Norwegian University of Life Sciences
Faculty of Biosciences
CIGENE

Philosophiae Doctor (PhD)
Thesis 2022:46

Characterization of genes and proteins related to chitin metabolism in Atlantic salmon

Karakterisering av gener og proteiner relatert til kitinmetabolisme i atlantisk laks

Matilde Mengkrog Holen

Characterization of genes and proteins related to chitin metabolism in Atlantic salmon

Karakterisering av gener og proteiner relatert til kitinmetabolisme i atlantisk laks

Philosophiae Doctor (PhD) Thesis

Matilde Mengkrog Holen

Norwegian University of Life Sciences
Faculty of Biosciences
CIGENE

Ås (2022)

Thesis number 2022:46
ISSN 1894-6402
ISBN 978-82-575-1918-6

Acknowledgments

The work presented in this thesis has been carried out at the Department of Biosciences at the Norwegian University of Life Sciences (NMBU).

Firstly, I will give a huge thanks to my supervisors and colleagues at NMBU for taking me under your wings and believing in me when I did not. A big thank you to my main supervisor, Matthew Kent, for always being available to work through methods, discuss results, and for asking the right questions. Another big thanks to my co-supervisors Simen Rød Sandve and Gustav Vaaje-Kolstad for great conversations and for including me in your research groups. I would not have managed this without you.

I would like to give a huge thanks to Yang Jin, Thomas Nelson Harvey and Jacob Torgersen for patiently teaching me how to dissect fish and cultivate salmon cells, and for your supervision and guidance on every step of RNA extraction, RNA-sequencing and RNA-seq data analysis. Thanks to Tina Rise Tuveng for helping me out with the methodology in the third manuscript, and for welcoming me to the PEP-group. A special thanks to Line Røsæg, the lab ladies, and everyone else at CIGENE that I haven't mentioned by name for creating a positive work environment and for making it fun to come to work.

Of course, I must also thank my family and friends for giving me the time and patience to complete this journey and for giving me the extra motivation needed to finish this.

I love you guys!

Hamar, May 2022

Matilde Mengkrog Holen

Table of contents

Summary

Chitin is an insoluble polysaccharide and important component in the extracellular matrix of invertebrates such as arthropods, algae, and fungi. In vertebrates, the role and significance of chitin have not been deeply explored, however, studies are beginning to provide evidence of the importance of chitin and enzymes that degrade chitin in diverse species. In humans, for example, the accumulation of chitin and chitin-degrading enzymes is associated with several different diseases, while in ray-finned fish, researchers have found evidence that the presence of chitin and chitin-degrading enzymes in the skin and gastrointestinal tract can help protect the fish from parasites. The enzymes that break down chitin are called chitinases. Before evidence emerged showing that ray-finned fish produce chitin, it was thought that chitinases mainly helped to degrade dietary chitin. With increasing evidence for the presence of chitin and genes encoding chitin-producing enzymes called chitin synthases, new questions about the role of the enzymes that synthesize and break down chitin have emerged.

We have used functional genomics and proteomics to characterize genes and enzymes related to the degradation and formation of chitin in Norway's most important aquacultural species, Atlantic salmon (*Salmo salar*). In addition to being economically important, the species has cultural significance, forms an important part of the ecosystem, and is a genetically fascinating species to study due to the whole-genome duplication that occurred in the ancestor of salmonids about 100 million years ago. Phylogenetic analysis showed that the whole-genome duplication event led to an increased number of genes encoding chitinases and chitin synthases. Comparative gene expression analyses showed that the chitinases are mainly expressed either in the stomach or in pyloric caeca and intestine, while the chitin synthases are exclusively expressed in pyloric caeca and intestine. The chitinases and chitin synthases expressed in the pyloric caeca of Atlantic salmon followed the same trend in gene expression; increasing as the intestine matures. While neither the expression levels nor protein activity of chitinases was affected by the addition of dietary chitin, we showed that chitinases isolated from the stomach can break down chitin to chitobiose at gastric-like conditions of Atlantic salmon. We therefore propose two different roles for the chitinases in Atlantic salmon; stomach chitinases mainly break down ingested chitin, and the chitinases and chitin synthases in the pyloric caeca and intestine break down and produce chitin in the intestinal mucosa.

Sammendrag

Kitin er et uløselig polysakkarid som er en viktig komponent i den ekstracellulære matriksen til flere virvelløse dyr som for eksempel leddyr, alger og sopp. Inntil nylig har det ikke vært forsket mye på hvilken rolle kitin spiller hos virveldyr, men flere resultater viser nå til at kitin og enzymer som bryter ned kitin også spiller en viktig rolle her. Hos mennesker er opphopning av kitin og kitinnedbrytende enzymer assosiert med flere ulike sykdommer, mens hos strålefinnefisk har forskere funnet bevis for at tilstedeværelse av kitin og kitinnedbrytende enzymer i hud og tarm kan være med på å beskytte fisken mot parasitter. Enzymene som bryter ned kitin kalles kitinaser, og før det fantes bevis for at strålefinnefisk også produserer kitin var det tenkt at kitinasene i hovedsak var med på å bryte ned kitin fra kosten. Med økt bevis for tilstedeværelse av kitin, samt gener som koder for kitin-produserende enzymer kalt kitin syntaser har nye spørsmål om rollen til enzymene som er med på å danne og bryte ned kitin meldt seg.

Vi har brukt funksjonell genomikk og proteomikk til å karakterisere gener og enzymer relatert til nedbrytning og dannelsen av kitin hos Norges viktigste art i oppdrettsnæringen, atlantisk laks (*Salmo salar)*. I tillegg til å være en viktig art både økonomisk, kulturelt og næringsmessig sett så er atlantisk laks en interessant art å studere grunnet helgenomduplikasjonen som skjedde hos stamfaren til laksefiskene for om lag 100 millioner år siden. Fylogenetiske analyser viste at denne helgenomduplikasjonen førte til et økt antall gener som koder for kitinaser og kitin syntaser. Videre viste komparative genuttrykksanalyser at kitinasene i hovedsak er uttrykt enten i magesekken eller i tarmen, mens kitin syntasene kun er uttrykt i tarmen. Kitinasene og kitin syntasene som er uttrykt i tarmen til atlantisk laks fulgte samme trend i genuttrykk; med økt genuttrykk ved modning av tarm. Mens verken genuttrykksnivåene eller proteinaktiviteten til kitinasene ble påvirket av tilføring av kitin i fôret viste vi at kitinaser isolert fra laksens magesekk kan bryte ned kitin og kitinholdige substrater til kitobiose ved forhold tilsvarende det som finnes i magesekken. Vi foreslår derfor to ulike roller for kitinasene i atlantisk laks hvor kitinasene i magesekken er tenkt å hovedsakelig bryte ned inntatt kitin, og kitinasene og kitin syntasene i tarmen er tenkt å henholdsvis bryte ned og produsere kitin i tarmoverflaten.

List of papers

Paper I

Holen MM, Kent MP, Vaaje-Kolstad G, Sandve SR. (2022). **Gene family expansion and functional diversification of chitinase and chitin synthase genes in Atlantic salmon (*Salmo salar*).** bioRxiv. doi: https://doi.org/10.1101/2022.05.05.490710 Manuscript.

Paper II

Holen MM, Sandve SR, Harvey TN, Jin Y, Angell IL, Rudi K, Kent MP. (2022). **The effect of dietary chitin on Atlantic salmon (*Salmo salar*) chitinase activity, gene expression, and microbial composition.** bioRxiv. doi: https://doi.org/10.1101/2022.05.05.490722. Manuscript.

Paper III

Holen MM, Tuveng TR, Kent MP, Vaaje-Kolstad G. (2022). **The gastric mucosa of Atlantic salmon (*Salmo salar*) is abundant in highly active chitinases.** Manuscript (will be available at bioRxiv after thesis submission).

1. Introduction

Chitin is one of the most abundant polysaccharides in nature. It is an insoluble and recalcitrant polysaccharide that consists of chains of β-1,4-linked *N*-acetyl-D-glucosamine (GlcNAc) residues and exists as a structural component in many invertebrates. It is found in fungi and algae cell walls, in the skeleton of sponges, in different body parts of mollusks and nematodes, and in the exoskeleton of arthropods such as insects and crustaceans where it works as a body armor (Figure 1).

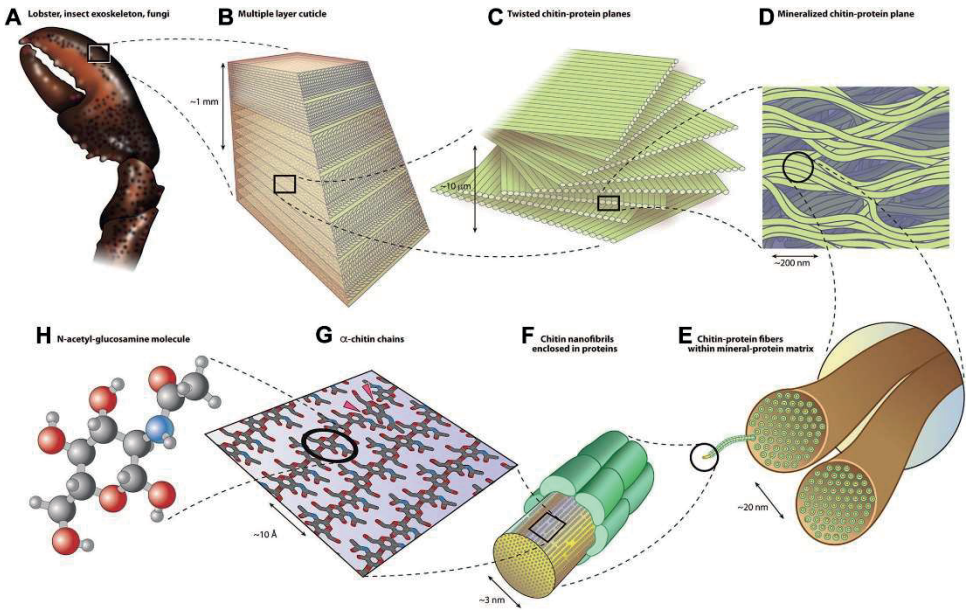

Figure 1. Schematic representation of chitin in exoskeleton structures (exemplified with the lobster exoskeleton; A). The chitin structures follow a hierarchal organization where multiple layers are organized in twisted chitin-protein planes consisting of mineralized chitin-protein fibers with linear polymeric chains of *N*-acetylglucosamine wrapped in protein (B-H). Linear polymeric chains of *N*-acetylglucosamine can either be internally cleaved by endochitinases or by exochitinases at free non-reducing ends (G; arrows). The figure was obtained from Van Dyken, S. J., & Locksley, R. M. (2018) [1], with permission from Elsevier.

Chitin makes up to 40% of the total body mass in shrimp and crab [2–4], and in some insects, chitin is also present in inner-body structures such as the peritrophic matrix (PM) which surrounds and protects the gut epithelium from abrasive particles and food-borne pathogens [5,6]. In nature, chitin polymers assemble into insoluble crystals with one of three different crystal forms; α-chitin, β-chitin, or γ-chitin, where the

neighboring chitin chains have antiparallel, parallel, or a mixture of parallel and antiparallel orientation respectively [7]. Exoskeletons of arthropods generally have α-chitin chains [8], while structures such as the squid pen, tube worms and possibly the PM of insects contain β-chitin [9]. The parallel orientation of the β-chitin chains results in weaker intermolecular bonds compared to α- and γ-chitin, making the β-chitin more flexible and susceptible to enzymatic degradation [10].

Despite its abundance in many plant and animal species, there has been a long-standing belief that chitin is not synthesized endogenously in vertebrates; however recent evidence now challenges this assertion [11]. Chitinous structures have been found in the forelimbs of axolotl larvae (*Ambystoma mexicanum*) [11], in the gut of zebrafish *(Danio rerio*) [11,12], Mozambique tilapia (*Oreochromis mossambicus*) [12], and rainbow trout (*Oncorhynchus mykiss*) [12], the blenny cuticle of *Paralipophrys trigloides* [13], the Ampullae of Lorenzini of Chondrichthyes [14] and in the scales of parrotfish (*Chlorurus sordidus*) [15], red snapper *(Lutjanus argentimaculatus*) [15], common carp (*Cyprinus carpio*) [16] and Atlantic salmon (*Salmo salar*) [11]. The chitinous gut structures in ray-finned fish are hypothesized to function as an immune barrier that has evolved from the PM found in insects (Figure 2) [12]. This structure most likely maintains gut homeostasis with endogenous microbiota and works as a first-line defense in the innate and adaptive immune system. Therefore, it is reasonable to suggest that chitin has variable functional roles in diverse species and is important for vertebrate physiology.

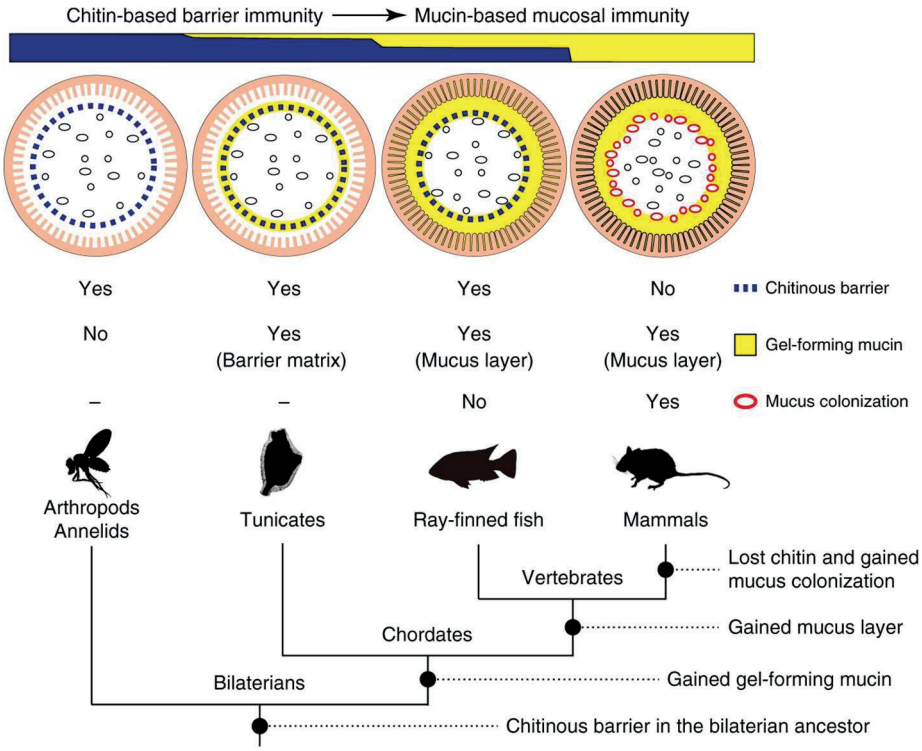

Figure 2. The hypothesized evolution of a mucin-based gut surface in mammals from a chitin-based membrane (peritrophic matrix, PM) in some insects such as arthropods and annelids. The PM is semi-permeable and allows nutrients to enter, but not luminal microbes (black ovals). Tunicates, invertebrates that are the closest relative to vertebrates, have a chitinous barrier surrounded by gel-forming mucins (yellow circle) which keep food microbes away from the epithelium. A similar barrier is found in ray-finned fish where the mucosal membrane consists of gel-forming mucin which is separated from the luminal space by a chitinous barrier making it harder for unfavorable bacteria to colonize in the mucosal membrane. The chitinous barrier is lost in mammals and bacteria colonize (red ovals) in the layer of gel-forming mucin surrounding the gut epithelium. The figure was obtained from Nakashima, K., Kimura, S., Ogawa, Y. et al. (2018) [12] according to the terms and conditions of the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/).

Chitin synthase (CHS), a member of the glycosyltransferase family 2 (BRENDA:EC 2.4.1.16), is highly conserved in chitin-containing organisms where it is responsible for endogenous chitin formation. CHS is a transmembrane protein that utilizes an activated molecule of uridine diphosphate *N*-acetylglucosamine (UDP-GlcNAc) to synthesize chitin by adding UDP-GlcNAc to the 4-hydroxyl group of the non-reducing sugar of the chitin chain that protrudes into the extracellular space [17,18]. The chitin chains are further modified and assembled into a chitin-matrix in the cuticular matrix or peritrophic matrices by the work of various hub proteins in arthropods [19]. Most

CHS contain two conserved sequence motifs in the catalytic domain, *EDR* and *QRRRW* which are predicted to be crucial for catalysis [20,21]. In addition, three motifs have shown to be conserved in the cataltytic domain of insect CHS, *CATMWHXT*, *QXFEY* and *WGTRE* [22,23]. While one CHS gene is found in amphibians, it seems that ray-finned fish have at least two CHS genes [11]. Knockout studies of CHS genes in zebrafish have shown to result in lower chitin-binding domain (CBD) staining signals and are thus hypothesized to be responsible for chitinous structures in ray-finned fish [11,12].

Although chitin synthase genes have not yet been identified in many vertebrate species, most vertebrates do possess chitinase (BRENDA:EC3.2.1.14) and chitobiase (BRENDA:EC3.2.1.52) activity [24]. Chitinase breaks down chitin polymers by the cleavage of internal β-1,4 glycosidic bonds to (mainly) dimers and trimers of *N*-acetylglucosamine, and chitobiase completes the degradation of dimers to monomers by releasing *N*-acetylglucosamine from non-reducing ends of the substrate. These enzymes belong to family 18 glycoside hydrolases (GH18), and one characteristic of GH18 enzymes is the catalytic domain with the conserved motif, *DXXDXDXE*, where glutamic acid (E) is the catalytic residue [25]. In vertebrate GH18 chitinases, the catalytic domain is often followed by a serine/threonine-rich linker region connected to one or multiple carbohydrate-binding modules (classified as CBM14 in the CAZy database) at the carboxyl-terminal region of the protein [26,27]. This domain has shown to be important for binding and facilitation of insoluble chitin degradation [28,29].

Mammals express three GH18 enzymes with demonstrated catalytic activity: two chitinases where the catalytic motif is intact; acidic mammalian chitinase (CHIA) [30] and chitotriosidase (CHIT) [31], and one chitobiase known to degrade asparagine-linked oligosaccharides on glycoproteins in lysosomes [32]. Mammals also express multiple chitinase-like lectins (chi-lectin; CHIL) with truncated GH18 domains that have rendered the enzymes catalytically inactive [33,34]. All mammalian GH18 chitinases (CHIA, CHIT, and CHILs) have been shown to modulate various immune responses in both mice and humans [35–42]. Environmental chitin polymers have for example shown to accumulate in the lungs of mice and humans deficient of CHIA (AMCase) causing lung disease [43] and in humans, an inactive splice-variant of CHIT has shown increased susceptibility to infections with chitin-containing pathogens causing filariasis [44]. Moreover, CHIA has shown to play a potential role in the digestion of chitinous organisms, ingested as feed or pathogens, as it is highly expressed in the stomach of a range of mammals [45–57]. The relative CHIA expression is lower in cattle (herbivore) and dogs (carnivore) compared to mice, pigs, and chickens (omnivores) which indicates that dietary habits affect chitinase expression [53].

Because of their pathophysiological relevance, mammalian chitinases have been extensively studied [37,40,42,58–60], but the role of chitinases in other vertebrates is less understood. It is known that fish express GH18 chitinases and chitobiases with strong sequence identity to mammalian orthologs [61], and chitinase activity in fish has been linked to a potential role in digestion because various fish prey on chitinous organisms [62]. Most fish chitinases examined to this date are expressed in the gut, and chitinases isolated from the stomach of greenling (*Hexagrammos otakii*), marbled rockfish (*Sebasticus marmoratus*), silver croaker (*Pennahia argentatus*), threeline grunt (*Parapristipoma trilineatum*), chub mackerel (*Scomber japonicus*), Japanese sardine (*Sardinops melanostictus*) and coelacanth (*Latimeria chalumnae)* have the potential to break down chitinous substrates from various sources naturally found in the diet [63]. These enzymes work best under gastric-like conditions which further suggests that the fish chitinases play a role in digestion.

Besides playing a role in digestion, other studies point to the possible roles of fish chitinases in defense. Zebrafish is a common model system for studies on pathogenesis and immune responses to infection in fish [64]. Infection studies have shown that zebrafish respond to *Salmonella* and *Mycobacterium marinu* infection with increased chitinase gene expression levels [65,66], and that chitinase 3 (*Chia.3*) is able to directly inhibit the fungal growth of *Canadia albicans* [67]. Turbot (*Scophthalmus maximus*), a major aquaculture species in southern Europe, is one of many fishes being susceptible to infection with *Vibrio anguillarum* and *Streptococcus iniae*. The turbot intestine most likely act as a portal of entry for these pathogens, and infection with either has been shown to induce the expression of turbot chitinases (*chit1-3*) [68]. These chitinases have also been shown to inhibit the fungal growth of the chitinous fungal pathogen *Mucor mucedo* [69], supporting a possible role in mucosal immunity where they prevent pathogen attachment. Other evidence for chitinases playing a role in immunity is provided by the orange-spotted grouper (*Epinephelus coioides*) where expression of two chitinase genes, chitinase1 (*gchi1*) and chitinase2 (*gchi2*), was observed to increase a few hours after injection of LPS (lipopolysaccharides), a strong activator of the immune response [70]. In addition, recombinant gChi1 had a positive effect on growth rate when used as an additive in feed containing 2% shrimp shell chitin. This was probably due to increased gene expression levels of various growth-regulating hormones and immune response genes.

In this thesis, we explore chitin genomics, metabolism, and biochemistry within Atlantic salmon (*Salmo salar*). This iconic species is of special interest for several reasons. Firstly, Atlantic salmon has a fascinating lifestyle, being an anadromous fish

that is born in freshwater but spends most of its time in seawater before it returns to freshwater to spawn and is thus dependent on genomic and physiological plasticity. Second, Atlantic salmon has undergone a relatively recent whole-genome duplication (WGD) event 100 – 80 million years ago (mya) [71] making it an interesting model species for understanding the fate of duplicated genes and gene products. Lastly, Atlantic salmon is one of the most important fish in the aquaculture industry where salmon aquaculture was worth approximately 8.5 billion GBP annually in 2017 [72]. To improve the fish health and sustainability of salmon aquaculture it is essential to understand the genes and gene products that have central roles in the basic biology of salmon. The high-quality reference genome assembly of Atlantic salmon in 2016 [73] has contributed to a large number of excellent genomic resources that have made it possible to improve our current knowledge of these genes. We used the 2016 version (ICSASG_v2) of the genome assembly in this thesis but wish to note that this has recently been superseded by a newer version (Ssal_v3.1) built with the help of long-read data.

Atlantic salmon express multiple chitinase- and chitin synthase-like (hereafter referred to as chitinase and CHS respectively) genes, but they have not yet been characterized. The genes and gene products have potential functions in at least three important aspects of salmon biology: i) maintenance of a chitin-based mucosal barrier (Figure 2), ii) digestion of chitin-containing organisms which is of importance if the diet is substituted with chitin-containing organisms such as krill, insects, yeast and algae, and iii) immune defense against infectious pathogens such as the salmon louse (*Lepeophtheirus salmonis*). Salmon louse is a chitin-containing parasite that has become a major challenge for farmed and wild salmonids, and increased chitinase expression has been associated with higher salmon louse resistance in pink salmon (*Oncorhynchus gorbuscha*) [74,75] and host response to louse attachment in Atlantic salmon [76].

1.1 Aims and objectives

Given the importance of chitin and chitin metabolism in many vertebrate species and the growing evidence for a role in ray-finned fish, we sought to characterize the chitinase- and chitin synthase genes and proteins to gain an understanding of their potential roles in Atlantic salmon biology. The work is presented in the following sections with the following objectives:

1. Determine the phylogenetic relationship between GH18 family proteins and CHS proteins in Atlantic salmon and related species and investigate the spatio-temporal regulation of these genes (paper I).

2. Investigate the influence of dietary chitin on the expression of genes and the presence of bacteria related to chitin metabolism in Atlantic salmon (paper II).

3. Isolate and characterize the *in vitro* behavior of chitin-binding proteins from the Atlantic salmon stomach (paper III).

2. Materials and Methods

To better characterize the genes and proteins involved with chitin metabolism in Atlantic salmon, a spectrum of functional genomic and proteomic methodologies was employed. Functional genomics aims to understand how genes or non-coding intergenic regions contribute to and/or regulate a biological process. Investigations of this type often involve assessing the expression of gene products under a given condition in a specific cell or tissue and relating this to chromatin structure and other epigenetic influences. To complement the information obtained from our functional genomics study, we made use of metagenomic and proteomics data. In metagenomics, collections of genomes; usually microbial communities in a sample of interest, are studied, while in proteomics, proteins from a biological system of interest are quantified and characterized. An overview of the main methods employed in the different papers is shown in Figure 3. The main concepts of these methods are explained in the following subsections.

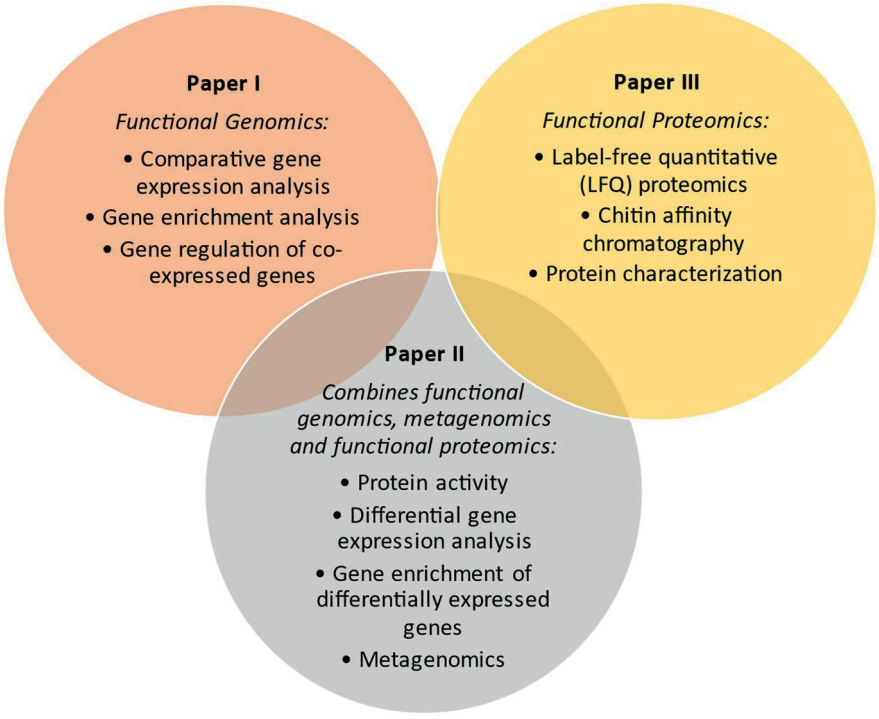

Figure 3. An overview of the main methods and approaches utilized in papers I, II, and III.

2.1 Functional genomics

The Atlantic salmon genome sequence released in 2016 by Lien et al [73] represents a solid foundation for application of functional genomics methods to analyze the gene expression and understand genome regulatory mechanisms. For some years now, Next Generation Sequencing (NGS) based technologies have been the method of choice for large-scale genomic and transcriptomic sequencing that is essential for conducting functional genomics research. Using a variety of sample preparation methods, NGS provides data that makes it possible to study how any gene or region in the genome is regulated in different cells, tissues, or environments, and characterize the composition of microbial DNA in biological samples. In this study, methods of NGS such as RNA (Ribonucleic acid) sequencing (RNA-seq), Assay for Transposase Accessible Chromatin with high-throughput sequencing (ATAC-seq), and 16S ribosomal RNA gene sequencing were employed to:

1) measure gene expression levels of genes related to chitin metabolism in Atlantic salmon,
2) classify genes that have correlated expression levels and could thus be regulated by the same transcription factors,
3) identify biological processes associated to particular genes through functional enrichment tests,
4) identify putative transcription factor binding motifs in promoters which could explain differences in gene regulation,
5) identify genes showing differential expression in response to the addition of dietary chitin,
6) investigate how bacterial populations in the intestine change as a response to dietary chitin.

2.1.1 Gene expression analysis using RNA-seq

RNA-sequencing makes it possible to describe qualitatively and quantitively the transcriptome of a cell or tissue at a specific time point. In my thesis project, total RNA was extracted from the stomach and pyloric caeca of Atlantic salmon using a RNeasy Mini Kit (Qiagen) to study the transcriptome in these tissues. See Figure 4 for a visual representation of the different segments of the gastrointestinal tract.

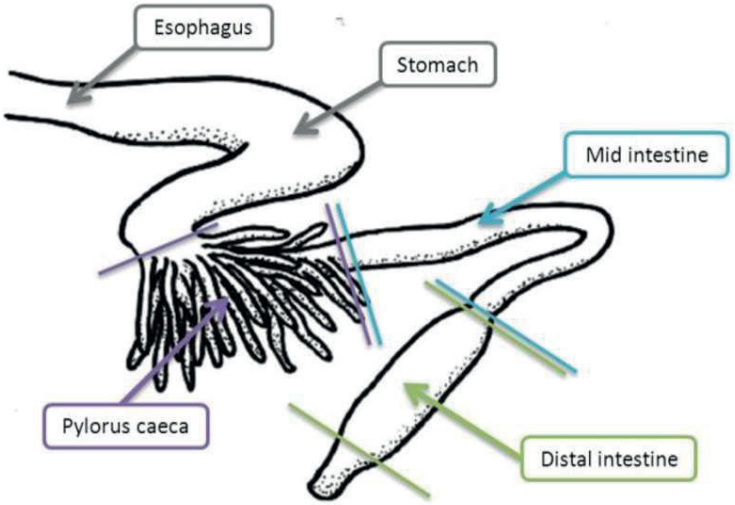

Figure 4. Illustration of the gastrointestinal tract of Atlantic salmon. The gastrointestinal tract consists of an esophagus followed by the cardiac stomach, the pyloric caeca (named pylorus caeca in this figure) with its characteristic finger-like structures, the mid intestine (midgut), and the distal intestine characterized by a larger diameter. The figure was obtained from Johansson (2014) [77].

