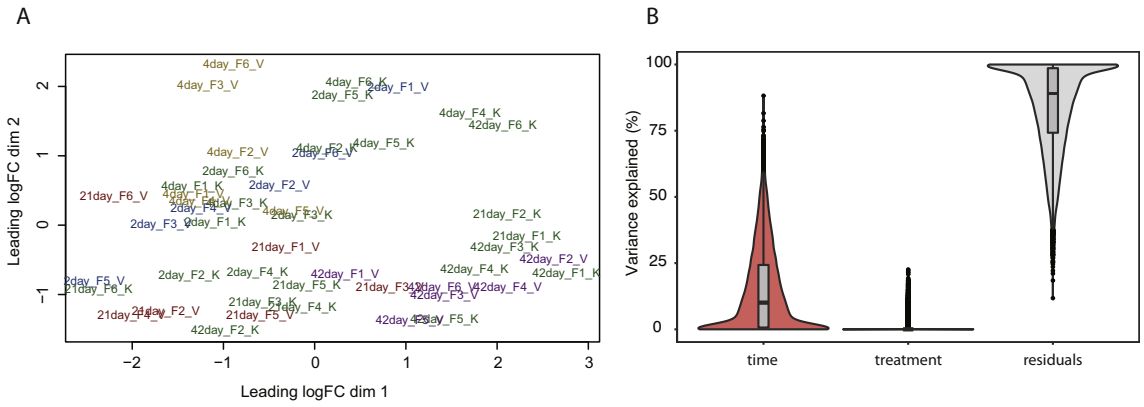

Fig. 1. A) MDS plot generated from the Trinity-derived counts using the edgeR MDSplot function plotting dimensions 1 and 2. All control samples are colored dark green. Day 2 samples are blue, day 4 samples are yellow, day 21 samples are red and day 42 samples are purple. There is clustering of the vaccinated samples with some overlap. In contrast the control samples do not cluster in a clear pattern and forms a background of noise in relation to the vaccinated samples. A similar plot from the cufflinks derived counts is available in the supplementary material demonstrating a similar pattern – SFig. 7. B) A variance partition plot made with the VariancePartition R package using counts from the *de novo* transcriptome. VariancePartition uses a linear mixed model to quantify the contribution of variance from different sources.

et al., 2011; Mikkelsen and Seppola, 2013; Nymo et al., 2016). By whole genome sequencing, it was demonstrated that Atlantic cod, as well as the entire Gadiformes lineage, lacks the MHCII pathway. This loss is further accompanied by an extreme expansion of *MHCI* genes with > 100 copies in some species (Star et al., 2011; Malmstrøm et al., 2016; Tørresen et al., 2018), and different repertoires of innate immune genes related to pattern recognition (Solbakken et al., 2016a; Solbakken et al., 2016b; Solbakken et al., 2017; Tørresen et al., 2018). From this, the question arose if and how Atlantic cod establishes immunity. Studies investigating the vaccine response of single genes in Atlantic cod describe overall inflammation, antimicrobial peptides, acute phase proteins as well as some interferon-related genes and cell-mediated cytotoxicity depending on the vaccination method (Lund et al., 2007; Caipang et al., 2008; Lund et al., 2008; Caipang et al., 2009; Gudmundsdottir et al., 2009; Magnadottir et al., 2009; Schrøder et al., 2009; Bakkemo et al., 2011; Ellingsen et al., 2011; Mikkelsen et al., 2011; Mikkelsen and Seppola, 2013; Vestvik et al., 2013; Seppola et al., 2015; Eslamloo et al., 2016; Nymo et al., 2016). However, there are no studies currently describing the global gene expression profile underlying proven post vaccination immunity in these teleost species lacking the MHCII pathway. The rapid development of sequencing technology, combined with improved genomic resources and development of bioinformatical tools, now enables us to perform large-scale transcriptomics in non-model species. Thus, the need for more overarching approaches to characterize the specific pathways used by this particular immune system can be fulfilled.

Here, we present a global transcriptome analysis of the immune response in Atlantic cod subsequent to a dip vaccination. The target disease, vibriosis, occurs in a range of teleost species both under farmed and wild conditions. Left untreated the infection causes fatal hemorrhagic septicemic disease (Frans et al., 2011). The fish used in this experiment was selected from family material provided by the Norwegian national breeding program of Atlantic cod, *i.e.* siblings from a family shown to be highly susceptible to vibriosis and proven to have an increased resistance against the pathogen post vaccination (Mikkelsen and Seppola, 2013). The vaccine consisted of gram negative *Vibrio anguillarum* (*Listonella anguillarum*) bacterin (i.e. suspension of attenuated bacteria). Our transcriptome analyses demonstrate a weak acute phase response and no evident inflammation. In contrast, highly profound transcriptome responses involving diffuse muscle- and neuron development as well as metabolic pathways were evident. Selecting enriched immune-related GO-terms only and investigating individual significantly differentially expressed immune genes we observe that MHCI antigen presentation is affected and that there is a potential presence of B- and T-cells, NK-like cells and monocytes/macrophages. Thus, even though these Atlantic cod showed increased resistance post vaccination, we could not detect a clear measurable response revealing which underlying immune-related mechanisms are responsible for immunity in Atlantic cod.

2. Results

In this study, we have chosen a multifaceted approach to detect differentially expressed genes. It consists of both *de novo* and reference genome based transcriptomics (Trinity (Haas et al., 2013) and Tuxedo (Trapnell et al., 2012), respectively) combined with EdgeR (Robinson et al., 2010) and Cufflinks/CummeRbund (Trapnell et al., 2012) for final differential expression analysis and result visualization. In addition, due to our experimental setup with pair-wise controls at all timepoints, we have applied a custom GLM-based analysis for clustering of genes by their expression pattern over time providing pre-defined patterns of interest (for details see Methods). Overall, the different analysis approaches applied to this dataset highlight different aspects of the vaccine response but collectively report a core set of significant differentially expressed genes.

Overall inspection of all analyses (EdgeR and Cufflinks) using MDS plots demonstrate that vaccinated individuals tend to cluster together, but the pairwise control samples form background of noise (Fig. 1A). A similar plot was generated from the Cufflinks-derived counts showing the same clustering pattern (SFig. 7). Furthermore, when attempting to localize the source of biological variation in our dataset we find that time is the major contributor, whereas the treatment effect is very small (Fig. 1B). This pattern does not appear to be related to technical issues in terms of RNA/library quality and quality of the reads (see STable 1 for experiment overview).

Due to the noisy background, the samples were subjected to both a pair-wise and a time:treatment GLM analysis in addition to the GLM-based expression profile clustering. The number of significantly differentially expressed genes/transcripts varies, but all analyses display similar trends. The number of reported genes is greatly reduced when taking advantage of the reference genome and applying a simple pair-wise analysis. However, we also find that the Trinity assembly supports the reference genome by providing better resolution in terms of missing/poorly annotated genes and novel transcripts, and that the GLM analyses detect subtle changes in expression not reported by the pair-wise analyses. Finally, the largest obstacle is the lack of annotation, especially for transcripts in the *de novo* transcriptome assembly, and we observe differential expression of a large amount of transcripts without

Table 1
Significantly differentially expressed genes reported for the different analyses applied to counts from both the transcriptome assembly (Trinity *de novo*) as well as the genome reference (cufflinks). The last column presents the number of genes with corresponding annotation, but duplications are not counted. These also represent the annotations used as input in the GO-term enrichment analysis (ClueGO). The custom script clustering of expression profiles reported genes in both in a control-dependent and independent manner. Here, we report only genes from the control-dependent analysis.

Time-point or expression pattern	Number of genes	Number of genes with annotation
Trinity *de novo*, edgeR pairwise contrasts		
2 day up	37	4
4 day up	0	0
21 day up	75	21
42 day up	1	0
2 day down	3	0
4 day down	7	4
21 day down	9	4
42 day down	0	0
Trinity *de novo*, edgeR GLM time-series		
2 day up	1927	597
4 day up	0	0
21 day up	0	0
42 day up	30	24
2 day down	1317	760
4 day down	0	0
21 day down	0	0
42 day down	80	37
Trinity *de novo*, custom script RSEM counts		
Increase time	600	124
Internal max time	121	45
Decrease time	545	210
Internal min time	900	231
Freestyle time	1360	640
Reference based, cufflinks, pairwise		
2 day up	116	67
4 day up	179	134
21 day up	269	204
42 day up	141	105
2 day down	47	32
4 day down	10	3
21 day down	91	60
42 day down	321	252
Reference based, custom script, cufflinks counts		
Increase time	180	98
Internal max time	36	22
Decrease time	261	201
Internal min time	62	28
Freestyle time	176	77

clear homology to genes with known function in either fish or mammals (Table 1).

All differentially expressed genes with annotation was subjected to a GO-term enrichment analysis using ClueGO (Bindea et al., 2009). Collectively, GO:terms significantly enriched from up-regulated genes on day 2 and day 4 describe muscle-related categories, metabolism, cellular organelles and some signaling. In contrast, down-regulated genes and their corresponding enriched GO-terms contains more structural GO:terms such as different cell components (Table 2).

Further, on day 21 and 42, the up-regulated genes mainly represent GO-terms related to neurons and synapses. In contrast, the down-regulated genes now reflect the muscle-related GO categories reported as positively enriched at day 2 and day 4 (Table 3).

