Author/s:
Thomas, ORB;Swearer, SE;Kapp, EA;Peng, P;Tonkin-Hill, GQ;Papenfuss, A;Roberts, A;Bernard, P;Roberts, BR

Title:
The inner ear proteome of fish

Date:
2019-01-01

Citation:
Thomas, O. R. B., Swearer, S. E., Kapp, E. A., Peng, P., Tonkin-Hill, G. Q., Papenfuss, A., Roberts, A., Bernard, P. & Roberts, B. R. (2019). The inner ear proteome of fish. FEBS Journal, 286 (1), pp.66-81. https://doi.org/10.1111/febs.14715.

Persistent Link:
https://hdl.handle.net/11343/284906

MR. OLIVER ROBERT BION THOMAS (Orcid ID : 0000-0002-8497-670X)

Received Date : 02-Sep-2018
Revised Date : 30-Oct-2018
Accepted Date : 03-Dec-2018
Article type : Regular Paper

Color : Fig 1-8

Suppinfo: Databases S1 to S3

The inner ear proteome of fish

Authors:

Oliver R. B. Thomas[a], Stephen E. Swearer[a], Eugene A. Kapp[b,c], Po Peng[a], Gerry Q. Tonkin-Hill[c], Anthony Papenfuss[c], Anne Roberts[b], Pascal Bernard[a], Blaine R. Roberts[b]

Author Affiliation:

[a]School of BioSciences, The University of Melbourne, Parkville Campus, Victoria, Australia 3010

[b]The Florey Institute of Neuroscience and Mental Health, The University of Melbourne, Parkville Campus, Victoria, Australia 3010

[c]The Walter and Eliza Hall Institute of Medical Research, Parkville, Australia 3010

Corresponding Author:

Oliver R.B. Thomas

School of BioSciences

The University of Melbourne,
Parkville Campus, Victoria,
Australia 3010
email: oliverrbthomas@gmail.com,
ORCID: https://orcid.org/0000-0002-8497-670X,
Ph: +61425704822

Running title: The inner ear proteome of fish

Abbreviations: OMP-1, Otolith Matrix Protein-1; Stm, Starmaker; OMM-64, Otolith Matrix Macromolecule-64; Stm-l, Starmaker-like; CA, Carbonic anhydrase; RP-HPLC, Reverse phase high performance liquid chromatography; LC-MS/MS, Liquid chromatography-tandem mass spectrometry; cDNA, Complementary deoxyribonucleic acid; mRNA, Messenger ribonucleic acid; RefSeq, Reference sequence database; Na^+/K^+-ATPase, Sodium/potassium-transporting ATPase; V-ATPase, V-type proton ATPase; B3AE, Band 3 anion exchange protein; Na^+/Ca^{2+}-exchanger 2, Sodium/calcium exchanger 2; CaM, Calmodulin; PMCA, Plasma membrane Ca^{2+}-ATPase; FAM20C, Extracellular serine/threonine protein kinase FAM20C; MMP 2, Matrix metalloproteinase 2; TIMP 2, Tissue inhibitor of matrix metalloproteinase 2; ApoA1, Apolipoprotein A1; DSPP, Dentin sialophosphoprotein; IDP, Intrinsically disordered protein ; CKII, Casein kinase II; tPA, Tissue-type plasminogen activator; uPA, Urokinase plasminogen activator; BLASTX, Basic local alignment search tool, nucleotide 6-frame translation-protein; TCA, Trichloroacetic acid; TFA, Trifluoroacetic acid; CNBr, Cyanogen bromide; TCEP, Tris(2-carboxyethyl)phosphine; ABC, Ammonium bicarbonate; ACN, Acetonitrile; APL, Andromeda Peak Lists; MGF, Mascot Generic Format; E-value, Expect-value; GRAVY, Grand average of hydropathy

Data deposition:

Proteomics data are available from the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier **PXD01102**1 and **10.6019/PXD011021**. Transcript data are

available from the NCBI BioSample Database with the accession numbers **SAMN10272979, SAMN10272980, SAMN10272981**.

Abstract:

The mechanisms that underpin the formation, growth and composition of otoliths, the biomineralized stones in the inner ear of fish, are largely unknown, as only a few fish inner ear proteins have been reported. Using a partial transcriptome for the inner ear of black bream (*Acanthopagrus butcheri*), in conjunction with proteomic data, we discovered hundreds of previously unknown proteins in the otolith. This allowed us to develop hypotheses to explain the mechanisms of inorganic material supply and daily formation of growth bands. We further identified a likely protein mediator of crystal nucleation and an explanation for the apparent metabolic inertness of the otolith. Due to the formation of both daily and annual increments, otoliths are routinely employed as natural chronometers, being used for age and growth estimation, fisheries stock assessments, and the reconstruction of habitat use, movement, diet, and the impacts of climate change. Our findings provide an unprecedented view of otolith molecular machinery, aiding in the interpretation of these essential archived data.

Keywords: Biomineralization, proteomics, transcriptomics, otolith, endolymph, inner ear, teleost, reverse-phase HPLC, LC-MS/MS, calcium carbonate, extracellular serine/threonine protein kinase FAM20C

Introduction

Otoliths are small, biomineralized earstones within the inner ear of fish that aid in gravity reception and balance [1]. Although understanding their functional anatomy and physiology is important to hearing and balance research, otoliths are especially valuable as biochronometers, archiving age and growth information essential for assessing fisheries stocks globally and predicting their responses to climate change, as well as elemental and stable isotope time series for reconstructing fish life histories of habitat use, movement, pollution exposure and diet [2]. All bony fish have three pairs of otoliths, termed *sagittae, lapilli* and *asterisci* [3]. These differ in size (*sagittae* are usually the largest), location in the inner ear, and calcium carbonate

polymorph. *Sagittae* and *lapilli* are typically aragonite (although calcite has been observed in some species [4]), and *asteriscii* are often composed of vaterite[3]. In this study, the term "otolith" refers to sagittal otoliths. Otoliths grow continuously from embryo, being encased within fluid-filled sacs that are the end organs of the inner ear. The fluid growth medium, endolymph, is rich in both inorganic materials as well as proteins and other accessory macromolecules necessary for otolith formation and growth [5].

Several hypotheses have been suggested to explain the supply of inorganic constituents from the saccular epithelial cells to the endolymph. It has been proposed that calcium is supplied through either a paracellular route or actively through ATPase activity, whereas bicarbonate is transferred through exchange with chloride ions [5-8]. However, the underpinning biochemical mechanisms for ion transfer, exchange and regulation within the endolymph have not been fully elucidated. Otoliths first form shortly after fertilization and grow incrementally, incorporating both inorganic and organic material from the endolymph. A consequence of this is that a physical record of endolymph change is created, allowing for a lifetime chronology to be established. The utility of otoliths as biochronometers is reliant upon a proper understanding of the protein-mediated mechanisms of both nucleation as well as increment formation. These mechanisms are currently poorly understood. Incremental otolith growth occurs via a daily cycle of deposition of alternating aragonite-rich (calcium carbonate) and protein-rich bands. Despite being first documented in 1971 by Panella [9], a suitable explanation for how this happens remains unresolved. Finally, a fundamental assumption of the use of otoliths as biochronometers is that they are metabolically inert and do not undergo any remodeling. Although a study has demonstrated that otoliths are only resorbed under extreme hypoxia [10], no mechanism has been identified for their apparent metabolic inertness. The reason that these fundamental questions remain unanswered is that fewer than a dozen otolith proteins have been identified. In comparison, the understanding of other biomineralizing systems is considerably more advanced: over 1300 proteins in mouse osteoblast extracellular matrix [11] and more than 200 in dental tissues (i.e. dentin, enamel and cementum), which also grow incrementally, are known [12-14].

Although the first constituent of the organic matrix of fish otoliths was described almost 50 years ago [15], to date only a handful of proteins have been successfully identified and characterized.

These include *Otolin-1* (sometimes called *saccular* or *inner-ear specific collagen*) [16], *Otolith Matrix Protein-1* (*OMP-1*) [17], *Starmaker* (*Stm*) [18], *Otolith Matrix Macromolecule-64* (*OMM-64*) [19], *Starmaker-like* (*Stm-l*) [20], an ortholog to mammalian *otoconin 90* (*Otoc1*) [21], *Osteonectin* (*Sparc*) [22], *Neuroserpin* [22], *Tectorin* [4], *Otogelin* [4], *Transferrin* [4], and *Myosin Light Chain 9* [4]. In addition to these structural constituents, other proteins thought to be necessary for biomineralization have also been identified in the inner ear. These include the Na^+/K^+ and Ca^{2+} *ATPases* [23], *GP96* [24], and *Carbonic anhydrase* (*CA*) isozymes [25]. Other macromolecules, such as carbohydrates, proteoglycans, and lipoproteins, have also been detected in the inner ear [26, 27].

To understand in greater detail the protein machinery that composes the otolith we conducted a bottom-up proteomic approach to develop a comprehensive picture of the inner ear for black bream (*Acanthopagrus butcheri*), a species which has been extensively studied in terms of otolith chemistry [28-39]. As otoliths archive proteins from all life stages, we constructed a library composed of a mixture of both adult and juvenile peptides for use in label-free quantitative Mascot searches. The inclusion of adult and juvenile tissues provides a more comprehensive dataset of the proteins likely to be present in an otolith regardless of life stage. This is important for future studies that may compare the proteins in otoliths between various life stages. Briefly, we produced deep sequenced proteomic mass spectra through high pH reverse phase high performance liquid chromatography (RP-HPLC) fractionation [40] followed by traditional low pH-LC tandem mass spectrometry (LC-MS/MS) of proteinaceous material from adult and juvenile otoliths as well as endolymph harvested from wild caught adults. The mass spectrometry peptide data were then matched with annotated cDNA libraries that had been created from homogenized adult brain tissue and juvenile heads and used to create a species-specific peptide library (Fig. 1, Database S1; Table 1, 2). This was then used to identify peptides present in the otolith for adult biological replicates (from the same environment) following a traditional quantitative LC-MS/MS approach. As otoliths are constantly bathed in endolymph, we also identified peptides in endolymph samples harvested at the same time as our otolith replicates. This allowed us to assess the likelihood of whether a given protein was present due to playing a functional role in biomineralization, or as an inescapable consequence of endolymph

contamination – being either trapped during increment formation or adhering to the surface of the otolith.

Results

High pH reverse phase fractionation of both otolith and endolymph samples for peptide library creation proved highly effective, yielding a roughly seven-fold increase in unique peptides when comparing fractions with starting materials (Table 1). Contig sequences from the cDNA FASTA files were successfully aligned with RefSeq proteins. The percentage of significantly aligned proteins differed from file to file, however, with 45.6%, 47.7% and 38.9% for Adult Brain and each of the Juvenile Heads, respectively (Table 2). The top scoring RefSeq protein was selected for each alignment, and the redundancies removed. Subsequently, the custom peptide library used for quantitative analysis contained 175,396 unique inner ear peptides from black bream, 944 proteins from *Sparidae*, 52,715 top hit proteins from RefSeq, *OMP-1* from *L. crocea*, and common contaminants (Database S1).

