Rainbow trout (Oncorhynchus mykiss) urea cycle and polyamine synthesis gene families show dynamic expression responses to inflammation

T. Clark, J. Tinsley, D.J. Macqueen, S.A.M. Martin

PII: S1050-4648(19)30234-7
DOI: https://doi.org/10.1016/j.fsi.2019.03.075
Reference: YFSIM 6043

To appear in: Fish and Shellfish Immunology

Received Date: 7 February 2019
Revised Date: 29 March 2019
Accepted Date: 29 March 2019

Please cite this article as: Clark T, Tinsley J, Macqueen DJ, Martin SAM, Rainbow trout (Oncorhynchus mykiss) urea cycle and polyamine synthesis gene families show dynamic expression responses to inflammation, Fish and Shellfish Immunology (2019), doi: https://doi.org/10.1016/j.fsi.2019.03.075.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Rainbow trout (*Oncorhynchus mykiss*) urea cycle and polyamine synthesis gene families show dynamic expression responses to inflammation

Clark T.  
Tinsley J.  
Macqueen D.J.  
Martin S.A.M.

1 School of Biological Sciences, University of Aberdeen, Tillydrone Avenue, Aberdeen, AB24 2TZ

2 BioMar Ltd, Grangemouth Docks, Grangemouth FK3 8UL, UK.

a Current address: The Roslin Institute and Royal (Dick) School of Veterinary Studies, The University of Edinburgh, Midlothian, EH25 9RG, UK.

*Corresponding author:
Prof Samuel A.M. Martin
School of Biological Sciences,
University of Aberdeen,
Tillydrone Avenue,
Aberdeen,
AB24 2TZ
UK
sam.martin@abdn.ac.uk
Abstract

The urea cycle is an endogenous source of arginine that also supports removal of nitrogenous waste following protein metabolism. This cycle is considered inefficient in salmonids, where only 10-15 percent of nitrogenous waste is excreted as urea. In rainbow trout, arginine is an essential amino acid that has attracted attention due to its many functional roles. These roles include the regulation of protein deposition, immune responses and polyamine synthesis; the latter is directly linked to the urea cycle and involved in tissue repair. The key enzymes used in the urea cycle, namely arginase, ornithine transcarbamylase, argininosuccinate synthase and argininosuccinate lyase, in addition to two rate limiting enzymes required for polyamine synthesis (ornithine decarboxylase and s-adenosylmethionine decarboxylase) are poorly studied in fishes, and their responses to inflammation remain unknown. To address this knowledge gap, we characterised these gene families using phylogenetics and comparative genomics, investigated their mRNA distribution among a panel of tissues and established their transcriptional responses to an acute inflammatory response caused by bacterial infection in liver and muscle. Gene duplicates (paralogues) were identified for arginase (ARG1a, 1b, 2a and 2b), ornithine decarboxylase (ODC1 and 2) and s-adenosylmethionine decarboxylase (SAMdc1 and 2), including paralogues retained from an ancestral salmonid-specific whole genome duplication. ARG2a and 2b were highly upregulated following bacterial infection in liver, whereas ARG1b was downregulated, while both paralogues of SAMdc and ODC were upregulated in both liver and muscle. Overall, these findings improve our understanding of the molecules supporting the urea cycle and polyamine synthesis in fish, highlighting major changes in the regulation of these systems during inflammation.

Key words: Urea cycle, polyamine, salmonids, genome duplication, immune, arginase
1. Introduction

The ornithine-urea cycle, first discovered by Krebs and Henseleit [1], is central to the metabolism of arginine and the excretion of nitrogenous waste. In fish, most nitrogenous waste is excreted as ammonia through the gills, with <10% excreted as urea in rainbow trout [2]. There are four main enzymes directly involved in the urea cycle, arginase, ornithine transcarbamylase, argininosuccinate synthase and argininosuccinate lyase, which metabolise and recycle arginine, ornithine, citrulline and argininosuccinate [3]. Arginine is a versatile amino acid with functional roles including the modulation of protein deposition, production of ornithine for polyamine synthesis, regulation of immune responses through nitric oxide (NO) production, and removal of nitrogenous waste [4]. Arginine also stimulates the release of insulin, glucagon and growth hormone in fishes, which may regulate metabolism and growth [5].

![Pathway diagram of urea cycle, nitric oxide cycle and polyamine synthesis.](image)

Figure 1. Pathway diagram of urea cycle, nitric oxide cycle and polyamine synthesis. Amino acids are coloured in red boxes, enzymes in green, polyamines in orange, co-substrates in pink and molecules as dark blue. Enzyme acronyms are as follows: ARG, arginase; OTC, ornithine transcarbamylase; ASS, argininosuccinate synthase; ASL, argininosuccinate synthase; CPS, carbamoyl phosphate synthase; INOS, nitric oxide synthase; SAMdc, S-adenylosylmethionine decarboxylase; ODC, ornithine decarboxylase; SPDS, spermidine synthase; SPMS, spermine synthase. Co-substrate acronyms are as follows: SAM, s-adenosylmethionine; DC-SAM, decarboxylated s-adenosylmethionine.

During the urea cycle (Fig. 1), arginine is converted to ornithine by the arginase (ARG) enzymes, resulting in urea as the by-product. Ornithine can then be converted into citrulline by ornithine transcarbamylase (OTC), used in proline synthesis by the action of ornithine aminotransferase, or for polyamine synthesis, where ornithine decarboxylase (ODC) converts ornithine into the polyamine putrescine [6]. Recycling citrulline back into arginine is
a two-step process that involves two further enzymes (ASS and ASL), as well as argininosuccinate as an intermediate [7]. Citrulline and argininosuccinate also form part of the NO cycle (Fig. 1), as citrulline is generated as a by-product following conversion of arginine to NO through the action of the NO synthase enzymes [8, 9].