Some of the major GO-term categories revealed underlying mechanisms with potential immune-related functionality. Genes positively differentially expressed over time derived from the *de novo* transcriptome revealed potential interferon gamma mediated signaling, negative regulation of innate immunity, formation of phagocytic vesicles and COPII vesicle transport and positive regulation of immune receptors. This was observed in parallel with blood homeostasis and regulation of apoptosis (Fig. 2).

Our overall analysis using GO-term enrichment did not reveal any stand-alone specific immune-related pathways. However, we detect several immune-related GO-clusters within the larger non-immune related clusters (Supplementary Figs. 8–30, Supplementary excel file 6). Therefore, we ran a GO-term enrichment selecting immune-related GO categories only to improve resolution. Here we found significant immune-related enrichment for both down-regulated and up-regulated genes. Terms related to antigen presentation are found to be down-regulated or have a negative quadratic expression profile (internal minimum) over time. In contrast, interferon gamma signaling and phagocytosis is up-regulated over time together with negative regulation of innate immunity (Fig. 3).

The GO-term enrichment analysis is restricted by Atlantic cod not being a model species, and fish-specific annotations are not readily recognized. Thus, we further explored the data by identifying individual significantly differentially expressed immune-related genes. We observe a modest activation of acute phase genes here represented by hemopexin (*HPX*) and transferrin in addition to some fibrinogen and apolipoproteins (latter two in Supplementary excel files 1, 2, 3, 4, 5). However, we find no clear signs of inflammation initiation and the genes transcribing inflammasome components (*PYCARD*, *CASP1*) are down-regulated. In terms of pattern recognition, we observe up-regulation of a few NOD-like receptor transcripts (*NLRC3* and *NLRP3*) and the Toll-like receptor adaptor protein *MYD88*. Finally, we observe a modest antibacterial response through the up-regulation of piscidin and *LYG1* genes. In terms of cell populations, a small set of putative cell markers are positively differentially expressed: T-cells (*CD8B*, *TRGC1*, *TNFSF11*, *EOMES*), NK-like cells (*LITR_NITR*, *B3GAT1*), B-cells (*IGLC2*), cells with granules and perforin activities (*PERF1*, *UNC13D*) and monocytes and macrophages (*IL34*). Of the overall 46 annotated *MH*CI Trinity transcripts, four were reported as significantly differentially expressed (three up-regulated, one down-regulated). Furthermore, we find up-regulation of *B2M* and *SEC61* together with several heat-shock proteins (heat shock proteins listed Supplementary excel files 1-5). The proteasomal subunits overall are down-regulated over time. Similarly, we observe down-regulation of complement and apoptosis over time/late in the experiment together with a small set of chemokines and cytokines – mainly chemoattractants (Table 4).

Additionally, to evaluate the transcriptome analysis conducted, we compared our expression data with already published qPCR analysis of selected immune genes performed on the same individuals. This data was published together with an overall vaccine trial experiment of which the current samples and corresponding data is derived from (Mikkelsen and Seppola, 2013). The contrasted qPCR findings to corresponding expression data from the *de novo* pairwise analysis supports our results, *i.e.* showing no significant differential expression of key immune genes in response to vaccination (Table 6).

3. Discussion

3.1. Subtle transcriptomic responses indicate an adaptive immune response

Vaccine responses in teleosts towards various pathogens, regardless of immunization protocol (mucosal or injection techniques), have mainly been characterized by using quantitative PCR methods or estimating antibody titers. These studies have demonstrated that most teleost display a more or less convention immune response with some degree of individual variation, which generally appears to be connected to vaccination method, vaccine content and/or adjuvant used (Schrøder et al., 2006; Corripio-Miyar et al., 2007; Caipang et al., 2008; Caipang et al., 2009; Gudmundsdottir et al., 2009; Mikkelsen et al., 2011; Ballesteros et al., 2012; Pridgeon et al., 2012; Sarropoulou et al., 2012; Mikkelsen and Seppola, 2013; Zhang et al., 2013; Chettri et al., 2015; Liu et al., 2015). Furthermore, functional studies of antibody responses

Table 2
All main GO-term categories reported for pairwise and GLM analyses, for both Cufflinks and Trinity, separated into up-regulated and down-regulated early on in the experiment *i.e.* day 2 and day 4. All GO-networks are available in the supplementary information.

	Upregulated GO:terms day 2 and 4		Downregulated GO:terms day 2 and 4
GO:0030049	Muscle filament sliding	GO:0016482	Cytoplasmic transport
GO:0031672	A band	GO:0005737	Cytoplasm
GO:0031674	I band	GO:0044444	Cytoplasmic part
GO:0016459	Myosin complex	GO:0060205	Cytoplasmic membrane-bounded vesicle lumen
GO:0043292	Contractile fiber	GO:0071682	Endocytic vesicle lumen
GO:0044449	Contractile fiber part	GO:0034774	Secretory granule lumen
GO:0044291	Cell-cell contact zone	GO:0043232	Intracellular non-membrane-bounded organelle
GO:0030018	Z disc	GO:0044446	Intracellular organelle part
GO:0060047	Heart contraction	GO:0072562	Blood microparticle
GO:0032982	Myosin filament	GO:0043933	Macromolecular complex subunit organization
GO:0030837	Negative regulation of actin filament polymerization	GO:0071339	MLL1 complex
GO:0035987	Endodermal cell differentiation	GO:0000278	Mitotic cell cycle
GO:0005773	Vacuole	GO:1902494	Catalytic complex
GO:0031090	Organelle membrane	GO:0005739	Mitochondrion
GO:0005829	Cytosol	GO:0016071	mRNA metabolic process
GO:0006996	Organelle organization	GO:0010605	Negative regulation of macromolecule metabolic process
GO:0044444	Cytoplasmic part	GO:0006412	Translation
GO:0044446	Intracellular organelle part	GO:0015630	Microtubule cytoskeleton
GO:0030193	Regulation of blood coagulation	GO:0044428	Nuclear part
GO:0050750	Low-density lipoprotein particle receptor binding	GO:0031072	Heat shock protein binding
GO:0031958	Corticosteroid receptor signaling pathway	GO:0044822	Poly(A) RNA binding
GO:0048583	Regulation of response to stimulus	GO:0044260	Cellular macromolecule metabolic process
GO:1900169	Regulation of glucocorticoid mediated signaling pathway	GO:0030529	Intracellular ribonucleoprotein complex
GO:0016620	Oxidoreductase activity, acting on the aldehyde or oxo group of donors, NAD or NADP as acceptor	GO:0016032	Viral process
GO:0044236	Multicellular organismal metabolic process	GO:0002223	Stimulatory C-type lectin receptor signaling pathway
GO:0006090	Pyruvate metabolic process		
GO:0044267	Cellular protein metabolic process		
GO:0006661	Phosphatidylinositol biosynthetic process		
GO:0031463	Cul3-RING ubiquitin ligase complex		
GO:0042383	Sarcolemma		
GO:0007586	Digestion		
GO:1990234	Transferase complex		
GO:0016301	Kinase activity		
GO:0016568	Chromatin modification		
GO:0019899	Enzyme binding		
GO:0044765	Single-organism transport		

report divergent results, *i.e.* from heterogenic repertoire of natural antibodies, through pathogen-lipopolysaccharide-specific antibodies to specific antibodies able to discriminate between pathogen serogroups (Corripio-Miyar et al., 2007; Lund et al., 2007; Gudmundsdottir et al., 2009; Magnadottir et al., 2009; Schrøder et al., 2009;Mikkelsen et al., 2011; Mikkelsen and Seppola, 2013). Mucosal vaccination (dip or immersion) towards vibriosis in Atlantic cod has been collectively shown to weakly, but significantly, induce the innate immune system through

Table 3
All main GO-term categories reported for pairwise and GLM analyses, for both Cufflinks and Trinity, separated into up-regulated and down-regulated later on in the experiment *i.e.* day 21 and day 42. All GO-networks are available in supplementary information.

	Upregulated GO:terms day 21 and 42		Downregulated GO:terms day 21 and 42
GO:0050804	Modulation of synaptic transmission	GO:0006939	Smooth muscle contraction
GO:0097458	Neuron part	GO:0006941	Striated muscle contraction
GO:0030424	Axon	GO:0016529	Sarcoplasmic reticulum
GO:0042552	Myelination	GO:0030016	Myofibril
GO:0043209	Myelin sheath	GO:0030048	Actin filament-based movement
GO:0048812	Neuron projection morphogenesis	GO:0030049	Muscle filament sliding
GO:0048857	Neural nucleus development	GO:0031672	A band
GO:0031103	Axon regeneration	GO:0042641	Actomyosin
GO:0044304	Main axon	GO:0043292	Contractile fiber
GO:0070252	Actin-mediated cell contraction	GO:0051015	Actin filament binding
GO:0030705	Cytoskeleton-dependent intracellular transport	GO:0005578	Proteinaceous extracellular matrix
GO:0015672	Monovalent inorganic cation transport	GO:0005912	Adherens junction
GO:0030426	Growth cone	GO:0031018	Endocrine pancreas development
GO:0030534	Adult behavior	GO:0031589	Cell-substrate adhesion
GO:0044297	Cell body	GO:0035987	Endodermal cell differentiation
GO:0045121	Membrane raft	GO:0044853	Plasma membrane raft
GO:0015631	Tubulin binding	GO:0048589	Developmental growth
GO:0030913	Paranodal junction assembly	GO:0048646	Anatomical structure formation involved in morphogenesis
GO:0010770	Positive regulation of cell morphogenesis involved in differentiation		
GO:1901016	Regulation of potassium ion transmembrane transporter activity		
GO:0010642	Negative regulation of platelet-derived growth factor receptor signaling pathway		

