Population-specific renal proteomes of marine and freshwater three-spined sticklebacks

Dietmar Kültz¹, Johnathon Li¹, Darlene Paguio¹, Tuan Pham¹, Marius Eidsaa², and Eivind Almaas²

¹Department of Animal Science, University of California, Davis
Davis, CA 95616, USA

²Department of Biotechnology, NTNU – Norwegian University of Science and Technology, 7491
Trondheim, Norway

Correspondence to:

Dietmar Kültz
Professor of Physiological Genomics
Department of Animal Science
One Shields Avenue, Meyer Hall
UC Davis
Davis, CA 95616
dkueltz@ucdavis.edu

Abbreviations: AC UniProt accession number, BPC base peak chromatogram, CARD caspase activation and recruitment domain, DIA data-independent acquisition, FW freshwater, ID identification, LFQ label-free quantitation, Nt-Ac, PDI protein disulfide isomerase, P-Hox, SW seawater, Rt retention time, TMAO trimethylamine oxide, wTO weighted topological overlap

Funding / Acknowledgements: This study was supported by National Science Foundation grant IOB-13355098 to DK. ME was supported by a strategic research grant from NTNU Faculty of Natural Sciences and Technology.

Abstract

Quantitative proteomics was used to reveal biochemical differences in kidneys of marine and freshwater three-spined sticklebacks. More than 1500 unambiguous proteins were identified, 106 of which are robustly co-translationally modified. Amino-terminal acetylation sites for 94 and proline hydroxylation sites for 12 proteins, including 4 protein disulfide isomerases having the consensus motif **APWCGHCK**, were determined. More than 1500 proteins were quantified by LC-MS/MS yielding 120 proteins with consistent population-specific abundance differences. Twenty-five of these were selected for validation by data-independent acquisition (DIA) and spectral library based MS2 quantitation. A dense biochemical network was revealed, which promotes the synthesis of the organic osmolytes betaine, sorbitol, trimethylamine oxid (TMAO), and urea. It contains 33 of 49 proteins that are elevated in marine compared to freshwater sticklebacks, including the most highly elevated proteins (dimethylaniline monooxygenase, alanine-glyoxylate aminotransferase, glycine N-methyltransferase). Freshwater stickleback kidneys contain elevated levels of proteolytic, cytoskeletal, extracellular matrix, and calcium signaling proteins. Proteins that are most elevated in freshwater sticklebacks are ES1 protein homolog, apoptosis-associated speck-like protein containing a CARD and caspase 1. Protein-abundance network analysis demonstrates significantly higher levels of synchronized abundance control in marine sticklebacks. The significance of these findings for biochemical diversification of renal function in marine and FW sticklebacks is discussed.

Introduction

Three-spined sticklebacks (*Gasterosteus aculeatus*) have captured the interest of biologists for many decades due to their ability to adapt to diverse environments, their interesting behaviors and morphologies, and their great abundance in aquatic ecosystems throughout the Northern hemisphere [1-4]. The stickleback genome has been sequenced and annotated and currently an increasing number of studies are addressing the genetic basis of morphological differences between marine (seawater, SW) and freshwater (FW) stickleback ecotypes [5-7]. However, large-scale studies directed at biochemical differences (e.g. ecotype-specific proteomes and metabolomes) are still very rare. Nevertheless, proteins and biochemical networks represent the link that determines how genetic information and environmental contexts interact to yield specific higher-order phenotypes such as distinct physiologies, morphologies, and behaviors [8, 9]. In addition, proteins represent the molecular building blocks giving rise to phenotypic variability that is acted upon by natural selection. Therefore proteome comparisons of conspecific populations that are genetically isolated and well adapted to different environments (ecotypes) have the potential to provide deep mechanistic insight into biological processes and molecular functions that are under strong selection pressure in particular environments [10].

Population proteomics is not a new field but its application to evolutionary ecology is still in its infancy [11]. In fact, we argue that the August Krogh principle, which is frequently referred to as a major guiding principle in Ecological and Evolutionary Biology ("For many problems there is an animal on which it can be most conveniently studied") [12] does just as well apply to proteins as it does to "problems". For many proteins there surely must be an organism, a tissue, and a specific context in which it is most highly abundant and, therefore, conveniently tractable under physiologically relevant conditions. Proteomics reveals which proteins are tractable with regard to dynamic changes in their abundance, covalent modification, and interaction with other macromolecules in a particular organism, tissue, and biological context. Moreover, proteomics has the potential to uncover entire biochemical pathways or networks that are most conveniently studied in organisms in which they are highly prominent and responsive to ecological and evolutionary challenges. Such protein networks may be greatly altered when ecological or evolutionary contexts change, even if none of the alleles encoding the proteins constituting the network differ in sequence or frequency. The dynamics of such networks also depends on substrate availability and covalent modification of regulatory proteins that control their activity. Both, substrate availability and covalent protein modification, are dynamic properties that are strongly influenced by the environment. In addition, population-specific horizontal gene transfer could give rise to differences in ancillary proteins and or enzyme products (cofactors) that may

profoundly affect such networks [13, 14]. For these reasons biochemical ecology, the study of biochemical pathways and networks in specific ecological contexts, represents an attractive approach for understanding the mechanisms of evolutionary diversification in conspecific populations or closely related congeneric species [15, 16].

Although biochemical studies comparing SW/ anadromous with FW ecotypes of three-spined sticklebacks are much less common than genetic, morphological or behavioral analyses, several previous investigations have documented ecotype-specific biochemical differences. Most previous biochemical studies of three-spined sticklebacks focus on enzyme activities in liver, muscle, and gill [17-22]. The stickleback kidney has been studied mostly with regard to hormonal, parasite, and pollutant effects on ultrastructural and mRNA expression changes that occur during the breeding season [23-27]. However, protein biochemical studies of stickleback kidneys are rare despite the pivotal function of this organ in body water and salt balance [28, 29].

Studies of teleost osmoregulation in other species have shown that their kidneys are responsible for maintaining body fluid and divalent ion homeostasis [30, 31]. They counteract passive movements of water and ions between body compartments (maintained at osmotic homeostasis of about 300 mosmol/kg) and the environment (<40 mosmol/ kg for FW and >1000 mosmol/ kg for SW). Therefore, kidney function greatly differs in fish inhabiting SW versus FW environments and in euryhaline fish acclimated to either FW or seawater (SW). In SW, teleost kidneys secrete a very small amount of urine containing highly concentrated metabolic waste and excess divalent ions. In FW, they secrete large amounts of dilute urine and reabsorb required ions [32, 33]. In addition, teleost kidneys are responsible for the reabsorption of calcium and phosphate, which is essential for supporting bone morphogenesis [34]. Therefore, differences in kidney biochemistry and physiology could potentially provide mechanistic insight into body armor and bone morphology phenotypes that distinguish SW and FW three-spined sticklebacks.

In the current study we have utilized the well-annotated *G. aculeatus* proteome and recent progress in quantitative proteomics to comprehensively characterize the renal proteome of SW and FW sticklebacks. Both sexes were collected from resident SW (Bodega Harbor) and FW (Lake Solano) populations after the breeding season. Unbiased label-free quantitative profiling [35, 36] was combined with hypothesis-driven validation of quantitative differences in key proteins by data-independent acquisition (DIA) and quantitative monitoring of selected MS2 transitions, a method that recently emerged from selected reaction monitoring (SRM)

approaches [37-39]. Using these approaches we uncover distinct biochemical networks and signatures that provide mechanistic insight into renal processes and functions, which are utilized differently in sticklebacks from SW and FW populations.

Materials and Methods

Animal collection and sample preparation: Adult sticklebacks were collected by seining from Bodega Harbor (34 g/kg salinity) and by using unbaited minnow traps from Lake Solano (0.4 g/kg salinity). Three independent experiments were performed with females collected in the summer of 2012 (2012F), females collected in the summer of 2014 (2014F), and males collected in the summer of 2014 (2014M). Each experiment analyzed six samples from each population (36 fish in total for all experiments). Both populations breed in spring and collections were, thus, done after the breeding season. All fish were sacrificed in accordance with UC Davis IACUC protocol 18010, sexed, and kidneys dissected within 2h of collection. Dissected tissue was snap-frozen in liquid nitrogen and stored at -80°C until protein extraction. Protein extraction, protein assay, and in solution trypsin digestion were performed as previously reported [35].

Protein identification and quantitative profiling: Tryptic peptides from each sample (200 ng total) were injected with a nanoAcquity sample manager (Waters, Milford, MA USA), trapped for 1 min at 15 µL/min on a Symmetry trap column (Waters 186003514), and separated on a 1.7µm particle size BEH C18 column (250mm x 75µm, Waters) by reversed phase LC using a nanoAcquity binary solvent manager (Waters). A 125 min linear gradient with 3-35% ACN was used. Peptides were ionized by nano-ESI using a pico-emitter tip (New Objective, Woburn, MA USA) and analyzed by an ImpactHD UHR-QTOF mass spectrometer (Bruker Daltonics, Bremen, Germany). Batch-processing of samples was controlled with Hystar 3.2 (Bruker). Charge state deconvolution and deisotoping was performed and peak lists generated with DataAnalysis 4.2 (Bruker) using default parameters. Peak lists were imported into PEAKS 7.5 (BSI, Waterloo, Canada) and PEAKS, Mascot 2.2 (Matrix Science, London, UK), and X!Tandem Cyclone (<http://www.thegpm.org/tandem/>) search engines were used to identify proteins from MS/MS spectra. The following parameters were used: enzyme specificity = trypsin; missed cleavages permitted = 2; variable modifications = Cys carbamidomethylation, Met oxidation, P-Hox, N-terminal carbamylation, Nt-Ac; precursor ion mass tolerance = 20 ppm (100 ppm for X!Tandem); fragment ion mass tolerance = 0.02 amu. The inclusion of Nt-Ac and P-Hox as variable modifications enabled an assessment of the frequency of these co-translational

modifications in the renal proteome. The probability that these modifications are localized to a particular amino acid within a peptide sequence was assessed with PEAKS 7.5 based on the corresponding MS/MS ions using the Ascore method [40]. Consensus motifs surrounding these modifications were identified with ScaffoldPTM 2.1 (Proteome Software Inc., Portland, OR).

The three-spined stickleback proteome (27,248 sequences) was downloaded on Feb. 07, 2014 from UniProt and annotated using BLAST2GO (BioBam Bioinformatics, Valencia, Spain) by searching against SwissProt with default parameters. Randomly scrambled decoy entries were added for each sequence by PEAKS 7.5 and this database containing 54,496 total entries was used for all searches to determine protein FDR. Search results from all three engines were consolidated in Scaffold 4.4 (Proteome Software Inc., Portland, OR) with peptide FDR < 0.1%, protein FDR < 1.0%, and peptides per protein => 2. Proteins sharing significant peptide evidence that could not be differentiated based on MS/MS analyses alone were grouped into clusters to satisfy the principles of parsimony using the Protein prophet algorithm in Scaffold [41]. Label-free quantitative profiling of peptide intensities and calculation of relative protein abundances in each sample was performed with PEAKS 7.0 (BSI). The following parameters were used: FDR <1%, the peptide needed to be identified in at least 4 samples, mass error tolerance < 30 ppm, retention time shift tolerance < 3 min, -log₁₀ significance p value => 20 for females and 7.5 for males, protein fold change => 2.

Targeted quantitative proteomics: Population-specific quantitative differences in protein abundances revealed with PEAKS profiling were independently validated for multiple key proteins by data-independent acquisition (DIA) and targeted proteomics with Skyline [42]. The same sample preparation and chromatography conditions as above were applied but the mass spectrometer was operated in MS/MS only (DIA) mode. The mass range was set to 490 – 910 m/z at 25 Hz scan rate with an isolation width of 11 m/z (1 m/z overlap). Spectral libraries were created from the precursor-dependent LC-MS/MS data for differentially abundant proteins identified by the initial PEAKS profiling. Only fully matching peptides with no sequence variation were included for quantitative proteomics to avoid confounding effects of potentially variable allele frequencies on the calculation of protein quantity. Spectral libraries were constructed with Skyline 3.1 [42] and used for identification of peptides that are associated with targeted proteins. Population-specific protein differences were then validated by targeted quantitation of 5 MS/MS transitions for 3 peptides of each protein. A small number of targeted proteins were represented by fewer than 15 (but always at least 5) suitable transitions. Quantitative data were statistically analyzed using Skyline 3.1 to independently validate significant population-specific abundance

differences. The Skyline dataset and corresponding spectral library were submitted to the public Panorama server [43] and can be accessed at <https://panoramaweb.org/labkey/Kultz-et-al-2015.url>.

Statistics and biochemical network construction: PEAKS statistical analysis (BSI, Waterloo, Canada) was used to identify renal proteins that are differently abundant in the SW and FW populations during the first-tier profiling step. This approach is based on quantifying MS1 isotopic chromatogram peak areas of the three most intense and unique peptide ions for each protein. Statistically significant protein differences are then expressed as the cumulative $-\lg_{10}$ t-test p value for all three peptides. Because no correction for multiple testing is applied the cutoff was set very high (>7.5). The entire comparison was replicated three times with independent samples from females collected in summer 2012, males collected in summer 2014, and females collected in summer 2014. Protein abundances calculated in PEAKS for each sample were exported to Excel for all proteins exceeding the PEAKS significance threshold in any of the three experiments. Based on these raw data, population-specific abundance ratios and t-test p values were calculated after testing for homoscedasticity using the F-test. If the t-test P value was <0.05 for at least two of the three experiments then the corresponding renal proteins were considered significantly different between populations. No correction for multiple testing was applied since the number of proteins quantified was approximately 1600 and the FDR was only 0.25% based on the condition that the P values had to be <0.05 in at least two independent experiments ($0.05 \times 0.05 = 0.0025$). Thus, less than 4 proteins (1600×0.0025) are expected random false positives. Targeted quantitation of MS/MS transitions was performed to validate abundance differences of key proteins identified during the first-tier profiling quantitation. Skyline t-tests comparing quantities of MS/MS transitions for Bodega Harbor with those for Lake Solano samples were performed for each protein and corrected for multiple testing as previously described [44]. A biochemical network connecting many of the proteins that are significantly more abundant in kidneys of SW sticklebacks was manually constructed using CellDesigner 4.0 [45] based on enzyme functional information contained in the UniProt [46], BRENDA [47] and Genecards [48] databases.

Protein abundance network construction and weighted network core analysis: For protein abundance measurements of each population (Bodega Harbor and Lake Solano) separately, Spearman correlations $corr(i,j)$ between all pairs of proteins i and j were calculated. The protein adjacency matrix $A=[a_{ij}]$ was defined using the adjacency function $a_{ij}=(corr(i,j))^b$ [49], where the power exponent $b=5$ was chosen in order to maintain the sign of the correlations. Next, the

weighted topological overlap (wTO) matrix [49-51] was calculated using the sign-sensitive version [52]

$$\omega_{ij} = \frac{\sum_u a_{iu} a_{uj} + a_{ij}}{\min(s_i, s_j) + 1 - |a_{ij}|}, \quad (1)$$

where the node strength (weighted connectivity) is defined as $s_i = \sum_j |a_{ij}|$. In order to assess the cut-off level for statistically significant wTO links, ω_{ij}^c , the sequence of protein abundance measurements was randomized for each protein individually before a (random) wTO network was generated using the above described construction procedure. Based on the generation of 100 separate random versions of wTO from the Bodega Harbor and Lake Solano data, the cut-off level was determined as $\omega_{ij}^c = 0.15$.

