

Open Access

Identification of Stress Related Molecular Biomarkers in Zebrafish Employing an In-Silico Approach to Access Toxicity based Risks in Aquaculture

Md Mursalin Khan*, Ashiqul Islam SM and Ahmed Mustafa

Department of Biology, Indiana-Purdue University, Fort Wayne, IN 46805, USA

Abstract

Toxicity induced stress on fishes has recently become a major concern in aquaculture. Monitoring risk factors in a routine manner can save aqua-culturists from a substantial amount of losses. Meta-analysis gives a successful way for high-throughput screening of stress inductive chemicals in fish habitats by firmly exploiting molecular biomarkers. Previously reported meta-analysis revealed Top-21 proteins that are frequently related to stress-induced proteome profiles. We performed a further computational analysis on specific functional categorization, proteinprotein interaction (PPI) network and co-expression profiles for these Top-21 proteins in search of more consistent and robust molecular biomarkers. The analysis identified functional class specific molecular biomarkers: for heat shock proteins the proposed biomarkers are hsp90b1, hsp90ab1, hspa5, hspa8, hspa9, and ptges3a; for cytoskeletal proteins the proposed biomarkers are actb1, myhz2 and zgc:55461, rac1, tuba1/2, tuba1, and zgc:136799; for keratin proteins the proposed biomarkers are Krt 4, Krt5, krt23, krt17, sc:d0144, cldnb, icn2, and pkp3; for oxidative damage defense proteins the proposed biomarkers are txn, prdx5, adh2a and prdx2. These predicted molecular biomarkers can be considered potential candidates for identifying the stressors present in fish habitats since the biomarkers are integral to stress modulation. In addition, as a consequence of the study, the reported proteins can also be applied as stress regulators in fish. Our study will help researchers to find suitable proteins for stress biomarkers and regulators. Further wet lab validation of those proposed proteins as biomarkers and regulators will be useful to lead to practical implementation of our proposal.

Keywords: Aquaculture; Stress; Zebrafish; Molecular biomarkers; Protein-protein network

Abbreviation of Proteins: hsp90b1: Heat Shock Protein 90, Beta, Member 1; hsp90ab1: Heat Shock Protein 90, Alpha, Class B Member 1; hspa5: Heat Shock Protein 5; hspa8: Heat Shock Protein 8; hspa9: Heat Shock Protein 9; ptges3a: Prostaglandin E Synthase 3a; actb1: Actin, Beta 1; myhz2: Myosin, Heavy Polypeptide 2; zgc: 55461 Tubulinlike; rac1: Ras-related C3 Botulinum Toxin Substrate 1; tuba1/2: Tubulin, Alpha 1, like 2; tuba1: Tubulin, Alpha 1; zgc: 136799 Ras-related C3 Botulinum Toxin Substrate 1-like; Krt 4: Keratin 4; Krt5: Keratin 5; krt23: Keratin 23; krt17: Keratin 17; sc: d0144 Glycoprotein A33; cldnb: Claudin b; icn2: Ictacalcin 2; pkp3: Plakophilin 3; txn: Thioredoxin; prdx5: Peroxiredoxin 5; adh2a: Aldehyde Dehydrogenase 2 Family a; prdx2: Peroxiredoxin 2

Introduction

The demand for high protein foods is rising rapidly with the increasing world population [1]. A push for healthier sources of protein in industrialized nations is increasing the demand as well. Sources of protein such as cattle and other livestock are not adequate to meet this necessity due to a global contraction of grazing area, hostile climate, extensive harvesting time and other limitations [2]. Wild fish populations are also not large enough to meet this demand, and many populations are showing signs of overharvesting [3]. As a result, prevalence and importance of aquaculture is rising. Aquaculture provides a high protein food that requires less space to farm, a faster harvest time, lower cost to produce and is less detrimental to the environment [4]. Meanwhile, ecotoxicology is also becoming a major concern in farming industries nowadays. The steady expansion of bioactive chemical compounds at the industrial level is achieving successes in synthetic industries, but these chemicals are a growing threat to farming industries. Aquaculture based fish farming is highly susceptible to the toxicity-related issues raised from abundance and emission of toxic chemicals in the environment. The problem is compounded by current aquaculture techniques that create a significant number of problems of stress leading to physiological and immunological changes. Fish, in aquaculture, are susceptible to stressors that over time lead to a reduction in immune response and an increase in susceptibility to diseases [5].