An important set of molecules derived from the urea cycle are the polyamines, containing two or more amine groups (-\(\text{NH}_2\)). The diverse functions of polyamines include regulation of protein synthesis [10], modulation of ion channels [11] and DNA and RNA binding [12]. They are also crucial for cellular proliferation and inflammatory responses [13], acting as antioxidants and offering cellular protection through inhibition of inflammatory mediators [14]. Polyamines are derived from ornithine by the activity of ODC and s-adenosylmethionine decarboxylase (SAMdc), with the simplest polyamine being putrescine, which can be further processed to spermidine and spermine (Fig. 1) [7]. During the immune response, high polyamine levels are found in rapidly proliferating cells and tissues [15, 16] supporting wound and tissue healing following infection or injury [6, 17]. The role of arginine and urea cycle products is attracting attention due to their roles in mediating immune functions. Inflammatory responses are associated with polarising T\(_{\text{H}}\)1 cytokines including IFN-\(\gamma\)- or TNF\(\alpha\)-activating M1 macrophages (‘kill’ macrophages), whereas anti-inflammatory processes activate M2 macrophages (‘healing’ macrophages) associated with T\(_{\text{H}}\)2 cytokines including interleukin 4 and 10 (IL-4, IL-10) [18]. M1 macrophages metabolise arginine into NO via inducible nitric oxide synthase (iNOS), resulting in a macrophage population with increased microbicidal activity [19]. On the other hand, anti-inflammatory responses and healing is associated with M2 cells, where arginine is converted to ornithine and subsequently metabolised to polyamines through ODC [20]. As M1 and M2 macrophages compete for arginine, the expression of either iNOS or arginase has a reciprocal regulatory effect [21]. Arginine is the sole precursor of NO and supplementation is known to increase NO synthesis in mammals [22].

Arginine is a functional feed that can modulate health and performance parameters in farmed fish. For example, arginine supplementation may enhance growth in Atlantic salmon during the transition from fresh to seawater [23], improve immune status in carp [24] and when combined with glutamine, enhance growth and feeding efficiency in Nile tilapia [25]. Despite such recent interest, little is known about the gene families encoding the enzymes involved in arginine metabolism and their response to disease and inflammation in fishes.

The first objective of this study was to identify and characterise the genes encoding the main urea cycle enzymes and the two rate limiting enzymes in polyamine synthesis in rainbow trout, including any paralogues retained during salmonid evolutionary history. The second
objective was to establish the mRNA expression of these genes in rainbow trout, performed
in a panel of tissues under control conditions, and following a bacterial pathogen challenge
in liver and muscle. The resultant data implies an important role for arginine in both
inflammation and tissue repair.

2. Materials and Methods

2.1. Animal work

All procedures described hereafter were carried out in compliance with the Animals
(Scientific Procedures) Act 1986 under UK Home Office license PPL number 70/8071 and
approved by the ethics committee at the University of Aberdeen. Juvenile rainbow trout were
purchased from College Mill Trout Farm (Perthshire, U.K.). The fish were kept at the
University of Aberdeen aquarium facility (School of Biological Sciences) in 400 L tanks at a
stocking density of <20 kg/m³. Tanks were supplied with recirculating freshwater with a flow
rate of 1.5 L/s. Fish were kept at a temperature of 14 ± 1°C and a photoperiod of 12:12
light:dark. A computerised control system was used to monitor pH, ammonia concentration
and oxygen levels. Fish were fed ad libitum daily with commercial pellets.

To assess candidate gene expression responses across tissues, n = 4 adult rainbow trout
(499 ± 54 g mean ± SEM) were used to sample a standard panel of tissues; gill, distal
intestine, heart, head kidney, liver, fast-twitch skeletal muscle and spleen (within 5 minutes
of death). Tissues were stored in 1.5 ml RNA later at 4°C for 24 h followed by long term
storage at -80°C.

For the bacterial immunological stimulation, fish (as described above) were anaesthetised by
immersion in 2-phenoxyethanol and then injected intraperitoneally (ip) with either phosphate
buffered saline (PBS) (0.5 ml/fish) or the pathogenic Hooke strain of the live Gram-negative
bacterium *Aeromonas salmonicida* (AS) (1.6 x 10⁶ ml⁻¹ cells, 0.5 ml/fish). After ip injection,
the fish were maintained in 400 L tanks in the University of Aberdeen’s freshwater challenge
facility for 48 h. After 48 h, n=10 fish from both PBS and AS groups were randomly sampled
and killed as previously described and both liver and fast-twitch skeletal muscle tissue
sampled and stored in RNA later as described above.

