

RESEARCH ARTICLE

# Multi-tissue RNA-seq and transcriptome characterisation of the spiny dogfish shark (*Squalus acanthias*) provides a molecular tool for biological research and reveals new genes involved in osmoregulation

Andres Chana-Munoz<sup>1</sup>, Agnieszka Jendroszek<sup>2na</sup>, Malene Sønnichsen<sup>2nb</sup>, Rune Kristiansen<sup>3</sup>, Jan K. Jensen<sup>2</sup>, Peter A. Andreasen<sup>2†</sup>, Christian Bendixen<sup>1</sup>, Frank Panitz<sup>1\*</sup>

**1** Department of Molecular Biology and Genetics, Aarhus University, Tjele, Denmark, **2** Department of Molecular Biology and Genetics, Aarhus University, Aarhus, Denmark, **3** Kattegatcentret, Grenaa, Denmark

† Deceased.

na Current address: Institute for Bioscience – Zoophysiology, Aarhus University, Aarhus, Denmark

nb Current address: Interdisciplinary Nanoscience Center - INANO-MBG, Aarhus University, Aarhus, Denmark

\* [frank.panitz@mbg.au.dk](mailto:frank.panitz@mbg.au.dk)



OPEN ACCESS

**Citation:** Chana-Munoz A, Jendroszek A, Sønnichsen M, Kristiansen R, Jensen JK, Andreasen PA, et al. (2017) Multi-tissue RNA-seq and transcriptome characterisation of the spiny dogfish shark (*Squalus acanthias*) provides a molecular tool for biological research and reveals new genes involved in osmoregulation. PLoS ONE 12(8): e0182756. <https://doi.org/10.1371/journal.pone.0182756>

**Editor:** Peng Xu, Xiamen University, CHINA

**Received:** April 28, 2017

**Accepted:** July 24, 2017

**Published:** August 23, 2017

**Copyright:** © 2017 Chana-Munoz et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** The datasets generated during the current study including raw reads and transcriptome assemblies are available in the European Nucleotide Archive (ENA) database at <http://www.ebi.ac.uk/ena/data/view/PRJEB14721>.

**Funding:** This study was funded by the Danish Council for Independent Research | Natural Sciences [<http://ufm.dk/en/research-and->

## Abstract

The spiny dogfish shark (*Squalus acanthias*) is one of the most commonly used cartilaginous fishes in biological research, especially in the fields of nitrogen metabolism, ion transporters and osmoregulation. Nonetheless, transcriptomic data for this organism is scarce. In the present study, a multi-tissue RNA-seq experiment and *de novo* transcriptome assembly was performed in four different spiny dogfish tissues (brain, liver, kidney and ovary), providing an annotated sequence resource. The characterization of the transcriptome greatly increases the scarce sequence information for shark species. Reads were assembled with the Trinity *de novo* assembler both within each tissue and across all tissues combined resulting in 362,690 transcripts in the combined assembly which represent 289,515 Trinity genes. BUSCO analysis determined a level of 87% completeness for the combined transcriptome. In total, 123,110 proteins were predicted of which 78,679 and 83,164 had significant hits against the SwissProt and Uniref90 protein databases, respectively. Additionally, 61,215 proteins aligned to known protein domains, 7,208 carried a signal peptide and 15,971 possessed at least one transmembrane region. Based on the annotation, 81,582 transcripts were assigned to gene ontology terms and 42,078 belong to known clusters of orthologous groups (eggNOG). To demonstrate the value of our molecular resource, we show that the improved transcriptome data enhances the current possibilities of osmoregulation research in spiny dogfish by utilizing the novel gene and protein annotations to investigate a set of genes involved in urea synthesis and urea, ammonia and water transport, all of them crucial in osmoregulation. We describe the presence of different gene copies and isoforms of key enzymes involved in this process, including arginases and transporters of urea and ammonia, for which sequence information is currently absent in the databases for this

innovation/councils-and-commissions/the-danish-council-for-independent-research/the-council/the-danish-council-for-independent-research-natural-sciences] under the grant number 0602-02216B to P.A. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

model species. The transcriptome assemblies and the derived annotations generated in this study will support the ongoing research for this particular animal model and provides a new molecular tool to assist biological research in cartilaginous fishes.

## Background

Marine cartilaginous fishes such as *Squalus acanthias* (spiny dogfish) are ureotelic organisms which synthesize and excrete urea as the final product of the nitrogen metabolism. Besides, they are also ureosmotic organisms with urea being the main osmolyte. Consequently, they retain large amounts of urea in their plasma in order to increase the osmolarity and maintain their body fluids isosmotic or slightly hyperosmotic with respect to sea water. To achieve such high concentration of urea in their body, efficient pathways of urea synthesis and retention are crucial, with liver, kidneys, intestine and gills being the primary tissues involved in this process [1–10] reviewed in [11, 12].

There are two main biosynthetic pathways of urea synthesis in cartilaginous fishes: the ornithine urea cycle (O-UC) and the purine degradation pathway, the first being the predominant pathway [1, 7, 13]. Cartilaginous fishes reabsorb most of the urea in the nephrons from the primary urine and transport it to the blood vessels [2, 4, 12] while also reducing the loss of ammonia in the gills and kidney in order to synthesize urea [12, 14, 15]. This mechanism is in part facilitated by their unique kidney morphology coupled with the presence of several metabolite transporters such as urea transporters (UTs), Rh-glycoproteins, aquaporins as well as the  $\text{Na}^+$ ,  $\text{K}^+$  and  $\text{Cl}^-$  reabsorption levels [2, 12, 14, 16, 17].

In addition to urea, other osmolytes including  $\text{Na}^+$  and  $\text{Cl}^-$ , are also important to maintain the osmolarity in cartilaginous fishes. Therefore, euryhaline cartilaginous fishes such as spiny dogfish are able to modify their internal urea,  $\text{Na}^+$  and  $\text{Cl}^-$  concentration as a response to changes in the salinity of the environment [5, 18, 19]. This process is highly regulated with the involvement of several tissues including the rectal gland, kidneys, liver, intestine and the gills, reviewed in [19]. The rectal gland in particular has been studied in great detail, due to its capacity of excreting large amounts of NaCl. Its simple tissue structure, the presence of several ion transporters and an early development of standardized experimental protocols have made rectal gland a model system for transport of chloride and ion exchange [20–26].

A significant part of the current knowledge about osmoregulation comes from studies in spiny dogfish, confirming this shark species as a valuable animal model for osmoregulation in cartilaginous fishes [1, 4, 10, 14, 17, 27]. Additionally, spiny dogfish has been used in other scientific fields for example in the study of choroid plexus transport [28–30] and drug discovery, where squalamine, a compound discovered in the liver of this shark species with antimicrobial activity [31, 32] has been more recently reported to have additional therapeutic activities [33–35]. In addition, to improve the research in this species, embryo-derived cell lines have been established a decade ago [36–38].

Although spiny dogfish is used as a model in a variety of biological research studies, only limited sequence resources providing transcriptome data are publicly available [36, 39, 40]. These previous experiments relied on Sanger sequencing technology and as a consequence, at the time of writing this manuscript the existing data comprises only 32,562 ESTs for *S. acanthias* on record at the National Center for Biotechnology Information (NCBI). Apart from rectal gland tissue [39], transcript information on other individual tissues is lacking for spiny dogfish. Genomic and transcriptomic resources for this animal group and particularly

for sharks remain scarce even though there have been previous studies generating sequence data for cartilaginous fishes [39, 41–45]. To overcome these limitations, we performed extensive transcriptome sequencing of this shark using next generation sequencing (NGS) technology.

The aim of the present work was to establish a comprehensive characterization of the transcriptome of spiny dogfish based on high quality RNA-seq reads from kidney, liver, brain and ovary. One objective was to provide an extended molecular resource for the shark spiny dogfish while the other goal was to demonstrate how this novel data can be applied to investigate and improve insight into osmoregulation in this species. To ensure that both aspects, the more general one of generating a resource as well as the more focused one of creating data for specific interests like osmoregulation, were met we performed RNA-seq from kidney and liver with regard to the latter aspect while brain and ovary were chosen to provide a better description of the transcripts in this shark species. Following *de novo* assembly with Trinity [46, 47], the resulting assemblies were functionally annotated using the Trinotate pipeline (<http://trinotate.github.io/>). Finally, the utility of the novel sequence and annotation resource to improve knowledge in spiny dogfish sharks was shown by investigation of the presence of genes involved in osmoregulation.

## Results

### RNA-seq and data processing

About 600 million raw reads were generated for spiny dogfish using the Illumina HiSeq2000 platform. Ribosomal and mitochondrial reads were removed and after trimming about 380 million high quality paired-end reads longer than 50bp remained; see Table 1 for detailed reports.

### *De novo* assembly and assessment of the transcriptome assemblies

Using the Trinity assembler on the cleaned RNA-seq of the individual tissues, a total of 346,582 (brain); 316,366 (kidney); 226,007 (liver) and 291,104 (ovary) transcripts were generated. The combined assembly on the merged reads from all four tissues yielded in 705,797 transcripts. Aligning the read pairs back to the respective assemblies resulted in at least 70% of the reads mapping as proper pairs. In order to evaluate the completeness of the assemblies, the BUSCO pipeline was performed against a dataset of 3,023 conserved genes in vertebrates [48]. This analysis showed that the completeness varies between the assemblies reaching 87% completeness in the combined assembly. Finally, using blastx [49] the assemblies were aligned against the proteome of the holocephalan elephant shark (*Callorhynchus milii*), the only cartilaginous fish for which a genome assembly is available [41]. The number of significant unique hits (e-value cut-off  $1e-20$ ) in elephant shark that were covered by 70% of their length by a

**Table 1. Raw data processing.**

Tissue	Raw	non-rRNA	non rRNA nor mitochondrial	Clean
Brain	103,564,548	103,374,482	72,383,520	63,079,760
Kidney	184,599,338	184,054,918	122,041,030	107,029,448
Liver	244,987,738	244,816,454	190,393,898	162,133,640
Ovary	63,688,000	63,387,924	57,334,040	50,820,452
Total	596,839,624	595,633,778	442,152,488	383,063,300

Total number of paired-end reads in the four RNA-seq experiments during the different filtering steps

<https://doi.org/10.1371/journal.pone.0182756.t001>

**Table 2. Assembly metrics and quality assessment of the transcriptome assemblies.**

Feature	Brain	Kidney	Liver	Ovary	Combined	Previously known ESTs
Trinity transcripts	346,582	316,366	226,007	291,104	705,797	33,267 <sup>a</sup>
Trinity genes	295,242	274,538	194,853	254,001	586,693	22,410 <sup>b</sup>
N50	925 bp	778 bp	697 bp	699 bp	866 bp	651 bp
Average length	605 bp	570 bp	560 bp	555 bp	616 bp	601 bp
Mapped reads <sup>c</sup>	78.79%	69.97%	76.08%	71.66%	71.77%	-
BUSCO <sup>d</sup>	78% (10%)	76% (10%)	62% (10%)	67% (10%)	87% (7%)	18% (10%)
Elephant shark coverage <sup>e</sup>	36,642 (52.6%)	36,132 (49.6%)	33,729 (44.9%)	35,691 (46.9%)	27,019 (59%)	8,301 (17.2%)

<sup>a</sup> Total number of ESTs (recorded at NCBI for *S. acanthias*, plus sequences reported by [39]).