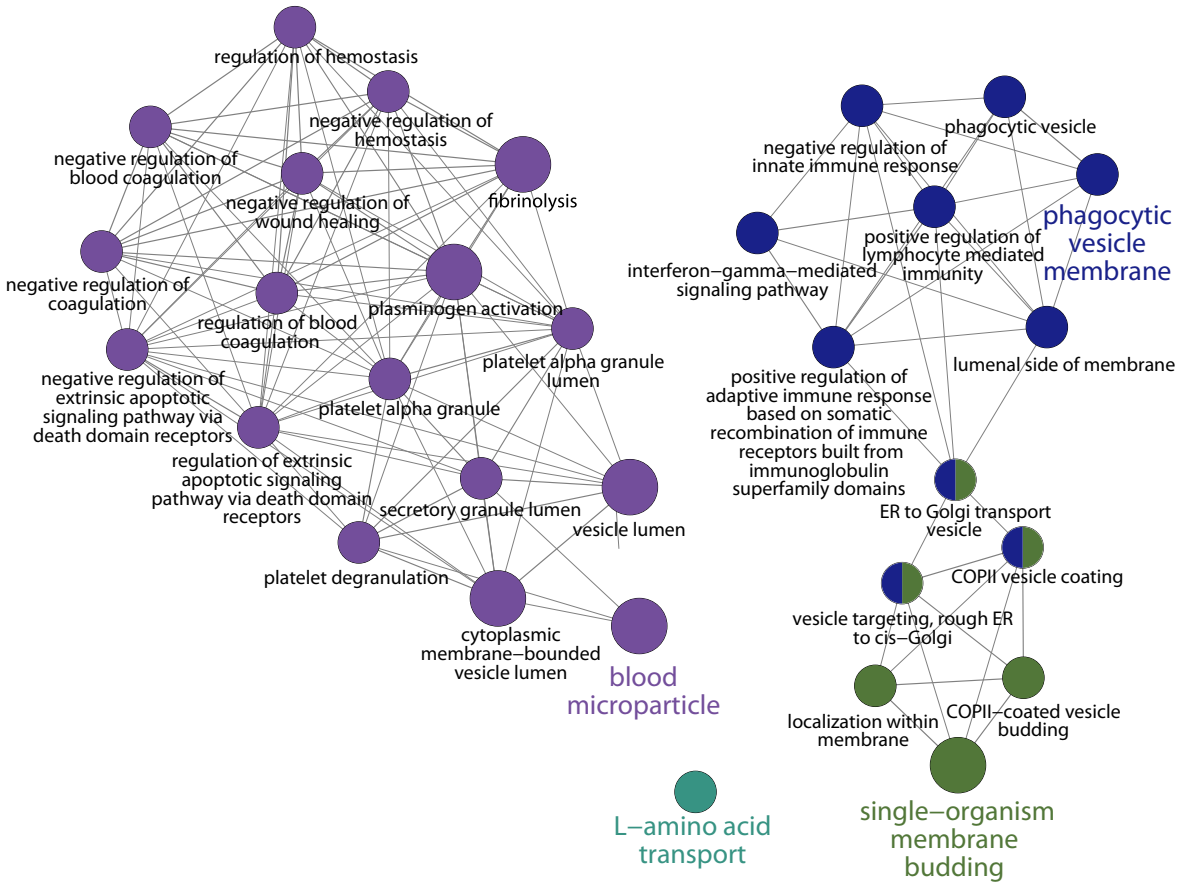

Fig. 2. GO-term enrichment network made with ClueGO, p-value cutoff = 0.05. Underlying genes are significantly differentially expressed with an increasing expression profile over time in a control-dependent manner. Counts derived from the *de novo* transcriptome.

up-regulation of *IL1B*, *IL10* and *IL12B* (the latter is a subunit of both *IL12* and *IL23*) together with hepcidin (*HAMP*). Induction of interferon gamma (*IFNG*) and lysozyme (*LYG2*) varies somewhat and appears to be correlated with vaccine content (Mikkelsen et al., 2011; Mikkelsen and Seppola, 2013). In contrast, injection-based strategies report a clear systemic up-regulation of *IL1B*, lysozyme *LYG2*, transferrin (*TF*), interferon regulatory factor 1 (*IRF1*), granzyme, catalase, *IL8* (*CXCL8*) and apolipoproteins to mention some (Caipang et al., 2008; Caipang et al., 2009; Feng et al., 2009). By in-depth transcriptome profiling, we find that the innate defenses in Atlantic cod are only marginally affected by dip vaccination against *V. anguillarum* with limited up-regulation of acute phase protein genes and no evident gene expression indicating inflammation or regulation thereof. These results are concordant with the weak differential expression of immune genes previously reported using qPCR on the same family material and experimental setting (Table 6) (Mikkelsen and Seppola, 2013). In terms of adaptive responses, teleosts harboring a classic adaptive immune system vaccination is shown to affect expression of *MHCI* and *MHCII* as well as the corresponding co-receptors on T-cells (*CD4* and *CD8*) together with an increase of immunoglobulin expression (Jiang et al., 2014; Liu et al., 2015; Makesh et al., 2015; Hwang et al., 2017). Here, we observe both significant up- and down regulation of different *MHCI* transcripts as well as a few immunoglobulin light chain fragments (Table 4, Supplementary table 2). In addition, we found B2M and SEC61 up-regulation, up-regulation of heat-shock proteins and down-regulation of several proteasomal subunits (Table 4, supplement). The divergent expression pattern of the specific MHCI transcripts could be correlated with the reported MHCI gene expansion and presence of different endosomal sorting signals in Atlantic cod (Malmstrøm et al., 2013). These signals have been suggested involved in alternative loading of antigen on MHCI (cross presentation) (Malmstrøm et al., 2013). The observed, though small, significant changes to the phagosomal pathway (Table 4) together with different *MHCI* expression patterns and up-regulation of *SEC61* could indicate use of cross-presentation on specific MHCI molecules as part of the vaccine response (Neefjes et al., 2011; Zehner and Burgdorf, 2015). We further observed the up-regulation of genes related to cytotoxic and lysosome functions such as perforin (*PRF1*) and cathepsin (*CTSC/CTSK/CTSS*) (Table 4) indicating some increase in phagocytosis and cellular cytotoxicity defense mechanisms (Trapani and Smyth, 2002; Neefjes et al., 2011).

3.2. Vaccination and the enriched muscle and neuron development GO-terms

The rather modest number of significantly differentially expressed immune genes in this study indicates that the immune system in Atlantic cod is not heavily influenced by dip vaccination. However, major responses from other non-immune system related pathways were observed, including large changes in muscle- and neuron-related pathways (Table 1). Similar findings have been reported in other teleost studies where gene regulation post vaccination has been investigated using microarrays. These studies reported that the contribution by the immune system was overshadowed by other pathways such as metabolism (Pridgeon et al., 2012; Zhang et al., 2013; Jiang et al., 2014). Further, vaccination has been shown to have an effect on growth rate in salmon as well as the development of the heart muscle (Berg et al., 2006; Fraser et al., 2015). There is also a study demonstrating that heart muscle cells from Atlantic cod remove plasmid DNA from the blood by endocytosis *in vitro* (Seternes et al., 2007). Thus, both development and growth combined with a likely immune-related function