2.2. Sequence, phylogenetic and genomic analysis

Putative protein-coding nucleotide sequences (cds) for candidate genes were originally
obtained from the rainbow trout genome hosted at https://www.genoscope.cns.fr/trout/ (NCBI
This was achieved using BLASTn searches with human orthologues downloaded from NCBI as the query: ODC (AH002917.2), SAMdc (NM_001634.5), ARG1 (NM_001244438.1), ARG2 (NM_001172.3), OTC (NM_0005315), ASS (AH002610.2) and ASL (M14218.1). These sequences were also consistent when compared to the rainbow trout genome deposited on NCBI (GCF_002163505.1). For additional phylogenetic analysis, further vertebrate (coding) CDS sequences were retrieved from NCBI and/or Ensembl [27] databases. Protein sequences from Atlantic salmon, rainbow trout, northern pike, zebrafish, spotted gar, chicken, mouse and human were retrieved and accession numbers for all protein sequences are displayed in Supplementary Table 1. MatGat [28] was used to predict amino acid identity/similarity between all vertebrate proteins for each gene. Protein sequences were aligned using ClustalW in the MEGA7 software [29]. A phylogenetic tree was constructed using the Jones-Taylor-Thornton model in the maximum likelihood method in MEGA 7, bootstrapped 500 times.

Intron-exon structure and gene synteny analysis was carried out for all candidate gene families. To determine the genomic neighbourhood around candidate genes and the conservation of gene order across the same species mentioned earlier, genes were manually examined in NCBI's genomic region browser. Intron-exon structures were determined from the same databases used to retrieve CDS sequences.

2.3. Primer design for quantitative PCR (qPCR)

Due to the duplicated nature of salmonid genomes, care was taken to design paralogue-specific primers for genes with more than one copy within the genome (details of all primers in Table 1). Nucleotide mRNA sequences were therefore aligned with ClustalOmega [30] in order to compare paralogues and identify distinguishing regions of sequence. To avoid amplification of genomic DNA, primers were designed to span an intron-exon junction or were placed in different exons. Primers were also designed to have an annealing temperature of ~64°C judged from OligoCalc [31] and a product length of between 100 and 330 bps. Prior to qPCR analysis (section 2.5), confirmation of PCR products generated using the paralogue-specific primers was carried out by cloning and sequencing. For confirmatory sequencing the PCR products were ligated into pGEM-T easy cloning vector (Promega) and then transformed into competent Escherichia coli cells (JM109). Plasmid DNA was isolated by Qiagen mini prep kits as described by manufacturer. A minimum of 5 clones per paralogue were sent for Sanger sequencing, carried out by Eurofins.
### Table 1. Rainbow trout primer sequences used for qPCR with NCBI accession numbers

<table>
<thead>
<tr>
<th>Gene</th>
<th>Sense</th>
<th>Primer 5'-3'</th>
<th>Product size</th>
<th>Annealing temperature</th>
<th>Accession</th>
</tr>
</thead>
</table>

1. References: [32], [33], [34]
### 2.4. RNA extraction and reverse transcription

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward</th>
<th>Reverse</th>
<th>Length</th>
<th>Accession</th>
</tr>
</thead>
<tbody>
<tr>
<td>EF-1α</td>
<td>CAAGGATATCCGTGTGGCA</td>
<td>ACAGCGAAACGACAAAGG</td>
<td>327</td>
<td>NM_001124339.1</td>
</tr>
<tr>
<td>β-actin</td>
<td>ATGGAGATGAAATCGGCCC</td>
<td>TGCCAGATCTTCTCCATG</td>
<td>260</td>
<td>XM_021595779.1</td>
</tr>
<tr>
<td>HPRT</td>
<td>CCGCCTAAGAGCTAAT</td>
<td>GTCTGGAAGCTGTCACCTA</td>
<td>237</td>
<td>XM_021583468.1</td>
</tr>
<tr>
<td>RPS29</td>
<td>GGTCATAGCAGCTCTATGG</td>
<td>AGTCAGCTTAAACAAACGAT</td>
<td>167</td>
<td>XM_021612450.1</td>
</tr>
<tr>
<td>SAA</td>
<td>TATGGAGCTGAGCAAGGAC</td>
<td>CGTCCCAGTGTTAGCTT</td>
<td>137</td>
<td>NM_001124436.1</td>
</tr>
<tr>
<td>HAMP</td>
<td>AGGAGGTGGAAGCATTGACAG</td>
<td>GTGGCTCTGACGCTTGGACCTA</td>
<td>101</td>
<td>XM_021595153.1</td>
</tr>
<tr>
<td>ODC1</td>
<td>CGTGTGCGACGCTAGTGC</td>
<td>CCATGTCAAAGACACAGCG</td>
<td>179</td>
<td>XM_021574142.1</td>
</tr>
<tr>
<td>ODC2</td>
<td>TGGTGCCACCTGAAGGCC</td>
<td>ATGGGCGCTTGCTGTAAGT</td>
<td>128</td>
<td>XM_021585068.1</td>
</tr>
<tr>
<td>SAMdc1</td>
<td>GCAAGGACAAGCTAATT</td>
<td>AACCTTGGGATGTTAGCAG</td>
<td>185</td>
<td>XM_021600286.1</td>
</tr>
<tr>
<td>SAMdc2</td>
<td>AACCTGAGGATGCTCGGAG</td>
<td>AACCTTGGGATGTTAGCAG</td>
<td>121</td>
<td>XM_021611778.1</td>
</tr>
<tr>
<td>ARG 1A</td>
<td>AGCACATATCCCTGACGTTG</td>
<td>CATCGATGTACATACAGCG</td>
<td>147</td>
<td>XM_021564871.1</td>
</tr>
<tr>
<td>ARG 1B</td>
<td>GTGGATGCGCTTCATGAC</td>
<td>CATCGATGTCATAGGTCGG</td>
<td>179</td>
<td>KX99866.1</td>
</tr>
<tr>
<td>ARG 2A</td>
<td>TCCAGAGAGTCATGGAAGTTTCTTGT</td>
<td>CCATGACTGACAAACCTGCTG</td>
<td>198</td>
<td>KX998967.1</td>
</tr>
<tr>
<td>ARG 2B</td>
<td>CTTGGTTAGGTCAAACCCAG</td>
<td>GTGAATGCGTTCGTTCG</td>
<td>163</td>
<td>KX998968.1</td>
</tr>
<tr>
<td>OTC</td>
<td>CACAGGCCAGGTGCTCCTTC</td>
<td>CATCGATGTCATAGGTCGG</td>
<td>116</td>
<td>XM_021597830.1</td>
</tr>
<tr>
<td>ASS</td>
<td>TGAGATTGAGGAGGAGCAGT</td>
<td>GCCCTGTGTTAGCTCCTGA</td>
<td>172</td>
<td>XM_021590913.1</td>
</tr>
<tr>
<td>ASL</td>
<td>ACAGCTTCCAACCTACACA</td>
<td>ACCGCATGACTGAGAATCCA</td>
<td>129</td>
<td>XM_021563243.1</td>
</tr>
</tbody>
</table>
Total RNA was extracted from 100mg of tissue homogenised in 1ml of TRI Reagent (Sigma-Aldrich) following the manufacturer’s instructions. The concentration and purity of RNA was estimated using a NanoDrop 1000 Spectrophotometer (Thermo Scientific). The integrity of RNA was assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies). First-strand cDNA was synthesized from 1 µg total RNA using a QuantiTech Reverse Transcription kit (QIAGEN), with an integrated genomic DNA elimination step, as per the manufacturer’s guidelines in a total volume of 20 µl. First strand cDNA samples were diluted 20-fold (working stock) with RNase/ DNase free water (Sigma-Aldrich) and stored at -20°C until use.