<sup>b</sup> Clustering of (a) with cd-hit at 90% identity.

<sup>c</sup> Percentage of reads properly mapped back to the transcriptome.

<sup>d</sup> Proportion of complete and fragmented (in brackets) conserved vertebrate core genes according to the total number of core genes in the different datasets as derived by BUSCO analysis.

<sup>e</sup> Number of unique hits in the elephant shark proteome; in brackets the percentage of best unique hits to elephant shark proteins covered by at least 70% in length by a transcript from the different assemblies or ESTs.

<https://doi.org/10.1371/journal.pone.0182756.t002>

transcript from the different assemblies varied from 45% to 53%. A summary of the quality assessment of the assemblies is shown in Table 2.

We filtered the transcripts by discarding potential contaminants, fragment sequences shorter than 300 bp and transcripts representing potential assembly artefacts (indicated by a FPKM value of zero) yielding a final set of 155,274 (brain), 140,336 (kidney), 110,527 (liver), 125,806 (ovary) and 362,690 (combined) transcripts, respectively. The stringent filtering increased the N50 values by about 50% while the completeness of the assemblies decreased only minimally as determined by BUSCO analysis using a set of conserved vertebrate genes. The number of filtered full-length cDNAs (transcripts classified by Transdecoder as ‘complete’) in the spiny dogfish transcriptomes amount to 24738 in brain (S1 File), 21129 in kidney (S2 File), 14650 in liver (S3 File) and 18109 in ovary (S4 File), respectively. An overview of the metrics of the different assemblies after filtering is shown in Table 3.

### Annotation of the *de novo* assembled transcriptomes

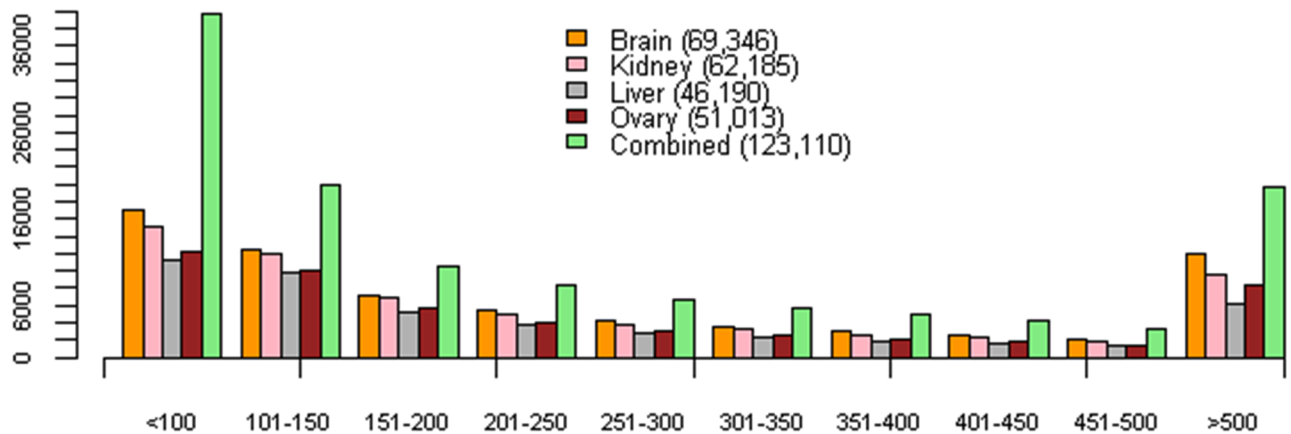
Using TransDecoder (<https://transdecoder.github.io/>), a total of 69,346 (brain), 62,185 (kidney), 46,198 (liver), 51,013 (ovary) and 123,110 (combined) peptides longer than 60 amino acids were predicted (Fig 1 and Table 3). In order to estimate the overlap with known proteins, the predicted peptides were compared against the SwissProt, UniRef90 and Pfam databases.

**Table 3. Transcriptome assembly metrics after filtering.**

	Brain	Kidney	Liver	Ovary	Combined
Trinity transcripts	155,274	140,336	110,527	125,806	362,690
Trinity genes	126,217	115,179	94,567	105,738	289,515
N50	1,586	1,439	1,196	1,337	1,302
Average length	950 bp	901 bp	823 bp	872 bp	866 bp
BUSCO	77%	74%	61%	67%	87%
Predicted proteins	69,346	62,185	46,198	51,013	123,110
Average length	310 aa	291 aa	270 aa	300 aa	287 aa

Assembly metrics for the assemblies after filtering.

<https://doi.org/10.1371/journal.pone.0182756.t003>



**Fig 1. Length range and tissue distribution of the predicted proteins in the different assemblies.** The total number of predicted proteins is shown in brackets. The x-axis represents the length range in amino acids while the y-axis corresponds to the total counts of the different predicted proteins.

<https://doi.org/10.1371/journal.pone.0182756.g001>

The percentage of predicted proteins with a blast hit in SwissProt was in the range of 71–72% in the assemblies of the individual tissues while this percentage in the combined assembly was 61%. Accordingly, the use of the larger protein database UniRef90 revealed a higher percentage of predicted proteins with a hit, ranging from at least 75% in the individual assemblies to 67% in the combined assembly. In addition, almost 50% of the predicted proteins had a hit in the Pfam database above the per domain noise score cutoff. SignalP predicted around 5% of the proteins having a signal peptide while TMHMM found that at least 11% of the proteins of each assembly carry at least one transmembrane region. Details of the annotation results are shown in Table 4. A graphical summary of the annotation results is shown in S1 Fig.

### Identification of orthologues to elephant shark

After protein prediction and redundancy removal (see Methods), the percentage of non-redundant elephant shark proteins with a homolog in the different spiny dogfish assemblies was 84% (brain), 82% (kidney), 72% (liver), 76% (ovary) and 90% (combined), respectively. From those, 9,036 (brain), 8,457 (kidney), 6,800 (liver), 7,550 (ovary) and 10,623 (combined assembly) corresponded to one-to-one orthologous pairs. (Fig 2). As expected, the encoded proteins unique to the respective assembly are related to their specific tissue function in sharks: For instance, neurosecretory proteins and myelin basic protein in brain, several members of the cystatin family protein, ferritin and several zymogens in liver and different solute carriers in kidney. In ovary, apart from genes associated with reproduction (such as fibrous sheet CABYR-binding protein like and aromatase among others) we note the presence of genes involved in the immune system such as cysteinyl leukotriene receptor 1-like, protein BTG1 and immunoglobulin heavy chain IgH, which is explained by the specific shark anatomy where the ovaries are enclosed in the epigonal organ, a part of the lymphomyeloid system in sharks [50, 51].

### Spiny dogfish shark annotated transcriptome assemblies: A molecular tool

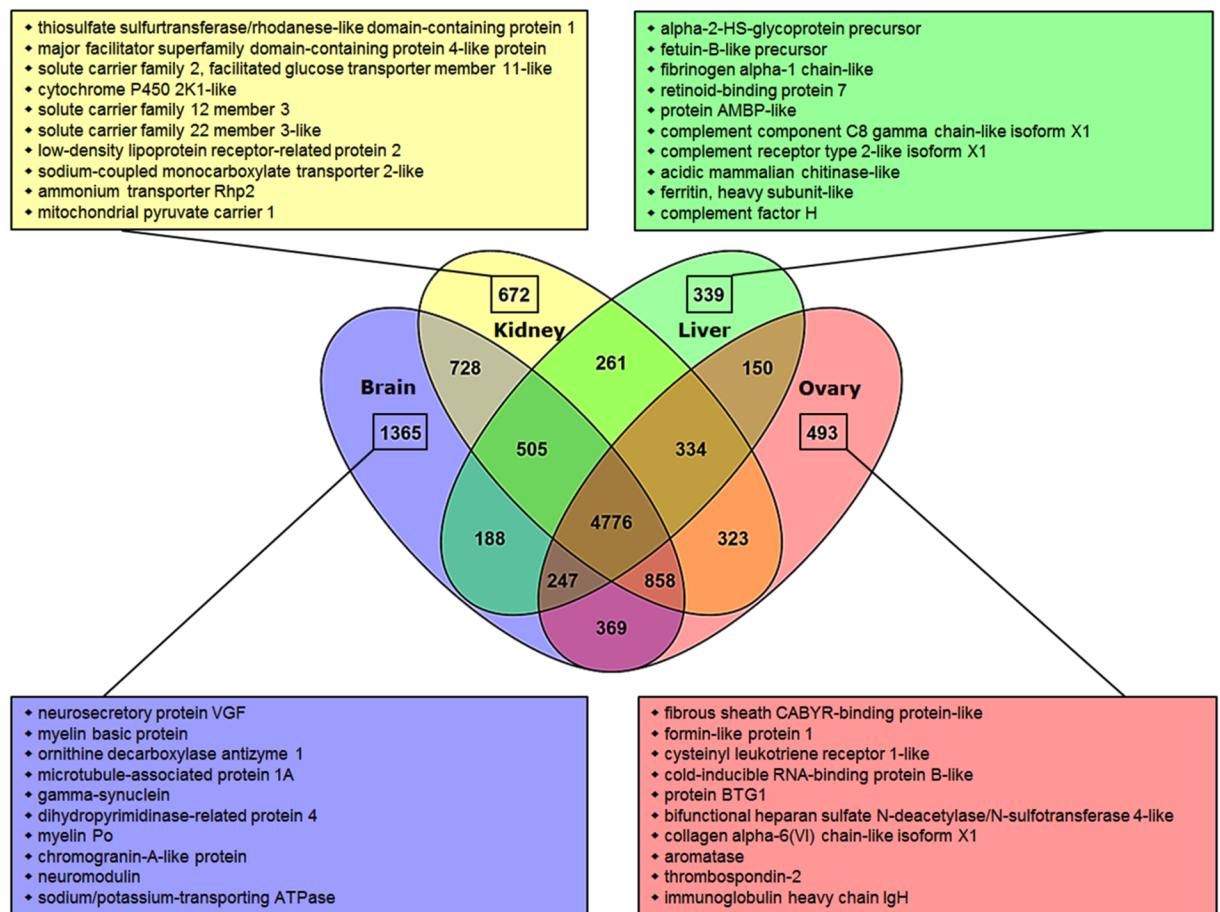
In order to assess the usability of our transcriptome assemblies as a molecular tool, we investigated the presence of genes involved in osmoregulation, a field where the spiny dogfish is of

**Table 4. Trinotate annotation of the transcriptome assemblies.**

Anotation feature	Brain	Kidney	Liver	Ovary	Combined
Predicted proteins	69,346	62,185	46,198	51,013	123,110
SwissProt blastx	60,250	55,074	41,614	44,680	100,636
SwissProt blastp	49,522	44,782	33,288	36,813	78,679
Uniref90 blastx	67,066	61,204	46,371	49,763	115,108
Uniref90 blastp	52,177	47,380	35,088	38,809	83,164
Pfam	38,256	34,073	25,162	28,578	61,215
SignalP	3,700	2,979	2,212	2,808	7,208
TMHMM	8,793	7,443	5,227	6,189	15,971
eggNOG	24,954	23,905	18,635	19,429	42,078
GO	49,256	45,594	35,910	38,770	81,582

Number of transcripts and predicted proteins with a hit in the different databases searched: SwissProt, UniRef90 and Pfam. Signal peptide and transmembrane regions predictions were performed with SignalP and TMHMM software. GO and eggNog annotation were retrieved from blast and Pfam results.

