Metabolomic strategies for aquaculture: A primer

Tim Young and Andrea C. Alfaro

Institute for Applied Ecology New Zealand, Auckland University of Technology, New Zealand

Abstract

Metabolomics is a fast-evolving field that provides qualitative and quantitative analyses of metabolites within cells, tissues or biofluids. Recent applications of metabolomics approaches in aquaculture research have highlighted the huge potential for solving problems within all aspects of the production line, from hatchery production to post-harvest quality control. To assist with the growing application of metabolomics in aquaculture research, this contribution provides a review of techniques and steps necessary to conduct metabolomics research, from experimental design to data interpretation. Specifically, we target scientists who are new to the field of metabolomics, and we offer simple, but comprehensive steps and strategies to conduct this type of research. We conclude this primer with some advice on how to access relevant expertise and facilities for metabolomics-based aquaculture research.

Keywords: Metabolomics · Profiling · Fingerprinting Aquaculture · Fish · Shellfish · Bivalves · Molluscs · Marine Bioinformatics · Biomarkers · Metabolism · Physiology NMR ·GC-MS · LC-MS · FT-IR · Multivariate data analysis Chemometrics

Correspondence

Prof. Andrea C. Alfaro, Institute for Applied Ecology New Zealand, School of Applied Sciences, Faculty of Health and Environmental Sciences, Auckland University of Technology, Private Bag 92006, Auckland 1142, New Zealand.

E-mail: aalfaro@aut.ac.nz

Introduction

Within the last decade, the field of metabolomics (the study of metabolites within cells, tissues or biofluids) has expanded, with a number of applications across the life sciences. In aquaculture alone, metabolite patterns have been successfully used to identify and resolve issues related to hatchery production (Young et al. 2015a, 2015b), nutrition and diet (Castro et al. 2015; Cheng et al. 2015), disease and immunology (Liu et al. 2015; Peng et al. 2015), and post-harvest quality control (e.g. Melis 2014; Chen et al. 2015), among others (reviewed by Alfaro & Young 2016). There are some reasons why this approach has been so successful in such a relatively short amount of time. To begin with, metabolomics is an approach that can generate comprehensive datasets of metabolites to describe complex biological systems. Furthermore, the same analytical and computational tools used to generate and interpret data can be performed on any living organism, since metabolites are highly conserved in structure and function across species (in contrast to genes). With recent advances in analytical techniques and computational analysis, complete datasets that describe changes and/or differences in biological systems can be carried out in a rapid and cost-effective manner. However, results stemming from this approach do not necessarily provide mechanistic and/or causal information regarding the patterns observed. In other words, the exploratory nature of this approach is likely to generate new hypotheses, and further targeted experiments may lead to validation of the resultant biological markers. This process allows for unexpected information to be revealed, leading to innovation and discovery in a very efficient manner.

Compared to other areas of research, such as agriculture, food science, and medical science, the

application of metabolomics to aquaculture research has only recently been realised. These applications have been reviewed in a companion paper (Alfaro & Young 2016). Thus, we limit the scope of this contribution to a review of the techniques and steps necessary to conduct metabolomics research in aquaculture. Specifically, the purpose of this review is to provide aquaculture researchers and other aquatic scientists who are new to the field of metabolomics with a simple, but comprehensive, primer on the various strategies that are involved in conducting a metabolomics-based investigation for the first time. This primer summarises information on experimental design, sample collection and preparation, choice of analytical platform, bioinformatics processing, statistical analyses, biological interpretation of the data, and reporting guidelines. We outline several aspects which require careful consideration, specifically for experiments involving aquatic organisms, and we direct readers to a range of specific aquaculture-related research studies to showcase the relevance of these topics.

We conclude this review with some advice on how researchers can access the relevant expertise and facilities for conducting a metabolomics-based project, and we provide some perspectives on the development of future technological strategies for assessing the health and welfare of wild and cultured aquatic organisms.

Metabolomic strategies

There are generally six steps involved in a metabolomics study: (i) robust experimental design, (ii) sample collection and preparation, (iii) analytical measurement and data acquisition, (iv) bioinformatics (data integrity checking and metabolite identifications), (v) statistical analyses, and (vi) biological interpretation and/or biomarker validation (Fig. 1). Due to the wide range of fields encompassed by metabolomics studies (biology, biochemistry, analytical chemistry, bioinformatics and statistics), it is highly recommended that consultation with a metabolomics specialist is carried out in the early stages of experimental design.

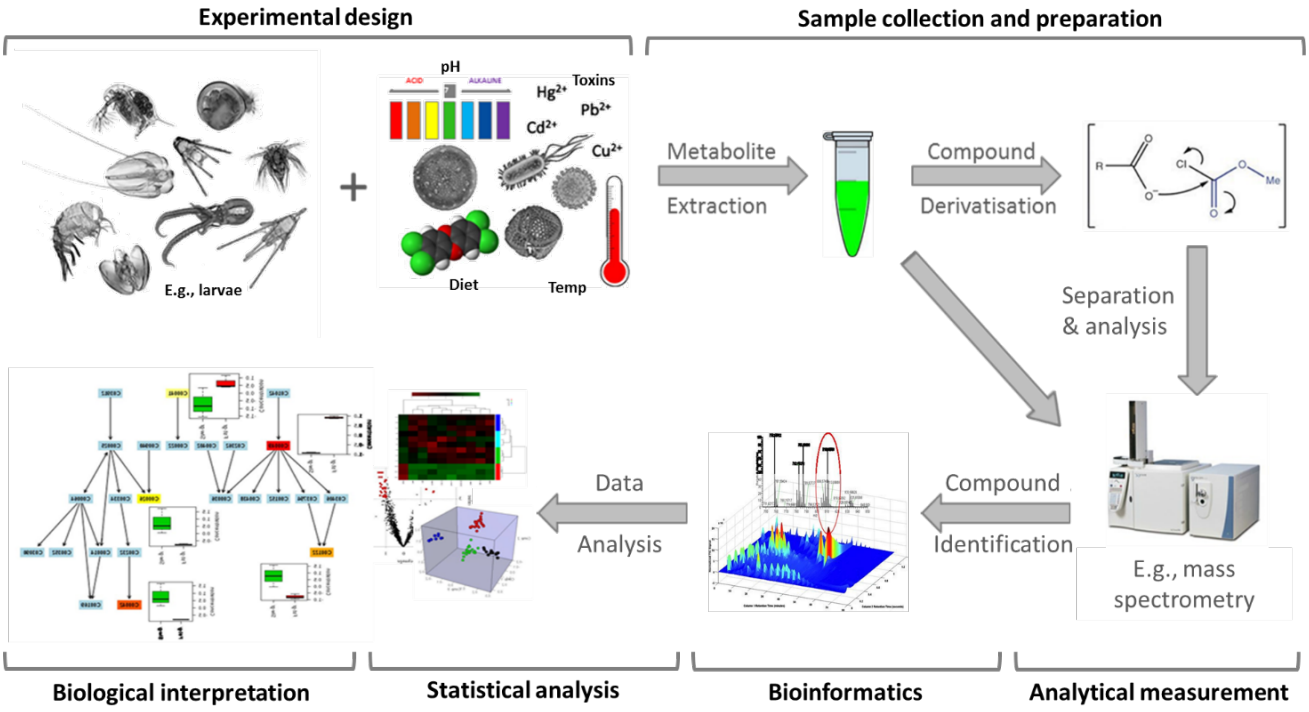

Figure 1. General workflow involved in a metabolomics study outlining the six main steps.

Experimental design & sampling

Along with good standard experimental design practices, there are a number of special considerations to keep in mind when planning a strategy for a metabolomics investigation. There are particular requirements, which mandate that samples be taken only after an experiment has been specifically designed and performed with a metabolomic-based analysis in mind. Samples which have previously been collected and stored for another purpose will unlikely be suitable for incorporation into a metabolomics-based study.

Collected samples must reflect and represent the biology in question, and be appropriate for the particular research questions of the study. It is critical that biological, technical and experimental variability be minimized, since the metabolome can change very rapidly in response to subtle changes in the environment. For example, the metabolic signatures of aquatic organisms can be affected by handling stress and air exposure (Karakach et al. 2009; Connor & Gracey 2012; Young et al. 2015a), so this should be kept to an absolute minimum - even experimentally characterized if possible. The acute stress of transferring fish and crustaceans between culture or storage tanks is reflected in the metabolome and, if not controlled, may influence results of a study (Schock et al. 2013; Mushtaq et al. 2014a). In the case of shellfish, metabolic responses to treatments can be masked when organisms are taken from their natural environments and into the laboratory for a period of acclimatization (Hines et al. 2007). Therefore, sampling and tissue dissections should be performed *in situ*, when possible. Furthermore, in the case of time-course experiments, sampling at the same time of day can be important due to inherent effects associated with circadian rhythms (Gooley 2014; Li et al. 2015).

Selection of adequate control animals is crucial in all omics-based investigations. In most cases, controls and treatment groups should have the same genetic background and should be matched for gender, age, size-class and/or development stage. For example, male and female mussels from a homogenous population can easily be discriminated based on their metabolite profiles (Cubero-Leon et al. 2012), and have sex-specific physiological responses to environmental stressors, toxin exposures and pathogen infections (Ji et al. 2013; Liu et al. 2014a; Ellis et al. 2014). The metabolome is so sensitive that differences in the age of fish larvae can be detected within samples that are only a few hours apart in developmental stage (Huang et al. 2013), and marine invertebrate larvae of the same age but different size-class can be discriminated based on their metabolite profiles (Young et al. 2015b). Thus, these features should be carefully managed to avoid potential experimental bias, unless they are the specific biological aspect under investigation.

Correct selection of sample material is also important. Different tissues (e.g., muscle, gills, liver, pancreas) undergo specific metabolic processes by virtue of their distinct functional purpose. Recent studies of tissue-specific metabolism in aquatic organisms include digestive gland *vs*. gill response differences during pathogen infection in mussels, and differences measured under future climate change scenarios in oysters (Liu et al. 2014b; Wei et al. 2015). In the case of biofluids, the serum and plasma components of blood contain significant chemical differences due to the way in which they are prepared (Yin et al. 2015). Thus, prior knowledge of the biological system is favourable in order to assess the suitability of particular tissues or biofluids for a given experiment.

Once the sample type has been decided, protocols for sampling should be developed. While there is limited information on how the speed of sampling affects the metabolite profile, we suggest that samples be taken rapidly and in a highly reproducible manner to minimize biological and technical variation. For example, if liver samples are to be taken from a number of fish, it would be prudent to make sure that the timing and procedures used to immobilise the organisms and to dissect the tissue be very similar between each animal. Application of anaesthetics during this process should be used with caution since they may disturb the metabolic baseline signature (Bando et al. 2010). The highly dynamic state of the metabolome continues in tissues and biological fluids even after they have been extracted from the organism. Therefore, in almost all metabolomics investigations it is vital that metabolic processes within samples be stopped, or quenched, as soon as possible during collection (reviewed by van Gulik et al. 2012). While other options exist, a typical method to quench metabolism in animal tissues involves snap-freezing samples in liquid nitrogen. Special considerations may need to be made for this, especially if sampling in the field. Furthermore,

it is recommended that samples be stored at or below -80°C until metabolite extraction in order to maintain inactivation of enzymatic and chemical processes, which may influence the metabolite profile. Immediate access to appropriate facilities for sampling and storage is essential. The choice of containers in which the samples will be stored also requires attention due to potential introduction of contaminants, such as surfactants and plasticizers, which may cause severe interferences during analysis (Courant et al. 2014). See Álvarez-Sánchez et al. (2010a) for additional information regarding appropriate selection of biological samples and a review of some practical aspects, which require consideration prior to sample preparation.

Techniques for preparing samples for analysis strongly depend on the type of biological material collected, and the analytical platform to be employed. Regardless of the approach, the metabolite extraction process should be rapid and robust, while minimizing the potential for sample degradation and metabolite modification (Allwood et al. 2013). Special considerations may also be required for processing marine samples due to potential interferences from salts within the sample matrices (Keller et al. 2008), or presence of complex polysaccharides in the case of macroalgae (Goulitquer et al. 2012). Approaches are numerous and constant method development by chemists provide an array of options. These range from simple one-step solvent extraction processes to more complicated procedures involving multiple stages and/or organic synthesis reactions (derivatization).

In general, the most commonly applied solvent extraction methods include: 1) extraction of polar and/or non-polar metabolites with a mixture of methanol, water and chloroform, 2) extraction of polar metabolites with methanol alone or in combination with water, and 3) extraction of polar metabolites with perchloric acid. There are many variations as to the solvent ratios which can be used, the temperature of extraction, the extraction duration, and the mechanical techniques used to disrupt tissue samples and lyse cells. Due to the diversity of possible techniques and wealth of excellent information already available in the literature, method particulars regarding sample preparation are outside the scope of this review. However, we have provided a sizeable table (Table I) containing references to primary literature which have an aquatic metabolomics-based focus, and we highlight the various strategies employed by each study, including the extraction technique used. These studies may be useful to readers as guiding exemplars for many of the strategies discussed in this article. For further details on the preparation of biological samples prior to metabolite detection, see Álvarez-Sánchez et al. (2010b). For comprehensive information on platform-specific sample preparation techniques for general biofluids and animal tissues, see Beckonert et al. (2007), Nováková & Vlčková (2009), Liebeke & Bundy (2012), Römisch-Margl et al. (2012), Vuckovi (2012) and Mushtaq et al. (2014b). For sample preparation techniques with a particular focus on fish and marine invertebrates, see Lin et al. (2007), Wu et al. (2008), del Carmen Alvarez et al. (2010), and Fernández-Varela et al. (2015).

Analytical platforms

A clear understanding of the analytical platform/s to be used is necessary before starting an experiment. Certain platforms have special requirements and may or may not be able to deliver the desired data/information. For example, to obtain broad metabolite coverage, including low abundance compounds, some procedures may require a tissue sample of only 2 mg wet weight, whereas others may require >100 mg. Unfortunately, there is not yet a single platform which can analyse all metabolites within a sample, and some instruments are better-suited for the analysis of particular metabolite classes than others. Hence, multiple platforms may need to be used depending on the aims and scope of the investigation. The costs associated with employing different analytical platforms vary widely, and access to appropriate facilities for sample analysis may limit the decision making process. Therefore, selection of the most appropriate instrument for a given metabolomics-based study will depend largely on the type of sample material collected, the available sample mass, the accessibility of analytical platforms, the end-goals of the researchers, and the budget of the project.

The most commonly applied, high-throughput and high-resolution platforms to analyse samples in metabolomics studies are nuclear magnetic resonance (NMR) and mass spectrometry (MS). In certain circumstances, lower resolution vibrational spectroscopy can also be used. See Figure 2 for

usage trends of the various platforms employed over the past decade. The selection of which platform to apply for a particular metabolomics study is always a compromise between cost, sensitivity, speed, chemical selectivity, and metabolite coverage (Table II). However, realistically, the choice of platform most-often comes down to the availability of analytical facilities and technical expertise through commercial or academic collaborations.

Nuclear magnetic resonance

Nuclear magnetic resonance (NMR) detects the characteristic spin properties of atomic nuclei. When nuclei with particular magnetic attributes are immersed in an external magnetic field, they align themselves with (low energy state) or against (high energy state) that field. Application of very specific radio frequency pulses to the nuclei induces a change in the energy state called a 'spin flip' (Savorani et al. 2013). The presence of other nuclei and chemical bonds in the immediate vicinity of a nucleus changes the intensity of the applied magnetic field by a small amount called nuclear shielding. As a result of this shielding, nuclei within a metabolite will absorb energy at slightly different frequencies, known as a chemical shift. The combination of all of these different frequencies produces a characteristic spectrum, or 'fingerprint' of the sample (Fig. 3A–D). In addition, more complex interactions of the spins under various pulse conditions can provide rich sets of information about the chemical bonding and composition of a molecule or mixture.

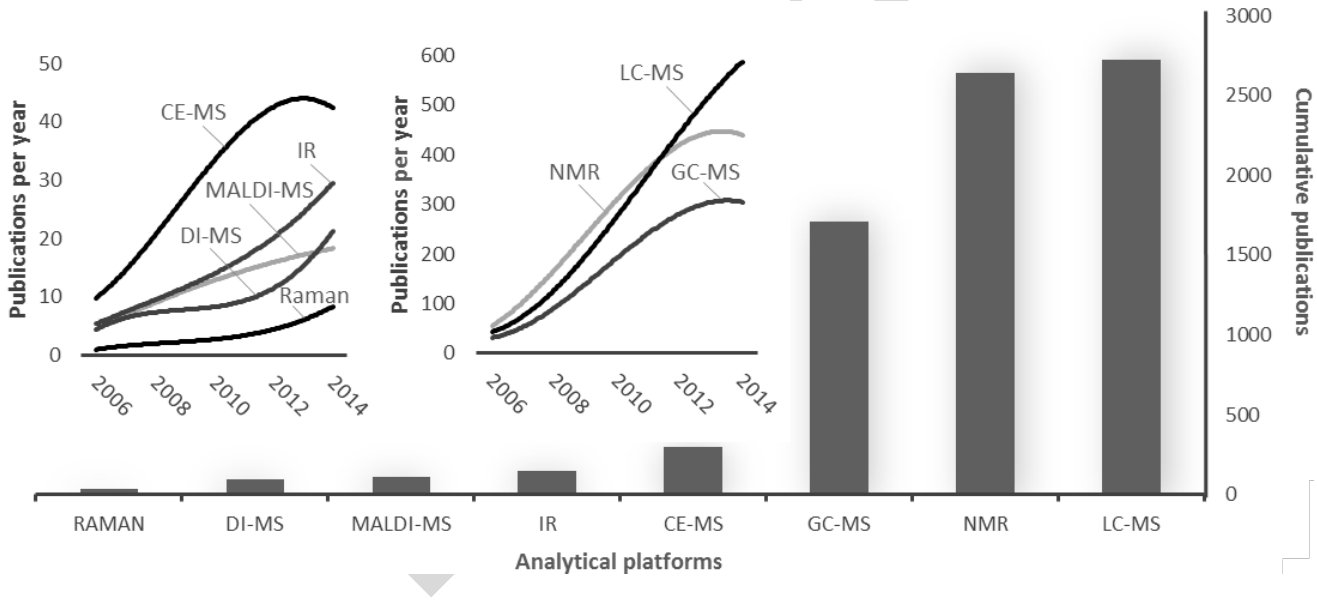

Figure 2. Bibliometric analysis in SciVerse Scopus abstract and citation database (March 24th, 2015). "Metabolom*" was used as a primary keyword in all searches and was combined with keywords for each analytical platform. Searches were limited to terms found within 'abstracts, titles and keywords', between the years 2006–2014, and to research articles only. The bar graph shows the cumulative number of peer-reviewed articles which performed metabolomics-based analyses using various analytical platforms. The inset line graphs show the general usage trends for each platform. These trends also reveal how quickly metabolomics has evolved into such a well-established field of biological research over this short period of time.

All isotopes that contain an odd number of protons and/or neutrons can theoretically be assessed by NMR approaches. However, if they are not found in biological molecules, or have low NMR sensitivities or low natural abundances, they are not often used for metabolomic studies. ^{1}H NMR is frequently applied in metabolomics investigations to probe the molecular arrangements of hydrogen atoms. The ^{1}H isotope is highly abundant in nature (>99.98%) and has a very high NMR sensitivity. See Schock et al. (2012) for an applied example of how ^{1}H NMR was employed to monitor the health of cobia in response to reduced fishmeal-based protein diets, and to identify differential regulation of metabolism

indicative of thyroid disruption and variations in the composition of gut microflora. ^{13}C NMR can also be used, but is much less abundant (1.1%) and less sensitive. However, ^{13}C NMR has special applications in tracer studies to investigate metabolite transformations and metabolic flux (Tikunov et al. 2014).