2.5. Quantitative gene expression analysis

qPCR analyses were performed with SYBR Green I dye chemistry using an Mx3005P System (Agilent Technologies). All assays were carried out in duplicate within 96 well plates using 15 µl reactions containing 5 µl of the 1:20-diluted cDNA (corresponding to 2.5 ng of reverse-transcribed total RNA), 500 nM sense/antisense primers and 7.5 µl Brilliant III Ultra-Fast SYBR Green (Agilent Technologies). The PCR cycling conditions were 1 cycle of 95°C for 3 minutes, followed by 40 cycles of 95°C for 20 seconds then 64°C for 20 seconds (two step PCR). Melting curve analysis (thermal gradient from 55°C to 95°C) was used to confirm the amplification of a single product. Each plate also included no-template controls in duplicate (cDNA replaced with water). Threshold fluorescence was set at 2500 during the linear phase of amplification. The efficiency of each qPCR assay was assessed using LinRegPCR quantitative PCR data analysis program (download: http://LinRegPCR.HFRC.nl) following Ruijter et al [35] recommendations. Expression data was then imported and analysed in Genex 5.4.3 (MultiD Analysis). Candidate gene expression was normalised to two reference genes for the tissue distribution (EF-1α and HPRT) and three genes for the infection study on muscle (EF-1α, HPRT and RPS29) and liver (EF-1α, ACTB, and HPRT). ACTB was replaced with RPS29 in muscle due to ACTB’s instability in muscle determined from the dissociation graph following qPCR. All reference gene primers used in the study are presented in Table 1.

2.6. Statistical analysis

Statistical analysis of qPCR data was performed in R (v3.4.0). A linear model (lm) was first constructed in R and the diagnostic plots (qq plot and residuals versus fitted values) were assessed in order to ensure both normality and equal variance. If data met the assumptions, the one-way ANOVA results from R’s lm function could then be interpreted. If data was not
normal, a log transformation was first performed and the diagnostics plots then reassessed.

If the data still did not meet the models assumptions following the transformation, a non-parametric test (Kruskal-Wallis) was then performed on the data.
3. Results

3.1 Comparative analysis of the urea cycle genes

The phylogenetic relationships of all genes were carried out in comparison to a range of representative vertebrate lineages. Two copies of \textit{ARG1}, \textit{ARG2}, \textit{SAMdc} and \textit{ODC} were identified, along with single copies of \textit{OTC}, \textit{ASS} and \textit{ASL}. Gene intron-exon structure, amino acid sequence analysis and synteny were performed on \textit{ODC} as a representative example.