Mascot searches against the custom library for label-free quantitative assessment of inner ear protein abundances resulted in the identification of 1,065 otolith peptides and 2,685 endolymph peptides. There were 997 total protein types present in otolith and endolymph tissue, with 25.3% occurring in the otolith, 61.1% in the endolymph and only 13.6% of these occurring in both sample types (Fig. 2). Of those proteins occurring in the otolith, 54.4% were common to all three replicates. We were able to identify the presence and relative abundance, at the peptide level, of over 380 proteins present in the otolith, including all previously identified otolith matrix proteins as well as 42 uncharacterized proteins (Fig. 3, Database S2, Database S3). Please note here that feature intensity, is a guideline to sample protein abundance. This can be affected by several factors including how well a given protein is extracted from the otolith, the effectiveness of it undergoing trypsin digestion, and the readiness of individual peptides to ionize within the MS/MS. A truncated list of inner ear proteins is presented, along with their ratios of otolith:endolymph abundance, in Fig. 4. Proteins presented are either known fish inner ear proteins or known to be involved in vertebrate biomineralization.

We identified several proteins from our quantitative proteomic data that are likely to be involved in inorganic material secretion; *Sodium/potassium-transporting ATPase* (*Na^+/K^+-ATPase*), *V-type proton ATPase* (*V-ATPase*), *Band 3 anion exchange protein* (*B3AE*), *Sodium/calcium exchanger 2* (*Na^+/Ca^{2+}-exchanger 2*) and *Calmodulin* (*CaM*). We also found peptides aligning with *Plasma membrane Ca^{2+}-ATPase* (*PMCA)* in our deep sequence data. We found two isozymes of *CA; CA 1* and *CA 6*. Interestingly, the former was restricted to the endolymph and the latter to the otolith.

We found, enriched in the otolith, two key proteins in nucleation and polymorph control: a phosphorylated putative *Stm* homolog as well as a kinase known to be involved in biomineralization, *Extracellular serine/threonine protein kinase FAM20C* (*FAM20C*) *[41]*. *Stm* and its homologs are intrinsically disordered proteins, and thus have low sequence homology, despite playing similar functional roles [42-45]. Consequently, we identified the black bream *Stm* homolog through Ward's minimum variance hierarchical clustering method (Fig. 5, Materials and Methods). Our constellation plot revealed high levels of conservation between all otolith proteins (except the *Stm* homologs) across taxa. Additionally, we undertook an error tolerant search of our proteomics data which identified that our putative *Stm* homolog has at least 4 phosphorylated serine residues (Database S3).

We, like others, also found considerable levels in the otolith of *Otolin-1*, *OMP-1*, and *Neuroserpin*, a member of the serine protease inhibitor family [4, 22, 46, 47]. In addition to this, we identified *Sparc*, which plays a major role in the biomineralization of bone [22]. Unlike previous studies, however, we also found the presence of other substrates to *Neuroserpin* and *Sparc* in the otolith and/or the endolymph: *Plasminogen*, *Matrix metalloproteinase 2* (*MMP 2*), and *Tissue inhibitor of metalloproteinase 2* (*TIMP 2*) [48].

In addition to the 'usual' otolith proteins, we found high abundances of bone and cartilage type proteoglycans (*Mimecan* (*Osteoglycin*), *Epiphycan, Biglycan, Decorin, Aggrecan*) , *Collagen types I* and *II*, and osteoblast regulating proteins (e.g. *E3 ligase itchy-like*) [11, 49-52]. Importantly, we found no evidence of protein markers for osteoclastic function. Our data set also reveals considerable levels of *Apolipoprotein A1* (*ApoA1*) and *Transferrin* as well as five key

proteoglycans, although the former two, being prime components of the endolymph, are probably the result of endolymph fluid trapping during increment deposition rather than their playing a functional role. It is possible, however, that *ApoA1* may be involved in initiating increment formation through binding a product of polyketide synthase, an enzyme recently implicated in proper otolith biomineralization [53]. Similarly, it cannot be ruled out that *Transferrin*, an iron transporter, plays a role in otolith growth [54].

Discussion

Supply of inorganic materials

The necessary organic and inorganic materials for otolith growth are secreted from saccular epithelial cells into the endolymph. For aragonite growth to occur, a sufficient proton gradient must be present along with bicarbonate and calcium ions [5]. It is likely that, at the epithelial cell plasma membrane, bicarbonate ions are being supplied to the endolymph by *B3AE* whereas calcium is transported by a combination of *PMCA* and Na^{+}/Ca^{2+}*-exchanger 2*. *B3AE* is a transmembrane protein that has been previously observed to regulate calcification in kidney duct cells [55]. *Carbonic anhydrase* (*CA*) has been implicated in otolith growth, its expression following a diurnal pattern [56]. It is likely that *CA 1* acts to produce bicarbonate ions from carbon dioxide in the epithelial cells, which is then supplied to the endolymph. At the growing edge of the otolith, *CA 6* is likely to be cooperating with *V-ATPase* in the plasma membrane of epithelial cells to maintain pH while also maintaining the observed constant supersaturation of bicarbonate in the endolymph [47]. These two proteins have been suggested to provide the same function during enamel maturation in teeth [57]. In Fig. 6, we present our hypothesis for the supply of inorganic components to the endolymph, expanding on previous models [5].

The importance of phosphorylation in nucleation and polymorph control

The nucleus (or primordium) of the otolith can be described as a dense organic core surrounded by a rigid calcified layer. The very early expression patterns of *Stm* in the inner ear [18] suggests it is involved in initial otolith seeding. *Stm, Stm-l, OMM-64,* and the functionally similar human *Dentin sialophosphoprotein* (*DSPP*) are intrinsically disordered proteins (IDPs), having low sequence homology but comparable functional roles in biomineral growth. As *Stm* and *OMM-64*

regulate crystal polymorph selection in the otoliths of zebrafish (*Danio rerio*) and rainbow trout (*Oncorhynchus mykiss*) respectively, it is likely that our *Stm* homolog performs the same function in black bream. Söllner [18] demonstrated in the seminal knockout studies of *Stm,* that a lack of this protein invariably resulted in calcite and malformed "star-shaped" otoliths. From this it was inferred that *Stm* both coordinates calcium ions for proper precipitation as aragonite as well as restricts the plane of crystal growth. Similarly, an *in vitro* study of *OMM-64* and *Otolin-1* suggested a cooperative role where absence of the former led to calcite being formed, whereas absence of the latter resulted in vaterite [58]. Considering that calcite is the naturally occurring inorganic form of calcium carbonate, the fact that the protein scaffold facilitates the precipitation of aragonite, which is normally only stable at much higher temperatures, is of prime interest.

Wojtas et al. [59] showed that the ability of *Stm* to act as a crystal nucleation factor is directly related to its degree of phosphorylation. In that study, the authors phosphorylated *Stm in vitro* through action with *Casein Kinase II* (*CK II*) and examined its effectiveness as a crystal nucleator and regulator. The authors suggest that the increased negative charge due to phosphorylation allows greater binding of calcium, and later allows for inhibition and control of crystallization following nucleation. A similar regulation of nucleation and mineral growth through phosphorylation has been shown in *amelogenin* in enamel [60, 61]. It has also been shown that, with IDPs, an increased degree of calcium binding leads to greater compaction of the protein, allowing for closer packing and thereby regulating the extent of crystal growth [45, 59].

In inorganic chemical analyses of otolith cores, the primordium has been found to be enriched in both manganese and phosphorus [62]. Until now, no suitable explanation has been given for this observation. In a previous study, we identified the presence of phosphorus exclusively in the high molecular weight fractions of endolymph [63]. The kinase *FAM20C* has two manganese (II) ion cofactors and has been found to regulate biomineralization pathways through phosphorylation of the small integrin-binding ligand, N-linked glycoprotein (SIBLING) family, which are found in the extracellular matrices of dentin and bone [41]. The ortholog of *Stm*, *DSPP*, is a member of this family. *FAM20C* acts by targeting a Serine-X-glutamic acid/phosphoserine (S-X-E/pS) motif [41].

We suggest, considering *FAM20C's* presence in our data (as well as the absence of other suitable kinases such as *CK II*), coupled with the evidence of manganese and phosphorus enrichment at the otolith core, that it is acting to phosphorylate the black bream *Stm* homolog during the crucial early period of nucleation. Only with effective and complete phosphorylation of *Stm* can the proper polymorph, aragonite, be selected during seeding, thereby setting the template for crystal growth. Furthermore, Reimer et al. [64] observed an increased prevalence of vaterite among otoliths from farmed salmon and hypothesized that this was due to a shift in the *Otolin-1*/*OMM-64* ratio and Gauldie [65] detected increased levels of phosphorus in aragonite otoliths when compared with vaterite otoliths. We hypothesise that it is not the ratio of these proteins that is of importance to polymorph selection, but rather the degree of phosphorylation of the *Stm* homolog that is present both during nucleation as well as later, during increment deposition.

These hypotheses (i.e. the roles of *Stm* and *FAM20C* in nucleation and polymorph selection) can be tested in two ways. Firstly, one can investigate the levels of expression of these proteins in embryonic fish. Secondly, to further understand their roles in polymorph selection, one could compare the proteins and associated degrees of phosphorylation between *asterisci* (vaterite) and *sagittae* (aragonite) otoliths.

Diel entrainment of increment deposition

Increment deposition is well characterized as occurring diurnally, with the main organic and mineral components of an increment being deposited during the crepuscular periods of the night:day cycle. As the endolymph is constantly supersaturated with respect to aragonite, Takagi et al. speculated that the timing of increment deposition is purely a function of organic molecules [47]. The main structural proteins in the organic matrix are *Otolin-1*, which has sequence homology to *Collagen X* as well as a C-terminus C1q domain, and *OMP-1*, which contains a transferrin-like domain [17, 66]. *Otolin* has been found to interact with *Otoconin 90* in the formation of otoconia in mammals [67]. In fish it may act as a general interaction hub for otolith organic matrix proteins (such as *OMP-1*) through trimerization of its C1q domain [68]. Consequently, we suggest that it specifically acts as a stabilizing molecule for the organic scaffold [69], providing a link between the various collagens and abundant *OMP-1* molecules in the matrix. This hypothesis is consistent with previous studies demonstrating no effect from

knockdown of *OMP-1* or *Otolin-1* on nucleation [69], and that the timing of expression of both proteins coincided with increment deposition [46, 70].

Considering this underpinning role for *Otolin-1*, any circadian regulation of the organic matrix must occur through interaction with this crucial molecule. We found both *MMP 2* and *TIMP 2*, proteins whose expression likely follows a circadian pattern as knockout studies of the clock gene PER2 has resulted in a shift in their levels [71]. *MMP 2* catalyzes the cleavage of several types of collagen, including type X. Considering that *MMP 2* is also able to cleave those collagen types occurring in the macular epithelium, organic matrix, otolith membrane and kinocilia, we suggest that regulation of *MMP 2* activity – specifically cleavage of *Otolin-1* – is essential for the cyclical growth pattern of the organic matrix. As *Otolin-1* is likely to be an interaction hub for the otolith matrix proteins, any cessation of organic growth can be achieved through its proteolysis.

Mineralization likely occurs at dawn, with *Otolin-1* expression occurring throughout the night [56, 72-74]. Our adult fish were caught in the late afternoon. This is reflected in the relatively high abundance of *Otolin-1* in our endolymph samples (Fig. 4). We theorize that at dawn, organic matrix growth is halted through a proteolytic cascade. Specifically, the inactive protease *Plasminogen* is coordinated and targeted to *Otolin-1* by *Sparc*. *Sparc* is multi-functional, being able to bind to calcium as well as collagens, and is therefore likely to have roles in both chaperoning collagen to the organic matrix and aiding mineral precipitation [22]. *Sparc* has been additionally shown to be able to enhance the activity of *Plasmin* (the active form of *Plasminogen*) [22]. *Plasminogen* is converted to its active form, the protease *Plasmin*, by either *Tissue-type plasminogen activator* (*tPA*) or *Urokinase plasminogen activator (uPA)*. Indeed, examination of our deep-sequenced proteomic data indicates the presence of *tPA* in the otolith, suggesting that its expression may also be circadian in nature, with its absence from the endolymph likely because we collected fish in the late afternoon. *Plasmin* then activates *MMP 2* [48], which cleaves *Otolin-1*, halting growth of the organic matrix and allowing aragonite mineralization to begin. At dusk, we suggest that this proteolytic cascade is inhibited by *TIMP 2* and *Neuroserpin*, allowing growth of the organic matrix to recommence. *TIMP* 2 inhibits any active *MMP 2* in the endolymph, whereas *Neuroserpin* halts any further activation of *MMP 2*.