In order to determine nodes central to the generated wTO networks, a modified version of the s-core network decomposition method [53] was introduced. Briefly, the s-core network decomposition is a method for peeling a network, where all nodes with strengths smaller or equal to a threshold value $s_i \leq s_T$ are iteratively removed. Note that all of the node strengths s_i are updated whenever a node is removed from the network. Thus, all of the nodes in the s_T -core possess strengths exceeding s_T . In the systematic s-core peeling process, zero is chosen as the first threshold value, and all subsequent threshold choices are determined by $s_T = \min_i (s_i)$. Instead of referring to s-cores based on their threshold value, we will use their sequence index: the whole network is the s_0 -core, the s_1 -core consists of all nodes with at least 1 nearest neighbor, the s_2 -core consists of all nodes left after setting the threshold value to the weakest among the connected nodes, etc. The innermost core is identified by the last determined threshold value before all nodes are removed. In correlation networks where the strength of the individual links $\omega_{ij} \in [-1, 1]$, the s-core method displays a tendency to identify innermost cores consisting of ~20% or more of the nodes in the full network (unpublished data, Eidsaa & Almaas, 2015). Thus, we introduce a version of s-core, the s-core+ method, which is able to sequentially identify the most central nodes also for these situations by perturbing the innermost s-core:

1. Conduct standard s-core analysis until the innermost core is identified: this is the s_n -core with threshold value s_T .
2. Remove the weakest magnitude link in the remaining network $\omega_{kl} = \min_{ij}(|\omega_{ij}|)$, connecting nodes k and l .
3. If the removal causes either $s_k \leq s_T$ or $s_l \leq s_T$, set a temporary threshold $s_{Tx} = \min(s_k, s_l)$ and proceed to 4. If not, repeat from 2.
4. If a new s-core can be identified from the current network, set the threshold value for the s_{n+1} -core to $s_T = s_{Tx}$ and repeat from 1. If not, repeat from 2.

Results

Identification of the stickleback renal proteome: Three independent comparisons of SW (Bodega Harbor) and FW (Lake Solano) sticklebacks were performed to robustly identify differences in renal proteomes that are population-specific and evident in different cohorts of fish and different years (2012 and 2014), as well as in both sexes. In experiment 2012F individual samples yielded 6224 ± 152 (mean $\pm$ SEM) unique peptides that were matched to 1126 ± 31 proteins per sample (corresponding to 779 ± 17 unambiguous protein groups) with at least 2 unique peptides per protein. In addition, 1534 ± 52 high-quality de novo sequenced spectra that did not match any protein in the database were identified per sample. Pooling all 12 samples in Scaffold yielded 1635 distinct protein clusters (non-redundant IDs). The dynamic range of detected proteins spans 5 orders of magnitude based on weighted spectral counts reported by Scaffold (Figure 1a). Experiment 2014F identified 7320 ± 475 unique peptides matched to 1239 ± 81 proteins per sample with at least 2 unique peptides per protein (846 ± 56 unambiguous protein groups). In addition, 1641 ± 115 high quality peptides that did not match any protein in the database were derived by de novo sequencing. Pooling all 12 samples in Scaffold yielded 1725 non-redundant protein clusters. Experiment 2014M identified 6503 ± 203 unique peptides matched to 1118 ± 37 proteins per sample with at least 2 unique peptides per protein (768 ± 26 unambiguous protein groups), and 1907 ± 65 unmatched de novo sequenced spectra. Pooling all 12 samples in Scaffold yielded 1575 non-redundant protein clusters.

The Molecular Tissue-specific Phenotypes (MoTiPhe) of renal proteomes in females was very similar in fish collected in 2012 and 2014 (Figure 1). The same set and a comparable number of proteins (and corresponding peptides) were identified (Supplementary tables 1 and 2).

Moreover, the dynamic range distribution of proteins is nearly identical in kidneys from females collected in 2012 and 2014. Of the 25 most highly abundant proteins 19 are shared between these two experiments. The other 6 proteins (triose phosphate isomerase, eosinophil peroxidase, alanine-glyoxylate aminotransferase 2, desmoykin, alpha-enolase, and protein disulfide isomerase A3) vary only slightly in abundance such that they are not consistently included in the top 25. Collagen alpha 1 type I is represented by two protein species in all experiments (see below) but, for unknown reasons, the 2 protein species have been clustered together by the Scaffold software for the 2012F experiment but not for the other 2 experiments. Experiment 2014M conducted on males yielded a nearly identical dynamic range distribution and the same share of the top 25 proteins to overall renal protein content as compared to the two experiments conducted on females (21%, Figure 1). However, notable differences between the top 25 abundant proteins in male versus female kidneys were also evident. Two spiggin 1 protein species (Q1XHH5 and Q2WFH7) and several molecular chaperones necessary for de novo protein synthesis (heat shock 70kDa protein isoform 5, heat shock 90kDa protein beta 1, prolyl 4 hydroxylase subunit beta, protein disulfide isomerase A4, transitional endoplasmic reticulum ATPase, and also elongation factor 1 alpha) were among the top 25 highly abundant proteins in males but not females. Conversely, cadherin 17 and vitellogenin 2 were highly abundant in females but not in males (Figure 1).

Identification of co-translationally modified renal proteins: Two common co-translational modifications that occur during protein synthesis are Nt-Ac and P-Hox. Including them as variable modifications in the database search criteria allowed an assessment of their frequency in the renal proteome of sticklebacks. Ninety-four amino-terminally acetylated (Nt-Ac) proteins were detected in 4 or more of the 12 samples of at least one experiment. The site assignment confidence of this modification was consistently maximal as evidenced by the highest possible localization Ascore [40] of 1000 in all instances and samples when Nt-Ac was detected. Although quantitation of Nt-Ac peptides taking into account the stoichiometry of the modification exceeds the scope of this paper, population-specific differences in the number of Nt-Ac peptides suggest divergence in co-translational processing of the corresponding proteins (Table 1). Such differences were only considered robust if they were observed in all experiments and the difference was ≥ 2 -fold between SW and FW samples in each experiment. Even though these criteria are likely overly conservative they were met by 4 proteins. Ubiquitin-conjugating enzyme E2 variant 1, 40S ribosomal protein S15, 14-3-3 protein beta/alpha, and tumor protein D54 were identified as much less amino-terminally acetylated in Bodega Harbor than in Lake Solano fish

(Table 1). The predominant Nt-Ac amino acid in renal proteins is Ala (36 proteins), followed by Met (30 proteins), Ser (19 proteins), T (7 proteins), E (1 protein), and C (1 protein) (Table 1).

Three major collagens (see Figure 1) are present in stickleback kidneys and all of them are heavily hydroxylated on many proline residues. P-Hox was also evident for several minor renal collagens (Supplementary table 3). Although stoichiometric quantitation of these P-Hox sites exceeds the scope of the current paper, large differences in P-Hox were detected for P426 in collagen alpha-1 type 1 (G3NNI9/ G3NNJ9) and for P639 and P699 in collagen alpha-2 type 1 (G3NTK4). In all three experiments and for all three of these proline residues, P-Hox was much more prominent in Bodega Harbor fish than in Lake Solano fish (Supplementary table 3). The most common motif surrounding hydroxylated prolines in stickleback renal collagens, which applies to 466 identified sites, was determined using ScaffoldPTM 2.1 (Proteome Systems, Portland, OR) based on a previously published algorithm [54]. This motif shows regular spacing of glycine residues separated by two other amino acids, the most common of which are proline and alanine (Figure 2). We also identified P-Hox sites of four protein disulfide isomerases (PDIs). Protein disulfide-isomerase A3 (G3NVG1) was hydroxylated on P395. Prolyl 4-hydroxylase subunit beta (G3Q2E1) was hydroxylated on P393. Thioredoxin domain-containing protein 5 (PDI ERp46, G3NCG1) was hydroxylated on P180. Protein disulfide-isomerase A4 (G3ND88) was hydroxylated on P478. In all cases, the modification was only detected in experiment 2014M, perhaps because PDIs, which are necessary for processing newly synthesized proteins, are present at higher levels in male than in female kidneys (see above, Figure 1). In addition, P-Hox of all four PDIs was more apparent in SW than in FW fish (Table 2).

Quantitative profiles of population-specific renal proteomes: The Top3 approach inherent in PEAKS and based on quantitation of MS1 isotopic chromatogram peak areas of the three most highly abundant and unique peptides per protein was used for label-free quantitation [55-57]. We quantified 1618, 1691, and 1592 (experiments 1 – 3, respectively) of the identified renal proteins in all samples. Using a significance threshold (PEAKS $-\log_{10}$ P value) of 20 for females and 7.5 for males we identified 94 (2012 females), 42 (2014 females), and 75 (2014 males) proteins that were at least twofold more abundant in Lake Solano than Bodega Harbor fish. We also identified 49 (2012 females), 68 (2014 females), and 32 (2014 males) proteins that were at least twofold less abundant in Lake Solano than Bodega Harbor fish (Figure 3). The reason for lowering the significance threshold for males was the overall lesser degree of proteome abundance diversity in males compared to females (Figure 3a-c). Thus, lowering the

PEAKS significance threshold for males was necessary to be able to compare the proteins that are most different between populations in males and females.

Renal proteins that are significantly different in abundance between FW Lake Solano and SW Bodega Harbor fish in multiple experiments clearly cluster into discrete functional entities (Table 3). Twenty proteins involved in amino acid and glyoxylate metabolism are consistently and significantly less abundant in kidneys from FW fish compared to those from SW sticklebacks. This is also the case for 8 proteins associated with the citric acid (TCA) cycle, 1 protein catalyzing trimethylamine oxid (TMAO) synthesis, and 2 proteins of the urea cycle (with the exception of arginase 1 in experiment 2014M). In addition, three enzymes that promote gluconeogenesis and sorbitol synthesis are consistently more abundant in kidneys from SW sticklebacks. Conversely, four glycolytic enzymes show the opposite trend, indicating that gluconeogenesis is favored over glycolysis in SW fish kidneys relative to FW fish kidneys. Interestingly, the use of glyceraldehyde-3-phosphate dehydrogenase isozymes also differs significantly in SW vs. FW populations. Five proteins with antioxidant/ xenobiotic functions are always less abundant while two proteins in this functional category (including thioredoxin) are consistently more abundant in kidneys from FW sticklebacks (Table 3). The majority of proteins that are consistently and (in most cases) significantly more abundant in kidneys from FW fish cluster into four categories: cytoskeleton structure (16 proteins), energy metabolism (2 creatine kinase protein species), proteolysis (21 proteins), calcium signaling (8 proteins), pattern recognition (2 proteins), and mitochondrial oxidative phosphorylation (4 proteins) (Table 3).

The remaining functional categories are more heterogeneous with regard to the population-specific protein abundance differences between FW and SW sticklebacks. The Na⁺/K⁺-ATPase beta-233 subunit (G3PKB5) is consistently significantly more abundant in FW fish kidneys while sideroflexin (a putative mitochondrial iron transport protein) is always significantly lower in FW fish kidneys. Another mitochondrial protein, ES1 protein homolog, is much more abundant in kidneys of FW sticklebacks than in those from SW fish but other proteins related to lipid or fatty acid metabolism are more abundant in SW fish kidneys. Likewise, lipopolysaccharide-binding protein and VAT-1 protein homolog, which are both involved in membrane trafficking, are more abundant in kidneys of FW fish while another membrane trafficking protein (NipSnap homolog 3A) is more highly abundant in SW fish kidneys. Eight signal transduction proteins are significantly more abundant in FW fish kidneys while 4 are consistently more abundant in SW fish kidneys (Table 3). Two of these signal transduction proteins are adenylate kinases (AKs),

one of which (AK1) is significantly higher and the other (AK4) significantly lower in kidneys from FW fish compared to those from SW fish.

Targeted validation of population-specific renal proteome differences: Tissues collected for experiments 2014F and 2014M were processed and analyzed by data-independent acquisition (DIA), an approach allowing targeted proteomics with Skyline. DIA resembles SRM in that only MS/MS spectra (and no precursor MS1 spectra) are acquired by the mass spectrometer [58]. However, MS/MS spectral acquisition takes place in regular time intervals (every 1.7s in our study) enabling highly accurate quantitation of peptides based on multiple independent measures per peptide. Transitions (MS/MS chromatographic peak areas) of the five most abundant y+ or b+ ions of the corresponding MS/MS library spectrum are being monitored simultaneously. This approach is illustrated for glyoxylate reductase (G3P0Y2) peptide AAALDVTHPEPLPR ($m/z = 496.2718$, $z = 3+$) in figure 4. Note that in this figure the scales of the y axes in panels D and F (Lake Solano samples) are much lower than the Y axes scales in panels C and E (Bodega Harbor samples). Using this fully independent targeted approach, the approximately 4-fold higher levels of glyoxylate reductase (G3P0Y2) in kidneys of Bodega Harbor fish compared to those of Lake Solano fish were confirmed.

In addition to glyoxylate reductase, twenty-four other proteins that represent the population-specific functional categories identified during the initial quantitative profiling step were selected for targeted proteomics. Their abundance in kidneys from fish of the two populations was analyzed using Skyline in the same manner as outlined above for glyoxylate reductase. This approach showed that all 25 proteins are significantly different in abundance between populations, which confirmed the results obtained in the initial profiling step (Figure 5). The only protein that does not differ in abundance in males of the two populations was histidine ammonia lyase (G3NKY0). This protein was only detected and quantified during the initial profiling step in females but not in males (Table 3).

Renal proteins with the highest SW/ FW abundance ratios are dimethylaniline monooxygenase [N-oxide-forming] (G3PP04, 42.6 in females and 49.1 in males), alanine-glyoxylate aminotransferase 2 (G3PVW6, 44.1 in females and 4.5 in males) and glycine N-methyltransferase (G3NQA8, 6.5 in females and 31.3 in males), which is consistent with the results obtained for the initial quantitative profiling. Proteins at the opposite end of the spectrum having the highest FW/ SW abundance ratios during the initial quantitative profiling and second-tier targeted quantitation steps were also identical (Figure 5, Table 3). These proteins are ES1 protein homolog (G3Q2N0, in 7.5 females and 16.9 in males), apoptosis-associated speck-like

protein containing a CARD (G3PEF4, 7.1 in females and 13.5 in males) and caspase 1 (G3P6H8, in 5.5 females and 15.0 in males).

After confirming the population-specific abundance differences of all 25 targeted proteins we constructed a biochemical network consisting of 33 enzymes that are consistently higher in SW stickleback kidneys relative to kidneys from FW fish (Figure 6). In addition, 5 enzymes that are systematically lower in kidneys from SW sticklebacks but support the same metabolite fluxes through the network are also included. Enzymatic reactions driving this network support metabolite fluxes that promote the synthesis of the organic osmolytes betaine, sorbitol, TMAO, and urea (Figure 6). Key nodes in this biochemical network are represented by glyoxylate and glycine and its main modules consist of amino acid metabolism, the TCA cycle, gluconeogenesis, TMAO metabolism, and the urea cycle.

Protein abundance network analysis: Based on the abundance data collected from the SW and FW stickleback kidneys, two separate networks were generated using the weighted topological overlap (wTO) approach [49-52]. The wTO network construction procedure will induce stronger links between pairs of proteins that not only are significantly correlated, but also share many of their correlated partner nodes. As a consequence, the resulting network contains clusters and network modules with clearer biological significance than those simply generated from a correlation matrix [49, 59]. When introducing a link strength cut-off to only retain statistically significant connections, the giant component of the FW network consists of 1,126 proteins and 17,429 connections. In contrast, the SW network consists of 1,379 proteins and 157,960 connections. Thus, whereas the average protein in the FW network has 31.0 connections, the average SW protein has 229.1 connections. The difference in number of connections between the two networks is striking, demonstrating that the proteins in SW stickleback kidneys are strongly correlated.