Sequencing was done using Illumina short-read technology. The steps in standard messenger RNA (mRNA) library construction and RNA-sequencing to assess gene expression levels are as follows [78]: mRNA species are enriched and randomly fragmented to the desired size. The fragments are reverse transcribed to blunt, double-stranded complementary DNA (cDNA) before sequencing adaptors are ligated to each end of the fragments. The library is introduced to a flow cell where each fragments adapter hybridizes into an immobilized single-stranded DNA capture probe before a bridge polymerase chain reaction (bridge-PCR) creates multiple copies of identical sequences in a tight cluster. A sequencing primer, polymerase, and fluorescently labeled ddNTPs (dideoxynucleoside triphosphates) are introduced allowing for single-base extension. A laser activates the fluorescently labeled ddNTPs and the fluorescent signal is detected with a camera. The termination chemistry on the ddNTP base is reversed allowing for a new single base to be added and the single-base extension process to be repeated. The labeled ddNTPs emit different colors making it possible for a computer to translate the different color signals to a specific order of bases, and finally, a sequence for each cluster (which originates from a single unique cDNA molecule) can be deducted [79].

After sequencing, quality control is conducted before the raw sequences are aligned to a reference genome, in this case, the Atlantic salmon genome (ICSASG_v2). The number of sequences (reads) that are aligned to a specific gene is counted and the counts are normalized. The normalization involves normalizing for differences in read number between samples and normalizing for reads per gene relative to transcript length. This normalization is necessary because the total amount of sequenced mRNA can vary between samples and because longer genes tend to get more reads even though the gene expression level is the same as shorter genes [80]. Gene expression levels may be expressed as transcripts per million (TPM) where the read counts first are divided by the (longest) transcript length in kilobases (kb) and then adjusted to the sum of all reads divided by 1,000,000. To be able to make statistical comparisons and further analyses, the TPM values are log-transformed.

Once the RNA-sequencing data is generated, there are a lot of different ways to analyze the data. In my thesis project, we compared the gene expression levels of genes predicted to be involved in chitin metabolism. This could be done with gene-specific methods such as quantitative PCR but, with a more comprehensive dataset in hand, we were able to make comparisons to how the rest of the genome was transcribed. In paper I, we performed an enrichment of genes that behaved similar to our genes of interest, and in paper II, we explored differential expression analysis (DEA) to get a better understanding of how Atlantic salmon respond to the addition of dietary chitin. I will explain the basic principles behind these analyses in the following sections.

Genes with correlated gene expression levels over a group of samples are more likely to be involved in connected biological processes controlled by the same regulators [81]. We used a biweight midcorrelation approach to find genes with correlated expression levels. Biweight midcorrelation gives a pairwise measure of how "similar" the gene expression profiles are. It is often used in co-expression analysis instead of the standard Pearson correlation because it is median-based rather than mean-based and thus less sensitive to outliers [82].

Another way to analyze genes in connected biological processes is to observe how they respond to a certain condition of interest. DEA was carried out to find which genes were expressed at different levels as a response to dietary chitin, given a significance threshold (false discovery rate (FDR) < 0.05). It is essential to normalize counts before a DEA analysis, and we used EdgeR's trimmed mean of M-values (TMM) normalization, where the counts are normalized with a scaling factor based on the

assumption that most genes are not differentially expressed between conditions [83]. To find differentially expressed genes (DEG), a dispersion parameter is estimated in the EdgeR pipeline before the test for differential expression is conducted [84]. The dispersion parameter represents the variation in gene expression levels between biological replicates and this is calculated using the quantile-adjusted conditional maximum likelihood method (qCML). After the dispersion parameter is estimated the differentially expressed genes are found using an exact test.

Once a group of genes has been identified as having correlated expression profiles or being differentially expressed in response to a given condition, it is possible to conduct enrichment analysis to relate the genes to biological processes. Here we tested for gene function enrichment in two different ways; 1) using KEGG (Kyoto Encyclopedia of Genes and Genomes) enrichment analysis [85], and 2) using GO (Gene Ontology) enrichment analysis [86]. With KEGG enrichment analysis, a set of genes are mapped to manually curated biochemical pathways with enzymes and molecules of known function. If a higher number of genes in a gene-set is mapped to a certain pathway than what is predicted by pure chance, we say that the genes are enriched for that pathway. GO, on the other hand, does not have genes/molecules/enzymes linked to pathways, but has sets of genes linked to specific biological functions called GO terms. There are three different types of terms; terms that describe the molecular function of the gene (e.g. catalytic activity, oxidoreductase activity), terms that describe the biological process of the gene (e.g. glucosamine biosynthetic process, oxidative phosphorylation), and terms that describe the cellular component of the gene (e.g. extracellular matrix, mitochondrial matrix). By analyzing genes with correlated expression profiles in the context of biochemical pathways (KEGG) and, independently, biochemical function (GO) it is possible to reveal associations that might not have been detected with either approach alone [87]. For example, in some situations, the obtained gene set includes a significantly large collection of genes known to work in a particular pathway, while in other situations the gene set contains a collection of genes not known to work in a particular pathway even though they are somewhat related to a given phenotype as a group. Therefore, combining the two strategies gives more information than using only one of them.

2.1.2 Measuring chromatin accessibility with ATAC-seq to find putative regulators of expression

ATAC-seq is a technique used in functional genomics to measure the chromatin accessibility in the genome [88]. Briefly, chromatin consists of repeating units of nucleosomes where DNA is wrapped around histone proteins, and areas of DNA that are accessible are more likely to be transcriptionally active. Transcriptional regulation is dependent on accessible regions close to the transcription start site allowing transcription factors (TFs) to bind and ATAC-seq is therefore used to predict areas where transcriptional regulation occurs. ATAC-seq exploits the properties of a hyperactive transposase (Tn5) which makes random cuts in open-chromatin regions of DNA and simultaneously ligates sequencing adapters resulting in a sequence-ready library.

To identify open chromatin based on ATAC-seq a software algorithm searches for local pileups (higher signals) of ATAC-seq reads in a genomic region. If the read depth is high enough, it is also possible to identify and quantify transcription factor binding as this will result in a local drop in ATAC-seq signals (due to less transposase accessibility) at the predicted site of transcription factor binding [89]. We used ATAC-seq data to search for open chromatin regions near the transcription start site and scanned these regions for transcription factor binding motifs, testing for enrichment of specific motifs. Putative transcription factor binding motifs was provided by SalMotifDB, a database for analyzing transcription factor binding sites in salmonid genomes [90].

2.1.3 Metagenomic evaluation of bacterial populations with 16S ribosomal RNA gene sequencing

16S ribosomal RNA (rRNA) sequencing is commonly used to identify the taxonomic distribution of bacteria in an environmental or biological sample. Using this approach, the highly conserved bacterial 16S rRNA gene regions are amplified with PCR using bacteria-specific primers before adapters are added and the library is sequenced. The resulting reads are compared against a 16S reference database and assigned to a phylogenetic rank. Reads of similar 16S gene sequences are clustered in the same OTUs (operational taxonomic units) and different OTUs usually represent distinct bacteria taxa at the genus level. Shotgun metagenomic sequencing is another method that makes it possible to sequence all DNA in a sample. This technology is often used

if a strain-level resolution is wanted and makes it possible to carry out metabolic profiling [91]. In our case, the samples obtained from the distal intestine of Atlantic salmon contained a very high proportion of host-DNA that would "eat up" reads from exogenous DNA if we were to carry out shotgun metagenomic sequencing and the amount of exogenous DNA after depletion of host DNA was not sufficient for sequencing. Moreover, for some species of fish it has been reported that cultivable microorganisms in the gastrointestinal tract only represent < 0.1% of the total microbial community [92]. For this reason, we decided that 16S rRNA gene sequencing of rRNA obtained from the distal intestine of Atlantic salmon was sufficient to say something about how the bacterial populations in the distal intestine differed because of dietary chitin.

2.2 Functional proteomics

Developing a more complete picture of biological events occurring in a cell or tissue requires not only functional genomics data but also an understanding of the proteome. To elucidate the biological properties of the proteins involved in chitin metabolism we complemented our genomics data with functional proteomics data, with a focus on characterizing the chitinases hypothesized to break down chitin polymers. Liquid chromatography with tandem mass spectrometry (LC-MS/MS) was employed to perform relative quantification of stomach chitinases in Atlantic salmon, the stomach chitinases were isolated with chitin affinity chromatography and the enzyme activity was characterized using both fluorogenic substrate analogs and natural chitin substrates. The methods are briefly explained in the subsequent sections.

2.2.1 Label-free quantitative (LFQ) proteomics

Label-free quantitative proteomics is a mass spectrometry-based method to identify and quantify proteins in a biological sample of interest [93]. This way of quantifying proteins is often used instead of isotope-labeled protein quantification because it does not require labeling of proteins, it is cheaper, and the sample preparation is more efficient. In brief, the proteins of interest are digested with a protease, e.g. trypsin, and the resulting peptides are separated with liquid chromatography (LC) and detected with tandem mass spectrometry (MS/MS) [94]. In tandem mass spectrometry two mass analyzers are coupled together. The total mass-to-charge

ratio (m/z) of the precursor peptide ions are first measured (MS1) before peptide ions with specific m/z-ratios are selected, fragmented, and re-measured (MS2) to obtain sequence information. This allows for more precise and accurate detection of peptides. The resulting peptide spectra are matched against a database consisting of theoretical spectra made *in silico* using protein sequences of interest, in this case, the proteome of Atlantic salmon (UP000087266), and the peptides are quantified based on the area under the curve of MS1 peptide ions. The resulting dataset of protein intensities is typically cleared for common contaminants and proteins only present in one of the biological replicates before the intensities are log-transformed for analysis.

2.2.2 Chitin affinity chromatography

Affinity chromatography using a chitin matrix was used to capture chitin-binding proteins in the stomach of Atlantic salmon. A crude solution of proteins is added to a column consisting of a chitin-matrix in a buffer solution that facilitates interaction between the chitin-binding domain of soluble chitinase proteins and the immobilized chitin-matrix. After washing to remove unbound proteins, the captured proteins are eluted from the column by changing the pH and/or ionic strength of the buffer.

2.2.3 Characterization of chitinase activity

In this study, chitinase activity was measured using two different approaches. The first of these uses 4-methylumbelliferyl (4MU) chitooligosaccharides which are soluble, synthetic analogs of chitin with a fluorescent group (4MU) that is released upon cleavage by chitinase. The fluorescence of liberated 4MU can be measured with a fluorescence spectrophotometer and the intensities are used as a measure of chitinase activity. This method requires low enzyme concentration to detect a signal (high sensitivity) and the analysis takes only a couple of minutes (usually 15-30 minutes of incubation time). However, the activity measurements from this approach do not always correspond with how the chitinase enzymes work on more biologically relevant chitin substrates [95]. An alternative approach is therefore to use insoluble chitin substrates found in nature. Chitin occurs mainly as α-chitin or β-chitin in nature [7], where α-chitin is the most stable crystalline form because of the strong intermolecular hydrogen bonding resulting from the antiparallel orientation of the chitin-chains. The complexity of natural chitin substrates makes activity measurements more cumbersome than those obtained when using short oligomeric

analogs for three reasons. Firstly, natural chitin is insoluble in most organic solvents and is usually embedded in protein-mineral complexes; this makes it difficult to correctly quantify the amount of chitin used in the experiment. Secondly, degradation products obtained from native chitin can be of different lengths and a flexible system for detection is needed. We used high-performance liquid chromatography to separate the oligosaccharides, disaccharides, and monosaccharides produced by chitinase and quantified the products based on the area under the curve. Finally, activity measurements with natural substrates are less sensitive than those obtained with fluorogenic chitooligosaccharides and require a higher concentration of enzyme or longer incubation times. For these reasons, we decided to assess activity independently using both native and analog substrates to obtain the most complete picture of activity.

3. Results in brief

3.1 Paper I

The first step in characterizing the genes and gene products related to chitin metabolism in Atlantic salmon was to identify the genes and study their evolution and expression patterns in different tissues of Atlantic salmon.

3.1.1 Atlantic salmon chitin metabolism genes are mainly expressed in the gastrointestinal tract

Atlantic salmon encodes multiple genes related to chitin metabolism. These genes are homologous to GH18 chitinases and when comparing against related fish species as well as human and mouse, we found that the salmon GH18 proteins were located in six major clades (Paper I, Figure 1B). Three chitinase proteins, Chia.3+4+7 shared the strongest sequence similarity to the acidic mammalian chitinase and the genes were exclusively expressed in the stomach, while another group, Chia.1+2+6+8+9+10, fell into two separate fish-specific clades showing lower sequence similarity to the acidic mammalian chitinase and the genes were mainly expressed in pyloric caeca and mid intestine (Paper I, Figure 2). The fourth clade contained the salmon chitinase protein named Chia.5 with a predicted catalytically inactive active site and the gene was lowly expressed. The last two clades contained ortholog proteins of chitobiase (CTBS) and stabilin-1 interacting protein (CHID1) and the genes were mainly expressed in the gastrointestinal tract (Paper I, Figure 2). Duplicated chitinase genes from the salmonid-specific whole-genome duplication event (Ss4R), including four chitinase genes (*chia.1+2* and *chia.9+10*) and two genes with conserved insect CHS domains (*chs1a* and *chs1b*), were mainly expressed in pyloric caeca and mid intestine (Paper I, Figure 2 & 3).

3.1.2 Distinct regulation of chitin metabolism genes indicates functional diversification

Five of the genes with pyloric caeca-specific expression (*chia.1+2+6+8* and *chs1b)* showed highly correlated gene expression profiles when the gene expression levels during the first ~100 days of ingesting external feed were compared (Paper I, Figure 4). These and other co-expressed genes were enriched for a Forkhead Box J2 (FOXJ2)

transcription factor binding motif in the gene promoter, whereas Caudal Type Homeobox 1 (CDX1) and Caudal Type Homeobox 2 (CDX2) transcription factor binding motifs were present in the gene promoters of most pyloric caeca-specific chitinases and CHS (*chia.1+2+6+9+10* and *chs1a+chs1b*). This indicates that Atlantic salmon has distinct regulation of chitin remodeling proteins; one group is exclusively expressed in the stomach with an unknown regulator, and two groups are mainly expressed in pyloric caeca, where CDX1, CDX2 and FOXJ2 are candidate regulators responsible for the distinct gene expression of gene duplicates.

3.2 Paper II

Given the strong bias in gastrointestinal expression of Atlantic salmon chitinases, we tested the effect of dietary chitin on GH18 and CHS gene expression and microbial composition.

3.2.1 Dietary chitin slightly affected the expression of chitin metabolism genes and microbiota

In paper I, we hypothesized that the chitinases with stomach-specific expression were mainly involved in the degradation of chitin-containing organisms, while the pyloric caeca- and mid-intestinal-specific chitinases were involved in the maintenance of a chitin-containing intestinal mucosa. To further assess this, we conducted a feeding trial where portions of the diet were substituted with cellulose (control), chitin extracted from shrimp shells (*Pandalus borealis*), or black soldier fly larvae (*Hermetia illucens*) meal and examined the chitinase gene expression- and activity. Neither chitinase- and CHS gene expression levels nor chitinase activity were significantly affected by the presence of dietary chitin (Paper II, Figure 1 & 2), although we observed a trend of a slightly lower chitinase activity of fish fed the black soldier fly larvae diet. This could be explained by a slight decrease in gene expression levels of three chitinases expressed in pyloric caeca (*chia.1+2+6*) possibly linked to the high abundance of potential chitin-degrading bacteria such as *Actinomyces* and *Bacillus* in the black soldier fly larvae diet and intestine of Atlantic salmon fed this diet (Paper II, Figure 4).

3.2.2 Dietary chitin caused downregulation of hexosamine pathway genes

The gene expression levels of the stomach-specific chitinases did not change when dietary chitin levels increased, and the most abundant transcripts in the stomach independent of diets originated from chitinase genes (Paper II, Figure 2A). However, the stomach transcriptome responded to the diet with shrimp shell, and gene enrichment of differentially expressed genes showed an upregulation of genes involved in the extracellular matrix formation and downregulation of key genes in the hexosamine pathway leading to the formation of UDP-GlcNAc, the substrate used by CHS for chitin synthesis (Paper II, Figure 3).

3.3 Paper III

In paper I and II we mainly examined the gene expression of genes related to chitin metabolism in Atlantic salmon and the next natural step was to characterize the proteins responsible for chitinase activity.

3.3.1 Chitinase proteins are highly abundant in the stomach mucosa of Atlantic salmon

In paper II we showed that the stomach chitinases were highly expressed independently of diets, and in paper III we observed that there is a high correlation between the abundance of proteins in stomach mucosa and its gene expression levels (Paper III, Figure 2). The stomach chitinases are among the most abundant proteins in the stomach mucosa of Atlantic salmon and are thus likely to play an important role.

3.3.2 Atlantic salmon stomach chitinases efficiently degrade chitin

To assess the functional properties of the Atlantic salmon stomach chitinases, Chia.3+4 from the stomach of Atlantic salmon was isolated (Paper III, Figure 3A). This mix of enzymes efficiently degrades various chitin substrates to chitobiose ($GlcNAc_2$) under gastric conditions. The enzymes are active at pH 2-6 for up to 24 hours at 14 °C and show a maximum activity at pH 2-3 when using fluorogenic substrates and pH 5

when using β-chitin as substrate (Paper III, Figure 3, 4 & 5). The activity on fluorogenic substrates is higher at seawater-like salinity (0.6 M NaCl) compared to the activity without the addition of NaCl. This indicates that the stomach chitinases can degrade chitin under stomach-like conditions in both fresh- and seawater.

The key findings and general conclusions from the results obtained in this thesis are presented in table 1.

Table 1. Summary of the key findings in paper I-III.

Paper	**Key findings**	**General conclusions**
I	• Atlantic salmon express multiple GH18 family and chitin synthase genes. • The gene families have increased in size due to multiple whole-genome duplication events. • One group of chitinases are mainly expressed in the stomach • Chitinases and chitin synthases expressed in pyloric caeca and mid intestine are hypothesized to be regulated by the same transcription factors.	Supports the idea of a chitin-based mucosal barrier in Atlantic salmon.
II	• Endogenous chitin does not affect the gene expression of GH18 family genes or chitin synthase.	
III	• Stomach chitinases are among the most abundant proteins in gastric mucosa. • Stomach chitinases (Chia.3+4) efficiently degrade chitin from various sources to chitobiose. • Stomach chitinases (Chia.3+4) are active at pH 2-6 over 24 hours.	Supports the idea that stomach chitinases in Atlantic salmon play a role in the degradation of ingested chitin-containing organisms (and possibly endogenous chitin).

4. Discussion

4.1 Increased gene family size and functional divergence of chitinase and CHS genes

In paper I, we show that Atlantic salmon encodes a large repertoire of proteins involved in chitin metabolism likely as a result of multiple whole-genome duplication (WGD) events. A very early vertebrate gene- or genome duplication event has been hypothesized to have resulted in functional specialization of the two active mammalian family 18 chitinase proteins, CHIA and CHIT [96]. Thereafter, the teleost-specific WGD (Ts3R) which occurred approximately 320 million years ago (Mya) [97–99] gave rise to a new group of fish chitinases evolving from CHIA and has previously been termed CHIO [61]. Finally, Atlantic salmon has undergone a more recent genome duplication event, the salmonid-specific whole-genome duplication event (Ss4R), approximately 80 Mya [71,100,101]. Half of the gene duplicates (55%) from the Ss4R are still predicted to be functional [73]. This includes the duplicated CHIO chitinases and CHSs in Atlantic salmon which are mainly expressed in pyloric caeca and mid intestine. The observed differences in duplicate gene expression levels after first feeding were linked to regulatory motif divergence and further imply functional diversification of these genes.

In addition to expressing chitinase- and CHS genes in pyloric caeca and mid intestine, Atlantic salmon also express two other groups of GH18 genes: one group being mainly expressed in the stomach and consists of three predicted-to-be active chitinases with strong sequence identity to the mammalian CHIA proteins, and one chitobiase (CTBS)-like gene being expressed in stomach, pyloric caeca, and mid intestine. This indicates that Atlantic salmon has the potential to degrade chitin into digestible monomers in the main digestive sections of the salmon gastrointestinal tract where nutrient uptake happens [102].

To investigate if the stomach-specific chitinases of Atlantic salmon had different enzymatic properties than the chitinases expressed in pyloric caeca, we tried to isolate chitinases from both stomach and pyloric caeca. Unfortunately, using affinity chromatography we were only able to isolate a mixture of two stomach chitinases and nothing from pyloric caeca. In paper III, we characterized the isolated enzymes and showed that they were able to degrade purified α-chitin, β-chitin, and chitin in pulverized crab shells, shrimp shells, and skin of black soldier fly pupae into chitobiose ($GlcNAc_2$) in conditions naturally found in the stomach of Atlantic salmon.

The stomach chitinases were among the most abundant proteins secreted into the stomach lumen together with proteases and proteins associated with immune system defenses and we hypothesize that the salmon stomach chitinases are involved in digestion and the first-line defense of feed-borne pathogens, like what has been proposed for the mammalian CHIA proteins. The reason why the other chitinases did not bind to the chitin column used for chitinase isolation needs further investigation, but it might be that the chitinases have evolved different substrate specificities or binding capacities as the unbound chitinases show less sequence similarity to human AMCase (CHIA) compared to the bound stomach chitinases. We know that the unbound chitinases efficiently hydrolyze fluorogenic chitin substrates as there was chitinase activity in the unbound protein solution after chitin column purification, but these substrates are very different from natural, crystalline chitin substrates, and activity on these substrates does not provide evidence that the unbound proteins are active on crystalline chitin. More effort is therefore needed to detect the natural chitin substrates for Atlantic salmon chitinases and a good place to start is to better characterize the chitin structures in the suggested chitin-based mucosal barrier.

4.2 Chitin metabolism in Atlantic salmon

In paper II, we found no association between the activity of chitin metabolism genes in Atlantic salmon and the amount of dietary chitin indicating that expression of chitin metabolism genes is not directly regulated by short-term exposure to exogenous chitin. Although the relatively low number of replicates (n = 3-5) could lead to low power to detect smaller shifts in gene expression, a study in Rainbow trout (*Oncorhynchus mykiss*) [103] reach the same conclusion as our study. These results, combined with the intriguing finding that Atlantic salmon and Rainbow trout seem to have low ability to digest chitin [104,105], begs the question: *why are the chitinases among the most abundant transcripts in the stomach and pyloric caeca?*

As mentioned in the introduction, Nakashima et al. (2018), proposed a chitin-based barrier in the gastrointestinal tract of ray-finned fish [12]. The chitin-based barrier is hypothesized to have evolved from invertebrate ancestors and resemble the PM found in the cardia and/or midgut of most insects. The insect PM is a semi-permeable, extracellular layer which is more flexible than for example the exoskeleton of insects and consists mainly of chitin fibers, glycoproteins, proteoglycans, and proteins [106]. It acts both to protect against abrasive food particles and pathogens, and to increase the efficiency of digestion [107]. The mechanism by which the PM is formed involves the

coordinated activity of three major regulators of chitin synthesis; i) glutamine-fructose-6-phosphate aminotransferase (GFPT), ii) UDP-*N*-acetylglucosamine pyrophosphorylase (UAP) and iii) CHS [108], all being key enzymes where fructose-6-phosphate is converted to UDP-GlcNAc, the substate used by CHS for chitin synthesis. There are currently only a few studies pointing to the presence of chitin in the gut of ray-finned fish, but the following results presented in this thesis adds support to the idea of a chitin-based barrier in Atlantic salmon:

- In paper I, we show that CHS and chitinases are expressed in the same intestinal tissues of Atlantic salmon and the predicted CHS proteins have identical domains as CHS isolated from insects [23].
- An *uap*-like gene and genes involved in mucin production were found to have highly correlated expression patterns across development, including a similar increase in expression at the time of transition from endo- to exogenous feeding when the pyloric caeca and intestine of Atlantic salmon is known to mature [102].
- The co-expressed genes were enriched for a forkhead box (FOXJ2) transcription factor binding motif known to be important for intestinal maturation [109].
- The chitinase- and CHS genes seem to be constitutively expressed at relatively high levels in the gastrointestinal tract of adult fish and is not affected by short term exposure of exogenous chitin.

Figure 5 shows a summary of the proposed mechanisms of chitin metabolism in Atlantic salmon. As information about what happens to chitinase degradations products in the PM of insects is currently lacking [110], it is difficult to make hypotheses about related mechanisms in Atlantic salmon.

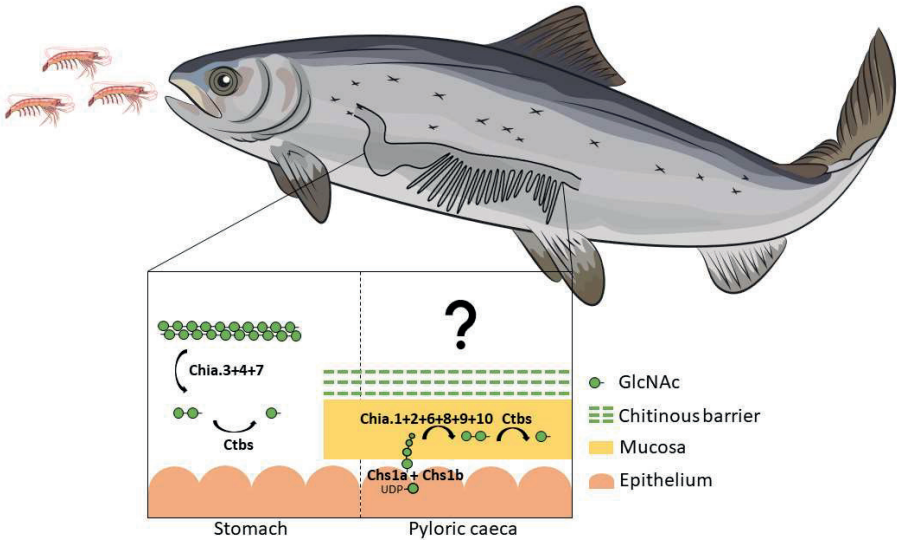

Figure 5. Proposed mechanism of chitin metabolism in Atlantic salmon. The stomach chitinases (Chia.3+4+7) break down ingested chitin to GlcNAc dimers (two green circles) which can further be broken down to monomers (one green circle) by chitobiase (Ctbs). Putative chitin synthases (Chs1a + Chs1b) in the pyloric epithelium synthesize chitin chains from UDP-GlcNAc which are incorporated into a chitinous mucosal barrier, while chitinases (Chia.1+2+6+8+9+10) and chitobiase (Ctbs) expressed in pyloric caeca degrade chitin structures surrounding the intestinal mucosa.

4.3 Future perspectives

The idea of a chitin-based mucosal barrier in ray-finned fish such as Atlantic salmon was proposed by Nakashima et al. (2018) during my work on this thesis. Before this, Tang et al. (2015) were the first (to our knowledge) to address the presence of endogenous chitin in ray-finned fish where they linked chitin biosynthesis in the gut of zebrafish to the activity of chitin synthase as gene knockdown of chitin synthase resulted in lower chitin staining signals [11]. This means that there is still a lot to uncover about the underlying mechanisms involved in chitin metabolism in ray-finned fish, and the work presented here provides a groundwork for further research on this. Confirming the presence of chitin and characterizing the proposed chitin structures in the intestine of Atlantic salmon and other ray-finned fish is among what remains to be done to prove that this barrier exists in ray-finned fish. This could be done using chitin-specific antibodies as have previously been used to detect chitin in the central nervous tissue of patients with Alzheimer's disease [111]. Successful production of chitin-specific antibodies makes it possible to apply ELISA (enzyme-

linked immunosorbent assay) to detect and quantify polysaccharides such as chitin at picogram levels [112]. Combining this with histology and immunohistochemistry methods could add useful information about how the chitin-based mucosal barrier is developed and interacts with for example exogenous microorganisms.

In paper III, we provided evidence that a mixture of two Atlantic salmon stomach chitinases can degrade various chitin substrates, but since we experienced difficulties with obtaining pure chitinase solutions directly from the host and through recombinant expression (results not shown) further efforts should be made to isolate the different chitinases and chitin synthases to better characterize their structure and mode of action. Furthermore, gene silencing, knockdown and/or protein inhibition of the candidate genes and gene products in this thesis would be a natural next step to analyze what impact they have on chitin synthesis, degradation, and gut integrity in general. When the genes, proteins, and substrates involved in the proposed chitin-based barrier have been characterized, comparative genomics and proteomics are necessary to be able to address unanswered questions such as the evolution of a chitin-based mucosal barrier in ray-finned fish.

5. Conclusions

By combining comparative and functional genomics with proteomics, we have provided novel insight into potential genes and proteins related to chitin metabolism in Atlantic salmon. This is a solid starting point to further assess the mechanisms behind the proposed chitin-based barrier in the salmon intestinal mucosa and ray-finned fish in general.

6. References

1. Van Dyken, S. J. & Locksley, R. M. Chitins and chitinase activity in airway diseases. *J. Allergy Clin. Immunol.* **142**, 364–369 (2018).