For example, molecules can be chemically labelled, or enriched, with the ^{13}C isotope and traced through metabolic processes, such as protein catabolism and lipid synthesis, to investigate the uptake and conversion of nutrients in fish (Conceição et al. 2007; Eckman et al. 2013).

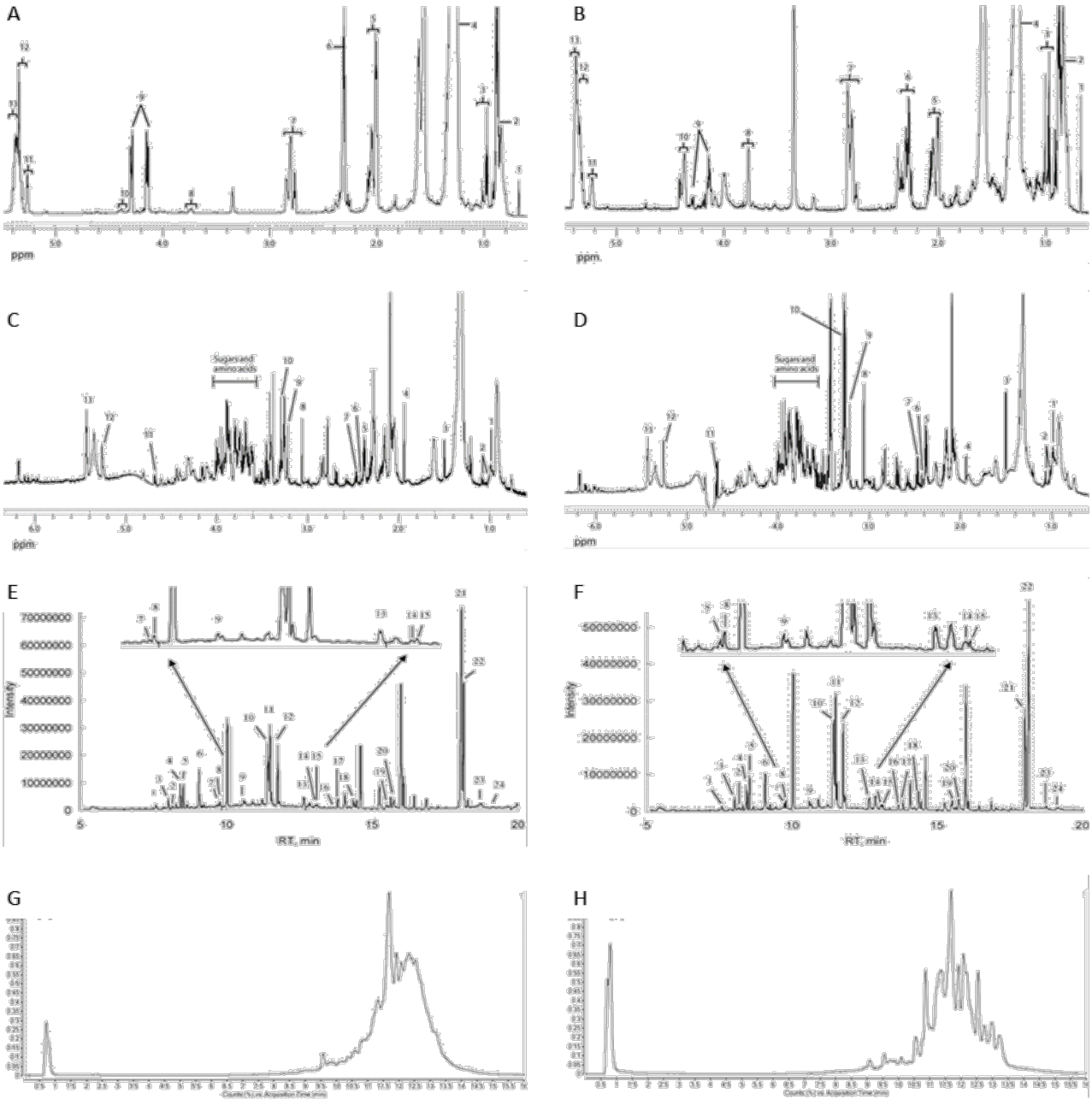

Figure 3. Multi-platform metabolomics-based analysis of fish (*Danio rario*) liver samples showing sex-specific differences in spectral fingerprints obtained from three platforms (NMR, CG-MS and LC-MS). ^{1}H NMR spectra of non-polar extracts from male (A) and female (B) fish. ^{1}H NMR spectra of polar extracts from male (C) and female (D) fish. GC-MS total ion chromatograms of non-polar extracts from male (E) and female (F) fish. LC-MS spectra of non-polar extracts from male (G) and female (H) fish. Numbered peaks represent reliably assigned metabolites after bioinformatics processing. Reprinted with permission from Ong et al. (2009).

Table I. A selection of studies using metabolomics-based approaches with relevance to aquaculture. Although not a complete list of all the available literature, we have provided a broad range of references which we think may be of interest to aquaculture researchers, and may be useful resources as guiding exemplars for the various strategies which can be employed, including: 1) methods for metabolite extraction in diverse organisms, tissues and biofluids; 2) use of different analytical platforms; 3) a range of primary bioinformatics software to process raw spectral data and assign metabolite identities; 4) a variety of databases for matching spectral signatures; 5) various pre-treatment techniques to prepare data for statistical analysis; 6) an array of univariate and multivariate statistical methods to identify sample group differences; and 7) some secondary bioinformatics software to aid interpretation of metabolite profiles within biologically meaningful contexts through use of global a priori knowledge stored in biochemical information databases.

Organism	Sample type	Experimental theme	Extraction method†	Metabolite component‡	Derivatisation method§	Analytical Platform/s¶	General approach	Metabolites Detected	Data pre-treatment methods applied\\\	Bioinformatics & statistical software used††	Statistical analyses & data visualisation‡‡	Databases used§§	Reference
Fish	**Embryos**	Baseline developmental metabolism: Multiplatform metabolomics	MeOH	P	MSTFA	GC-MS, LC-MS	Fingerprinting & profiling	55+	Autoscaled	MZmine, SIMCA-P, MEV	PCA, OPLS-DA, HCA, Kruskal-Wallis test heatmap	NIST library, HMDB	Huang et al. 2013
	Larvae	Nutritional & thermal influence on larval physiology & growth	MeOH/H_2O	P	-	^{1}H NMR	Fingerprinting & profiling	28	Spectral area normalisation, mean centered	TopSpin, Chenomx NMR Suite, MATLAB, PLS Toolbox	PCA	CRL	Chauton et al. 2015
	Flesh/muscle	Nutrition & alternative feed development	MeOH/$CHCl_3$	P, NP	-	^{1}H NMR	Profiling	45	Spectral intensity normalisation, Pareto scaled	TopSpin, AMIX, SAS, SIMCA-P	PCA, OPLS-DA, ANOVA	HMBD, literature search	Cheng K. et al. 2015
		Culture conditions & post-harvest storage	$HClO_4$	P	-	^{1}H NMR	Fingerprinting & profiling	11+	Spectral intensity normalisation, mean centered	MestReC, R	PCA, *t*-test	Undefined	Picone et al. 2011
		Optimisation of sample preparation techniques	$HClO_4$, MeCN/H_2O, MeOH/H_2O/$CHCl_3$, MeOH/H_2O	P, NP	-	^{1}H NMR	Fingerprinting & profiling	26	Bin area normalisation, glog transformation, mean centered	XWINNMR, Chenomx NMR Suite, MATLAB, PLS Toolbox	PCA	Chemical shift data from the literature, CRL	Lin et al. 2007
		Effects of salmon farming on wild fish populations	$HClO_4$	P	-	^{1}H NMR	Fingerprinting & profiling	23	Undefined	MATLAB	RPCA, PLS-LDA	Chemical shift data from the literature	Marhuenda-Egea et al. 2015
		Nutritional history prediction & alternative feeds	C_6H_{12}/H_2O	P, NP		DART-MS	Fingerprinting & profiling	59+	Spectral area normalisation, log transformation, Pareto scaling	MassCenter, SIMCA, Excel	PCA, OPLS-DA	Undefined	Cajka et al. 2013
		Monitoring compositional changes in fillets during post-harvest cold-storage	TCA	P	-	^{1}H-^{13}C NMR	Profiling	51	Untreated	TopSpin, MestReC	N/A	Chemical shift data from the literature, HMDB, BMRB, YMDB, ECMDB	Shumilina et al. 2015
		Food authentication, forensics, providence	MeOH/H_2O/$CHCl_3$	NP	-	^{13}C NMR	Fingerprinting	N/A	Peak maximum normalisation, vast stability scaling	AI Trilogy, Tiberius	PNN, SVM	N/A	Aursand et al. 2009
	Liver	Nutrition & alternative feed development	MeOH/$CHCl_3$	P, NP	-	^{1}H NMR	Profiling	49	Spectral intensity normalisation, Pareto scaled	TopSpin, AMIX, SAS, SIMCA-P	PCA, OPLS-DA, ANOVA	HMBD, literature search	Cheng K. et al. 2015
		Nutrition & alternative feed development	MeOH/H_2O/$CHCl_3$	P, NP	-	^{1}H NMR	Profiling	23+	Spectral area normalisation, Pareto scaled	TopSpin, Chenomx NMR Suite, AMIX, SIMCA-P	PCA, OPLS-DA, ANOVA	CRL, HMBD, literature search	Wagner et al. 2014
		Enhancing disease resistance via simple metabolic modulation	MeOH	P	MSTFA	GC-MS	Profiling	60	Peak height normalisation, median centered, quartile range scaled, log transformed	AMDIS, R, SIMCA-P, SPSS, Prism, MetaboAnalyst (MetPA)	HCA, heatmap, PCA, ICA, MPEA	KEGG, NIST library	Peng et al. 2015
		Enhancing disease resistance via simple metabolic modulation	MeOH	P	MSTFA	GC-MS	Profiling	60	Peak height normalisation, median centered, quartile range scaled, log transformed	AMDIS, R, SIMCA-P, SPSS, Prism	PCA, OPLS-DA, heatmap	KEGG, NIST library	Cheng Z. et al. 2015
		Enhancing disease resistance via simple metabolic modulation	MeOH	P	MSTFA	GC-MS	Profiling	58	Peak height normalisation, median centered, quartile range scaled, log transformed	AMDIS, R, SIMCA-P, SPSS, Prism, MetaboAnalyst (MetPA)	PCA, OPLS-DA, MPEA, HCA, heatmap	KEGG, NIST library	Ma et al. 2015
		Nutrition & alternative feed development	MeOH/H_2O/$CHCl_3$	P	-	^{1}H NMR	Fingerprinting & profiling	12+	Centered (undefined), Pareto scaled	TopSpin, AMIX, SIMCA-P	PCA, OPLS-DA, ANOVA	CRL, BMRB, HMDB	Abro et al. 2014
		Enhancing disease resistance via simple metabolic modulation	MeOH, MeCN/H_2O	P	MSTFA	GC-MS, LC-MS	Profiling	64+	Peak height normalisation, median centered, quartile range scaled, log transformed	AMDIS, R, MarkerLynx, SIMCA-P, SPSS, Prism, MetaboAnalyst (MetPA)	PCA, ICA, MPEA, HCA, heatmap,	KEGG, NIST library	Zhao et al. 2015
		Optimisation of extraction methods	MeOH/H_2O/$CHCl_3$ (various protocols)	P, NP	-	^{1}H NMR, FT-ICRMS	Fingerprinting & profiling	8+	Bin or spectral area normalisation, glog transformation, mean centered	TopSpin. Chenomx NMR Suite, PLS Toolbox, MATLAB	PCA, ANOVA	Chemical shift data from the literature, CRL	Wu et al. 2008
		Utilisation of dietary protein: Growth-metabolic interactions	$HClO_4$	P	ECF	GC-MS	Profiling	12+	Peak height normalisation	SPSS, SIMCA-P	PLS-DA, *t*-test	NIST library, in-house library	Jin et al. 2005
		Health biomarkers & stress evaluation: Multi-platform, large *n* features	MeOH	P		LC-MS, LC-MS/MS, FI-MS/MS	Profiling	95+	Mean centered, Pareto scaled	MetaboAnalyst	PCA, PLS-DA, Mann-Whitney *U* test	N/A	Benskin et al. 2014
		Health biomarkers: Tumor diagnostics	MeCN/H_2O, MeOH/H_2O/$CHCl_3$	P	-	FT-ICRMS	Fingerprinting	4+	Bin area normalisation, glog transformation, mean centered	MIDAS, MSCalc, MATLAB, PLS Toolbox	PCA, PLS-R, *t*-test	N/A	Stentiford et al. 2005
	Kidney	Symptoms of anaemia & health biomarker identification	MeOH/H_2O/$CHCl_3$	P	-	^{1}H NMR	Fingerprinting & profiling	37+	Spectral area normalisation, mean centered, Pareto scaled	TopSpin, MestReNova, Chenomx NMR Suite, AMIX, SPSS	PCA, *t*-test, Mann-Whitney *U* test, Wilcoxon test	HMDB, BML	Allen et al. 2015
	Intestine	Feed additives to enhance growth & metabolism	MeCN/$CHCl_3$	P	-	LC-MS	Profiling	78+	Biomass normalised, autoscaled	Chemstation, EasyLCMS, SPSS, Metaboanalyst	PCA, HCA, heatmap, ANOVA, *t*-test,	Undefined	Robles et al. 2013

Table I. Continued.

Organism	Sample type	Experimental theme	Extraction method†	Metabolite component‡	Derivatisation method§	Analytical Platform/s¶	General approach	Metabolites Detected	Data pre-treatment methods applied\\	Bioinformatics & statistical software used††	Statistical analyses & data visualisation‡‡	Databases used§§	Reference
Fish (cont.)	**Gut contents & faeces**	Effects of diet on microbial symbiosis & co-metabolism + non-invasive sampling	MeOH	P	-	^{1}H-^{13}C NMR	Fingerprinting & profiling	25+	Spectral area normalisation	TopSpin, SpinAssign, R, Gephi	PCA, PLS-DA	PRIMe DB, BMRB, in-house library	Asakura et al. 2014
	Humoral fluid	Mechanism of vaccine action against disease	H_2O	P	MSTFA	GC-MS	Profiling	65	Median centered, interquartile range scaled, Pereto scaled or autoscaled	XCalibur, NIST MS search, SPSS, SIMCA, MetaboAnalyst	OPLS-DA, Heatmap, Mann-Whitney *U* test, MPEA	NIST library, KEGG	Guo et al. 2015
	Haemolymph: Plasma (p) or serum (s)	Toxicological biomarkers & environmental monitoring	None required (p), MeOH/$CHCl_3$ for lipid fraction (p)	P, NP	-	^{1}H NMR	Fingerprinting (some limited profiling)	7+	Bin integral normalisation, mean centered, pareto scaled	NMR Processor, VNMR, SIMCA-P	PCA, PLS-DA	Chemical shift data from the literature	Sammuelsson et al. 2006
		Metabolic effects of food deprivation	None required (p)	P, NP	-	^{1}H-^{13}C NMR	Fingerprinting (some limited profiling)	4+	Bin integral normalisation, Pareto scaled	NMR Processor, SIMCA-P	PCA, OPLS-DA	Chemical shift data from the literature, HMDB	Kullgren et al. 2010
		Utilisation of dietary protein: Growth-metabolic interactions	None required (p)	P	BSTFA	GC-MS	Profiling	16+	Peak height normalisation	SPSS, SIMCA-P	PLS-DA, *t*-test	NIST library, in-house library	Jin et al. 2005
		Spawning-induced inappetence & stress: High resolution platform, large *n* features	MeCN (s)	P	MSTFA	2D GCxGC-MS	Profiling	137	Log transformed, autoscaled, quartile range filtering, KNN (missing variables)	ChromaTOF, MetPP, MetaboAnalyst	PCA, PLS-DA, *t*-test, MPEA	In-house MS library, KEGG	Cipriano et al. 2015
		Health & immunology: Predicting survival	None required (s)	P, NP	MSTFA	GC-MS	Profiling	67	Peal area normalisation, log transformation, autoscaled	XCalibur, SPSS, SIMCA, MetaboAnalyst	Kruskal-Wallis test, Mann-Whitney *U* test, OPLS-DA, MPEA,	NIST library, KEGG	Guo et al. 2014
		Health, nutrition & alternative feed development	None required (s)	P, NP	-	^{1}H-^{13}C NMR	Fingerprinting & profiling	34	Spectral area normalisation, mean centered, Pareto scaled	Chenomx NMR Suite, AMIX, Excel	PCA, Kruskal-Wallis test, ANOVA	CRL, BMRB, in-house library	Schock et al. 2012
	Primary cell culture: Cells (c), media (m)	Effects of plant-derived contaminants in fish feeds: A multiplatform study (c)	MeOH/H_2O/$CHCl_3$	P, NP	-	FT-ICR-MS, ^{1}H NMR	Fingerprinting & lipid profiling	40+	Probabilistic quotient normalisation, glog transformation, mean centered	ProMetab, MATLAB, PLS Toolbox, MI-Pack	PCA, PLS-DA, ANOVA	KEGG	Søfteland et al. 2014
		Nutritional supplementation & diet optimisation (m)	None required	P, NP	-	^{1}H NMR	Fingerprinting & profiling	17	Pareto scaled	Chenomx NMR Suite, SIMCA-P, Statistica	PCA, OPLS-DA, ANOVA	Undefined	Andersen et al. 2014
	Fin tissue	Identifying animal providence	None required	P, NP	-	FT-IR	Fingerprinting	N/A	Undefined	Undefined	PCA	N/A	Nurdalila et al. 2015
	Skin mucus	Minimally-invasive sampling & ecotox: Large *n* features	MeOH/H_2O	P	-	LC-MS/MS	Profiling	204	Peak height normalisation, autoscaled	SIEVE, SIMCA-P, Excel, Systat	PCA, PLS-DA, ANOVA, *t*-test	HMBD, Metlin, LipidMaps	Ekman et al. 2015
	Fish oil capsules	Food authentication, forensics & quality control	None required	NP	-	^{13}C NMR	Fingerprinting	N/A	Peak maximum normalisation	Undefined	PCA, KNNA, GTM, PNN, GRNN	Chemical shift data from the literature	Aursand et al. 2007
	Canned fish packing oil	Food authentication, forensics & quality control	None required	P, NP	-	FT-IR	Fingerprinting	N/A	Undefined	MATLAB, PLS Toolbox	PCA, PLS-DA	N/A	Dominguez-Vidal et al. 2016
Molluscs	**Larvae**	Identification of larval quality biomarkers during hatchery culture	MeOH/H_2O	P	MCF	GC-MS	Profiling	29	Peak height normalisation, metabolite ratios, log transformation, autoscaled	R, AMDIS, MetaboAnalyst	PLS-DA, HCA, heatmap, volcano plot, EBAM, SAM	In-house MS library	Young et al. 2015b
		Handling stress & culture conditions	MeOH/H_2O	P	MCF	GC-MS	Profiling	27	Peak height normalisation, autoscaled	R, AMDIS, MetaboAnalyst, SPSS	PCA, PLS-DA, HCA, heatmap, *t*-test,	In-house MS library	Young et al. 2015a
	Adductor muscle	Organ function & physiology: A multi organ study	$HClO_4$	P	-	^{1}H NMR	Profiling	37+	Peak area normalisation	NMR Processor, SpinWorks	N/A	HMBD	Tikunov et al. 2010
		Optimisation of extraction methods & animal providence: A multi organ study	$HClO_4$, MeCN, Ringer's solution	P	-	^{1}H NMR	Fingerprinting & profiling	32	Bin integral normalisation, Pareto scaled	TopSpin, NMR Processor, Chenomx NMR Suite, JMP, Excel	PCA, *t*-test	HMDB, BMRB	Hurley-Sanders et al. 2015a,b
	Mantle	Sex discrimination	MeOH/H_2O/$CHCl_3$	P	-	^{1}H NMR	Fingerprinting & profiling	16+	Spectral area normalisation, biomass normalisation, glog transformation, mean centered	TopSpin, MATLAB, Chenomx NMR Suite, Excel	PCA, LDA, *t*-test	CRL	Hines et al. 2007
		Ocean acidification, disease, thermal stress & sex differences	MeOH/H_2O/$CHCl_3$	P	-	^{1}H NMR	Fingerprinting & profiling	25	Probabilistic quotient normalisation, glog transformation	MATLAB, PRIMER	PERMANOVA, MDS, SIMPER	Undefined	Ellis et al. 2014
	Hepatopancreas	Health & immunology: Host responses to bacterial pathogens	MeOH/H_2O/$CHCl_3$	P	-	^{1}H NMR	Fingerprinting & profiling	27	Spectral area normailisation, biomass normalisation, glog transformation, mean centered	TopSpin, Chenomx NMR Suite, MATLAB, PLS Toolbox, Minitab	PCA, PLS-DA, OPLS-DA, ANOVA	Chemical shift data from the literature	Wu et al. 2013
		Ocean acidification: Integrated metabolomics & proteomics	MeOH/H_2O/$CHCl_3$	P	-	^{1}H NMR	Fingerprinting & profiling	32	Spectral area normalisation, glog transformation	TopSpin, Chenomx NMR Suite, MATLAB, SIMCA-P	PLS-DA, OPLS-DA	CRL	Wei et al. 2015
	Foot	Health: Biomarkers for toxicology, hypoxia & food limitation	MeCN/H_2O	P	-	^{1}H NMR	Fingerprinting & profiling	20	Box-Cox transformation	SpecManager, Excel, Minitab, GenStat	PCA, LDA, ANOVA	HMBD	Tuffnail et al. 2009
	Gonad	Sex discrimination & reproductive physiology	SPE & fractionation	P, NP	-	LC-MS	Fingerprinting & profiling	21+	Spectral area normalisation, mean centered	MarkerLynx, MassLynx, SIMCA-P	PCA, PLS-DA, OPLS-DA, Mann-Whitney *U* test	HMBD, KEGG LDB, BiGG DB, PubChem, NLDB, MBDB	Cubero-Leon et al. 2012

Table I. Continued.