Neuroserpin is well established as a regulator of proteolytic cascades through its inhibition of *tPA* and *uPA*. A flowchart of our proposed mechanism for circadian control of organic matrix growth is presented in Fig. 7. Whilst our proposed mechanism is hypothetical, we believe that it can be effectively tested through an immunohistochemistry study focused on the timing of expression of these proteins (and their interactions) over a 24 h period, in combination with genetic manipulation, in a model organism such as *D. rerio*. This would be especially valuable as very little is known about the daily timing of otolith proteins other than *Otolin-1* and *OMP-1*.

Otoliths are like bones but inert

We found the presence of many other proteins involved in the biomineralization of bone, cartilage or teeth in other vertebrates (Fig. 8) [11, 12, 49, 50]. This suggests that the macromolecules involved in biomineralization are highly conserved - a given protein may carry out the same function in separate systems, regardless of whether the mineral is hydroxyapatite (bone, cartilage or teeth) or calcium carbonate (otolith). Interestingly, while we found an abundance of osteoblast-associated proteins (e.g. *E3 ligase itchy-like*, *Collagen triple helix repeat-containing protein 1* [52, 75, 76]), we found no typical markers of osteoclast activity such as *Carbonic anhydrase 2*, *Cathepsin K*, *Calcitonin receptor*, *Cell surface aminopeptidase N/CD13*, or *Integrin β3 [77]*. *Cathepsin K*, particularly, is abundantly expressed by osteoclasts [78]. In contrast to this, evidence of osteoblast-like activity has been observed in previous studies of fish otoliths, notably with *Osterix* (a protein that differentiates osteoblasts) being identified in the piscine inner ear [79]. In a two-way biomineralizing system, which involves the alternating processes of growth and remodeling, the lack of one of these protein types would be unusual. The lack of osteoclastic pathways supports the long-held assumption that otolith formation is unidirectional and doesn't undergo remodeling. This provides further support to the assumption that otoliths are 'permanent' archives of fish life histories.

Conclusion

In conclusion, we identified over 380 proteins in the otoliths of adult black bream. As the majority of these are unlikely to be directly involved in biomineralization, it suggests that a diversity of proteins present in the endolymph are being trapped in the otolith during increment formation. This raises the intriguing possibility that the otolith is not only archiving elemental

markers of environmental history, but also protein markers of development and physiological change over an individual's lifetime. Depending on the lability and spatial resolution of these biomarkers, this has the potential to reconstruct the life histories of fishes to an unprecedented level, opening up new and exciting avenues of research in ecology and fisheries science.

Materials and Methods

Animals

Juvenile *Acanthopagrus butcheri* were caught *via* seine net either on the Werribee River, Australia (37°57'45.6"S, 144°40'02.9"E, cDNA library creation) in May 2016 or in the Gippsland Lakes (37°51'28.3572"S, 147°44'39.0984"E, peptide library creation) in July 2017. Juveniles were measured (5-7 cm fork length) and then euthanised. Following this, they were immediately placed on dry ice until their return to the University of Melbourne, and stored at -80 °C. Adult *A. butcheri* were caught *via* gill net on the Werribee River, Australia over a range of dates; August 2015, March 2016, February and June 2017. Adult fish were measured (30-38 cm fork length), euthanised and decapitated at the first cervical vertebra. Following decapitation, the otoliths, endolymph and brain tissue (which also included saccular tissue) were removed and individually stored on dry ice in labelled microcentrifuge tubes. Our method for careful removal of otoliths and endolymph has been described previously [63]. Upon return from the field, all samples were stored at -80 °C at the University of Melbourne. Fish were collected under Arthur Rylah Institute Animal Ethics Committee approval 14/12 and the State of Victoria, Australia Department of Environment and Primary Industries' Fisheries research permits #1135 and #1204.

Chemicals and standards

All standards, samples and buffers were prepared using ultra-pure Milli-Q H_2O (18.2MΩ; Merck Millipore, Australia). Unless otherwise indicated, all chemicals were purchased from Sigma Aldrich, Castle Hill, Australia and were of the highest available purity.

Transcriptome sequencing, assembly and annotation

Two juveniles were selected for transcriptome assembly. They were thawed and decapitated at the first cervical vertebra. Their heads were homogenised in *RNALater*™ and refrozen for

subsequent RNA extraction. A sample of adult tissue (consisting of the brain, as well as saccular tissue) was similarly homogenised in *RNALater*™. Total RNA was extracted and purified from each sample using a QIAGEN RNeasy® Plus Mini Kit, following the supplied protocol. All samples were mRNA-purified, converted to cDNA and prepared for sequencing as per the Low Sample protocol, Illumina® TruSeq® RNA Sample Preparation v2 Guide. cDNA libraries were sequenced on an Illumina® NextSeq 500® at the Walter and Eliza Hall Institute of Medical Research (WEHI). cDNA sequences were then assembled using Trinity (v2.2.0). FastQC (v0.11.5) was initially employed to generate quality control summaries for the input FASTQ files. The results from this suggested that there was a small amount of adapter contamination at the beginning of the reads. Consequently, three similar assembly pipelines were run. The first used the default Trinity pipeline whereas the second and third trimmed the input reads by removing the first 5 bases and running Trimmomatic (v 0.36) respectively. This ultimately produced three cDNA FASTA files – two for juvenile heads, and one for adult brain.

The three cDNA FASTA files were then annotated by using BLASTX (v2.6.0+) to compute alignments between the query cDNA contigs and the complete RefSeq protein database (retrieved on May 28 2017, ftp://ftp.ncbi.nlm.nih.gov/refseq/release/) with a threshold Expect value (E-value) of 1E-6, on a supercomputer (X99-PRO server, Xeon E7 processors, 32GB RAM, Plant Genome and Molecular Evolution Lab, Department of Life Science, Tunghai University, Taiwan). The output was parsed to a tabular format and the matches per query sequence were sorted hierarchically based on Bitscore, E-value, and Percentage Identity. The best matches for each query were extracted, compiled to a table, and imported into OpenRefine (v2.7).

Tissue preparation for peptide library creation

Otoliths from wild caught juveniles (n=10, 6-28 mg) and adults (n=6, 74-120 mg) were dissolved in an excess of 20% w/v trichloroacetic acid (TCA, Chem-Supply, Australia) for 2h. TCA at this concentration has been demonstrated to be optimal for otolith protein extraction, as it both dissolves the mineral phase as well as precipitates the entirety of the protein phase while minimizing protein denaturation [4]. The dissolved otoliths were then centrifuged for 30 min at 16000 x g (Eppendorf 5415 D), after which the supernatant was discarded, and the remaining

pellets washed with 100 µL of -20 °C acetone and recentrifuged for 5 min at 16000 x g. The pellets were air-dried before being resuspended in 1mL of 90% v/v trifluoroacetic acid (TFA) in the presence of 50 mg of cyanogen bromide (CNBr). These were then incubated, in the dark, at ambient conditions for 48 h. Finally, the suspensions were frozen at -80 °C and lyophilized (Christ Alpha 1-4 LD Plus). Protein content of each otolith was determined through the Bradford Assay indicating that 0.1-0.5% w/w protein was extracted from each otolith.

Endolymph samples from wild caught adults (n=4) and the lyophilized otolith material were suspended in 100 µL of 8 M Urea 100 mM ammonium bicarbonate (ABC). The protein content of each sample was determined through the Bradford Assay. Tris(2-carboxyethyl)phosphine (TCEP, ThermoFisher, Australia) was added to a final concentration of 10 mM and incubated at 60°C for 10 min. After this, the solution was alkylated through the addition of Iodoacetamide (IAA) to 40 mM and incubated at 37°C in the dark for 1 h. Each sample then underwent an eightfold dilution with 50 mM ABC/1 mM calcium chloride Finally, trypsin (Sigma T1426 bovine trypsin) was added at a weight ratio of 1:50 and the entire mixture allowed to incubate at 37°C, in the dark, for 12 h. Following trypsin digestion, the samples were acidified in 10% v/v TFA and underwent C18 clean-up on a Waters Positive Pressure-96 Processor, employing 99.9% ACN/0.1% TFA as the initial medium, 0.1% TFA for equilibration and washing, and eluting in 70% v/v ACN/0.1% Formic acid (FA). The cleaned material was vacuum dried (Savant SC100 SpeedVac) and then resuspended in 100 µL of 5% ACN/0.2% ammonium hydroxide (NH_4OH).

Samples were pooled to create both an otolith and an endolymph sample that each contained 200 µg of peptides. These pooled samples were fractionated, following the method of Gilar et al. [40], through high pH reverse phase high performance liquid chromatography (RP-HPLC) on an Agilent Technologies 1200 Series Infinity II fitted with a Waters Acquity UPLC BEH 300 Å column. The buffers employed were 0.2% NH_4OH and 90% ACN/0.2% NH_4OH. Fractions were collected every 0.7 min, resulting in 32 fractions collected from each starting material. These were concatenated to 16 samples by adding fractions that were 16 fractions apart (e.g. Fraction 1 + Fraction 17, Fraction 2 + 18 etc.). The starting material was retained. Finally, the 34 subsequent samples (16 concatenated otolith, 16 concatenated endolymph, and corresponding starting materials) were vacuum dried and resuspended in 50 µL of a 2% ACN/0.05% TFA

buffer. Each of the 34 samples was individually separated and sequenced on a Thermo Scientific Orbitrap Fusion Lumos Tribrid Mass Spectrometer (LC-MS/MS) at the Mass Spectrometry and Proteomics Facility at the Bio21 Institute, the University of Melbourne, Australia.

Custom peptide library creation

The resulting raw mass spectra from the fractionated endolymph and otolith material were run through MaxQuant (v1.5.8.3) to provide peptide feature intensities. Following this, the Andromeda Peak Lists (APL) were converted to Mascot Generic Format (MGF) using the WEHI APL to MGF converter, before being searched with Mascot (v2.4.0) on the WEHI Mascot server (https://sysbio-mascot.wehi.edu.au/mascot/home.html) against our annotated cDNA FASTA files. This search was conducted in six reading frames, and generated peptide-cDNA matches, while retaining accession information from Trinity. We employed the following thresholds for all searches: a Mascot score of 13 or higher and E-value of 0.01 or lower. These transcript-matched peptides were then combined, in a database, with all currently known *Sparidae* proteins (retrieved on January 3 2017, http://www.uniprot.org). In addition to this, to capture any missing peptide sequences, we also included all top scoring proteins from the results of the transcriptome BLASTX alignments with RefSeq. As a preliminary evaluation of our transcripts revealed the absence of the cDNA entry corresponding to *OMP-1*, we included the full sequence of this from giant yellow croaker, (*Larimichthys crocea*). We did this because, while *OMP-1* mRNA expression is constant in rainbow trout [46, 47], it is possible that it was either not present at the time of death or was otherwise not captured by our mRNA extraction. Finally, we also included common contaminant sequences (e.g. *trypsin, keratin, alpha-2-casein, serum albumin* and others). Our custom peptide library is available as Database S1.