2. Charoenvuttitham, P., Shi, J. & Mittal, G. S. Chitin Extraction from Black Tiger Shrimp (*Penaeus monodon*) Waste using Organic Acids. *Sep. Sci. Technol.* **41**, 1135–1153 (2006).

3. Kurita, K. Chitin and chitosan: functional biopolymers from marine crustaceans. *Mar. Biotechnol.* **8**, 203–226 (2006).

4. Yan, N. & Chen, X. Sustainability: Don't waste seafood waste. *Nature* **524**, 155–157 (2015).

5. Kelkenberg, M., Odman-Naresh, J., Muthukrishnan, S. & Merzendorfer, H. Chitin is a necessary component to maintain the barrier function of the peritrophic matrix in the insect midgut. *Insect Biochem. Mol. Biol.* **56**, 21–28 (2015).

6. Kato, N. *et al.* Regulatory mechanisms of chitin biosynthesis and roles of chitin in peritrophic matrix formation in the midgut of adult Aedes aegypti. *Insect Biochem. Mol. Biol.* **36**, 1–9 (2006).

7. Jang, M.-K., Kong, B.-G., Jeong, Y.-I., Lee, C. H. & Nah, J.-W. Physicochemical characterization of α-chitin, β-chitin, and γ-chitin separated from natural resources. *J. Polym. Sci. A Polym. Chem.* **42**, 3423–3432 (2004).

8. Minke, R. & Blackwell, J. The structure of α-chitin. *J. Mol. Biol.* **120**, 167–181 (1978).

9. Zhao, X., Zhang, J. & Zhu, K. Y. Chito-Protein Matrices in Arthropod Exoskeletons and Peritrophic Matrices. in *Extracellular Sugar-Based Biopolymers Matrices* (eds. Cohen, E. & Merzendorfer, H.) vol. 12 3–56 (Springer International Publishing, 2019).

10. Shigemasa, Y., Saito, K., Sashiwa, H. & Saimoto, H. Enzymatic degradation of chitins and partially deacetylated chitins. *Int. J. Biol. Macromol.* **16**, 43–49 (1994).

11. Tang, W. J., Fernandez, J., Sohn, J. J. & Amemiya, C. T. Chitin is endogenously produced in vertebrates. *Curr. Biol.* **25**, 897–900 (2015).

12. Nakashima, K. *et al.* Chitin-based barrier immunity and its loss predated mucus-colonization by indigenous gut microbiota. *Nat. Commun.* **9**, 3402 (2018).

13. Wagner, G. P., Lo, J., Laine, R. & Almeder, M. Chitin in the epidermal cuticle of a vertebrate (Paralipophrys trigloides, Blenniidae, Teleostei). *Experientia* **49**, 317–319 (1993).

14. Phillips, M. *et al.* Chitin is present within the Ampullae of Lorenzini of chondrichthyan fishes. *BioRxiv* (2019) doi:10.1101/687301.

15. I. F. M. Rumengan, P. Suptijah, S. Wullur & A. Talumepa. Characterization of chitin extracted from fish scales of marine fish species purchased from local markets in North Sulawesi, Indonesia. *IOP Conference Series: Earth and Environmental Science* **89**, 012028 (2017).

16. G Zaku, S., A Emmanuel, S., C Aguzue, O. & A Thomas, S. *Extraction and characterization of chitin; a functional biopolymer obtained from scales of common carp fish (Cyprinus carpio l.): A lesser known source*. vol. 5 (2011).

17. Glaser, L. & Brown, D. H. The synthesis of chitin in cell-free extracts of Neurospora crassa. *J. Biol. Chem.* **228**, 729–742 (1957).

18. Merzendorfer, H. Chitin synthesis inhibitors: old molecules and new developments. *Insect Sci.* **20**, 121–138 (2013).

19. Behr, M. & Ganesan, K. Improving Polysaccharide-Based Chitin/Chitosan-Aerogel Materials by Learning from Genetics and Molecular Biology. *Materials (Basel)* **15**, (2022).

20. Sinnott, M. L. Catalytic mechanism of enzymic glycosyl transfer. *Chem. Rev.* **90**, 1171–1202 (1990).

21. Breton, C., Mucha, J. & Jeanneau, C. Structural and functional features of glycosyltransferases. *Biochimie* **83**, 713–718 (2001).

22. Zhu, Y.-C. *et al.* Sequence of a cDNA and expression of the gene encoding a putative epidermal chitin synthase of Manduca sexta. *Insect Biochem. Mol. Biol.* **32**, 1497–1506 (2002).

23. Merzendorfer, H. & Zimoch, L. Chitin metabolism in insects: structure, function and regulation of chitin synthases and chitinases. *J. Exp. Biol.* **206**, 4393–4412 (2003).

24. Jeuniaux, C. & Cornelius, C. Distribution and activity of chitinolytic enzymes in the digestive tract of birds and mammals. (1997).

25. Terwisscha van Scheltinga, A. C., Hennig, M. & Dijkstra, B. W. The 1.8 A resolution structure of hevamine, a plant chitinase/lysozyme, and analysis of the conserved sequence and structure motifs of glycosyl hydrolase family 18. *J. Mol. Biol.* **262**, 243–257 (1996).

26. Funkhouser, J. D. & Aronson, N. N. Chitinase family GH18: evolutionary insights from the genomic history of a diverse protein family. *BMC Evol. Biol.* **7**, 96 (2007).

27. CAZy - CBM14. http://www.cazy.org/CBM14.html.

28. Tjoelker, L. W. *et al.* Structural and functional definition of the human chitinase chitin-binding domain. *J. Biol. Chem.* **275**, 514–520 (2000).

29. Renkema, G. H. *et al.* Synthesis, sorting, and processing into distinct isoforms of human macrophage chitotriosidase. *Eur. J. Biochem.* **244**, 279–285 (1997).

30. Boot, R. G. *et al.* Identification of a novel acidic mammalian chitinase distinct from chitotriosidase. *J. Biol. Chem.* **276**, 6770–6778 (2001).

31. Renkema, G. H., Boot, R. G., Muijsers, A. O., Donker-Koopman, W. E. & Aerts, J. M. Purification and characterization of human chitotriosidase, a novel member of the chitinase family of proteins. *J. Biol. Chem.* **270**, 2198–2202 (1995).

32. Kuranda, M. J. & Aronson, N. N. A di-N-acetylchitobiase activity is involved in the lysosomal catabolism of asparagine-linked glycoproteins in rat liver. *J. Biol. Chem.* **261**, 5803–5809 (1986).

33. Renkema, G. H. *et al.* Chitotriosidase, a chitinase, and the 39-kDa human cartilage glycoprotein, a chitin-binding lectin, are homologues of family 18 glycosyl hydrolases secreted by human macrophages. *Eur. J. Biochem.* **251**, 504–509 (1998).

34. Houston, D. R., Recklies, A. D., Krupa, J. C. & van Aalten, D. M. F. Structure and ligand-induced conformational change of the 39-kDa glycoprotein from human articular chondrocytes. *J. Biol. Chem.* **278**, 30206–30212 (2003).

35. Kanneganti, M., Kamba, A. & Mizoguchi, E. Role of chitotriosidase (chitinase 1) under normal and disease conditions. *J. Epithel. Biol. Pharmacol.* **5**, 1–9 (2012).

36. Hollak, C. E., van Weely, S., van Oers, M. H. & Aerts, J. M. Marked elevation of plasma chitotriosidase activity. A novel hallmark of Gaucher disease. *J. Clin. Invest.* **93**, 1288–1292 (1994).

37. Bucolo, C., Musumeci, M., Musumeci, S. & Drago, F. Acidic Mammalian chitinase and the eye: implications for ocular inflammatory diseases. *Front. Pharmacol.* **2**, 43 (2011).

38. Guan, S.-P., Mok, Y.-K., Koo, K.-N., Chu, K.-L. & Wong, W. S. F. Chitinases: biomarkers for human diseases. *Protein Pept. Lett.* **16**, 490–498 (2009).

39. Malaguarnera, L. *et al.* Chitotriosidase gene expression in Kupffer cells from patients with non-alcoholic fatty liver disease. *Gut* **55**, 1313–1320 (2006).

40. Shuhui, L., Mok, Y.-K. & Wong, W. S. F. Role of mammalian chitinases in asthma. *Int. Arch. Allergy Immunol.* **149**, 369–377 (2009).

41. Zhu, Z. *et al.* Acidic mammalian chitinase in asthmatic Th2 inflammation and IL-13 pathway activation. *Science* **304**, 1678–1682 (2004).

42. Kawada, M., Hachiya, Y., Arihiro, A. & Mizoguchi, E. Role of mammalian chitinases in inflammatory conditions. *Keio J. Med.* **56**, 21–27 (2007).

43. Van Dyken, S. J. *et al.* Spontaneous Chitin Accumulation in Airways and Age-Related Fibrotic Lung Disease. *Cell* **169**, 497-509.e13 (2017).

44. Choi, E. H. *et al.* Genetic polymorphisms in molecules of innate immunity and susceptibility to infection with Wuchereria bancrofti in South India. *Genes Immun.* **2**, 248–253 (2001).

45. Strobel, S., Roswag, A., Becker, N. I., Trenczek, T. E. & Encarnação, J. A. Insectivorous bats digest chitin in the stomach using acidic mammalian chitinase. *PLoS ONE* **8**, e72770 (2013).

46. Ohno, M. *et al.* Quantification of Chitinase mRNA Levels in Human and Mouse Tissues by Real-Time PCR: Species-Specific Expression of Acidic Mammalian Chitinase in Stomach Tissues. *PLoS ONE* **8**, e67399 (2013).

47. Tabata, E. *et al.* Protease resistance of porcine acidic mammalian chitinase under gastrointestinal conditions implies that chitin-containing organisms can be sustainable dietary resources. *Sci. Rep.* **7**, 12963 (2017).

48. Ohno, M. *et al.* Acidic mammalian chitinase is a proteases-resistant glycosidase in mouse digestive system. *Sci. Rep.* **6**, 37756 (2016).

49. Tabata, E. *et al.* Gastric and intestinal proteases resistance of chicken acidic chitinase nominates chitin-containing organisms for alternative whole edible diets for poultry. *Sci. Rep.* **7**, 6662 (2017).

50. Ohno, M., Tsuda, K., Sakaguchi, M., Sugahara, Y. & Oyama, F. Chitinase mRNA levels by quantitative PCR using the single standard DNA: acidic mammalian chitinase is a major transcript in the mouse stomach. *PLoS ONE* **7**, e50381 (2012).

51. Ma, J.-E. *et al.* Acidic mammalian chitinase gene is highly expressed in the special oxyntic glands of Manis javanica. *FEBS Open Bio* **8**, 1247–1255 (2018).

52. Suzuki, M. *et al.* Cellular expression of gut chitinase mRNA in the gastrointestinal tract of mice and chickens. *J. Histochem. Cytochem.* **50**, 1081–1089 (2002).

53. Tabata, E. *et al.* Chitin digestibility is dependent on feeding behaviors, which determine acidic chitinase mRNA levels in mammalian and poultry stomachs. *Sci. Rep.* **8**, 1461 (2018).

54. Goto, M., Fujimoto, W., Nio, J., Iwanaga, T. & Kawasaki, T. Immunohistochemical demonstration of acidic mammalian chitinase in the mouse salivary gland and gastric mucosa. *Arch. Oral Biol.* **48**, 701–707 (2003).

55. Tabata, E. *et al.* High expression of acidic chitinase and chitin digestibility in the stomach of common marmoset (Callithrix jacchus), an insectivorous nonhuman primate. *Sci. Rep.* **9**, 159 (2019).

56. Krykbaev, R. *et al.* Evolutionary and biochemical differences between human and monkey acidic mammalian chitinases. *Gene* **452**, 63–71 (2010).

57. Paoletti, M. G., Norberto, L., Damini, R. & Musumeci, S. Human gastric juice contains chitinase that can degrade chitin. *Ann. Nutr. Metab.* **51**, 244–251 (2007).

58. Hu, C. *et al.* Physiological and pathophysiological roles of acidic mammalian chitinase (CHIA) in multiple organs. *Biomed. Pharmacother.* **138**, 111465 (2021).

59. Xiao, W. *et al.* Human secreted stabilin-1-interacting chitinase-like protein aggravates the inflammation associated with rheumatoid arthritis and is a potential macrophage inflammatory regulator in rodents. *Arthritis Rheumatol.* **66**, 1141–1152 (2014).

60. Cho, S. J., Weiden, M. D. & Lee, C. G. Chitotriosidase in the pathogenesis of inflammation, interstitial lung diseases and COPD. *Allergy Asthma Immunol. Res.* **7**, 14–21 (2015).

61. Hussain, M. & Wilson, J. B. New paralogues and revised time line in the expansion of the vertebrate GH18 family. *J. Mol. Evol.* **76**, 240–260 (2013).

62. Lindsay, G. J. H. Distribution and function of digestive tract chitinolytic enzymes in fish. *J. Fish Biol.* **24**, 529–536 (1984).

63. Ikeda, M., Kakizaki, H. & Matsumiya, M. Biochemistry of fish stomach chitinase. *Int. J. Biol. Macromol.* **104**, 1672–1681 (2017).

64. Jørgensen, L. von G. Zebrafish as a model for fish diseases in aquaculture. *Pathogens* **9**, (2020).

65. Stockhammer, O. W., Zakrzewska, A., Hegedûs, Z., Spaink, H. P. & Meijer, A. H. Transcriptome profiling and functional analyses of the zebrafish embryonic innate immune response to Salmonella infection. *J. Immunol.* **182**, 5641–5653 (2009).

66. Meijer, A. H. *et al.* Transcriptome profiling of adult zebrafish at the late stage of chronic tuberculosis due to Mycobacterium marinum infection. *Mol. Immunol.* **42**, 1185–1203 (2005).

67. Teng, Z., Sun, C., Liu, S., Wang, H. & Zhang, S. Functional characterization of chitinase-3 reveals involvement of chitinases in early embryo immunity in zebrafish. *Dev. Comp. Immunol.* **46**, 489–498 (2014).

68. Gao, C. *et al.* Characterization and expression analysis of chitinase genes (CHIT1, CHIT2 and CHIT3) in turbot (Scophthalmus maximus L.) following bacterial challenge. *Fish Shellfish Immunol.* **64**, 357–366 (2017).

69. Manson, F. D. C., Fletcher, T. C. & Gooday, G. W. Localization of chitinolytic enzymes in blood of turbot, Scophthalmus maximus, and their possible roles in defence. *J. Fish Biol.* **40**, 919–927 (1992).

70. Zhang, Y. *et al.* Stimulatory effects of chitinase on growth and immune defense of orange-spotted grouper (Epinephelus coioides). *Fish Shellfish Immunol.* **32**, 844–854 (2012).

71. Macqueen, D. J. & Johnston, I. A. A well-constrained estimate for the timing of the salmonid whole genome duplication reveals major decoupling from species diversification. *Proc. Biol. Sci.* **281**, 20132881 (2014).

72. *World Food and Agriculture - Statistical Pocketbook 2018*. (FAO, 2018). doi:10.4060/CA1796EN.

73. Lien, S. *et al.* The Atlantic salmon genome provides insights into rediploidization. *Nature* **533**, 200–205 (2016).

74. Jones, S., Kim, E. & Bennett, W. Early development of resistance to the salmon louse, Lepeophtheirus salmonis (Krøyer), in juvenile pink salmon, Oncorhynchus gorbuscha (Walbaum). *J. Fish Dis.* **31**, 591–600 (2008).

75. Sutherland, B. J. G., Jantzen, S. G., Sanderson, D. S., Koop, B. F. & Jones, S. R. M. Differentiating size-dependent responses of juvenile pink salmon (Oncorhynchus gorbuscha) to sea lice (Lepeophtheirus salmonis) infections. *Comp. Biochem. Physiol. Part D Genomics Proteomics* **6**, 213–223 (2011).

76. Robledo, D., Gutiérrez, A. P., Barría, A., Yáñez, J. M. & Houston, R. D. Gene expression response to sea lice in atlantic salmon skin: RNA sequencing comparison between resistant and susceptible animals. *Front. Genet.* **9**, 287 (2018).

77. Johansson, G. S. Functional feed for Atlantic salmon (Salmo salar L.).Effect on gastrointestinal tract during parr-smolt transformation and susceptibility to infectious pancreatic necrosis virus (IPNV). *undefined* (2014).

78. Goodwin, S., McPherson, J. D. & McCombie, W. R. Coming of age: ten years of next-generation sequencing technologies. *Nat. Rev. Genet.* **17**, 333–351 (2016).

79. Bentley, D. R. *et al.* Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **456**, 53–59 (2008).

80. Dillies, M.-A. *et al.* A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. *Brief. Bioinformatics* **14**, 671–683 (2013).

81. Eisen, M. B., Spellman, P. T., Brown, P. O. & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* **95**, 14863–14868 (1998).

82. Zheng, C.-H., Yuan, L., Sha, W. & Sun, Z.-L. Gene differential coexpression analysis based on biweight correlation and maximum clique. *BMC Bioinformatics* **15 Suppl 15**, S3 (2014).

83. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, R25 (2010).

84. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).

85. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44**, D457-62 (2016).

86. Ashburner, M. *et al.* Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).

87. Chen, L. *et al.* Prediction and analysis of essential genes using the enrichments of gene ontology and KEGG pathways. *PLoS ONE* **12**, e0184129 (2017).

88. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013).

89. Bentsen, M. *et al.* ATAC-seq footprinting unravels kinetics of transcription factor binding during zygotic genome activation. *Nat. Commun.* **11**, 4267 (2020).

90. Mulugeta, T. D. *et al.* SalMotifDB: a tool for analyzing putative transcription factor binding sites in salmonid genomes. *BMC Genomics* **20**, 694 (2019).

91. Durazzi, F. *et al.* Comparison between 16S rRNA and shotgun sequencing data for the taxonomic characterization of the gut microbiota. *Sci. Rep.* **11**, 3030 (2021).

92. Zhou, Z. *et al.* Methodological approaches used to assess fish gastrointestinal communities. in *Aquaculture nutrition: gut health, probiotics and prebiotics* (eds. Merrifield, D. & Ringø, E.) 101–127 (John Wiley & Sons, Ltd, 2014). doi:10.1002/9781118897263.ch5.

93. Wang, W. *et al.* Quantification of proteins and metabolites by mass spectrometry without isotopic labeling or spiked standards. *Anal. Chem.* **75**, 4818–4826 (2003).

94. Coon, J. J., Syka, J. E. P., Shabanowitz, J. & Hunt, D. F. Tandem mass spectrometry for peptide and protein sequence analysis. *BioTechniques* **38**, 519, 521, 523 (2005).

95. Barad, B. A. *et al.* Differences in the chitinolytic activity of mammalian chitinases on soluble and insoluble substrates. *Protein Sci.* **29**, 966–977 (2020).

96. Bussink, A. P., Speijer, D., Aerts, J. M. F. G. & Boot, R. G. Evolution of mammalian chitinase(-like) members of family 18 glycosyl hydrolases. *Genetics* **177**, 959–970 (2007).

97. Jaillon, O. *et al.* Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. *Nature* **431**, 946–957 (2004).

98. Kasahara, M. *et al.* The medaka draft genome and insights into vertebrate genome evolution. *Nature* **447**, 714–719 (2007).

99. Nakatani, Y., Takeda, H., Kohara, Y. & Morishita, S. Reconstruction of the vertebrate ancestral genome reveals dynamic genome reorganization in early vertebrates. *Genome Res.* **17**, 1254–1265 (2007).

100. Allendorf, F. W. & Thorgaard, G. H. Tetraploidy and the evolution of salmonid fishes. in *Evolutionary genetics of fishes* (ed. Turner, B. J.) 1–53 (Springer US, 1984). doi:10.1007/978-1-4684-4652-4_1.

101. Near, T. J. *et al.* Resolution of ray-finned fish phylogeny and timing of diversification. *Proc Natl Acad Sci USA* **109**, 13698–13703 (2012).

102. Sahlmann, C. *et al.* Ontogeny of the Digestive System of Atlantic Salmon (Salmo salar L.) and Effects of Soybean Meal from Start-Feeding. *PLoS ONE* **10**, e0124179 (2015).

103. Lindsay, G. J. H. *et al.* The growth of rainbow trout (Salmo gairdneri) given diets containing chitin and its relationship to chitinolytic enzymes and chitin digestibility. *Aquaculture* **37**, 315–334 (1984).

104. Karlsen, Ø., Amlund, H., Berg, A. & Olsen, R. E. The effect of dietary chitin on growth and nutrient digestibility in farmed Atlantic cod, Atlantic salmon and Atlantic halibut. *Aquac. Res.* **48**, 123–133 (2017).

105. Buddington, R. K. Hydrolysis-Resistant Organic Matter as a Reference for Measurement of Fish Digestive Efficiency. *Trans. Am. Fish. Soc.* **109**, 653–656 (1980).

106. Lehane, M. J. Peritrophic matrix structure and function. *Annu. Rev. Entomol.* **42**, 525–550 (1997).

107. Hegedus, D., Erlandson, M., Gillott, C. & Toprak, U. New insights into peritrophic matrix synthesis, architecture, and function. *Annu. Rev. Entomol.* **54**, 285–302 (2009).

108. Muthukrishnan, S., Merzendorfer, H., Arakane, Y. & Kramer, K. J. Chitin metabolism in insects. in *Insect molecular biology and biochemistry* 193–235 (Elsevier, 2012). doi:10.1016/B978-0-12-384747-8.10007-8.

109. van der Sluis, M. *et al.* Forkhead box transcription factors Foxa1 and Foxa2 are important regulators of Muc2 mucin expression in intestinal epithelial cells. *Biochem. Biophys. Res. Commun.* **369**, 1108–1113 (2008).

110. Liu, X., Cooper, A. M. W., Zhang, J. & Zhu, K. Y. Biosynthesis, modifications and degradation of chitin in the formation and turnover of peritrophic matrix in insects. *J. Insect Physiol.* **114**, 109–115 (2019).

111. Pisa, D., Alonso, R., Rábano, A., Horst, M. N. & Carrasco, L. Fungal Enolase, β-Tubulin, and Chitin Are Detected in Brain Tissue from Alzheimer's Disease Patients. *Front. Microbiol.* **7**, 1772 (2016).

112. Tsurkan, M. V. *et al.* Progress in chitin analytics. *Carbohydr. Polym.* **252**, 117204 (2021).

Papers I - III

Paper I

Gene family expansion and functional diversification of chitinase and chitin synthase genes in Atlantic salmon (*Salmo salar*)

Matilde Mengkrog Holen[1], Gustav Vaaje-Kolstad[2], Matthew Peter Kent[1], Simen Rød Sandve[1]

[1] Centre for Integrative Genetics (CIGENE), Department of Animal and Aquacultural Sciences, Faculty of Biosciences, Norwegian University of Life Sciences, NO-1432 Ås, Norway

[2] Department of Chemistry, Biotechnology and Food Science (IKBM), Norwegian University of Life Sciences, NO-1432 Ås, Norway

Corresponding author: Simen Rød Sandve, simen.sandve@nmbu.no, +4794870082

Abstract

Background: Chitin is one of the most abundant polysaccharides in nature, forming important structures in insects, crustaceans, and fungal cell walls. Vertebrates on the other hand are generally considered "non-chitinous" organisms, despite having highly conserved chitin metabolism associated genes. Recent work has revealed that the largest group of vertebrates, the teleosts, have the potential to both synthesize and degrade endogenous chitin. Yet little is still known about the genes and proteins responsible for these dynamic processes. Here we used comparative genomics, transcriptomics, and chromatin accessibility data to characterize the repertoire, evolution, and regulation of genes involved in chitin-metabolism in teleosts, with a particular focus on Atlantic salmon.

Results: Reconstruction of gene family phylogenies provide evidence for an expansion of teleost and salmonid chitinase and chitin synthase genes after multiple whole-genome duplications. Analyses of multi-tissue gene expression data demonstrated a strong bias of gastrointestinal tract expression for chitin metabolism genes, but with different spatial and temporal tissue specificities. Finally, we integrated transcriptomes from a developmental time series of the gastrointestinal tract with chromatin accessibility data to identify putative transcription factors responsible for regulating chitin-metabolism gene expression (CDX1 and CDX2) as well as tissue-specific divergence in the regulation of gene duplicates (FOXJ2). These transcription factors are also potential regulators of multiple glycosyltransferases being co-expressed with the chitin remodeling genes.

Conclusion: The findings presented here add support to the hypothesis that chitin metabolism genes in teleosts play a role in developing and maintaining a chitin-based barrier in the teleost gut and provide a basis for further investigations into the molecular basis of this barrier.

Introduction

Chitin is one of the most abundant polysaccharides in nature, serving as the main building block in insect and crustacean exoskeletons as well as forming structural and protective components in fungi. Chitinases and chitin synthases (CHS) are the two major groups of enzymes that have evolved to degrade and synthesize chitin. Decades of work on these enzymes has revealed that bacterial genes encode chitinases that enable bacteria to degrade and utilize chitin as a nutrient source (Cohen-Kupiec and Chet 1998; Beier and Bertilsson 2013), while eukaryotes rich in chitin (i.e. insects, crustaceans and fungi) depend on endogenous chitinases and CHS for normal growth and development (Gooday 1992; Merzendorfer and Zimoch 2003; Zhu *et al.* 2008; Zhang, Zhang, *et al.* 2011; Zhang, Liu, *et al.* 2011; Eichner *et al.* 2015). Curiously, large and highly conserved repertoires of chitinase and CHS genes are also found in vertebrates that do not rely on chitin as a source of nutrition nor possess obvious chitinous body structures such as exoskeletons. Recent experimental work has shown that teleost fish produce chitin in the gastrointestinal tract (GIT) similar to those found in the insect gut epithelium (peritrophic matrix) (Tang *et al.* 2015; Nakashima *et al.* 2018). This realization contradicts the generally held belief that vertebrates are non-chitinous and questions the dogma that chitin does not play an important role in vertebrate physiology.

The function of chitinases in fish has received attention for several reasons. Firstly, chitin is a major component in the natural diets of many fish species, and speculation exists as to whether chitinases could aid in the degradation of chitin to digestible carbohydrates. While fish tissues are known to possess chitinase activity and several fish chitinases have been identified (Fänge *et al.* 1976; Lindsay 1984; Gutowska *et al.* 2004; Zhang *et al.* 2012; Koch *et al.* 2014; Teng *et al.* 2014; Gao *et al.* 2017; Ikeda *et al.* 2017), the activity of these enzymes does not seem to correlate with the ability of fish to digest chitin nor with the amount of chitin in their natural diet (Buddington 1980; Lindsay *et al.* 1984; Kono *et al.* 1987; Danulat 1987; Karlsen *et al*. 2017). Secondly, chitin is present in many fish tissues and structures, such as the developing gut of zebrafish (*Danio rerio*) (Tang *et al.* 2015), the blenny cuticle of *Paralipophrys trigloides* (Wagner *et al.* 1993), the Ampullae of Lorenzini of Chondrichthyes (Phillips *et al.* 2020) and in the scales of parrotfish (*Chlorurus sordidus*), red snapper *(Lutjanus argentimaculatus*), common carp (*Cyprinus carpio*)

and Atlantic salmon (*Salmo salar*) (Zaku *et al.* 2011; Tang *et al.* 2015; Rumengan *et al.* 2017), but the role of chitin in these structures is not known. Thirdly, salmonid fish chitinases have been linked to host-parasite interactions during an infestation of salmon louse (*Lepeophtheirus salmonis*), a small crustacean that feeds on the skin, mucus, and blood of salmonids. For example, resistance to salmon lice in Pink salmon (*Oncorhynchus gorbuscha*) has been suggested to be linked to an increased response of host chitinase in larger fish (Sutherland *et al.* 2011), and upregulation of chitinase gene expression together with genes involved in tissue repair and wound healing in lice-infected skin of Atlantic salmon (Robledo *et al.* 2018).

Our understanding of the repertoire and function of chitin degrading enzymes in vertebrates is mostly derived from studies of mammalian genes and proteins. These genes all belong to the glycoside hydrolase 18 family (GH18), which is an ancient multigene family with a conserved *DXXDXDXE* catalytic motif where glutamate represents the catalytic acid (Terwisscha van Scheltinga *et al.* 1996). In mammals, genes encoding these enzymes can be further subdivided into five main groups. Three of these groups have demonstrated enzymatic activity that enables them to break down chitin; chitotriosidase (CHIT1), acidic mammalian chitinase (CHIA), and di-*N*-acetyl-chitobiase (CTBS). CHIT1 and CHIA are hypothesized to have evolved from one common ancestor gene through whole-genome duplication (WGD) in a common vertebrate ancestor (Hussain and Wilson 2013) and can hydrolyze longer chains of chitin into shorter fragments (chitobiose and chitotriose) (Renkema *et al.* 1995; Boot *et al.* 2001). The CTBS group is more distantly related to CHIT1/CHIA and has evolved to hydrolyze shorter, soluble chitooligosaccharides into *N*-acetyl glucosamine monomers (GlcNAc) allowing for complete degradation of chitin. Chitinase-domain containing protein 1 (CHID1) is another chitinase-related group of proteins that is highly conserved in all vertebrates, although the sequence similarity to other GH18 chitinases is low. Human CHID1 (stabilin-1 interacting protein) lacks essential catalytic residues but contains conserved aromatic residues potentially important for saccharide binding (Meng *et al.* 2010). In mammals, but not in all vertebrates, other saccharide-binding chitinases are termed chitinase-like lectins (CHIL). CHIL are non-enzymatic chitinase-like proteins very similar to CHIA and CHIT1, but with active site mutations that render the proteins catalytically incompetent. Human CHIL (OVGP1, CHI3L1, and CHI3L2) are according to phylogenetic analyses of mammalian CHIL predicted to have evolved from gene duplications of ancestral CHIA and CHIT1 (Funkhouser and Aronson 2007; Bussink *et al.* 2007). A newly identified group of

vertebrate chitinases that does not fit into any of the five mammalian groups is a group called CHIO (Hussain and Wilson 2013). Like CHIL, CHIO is also hypothesized to have evolved from ancestral CHIA and/or CHIT1. Two rounds of whole-genome duplication events specific for teleost (Ts3R) and salmonid fish (Ss4R) have resulted in an amplification of genes that are closely related to this group. There is, however, a lack of systematic effort to characterize the potential for teleost genomes to encode chitin degrading and synthesizing enzymes.