Since both of the networks are highly connected, we employ the *s*-core+ network-peeling algorithm to identify sets of proteins that are most central in the networks (Figure 7). Here, a protein with large (small) BH/LS ratio is depicted in blue (green), and positive (negative) wTO values for the links are red (blue). Figure 7a shows the set of nodes resulting from *s*-core+ index 700 for FW stickleback kidneys. This network consists of 230 nodes with an average connectivity of 64.2 and is divided into two clearly separated modules. The smallest module contains 29 nodes, where 26 of these are associated with 40S or 60S ribosomal proteins. The bigger module contains a large number of metabolic enzymes involved in fatty acid metabolism and proteostasis. Figure 7b shows the set of 223 nodes from the *s*-core+ index 695 for SW

stickleback kidneys. In contrast to the FW core, this set consists of a single module with average connectivity of 218.0, i.e. a protein is directly connected to almost all of the others. While 91 of the proteins are shared between the FW and SW cores, there is only a single protein shared among the ten most central ones (G3PLE8 - ATP synthase subunit beta). When filtering for LS/BH abundance ratio > 2 before conducting the s-core+ selection, we find that the FW central cores are considerably reduced in both size and density (panel C). The network then consists of three sparsely connected and separated modules with varied protein functions. In contrast, when filtering the SW network for BH/LS abundance ratio > 2 before s-core+ analysis, we find 52 proteins with an average connection of 50.7 (Figure 7d).

Discussion

The present study represents the first to analyze the kidney proteome of *G. aculeatus*. In other fish species, proteomics has been used to study kidney function for at least a decade. We have previously applied this approach to analyze effects of salinity stress on the renal proteome of dogfish shark (*Squalus acanthias*) [60]. Bony fish species for which the renal proteome has been analyzed include zebrafish (*Brachidanio rerio*) [61, 62], Atlantic salmon [63-65], smallmouth bass (*Micropterus dolomieu*) [66], channel catfish (*Ictalurus punctatus*) [67, 68], gilthead sea bream (*Sparus aurata*) [69], and puffer fish (*Takifugu rubripes*) [70]. Most of these previous studies on teleosts focus on the effects of pathogens or toxins on the proteome of the head kidney. In the current study we did not separate head and trunk kidneys but analyzed the whole kidney. In teleost fishes the trunk kidney is much bigger than the head kidney and our proteome results are therefore mostly representative of trunk kidneys, which are responsible for renal regulation of water and electrolyte homeostasis [71].

The trunk kidney proteome has been analyzed in ayu (*Plecoglossus altivelis*) [72]. Ayu are euryhaline fish and the authors investigated effects of a salinity increase (transfer from FW to brackish water) on the kidney proteome. They separated renal proteins on 2D gels and prepared differentially abundant spots for MALDI-TOF/TOF mass spectrometry to identify the corresponding proteins. A much smaller number of proteins has been identified in the ayu study and none of them are identical to the population-specific renal proteins identified here in kidneys from three-spined sticklebacks. Nevertheless, the regulation of several renal proteins in ayu exposed to increased salinity (decrease of glycolytic enzymes and enzymes involved in oxidative phosphorylation, increase of S-adenosylhomocysteinase as salinity increases) is consistent with the network presented in figure 6. Thus, it is possible that organic osmolyte

synthesis pathways identified for sticklebacks are regulated by environmental salinity and also utilized in other species.

Amino-terminal acetylation of renal proteins in sticklebacks: Both Nt-Ac and P-Hox are cotranslational modifications that occur when nascent polypeptide chains are folded into functional proteins in the endoplasmic reticulum [73, 74]. Almost all renal stickleback proteins identified to be amino-terminally acetylated fit the consensus motif for amino-terminal acetyl transferase A (NatA) substrates [75]. The only exceptions are beta-actin (G3N8X9/ G3N8W8), which is acetylated on glutamic acid, and leucyl aminopeptidase, which is acetylated on cysteine (Table 1). This finding suggests that NatA is the predominant Nat present in stickleback kidneys.

Four proteins were much more prominently acetylated in kidneys from FW than in those from SW fish (G3PSC5 40S ribosomal protein S15, G3NHX0 14-3-3 protein beta/ alpha, G3PDF5 ubiquitin-conjugating enzyme Es variant 1, and G3P7U2 tumor protein D54). All of these proteins have functions in cellular proteostasis. This observation is intriguing since proteolysis is one of the main physiological processes that is strongly enriched in kidneys from FW relative to those from SW sticklebacks (see below). It is therefore possible that population-specific differences in proteolytic enzymes are accompanied by corresponding differences in Nt-Ac of regulatory proteins to alter proteostasis networks.

Population-specific Nt-Ac of the 4 proteins mentioned above could also contribute to the observed population-specific differences in cytoskeleton structure and membrane trafficking, which are processes that can be modulated by Nt-Ac of certain regulatory proteins [76-79]. Although Nt-Ac has been mostly associated with protein stabilization, it can also destabilize some proteins, have no effect on protein stability, or have species-specific effects on protein stability [80]. For instance, Nt-Ac alters the stability of teleost (pike) but not rat parvalbumin alpha [81]. In the current study we have identified parvalbumin beta as an amino-terminally acetylated protein. Parvalbumin beta Nt-Ac peptides were identified in exactly the same number of Lake Solano and Bodega Harbor samples. However, we also found that two parvalbumin protein variants are much more abundant in kidneys of FW sticklebacks than in kidneys of SW sticklebacks. These findings, taken together, do not support (but neither conclusively exclude) a role of Nt-Ac in controlling the stability of the calcium signaling protein parvalbumin in stickleback kidneys. In addition to its signaling functions, parvalbumin is also critical for renal handling of divalent cations (calcium and magnesium) [82]. Future studies focusing on this

cotranslational modification are needed to shed light on its regulation and functional significance in SW and FW sticklebacks.

Proline hydroxylation of renal collagens and protein disulfide isomerases: P-Hox of 12 proteins (8 collagens and 4 PDIs) is documented in this study. Three collagens (two variants of collagen alpha 1 type 1 and collagen alpha 2 type 1) are the predominant proteins present in stickleback kidneys. These highly abundant basement membrane proteins are heavily hydroxylated on proline, which is consistent with previous findings on other animal epithelia, cartilage, bone, and other tissues. In fact, collagen is the most highly abundant protein of most animals and its hydroxyproline residues represent a major source of essential amino acids for carnivorous animals [83]. We have identified the P-Hox pattern of these collagens and derived a consensus motif that fits the Gly-Xaa-Yaa motif common to animal collagens [84]. Moreover, the stickleback consensus motif shows that the amino acid triplet containing hydroxylated proline is flanked by at least one additional Gly-Xaa-Yaa motif on each side. Amino acids that are intercalated between glycine are most frequently proline, alanine, and arginine.

Only a fraction of the prolines present in stickleback collagens is hydroxylated, which is consistent with collagens of other species and indicates that a particular spacing between hydroxylated residues needs to be maintained for optimal stabilization of collagen triple helices [74, 84]. Our results indicate that at least one site in collagen alpha 1 type 1 (P426) and two adjacent sites in collagen alpha 2 type 1 (P639 and P699) are hydroxylated in kidneys of SW but not in kidneys of FW sticklebacks. The functional significance of this difference is currently unknown calling for further study and confirmation using quantitation approaches that take into account stoichiometry,. Nonetheless, these 3 candidate sites are the only ones out of hundreds of other sites that display a very robust population-dependent pattern of P-Hox, which greatly narrows the focus for future, more detailed and targeted studies of this cotranslational modification in stickleback kidneys.

P-Hox of proteins other than collagen has been reported previously for other species. In particular, proteins that harbor collagen-like sequence domains, conotoxins, elastin, and argonaute proteins that are essential parts of the RNA-induced silencing complex (RISC) are known to be hydroxylated on proline [74, 85]. In contrast to collagen, some proteins subject to this cotranslational modification are only hydroxylated on a single proline residue [74]. For instance, hypoxia-inducible factor (HIF) alpha variants are hydroxylated on one out of two conserved prolines, which promotes ubiquitin ligase binding and HIF alpha degradation [86].

533 This molecular mechanism is evolutionarily highly conserved and critical for intracellular oxygen
534 sensing [86].

535 The only renal proteins other than collagen that were detected as robustly proline hydroxylated
536 in stickleback kidneys were 4 protein disulfide isomerases, including prolyl-4-hydroxylase beta
537 (P4Hb) (Table 2). To our best knowledge P-Hox of protein disulfide isomerases (PDIs) has not
538 been previously reported. All 4 *G. aculeatus* PDIs are hydroxylated at the consensus motif
539 **A**PWCGHCK. PDI A3 and A4 share the exact same hydroxylated peptide sequence but the
540 hydroxylated peptide sequences for P4Hb and ERp46 PDIs are unique, illustrating that the
541 above motif represents a true consensus site.

542 The P-Hox of P4Hb on proline 393 is intriguing since this enzyme is the main protein
543 responsible for P-Hox of collagen and other proteins [74]. It is possible that another enzyme,
544 e.g. prolyl-3-hydroxylase [87], hydroxylates P4Hb or, alternatively, that PH4b autohydroxylates
545 under certain conditions. These conditions may be met in males but not females since P-Hox
546 was not detected for any of the 4 PDIs in kidneys from female sticklebacks. Alternatively, the
547 lack of detection in kidneys of female fish may simply be a result of the greater abundance of
548 these PDIs in male kidneys (Figure 1). Whatever the case, PDI P-Hox likely alters the properties
549 of these enzymes. Therefore, resolving how PDIs are hydroxylated and how this modification
550 affects their substrate specificity, enzyme activity, stability, and localization represent important
551 directions for future research.

552 **Renal organic osmolyte synthesis in SW sticklebacks:** The biochemical and wTO network
553 models depicted in figures 6 and 7 include almost all proteins that were identified as significantly
554 more abundant in kidneys of SW sticklebacks compared to those from FW fish. The
555 corresponding network connects biochemical reactions that promote the synthesis of the
556 organic osmolytes betaine, sorbitol, TMAO, and urea. Glyoxylate and glycine are centrally
557 positioned within this network and represent sources for betaine and sorbitol synthesis. Glycine
558 N-methyltransferase (GNMT, G3NQA8) and alanine glyoxylate aminotransferase 2 (AGT2,
559 G3PVW6) represent two of the top 3 proteins with the highest Bodega Harbor / Lake Solano
560 abundance (BH/LS) ratio, the third one being dimethylaniline monooxygenase (DAMO,
561 G3PP04), which catalyzes the synthesis of the organic osmolyte TMAO. Moreover, three
562 alanine aminotransferases with high BH/LS ratios (AATs G3Q116, G3PYW9, G3NLG2) have
563 the ability to transaminate glyoxylate to glycine, in addition to utilizing alanine for pyruvate
564 production [88]. Only the latter reaction is shown in figure 6 to retain clarity.

The best known substrates for glyoxylate in vertebrates are glycine, glyoxal, hydroxyproline, and glycolate and their relative concentrations (plus the presence of the proper enzymes) determine the direction of the glyoxylate to glycine conversion [89]. Based on our data we propose that in kidneys of SW sticklebacks glyoxylate is predominantly converted to glycine, which is then further utilized for the production of the organic osmolytes betaine and sorbitol (figure 6).

The most likely sources of glyoxylate in stickleback kidneys are dietary glycolate and dietary hydroxyproline. Glycolate is abundant in plant-based diets while hydroxyproline is abundant in collagen-rich animal-based diets [83, 90]. Since sticklebacks are omnivores and food supply is plentiful in Bodega Harbor (eelgrass beds, zooplankton) glycolate and hydroxyproline both represent feasible sources for renal glyoxylate production. An enzyme (delta-1-pyrroline-5-carboxylate dehydrogenase) that is involved in the conversion of hydroxyproline to glyoxylate is indeed more abundant in kidneys of Bodega Harbor fish compared to kidneys of Lake Solano sticklebacks (in both males and females) although it did not meet the significance criteria applied in this study. Another key enzyme in this pathway, 4-hydroxy-2-oxoglutarate aldolase (HOGA1) [90], was not detected in our study.

Glyoxylate reductases (glycolate:NAD⁺ oxidoreductases G3PUE4, G3P0Y2), which catalyze the conversion of glycolate to glyoxylate have consistently high BH/LS ratios indicating that glycolate may be the main source of glyoxylate. However, the direction of this enzymatic reaction cannot be deduced from protein abundance data alone and additional work is necessary to confirm this hypothesis. Two enzymes that connect folate with glyoxylate metabolism via the mitochondrial cleavage system are dihydrofoate reductase (G3PJP3) and aminomethyltransferase (G3NQQ2). These enzymes have significantly elevated BH/LS ratios (Table 3) suggesting that another potential source of glyoxylate could be dietary folate.

Glyoxylate-dependent accumulation of glycine via the pathways discussed above enables betaine accumulation and the production of pyruvate and hydroxypyruvate for fueling gluconeogenesis (Figure 6). Pyruvate is metabolized to phosphoenolpyruvate (PEP) by pyruvate carboxylase (G3Q870) and phosphoenolpyruvate carboxykinase (G3NC51), both of which have significantly elevated BH/LS ratios. Increased concentrations of PEP and hydroxypyruvate then drive the synthesis of 3-phosphoglycerate, which stimulates gluconeogenesis. Three enzymes that promote gluconeogenesis and sorbitol synthesis are consistently more abundant in kidneys from SW sticklebacks (Table 3). Conversely, four glycolytic enzymes show the opposite trend, indicating that gluconeogenesis is favored over glycolysis in SW fish kidneys relative to FW fish kidneys. Interestingly, the use of glyceraldehyde-3-phosphate dehydrogenase (GA3PDH)

variants differs significantly in SW vs. FW populations. This finding strengthens the argument that the ratio of metabolic fluxes through glycolysis and gluconeogenesis is different in kidneys of SW and FW sticklebacks.

Differential usage of GA3PDH variants has been documented to alter the preference for glycolysis versus gluconeogenesis in other species [91]. In addition, metabolite fluxes through glycolysis versus gluconeogenesis and TCA cycle versus glyoxylate pathways are also regulated by posttranslational modification, e.g. by lysine acetylation, of the corresponding enzymes [92]. In our study we have not analyzed internal lysine acetylation of proteins but the data presented and discussed here give merit to conducting such analyses in different stickleback ecotypes. Another example for differential usage of protein variants in SW and FW sticklebacks is adenylate kinase (AK). This enzyme is critical for monitoring the state of energy metabolism in cells [93] and represented predominantly by the AK4 variant in SW and the AK1 variant in FW sticklebacks. In other species, these AK variants preferentially localize to different cell compartments and interact with different proteins, which results in distinct downstream signaling outcomes [93, 94].

Why do SW sticklebacks have a much greater capacity for producing organic osmolytes in their kidneys than FW fish? The question is not as trivial as it may sound. Teleost fishes are osmoregulators and it has been proposed that they are unable to increase urine osmolality above that of plasma [95]. Thus, despite SW being greatly hyperosmotic and FW strongly hypoosmotic relative to plasma, there should not be an osmotic gap that needs to be accounted for by accumulation of organic osmolytes nor should there be a need to generate hyperosmotic conditions in the kidney. In reality, however, there are differences in plasma osmolality (exceeding 100 mosmol / kg in some cases) when euryhaline fish such as three-spined sticklebacks are acclimated to either FW or SW [32]. Hence, teleost osmoregulation is highly effective but not perfect.