Two \textit{ODC} paralogues were identified in the rainbow trout genome (\textit{ODC1} accession: XM\_021574142, Chr19, LOC110498573, \textit{ODC2} accession: XM\_021585068.1, Chr25, LOC110505682). The open reading frame for \textit{ODC1} and \textit{ODC2} encoded 457 amino acids (aa) and 456 aa respectively, and a conserved gene structure of 8 exons and 7 introns (Fig. 2) was evident between the paralogues. Across species, exons 2 and 3 were identical lengths for all species examined, while exon 4 was conserved in all teleosts, whereas in the tetrapods an additional intron is present resulting in tetrapod exons 4 and 5 being the same length as teleost exon 4 (Fig. 2). Similarly, within salmonid exon 5, an additional intron can be found in the remaining species analysed (Fig. 2). Exons 6 and 7 in salmonids and corresponding exons in the rest of the vertebrate species are highly conserved with tetrapods having 3 nucleotides less in relation to salmonid exon 6 and all species having identical length in salmonid exon 7 (Fig. 2).
Phylogenetic analysis of the ODC1 and 2 proteins revealed that each molecule has an orthologue in both Atlantic salmon and rainbow trout that shared ~98% aa identity (Table 2). Within species, both rainbow trout and salmon ODC1 compared to ODC2 has 93.4% and 94.7% identity respectively (Table 2). The presence of the genes on distinct chromosomes and the branching of northern pike as a sister group to both ODC1 and 2 suggests these genes are products of the salmonid specific WGD (ssWGD) (Fig. 3). ODC1 and 2 are highly conserved with other vertebrate ODC orthologues (Table 2), for example sharing ~73% aa level identity each with human.
<table>
<thead>
<tr>
<th></th>
<th>O. mykiss ODC1</th>
<th>S. salar ODC1</th>
<th>O. mykiss ODC2</th>
<th>S. salar ODC2</th>
<th>E. lucius ODC</th>
<th>D. rerio ODC</th>
<th>L. oculatus ODC</th>
<th>G. gallus ODC</th>
<th>M. musculus ODC</th>
<th>H. sapiens ODC</th>
</tr>
</thead>
<tbody>
<tr>
<td>O. mykiss ODC1</td>
<td>98.2</td>
<td>93.2</td>
<td>94.5</td>
<td>92.4</td>
<td>85</td>
<td>77.5</td>
<td>71.4</td>
<td>71.2</td>
<td>72.5</td>
<td></td>
</tr>
<tr>
<td>S. salar ODC1</td>
<td>99.3</td>
<td>93.4</td>
<td>94.7</td>
<td>92.8</td>
<td>84.6</td>
<td>77.9</td>
<td>70.8</td>
<td>71</td>
<td>72.3</td>
<td></td>
</tr>
<tr>
<td>O. mykiss ODC2</td>
<td>66.1</td>
<td>95.1</td>
<td>96.2</td>
<td>95.9</td>
<td>93.4</td>
<td>85.9</td>
<td>79</td>
<td>71.4</td>
<td>72.1</td>
<td>72.9</td>
</tr>
<tr>
<td>S. salar ODC2</td>
<td>96.9</td>
<td>97.1</td>
<td>98.9</td>
<td>94.5</td>
<td>92.6</td>
<td>85.5</td>
<td>79.4</td>
<td>71</td>
<td>72.1</td>
<td>72.5</td>
</tr>
<tr>
<td>E. lucius ODC</td>
<td>95.4</td>
<td>96.1</td>
<td>95.2</td>
<td>95.9</td>
<td>87</td>
<td>79.2</td>
<td>71.7</td>
<td>72.5</td>
<td>72.7</td>
<td></td>
</tr>
<tr>
<td>D. rerio ODC</td>
<td>90.7</td>
<td>91.8</td>
<td>92</td>
<td>92</td>
<td>93.1</td>
<td>80.1</td>
<td>73.7</td>
<td>73.9</td>
<td>74.1</td>
<td></td>
</tr>
<tr>
<td>L. oculatus ODC</td>
<td>87.6</td>
<td>88</td>
<td>88.2</td>
<td>87.3</td>
<td>87.6</td>
<td>89.7</td>
<td>79.4</td>
<td>76.8</td>
<td>78.3</td>
<td></td>
</tr>
<tr>
<td>G. gallus ODC</td>
<td>83.8</td>
<td>83.6</td>
<td>82.5</td>
<td>82.5</td>
<td>82.6</td>
<td>86.4</td>
<td>89.3</td>
<td>81.5</td>
<td>83.9</td>
<td></td>
</tr>
<tr>
<td>M. musculus ODC</td>
<td>84.6</td>
<td>85.2</td>
<td>85.2</td>
<td>84.8</td>
<td>84.2</td>
<td>87</td>
<td>88.4</td>
<td>91.2</td>
<td>90.7</td>
<td></td>
</tr>
<tr>
<td>H. sapiens ODC</td>
<td>83.9</td>
<td>84.6</td>
<td>83.5</td>
<td>83.7</td>
<td>83.3</td>
<td>85.5</td>
<td>89.5</td>
<td>92.9</td>
<td>94.1</td>
<td></td>
</tr>
</tbody>
</table>

Identity

Similarity

Table 2. Comparison of amino acid identities and similarities (%) for ODC using the similarity matrix in MatGat 2.02 software. Accession numbers for all proteins can be found on Supplementary Table 1.

Phylogenetic analysis further confirmed ODC1, ODC2 (Fig. 3) and SAMdc1, SAMdc2 (Supplementary Fig. 1) as products of ssWGD with higher relatedness to their salmonid relative than to their own species duplicate. In the case of the ARG1 and ARG2, these genes are present in all vertebrates and likely diverged before the evolution of vertebrates. However as there are 4 copies of arginase present in salmonids (ARG1a, ARG1b, ARG2a and ARG2b), it can be seen that the two paralogues for arginase 1 and 2 are likely products of the ssWGD (Supplementary Fig. 2). All other genes characterised in this study (OTC, ASS, ASL) have not retained a duplicate copy (Supplementary Figures 3-5).
Synthesis revealed highly conserved gene order in the genomic regions containing ODC1 and 2, both in relation to each other and also in comparison to single copy ODC orthologues in non-salmonid teleosts and tetrapod vertebrates (Fig. 4). At the 5' of ODC, the most proximal neighbouring gene (NOL10) was retained in the same location in all vertebrates. We also identified several other annotated genes with high synteny located near ODC including KCNA1 and ROCK2. Towards the 3' end of ODC, salmonid genes shared similarity of gene order with the northern pike, but with no other vertebrates including zebrafish, suggesting a