Quantitative assessment

In order to compare the presence of proteins in endolymph and otoliths extracted from individual fish, we undertook a label-free quantitative proteomic approach. Individual otoliths from wild caught adult bream (n=3, 37-59 mg) were manually abraded on all sides with sandpaper (3M Wetordry 1000 grit) to physically remove any adhering membranous tissue or endolymph, and then individually dissolved in excess 20% w/v TCA. Endolymph samples (n=3, 34-90 µL) were thawed to room temperature and similarly dissolved. Both tissue types were then put through the

CNBr/trypsin digest and C18 clean-up protocols as described above for otoliths. Following clean-up, each sample (representing either endolymph or an otolith from an individual fish) was resuspended in 2% ACN/0.05% TFA to give an estimated protein concentration of 0.35 mg/mL. Each sample was then individually run on a Thermo Scientific NanoLC-Q Exactive Plus Hybrid Quadrupole-Orbitrap Mass Spectrometer at the Mass Spectrometry and Proteomics Facility at the Bio21 Institute, University of Melbourne, Australia.

Mass spectra from each otolith (n=3) and endolymph (n=3) sample were similarly analysed with MaxQuant, converted to MGF and searched with Mascot on the WEHI Mascot server against our custom database (peptides from RP-HPLC, *Sparidae* proteins, top-scoring proteins from BLASTX, *L. crocea OMP-1*, contaminants). We employed the same E-value and Mascot score thresholds as described above. Additionally, we performed decoy searches (i.e. utilizing amino acid sequences that had been reversed, shuffled, or Markov chain-generated) to validate these thresholds as per Käll et al. [80]. The outputs from these searches were ranked by feature intensity, providing an estimate of the most to least abundant peptides (and therefore proteins) in each sample. The average feature intensity (FI) for each peptide was determined for the three otolith samples and the three endolymph samples and log 2 (x+1) transformed (Fig. 3, 4). In addition to this, to identify potential post-translational modifications such as phosphorylation and glycosylation, we also performed an error tolerance Mascot search on our data. A flowchart of our library-creation and quantitative analysis methods is presented in Fig. 1. The parameters employed are presented in Table 3.

Identification of a starmaker-like protein

Preliminary Mascot searches revealed peptides for a candidate starmaker-like protein, and its full primary sequence was inferred from its cDNA entry. This sequence was analysed using the protein parameter tool at ExPASy (https://web.expasy.org/protparam/) to provide molecular weight, theoretical isoelectric point, percentage amino acid composition, instability index, aliphatic index and grand average of hydropathy (GRAVY). In addition to the candidate *starmaker-like* protein, parameters were also extracted from the primary sequences of *Otolin-1* (or its homolog), *OMP-1*, *Transferrin*, *Neuroserpin*, *Starmaker/OMM-64/Dentin sialophosphoprotein* (*DSPP*), and *Sparc* for black bream as well as the following species: *Danio*

rerio, Lates calcarifer, *Larimichthys crocea*, *Oryzias latipes*, *Oncorhynchus mykiss*, *Oreochromis niloticus*, *Salmo salar*, *Seriola lalandi*, *Stegastes partitus, Takifugu rubipes* and *Homo sapiens.* These data were then analysed using Ward's minimum variance clustering method, and visualized as a constellation plot, labelling our candidate protein as an unknown (Fig. 5; JMP v 13, SAS Institute Inc).

Data deposition

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier **PXD011021** and **10.6019/PXD011021**. Transcript data have been deposited in NCBI BioSample Database with the accession numbers **SAMN10272979, SAMN10272980, SAMN10272981**.

Acknowledgments: We thank G. Jenkins, M. Sievers, K. Hassell, J. French, D. Chamberlain, and L. Barrett for assistance with sample collection and S. Mukherjee for assistance with laboratory work. We thank the Plant Genome & Molecular Evolution Lab, Department of Life Science, Tunghai University, Taiwan for use of their supercomputer, and Steven Wilcox at the Systems Biology Division at the Walter and Eliza Hall Institute of Medical Research for advice and sequencing of cDNA. We similarly thank Karin Limburg for her advice and comments. Funding was provided by the Australian Research Council (project LP140100087) and the University of Melbourne (John and Allan Gilmour Research Award). ORBT was supported by an Australian Postgraduate Award.

Author contributions: ORBT – Conceptualization, Formal Analysis, Investigation, Methodology, Visualization, Writing – original draft, review and editing;
SES – Conceptualization, Formal analysis, Funding acquisition, Methodology, Supervision, Validation, Writing – review and editing; **EAK** – Data curation, Formal analysis, Methodology, Software; **PP** – Data curation, Formal Analysis, Methodology, Resources, Software; **GQT** – Data curation, Formal analysis, Software; **AP** – Data curation, Formal analysis, Methodology, Resources, Software; **AR** – Investigation, Resources; **PB** – Investigation, Resources; **BRR** – Conceptualization, Formal analysis, Funding acquisition, Methodology, Project administration, Supervision, Validation, Writing – review and editing

References

1. Lowenstein, O. (1936) The equilibrium function of the vertebrate labyrinth, *Biol Rev Camb Philos Soc*. **11**, 113-145.

2. Izzo, C., Doubleday, Z. A., Grammer, G. L., Gilmore, K. L., Alleway, H. K., Barnes, T. C., Disspain, M. C. F., Giraldo, A. J., Mazloumi, N. & Gillanders, B. M. (2016) Fish as proxies of ecological and environmental change, *Rev Fish Biol Fish.* **26**, 265-286.

3. Campana, S. E. (1999) Chemistry and composition of fish otoliths: pathways, mechanisms and applications, *Mar Ecol Prog Ser.* **188**, 263-297.

4. Weigele, J., Franz-Odendaal, T. A. & Hilbig, R. (2016) Not all inner ears are the same: Otolith matrix proteins in the inner ear of sub-adult cichlid fish, oreochromis mossambicus, reveal insights into the biomineralization process, *Anat Rec Adv Integr Anat Evol Biol.* **299**, 234-245.

5. Payan, P., De Pontual, H., Boeuf, G. & Mayer-Gostan, N. (2004) Endolymph chemistry and otolith growth in fish, *C R Palevol.* **3**, 535-547.

6. Cruz, S., Shiao, J.-C., Liao, B.-K., Huang, C.-J. & Hwang, P.-P. (2009) Plasma membrane calcium ATPase required for semicircular canal formation and otolith growth in the zebrafish inner ear, *J Exp Biol.* **212**, 639-647.

7. Mugiya, Y. & Yoshida, M. (1995) Effects of calcium antagonists and other metabolic modulators on in vitro calcium deposition on otoliths in the rainbow trout Oncorhynchus mykiss, *Fish Sci.* **61**, 1026-1030.

8. Tohse, H. & Mugiya, Y. (2001) Effects of enzyme and anion transport inhibitors on in vitro incorporation of inorganic carbon and calcium into endolymph and otoliths in salmon Oncorhynchus masou, *Comparative Biochemistry and Physiology a-Molecular and Integrative Physiology*. **128**, 177-184.

9. Panella, G. (1971) Fish otoliths: daily growth layers and periodical patterns, *Science (New York, NY)*. **173**, 1124-1127.

10. Mugiya, Y. & Uchimura, T. (1989) Otolith resorption induced by anaerobic stress in the goldfish, *Carassius auratus*, *J Fish Biol.* **35**, 813-818.

11. Xiao, Z., Camalier, C. E., Nagashima, K., Chan, K. C., Lucas, D. A., De la Cruz, M. J., Gignac, M., Lockett, S., Issaq, H. J., Veenstra, T. D., Conrads, T. P. & Beck, G. R. (2007)

Analysis of the extracellular matrix vesicle proteome in mineralizing osteoblasts, *J Cell Physiol.* **210**, 325-335.
12. Salmon, C. R., Giorgetti, A. P. O., Leme, A. F. P., Domingues, R. R., Kolli, T. N., Foster, B. L. & Nociti, F. H. (2017) Microproteome of dentoalveolar tissues, *Bone.* **101**, 219-229.
13. Lima Leite, A., Silva Fernandes, M., Charone, S., Whitford, G. M., Everett, E. T. & Buzalaf, M. A. R. (2018) Proteomic Mapping of Dental Enamel Matrix from Inbred Mouse Strains: Unraveling Potential New Players in Enamel, *Caries Res.* **52**, 78-87.
14. Charone, S., De Lima Leite, A., Peres-Buzalaf, C., Silva Fernandes, M., Ferreira de Almeida, L., Zardin Graeff, M. S., Cardoso de Oliveira, R., Campanelli, A. P., Groisman, S., Whitford, G. M., Everett, E. T. & Buzalaf, M. A. (2016) Proteomics of Secretory-Stage and Maturation-Stage Enamel of Genetically Distinct Mice, *Caries Res.* **50**, 24-31.
15. Degens, E. T., Deuser, W. G. & Haedrich, R. L. (1969) Molecular structure and composition of fish otoliths, *Mar Biol.* **2**, 105-113.
16. Davis, J. G., Oberholtzer, J. C., Burns, F. R. & Greene, M. I. (1995) Molecular cloning and characterization of an inner ear-specific structural protein, *Science.* **267**, 1031-4.
17. Murayama, E., Okuno, A., Ohira, T., Takagi, Y. & Nagasawa, H. (2000) Molecular cloning and expression of an otolith matrix protein cDNA from the rainbow trout, *Oncorhynchus mykiss*, *Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology.* **126**, 511-520.
18. Söllner, C., Burghammer, M., Busch-Nentwich, E., Berger, J., Schwarz, H., Riekel, C. & Nicolson, T. (2003) Control of crystal size and lattice formation by starmaker in otolith biomineralization, *Science.* **302**, 282-6.
19. Tohse, H., Takagi, Y. & Nagasawa, H. (2008) Identification of a novel matrix protein contained in a protein aggregate associated with collagen in fish otoliths, *FEBS J.* **275**, 2512-2523.
20. Bajoghli, B., Ramialison, M., Aghaallaei, N., Czerny, T. & Wittbrodt, J. (2009) Identification of starmaker-like in medaka as a putative target gene of Pax2 in the otic vesicle, *Dev Dyn.* **238**, 2860-6.
21. Petko, J. A., Millimaki, B. B., Canfield, V. A., Riley, B. B. & Levenson, R. (2008) Otoc1: A novel otoconin-90 ortholog required for otolith mineralization in zebrafish, *Dev Neurobiol.* **68**, 209-222.