Organic osmolyte synthesis increases in parallel to environmental salinity in several euryhaline teleosts to buffer intracellular inorganic ion homeostasis against changes in plasma osmolality [96, 97]. A higher capacity for organic osmolyte synthesis in kidneys of SW – compared to FW - sticklebacks is consistent with those findings. Notably, sorbitol accumulation in mammalian kidneys is mediated mostly via aldose reductase [98] while our data suggest that in stickleback kidneys sorbitol dehydrogenase is the preferred enzyme for this task. Our data also suggest a greater capacity for synthesizing the organic osmolytes urea and TMAO in kidneys of SW

630 compared to FW sticklebacks. TMAO counteracts denaturing effects of urea and the
631 combination of these two organic osmolytes (usually in a 2:1 ratio) is cytoprotective [99].

632 A role of organic osmolytes for buffering plasma osmolality differences between SW and FW
633 fish is undisputed. Much more contentious is the question whether SW three-spined
634 sticklebacks have the ability to increase urine osmolality over that of plasma by accumulation of
635 organic osmolytes (and to what extent). It would seem beneficial for SW fish to accumulate high
636 levels of urea, TMAO and other organic osmolytes in the renal interstitium to aid in the urinary
637 concentration of metabolic waste and divalent ions while maximizing water reabsorption. And
638 yet, previous measurements of urine osmolality in other SW and FW fish do not support this
639 hypothesis [95, 100].

640 An alternative explanation for a greater capacity for synthesizing urea and TMAO in kidneys of
641 SW sticklebacks is their elevated amino acid metabolism which produces greater amounts of
642 nitrogenous waste than in FW sticklebacks. Elevated amino acid metabolism may fuel
643 gluconeogenesis, which our data also suggest to be elevated in kidneys of SW fish compared to
644 FW fish. However, many of the renal enzymes that are significantly different in abundance
645 between SW and FW fish belong to pathways that rely on glycine/ glyoxylate (amino acid)
646 metabolism but are distinct from gluconeogenesis (e.g. betaine synthesis and the urea cycle). In
647 addition, sorbitol dehydrogenase draws substrate from gluconeogenesis suggesting that
648 metabolite flux through gluconeogenesis serves (at least in part) to support the synthesis of the
649 compatible osmolyte sorbitol. Overall, increased amino acid metabolism in kidneys of SW
650 sticklebacks may serve the dual purpose of fueling glucose production and energy metabolism
651 via gluconeogenesis as well as increased production of organic osmolytes.

652 ***Proteolysis, calcium signaling, and renal structure in FW sticklebacks:*** Proteins involved in
653 proteolysis, cytoskeleton structure, and signal transduction are much more highly abundant in
654 kidneys of FW sticklebacks than in kidneys of SW fish. Many of these proteins are contained in
655 the innermost core of the wTO network. This observation suggests that protein turnover rates
656 are elevated in kidneys of FW sticklebacks, perhaps as a result of non-optimal conditions for
657 protein stability that are encountered during FW colonization. Renal function, in particular active
658 ion transport and water reabsorption, greatly differs in FW sticklebacks from that of their SW
659 ancestors (see above). Functional differences are supported by consistently and significantly
660 greater abundance of the Na⁺/K⁺-ATPase beta-233 subunit (G3PKB5) in FW stickleback
661 kidneys compared to SW stickleback kidneys. In agreement with this finding, renal Na⁺/K⁺-
662 ATPase activities of striped bass (*Morone saxatilis*) [101] and spotted green pufferfish

(*Tetraodon nigroviridis*) [102] are inversely correlated with salinity. On the contrary, in gilthead sea bream (*Sparus aurata*) Na⁺/K⁺-ATPase activity increases in parallel with salinity [103]. The reason for these species-specific differences is not known.

Differences in renal function are also reflected by markedly increased levels of many proteins associated with cytoskeletal structure. In addition, many signal transduction proteins, in particular proteins involved in calcium signaling, are greatly altered in their abundance in kidneys of FW – compared to SW - sticklebacks. The innermost core of the wTO network for the renal proteome of FW sticklebacks displays much less connectivity than that for SW sticklebacks. It consists of several discrete modules that reflect the functional diversity of the processes outlined above. Therefore, a detailed discussion of the many possible functional implications of altered levels for all these proteins exceeds the scope of the current paper. Nonetheless, several generalizations can be made.

First, abundance differences in many signaling and cytoskeletal proteins suggest extensive rearrangements of renal structure and function. Second, large differences in many renal calcium signaling proteins imply that renal calcium handling and reabsorptive processes for calcium and phosphate are greatly altered in sticklebacks that colonize FW. These processes are important for bone mineralization and morphogenesis and could contribute to differences in body armor [104]. Altered abundances of calcium signaling/ handling proteins in response to changes of environmental salinity have also been observed by proteomic analysis of whole larvae of European whitefish (*Coregonus lavaretus*) [105]. Third, proteome differences, in particular with regard to protein turnover and cytoskeletal structure, are mirrored by differences in renal ultrastructure. For instance, in FW female sticklebacks, glomeruli, podocyte nuclei, and renal tubular cells are larger than in SW females while the latter have a greater number of renal mesangial and juxtaglomerular cells [106]. Based on these ultrastructure data the author concluded that FW sticklebacks have elevated rates of proximal tubular reabsorption and lysosomal processing of nutrients that serve as building blocks for macromolecules [106]. Our data suggesting elevated protein turnover (e.g. higher abundances of lysosomal glycoproteinases and cathepsins) in kidneys of FW sticklebacks support this supposition.

Conclusions

Large differences in renal proteomes of SW (Bodega Harbor) and FW (Lake Solano) sticklebacks have been consistently identified in males and females collected in different years. Nt-Ac motifs of renal proteins suggest that NatA acetylase is responsible for the vast majority of

this cotranslational modification in stickleback kidneys. Identification of P-Hox patterns of renal collagens and a new consensus motif for P-Hox PDIs provide insight into possible cotranslational mechanisms for modulating extracellular matrix organization and renal proteostasis in different stickleback ecotypes. A biochemical network that is centered around glyoxylate/ glycine metabolism in the renal proteome of SW sticklebacks can readily be captured using a wTO network approach, which faithfully reproduces true biochemical connectivity. Therefore, this approach can be used in future analyses to reveal novel biochemical/ signaling networks. The SW stickleback network is proposed to promote the synthesis of four organic osmolytes and gluconeogenesis. Protein turnover, cytoskeleton structure, and calcium signaling are major processes that are elevated in kidneys of FW sticklebacks.

Finally, a recent transcriptomic analysis of three-spined stickleback kidneys from Bodega Harbor and a freshwater Lake at Davis, CA identified a number of transcripts that were either differently abundant in response to a salinity challenge or differently abundant in the two populations [107]. However, despite the use of comparable SW and FW populations there is virtually no overlap in differently abundant transcripts found in that study and differently abundant proteins found in the present study. This lack of correlation between transcriptome and proteome levels is not uncommon [108-110]. Since proteomes define higher-order phenotypes directly, their study will continue to make unique contributions to advance ecological and evolutionary biology.

References

- [1] Huntingford FA, Ruiz-Gomez ML. Three-spined sticklebacks *Gasterosteus aculeatus* as a model for exploring behavioural biology. *J Fish Biol.* 2009;75:1943-76.
- [2] Paepke H-J. *Die Stichlinge. Gasterosteidae*. 2nd ed. Magdeburg, Germany: Westarp-Wissenschaften, 175 pp.; 1996.
- [3] Tinbergen N, Vaniersel JJA. Displacement Reactions in the 3-Spined Stickleback. *Behaviour.* 1948;1:56-63.

723 [4] Bell MA. Stickleback fishes: Bridging the gap between population biology and paleobiology.
724 Trends Ecol Evol. 1988;3:320-4.

725 [5] Jones FC, Grabherr MG, Chan YF, Russell P, Mauceli E, Johnson J, et al. The genomic
726 basis of adaptive evolution in threespine sticklebacks. Nature. 2012;484:55-61.

727 [6] Kimmel CB, Ullmann B, Walker C, Wilson C, Currey M, Phillips PC, et al. Evolution and
728 development of facial bone morphology in threespine sticklebacks. Proceedings of the National
729 Academy of Sciences of the United States of America. 2005;102:5791-6.

730 [7] Colosimo PF, Peichel CL, Nereng K, Blackman BK, Shapiro MD, Schluter D, et al. The
731 genetic architecture of parallel armor plate reduction in threespine sticklebacks. Plos Biol.
732 2004;2:635-41.

733 [8] Diz AP, Martinez-Fernandez M, Rolan-Alvarez E. Proteomics in evolutionary ecology: linking
734 the genotype with the phenotype. Mol Ecol. 2012;21:1060-80.

735 [9] Feussner I, Polle A. What the transcriptome does not tell - proteomics and metabolomics are
736 closer to the plants' patho-phenotype. Curr Opin Plant Biol. 2015;26:26-31.

737 [10] Tomanek L. Environmental proteomics: changes in the proteome of marine organisms in
738 response to environmental stress, pollutants, infection, symbiosis, and development. Ann Rev
739 Mar Sci. 2011;3:373-99.

740 [11] Nedelkov D, Kiernan UA, Niederkofer EE, Tubbs KA, Nelson RW. Population proteomics:
741 the concept, attributes, and potential for cancer biomarker research. Mol Cell Proteomics.
742 2006;5:1811-8.

743 [12] Krebs HA. The August Krogh Principle: "For many problems there is an animal on which it
744 can be most conveniently studied". J Exp Zool. 1975;194:221-6.

745 [13] Koonin EV. Carl Woese's vision of cellular evolution and the domains of life. RNA Biol.
746 2014;11:197-204.

747 [14] Woese CR. Translation: in retrospect and prospect. RNA. 2001;7:1055-67.

- 748 [15] Somero GN. The physiology of global change: linking patterns to mechanisms. Ann Rev
749 Mar Sci. 2012;4:39-61.
- 750 [16] Somero GN. Unity in diversity: a perspective on the methods, contributions, and future of
751 comparative physiology. Annu Rev Physiol. 2000;62:927-37.
- 752 [17] Mueller IA, O'Brien KM. Nitric oxide synthase is not expressed, nor up-regulated in
753 response to cold acclimation in liver or muscle of threespine stickleback (*Gasterosteus*
754 *aculeatus*). Nitric Oxide. 2011;25:416 -22.
- 755 [18] Schaarschmidt T, Meyer E, Jürss K. A comparison of transport-related gill enzyme activities
756 and tissue-specific free amino acid concentrates of Baltic Sea (brackish water) and
757 freshwater threespine sticklebacks, *Gasterosteus aculeatus*, after salinity and temperature
758 acclimation. Mar Biol. 1999;135:689-97.
- 759 [19] Jürss K, Marin K, Kerstan T, Bastrop R. Population and gender-based differences in tissue-
760 specific enzyme activities in brackish water and freshwater threespine sticklebacks,
761 *Gasterosteus aculeatus*. Mar Biol. 1997;130:301-8.
- 762 [20] Jürss K, Bittorf T, Vökler T, Wacke R, Bastrop R. Experimental studies on biochemical and
763 physiological differences between the 3 morphs of the three-spined stickleback, *Gasterosteus*
764 *aculeatus* L .3. Liver somatic index and enzyme activities at different times of year. Zool Jahrb
765 Allg Zool. 1985;89:441-51.
- 766 [21] Jürss K, Bittorf T, Vökler T, Wacke R. Experimental studies on biochemical and
767 physiological differences between the 3 morphs of the three-spined Stickleback, *Gasterosteus*
768 *aculeatus* L . 2. Alanine aminotransferase, aspartate aminotransferase and glutamate
769 dehydrogenase activities of the liver. Zool Jahrb Allg Zool. 1983;87:1-7.
- 770 [22] Jürss K, Bittorf T, Vökler T, Wacke R. Experimental studies on biochemical and
771 physiological differences between the 3 morphs of the three-spined stickleback, *Gasterosteus*
772 *aculeatus* L . 1. Gill Na/K-ATPase, muscle alanine aminotransferase and muscle aspartate
773 aminotransferase Activities. Zool Jahrb Allg Zool. 1982;86:267-72.

774 [23] Svensson J, Fick J, Brandt I, Brunstrom B. Environmental concentrations of an androgenic
775 progestin disrupts the seasonal breeding cycle in male three-spined stickleback (*Gasterosteus*
776 *aculeatus*). *Aquat Toxicol*. 2014;147:84 -91.

777 [24] Hogan NS, Gallant MJ, van den Heuvel MR. Exposure to the pesticide linuron affects
778 androgen-dependent gene expression in the three-spined stickleback (*Gasterosteus*
779 *aculeatus*). *Environ Toxicol Chem*. 2012;31:1391 -5.

780 [25] Gao K, Brandt I, Goldstone JV, Jonsson ME. Cytochrome P450 1A, 1B, and 1C mRNA
781 induction patterns in three-spined stickleback exposed to a transient and a persistent inducer.
782 *Comp Biochem Physiol C Toxicol Pharmacol*. 2011;154:42-55.

783 [26] Hoffmann E, Walstad A, Karlsson J, Olsson PE, Borg B. Androgen receptor-beta mRNA
784 levels in different tissues in breeding and post-breeding male and female sticklebacks,
785 *Gasterosteus aculeatus*. *Reprod Biol Endocrinol*. 2012;10:23.

786 [27] Hogan NS, Wartman CA, Finley MA, van der Lee JG, van den Heuvel MR. Simultaneous
787 determination of androgenic and estrogenic endpoints in the threespine stickleback
788 (*Gasterosteus aculeatus*) using quantitative RT-PCR. *Aquat Toxicol*. 2008;90:269-76.

789 [28] Wendelaar Bonga SE, Greven JA, Veenhuis M. The relationship between the ionic
790 composition of the environment and the secretory activity of the endocrine cell types of *Stannius*
791 corpuscles in the teleost *Gasterosteus aculeatus*. *Cell Tissue Res*. 1976;175:297-312.

792 [29] Guderley H, Leroy PH. Family origin and the response of threespine stickleback,
793 *Gasterosteus aculeatus*, to thermal acclimation. *J Comp Physiol B*. 2001;171:91-101.

794 [30] Beyenbach KW. Kidneys sans glomeruli. *Am J Physiol Renal Physiol*. 2004;286:F811-27.

795 [31] Lahlou B, Henderson IW, Sawyer WH. Renal adaptations by *Opsanus tau*, a euryhaline
796 aglomerular teleost, to dilute media. *Am J Physiol*. 1969;216:1266-72.

797 [32] Kültz D. Physiological mechanisms used by fish to cope with salinity stress. *J Exp Biol*.
798 2015;218:1907-14.

799 [33] de Ruiter AJ. Changes in glomerular structure after sexual maturation and seawater
800 adaptation in males of the euryhaline teleost *Gasterosteus aculeatus* L. *Cell Tissue Res.*
801 1980;206:1-20.

802 [34] Barlet JP, Gaumet N, Coxam V, Davicco MJ. Calcitonine et stannioalcine. Particularites de
803 la regulation endocrinienne du metabolisme phospho-calcique chez les mammiferes et les
804 poissons. *Ann Endocrinol (Paris)*. 1998;59:281 -90.

805 [35] Kültz D, Li J, Gardell A, Sacchi R. Quantitative molecular phenotyping of gill remodeling in a
806 cichlid fish responding to salinity stress. *Mol Cell Proteomics*. 2013;12:3962-75.

807 [36] Lin H, He L, Ma B. A combinatorial approach to the peptide feature matching problem for
808 label-free quantification. *Bioinformatics*. 2013;29:1768-75.