An early detection of chemical components having potential toxic effects on fish could be effective so that the necessary steps can be taken against toxicity related problems. However, with their increasing number, it is becoming difficult to measure toxicity in fish. Moreover, proper risk assessment should consider the effects of a combination of chemical compounds as well. On the other hand, the paucity of toxicological data is making the risk assessment more difficult [6,7]. A tendency of the scientific society towards performing less toxicity test on the animals for an ethical reason is amplifying the challenge to reveal the risks of the compounds. A genomic approach to identify a toxic or stress condition using microarray or RNA-seq technology could be a successful alternative of phenotypic toxicology study [8,9].

Toxico-genomics assume that physiological changes due to toxic effect also reflect in the molecular level of the exposed organism. Accordingly, a proper understanding of the changes in molecular level is potentially an alternative way to distinguish toxic or stress related

*Corresponding author: Md Mursalin Khan, Department of Biology, Indiana-Purdue University, Fort Wayne, IN 46805, USA, Tel: 260-4816325; E-mail: khanm05@ipfw.edu

Received July 25, 2015; Accepted August 21, 2015; Published August 28, 2015

Citation: Khan M, Ashiqul Islam SM, Mustafa A (2015) Identification of Stress Related Molecular Biomarkers in Zebrafish Employing an *In-Silico* Approach to Access Toxicity based Risks in Aquaculture. Poult Fish Wildl Sci 3: 137. doi:10.4172/2375-446X.1000137

Copyright: © 2015 Khan M, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

conditions in fish farming. However, transcriptomic studies such as microarray or next generation sequencing dependence may not be appropriate to understand meaningful changes in the molecular level [10-12]. The above mentioned approaches merely depend on the mRNA expression level in the exposed organisms. The difference between protein levels and mRNA expression profile can be as much as 56% to 81% [13]. Therefore, a dedicated proteomic analysis is more preferred to investigate phenotypic changes due to changes in of environmental conditions [14,15]. A typical global proteomic experiment is conducted to compare a control and a sample condition [16]. Two-dimensional gel electrophoresis (2D-GE) [16,17] or Mass spectrometry [18] techniques are generally used to analyze the proteins after treatment, but these proteomic approaches are limited to the identification of the set of proteins that were present in the sampling moment only [19]. Metaanalysis of differential proteomics studies is becoming quite promising to solve this type of problem [20-22]. Furthermore, meta-analysis was used to find Top-21 stress related proteins from zebrafish, and those can be used as molecular biomarkers to access a toxic condition in zebrafish [23]. Certainly, the elevation of these Top-21 proteins can be used to predict cellular stresses in zebrafish due to a potential toxic condition. However, some of the members from the Top-21 protein families, for instance hsp90, occurs not only during stress response but also in a multitude of cellular function [24,25]. Therefore, it is imperative to investigate the stress induced network, protein-protein interaction and co-expressions of the top-21 protein in the orthologs to find out robust and consistent stress specific molecular biomarkers.