22. Kang, Y. J., Stevenson, A. K., Yau, P. M. & Kollmar, R. (2008) Sparc protein is required for normal growth of zebrafish otoliths, *J Assoc Res Otolaryngol.* **9**, 436-51.
23. Blasiole, B., Canfield, V. A., Vollrath, M. A., Huss, D., Mohideen, M., Dickman, J. D., Cheng, K. C., Fekete, D. M. & Levenson, R. (2006) Separate Na,K-ATPase genes are required for otolith formation and semicircular canal development in zebrafish, *Dev Biol.* **294**, 148-160.
24. Sumanas, S., Larson, J. D. & Bever, M. M. (2003) Zebrafish chaperone protein GP96 is required for otolith formation during ear development, *Dev Biol.* **261**, 443-455.
25. Tohse, H., Ando, H. & Mugiya, Y. (2004) Biochemical properties and immunohistochemical localization of carbonic anhydrase in the sacculus of the inner ear in the salmon *Oncorhynchus masou*, *Comparative Biochemistry and Physiology a-Molecular & Integrative Physiology.* **137**, 87-94.
26. Takagi, Y., Ishida, K. & Mugiya, Y. (2000) Carbohydrates of the otolith organ in the rainbow trout *Oncorhynchus mykiss* detected by lectins, *Fish Sci.* **66**, 933-939.
27. Borelli, G., Mayer-Gostan, N., De Pontual, H., Boeuf, G. & Payan, P. (2001) Biochemical relationships between endolymph and otolith matrix in the trout (*Oncorhynchus mykiss*) and turbot (*Psetta maxima*), *Calcif Tissue Int.* **69**, 356-364.
28. Morison, A. K., Coutin, P. C. & Robertson, S. G. (1998) Age determination of black bream, Acanthopagrus butcheri (Sparidae), from the Gippsland Lakes of south-eastern Australia indicates slow growth and episodic recruitment, *Mar Freshw Res.* **49**, 491-498.
29. Sarre, G. A. & Potter, I. C. (2000) Variation in age compositions and growth rates of Acanthopagrus butcheri (Sparidae) among estuaries: some possible contributing factors, *Fish Bull.* **98**, 785-799.
30. Elsdon, T. S. & Gillanders, B. M. (2003) Relationship between water and otolith elemental concentrations in juvenile black bream Acanthopagrus butcheri, *Mar Ecol Prog Ser.* **260**, 263-272.
31. Elsdon, T. S. & Gillanders, B. M. (2004) Fish otolith chemistry influenced by exposure to multiple environmental variables, *J Exp Mar Biol Ecol.* **313**, 269-284.
32. de Vries, M. C., Gillanders, B. M. & Elsdon, T. S. (2005) Facilitation of barium uptake into fish otoliths: Influence of strontium concentration and salinity, *Geochim Cosmochim Acta.* **69**, 4061-4072.

33. Elsdon, T. S. & Gillanders, B. M. (2005) Alternative life-history patterns of estuarine fish: barium in otoliths elucidates freshwater residency, *Can J Fish Aquat Sci.* **62**, 1143-1152.
34. Ranaldi, M. M. & Gagnon, M. M. (2008) Trace metal incorporation in otoliths of black bream (Acanthopagrus butcheri munro), an indicator of exposure to metal contamination, *Water Air Soil Pollut.* **194**, 31-43.
35. Partridge, G. J., Jenkins, G. I., Doupe, R. G., De Lestang, S., Ginbey, B. M. & French, D. (2009) Factors affecting mark quality of alizarin complexone-stained otoliths in juvenile black bream Acanthopagrus butcheri and a prescription for dosage, *J Fish Biol.* **75**, 1518-1523.
36. Disspain, M., Wallis, L. A. & Gillanders, B. M. (2011) Developing baseline data to understand environmental change: a geochemical study of archaeological otoliths from the Coorong, South Australia, *J Archaeol Sci.* **38**, 1842-1857.
37. Sakabe, R., Lyle, J. M. & Crawford, C. M. (2011) The influence of freshwater inflows on spawning success and early growth of an estuarine resident fish species, Acanthopagrus butcheri, *J Fish Biol.* **78**, 1529-1544.
38. Webb, S. D., Woodcock, S. H. & Gillanders, B. M. (2012) Sources of otolith barium and strontium in estuarine fish and the influence of salinity and temperature, *Mar Ecol Prog Ser.* **453**, 189-199.
39. Williams, J., Jenkins, G. P., Hindell, J. S. & Swearer, S. E. (2018) Fine-scale variability in elemental composition of estuarine water and otoliths: Developing environmental markers for determining larval fish dispersal histories within estuaries, *Limnol Oceanogr.* **63**, 262-277.
40. Gilar, M., Olivova, P., Daly, A. E. & Gebler, J. C. (2005) Two-dimensional separation of peptides using RP-RP-HPLC system with different pH in first and second separation dimensions, *J Sep Sci.* **28**, 1694-1703.
41. Tagliabracci, V. S., Engel, J. L., Wen, J. Z., Wiley, S. E., Worby, C. A., Kinch, L. N., Xiao, J. Y., Grishin, N. V. & Dixon, J. E. (2012) Secreted kinase phosphorylates extracellular proteins that regulate biomineralization, *Science.* **336**, 1150-1153.
42. Kaplon, T. M., Rymarczyk, G., Nocula-Lugowska, M., Jakob, M., Kochman, M., Lisowski, M., Szewczuk, Z. & Ozyhar, A. (2008) Starmaker exhibits properties of an intrinsically disordered protein, *Biomacromolecules.* **9**, 2118-25.
43. Uversky, V. N. & Dunker, A. K. (2010) Understanding protein non-folding, *Biochimica Et Biophysica Acta-Proteins and Proteomics.* **1804**, 1231-1264.

44. Rozycka, M., Wojtas, M., Jakob, M., Stigloher, C., Grzeszkowiak, M., Mazur, M. & Ozyhar, A. (2014) Intrinsically disordered and pliable Starmaker-like protein from medaka (Oryzias latipes) controls the formation of calcium carbonate crystals, *PLoS One.* **9**, e114308.
45. Poznar, M., Hołubowicz, R., Wojtas, M., Gapiński, J., Banachowicz, E., Patkowski, A., Ożyhar, A. & Dobryszycki, P. (2017) Structural properties of the intrinsically disordered, multiple calcium ion-binding otolith matrix macromolecule-64 (OMM-64), *Biochim Biophys Acta.* **1865**, 1358-1371.
46. Murayama, E., Takagi, Y. & Nagasawa, H. (2004) Immunohistochemical localization of two otolith matrix proteins in the otolith and inner ear of the rainbow trout, *Oncorhynchus mykiss*: comparative aspects between the adult inner ear and embryonic otocysts, *Histochem Cell Biol.* **121**, 155-166.
47. Takagi, Y., Tohse, H., Murayama, E., Ohira, T. & Nagasawa, H. (2005) Diel changes in endolymph aragonite saturation rate and mRNA expression of otolith matrix proteins in the trout otolith organ, *Mar Ecol Prog Ser.* **294**, 249-256.
48. Mazzieri, R., Masiero, L., Zanetta, L., Monea, S., Onisto, M., Garbisa, S. & Mignatti, P. (1997) Control of type IV collagenase activity by components of the urokinase-plasmin system: A regulatory mechanism with cell-bound reactants, *EMBO J.* **16**, 2319-2332.
49. Iozzo, R. V. (1998) Matrix proteoglycans: From molecular design to cellular function, *Annu Rev Biochem.* **67**, 609-652.
50. Zhen, E. Y., Brittain, I. J., Laska, D. A., Mitchell, P. G., Sumer, E. U., Karsdal, M. A. & Duffin, K. L. (2008) Characterization of metalloprotease cleavage products of human articular cartilage, *Arthritis Rheum.* **58**, 2420-2431.
51. Ye, Y. P., Hu, W. H., Guo, F. J., Zhang, W. K., Wang, J. & Chen, A. M. (2012) Glycosaminoglycan chains of biglycan promote bone morphogenetic protein-4-induced osteoblast differentiation, *Int J Mol Med.* **30**, 1075-1080.
52. Liu, J., Li, X., Zhang, H., Gu, R., Wang, Z., Gao, Z. & Xing, L. (2017) Ubiquitin E3 ligase Itch negatively regulates osteoblast function by promoting proteasome degradation of osteogenic proteins, *Bone & Joint Research.* **6**, 154-161.
53. Hojo, M., Omi, A., Hamanaka, G., Shindo, K., Shimada, A., Kondo, M., Narita, T., Kiyomoto, M., Katsuyama, Y., Ohnishi, Y., Irie, N. & Takeda, H. (2015) Unexpected link between polyketide synthase and calcium carbonate biomineralization, *Zool Lett.* **1**, 16.

54. Lambert, L. A. (2012) Molecular evolution of the transferrin family and associated receptors, *Biochimica Et Biophysica Acta-General Subjects.* **1820**, 244-255.
55. Drenckhahn, D., Schluter, K., Allen, D. & Bennett, V. (1985) Colocalization of band 3 with ankyrin and spectrin at the basal membrane of intercalated cells in the rat kidney, *Science.* **230**, 1287-1289.
56. Tohse, H., Murayama, E., Ohira, T., Takagi, Y. & Nagasawa, H. (2006) Localization and diurnal variations of carbonic anhydrase rnRNA expression in the inner ear of the rainbow trout *Oncorhynchus mykiss*, *Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology.* **145**, 257-264.
57. Yin, K. & Paine, M. L. (2017) Bicarbonate transport during enamel maturation, *Calcif Tissue Int.* **101**, 457-464.
58. Tohse, H., Saruwatari, K., Kogure, T., Nagasawa, H. & Takagi, Y. (2009) Control of polymorphism and morphology of calcium carbonate crystals by a matrix protein aggregate in fish otoliths, *Crystal Growth & Design.* **9**, 4897-4901.
59. Wojtas, M., Wolcyrz, M., Ozyhar, A. & Dobryszycki, P. (2012) Phosphorylation of intrinsically disordered starmaker protein increases its ability to control the formation of calcium carbonate crystals, *Crystal Growth & Design.* **12**, 158-168.
60. Wiedemann-Bidlack, F. B., Kwak, S. Y., Beniash, E., Yamakoshi, Y., Simmer, J. P. & Margolis, H. C. (2011) Effects of phosphorylation on the self-assembly of native full-length porcine amelogenin and its regulation of calcium phosphate formation in vitro, *J Struct Biol.* **173**, 250-260.
61. Le Norcy, E., Lesieur, J., Sadoine, J., Rochefort, G. Y., Chaussain, C. & Poliard, A. (2018) Phosphorylated and Non-phosphorylated Leucine Rich Amelogenin Peptide Differentially Affect Ameloblast Mineralization, *Front Physiol.* **9**, 10.
62. Shima, J. S. & Swearer, S. E. (2009) Larval quality is shaped by matrix effects: implications for connectivity in a marine metapopulation, *Ecology.* **90**, 1255-1267.
63. Thomas, O. R. B., Ganio, K., Roberts, B. R. & Swearer, S. E. (2017) Trace element-protein interactions in endolymph from the inner ear of fish: implications for environmental reconstructions using fish otolith chemistry, *Metallomics.* **9**, 239-249.