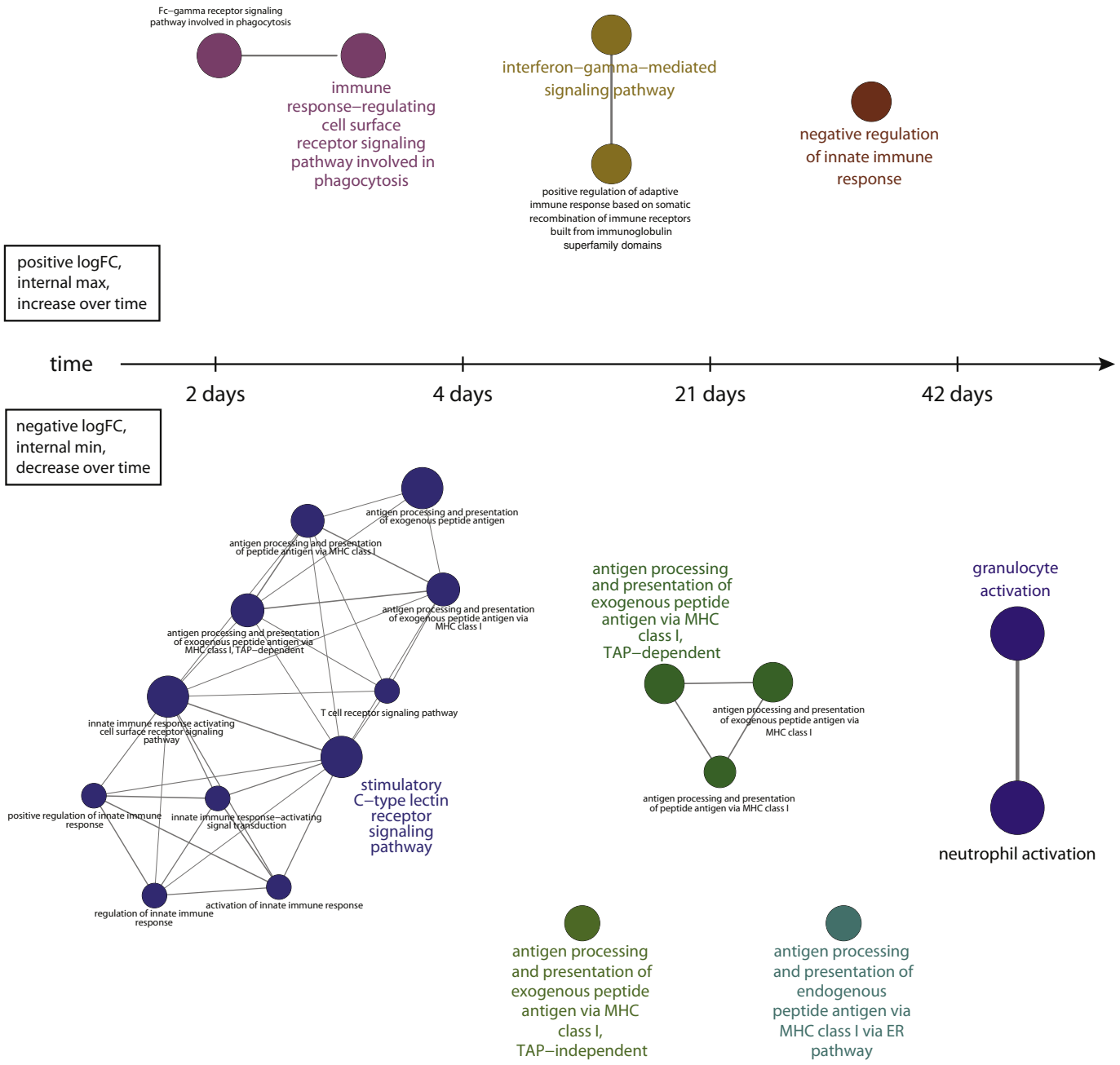

Fig. 3. Selected GO-term networks from the GO-term enrichment in ClueGO put together over a drawn time-line. The GO-networks are derived from significantly differentially expressed genes from pairwise, GLM and custom script analyses. All GO-networks are available in supplementary information.

for heart cells can explain the GO enrichment of these pathways in our dataset. Dip vaccination is a mucosal vaccination strategy and targets mucosal linings such as gills, skin, oral and anal openings. Mucosal tissues in fish are physical barriers that are considered active immune tissues that execute multiple functions like osmoregulation, oxygenation, nutrient uptake and environmental censoring. In addition to epithelial cells and mucus-producing cells, fish skin also contains neuroendocrine and neuroepithelial cells. Their function and interaction with the immune system is still unclear but it is suggested they regulate and maintain homeostasis, both in terms immunity and physiology, by secreting hormones, neuropeptides and neurotransmitters (Parra et al., 2015 and references therein) and thus could explain the strong presence of neuronal GO-terms in our study.

3.3. Establishment of memory

Establishment of memory in species with a classical adaptive immune system – including the MHCII pathway – is generally described as an interaction between a CD4+ T-cell and an antigen presenting cell where the T-cell has αβ T-cell receptors (TcRs) and the antigen presenting cells has MHCII. However, memory is also known to be established in CD8+ T-cell lineages which are present both in mucosal tissues as well as systemically in mammals (Jameson and Masopust, 2009; Lauvau et al., 2016). As Atlantic cod lacks CD4 (and MHCII), and thus does not harbor the conventional route for antigen presentation *via* CD4+ T-cells, CD8+ T-cell memory lineages could be an alternative mechanism leading to immunity. In this regard, we do observe an up-regulation of *CD8B* as well as a T-cell co-stimulator *CD46*. However, we do not detect any of the cytokines involved in differentiation of the

Table 4
Overview of individual immune-related genes and their overall expression pattern across all analyses. Multi-copy/multi-transcript genes have been collapsed and their variable significant expression patterns are evident by the multiple ticked boxes. LogFC values are added for genes reported as up-regulated in pairwise or glm time:treatment analyses in Table 5.

		Up early	Increase	Int.max	Up late	Down early	Decrease	Int.min	Down late
Acute phase, inflammation, antibacterial									
Activates serine proteases of the immune system	CTSC		x						
Acute phase protein, hemopexin	HPX		x		x	x			
Lysosomal antigen processing/antibacterial activity	IL4I1								x
Lectin, acute phase response, neutrophil activation	LGALS3								x
Lysozyme	LYG1		x		x				
Antibacterial, fish specific	Piscidin	x							
Acute phase protein, plasminogen	PLG					x			
Inflammasome	PYCARD					x			
Acute phase protein, transferrin	TF		x		x	x			
T-cells/NK-cells /B-cells/dendritic cells									
NK cell marker	B3GAT1			x					
T-cells	CD8B	x							
KLRB1 receptor, inhibits NK-cell mediated lysis	CLEC2D						x		
T-cell differentiation	EOMES								x
Immunoglobulin, light chain	IGLC2	x	x				x	x	x
Leukocyte/novel immune-type receptors, fish specific	LITR_NITR_2		x				x		
Perforin, granule-dependent cell death	PRF1	x			x				x
Cytokine, dendritic cell - > T-cell stimulation	TNFSF11						x		x
T-cell receptor gamma	TRGC1	x	x						
Granule maturation and docking, lymphocytes	UNC13D	x						x	
MHCI/MHCII									
MHCI	B2M		x						
Protease/removes invariant chain from MHCII	CTSK/CTSS							x	
MHCI antigen trimming	ERAP1					x			
MHCI	HLA-A/B	x	x			x		x	x
MHCI, antigen transport	TAP2							x	
Endoplasmatic reticulum/peptide processing									
Endosomal marker	EEA1	x							
Lysosome marker	LAMP1					x			
Proteasome subunit	PSMB1						x		
Proteasome subunit	PSMB2					x			
Proteasome subunit	PSMB5	x					x		
Proteasome subunit	PSMB6					x	x		
Proteasome subunit	PSMB7	x						x	
Proteasome subunit	PSMC6					x	x		
Proteasome subunit	PSMD11					x	x		
Proteasome subunit	PSMD6						x		
Proteasome subunit	PSMD7					x	x		
Proteasome subunit	PSMD8						x		
Protein transport into endoplasmatic reticulum	SEC61A1/62	x					x		
Pattern recognition receptors (PRRs)									
PRR on dendritic cells.	CD209					x			
Regulates PRR DDX58/RIG-I and IFIH1/MDA signaling	DHX58/IFIH1				x			x	
PRR, toll-like receptor adaptor protein	MYD88	x							
PRR, NLR family CARD domain containing	NLRC3	x			x	x	x	x	x
PRR, NACHT, LRR and PYD domain containing	NLRP12						x		
PRR, NACHT, LRR and PYD domain containing	NLRP3		x	x				x	
Complement, caspases, cell death									
Cell death	BCL2								x
Cell death inhibitor	BCL2L1	x							
Complement component	C1QBP					x			
Complement component	C1QL4				x				
Complement component	C6								x
Caspase cascade, cleaves IL1	CASP1					x			
Caspase cascade, apoptosis	CASP6								x
Caspase cascade, cell death	CASP8					x		x	
Complement cofactor, co-stimulatory for T-cells	CD46		x						
Complement inhibitor	CD59		x		x				
Complement C1q receptor	CD93								x
Complement factor	CFD						x		x
Apoptosis regulator	FADD					x			
Lectin, strong inducer of T-cell apoptosis	LGALS1					x	x		x
Apoptosis, regulation	MCL1							x	
Tumor necrosis factor family, apoptosis	TNFSF10								x
Chemokines and cytokines									
Chemoattractant, monocytes	CCL2						x		
Chemoattractant, neutrophils	CXCL1								x

(*continued on next page*)

Table 4 (*continued*)

		Up early	Increase	Int.max	Up late	Down early	Decrease	Int.min	Down late
Chemoattractant, T-cells, monocytes	CXCL12								x
Chemoattractant, neutrophils, basophils and T-cell	CXCL8						x		x
Cytokine, proliferation, survival, differentiation of monocytes and macrophages	IL34				x				x

CD8+ memory lineages (Raeber et al., 2018). There is also down-regulation of EOMES – a CD8+ memory T-cell marker (Lauvau et al., 2016). There are two classes of T-cells present in teleost fish expressing either αβ or γδ TcRs (Scapigliati et al., 2018) and we observe significant up-regulation of two TcR gamma chain transcripts. A subset of mucosal T-cells uses γδ TcRs instead of the more common αβ. In mammals, the γδ T-cells are phenotypically different and have TcRs with little diversity (Salinas, 2015). γδ T-cells function in recognizing unconventional antigens such as microbial metabolites and lipids without the help of MHCI or MHCII. Furthermore, their TcRs can function directly as a pattern recognition receptor (PRRs) (Wan et al., 2016). However, they do perform more classic functions, *i.e.* being antigen presenting cells and provide CD4+ and CD8+ like T-cell help. Thus, they can act as a potential link between adaptive and innate immunity (Wan et al., 2016). In seabass, trout, salmon and zebrafish these cells are generally CD8+ (Rombout et al., 2014; Wan et al., 2016). γδ T-cells have been connected to the production of IgT (and implicated in IgM) at the mucosal surface in teleost fish (Wan et al., 2016) and do display memory functionality in mammals (Sheridan et al., 2013). As we find a TcR gamma chain, *CD8B* and immunoglobulin light chains significantly up-regulated in our data (Table 4) and have constitutive expression of *IgM*, *IgD* and *TcR* delta (Supplementary Figs. 3–7) the overall genetic repertoire is present to enable a CD8+ γδ TcR memory cell population. However, we cannot rule out other CD8+ T-cell lineages playing a major role in Atlantic cod immunity. Furthermore, considering that Atlantic cod lacks *CD4* (Star et al., 2011), and that fish do not form germinal centers (Magor, 2015), T-cell independent activation of B-cells could explain the differentially expressed immunoglobulin light chains (Vinuesa and Chang, 2013). However, there was no further indication of B-cell activation (Table 4). Collectively, we find no clear transcriptomic signals indicating establishment of memory using conventional pathways in Atlantic cod.