<https://doi.org/10.1371/journal.pone.0182756.t004>



**Fig 2. Relationship between elephant shark and spiny dogfish orthologs in the brain, kidney, liver and ovary transcriptome assemblies.** Colored squares show the elephant shark annotation for the ten most expressed spiny dogfish one-to-one orthologs (measured in FPKM) unique to each assembly.

<https://doi.org/10.1371/journal.pone.0182756.g002>

great interest as a model system. To this end the Trinotate annotation reports (S1–S5 Tables) were searched for genes, which are involved in urea synthesis as well as urea and ammonia transport as the main processes participating in osmoregulation in cartilaginous fishes.

**Urea synthesis.** To identify transcripts, which were assigned to the known gene ontology terms related to urea synthesis biosynthetic pathways, we queried the Trinotate annotation reports for the gene ontology terms “urea cycle” (GO:0000050), “allantoin catabolic process” (GO:0000256) and “urate oxidase activity” (GO:0004846). Since the transporter CPSIII in cartilaginous fishes uses glutamine instead of ammonia as substrate [3, 52], the term “glutamate-ammonia ligase activity” (GO:0004356) was also inquired. The best scoring hits to SwissProt and Uniref90 databases for these transcripts showed proteins involved in urea biosynthetic pathways. To enable comprehensive gene identification, we also compared the transcripts found in the different assemblies with the elephant shark, the closest organism with an available genome assembly [41]. To this end, the proteome of the elephant shark was used as reference and elephant shark—spiny dogfish orthologous groups were identified. We found all the genes involved in purine degradation pathways and ornithine-urea cycle to be present in the liver assembly with the exception of one arginase gene (Arg 1-like). However, some other tissues such as kidney present many of the genes involved in these urea biosynthetic pathways. The results are summarized in Table 5. Detailed annotation for the genes and isoforms involved in urea biosynthesis that were identified from the annotation reports is provided in S6 Table, while the elephant shark—spiny dogfish orthologous groups of urea synthesis genes are listed in S8 Table; spiny dogfish predicted protein sequences involved in the urea cycle are given in S5 File. Importantly, blast searches against NCBI *Squalus acanthias* nucleotide, ESTs and protein databases revealed that in the case of Uox, Ornt1, Gs2, Aln, Allc, and Arg2 the transcripts identified code for proteins for which our sequences represent novel information which has not been described before. Moreover, with the exception of CPSIII, the remaining spiny dogfish transcripts identified for these biosynthetic pathways either code for a longer version of the protein or are not fully covered by the existing spiny dogfish sequences indicating that our RNA-seq and transcriptome resources provide additional novel sequence information for these genes (see S9 Table).

**Urea transport.** Annotation reports were searched under the gene ontology terms “urea transmembrane transporter activity” (GO:0015204), “urea transport” (GO:0015840) and “urea transmembrane transport” (GO:0071918). The resulting transcripts code for proteins that are orthologues to the elephant shark urea transporter efUT1 as well as to two different aquaporin genes (Aqp3-like) proteins. In the case of the urea transporter, two isoforms of *ut1* were identified in the kidney assembly: the first one encodes an identical protein sequence to the previously described urea transporter in spiny dogfish shUT [53] while the second one codes for a longer protein and possesses an extended carboxyl terminal part similar to the one identified in the Atlantic stingray *Dasyatis sabina* [54, 55], see Fig 3. In addition, three more urea transporter transcripts were found in the spiny dogfish brain transcriptome belonging to the same gene assembled by Trinity. However, the transcript coding for the longest peptide with a partial length encoding 272 amino acids (TR95738|c0\_g3\_i3|m.93774; see S6 File) is also an orthologue of efUT1, lacks strong similarity to the Ut1 isoforms identified in the spiny dogfish kidney assembly and does not cluster in the same group as the other Ut1 proteins (Fig 4), making it difficult to assess whether it is an Ut1 isoform or a protein encoded by a novel gene. Finally, the two transcripts coding for Aqp3-like proteins were only found in the spiny dogfish kidney assembly (Table 5). Blast searches confirmed that their best hit in the existing NCBI *S. acanthias* databases is Aqp4 but with only 27% sequence identity, thus indicating that of these two genes identified have not previously been described for spiny dogfish.

**Table 5. Genes involved in osmoregulation identified in the different spiny dogfish transcriptome assemblies based on orthologous proteins between elephant shark NCBI proteins and spiny dogfish assemblies.**

	Gene products	Brain	Kidney	Liver	Ovary
Purine degradation pathway	Aln	-	-	+	-
	Allc	-	+	+	-
	Uox	-	-	+	-
O-UC	Arg-1-like	-	+	-	-
	Arg-2	*	+	+	+
	Otc	*	+	+	+
	Ass1	+	+	+	+
	Asl	+	+	+	+
	CpsIII	-	-	+	-
	Gs1	+	+	+	+
	Gs2	-	+	+	+
Accessory O-UC	Ornt1	+	-	+	+
	Nags	-	-	+	-
Urea transport	Ut-1	*	+	-	-
	Aqp3-like	-	+	-	-
	Aqp3-like	-	+	-	-
Water transport	Aqp1-like	+	-	-	-
	Aqp9	-	+	+	-
	Aqp4	-	+	-	-
	Aqp15	-	+	-	-
Rh-glycoproteins	Rhag	+	+	+	+
	Rhbg	+	+	+	-
	Rhcg <sup>a</sup>	*	*	*	-
	Rhp2	-	+	*	-
	Rh30-like	*	-	-	-

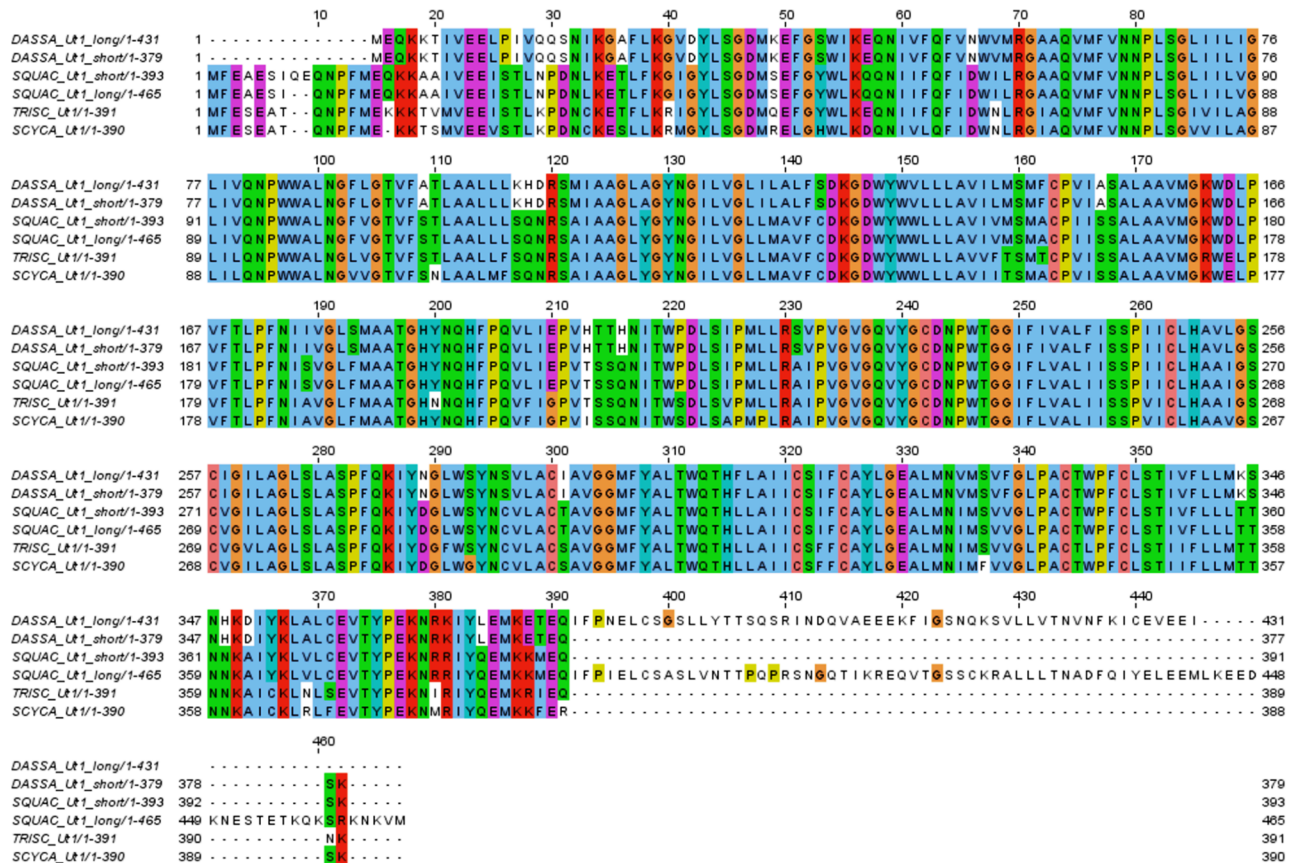
Aln (allantoinase), Allc (allantoicase), Uox (urate oxidase), Arg (arginase), Otc (ornithine transcarbamylase), Ass (argininosuccinate synthase), Asl (argininosuccinate lyase), CpsIII (carbamoyl phosphate synthetase III) and Gs (glutamine synthetase), Ornt1 (ornithine-citrulline mitochondrial transporter) and Nags (N-acetyl synthase) which likely acts as cofactor in CPSIII in cartilaginous fishes [11, 52], Ut-1 (urea transporter) and Aqp (aquaporins).

<sup>a</sup> Possible orthologue to elephant shark. (+) Spiny dogfish predicted proteins covering >70% or <70% (\*) of the length of the hit to the elephant shark orthologue. (-) No gene product identified. Accession numbers and transcript IDs for elephant shark proteins and the predicted proteins of the *de novo* assemblies are found in S8 Table.