In this paper, we have characterized the evolution and diversification of genes involved in chitin breakdown and synthesis in Atlantic salmon. Using a comparative approach that combines both comparative and functional genomics, we provide an improved understanding of putative protein functions and gene regulation of chitinase and CHS genes in Atlantic salmon. Our results provide a knowledge base for further functional studies of chitin-biology in teleost fish and support the idea that chitin plays a major role in GIT function and physiology.

Materials and methods

Phylogenetic analysis

Orthofinder (v.0.3.1) was used to construct orthogroups using the longest protein isoform sequence from gene. Species included in the orthogroups computation were spotted gar (*Lepisosteus oculatus*, LepOcu1), zebrafish (*Danio rerio*, GRCz10), stickleback (*Gasterosteus aculeatus*, BROADS1), Japanese medaka (*Oryzias latipes*, HdrR), pike (*Esox Lucius*, Eluc_V3), rainbow trout (*Oncorhynchus mykiss*, Omyk_1.0), coho salmon (*Oncorhynchus kisutch*, Okis_V1), Atlantic salmon (*Salmo salar*, ICSASG_v2), human (*Homo sapiens*, GRCh38), and house mouse (*Mus musculus*, GRCm38). For each orthogroup, protein sequences were then aligned using MAFFT (v.7) (Katoh and Standley 2013). A maximum likelihood phylogenetic tree was constructed in MEGA7 (Kumar *et al.* 2016) using a neighbor-joining algorithm with a JTT substitution model and 100 bootstrap replicates.

Tissue expression profiles

See Supplementary Table 1 for more information about species, tissues examined, number of individuals, and where the data are available. Tissue expression profiles from Atlantic salmon (except stomach, pyloric caeca, and midgut), rainbow trout,

zebrafish and pike (n = 1 for all tissues except the liver where n = 3 for rainbow trout and n = 4 for zebrafish) were generated from RNA-sequencing (RNA-seq) data as described previously (Lien *et al.* 2016; Gillard *et al.* 2021). In brief, the STAR aligner with default settings (Dobin *et al.* 2013) was used to map RNA-seq reads to the annotated reference genomes and RSEM (Li and Dewey 2011) was used to estimate read counts. Tissue expression data from the stomach, pyloric caeca and midgut of Atlantic salmon (n = 15 for stomach and pyloric caeca, n = 167 for midgut) were generated from previously published RNA-seq data (Gillard *et al.* 2018; Jin *et al.* 2018) following the described method. The RNA-seq data were mapped to the annotated genome (ICSASG_v2) using the STAR aligner, and the read counts were estimated with HTSeq-count (Anders *et al.* 2015). The read counts were transformed to Transcript Per Million Reads (TPM) values normalized for average transcript length and sample size. To get TPM values, the raw gene counts were first divided by the transcript length before dividing by the total library count number. The mean gene expression value was used for the liver, and the median gene expression value was used for the stomach, pyloric caeca, and midgut. The gene expression values were log-transformed (Log_2(TPM + 1) before further analysis. The RNA-seq data analysis was performed using R (v.3.6.0).

Gene expression in pyloric caeca of Atlantic salmon spanning the transition from endogenous to exogenous feeding

The expression profiles of chitinases and CHS in pyloric caeca of Atlantic salmon spanning the developmental transition to external feeding were obtained from an RNA-seq dataset available through ArrayExpress under the project number E-MTAB-8306 and was generated as described previously (Jalili *et al.* 2019). Differences in expression levels compared to day 0 was tested by comparing means of expression using a Wilcoxon test with the function "stat_compare_means" in the R-package "Ggpubr" using the default "wilcox.test" parameter. The p-values were adjusted for multiple testing. Genes with low expression (TPM < 1) were removed before the co-expression analysis and quality control of the resulting genes was conducted using the function "goodSamplesGenes" in the "WGCNA" package in R (Langfelder and Horvath 2008) with the argument "verbose = 3". The co-expression analysis was carried out using the minimum biweight midcorrelation ("bicor") function in the "WGCNA" package with the argument «maxPOutliers = 0.05» and genes with a correlation value above 0.69 was referred to as co-expressed genes. Gene

enrichment of the co-expressed genes was done using KEGG enrichment with the function “kegga” from the “limma” package in R (Ritchie *et al.* 2015) and the argument «species.KEGG = "sasa"» and the universe specified to be only expressed genes. The p-values returned by “kegga” were not adjusted for multiple testing.

Chromatin accessibility in pyloric caeca

ATAC-seq reads from pyloric caeca of Atlantic salmon were downloaded from ArrayExpress (E-MTAB-9001). Read mapping and ATAC-peak calling was done using BWA (v.0.7.17) (Li and Durbin 2009) and Genrich v.06 (https://github.com/jsh58/Genrich) as described in detail in Bertolotti *et al.* (2020).

Transcription factor motif enrichment

DNA sequences from open chromatin (i.e. within ATAC-seq peaks) around TSS (1000bp upstream to 200bp downstream) of chitinase- and CHS genes were used for transcription factor motif scan and enrichment. The scan and enrichment was carried out using SalMotifDB (Mulugeta *et al.* 2019), a tool for analyzing putative transcription factor binding sites in Atlantic salmon. Consensus motifs were obtained using the “ggseqlogo” package in R.

Results

Phylogenetic analysis of chitinase protein sequences

The annotated Atlantic salmon genome (ICSASG_v2) includes 12 genes with strong homology to mammalian chitinase genes (see Supplementary Table 2 for gene IDs, proteins accession numbers, and names given in this paper) belonging to the family 18 of the glycoside hydrolases, as classified by the carbohydrate-active enzyme (CAZy) database (Drula *et al.* 2022). To investigate the evolutionary history of the chitinase gene family in fish we reconstructed phylogenetic trees of genes within the glycoside hydrolase family 18 orthogroup. The species selection was designed to include vertebrates that have experienced different numbers of whole-genome duplications. All species share the two whole-genome duplications occurring in the ancestor of all vertebrates, and except gar, all fish species share an additional whole genome duplication at the base of the teleost lineage (Ts3R), while Atlantic salmon,

rainbow trout and Coho salmon share an additional fourth salmonid-specific whole-genome duplication event (Ss4R) (Figure 1A).

Our analysis revealed that the 12 Atlantic salmon chitinase proteins are distributed among six major clades (Figure 1B). These six clades formed two major "superclades" (supported by a high bootstrap value of 83), one containing CTBS and CHID1 type proteins (two from Atlantic salmon) and the other containing CHIT1, CHIA, and CHIO (10 from Atlantic salmon). Of the 10 Atlantic salmon proteins annotated as acidic mammalian chitinases (CHIA) in the NCBI RefSeq annotation (release 100) all salmon proteins share the following chitinase characteristics: a signal peptide, a glycoside hydrolase 18 family catalytic domain, and a chitin-binding domain (CBM14; identified by dbCAN2 annotation (Zhang *et al.* 2018)) at the carboxyl-terminus. Three salmon CHIA proteins (namely Chia.3, Chia.4, and Chia.7) fell into a monophyletic clade (descending from a common ancestor) containing the human acidic mammalian protein AMCase, whereas the remaining seven CHIA protein sequences were distributed among two teleost-specific monophyletic clades. One salmon CHIA protein (Chia.5) will hereafter be referred to as a CHIT1-member. This protein is the only salmon chitinase protein with loss-of-function mutations in the catalytic motif and a truncated chitin-binding domain. The remaining six salmon chitinase proteins (Chia.1, Chia.2, Chia.6, Chia.8, Chia.9, and Chia.10) belong to two clades (termed CHIO I and II) forming a larger monophyletic group.

To make interferences about how ancient whole-genome duplication and other duplication events have contributed to the present diversity of chitinase proteins in teleost fish, the protein sequence phylogeny was compared with the species tree topology. The CHID1 and CTBS clades only contain one protein sequence per species, and the protein trees resemble to a large extent the species topology except for the polytomy in the CTBS clade that fails to place the mammals as a sister clade to the teleost species. This is in agreement with the hypothesis that CTBS and CHID1 genes resulted from an ancient gene duplication before the vertebrate diversification (Funkhouser and Aronson 2007; Hussain and Wilson 2013), possibly the whole genome duplication at the base of all vertebrates. The two distinct fish-specific CHIA subclades are more closely related to each other than to their sister subclade containing the mammalian CHIA proteins. Furthermore, since both fish subclades contain a predicted gar protein it is likely that these fish-specific duplicates arose through a duplication event prior to the divergence of teleosts. In the fish-specific CHIO clade, comprising 39 protein sequences, the three gar-specific proteins cluster closely together and, because of low bootstrap values (< 70) for key splits in the tree,

we cannot firmly place these in relation to the remaining teleosts. However, based on the sequence relationships between the teleost CHIO species it is likely that Ts3R has contributed to at least one CHIO-duplication event as previously hypothesized (Hussain and Wilson 2013). Two nodes reflecting the Ss4R event can be inferred in the CHIO I and II clades, including the branches containing Chia.1+10, and Chia.2+9. These proteins are located on homologous regions of different chromosomes (22 and 12) in Atlantic salmon (Supplementary Figure 1).

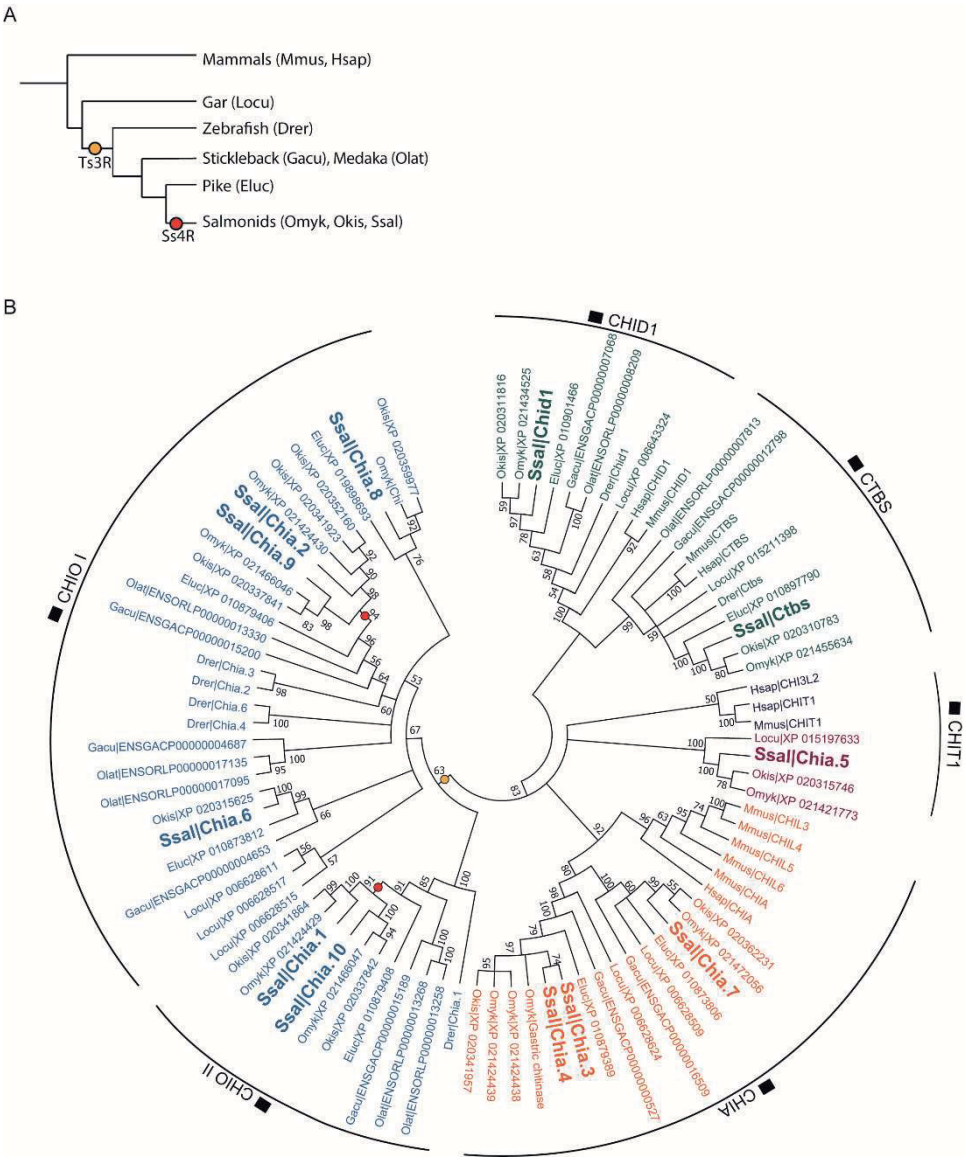

Figure 1. Comparative analysis of GH18 proteins. A) Whole-genome duplication (WGD) events experienced by species included in the phylogenetic comparison. B) Phylogenetic tree of GH18 chitinase proteins in spotted gar (locu), zebrafish (drer), stickleback (gacu), Japanese medaka (olat), pike (eluc), rainbow trout (omyk), coho salmon (okis), Atlantic salmon (ssal), human (hsap) and house mouse (mmus). The colors represent different monophyletic clades in the phylogenetic tree.

Tissue-specific regulatory divergence of chitinase genes in the gastrointestinal tract

A comparative analysis of tissue expression in zebrafish, pike, rainbow trout, and Atlantic salmon (Figure 2) was performed to characterize the divergence of gene regulation encoding chitinase enzymes. Members of the CHIT1 group were not included in this analysis as their expression levels were low, indicating that they may represent pseudogenes encoding non-functional enzymes. Across all species, the results showed a clear bias towards gene expression in GIT and revealed both conserved expression divergence among orthologs in different species as well as lineage-specific regulatory divergence.

CHIA genes displayed the most conserved tissue expression regulation across all species with stomach-specific expression. A similar stomach bias is also observed for CHIA in mice, bats, pigs, chickens, and humans indicating that CHIA enzymes share an important gastric function that is conserved across fish, mammals, and birds (Boot *et al.* 2001; Strobel *et al.* 2013; Ohno *et al.* 2016; Tabata, Kashimura, Wakita, O hno, Sakaguchi, Sugahara, Kino, *et al.* 2017; Tabata, Kashimura, Wakita, Ohno, Sakag uchi, Sugahara, Imamura, *et al.* 2017; Tabata *et al.* 2019). Notably, the agastric (stomach-less) zebrafish do not express genes related to gastric functions, including CHIA genes.

The tissue expression profiles of CHIO- I and II genes show different patterns compared to the CHIA genes. Although CHIO genes also display GIT expression dominance, these genes are not stomach-specific but rather expressed in other GIT sections such as pyloric caeca and midgut. Additionally, CHIO gene expression is generally less tissue-specific and has larger inter- and intra-tissue specific variations in tissue expression patterns. For example, while both salmon CHIO II genes are expressed almost exclusively in the GIT, the CHIO I clade contains salmon genes expressed in the GIT and a gene (*chia.8*) that is lowly expressed in all tissues examined. We also observe some less striking, but clear, cases of regulatory divergence of CHIO II genes following the more recent Ss4R, with one duplicate being mostly expressed in pyloric caeca (*chia.10*), while the other is expressed in both pyloric caeca and gut (*chia.1*).

Finally, the CTBS and CHID1 gene groups show a different tissue regulation pattern than the other chitinases. Both gene groups are generally expressed at low levels compared to their CHIA and CHIO counterparts. CHID1 orthologs in zebrafish, pike, and rainbow trout are ubiquitously expressed in all tissues, but the salmon *chid1* gene

is expressed at two-fold higher levels in the midgut compared to the stomach and pyloric caeca, and four-fold higher in the midgut compared to non-GIT tissues indicating some tendency to GIT specific expression. Similar to CHID1, fish CTBS are expressed across all tissues, but with two-fold higher expression in the GIT of Atlantic salmon.

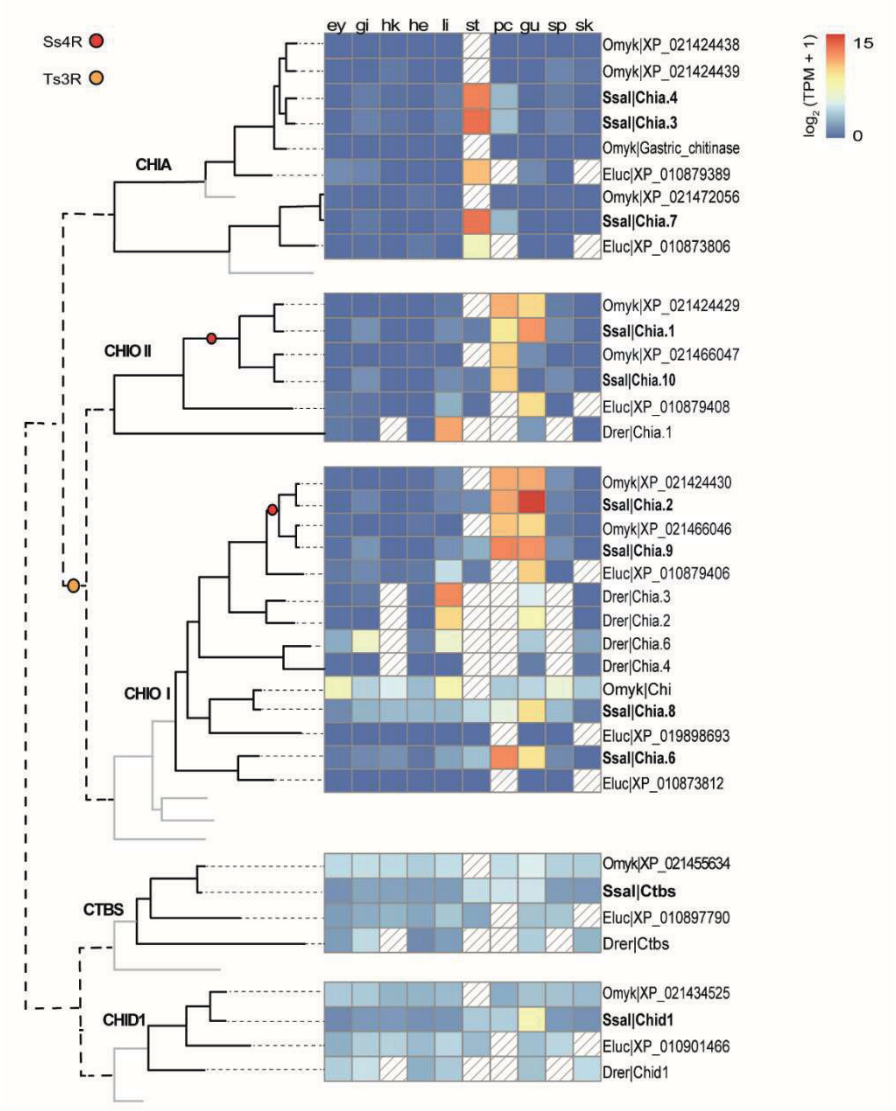

Figure 2. Comparative tissue expression of GH18 genes in zebrafish (drer), pike (eluc), rainbow trout (omyk), and Atlantic salmon (ssal). n≥1. The phylogenetic tree is a simplified version of figure 1B and the lines do not represent real evolutionary distances. Yellow and red circles represent the teleost-specific whole-genome duplication (Ts3R) and salmonid-specific whole-genome duplication (Ss4R), respectively. Solid light grey lines indicate the phylogenetic

position of a spotted gar, but expression data were not analyzed for this outgroup. The tissue expression panel shows gene expression of GH18 genes in the following tissues: ey = eye, gi = gill, hk = head kidney, he = heart, li = liver, st = stomach, pc = pyloric caeca, gu = midgut, sp = spleen and sk = skin. Colored boxes indicate gene expression in the range of 0 to 15 $\log_2$(TPM + 1) values, while diagonal lines represent missing data.

Chitin synthase genes are mainly expressed in pyloric caeca and midgut of teleost fish

The phylogenetic analysis of the gene family containing genes encoding CHS proteins showed a split into two major subclades (I and II) which, since gar has a single gene copy, likely arose in the Ts3R whole-genome duplication event (Figure 3). Furthermore, the salmonid-specific gene copies in subclade II (i.e. *chs1a* and *chs1b* from Atlantic salmon) likely originate from the Ss4R as they are located on chromosomes (28 and 1 respectively) matching the well described synteny within the duplicated Atlantic salmon genome (Lien *et al.* 2016) (Supplementary Figure 1). The tissue expression pattern shows that the CHS genes in subclade I are expressed in low abundance in all tissues, whereas CHS genes in subclade II follow the same expression pattern as CHIO- I and II genes (Figure 2), with expression specific to pyloric caeca and gut (Figure 3).

Notably, predicted protein sequences of teleost CHS genes contain similar conserved amino acid sequence motifs as found in insect CHS proteins. The motifs *EDR* and *QRRRW* are common for all CHS, while *CATMWHXT* and *QKFEY* are signatures of insect CHS (Merzendorfer and Zimoch 2003). *EDR*, *QRRW,* and *QKFEY* are motifs found in all predicted fish CHS protein sequences examined, but the *CATMWHXT* motif is present in the CHS subclade II only.

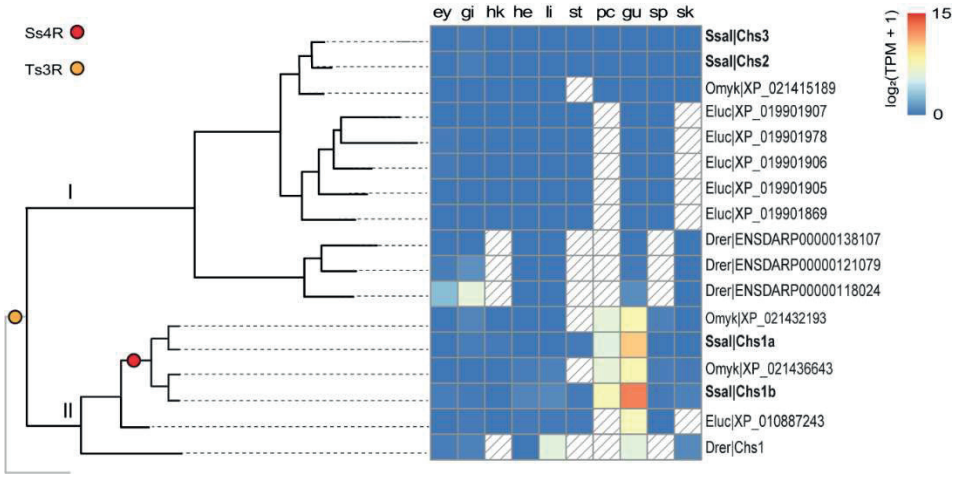

Figure 3. Comparative tissue expression of CHS (chitin synthase) genes in zebrafish (drer), pike (eluc), rainbow trout (omyk), and Atlantic salmon (ssal). n≥1. Yellow and red circles in the illustrative phylogenetic tree represent the teleost-specific whole-genome duplication (Ts3R) and salmonid-specific whole-genome duplication (Ss4R), respectively. The lines do not represent real evolutionary distances. Solid light grey lines indicate the phylogenetic position of a spotted gar, but expression data were not analyzed for this outgroup. The tissue expression panel shows gene expression of CHS genes in the following tissues: ey = eye, gi = gill, hk = head kidney, he = heart, li = liver, st = stomach, pc = pyloric caeca, gu = midgut, sp = spleen and sk = skin. Colored boxes indicate gene expression in the range of 0 to 15 $\log_2$(TPM + 1) values. Boxes with diagonal lines represent missing data.

Gene regulation of chitinases and chitin synthases in pyloric caeca of Atlantic salmon

To be able to better understand the regulation of chitin metabolism genes we leveraged a RNA-seq dataset from pyloric caeca of Atlantic salmon (ArrayExpress, E-MTAB-8306) that spans the developmental transition from endogenous to exogenous nutrition (Jalili *et al.* 2019). The changes in gene expression observed across the developmental time series show two major trends; 5 genes significantly increased expression following external feed intake ($p < 0.01$), while 3 genes did not (Figure 4A).

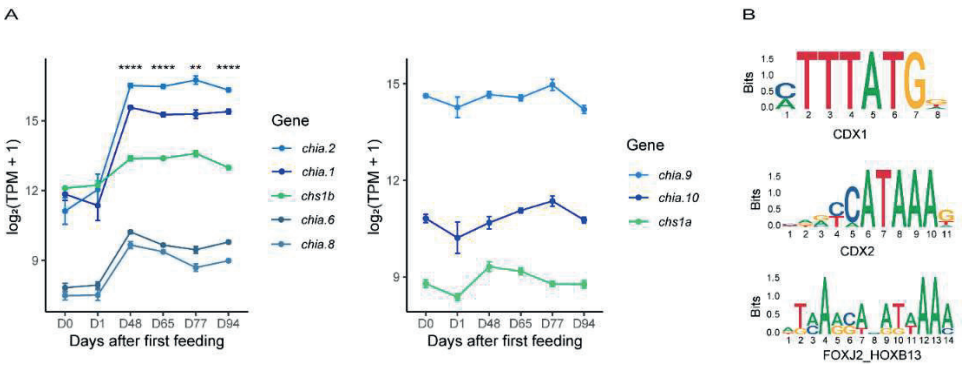

Figure 4. A) Gene expression levels of *chia.1*, *chia.2*, *chia.6*, *chia.*8, *chia.9*, *chia.10*, *chs1a* and *chs1b* before and after external feeding. Chitinase and CHS gene expression in the pyloric caeca of Atlantic salmon days before (D0) and days after external feeding (D1-D94). Please note that the y-axis does not extend to 0. The asterisks (**, ****) indicate a significant difference in expression compared to D0 (p.adj<0.01 and p.adj<0.0001, respectively, n≥4. B) Consensus motifs for binding of transcription factors in promoters of chitinase- and CHS genes. CDX1 and CDX2 motifs are found in promoters of all chitinase- and CHS genes. The FOXJ2_HOXB13 motif is only present in promoters of genes with a significant increase in expression upon transition to external feeding. The numbers indicate the consensus site position of each base, and the vertical axes (Bits) indicate the information content of the base frequency at the given base position.

Co-expressed genes (genes with correlated expression profiles) are often controlled by the same regulators and involved in the same biological processes. To better understand the mechanisms underlying the regulation of genes involved in chitin metabolism, and particularly drive increased expression of some chitinases and CHS following feed ingestion we used a co-expression approach. We first used biweight midcorrelation (bicor) to identify genes with similar expression patterns in the pyloric caeca across the developmental time series. Based on KEGG (Kyoto Encyclopedia of Genes and Genomes) gene enrichment analysis, co-expressed (bicor > 0.69, n = 36) with salmon *chia.1*, *chia.2*, *chia.6*, *chia.8,* and *chs1b* were genes involved in metabolic processes like amino sugar and nucleotide sugar metabolism (p-value = $8.79 \cdot 10^{-10}$) and glycosphingolipid biosynthesis (p-value = $5.92 \cdot 10^{-4}$) (Supplementary Table 3). We found the chitinase and CHS genes to be associated with the amino sugar and nucleotide sugar metabolism KEGG pathway, together with a UDP-*N*-acetylhexosamine pyrophosphorylase-like gene (*uap1*) which most likely codes for an enzyme that converts uridine triphosphate (UTP) and *N*-acetylglucosamine-1-phosphate (GlcNAc-1-P) into uridine diphosphate *N*-acetylglucosamine (UDP-GlcNAc) (Mio *et al.* 1998). This is known to result in an activated substrate required for chitin synthesis by CHS. UDP-GlcNAc can also be

transferred by beta-1,3-galactosyl-O-glycosyl-glycoprotein beta-1,6-*N*-acetylglucosaminyltransferase (GCNT1) to form mucin-type O-glycan structures. *Gcnt1* is one of the multiple glycosyltransferases being co-expressed with chitinase- and CHS genes. Associated with the glycosphingolipid biosynthetic pathway we found two additional glycosyltransferase genes coding for alpha-2,8-sialyltransferase-like proteins, involved in the transfer of sialic acid to produce sialyl glycoconjugates.

To further dissect out putative transcription factors involved in the regulation of chitinase and CHS genes, we performed a transcription factor binding (TFBS) scan for two classes of genes: (1) all CHIO- and CHS genes being highly expressed in pyloric caeca, and (2) those that we find induced during the developmental transition to external feeding exclusively (Figure 4A). We based the scan on open-chromatin regions of promoter sequences of chitinase– and CHS genes. The data show that all chitinase- and CHS genes had open chromatin regions spanning the TSS in pyloric caeca except CHIA-genes and *chia.8*. Furthermore, the transcription factor motif scan revealed that two motifs were common for all CHIO- and CHS genes. These motifs were two homeodomain (HOX) related motifs: a caudal type homeobox 1 (CDX1) motif and a caudal type homeobox 2 (CDX2) motif (Figure 4B). For the CHIO-, CHS- and co-expressed genes induced during the transition to external feeding the FOXJ2_HOXB13 motif associated with binding of the forkhead box (FOX) transcriptional factor family (Figure 4B) were enriched (p-value < 0.01).