Organism	Sample type	Experimental theme	Extraction method†	Metabolite component‡	Derivatisation method§	Analytical Platform/s¶	General approach	Metabolites Detected	Data pre-treatment methods applied\\	Bioinformatics & statistical software used††	Statistical analyses & data visualisation‡‡	Databases used§§	Reference
Molluscs (cont.)	**Gills**	Identification of thermal stress biomarkers	$MeOH/H_2O$	P	Undefined	GC-MS	Profiling	52	Biomass normalisation	JMP	DFA	Undefined	Dunphy et al. 2015
		Toxicology: Integrated metabolomics & proteomics	$MeOH/H_2O/CHCl_3$	P	-	^{1}H NMR	Fingerprinting & profiling	28	Spectral area normalisation, glog transformation, mean centered	TopSpin, Chenomx NMR Suite, MATLAB, SIMCA-P	PLS-DA, OPLS-DA	Chemical shift data from the literature	Ji et al. 2013
		Toxicology: Integrated metabolomics & proteomics	$MeOH/H_2O/CHCl_3$	P	-	^{1}H NMR	Fingerprinting & profiling	25+	glog transformation	TopSpin, Chenomx NMR Suite, SIMCA-P, MATLAB	PCA, PLS-DA, OPLS-DA	CRL, KEGG	Song et al. 2016
		High resolution NMR: Coastal marine pollution & toxicology	$MeOH/H_2O/CHCl_3$	P	-	HR-MAS ^{1}H NMR	Fingerprinting & profiling	27+	Spectral area normalisation, glog transformation, mean centered	XWIN-NMR, Chenomx NMR Suite, MATLAB, Unscrambler X, Excel	PCA, *t*-test	HMBD, CRL	Cappello et al. 2013
		Toxicology: Integrated metabolomics & proteomics	$MeOH/H_2O/CHCl_3$	P	-	^{1}H NMR	Fingerprinting & profiling	9+	Spectral area normalisation, glog transformation, mean centered	TopSpin, Chenomx NMR Suite, MATLAB, SIMCA-P	PLS-DA, OPLS-DA	Chemical shift data from the literature	Ji et al. 2016
	Central nervous system & glands	Baseline molecular phenotyping: Multiplatform metabolomics	$MeOH/H_2O/CHCl_3$	P, NP	BF_3MeOH (lipids)	GC-MS, LC-MS (RPLC & HILIC)	Fingerprinting & profiling	73+	Selected peak normalisation, total protein content normalised	DataAnalysis, ProfileAnalysis, SPSS, Excel, SIMCA-P	PCA, PLS-DA, OPLS-DA, HCA, ANOVA	Undefined	Tufi et al. 2015a
		Neurotoxicity of pesticides in aquatic environments: Multiplatform metabolomics	$MeOH/H_2O/CHCl_3$	P, NP	BF_3MeOH (lipids)	GC-MS, LC-MS	Profiling	73	Total protein content normalised	Compass DataAnalysis, DataAnalysis, PathwayScreener, ProfileAnalysis, IMPaLA, SPSS, Metamapp, Cytoscape	ORA, BNM, *t*-test	Reactome, EHMN, KEGG, Wikipathways, SMPDB, HumananCyc	Tufi et al. 2015b
	Gastrointestinal tract and/or digestive gland	Bioindicator species for pollution monitoring	$MeOH/H_2O/CHCl_3$	P	-	^{1}H NMR	Fingerprinting & profiling	19+	Spectral area normalisation, glog transformation, mean centered	MATLAB, PLS Toolbox	ANOVA, PCA, PLS-DA	Chemical shift data from the literature	Liu et al. 2011
		Metabolic effects of food deprivation & extraction method optimisation	$MeOH/H_2O/CHCl_3$	P	-	^{1}H NMR	Profiling	28	Log transformation, median centered	Chenomx NMR Suite, TopSpin, R, Unscrambler	PLS-DA, Mann-Whitney *U* test	CRL	Sheedy et al. 2015
	Whole soft tissue	Dual platform metabolomics: Toxicological mechanisms	$MeOH/H_2O/CHCl_3$	P	MSTFA	GC-MS, ^{1}H NMR	Fingerprinting & profiling	NMR: 17+ GC-MS: 24+	Spectral area normailsation, autoscaled (GC-MS), mean centered & Pareto scaled (NMR)	SpecManager, SIMCA-P	PCA, PLS-DA	Chemical shift data from the literature, NIST library	Spann et al. 2011
		Toxicological mechanisms	$MeOH/H_2O/CHCl_3$	P	-	^{1}H NMR	Fingerprinting & profiling	25+	Bin integral normalisaiton, glog transformation, mean centered	TopSpin, MATLAB, PLS Toolbox	PCA, PLS-DA, ANOVA	Undefined	Wu & Wang 2010
		Unique extraction & platform: Method assessment	SPME	P, NP	-	2D GCxGC-MS	Fingerprinting & profiling	63+	Fourth root transformation	ChromaTOF, PRIMER	PERMANOVA, PCoA, HCA	WMSL, NIST library	Rocha et al. 20013
		Coastal marine pollution & toxicology	$MeOH/H_2O/CHCl_3$	P	-	^{1}H NMR	Fingerprinting & profiling	24	Pareto scaled	Chenomx NMR Suite, MATLAB, SIMCA-P, R	PCA, OPLS-DA, PLS-DA	CRL	Kwon et al. 2012
	Haemolymph: whole blood (wb) or plasma (p)	Mechanical shaking & salinity stress	None required (wb)	P, NP	-	FT-IR	Fingerprinting	N/A	Undefined	Undefined	MANOVA, PCA, CVA	N/A	Bussell et al. 2008
		Developing strategies for identifying stress	None required (wb)	P, NP	-	FT-IR	Fingerprinting	N/A	Untreated	Undefined	MANOVA, PCA, CVA	N.A	Gidman et al. 2007
		Health/stress biomarkers: Toxicology	None required (p)	P, NP	-	^{1}H NMR	Fingerprinting & profiling	18	Spectral area normalisation, glog transformation, mean centered	MATLAB, PLS Toolbox, Chenomx NMR Suite	ANOVA, PCA	CRL	Zhou et al. 2015
Crustaceans	**Claw muscle**	Nutritional composition & quality assessment, 3 spp.	$MeOH/H_2O/CHCl_3$	P, NP	-	^{1}H NMR	Fingerprinting & profiling	24+	Spectral area normalisation, mean centered, areto scaled	TopSpin, AMIX, SIMCA-P	PCA, PLS-DA	Chemical shift data from the literature	Zotti et al. 2016
	Tail muscle & other	Identification of health & stress biomarkers in shrimp during intensive culture	$MeOH/H_2O/CHCl_3$	P	-	^{1}H-^{13}C NMR	Fingerprinting & profiling	50+	Spectral area normalisation, mean centered, Pareto scaled	AMIX, Chenomx NMR Suite, Excel	PCA, *t*-test	HMDB, BMRB, CRL, in-house library	Schock et al. 2013
	Hepatopancreas	Mechanisms of white spot virus syndrome in shrimp	Undefined	Undefined	-	^{1}H NMR	Fingerprinting & profiling	27+	Undefined	TopSpin, SIMCA-P	PCA, OPLS-DA	Chemical shift data from the literature	Liu et al. 2015
	Haemolymph	Pathogen-induced oxidative stress responses in crabs: New biochemical insights	None required	P, NP	-	^{1}H-^{13}C NMR	Fingerprinting & profiling	20	Spectral area normalisation, mean centered	AMIX, SIMCA-P	PCA, sPCA, PLS-DA, ANOVA	Chemical shift data from the literature, in-house library	Schock et al. 2010
Echinoderms	**Muscle**	Thermal stress responses	$MeOH/H_2O/CHCl_3$	P	-	^{1}H NMR	Fingerprinting & profiling	31	Spectral area normalisation, glog transformation, mean centered	TopSpin, Chenomx NMR Suite, MATLAB, SIMCA-P	PCA, OPLS-DA	Chemical shift data from the literature	Shao et al. 2015
Macroalgae	**Thallus**	Effects of food processing on nutrient composition	$MeCN/H_2O$	P	-	^{1}H NMR	Fingerprinting & profiling	32+	Bin integral normalisation	AMIX, SIMCA-P, MATLAB, SPSS	PCA, OPLS-DA, heatmap, *t*-test	Chemical shift data from the literature	Ye et al. 2014
		New insights into metabolism	MeOH, H_2O	P	-	^{1}H NMR	Profiling	27	N/A	Chenomx NMR Suite	N/A	PRIMe DB, HMDB, BMRB	Gupta et al. 2013
	Stipe & blades	Seasonal variations in metabolism: Multiplatform metabolomics	None required (FT-IR), MeOH, H_2O (NMR)	P, NP	-	FT-IR, ^{1}H-^{13}C NMR	Fingerprinting & profiling	51	Specific peak intensity normalisation (FT-IR), spectral area normalisation (NMR)	Excel, TopSpin, SpinAssign, OMNIC, R, Amos, Gephi, Fityk	PCA, ICA, SOMs, CNA, SEM, MCR-ALS	PRIMEe DB	Ito et al. 2014

Table I. Continued.

Organism	Sample type	Experimental theme	Extraction method†	Metabolite component‡	Derivatisation method§	Analytical Platform/s¶	General approach	Metabolites Detected	Data pre-treatment methods applied\\	Bioinformatics & statistical software used††	Statistical analyses & data visualisation‡‡	Databases used§§	Reference
Microalgae	**Cells or extracts**	Screening microalgae to identify commercially useful mutants	None required (FT-IR), MeOH/H_2O/$CHCl_3$ (LC-MS)	P, NP	-	FT-IR, LC-MS	Fingerprinting & lipid profiling	11+ lipid classes	Biomass normalised	MATLAB, Xcalibur, Unscrambler	PCA, PC-DFA, PLS-DA, PLS-R	MMD, HMDB, KEGG, BioCyc, LIPIDMAPS, DrugBank	Bajhaiya et al. 2016
		Profiling diatoms for bioenergy & feedstock	MTBE/MeOH/H_2O	P, NP	MSTFA	GC-MS, LC-MS	Profiling	96+	Cell density normalised, median scaled, log transformed	Expressionist Refiner MS, ChromaTOF, R	PCA, HCA, heatmap, *t*-test	In-house MS library	Bromke et al. 2015
		Dual platform metabolomics: Natural products research	MeOH/H_2O/$CHCl_3$ (GC-MS), MeOH (LC-MS)	P	MSTFA	GC-MS, LC-MS	Profiling	128	Peak area & cell density normalisation	AMDIS, SIMCA-P, MeV	PLS-DA, HCA, heatmap, MPEA, WGCNA	KEGG	Yu et al. 2015
		Photobioreactor culture conditions & bioresource development	None required (FT-IR), MeOH & EtOH/H_2O (LC-MS)	P, NP	-	FT-IR, LC-MS	Fingerprinting & profiling	13+	Second derivative calculation (FR-IR), log transformed, Pareto scaled	OpusLab, XCMS, Xcalibur, R, Spectrum Database, Statistica Data Miner, MetaboAnalyst, SIMCA-P	Kruskal-Wallis test, HCA, Mann-Whitney *U* test, heatmap, PLS-DA	In-house MS library	Courant et al. 2013
		Chemical interactions between bacteria & diatoms	MeOH/EtOH,$CHCl_3$ (GC-MS)	P	MSTFA	GC-MS	Fingerprinting & profiling	19+	Cell density normalised	MassLynx, AMDIS, MET-IDEA, SigmaPlot, Excel	RM-ANOVA, PCoA, CAP, heatmap	NIST library, GMDB	Paul et al. 2013
Bacteria	**Cell extract**	Mechanisms of white spot syndrome virus in shrimp: Nutritional treatment	MeOH, H_2O	P	MSTFA	GC-MS	Fingerprinting & profiling	3+	Untreated	Xcalibur, NIST MS search	Wilcoxon rank-sum test,	NIST library	Zhu & Jin 2015
		Restoring pathogen susceptibility to antibiotics: Simple metabolic modulation	MeOH	P	MSTFA	GC-MS	Profiling		Undefined	Xcalibur, NIST MS Search, MetaGeneAlyse	ICA	NIST library	Su et al. 2015

† **Extraction solvents:** MeCN (acetonitrile), MeOH (methanol), EtOH (ethanol), MTBE (methyl-tert-butyl ether), H_2O (water), $CHCl_3$ (chloroform), $HClO_4$ (perchloric acid), C_6H_{12} (cyclohexane), TCA (trichloroacetic acid), SPME (solid phase microextraction), SPE (solid phase extraction)

‡ **Metabolite components:** P = polar component, NP = non-polar component

§ **Derivatisation:** MSTFA = silylation with *N*-Methyl-*N*-(trimethylsilyl)trifluoroacetamide, BSTFA = silylation with *N*,*O*-Bis(trimethylsilyl)trifluoroacetamide, MCF = alkylation with methyl chloroformate, ECF = alkylation with ethyl chloroformate

¶ **Analytical platforms:** FI-MS/MS (flow injection tandem mass spectrometry), FT-IR (Fourier transform infrared spectroscopy), ^{1}H NMR (proton nuclear magnetic resonance), ^{1}H-^{13}C NMR (two dimensional proton and carbon NMR for assisting metabolite identifications), HR-MAS NMR (high resolution magic angle spinning NMR), FT-ICRMS (Fourier transform ion cyclotron resonance mass spectrometry), DART-MS (direct analysis in real time mass spectrometry), GC-MS (gas chromatography mass spectrometry), LC-MS (liquid chromatography mass spectrometry), RPLC (reverse phase liquid chromatography), HILIC (hydrophilic interaction liquid chromatography)

\\ **Note:** Whilst not explicitly stated within the table, most metabolomics-based investigations will also include normalisation of data to an internal standard as a data pre-treatment method to compensate for potential technical variations (e.g., variable metabolite recoveries during sample preparation and processing).

†† **Software:** AI Trilogy (Ward Systems Group Inc., US), AMDIS (Automated Mass Deconvolution and Identification System [The National Institute of Standards and Technology, US]), Amos (IBM Corp., US), AMIX (Bruker Corp., Germany), Chemstation (Agilent, US), Chenomx NMR Suite (Chenomx Inc., Canada), ChromaTOF (LECO Corp., US), Compass DataAnalysis (Bruker Corp., Germany), Cytoscape (Shannon et al. 2003), DataAnalysis (Bruker Corp., Germany), EasyLCMS (Fructuoso et al. 2012), Excel (Microsoft, US), Expressionist Refiner MS (Genedata, Switzerland), FityK (Wojdyr 2010), Gephi (Bastian et al. 2009), GenStat (VSN Internaitonal, UK), IMPaLA (Integrated Molecular Pathway Level Analysis [Kamburov et al. 2011]), JMP (SAS Institute Inc., US), MarkerLynx (Waters Corp., US), MassCenter (JEOL, Japan), MassLynx (Waters Corp, US), MATLAB (Mathworks, US), MestReC (Mestrelab Research, Spain), MestReNova (Mestrelab Research, Spain), Metaboanalyst (Xia et al. 2015), MetaGeneAlyse (Daub et al. 2003), MetaMapp (Barupal et al. 2012), MET-IDEA (Lei et al. 2012), MetPA (Xia & Wishart 2010 [now a component of MetaboAnalyst]), MetPP (Wei et al. 2013), MEV (Multi Experiment Viewer [Howe et al. 2010]), MIDAS (Wang et al. 2014), Minitab (Minitab Inc., US), MSCalc (SoftShell International Ltd., US), MZmine (Katajamaa et al. 2006), NMR Processor (ACD/Labs, Canada), OMNIC (Thermo Fisher Scientific, US), PathwayScreener (Bruker Corp., Germany), PLS Toolbox (Eigenvector Research Inc., US), PRIMe DB (Platform for RIKEN Metabolomics Database, RIKEN Yokohama Institute, Japan), PRIMER (Plymouth Routines In Multivariate Ecological Research [PRIMER-E Ltd, UK]), Prism (GraphPad Software Inc., US), ProfileAnalysis (Bruker Corp., Germany), ProMetab (Parson et al. 2007), R (R Core Team 2014), SAS (Statistical Analysis System [SAS Institute Inc., US]), SIEVE (Thermo Fisher Scientific, US), SIMCA-P (MKS Unimetrics, Sweden), SpecManager (ACD/Labs, Canada), Spectrum Database (ACD/Labs, Canada), SpinAssign (Platform for RIKEN Metabolomics, RIKEN Yokohama Institute, Japan), SpinWorks (University of Manitoba, US), SPSS (IBM Corp., US), Statistica Dataminer (StatSoft, US), Systat (Systat Software, US), Tiberius (Tiberius Data Mining, Australia), TopSpin (Bruker Corp., Germany), Unscrambler (CAMO, Norway), VNMR (Varian Inc., US), Xcalibur (Thermo Fisher Scientific, US), XWINNMR (Bruker Corp., Germany).