64. Reimer, T., Dempster, T., Wargelius, A., Fjelldal, P. G., Hansen, T., Glover, K. A., Solberg, M. F. & Swearer, S. E. (2017) Rapid growth causes abnormal vaterite formation in farmed fish otoliths, *J Exp Biol.* **220**, 2965-2969.
65. Gauldie, R. W. (1986) Vaterite otoliths from chinook salmon (*Oncorhynchus tshawytscha*), *N Z J Mar Freshwat Res.* **20**, 209-217.
66. Murayama, E., Takagi, Y., Ohira, T., Davis, J. G., Greene, M. I. & Nagasawa, H. (2002) Fish otolith contains a unique structural protein, otolin-1, *Eur J Biochem.* **269**, 688-696.
67. Deans, M. R., Peterson, J. M. & Wong, G. W. (2010) Mammalian Otolin: A multimeric glycoprotein specific to the inner ear that interacts with otoconial matrix protein otoconin-90 and cerebellin-1, *PLoS One.* **5**.
68. Holubowicz, R., Wojtas, M., Taube, M., Kozak, M., Ozyhar, A. & Dobryszycki, P. (2017) Effect of calcium ions on structure and stability of the C1q-like domain of otolin-1 from human and zebrafish, *FEBS J.* **284**, 4278-4297.
69. Murayama, E., Herbomel, P., Kawakami, A., Takeda, H. & Nagasawa, H. (2005) Otolith matrix proteins OMP-1 and Otolin-1 are necessary for normal otolith growth and their correct anchoring onto the sensory maculae, *Mech Dev.* **122**, 791-803.
70. Weigele, J., Franz-Odendaal, T. A. & Hilbig, R. (2016) Formation of the inner ear during embryonic and larval development of the cichlid fish (*Oreochromis mossambicus*), *Connect Tissue Res*, 1-24.
71. Su, X., Chen, D., Yang, K., Zhao, Q., Zhao, D., Lv, X. & Ao, Y. (2017) The circadian clock gene PER2 plays an important role in tumor suppression through regulating tumor-associated genes in human oral squamous cell carcinoma, *Oncol Rep.* **38**, 472-480.
72. Mugiya, Y. (1987) Phase Difference between Calcification and Organic Matrix Formation in the Diurnal Growth of Otoliths in the Rainbow-Trout, Salmo-Gairdneri, *Fish Bull.* **85**, 395-401.
73. Zhang, Z. & Runham, N. W. (1992) Temporal deposition of incremental and discontinuous zones in the otoliths of *Oreochromis niloticus* (L.), *J Fish Biol.* **40**, 333-339.
74. Borelli, G., Guibbohni, M. E., Mayer-Gostan, N., Priouzeau, F., De Pontual, H., Allemand, D., Puverel, S., Tambutte, E. & Payan, P. (2003) Daily variations of endolymph composition: relationship with the otolith calcification process in trout, *J Exp Biol.* **206**, 2685-2692.

75. Jeong, J. A., Ko, K. M., Bae, S., Jeon, C. J., Koh, G. Y. & Kim, H. (2007) Genome-wide differential gene expression profiling of human bone marrow stromal cells, *Stem Cells.* **25**, 994-1002.

76. Siebers, M. C., ter Brugge, P. J., Walboomers, X. F. & Jansen, J. A. (2005) Integrins as linker proteins between osteoblasts and bone replacing materials. A critical review, *Biomaterials.* **26**, 137-46.

77. Castillo, L. M., Guerrero, C. A. & Acosta, O. (2017) Expression of typical osteoclast markers by PBMCs after PEG-induced fusion as a model for studying osteoclast differentiation, *J Mol Histol.* **48**, 169-185.

78. Drake, F. H., Dodds, R. A., James, I. E., Connor, J. R., Debouck, C., Richardson, S., LeeRykaczewski, E., Coleman, L., Rieman, D., Barthlow, R., Hastings, G. & Gowen, M. (1996) Cathepsin K, but not cathepsins B, L, or S, is abundantly expressed in human osteoclasts, *J Biol Chem.* **271**, 12511-12516.

79. Renn, J. & Winkler, C. (2014) Osterix/Sp7 regulates biomineralization of otoliths and bone in medaka (Oryzias latipes), *Matrix Biol.* **34**, 193-204.

80. Kall, L., Storey, J. D., MacCoss, M. J. & Noble, W. S. (2008) Posterior error probabilities and false discovery rates: two sides of the same coin, *J Proteome Res.* **7**, 40-4.

81. Kohfeldt, E., Sasaki, T., Gohring, W. & Timpl, R. (1998) Nidogen-2: a new basement membrane protein with diverse binding properties, *J Mol Biol.* **282**, 99-109.

	Peptides, starting material	**Peptides, RP-HPLC fractions**	**Total**
Otoliths	459	3216	3261
Endolymph	2743	21080	21355

Table 1.

Comparison of unique peptides occurring in starting material versus RP-HPLC fractions for pooled otolith and endolymph samples from *Acanthopagrus butcheri*. All peptides are Mascot Score ≥13 and E-

value ≤ 0.01. Mass spectra were produced on a Thermo Scientific Orbitrap Fusion Lumos Tribrid Mass Spectrometer.

FASTA file	No. of cDNA contigs	No. of successful alignments	No. of best scoring alignments
Adult Brain	92273	17963561	42033
Juvenile Head 1	70334	14465474	33553
Juvenile Head 2	44402	6997060	17258

Table 2.

Results of BLASTX v2.60+ alignment of *Acanthopagrus butcheri* cDNA FASTA with RefSeq proteins. RefSeq proteins were retrieved on May 28 2017, ftp://ftp.ncbi.nlm.nih.gov/refseq/release/, and totalled 84,756,971. An E-value of 1E-6 was employed as a threshold. All search results were sorted hierarchically based on Bitscore, E-value and Percentage Identity.

Parameter	RP-HPLC, Endolymph	RP-HPLC, otolith	Quantitative, Endolymph	Quantitative, Otolith
Missed cleavages	2	2	2	2
Fixed modifications	CAM	CAM	CAM	CAM
Variable modifications	MetOx, N-acetyl, N-pyroGlu	MetOx, N-acetyl, N-pyroGlu, Hse, Hsl	MetOx, N-acetyl, N-pyroGlu, Hse, Hsl	MetOx, N-acetyl, N-pyroGlu, Hse, Hsl
Precursor mass tolerance	20 ppm	20 ppm	20 ppm	20 ppm
Fragment ion tolerance	0.3 Da	0.3 Da	20 mmu	20 mmu
Database used	cDNA	cDNA	Custom	Custom
Mascot score threshold	13	13	13	13

E-value threshold	0.01	0.01	0.01	0.01

Table 3.

Parameters employed for Mascot searches of proteomic data. Reverse-phase high pH liquid chromatography fractionated samples (RP-HPLC) were run on a Thermo Scientific Orbitrap Fusion Lumos Tribrid Mass Spectrometer. Quantitative samples were run on a Thermo Scientific NanoLC-Q Exactive Plus Hybrid Quadrupole-Orbitrap Mass Spectrometer. Abbreviations are as follows: CNBr – cyanogen bromide; CAM – carbamidomethyl; MetOx – methionine oxidation; N-acetyl – N-terminus protein acetylation; N-pyroGlu – N-terminus pyroglutamic acid; Hse – Homoserine; Hsl – Homoserine lactone.

Figure 1.

Schematic of workflow for creation of a custom species-specific proteomics sequence database (Database S1).

Figure 2.

Comparison of protein distributions, based on unique peptides, occurring in otolith and endolymph samples from adult *Acanthopagrus butcheri* (n=3). All samples underwent identical CNBr/Trypsin digestion protocols and were sequenced on a Thermo Scientific NanoLC-Q Exactive Plus Hybrid Quadrupole-Orbitrap Mass Spectrometer. All peptides are Mascot Score ≥13 and E-value ≤ 0.01. Venn diagram created with http://bioinfogp.cnb.csic.es/tools/venny/

Figure 3.

All proteins, based on proteotypical peptides, found in the otoliths of adult black bream (*Acanthopagrus butcheri*, n=3). Previously identified inner ear proteins are highlighted in red and labelled. Heights of bars are indicative of protein abundance in the otolith, based on logarithm+1 transformed average feature intensity. All data were collected on a Thermo Scientific NanoLC-Q Exactive Plus Hybrid Quadrupole-Orbitrap Mass Spectrometer. Feature intensities were calculated using MaxQuant v1.5.8.3. The full list of peptides with associated feature intensities is available as Database S2.

Figure 4.

Heat map of selected proteins occurring in the inner ear of black bream (*Acanthopagrus butcheri*). Relative protein abundances are represented as the logarithm ratio of feature intensities (FI) of peptides in otolith relative to endolymph (n=3). All data were collected on a Thermo Scientific NanoLC-Q Exactive Plus Hybrid Quadrupole-Orbitrap Mass Spectrometer. Feature intensities were calculated using MaxQuant v1.5.8.3. Proteins presented have either been previously identified in the fish inner ear or known to be involved in biomineralization. The full list of peptides with associated feature intensities is available as Database S2.

Figure 5.

Constellation plot illustrating the clustering of inner ear proteins from a range of fish species; *Acanthopagrus butcheri*, *Danio rerio*, *Lates calcarifer*, *Larimichthys crocea*, *Oryzias latipes*, *Oncorhynchus mykiss*, *Oreochromis niloticus*, *Salmo salar*, *Seriola lalandi*, *Stegastes partitus*, *Takifugu rubipes*, as well as *Homo sapiens*. The putative *Starmaker* homolog from *A. butcheri* ("Unknown") clusters closely with *Otolith Matrix Macromolecule-64* (*OMM-64*), the *Starmaker* functional homolog in *O. mykiss*. The distribution of *Dentin sialophosphoprotein* (*DSPP*), *Starmaker* and their homologs is as expected for Intrinsically Disordered Proteins. In contrast, *Transferrin*, *Neuroserpin*, *Otolin-1* (*Otolin*), *Otolith Matrix Protein-1* (*OMP-1*), and *Osteonectin* (*SPARC*) show high levels of conservation among species. The following data were analysed, for each protein, using the protein parameter tool at ExPASy (https://web.expasy.org/protparam/); molecular weight, theoretical isoelectric point, percentage amino acid composition, instability index, aliphatic index and grand average of hydropathy (GRAVY). The plot was constructed, in JMP v 13 (SAS Institute Inc.) using Ward's minimum variance hierarchical clustering method on standardised data. In this approach, the distance between two clusters (represented by the length of the line in the plot) is the ANOVA sum of squares between clusters summed over all variables. Clustering at each subsequent level of the hierarchy is achieved by merging clusters from the previous level, which minimizes the within-cluster sum of squares relative to all possible clusters.

Figure 6.

Schematic of hypothetical ion transport mechanisms in the inner ear of black bream (*Acanthopagrus butcheri*). The sagittal otolith pictured is in the sacculum, a fluid-filled sac in the inner ear. The saccula are located in the cranium – behind the brain and approximately level, vertically, with the eyes. In the epithelial cell, *Carbonic anhydrase 1* (*CA 1*) converts carbon dioxide to bicarbonate ions and protons. The