809 [37] Gillet LC, Navarro P, Tate S, Rost H, Selevsek N, Reiter L, et al. Targeted Data Extraction
810 of the MS/MS Spectra Generated by Data-independent Acquisition: A New Concept for
811 Consistent and Accurate Proteome Analysis. *Mol Cell Proteomics*. 2012;11.

812 [38] Egertson JD, MacLean B, Johnson R, Xuan Y, MacCoss MJ. Multiplexed peptide analysis
813 using data-independent acquisition and Skyline. *Nat Protoc*. 2015;10:887-903.

814 [39] Picotti P, Aebersold R. Selected reaction monitoring-based proteomics: workflows,
815 potential, pitfalls and future directions. *Nat Methods*. 2012;9:555-66.

816 [40] Beausoleil SA, Villen J, Gerber SA, Rush J, Gygi SP. A probability-based approach for
817 high-throughput protein phosphorylation analysis and site localization. *Nat Biotechnol*.
818 2006;24:1285-92.

819 [41] Nesvizhskii AI, Keller A, Kolker E, Aebersold R. A statistical model for identifying proteins
820 by tandem mass spectrometry. *Anal Chem*. 2003;75:4646-58.

821 [42] MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, et al. Skyline:
822 an open source document editor for creating and analyzing targeted proteomics experiments.
823 *Bioinformatics*. 2010;26:966-8.

824 [43] Sharma V, Eckels J, Taylor GK, Shulman NJ, Stergachis AB, Joyner SA, et al. Panorama: a
825 targeted proteomics knowledge base. *J Proteome Res*. 2014;13:4205-10.

826 [44] Hochberg Y, Benjamini Y. More powerful procedures for multiple significance testing. Stat
827 Med. 1990;9:811-8.

828 [45] Matsuoka Y, Funahashi A, Ghosh S, Kitano H. Modeling and simulation using CellDesigner.
829 Methods in molecular biology. 2014;1164:121-45.

830 [46] Magrane M, Consortium U. UniProt Knowledgebase: a hub of integrated protein data.
831 Database : the journal of biological databases and curation. 2011;2011:bar009.

832 [47] Chang A, Schomburg I, Placzek S, Jeske L, Ulbrich M, Xiao M, et al. BRENDA in 2015:
833 exciting developments in its 25th year of existence. Nucleic acids research. 2015;43:D439-46.

834 [48] Stelzer G, Dalah I, Stein TI, Satanower Y, Rosen N, Nativ N, et al. In-silico human
835 genomics with GeneCards. Hum Genomics. 2011;5:709-17.

836 [49] Zhang B, Horvath S. A general framework for weighted gene co-expression network
837 analysis. Stat Appl Genet Mol Biol. 2005;4:Article17.

838 [50] Carlson MR, Zhang B, Fang Z, Mischel PS, Horvath S, Nelson SF. Gene connectivity,
839 function, and sequence conservation: predictions from modular yeast co-expression networks.
840 BMC Genomics. 2006;7:40.

841 [51] Ravasz E, Somera AL, Mongru DA, Oltvai ZN, Barabasi AL. Hierarchical organization of
842 modularity in metabolic networks. Science. 2002;297:1551-5.

843 [52] Nowick K, Gernat T, Almaas E, Stubbs L. Differences in human and chimpanzee gene
844 expression patterns define an evolving network of transcription factors in brain. Proceedings of
845 the National Academy of Sciences of the United States of America. 2009;106:22358-63.

846 [53] Eidsaa M, Almaas E. s-core network decomposition: a generalization of k-core analysis to
847 weighted networks. Phys Rev E Stat Nonlin Soft Matter Phys. 2013;88:062819.

848 [54] Schwartz D, Gygi SP. An iterative statistical approach to the identification of protein
849 phosphorylation motifs from large-scale data sets. Nat Biotechnol. 2005;23:1391-8.

850 [55] Distler U, Schmeisser MJ, Pelosi A, Reim D, Kuharev J, Weiczner R, et al. In-depth protein
851 profiling of the postsynaptic density from mouse hippocampus using data-independent
852 acquisition proteomics. *Proteomics*. 2014;14:2607-13.

853 [56] Qi D, Brownridge P, Xia D, Mackay K, Gonzalez-Galarza FF, Kenyani J, et al. A software
854 toolkit and interface for performing stable isotope labeling and top3 quantification using
855 Progenesis LC-MS. *OMICS*. 2012;16:489-95.

856 [57] Silva JC, Gorenstein MV, Li GZ, Vissers JP, Geromanos SJ. Absolute quantification of
857 proteins by LCMSE: a virtue of parallel MS acquisition. *Mol Cell Proteomics*. 2006;5:144-56.

858 [58] Chen K, Egertson J, MacCoss M. Targeted Data Independent Acquisition for Mass
859 Spectroscopy Proteomics. *J Invest Med*. 2013;61:204-.

860 [59] Li A, Horvath S. Network neighborhood analysis with the multi-node topological overlap
861 measure. *Bioinformatics*. 2007;23:222-31.

862 [60] Lee J, Valkova N, White MP, Kültz D. Proteomic identification of processes and pathways
863 characteristic of osmoregulatory tissues in spiny dogfish shark (*Squalus acanthias*). *Comp*
864 *Biochem Physiol Part D Genomics Proteomics*. 2006;1:328-43.

865 [61] Saxena S, Singh SK, Lakshmi MG, Meghah V, Sundaram CS, Swamy CV, et al. Proteome
866 profile of zebrafish kidney. *Journal of proteomics*. 2011;74:2937-47.

867 [62] Xiong XP, Dong CF, Xu X, Weng SP, Liu ZY, He JG. Proteomic analysis of zebrafish
868 (*Danio rerio*) infected with infectious spleen and kidney necrosis virus. *Dev Comp Immunol*.
869 2011;35:431-40.

870 [63] Nostbakken OJ, Martin SA, Cash P, Torstensen BE, Amlund H, Olsvik PA. Dietary
871 methylmercury alters the proteome in Atlantic salmon (*Salmo salar*) kidney. *Aquat Toxicol*.
872 2012;108:70-7.

873 [64] Nostbakken OJ, Goksoyr A, Martin SA, Cash P, Torstensen BE. Marine n-3 fatty acids alter
874 the proteomic response to methylmercury in Atlantic salmon kidney (ASK) cells. *Aquat*
875 *Toxicol*. 2012;106-107:65-75.

876 [65] Booy AT, Haddow JD, Ohlund LB, Hardie DB, Olafson RW. Application of isotope coded
877 affinity tag (ICAT) analysis for the identification of differentially expressed proteins following
878 infection of atlantic salmon (*Salmo salar*) with infectious hematopoietic necrosis virus
879 (IHNV) or *Renibacterium salmoninarum* (BKD). *J Proteome Res.* 2005;4:325-34.

880 [66] Ripley J, Iwanowicz L, Blazer V, Foran C. Utilization of protein expression profiles as
881 indicators of environmental impairment of smallmouth bass (*Micropterus dolomieu*) from the
882 Shenandoah River, Virginia, USA. *Environ Toxicol Chem.* 2008;27:1756-67.

883 [67] Booth NJ, Beekman JB, Thune RL. *Edwardsiella ictaluri* encodes an acid-activated urease
884 that is required for intracellular replication in channel catfish (*Ictalurus punctatus*)
885 macrophages. *Appl Environ Microbiol.* 2009;75:6712-20.

886 [68] Booth NJ, Bilodeau-Bourgeois AL. Proteomic analysis of head kidney tissue from high and
887 low susceptibility families of channel catfish following challenge with *Edwardsiella ictaluri*. *Fish*
888 *Shellfish Immunol.* 2009;26:193-6.

889 [69] Addis MF, Cappuccinelli R, Tedde V, Pagnozzi D, Viale I, Meloni M, et al. Influence of
890 *Moraxella* sp. colonization on the kidney proteome of farmed gilthead sea breams (*Sparus*
891 *aurata*, L.). *Proteome Sci.* 2010;8:50.

892 [70] Lu J, Chen H, Xu Q, Zheng J, Liu H, Li J, et al. Comparative proteomics of kidney samples
893 from puffer fish *Takifugu rubripes* exposed to excessive fluoride: an insight into molecular
894 response to fluorosis. *Toxicol Mech Methods.* 2010;20:345-54.

895 [71] Hentschel H, Elger M, Dawson M, Renfro L. Chapter 12: Urinary tract. In: Ostrander GK,
896 editor. *The laboratory fish.* San Diego: Academic Press; 2000. p. 181-8.

897 [72] Chen J, Wu HQ, Shi YH, Li CH, Li MY. The effects of environmental salinity on trunk kidney
898 proteome of juvenile ayu (*Plecoglossus altivelis*). *Comp Biochem Physiol Part D Genomics*
899 *Proteomics.* 2009;4:263-7.

900 [73] Plevoda B, Sherman F. Nalpa -terminal acetylation of eukaryotic proteins. *J Biol Chem.*
901 2000;275:36479-82.

902 [74] Gorres KL, Raines RT. Prolyl 4-hydroxylase. *Crit Rev Biochem Mol Biol.* 2010;45:106-24.

903 [75] Starheim KK, Gevaert K, Arnesen T. Protein N-terminal acetyltransferases: when the start
 904 matters. *Trends Biochem Sci.* 2012;37:152-61.

905 [76] Behnia R, Panic B, Whyte JR, Munro S. Targeting of the Arf-like GTPase Arl3p to the Golgi
 906 requires N-terminal acetylation and the membrane protein Sys1p. *Nature cell biology.*
 907 2004;6:405-13.

908 [77] Forte GM, Pool MR, Stirling CJ. N-terminal acetylation inhibits protein targeting to the
 909 endoplasmic reticulum. *Plos Biol.* 2011;9:e1001073.

910 [78] Mannherz HG, Hannappel E. The beta-thymosins: intracellular and extracellular activities of
 911 a versatile actin binding protein family. *Cell motility and the cytoskeleton.* 2009;66:839-51.

912 [79] Coulton AT, East DA, Galinska-Rakoczy A, Lehman W, Mulvihill DP. The recruitment of
 913 acetylated and unacetylated tropomyosin to distinct actin polymers permits the discrete
 914 regulation of specific myosins in fission yeast. *J Cell Sci.* 2010;123:3235-43.

915 [80] Chen X, Chi Y, Bloecher A, Aebersold R, Clurman BE, Roberts JM. N-acetylation and
 916 ubiquitin-independent proteasomal degradation of p21(Cip1). *Molecular cell.* 2004;16:839 -47.

917 [81] Permyakov SE, Vologzhannikova AA, Emelyanenko VI, Knyazeva EL, Kazakov AS,
 918 Lapteva YS, et al. The impact of alpha-N-acetylation on structural and functional status of
 919 parvalbumin. *Cell calcium.* 2012;52:366-76.

920 [82] Olinger E, Schwaller B, Loffing J, Gailly P, Devuyst O. Parvalbumin: calcium and
 921 magnesium buffering in the distal nephron. *Nephrol Dial Transplant.* 2012;27:3988-94.

922 [83] Wang W, Wu Z, Dai Z, Yang Y, Wang J, Wu G. Glycine metabolism in animals and
 923 humans: implications for nutrition and health. *Amino acids.* 2013;45:463-77.

924 [84] Ramshaw JA, Shah NK, Brodsky B. Gly-X-Y tripeptide frequencies in collagen: a context for
 925 host-guest triple-helical peptides. *J Struct Biol.* 1998;122:86-91.

926 [85] Lopez-Vera E, Walewska A, Skalicky JJ, Olivera BM, Bulaj G. Role of hydroxyprolines in
 927 the in vitro oxidative folding and biological activity of conotoxins. *Biochemistry.* 2008;47:1741-
 928 51.

929 [86] Kaelin WG. Proline hydroxylation and gene expression. Annual review of biochemistry.
930 2005;74:115-28.

931 [87] Hudson DM, Eyre DR. Collagen prolyl 3-hydroxylation: a major role for a minor post-
932 translational modification? Connect Tissue Res. 2013;54:245-51.

933 [88] Donini S, Ferrari M, Fedeli C, Faini M, Lamberto I, Marletta AS, et al. Recombinant
934 production of eight human cytosolic aminotransferases and assessment of their potential
935 involvement in glyoxylate metabolism. Biochem J. 2009;422:265-72.

936 [89] Pey AL, Albert A, Salido E. Protein homeostasis defects of alanine-glyoxylate
937 aminotransferase: new therapeutic strategies in primary hyperoxaluria type I. Biomed Res Int.
938 2013;2013:687658.

939 [90] Hoppe B. An update on primary hyperoxaluria. Nat Rev Nephrol. 2012;8:467-75.

940 [91] Ito F, Chishiki H, Fushinobu S, Wakagi T. Comparative analysis of two glyceraldehyde-3-
941 phosphate dehydrogenases from a thermoacidophilic archaeon, Sulfolobus tokodaii. FEBS Lett.
942 2012;586:3097-103.

943 [92] Wang Q, Zhang Y, Yang C, Xiong H, Lin Y, Yao J, et al. Acetylation of metabolic enzymes
944 coordinates carbon source utilization and metabolic flux. Science. 2010;327:1004-7.

945 [93] Zhang S, Nermutlu E, Terzic A, Dzeja P. Adenylate kinase signaling network: a major hub
946 in cell energetics and metabolic signaling. In: Aon MA, et al. e, editors. Systems biology of
947 metabolic and signaling networks. Berlin: Springer Verlag; 2014. p. 145-62.

948 [94] Dzeja P, Terzic A. Adenylate kinase and AMP signaling networks: metabolic monitoring,
949 signal communication and body energy sensing. Int J Mol Sci. 2009;10:1729-72.

950 [95] Jones IC, Chan DK, Rankin JC. Renal function in the European eel (*Anguilla anguilla* L.):
951 changes in blood pressure and renal function of the freshwater eel transferred to sea-water. J
952 Endocrinol. 1969;43:9-19.

953 [96] Gardell AM, Yang J, Sacchi R, Fangue NA, Hammock BD, Kültz D. Tilapia (*Oreochromis*
954 *mossambicus*) brain cells respond to hyperosmotic challenge by inducing myo-inositol
955 biosynthesis. JExpBiol. 2013;submitted.

956 [97] Kalujnaia S, Gellatly SA, Hazon N, Villasenor A, Yancey PH, Cramb G. Seawater
 957 acclimation and inositol monophosphatase isoform expression in the European eel (*Anguilla*
 958 *anguilla*) and Nile tilapia (*Oreochromis niloticus*). *Am J Physiol Regul Integr Comp Physiol*.
 959 2013;305:R369-84.

960 [98] Burg MB, Kwon ED, Kültz D. Osmotic regulation of gene expression. *FASEB J*.
 961 1996;10:1598-606.

962 [99] Yancey PH, Clark ME, Hand SC, Bowlus RD, Somero GN. Living with water stress:
 963 evolution of osmolyte systems. *Science*. 1982;217:1214-22.

964 [100] Genz J, McDonald MD, Grosell M. Concentration of MgSO₄ in the intestinal lumen of
 965 *Opsanus beta* limits osmoregulation in response to acute hypersalinity stress. *Am J Physiol*
 966 *Regul Integr Comp Physiol*. 2011;300:R895-909.

967 [101] Madsen SS, McCormick SD, Young G, Endersen JS, Nishioka RS, Bern HA. Physiology
 968 of seawater acclimation in the striped bass, *Morone saxatilis* (Walbaum). *Fish Physiol*
 969 *Biochem*. 1994;13:1-11.