To identify the molecular biomarkers of stress, meta-analysis was performed for zebrafish, Danio rerio, from which, Top-21 stress-related proteins have been identified [23]. We have tried to analyze those Top-21 proteins and their interaction networks via different In-Silico approaches to find out the potential molecular biomarkers for stress. Specifically, our goals were to identify specific biomarkers in those Top-21 proteins and to propose new biomarkers for the stress as well as stress regulator proteins. To achieve these goals, we have (1) analyzed multiple sequence of stress proteins to generate multiple sequences alignment (MSA) and phylogeny; (2) visualized the distribution of those proteins using sub-cellular localization; (3) classified functional proteins using domain based analysis (4) developed a PPI network to show their activation, inhibition, binding, phenotype, catalysis, posttranslational modification, reaction, and expression; and (5) identified their co-expressions with each other (6) developed a PPI network and generated their co-expressions using orthologs based COGs mode to visualize the interaction and expression in other species, and (7) calculated the homology-based conserveness of those proteins in other organisms. Finally, using all those In-Silico approach we proposed potential molecular biomarker as well as candidates for stress regulation of zebrafish.

Methods

Sequence extraction

SeqinR package of the R 3.2.1 program was used to retrieve sequences from National Center for Biotechnology Information (NCBI) database [26]. To extract the sequences and their information from NCBI, function "getSequence" of SeqinR package of R 3.2.1 was used. For database selection function "choosebank" was applied and NCBI "genbank" was selected for sequence extraction. To get the amino acid sequence in fasta format function "as.SeqFastaAA" was used. For Page 2 of 8

the annotation of the sequences "getAnnot" function was applied.

Multiple sequence analysis

Sequences of Top-21 proteins were subjected to multiple sequence alignment using "clustal" function of SeqinR package of R 3.2.1 [26,27]. "clustal" function gave the CLUSTAL format (*.aln) output which is the format of the ClustalW multi-alignment tool. Multiple alignment visualizes the amino acid similarity among query sequences of on the basis of homology. To read the inferred homology of the sequences "read.alignment" function of SeqinR package was used.

Phylogenetic tree construction

The MEGA6 software was used for determining evolutionary model, phylogenetic tree construction and estimation of evolutionary divergence [28]. The CLUSTAL format (*.aln) of the multiple sequence alignment was applied to create the phylogeny. The phylogeny tree option of MEGA6 was used to construct a neighbor-joining (NJ) tree [29]. With the NJ method, the S value was not computed for all or many topologies. The examination of different topologies was imbedded in the algorithm. MEGA6 displayed NJ trees in a manner similar to rooted trees [29]. For constructing the NJ tree, MEGA6 was commanded that it specify the distance estimation method and subset of sites which need to include.

Sub-cellular localization

To identify the localization of a protein in cellular compartments CELLO2GO predictor was used. CELLO combines SVM-based prediction and homology search to predict localization for both Gram-positive, Gram-negative bacteria, Archaea and Eukaryotes [30]. Eukaryotes option along with E-value 0.001 was selected to understand the subcellular distribution of Top-21 proteins. CELLO2GO predictor gave individual score for each subcellular location and the highest scored location was chosen as final localization.

Domain based functional classification

Top-21 proteins were classified on the basis of their associated biological processes, cellular components, and molecular functions. Gene Ontology (GO) [31] and InterPro [32] were used for the domain based classification for functional group determination. Depending on functional domains and the clusters of PPI network, Top-21 proteins were divided in six clusters.

Protein-protein network, co-expression, and conservation

To determine the interaction between proteins, the proteinprotein interaction network (PPI) was created. STRING10 was used for PPI networking, co-expressions, and determination of conservation [33]. STRING 10 show interactions among proteins along with their activation, inhibition, binding, phenotype, catalysis, post translational modification, reaction, and expression. STRING database uses a combination of prediction approaches and an integration of other information (neighborhood, transferred neighborhood, gene fusion, text mining, databases, homology transfer, co-occurrence, experiments and so on). STRING PPIs come from a mix of experimental data; PPIs copied from public databases (e.g. KEGG and BioGRID) and predicted PPIs [33]. Protein-protein network and Co-expressions were also created with orthologs of Top-21 proteins using COGs mode of STRING10. The interaction networks for proteins were merged using Advanced Network Merge plug-in in the "Cytoscape" platform [34].

The abbreviation Top-21 will be used throughout the article instead

of the "Top-21 reported stressor induced protein in zebrafish identified by meta-analysis".