3.4. Alternative memory strategies

In mammals, innate memory has also been attributed to several

Table 5
Expression intensity in logFC of transcripts reported as significantly differentially expressed in any of the pairwise analyses as listed in Table 4. * indicates genes with multiple annotated transcripts. If a transcript is differentially expressed at several time points, the time points and corresponding logFC values are listed together separated by /. Overview of all transcripts and their corresponding logFC or expression pattern can be found in the supplementary excel files.

Gene identifier	Timepoint	logFC	Gene identifier	Timepoint	logFC
BCL2	Day42	−1.73	IgL(IGKC/IGLC2)*	Day2	1.39
BCL2L1	Day2	0.58	IgL(IGKC/IGLC2)*	Day2	2.18
C1QBP	Day2	−0.39	IgL(IGKC/IGLC2)*	Day42	−1.17
C1QL4	Day21	2.84	IL34	Day21/42	1.64/−1.49
C6	Day42	−5.9	IL4L1	Day42	−1.53
CASP1	Day2	−0.77	LAMP1	Day2	−0.47
CASP6	Day21	−3.18	LGALS1	Day4/42	−1.98/−0.8
CASP8	Day2	−0.65	LGALS3	Day21	−1.87
CD209	Day2	−0.41	LYG1	Day21	0.82
CD59	Day21	3.44	MYD88	Day2	0.51
CD8B	Day2	0.74	NLRC3*	Day2	1.50
CD93	Day21/42	−1.01/−1.21	NLRC3*	Day2	0.87
CFD	Day42	−7.31	NLRC3*	Day42	0.95
CXCL1	Day42	−4.96	NLRC3*	Day4	−7.68
CXCL12	Day42	−1.53	NLRC3*	Day4	−9.09
CXCL8	Day42	−5.1	NLRC3*	Day2	−0.78
DHX58/IFIH1	Day42	1.27	NLRC3*	Day4	−6.18
EEA1	Day2	0.45	NLRC3*	Day21	−1.94
EOMES	Day42	−1.57	NLRC3*	Day21	−2.49
ERAP1	Day2	−0.26	Piscidin	Day2	0.84
FADD	Day2	−0.24	PLG	Day2	−2.42
HLA-A/HLA-B*	Day2	1.70	PRF1*	Day4/21/42	2.26/3.09/−0.84
HLA-A/HLA-B*	Day2	0.71	PRF1*	Day42	−2.81
HLA-A/HLA-B*	Day2	−0.59	PSMB2	Day2	−0.3
HLA-A/HLA-B*	Day21	−1.69	PSMB5	Day2	0.55
IgL(IGKC/IGLC2)*	Day2	1.83	PSMB6	Day2	−0.43
IgL(IGKC/IGLC2)*	Day2	1.56	PSMB7	Day2	0.51
IgL(IGKC/IGLC2)*	Day2	1.34	PSMC6	Day2	−0.29
IgL(IGKC/IGLC2)*	Day2	1.27	PSMD11	Day2	−0.24
IgL(IGKC/IGLC2)*	Day2	1.16	PSMD7	Day2	−0.43
IgL(IGKC/IGLC2)*	Day2	1.12	PYCARD	Day2	−0.52
IgL(IGKC/IGLC2)*	Day2	1.10	SEC61A1/62	Day2	0.51
IgL(IGKC/IGLC2)*	Day2	1.01	TF	Day2/21/42	−2.23/2.94/3.17
IgL(IGKC/IGLC2)*	Day2	0.98	TNFSF10	Day42	−1.04
IgL(IGKC/IGLC2)*	Day2	0.79	TNFSF11	Day42	−2.36
IgL(IGKC/IGLC2)*	Day2	0.58	TRGC1	Day2	0.85
IgL(IGKC/IGLC2)*	Day2	0.50	UNC13D	Day2	1.91

Table 6
Contrast of previously performed qPCR analysis on selected immune genes (Mikkelsen and Seppola, 2013) with the current dataset derived from the *de novo* transcriptome edgeR pairwise differential expression analysis. The qPCR data is displayed as fold change (FC) calculated against house-keeping genes in contrast to calculated log fold change between vaccinated and control samples in the current dataset. IL8, IL12p40 and IFNg had several transcripts available in our *de novo* assembly. For IL12p40 and IFNg only one transcript was expressed above filtering thresholds. For IL8 four transcripts were expressed above filtering thresholds.

		Mikkelsen et al., 2013 CC-L2 family			Current dataset		
Gene	Trinity ID	FC day 2	FC day 4	Significant?	logFC day 2	logFC day 4	Significant?
IL1B	GG19241|c0_g1	0.5	0.5	No	−0.79	0.08	No
IL8	qPCR	0.7	0.8	No	NA	NA	NA
IL8	GG12684|c49_g1	NA	NA	NA	−0.09	−0.63	No
IL8	GG23237|c74_g2	NA	NA	NA	0.16	0.08	No
IL8	GG2424|c0_g1	NA	NA	NA	0.009	−0.57	No
IL8	GG4669|c0_g1	NA	NA	NA	0.009	−0.44	No
IL-10	GG30290|c0_g1	0.9	0.5	No	1.48	3.13	No
IL-12p40	GG20342|c1_g1	2.6	1.8	Yes, day 2	−0.01	−0.24	No
IFNg	GG15905|c1_g2	1.4	1.1	No	0.04	0.25	No
Cathelicidin	GG23028|c13_g2	0.6	0.4	No	0.57	−0.07	No
Hepcidin	GG19294|c68_g1	1.7	1.2	No	0.02	0.25	No

other cell lineages of the immune system such as monocytes, macrophages, natural-killer (NK) cells and innate lymphoid cells (ILCs) (Gardiner and Mills, 2016; Netea et al., 2016). Innate lymphoid cells (ILCs) are a relatively recent discovery of non-T and non-B cell populations and consist of three ILCs termed ILC1, ILC2 and ILC3 in addition to natural killer (NK) cells and lymphoid-tissue-inducer (LTi) cells in mammals. These cells are not well described in teleosts, but genetic markers for ILC2 and NK-cells have been described such as perforin, granzymes, Tbet (TBX21), EOMES and receptors of novel immune-type receptor (NITR) and leukocyte immunoglobulin-like receptor families (LITR). ILCs contribute to host defense, but do not have antigen-specific receptors. Markers commonly associated with T-cell subsets are often found in ILCs and thus they are viewed as the innate immune systems' version of T-cells (Vivier et al., 2016). In this regard, we observe some up-regulation of T- and NK-cell-like related markers, active down-regulation of neutrophil and monocyte recruitment and down-regulation of a single interleukin *IL34* (Table 4) indicating that a potential innate memory mechanism in Atlantic cod could be related to innate lymphoid cells/NK-like cells. However, we find no candidate cytokines. Most likely, this is related to the greater cytokine diversity in fish compared to mammals, making annotation of cytokines problematic (Zhu et al., 2013; Buchmann, 2014; Bird and Tafalla, 2015).

3.5. Subtle expression responses and large individual variations

Mucosal vaccination establishes a response locally at the mucosal membranes in fish in contrast to injection vaccinations where more systemic patterns appear in blood and blood-rich tissues (Rombout et al., 2014; Salinas, 2015). The lack of a clear systemic response in this experiment indicates that the mode of vaccination results in more local responses and is concordant with previous studies in the same/similar system showing small to no response in terms of systemic inflammation and regulation (Mikkelsen et al., 2011; Mikkelsen and Seppola, 2013). Furthermore, the time-points selected for sequencing might have missed the initial response to the vaccine leading to poor transcriptomic resolution of the vaccine response. Finally, the MDS clustering of samples included in this experiment did not resolve properly. This suggests that the vaccine response in Atlantic cod juveniles is subject to large individual variation even though we used family material in our study. Alternatively, the vaccine response produces very subtle changes in gene expression – regardless of tissues being sampled – that are too small to be detected by the statistical methods applied here.

3.6. Concluding remarks

We do not observe a clear non-conventional adaptive immune response to dip vaccination considering the loss of the MHCII pathway in Atlantic cod. However, we do find subtle changes in gene expression related to MHCI and T-cells. We thus explore the possibility of immunological memory established by CD8+ T-cells and then by a γδ TcR subset instead of the classical αβ T-cells. We do not find strong evidence of T-cell independent B-cell stimulation. However, there is some evidence indicating possible memory established by innate lymphoid cells/NK-like cells, but these cell types are still poorly defined in teleosts and needs further study. As the vaccination strategy applied here likely establishes a response mainly in the mucosal linings, future studies on mucosal vaccine responses in Atlantic cod should consists of both mucosal and blood rich tissue samples to disentangle systemic and local responses. Furthermore, as the responses observed here are subtle, and affected by large individual variation, any future experiments should also contain a re-challenge towards the pathogen post vaccination to amplify the transcriptional signals underlying establishment of immunity.

4. Methods

Please see GitHub repository for details: https://github.com/uio-cels/Solbakken_RNAseq.