970 [102] Lin CH, Tsai RS, Lee TH. Expression and distribution of Na, K-ATPase in gill and kidney
 971 of the spotted green pufferfish, *Tetraodon nigroviridis*, in response to salinity challenge. *Comp*
 972 *Biochem Physiol A Mol Integr Physiol*. 2004;138:287-95.

973 [103] Sangiao-Alvarellos S, Laiz-Carrión R, Guzmán JM, Martín del Río MP, Míguez JM,
 974 Mancera JM, et al. Acclimation of *S. aurata* to various salinities alters energy metabolism of
 975 osmoregulatory and nonosmoregulatory organs. *Am J Physiol Regul Integr Comp Physiol*.
 976 2003;285:R897-907.

977 [104] Guerreiro PM, Renfro JL, Power DM, Canario AV. The parathyroid hormone family of
 978 peptides: structure, tissue distribution, regulation, and potential functional roles in calcium and
 979 phosphate balance in fish. *Am J Physiol Regul Integr Comp Physiol*. 2007;292:R679-96.

980 [105] Papakostas S, Vasemagi A, Vaha JP, Himberg M, Peil L, Primmer CR. A proteomics
 981 approach reveals divergent molecular responses to salinity in populations of European whitefish
 982 (*Coregonus lavaretus*). *Mol Ecol*. 2012;21:3516-30.

- [106] Wendelaar Bonga SE. Morphometrical Analysis with Light and Electron-Microscope of Kidney of Anadromous 3-Spined Stickleback *Gasterosteus-Aculeatus*, Form *Trachurus*, from Fresh Water and from Sea-Water. *Z Zellforsch Mik Ana*. 1973;137:563-88.
- [107] Wang G, Yang E, Smith KJ, Zeng Y, Ji G, Connon R, et al. Gene expression responses of threespine stickleback to salinity: implications for salt-sensitive hypertension. *Front Genet*. 2014;5:312.
- [108] Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, et al. Global quantification of mammalian gene expression control. *Nature*. 2011;473:337-42.
- [109] Evans TG. Considerations for the use of transcriptomics in identifying the 'genes that matter' for environmental adaptation. *J Exp Biol*. 2015;218:1925-35.
- [110] Vogel C, Marcotte EM. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat Rev Genet*. 2012;13:227-32.

Figure legends

Figure 1: Molecular Tissue-specific Phenotypes (MoTiPhe) of three-spined stickleback kidneys based on proteomes analyzed in experiments **a)** 2012F, **b)** 2014F, and **c)** 2014M. The pie charts in the upper left display the share of the top 25 most highly abundant proteins to overall renal protein content. The column graphs in the lower left display the number of proteins detected relative to their placement within the dynamic range of detection. X axis values denote the percentage of protein abundance relative to overall renal protein content based on Scaffold weighted spectral counts. Note that the y axis is plotted on a log2 scale. The bar graphs on the right display the 25 most abundant proteins and their contribution to overall renal protein content.

Figure 2: Most common consensus motif surrounding proline hydroxylation (P-Hox) sites in stickleback collagens. This motif was identified with ScaffoldPTM (Proteome Systems, Portland, OR) based on a previously published motif search algorithm [54].

Figure 3: Results of label-free quantitative profiling of renal proteomes from Lake Solano freshwater (LS) and Bodega Harbor seawater (BH) sticklebacks. **A)** Volcano plots depicting the

abundance ratio versus PEAKS significance score distribution of all renal proteins for all three experiments. Dashed lines indicate the thresholds that were applied to identify proteins that are significantly differently abundant in kidneys of freshwater vs. seawater sticklebacks (upper left and upper right quadrants in each volcano plot). **B**) Heat maps depicting the abundances of all proteins clustering in the upper left and upper right quadrants of the volcano plots shown in panel A. Each column represents a single individual and each row a unique protein. Note that heat maps are displayed on a log₂ scale with bright red representing proteins with an LS/BH ratio of 16 or higher and bright green representing proteins with an LS/BH ratio of 1/16 or lower. A legend on the lower right displays the color key for intermediate ratios.

Figure 4: Targeted MS/MS transition-based quantitation of peptide AAALDVTHPEPLPR (m/z = 496.2718, z = 3+) for protein G3P0Y2 (Glyoxylate reductase) for samples collected in experiments 2014F and 2014M and analyzed by data-independent acquisition. **A**) Library spectrum generated from prior data-dependent LCMS runs representing this peptide in Skyline. Singly-charged MS/MS fragment ions are ranked by intensity (y₆ > y₄ > b₂ > b₃ > y₇). **B**) Cumulative peak area quantitation graph generated by Skyline. Each sample is represented by a bar that is color-coded according to the five most intense MS/MS fragment ions in the spectral library. The dotp values (0 – 1) indicate the quality of the match in relative intensities of each of the fragment ions to that of the spectral library (first bar on the left). Samples 2 – 7 are for Bodega Harbor (BH) females, 8 – 13 for BH males, 14 – 19 for LS females, and 20 – 25 for LS males. **C**) Five transitions representing this peptide for each of the samples analyzed. Transition traces are color-coded as indicated in panel B. These traces illustrate the high quality and reproducibility of chromatographic separation (narrow peak width, symmetrical peak shape, low background, good retention time reproducibility) and high mass accuracy (< 27 ppm error).

Figure 5: Validation of population-specific abundance differences in renal proteins targeted for quantitative Skyline analysis of the corresponding MS/MS transitions. Abundance ratios (mean Bodega Harbor/ mean Lake Solano) for all targeted proteins are depicted (upper panel: LS/BH > 1; lower panel: LS/BH < 1). Error bars illustrate S.E.M. (n=6). The data for glyoxylate reductase (G3P0Y2) shown in position 8 of the lower panel correspond to those shown in figure 5. An asterisk denotes that there is no significant population difference in the abundance of the corresponding protein.

Figure 6: Biochemical network consisting of enzymes that are significantly more abundant (cyan) or less abundant (orange) in kidneys from seawater (Bodega Harbor) sticklebacks compared to kidneys from freshwater (Lake Solano) sticklebacks. Enzymes are depicted as

rounded rectangles, metabolites are represented as ovals. Amino acids are shown in green and organic osmolytes in yellow. Blue arrows denote enzymatic reactions that support the synthesis of the organic osmolytes sorbitol, betaine, TMAO, and urea. Red arrows denote enzymatic reactions that oppose metabolite fluxes supporting organic osmolyte synthesis. Note that folate metabolism is directly connected to serine and glycine metabolism (e.g. via serine hydroxymethyl transferase) but this connection is not shown to retain clarity of the diagram. PPP = pentose phosphate pathway.

Figure 7: Central cores in the protein abundance network resulting from weighted network analysis. Blue (green) shading of nodes indicates strength of BH/LS ratio. A red (blue) link color indicates a positive (negative) interaction. Conducting s-core+ analysis of the weighted protein abundance network of **A)** Lake Solano displaying the 230 most central proteins, and **B)** Bodega Harbor displaying the 223 most central proteins. When filtering on $LS/BH > 2$ before s-core+ analysis, the **C)** Lake Solano network consists of 40 nodes and **D)** the Bodega Harbor of 52 nodes (with $BH/LS > 2$). All proteins are annotated with their UniProtKB accession number.

1058 **Table 1:** Identification of amino-terminally acetylated renal proteins.

UniProt AC	Protein annotation	Pos.	Peptide	LS12 F	BH12 F	LS14 F	BH14 F	LS14 M	BH14 M
G3PSC5	40S ribosomal protein S15	A1	A(+42.01)DTEIKK.K	3	0	5	1	3	0
G3NMQ4	60S ribosomal protein L4	A1	A(+42.01)CSRPLISVYSEK.G	5	0	0	0	1	0
G3NHX0, G3NHX7	14-3-3 protein beta/alpha	A2	M.A(+42.01)DKEEQIQK.A	6	3	2	0	4	1
G3PA34	40S ribosomal protein S12	A2	M.A(+42.01)EEGSSVATGGVMDVNTALPEVLK.T	4	3	3	4	5	4
G3NMS7, G3NMT5	60S ribosomal protein L4	A2	M.A(+42.01)CSRPLISVYSEK.G	4	4	3	3	5	4
G3P3C4	Actin-capping protein CapZ subunit alpha	A2	M.A(+42.01)DHEEQSLDEEK.V	4	5	3	4	5	6
G3PWK6	Actin-related protein 2/3 complex subunit 1B	A2	M.A(+42.01)YHSFLLEPISC(+57.02)HAWNK.D	4	3	0	0	0	0
G3PWK6	Actin-related protein 2/3 complex subunit 1B	A2	M.A(+42.01)YHSFLLEPISCHAWNK.D	4	1	1	0	1	0
G3NDS5	Adenylyl cyclase-associated protein 1	A2	M.A(+42.01)DLASLVQR.L	5	2	4	4	4	4
G3NN97	Calmodulin	A2	M.A(+42.01)DQLTEEQIAEFK.E	3	1	0	0	0	0
G3N5N8	Coactosin-like protein	A2	M.A(+42.01)TQIDKEACR.D	6	6	5	6	6	6
G3NJF0	Cofilin 2	A2	M.A(+42.01)SGVTVNDEVIR.V	6	6	0	0	0	0
G3NJ65, G3NJ74	Cytosolic non-specific dipeptidase 2	A2	M.A(+42.01)HLPALFK.Y	3	2	0	1	1	0
G3Q0K3	Elongation factor 1-gamma-A	A2	M.A(+42.01)AGTLTYTPENWR.A	1	3	0	2	0	0
G3NKU8	Eukaryotic translation initiation factor 4H	A2	M.A(+42.01)DFDNYDER.D	1	3	0	2	0	0
G3PKE0, G3PKE4	Galectin-8	A2	M.A(+42.01)LNQQQPFYNIPIPTGSIHGGGLQEGK.S	6	6	0	0	0	0
G3P2D0	Glutaredoxin-1	A2	M.A(+42.01)QQFVQTK.I	1	1	1	0	0	2
G3NDH2, G3NEJ1	Histone H1	A2	M.A(+42.01)EVAPAPAAAAPK.A	6	5	3	4	4	6
G3NY66	Maleylacetoacetate isomerase	A2	M.A(+42.01)SQTKPVLHGYFR.S	4	0	1	2	0	1
G3PQB1, G3PQB3	Muscleblind-like protein 1	A2	M.A(+42.01)MNMAHIR.D	0	0	0	0	0	2
G3NU76	N[G]N[G]-dimethylarginine dimethylaminohydrolase 1	A2	M.A(+42.01)GLMSGFGHYTHAVVR.G	6	5	5	5	4	4
G3PZH5	NADH dehydrogenase [ubiquinone] 1 alpha 12	A2	M.A(+42.01)QYANIVR.R	5	5	4	2	5	4
G3PET2	Parvalbumin beta	A2	M.A(+42.01)FASVGLTDPDVAK.A	6	5	0	0	0	1
G3N963	Prohibitin-2	A2	M.A(+42.01)NKDPGGFLQQLR.Q	6	6	3	2	4	6
G3NXC1	Prothymosin alpha	A2	M.A(+42.01)DTAIEK.T	5	1	3	2	4	5
G3PJ29	Rho GDP-dissociation inhibitor 1	A2	M.A(+42.01)EQEPTPEQLEELAAANR.E	6	6	6	6	6	6
G3PCP4	SAP domain-containing ribonucleoprotein	A2	M.A(+42.01)EVIELQK.L	2	3	0	0	2	1
G3NZG1	S-formylglutathione hydrolase	A2	M.A(+42.01)LTQVSSNR.C	0	2	0	0	0	2
G3PA75, G3PA76	Thioredoxin domain-containing protein 17	A2	M.A(+42.01)HYEEVAVR.G	5	0	0	0	0	0
G3PJ42	THO complex subunit 4 / Transcriptional coactivator Aly-REF / bZIP-enhancing factor BEF	A2	M.A(+42.01)DKMDMSLDDIIK.Q	5	0	0	0	0	0
G3NRF8, G3NRH8	Transgelin	A2	M.A(+42.01)NKGPAPFGMSR.H	6	6	1	1	0	0
G3QAJ2	Transgelin	A2	M.A(+42.01)NVGPSFGLSR.E	5	5	1	1	0	0
G3PBF5	Tropomyosin alpha-4 chain	A2	M.A(+42.01)GLTSLDAVK.R	6	6	0	0	0	0
G3PDF5	Ubiquitin-conjugating enzyme E2 variant 1	A2	M.A(+42.01)VAAGSVVK.V	6	2	3	0	2	0
G3Q069, G3Q071	UBX domain protein 1 / SAPK substrate protein 1	A2	M.A(+42.01)ELTTLESLLMFGFER.N	0	0	0	0	0	2
G3QCJ5	Voltage-dependent anion-selective channel protein 1	A2	M.A(+42.01)APPTYVDLGK.S	3	4	0	0	0	2
G3P2M1	Voltage-dependent anion-selective channel protein 2	A2	M.A(+42.01)VPPCYADLGK.S	5	5	1	1	2	1
G3PS29	V-type proton ATPase subunit G 1	A2	M.A(+42.01)SQSQGIQQLLQAEK.R	5	5	3	4	3	2
G3NQS8	Leucyl aminopeptidase	C2	M.C(+42.01)(+57.02)TSIQPVEWTADLK.N	4	1	3	1	2	0
G3NQS8	Leucyl aminopeptidase	C2	M.C(+42.01)TSIQPVEWTADLK.N	5	6	5	5	2	5
G3N8X9, G3N8W8	Actin, beta	E2	M.E(+42.01)DDIAALVVDNGSGMCK(+57.02)K.A	0	0	1	0	5	1
G3N8X9, G3N8W8	Actin, beta	E2	M.E(+42.01)DDIAALVVDNGSGMCK(+43.01).A	1	0	0	0	2	0
G3N8X9, G3N8W8	Actin, beta	E2	M.E(+42.01)DDIAALVVDNGSGMCK.A	1	1	0	4	0	0