‡‡ **Statistical analyses:** ANOVA (analysis of variance), RM-ANOVA (repeated measures ANOVA), MANOVA (multivariate ANOVA), PERMANOVA (permutation MANOVA), BNM (biochemical network mapping), CAP (canonical analysis of principal coordinates), CNA (correlation network analysis), CVA (canonical variates analysis), DFA (discriminant function analysis), EBAM (empirical Bayes analysis of metabolites), GRNN (general regression neural networks), GTM (generative topographic mapping), HCA (hierarchical cluster anlaysis), ICA (independent component analysis), KNNA (Kohonen neural network analysis), MDS (multi-dimensional scaling), MCR-ALS (multivariate curve resolution-alternating least squares), MPEA (metabolite pathway enrichment analysis), ORA (over-representation analysis), PNN (probabilistic neural networks), PCA (principal components analysis), SPCA (supervised PCA), RPCA (robust PCA), PC-DFA (principal component discriminant function analysis), PCoA (principal coordinate analysis), PLS-DA (projection to latent structures discriminant analysis), OPLS-DA (orthogonal projection to latent structures discriminant analysis), PLS-LDA (projection to latent structures linear discriminant analysis), PLS-R (projection to latent structures regression), SAM (significant analysis of metabolites), SEM (structural equation modelling), SIMPER (Similarity Percentage analysis), SOMs (self-organising maps), WGCNA (weighted gene correlation network analysis).

§§ **Databases:** BiGG DB (Database for Biochemically, Genetically and Genomically structured genome-scale metabolic network reconstructions), BML (Birmingham Metabolite Library), BMRB (Biological Magnetic Resonance Data Bank), CRL (Chenomx Reference Library), ECMDB (E. coli Metabolome Database), EHMN (Edinburgh Human Metabolic Network), GMDB (Golm Metabolome Database), HMDB (Human Metabolome Database), KEGG (Kyoto Encyclopaedia of Genes and Genomes), KEGG LDB (KEGG Ligand Database), NIST (The National Institute of Standards and Technology) library, MBDB (Massbank Database), MMD (Manchester Metabolomics Database), NLDB (Nature Lipidomics Database), SMPDB (The Small Molecule Pathway Database), WMSL (Wiley Mass Spectral Libraries), YMDB (Yeast Metabolome Database).

A range of other NMR-based techniques are also available which have various applications and levels of analytical sensitivity, such as two-dimensional and hyphenated platform approaches (reviewed by Simpson & Bearden 2013; Bharti & Roy 2014; Larive et al. 2014). NMR was originally the workhorse of metabolite profiling in the early days, but recent advances in mass spectrometry based approaches offer alternative methods of analyses. These platforms are often used in combination, since they have their own individual merits (Ong et al. 2009; Zhang et al. 2012). For example, NMR is a non-destructive technique and acquires highly robust and reproducible measurements. Separation of metabolites prior to detection is not necessary and minimal sample preparation is required. NMR is generally cheaper to perform, but unfortunately has comparatively low sensitivity in relation to MS-based platforms, which means only metabolites that are present in significant quantities can be detected.

Mass spectrometry

Mass spectrometry (MS) is a method which involves the measurement of molecular weights of molecules (reviewed by El-Aneedet al. 2009; Viant & Sommer 2013). There are three components to a mass spectrometer: the ion source, the mass analyser, and the detector (Glish & Vachet 2003). There are many different types of these components (see Dettmer et al. 2007; El-Aneed et al. 2009; Junot et al. 2014). At the ion source, metabolites within a sample are ionized by a variety of processes. For metabolomics work, the most commonly used ionisation techniques are electron ionisation and electrospray ionisation (Lei et al. 2011). In most cases, the molecules become sufficiently excited to fragment into a number of electrically charged ions. These ions move into the mass analyser where they are separated based on their mass to charge (m/z) ratio by accelerating them and subjecting them to various combinations of electric, magnetic or electromagnetic fields or in a 'time of flight' mass spectrometer which assesses how fast they are travelling. Fragments with different m/z ratios travel at different speeds and are deflected from their forward trajectory to different degrees; lighter ions deflect more than heavier ions, and the higher the ionic charge, the greater the deflection. This allows the various types of mass analysers to filter the ions. The ions are then directed into a device that counts the number of ions at each different mass. This information is plotted in a spectrum of the ion abundance as a function of the m/z ratio. The identity of a metabolite can be putatively elucidated by comparing the fragmentation patterns against open access and/or proprietary databases which contain mass spectra of known compounds. Depending on the particular instrument, some (high resolution mass spectrometers) are capable of determining the actual elemental composition of each ion, thus providing an extra dimension of information for validation of metabolite identity.

MS-based methods are becoming highly sophisticated and newly-developed platform variations are increasingly being showcased in the scientific literature. MS can be performed directly on samples without pre-separation of metabolites (reviewed by Ibáñez et al. 2014). While direct MS techniques are rapid, they also suffer from low ionization efficiencies and ion suppression. Thus, to decrease the complexity of the sample matrix and enhance the sensitivity and selectivity of the analysis, MS-based metabolomic approaches usually involve separation of metabolites via chromatography or electrophoresis prior to MS detection. The benefits of pre-separation are that a significant amount of information is available from the pre-separation process, and metabolites with the same mass can easily be distinguished since they are introduced into the MS system at different times. In addition, higher quantitative accuracies can be achieved since problems associated with ion suppression and other interferences are greatly reduced. Gas chromatography (GC), liquid chromatography (LC), and capillary electrophoresis (CE) are the most commonly applied methods for this purpose. When coupled, these instruments are called hyphenated platforms (GC-MS, LC-MS and CE-MS). Each of these platforms have their own unique advantages, and can be used in combination to obtain very broad coverage of the metabolome (Lei et al. 2011).

Gas Chromatography Mass Spectrometry

Gas Chromatography (GC) separates metabolites which are volatile and thermally stable, or which become volatile and thermally stable after functional group modifications (e.g., alkylation or

silylation via chemical derivatization [Villas-Bôas et al. 2011]) (reviewed by Garcia & Barbas 2011). Once the sample extract has been prepared, it is injected into a hot gas stream flowing through a long and very small diameter tube in the GC instrument. The inside walls of the tube, called a column for historical reasons, are coated with material that has some affinity for the various components in the mixture. The different interactions of the metabolites with the gas stream and the column walls result in differential flow speeds through the column and they exit the column at different times (producing a chromatogram [Fig. 3E,F]); thus entering the mass spectrometer in a unique sequence. This combination of unique entrance times and associated information on the physicochemical properties of metabolites provides an enhanced means of profiling and identification. See Figure 4 for an illustrated overview of the analytical processes involved using pre-separation techniques combined with MS for the metabolomic analysis of complex sample matrices. GC-MS has the advantage that it produces very stable metabolite retention times within the column, does not have drawbacks associated with ion suppression, and generates highly reproducible fragmentation patterns. These features mean that metabolite identifications can more easily be authenticated by matching spectra against those contained within numerous open-access spectral libraries. However, if samples need to be derivatized, are thermally unstable, or have too high a molecular weight, GC-MS may not be suitable. See Zhao X. et al. (2015) for an applied example of how GC-MS was used to identify biomarkers for temperature stress in tilapia, and to discover that an exogenous supply of L-proline into the culture water led to higher disease resistance against bacterial pathogens.

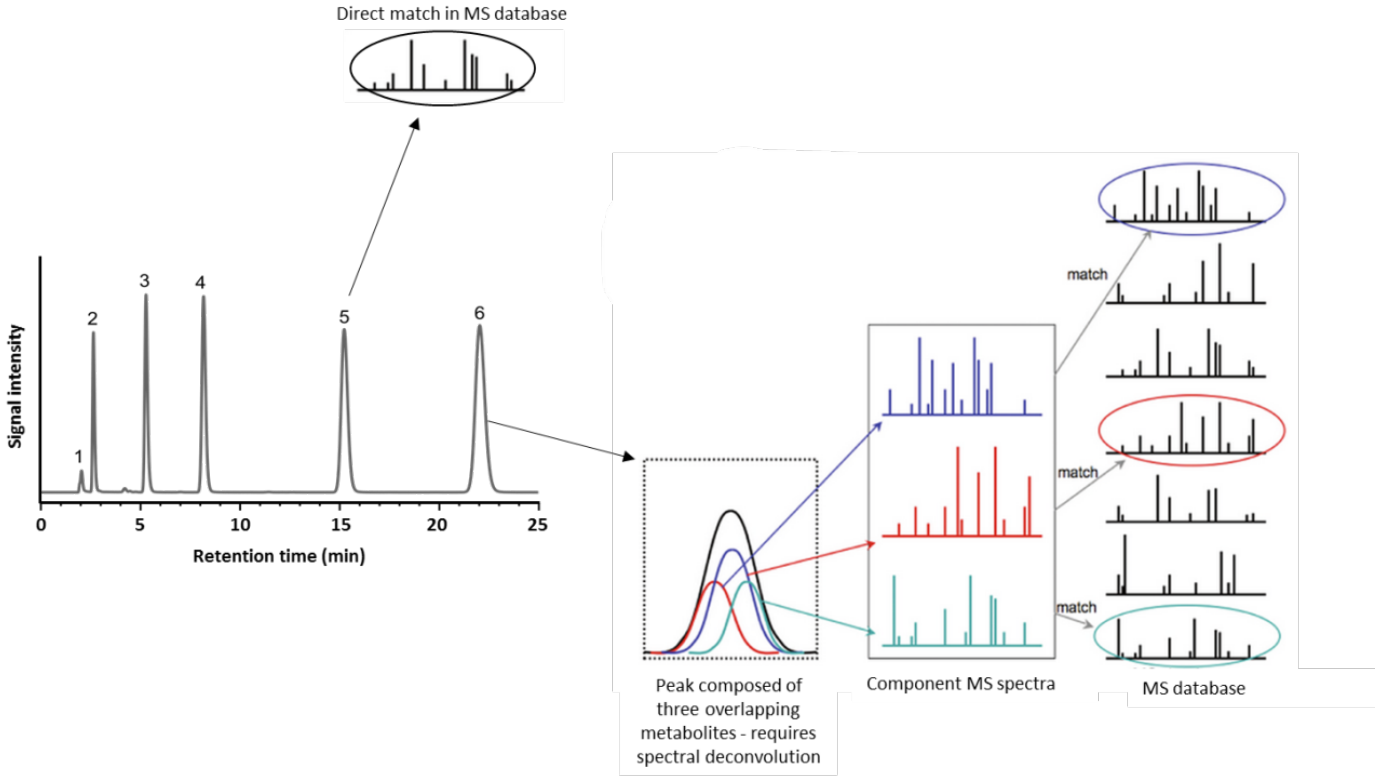

Figure 4. Overview of the processes involved using pre-separation techniques combined with mass spectrometry. Gas chromatography, liquid chromatograph or capillary electrophoresis is used to separate metabolites in the sample extract to produce a chromatogram. Compounds within the peaks are then sequentially analysed by mass spectrometry, and their ion m/z ratios are compared to those stored in mass spectral databases for identification. In some cases, peaks may comprise multiple metabolites with similar physicochemical properties which are unable to be separated by the pre-separation device and deconvolution of the spectra is required.

Table II. Comparisons between different analytical platforms for processing metabolomics samples.

Platform	Advantages	Disadvantages	Processing Cost†
NMR	Rapid analysis time (5–10 mins) Simple sample preparation No derivatisation needed Provides detailed structural information Low chemical bias Very reproducible Can be high resolution Excellent metabolite recovery (no suppression) Highly quantitative (without standards)	Low sensitivity Convoluted Spectra Libraries of limited use due to complex matrix More than one peak per component Peak overlap common pH adjustment required	Cheap $30–100 USD per sample
GC-MS	Very sensitive Very robust Large linear range MS provides some structural information Many available libraries for metabolite identification Pre-separation provides additional information Does not suffer from ion suppression Reproducible retention times Quantitative (with appropriate standards)	Slow analysis time (30–60 mins) Extensive sample preparation Derivatisation required Destructive to sample Some metabolites cannot be made volatile Some metabolites are too large for analysis Cannot detect some thermally unstable metabolites	Expensive $100–200 USD per sample
LC-MS	Very sensitive Can detect a very wide range of metabolites MS provides some structural information High mass accuracy Many modes of pre-separation available Pre-separation provides additional information Quantitative (with appropriate standards)	Analysis time can be slow (10–60 mins) Lack of comprehensive spectral libraries Ion suppression & adduct formation problems Destructive Metabolite identification is difficult Low retention time reproducibility	Very expensive $150–400 USD per sample
FT-IR, NIR, Raman	Very rapid analysis time (10–60 secs) Low chemical bias Can be used directly on samples No derivatisation required Complete fingerprint of sample composition Useful for identifying functional groups	Extremely convoluted spectra More than one peak per component Metabolite identification almost impossible Often requires sample drying	Very Cheap $10–50 per sample

† Per sample processing costs vary between service providers and depend highly on the number of samples to be analysed within a particular project, and the resolution of the specific platform to be employed. The price ranges displayed are not strictly defined limits but are typical of the current rates charged for commercial samples, and generally will include metabolite extraction (and derivatisation when required), instrumental analysis, and some primary bioinformatics processing. Additional statistical analysis, secondary bioinformatics processing and/or biological interpretation of the data can usually be provided as an extra service by most facilities e.g. $50–200 USD per hour for a metabolomics specialist. Academic-based metabolomics service providers may offer discounted rates for collaborative projects.

Liquid Chromatography Mass Spectrometry

High Performance Liquid Chromatography (HPLC or LC) is based on similar chromatographic principals as GC, but the sample is not heated to high temperatures (reviewed by Xiao et al. 2012). The most important distinction between GC and LC is that GC largely separates metabolites based on their boiling points with secondary retention by polarity, whereas metabolite size or polarity are the main mechanisms of LC. The column in this case is a tube a few millimetres wide and a few centimetres long and packed with extremely fine powder coated with material that has some affinity for the various metabolite components in the sample. A suitable solvent mixture is pumped at very high pressure through the column. The sample is introduced into the column as a solution and, like GC, the various components of the mixture travel through the column at different speeds and exit in sequence to produce a chromatogram (Fig. 3G,H). From there the outlet stream is directed into a suitable detector which, for metabolomic work, is typically a high-resolution mass spectrometer. Therefore, LC-MS can analyse a wider range of metabolites since non-volatile and thermally sensitive compounds can be separated in the liquid phase. However, LC-MS suffers from greater ionization and matrix effects,

and lower chromatographic reproducibility. These features make assigning metabolites through spectral library matching considerably more difficult. Nevertheless, LC-MS is a very popular metabolomic platform and the analytical technology, spectral libraries, and software for processing spectra are continually being updated to improve metabolite identifications. See Yan et al. (2012) for an applied example of how LC-MS was used to identify species-specific metabolic stress responses of fish immediately after a tropical cyclone at various cage-farming sites, and to determine the physiological mechanisms which resulted in high mortalities during the following month of grow out.

Capillary Electrophoresis Mass Spectrometry

Capillary electrophoresis (CE) is an alternative pre-separation technique which separates metabolites based on their ionic charge characteristics, or electrophoretic mobility (reviewed by Ramautar et al. 2009). In many respects this technique is similar to LC, but molecules are separated based on their ionic affinities and size rather than on their solid phase solubilities (Hiryama et al. 2014). In CE, sample extracts enter a column which contains electrolytes. Charged metabolites migrate through the column and exit at different times under the influence of an electric field, and can be further concentrated using gradients in conductivity and pH. CE-MS is an efficient platform that does not require rigorous sample pre-treatment, is useful for small samples, is good at separating highly polar metabolites, has separation power and sensitivities which are comparable to GC-MS and LC-MS, and can quantify certain metabolites that other hyphenated MS platforms cannot. On the other hand, CE is unable to separate non-charged compounds, and suffers more than GC or LC from poor reproducibility. However, recent advances in CE-MS technologies are contributing to the increasing usage of the technique in metabolomics studies (Ramautar et al. 2015). See Koyama et al. (2015) for an applied example of how CE-MS was used to gain detailed metabolic insights into salinity adaption of brackish-water clams from four commercial fishery grounds in Japan with different water chemistries.

The sensitivity, or at least the detection limits, of MS techniques can be extremely high. As long as a substance can be separated, detection limits in parts per million or even better are possible. If pre-concentration techniques are used, molecules can easily be detected in concentrations of parts per trillion or better.

Vibrational spectroscopy

The analysis of complex sample matrices can also be performed using lower resolution instruments which measure the vibrational signatures of broad metabolite functional groups (Moore et al. 2014). Such analyses generally do not provide detailed information for identifying particular metabolites, but can still be very useful for obtaining an overall 'metabolite fingerprint' of a sample. This fingerprint is based on the holistic composition of functional group chemistries across all metabolites within the sample, and can be used to classify samples from different conditions when significant variations are observed. However, the drawback is that biological interpretation of spectra can be difficult because of this non-specificity. The application of vibrational-based technologies, such as Fourier transform infrared (FT-IR), near infrared (NIR), and Raman spectroscopy are growing in popularity due to their rapid and high through-put analysis capabilities, their ability to work with very small samples, and their very low cost compared to other platforms.

Infrared Spectroscopy

Infrared techniques work on the principal that when a sample is exposed to light, or electromagnetic radiation, the different chemical bonds within metabolite functional groups absorb energy at different wavelengths and vibrate in characteristic ways. A plot of the absorbance or transmittance of light at different wavelengths produces a spectrum which represents the overall metabolite composition of the sample to provide a snapshot, or fingerprint, of the organism's metabolome (Fig 5.). Infrared platforms are categorized into Near-Infrared (NIR) 0.78–3 µm and Mid-Infrared (MIR) 3–50 µm depending on the wavelength of light used to analyse the samples. Modern instruments commonly use Fourier transform techniques (a mathematical process which converts the raw data from the instrument into a spectrum) so the expression FT-IR is often seen when discussing MIR spectroscopy. MIR analysis examines the absorptions of bond vibrations and other molecular movements, whereas NIR evaluates the overtones

and combinations of strong MIR absorptions, which, while not as specific as the sharper, stronger MIR absorptions, can be characteristic and more easily quantified for a range of biologically important functional groups such as sugars, fats, and proteins. Unlike MIR, NIR can penetrate many millimetres through water and the instruments can use glass optics. Infrared platforms have proven useful for a range aquaculture-related purposes. For example, to identify pathogenic bacteria responsible for disease in farmed salmon (Wortberg et al. 2012), to determine the causation for post-harvest variations in shrimp quality based on the methods used for culling (Fu et al. 2014), to identify fraudulently marketed fish from different origins (Vidal et al. 2014), to assess the meat quality of various fish species (Cheng et al. 2013; Qu et al. 2015), to develop new food safety and authentication techniques for classifying shelled shrimp based on their post-harvest storage conditions (Qu et al. 2015), and to develop fast and cost-effective methods for proximate chemical analysis of cultured shellfish for the purposes of monitoring animal condition and assisting in selective breeding programs (Brown et al. 2012), among others.

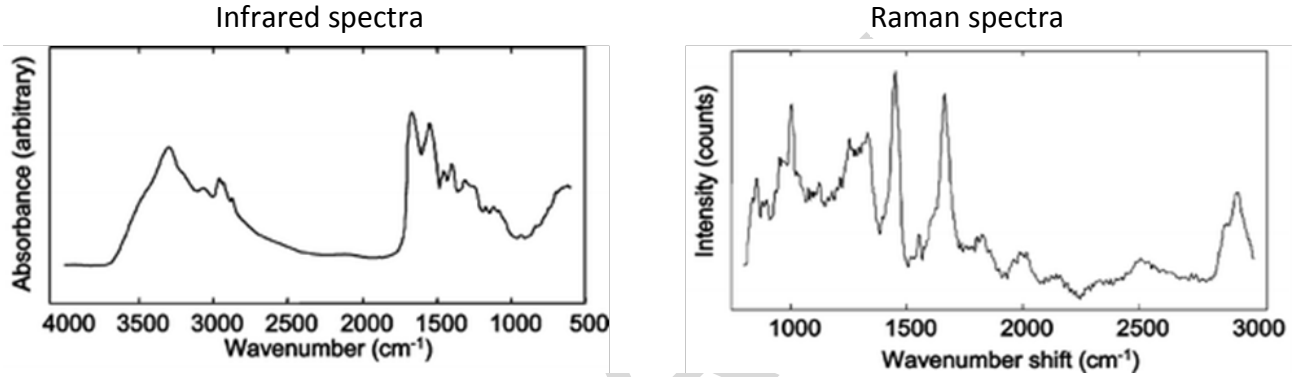

Figure 5. An example of comparative IR and Raman spectra obtained from the analysis of blood serum (reproduced from Ellis & Goodacre 2006).