Results and Discussions

In our analyses, we tried to find out the candidate proteins for the stress related molecular biomarkers to identify the level and type of stressors according to their expression. The identified molecular biomarkers may be used for stress modulation by regulating their expression in cellular level. Primarily, we used Top-21 zebrafish stress proteins for biomarker selection as reported by meta-analysis of different expression profiling experiments (Supplement Table 1) [23].

Top-21 proteomic analyses

We retrieved 21 protein sequences from NCBI protein data base, based on the meta-analyses of Top-21 stress proteins of zebrafish reported by Groh and Sutter in various expression profiles which showed alteration in the presence of stressors [23]. We used the multiple sequence alignment tool ClustalW in order to find the relationship among the sequences Top-21 stress proteins. The relationship found among the Top-21 sequences was significant. To further analyze the sequences for evolutionary distance, we used the MSA data of proteins as inputs for MEGA6.0 software. We constructed the Neighbor Joining phylogenetic tree to identify the distance among 21 protein sequences. In the phylogeny, we found most of the proteins are closely related in the tree (Figure 1). In Figure 1, the horizontal dimension gives the amount of genetic change. The horizontal lines represent evolutionary lineages changing over time. The longer the branch, the larger the amount of change in the horizontal dimension. The bar at the bottom provides a scale of "0.2". In Figure 1 the line segment with the number shows the length of branch that represents amount of genetic changes. Some pairs of sequences are noticeably related in both MSA and phylogeny so that their interactions among them can be predicted. Our analyses showed that related proteins are in proximal branches and similar functional proteins are in nearby leaves of the tree.

Sub-cellular localization

To visualize the subcellular localization of Top-21 proteins, we used the CELLO2GO system for screening various properties of a targeted protein and their subcellular localization of the queried proteins by combining the CELLO localization-predicting and BLAST homologysearching approaches [30]. Given our query protein sequence, CELLO2GO uses BLAST to search for homologous sequences that are Gene Ontology (GO) annotated in an in-house database derived from the UniProt Knowledge Base database [30,31]. Our analyses show that out of 21 proteins, 11 proteins are cytoplasmic, 2 are nucleic, 4 are mitochondrial, 2 are cytoskeletal, and 1 each is endoplasmic reticular and extra-cellular, respectively. Diversified localization of those proteins indicates the diverse cellular function on different cellular organelles (Table 1).

Protein-protein network, co-expression, and protein conservation

In our analyses of the MSA we found some conserved sequences among Top-21 proteins. To further investigate in details we created the phylogenetic tree which showed evolutionary relations among Top-21 stressor proteins. The subcellular localization showed the divergence of the localization which gave the idea about functional diversity as well. To analyze the functional relationship of those proteins, we used PPI network based on the Neighborhood, Gene Fusion, Cooccurrence, Coexpression, Experiments Databases, and Text-mining. In PPI network of Top-21 proteins we found crucial interactions among them (Figure 2). These interactions showed some clusters of proteins such as, actin, tubulin, and keratin in one cluster and hspa5, hspa8, hspa9, and hsp90ab1 in another cluster. GO and InterPro based protein sequence analysis and classification showed that those clustering proteins also have similar function (Table 1).

After observing the clusters of Top-21 PPI-Network we wanted to see the co-expression of those proteins. In co-expression analysis for zebrafish, we observed much less expressional relationship among Top-