<https://doi.org/10.1371/journal.pone.0182756.t005>

To further investigate the presence of additional aquaporin genes, the transcriptome assemblies were searched for transcripts coding for the aquaporin family protein domain “major intrinsic protein” (PF00230.15). We observed only one Trinity gene coding for that protein domain in each of the brain, liver and ovary transcriptomes while we found 9 different in kidney (S7 Table). Elephant shark—spiny dogfish orthologous pairs revealed that these transcripts code for proteins that are orthologues to elephant shark Aqp1-like, Aqp9, Aqp4, Aqp-fa-chip, Aqp0 (lens fiber major intrinsic protein-like) and to the two different Aqp3-like proteins above described (S8 Table; spiny dogfish predicted protein sequences for aquaporins are provided in S7 File). An *aqp1-like* transcript was only found in brain, while *aqp9* transcripts were identified in kidney and liver. On the other hand, the remaining aquaporin transcripts were only identified in the kidney assembly (Table 5). Blast searches against nucleotide and protein databases for *S. acanthias* showed that the partial transcript, which is orthologous to elephant shark aquaporin-FA-CHIP encoded a protein identical to Aqp15, previously described in





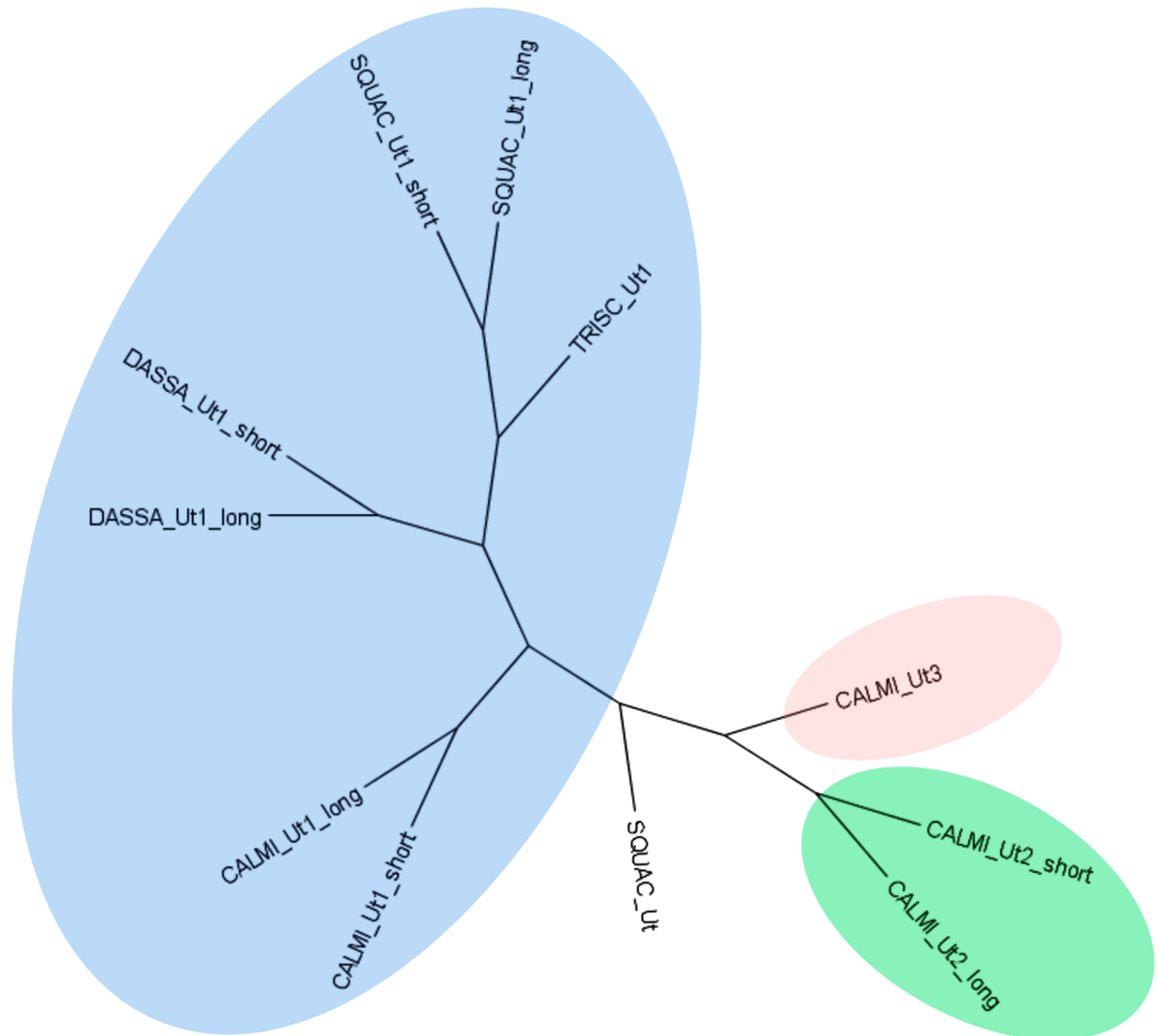
**Fig 3. Multiple alignments of Ut1 proteins in selected elasmobranchs.** Atlantic stingray (*Dasyatis sabina*; DASSA), spiny dogfish (*Squalus acanthias*; SQUAC) kidney isoforms, banded houndshark (*Triakis scyllium*; TRYSC) and lesser spotted catshark (*Scyliorhinus canicula*; SCYCA). Alignment performed with MUSCLE showing ClustalX color scheme at 30% conservation threshold. Accession numbers for DASSA\_Ut1\_long (AAM46683.1), DASSA\_Ut1\_short (AAQ07592.1), TRISC\_Ut1 (BAC75980.1) and SCYCA\_Ut1 (AEH59797.1). Protein IDs for Ut1\_long (TR103653|c5\_g2\_i2|m.138919), Ut1\_short (TR103653|c5\_g2\_i1|m.138917) identified in the *de novo* assemblies.

<https://doi.org/10.1371/journal.pone.0182756.g003>

spiny dogfish [56]. Even though the identified Aqp1-like, Aqp4, Aqp9 and Aqp15 align to the known ESTs, nucleotides or proteins for spiny dogfish, the resulting predicted proteins improve the existing spiny dogfish sequences and refine information in the NCBI databases (see S9 Table). Taken together, completely novel sequences are presented for Aqp3-like and Aqp0 while the sequence information for Aqp1-like and Aqp9 is also improved based on the *de novo* assemblies. In the case of Aqp15 and Aqp4 [57], however, the predicted proteins were identical to the ones already described.

Finally, the presence of several sodium, potassium and chloride transporters was assessed in the kidney assembly by means of transcripts assigned to the gene ontology terms listed in Table 6 and S7 Table. The elephant shark—spiny dogfish orthologue relationship revealed 82 orthologues using these gene ontology terms in the kidney assembly (data not shown). However, the functional roles of these genes in shark kidney will have to be investigated in further studies.

**Ammonia transport.** Transcriptome annotation revealed several genes coding for ammonium transporter protein domains in brain (5), kidney (4), liver (6) and ovary (2), which align to known Rh-glycoproteins such as Rhag, Rbhg, Rhcg and Rhp2 in SwissProt and UniRef 90 databases (S7 Table). Orthologous pairs between elephant shark and our *de novo* assembled



**Fig 4. Phylogenetic tree of multiple urea transporter proteins in selected elasmobranchs and elephant shark.** Ut1 (blue), Ut2 (green) and Ut3 (pink). *Dasyatis sabina* (DASSA), *Squalus acanthias* (SQUAC), *Triakis scyllium* (TRYSC), *Scyliorhinus canicula* (SCYCA) and *Callorhynchus milii* (CALMI). SQUAC\_Ut1\_long and SQUAC\_Ut1\_short correspond to the Ut isoforms identified in the kidney assembly (TR103653|c5\_g2\_i2|m.138919 and TR103653|c5\_g2\_i1|m.138917, respectively). SQUAC\_Ut corresponds to the isoform identified in brain (TR95738|c0\_g3\_i3|m.93774). Accession numbers for DASSA\_Ut1\_long (AAM46683.1), DASSA\_Ut1\_short (AAQ07592.1), TRISC\_Ut1 (BAC75980.1), SCYC\_Ut1 (AEH59797.1), CALM\_Ut1\_long (BAH58773.1), CALMI\_Ut1\_short (BAH58774.1), CALMI\_Ut2\_long (BAH58775.1), CALMI\_Ut2\_long (BAH58776.1) and CALMI\_Ut3 (BAH58777.1).

<https://doi.org/10.1371/journal.pone.0182756.g004>

spiny dogfish transcriptomes confirmed the identity of those Rh-glycoproteins in spiny dogfish (see [Table 5](#) and [S8 Table](#)). From those, Rhag, Rhbg and Rhp2 exhibit 100% identity to the same proteins previously identified in spiny dogfish [14]. Moreover, transcripts coding for orthologues to elephant shark Rh30-like protein were found. The longest isoform of these transcripts is found in the brain assembly, which codes for an open reading frame (ORF) of 304 amino acids length ([S8 File](#)) with 64% identity to the elephant shark Rh30-like protein. Additionally, we found a partial transcript in the brain assembly encoding an open of 337 amino acids (TR88392|c0\_g1\_i1|m.62153; see [S8 File](#)) whose best Uniref90 hit is the elephant shark protein Rhcg. Although this protein was neither included in any orthologue group between

**Table 6. Gene ontology (GO) terms searched in the kidney assembly related with Na<sup>+</sup>, K<sup>+</sup> and Cl<sup>-</sup> transport.**

GO term	Description
GO:0008511	Sodium: potassium: chloride symporter activity
GO:0006813	Potassium ion transport
GO:0070294	Renal sodium ion absorption
GO:0035725	Sodium ion transmembrane transport
GO:1902476	Chloride transmembrane transport
GO:0006821	Chloride transport
GO:0006814	Sodium ion transport
GO:0002028	Regulation of sodium ion transport
GO:0003096	Renal sodium ion transport

<https://doi.org/10.1371/journal.pone.0182756.t006>

elephant shark and spiny dogfish, it shares 72% identity with the same elephant shark protein. Altogether, our transcriptome resources provided evidence for the putative Rhcg sequence being completely new for spiny dogfish in addition to longer predicted peptide sequences, e.g. the predicted Rhag is 83 amino acids longer in the carboxyl-terminal part and the predicted Rh30-like is 101 amino acids longer in the amino-terminal part. The predicted proteins identified as orthologues to elephant shark Rhbg and Rhp2, however, are fully covered in the existing *S. acanthias* databases. The phylogenetic tree of the Rh-glycoproteins identified (Fig 5) shows, as expected, that all of the identified Rh-glycoproteins cluster with their respective orthologs in the elephant shark.

## Discussion

The high number of the spiny dogfish predicted proteins align to the different databases searched provide a more complete annotation than obtained by previous studies involving other shark transcriptomes [43–45]. Although the activities of all the enzymes involved in urea synthesis in spiny dogfish have been biochemically studied since the 1960s [1], sequence information of the genes that code for those proteins is still incomplete in this shark species. The analysis of the spiny dogfish transcriptome reveals the presence of all the genes involved in the two main biosynthetic pathways for urea synthesis in cartilaginous fishes (Table 5) and as expected, all of these genes are expressed in the liver, the main tissue involved in urea synthesis. However, in accordance with previous studies [6–8, 58] some of the urea synthesis genes were identified in extra-hepatic tissues as well.