Raman Spectroscopy

Raman spectroscopy is a technique closely related to MIR. When a laser beam hits a molecule, approximately 1 in 10^7 photons will interact with electrons in the chemical bonds resulting in the scattered laser light having extra wavelengths (a few nm) added to it and subtracted from it which correspond to the vibration frequencies of the bonds in the molecule. These shifts in photon wavelengths are called the 'Raman effect'. The original laser colour can be subtracted using filters and the remaining frequencies provide information about the vibrational, rotational and other low frequency transitions within metabolites. Raman spectra are closely related to MIR spectra and look very similar (Fig. 5). The principal difference is that the sorts of chemical bonds that give weak MIR absorptions are usually very strong in the Raman spectrum, and vice versa, so the two techniques are complementary. Although its big drawback is the very weak Raman signal, Raman spectroscopy has several major advantages. Glass optics can be used and since Raman spectroscopy is based on the scattering of incident light rather than on absorption, it does not suffer from interferences caused by water. Thus, measurements can be made directly on biofluids and aqueous extracts, minimal to no sample preparation is required, and spectra can be obtained very quickly. See Ishigaki et al. (2014) for an applied example of how non-invasive Raman spectroscopy was used on live fish eggs to predict and monitor their quality and viability to ensure successful fertilizations.

Metabolite fingerprinting vs. profiling

There are generally two approaches to generation and examination of metabolomics data – metabolite fingerprinting and metabolite profiling. The approach utilized depends largely on the objectives of the investigation and the facilities available.

Metabolite fingerprinting compares the overall nature of samples based on the entire set of signals generated by the analytical platform. These signals, or features, are analysed using statistical techniques to discern patterns in the data for the purpose of sample classification. This approach usually involves the analysis of a very large number of signals, which represents the total compositions of metabolites, and does not necessarily require metabolite identification. Data obtained from NMR and vibrational spectroscopy platforms are particularly well-suited for metabolite fingerprinting. However, interpretations of results within particular biological frameworks are limited unless further analyses of relevant features within the fingerprints are performed. Nevertheless, metabolite fingerprinting can be convenient for situations when only sample-class discrimination is required. For example, Aursand et al. (2007) used metabolite fingerprinting to reliably identify fraudulently mislabelled fish oil products to ensure food safety and develop novel techniques for food traceability and quality assurance.

Metabolite profiling evaluates all of the signals generated by the analytical platform so that they may be characterized and matched to spectra of known metabolites in reference libraries. Once identified, data analyses are then performed on the abundances of the metabolites within the samples. This approach provides data which can be more easily interpreted across various biological frameworks since features are ascribed an identity with often well-known biochemical roles. Metabolite profiling frequently leads to discovery of biomarkers and development of novel and testable hypotheses. For example, Guo et al. (2014) used metabolite profiling to identify early-warning biomarkers to predict fish health, and to better-understand the mechanisms of defence against bacterial infection.

Metabolite fingerprinting combined with profiling is sometimes used when very large numbers of signals are present within the raw spectral data so that only those features statistically different between samples, or otherwise deemed important, are subsequently identified. This approach can be used to reduce the computational and resource demands of processing noisy or large and complex datasets. In most metabolite fingerprinting applications, signals that are different between samples are usually identified to aid interpretation of the data. For example, Savorani et al. (2010) and Picone et al. (2011) used combined approaches to identify factors responsible for meat quality variation in fish reared under different culture environments, and stored under different post-harvest conditions.

NMR-based metabolomics usually involves fingerprinting as an initial step, whereas use of MS-based hyphenated platforms may involve profiling only. It is important to note that the definitions and term usage for these two approaches tend to be flexibly applied in the literature and as yet there are no standardized descriptions.

Data Analysis

Although metabolomic datasets are often very large and complex, recent advances in bioinformatics and streamlined statistical workflows provide simple strategies for coping with the high dimensional data (Johnson et al. 2015). Bioinformatics is an interdisciplinary field incorporating computer science, database management, mathematics and statistics. Primary bioinformatics processing involves analysis of the raw data obtained from the analytical platform and incorporates all procedures which are required to generate a list of features or metabolites. The resulting data can then be analysed by a range of classical and applied statistical procedures.

A number of steps are involved in the primary bioinformatics processing and usually includes data conversion, spectral processing (e.g., deconvolution, alignment, noise reduction), feature selection, metabolite identification *via* database matching, metabolite quantification, and quality control procedures. While a variety of freely or commercially available software packages exist to perform these tasks, many laboratories employ their own proprietary programs and algorithms, which have been custom designed for their unique situations and analytical set-ups. The methods used for primary bioinformatics processing vary widely and depend on data type and the analytical platform employed. Thus, it is impossible to provide general advice. However, at the end of this section we direct readers to a wide range of relevant literature, which covers these topics in more depth.

Prior to statistical analysis of metabolite profile/fingerprint data, data scaling, normalization and/or transformations are often performed to enhance extractability of biologically relevant

information from the dataset. Metabolite concentrations have huge dynamic ranges, and variance is typically larger at higher concentrations. Because many statistical procedures rely on homoscedasticity or distributional assumptions, it is important to alleviate the dependency of the variance on the concentration through variance-stabilizing transformation or transformation to normality. Furthermore, the relative abundances of different metabolites are not proportional to the biological importance that they may represent, and many data analysis techniques fail to take this into consideration. Some of the more commonly applied pre-treatment methods for metabolomics data include centering, autoscaling, pareto scaling, range scaling, log transformation, and power transformation (reviewed by van den Berg et al. 2006). The pre-treatment method chosen may vary between different metabolomics datasets; hence, a solid understanding of how the implemented method affects the outcome of subsequent statistical analyses is essential for reliable interpretations.

Statistical data analysis can be achieved using general statistical software (e.g., Minitab [Minitab Inc., PA, USA], SIMCA [Umetrics, Umea, Sweden], SPSS [IBM Corp., NY, USA], STATISTICA [Statsoft Inc., OK, USA]) or dedicated metabolomics-based data analysis packages (e.g., DeviumWeb [Grapov 2014], MeltDB [Kessler et al. 2013], Metaboanalyst [Xia et al. 2015]). A basic knowledge of programming is useful for employing, modifying, or writing script in certain data analysis environments (e.g., Matlab [Mathworks Inc., MA, USA], R [R Core Team 2014]). However, recent development of easy-to-use graphical user interfaces for these environments have substantially reduced the need for advanced programming skills. See Mishra and Van der Hoot (2016) for further information regarding the latest advances in available computational tools and resources for the analysis of metabolomics data. OMICtools (http://omictools.com/) is also a useful and growing online repository of web-accessible tools related to omics-based data analysis.

Similar to other –omics disciplines, it is common for the number of measured variables (genes, proteins or metabolites) within each sample to far exceed the number of samples analysed. Metabolomics data are by their very nature multivariate in design and lend themselves particularly well to multivariate statistical analyses. However, univariate techniques can also be employed to extract valuable information from the data. Use of both approaches in combination is routinely performed and recommended because they can expose different characteristics of the samples (Sugimoto et al. 2012).

Univariate methods

Univariate methods involve analysis of single variables (metabolites) at a time. T-tests and ANOVA's (analysis of variance), or their non-parametric equivalents (e.g. Mann-Whitney *U* test, Kruskal-Wallis test), are the most commonly applied univariate techniques to identify differences in metabolite abundances between samples. However, due to the high number of variables, it is important to correct for multiple hypothesis testing to protect against the likelihood of identifying false-positives (Broadhurst & Kell 2006). SAM (significant analysis of microarrays/metabolites) is an example of a univariate method which is able to account for correlations between metabolites and does not assume independence, unlike the T-test and ANOVA. Volcano plots are often used for the univariate analysis of gene, protein and metabolite expression data (Li 2012). Volcano plots are scatterplots which incorporate a measure of statistical significance (T-test p-values) with information about the magnitude of metabolite change (fold-change) (Fig. 6).

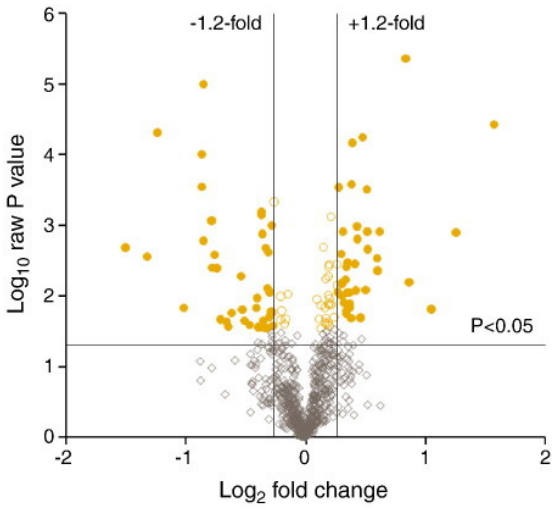

Figure 6. Example of a volcano plot. Solid yellow circles represent metabolites which are significantly different between sample groups (P<0.05), as well as have large variation (> 1.2 fold-change) in their mean abundances.

They allow quick identification of metabolites, which are not only statistically different between two sample conditions, but which also co-display large variations in abundance. See Young et al. (2015) for an applied example of how a volcano plot and SAM was used to assist construction of a multivariate classification model for assessing the quality of hatchery-reared mussel larvae.

Univariate methods are attractive because they are generally simple to apply and the results are easily interpreted and communicated across various levels of expertise. However, they cannot detect group differences when only minor variations exist on a single molecule level. Associations between metabolites and low variations in abundance can be highly important on a systems level due to the orchestrated flux of metabolites within common biochemical networks.

Multivariate methods

Since univariate techniques may not account for interrelations between metabolites, multivariate methods are applied to compensate, and to provide additional and complementary information for assisting interpretation of the data. Multivariate techniques can be used to reduce complexity and identify patterns, group structure, and relationships among metabolites and samples (Worley & Powers 2013). Commonly applied procedures include Principal Components Analysis (PCA), Projection to Latent Structures Discriminant Analysis (PLS-DA), and clustering.

Principal Components Analysis

PCA is a mathematical procedure that aims to capture and extract most of the important information in a high-dimensional data matrix and re-express it in fewer dimensions (Abdi & Williams 2010). In doing so, the data can be more easily visualized, described, and analysed. PCA does this by combining the multiple correlated variables into a number of smaller uncorrelated variables called principal components. A different data matrix is constructed in which the first 2–3 new variables account for the vast majority of the total variance in the original data. The samples can then be projected and visualized on a 2D or 3D score-plot (Fig. 7). PCA is an unsupervised statistical technique which incorporates only the independent metabolite information. Dependant variables are not required for modelling and information of sample class membership is not included in the analysis. As an unsupervised technique, patterns among the independent variables are discerned and groups of samples are formed based solely on the structure of the metabolite data. The PCA algorithm therefore achieves unbiased dimensionality reduction and only exposes group structure when within-group variation is substantially less than between-group variation. PCA is very useful for visualizing multi-dimensional data, identifying outliers, conducting classification studies, identifying a subset of original variables which explain most of the variation between samples, and for exploratory data analysis before building predictive models. See Kokushi et al. (2015) for an applied example of how PCA was used to identify differential regulation of metabolic pathways due to insecticide exposure in freshwater carp.

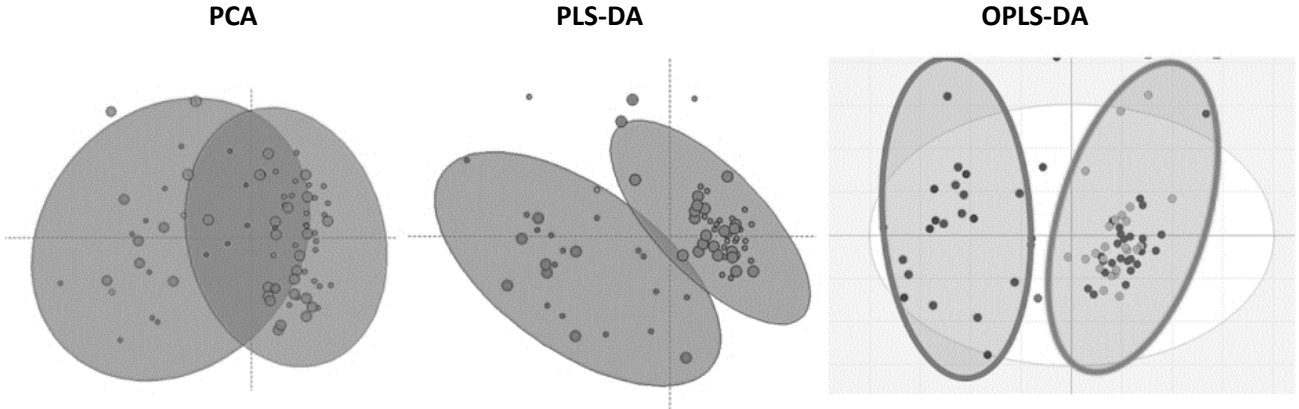

Figure 7. Comparison of multivariate data reduction techniques for assessing sample groupings using non-supervised principal component analysis (PCA) and supervised projection to latent structures discriminant analysis (PLS-DA), and its orthogonal extension (OPLS-DA). Discrimination power: PCA < PLS-DA < OPLS-DA.

Projection to Latent Structures Discriminant Analysis

Similar to PCA, PLS-DA is a technique, which can be used to reduce dimensionality and help visualize and analyze multivariate data (Worley & Powers 2013). However, PLS-DA is a supervised statistical technique, which incorporates information about the sample classes. Using this information, PLS-DA rotates the data within the newly created latent variable subspace in a way that maximizes separation between groups of samples. This can result in much clearer separations than when PCA is applied (Fig. 7). PLS-DA can be very useful for identifying and ranking metabolites which contribute most towards sample group separations and, when applied correctly, to assist construction of predictive classification models. Orthogonal PLS-DA is a related technique, which can further enhance separations due to its ability to distinguish between predictive and non-predictive (orthogonal) variation (Bylesjö et al. 2006) (Fig. 7).

PLS-DA and its extensions have a tendency to over-fit the model to the data. Therefore, validation is important when using these algorithms in predictive capacities. Model validation is the process of defining a model's performance and is a critical requirement for predictive modelling (Szymańska et al. 2012). This ensures that the model's internal variable rankings are truly informative. Commonly used methods to test a model's performance include permutation-based tests and cross validation (Worley & Powers 2013). The ideal scenario involves the use of a training dataset to build the model, and a separate validation dataset to assess its predictive capacity. See Liu et al. (2015) for an applied example of how PLS-DA and OPLS-DA were used to identify metabolites associated with white spot syndrome virus infection in shrimp, and provide preliminary information for developing biomarkers for diagnosing the pathophysiology of the disease.

Clustering

Clustering is a collection of statistical procedures, which aims to group samples together that are most similar in their metabolite profile (reviewed by Andreopoulos et al. 2009). Like PCA, most clustering techniques involve unsupervised approaches to group samples and the goal of clustering is to identify the actual groups based on the underlying structure of the data. Where PCA selects the variables with the most variation to form a reduced data matrix for partitioning samples, cluster analysis algorithms do not lose variance through dimensionality reduction in the same way and generally use all variables equally to display sample similarity/dissimilarity. Although clustering can be used to discover structures within the data irrespective of sample-class membership, it does not explain why they exist. Nevertheless, clustering is a very useful exploratory technique for uncovering patterns, finding natural groupings, confirming known groupings, identifying outliers, and discovering groups of metabolites with similar expression patterns across a wide range of biological conditions by clustering the variables rather than the samples. The most commonly applied clustering algorithms in metabolomics-based investigations are Hierarchical Cluster Analysis (HCA) and k-means clustering.

HCA is a method which seeks to construct a hierarchy of clusters and arrange them into a binary tree-structured graph called a dendrogram (Meunier et al. 2007). HCA does this by successively merging comparable groups based on the similarity/dissimilarity, or distance, between them. Visualizing this tree provides a useful summary of the data. HCA can be combined with data visualization techniques to provide new ways of looking at the data, and to enhance the extraction of important information (Fig.8). See Courant al. (2013) for an applied example of how HCA was combined with heatmap analysis to assist visualization of metabolite-group expressions, and to identify biomarkers for fine-scale monitoring during continuous culture of microalgae under different nitrogen regimes.

K-means is a non-hierarchical, unsupervised, partitional clustering approach. Although sample class membership information is not incorporated into the analysis, the researcher must initially define how many clusters (k number of clusters) into which the samples are to be partitioned. Like other clustering techniques, the aim is to gather samples into groups so that those in the same group are most similar to one another, and those in different groups are as different as possible. Working within an *n*-dimensional subspace of true vectors (number of variables), the algorithm performs this task through an iterative sequence of minimizing the sample distances to a centroid point

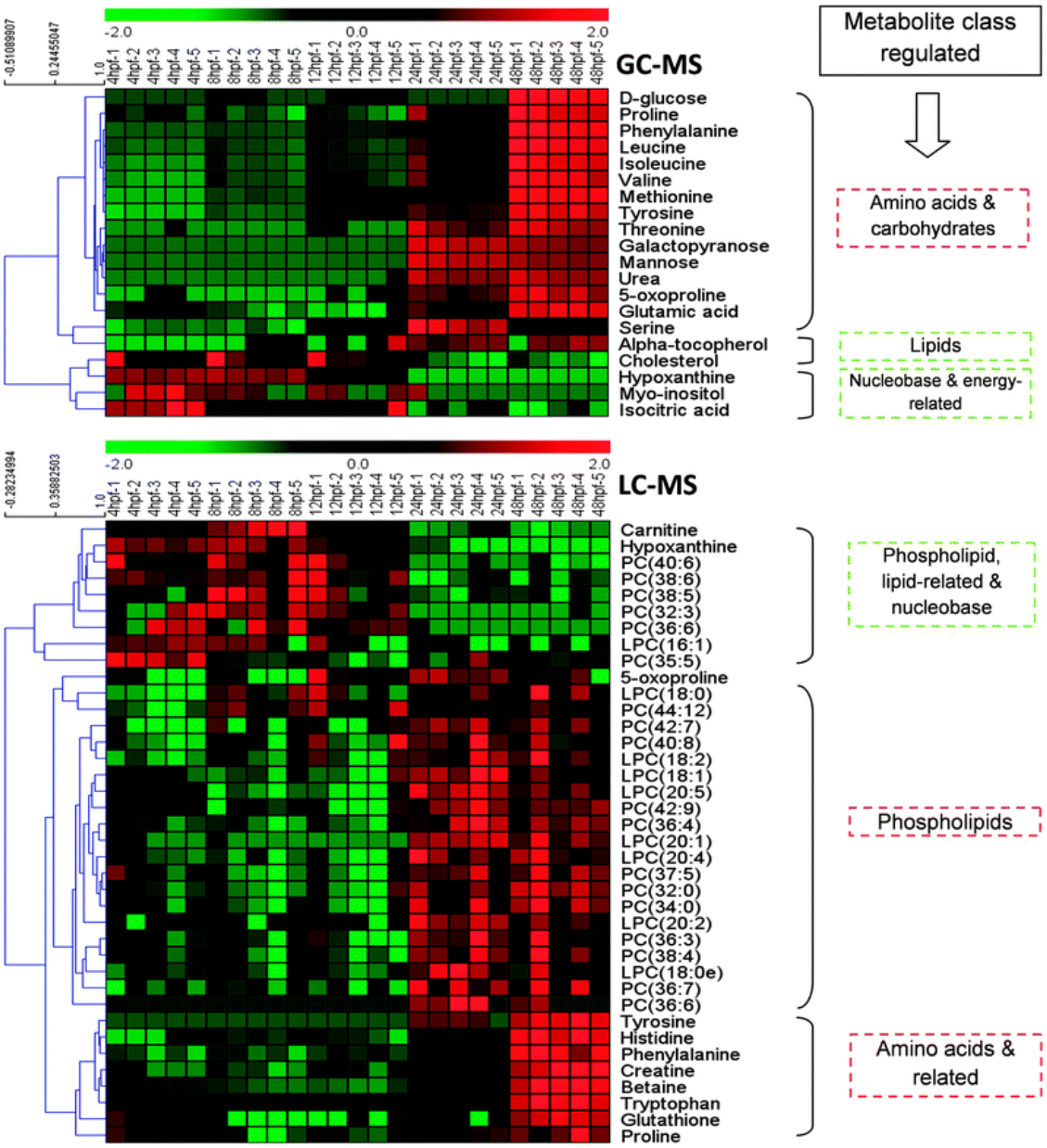

Figure 8. Combined heatmap and hierarchical cluster analysis of metabolites in developing zebrafish during embryogenesis *via* GC/MS- and LC/MS-based metabolomics (reproduced from Huang et al. 2013). Each column represents a sample (five biological replicates for each of the five development stages). Each row represents the abundance of a particular metabolite (red = high abundance, green = low abundance). Metabolites cluster naturally into groups which, in this case, have functional relationships (labelled metabolite classes in dotted boxes).

within each of the k number of clusters, and reallocating the samples to the cluster with the closest centroid so as to minimize the within-cluster sum of squares. Initially, the first centroid points are randomly placed and samples are assigned to a cluster. Then, the true centroid points of those clusters are calculated and repositioned, samples are reassigned, and the clusters are redefined. This is performed repeatedly until convergence is found. Use of k-means clustering in aquaculture-related metabolomics research is limited thus far. However, see Yu et al. (2013) for a relevant example of how k-means was used to identify groups of genes with similar expression profiles in fish which had been

fed a diet contaminated with the persistent organic pollutant BDE-47, and to determine potential enzymatic and metabolic mechanisms of toxicity defence.