Figure 3. Maximum likelihood phylogenetic tree showing the evolutionary relationship of ODC amino acid sequences in vertebrates. The branch support values were gained by non-parametric bootstrapping (500 replicates). The scale bar represents the number of modelled substitutions per site. Accession numbers from NCBI are as follows: O. mykiss ODC1 (XP_021429817.1), O. mykiss ODC2 (XP_021440743.1), S. salar ODC1 (XP_014047562.1), S. salar ODC2 (XP_014066501.1), E. lucius ODC (XP_010890677.1), L. oculatus ODC (XP_006626107.1), D. rerio ODC (NP_571876.1), G. gallus ODC (NP_001161238.1), M. musculus ODC (NP_038642.2), H. sapiens ODC (AAA59968.1)

Figure 4. Phylogenetic tree and gene synteny of ODC in vertebrates. The tree was constructed using the maximum likelihood method in MEGA7 and bootstrapped 500 times. The syntenically conserved gene blocks are shown in matching colours. The arrows represent transcriptional direction. Gene synteny was compiled from up and downstream locations relative to each species ODC taken from NCBI, ODC protein accession numbers on supplementary table 1, chromosome number and range (from left of the diagram to right) as follows: O. mykiss ODC1 (Chr 19, 34,937,046 > 35,209,567), O. mykiss ODC2 (Chr 25, 63,402,440 > 63,152,107), S. salar ODC1 (Chr ssa01, 29,382,254 > 29,167,678 ), S. salar ODC2 (Chr ssa09, 27,309,511 > 27,043,274), E. lucius ODC (Chr LG15 , 14,186,061 > 14,321,347), D. rerio ODC (Chr 17, 51,655,239 > 51,757,548), L. oculatus ODC (Chr LG1, 43,496,860 > 43,340,082), G. gallus ODC (Chr 3, 97,175,730 > 96,773,203), M. musculus ODC (Chr 12, 16,894,978 > 17,791,944), H. sapiens ODC (Chr 2, 11,179,759 > 10,427,617).
chromosomal rearrangement prior to the salmonid / pike divergence from other teleosts. As for the genes downstream of ODC, we identified some annotated genes in two copies on both the ODC1 and 2 chromosomes, barring the YIPF family of genes. Towards the 3' of rainbow trout ODC1 there was no YIPF gene annotated, but in Atlantic salmon YIPF was identified downstream of both ODC1 and 2.

3.2. Urea cycle and polyamine synthesis gene expression

3.2.1 Tissue distribution

The relative mRNA expression levels of the characterized urea cycle (ARG1a, ARG1b, ARG2a, ARG2b, OTC, ASS and ASL) and rate limiting enzymes of polyamine synthesis (ODC1, ODC2, SAMdc1 and SAMdc2) were quantified by qPCR in seven tissues in healthy rainbow trout under control conditions (Fig. 5 for ODC1 and 2; all other genes: supplementary figures 6-8). Both ODC paralogues were expressed in all tissues examined, with ODC2 less abundant than ODC1 in all tissues barring gill (Fig. 5). ODC1 was most highly expressed in head kidney followed by spleen and gill. ODC2 had the highest expression in gill followed by head kidney.

The four arginase paralogues showed a more variable expression distribution (Supplementary Fig. 6). ARG1a and ARG1b both showed highest expression in the liver (Supplementary Fig. 6). ARG1a was expressed in all tissues examined, with highest expression in liver and gill. ARG1b showed highest expression in liver, with lower expression levels in gill, heart and muscle, and no detectable expression in gut, head kidney or spleen expression. ARG2a and ARG2b were expressed in all tissues, both showing highest expression in the muscle and lowest expression in liver.

Figure 5. Tissue distribution of ODC1 (A) and ODC2 (B) in rainbow trout. The relative expression in each tissue was normalised with the expression of two housekeeping genes EF-1α and HPRT. Bars represent mean (± SEM), n=4.
The single copy urea cycle genes *OTC*, *ASS* and *ASL* were expressed in all the tissues examined, with *OTC* and *ASL* having highest expression in muscle and *ASS* in gill (Supplementary Fig. 7). *OTC* was expressed at a very low level in the liver (Supplementary Fig. 7). The two *SAMdc* paralogues were expressed in all tissues, with both genes having highest expression in heart and muscle (Supplementary Fig. 8).

### 3.2.2. Modulation of mRNA expression following bacterial infection

Expression of the urea cycle enzymes and polyamine synthesis genes was examined in adult rainbow trout liver and muscle tissue sampled following an experimental infection with a pathogenic strain of the bacterium *Aeromonas salmonicida* (AS). To confirm the fish were undergoing an inflammatory response to the infection, the expression of two marker genes for the acute phase response (APR), serum amyloid A (*SAA*) and hepcidin (*HAMP*), were examined in control and AS-infected fish. Both genes showed highly significant upregulation in infected fish (Fig. 6) indicating a strong immune response.

Following AS infection both *ODC* paralogues were significantly upregulated in the liver (Fig. 7). In muscle, *ODC1* and 2 expression did not change following infection (Fig. 7). Both *SAMdc1* and 2 were significantly upregulated in liver but not muscle (Fig. 7). *ARG1a* was significantly downregulated in liver and unchanged in muscle whereas the opposite pattern
was observed for ARG1b (Fig. 8). ARG2a and ARG2b expression was significantly increased in liver, with ARG2b also significantly increased in muscle but not ARG2a (Fig. 8).