Earlier studies have described one glutamine synthetase (Gs) in spiny dogfish [59], being orthologous to elephant shark Gs1. In the present work we were able to identify two different glutamine synthetases in the spiny dogfish that are orthologues to the two glutamine synthetases previously identified in the elephant shark [8] termed Gs1 and Gs2 (Fig 6).

Comparison with public resources for spiny dogfish failed to retrieve a sequence similar to the *gs2* transcripts identified in the present work, therefore, to our knowledge this is the first study in which a second *gs* gene in this species was predicted. In accordance with the findings in elephant shark [8] and the existing sequence information for *S. acanthias*, we predict several *gs1* isoforms coding for alternative N-terminal parts carrying and lacking mitochondrial localization signals. On the other hand, unlike the in the elephant shark we do find several *gs2* isoforms in spiny dogfish for which in all cases the predicted protein lacks the mitochondrial sequence signal. Previous studies have described that the expression of these multiple genes varies among tissues and developmental stages in the elephant shark [8], suggesting

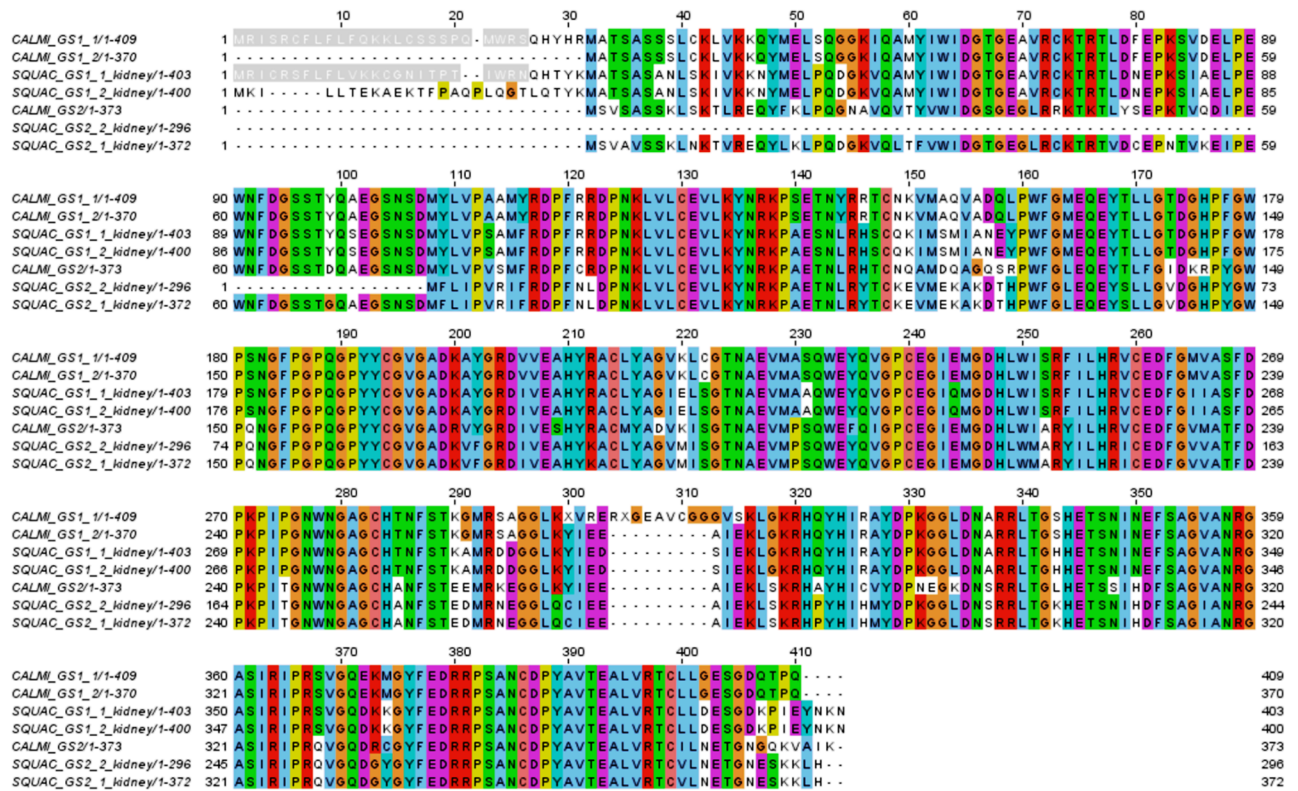


**Fig 5. Phylogenetic tree of the different Rh-glycoproteins found with their elephant shark orthologues.** Accession numbers: CALMI\_Rhag (AFO95352.1), CALMI\_Rhb (AFP03342.1), CALMI\_Rhcg (XP\_007906553.1), CALMI\_Rhp2 (AFP00304.1), CALMI\_Rh30-like (NP\_001279904.1). Protein IDs: SQUAC\_Rhag (TR44094|c0\_g1\_i1|m.17165), SQUAC\_Rhb (TR9088|c0\_g1\_i1|m.3455), SQUAC\_Rhcg (TR88392|c0\_g1\_i1|m.62153), SQUAC\_Rhp2 (TR33065\_c0\_g1\_i1\_m.11704), SQUAC\_Rh30-like (TR93310|c0\_g1\_i2|m.80440). Colors showing groups of Rhag (grey), Rhb (brown), Rhcg (dark blue), Rhp2 (light blue) and Rh-30-like (green).

<https://doi.org/10.1371/journal.pone.0182756.g005>

differentiation of the role of the different isoforms. The biological implication of these findings in spiny dogfish and in elasmobranchs needs further investigation.

Regarding the transcripts coding for arginase proteins, we were able to identify transcripts coding for different arginases, Arg1 and Arg2 as described in the elephant shark [8]. Transcripts coding for Arg2 are found in the four tissues and their predicted proteins have not been described before in the spiny dogfish while the ones coding for Arg1-like are only found in the



**Fig 6. Multiple alignments of the different glutamine synthetase (Gs) in elephant shark (CALMI) and spiny dogfish (SQUAC).** Mitochondrial localization signals are highlighted in grey. Alignment performed with MUSCLE showing ClustaX color scheme at 30% conservation threshold. Protein IDs and accession numbers for SQUAC\_GS1\_1\_kidney (TR71041|c0\_g1\_i1|m.31201), SQUAC\_GS1\_2\_kidney (TR71041|c0\_g1\_i2|m.31207), SQUAC\_GS2\_1\_kidney (TR91001|c0\_g1\_j3|m.57038), SQUAC\_GS2\_2\_kidney (TR91001|c0\_g1\_i1|m.57036), CALMI\_GS1\_1 (XP\_007901126.1) and CALMI\_GS1\_2 (BAK96221.1), CALMI\_GS2 (BAK96222.1).

<https://doi.org/10.1371/journal.pone.0182756.g006>

spiny dogfish kidney assembly and their predicted protein is identical to the previously identified arginase. The predicted Arg2 protein seems to be the predominant arginase in the O-UC although both arginases carry mitochondrial localization signals which is in accordance with the elasmobranch pathway of O-UC where the last step of urea synthesis occurs in the mitochondria [8, 10, 11].

With reference to urea transport and reabsorption genes, three different urea transporters (UT) have been described in the elephant shark, termed efUT1, efUT2, and efUT3 [60]. In comparison, only one gene has been reported for the spiny dogfish (shUT) which is the ortholog of efUT1 [53]. In the case of the Atlantic stingray the efUT1 ortholog codes for two isoforms, with the second one showing an extended carboxyl terminal part [54, 55]. Analogously to the Atlantic stingray, we identified a novel and longer isoform of shUT in spiny dogfish which has been suggested to play a role in the acclimation to different salinities for euryhaline elasmobranchs. Since spiny dogfish has been described euryhaline future studies are needed to investigate whether this second isoform participates in such acclimation.

Water retention has been suggested as an important step in urea reabsorption process in which aquaporins (AQP) play an active role [12]. Although the precise role of aquaporins in the transport of urea and ammonia is not well understood, human AQP3, AQP7 and AQP9 are known to be permeable to both ammonia and urea while AQP8 and AQP10 only to ammonia [61, 62]. We identified transcripts coding for Aqp1, Aqp4, Aqp9, and Aqp3 (two copies),

all of which were found in the spiny dogfish kidney. The presence of two different Aqp3 genes in the kidney assembly is noteworthy; however, to determine the potential roles of these aquaporins in urea and ammonia reabsorption in cartilaginous fishes does require further investigation.

Apart from minimizing urea loss, sharks are able to diminish leakage of ammonia through the gills and the kidney [12, 14, 15]. Rh-glycoproteins are a group of proteins belonging to the ammonium transporter superfamily being able to transport ammonia [63]. They are believed to play a crucial role in the urea metabolism of cartilaginous fishes [12, 14, 15]. This protein family has five different members in elephant shark: Rhag, Rhbg, Rhcg, Rhp2 and Rh-30-like, of which only three, Rhag, Rhbg and Rhp2, have been described in spiny dogfish so far [14]. Our transcriptome analysis predicted proteins with complete sequence identity to these three proteins. In addition, we found partial length contigs coding for Rh-30 and a transcript that encoded a putative Rhcg, which sequence has not been previously described in spiny dogfish.

## Conclusions

This is the first study to generate large-scale NGS data and single tissue transcriptome characterization for the model species *Squalus acanthias*. We present an extensive multi-tissue transcriptome analysis and the annotation of our assemblies showed that the predicted proteins have a high percentage of significant hits in the SwissProt and UniRef90 databases. The completeness of the assemblies using BUSCO revealed that although we only sequenced four tissues, the combined assembly has a completeness of 87%. Querying annotation reports provided important new sequence information in genes involved in urea-based osmoregulation, a scientific field where elasmobranchs and particularly spiny dogfish as an animal model are of great interest.

Taken together, our data collected provides a new resource as well as an extended catalogue of genes involved in osmoregulation in spiny dogfish. Due to the detailed annotation generated, the same approach would be suitable to extract sequence information related to other areas of biological interests. Consequently, we believe the data generated will be of great support for the ongoing research in elasmobranchs and in spiny dogfish shark, particularly.

## Materials and methods

### Tissue sampling and ethics statement

The shark tissue samples were opportunistically obtained from the public aquarium Kattegatcenteret, Grenå, Denmark independent of our study. Due to veterinary reasons the specimen was to be euthanized and we were contacted to gain access to the tissue material as a collaborative effort by the Kattegatcenter; following dissection samples were snap-frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$ . It is important to emphasize that the shark did not get euthanized as part of this investigation and that the tissue material became opportunistically available for research. Thus, no ethical approval or permit for animal experimentation was required, as the individual was not sacrificed specifically for this study.