While we have discussed three widely used multivariate techniques to analyse metabolomics data, one should be aware that an array of other procedures are available which may be better-suited for the analysis of particular datasets in some cases. These include: multivariate analysis of variance, linear discriminant analysis, partial least squares regression, support vector machines, k-nearest neighbour, random forests, soft independent modelling of class analogies, and self-organizing maps, among others. For further information on these alternative data analysis approaches, see Steuer et al. (2007), Liland (2011), and Xi et al. (2014). For platform-specific reviews on various bioinformatics processes see Smolinska et al. (2012), Sugimoto et al. (2012), Du & Zeisei (2013), Engel et al. (2013), and Wei et al. (2012 & 2014).

Biomarker discovery and validation

The aim of many metabolomics-based investigations is to discover novel metabolite biomarkers, which correlate with specific diseases or health states. These molecules can then be used as early diagnostic tools, or in conjunction with other assessments for confirmation of pathology. Metabolite biomarkers are also very useful within the aquaculture industry to assist in the evaluation of pre- and post-harvest meat quality, and for food safety and traceability purposes (Alfaro & Young 2015). Biomarkers may be single metabolites, multiple metabolites, ratios of metabolite pairs, particular features (e.g., ion fragments), or entire unannotated spectral fingerprints.

The initial step in biomarker discovery is often to perform an exploratory experiment with different treatments or animal conditions, and to identify features which are substantially different between the sample groups using one or more statistical procedures outlined in the previous section. Once determined, these features can be considered as 'candidate biomarkers'. The purpose of initial biomarker discovery is to identify the most salient features for further investigation, and may involve low biological sample replication ($n < 10$), although higher replication is usually preferred. The results of these studies can be very useful for generating hypotheses, and gaining preliminary mechanistic insights into metabolic factors responsible for, or involved in, particular health states or other conditions. However, when the ultimate goal is to develop practical biomarkers with useable applications and minimal risks for Type II errors occurring, they must have extremely reproducible performances. Thus, in order to ensure that the identified candidate biomarkers have high sensitivity and specificity for the particular condition under investigation, it is important to validate them.

The process of biomarker validation was born from the medical research field where the misdiagnosis of a health condition might result in a disastrous outcome for a patient. Biomarker validation is a quality assurance process of defining the performance of a biomarker within acceptable limits, whilst understanding and minimizing the rate of false discovery. Biomarker validation usually involves one or more additional experiments where the candidate metabolite/feature is targeted more specifically for quantification using complementary or alternative analytical platforms with a high selectivity for that analyte. Such experiments will typically also involve much higher sample replication ($n = 100–1000$), experimental replication, a broadening of scope in some cases (e.g., incorporation of multiple sexes, development stages, and environmental conditions), and rigorously refined statistical approaches (e.g. permutation based cross-validation or using different sub-sets of the samples to construct, validate, and test the performance of a predictive model [Westerhuis et al. 2008; Szymańska et al. 2012; Xia et al. 2013]). Accordingly, biomarker validation within the framework of a high-quality clinical study can involve substantial costs and time, which may not be viable for an environmental study or commercial aquaculture exercise. However, what constitutes validation is a subjective measure and is scalable within the confines of the researcher.

In a practical scenario where time and funds are limited, an alternative approach might be to employ a particular candidate biomarker whilst accepting its potential vulnerability, and continually adding new data to the predictive model as it becomes available. In this way, quality control limits on the model's performance can be set and monitored as it is updated. Thus, expenses are diluted over time

and performance-based milestones can be implemented based on cost-benefit analyses to guide management decisions in an empirically data driven context. Validation is an important concept when identifying and implementing new biomarkers, and should be a carefully considered component within the general strategy of a metabolomics-based biomarker discovery and development project. For further information on biomarker discovery and validation procedures, see Xia et al. (2013).

Biological interpretation & secondary bioinformatics

The discovery and identification of biomarkers do not always necessitate in-depth functional explanations for their presence and/or roles. For example, a simple metabolomics-based study using nuclear magnetic resonance and a fingerprinting approach combined with pattern-recognition tools (e.g., PCA, PLS-DA) could be used for food authentication purposes to identify biomarkers to classify an adulterated product, or determine its provenance (reviewed by Cubero-Leon et al. 2014). However, for many investigations, more detailed insights into the reasons for sample group separations are required, and the procurement of mechanistic biochemical explanations are highly desirable. In such cases, it becomes necessary to interpret the data within biologically meaningful frameworks.

The past 100 years of biological research has provided us with an amazing wealth of knowledge concerning cellular metabolism across a wide range of taxa. Rigorous empirical experimentation by a multitude of pioneers during this period has established the major biochemical pathways. Not only do we know which genes, enzymes, cofactors, substrates, products, and intermediates are involved in these pathways, in many cases we also know about individual enzyme kinetics and have detailed information about vast arrays of endogenous and exogenous factors which influence their pathway flux (German et al. 2005). Information such as this provides us with a rich source of knowledge which can be used to assist the interpretation of biochemical data. Nevertheless, interpretation of metabolite expression data can be one of the most challenging aspects of a metabolomics study.

In many cases, concentrations of particular metabolites within a tissue, biofluid, or organism may correlate very well with our current understanding of biochemical networks and the functional relationships among metabolites, enzymes, and genes within normal or perturbed systems. For example, classic signs of stress caused by pathogen or toxin exposure in aquatic animals include increased levels of reactive oxygen species (ROS), and differential co-expression of metabolites (e.g., glutathione, NADPH) and enzymes (e.g., glutathione reductase, superoxide dismutase, catalase) involved in regulating excess ROS production in order to maintain redox homeostasis (Parrilla-Taylor et al. 2013; Macías-Mayorga et al. 2015). The results of recent omics-based investigations provide data which corroborate the presence of such mechanisms in various taxa, as well as offer new information on associated regulatory pathways (Srivastava et al. 2013; Barth et al. 2014; Shi et al. 2015). On the other hand, a number of studies (particularly those containing metabolomics-based components) are providing data which are shedding light on unfamiliar biochemical associations which cannot be explained by our current theses of molecular biology and biochemistry (Steuer 2006). For example, unlike genes and proteins, it is relatively common for metabolite levels within a particular pathway to be highly correlated with metabolites from other pathways for which a mechanistic connection is not currently known. It is intriguing and unexpected results like these that are starting to deliver new information that is helping to push forward our understanding of metabolic networks at an astonishing rate, and also highlights the usefulness and efficiency of omics-based approaches for generating novel data to assist new interpretations. The continual development of metabolomic techniques to characterize larger and larger sets of metabolites requires new methods to analyse these data in order to obtain biologically meaningful information. Here, we briefly outline a few methods that can be used to help researchers interpret their metabolomics data beyond more conventional scenarios, involving assessments of single metabolite variations based on *a priori* biochemical knowledge.

If biological replication is sufficient, a simple method involves correlation analysis in which construction of a correlation matrix of pairwise

metabolite level comparisons are made. Such matrices can be useful for identifying potentially important relationships requiring further investigation. For example, consider the following hypothetical situation where levels of metabolite X and metabolite Y are not significantly different between control and treatment groups. However, metabolite X is positively correlated with metabolite Y in the control group, and negatively correlated in the treatment group (Fig. 9). Such a scenario would indicate that some major perturbation of the underlying network was taking place, and would have gone undetected had the correlations not been investigated. Differential non-linear correlation patterns may also be present which would require alternative methods of detection. For further information on the interpretation of linear and non-linear correlations in metabolomics data, see Camacho et al. (2005) and Steuer (2006).

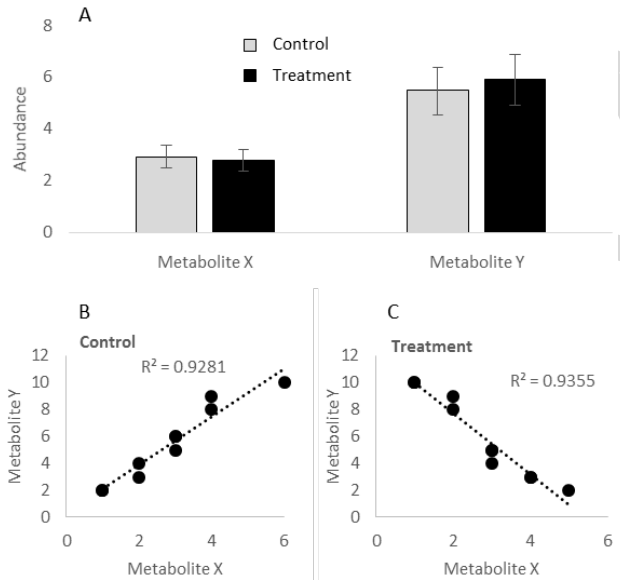

Figure 9. An example of a situation where mean levels of metabolite X and metabolite Y are not significantly different between groups of samples (A), but are differentially correlated within each group (B & C).

While our knowledge is relatively comprehensive compared to only a few decades ago, much of our understanding to date has come from highly targeted analyses of specific pathway components, and it is increasingly becoming clear that there are many gaps to be filled. With more of a focus on the interconnections between pathway components, we are starting to uncover new insights into metabolism which are much more integrated than ever before. An alternative method for identifying metabolite association patterns is called correlation network analysis.

Correlation networks are increasingly being used in omics-based applications to visually capture the overall network of interconnections between biomolecules and to describe the correlation patterns, to identify relationships between entire biochemical pathways, to discover new modules or clusters of relationships, and to assist data interpretation (Langfelder & Horvath 2008; Hero & Rajaratnam 2015). Applied to metabolomics, correlation network analysis is a technique that maps the relationships between every metabolite pair onto a metabolite network. Lines between metabolites typically are descriptive of the relationship between them (e.g., a solid line for a positive correlation and a dotted line for a negative correlation), and may also be quantitative (e.g., defined by the width of the line). The positions of the metabolites within the network map may be placed manually to enhance visualisation, or for additional interpretive purposes they may be positioned using algorithms to identify and define metabolite modules that cluster together. For an applied example of a study involving an aquatic organism, see Southam et al. (2008), who used a combination of correlation analysis techniques to identify key metabolic differences in hepatic tumors of flatfish compared to control tissues, and to assist detection and interpretation of the underlying mechanisms involved in the diseased phenotype (Fig. 10).

Correlation network analysis can additionally be used to integrate transcriptomic, proteomic and metabolomic datasets to help identify functional roles at different biochemical levels (e.g., gene-gene/protein interactions and relationships between enzymes and metabolites) (Higashi & Saito 2013). There are a number of software packages available to perform correlation network analysis, such as DPClus (Altaf-Ul-Amin et al. 2006), Metscape (Karnovsky et al. 2012), COVAIN (Sun & Weckworth 2012), 3Omics (Kuo et al. 2013), and MetaMapR (Grapov et al. 2015). For further information on correlation network analysis and various applications, see Steuer (2006), Adourian et al. (2008), Hüning et al. 2013, and Kotze et al. (2013).

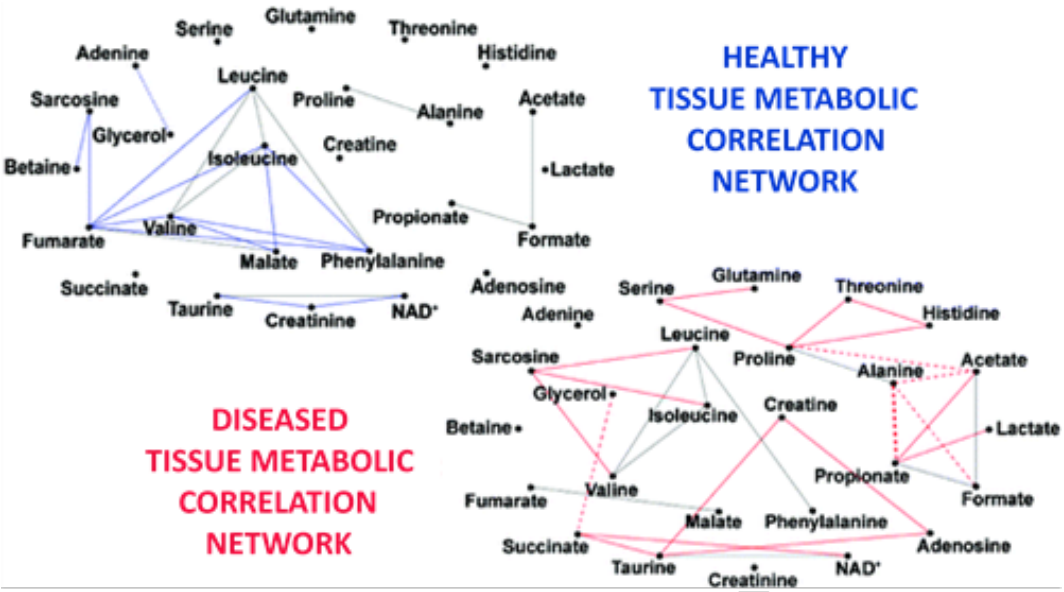

Figure 10. An example of two correlation networks constructed using NMR-based metabolomics data from samples of healthy and diseased fish livers (reproduced from Southam et al. 2008). Solid lines represent positive correlations between metabolites, and dotted lines represent negative correlations. Grey lines represent similarly shared relationships between the healthy and diseased phenotypes, and coloured lines represent those which are dissimilar. Clear differences in the underlying biochemical networks are easily visualised using this technique.

Other procedures useful for supporting data interpretation include a variety of 'pathway enrichment analysis' techniques. Rather than focusing on individual metabolites which may be responsible for discriminating groups of samples, pathway enrichment analysis techniques aim to discover predefined metabolic pathways or biological networks that are altered in an orchestrated manner. Such analyses make use of large amounts of biochemical information collated over decades and stored in publicly accessible depositories, such as the Kyoto Encyclopaedia of Genes and Genomes (KEGG), and can be considered as secondary bioinformatics processes. There may be cases where levels of individual compounds are not identified as being statistically different between samples using conventional statistical approaches. However, when analysed together as functional groups, or metabolite sets within their known pathways, it might be revealed that particular pathways as a whole are being differentially regulated under certain experimental conditions. The recent development of secondary bioinformatics tools (reviewed by Booth et al. 2013) to analyse biochemical data within the context of predefined metabolite sets are changing the way that the results of metabolomics projects are interpreted. Pathway Activity Profiling (PAPi) is one example of such a technique (Aggio et al. 2010).

PAPi is an algorithm developed into an R package which can be used to analyse sets of functionally-related metabolites, and quantitatively compare the activity of metabolic pathways between different groups of samples. PAPi performs this task by calculating 'activity scores' based on the number of metabolites identified from each pathway and their relative abundances. Pathways for which each detected metabolite is involved in is collected from KEGG, and each is given a score based on the absolute abundance/relative abundance of the metabolite to which it is linked. The pathways are ranked by the total number of metabolites they comprise, and the percentage of detected compounds within them are calculated. The sum of the scores for each pathway are then calculated and normalized (dividing by the proportion of metabolites detected from within the respective pathway) (Aggio et al. 2010). This simple yet effective method can be used to help determine the likelihood of a particular biochemical process being up- or down-regulated under certain circumstances. To our knowledge, PAPi has not yet

been applied to studies involving aquatic organisms. However, it has been successfully applied in a number of other biological systems (Han et al. 2012; Portella et al. 2014; Zhao C. et al. 2015).

Another useful pathway analysis tool is called Metabolite Set Enrichment Analysis (MSEA) (Xia & Wishart 2010; Kankainen et al. 2011; Persicke et al. 2012). MSEA is an algorithm designed to detect subtle, but consistent changes among groups of metabolites within the same biological pathway. Using an analysis package with MSEA capabilities (e.g., MarVis-Pathway [Kaever et al. 2014], MeltDB [Kessler et al. 2013], Metaboanalyst [Xia et al. 2015]), a quantitative dataset of annotated metabolites can be cross-referenced with information in the KEGG database, and metabolite sets belonging to reference pathways from various model organisms (e.g., human, mouse, zebrafish, drosophila, nematode) can be analysed together as a group. The ability to examine biochemical information for different animal models is a key advantage of MSEA, and options also exist to use proprietary/customised background sets of data from any organism. Pathway enrichment analysis techniques, which use software to interrogate databases that contain global biochemical knowledge, are tremendously powerful data interpretation tools. For applied examples see Zhao X. et al. (2015) and Ma et al. (2015), who utilized MSEA to identify differentially enriched pathways in Tilapia infected with two pathogenic *Streptococcus* species, and to develop remedial strategies to enhance disease resistance.

Table III. Summary of sample-specific topics (prior to chemical analysis) which should be described in detail when reporting the results of a metabolomics project.