Page 4 of 8



Rank position	Protein ID	Protein name	Subcellular Localization	Intetpro ID
1	NP001103373	Heat shock protein 8	Cytoplasmic	IPRO13126
2	NP571106	Actin. cytoplasmic 1 (beta-actin. bactin 1)	Cytoplasmic	IPRO04000
3-8	NP998223	Heat shock protein	ER	IPR013126
3-8	NP571385	Heat shock protein 90B	Cytoplasmic	IPR001404
3-8	NP848524	Keratin 18	Cytoplasmic	1PR002957
3-8	NP001015057	Transferrin alpha	IL-	IPR016357
3-8	NP571180	Tropomyosin alpha	Cytoskeletal	1PR000533
3-8	NP001038362	Viteilogenin 1	Mitochondrial	IPR015816
9-21	NP957346	40S ribosomal protein SA	Cytoplasmic	IPR027504
9-21	NP998466.2*	Aldehyde dehydrogenase 2B	Mitochondrial	1PR016161
9-21	NP861423	Annexin IA	Cytoplasmic	IPR001464
9-21	NP001019600	ATP synthase. beta polypeptide 5	Mitochondrial	IPR005722
9-21	NP999888	Enolase 3	Cytoplasmic	IPR000941
9-21	NP958483	Heat shock protein 9	Mitochondrial	IPR013126
9-21	NP571231	Keratin 5	Nuclear	EPR003054
9-21	NP956374	Keratin 8	Nuclear	1PR003054
9-21	NP694514	Myosin. hezny polypeptide 1.2	Cytoskeletal	1PRO04009
9-21	NP571002	Nucleoside diphosphate kinase B	Cytoplasmic	IPR001564
9-21	NP001002468	Peroxiredoxin	Cytoplasmic	IPR024706
9-21	NP705954	Triosephosphate isomerase IB	Cytoplasmic	IPR000652
9-21	NP998655	Tubulin, beta 2C	Cytoplasmic	IPR000217

(i totelli iD. changed in NODI)

Source: This table has been generated based on the finding of Groh and Sutter [19].

 Table 1: Subcellular localization of Top-21 proteins of Zebrafish found in meta-analysis of stress proteins.

21 proteins. For NP957346 protein (rpsa), we added identical proteins to get the co-expression according to domain analysis of InterPro. The co-expression analysis software used text-mining and experimental databases to determine the expressional relationship so it might happen due to limited number of data about zebrafish proteins. We therefore used "COG-mode" in which predicted interactions are propagated to proteins in other organisms for which interaction has been described by inference of orthology [35]. In the STRING10 based COGs analysis, we found a higher number of evidence-based interactions for COGs based PPI Network of Top-21 proteins (Supplement Table 2) (Figure 3). We also found a significant change in co-expression of proteins compared to those of zebrafish, which showed a much higher association score

for their orthologs (Figure 4). Hence, COGs based PPI network and co-expression analysis proved that Top-21 proteins actually correlated with each other in stress condition so they can be used as the molecular biomarkers to identify the levels of stress and their types.

Because of these expressions and interactions of zebrafish Top-21 proteins with other organisms, we tried to see the conservation of those proteins using homology based conservation study (Figure 5). The results showed that hspa5, hspa9, prdx2, tpi1b, ATP5B, nme2b.1, aldh2b, hspa8, and eno3 proteins are conserved in all prokaryotic and eukaryotic organisms while krt18, tfa, tpma, vtg1, actb1, zgc:55461, krt8, anax1a, myh22, krt5, and rpsa proteins are conserved for



Figure 3: Ortholog based COGs mode generated Top-21 PPI-network demonstrating evidence.



eukaryotes only.

Based on the InterPro protein functional analysis and Top21-PPI-network clusters we identified some stress protein functional groups. We divided them into six functional groups: 1) heat shock, 2) cytoskeleton, 3) Keratin, 4) oxidative damage defense, 5) energy metabolism, and 6) miscellaneous proteins. Later, we tried to identify the group-wise molecular biomarkers and stress regulatory targets using proteomic approaches. For the interaction and expression based detection of the molecular biomarkers and stress regulator targets we used PPI network, phylogeny, and co-expression. For each group of proteins we added their closely related interactive proteins to draw out the total interactive networks. Finally, we proposed the detailed PPI networks and co-expression analysis for each functional groups along with their interactive proteins.