Discussion

Our results suggest that Ts3R and Ss4R duplication events resulted in the expansion of chitinase- and CHS genes in fish (Figure 1) and that these genes generally encode proteins with conserved residues in active motifs. This is strikingly different from mammals, which have lost their genes for CHS and where mutations in the active site of chitinases followed by mammal specific gene duplications have resulted in the expansion of non-enzymatic chitinase-like lectins. In general, teleost and salmonid chitinase- and CHS genes share a clear expression bias towards the gastrointestinal tract (Figure 2 & 3). This expression bias may explain the presence of chitin in the gut of zebrafish (Tang *et al.* 2015) and rainbow trout (Nakashima *et al.* 2018), and adds support to a chitin-based mucosal barrier in the gut of teleost fish previously hypothesized to have evolved from the chitin-based barrier we find in invertebrates (Nakashima *et al.* 2018). This is also in line with the increased expression of chitin metabolism genes during development (Figure 4), which coincides with the

development of a pyloric caeca with increased complex intestinal mucosal structures (Sahlmann *et al.* 2015). The transition from endogenous to external feeding involves exposure to both larger food particles and new microbial communities. We predict that this exposure boosts the expression of chitinases and CHS needed to synthesize and remodel a chitinous layer that surrounds the intestinal mucosa and protects the intestinal epithelium, in addition to other genes related to intestinal differentiation and mucus production. Co-expression of *uap1* and glycosyltransferases like *gcnt1* supports this assumption, as UAP1 can produce the activated UDP-GlcNAc used by both CHS to produce chitin and by GCNT1. GCNT1 is important for the production of mucins and mice deficient in related genes have been shown to have increased permeability of the mucosal barrier which can alter the mucosal immune homeostasis (Stone *et al.* 2009). Furthermore, the co-expressed alpha-2,8-sialyltransferase orthologs are, in humans, linked to the production of gangliosides; glycosphingolipids that contain one or multiple residues of sialic acid and that is localized in the brush border membrane of intestinal enterocytes. Such gangliosides are known to be important for maintaining intestinal integrity and reducing inflammation (Miklavcic *et al.* 2012).

Little is known about the regulatory networks of chitinases and chitin synthetases. The presence of FOX, CDX1, and CDX2 motifs in the promoter regions of CHIO- and CHS genes presented here support the hypothesis of a possible role in the formation of a chitin-based mucosal barrier. Both CDX1 and CDX2 are known to be major regulators of intestine-specific genes and are crucial for intestinal differentiation. In zebrafish, for example, knockdown and overexpression experiments show that CDX1B, homologous to mammalian CDX1, is responsible for terminal differentiation of goblet cells, cells that are responsible for the secretion of mucins into the intestinal mucosa (Chen *et al.* 2009). Moreover, it is plausible that the FOXJ2_HOXB13 motif could be bound by intestinal FOX proteins like FOXA1 and FOXA2. These proteins are also linked to intestinal goblet cell mucus production (van der Sluis *et al.* 2008). Synergetic transcription factor binding of CDX and FOXA transcription factors has previously been shown to regulate intestine-specific mucins (Jonckheere *et al.* 2007). Similar synergetic effects can possibly explain the difference in gene expression of the chitinase- and CHS genes lacking the FOX motif in their promoter. That said, the chromatin accessibility is likely to change during development and it is important to take into consideration that the ATAC-seq data used to guide the search for transcription factor binding sites in this study was derived from adult fish.

Nevertheless, CDX1, CDX2 and FOX are interesting candidates for futures studies into the regulation of chitin metabolism genes in fish GIT.

Functional diversification is observed for the different fish chitinases. We find that, unlike the CHIO- and CHS genes, CHIA genes are exclusively expressed in the stomach and have no open chromatin regions in their promoters in the pyloric caeca of Atlantic salmon. Thus, these genes are likely to be regulated by other transcription factors than the CHIO- and CHS genes. The high degree of sequence similarity to mammalian stomach chitinases suggests that the fish-specific CHIA proteins share an ancestral function related to the presence of an acidic stomach (Krogdahl *et al.* 2015). This hypothesis is strengthened by the loss of CHIA genes in agastric zebrafish. As various fish stomach CHIA proteins have shown to be able to break down chitin structures typically found in the natural diet of Atlantic salmon, like shrimp, squid, and insects (Ikeda *et al.* 2017), we cannot rule out a possible role of teleost CHIA proteins in digestion. The results presented here thus imply that some fish species like Atlantic salmon do have the genetic toolbox needed to tolerate and digest chitin-containing feed.

Conclusion

There has been an expansion of chitinase and chitin-synthase like proteins in Atlantic salmon and different groups of chitinases have evolved to be expressed in different tissues. Based on our results we hypothesize two different roles of Atlantic salmon chitinases and CHS: (1) that stomach-related CHIA proteins aid in the digestion of chitin, whereas (2) CHIO- and CHS proteins are involved in remodeling of chitinous structures surrounding mucosal membranes of pyloric caeca and gut. To verify this, functional characterization of chitinous structures and the enzymes that remodel these are needed. However, this work provides a basis for future functional studies identifying the underlying mechanisms for the presence of chitinous structures in Atlantic salmon.

Data availability statement

The authors declare that the data supporting the results in this study are accessible in the paper and the supplementary file.

Acknowledgements

We want to acknowledge Dr. Gareth Benjamin Gillard, Dr. Yang Jin, and Dr. Thomas Nelson Harvey for their help with bioinformatics.

References

Anders, S., P. T. Pyl, and W. Huber, 2015 HTSeq — a Python framework to work with high-throughput sequencing data. Bioinformatics 31: 166–169.

Beier, S., and S. Bertilsson, 2013 Bacterial chitin degradation-mechanisms and ecophysiological strategies. Front. Microbiol. 4: 149.

Bertolotti, A. C., R. M. Layer, M. K. Gundappa, M. D. Gallagher, E. Pehlivanoglu *et al.*, 2020 The structural variation landscape in 492 Atlantic salmon genomes. Nat. Commun. 11: 5176.

Boot, R. G., E. F. Blommaart, E. Swart, K. Ghauharali-van der Vlugt, N. Bijl *et al.*, 2001 Identification of a novel acidic mammalian chitinase distinct from chitotriosidase. J. Biol. Chem. 276: 6770–6778.

Buddington, R. K., 1980 Hydrolysis-Resistant Organic Matter as a Reference for Measurement of Fish Digestive Efficiency. Trans. Am. Fish. Soc. 109: 653–656.

Bussink, A. P., D. Speijer, J. M. F. G. Aerts, and R. G. Boot, 2007 Evolution of mammalian chitinase(-like) members of family 18 glycosyl hydrolases. Genetics 177: 959–970.

Chen, Y.-H., Y.-F. Lu, T.-Y. Ko, M.-Y. Tsai, C. Y. Lin *et al.*, 2009 Zebrafish cdx1b regulates differentiation of various intestinal cell lineages. Dev. Dyn. 238: 1021–1032.

Cohen-Kupiec, R., and I. Chet, 1998 The molecular biology of chitin digestion. Curr. Opin. Biotechnol. 9: 270–277.

Danulat, E., 1987 Digestibility of chitin in cod,Gadus morhua, in vivo. Helgoländer Meeresunters. 41: 425–436.

Dobin, A., C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski *et al.*, 2013 STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29: 15–21.

Drula, E., M.-L. Garron, S. Dogan, V. Lombard, B. Henrissat *et al.*, 2022 The carbohydrate-active enzyme database: functions and literature. Nucleic Acids Res. 50: D571–D577.

Eichner, C., E. Harasimczuk, F. Nilsen, S. Grotmol, and S. Dalvin, 2015 Molecular characterisation and functional analysis of LsChi2, a chitinase found in the salmon louse (Lepeophtheirus salmonis salmonis, Krøyer 1838). Exp. Parasitol. 151–152: 39–48.

Funkhouser, J. D., and N. N. Aronson, 2007 Chitinase family GH18: evolutionary insights from the genomic history of a diverse protein family. BMC Evol. Biol. 7: 96.

Fänge, R., G. Lundblad, and J. Lind, 1976 Lysozyme and chitinase in blood and lymphomyeloid tissues of marine fish. Mar. Biol. 36: 277–282.

Gao, C., X. Cai, Y. Zhang, B. Su, H. Song *et al.*, 2017 Characterization and expression analysis of chitinase genes (CHIT1, CHIT2 and CHIT3) in turbot (Scophthalmus maximus L.) following bacterial challenge. Fish Shellfish Immunol. 64: 357–366.

Gillard, G. B., L. Grønvold, L. L. Røsæg, M. M. Holen, Ø. Monsen *et al.*, 2021 Comparative regulomics supports pervasive selection on gene dosage following whole genome duplication. Genome Biol. 22: 103.

Gillard, G., T. N. Harvey, A. Gjuvsland, Y. Jin, M. Thomassen *et al.*, 2018 Life-stage-associated remodelling of lipid metabolism regulation in Atlantic salmon. Mol. Ecol. 27: 1200–1213.

Gooday, G., 1992 What are the roles of chitinases in the growing fungus? FEMS Microbiol. Lett. 100: 387–391.

Gutowska, M. A., J. C. Drazen, and B. H. Robison, 2004 Digestive chitinolytic activity in marine fishes of Monterey Bay, California. Comp Biochem Physiol, Part A Mol Integr Physiol 139: 351–358.

Hussain, M., and J. B. Wilson, 2013 New paralogues and revised time line in the expansion of the vertebrate GH18 family. J. Mol. Evol. 76: 240–260.

Ikeda, M., H. Kakizaki, and M. Matsumiya, 2017 Biochemistry of fish stomach chitinase. Int. J. Biol. Macromol. 104: 1672–1681.

Jalili, M., Y. Jin, A. M. Bones, Y. Olsen, O. Vadstein *et al.*, 2019 Dietary fatty acid source has little effect on the development of the immune system in the pyloric caeca of Atlantic salmon fry. Sci. Rep. 9: 27.

Jin, Y., R. E. Olsen, M.-A. Østensen, G. B. Gillard, S. A. Korsvoll *et al.*, 2018 Transcriptional development of phospholipid and lipoprotein metabolism in different intestinal regions of Atlantic salmon (Salmo salar) fry. BMC Genomics 19: 253.

Jonckheere, N., A. Vincent, M. Perrais, M.-P. Ducourouble, A. K. Male *et al.*, 2007 The human mucin MUC4 is transcriptionally regulated by caudal-related homeobox, hepatocyte nuclear factors, forkhead box A, and GATA endodermal transcription factors in epithelial cancer cells. J. Biol. Chem. 282: 22638–22650.

Karlsen, Ø., H. Amlund, A. Berg, and R. E. Olsen, 2017 The effect of dietary chitin on growth and nutrient digestibility in farmed Atlantic cod, Atlantic salmon and Atlantic halibut. Aquac. Res. 48: 123–133.

Katoh, K., and D. M. Standley, 2013 MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30: 772–780.

Koch, B. E. V., J. Stougaard, and H. P. Spaink, 2014 Spatial and temporal expression patterns of chitinase genes in developing zebrafish embryos. Gene Expr. Patterns 14: 69–77.

Kono, M., T. Matsui, and C. Shimizu, 1987 Effect of chitin, chitosan, and cellulose as diet supplements on the growth of cultured fish. NSUGAF 53: 125–129.

Krogdahl, Å., A. Sundby, and H. Holm, 2015 Characteristics of digestive processes in Atlantic salmon (Salmo salar). Enzyme pH optima, chyme pH, and enzyme activities. Aquaculture 449: 27–36.

Kumar, S., G. Stecher, and K. Tamura, 2016 MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33: 1870–1874.

Langfelder, P., and S. Horvath, 2008 WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 9: 559.

Li, B., and C. N. Dewey, 2011 RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12: 323.

Li, H., and R. Durbin, 2009 Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25: 1754–1760.

Lien, S., B. F. Koop, S. R. Sandve, J. R. Miller, M. P. Kent *et al.*, 2016 The Atlantic salmon genome provides insights into rediploidization. Nature 533: 200–205.

Lindsay, G. J. H., M. J. Walton, J. W. Adron, T. C. Fletcher, C. Y. Cho *et al.*, 1984 The growth of rainbow trout (Salmo gairdneri) given diets containing chitin and its relationship to chitinolytic enzymes and chitin digestibility. Aquaculture 37: 315–334.

Lindsay, G. J. H., 1984 Distribution and function of digestive tract chitinolytic enzymes in fish. J. Fish Biol. 24: 529–536.

Meng, G., Y. Zhao, X. Bai, Y. Liu, T. J. Green *et al.*, 2010 Structure of human stabilin-1 interacting chitinase-like protein (SI-CLP) reveals a saccharide-binding cleft with lower sugar-binding selectivity. J. Biol. Chem. 285: 39898–39904.

Merzendorfer, H., and L. Zimoch, 2003 Chitin metabolism in insects: structure, function and regulation of chitin synthases and chitinases. J. Exp. Biol. 206: 4393–4412.

Miklavcic, J. J., K. L. Schnabl, V. C. Mazurak, A. B. R. Thomson, and M. T. Clandinin, 2012 Dietary ganglioside reduces proinflammatory signaling in the intestine. J. Nutr. Metab. 2012: 280286.

Mio, T., T. Yabe, M. Arisawa, and H. Yamada-Okabe, 1998 The eukaryotic UDP-N-acetylglucosamine pyrophosphorylases. Gene cloning, protein expression, and catalytic mechanism. J. Biol. Chem. 273: 14392–14397.

Mulugeta, T. D., T. Nome, T.-H. To, M. K. Gundappa, D. J. Macqueen *et al.*, 2019 SalMotifDB: a tool for analyzing putative transcription factor binding sites in salmonid genomes. BMC Genomics 20: 694.

Nakashima, K., S. Kimura, Y. Ogawa, S. Watanabe, S. Soma *et al.*, 2018 Chitin-based barrier immunity and its loss predated mucus-colonization by indigenous gut microbiota. Nat. Commun. 9: 3402.

Ohno, M., M. Kimura, H. Miyazaki, K. Okawa, R. Onuki *et al.*, 2016 Acidic mammalian chitinase is a proteases-resistant glycosidase in mouse digestive system. Sci. Rep. 6: 37756.

Phillips, M., W. J. Tang, M. Robinson, D. O. Daza, K. Hassan *et al.*, 2020 Evidence of chitin in the ampullae of Lorenzini of chondrichthyan fishes. Curr. Biol. 30: R1254–R1255.

Renkema, G. H., R. G. Boot, A. O. Muijsers, W. E. Donker-Koopman, and J. M. Aerts, 1995 Purification and characterization of human chitotriosidase, a novel member of the chitinase family of proteins. J. Biol. Chem. 270: 2198–2202.

Ritchie, M. E., B. Phipson, D. Wu, Y. Hu, C. W. Law *et al.*, 2015 limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 43: e47.

Robledo, D., A. P. Gutiérrez, A. Barría, J. M. Yáñez, and R. D. Houston, 2018 Gene expression response to sea lice in atlantic salmon skin: RNA sequencing comparison between resistant and susceptible animals. Front. Genet. 9: 287.

Sahlmann, C., J. Gu, T. M. Kortner, I. Lein, Å. Krogdahl *et al.*, 2015 Ontogeny of the Digestive System of Atlantic Salmon (Salmo salar L.) and Effects of Soybean Meal from Start-Feeding. PLoS ONE 10: e0124179.

van der Sluis, M., A. Vincent, J. Bouma, A. Korteland-Van Male, J. B. van Goudoever *et al.*, 2008 Forkhead box transcription factors Foxa1 and Foxa2 are important regulators of Muc2 mucin expression in intestinal epithelial cells. Biochem. Biophys. Res. Commun. 369: 1108–1113.

Stone, E. L., M. N. Ismail, S. H. Lee, Y. Luu, K. Ramirez *et al.*, 2009 Glycosyltransferase function in core 2-type protein O glycosylation. Mol. Cell. Biol. 29: 3770–3782.

Strobel, S., A. Roswag, N. I. Becker, T. E. Trenczek, and J. A. Encarnação, 2013 Insectivorous bats digest chitin in the stomach using acidic mammalian chitinase. PLoS ONE 8: e72770.

Sutherland, B. J. G., S. G. Jantzen, D. S. Sanderson, B. F. Koop, and S. R. M. Jones, 2011 Differentiating size-dependent responses of juvenile pink salmon (Oncorhynchus gorbuscha) to sea lice (Lepeophtheirus salmonis) infections. Comp. Biochem. Physiol. Part D Genomics Proteomics 6: 213–223.

Tabata, E., A. Kashimura, S. Wakita, M. Ohno, M. Sakaguchi *et al.*, 2017 Gastric and intestinal proteases resistance of chicken acidic chitinase nominates chitin-containing organisms for alternative whole edible diets for poultry. Sci. Rep. 7: 6662.

Tabata, E., A. Kashimura, S. Wakita, M. Ohno, M. Sakaguchi *et al.*, 2017 Protease resistance of porcine acidic mammalian chitinase under gastrointestinal conditions implies that chitin-containing organisms can be sustainable dietary resources. Sci. Rep. 7: 12963.

Tabata, E., S. Wakita, A. Kashimura, Y. Sugahara, V. Matoska *et al.*, 2019 Residues of acidic chitinase cause chitinolytic activity degrading chitosan in porcine pepsin preparations. Sci. Rep. 9: 15609.

Tang, W. J., J. Fernandez, J. J. Sohn, and C. T. Amemiya, 2015 Chitin is endogenously produced in vertebrates. Curr. Biol. 25: 897–900.

Teng, Z., C. Sun, S. Liu, H. Wang, and S. Zhang, 2014 Functional characterization of chitinase-3 reveals involvement of chitinases in early embryo immunity in zebrafish. Dev. Comp. Immunol. 46: 489–498.

Terwisscha van Scheltinga, A. C., M. Hennig, and B. W. Dijkstra, 1996 The 1.8 A resolution structure of hevamine, a plant chitinase/lysozyme, and analysis of the conserved sequence and structure motifs of glycosyl hydrolase family 18. J. Mol. Biol. 262: 243–257.

Wagner, G. P., J. Lo, R. Laine, and M. Almeder, 1993 Chitin in the epidermal cuticle of a vertebrate (Paralipophrys trigloides, Blenniidae, Teleostei). Experientia 49: 317–319.

Zhang, Y., S. Feng, J. Chen, C. Qin, H. Lin *et al.*, 2012 Stimulatory effects of chitinase on growth and immune defense of orange-spotted grouper (Epinephelus coioides). Fish Shellfish Immunol. 32: 844–854.

Zhang, H., M. Liu, Y. Tian, and X. Hu, 2011 Comparative characterization of chitinases from silkworm (Bombyx mori) and bollworm (Helicoverpa armigera). Cell Biochem. Biophys. 61: 267–275.

Zhang, H., T. Yohe, L. Huang, S. Entwistle, P. Wu *et al.*, 2018 dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. Nucleic Acids Res. 46: W95–W101.

Zhang, J., X. Zhang, Y. Arakane, S. Muthukrishnan, K. J. Kramer *et al.*, 2011 Comparative genomic analysis of chitinase and chitinase-like genes in the African malaria mosquito (Anopheles gambiae). PLoS ONE 6: e19899.

Zhu, Q., Y. Arakane, R. W. Beeman, K. J. Kramer, and S. Muthukrishnan, 2008 Functional specialization among insect chitinase family genes revealed by RNA interference. Proc Natl Acad Sci USA 105: 6650–6655.

Paper II

The effect of dietary chitin on Atlantic salmon (*Salmo salar*) chitinase activity, gene expression, and microbial composition

Matilde Mengkrog Holen[1], Simen Rød Sandve[1], Thomas Nelson Harvey[1], Yang Jin[1], Inga Leena Angell[2], Knut Rudi[2], Matthew Peter Kent[1]

[1] Centre for Integrative Genetics (CIGENE), Department of Animal and Aquacultural Sciences, Faculty of Biosciences, Norwegian University of Life Sciences, NO-1432 Ås, Norway

[2] Department of Chemistry, Biotechnology and Food Science (IKBM), Norwegian University of Life Sciences, NO-1432 Ås, Norway

Corresponding author: Matthew Peter Kent, matthew.peter.kent@nmbu.no, +4747362026

Abstract

Background: Chitin is a common component in the natural diet of many fish, and a range of chitinases with the potential to down chitin have been identified. Yet whether chitin is metabolized in fish is still unclear. Here we used a combination of chitinase activity assay, transcriptomics, and 16S rRNA bacterial analysis to assess the effect of chitin supplementation on Atlantic salmon gene expression and microbial community.

Results: Atlantic salmon express multiple genes associated with chitin metabolism, and we show that the expression and activity of Atlantic salmon chitinases are not affected by the addition of dietary chitin. We do, however, demonstrate an association between gut microbial composition, chitinase activity in the gut, and host chitinase expression.

Conclusion: The findings presented here support the idea that chitin metabolism genes are linked to the maintenance of a chitin-based barrier in the teleost gut. These results contribute to a greater understanding of chitin metabolism in fish.

Introduction

Chitin is a tough and insoluble polymer consisting of β-1,4-linked *N*-acetyl-D-glucosamine (GlcNAc) residues that serves as a protective and structural component in arthropod exoskeletons such as insects and crustaceans (Kramer *et al.* 1995; Kurita 2006). This complex carbohydrate is commonly found in the natural diet of many fish, including Atlantic salmon (*Salmo salar*), and chitinous organisms like Antarctic krill (*Euphausia superba*) and black soldier fly (*Hermetia illucens*) are considered two sustainable feed sources for Atlantic salmon farming (Hansen *et al.* 2010; Belghit *et al.* 2019). Supplementation of chitin in Atlantic salmon feed has been shown to promote the growth of potentially beneficial gut microbes that *in vitro* inhibit the growth of common fish pathogens (Askarian *et al.* 2012). However, some studies have reported a reduced growth rate in salmon fed chitin-rich diets and it has been hypothesized that chitin acts as an energy sink when fish are not able to digest and utilize this polysaccharide properly (Karlsen *et al.* 2017; Renna *et al.* 2017).

Chitinase and chitobiase are the enzymes responsible for the hydrolysis of chitin and the production of GlcNAc monomers. Their activity has been detected in the digestive tract of a variety of fish species (Jeuniaux 1961; Fänge *et al.* 1979; Lindsay *et al.* 1984;

Lindsay 1984; Kono *et al.* 1987; Gutowska *et al.* 2004; Krogdahl *et al.* 2005; Abro *et al.* 2014), including salmonids such as rainbow trout (*Oncorhynchus mykiss*) (Lindsay *et al.* 1984), and a range of fish chitinases have been identified and biochemically characterized (Ikeda *et al.* 2009, 2012, 2013, 2017; Zhang *et al.* 2012; Koch *et al.* 2014; Teng *et al.* 2014; Kakizaki *et al.* 2015; Kawashima *et al.* 2016; Pohls *et al.* 2016; Gao *et al.* 2017). Most of these enzymes are detected in the gastrointestinal tract and show acid-resistant activities toward insoluble chitin substrates, indicating that they could have the ability to digest chitin (Ikeda *et al.* 2017). As far as we know, no one has characterized chitinase proteins in Atlantic salmon.

Another family of enzymes involved in chitin metabolism is chitin synthases (CHS), which synthesize chitin from GlcNAc. Since chitin is present in their scales (Tang *et al.* 2015), it is hypothesized that salmon express CHS and that these are active, however very little is known about the role of these enzymes in metabolic processes. In this study we investigated the effect of chitin supplementation on the expression of host chitinases, chitobiase, and CHS as well as on the microbiome of Atlantic salmon.

Materials and methods

Feeding trial

The fish used in this experiment were Atlantic salmon post-smolts obtained as fry from AquaGen Breeding Centre, Kyrksæterøra, and reared in fresh water at the Centre for Fish Research, NMBU. The fish (n = 32) were acclimatized to experimental conditions for 41 days before the trial and were fed a standard commercial diet. Groups of six fish were placed in each tank with an average weight of 812 ± 113 g and diets were tested in two replicate tanks. The fish were fed one of three different diets (Table 1) over 29 days through an automatic feeding system. Starting with a standard core feed composition, Diet 1 – "control" was supplemented with 6 % cellulose bought from FôrTek, NMBU, Diet 2 – "shrimp shell" was supplemented with 6 % shrimp (*Pandalus borealis*) shell chitin bought from Primex (ChitoClear Chitin, Iceland) and in Diet 3, a portion of the fish meal was substituted with insect meal from defatted black soldier fly larvae (*Hermetica illucens*) obtained from Protix Biosystems BV (Dongen, Netherlands) to a final concentration of approximately 6% chitin. After 29 days fish were euthanized by a blow to the head in accordance with the national regulations of animal (Dyrevelferdsloven 2015). Tissue samples and the contents

from the stomach and pyloric caeca were collected separately, flash-frozen in liquid nitrogen, and stored at -80 °C until analysis of gene expression and chitinase activity. Contents from the distal intestine (DI), the most distal compartment of the gut, were sampled for 16S rRNA sequencing as previously described (Rudi *et al.* 2018).

Table 1. Ingredients and proximal composition of the experimental diets.

	Experimental diets		
	Control	*Shrimp shell*	*Fly larvae*
Ingredients (g kg^{-1})			
Cellulose	**4.7**	0.0	0.0
Fish meal	32.2	32.2	9.1
Wheat gluten	9.4	9.4	9.4
Pregel potato starch	8.6	8.6	8.6
Defatted black soldier fly larvae meal	0.0	0.0	**32.4**
Shrimp shell chitin	0.0	**4.7**	0.0
Gelatine	5.8	5.8	5.8
Choline chloride	0.1	0.1	0.1
Fish oil	10.8	10.8	6.1
Vitamin/mineral mixture	0.5	0.5	0.5
Proximate composition (%)			
Ash	9.17	8.79	6.38
Moisture	6.91	6.58	4.71
Lipid	19.64	19.64	19.25
Crude protein (N · 6.25)	46.9	46.9	47.36
Chitin	**0**	**6.11**	**6.35***
Cellulose	6.11	0	0

***** This is a proximate measure using the following formula to calculate chitin in insect meal: chitin (%) = ash-free ADF (%) - ADIP (%) as previously done by Marono S. *et al* (2015) (Marono *et al.* 2015)

RNA-sequencing

Total RNA was extracted from the stomach and pyloric caeca tissue samples stored at -80 °C (n = 4 for control and fly larvae, n = 5 for shrimp shell) using the RNeasy Plus Universal Kit (QIAGEN). RNA quality was assessed using a 2100 Bioanalyzer with the RNA 6000 nano kit (Agilent) and the concentration was measured using a Nanodrop 8000 spectrophotometer (Thermo Fisher Scientific). Extracted RNA with RNA integrity number (RIN) ≥7 was used as input for the TruSeq Stranded mRNA HT

Sample Prep Kit (Illumina) according to the manufacturer's recommendations. Mean library length was measured on 2100 Bioanalyzer using the DNA 1000 kit (Agilent) and the library concentration was quantified with the Qbit Broad Range kit (Thermo Fisher Scientific). Single-end sequencing (100bp reads) was performed at the Norwegian Sequencing Center (Oslo, Norway) using a HiSeq4000 instrument (Illumina).

Gene expression

Trimming, mapping, and counting of reads were done using the bcbio-nextgen pipeline (https://github.com/bcbio/bcbio-nextgen). The sequencing reads were aligned to the Atlantic salmon genome (ICSASG_v2) (Lien *et al.* 2016) using the STAR aligner with default settings (Dobin *et al.* 2013) after adapter trimming. FeatureCounts was then used to count reads aligned to genes (Liao *et al.* 2014). Raw gene counts were transformed to transcripts per million (TPM) values to normalize for gene length before comparison of chitinase gene expression levels (Welch's t-test, $\alpha = 0.05$). Gene expression values were normalized by library size (see TMM normalization in EdgeR user guide (Robinson and Oshlack 2010)) before differential expression analysis. Based on PCA analysis, two samples (X8.CFE.7.F4.PC and X9.CFE.13.F1.PC) showed unusual gene expression patterns and were removed from further analysis of pyloric caeca for fish fed control and fly larvae. All statistical analysis was done in R (v.3.6.0).

Differential expression analysis

Lowly expressed genes ($\log_2(TPM + 1)$ values < 1) were filtered out prior to differential expression analysis. The analysis was carried out using the standard EdgeR protocol (Robinson *et al.* 2010) where an exact test of expression values between the experimental diet (shrimp shell or fly larvae) and control diet gave a log2-fold change, p-value, and false discovery rate (FDR) for each gene. Genes with an FDR < 0.05 and absolute log2 fold change (logFC) > 0.5 were defined as differentially expressed.

Gene enrichment analysis

Gene enrichment analysis of the differential expressed genes was performed using KEGG (Kyoto Encyclopedia of Genes and Genomes) and GO (Gene Ontology). KEGG analysis was carried out using the "kegga" function in the limma-package (Ritchie *et al.* 2015) with the argument "species.KEGG = "sasa"" and p-values < 0.05. GO analysis was carried out using the package GOstats (Falcon and Gentleman 2007), following the steps outlined here (https://www.bioconductor.org/packages/release/bioc/vignettes/GOstats/inst/doc/GOstatsForUnsupportedOrganisms.pdf), using the argument "ontology = "BP"" (biological processes) and Bonferroni adjusted p-values (q) < 0.05.