Focus area	Descriptions
Sample	Taxonomic classifications, common name/s, genotype/s, ecotype/s, sample composition, sample type, specimen condition (phenotypic characteristics, weight, age, sex, development stage, health)
Environment	Any field environment: Geographic location, habitat, depth, meteorological conditions (e.g. precipitation, wind speed/direction, humidity), lunar/solar phase, other measured parameters (e.g. pollutant concentrations)
	Any aquatic environment: Water temperature, tidal phase (or submergence/emergence information), other measured parameters (e.g. salinity, pH, dissolved inorganic/organic content, oxygen concentration)
	Any laboratory environment: Details not covered elsewhere, laboratory address and contact information
Process (biological)	Maintenance and acclimation of organisms: Procedure and means (e.g. cage, aquaria, static/flowthrough tanks, continuous culture), reasons for maintenance/acclimation, other parameters (e.g. feeding regime, lighting regime, tank/cage dimensions)
	Manipulation of organisms/samples: Controlled manipulation as part of the study (e.g. exposure to a toxicant, environmental perturbation or dietary manipulation etc), dissection of a specific organ/tissue, capture/sampling means and procedures (e.g. netted, electrically stunned, anaesthetised, razor cut), reason for capture, other capture parameters (e.g. handling/stress aspects, time to capture, air exposure duration)
	Sample handling and storage/preservation: Procedure and means (e.g. snap frozen and stored in liquid nitrogen or on dry ice, sample container material), reasons for storage/preservation, temperature and duration of storage
	Organism/sample transportation: Procedures and means (e.g. live/dead, submerged/emerged, refrigerated container, dry shipper, temperature, transport duration)

Reporting guidelines in metabolomics

The final stage of a metabolomics project is to disseminate the findings, either internally through technical reports, or externally through peer-reviewed publication. Whichever route is taken, it is advised that researcher's follow to the best of their abilities a number of 'minimum reporting standards' which have been developed over the past decade by the wider metabolomics community (the Metabolomics Standards Initiative [http://www.metabolomics-msi.org/]). These readily available standards are 'highly recommended' guidelines for the reporting of various aspects of a metabolomics project, and provide a framework to ensure scientific rigour, allow study replication, support data sharing, and enable a better-informed process of assessment

and interpretation. In accordance with other biological science investigations, typical areas of focus include detailed descriptions of the biological sample/s involved in the study, descriptions of the environment/s involved in the study, and descriptions of biologically-relevant processes involved in the study (Table III) (see Morrison et al. 2007). Additional aspects to consider when reporting metabolomics-derived information include prescribing to the use of specific standard terms (ontology), providing particular details of a wide range of platform-specific instrumental parameters and data processing methods (computational, bioinformatics, statistical), quality-scoring metabolite identifications, and, among others, participating in the standards initiative to advance the future of the field by reporting and exchanging various levels of metadata with others. We strongly advise that all metabolomics researchers, from aspiring to seasoned investigators, become familiar with the recommended reporting guidelines for each component of a metabolomics project (summarised in Figure 11).

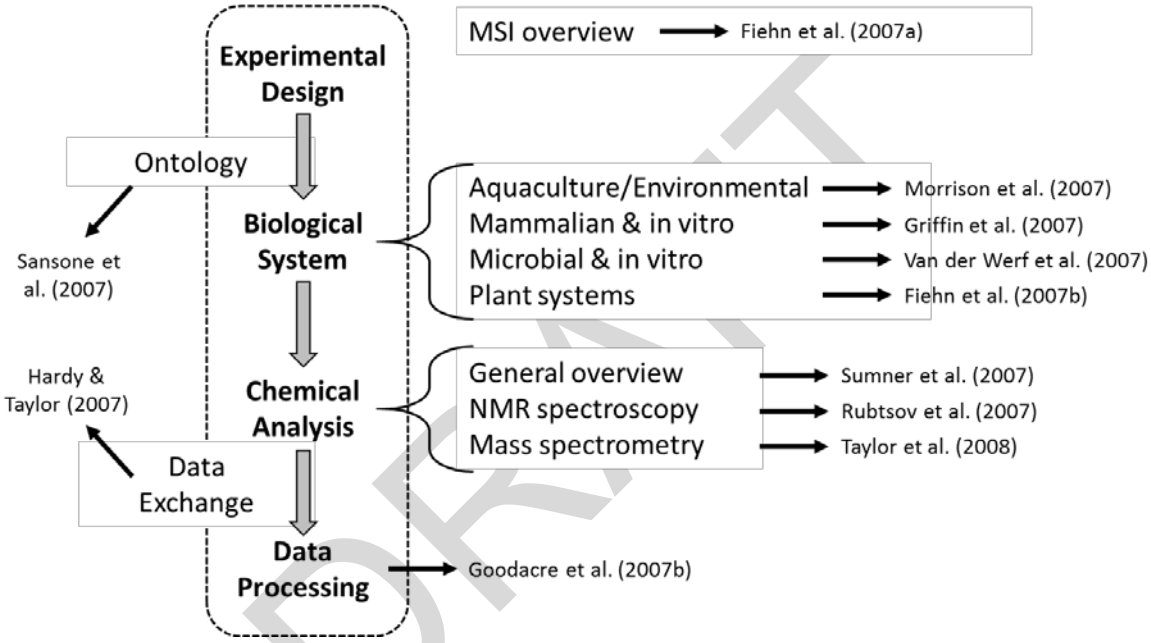

Figure 11. Overview of the metabolomics workflow showing the different components for which the Metabolomics Standards Initiative (MSI) have developed recommended minimum reporting standards (modified from Goodacre 2014).

Incorporating metabolomics

Two main avenues exist for researchers who wish to conduct metabolomics investigations, or add a metabolomics component to an existing research project. There are a number of commercial metabolomics laboratories worldwide that offer streamlined services. Core facilities at various universities and centres house a combination of infrastructure and expertise to carry out a range of advanced metabolomics studies. These organisations can provide excellent support from consultation on experimental design to data analysis and interpretation of results. Inevitably, significant costs are usually associated with such commercial services. Alternatively, access to metabolomics facilities can be gained through academic institutions for substantially reduced charges based on collaborative agreements. For scientists wanting to conduct metabolomics research for the first time, it is important to note that running a successful metabolomics project requires an adequate experience in chemistry, statistics, bioinformatics and the advice from a

metabolomics expert on hand. For researchers with sufficient chemistry knowledge and access to appropriate equipment and facilities, extraction and initial identification of metabolites may be relatively easy. However, there are some specific constraints in sample collection/preparation and experimental design that need to be considered. In addition, the bioinformatics required for data analysis and interpretation are significantly complex and may require the involvement of a bioinformatics expert. Regardless of the approach, we suggest that new metabolomics projects incorporate the appropriate expertise from the start. Furthermore, we urge scientists to give appropriate consideration to the expected results and implications of findings, since this approach is exploratory by nature.

Summary

In summary, metabolomics is a relatively new approach that has the potential to make a huge contribution to the field of aquaculture. With a wide range of analytical platforms available today and the rapidly evolving computational and bioinformatics capabilities, we are likely to see a growing number of studies using metabolomics in all aspects of cultivating aquatic organisms.

However, it is important to be aware of the potential limitations of this approach, especially with regard to sensitivity to external influences during sample collection and complex bioinformatics procedures required to obtain meaningful biological interpretations. We are still at an early stage in the application of metabolomics in aquaculture, and it is envisioned that more streamlined procedures and strategies will be generated in the coming years to facilitate implementation of this powerful approach. Some of those advances will involve the development of extensive metabolite biomarker libraries, easy-to-use bioinformatics packages, small robust analytical platforms for use in the field, and improvements in analytical sensitivities and metabolite coverage. But more importantly, our future challenge will no doubt be to translate the clear potential of this approach into practical solutions to significantly improve the commercial aquaculture sector.

Aknowledgements

We are thankful for the valuable comments of J. Robertson, M. Rohan, and K. Campbell on an earlier version of this manuscript. We also acknowledge support from an Auckland University of Technology Vice Chancellor Doctoral Scholarship to T. Young under the supervision of A.C. Alfaro.

REFERENCES

Abdi H, Williams LJ (2010) Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics* 2(4): 433–459.

Abro R, Moazzami AA, Lindberg JE, Lundh T (2014) Metabolic insights in Arctic charr (*Salvelinus alpinus*) fed with zygomycetes and fish meal diets as assessed in liver using nuclear magnetic resonance (NMR) spectroscopy. *International Aquatic Research* 6(2): 1–11.

Adourian A, Jennings E, Balasubramanian R, Hines WM, Damian D, Plasterer TN, et al. (2008) Correlation network analysis for data integration and biomarker selection. *Molecular BioSystems* 4(3): 249–259.

Aggio RB, Ruggiero K, Villas-Bôas SG (2010) Pathway Activity Profiling (PAPi): From the metabolite profile to the metabolic pathway activity. *Bioinformatics* 26(23): 2969–2976.

Alfaro AC, Young T (2016) Showcasing metabolomic applications in aquaculture: A review (submitted). *Reviews in Aquaculture.*

Allen PJ, Wise D, Greenway T, Khoo L, Griffin MJ, Jablonsky M (2015) Using 1-D ^{1}H and 2-D ^{1}H J-resolved NMR metabolomics to understand the effects of anemia in channel catfish (*Ictalurus punctatus*). *Metabolomics* 11(5): 1131–1143.

Allwood JW, Winder CL, Dunn WB, Goodacre R (2013) Considerations in sample preparation, collection, and extraction approaches applied in microbial, plant, and mammalian metabolic profiling. In *Methodologies for Metabolomics: Experimental Strategies and Techniques* (Eds. Lutz NW., Sweedler JV, Wevers RA). pp 79–118. Cambridge University Press, Cambridge, United Kingdom. ISBN: 9780521765909.

Altaf-Ul-Amin M, Tsuji H, Kurokawa K, Asahi H, Shinbo Y, Kanaya S (2006) DPClus: A density-periphery based graph clustering software mainly focused on detection of protein complexes in interaction networks. *Journal of Computer Aided Chemistry* 7:150–156.

Álvarez-Sánchez B, Priego-Capote F, de Castro ML (2010a) Metabolomics analysis I. Selection of biological samples and practical aspects preceding sample preparation. *TrAC Trends in Analytical Chemistry* 29(2): 111–119.

Álvarez-Sánchez B, Priego-Capote F, de Castro ML (2010b) Metabolomics analysis II. Preparation of biological samples prior to detection. *TrAC Trends in Analytical Chemistry* 29(2): 120–127.

Anderson SM, Taylor R, Holen E, Aksnes A, Espe M (2014) Arginine supplementation and exposure time affects polyamine and glucose metabolism in primary liver cells isolated from Atlantic salmon. *Amino Acids* 46(5): 1225–1233.

Andreopoulos B, An A, Wang X, Schroeder M (2009) A roadmap of clustering algorithms: Finding a match for a biomedical application. *Briefing in Bioinformatics* 10(3): 297–314.

Asakura T, Sakata K, Yoshida S, Date Y, Kikuchi J (2014) Noninvasive analysis of metabolic changes following nutrient input into diverse fish species, as investigated by metabolic and microbial profiling approaches. *PeerJ* 2:e550. http://dx.doi.org/10.7717/peerj.550

Aursand M, Standal IB, Axelson DE (2007) High-resolution ^{13}C nuclear magnetic resonance spectroscopy pattern recognition of fish oil capsules. *Journal of Agricultural and Food Chemistry*, 55(1): 38–47.

Aursand M, Standal IB, Prael A, McEvoy L, Irvine J, Axelson DE (2009) ^{13}C NMR pattern recognition techniques for the classification of Atlantic salmon (*Salmo salar* L.) according to their wild, farmed, and geographical origin. *Journal of Agricultural and Food Chemistry* 57(9): 3444–3451.

Bajhaiya AK, Dean AP, Driver T, Trivedi DK, Rattray NJ, Allwood JW, et al. (2016) High-throughput metabolic screening of microalgae genetic variation in response to nutrient limitation. *Metabolomics* 12(1): 1-14.

Bando K, Kawahara R, Kunimatsu T, Sakai J, Kimura J, Funabashi H, et al. (2010) Influences of biofluid sample collection and handling procedures on GC–MS based metabolomic studies. *Journal of Bioscience and Bioengineering* 110(4): 491–499.

Bartel J, Krumsiek J, Theis FJ (2013) Statistical methods for the analysis of high-throughput metabolomics data. *Computational and Structural Biotechnology Journal* 4(5): 1–9.

Barth J, Bergner SV, Jaeger D, Niehues A, Schulze S, Scholz M, et al. (2014) The interplay of light and oxygen in the reactive oxygen stress response of *Chlamydomonas reinhardtii* dissected by quantitative mass spectrometry. *Molecular & Cellular Proteomics* 13(4): 969–989.

Barupal DK, Haldiya PK, Wohlgemuth G, Kind T, Kothari SL, Pinkerton KE, et al. (2012) MetaMapp: Mapping and visualizing metabolomic data by integrating information from biochemical pathways and chemical and mass spectral similarity. *BMC Bioinformatics* 13(1): 99–114.

Bastian M, Heymann S, Jacomy M (2009) *Gephi: An Open Source Software for Exploring and Manipulating Networks*. International AAAI Conference on Weblogs and Social Media: San Jose, California.

Beckonert O, Keun HC, Ebbels TM, Bundy J, Holmes E, Lindon JC, et al. (2007) Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. *Nature Protocols* 2(11): 2692–2703.

Benskin JP, Ikonomou MG, Liu J, Veldhoen N, Dubetz C, Helbing CC, et al. (2014) Distinctive metabolite profiles in in-migrating Sockeye salmon suggest sex-linked endocrine perturbation. *Environmental Science & Technology*, 48(19): 11670–11678.

Bharti SK, Roy R (2014) Metabolite identification in NMR-based metabolomics. *Current Metabolomics* 2(3): 163–173.

Booth SC, Weljie AM, Turner RJ (2013) Computational tools for the secondary analysis of metabolomics experiments. *Computational and Structural Biotechnology Journal* 4(5): 1–13.

Broadhurst DI, Kell DB (2006) Statistical strategies for avoiding false discoveries in metabolomics and related experiments. *Metabolomics* 2(4): 171–196.

Bromke MA, Sabir JS, Alfassi FA, Hajarah NH, Kabli SA, Al-Malki AL, et al. (2015) Metabolomic profiling of 13 diatom cultures and their adaptation to nitrate-limited growth conditions. *PloS ONE* 10(10): e0138965.

Brown MR, Kube PD, O'Connor S, Cunningham M, King H (2012) Application of near-infrared reflectance spectroscopy for the rapid chemical analysis of Sydney rock oyster (*Saccostrea Glomerata*) and Pacific oyster (*Crassostrea gigas*). *Journal of Shellfish Research* 31(4): 1051–1060.

Bussell JA, Gidman EA, Causton DR, Gwynn-Jones D, Malham SK, Jones MLM, et al. (2008) Changes in the immune response and metabolic fingerprint of the mussel, *Mytilus edulis*

(Linnaeus) in response to lowered salinity and physical stress. *Journal of Experimental Marine Biology and Ecology* 358(1): 78–85.

Bylesjö M, Rantalainen M, Cloarec O, Nicholson JK, Holmes E, Trygg J (2006) OPLS discriminant analysis: Combining the strengths of PLS-DA and SIMCA classification. *Journal of Chemometrics* 20(8-10): 341–351.

Cajka T, Danhelova H, Vavrecka A, Riddellova K,Kocourek V, Vacha F, et al. (2013) Evaluation of direct analysis in real time ionization-mass spectrometry (DART-MS) in fish metabolomics aimed to assess the response to dietary supplementation. *Talanta* 115: 263–270.

Camacho D, de la Fuente A, Mendes P (2005) The origin of correlations in metabolomics data. *Metabolomics* 1(1): 53–63.

Cappello T, Mauceri A, Corsaro C, Maisano M, Parrino V, Lo Paro G, et al. (2013) Impact of environmental pollution on caged mussels *Mytilus galloprovincialis* using NMR-based metabolomics. *Marine Pollution Bulletin* 77(1–2): 132–139.

Castro C, Corraze G, Panserat S, Oliva-Teles A (2015) Effects of fish oil replacement by a vegetable oil blend on digestibility, postprandial serum metabolite profile, lipid and glucose metabolism of European sea bass (*Dicentrarchus labrax*) juveniles. *Aquaculture Nutrition* 21(5): 592–603.

Chauton MS, Galloway TF, Kjørsvik E, Størseth TR, Puvanendran V, van der Meeren T, et al. (2015) ^{1}H NMR metabolic profiling of cod (Gadus morhua) larvae: Potential effects of temperature and diet composition during early developmental stages. *Biology Open* DOI: 10.1242/bio.014431

Chen S, Zhang C, Xiong Y, Tian X, Liu C, Jeevithan E, Wu W (2015) A GC-MS-based metabolomics investigation on scallop (*Chlamys farreri*) during semi-anhydrous living-preservation. *Innovative Food Science & Emerging Technologies* 31: 185–195.

Cheng JH, Dai Q, Sun DW, Zeng XA, Liu D, Pu HB (2013) Applications of non-destructive spectroscopic techniques for fish quality and safety evaluation and inspection. *Trends in Food Science & Technology* 34(1): 18–31.

Cheng K, Wagner L, Moazzami AA, Gómez-Requeni P, Vestergren AS, Brännäs E, et al. (2015a) Decontaminated fishmeal and fish oil from the Baltic Sea are promising feed sources for Arctic char (*Salvelinus alpinus* L.)—studies of flesh lipid quality and metabolic profile. *European Journal of Lipid Science and Technology* DOI: 10.1002/ejlt.201500247

Cipriano RC, Smith ML, Vermeersch KA, Dove AD, Styczynski MP (2015) Differential metabolite levels in response to spawning-induced inappetence in Atlantic salmon *Salmo salar*. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics* 13: 52–59.

Conceição LE, Morais S, Rønnestad I (2007) Tracers in fish larvae nutrition: A review of methods and applications. *Aquaculture* 267(1): 62–75.

Connor KM, Gracey AY (2012) High-resolution analysis of metabolic cycles in the intertidal mussel *Mytilus californianus*. *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology* 302(1): R103–R111.

Courant F, Antignac JP, Dervilly-Pinel G, Le Bizec B (2014) Basics of mass spectrometry based metabolomics. *Proteomics* 14(21-22): 2369–2388.

Courant F, Martzolff A, Rabin G, Antignac JP, Le Bizec B, Giraudeau P, et al. (2013) How metabolomics can contribute to bio-processes: A proof of concept study for biomarkers discovery in the context of nitrogen-starved microalgae grown in photobioreactors. *Metabolomics* 9(6): 1286–1300.

Cubero-Leon E, Minier C, Rotchell JM, Hill EM (2012) Metabolomic analysis of sex specific metabolites in gonads of the mussel, *Mytilus edulis*. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics* 7(2): 212–219.

Cubero-Leon E, Peñalver R, Maquet A (2014) Review on metabolomics for food authentication. *Food Research International* 60: 95–107.

Daub CO, Kloska S, Selbig J (2003) MetaGeneAlyse: Analysis of integrated transcriptional and metabolite data. *Bioinformatics* 19: 2332–2333

del Carmen Alvarez M, Donarski JA, Elliott M, Charlton AJ (2010) Evaluation of extraction methods for use with NMR-based metabolomics in the marine polychaete ragworm, *Hediste diversicolor*. *Metabolomics* 6(4): 541–549.

Dettmer K, Aronov PA, Hammock BD (2007) Mass spectrometry-based metabolomics. Mass Spectrometry Reviews 26(1): 51–78

Dominguez-Vidal A, Pantoja-de la Rosa J, Cuadros-Rodríguez L, Ayora-Cañada MJ (2016) Authentication of canned fish packing oils by means of Fourier transform Infrared spectroscopy. *Food Chemistry* 190: 122–127.

Dunphy BJ, Watts E, Ragg NL (2015) Identifying thermally-stressed adult green-lipped mussels (*Perna canaliculus* Gmelin, 1791) via metabolomic profiling. *American Malacological Bulletin* 33(1): 127–135.

Ekman DR, Skelton DM, Davis JM, Villeneuve DL, Cavallin JE, Schroeder A, et al. (2015) Metabolite profiling of fish skin mucus: A novel approach for minimally-invasive environmental exposure monitoring and surveillance. *Environmental Science & Technology* DOI: 10.1021/es505054fEkmann KS, Dalsgaard J, Holm J, Campbell PJ, Skov PV (2013) Effects of dietary energy density and digestible protein: energy ratio on de novo lipid synthesis from dietary protein in gilthead sea bream (*Sparus aurata*) quantified with stable isotopes. *British Journal of Nutrition* 110(10): 1771–1781.