Identification of molecular biomarker and regulator based on functional cluster

According to the action based PPI networks of Top-21 Heat Shock Proteins (HSP) - hsp90b1, hsp90ab1, hspa5, hspa8, and hspa9 are central regulators (Figure 6) (Supplement Table 3). We used all Top-21 HSP-PPI network proteins to observe their phylogenetic relation. In the phylogeny, we found most interactive proteins are in the proximal branch of the tree (Supplement Figure 1). Interestingly, in the nearby leaves of hsp90b1 we found ptges3a protein which is not located in the central regulatory position of the Top-21 HSP-PPI network. So we assumed that ptges3a may have a similar role in stress regulation. When we analyze the co-expressions of Top-21 HSP-PPI network proteins, ptges3a showed the co-expression with some proteins of zebrafish, but in other organisms it showed expression with several other proteins like hspa9, hspa90ab1, sgut1, hsp90b1, FP236598.1, stip1, TTC31, ttc12 and zgc:123010 with a higher association score (Figure 7). This supports our assumption of its role in stress modulation. Hence, we identified ptges3a as a stress biomarker protein.

In case of cytoskeletal proteins we found actin, tubulin, and myosin which play vital roles in stress modulation [36]. According to Top-21 Cytoskeleton Protein-PPI network (Figure 8) (Supplement Table 4) actb1, myhz2 and zgc: 55461 (tubulin like), rac1, tuba1/2, tuba1, and zgc: 136799 are crucial in stress modulation. Phylogeny and co-expression also provided the basis of their functional relationships and expressions (Supplement Figures 2 and 3) to select those as molecular biomarker and regulators of stress.

Keratin cluster analysis presented an interesting observation. Unlike heat shock and cytoskeletal proteins in Top-21 Keratin-PPI network (Figure 9) (Supplement Table 5) top reported keratins proteins Krt18 and krt8 were not located in the central regulatory position. However, other keratin proteins of the network such as krt4 and krt23 (zgc: 92533 and zgc: 92061) showed higher interactions. Along with Top-21 Keratin-PPI network, co-expression and phylogeny results for Top-21 Keratin-PPI network also showed that Krt5, krt23 (zgc:92533











and zgc:92061), krt17, sc:d0144, cldnb, icn2, and pkp3 are potential biomarkers and regulators for stress modulation (Supplement Figures 4 and 5).

Oxidative stress is one of the most deleterious stress for fish. Reactive oxygen species destroy homeostasis of organisms and cause disease [37]. Two kinds of oxidative defense proteins are found in the Top-21 reported stress proteins- Aldehyde dehydrogenase 2B (aldh2b) and Peroxiredoxin 2 (prdx2). In Top-21 Oxidative Defense Protein -PPI network (Figure 10) (Supplement Table 6) of these two proteins along with their proximal interactive proteins, we observed that prdx2 is interactive with six more proteins sod1,txn, gpx1b, ar, myca and urod but no interaction mostly reported aldh2b. On the other hand Aldehyde dehydrogenase 2B (aldh2b) interacts with some other Aldehyde dehydrogenase proteins, but does not interact with prdx2. Therefore, aldh2a and prdx5 are the connectors between the two sub clusters of the network. In the Top-21 Oxidative Defense Protein -PPI network txn, prdx5, and adh2a along with prdx2 are seen to be more interactive while aldh2b is less for regulating stress even though it is highly expressed in expression profiling. Phylogenetic trees also support this finding by showing aldh2b, aldh2a and aldh2l proteins in same internal node which is away from other oxidative damage defense proteins (Supplement Figure 6). Co-expression of the Top-21 Oxidative Defense Protein -PPI network proteins further supported the evidence that aldh is not expressed with other oxidative damage defense proteins (Supplement Figure 7). Therefore, it is clear that aldh and prdx group work in a separate manner to modulate stress.

Energy metabolism proteins are housekeeping proteins for most cells [38]. It is common to observe many energy metabolism proteins in the expression profiling. Meta-analysis of the expression profiling





of zebrafish stressor induced proteins provided a number of energy metabolism proteins such as ATP synthase 5(ATP5B), enolase 3(eno3) and Triosephosphate isomerase 1B (tpi1b). The Energy metabolism-PPI-network showed promising interactions among the ATP synthases but their roles in the cell as housekeeping proteins prevent us from making a solid conclusion on the basis of expression profiles (Figure 11) (Supplement Table 7).