Chitinase activity test

We measured the chitinase activity of fish analyzed for gene expression that contained stomach and pyloric caeca content (n = 3) as described previously (Ohno *et al.* 2013). Approximately 200 mg of contents were homogenized using TissueLyser II (QIAGEN) in 900 uL of 50 mM sodium acetate (pH = 5.5) containing 1X Halt protease inhibitor cocktail (Thermo Fisher Scientific). The samples were centrifuged at 14.000 g for 20 min at 4°C to pellet particulates, and the supernatant was collected for protein quantification and chitinase activity. Protein concentration was measured using Bradford protein assay (Quick Start™ Bradford Assay, BioRad) with BSA (Bovine Serum Albumin) used as standard, according to manufacturer's instructions. Chitinase activity was measured by monitoring the hydrolysis of 4-Methylumbelliferyl β-D-N,N′,N′′-triacetylchitotrioside (Sigma-Aldrich, M5639), a fluorogenic chitin substrate suitable for measuring endochitinase activity, according to the Fluorimetric Chitinase Assay Kit (Sigma-Aldrich, CS1030) with the following modifications. Briefly, 10 μg of pyloric caeca proteins and 1 μg of stomach proteins were incubated with 200 μM 4-Methylumbelliferyl β-D-N,N′,N′′-triacetylchitotrioside (4-MU-$GlcNAc_3$) in McIlvaine's buffer (0.1 M citric acid and 0.2 M Na_2HPO_4, pH 6) in a volume of 100 μL at 28 °C. The pH and temperature were determined from pilot experiments on a recombinant Atlantic salmon chitinase (rChia.8, unpublished) where pH 6 and 28 °C were optimal conditions for activity of this enzyme which is expressed in both stomach and pyloric caeca of Atlantic salmon. The reaction was terminated after 30 min by adding 400 mM sodium carbonate and the fluorescence of the released 4-Methylumbelliferone (4-MU) was measured at an excitation wavelength of 360 nm and an emission wavelength of 450 nm using a

SpectraMax M2 plate reader (Molecular Devices) no later than 30 minutes after stopping the reaction. The assay was performed in triplicates and a 4-MU standard curve was used to quantify 4-MU resulting from the hydrolytic reaction. The measured fluorescence was corrected for hydrolysis of the substrate without enzyme added. The chitinase activity was expressed as unit/mg of total protein in the sample where 1 unit is defined as 1 umol of 4-MU formed per minute. Statistical comparisons were done using Student's t-test ($\alpha = 0.05$ and $\alpha = 0.1$).

Illumina sequencing of the 16S rRNA gene

DNA was extracted from ground samples of Diet 1-3 fish feed (n=2 per feed) and the contents collected from the distal intestine (DI) from all fish (n=12 per feed) using the Mag midi kit (LGC Genomics, UK) following the manufactures instructions. Preparation of amplicon library using the primer pair 341F/806R (Yu *et al.* 2005) and sequencing was done as previously described (Rudi *et al.* 2018). The resulting reads were processed as previously described (Angell *et al.* 2020) using a sequence depth of 10,000 sequences per sample.

Results

Impact of dietary chitin on chitinase activity and expression

We performed an *in vitro* quantification of chitinase activity from crude extracts of total soluble material collected from the stomach and pyloric caeca of fish fed one of three diets differing in chitin content. Chitinase activity relative to the amount of protein was consistently much higher (ranging from 5 to11-fold difference) in the stomach than in pyloric caeca irrespective of diets. Chitinase activity was unaffected by the inclusion of dietary chitin when compared to control (Student's t-test, $p > 0.05$; Figure 1). However, a trend towards lower activity in fish fed fly larvae compared to the fish fed control and shrimp shell was observed (Student's t-test, $p < 0.1$).

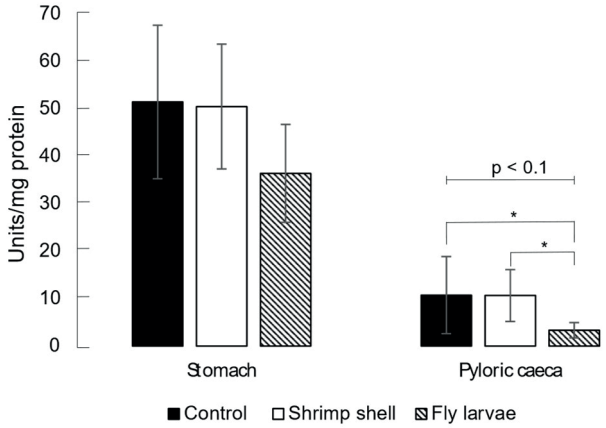

Figure 1. Chitinase activity (units/mg protein) in the stomach- and pyloric caeca contents of fish (n=3) fed control, shrimp shell, and fly larvae using 4-MU-GlcNAc$_3$. Crude soluble protein solutions from stomach (1 μg per reaction) and pyloric caeca (10 μg per reaction) was incubated with 200 μM 4-MU-GlcNAc$_3$ for 30 minutes at 28 °C in McIlvaine's buffer, pH 6. 1 unit is defined as 1 μmol of 4-MU formed per minute. The values shown are means of triplicates ± SD.

According to the NCBI Salmo Salar RefSeq annotation (GCF_000233375.1; release 100), Atlantic salmon codes for 10 chitinase-like genes from the glycosyl hydrolase family 18 (hereafter named *chia.1-10*; see Supplementary Table 1 for the corresponding gene IDs), a single chitobiase-like gene from the glycosyl hydrolase family 18 (hereafter named *ctbs*), and four genes with putative chitin-synthase domains (hereafter named *chs1a*, *chs1b*, *chs2,* and *chs3*). Eight chitinase genes and two CHS genes showed tissue-specific expression (stomach; *chia.3*, *4, 5* and *7*, pyloric caeca; *chia.1*, *2*, *6*, *9* and *10, chs1a* and *chs1b*) with *chia.8* and *ctbs* being ubiquitously expressed in both tissues (Figure 2A). Notably, three of the chitinase genes expressed in the stomach (*chia.3*, *4,* and *7*) were among the most abundant transcripts present in the tissue, whereas *chia.1*, *chia.2,* and *chia.6* were among the most abundant transcripts in pyloric caeca. Two CHS genes, *chs2* and *chs3* were not expressed in any of the two tissues. To assess the response of chitinase-, chitobiase- and CHS genes to the inclusion of dietary chitin, we compared their expression levels in the stomach (Figure 2B) and pyloric caeca (Figure 2C). One chitinase, *chia.5*, was lowly expressed ($\log_2$(TPM) < 2) and is therefore not shown in Figure 2B. The inclusion of dietary chitin had no significant effect on chitinase, chitobiase, and CHS gene expression (Welch's t-test, $p > 0.05$), although the expression of *chia.1, chia.2,* and *chia.6* in pyloric caeca appeared to be slightly lower in fish fed fly larvae ($p < 0.13$).

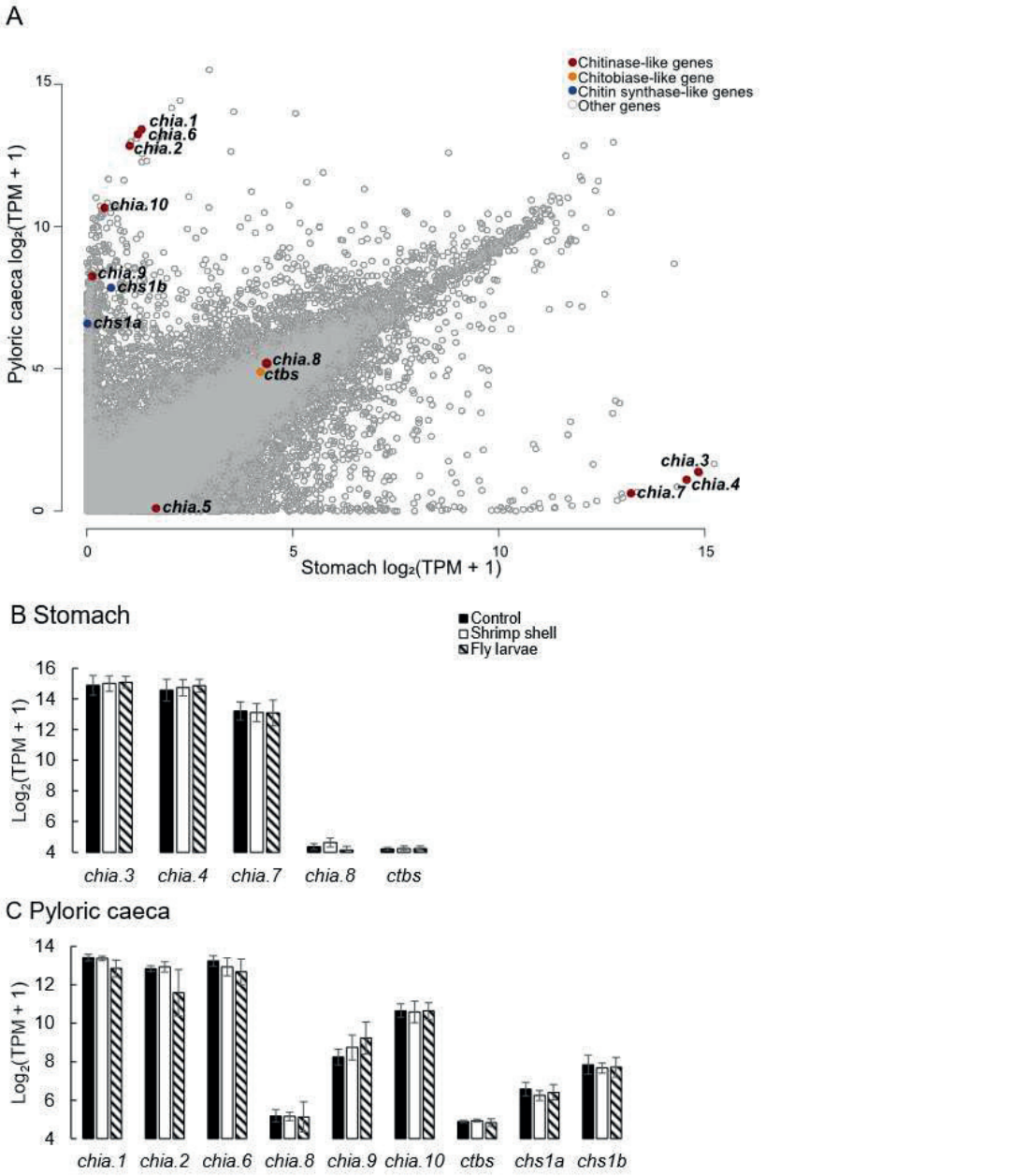

Figure 2. A) Gene expression levels (log_2(TPM + 1) of chitinase-, chitobiase- and CHS genes compared to the expression levels of all genes in the stomach and pyloric caeca for fish fed the control diet (n=4 for stomach and n=3 for pyloric caeca). B) Gene expression levels (log_2(TPM + 1) of chitinase and chitobiase genes in the stomach (n=4 for control and fly larvae, n=5 for shrimp shell). C) Gene expression levels (log_2(TPM + 1) of chitinase and CHS genes in pyloric caeca (n=3 for control and fly larvae, n=5 for shrimp shell). Please note that the y-axis does not extend to 0.

Gene enrichment analysis

In addition to focusing on genes related to chitin metabolism, we examined how the transcriptome more generally responded to the inclusion of dietary chitin as a replacement to cellulose. Beginning with the stomach, feeding fish a diet including fly larvae had very little effect on the transcriptome compared to the control diet with only one differentially expressed gene (DEG), namely dual-specificity protein phosphatase 1-like (*dusp1*). In contrast, a shrimp shell containing diet provoked a larger effect with 889 upregulated and 570 downregulated genes. The most enriched GO (biological process) and KEGG terms among the upregulated genes were processes and pathways involved in cell organization including muscle structure development, focal adhesion, and extracellular matrix (ECM) receptor interactions (Table 2).

Table 2. Top GO terms and KEGG pathways of upregulated (up) and downregulated (down) DEGs in the stomach of fish fed shrimp shells. The rich factor is the ratio of DEG number and the total number of genes annotated to this pathway.

Enriched term	DEGs	Rich factor	P-value	Method	Mode
Muscle structure development	184	0.108	7.69E-35	GO_BP	Up
Anatomical structure morphogenesis	396	0.063	3.18E-27	GO_BP	Up
Cell adhesion	189	0.087	5.58E-25	GO_BP	Up
Focal adhesion	41	0.064	1.25E-12	KEGG	Up
ECM-receptor interaction	19	0.067	8.75E-07	KEGG	Up
Regulation of actin cytoskeleton	30	0.048	9.88E-07	KEGG	Up
Synaptic vesicle recycling	19	0.116	2.91E-05	GO_BP	Down
Glycoprotein biosynthetic process	38	0.066	3.29E-05	GO_BP	Down
Protein glycosylation	30	0.071	5.77E-05	GO_BP	Down
UDP-*N*-acetylglucosamine biosynthetic process	5	0.42	3.45E-04	GO_BP	Down
Metabolic pathways	80	0.025	4.59E-10	KEGG	Down
Amino sugar and nucleotide sugar metabolism	11	0.085	1.20E-06	KEGG	Down
Glutathione metabolism	9	0.084	1.20E-05	KEGG	Down
Mucin type O-glycan biosynthesis	8	0.09	2.28E-05	KEGG	Down

Among the downregulated genes in the stomach, the most enriched processes and pathways were those where glycosylation plays a central role, including synaptic vesicle recycling, glycoprotein biosynthetic processes, and amino sugar and nucleotide sugar metabolism. Notably, several of the downregulated genes are central in the pathway whereby fructose-6-phosphate is converted to uridine diphosphate *N*-acetylglucosamine (UDP-*N*-acetylglucosamine biosynthetic process), the substrate molecule used by CHS to synthesize chitin (Figure 3). This includes glutamine-fructose-6-phosphate aminotransferase 2-like (*gfpt2*), glucosamine-phosphate *N*-acetyltransferase 1 (*gnpnat1*) and UDP-*N*-acetylhexosamine pyrophosphorylase-like (*uap1*).

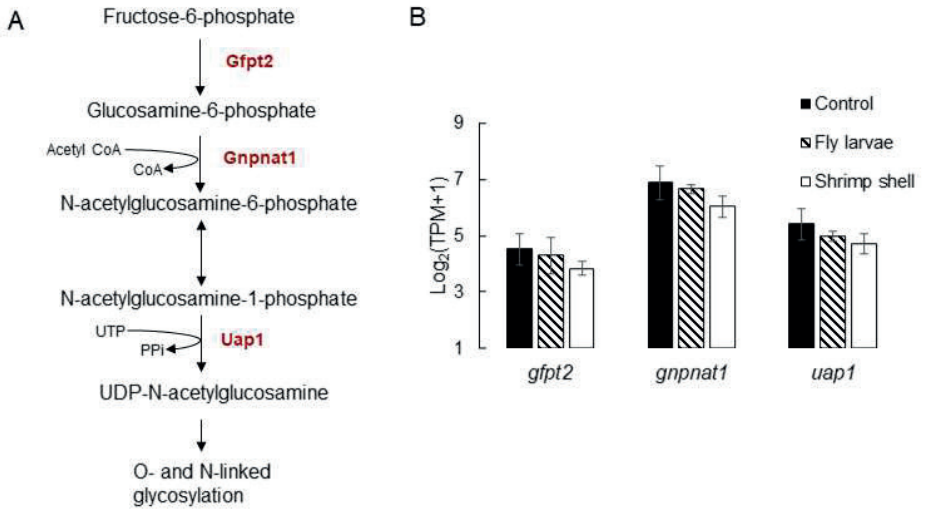

Figure 3. A) The hexosamine biosynthetic pathway leading to the production of UDP-*N*-acetylglucosamine (UDP-GlcNAc) where central enzymes discussed in the text are marked in red. B) Gene expression levels (log_2(TPM + 1) of differentially expressed hexosamine biosynthetic pathway genes in the stomach (n=4 for control and fly larvae, n=5 for shrimp shell). Please note that the y-axis does not extend to 0.

When analyzing data from the pyloric caeca only two genes (endonuclease domain-containing 1 protein-like; *endd1*, and ribonuclease P protein subunit p30-like; *rpp30*) were differentially expressed when replacing cellulose with shrimp shell chitin, whereas 53 genes were upregulated and eight downregulated in response to feeding fly larvae. GO enrichment of the 53 upregulated genes revealed that many of these genes (n=22) are involved in the biosynthesis of cholesterol including HMG-CoA

reductase (*hmgr*) which is responsible for the rate-limiting step in the cholesterol biosynthesis pathway. No common enriched term was found for the downregulated genes, but the 8 genes detected included acid-sensing (proton gated) ion channel 1 (*asic1*), collagen alpha-1(XXIV) chain-like (*col24a1*), solute carrier family 25 member 48-like (*slc25a48*), peptidase inhibitor R3HDML-like (*r3hdml*), lecithin retinol acyltransferase-like (*lrat*) and three uncharacterized genes.

The bacterial composition of distal intestine contents

To investigate the impact of dietary chitin on the microbial community, we analyzed the bacterial metapopulation of feed and DI using 16S rRNA gene sequencing. We observed a clear relationship whereby bacteria present in the raw feedstuff were also present in the DI (Figure 4). The microbial profiles of DI contents from fish fed control and shrimp shell diets were almost identical, at the genus level the taxa dominating the distal intestinal contents of fish fed control and shrimp shell were *Lactobacillus* (45.7% and 50.7%), *Streptococcus* (9.4% and 9.3%) and *Weissella* (6.4% and 6.5%). For fish fed fly larvae, the dominant taxa were *Actinomyces* (28.2%), *Bacillus* (21.6%), and *Enterococcus* (16.9%). All prevalent taxa in distal intestinal contents were also present in the feed samples, although the relative amount of *Actinomyces* in fish fed fly larvae was substantially higher than in the feed (28.2% vs 2.18%).

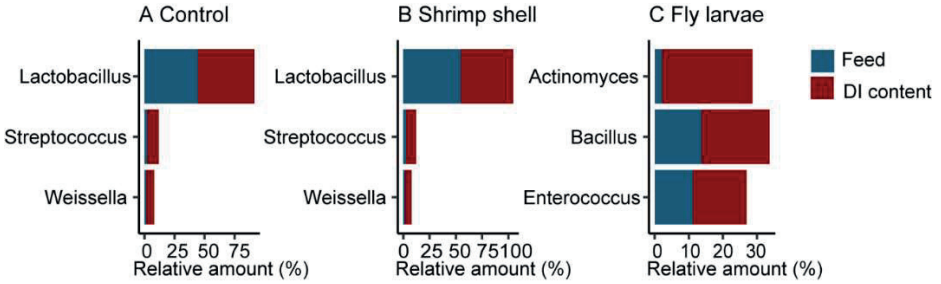

Figure 4. The three most abundant bacteria found in distal intestine (DI content) and fish feed as assessed by 16S rRNA gene sequencing of fish fed A. Control diet (mean value of n=11 for DI content and n=2 for feed), B. Shrimp shell diet (mean value of n=9 for DI content and n=2 for feed), and C. Fly larvae diet (mean value of n=10 for DI content and n=2 for feed).

Discussion

Transcriptome analyses revealed that the stomach and pyloric caeca react differently to the diets given in this trial. The largest changes in terms of DEGs were observed in the stomach when fish were fed a diet supplemented with chitin from shrimp shells. GO and KEGG analysis lead us to conclude that these changes involve the upregulation of genes involved in tissue organization and extracellular matrix-cell interactions and are linked to structural changes in the gastric tissue. The mucosal barrier that covers the gastrointestinal tract of Atlantic salmon consists of highly glycosylated mucins (Jin *et al.* 2015) and the downregulation of genes involved in glycosylation of these proteins further implies a structural change or a response to the extracellular

environment. We hypothesize that this change is partly caused by chitin degradation, with the release of products such as GlcNAc as the result of the activity of highly expressed chitinases in the stomach (Figure 2A-B). In line with what we observe, higher GlcNAc concentrations are associated with the downregulation of genes involved in the biosynthesis of UDP-GlcNAc (Figure 3). UDP-GlcNAc is a substrate required for O – and N-glycosylation of proteins, including mucins that are heavily glycosylated. In mice, orally administered GlcNAc was shown to enter the hexosamine biosynthetic pathway and increase the abundance of UDP-GlcNAc (Ryczko *et al.* 2016). UDP-GlcNAc has shown to act as an end-product inhibitor of this pathway (Chiaradonna *et al.* 2018) and regulates the gene expression of *gfpt2*, which converts fructose-6-phosphate and glutamine to glucosamine-6-phosphate and glutamate, eventually modulating glycosylation homeostasis.

There is evidence for the presence of chitinous structures surrounding the mucosal barrier of ray-finned fish (Tang *et al.* 2015; Nakashima *et al.* 2018) and the results presented here are consistent with the evolutionary conservation of host chitinases and chitobiase to participate in the remodeling of these structures. Previous studies show that Atlantic salmon is not able to utilize chitin to a significant extent (Olsen *et al.* 2006). In line with our results, increasing the dietary chitin content has previously been shown not to correlate with increased chitinase activity in fish (Lindsay *et al.* 1984; Kono *et al.* 1987; Abro *et al.* 2014), and it seems as if the chitinase activity is always present independent of dietary chitin. There could be two possible reasons for this: 1) gut chitinase activity is not regulated by the addition of chitin because Atlantic salmon is exposed to a relatively constant supply of chitin during its life cycle, and/or 2) the chitinase and chitobiase genes are constitutively expressed because of their role as chitin remodelers in the intestinal mucosa. Relatively high CHS gene expression levels in the same intestinal segments of Atlantic salmon as chitinases and chitobiase genes are expressed favor the second hypothesis.

Compared to stomach, few DEGs were detected in pyloric caeca with the greatest changes observed in fish fed fly larvae; this included the significant upregulation of cholesterol biosynthetic genes. Such an upregulation could be expected as a response to lower cholesterol levels in the feed, as insect lipids usually contain low cholesterol levels, but a substantial amount of phytol sterol (Secci *et al.* 2018). This has previously been shown to induce the cholesterol biosynthetic pathway in the pyloric caeca of Atlantic salmon (Jin *et al.* 2018).

Chitinase activity in the stomach and pyloric caeca contents of Atlantic salmon did not seem to be significantly affected by the addition of dietary chitin, but fish fed fly larvae

had slightly lower chitinase activity in pyloric caeca contents than fish fed control and shrimp shell diets (Figure 1). This is in line with the slight decrease in gene expression levels of three chitinase genes of fish fed fly larvae; *chia.1*, *chia.2*, and *chia.6,* all being exclusively expressed in pyloric caeca (Figure 2C). In general, the transcriptome of Atlantic salmon did not seem to respond strongly to changing the standard commercial diet with fly larvae diet, but the bacterial composition of fish fed fly larvae was different from the composition of fish fed control and shrimp shell diets (Figure 4). In accord with our findings, an increase in the relative abundance of *Actinomyces*, *Bacillus,* and *Enterococcus* when Atlantic salmon is fed fly larvae have previously been reported (Li *et al.* 2021). Since *Actinomyces* and *Bacillus* are potential chitin degraders (Askarian *et al.* 2012; Beier and Bertilsson 2013; Wang *et al.* 2018), we hypothesize that the observed decrease in Atlantic salmon chitinase gene expression levels and activity under fly larvae diet is an effect of the activity of bacterial chitinases, reducing the level of dietary and host-derived chitin available as a substrate to endogenous chitinases in the gastrointestinal tract.

Conclusion

We show that the stomach and pyloric caeca transcriptome of Atlantic salmon did not respond to a great extent to the presence of dietary chitin, in support of the idea that evolutionary conservation of host chitinases is mostly linked to remodeling of chitin as a structural element in the gut lining (Nakashima *et al.* 2018). Furthermore, we demonstrate an association between gut microbial composition, chitin activity in the gut, and host chitinase gene expression, and hypothesize functional interconnection between chitinase-secreting gut bacteria (e.g. *Actinomyces*) and chitinase gene regulation in the host. These results contribute to a greater understanding of chitin metabolism in fish in general.

Data availability statement

Raw RNA-seq data are available on ArrayExpress with the accession number E-MTAB-10594. The raw 16S sequence data are available at NCBI's Sequence Read Archive (SRA) with accession number PRJNA820557.

Welfare statement

The experiment was conducted in accordance with Norwegian and European regulations related to animal research. Formal approval of the experiment by the Norwegian Animal Research Authority (NARA) was not required as the experimental conditions were following routine practices at the Centre for Fish Research at NMBU and no compromised welfare was expected.

Acknowledgements

We want to acknowledge Protix Biosystems BV for providing us with insect meal from defatted black soldier fly larvae (*Hermetica illucens*) and Dr. Jon Øvrum Hansen for helping with feed composition calculations and feed pellet processing.

References

Abro, R., K. Sundell, E. Sandblom, H. Sundh, E. Brännäs *et al.*, 2014 Evaluation of chitinolytic activities and membrane integrity in gut tissues of Arctic charr (Salvelinus alpinus) fed fish meal and zygomycete biomass. Comp Biochem Physiol B, Biochem Mol Biol 175: 1–8.

Angell, I. L., L. Bergaust, J. F. Hanssen, E. M. Aasen, and K. Rudi, 2020 Ecological Processes Affecting Long-Term Eukaryote and Prokaryote Biofilm Persistence in Nitrogen Removal from Sewage. Genes (Basel) 11:.

Askarian, F., Z. Zhou, R. E. Olsen, S. Sperstad, and E. Ringø, 2012 Culturable autochthonous gut bacteria in Atlantic salmon (Salmo salar L.) fed diets with or without chitin. Characterization by 16S rRNA gene sequencing, ability to produce enzymes and in vitro growth inhibition of four fish pathogens. Aquaculture 326–329: 1–8.

Beier, S., and S. Bertilsson, 2013 Bacterial chitin degradation-mechanisms and ecophysiological strategies. Front. Microbiol. 4: 149.

Belghit, I., N. S. Liland, P. Gjesdal, I. Biancarosa, E. Menchetti *et al.*, 2019 Black soldier fly larvae meal can replace fish meal in diets of sea-water phase Atlantic salmon (Salmo salar). Aquaculture 503: 609–619.

Chiaradonna, F., F. Ricciardiello, and R. Palorini, 2018 The Nutrient-Sensing Hexosamine Biosynthetic Pathway as the Hub of Cancer Metabolic Rewiring. Cells 7:.

Dobin, A., C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski *et al.*, 2013 STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29: 15–21.

Dyrevelferdsloven, 2015 *Forskrift om bruk av dyr i forsøk, § 16*.

Falcon, S., and R. Gentleman, 2007 Using GOstats to test gene lists for GO term association. Bioinformatics 23: 257–258.

Fänge, R., G. Lundblad, J. Lind, and K. Slettengren, 1979 Chitinolytic enzymes in the digestive system of marine fishes. Mar. Biol. 53: 317–321.

Gao, C., X. Cai, Y. Zhang, B. Su, H. Song *et al.*, 2017 Characterization and expression analysis of chitinase genes (CHIT1, CHIT2 and CHIT3) in turbot (Scophthalmus maximus L.) following bacterial challenge. Fish Shellfish Immunol. 64: 357–366.

Gutowska, M. A., J. C. Drazen, and B. H. Robison, 2004 Digestive chitinolytic activity in marine fishes of Monterey Bay, California. Comp Biochem Physiol, Part A Mol Integr Physiol 139: 351–358.

Hansen, J. Ø., M. Penn, M. Øverland, K. D. Shearer, Å. Krogdahl *et al.*, 2010 High inclusion of partially deshelled and whole krill meals in diets for Atlantic salmon (Salmo salar). Aquaculture 310: 164–172.

Ikeda, M., H. Kakizaki, and M. Matsumiya, 2017 Biochemistry of fish stomach chitinase. Int. J. Biol. Macromol. 104: 1672–1681.

Ikeda, M., Y. Kondo, and M. Matsumiya, 2013 Purification, characterization, and molecular cloning of chitinases from the stomach of the threeline grunt Parapristipoma trilineatum. Process Biochemistry 48: 1324–1334.

Ikeda, M., K. Miyauchi, and M. Matsumiya, 2012 Purification and characterization of a 56 kDa chitinase isozyme (PaChiB) from the stomach of the silver croaker, Pennahia argentatus. Biosci. Biotechnol. Biochem. 76: 971–979.

Ikeda, M., K. Miyauchi, A. Mochizuki, and M. Matsumiya, 2009 Purification and characterization of chitinase from the stomach of silver croaker Pennahia argentatus. Protein Expr. Purif. 65: 214–222.

Jeuniaux, C., 1961 Chitinase: an addition to the list of hydrolases in the digestive tract of vertebrates. Nature 192: 135–6.

Jin, Y., R. E. Olsen, G. B. Gillard, M.-A. Østensen, S. A. Korsvoll *et al.*, 2018 A systemic study of lipid metabolism regulation in salmon fingerlings and early juveniles fed plant oil. Br. J. Nutr. 120: 653–664.

Jin, C., J. T. Padra, K. Sundell, H. Sundh, N. G. Karlsson *et al.*, 2015 Atlantic Salmon Carries a Range of Novel O-Glycan Structures Differentially Localized on Skin and Intestinal Mucins. J. Proteome Res. 14: 3239–3251.

Kakizaki, H., M. Ikeda, H. Fukushima, and M. Matsumiya, 2015 Distribution of Chitinolytic Enzymes in the Organs and cDNA Cloning of Chitinase Isozymes from the Stomach of Two Species of Fish, Chub Mackerel (*Scomber japonicus*) and Silver Croaker (*Pennahia argentata*). OJMS 05: 398–411.

Karlsen, Ø., H. Amlund, A. Berg, and R. E. Olsen, 2017 The effect of dietary chitin on growth and nutrient digestibility in farmed Atlantic cod, Atlantic salmon and Atlantic halibut. Aquac. Res. 48: 123–133.

Kawashima, S., H. Ikehata, C. Tada, T. Ogino, H. Kakizaki *et al.*, 2016 Stomach Chitinase from Japanese Sardine Sardinops melanostictus: Purification,

Characterization, and Molecular Cloning of Chitinase Isozymes with a Long Linker. Mar. Drugs 14: 22.