G3NH29, G3NE92	14-3-3 protein beta/alpha	M1	M(+42.01)DKNDLVQK.A	3	1	1	0	2	0
G3P1S2, G3P1S3	40S ribosomal protein S24	M1	M(+42.01)NDTVTVR.T	6	5	0	0	0	0
G3PSE7	40S ribosomal protein S28	M1	M(+42.01)DASR.V	5	6	3	3	2	1
G3P2Y2	Acetyl-CoA acetyltransferase, cytosolic	M1	M(+42.01)NPEPVVIVSAAR.T	3	1	0	0	0	0
G3N7V4	Actin-capping protein CapZ subunit beta	M1	M(+42.01)NDQQLDC(+57.02)ALDLMR.R	0	1	0	1	1	4
G3N7V4, G3N7W0	Actin-capping protein CapZ subunit beta	M1	M(+42.01)NDQQLDCALDLMR.R	2	3	0	1	0	0
G3QBU5	ADP/ATP translocase 2	M1	M(+42.01)NETAVSFAK.D	6	6	1	1	1	1
G3QBU5	ADP/ATP translocase 2	M1	M(+42.01)NETAVSFAKDFLAGGISAAISK.T	6	6	0	0	0	0
G3NGX3	Alcohol dehydrogenase [NADP+] B	M1	M(+42.01)NDFAVLNTGR.K	3	1	1	0	3	2
G3PVW1, G3PVW3	Alcohol dehydrogenase 3	M1	M(+42.01)ETAGK.V	5	3	1	0	2	0
G3Q0C8, G3Q0D0	Aspartoacylase 2B	M1	M(+42.01)EHVSFLPLSR.F	1	1	1	1	3	0
G3Q1V4	Delta-aminolevulinic acid dehydratase	M1	M(+42.01)QTPAESIIHSGYFHPTLR.Y	1	2	3	2	4	1
G3PGF0	Galectin-1	M1	M(+42.01)MNGMIK.N	1	4	0	0	2	0
G3PPH3	Importin / karyopherin subunit beta-1	M1	M(+42.01)ELITILEK.T	6	5	1	1	0	0
G3P455, G3P467	LIM and SH3 domain protein 1 [LASP1]	M1	M(+42.01)NPIC(+57.02)SR.C	4	1	0	0	0	1
G3P455, G3P467	LIM and SH3 domain protein 1 [LASP1]	M1	M(+42.01)NPICSR.C	5	5	6	6	4	0
G3NU80	Poly[rC]-binding protein 2	M1	M(+42.01)ENQSHLGDQR.A	1	2	1	2	1	1
G3NKS5	Protein S100-A16	M1	M(+42.01)ESAIQTIVQVFLK.S	4	4	0	0	1	1
G3NFD8	Purine nucleoside phosphorylase	M1	M(+42.01)EAAPSTSPC(+57.02)R.Y	1	3	0	0	2	3
G3NFD8	Purine nucleoside phosphorylase	M1	M(+42.01)EAAPSTSPCR.Y	0	0	1	0	2	2
G3PR91	Splicing factor 3A subunit 2	M1	M(+42.01)DFQHRA.A	3	5	0	0	0	0
G3NYT4	Splicing factor 3A subunit 3	M1	M(+42.01)ETILEQQR.R	1	0	5	3	4	2
G3Q8M1	T-complex protein 1 chaperonin subunit eta	M1	M(+42.01)MPAQVILLK.E	6	4	2	1	4	2
G3P7N4, G3P7I0	T-complex protein 1 subunit gamma	M1	M(+42.01)MGQQVLVLNHNMK.R	4	3	4	4	3	3
G3PJZ4	Tetratricopeptide repeat protein 38	M1	M(+42.01)LASSFR.D	0	0	0	0	3	1
G3P8Z9	Tropomyosin alpha-1 chain	M1	M(+42.01)DAIK.K	5	0	1	0	0	0
G3P7U2, G3P7U9	Tumor protein D54	M1	M(+42.01)DPASQDINLNSPNK.G	3	0	3	0	3	0
G3PYB9, G3PYC5	Ubiquitin carboxyl-terminal hydrolase isozyme L1	M1	M(+42.01)EWTPMEINPEMLNK.M	6	6	0	0	0	0
G3NIR4	Ubiquitin fusion degradation protein 1 homolog	M1	M(+42.01)FSFHVFDHPMSR.G	5	1	0	1	1	0
G3NB29	UPF0317 protein C14orf159	M1	M(+42.01)EQSNGLHFR.S	0	0	0	1	0	4
G3Q6B5	UPF0585 protein C16orf13 homolog A	M2	R.M(+42.01)LNAAAAER.N	5	1	0	0	0	0
G3P5V0	40S ribosomal protein S18	S2	M.S(+42.01)LVIEPKFQHILR.V	6	6	4	4	4	6
G3NRK0	40S ribosomal protein SA / 37/67 kDa laminin receptor	S2	M.S(+42.01)GGLDVLQMK.E	6	5	6	6	6	1
G3PUE2	60S ribosomal protein L28	S2	M.S(+42.01)SHLQWMVIR.N	0	0	3	1	0	1
G3PY86, G3PY99	ADP/ATP translocase 3	S2	M.S(+42.01)ETAIISFAK.D	6	4	0	0	1	0
G3NMV1	Carbonic anhydrase 1	S2	M.S(+42.01)HAWGYAANNGPDK.W	5	1	1	0	1	0
G3NMV1	Carbonic anhydrase 1	S2	M.S(+42.01)HAWGYAANNGPDKWADNFPIANGPR 2 .Q	0	2	2	4	1	
G3NG23	Cystatin B	S2	M.S(+42.01)AMCGGLGAISEANEEIQK.I	3	0	4	1	4	5
G3P8F3	Cytochrome c oxidase subunit 6B	S2	M.S(+42.01)DTVEEK.I	2	3	0	0	0	0
G3NP05, G3NP10	Enhancer of rudimentary homolog	S2	M.S(+42.01)HTILLVQPTK.R	4	3	2	0	6	5
G3NU46	Small nuclear ribonucleoprotein Sm D2	S2	R.S(+42.01)LLTKPK.S	2	2	0	0	0	0
G3P5F8	Alcohol dehydrogenase (NADP+) A	S2	M.S(+42.01)AFVTLSTGQR.M	6	5	6	3	5	5
G3PNE8	Hemoglobin subunit alpha-2	S2	M.S(+42.01)LTDKDK.A	3	3	0	0	0	0
G3PNE8	Hemoglobin subunit alpha-2	S2	M.S(+42.01)LTDKDKATVK.A	6	4	1	0	1	3
G3NLD8	High density lipoprotein [HDL] binding protein vigilin	S2	M.S(+42.01)SVAVLTKESFNEHR.S	5	0	0	0	0	0
G3P2I6	Myosin light chain 6	S2	M.S(+42.01)DFSEDQILEFK.E	5	0	0	0	0	0

G3PWF4	Peptidyl-prolyl cis-trans isomerase D / Cyclophilin D	S2	M.S(+42.01)HPVPSEKPSNPANPR.V	5	5	4	5	1	0
G3NZU2	Putative ATP-dependent RNA helicase an3	S2	M.S(+42.01)HVVVDNPHGLDQQLSALDLNSADGQ GGGTGR.R	3	0	4	0	0	0
G3PPW2	Serine/arginine-rich splicing factor 2	S2	M.S(+42.01)YGRPPPDVDGMTSLK.V	4	5	3	2	3	0
G3PPW2	Serine/arginine-rich splicing factor 2	S2	M.S(+42.01)YGRPPPDVDGMTSLKVDNLTYR.T	3	5	0	0	0	0
G3N9B7, G3N9B8	Serine/arginine-rich splicing factor 3	S2	M.S(+42.01)QTDQFPLDCK.V	2	4	4	5	3	2
G3NFD3	Tropomyosin alpha-3 chain	S2	M.S(+42.01)GGLNSIDAVK.K	4	0	0	0	0	0
G3NIZ7	Twinfilin-1	S2	M.S(+42.01)HQTGIQAGNDVK.D	6	6	6	6	6	6
G3NEJ8	40S ribosomal protein S5	T2	M.T(+42.01)EWETAPAVAESPEIK.L	5	3	6	5	6	6
G3NJP8	Fructose-bisphosphate aldolase A	T2	M.T(+42.01)HAYPFLSTEQK.K	1	2	1	4	2	0
G3NJP8	Fructose-bisphosphate aldolase A	T2	M.T(+42.01)HAYPFLSTEQK.K	0	0	0	0	3	0
G3Q1L3	Fructose-bisphosphate aldolase B	T2	M.T(+42.01)HQFPSLSQEQK.K	1	1	2	3	3	1
G3Q1L3	Fructose-bisphosphate aldolase B	T2	M.T(+42.01)HQFPSLSQEQK.K	4	5	0	1	0	0
G3QBK7	Fructose-bisphosphate aldolase B	T2	M.T(+42.01)HQYPALTPEQK.K	6	6	3	4	5	0
G3QBK7	Fructose-bisphosphate aldolase B	T2	M.T(+42.01)HQYPALTPEQK.K	3	0	3	0	3	0
G3P4V7	NHP2-like protein 1	T2	M.T(+42.01)EAQVNP.K	6	6	6	5	6	6
G3NYX9	Plastin-2	T2	M.T(+42.01)TPVPISPEEQEDLR.E	1	0	3	2	0	0
G3P6R2	Transketolase-like protein 2	T2	M.T(+42.01)SYHKPDEK.T	6	6	0	0	0	0

1059 Pos. indicates the modified amino acid in single letter code followed by its position in the protein
1060 sequence. The numbers below the headers LS12F (Lake Solano females from 2012), BH12F
1061 (Bodega Harbor females from 2012), LS14F (Lake Solano females from 2014), BH14F (Bodega
1062 Harbor females from 2014), LS14M (Lake Solano males from 2014), and BH14M (Bodega
1063 Harbor males from 2014) indicate the numbers of samples in which the amino-terminally
1064 acetylated peptide has been detected (out of maximal 6). The confidence score for correct
1065 localization of the modification to the corresponding amino acid (Ascore) was maximal (1000) for
1066 all proteins shown and all replicates indicated. Rows printed in bold italic designate proteins for
1067 which the detection frequency of amino-terminally acetylated peptides was consistently much
1068 higher for freshwater than for seawater sticklebacks.

1069

1070 **Table 2:** Identification of proline-hydroxylated protein disulfide isomerases.

UniProt AC	Annotation	Pos.	Peptide	Lake Solano		Bodega Harbor	
				Ascore	n	Ascore	n
G3NVG1	Protein	P39					
	disulfide- isomerase A3 (PDI A3)	5 P39 5	K.DVLIEFYAP(+15.99)WCGHCK(+43.01).S K.DVLIEFYAP(+15.99)WCGHCK.S	1000 1000	1 1	1000 1000	4 4
G3Q2E1	Prolyl 4- hydroxylas e subunit beta (PDI)	P39 3	K.NVFVEFYAP(+15.99)WCGHCK.Q	1000	1	1000	3
G3NCG1	Thioredoxi n domain- containing protein 5 (PDI ERp46)	P18 0 P18 0	K.FFAP(+15.99)WC(+57.02)GHC(+57.02)K.A K.FFAP(+15.99)WCGHC(+57.02)K.A	1000 1000	1 1	1000 1000	1 2
G3ND88	Protein	P47	K.DVLIEFYAP(+15.99)WC(+57.02)GHC(+57.02)K.K	1000	1	1000	2
	disulfide- isomerase A4 (PDI A4)	8 P47 8 P47 8	K.DVLIEFYAP(+15.99)WCGHCK(+43.01).K K.DVLIEFYAP(+15.99)WCGHCK.K	1000 1000	1 1	1000 1000	4 4

1071

1072 The peptides shown were only detected in experiment 2014M (performed on Bodega Harbor

1073 males collected in 2014). Pos. indicates the modified amino acid in single letter code followed

1074 by its position in the protein sequence. The Ascore indicates the confidence of assigning the

1075 modification to the corresponding amino acid (maximum Ascore = 1000) [40]; n = numbers of

1076 samples in which the corresponding peptide was detected (out of maximal 6).

1077

1078 **Table 3:** Proteins that are significantly differently abundant in kidneys from freshwater and
1079 seawater sticklebacks.

UniProt AC Protein annotation		2012 Females					2014 Females					2014 Males				
		Cov.	Pep.	Uniq.	LS/BH	sign. P	Cov.	Pep.	Uniq.	LS/BH	sign. P	Cov.	Pep.	Uniq.	LS/BH	sign. P
Amino acid metabolism																
G3NQA8	Glycine N-methyltransferase	21	5	5	0.0	0.0191	36	8	8	0.0	0.0083	44	10	10	0.1	0.0027
G3Q116	Alanine aminotransferase 1	15	5	5	0.2	0.0300	19	7	7	0.1	0.0317	2	1	1	0.7	0.5793
G3PKH2, G3PKG4	Serine-pyruvate aminotransferase	42	16	16	0.1	0.0102	42	17	17	0.1	0.0062	17	8	8	0.3	0.0627
G3PVW6	Alanine-glyoxylate aminotransferase 2	69	23	13	0.0	BH only	80	30	18	0.0	0.0108					
G3NXE2	Aspartate aminotransferase cytoplasmic	29	11	9	0.4	0.0055	33	11	9	0.4	0.0304	19	6	6	0.4	0.0382
G3NWI4, G3NWI0	Histamine N-methyltransferase	8	2	2	0.0	0.0084	22	5	5	0.0	0.0027	5	1	1	0.6	0.1443
G3PUE4	Glyoxylate reductase	24	6	6	0.1	0.0166	29	9	8	0.1	0.0298					
G3P0Y2	Glyoxylate reductase	24	5	3	0.3	0.1025	34	7	6	0.2	0.0006	24	6	5	0.4	0.0129
G3PYW9	Alanine aminotransferase 2-like	25	11	5	0.1	0.0119	45	20	19	0.1	0.0007	14	7	7	0.9	0.8301
G3NLG2, G3PME0	Alanine aminotransferase 2	5	2	2	0.2	0.0897	11	6	6	0.0	0.0212	2	1	1	0.4	0.1651
G3PZF6	Sarcosine dehydrogenase	13	10	10	0.7	0.3006	16	12	12	0.0	BH only	11	8	8	0.4	0.0020
G3PKG8	Peroxisomal sarcosine oxidase	9	3	3	0.4	0.1381	13	4	4	0.1	0.0151					
G3P497	Amine oxidase [flavin-containing]	5	2	2	0.1	0.0483	9	4	4	0.3	0.0267					
G3N6Z7	Urocanate hydratase	15	8	8	0.2	0.0385	19	10	10	0.1	0.0265	2	1	1	0.6	0.4196
G3NKZ1	Histidine ammonia-lyase	24	12	12	0.5	0.0622	20	11	11	0.0	0.0042					
G3Q9D1	Probable imidazolonepropionase	4	2	2	0.1	0.0741	15	6	6	0.1	0.0135					
G3NHJ8	Adenosylhomocysteinase	44	19	18	0.5	0.0630	51	22	22	0.6	0.0526	49	23	22	0.5	0.0271
G3PJP3	Dihydrofolate reductase	13	2	2	0.2	0.0569						17	3	3	0.1	0.0003
G3NQQ2	Aminomethyltransferase	6	2	2	0.1	0.0129	10	4	4	0.1	0.0204	8	3	3	0.5	0.1390
G3PGE5	Methylmalonate-semialdehyde dehydrogenase [acylating]	58	18	18	0.4	0.0396	64	24	23	0.5	0.0313	47	17	16	1.2	0.6155
Antioxidant/ xenobiotic metabolism																
G3Q613	Thioredoxin	55	5	5	3.2	0.0270	60	6	6	2.0	0.0256	55	5	5	2.6	0.0129
G3Q3S1	SH3 domain-binding glutamic acid-rich-like protein	59	7	7	5.2	0.0028	39	4	4	2.0	0.1647	59	4	4	25.2	0.0339
G3QA68	Ester hydrolase C11orf54 homolog	50	11	11	0.3	0.0081	53	11	11	0.4	0.0400	40	10	10	0.8	0.5237
G3NA48	Aflatoxin B1 aldehyde reductase 2	31	7	7	0.3	0.0369	41	9	9	0.2	0.0003	38	8	8	0.6	0.0256
G3NXL1	Pyridine nucleotide-disulfide oxidoreductase domain-containing protein 2	28	10	10	0.2	0.0027	29	11	11	0.0	BH only					
G3N8Q7	Abhydrolase domain-containing protein 14B	61	9	9	0.6	0.0218	61	9	9	0.4	0.0076	32	5	5	0.8	0.4508
G3PIX6	Glutathione peroxidase	11	1	1	0.3	0.0047	26	3	3	0.2	0.0048					
TCA cycle																
G3NC51	Phosphoenolpyruvate carboxykinase [GTP]	59	32	28	0.4	0.0099	62	38	35	0.2	0.0026	50	29	26	0.7	0.2390
G3NLL5	4-aminobutyrate aminotransferase	51	18	18	0.1	0.0098	52	20	19	0.0	0.0017	35	12	12	0.4	0.0042
G3NEX2	Succinate-semialdehyde dehydrogenase	13	3	3	0.3	0.0455	22	7	6	0.2	0.0067	5	2	2	0.6	0.2022
G3PLG0, G3PLG5	Fumarylacetoacetase	9	4	4	0.4	0.0371	17	5	5	0.2	0.0049	3	1	1	1.2	0.6032
G3P457	Succinyl-CoA ligase [GDP-forming] subunit beta	29	12	12	0.4	0.0376	33	16	16	0.3	0.0216	6	3	3	1.1	0.6468
G3P3D2	Isocitrate dehydrogenase [NADP]	17	7	4	0.6	0.1519	28	11	8	0.4	0.0060	19	7	5	0.8	0.4777
G3Q870	Pyruvate carboxylase	26	24	23	0.7	0.1133	32	30	29	0.3	0.0001	15	15	15	0.6	0.1392
G3NF09	Putative hydroxypyruvate isomerase						21	4	4	0.1	0.0130					
Gluconeogenesis/ Glycolysis																
G3PTJ4	Fructose-1,6-bisphosphatase 1	39	10	10	0.1	0.0041	44	15	15	0.1	0.01631	12	3	3	0.5	0.0578
G3PSH5	Sorbitol dehydrogenase	12	4	4	0.2	0.0032	8	2	2	0.2	0.0389	8	2	2	0.3	0.0872
G3Q4G6, G3Q4G2	Glyceraldehyde-3-phosphate dehydrogenase	23	10	1	0.3	0.0126	45	14	7	1.0	0.8545	45	12	6	0.6	0.7138
G3NR74, G3NRB0	Glyceraldehyde 3-phosphate dehydrogenase 1	44	13	9	2.4	0.0128	45	15	12	1.9	0.0087	45	14	10	1.4	0.0397
G3P5F8	Alcohol dehydrogenase [NADP+] A	46	16	16	2.2	0.0021	42	14	14	1.7	0.1187	52	17	17	2.5	0.0182
G3QBK7	Fructose-bisphosphate aldolase B	49	19	15	2.1	0.0042	63	22	19	4.1	0.0342	60	22	18	0.9	0.9193