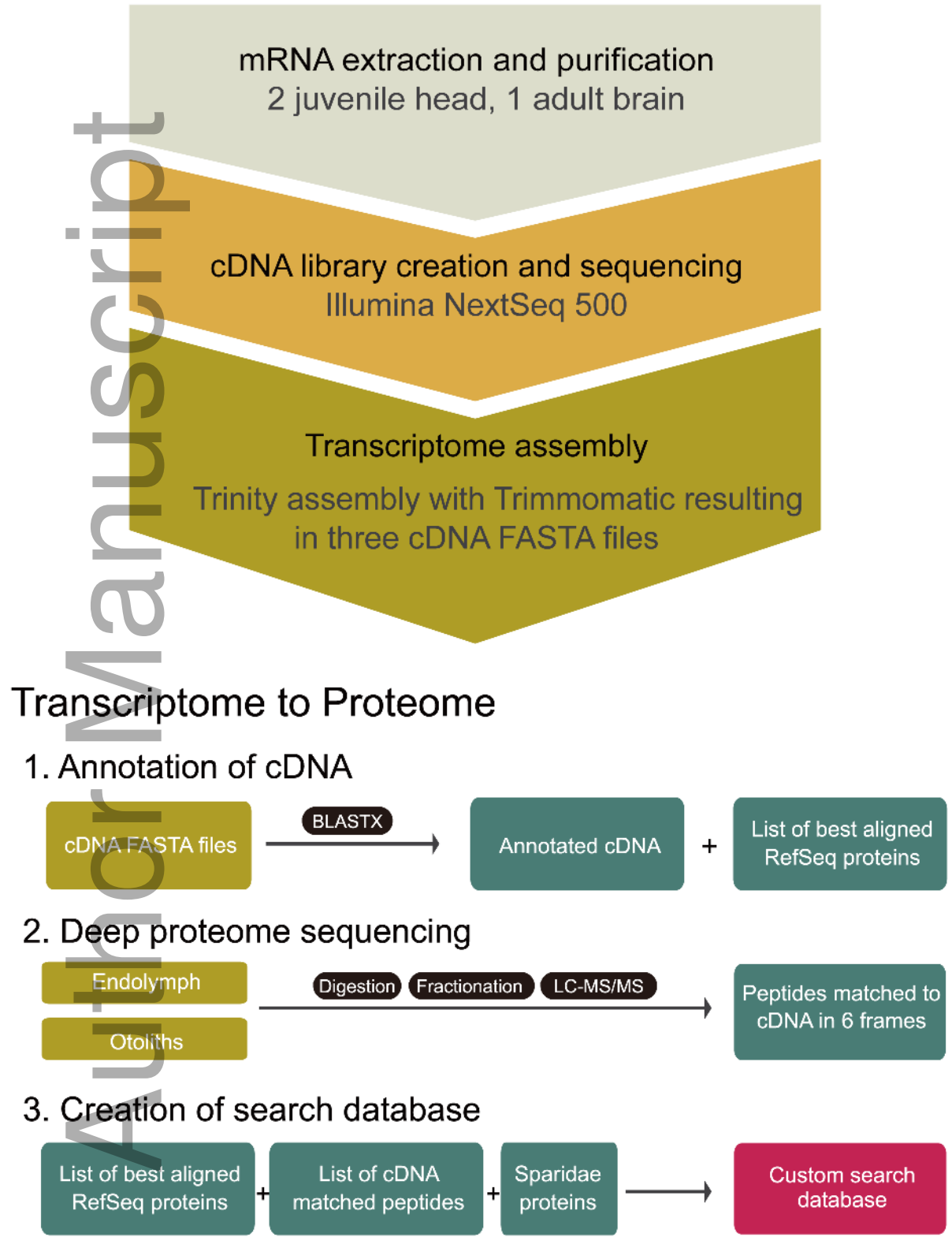
bicarbonate ions are transferred across the cell membrane by *Band 3 anion exchange protein* (*B3AE*), which exchanges them for chloride ions in the endolymph. *Calmodulin* (*CaM*) transfers intracellular calcium to *Na^+/Ca^{2+}-exchanger 2* and *Plasma membrane Ca^{2+}-ATPase* (*PMCA*). *Na^+/Ca^{2+}-exchanger 2* and *PMCA* transfer calcium into the endolymph. At the growing surface of the otolith, bicarbonate ions precipitate with calcium as aragonite needles. The carbon dioxide created as a by-product of this reaction is converted to protons and bicarbonate *in situ* by *Carbonic anhydrase 6* (*CA 6*). The protons are pumped out of the endolymph by *V-type proton ATPase* (*V-ATPase*). The combined action of *V-ATPase* and *CA 6* maintain the necessary pH for aragonite formation while at the same time ensuring that the endolymph is supersaturated with bicarbonate ions. Sodium/potassium ATPase (*Na^+/K^+ ATPase*) generates the electrochemical potential necessary for all active transport processes.

Figure 7.

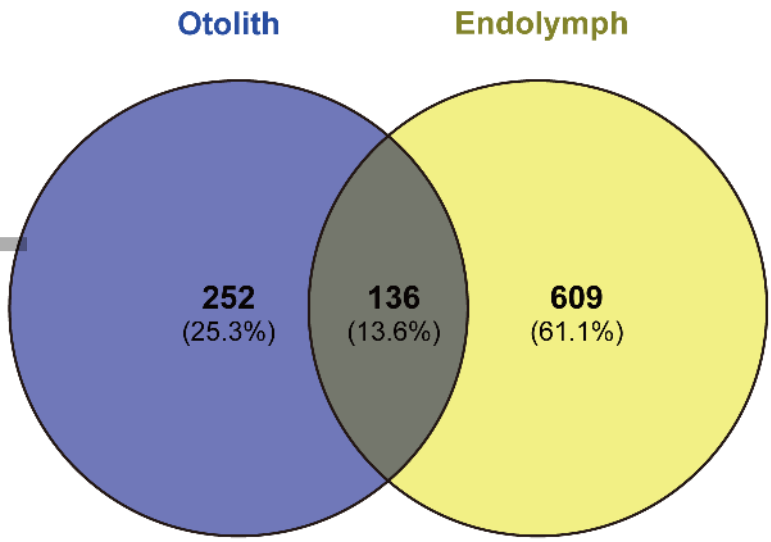
Proposed regulation of diel-entrained organic matrix deposition. It is believed that the organic phase is deposited at dusk and that mineralization occurs at dawn [47, 72-74]. At dusk, Otolin-1 is expressed, forming interaction hubs with other structural proteins [68]. The organic matrix continues to be deposited until dawn, where the action of Matrix metalloproteinase 2 (MMP 2) cleaves Otolin-1 hubs, halting organic deposition, and allowing mineral precipitation to commence. The activation of MMP 2 is through a proteolytic cascade. At dusk, this cascade is inhibited from activating further MMP 2 by Neuroserpin. Any still active MMP 2 are inhibited by the action of Tissue Inhibitor of Metalloproteinase 2 (TIMP 2). MMP2 and TIMP 2 levels have been shown to be affected by the clock gene, PER2 [71].

Figure 8.

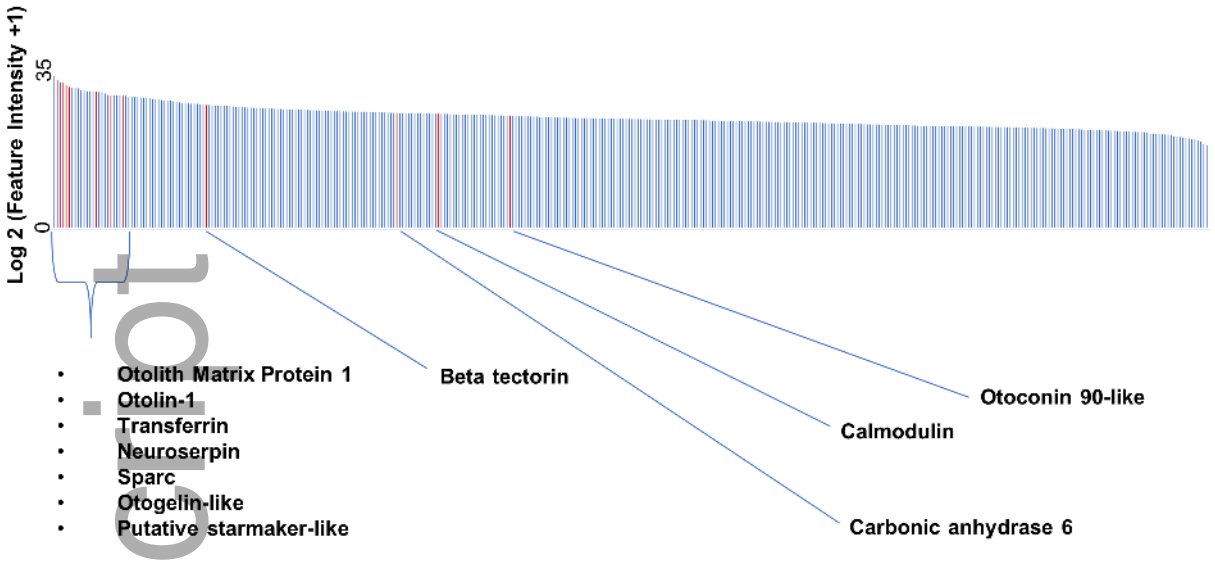
Proteins common to bone and other biomineralization systems found in the inner ear of adult black bream (*Acanthopagrus butcheri*, n=3) [11, 12, 49, 50]. In addition to structural collagens and proteoglycans, we also found proteins that are associated with osteoblast function or otherwise act as substrates for extracellular matrix formation [52, 75, 76, 81].