Koch, B. E. V., J. Stougaard, and H. P. Spaink, 2014 Spatial and temporal expression patterns of chitinase genes in developing zebrafish embryos. Gene Expr. Patterns 14: 69–77.

Kono, M., T. Matsui, and C. Shimizu, 1987 Effect of chitin, chitosan, and cellulose as diet supplements on the growth of cultured fish. NSUGAF 53: 125–129.

Kramer, K. J., T. L. Hopkins, and J. Schaefer, 1995 Applications of solids NMR to the analysis of insect sclerotized structures. Insect Biochem. Mol. Biol. 25: 1067–1080.

Krogdahl, A., G. I. Hemre, and T. P. Mommsen, 2005 Carbohydrates in fish nutrition: digestion and absorption in postlarval stages. Aquac. Nutr. 11: 103–122.

Kurita, K., 2006 Chitin and chitosan: functional biopolymers from marine crustaceans. Mar. Biotechnol. 8: 203–226.

Li, Y., L. Bruni, A. Jaramillo-Torres, K. Gajardo, T. M. Kortner *et al.*, 2021 Differential response of digesta- and mucosa-associated intestinal microbiota to dietary insect meal during the seawater phase of Atlantic salmon. Anim Microbiome 3: 8.

Liao, Y., G. K. Smyth, and W. Shi, 2014 featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30: 923–930.

Lien, S., B. F. Koop, S. R. Sandve, J. R. Miller, M. P. Kent *et al.*, 2016 The Atlantic salmon genome provides insights into rediploidization. Nature 533: 200–205.

Lindsay, G. J. H., M. J. Walton, J. W. Adron, T. C. Fletcher, C. Y. Cho *et al.*, 1984 The growth of rainbow trout (Salmo gairdneri) given diets containing chitin and its relationship to chitinolytic enzymes and chitin digestibility. Aquaculture 37: 315–334.

Lindsay, G. J. H., 1984 Distribution and function of digestive tract chitinolytic enzymes in fish. J. Fish Biol. 24: 529–536.

Marono, S., G. Piccolo, R. Loponte, C. Di Meo, Y. A. Attia *et al.*, 2015 In vitro crude protein digestibility of Tenebrio molitor and Hermetia illucens insect meals and its correlation with chemical composition traits. Ital. J. Anim. Sci. 14:.

Nakashima, K., S. Kimura, Y. Ogawa, S. Watanabe, S. Soma *et al.*, 2018 Chitin-based barrier immunity and its loss predated mucus-colonization by indigenous gut microbiota. Nat. Commun. 9: 3402.

Ohno, M., Y. Togashi, K. Tsuda, K. Okawa, M. Kamaya *et al.*, 2013 Quantification of Chitinase mRNA Levels in Human and Mouse Tissues by Real-Time PCR: Species-Specific Expression of Acidic Mammalian Chitinase in Stomach Tissues. PLoS ONE 8: e67399.

Olsen, R. E., J. Suontama, E. Langmyhr, H. Mundheim, E. Ringo *et al.*, 2006 The replacement of fish meal with Antarctic krill, Euphausia superba in diets for Atlantic salmon, Salmo salar. Aquac. Nutr. 12: 280–290.

Pohls, P., L. González-Dávalos, O. Mora, A. Shimada, A. Varela-Echavarria *et al.*, 2016 A complete chitinolytic system in the atherinopsid pike silverside Chirostoma estor: gene expression and activities. J. Fish Biol. 88: 2130–2143.

Renna, M., A. Schiavone, F. Gai, S. Dabbou, C. Lussiana *et al.*, 2017 Evaluation of the suitability of a partially defatted black soldier fly (Hermetia illucens L.) larvae meal as ingredient for rainbow trout (Oncorhynchus mykiss Walbaum) diets. J. Anim. Sci. Biotechnol. 8: 57.

Ritchie, M. E., B. Phipson, D. Wu, Y. Hu, C. W. Law *et al.*, 2015 limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 43: e47.

Robinson, M. D., D. J. McCarthy, and G. K. Smyth, 2010 edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26: 139–140.

Robinson, M. D., and A. Oshlack, 2010 A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol. 11: R25.

Rudi, K., I. L. Angell, P. B. Pope, J. O. Vik, S. R. Sandve *et al.*, 2018 Stable Core Gut Microbiota across the Freshwater-to-Saltwater Transition for Farmed Atlantic Salmon. Appl. Environ. Microbiol. 84:.

Ryczko, M. C., J. Pawling, R. Chen, A. M. Abdel Rahman, K. Yau *et al.*, 2016 Metabolic Reprogramming by Hexosamine Biosynthetic and Golgi N-Glycan Branching Pathways. Sci. Rep. 6: 23043.

Secci, G., F. Bovera, S. Nizza, N. Baronti, L. Gasco *et al.*, 2018 Quality of eggs from Lohmann Brown Classic laying hens fed black soldier fly meal as substitute for soya bean. Animal 12: 2191–2197.

Tang, W. J., J. Fernandez, J. J. Sohn, and C. T. Amemiya, 2015 Chitin is endogenously produced in vertebrates. Curr. Biol. 25: 897–900.

Teng, Z., C. Sun, S. Liu, H. Wang, and S. Zhang, 2014 Functional characterization of chitinase-3 reveals involvement of chitinases in early embryo immunity in zebrafish. Dev. Comp. Immunol. 46: 489–498.

Wang, D., A. Li, H. Han, T. Liu, and Q. Yang, 2018 A potent chitinase from Bacillus subtilis for the efficient bioconversion of chitin-containing wastes. Int. J. Biol. Macromol. 116: 863–868.

Yu, Y., C. Lee, J. Kim, and S. Hwang, 2005 Group-specific primer and probe sets to detect methanogenic communities using quantitative real-time polymerase chain reaction. Biotechnol. Bioeng. 89: 670–679.

Zhang, Y., S. Feng, J. Chen, C. Qin, H. Lin *et al.*, 2012 Stimulatory effects of chitinase on growth and immune defense of orange-spotted grouper (Epinephelus coioides). Fish Shellfish Immunol. 32: 844–854.

Paper III

The gastric mucosa of Atlantic salmon (*Salmo salar*) is abundant in highly active chitinases

Matilde Mengkrog Holen[1], Tina Rise Tuveng[2], Matthew Peter Kent[1], Gustav Vaaje-Kolstad[2]

[1] Centre for Integrative Genetics (CIGENE), Department of Animal and Aquacultural Sciences, Faculty of Biosciences, Norwegian University of Life Sciences, NO-1432 Ås, Norway

[2] Department of Chemistry, Biotechnology and Food Science (IKBM), Norwegian University of Life Sciences, NO-1432 Ås, Norway

Corresponding author: Gustav Vaaje-Kolstad, gustav.vaaje-kolstad@nmbu.no, +4767232573

Abstract

The Atlantic salmon (*Salmo salar*) genome contains 10 chitinase encoding genes, but little is known about the function of these chitinases. We show that the protein products of three genes, the family 18 glycoside hydrolase (GH18) chitinases Chia.3, Chia.4, and Chia.7 are secreted in the stomach mucosa and are amongst the most abundant proteins in this matrix. Chia.3 and Chia.4, sharing 95% sequence identity, were not possible to separate by standard chromatographic methods and were thus purified as a chitinase pair. Biochemical analysis revealed chitinolytic activity towards β-chitin for up to 24 hours at pH 2-6. Further *in vitro* analysis showed that this chitinase pair efficiently degraded various chitin-containing substrates to chitobiose ($GlcNAc_2$) suggesting that Atlantic salmon has the potential to utilize novel chitin-containing feed sources.

Introduction

Chitin is an insoluble polysaccharide, consisting of β-1,4-linked *N*-acetyl-D-glucosamine residues. It is one of the most common biopolymers in nature and is found in three distinct allomorphic forms depending on the orientation of the chitin chains [1]: α-chitin, β-chitin, and γ-chitin. In α-chitin, the most abundant form of chitin, the polymer chains have antiparallel orientation resulting in stronger intermolecular forces compared to β-chitin and γ-chitin with a parallel and a mixture of parallel and antiparallel orientation of the polymer chains respectively. Chitin functions as a structural component in algae and fungi cell walls [2,3], in the exoskeleton of arthropods such as crustaceans and insects [4,5], and is even hypothesized to be present in the scales and gut lining of some vertebrates, including ray-finned fish [6,7].

Chitinases are thought to play a role in at least three different processes [8], firstly in the breakdown of chitinous body structures during development, secondly they can be deployed to defend against infection of chitinous pathogens, and thirdly they can be involved in the digestion of chitin for nutrient absorption and energy production. Fish are known to express chitinases from the glycoside hydrolase 18 family (GH18), a multigene family with a conserved catalytic motif; DXXDXDXE [9], but little is known about the role of these chitinases. To our knowledge, Jeuniaux (1961) was the first to report endogenous chitinase activity in the fish gut using a β-chitin suspension from squid pen as substrate [10]. Such gut chitinases are hypothesized to be secreted by the gastric mucosa [11]. These enzymes have been shown to have an acidic pH optimum, while fish that lacks an acidic stomach (e.g. zebrafish) have comparable activities at neutral pH [12]. Chitinase activity in fish intestines has not been shown to correlate with the amount of dietary chitin, but fish that swallow prey whole have shown higher chitinase activity relative to the fish that have pharyngeal teeth or other buccal cavity modifications [13]. Furthermore, fish chitinases are mainly expressed in the stomach, and stomach chitinases have different activities on various insoluble chitin substrates depending on diet [14]. This indicates that fish chitinases could aid in the digestion of chitinous feed.

Wild Atlantic salmon (*Salmo salar*) is known to prey on chitinous organisms such amphipods, euphausiids, shrimp, and insects [15,16] and possess 10 genes encoding family GH18 chitinases [17]. A better understanding of the biological functions of these proteins may be of value to the salmon industry as it searches for new, alternative feed sources. Here, we quantify the relative amount of stomach chitinases in the gastric mucosa of Atlantic salmon and isolate and characterize two of these. The

results provide evidence that chitin can be degraded by these enzymes within the salmon gut.

Results

Sequence analysis

The Atlantic salmon genome encodes 10 chitinase-like genes according to the NCBI RefSeq annotation (GCF_000233375.1; release 100) that all belong to family 18 of the glycoside hydrolases, as classified by the carbohydrate active enzyme (CAZy) database [18]. Three of these chitinase genes (hereby named *chia.3*, *chia.4*, and *chia.7*; proteins named Chia.3, Chia.4, and Chia.7, respectively) have shown stomach-specific expression [17]. The sequence identity of the amino acid sequences of the chitinases range from 60-95% when aligned with each other and 61-65% when aligned with the ortholog human acidic mammalian chitinase, AMCase (Table 1).

Table 1. Sequence identity (%) between Chia.3, Chia.4, and Chia.7 and AMCase (the UniProt ID is given in parenthesis) after removing the predicted signal peptide sequence (identified with SignalP v.5.0 [19])

	Chia.3 (A0A1S3L8D8)	**Chia.4 (A0A1S3L8T9)**	**Chia.7 (A0A1S3MFN1)**	**AMCase (Q9BZP6)**
Chia.3 (A0A1S3L8D8)	100	95	60	65
Chia.4 (A0A1S3L8T9)	95	100	60	65
Chia.7 (A0A1S3MFN1)	60	60	100	61

All three chitinases share an N-terminal signal peptide, the GH18 catalytic motif DGLDXDWE, multiple putative *N*-Acetylgalactosamine (GalNAc) O-glycosylation sites, and a C-terminal domain identified as a family 14 carbohydrate-binding module (CBM14; identified by dbCAN2 annotation [20]) (Figure 1). In addition, AMCase, Chia.3 and Chia.4 share three residues hypothesized to be important for acidic activity [21,22]: an arginine (R) at position 145 and histidine (H) residues at positions 208 and 269. Chia.7 shows the least sequence identity to AMCase and has asparagine (N) residues at position 208 and 269.

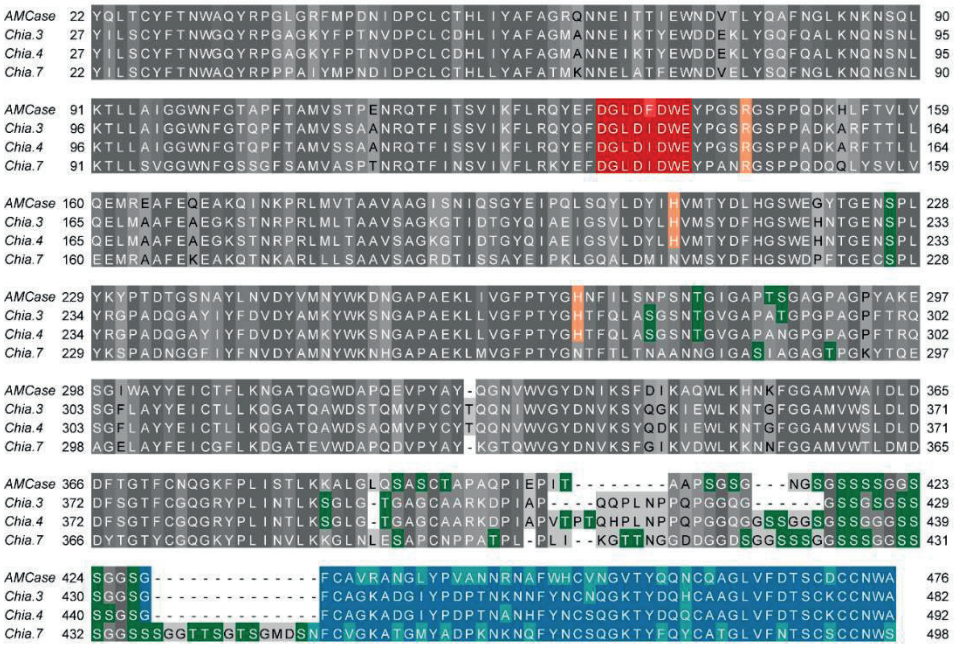

Figure 1. Multiple sequence alignment of stomach-specific chitinases from Atlantic salmon (Chia.3, Chia.4 and Chia.7) with human acidic mammalian chitinase (AMCase). The N-terminal signal peptides (identified with SignalP v.5.0 [19], default settings) were removed from the alignment. The catalytic motif residues are highlighted in red. Residues in the catalytic domain that may be important for activity in acidic environments are highlighted in orange. Predicted GalNAc O-glycosylation sites (identified with NetOGlyc v.4.0 [23], default settings) are highlighted in green and the C-terminal CBM14 (identified with dbCAN2 annotation [20]) residues are highlighted in blue (cyan if the sequence identity is low).

Chitinase abundance correlate with gene expression levels

Label-free quantitative (LFQ) proteomics was used to determine the relative amount of proteins in the gastric mucosa of Atlantic salmon collected after one day without feed. Strikingly, Chia.3, Chia.4, and Chia.7 were among the most abundant proteins in the gastric mucosa (Figure 2; Supplementary Table 1). We took advantage of published RNA-seq data generated from Atlantic salmon stomach tissue (ArrayExpress, E-MTAB-10594) to determine the correlation between the relative amount of the top 20 secreted stomach proteins and the gene expression levels of the genes coding for these proteins. The Spearman correlation (ρ) between the relative amount of proteins ($\log_2$(LFQ intensity + 1)) and gene expression levels ($\log_2$(TPM + 1), where TPM stands for transcripts per million) was 0.68 (p-value = 0.0014), indicating a positive correlation between gene expression levels and protein

abundance of the secreted stomach proteins (Figure 2). Eight genes were both the most highly expressed transcripts and the most abundant proteins in the stomach including the three chitinases (Chia.3, Chia.4, and Chia.7), two Pepsin-A-like proteases, one IgGFc-binding protein-like, a cysteine protease inhibitor; cystatin, and a lectin; fish-egg lectin. The full proteomic dataset is available in Supplementary Table 2.

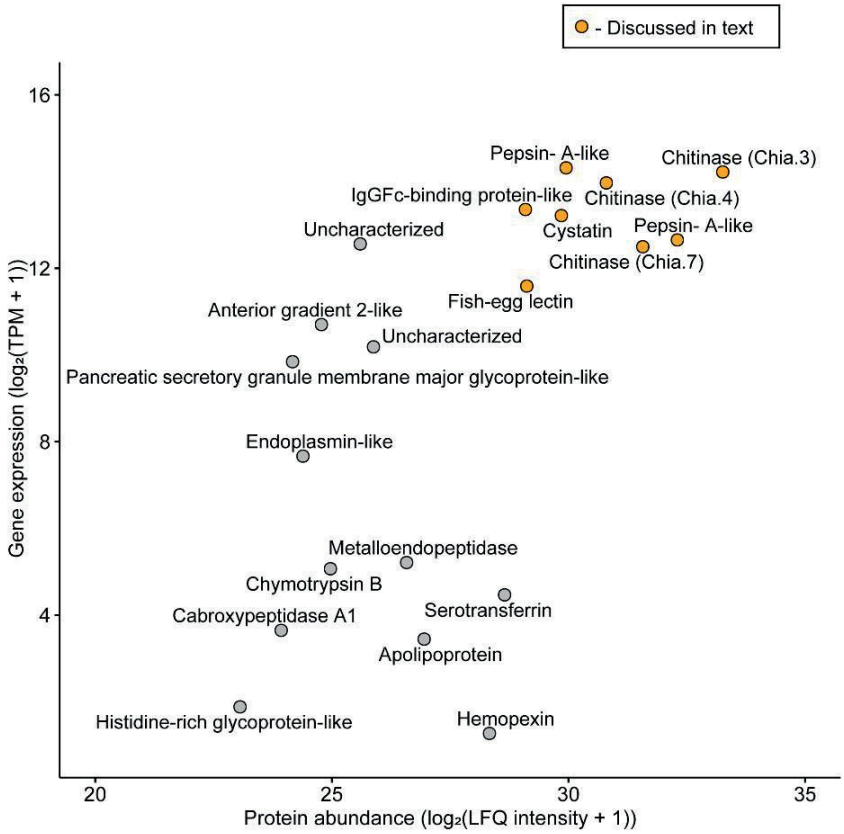

Figure 2. Spearman correlation between protein abundance of proteins (mean of $\log_2$(LFQ intensity + 1), n = 3) and gene expression levels (mean of $\log_2$(TPM + 1), n = 13) in Atlantic salmon stomach. The points highlighted in orange are discussed in the text. The corresponding Uniprot IDs can be found in Supplementary Table 1.

Purification and characterization of Atlantic salmon stomach chitinases

To capture the Atlantic salmon stomach chitinases, a stomach tissue homogenate was passed over a chitin-affinity column and the bound proteins were eluted by pH reduction from 7 (binding buffer) to 3 (elution buffer). The eluate contained two proteins represented by two distinct bands at approximately 50 and 45 kDa (Figure 3). The identity of the proteins was determined by LC-MS/MS, showing the presence of Chia.3 and Chia.4 in both bands. The predicted molecular weights of Chia.3 and Chia.4 after removal of the N-terminal signal peptide are 49.6 and 50.4 kDa respectively, and the 5 kDa difference of the two bands corresponds to the molecular mass of the CBM14 domain. The proteins were not possible to separate by standard chromatographic techniques, most likely due to their highly similar protein sequence (95% alike).

Chia.7 was not detected in the protein eluted from the column, but the fact that we know the protein is abundantly expressed suggests that conditions used in this affinity purification are not suited to capture the Chia.7 protein. All subsequent enzyme assays were performed on the Chia.3 + 4 eluate.

To determine the influence of pH and NaCl on the activity of the chitinases, activity was determined using the chitotriose analog 4MU-chitobioside (4MU-$GlcNAc_2$) or chitotetraose analog 4MU-chitotrioside (4MU-$GlcNAc_3$). The highest activity was observed at pH 2 and 3 followed by a gradual decrease to pH 8 where essentially no activity could be measured (Figure 3B). The addition of NaCl to the reaction mixture yielded a chitinase activity that increased with increasing salt concentrations, showing almost a doubling in activity from 0 to 0.6 M NaCl (Figure 3C). The latter salt concentration approximately represents the salinity of seawater.

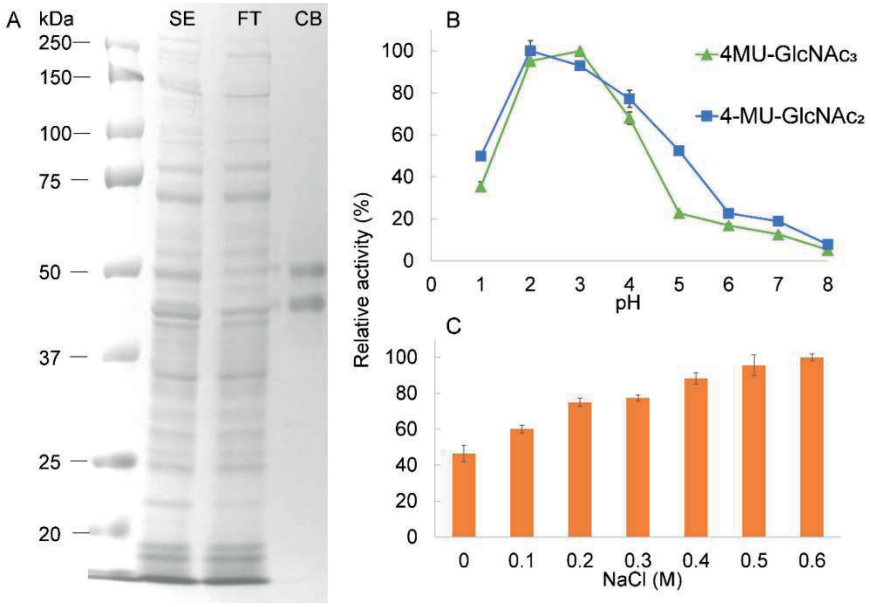

Figure 3. Purification and characterization of stomach chitinases from Atlantic salmon. (**A**) SDS-PAGE analysis of Atlantic salmon stomach chitinases were purified using chitin affinity chromatography; (1) ladder (Precicion Plus Protein, BioRad), (2) SE: total stomach extract, (3) FT: column flow-through, (4) CB: proteins eluted from the washed chitin affinity column. (**B**) Enzymatic activity at different pH values, using 100 µM 4MU-$GlcNAc_2$ (blue) and 100 µM 4MU-$GlcNAc_3$ (green) as substrate, with enzyme concentrations of 31 and 7.8 nM respectively. The reactions were incubated for 15 minutes at 37 °C in 0.1 M Gly-HCL buffer (pH 1.0 – 3.0) or McIlvaine's buffer (0.1 M citric acid and 0.2 M Na_2HPO_4, pH 4.0 – 8.0). Data points are joined by lines to indicate the shape of the pH activity relationship (**C**) Enzymatic activity with increasing NaCl concentration (0-0.6 M) using 100 µM 4MU-$GlcNAc_2$ as substrate and an enzyme concentration of 31 nM, with conditions identical as in (B), incubating the reaction in McIlvaine's buffer at pH 5. The activity was assessed in triplicate, values shown are means ± SD.

Activity towards insoluble chitin

Information about the pH optimum of an enzyme is sometimes not predictive of the enzyme performance over time (operational stability). To determine the hydrolytic potential of the chitinase pair in relevant pH conditions using a relevant substrate, the progress of a chitin degradation reaction was followed for 24 h in reactions having pH ranging from 2-6 (Figure 4). Interestingly, the highest yield after 24 h was obtained at pH 5, for which the pH optimum was less than 50% for that of pH 2 and 3 (Figure 3, panel B). The major product arising from the reactions was chitobiose ($GlcNAc_2$), but some GlcNAc was also observed (<10% of the total products).

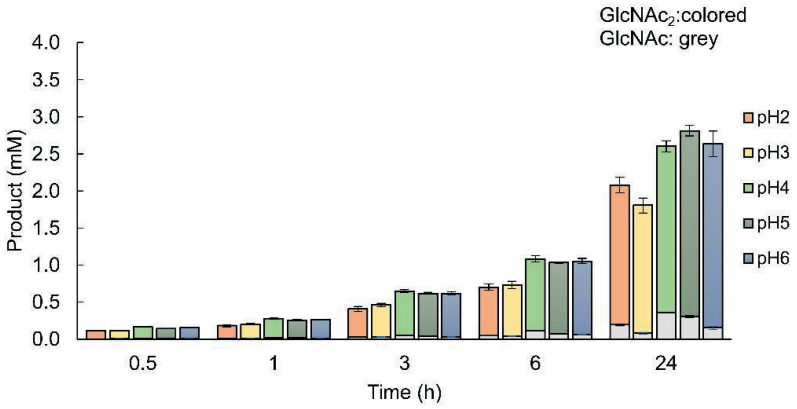

Figure 4. Operational pH stability. Concentrations of degradations products, GlcNAc dimers ($GlcNAc_2$, colored boxes) and monomers (GlcNAc, grey boxes), produced by Chia.3 + 4 when incubated with β-chitin in pH 2-6 over 24 hours are shown in a bar chart representation. The amount of degradation products was analyzed after 0.5, 1, 3, 6, and 24 hours of incubation at 14 °C. The reactions were performed with an enzyme concentration of 0.2 μM and a substrate concentration of 10 mg/mL. All reactions were run in triplicates, values are means ± SD.

The activity of Atlantic salmon stomach chitinases and ChiB from *Serratia marcescens* on α-chitin-containing organisms

The ability of the chitinase pair to depolymerize insoluble β-chitin prompted us to investigate whether the enzymes were able to break down α-chitin-containing organisms commonly found in the diet of Atlantic salmon using the shell from shrimp and crab, and skin from black soldier fly pupae. To put the activity of the chitinase pair in a metabolic context, the well-characterized chitinase from the soil bacterium

*Serratia marcescens, Sm*ChiB was included for comparison. The experiment was done at 14 °C and pH 4.8 which are conditions similar to the Atlantic salmon stomach environment [24].

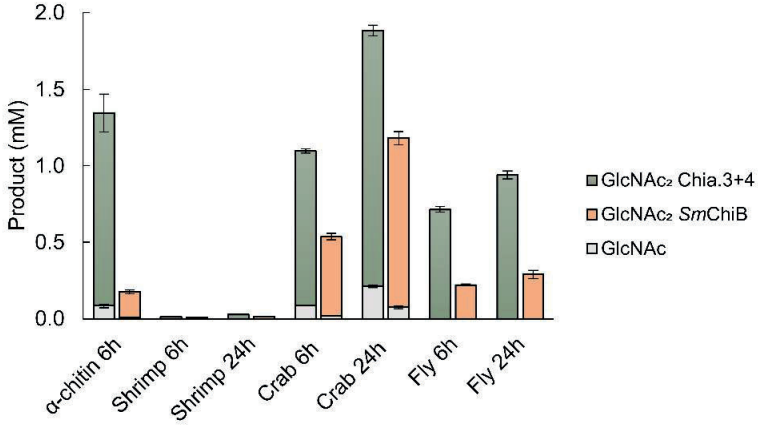

Figure 5. Degradation of natural α-chitin substrates. The bar chart displays the amount of degradation products, GlcNAc dimers ($GlcNAc_2$, colored boxes) and monomers (GlcNAc, grey boxes), produced by Atlantic salmon stomach chitinases (Chia.3 and Chia.4) compared to ChiB from *Serratia marcescens (Sm*ChiB*)* when incubated with α-chitin for 6 hours and shrimp shell (shrimp), crab shell (crab) and black soldier fly pupae skin (fly) for 6 and 24 hours in 0.1 M sodium acetate buffer (pH 4.8) at 14 °C. The reactions were performed with an enzyme concentration of 0.5 µM and a substrate concentration of 10 mg/mL for all substrates except black soldier fly pupae skin with a substrate concentration of 25 mg/mL. All reactions were run in triplicates, values are means ± SD.

The results show that the chitinase pair and *Sm*ChiB were all able to partly depolymerize the chitin-containing substrates tested (Figure 5). The Atlantic salmon chitinase pair showed substantial activity to all substrates investigated except the dried shrimp shell and showed higher chitinolytic activity on all substrates compared to *Sm*ChiB under Atlantic salmon gastric-like conditions, i.e. not optimal conditions for *Sm*ChiB [25]. Interestingly, *Sm*ChiB was less active on purified α-chitin than crab shell, while the opposite was observed for the chitinase pair after 6 hours of incubation.

Discussion

Atlantic salmon possess multiple genes encoding chitinases including three variants that are abundant as both gene transcripts and proteins in the stomach. The active proteins are secreted into the gastric mucosa but their functional role there is

unknown. The stomach of Atlantic salmon works both as a "feed-grinder" and as a first-line defense against water- and feed-borne pathogens. In theory, chitinases could play a role in the digestion of chitin-containing feed and/or chitin-containing pathogens such as fungi. The results show that the chitinases are present together with high levels of IgGFc-binding-like protein, cystatin and fish-egg lectin (Figure 2), three proteins shown to play a potential role in the teleost immune system against pathogens [26–30]. Moreover, the Atlantic salmon stomach chitinases share many features with human AMCase [31]; they share the same conserved domains, they are abundant in the gastrointestinal tract, they are acid-stable, and they are able to degrade crab shell chitin. AMCase has shown to play potential roles in digestion and the immune response against chitin-containing organisms, a role that could be similar for the Atlantic salmon gastric chitinases.