### RNA-seq: Library preparation and sequencing

Total RNA was isolated from brain, kidney, liver and ovary applying the procedure for isolation of total RNA as provided by the mirVana miRNA Isolation Kit (Ambion) according to the manufacturer's instructions. RNA was treated with Ribo-Zero rRNA removal kit (Epicentre) in order to remove rRNA from the RNA population. Libraries were prepared with a selected insert size of 300 bp using ScriptSeq v2 RNA-Seq Library Preparation Kit (Epicentre)

for strand-specific and multiplexed libraries following manufacturer instructions. Paired-end sequencing was performed on the Illumina HiSeq2000 platform to a read length of 150 bp. De-multiplexing was performed with Illumina software allowing zero mismatches in the barcode. RNA-seq raw reads after de-multiplexing and removal of technical sequences by Illumina software have been deposited in the European Nucleotide Archive (ENA) database under the study accession PRJEB14721.

### Filtering of raw data

FastQC (<http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc>) was used in order to get a visual overview of the read quality. Before assembly, possible remaining rRNA as well as mitochondrial reads were removed from the raw data based on hits against the LSU\_Ref and SSU\_Ref Silva databases version 119 [64] and the *Squalus acanthias* mitochondrial genome (accession NC\_002012): Raw reads were mapped to both databases using Bowtie2 with the wrapper script provided by FastQScreen version v0.4.4 with default parameters ([http://www.bioinformatics.babraham.ac.uk/projects/fastq\\_screen/](http://www.bioinformatics.babraham.ac.uk/projects/fastq_screen/)). The reads that did not map to these indexes were considered non-ribosomal and non-mitochondrial and were used for downstream filtering steps.

Trimmomatic version 0.32 [65] was used to remove adapter sequences and trim low quality bases using the parameters `-phred33 ILLUMINACLIP : Scriptseq2_adapters:2:30:10 LEADING : 3 TRAILING : 3 SLIDINGWINDOW : 4:15 MINLEN : 50`.

### De novo assembly

High quality paired-end reads were assembled using Trinity software version 2.0.6 with default parameters except `-max_memory 15G -CPU 20 -SS_lib_type FR -min_kmer_cov 2 -KMER_SIZE 25 -min_contig_length 200` for the single tissue assemblies. For the combined assembly, reads from the four tissues were merged into a single fastq file and assembled together. The parameters applied were the same as for the individual assemblies except for the added option `-normalize_reads`. In both cases, the assembly metrics were obtained using the *Trinity\_stats.pl* script from the Trinity software package.

### Back mapping of reads and abundance estimation

Reads were mapped back to the transcripts using the script *bowtie\_PE\_separate\_then\_join.pl* from the Trinity software 2.1.1; Bowtie [66] was run with the parameters `-p 20 -all-best-strata -m 300`. The percentage of reads that map to the assembly as properly paired, improper pairs, left only and right only was calculated with the script *SAM\_nameSorted\_to\_uniq\_count\_stats.pl*. Abundance of each transcript and gene was calculated using the *align\_and\_estimate\_abundance.pl* script from Trinity software 2.0.6. Default settings were used except for the options `-est_method RSEM -aln_method bowtie -trinity_mode -prep_reference -SS_lib_type FR`. RSEM version 1.2.19 [67] was used to estimate the abundance of each transcript.

### Assessment of full-length coverage of transcripts and completeness of the assemblies

Assemblies were aligned with blastx version ncbi-blast-2.2.30+ [49] to the elephant shark proteome (downloaded from NCBI protein database, as of January 28, 2016). Blastx output was filtered to an e-value of at least  $1e-20$ . The percentage of the assembly query that covers a

database hit was calculated using the Perl script bound to Trinity suite *analyze\_blastPlus\_to-pHit\_coverage.pl*.

To check the completeness of the assemblies BUSCO software version 1.1 [48] with `trans` option against the vertebrata dataset was used. BUSCO software and the vertebrata dataset were downloaded from <http://busco.ezlab.org/>. To run BUSCO, the programs ncbi-blast-2.2.30+ [49], HMMER version 3.1 [68] and EMBOSS 6.3.1 [69] were used.

## Filtering of assembled transcripts

Transcripts with zero fragments per kilo base per million reads (FPKM), transcripts per million (TPM), and IsoPct (percentage of the reads that align to each isoform over the reads that aligned to all the gene isoforms) were removed. A second filter was performed to remove transcripts shorter than 300 bp. Tabulated expression values for the genes of each tissue are available to allow more stringent filtering (S9, S10, S11 and S12 Files). In addition, the assemblies were screened for possible vector contamination by aligning them to the Emvec databank <ftp://ftp.ebi.ac.uk/pub/databases/emvec/> using blastn with a cutoff of 90% identity and an e-value of 1e-10.

## Functional annotation of the transcriptome

Protein-coding transcripts were predicted with TransDecoder\_r20140704 (<https://transdecoder.github.io/>) with `-S-m 60` options and the rest parameters set as default. Trinotate version 2.0.0 was used to retrieve the functional annotation of the *de novo* assembled transcriptomes of spiny dogfish. In order to annotate the transcriptome, SwissProt, Uniref 90, Pfam databases and the Trinotate boilerplate compatible with those databases and Trinotate v2.0.0 were downloaded from the Trinotate webpage <http://trinotate.github.io/> on October 23, 2015. Additionally, ncbi-blast-2.2.30+ [49] TMHMM software version 2.0 [70], HMMER version 3.1 [68] and SignalP version 4.1 [71] were used with the parameters recommended for Trinotate annotation. Finally, the annotation reports for the assemblies were generated using an e-value cutoff of 1e-05 and the per domain noise score in Pfam as cutoff. Gene Ontology (GO) [72] and eggNOG [73] annotations were retrieved from the blast results. In the case of GO, additional GO annotation was retrieved from the Pfam database results.

## Elephant shark orthologue assignment

The elephant shark proteome (downloaded from NCBI protein database, as of January 28, 2016) and the predicted proteins from the *de novo* assemblies were used for orthologue assignment after removing redundancy at 90% identity with `cd-hit` [74, 75]. In order to get a reliable set of orthologues between both datasets, only proteins with at least 100 amino acids of length were kept. Finally, Inparanoid version 4.1 [76] was used with default parameters and the options `bootstrap` and `BLOSUM 80`. The Venn diagram was generated using Venny 2.0.2 [77].

## Comparison with existing *Squalus acanthias* resources

*S. acanthias* ESTs, nucleotides and protein sequences were downloaded from NCBI on June 29, 2016. Regarding the EST dataset for spiny dogfish 5,824 sequences correspond to an embryo-derived cell line sample [36] (accession SAMN00176998), 15,078 from a study that generated ESTs from a pool of multiple tissues (accession SAMN00175664) and 11,660 correspond to the ESTs generated in two independent studies from the rectal gland (accessions SAMN00150616 and SAMN00154362). In addition, 705 ESTs previously generated in [39] were also downloaded and merged with the ESTs mentioned above and the *S. acanthias*



nucleotide sequences from NCBI. A non-redundant set of the proteins identified was produced using cd-hit at 90% identity which was then used as query. Tblastn and blastp were run with an e-value cutoff of 1e-06 and 90% identity and the rest of the parameters as default.

## Generation of multiple sequence alignments, phylogenetic trees and mitochondrial signal prediction

Multi-sequence alignments were performed with MUSCLE v3.8.31 [78] using default parameters and visualized with Jalview 2.10.1 [79]. Mitochondrial signals were predicted with TargetP [80] at <http://www.cbs.dtu.dk/services/TargetP/> with a cutoff of 0.75.

Multi-sequence alignments were trimmed with trimAl [81] contained in ETE Toolkit 3.0.0b34 [82] with `-automated 1` option and the remaining parameters as default. Trimmed alignments were used as input for RaXML version 8.2.9 [83] using the JTT gamma model under the parameters `-m PROTGAMEJTT -f a -x 12345 -N 1000 -T 60 -p 12345` and the remaining options as default. The resulting best trees were visualized with FigTree version 1.4.2 <http://tree.bio.ed.ac.uk/software/figtree/>.

## Supporting information

**S1 Fig. Graphical summary of annotation results in the different assemblies.**

(TIF)

**S1 Table. Trinotate annotation report for brain assembly.**

(ZIP)

**S2 Table. Trinotate annotation report for kidney assembly.**

(ZIP)

**S3 Table. Trinotate annotation report for liver assembly.**

(ZIP)

**S4 Table. Trinotate annotation report for ovary assembly.**

(ZIP)

**S5 Table. Trinotate annotation report for the combined assembly.**

(ZIP)

**S6 Table. Urea synthesis transcripts identified in the different assemblies.**

(XLSX)

**S7 Table. Urea transport, aquaporins, Rh-glycoproteins and ion channels transcripts identified in the different assemblies.**

(XLSX)

**S8 Table. Orthologous groups involved in osmoregulation.** Groups were identified by Inparanoid after redundancy removal between elephant shark and spiny dogfish predicted proteins in the different assemblies.

(XLSX)

**S9 Table. Comparison of the genes identified against the NCBI *Squalus acanthias* nucleotide and protein databases.**

(XLSX)

- S1 File. Spiny dogfish full-length cDNA fasta sequences from brain transcriptome (24738 filtered transcripts).**  
(ZIP)
- S2 File. Spiny dogfish full-length cDNA fasta sequences from kidney transcriptome (21129 filtered transcripts).**  
(ZIP)
- S3 File. Spiny dogfish full-length cDNA fasta sequences from liver transcriptome (14650 filtered transcripts).**  
(ZIP)
- S4 File. Spiny dogfish full-length cDNA fasta sequences from ovary transcriptome (18109 filtered transcripts).**  
(ZIP)
- S5 File. Spiny dogfish predicted protein sequences involved in urea cycle.**  
(TXT)
- S6 File. Spiny dogfish predicted protein sequences involved in urea transport.**  
(TXT)
- S7 File. Spiny dogfish predicted protein sequences for in aquaporins.**  
(TXT)
- S8 File. Spiny dogfish predicted protein sequences for in Rh-glycoproteins.**  
(TXT)
- S9 File. Tabulated expression values for spiny dogfish brain transcriptome.**  
(ZIP)
- S10 File. Tabulated expression values for spiny dogfish kidney transcriptome.**  
(ZIP)
- S11 File. Tabulated expression values for spiny dogfish liver transcriptome.**  
(ZIP)
- S12 File. Tabulated expression values for spiny dogfish ovary transcriptome.**  
(ZIP)

## Acknowledgments

We thankfully acknowledge Mette Jeppesen, Bente Flügel Majgren and Hanne Jørgensen for their excellent work in the sequencing laboratory.

## Author Contributions

**Conceptualization:** Christian Bendixen, Frank Panitz.

**Funding acquisition:** Peter A. Andreasen.

**Investigation:** Andres Chana-Munoz.

**Resources:** Rune Kristiansen, Peter A. Andreasen.

**Writing – original draft:** Andres Chana-Munoz.

**Writing – review & editing:** Agnieszka Jendroszek, Malene Sønnichsen, Rune Kristiansen, Jan K. Jensen, Christian Bendixen, Frank Panitz.