4.1. Fish and experiment setup

The fish selected for this study were part of a larger vaccine investigation setup published in 2013 (Mikkelsen and Seppola, 2013) and were derived from family material maintained by the Norwegian national breeding program of Atlantic cod (www.nofima.no). We selected individuals that were siblings from a family deemed highly susceptible towards vibriosis in the 2013 vaccine investigation (Mikkelsen and Seppola, 2013). The fish were transported (at approx. 1.6 g) to the Aquaculture Research Station (Tromsø, Norway) for grow-out in seawater of 3.4% salinity at 10 °C, 24 h light and fed with commercial feed (BioMar, Norway). The rates of water inflow were adjusted to an oxygen saturation of 90–100% in the outlet water. See Mikkelsen et al. for further details (Mikkelsen and Seppola, 2013). The fish were reported to be healthy without any history of diseases and the experiment was approved by the National Animal Research authority in Norway. All methods were in accordance with the approved guidelines.

The original experiment had the following setup: An experimental dip vaccine produced by PHARMAQ AS (Norway) contained bacterin of *V. anguillarum* serotype O2b isolate 4299. Atlantic cod (approx. 2.5 g) were dip vaccinated by immersion for 30 s in diluted vaccine (1:10 in seawater), according to the manufacturer's instruction. Controls were mock vaccinated by dipping in sea water without vaccine. The fish, 10

vaccinated and 10 control groups, were distributed in 20 parallel, circular, centrally drained, fiberglass tanks (100*L*) with approx. 100–125 fish in each tank (density < 20 kg/dm^3). Fish for pre-challenge (n = 72) were left untreated and kept in a separate tank. One week prior to challenge the fish were anaesthetized with Metacainum (Norsk Medisinaldepot, Norway) (70 mg L^{-1}) and marked at the operculum Visible Implant Fluorescent Elastomer (Northwest Marine Technology Inc. US) before being distributed in 4 × 500 L tanks (2 tanks with coastal cod and North East Arctic cod families, respectively) with 80 fish from each family (40 vaccinated and 40 controls) in each tank (Table 1). Samples used for RNA was collected at 2 days, 4 days, 21 days and 42 days post vaccination. Six individuals were collected from each vaccinated and control group (n = 48). Head kidney and spleen were aseptically removed and transferred to RNA-Later (Ambion), and kept at 4 °C overnight before being stored at −80 °C.

4.2. RNA isolation, library preparation and sequencing

The Vibrio vaccinated treated and control sampled tissues were homogenized in 1× lysis buffer using MagNALyser Green Beads and the MagNaLyser Instrument (Roche Diagnostics). Total RNA was purified using an ABI Prism 6100 Nucleic Acid Prep Station (Applied Biosystems) with the recommended on-column DNase treatment.

All RNA isolates were quality controlled using Agilent 2100 Bioanalyzer (BioRad) before library preparation.

Libraries for RNAseq were prepared using the TruSeq™ RNA low-throughput (LT) protocol (Illumina). All samples were fragmented for 4 min to obtain the size distribution desired according to the TruSeq protocol. Sample and library overviews are presented in STable 1.

All libraries were sequenced 100 bp paired-end (PE) at the Norwegian Sequencing Centre on the Illumina HiSeq 2000 (www.sequencing.uio.no).

The obtained sequences were trimmed using Sickle with a 40 bp minimum remaining sequence length, a Sanger quality threshold of 20 and no 5′ end trimming (Joshi, n.d.). Remaining sequencing adapters were removed using Cutadaptv1.0 (Martin, 2012). Results were quality controlled using FastQCv0.9.2 to ensure improvement compared to raw data (Andrews, 2011).

Sequencing data is available through ENA PRJEB31369. Sequencing data derived from an *Francisella noatunensis* infection in Atlantic cod juveniles (Solbakken et al., 2019) used to supplement the transcriptome assembly is available through ENA PRJEB31396.

4.3. Methodological considerations for large scale RNAseq analysis on a non-model species

The Atlantic cod genome was first published in 2011 (Star et al., 2011) and consisted on only 454 sequencing using both shotgun and paired-end libraries to an average coverage of 40×. In 2017 Tørresen et al. released an improved version of the cod genome (gadMor2) adding BAC-ends, Illumina and PacBio sequencing data to improve genome resolution. The genome is a reconciled assembly where the best features of 4 subassemblies were combined. Furthermore, it also contains an unusual amount of tandem repeats (Tørresen et al., 2017). Our observation is that immune-related genes often are located to genomic regions with repeats or poor resolution and often without annotation. Thus, we opted to analyze the RNAseq data using both a reference-based (TopHat/Cufflinks) and a *de novo* (Trinity) strategy to better capture unassembled, partially or un-annotated (immune) genes providing the best possible resolution of the genetic background in Atlantic cod before performing differential expression analysis.

In terms of differential expression, the data have been analysed in a pairwise manner, with a general linearized model (GLM) time:-treatment as well as a custom script clustering genes according to five given expression profiles in a control-dependent manner. Details are given below.

4.4. Reference-genome based approach using Tuxedo

The second version of the Atlantic cod genome (Tørresen et al., 2017) was used as the genomic reference for a Tophat/Cufflinks pipeline according to the workflow described in (Trapnell et al., 2012). Mapping of samples towards the reference-genome GFF3 file was performed with Tophat v2.0.14 with default settings. Sample-specific transcriptomes were generated with Cufflinks v2.1.1. Cuffmerge was used to concatenate all the individual transcriptomes. Differential expression analysis was performed with Cuffdiff in a pair-wise manner between treated and control for each time-point. The output from Cuffdiff was further explored using CummeRbund v2.8.2 in R v3.1.3 for presentation purposes (Goff et al., 2013; R-Core-Team, 2015). An overview of the analysis is presented in the supplementary information (SFig. 6 and 7). Significantly differentially expressed genes are listed in Supplementary excel file 1. Read counts were also extracted from the Cuffdiff experiment and analysed using the custom script as described below. Significant genes from this analysis are listed in Supplementary excel file 2.

4.5. Reference-genome-guided approach using Trinity

Two RNAseq studies provided reads for the transcriptome assembly used here – the reads derived from the *Vibrio* vaccination described above and the reads derived from a *Francisella* challenge study with the same number of samples (Solbakken et al., 2019) (In total, the 96 libraries (48 from each experiment) provided on average 20.51 million trimmed read-pairs resulting in 1969.31 million reads in total.

We applied the Trinity transcriptome assembler v 2.0.6 using the genome-guided option with the second version of the Atlantic cod genome (Tørresen et al., 2017). The genome was indexed using Bowtie1 (v1.0.0) and then mapped using Tophat (v2.0.9) and sorted with Samtools (v0.1.19). The built-in normalization step of Trinity was applied reducing the trimmed read dataset to approximately 45 million read pairs (Grabherr et al., 2011; Haas et al., 2013). The following parameters were changed for the Trinity run: genome-guided, max intron 10,000, max memory 150 Gb, bflyHeapSpaceMax 10G, bflyCPU 12 and CPU 10.

The assembly was evaluated with the trinity_stats.pl and align_and_estimate_abundance.pl scripts – the latter with RSEM estimation method and bowtie aligner. The abundance estimation output was further used to filter the assembly on transcript level with FPKM = 1 using filter_fasta_by_rsem_values.pl. This resulted in 44,543 transcripts with an overall contig N50 of 2568 bp, median contig length of 1132 bp and a total of ~73.3 million assembled bases. Based on the longest open reading frames (ORFs) the transcript dataset was reduced to 32,934 “genes” with an overall contig N50 of 2490 bp and median contig length of 1014 bp. Assembly statistics are presented in STable 2.

Overall annotation was performed using Trinotate v2.0.1 following all mandatory steps with default parameters on the non-filtered assembly and transferred to the filtered assembly transcripts. The annotation of genes specifically discussed in this study have been verified through reciprocal BLAST by extracting the longest isoform of the gene in question and subjecting it to a BLASTX towards all UniProt entries using the UniProt BLAST tool with default settings (UniProt, 2015).

In addition, a list of specific immune genes was annotated manually (Supplementary table 3). Queries from human and at least 3 fish species was downloaded from UniProt and/or Genbank and aligned in MEGA7 (Kumar et al., 2016) to ensure sequence homology. The queries were used in a TBLASTN (Camacho et al., 2009) search towards the Trinity assembly using tabular output, default parameters and an e-value cutoff of 1e-1. All transcript hits were extracted and the longest isoform selected for a reciprocal BLASTX using the NCBI online BLAST tool. When needed, the extracted isoform was also aligned towards the queries in MEGA7 to evaluate hit accuracy. A list of conservatively annotated transcripts is presented in STable 3.

4.6. Sample mapping, read count extraction

The trimmed reads from all Vibrio-related samples were mapped against the filtered Trinity assembly using the built-in align_and_estimate_abundance script in Trinity with RSEM estimation method and bowtie aligner, before extracting raw counts using abundance_estimates_to_matrix.pl again with RSEM as the estimation method. For the Trinity-generated read-counts, the variance within the experiment was explored using the R-package VariancePartition (Hoffman and Schadt, 2016). Differential expression analysis was first analysed in a pair-wise manner using an edgeR script included in the Trinity program. Then, the same counts were subjected to a GLM-based analysis time:treatment using glmQLFit and glmQLFTest. Finally, the read counts were analysed using the custom script as described below. An overview of the analysis is presented in SFig. 1–5). All significantly differentially expressed genes are reported in Supplementary excel file 3 (default pairwise analysis), Supplementary excel file 4 (time:treatment GLM) and Supplementary excel file 5 (custom script).