G3NY35	Fructose-bisphosphate aldolase A	30	14	1	9.3	0.0501	38	16	2	3.4	0.1190	45	17	5	54.1	0.2147
TMAO metabolism																
G3PP04	Dimethylaniline monooxygenase [N-oxide-forming]	28	15	15	0.0	0.0006	35	14	14	0.0	0.0158	16	8	8	0.1	0.0101
Urea cycle																
G3P780	Arginase 1	20	6	5	0.3	0.0367	25	8	8	0.2	0.0111	3	1	1	2.6	0.0796
G3NU76	N[G]N[G]-dimethylarginine dimethylaminohydrolase 1	54	12	12	0.4	0.0108	65	15	15	0.3	0.0026	46	9	9	0.7	0.2078
Cytoskeleton structure																
G3N5N8	Coactosin-like protein	61	8	8	2.7	0.0032	75	12	12	2.6	0.0390	62	10	10	3.0	0.0096
G3PI97	Filamin-A	23	22	17	3.0	0.0211	25	22	17	1.5	0.1591	23	22	17	1.8	0.0050
G3NRF8	Transgelin	87	20	20	2.4	0.0077	92	24	23	1.9	0.2044	92	22	21	2.1	0.0347
G3PXD9	PDZ and LIM domain protein 3	35	10	10	4.4	0.0069	23	6	6	3.8	0.0088	26	6	6	7.8	0.0057
G3PWK6	Actin-related protein 2/3 complex subunit 1B	28	8	8	3.0	0.0000	26	7	7	3.0	0.0360	28	8	7	2.5	0.0426
G3NWQ0	Actin-related protein 2/3 complex subunit 5	46	5	5	3.4	0.0001	31	4	4	3.5	0.0073	46	4	4	2.6	0.0244
G3Q087	Actin-related protein 2/3 complex subunit 3	24	4	4	1.6	0.0077	24	4	4	2.6	0.0107	24	4	4	3.8	0.0222
G3P2I6	Myosin light chain 6	96	14	3	2.0	0.0346	96	15	12	1.6	0.2079	96	14	11	2.5	0.0187
G3NZN3	Myosin light chain 9	40	7	1	2.9	0.0002	44	7	2	1.5	0.2335	33	6	1	4.0	0.0110
G3PBF5	Tropomyosin alpha-4 chain	64	22	4	2.3	0.0434	52	13	2	1.6	0.5394	50	16	3	4.0	0.0316
G3Q2Y3	Desmoplakin-3 / Junction plakoglobin	9	6	5	2.1	0.0000	21	11	11	1.1	0.6837	13	7	7	1.8	0.0379
G3Q946	Src substrate protein p85	22	9	9	3.9	0.0000	12	5	5	4.2	0.0500	22	7	7	15.0	0.0211
G3PJ29	Rho GDP-dissociation inhibitor 1	70	15	15	2.0	0.0017	76	17	17	1.8	0.0205	74	15	15	2.0	0.0599
G3N821	Profilin-1	78	7	7	4.5	0.0013	85	10	10	3.8	0.0027	78	7	7	7.4	0.0180
G3P3D5	Profilin-2	58	8	4	2.9	0.0402	90	13	6	1.8	0.0049	89	11	4	1.1	0.8302
G3PK01	Coronin 1A	44	18	17	2.3	0.0027	44	18	17	3.0	0.0077	42	17	16	5.6	0.0408
G3PJZ8																
Energy metabolism																
G3NLT5	Creatine kinase B-type	50	18	16	6.0	0.0000	55	20	14	3.4	0.0030	48	17	13	6.4	0.0050
G3NVC1	Creatine kinase U-type	67	21	2	11.8	0.0048	70	26	2	11.5	0.0137	33	11	7	4.7	0.0427
G3NVD0																
Proteolysis																
G3PFE4	Apoptosis-associated speck-like protein containing a CARD	16	2	2	891.5	0.0003	20	3	2	171.6	0.0026	23	3	2	769.9	0.0650
G3P6H8	Caspase 1 / Interleukin 1 beta convertase	2	1	1	109.5	0.0037	5	2	2	68.4	0.0148	5	2	2		LS only
G3PTU9	Programmed cell death protein 5	30	3	3	42.2	0.0128	10	1	1	3.8	0.1128	30	3	3		LS only
G3NYX9	Plastin-2	38	17	13	3.1	0.0644	38	20	15	2.3	0.0528	42	20	16	4.7	0.0048
G3NA52	RAC2 / Ras-related C3 botulinum toxin substrate 2	14	3	2	4.3	0.0001	18	4	2	4.2	0.0284	23	5	3	17.2	0.0584
G3PKE0	Galectin-8	42	9	9	2.2	0.0019	47	11	10	2.5	0.0269	37	9	9	4.5	0.0187
G3NIT2	Galectin-3	28	6	6	4.1	0.0199	11	2	2	5.5	0.0134	28	4	4	17.7	0.0121
G3Q6I1	Beta-2-microglobulin	13	2	2	22.9	0.0357	19	3	3	509.6	0.0158	19	3	3		LS only
G3NZ72	Dipeptidyl peptidase 1	18	6	6	3.9	0.0041	14	4	4	4.9	0.0088	18	5	5	4.4	0.0053
G3P148	Tripeptidyl-peptidase 1	22	7	7	2.7	0.0062	17	5	5	1.6	0.1751	12	3	3	5.7	0.0073
G3PUV2	Cathepsin B	22	6	6	2.9	0.0129	25	6	6	2.6	0.0215	19	5	5	5.6	0.0106
G3PJR0	Cathepsin S	27	6	6	2.1	0.0498	13	3	3	7.8	0.0137	21	4	4	13.6	0.0079
G3PYU1	Cathepsin D	30	9	9	1.7	0.1216	28	9	9	1.6	0.0367	24	7	7	4.4	0.0387
G3NC06	Proteasome activator complex subunit 1 /	19	4	4	5.9	0.0106	32	7	7	2.3	0.2684	8	2	2	7.3	0.0164
G3NC03	Proteasome activator 28 alpha															
G3NA57	Proteasome activator complex subunit 2 /	43	10	10	1.6	0.0295	53	12	12	2.4	0.0476	35	8	8	8.3	0.0217
	Proteasome activator 28 beta															
G3NQQ4	Proteasome subunit alpha	13	3	1	36.3	0.0077	12	3	1	6.9	0.0482	23	5	3	25.7	0.0094
G3NQQ6																
G3QA17	Heat shock 27 kDa protein beta-1	12	2	2	2.7	0.0186	12	2	2	3.2	0.0371	10	2	2	4.1	0.1621
G3PS82	Gamma-interferon-inducible lysosomal thiol reductase	24	3	3	6.7	0.0392	24	3	3	3.6	0.0510	24	3	3	26.2	0.0301
G3N4S6	Alpha-mannosidase lysosomal	7	6	6	2.7	0.0220	5	4	4	4.0	0.0197	8	9	9	0.1	0.0754
G3N704	Nephrosin 1	46	12	1	0.0	0.0462	50	14	1	0.0	BH only	46	11	1	0.5	0.4282
G3PJB8	Alpha-N-acetylgalactosaminidase	9	3	3	0.3	0.0007	10	4	4	0.2	0.0000	5	2	2	0.4	0.0319
Calcium signaling																
G3PEU7	Parvalbumin thymic CPV3	60	8	7	54.1	0.0011	60	6	5	114.5	0.1168	42	5	4	332.4	0.0261

G3PET2	Parvalbumin beta	47	5	4	5.7	0.1452	53	6	5	3.9	0.0035	59	6	5	99.0	0.2722
G3NDQ7	EPCAM Tumor-associated calcium signal transducer	23	7	7	4.2	0.0024	33	12	10	2.4	0.0543	21	6	6	4.8	0.0056
G3NJM9	Caldesmon 1	18	8	8	5.4	0.0031	15	6	6	2.7	0.18231	9	3	3	11.4	0.0230
G3P5I1	Caldesmon 3	27	3	1	10.1	0.0165						33	3	1	6.6	0.0213
G3PE15	Protein S100-A11 / Calgizzarin	22	3	2	3.4	0.0016	32	4	3	2.2	0.0596	15	2	2	5.4	0.0220
G3NGI8	Protein S 100-A4 / Ictacalcin	40	4	4	2.3	0.0329	35	4	4	4.6	0.1481	45	6	6	3.2	0.2535
G3NGJ4	Protein S100-A1	25	2	2	2.7	0.0155	28	3	3	3.6	0.0359	25	2	2	10.9	0.0494
Osmoregulation																
G3PKB5	Sodium/potassium-transporting ATPase beta 233	22	7	4	4.3	0.0057	16	6	3	3.3	0.0082	16	5	3	6.9	0.0007
Signal transduction																
G3Q0J6	Desmoykin	26	34	16	2.7	0.0007	20	26	7	1.3	0.2088	24	29	7	2.3	0.0246
G3NGG6	RNA-binding protein 39	4	2	1	LS only							4	2	1	249.8	0.0024
G3PVS6	Heterogeneous nuclear ribonucleoprotein A0	31	10	10	2.7	0.0108	31	9	9	1.4	0.3424	27	8	8	1.9	0.0395
G3PVY2	High mobility group protein B2	31	10	9	2.3	0.0030	31	7	7	1.3	0.3861	29	10	8	3.9	0.0208
G3PWB7	Transcription factor BTF3 / Basal transcription factor 3	31	3	3	3.0	0.0086	36	4	4	7.1	0.0161	31	4	3	39.2	0.0024
G3PZ17	FYN-binding protein	3	1	1	3.4	0.0001	3	1	1	3.4	0.0023	6	2	2	5.8	0.0212
G3PAN1	Serine/threonine-protein kinase PAK 2	3	1	1	3.5	0.0212	8	3	1	1.9	0.0167	3	1	1	4.8	0.0105
G3PZP5	GTP:AMP phosphotransferase	70	15	15	1.6	0.0328	45	10	10	1.1	0.8267	70	8	8	3.8	0.0084
G3PWCS	Adenylate kinase 1	43	9	5	2.6	0.0021	42	8	7	1.7	0.0254	53	10	4	2.7	0.0161
G3NVX9	Adenylate kinase 4	52	8	8	0.2	0.0029	55	10	10	0.3	0.0020	31	5	5	0.4	0.1272
G3NYE2	Casein kinase II subunit alpha	3	1	1	0.3	0.0051	3	1	1	0.3	0.0245	3	1	1	0.8	0.3266
G3NRD6	Protein NDRG1, N-myc downstream-regulated 1	24	7	7	0.5	0.0229	16	4	4	0.4	0.0376	21	5	5	0.6	0.0754
G3N9S3	Mesencephalic astrocyte-derived neurotrophic factor	32	6	6	0.1	0.3154	12	3	3	0.4	0.5442	34	7	7	0.2	0.0002
Lipid/ fatty acid metabolism																
G3Q2N0	ES1 protein homolog	65	15	15	54.0	0.0032	65	14	14	9.4	0.0156	62	14	14	35.5	0.0030
G3N9H8	Acyl-CoA-binding protein	70	8	7	2.7	0.0245	91	10	9	1.6	0.0609	70	8	7	2.4	0.0289
G3P9E7	Acyl-CoA synthetase 2						35	17	4	0.1	0.0224		21	5	0.1	0.0137
G3PYY3	Phosphotriesterase-related protein	54	16	16	0.7	0.0232	67	18	18	0.3	0.0006	32	9	9	0.6	0.0576
G3Q664	Dehydrogenase/reductase SDR family member 4	18	3	3	0.2	0.0350	14	3	3	0.1	0.0124					
Mitochondrial proteins																
G3Q4U3	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 7	19	3	3	4.9	0.0069	19	2	2	0.8	0.6422	19	2	2	LS only	
G3PCY9	Cytochrome c	68	9	8	2.2	0.0164	69	11	10	1.9	0.0689	61	8	7	3.1	0.0390
G3P173	Cytochrome c oxidase subunit 6B	56	6	6	2.8	0.0093	56	5	5	0.6	0.2156	45	4	4	21.9	0.0106
G3PP79	Cytochrome c oxidase subunit 5A	63	9	9	1.6	0.0104	67	10	9	1.0	0.8680	63	9	8	2.7	0.0073
G3Q298	Sideroflexin-1	22	5	5	0.3	0.0009	21	5	5	0.1	0.0152	16	4	4	0.3	0.0068
Membrane trafficking																
G3PV95	Protein NipSnap homolog 3A	18	4	4	0.3	0.0307	36	8	8	0.3	0.0001	7	2	2	0.8	0.3253
G3NSE0	Lipopolysaccharide-binding protein	5	2	2	13.9	0.0086	2	1	1	8.0	0.0561	5	2	2	24.4	0.0075
G3P331	Synaptic vesicle membrane protein VAT-1 homolog	12	3	3	2.5	0.0196	18	6	6	13.0	0.2312	18	5	5	1.4	0.7405
Immunity/ Pattern recognition																
G3PH18	Ig mu chain C region membrane-bound form	11	6	1	5.2	0.0153	14	8	1	2.5	0.2272	11	5	1	LS only	LS only
G3N4E3	Major histocompatibility complex class I-related gene protein	12	4	4	55.8	0.0018	7	3	1	LS only		9	3	3	LS only	

1080

1081 Proteins are grouped according to their predominant functions. Results for all three experiments
1082 are shown; Cov. = sequence coverage in %, Pep. = number of identified peptides; Uniq. =
1083 number of identified peptides that are unique for the particular protein, LS/BH = Ratio of mean
1084 protein abundance in Lake Solano versus Bodega Harbor fish (n=6 for each population), sign. P
1085 = t-test P value.

Figure 1

Figure 2

Figure 3

Figure 4

Figure 5

Figure 6

Figure 7