Conclusion

In our work we identified biomarkers for stress by using *In-Silico* analysis of the Top-21 reported stressor induced proteins of zebrafish found in the expression profile. We found out new potential candidate proteins for molecular biomarkers of stress to determine the levels and the types of stress. The molecular biomarkers are also candidates for stress modulation by up or down regulation of their protein expression at the cellular level. Our study will help researchers identify suitable proteins for stress biomarkers and regulators. Further, wet lab experiments will be needed to validate those proposed proteins as biomarkers and regulators.

References

- Shannon R, Mustafa A (2015) Stress susceptibility of echinoderms under aquaculture conditions. Global Aquaculture Alliance Magazine- The Advocate 26-27.
- Ibrahim M, Khan M, Mustafa A (2015) Can saline aquaponics be an alternative method for sea products? Proceedings of the Marine Conservation and Blue Economy Symposium, Dhaka, Bangladesh 8: 62-63.
- Mustafa A, Randolph L, Dhawale S (2013) Effects of Nutritional Supplements Phosphatidylcholine and Beta-Carotene on Growth and Selected Stress and Immune Parameters in Nile Tilapia, *Oreochromis niloticus* (L.). Advances in Zoology and Botany 1: 57-61.
- Hough T, Qasim M, Glaze C, Blumenthal E, Mustafa A (2015) Effects of Omega-3 Fatty Acid Supplementation on Aquaponics-System Raised Tilapia. Bioengineering and Biosciences (In Print).
- Hossain MR, Mustafa A (2014) Effects of puerarin on the modulation of health in chronically stressed Chinook salmon (*Oncorhynchus tshawytscha*). Journal of Applied Aquaculture 26: 119-131.
- Eggen RIL, Suter MJF (2007) Analytical chemistry and ecotoxicology -tasks, needs and trends. J Toxicol Environ Health A 70: 724-726.