4.7. Error distributions

Most RNAseq analysis packages assume that such data follows a negative binomial distribution of variability. We tested this assumption using a custom script testing the fit of the Poisson distribution, the negative binomial distribution and the zero-inflated negative binomial distribution (using the pscl package in R, script available in the GitHub repository). About 90% of all genes were classified as having negative binomial distribution and thus, in all cases, the negative binomial distribution was used for all down-stream analyses (data not shown, available in Github repository).

4.8. Custom script approach for gene expression pattern clustering

We wanted to further characterize the behavior of the dataset outside of what the most common “clustering by expression” RNAseq analysis packages could provide. In this way we could take into account our continuous control samples (contrary to only having a single time 0 control group) and define the expression patterns beforehand. This categorization was performed with a set of GLM regression models; no time dependency, linear time dependency, quadratic time dependency, factorial time dependency, pure treatment effect (no time dependency), treatment combined with linear time (interaction), treatment combined with quadratic time (interaction) and treatment combined with factorial time (interaction). Estimated regression coefficients were then used to determine in which category each gene's expression was to be assigned using the conservative BIC model selection criterion. In this manuscript the expression profiles are described as ‘increasing over time’, ‘internal maximum’, ‘decrease over time’, ‘internal minimum’ and ‘freestyle’. Only data from the control-dependent analyses are shown. (Note that if a quadratic effect was found but with minima/maxima outside the data material, it would be classified as either increasing or decreasing, depending on the estimated quadratic effect).

4.9. GO and gene network analyses

For all significant differentially expressed genes the (if available) corresponding annotation was translated to the “human” gene name in cases of clear orthology. In cases where orthology was questionable the automated annotation remained. In the *de novo* transcriptome assembly several genes had obtained multiple annotations. Here, the longest isoform was manually subjected to a BLASTX at the NCBI BLAST server with default settings and restricting the database to teleostei only. The multi-annotation was converted to a single annotation whenever possible. All significantly differentially expressed genes with corresponding annotation reported from all analyses above were analysed in Cytoscape (Shannon et al., 2003) using the plugin ClueGO (Bindea et al., 2009). ClueGO performs a GO-term enrichment (two-sided hypergeometric test for both enrichment and depletion) and was run selecting a full GO-term analysis with otherwise default parameters. The Bonferroni step-down *p*-value cutoff was set to $p = 0.05$ unless otherwise stated. GO-term fusion (related terms that share similar genes are fused) was applied to the larger networks to improve readability (specifically stated). In some cases, a GO-term analysis specifying only immune-related GO-terms was run (specifically stated). The layout of the networks was manually rearranged to improve readability. Size of the nodes corresponds to the corrected p-value. The thickness of the edges (lines) connecting the nodes represents the kappa score – the degree of connectivity between two nodes based on their overlapping genes.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cbd.2019.100597.

Declaration of Competing Interest

All authors have read and approved the manuscript. We declare no conflicts of interest. We also declare that all handling of animals was done according to regulations in Norway and that the experiment has been approved by local animal welfare authorities.

Acknowledgements

This work was supported by The Research Council of Norway (Grant numbers 199806/S40 and 222378/F20 to KSJ/SJ and grant number 199672/E40 to MS). The reference genome assembly was made using the Abel Cluster, owned by the University of Oslo and the Norwegian metacenter for High Performance Computing (NOTUR), and operated by the Department for Research Computing at USIT, the University of Oslo IT-department. http://www.hpc.uio.no/. The sequencing service was provided by the Norwegian Sequencing Centre (www.sequencing.uio.no), a national technology platform hosted by the University of Oslo.

Additional information

This manuscript has a GitHub repository providing all data: https://github.com/uio-cels/Solbakken_RNAseq. Sequencing data is available through ENA: PRJEB31369 (additional sequences included in the transcriptome PRJEB31396).

Competing financial interests

The authors declare no competing financial interests.

References

Andrews, S., 2011. The FastQC Project. Available from. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Bakkemo, K.R., et al., 2011. Intracellular localisation and innate immune responses following *Francisella noatunensis* infection of Atlantic cod (*Gadus morhua*) macrophages. Fish Shellfish Immunol. 31 (6), 993–1004.

Ballesteros, N.A., et al., 2012. Oral immunization of rainbow trout to infectious pancreatic necrosis virus (Ipnv) induces different immune gene expression profiles in head kidney and pyloric ceca. Fish Shellfish Immunol. 33 (2), 174–185.

Berg, A., et al., 2006. Time of vaccination influences development of adhesions, growth and spinal deformities in Atlantic salmon *Salmo salar*. Dis. Aquat. Org. 69 (2–3), 239–248.

Bindea, G., et al., 2009. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics 25 (8), 1091–1093.

Bird, S., Tafalla, C., 2015. Teleost chemokines and their receptors. Dev. Biol. (Basel) 4 (4), 756–784.

Buchmann, K., 2014. Evolution of innate immunity: clues from invertebrates via fish to mammals. Front. Immunol. 5, 459.

Caipang, C.M., et al., 2008. Intraperitoneal vaccination of Atlantic cod, *Gadus morhua* with heat-killed *Listonella anguillarum* enhances serum antibacterial activity and expression of immune response genes. Fish Shellfish Immunol. 24 (3), 314–322.

Caipang, C.M., Brinchmann, M.F., Kiron, V., 2009. Profiling gene expression in the spleen of Atlantic cod, *Gadus morhua* upon vaccination with *Vibrio anguillarum* antigen. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 153 (3), 261–267.