Both OTC and ASL’s expression was significantly increased in muscle, but not liver (Fig. 9), whereas ASL expression increased significantly in infected liver but not in muscle (Fig. 9).

Figure 7. Relative expression of the rate limiting enzymes in polyamine synthesis, SAMdc and ODC, in rainbow trout muscle and liver tissue following a bacterial infection. Fish were injected intraperitoneally with either phosphate buffered saline (PBS) or Aeromonas salmonicida (AS). The top row of graphs show gene expression in liver tissue: ODC1 (A), ODC2 (B), SAMdc1 (C), SAMdc2 (D), expression was normalised to housekeeping genes EF-1α, ACTB and HPRT. The bottom row of graphs show gene expression in muscle tissue: ODC1 (E), ODC2 (F), SAMdc1 (G), SAMdc2 (H), expression was normalised to housekeeping genes EF-1α, RPS29 and HPRT. Bars represent mean (± SEM), n=10, * = p < 0.05, ** = p < 0.01, *** = p < 0.001.

Figure 8. Relative expression of rainbow trout Arginase paralogues following a bacterial infection. Fish were injected intraperitoneally with either phosphate buffered saline (PBS) or Aeromonas salmonicida (AS). Top row represents gene expression of ARG1a (A), ARG1b (B), ARG2a (C) and ARG2b (D) in liver tissue, expression was normalised to the housekeeping genes EF-1α, ACTB and HPRT. Bottom row represents gene expression of ARG1a (E), ARG1b (F), ARG2a (G) and ARG2b (H) in muscle tissue, expression was normalised to the housekeeping genes EF-1α, RPS29 and HPRT. Bars represent mean (± SEM), n=10, * = p < 0.05, ** = p < 0.01, *** = p < 0.001.
keeping genes EF-1α, RPS29 and HPRT. Bars represent mean (± SEM), n=10, * = p < 0.05, ** = p < 0.01, *** = p < 0.001

Figure 9. Relative expression of rainbow trout OTC (A), ASS (B) and ASL (C) in liver tissue following a bacterial infection. Second row represents relative expression of rainbow trout OTC (D), ASS (E) and ASL (F) in muscle tissue. Fish were injected intraperitoneally with either phosphate buffered saline (PBS) or Aeromonas salmonicida (AS). For the genes examined in the liver, expression was normalised to housekeeping genes EF-1α, ACTB and HPRT. For the bottom panel where muscle tissue was examined, expression was normalised to housekeeping genes EF-1α, RPS29 and HPRT. Bars represent mean (± SEM), n=10, * = p < 0.05, ** = p < 0.01, *** = p < 0.001
4. Discussion

Arginine metabolism and the urea cycle are major components of nitrogen metabolism, the inflammatory response and subsequent tissue repair. However, there is a lack of information available in salmonid fish regarding the genes regulating these metabolic pathways. Here we address this knowledge gap by characterising the key enzymes in the urea pathway and by documenting their transcriptional responses to a bacterial infection.

In fish, arginine is often regarded as an essential amino acid required for efficient protein synthesis, in addition to other functional processes influencing health status [5, 36]. For arginine to be synthesised endogenously, carbamoylphosphate, an intermediate molecule in the urea cycle, is combined with ornithine to form citrulline through the action of OTC (Fig. 1). In ureotelic mammals and amphibians, carbamoylphosphate synthetase (CPS) catalyses the formation of carbamoylphosphate, allowing its use in the urea cycle [37, 38]. In rainbow trout, CPS is expressed at early life stages but not in the liver of adults [37] with low levels in muscle [39]. This lack of hepatic CPS activity in salmonids could be the reason for an incomplete urea cycle and inability to synthesise arginine endogenously. Some teleosts including the toadfish Opsanus beta [40], the catfish Clarias batrachus [41], and the lungfish Protopterus aethiopicus [42], have detectable CPS activity and a functional urea cycle. However, our gene expression data suggests that the urea cycle enzymes are functional in adult rainbow trout and the modulation we observed following AS infection indicates a role during the inflammatory response.

Our phylogenetic analyses revealed that ODC, SAMdc and ARG have multiple paralogues retained from the salmonid-specific whole genome duplication (ssWGD), which occurred 88-103 Mya [43, 44]. In the case of ARG, we observed four salmonid copies, with two paralogues retained for both ARG1 and ARG2, which was also recently shown for Atlantic salmon [20], which are conserved in all vertebrates and presumably the result of an early gene duplication event, perhaps past WGD events in the vertebrate ancestor [45]. However, we did not identify any duplicated paralogues of the key urea cycle genes retained from the teleost-specific WGD event [45]. The duplicated copies that are retained either specialise in function (subfunctionilization) or develop a novel function (neofunctionilization) [46, 47]. Following a WGD event, the resulting duplicated genome eventually only retains a small percentage of duplicated genes, as the redundant genes are inactivated by a process termed gene fractionation [26]. The ssWGD is relatively recent when compared to the teleost specific WGD (~100Mya compared to ~300Mya) and this is evident from the large number of duplicated genes still present in the genome (48% of genes with retained ohnologs) [26].
The genes encoding *OTC*, *ASS* and *ASL* were found to be present as single copy genes suggesting one duplicated copy was lost in the ancestor of the trout and salmon lineage. The genes encoding the urea cycle molecules *ARG*, *OTC*, *ASS* and *ASL* were expressed in all tissues examined, with *OTC* showing negligible levels in liver; this latter observation may contribute to a low functioning urea cycle in salmonids, and other teleost's [37]. However, genes encoding the other urea cycle enzymes (*ARG*, *ASS* and *ASL*) were transcribed in the liver, indicating components of the urea cycle may have other metabolic roles beyond nitrogenous waste excretion. The expression of *ASS* and *ASL*, may indicate an efficient conversion of citrulline to arginine in adult trout, but the lack of the carbamoyl substrate is a limiting factor. If the urea cycle is fully functioning, then citrulline could be used to bolster arginine levels, as is found in mammalian species [48, 49]. This idea is also supported by the functioning NOS cycle in salmonids as they are able to produce NO in the innate immune response against pathogens [50]. Citrulline is generated as a by-product of the iNOS reaction and could be recycled back into arginine by ASS and ASL. The arginase genes show an interesting expression profile, *ARG 1a* and *1b* are expressed at high levels in the liver whereas *ARG 2a* and *2b* are virtually absent in liver tissue. For genes encoding polyamine enzymes there is low level of expression in non-stimulated liver and small difference between paralogue expression.