## References

- Schooler JM, Goldstein L, Hartman SC, Forster RP. Pathways of urea synthesis in the elasmobranch, *Squalus acanthias*. *Comp Biochem Physiol*. 1966; 18(2):271–81. PMID: [5964728](#)
- Boylan JW. A model for passive urea reabsorption in the elasmobranch kidney. *Comp Biochem Physiol A Comp Physiol*. 1972; 42(1):27–30. PMID: [4402716](#)
- Webb JT, Brown GW, Jr. Glutamine synthetase: assimilatory role in liver as related to urea retention in marine chondrichthyes. *Science*. 1980; 208(4441):293–5. PMID: [6102799](#)
- Part P, Wright PA, Wood CM. Urea and water permeability in dogfish (*Squalus acanthias*) gills. *Comp Biochem Physiol A Mol Integr Physiol*. 1998; 119(1):117–23. PMID: [11253775](#)
- Steele SL, Yancey PH, Wright PA. The little skate *Raja erinacea* exhibits an extrahepatic ornithine urea cycle in the muscle and modulates nitrogen metabolism during low-salinity challenge. *Physiol Biochem Zool*. 2005; 78(2):216–26. <https://doi.org/10.1086/427052> PMID: [15778941](#)
- Anderson WG, Good JP, Pillans RD, Hazon N, Franklin CE. Hepatic urea biosynthesis in the euryhaline elasmobranch *Carcharhinus leucas*. *J Exp Zool A Comp Exp Biol*. 2005; 303(10):917–21. <https://doi.org/10.1002/jez.a.199> PMID: [16161010](#)
- Kajimura M, Walsh PJ, Mommsen TP, Wood CM. The dogfish shark (*Squalus acanthias*) increases both hepatic and extrahepatic ornithine urea cycle enzyme activities for nitrogen conservation after feeding. *Physiol Biochem Zool*. 2006; 79(3):602–13. <https://doi.org/10.1086/501060> PMID: [16691526](#)
- Takagi W, Kajimura M, Bell JD, Toop T, Donald JA, Hyodo S. Hepatic and extrahepatic distribution of ornithine urea cycle enzymes in holocephalan elephant fish (*Callorhynchus milii*). *Comp Biochem Physiol B Biochem Mol Biol*. 2012; 161(4):331–40. <https://doi.org/10.1016/j.cbpb.2011.12.006> PMID: [22227372](#)
- Wood CM, Liew HJ, De Boeck G, Walsh Patrick J. A perfusion study of the handling of urea and urea analogues by the gills of the dogfish shark (*Squalus acanthias*). *PeerJ*. 2013; 1:e33. <https://doi.org/10.7717/peerj.33> PMID: [23638369](#)
- Casey CA, Anderson PM. Subcellular location of glutamine synthetase and urea cycle enzymes in liver of spiny dogfish (*Squalus acanthias*). *J Biol Chem*. 1982; 257(14):8449–53. PMID: [6123510](#)
- Anderson PM. Urea and glutamine synthesis: Environmental influences on nitrogen excretion. *Fish Physiology*. Volume 20: Academic Press; 2001. p. 239–77.
- Hyodo S, Kakumura K, Takagi W, Hasegawa K, Yamaguchi Y. Morphological and functional characteristics of the kidney of cartilaginous fishes: with special reference to urea reabsorption. *Am J Physiol Regul Integr Comp Physiol*. 2014; 307(12):R1381–95. <https://doi.org/10.1152/ajpregu.00033.2014> PMID: [25339681](#)
- LeMoine CM, Walsh PJ. Evolution of urea transporters in vertebrates: adaptation to urea's multiple roles and metabolic sources. *J Exp Biol*. 2015; 218(Pt 12):1936–45. <https://doi.org/10.1242/jeb.114223> PMID: [26085670](#)
- Nawata CM, Walsh PJ, Wood CM. Physiological and molecular responses of the spiny dogfish shark (*Squalus acanthias*) to high environmental ammonia: scavenging for nitrogen. *J Exp Biol*. 2015; 218(Pt 2):238–48. <https://doi.org/10.1242/jeb.114967> PMID: [25609784](#)
- Nakada T, Westhoff CM, Yamaguchi Y, Hyodo S, Li X, Muro T, et al. Rhesus glycoprotein p2 (Rhp2) is a novel member of the Rh family of ammonia transporters highly expressed in shark kidney. *J Biol Chem*. 2010; 285(4):2653–64. <https://doi.org/10.1074/jbc.M109.052068> PMID: [19926789](#)
- Kakumura K, Takabe S, Takagi W, Hasegawa K, Konno N, Bell JD, et al. Morphological and molecular investigations of the holocephalan elephant fish nephron: the existence of a countercurrent-like configuration and two separate diluting segments in the distal tubule. *Cell Tissue Res*. 2015; 362(3):677–88. <https://doi.org/10.1007/s00441-015-2234-4> PMID: [26183720](#)
- Schmidt-Nielsen B, Truniger B, Rabinowitz L. Sodium-linked urea transport by the renal tubule of the spiny dogfish *Squalus acanthias*. *Comp Biochem Physiol A Comp Physiol*. 1972; 42(1):13–25. PMID: [4402703](#)
- Deck CA, Bockus AB, Seibel BA, Walsh PJ. Effects of short-term hyper- and hypo-osmotic exposure on the osmoregulatory strategy of unfed North Pacific spiny dogfish (*Squalus suckleyi*). *Comp Biochem Physiol A Mol Integr Physiol*. 2016; 193:29–35. <https://doi.org/10.1016/j.cbpa.2015.12.004> PMID: [26686463](#)

19. Hazon N, Wells A, Pillans RD, Good JP, Gary Anderson W, Franklin CE. Urea based osmoregulation and endocrine control in elasmobranch fish with special reference to euryhalinity. *Comp Biochem Physiol B Biochem Mol Biol*. 2003; 136(4):685–700. PMID: [14662294](#)
20. Epstein FH. The shark rectal gland: a model for the active transport of chloride. *Yale J Biol Med*. 1979; 52(6):517–23. PMID: [231864](#)
21. Silva P, Solomon RJ, Epstein FH. The rectal gland of *Squalus acanthias*: a model for the transport of chloride. *Kidney Int*. 1996; 49(6):1552–6. PMID: [8743453](#)
22. Riordan JR, Forbush B 3rd, Hanrahan JW. The molecular basis of chloride transport in shark rectal gland. *J Exp Biol*. 1994; 196:405–18. PMID: [7529818](#)
23. Claiborne JB, Choe KP, Morrison-Shettler AI, Weakley JC, Havird J, Freiji A, et al. Molecular detection and immunological localization of gill Na<sup>+</sup>/H<sup>+</sup> exchanger in the dogfish (*Squalus acanthias*). *Am J Physiol Regul Integr Comp Physiol*. 2008; 294(3):R1092–102. <https://doi.org/10.1152/ajpregu.00718.2007> PMID: [18094061](#)
24. Lehrich RW, Aller SG, Webster P, Marino CR, Forrest JN Jr. Vasoactive intestinal peptide, forskolin, and genistein increase apical CFTR trafficking in the rectal gland of the spiny dogfish, *Squalus acanthias*. Acute regulation of CFTR trafficking in an intact epithelium. *J Clin Invest*. 1998; 101(4):737–45. <https://doi.org/10.1172/JCI803> PMID: [9466967](#)
25. Burger JW, Hess WN. Function of the rectal gland in the spiny dogfish. *Science*. 1960; 131(3401):670–1. PMID: [13806061](#)
26. Evans DH. A brief history of the study of fish osmoregulation: the central role of the Mt. Desert Island Biological Laboratory. *Front Physiol*. 2010; 1:13. <https://doi.org/10.3389/fphys.2010.00013> PMID: [21423356](#)
27. Wood CM, Walsh PJ, Kajimura M, McClelland GB, Chew SF. The influence of feeding and fasting on plasma metabolites in the dogfish shark (*Squalus acanthias*). *Comp Biochem Physiol A Mol Integr Physiol*. 2010; 155(4):435–44. <https://doi.org/10.1016/j.cbpa.2009.09.006> PMID: [19782147](#)
28. Villalobos AR, Miller DS, Renfro JL. Transepithelial organic anion transport by shark choroid plexus. *Am J Physiol Regul Integr Comp Physiol*. 2002; 282(5):R1308–16. <https://doi.org/10.1152/ajpregu.00677.2001> PMID: [11959670](#)
29. Reichel V, Miller DS, Fricker G. Texas Red transport across rat and dogfish shark (*Squalus acanthias*) choroid plexus. *Am J Physiol Regul Integr Comp Physiol*. 2008; 295(4):R1311–9. <https://doi.org/10.1152/ajpregu.90373.2008> PMID: [18650317](#)
30. Baehr CH, Fricker G, Miller DS. Fluorescein-methotrexate transport in dogfish shark (*Squalus acanthias*) choroid plexus. *Am J Physiol Regul Integr Comp Physiol*. 2006; 291(2):R464–72. <https://doi.org/10.1152/ajpregu.00814.2005> PMID: [16914433](#)
31. Moore KS, Wehrli S, Roder H, Rogers M, Forrest JN, McCrimmon D, et al. Squalamine: an aminosterol antibiotic from the shark. *Proc Natl Acad Sci U S A*. 1993; 90(4):1354–8. PMID: [8433993](#)
32. Rao MN, Shinnar AE, Noecker LA, Chao TL, Feibush B, Snyder B, et al. Aminosterols from the dogfish shark *Squalus acanthias*. *J Nat Prod*. 2000; 63(5):631–5. PMID: [10843574](#)
33. Sills AK Jr, Williams JI, Tyler BM, Epstein DS, Sipsos EP, Davis JD, et al. Squalamine inhibits angiogenesis and solid tumor growth in vivo and perturbs embryonic vasculature. *Cancer Res*. 1998; 58(13):2784–92. PMID: [9661892](#)
34. Li D, Williams JI, Pietras RJ. Squalamine and cisplatin block angiogenesis and growth of human ovarian cancer cells with or without HER-2 gene overexpression. *Oncogene*. 2002; 21(18):2805–14. <https://doi.org/10.1038/sj.onc.1205410> PMID: [11973639](#)
35. Perni M, Galvagnion C. A natural product inhibits the initiation of alpha-synuclein aggregation and suppresses its toxicity. 2017; 114(6):E1009–e17.
36. Parton A, Forest D, Kobayashi H, Dowell L, Bayne C, Barnes D. Cell and molecular biology of SAE, a cell line from the spiny dogfish shark, *Squalus acanthias*. *Comp Biochem Physiol C Toxicol Pharmacol*. 2007; 145(1):111–9. <https://doi.org/10.1016/j.cbpc.2006.07.003> PMID: [16949345](#)
37. Barnes DW. Cell and molecular biology of the spiny dogfish *Squalus acanthias* and little skate *Leucoraja erinacea*: insights from in vitro cultured cells. *J Fish Biol*. 2012; 80(5):2089–111. <https://doi.org/10.1111/j.1095-8649.2011.03205.x> PMID: [22497417](#)
38. Forest D, Nishikawa R, Kobayashi H, Parton A, Bayne CJ, Barnes DW. RNA expression in a cartilaginous fish cell line reveals ancient 3' noncoding regions highly conserved in vertebrates. *Proc Natl Acad Sci U S A*. 2007; 104(4):1224–9. <https://doi.org/10.1073/pnas.0610350104> PMID: [17227856](#)
39. Deck CA, McKay SJ, Fiedler TJ, LeMoine CM, Kajimura M, Nawata CM, et al. Transcriptome responses in the rectal gland of fed and fasted spiny dogfish shark (*Squalus acanthias*) determined by suppression subtractive hybridization. *Comp Biochem Physiol Part D Genomics Proteomics*. 2013; 8(4):334–43. <https://doi.org/10.1016/j.cbd.2013.09.003> PMID: [24145117](#)