Camacho, C., et al., 2009. BLAST+: architecture and applications. BMC Bioinf. 10, 421.
Chettri, J.K., et al., 2015. Comparative evaluation of infection methods and environmental factors on challenge success: *Aeromonas salmonicida* infection in vaccinated rainbow trout. Fish Shellfish Immunol. 44 (2), 485–495.
Corripio-Miyar, Y., et al., 2007. Vaccination experiments in the gadoid haddock, *Melanogrammus aeglefinus* L., against the bacterial pathogen *Vibrio anguillarum*. Vet. Immunol. Immunopathol. 118 (1–2), 147–153.
Ellingsen, T., et al., 2011. *Francisella noatunensis* in Atlantic cod (*Gadus morhua* L.); waterborne transmission and immune responses. Fish Shellfish Immunol. 31 (2), 326–333.
Eschmeyer, W.N., Fricke, R., 2015. Catalog of Fishes: Genera, Species, References.
Eslamloo, K., et al., 2016. Transcriptome profiling of the antiviral immune response in Atlantic cod macrophages. Dev. Comp. Immunol. 63, 187–205.
Feng, C.Y., et al., 2009. Identification and analysis of differentially expressed genes in immune tissues of Atlantic cod stimulated with formalin-killed, atypical *Aeromonas salmonicida*. Physiol. Genomics 37 (3), 149–163.
Frans, I., et al., 2011. *Vibrio anguillarum* as a fish pathogen: virulence factors, diagnosis and prevention. J. Fish Dis. 34 (9), 643–661.
Fraser, T.W., et al., 2015. Vaccination and triploidy increase relative heart weight in farmed Atlantic salmon, *Salmo salar* L. J. Fish Dis. 38 (2), 151–160.
Gao, Y., et al., 2014. A live attenuated combination vaccine evokes effective immune-mediated protection against *Edwardsiella tarda* and *Vibrio anguillarum*. Vaccine 32 (45), 5937–5944.
Gardiner, C.M., Mills, K.H., 2016. The cells that mediate innate immune memory and their functional significance in inflammatory and infectious diseases. Semin. Immunol. 28 (4), 343–350.
Goff, L., Trapnell, C., Kelley, D., 2013. cummeRbund: Analysis, Exploration, Manipulation, and Visualization of Cufflinks High-Throughput Sequencing Data. https://bioconductor.org/packages/release/bioc/html/cummeRbund.html.
Grabherr, M.G., et al., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29 (7), 644–652.
Gudmundsdottir, S., et al., 2009. Specific and natural antibody response of cod juveniles vaccinated against *Vibrio anguillarum*. Fish Shellfish Immunol. 26 (4), 619–624.
Haas, B.J., et al., 2013. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. 8 (8), 1494–1512.
Haase, D., et al., 2016. Comparative transcriptomics of stickleback immune gene responses upon infection by two helminth parasites, *Diplostomum pseudospathaceum* and *Schistocephalus solidus*. Zoology (Jena) 119 (4), 307–313.
Hoare, R., et al., 2017. Efficacy of a polyvalent immersion vaccine against *Flavobacterium psychrophilum* and evaluation of immune response to vaccination in rainbow trout fry (*Onchorynchus mykiss* L.). Vet. Res. 48 (1), 43.
Hoffman, G.E., Schadt, E.E., 2016. variancePartition: interpreting drivers of variation in complex gene expression studies. BMC Bioinf. 17 (1), 483.
Hwang, J.Y., et al., 2017. RNA-Seq transcriptome analysis of the olive flounder (*Paralichthys olivaceus*) kidney response to vaccination with heat-inactivated viral hemorrhagic septicemia virus. Fish Shellfish Immunol. 62, 221–226.
Jameson, S.C., Masopust, D., 2009. Diversity in T cell memory: an embarrassment of riches. Immunity 31 (6), 859–871.
Jiang, J., et al., 2014. Differential transcriptomic response in the spleen and head kidney following vaccination and infection of Asian seabass with *Streptococcus iniae*. PLoS One 9 (7), e99128.
Jiang, Y., et al., 2016. Transcriptome signatures in common carp spleen in response to *Aeromonas hydrophila* infection. Fish Shellfish Immunol. 57, 41–48.
Joshi, N. Sickle - a Windowed Adaptive Trimming Tool for FASTQ Files Using Quality. https://github.com/ucdavis-bioinformatics/sickle.
Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33 (7), 1870–1874.
Lauvau, G., et al., 2016. Memory CD8(+) T cells: innate-like sensors and orchestrators of protection. Trends Immunol. 37 (6), 375–385.
Liu, X., et al., 2015. Profiling immune response in zebrafish intestine, skin, spleen and kidney bath-vaccinated with a live attenuated *Vibrio anguillarum* vaccine. Fish Shellfish Immunol. 45 (2), 342–345.
Lund, V., Bordal, S., Schrøder, M.B., 2007. Specificity and durability of antibody responses in Atlantic cod (*Gadus morhua* L.) immunised with *Vibrio anguillarum* O2b. Fish Shellfish Immunol. 23 (4), 906–910.
Lund, V., et al., 2008. Atypical furunculosis vaccines for Atlantic cod (*Gadus morhua*); vaccine efficacy and antibody responses. Vaccine 26 (52), 6791–6799.
Maekawa, S., et al., 2017. Transcriptome analysis of immune response against *Vibrio harveyi* infection in orange-spotted grouper (*Epinephelus coioides*). Fish Shellfish Immunol. 70, 628–637.
Magnadottir, B., et al., 2009. Natural antibodies of cod (*Gadus morhua* L.): specificity, activity and affinity. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 154 (3), 309–316.
Magor, B.G., 2015. Antibody affinity maturation in fishes-our current understanding. Dev. Biol. (Basel) 4 (3), 512–524.
Makesh, M., Sudheesh, P.S., Cain, K.D., 2015. Systemic and mucosal immune response of rainbow trout to immunization with an attenuated *Flavobacterium psychrophilum* vaccine strain by different routes. Fish Shellfish Immunol. 44 (1), 156–163.
Malmstrøm, M., et al., 2013. Unraveling the evolution of the Atlantic cod's (*Gadus morhua* L.) alternative immune strategy. PLoS One 8 (9), e74004.
Malmstrøm, M., et al., 2016. Evolution of the immune system influences speciation rates in teleost fishes. Nat. Genet. 48 (10), 1204–1210.
Martin, M., 2012. Cutadapt removes adapter sequences from high-throughput sequencing reads. Bioinform. Act. 17 (1), 10–12.
Mikkelsen, H., Seppola, M., 2013. Response to vaccination of Atlantic cod (*Gadus morhua* L.) progenies from families with different estimated family breeding values for vibriosis resistance. Fish Shellfish Immunol. 34 (1), 387–392.
Mikkelsen, H., et al., 2011. Vibriosis vaccines based on various sero-subgroups of *Vibrio anguillarum* O_2 induce specific protection in Atlantic cod (*Gadus morhua* L.) juveniles. Fish Shellfish Immunol. 30 (1), 330–339.
Neefjes, J., et al., 2011. Towards a systems understanding of MHC class I and MHC class II antigen presentation. Nat. Rev. Immunol. 11 (12), 823–836.
Netea, M.G., et al., 2016. Trained immunity: a program of innate immune memory in health and disease. Science 352 (6284), aaf1098.
Nymo, I.H., et al., 2016. Experimental challenge of Atlantic cod (*Gadus morhua*) with a *Brucella pinnipedialis* strain from hooded seal (*Cystophora cristata*). PLoS One 11 (7), e0159272.
Parra, D., Reyes-Lopez, F.E., Tort, L., 2015. Mucosal immunity and B cells in teleosts: effect of vaccination and stress. Front. Immunol. 6, 354.
Pridgeon, J.W., et al., 2012. Global gene expression in channel catfish after vaccination with an attenuated *Edwardsiella ictaluri*. Fish Shellfish Immunol. 32 (4), 524–533.
Raeber, M.E., et al., 2018. The role of cytokines in T-cell memory in health and disease. Immunol. Rev. 283 (1), 176–193.
R-Core-Team, 2015. R: A Language and Environment for Statistical Computing. Available from. http://www.R-project.org.
Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26 (1), 139–140.
Rombout, J.H., Yang, G., Kiron, V., 2014. Adaptive immune responses at mucosal surfaces of teleost fish. Fish Shellfish Immunol. 40 (2), 634–643.
Salinas, I., 2015. The mucosal immune system of teleost fish. Dev. Biol. (Basel) 4 (3), 525–539.
Sarropoulou, E., et al., 2012. Characterization of European sea bass transcripts by RNA SEQ after oral vaccine against *V. anguillarum*. Mar. Biotechnol. (N.Y.) 14 (5), 634–642.
Scapigliati, G., Fausto, A.M., Picchietti, S., 2018. Fish lymphocytes: an evolutionary equivalent of mammalian innate-like lymphocytes? Front. Immunol., vol. 9, 971.
Schrøder, M.B., et al., 2006. Early vaccination and protection of Atlantic cod (*Gadus morhua* L.) juveniles against classical vibriosis. Aquaculture 254 (1–4), 46–53.
Schrøder, M.B., et al., 2009. Comparison of antibody responses in Atlantic cod (*Gadus morhua* L.) to *Vibrio anguillarum*, *Aeromonas salmonicida* and *Francisella* sp. Fish Shellfish Immunol. 27 (2), 112–119.
Seppola, M., et al., 2015. Ultrapure LPS induces inflammatory and antibacterial responses attenuated *in vitro* by exogenous sera in Atlantic cod and Atlantic salmon. Fish Shellfish Immunol. 44 (1), 66–78.
Seternes, T., et al., 2007. Specific endocytosis and degradation of naked DNA in the endocardial cells of cod (*Gadus morhua* L.). J. Exp. Biol. (210), 2091–2103 Pt 12.
Shannon, P., et al., 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13 (11), 2498–2504.
Sheridan, B.S., et al., 2013. Gammadelta T cells exhibit multifunctional and protective memory in intestinal tissues. Immunity 39 (1), 184–195.
Solbakken, M.H., et al., 2016a. Evolutionary redesign of the Atlantic cod (*Gadus morhua* L.) toll-like receptor repertoire by gene losses and expansions. Sci. Rep. 6, 25211.
Solbakken, M.H., et al., 2016b. Successive losses of central immune genes characterize the Gadiformes' alternate immunity. Genome Biol. Evol. 8 (11), 3508–3515.
Solbakken, M.H., et al., 2017. Linking species habitat and past palaeoclimatic events to evolution of the teleost innate immune system. Proc. Biol. Sci. 284 (1853).
Solbakken, M.H., et al., 2019. Disentangling the immune response and host-pathogen interactions in *Francisella noatunensis* infected Atlantic cod. Comp. Biochem. Physiol. Part D Genomics Proteomics 30, 333–346.
Star, B., et al., 2011. The genome sequence of Atlantic cod reveals a unique immune system. Nature 477 (7363), 207–210.
Tang, X., et al., 2018. Recombinant NADP-dependent isocitrate dehydrogenase of *Edwardsiella tarda* induces both Th1 and Th2 type immune responses and evokes protective efficacy against edwardsiellosis. Vaccine 36 (17), 2337–2345.
Tørresen, O.K., et al., 2017. An improved genome assembly uncovers prolific tandem repeats in Atlantic cod. BMC Genomics 18 (1), 95.
Tørresen, O.K., et al., 2018. Genomic architecture of haddock (*Melanogrammus aeglefinus*) shows expansions of innate immune genes and short tandem repeats. BMC Genomics 19 (1), 240.
Trapani, J.A., Smyth, M.J., 2002. Functional significance of the perforin/granzyme cell death pathway. Nat. Rev. Immunol. 2 (10), 735–747.
Trapnell, C., et al., 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and cufflinks. Nat. Protoc. 7 (3), 562–578.
UniProt, C, 2015. UniProt: a hub for protein information. Nucleic Acids Res. 43 (Database issue), D204–D212.
Vestvik, N., et al., 2013. *Francisella noatunensis* subsp. *noatunensis* replicates within Atlantic cod (*Gadus morhua* L.) leucocytes and inhibits respiratory burst activity. Fish Shellfish Immunol. 35 (3), 725–733.
Vinuesa, C.G., Chang, P.P., 2013. Innate B cell helpers reveal novel types of antibody responses. Nat. Immunol. 14 (2), 119–126.
Vivier, E., et al., 2016. The evolution of innate lymphoid cells. Nat. Immunol. 17 (7), 790–794.
Wan, F., et al., 2016. Characterization of gammadelta T cells from zebrafish provides insights into their important role in adaptive humoral immunity. Front. Immunol. 7, 675.
Zehner, M., Burgdorf, S., 2015. Sec61 in antigen cross-presentation. Oncotarget 6 (24), 19954–19955.
Zhang, H., et al., 2013. Transcriptome profiling reveals Th17-like immune responses induced in zebrafish bath-vaccinated with a live attenuated *Vibrio anguillarum*. PLoS One 8 (9), e73871.
Zhang, D., et al., 2017. More than just antibodies: protective mechanisms of a mucosal vaccine against fish pathogen *Flavobacterium columnare*. Fish Shellfish Immunol. 71, 160–170.
Zhu, L.Y., et al., 2013. Advances in research of fish immune-relevant genes: a comparative overview of innate and adaptive immunity in teleosts. Dev. Comp. Immunol. 39 (1–2), 39–62.