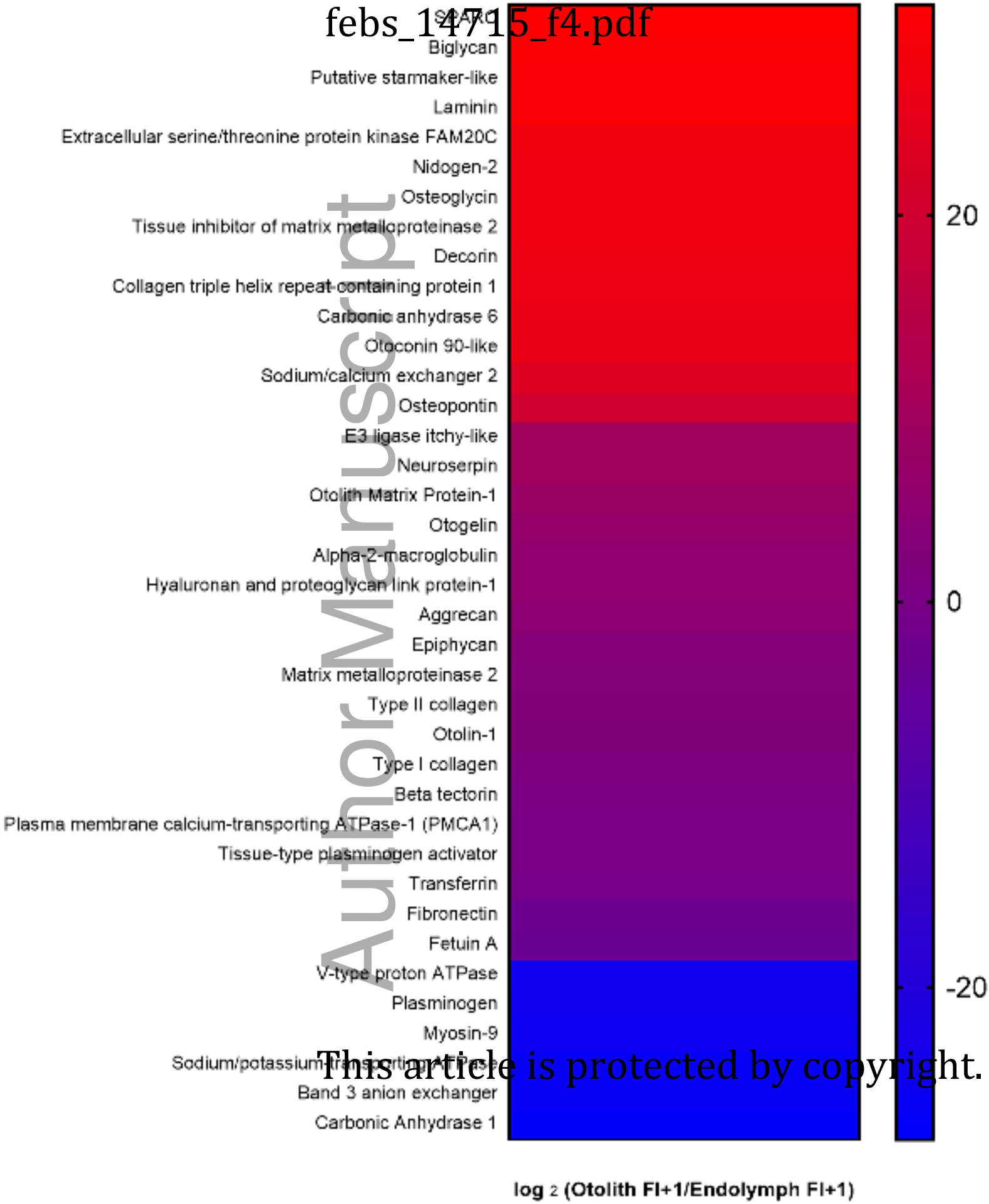
febs_14715_f1.tif

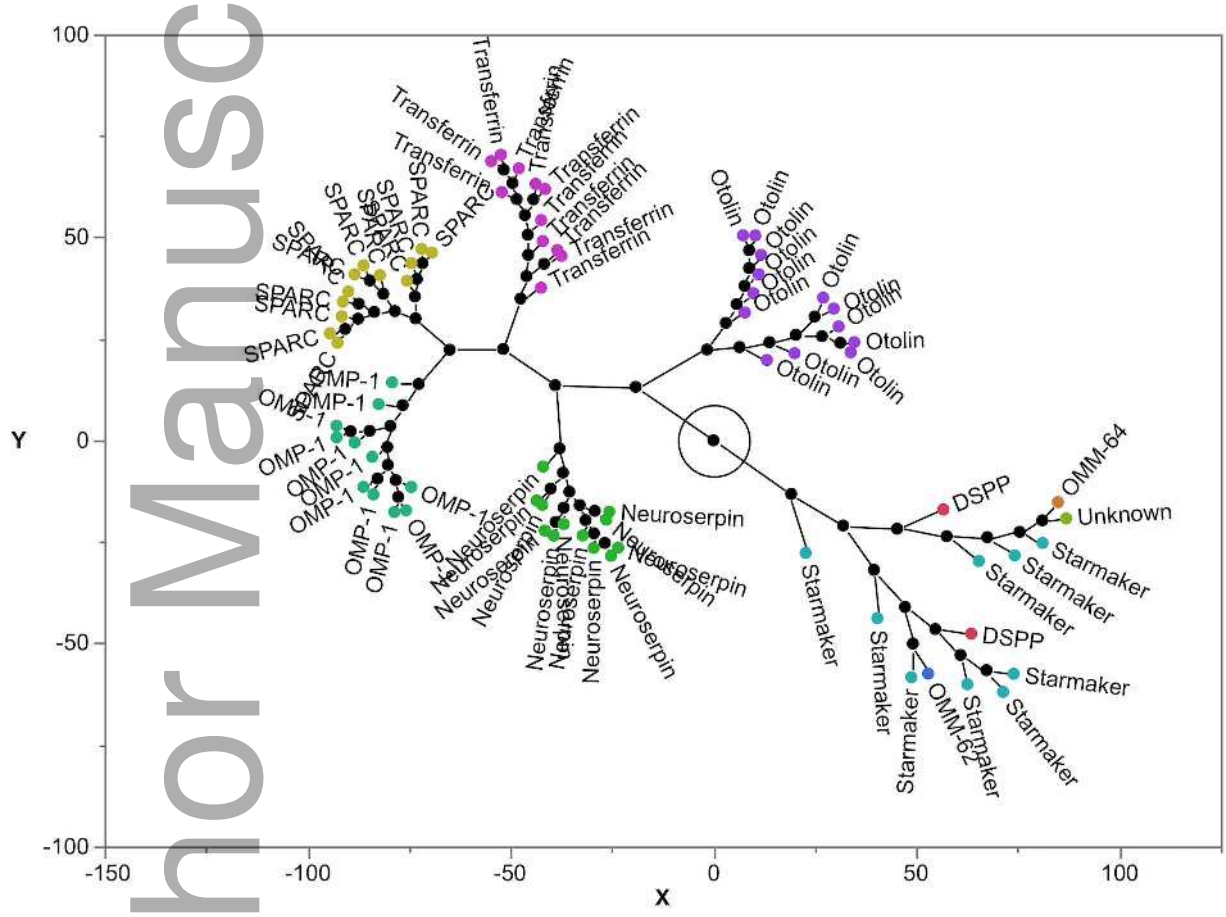


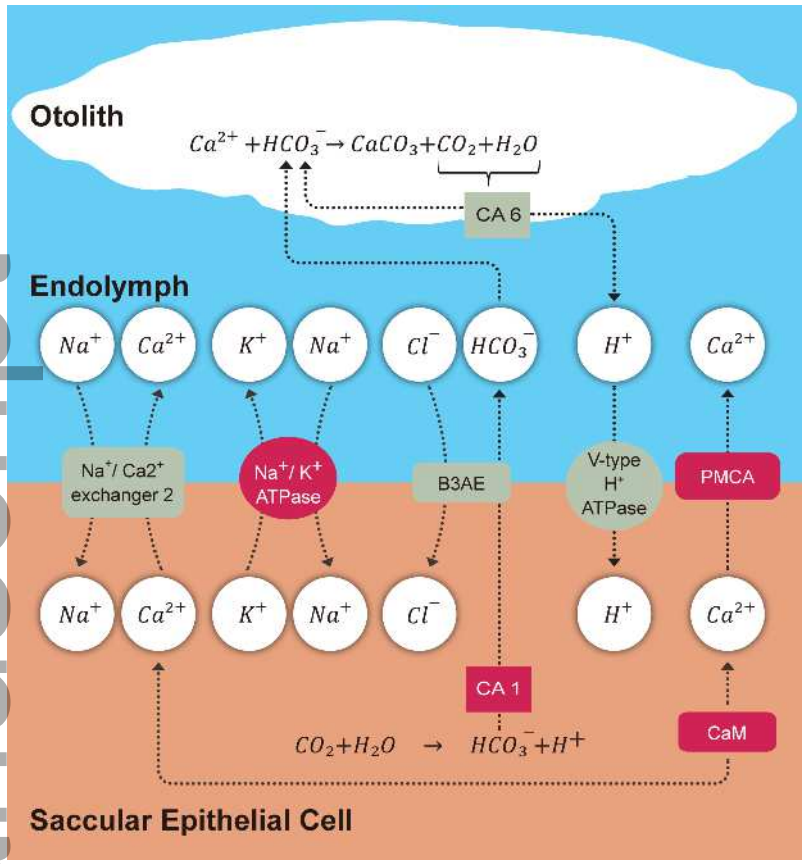
febs_14715_f2.tif



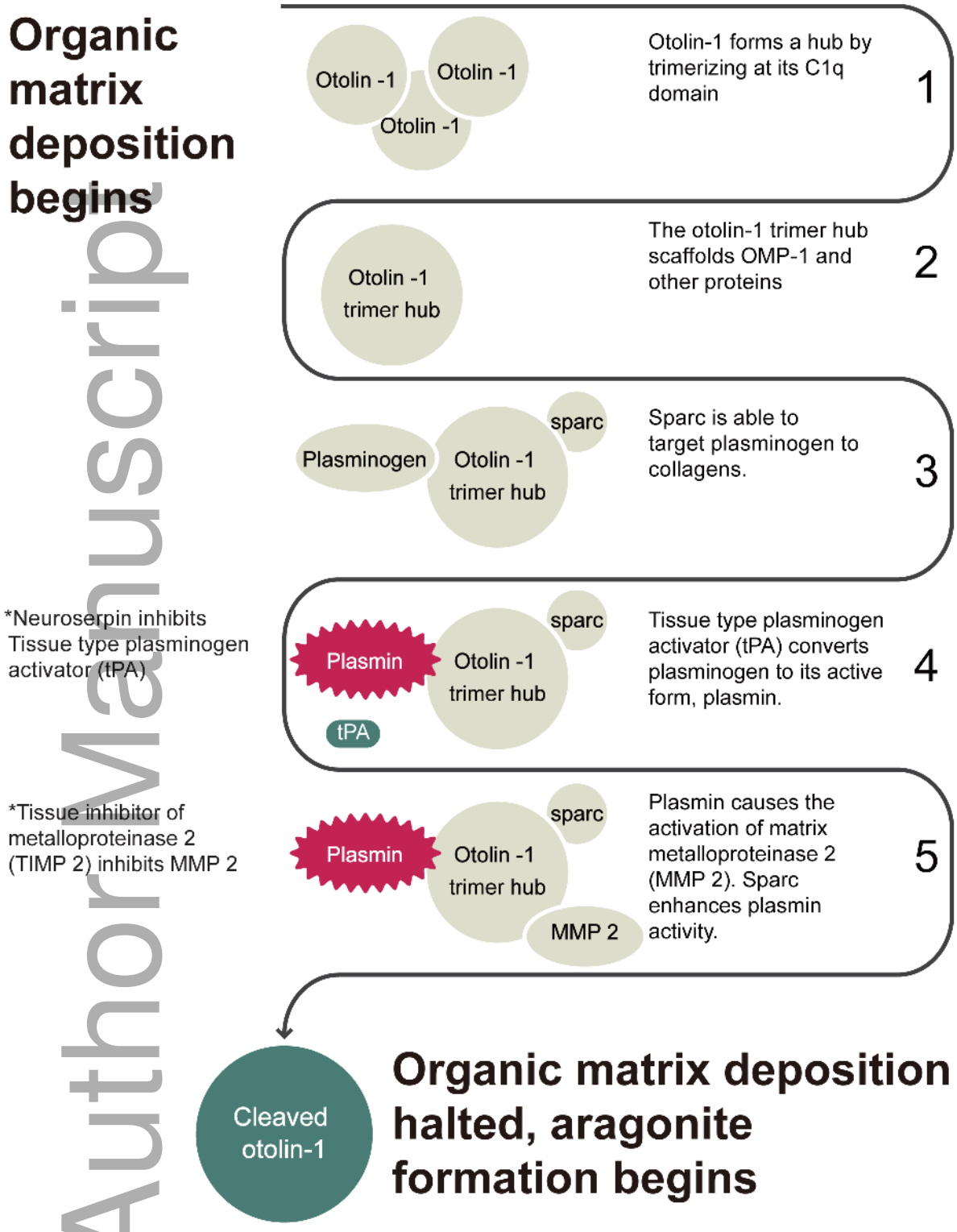
febs_14715_f3.tif



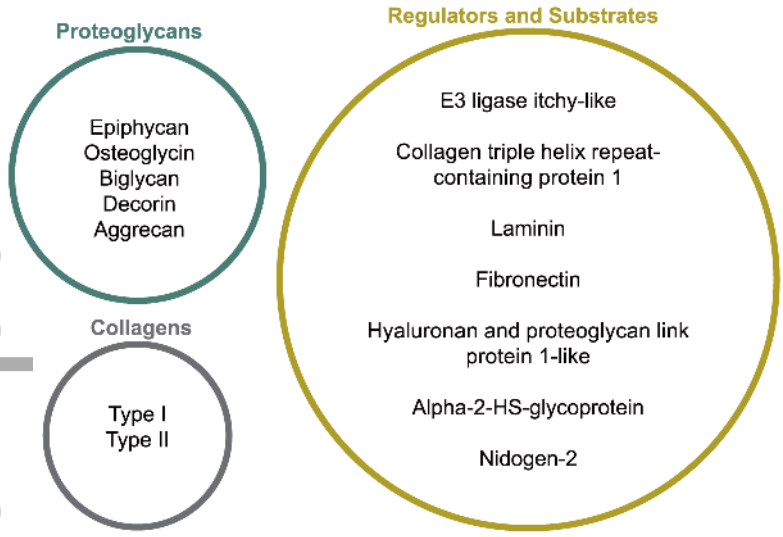




febs_14715_f6.tif



febs_14715_f7.tif



febs_14715_f8.tif