To gain information on the expression of the urea cycle genes/paralogues, we quantified their mRNA expression and transcriptional responses following bacterial infection in muscle and liver. We sampled liver as a key indicator of the acute phase response [51, 52] that shows a well-established response to bacterial infection, while also acting as the main site for amino acid metabolism and the urea cycle. Skeletal muscle was also selected for analysis, due to high transcript levels identified in many of the genes in the tissue distributions (see Results section 3.2.1). Fish were sampled 48h after infection to represent the early immune response before physiological changes due to disease could occur [51]. There were significant increases in liver expression for the polyamine genes *ODC1, 2* and *SAMdc1 and 2* suggesting upregulation of the polyamine pathway and subsequent production of putrescine. This could be related to cellular repair and also potential increased availability of ornithine from increased *ARG2* activity following infection.

Both the *ARG2* paralogues increased in expression following infection, whilst *ARG1a* was significantly decreased showing potential subfunctionalisation of the duplicated genes. Recent research has suggested that *ARG1* is involved as a major metaboliser of hepatic arginine whereas *ARG2* may be more involved with the immune response in the form of healing M2 macrophages [19]. During an immune response M2 macrophages demonstrate
elevated levels of arginase activity and also play an important role in the innate immune
defence against various pathogens in both a bactericidal and healing sense [18]. The two
major types of macrophages, M1 and M2, both depend on the same substrate (arginine) for
either healing (M2) or bactericidal activity (M1). The enzymes iNOS and arginase have been
described as useful markers for M1 and M2 macrophages (respectively) in both mammals
[53] and some fish species [54]. As both iNOS and arginase compete for arginine they can
regulate each other’s expression either driving an inflammatory response via the nitric oxide
cycle or wound healing from polyamine synthesis by ODC [53]. The increased expression of
ARG2 along with ODC and SAMdc following infection shown in this study suggests that the
conversion of arginine into polyamines is taking place for tissue repair. This is in agreement
with studies in humans where arginine derived from ornithine also plays a role in tissue
remodelling as high levels of arginase can be observed in fibroblasts of patients suffering
from pulmonary fibrosis [55]. Further evidence showing the regulatory effect ARG and iNOS
have on each other can be seen in mice when infected with Helicobacter pylori, arginase 2
knockout lead to increased M1 macrophage activation [56].

The transcriptional changes seen in the liver from ARG2 and the polyamine synthesis genes
ODC and SAMdc suggests a signature activation of wound healing M2 macrophages [21].
Although macrophages are virtually present in all tissues [20], we find in muscle, that during
inflammation there is a less dramatic response compared to liver. There were no significant
changes in the muscle for genes involved in polyamine synthesis suggesting liver is a major
source for these molecules. For ARG genes there was a similar pattern in muscle and liver
indicating conserved regulation between tissues and a conserved inflammatory response.
The urea cycle enzymes (OTC, ASS and ASL) displayed some variation between tissue
where OTC and ASL were significantly upregulated in muscle but unchanged in liver while
ASS was significantly increased in liver but unchanged in muscle.

5. Conclusion
In summary, the genes encoding the enzymes of the urea cycle and the two rate limiting
enzymes in polyamine synthesis have been characterised and their response to infection
investigated. Our findings demonstrate that ARG 1 and 2, SAMdc and ODC genes have
retained functional paralogues from the salmonid-specific WGD, with several of the
duplicated copies showing different regulation across tissues. The nutritional requirement of
arginine in the diet in salmonids is likely to be due to a lack of activity from CPS and OTC
enzymes in adult liver. It is likely that half of the urea cycle is functional and the enzymes
responsible for the conversion of citrulline to arginine are active, especially due to the
functioning NO cycle. We also observed significant changes in the urea/polyamine pathways following bacterial challenge, suggesting enhanced recycling and metabolism of arginine for both inflammatory and tissue healing roles following infection.

Acknowledgements

The research was supported by a studentship to T. Clark funded between the University of Aberdeen and BioMar Ltd.

References


mRNA accompanies activation of human and mouse monocytes/macrophages.

*FEBS letters*, 268(1), pp. 32-34.


Highlights

- We characterised the urea cycle and polyamine synthesis enzymes in rainbow trout
- Both pathways are significantly modulated following *Aeromonas salmonicida* infection
- Polyamine genes upregulated in liver following challenge
- Duplicated arginase paralogues have differential responses to infection
- Arginine is recycled and metabolised for bactericidal and tissue healing roles