In this study, Chia.3 and Chia.4 were purified from Atlantic salmon stomachs using chitin affinity chromatography, but we were unable to purify Chia.7 despite its identification by proteomic analysis. Chia.3 and Chia.4 are almost identical, while Chia.7 shows less sequence identity to the other stomach chitinases and differences in chitin-binding capacities between the different chitinases could be the reason for not being able to isolate Chia.7 in the same conditions. The combined fraction of pure Chia.3 and Chia.4 showed a pH optimum of pH 2 and 3 respectively when using fluorogenic substrates (4MU-chitobioside and 4MU-chitotriose). This pH optimum was different from the operational pH stability optimum of pH 5 determined towards β-chitin, a more realistic substrate for the chitinases. Differences in substrate length and composition have previously been shown to affect the chitinolytic pH optimum [32], and this underscores the need to carefully consider the substrates used when characterizing proteins. Also, the non-natural aglycon leaving group of the 4MU-conjugated substrates and the insoluble nature of β-chitin may contribute to the pH preference of the enzyme. Atlantic salmon has a gastric stomach with an average pH of 4.8 depending on both feed type and time since ingestion [24], which corroborates the operational pH stability data. Further, the pH in the contents of the most distal part of the stomach is generally higher (pH 5) than the middle part and increases to a more neutral pH (pH 8) in the pyloric, mid, and distal intestine. Our results show that the Atlantic salmon stomach chitinase pair were highly active and stable at pH 2.0-6.0 over 24 hours. This is biologically relevant as it may take up to somewhere between 24 to 48 hours for the fish to empty its stomach upon ingestion [33]. A pH optimum of pH 2 when using fluorogenic substrates has previously been reported for stomach chitinases isolated from humans, mice, chicken, and pigs [31,34–36]. Moreover, the optimum pH of stomach chitinases from Silver croaker (*Pennahia argentatus*),

Marbled rockfish (*Sebastiscus marmoratus),* Red sea bream (*Pagrus major*), Japanese eel (*Anguilla japonica*), and Red scorpionfish (*Scorpaena scrofa*) have been reported to be pH 4.0-5.0, 4.0-4.5, 5.5, 4.4, and 5.0 respectively when using colloidal chitin as a substrate [14]. This is in line with our observations.

Salmon migrate from the rivers into the sea and must adapt to a change in diet and increasing salt concentrations. The results show that the stomach chitinases of Atlantic salmon were substantially more active at 0.6 M NaCl, equivalent to seawater salinity, than without any NaCl. The NaCl concentration of seawater corresponds to an osmolality of 1200 mOsm and the stomach chyme osmolality of rainbow trout (*Oncorhynchus mykiss*) has been reported to reach a maximum of 775 mOsm two hours after feeding [37]. This value was reported for fish living in freshwater, and it is probably higher in seawater salmonids since it is known that salmon drink more external water for osmoregulation in seawater than in freshwater and because the gut is an important osmoregulatory organ in seawater teleost [38,39]. Furthermore, our results show that the Atlantic salmon stomach chitinases degraded chitin from shrimp shells, crab shells and black soldier fly pupae more efficiently than ChiB from *S. marcescens*. The latter bacterium utilizes chitin as an energy source and is one of the most efficient chitin degraders out of 100 tested microorganisms [40,41]. The higher efficiency of salmon chitinases could be a result of non-optimal conditions for *Sm*ChiB which works best at pH 5.0-6.0 with a temperature optimum of 58 °C [25], and/or a synergy effect of the two salmon chitinases working together.

Recent studies show that insect meal from black soldier fly has the potential to replace fish meal in the aquaculture industry [42,43] and that chitin and chitin degradation products can act as immunostimulants [44,45]. Altogether, our results show that using chitin-containing organisms as novel feed sources for farmed Atlantic salmon can be of nutritional value.

Conclusion

Our results show that some of the most dominant proteins in the stomach of Atlantic salmon are chitinases that are capable of effectively degrading chitin or chitin-containing substrates from various sources. The stomach chitinases are active and stable in the gastric-like conditions of Atlantic salmon and are therefore likely to play a role in the digestion of chitin-containing organisms commonly found in the natural diet of salmon. The results presented here can be taken into consideration when searching for novel feed ingredients in the aquaculture industry.

Acknowledgments

We want to thank Morten Skaugen at the Proteomics Core Facility at NMBU for helping with sample preparation and mass spectrometry analysis. We are also grateful to the people at the Center for Sustainable Aquaculture at NMBU for providing us with the fish used in the experiments and Sophanit Mekasha for providing us with *Sm*ChiB.

Methods

Multiple sequence alignment

The amino acid sequences of Chia.3, Chia.4, Chia.7, and AMCase were downloaded from Uniprot and the alignment was calculated using Mafft with default settings in Jalview v. 2.11.14. The signal peptides were predicted with SignalP v.5.0 [19] using default settings. The GalNAc O-glycosylation sites were identified with NetOGlyc v.4.0 [23] using default settings and the CBM14 domain was identified with dbCAN2 annotation [20].

Chitinous substrates

β-chitin (extracted from squid pen, Batch 20140101, France Chitin, Orange, France) was pulverized with a bead mill (Planetary Ball Mill PM 100, Retsch) to approx. 200 µm particle size. Shrimp shell was peeled off shrimps (*Pandalus Borealis*, Polar Seafood Norway) and the filling was removed from crab shell (*Cancer pagurus*, Lofotprodukt). Both products were bought at the local food market. Shrimp shell, crab shell, and black soldier fly pupae skin (*Hermetica illucens*, a kind gift from Fraunhofer-Gesellschaft, Munich, Germany) were dried at 105 °C overnight before the experiments were run. Shrimp- and crab shells were first crushed with mortar and pestle before the shells, black soldier fly pupae skin, and α-chitin (extracted from *Pandalus borealis*, Seagarden, Avaldsnes, Norway) were pulverized with a bead mill (Planetary Ball Mill PM 100, Retsch) to approx. 200 µm particle size.

Enzymes

Chitinases from Atlantic salmon were isolated from stomach tissue and purified as described below. Chitinase B from *Serratia marcescens* (*Sm*ChiB) was overexpressed in *Escherichia coli* and purified as previously described [46].

Proteomic analysis of Atlantic salmon gastric mucosa

Gastric mucosa of adult Atlantic salmon (n = 3, two male and one female, average fish weight 2245 g) was obtained from the process plant for fish farming laboratory at The Norwegian University of Life Sciences (NMBU). The fish was starved for one day before they were euthanized by a blow to the head. The stomach was dissected from

the fish, the gastric mucosa was scraped off and mixed with 1 mL ice-cold phosphate buffer (20 mM sodium phosphate buffer pH 7.0, 150 mM NaCl) with 1X protease inhibitor cocktail (complete, EDTA-free Protease Inhibitor Cocktail, Roche) by pipetting and gentle vortexing. The homogenate was centrifuged at 12,000 rpm for 10 minutes at 4 °C and the supernatant was filtered through a 40- μm cell strainer. The total protein concentration was determined with Bradford Protein Assay (Bio-Rad) using Bovine Serum Albumin as standard. A total amount of 2 μg protein was loaded on an SDS-PAGE gel. The proteins were allowed to enter the gel, but without full separation of the proteins in the gel. The gel was stained with Coomassie Brilliant Blue R-250 (Bio-Rad) and de-stained before the region of the gel containing proteins were cut out for in-gel digestion, essentially performed as described by Shevchenko et al [47]. In brief, proteins were reduced with DTT and alkylated with iodoacetamide before trypsinization. ZipTip C18 pipette tips (Merck Millipore) were used to purify peptides, followed by drying under vacuum. The peptides were dissolved in 10 μl of 2% (v/v) acetonitrile, 0.1% (v/v) trifluoroacetic acid and the peptide concentration was measured using av NanoDrop One and used to normalize the amount of peptides injected for LC-MS/MS analysis. The LC-MS/MS analysis was performed as described by Tuveng et al [48].

MS Raw files were analyzed using MaxQuant [49] version 1.6.17.0, and proteins were identified and quantified using the MaxLFQ algorithm [50]. Samples were searched against the proteome of *Salmo salar* downloaded from Uniprot (UP000087266), and a list of common contaminants (included in the MaxQuant software package). As variable modifications protein N-terminal acetylation, oxidation of methionine, conversion of glutamine to pyroglutamic acid, and deamination of asparagine and glutamine were used, while carbamidomethylation of cysteine residues was used as a fixed modification. Trypsin was used as a digestion enzyme and two missed cleavages were allowed. The feature 'Match between runs' in MaxQuant, which enables identification transfer between samples based on accurate mass and retention time, was applied with default settings [50]. The results from MaxQuant were further processed using Perseus (version 1.6.15.0) Proteins categorized as 'only identified by site', 'reverse' or as 'contaminant' were removed from the dataset. As an additional cut-off criterium, proteins were only considered present if they were detected in at least two of three replicates. The LFQ (label-free quantification) intensities were log2-transformed and averaged before analysis. The downstream analysis focused on proteins predicted to have a signal peptide using the *Salmo Salar* proteome (UniProt id: UP000087266).

Comparison of protein abundance with gene expression levels

RNA-sequencing data from the stomach of Atlantic salmon was downloaded from ArrayExpress under project number E-MTAB-10594. The bcbio-nextgen pipeline (https://github.com/bcbio/bcbio-nextgen was used to trim, map and count raw reads before aligning to the Atlantic salmon genome (ICSASG_v2) [51] using STAR [52]. Reads aligned to genes were counted with FeatureCounts [53] and transformed to transcripts per million (TPM) values to normalize for gene length. The TPM values were log2-transformed and averaged (n = 13) before analysis. A subset of genes coding for the 20 most abundant proteins secreted in stomach mucosa was used to calculate the Spearman correlation between the gene expression levels (log2(TPM + 1)) and the relative protein abundance (log2(LFQ intensity + 1)). The statistical analysis was done using the "ggpubr" package in R v.4.0.3.

Purification of chitinases from stomach tissue

Stomach tissue (n = 2, average 7.2 g per purification, two rounds of purification) from adult Atlantic salmon (one female, one male, average weight 2230 g) was obtained from the process plant for the fish farming laboratory at NMBU. The fish was euthanized by a blow to the head and the stomach was dissected from the fish. Stomach content was removed before the tissue was cut into small pieces and stored in ice-cold phosphate buffer (20 mM sodium phosphate buffer pH 7.0, 150 mM NaCl) with 2X protease inhibitor cocktail (complete, EDTA-free Protease Inhibitor Cocktail, Roche). The tissue was homogenized directly after dissection.

Stomach tissue was homogenized in 5 volumes of ice-cold phosphate buffer (20 mM sodium phosphate buffer pH 7.0, 150 mM NaCl) with 2X protease inhibitor cocktail (complete, EDTA-free Protease Inhibitor Cocktail, Roche) using TissueRuptor II (QIAGEN, 20 sec on/off). The homogenate was filtered through a 40-μm cell strainer and centrifuged at 12,000 rpm for 20 minutes at 4 °C. NaCl was added to the supernatant to get a final concentration of 1.0 M NaCl. The supernatant was filtered through a 0.22 μm sterile filter and used for chitin affinity chromatography.

The stomach extract was purified on a 1.5 cm diameter, 10 mL column packed with chitin resin slurry (New England Biolabs). The column was pre-equilibrated with phosphate buffer (20 mM sodium phosphate buffer pH 7.0, 1.0 M NaCl) before the stomach extract was applied to the column. After washing with phosphate buffer (20

mM sodium phosphate buffer pH 7.0, 1.0 M NaCl), the chitinases were eluted with 100 mM acetic acid. A flow rate of 1 mL/min was used at all steps. Concentration and buffer exchange of the eluted chitinases to phosphate buffer (20 mM sodium phosphate buffer pH 7.0, 150 mM NaCl) was done using a 10 kDa centrifugal filter (Macrosep Advance Centrifugal Device, 10 kDa cutoff, Pall corporation). The purity of the eluted chitinases was examined with SDS-PAGE and Coomassie Brilliant Blue R-250 staining, and the proteins in the appearing bands were further identified by LC-MS at the local Proteomics Core Facility (NMBU). The protein concentration was determined with Quick Start Bradford Protein Assay (Bio-Rad) using Bovine Serum Albumin as standard.

Confirmation of chitinase proteins with mass spectrometry

The protein bands in the de-stained gel were cut out using a clean scalpel blade. After trimming away unstained gel, the bands were further divided into 1-2 mm cubes and transferred to clean 0.2 mL PCR tubes. 100 µl of 50 % acetonitrile (ACN), 50 mM ammonium bicarbonate (ABC) was added to each tube, which was then incubated at room temperature with shaking for 10 minutes. After brief centrifugation, the liquid was aspirated and replaced with 200 µl 100 % ACN. The tubes were incubated at room temperature for 15 minutes and the liquid was removed by aspiration.

The in-gel reduction was performed by adding 50 µl 10 mM DTT, 50 mM ABC to the dried gel pieces, and incubating for 30 minutes at 56 °C in a thermocycler. Alkylation was performed by replacing the solution with 50 µl of 50 mM iodoacetamide, 50 mM ABC, and incubating in the dark for 20 min at room temperature.

After having removed the alkylation solution, 200 µl of 100% ACN was added and the tubes were incubated at room temperature for 15 minutes, followed by liquid removal and brief air drying of the gel pieces. The tubes were put on ice, and 30 µl ice-cold trypsin solution (13 ng/µl, in 50 mM ABC) was added to each tube. The gel pieces were allowed to swell for a total of 90 minutes on ice, with occasional checks to ensure that they were completely covered with the digestion solution. Finally, the tubes were transferred to a thermocycler and incubated overnight at 37 °C.

Trypsin digestion was terminated by adding 50 µl TFA solution (final concentration 0.2%), and the tubes were sonicated for 10 min in a water bath sonicator. After brief centrifugation, the liquid was transferred to a clean tube. 50 µl of 0.1% TFA was added to the gel pieces and the sonication step was repeated. After combining the two

extracts, peptides were purified using STAGE spin-tips, as previously described [54]. Eluted peptides were dried in an Eppendorf Concentrator Plus vacuum centrifuge and dissolved in loading solution (0.05% TFA, 2% ACN in Milli-Q water) before LC-MS/MS analysis.

Samples were loaded onto a trap column (Acclaim PepMap100, C_{18}, 5 µm, 100 Å, 300 µm i.d. x 5 mm, Thermo Scientific) and backflushed onto a 50 cm analytical column (Acclaim PepMap RSLC C_{18}, 2 µm, 100 Å, 75 µm i.d., Thermo Scientific). Starting conditions were 96 % solution A [0.1 % (v/v) formic acid], 4% solution B [80 % (v/v) ACN, 0.1 % (v/v) formic acid]. Peptides were eluted using a flow rate of 300 nl/min using a 70 min method, with the following gradient: from 3.2 to 10 % B in 3 minutes, 10 to 35 % B in 44 minutes and 35 to 60% B in 3 minutes, followed by a 5 min wash at 80 % B and a 15 min equilibration at 4% B. The Q-Exactive mass spectrometer was operated in data-dependent acquisition (DDA) mode using a Top10 DDA method, where acquisition alternates between orbitrap-MS and higher-energy collisional dissociation (HCD) orbitrap-MS/MS acquisition of the 10 most intense precursor ions. Only charge states 2-5 were selected for fragmentation, and the normalized collision energy (NCE) was set to 28. The selected precursor ions were excluded for repeated fragmentation for 20 seconds. The resolution was set to R=70,000 and R=35,000 for MS and MS/MS, respectively. Automatic gain control values were set to $3x10^6$ and $5x10^4$ for MS and MSMS, respectively, with a maximum injection time of 100 and 128 ms. The MS Raw files were analyzed using MaxQuant as described in the previous section.

Chitinase assays with fluorogenic substrates

Chitinolytic activity was determined using a Chitinase Assay Kit (Fluorimetric, Sigma-Aldrich) with 4-Methylumbelliferyl N,N'-diacetyl- β -D-chitobioside (4-MU-$GlcNAc_2$) and 4-Methylumbelliferyl β-D-N,N',N''-triacetylchitotriose (4-MU-$GlcNAc_3$) according to the manufacturer's protocol. To determine the activity in different salt concentrations 31 nM of purified chitinase was incubated with 100 µM 4-MU-$GlcNAc_2$ in McIlvaine's buffer (0.1 M citric acid and 0.2 M Na_2HPO_4, pH 5) with different concentrations of NaCl (0 – 0.6 M) in a volume of 100 µL at 37 °C for 15 minutes. To determine pH optimum, 31 nM (for assay with 4-MU-$GlcNAc_2$) or 7.8 nM (for assay with 4-MU-$GlcNAc_3$) of purified chitinase was incubated with 100 µM substrate in 0.1 M Gly-HCL buffer (pH 1.0 – 3.0) or McIlvaine's buffer (0.1 M citric acid and 0.2 M Na_2HPO_4, pH 4.0 – 8.0) in a volume of 100 µL at 37 °C for 15 minutes. The reaction

was stopped by the addition of 200 µL 0.4 M sodium carbonate. The fluorescence of the released 4-Methylumbelliferone (4-MU) was measured using a fluorimeter with excitation at 360 nm and emission at 450 nm and a 4-MU standard curve was used to quantify 4-MU resulting from the hydrolytic reaction. The measured fluorescence was corrected for hydrolysis of the substrate without the addition of an enzyme (blank). Each reaction was performed in triplicate.

pH and stability assay

Purified chitinase enzyme or phosphate buffer (blank) was mixed with β-chitin in appropriate volumes of the following buffer solutions: 20 mM glycine-HCL (pH 2 and 3), 20 mM sodium acetate (pH 4 and 5), 20 mM sodium phosphate (pH 6), to yield final concentrations of 0.2 µM (chitinase) and 10 mg/mL (β-chitin). The reactions mixtures were incubated at 14 °C in an Eppendorf thermomixer at 1000 rpm and samples were taken after 0.5, 1, 3, 6 and 24 hours of incubation and filtered to remove β-chitin particles and thereby stop the reaction (0.45 µm vacuum filter, Merck Millipore). To adjust the samples for chromatography and to inactivate enzymes, H_2SO_4 was added to a final concentration of 20 mM in the filtrate. All reactions were run in triplicates. The end products were analyzed by HPLC (see below).

Chitinase assay with α-chitin, shrimp shells, crab shells and black soldier fly pupae skin

Purified chitinase from Atlantic salmon (0.5 µM), purified ChiB from *Serratia marscecens* (0.5 µM) or phosphate buffer (blank) was mixed with α -chitin (10 mg/mL), shrimp shell (10 mg/mL), crab shell (10 mg/mL) and black soldier fly pupae skin (25 mg/mL) in 0.1 M sodium acetate (pH 4.8), and incubated at 14 °C in an Eppendorf thermomixer at 1000 rpm (all concentrations noted in parenthesis indicate final concentrations in the reaction mixtures). Samples were withdrawn after 6 and 24 hours of incubation, filtered by a 0.45 µm filter to remove the substrate particles from the reaction mixture and thereby stop the reaction. The enzyme activity was inactivated by addition H_2SO_4 to a final concentration of 20 mM in the filtrate. All reactions were run in triplicates. The end products were analyzed by HPLC (see below).

High-performance liquid chromatography (HPLC)

Concentrations of mono- and disaccharides of *N*-acetylglucosamine (GlcNAc and $GlcNAc_2$) were determined as previously described [55].

Availability of data

The proteomics data have been deposited to the ProteomeXchange consortium (http://proteomecentral.proteomexchange.org) via the PRIDE [56] partner repository with the dataset identifier PXD030291.

References

1. Jang, M.-K., Kong, B.-G., Jeong, Y.-I., Lee, C. H. & Nah, J.-W. Physicochemical characterization of α-chitin, β-chitin, and γ-chitin separated from natural resources. *J. Polym. Sci. A Polym. Chem.* **42**, 3423–3432 (2004).

2. Bowman, S. M. & Free, S. J. The structure and synthesis of the fungal cell wall. *Bioessays* **28**, 799–808 (2006).

3. Pearlmutter, N. L. & Lembi, C. A. Localization of chitin in algal and fungal cell walls by light and electron microscopy. *J. Histochem. Cytochem.* **26**, 782–791 (1978).

4. Kramer, K. J., Hopkins, T. L. & Schaefer, J. Applications of solids NMR to the analysis of insect sclerotized structures. *Insect Biochem. Mol. Biol.* **25**, 1067–1080 (1995).

5. Kurita, K. Chitin and chitosan: functional biopolymers from marine crustaceans. *Mar. Biotechnol.* **8**, 203–226 (2006).

6. Tang, W. J., Fernandez, J., Sohn, J. J. & Amemiya, C. T. Chitin is endogenously produced in vertebrates. *Curr. Biol.* **25**, 897–900 (2015).

7. Nakashima, K. *et al.* Chitin-based barrier immunity and its loss predated mucus-colonization by indigenous gut microbiota. *Nat. Commun.* **9**, 3402 (2018).

8. Rathore, A. S. & Gupta, R. D. Chitinases from Bacteria to Human: Properties, Applications, and Future Perspectives. *Enzyme Res.* **2015**, 791907 (2015).

9. Hussain, M. & Wilson, J. B. New paralogues and revised time line in the expansion of the vertebrate GH18 family. *J. Mol. Evol.* **76**, 240–260 (2013).

10. Jeuniaux, C. Chitinase: an addition to the list of hydrolases in the digestive tract of vertebrates. *Nature* **192**, 135–6 (1961).

11. Dandrifosse, G. & Schoffeniels, E. Existence of a mechanism of chitinase transport across cell membranes. *Biochim Biophys Acta* **94**, 165–74 (1965).

12. Teng, Z., Sun, C., Liu, S., Wang, H. & Zhang, S. Functional characterization of chitinase-3 reveals involvement of chitinases in early embryo immunity in zebrafish. *Dev. Comp. Immunol.* **46**, 489–498 (2014).

13. Lindsay, G. J. H. Distribution and function of digestive tract chitinolytic enzymes in fish. *J. Fish Biol.* **24**, 529–536 (1984).

14. Ikeda, M., Kakizaki, H. & Matsumiya, M. Biochemistry of fish stomach chitinase. *Int. J. Biol. Macromol.* **104**, 1672–1681 (2017).

15. Jacobsen, J. Feeding habits of wild and escaped farmed Atlantic salmon, Salmo salar L., in the Northeast Atlantic. *ICES J. Mar. Sci* **58**, 916–933 (2001).

16. Rikardsen, A. H. *et al.* Geographical differences in marine feeding of Atlantic salmon post-smolts in Norwegian fjords. *J. Fish Biol.* **64**, 1655–1679 (2004).

17. Holen, M. M., Kent, M. P., Vaaje-Kolstad, G. & Sandve, S. R. Gene family expansion and functional diversification of chitinase and chitin synthase genes in Atlantic salmon (Salmo salar). *BioRxiv* (2022) doi:10.1101/2022.05.05.490710.

18. Drula, E. *et al.* The carbohydrate-active enzyme database: functions and literature. *Nucleic Acids Res.* **50**, D571–D577 (2022).

19. Nielsen, H., Engelbrecht, J., Brunak, S. & von Heijne, G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* **10**, 1–6 (1997).

20. Zhang, H. *et al.* dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* **46**, W95–W101 (2018).

21. Olland, A. M. *et al.* Triad of polar residues implicated in pH specificity of acidic mammalian chitinase. *Protein Sci.* **18**, 569–578 (2009).

22. Bussink, A. P., Vreede, J., Aerts, J. M. F. G. & Boot, R. G. A single histidine residue modulates enzymatic activity in acidic mammalian chitinase. *FEBS Lett.* **582**, 931–935 (2008).

23. Steentoft, C. *et al.* Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO J.* **32**, 1478–1488 (2013).

24. Krogdahl, Å., Sundby, A. & Holm, H. Characteristics of digestive processes in Atlantic salmon (Salmo salar). Enzyme pH optima, chyme pH, and enzyme activities. *Aquaculture* **449**, 27–36 (2015).

25. Brurberg, M. B., Nes, I. F. & Eijsink, V. G. H. Comparative studies of chitinases A and B from Serratia marcescens. *Microbiology (Reading, Engl)* **142 (Pt 7)**, 1581–1589 (1996).

26. Elvitigala, D. A. S. & Lee, J. Identification and molecular profiling of a novel homolog of cystatin C from rock bream (Oplegnathus fasciatus) evidencing its transcriptional sensitivity to pathogen infections. *Mol. Biol. Rep.* (2021) doi:10.1007/s11033-021-06415-9.

27. Kim, B.-S. *et al.* Molecular characterisation and expression analysis of a fish-egg lectin in rock bream, and its response to bacterial or viral infection. *Fish Shellfish Immunol.* **31**, 1201–1207 (2011).

28. Peatman, E., Lange, M., Zhao, H. & Beck, B. H. Physiology and immunology of mucosal barriers in catfish (Ictalurus spp.). *Tissue Barriers* **3**, e1068907 (2015).

29. Wang, Y., Bu, L., Yang, L., Li, H. & Zhang, S. Identification and functional characterization of fish-egg lectin in zebrafish. *Fish Shellfish Immunol.* **52**, 23–30 (2016).

30. Zhang, K. *et al.* Antibacterial functions of a novel fish-egg lectin from spotted knifejaw (Oplegnathus punctatus) during host defense immune responses. *Dev. Comp. Immunol.* **111**, 103758 (2020).

31. Boot, R. G. *et al.* Identification of a novel acidic mammalian chitinase distinct from chitotriosidase. *J. Biol. Chem.* **276**, 6770–6778 (2001).

32. Barad, B. A. *et al.* Differences in the chitinolytic activity of mammalian chitinases on soluble and insoluble substrates. *Protein Sci.* **29**, 966–977 (2020).

33. Aas, T. S. *et al.* Measurement of gastrointestinal passage rate in Atlantic salmon (Salmo salar) fed dry or soaked feed. *Aquaculture Reports* **8**, 49–57 (2017).

34. Ohno, M. *et al.* Acidic mammalian chitinase is a proteases-resistant glycosidase in mouse digestive system. *Sci. Rep.* **6**, 37756 (2016).

35. Tabata, E. *et al.* Gastric and intestinal proteases resistance of chicken acidic chitinase nominates chitin-containing organisms for alternative whole edible diets for poultry. *Sci. Rep.* **7**, 6662 (2017).

36. Tabata, E. *et al.* Protease resistance of porcine acidic mammalian chitinase under gastrointestinal conditions implies that chitin-containing organisms can be sustainable dietary resources. *Sci. Rep.* **7**, 12963 (2017).

37. Bucking, C. & Wood, C. M. Water dynamics in the digestive tract of the freshwater rainbow trout during the processing of a single meal. *J. Exp. Biol.* **209**, 1883–1893 (2006).

38. Usher, M. L., Talbot, C. & Eddy, F. B. Effects of transfer to seawater on digestion and gut function in Atlantic salmon smolts (Salmo salar L.). *Aquaculture* **90**, 85–96 (1990).

39. Kirsch, R. & Meister, M. F. Progressive processing of ingested water in the gut of sea-water teleosts. *J. Exp. Biol.* **98**, 67–81 (1982).

40. Vaaje-Kolstad, G., Horn, S. J., Sørlie, M. & Eijsink, V. G. H. The chitinolytic machinery of Serratia marcescens--a model system for enzymatic degradation of recalcitrant polysaccharides. *FEBS J.* **280**, 3028–3049 (2013).

41. Monreal, J. & Reese, E. T. The chitinase of Serratia marcescens. *Can. J. Microbiol.* **15**, 689–696 (1969).

42. Belghit, I. *et al.* Black soldier fly larvae meal can replace fish meal in diets of sea-water phase Atlantic salmon (Salmo salar). *Aquaculture* **503**, 609–619 (2019).

43. Li, Y. *et al.* Total replacement of fish meal with black soldier fly (Hermetia illucens) larvae meal does not compromise the gut health of Atlantic salmon (Salmo salar). *Aquaculture* 734967 (2020) doi:10.1016/j.aquaculture.2020.734967.

44. Ringø, E., Zhou, Z., Olsen, R. E. & Song, S. K. Use of chitin and krill in aquaculture - the effect on gut microbiota and the immune system: a review. *Aquac. Nutr.* **18**, 117–131 (2012).

45. Ngo, D.-H. & Kim, S.-K. Antioxidant effects of chitin, chitosan, and their derivatives. in *Marine carbohydrates: fundamentals and applications, part B* vol. 73 15–31 (Elsevier, 2014).

46. Mekasha, S. *et al.* Development of enzyme cocktails for complete saccharification of chitin using mono-component enzymes from Serratia marcescens. *Process Biochemistry* **56**, 132–138 (2017).

47. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V. & Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **1**, 2856–2860 (2006).

48. Tuveng, T. R. *et al.* Proteomic investigation of the secretome of Cellvibrio japonicus during growth on chitin. *Proteomics* **16**, 1904–1914 (2016).

49. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).

50. Cox, J. *et al.* Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteomics* **13**, 2513–2526 (2014).

51. Lien, S. *et al.* The Atlantic salmon genome provides insights into rediploidization. *Nature* **533**, 200–205 (2016).

52. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

53. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).

54. Sprotocols. A Spinnable And Automatable Stagetip For High Throughput Peptide Desalting And Proteomics. *Zenodo* (2014) doi:10.5281/zenodo.13643.

55. Hamre, A. G. *et al.* Treatment of recalcitrant crystalline polysaccharides with lytic polysaccharide monooxygenase relieves the need for glycoside hydrolase processivity. *Carbohydr. Res.* **473**, 66–71 (2019).

56. Vizcaíno, J. A. *et al.* The PRoteomics IDEntifications (PRIDE) database and associated tools: status in 2013. *Nucleic Acids Res.* **41**, D1063-9 (2013).

ISBN: ISBN 978-82-575-1918-6
ISSN: 1894-6402

Norwegian University
of Life Sciences

Postboks 5003
NO-1432 Ås, Norway
+47 67 23 00 00
www.nmbu.no

114008 / ANDVORDGRAFISK.NO