- Eggen RIL, Behra R, Burkhardt-Holm P, Escher BI, Schweigert N (2004) Peer reviewed: Challenges in ecotoxicology. Environ Science and Technology 38: 58A-64A.
- 8. Boverhof DR, Zacharewski TR (2006) Toxicogenomics in risk assessment: Applications and needs. Toxicol Sci 89: 352-360.
- 9. Van AG, Ankley GT, Baldwin WS, Bearden DW, Benson WH, et al. (2010) Integrating omic technologies into aquatic ecological risk assessment and environmental monitoring: Hurdles, achievements, and future outlook. Environmental Health Perspective 118: 1-5.
- Link V, Carvalho L, Castanon I, Stockinger P, Shevchenko A (2006) Identification of regulators of germ layer morphogenesis using proteomics in zebrafish. Journal of Cell Science 119: 2073-2083.
- Wei Y, Chan LL, Wang D, Zhang H, Wang J, et al. (2008) Proteomic analysis of hepatic protein profiles in rare minnow (Gobiocypris rarus) exposed to perfluorooctanoic acid. Journal of Proteome Research 7: 1729-1739.
- Washburn MP, Koller A, Oshiro G, Ulaszek RR, Plouffe D, et al. (2003) Protein pathway and complex clustering of correlated mRNA and protein expression analyses in Saccharomyces cerevisiae. Proceedings of the National Academy of Science 100: 3107-3112.
- Li JJ, Bickel PJ, Biggin MD (2014) System wide analyses have underestimated protein abundances and the importance of transcription in mammals. Peer Journal 2: e270.
- Rees BB, Andacht T, Skripnikova E, Crawford DL (2011) Population proteomics: quantitative variation within and among populations in cardiac protein expression. Molecular Biology and Evolution 28: 1271-1279.
- Diz AP, Martínez-Fernández M, Rolán-Alvarez E (2012) Proteomics in evolutionary ecology: Linking the genotype with the phenotype. Molecular Ecology 21: 1060-1080.
- Görg A, Weiss W, Dunn MJ (2004) Current two-dimensional electrophoresis technology for proteomics. Proteomics 4: 3665-3685.
- Gygi SP, Corthals GL, Zhang Y, Rochon Y, Aebersold R (2000) Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. Proceedings of the National Academy of science USA 97: 9390-9395.
- Syka JEP, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF (2004) Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. Proceedings of the National Academy of science USA 101: 9528-9533.
- Groh KJ, Nesatyy VJ, Sutter MJF (2011) Proteomics for the analysis of environmental stress responses in prokaryotes. John Wiley & Sons, Hoboken 603-625.
- Chan SK, Griffith OL, Tai IT, Jones SJM (2008) Meta-analysis of colorectal cancer gene expression profiling studies identifies consistently reported candidate biomarkers. Cancer Epidemiol Biomark Cancer Epidemiol Biomarkers Prev 17: 543-552.
- Anderson K, Taylor DA, Thompson EL, Melwani AR, Nair SV, et al. (2015) Metaanalysis of studies using suppression subtractive hybridization and microarrays to investigate the effects of environmental stress on gene transcription in oysters. PLOS ONE 10: e0118839.
- Schwartz D, Chou MF, Church GM (2008) Predicting protein post-translational modifications using meta-analysis of proteome scale data sets. Mol Cell Proteomics 8: 365-379.
- 23. Groh KJ, Suter MJF (2015) Stressor-induced proteome alterations in zebrafish: A meta-analysis of response patterns. Aquat Toxicol 159: 1-12.
- Haslbeck V, Kaiser CJO, Richter K (2012) Hsp90 in non-mammalian metazoan model systems. Biochimica et Biophysica Acta 1823: 712-721.
- Kaplan KB, Li R (2012) A prescription for "stress"-the role of Hsp90 in genome stability and cellular adaptation. Trends in Cell Biology 22: 576-583.
- 26. Charif D, Lobry JR (2007) SeqinR 1.0-2: A Contributed Package to the R Project for Statistical Computing Devoted to Biological Sequences Retrieval and Analysis. Structural Approaches to Sequence Evolution, Biological and Medical Physics. Biomedical Engineering 207-232.
- 27. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673-4680.

Page 8 of 8

- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular Biology and Evolution 30: 2725-2729.
- Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4: 406-425.
- Yu CS, Cheng CW, Su WC, Chang KC, Huang Sw, et al, (2014) CELLO2GO: A Web Server for Protein subCELlular LOcalization Prediction with Functional Gene Ontology Annotation. PLoS ONE 9: 0099368.
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, et al. (2000) Gene ontology: tool for the unification of biology. Nature Genetics 25: 25-29.
- Hunter S, Apweiler R, Attwood TK, Bairoch A, Bateman A, et al. (2009) InterPro: the integrative protein signature database. Nucleic Acids Research 37: D211-D215.
- 33. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, et al, (2015)

STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Research 43 (Database issue): D447-D452.

- 34. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, et al. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Research 13: 2498-2504.
- Mering CV, Jensen LJ, Senel B, Hooper SD, Krupp M, et al. (2004) STRING: known and predicted protein–protein associations, integrated and transferred across organisms. Nucleic Acid Research 33: D433-D437.
- Tojkander S, Gateva G, Lappalainen P (2012) Actin stress fibers assembly, dynamics and biological roles. Journal of Cell Science 125: 1855-1864.
- Uttara B, Singh AV, Zamboni P, Mahajan, RT (2009) Oxidative Stress and Neurodegenerative Diseases: A Review of Upstream and Downstream Antioxidant Therapeutic Options. Current Neuropharmacology 7: 65-74.
- Eisenberg E, Levanon EY (2013) Human housekeeping genes, revisited. TRENDS in Genetics 29: 569-574.