40. Parton A, Bayne CJ, Barnes DW. Analysis and functional annotation of expressed sequence tags from in vitro cell lines of elasmobranchs: Spiny dogfish shark (*Squalus acanthias*) and little skate (*Leucoraja erinacea*). *Comp Biochem Physiol Part D Genomics Proteomics*. 2010; 5(3):199–206. <https://doi.org/10.1016/j.cbd.2010.04.004> PMID: 20471924
41. Venkatesh B, Lee AP, Ravi V, Maurya AK, Lian MM, Swann JB, et al. Elephant shark genome provides unique insights into gnathostome evolution. *Nature*. 2014; 505(7482):174–9. <https://doi.org/10.1038/nature12826> PMID: 24402279
42. King BL, Gillis JA, Carlisle HR, Dahn RD. A natural deletion of the HoxC cluster in elasmobranch fishes. *Science (New York, Ny)*. 2011; 334(6062):1517–.
43. Mulley JF, Hargreaves AD, Hegarty MJ, Heller RS, Swain MT. Transcriptomic analysis of the lesser spotted catshark (*Scyliorhinus canicula*) pancreas, liver and brain reveals molecular level conservation of vertebrate pancreas function. *BMC Genomics*. 2014; 15:1074. <https://doi.org/10.1186/1471-2164-15-1074> PMID: 25480530
44. Richards VP, Suzuki H, Stanhope MJ, Shivji MS. Characterization of the heart transcriptome of the white shark (*Carcharodon carcharias*). *BMC Genomics*. 2013; 14:697. <https://doi.org/10.1186/1471-2164-14-697> PMID: 24112713
45. Krishnaswamy Gopalan T, Gururaj P, Gupta R, Gopal DR, Rajesh P, Chidambaram B, et al. Transcriptome Profiling Reveals Higher Vertebrate Orthologous of Intra-Cytoplasmic Pattern Recognition Receptors in Grey Bamboo Shark. *PLoS ONE*. 2014; 9(6):e100018. <https://doi.org/10.1371/journal.pone.0100018> PMID: 24956167
46. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol*. 2011; 29(7):644–52. <https://doi.org/10.1038/nbt.1883> PMID: 21572440
47. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc*. 2013; 8(8):1494–512. <https://doi.org/10.1038/nprot.2013.084> PMID: 23845962
48. Simao FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*. 2015; 31(19):3210–2. <https://doi.org/10.1093/bioinformatics/btv351> PMID: 26059717
49. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990; 215(3):403–10. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2) PMID: 2231712
50. Hamlett WC, Jezior M, Spieler R. Ultrastructural analysis of folliculogenesis in the ovary of the yellow spotted stingray, *Urolophus jamaicensis*. *Ann Anat*. 1999; 181(2):159–72. [https://doi.org/10.1016/S0940-9602\(99\)80003-X](https://doi.org/10.1016/S0940-9602(99)80003-X) PMID: 10332520
51. Li R, Wang T, Bird S, Zou J, Dooley H, Seccombe CJ. B cell receptor accessory molecule CD79alpha: characterisation and expression analysis in a cartilaginous fish, the spiny dogfish (*Squalus acanthias*). *Fish Shellfish Immunol*. 2013; 34(6):1404–15. <https://doi.org/10.1016/j.fsi.2013.02.015> PMID: 23454429
52. Anderson PM. Glutamine- and N-acetylglutamate-dependent carbamoyl phosphate synthetase in elasmobranchs. *Science*. 1980; 208(4441):291–3. PMID: 6245445
53. Smith CP, Wright PA. Molecular characterization of an elasmobranch urea transporter. *Am J Physiol*. 1999; 276(2 Pt 2):R622–6. PMID: 9950946
54. Janech MG, Fitzgibbon WR, Chen R, Nowak MW, Miller DH, Paul RV, et al. Molecular and functional characterization of a urea transporter from the kidney of the Atlantic stingray. *Am J Physiol Renal Physiol*. 2003; 284(5):F996–f1005. <https://doi.org/10.1152/ajprenal.00174.2002> PMID: 12388386
55. Janech MG, Fitzgibbon WR, Nowak MW, Miller DH, Paul RV, Ploth DW. Cloning and functional characterization of a second urea transporter from the kidney of the Atlantic stingray, *Dasyatis sabina*. *Am J Physiol Regul Integr Comp Physiol*. 2006; 291(3):R844–53. <https://doi.org/10.1152/ajpregu.00739.2005> PMID: 16614049
56. Finn RN, Chauvigne F, Hlidberg JB, Cutler CP, Cerda J. The lineage-specific evolution of aquaporin gene clusters facilitated tetrapod terrestrial adaptation. *PLoS ONE*. 2014; 9(11):e113686. <https://doi.org/10.1371/journal.pone.0113686> PMID: 25426855
57. Cutler CP, Maciver B, Cramb G, Zeidel M. Aquaporin 4 is a Ubiquitously Expressed Isoform in the Dogfish (*Squalus acanthias*) Shark. *Front Physiol*. 2011; 2:107. <https://doi.org/10.3389/fphys.2011.00107> PMID: 22291652
58. Takagi W, Kajimura M, Tanaka H, Hasegawa K, Bell JD, Toop T, et al. Urea-based osmoregulation in the developing embryo of oviparous cartilaginous fish (*Callorhynchus milii*): contribution of the extraembryonic yolk sac during the early developmental period. *J Exp Biol*. 2014; 217(Pt 8):1353–62. <https://doi.org/10.1242/jeb.094649> PMID: 24363418

59. Matthews GD, Gould RM, Vardimon L. A single glutamine synthetase gene produces tissue-specific subcellular localization by alternative splicing. *FEBS Lett.* 2005; 579(25):5527–34. <https://doi.org/10.1016/j.febslet.2005.08.082> PMID: 16213501
60. Kakumura K, Watanabe S, Bell JD, Donald JA, Toop T, Kaneko T, et al. Multiple urea transporter proteins in the kidney of holocephalan elephant fish (*Callorhynchus milii*). *Comp Biochem Physiol B Biochem Mol Biol.* 2009; 154(2):239–47. <https://doi.org/10.1016/j.cbpb.2009.06.009> PMID: 19559810
61. Litman T, Sogaard R, Zeuthen T. Ammonia and urea permeability of mammalian aquaporins. *Handb Exp Pharmacol.* 2009;( 190):327–58. [https://doi.org/10.1007/978-3-540-79885-9\\_17](https://doi.org/10.1007/978-3-540-79885-9_17) PMID: 19096786
62. Li C, Wang W. Urea transport mediated by aquaporin water channel proteins. *Subcell Biochem.* 2014; 73:227–65. [https://doi.org/10.1007/978-94-017-9343-8\\_14](https://doi.org/10.1007/978-94-017-9343-8_14) PMID: 25298348
63. Wright PA, Wood CM. A new paradigm for ammonia excretion in aquatic animals: role of Rhesus (Rh) glycoproteins. *J Exp Biol.* 2009; 212(Pt 15):2303–12. <https://doi.org/10.1242/jeb.023085> PMID: 19617422
64. Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, et al. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res.* 2007; 35(21):7188–96. <https://doi.org/10.1093/nar/gkm864> PMID: 17947321
65. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014; 30(15):2114–20. <https://doi.org/10.1093/bioinformatics/btu170> PMID: 24695404
66. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 2009; 10(3):R25. <https://doi.org/10.1186/gb-2009-10-3-r25> PMID: 19261174
67. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics.* 2011; 12:323. <https://doi.org/10.1186/1471-2105-12-323> PMID: 21816040
68. Eddy SR. Profile hidden Markov models. *Bioinformatics.* 1998; 14(9):755–63. E PMID: 9918945
69. Rice P, Longden I, Bleasby A. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet.* 2000; 16(6):276–7. PMID: 10827456
70. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol.* 2001; 305(3):567–80. <https://doi.org/10.1006/jmbi.2000.4315> PMID: 11152613
71. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods.* 2011; 8(10):785–6. <https://doi.org/10.1038/nmeth.1701> PMID: 21959131
72. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet.* 2000; 25(1):25–9. <https://doi.org/10.1038/75556> PMID: 10802651
73. Powell S, Szklarczyk D, Trachana K, Roth A, Kuhn M, Muller J, et al. eggNOG v3.0: orthologous groups covering 1133 organisms at 41 different taxonomic ranges. *Nucleic Acids Res.* 2012; 40(Database issue):D284–9. <https://doi.org/10.1093/nar/gkr1060> PMID: 22096231
74. Li W, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics.* 2006; 22(13):1658–9. <https://doi.org/10.1093/bioinformatics/btl158> PMID: 16731699
75. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics.* 2012; 28(23):3150–2. <https://doi.org/10.1093/bioinformatics/bts565> PMID: 23060610
76. Remm M, Storm CE, Sonnhammer EL. Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. *J Mol Biol.* 2001; 314(5):1041–52. <https://doi.org/10.1006/jmbi.2000.5197> PMID: 11743721
77. Oliveros JC. VENNY. An interactive tool for comparing lists with Venn Diagrams. <http://bioinfogp.cnb.csic.es/tools/venny/index.html2007>.
78. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004; 32(5):1792–7. <https://doi.org/10.1093/nar/gkh340> PMID: 15034147
79. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics.* 2009; 25(9):1189–91. <https://doi.org/10.1093/bioinformatics/btp033> PMID: 19151095
80. Emanuelsson O, Nielsen H, Brunak S, von Heijne G. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol.* 2000; 300(4):1005–16. <https://doi.org/10.1006/jmbi.2000.3903> PMID: 10891285

81. Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*. 2009; 25(15):1972–3. <https://doi.org/10.1093/bioinformatics/btp348> PMID: 19505945
82. Huerta-Cepas J, Serra F, Bork P. ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. *Mol Biol Evol*. 2016; 33(6):1635–8. <https://doi.org/10.1093/molbev/msw046> PMID: 26921390
83. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 2014; 30(9):1312–3. <https://doi.org/10.1093/bioinformatics/btu033> PMID: 24451623