El-Aneed A, Cohen A, Banoub J (2009) Mass spectrometry, review of the basics: Electrospray, MALDI, and commonly used mass analyzers. *Applied Spectroscopy Reviews* 44(3): 210–230.

Ellis RP, Spicer JI, Byrne JJ, Sommer U, Viant MR, White DA, et al. (2014) ^{1}H NMR metabolomics reveals contrasting response by male and female mussels exposed to reduced seawater pH, increased temperature, and a pathogen. *Environmental Science & Technology* 48(12): 7044–7052.

Engel J, Gerretzen J, Szymańska E, Jansen JJ, Downey G, Blanchet L, et al. (2013) Breaking with trends in pre-processing? *TrAC Trends in Analytical Chemistry* 50: 96–106.

Fernández-Varela R, Tomasi G, Christensen JH (2015) An untargeted gas chromatography mass spectrometry metabolomics platform for marine polychaetes. *Journal of Chromatography A* 1384: 133–141.

Fiehn O, Robertson D, Griffin J, van der Werf M, Nikolau B, Morrison N, et al. (2007a) The metabolomics standards initiative (MSI). *Metabolomics* 3: 175–178.

Fiehn O, Sumner LW, Rhee SY, Ward J, Dickerson J, Lange BM, et al. (2007b). Minimum reporting standards for plant biology context information in metabolomic studies. *Metabolomics* 3 195–201.

Fructuoso S, Sevilla Á, Bernal C, Lozano AB, Iborra JL, Cánovas M (2012) EasyLCMS: An asynchronous web application for the

automated quantification of LC-MS data. *BMC Research Notes* 5(1): 428–439.

Fu LL, Chen X, Wang Y (2014) Quality evaluation of farmed whiteleg shrimp, *Litopenaeus vannamei*, treated with different slaughter processing by infrared spectroscopy. *Food chemistry* 151: 306–310.

Garcia A, Barbas C (2011) Gas chromatography-mass spectrometry (GC-MS)-based metabolomics. *Methods in Molecular Biology* 708: 191–204.

German JB, Hammock BD, Watkins SM (2005) Metabolomics: Building on a century of biochemistry to guide human health. *Metabolomics* 1(1): 3–9.

Gidman EA, Jones MLM, Bussell JA, Malham SK, Reynolds B, Seed R, et al. (2007) A methodology for screening haemolymph of intertidal mussels, *Mytilus edulis*, using FT-IR spectroscopy as a tool for environmental assessment. *Metabolomics* 3(4): 465–473.

Glish GL, Vachet RW (2003) The basics of mass spectrometry in the twenty-first century. *Nature Reviews Drug Discovery* 2(2): 140–150.

Goodacre R. (2014) Water, water, every where, but rarely any drop to drink. *Metabolomics* 10: 5–7.

Goodacre R, Broadhurst D, Smilde A, Kristal BS, Baker JD, Beger R, et al. (2007) Proposed minimum reporting standards for data analysis in metabolomics. *Metabolomics* 3: 231–241.

Gooley JJ (2014) Applications of circadian metabolomics. *Current Metabolomics*, *2*(1): 2–14.

Goulitquer S, Potin P, Tonon T (2012) Mass spectrometry-based metabolomics to elucidate functions in marine organisms and ecosystems. *Marine Drugs* 10(4): 849–880.

Grapov D (2014) DeviumWeb: Dynamic Multivariate Data Analysis and Visualization Platform v0.4. DOI: 10.5281/zenodo.12879,https://github.com/dgrapov/DeviumWebGrapov D, Wanichthanarak K, Fiehn O (2015) MetaMapR: pathway independent metabolomic network analysis incorporating unknowns. *Bioinformatics* DOI: 10.1093/bioinformatics/btv194

Griffin JL, Nicholls AW, Daykin C, Heald S, Keun H, Schuppe-Koistinen I, et al. (2007) Standard reporting requirements for biological samples in metabolomics experiments: Mammalian/in vivo experiments. *Metabolomics* 3: 179–188.

Guo C, Huang X-Y, Yang M-J, Wang S, Ren S-T, Peng X-X (2014) GC/MS-based metabolomics approach to identify biomarkers differentiating survivals from death in crucian carps infected by *Edwardsiella tarda*. *Fish & Shellfish Immunology* 39(2): 215–222.

Guo C, Peng B, Song M, Wu CW, Yang MJ, Zhang JY, et al. (2015) Live *Edwardsiella tarda* vaccine enhances innate immunity by metabolic modulation in zebrafish. *Fish & Shellfish Immunology* 47(2): 664–673.

Gupta V, Thakur R S, Reddy CRK, Jha B (2013) Central metabolic processes of marine macrophytic algae revealed from NMR based metabolome analysis. *RSC Advances* 3(19): 7037–7047.

Han TL, Cannon RD, Villas-Bôas SG (2012) Metabolome analysis during the morphological transition of *Candida albicans*. *Metabolomics* 8(6): 1204–1217.

Hardy NW, Taylor CF (2007) A roadmap for the establishment of standard data exchange structures for metabolomics. *Metabolomics* 3: 243–248.

Hero A, Rajaratnam B (2015) *Large Scale Correlation Mining for Biomolecular Network Discovery*. Technical Report No. 2015-02, Department of Statistics, Stanford University, California, USA.

Higashi Y, Saito K (2013) Network analysis for gene discovery in plant-specialized metabolism. *Plant, Cell & Environment* 36(9): 1597–1606.

Hines A, Oladiran GS, Bignell JP, Stentiford GD, Viant MR (2007) Direct sampling of organisms from the field and knowledge of their phenotype: Key recommendations for environmental metabolomics. *Environmental Science & Technology* 41(9): 3375–3381.

Hirayama A, Wakayama M, Soga T (2014) Metabolome analysis based on capillary electrophoresis-mass spectrometry. *TrAC Trends in Analytical Chemistry* 61: 215–222.

Howe E, Holton K, Nair S, Schlauch D, Sinha R, Quackenbush J (2010) MeV: Multiexperiment Viewer. In *Biomedical Informatics for Cancer Research* (pp. 267–277). Springer US.

Huang SM, Xu F, Lam SH, Gong Z, Ong CN (2013) Metabolomics of developing zebrafish embryos using gas chromatography- and liquid chromatography-mass spectrometry. *Molecular BioSystems* 9(6): 1372–1380.

Hüning AK, Melzner F, Thomsen J, Gutowska MA, Krämer L, Frickenhaus S, et al. (2013) Impacts of seawater acidification on mantle gene expression patterns of the Baltic Sea blue mussel: Implications for shell formation and energy metabolism. *Marine Biology* 160(8): 1845–1861.

Hurley-Sanders JL, Stoskopf MK, Nelson SA, Showers W, Law JM, Gracz HS, et al. (2015a) Tissue extraction methods for metabolic profiling of a freshwater bivalve, *Elliptio complanata* (Lightfoot, 1786). *American Malacological Bulletin* 33(2): 1–10.

Hurley-Sanders JL, Levine JF, Nelson SA, Law JM, Showers WJ, Stoskopf MK (2015b) Key metabolites in tissue extracts of *Elliptio complanata* identified using ^{1}H nuclear magnetic resonance spectroscopy. *Conservation Physiology* 3(1): cov023.

Ibáñez C, García-Cañas V, Valdés A, Simó C (2014) Direct mass spectrometry-based approaches in metabolomics. In Barcelo D. (Ed.) *Fundamentals of Advanced Omics Technologies: From Genes to Metabolites* (pp. 235–254): Elsevier, Oxford, UK.

Ishigaki M, Taketani A, Sato H (2014) Discrimination of fish egg quality and viability by Raman spectroscopy. *Analytical Methods* 6(23): 9206–9211.

Ito K, Sakata K, Date Y, Kikuchi J (2014) Integrated analysis of seaweed components during seasonal fluctuation by data mining across heterogeneous chemical measurements with network visualization. *Analytical Chemistry* 86(2): 1098–1105.

Ji C, Li F, Wang Q, Zhao J, Sun Z, Wu H (2016) An integrated proteomic and metabolomic study on the gender-specific responses of mussels *Mytilus galloprovincialis* to tetrabromobisphenol A (TBBPA). *Chemosphere* 144: 527–539.

Ji C, Wu H, Wei L, Zhao J, Wang Q, Lu H (2013) Responses of *Mytilus galloprovincialis* to bacterial challenges by metabolomics and proteomics. *Fish & Shellfish Immunology* 35(2): 489–498.

Ji C, Wu H, Wei L, Zhao J, Yu J (2013) Proteomic and metabolomic analysis reveal gender-specific responses of mussel *Mytilus galloprovincialis* to 2,2',4,4'-tetrabromodiphenyl ether (BDE 47). *Aquatic Toxicology* 140: 449–457.

Jin Y, Tian LX, Xie SW, Guo DQ, Yang HJ, Liang GY, et al.(2015) Interactions between dietary protein levels, growth

performance, feed utilization, gene expression and metabolic products in juvenile grass carp (*Ctenopharyngodon idella*). *Aquaculture* 437: 75–83.

Johnson CH, Ivanisevic J, Benton HP (2015) Bioinformatics: The next frontier of metabolomics. *Analytical Chemistry* 87(1): 147–156.

Junot C, Fenaille F, Colsch B, Bechner F (2014) High resolution mass spectrometry based techniques at the crossroads of metabolic pathways. *Mass Spectrometry Reviews* 33(6): 471–500.

Kaever A, Landesfeind M, Feussner K, Mosblech A, Heilmann I, Morgenstern B, et al. (2014) MarVis-Pathway: Integrative and exploratory pathway analysis of non-targeted metabolomics data. *Metabolomics* 1–14. DOI: 10.1007/s11306-014-0734-y

Kamburov A, Cavill R, Ebbels TM, Herwig R, Keun HC (2011) Integrated pathway-level analysis of transcriptomics and metabolomics data with IMPaLA. *Bioinformatics* 27(20): 2917–2918.

Kankainen M, Gopalacharyulu P, Holm L, Orešič M (2011) MPEA—metabolite pathway enrichment analysis. *Bioinformatics* 27(13): 1878–1879.

Karakach TK, Huenupi EC, Soo EC, Walter JA, Afonso LO (2009) ^{1}H-NMR and mass spectrometric characterization of the metabolic response of juvenile Atlantic salmon (*Salmo salar*) to long-term handling stress. *Metabolomics* 5(1): 123–137.

Karnovsky A, Weymouth T, Hull T, Tarcea VG, Scardoni G, Laudanna C, et al. (2012) Metscape 2 bioinformatics tool for the analysis and visualization of metabolomics and gene expression data. *Bioinformatics* 28(3): 373–380.

Katajamaa M, Miettinen J, Orešič M (2006) MZmine: Toolbox for processing and visualization of mass spectrometry based molecular profile data. *Bioinformatics* 22(5): 634–636.

Keller BO, Sui J, Young AB, Whittal RM (2008) Interferences and contaminants encountered in modern mass spectrometry. *Analytica Chimica Acta* 627(1): 71–81.

Kessler N, Neuweger H, Bonte A, Langenkämper G, Niehaus K, Nattkemper TW, et al. (2013) MeltDB 2.0–advances of the metabolomics software system. *Bioinformatics* 29(19): 2452–2459.

Kokushi E, Uno S, Pal S, Koyama J (2015) Effects of chlorpyrifos on the metabolome of the freshwater carp, *Cyprinus carpio*. *Environmental Toxicology* 30(3): 253–260.

Kotze HL, Armitage EG, Sharkey KJ, Allwood JW, Dunn WB, Williams KJ, et al. (2013) A novel untargeted metabolomics correlation-based network analysis incorporating human metabolic reconstructions. *BMC Systems Biology* 7(1), 107.

Koyama H, Okamoto S, Watanabe N, Hoshino N, Jimbo M, Yasumoto K, et al. (2015) Dynamic changes in the accumulation of metabolites in brackish water clam *Corbicula japonica* associated with alternation of salinity. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* 181: 59–70.

Kullgren A, Samuelsson LM, Larsson DG, Björnsson BT, Bergman EJ (2010) A metabolomics approach to elucidate effects of food deprivation in juvenile rainbow trout (*Oncorhynchus mykiss*). *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology* 299(6): R1440-8.

Kuo TC, Tian TF, Tseng YJ (2013) 3Omics: A web-based systems biology tool for analysis, integration and visualization of human transcriptomic, proteomic and metabolomic data. *BMC Systems Biology* 7(1): 64.

Kwon YK, Jung YS, Park JC, Seo J, Choi MS, Hwang GS (2012) Characterizing the effect of heavy metal contamination on marine mussels using metabolomics. Marine Pollution Bulletin 64(9): 1874–1879.

Langfelder P, Horvath S (2008) WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics* 9(1): 559.

Larive CK, Barding GA, Dinges MM (2014) NMR spectroscopy for metabolomics and metabolic profiling. *Analytical Chemistry* 87(1): 133–146.

Lei Z, Huhman DV, Sumner LW (2011) Mass spectrometry strategies in metabolomics. *Journal of Biological Chemistry* 286(29): 25435–25442.

Lei Z, Li H, Chang J, Zhao PX, Sumner LW (2012) MET-IDEA version 2.06; improved efficiency and additional functions for mass spectrometry-based metabolomics data processing. *Metabolomics* 8(1): 105–110.

Li W (2012) Volcano plots in analyzing differential expressions with mRNA microarrays. *Journal of Bioinformatics and Computational Biology* 10(06): 1231003.

Li Y, Li G, Görling B, Luy B, Du J, Yan J (2015) Integrative analysis of circadian transcriptome and metabolic network reveals the role of de novo purine synthesis in circadian control of cell cycle. *PLoS Computational Biology* 11(2): e1004086

Liebeke M, Bundy JG (2012) Tissue disruption and extraction methods for metabolic profiling of an invertebrate sentinel species. *Metabolomics* 8(5): 819–830.

Liland KH (2011) Multivariate methods in metabolomics - from pre-processing to dimension reduction and statistical analysis. *TrAC Trends in Analytical Chemistry* 30(6): 827–841.

Lin CY, Wu H, Tjeerdema RS, Viant MR (2007) Evaluation of metabolite extraction strategies from tissue samples using NMR metabolomics. *Metabolomics* 3(1): 55–67.

Liu PF, Liu QH, Wu Y, Jie H (2015) A pilot metabolic profiling study in hepatopancreas of *Litopenaeus vannamei* with white spot syndrome virus based on ^{1}H NMR spectroscopy. *Journal of Invertebrate Pathology* 124: 51–56.

Liu X, Ji C, Zhao J, Wang Q, Li F, Wu H (2014b) Metabolic profiling of the tissue-specific responses in mussel *Mytilus galloprovincialis* towards *Vibrio harveyi* challenge. *Fish & Shellfish Immunology* 39(2): 372–377.

Liu X, Sun H, Wang Y, Ma M, Zhang Y (2014a) Gender-specific metabolic responses in hepatopancreas of mussel *Mytilus galloprovincialis* challenged by *Vibrio harveyi*. *Fish & Shellfish Immunology* 40(2): 407–413.

Liu X, Zhang L, You L, Yu J, Cong M, Wang Q, et al. (2011) Assessment of clam *Ruditapes philippinarum* as heavy metal bioindicators using NMR-based metabolomics. *CLEAN–Soil, Air, Water* 39(8): 759–766.

Ma YM, Yang MJ, Wang S, Li H, Peng XX (2015) Liver functional metabolomics discloses an action of L-leucine against *Streptococcus iniae* infection in tilapias. *Fish & Shellfish Immunology* 45(2): 414–421.

Macías-Mayorga D, Laiz I, Moreno-Garrido I, Blasco J (2015) Is oxidative stress related to cadmium accumulation in the mollusc *Crassostrea angulata*? *Aquatic Toxicology* 161: 231–241.

Marhuenda-Egea FC, Toledo-Guedes K, Sanchez-Jerez P, Ibanco-Cañete R, Uglem I, Saether B-S (2015) A metabolomic approach to detect effects of salmon farming on wild saithe (*Pollachius virens*) populations. *The Journal of Agricultural and Food Chemistry* DOI: 10.1021/acs.jafc.5b04765

Melis R, Cappuccinelli R, Roggio, Anedda R (2014) Addressing marketplace gilthead sea bream (*Sparus aurata* L.) differentiation by ^{1}H NMR-based lipid fingerprinting. *Food Research International* 63(B): 258–264.

Meunier B, Dumas E, Piec I, Bechet D, Hebraud M, Hocquette JF (2007) Assessment of hierarchical clustering methodologies for proteomic data mining. *Journal of Proteome Research* 6(1): 358–366.

Mishra BB, van der Hooft JJJ (2016) Updates in metabolomics tools and resources: 2014–2015. *Electrophoresis* 37(1): 86–110.

Moore DS, Jepsen PU Volka K (2014) Principles of vibrational spectroscopic methods and their application to bioanalysis. In *Handbook of Spectroscopy* (Eds. Gauglitz G, Moore DS), Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany. doi: 10.1002/9783527654703.ch27

Morrison N, Bearden D, Bundy JG, Collette T, Currie F, Davey MP, et al (2007) Standard reporting requirements for biological samples in metabolomics experiments: Environmental context. *Metabolomics* 3: 203–210.

Mushtaq MY, Choi YH, Verpoorte R, Wilson EG (2014b) Extraction for metabolomics: Access to the metabolome. *Phytochemical Analysis* 25(4): 291–306.

Mushtaq MY, Marçal RM, Champagne DL, van der Kooy F, Verpoorte R, Choi YH (2014a) Effect of acute stresses on zebra fish (*Danio rerio*) metabolome measured by NMR-based metabolomics. *Planta Medica* 80(14): 1227–1233.

Nováková L, Vlčková H (2009) A review of current trends and advances in modern bio-analytical methods: Chromatography and sample preparation. *Analytica Chimica Acta* 656(1): 8–35.

Nurdalila AA, Bunawan H, Kumar SV, Rodrigues KF, Baharum SN (2015) Homogeneous nature of Malaysian marine fish *Epinephelus fuscoguttatus* (Perciformes; Serranidae): Evidence based on molecular markers, morphology and Fourier transform infrared analysis. *International Journal of Molecular Sciences* 16(7): 14884–14900.

Ong ES, Chor CF, Zou L, Ong CN (2009) A multi-analytical approach for metabolomic profiling of zebrafish (*Danio rerio*) livers. *Molecular BioSystems* 5(3): 288–298.

Parrilla-Taylor DP, Zenteno-Savín T, Magallón-Barajas FJ (2013) Antioxidant enzyme activity in pacific whiteleg shrimp (*Litopenaeus vannamei*) in response to infection with white spot syndrome virus. *Aquaculture* 380–383: 41–46.

Parsons HM, Ludwig C, Günther UL, Viant MR (2007) Improved classification accuracy in 1- and 2-dimensional NMR metabolomics data using the variance stabilising generalised logarithm transformation. *BMC Bioinformatics* 8(1): 234–250.

Paul C, Mausz MA, Pohnert G (2013) A co-culturing/metabolomics approach to investigate chemically mediated interactions of planktonic organisms reveals influence of bacteria on diatom metabolism. *Metabolomics* 9(2): 349–359.

Peng B, Ma YM, Zhang JY, Li H (2015) Metabolome strategy against *Edwardsiella tarda* infection through glucose-enhanced metabolic modulation in tilapias. *Fish & Shellfish Immunology* 45(2): 869–876.

Persicke M, Rückert C, Plassmeier J, Stutz LJ, Kessler N, Kalinowski J, Goesmann A, Neuweger H (2012) MSEA: Metabolite set enrichment analysis in the MeltDB metabolomics software platform: Metabolic profiling of *Corynebacterium glutamicum* as an example. *Metabolomics* 8(2): 310–322.

Picone G, Engelsen SB, Savorani F, Testi S, Badiani A, Capozzi F (2011) Metabolomics as a powerful tool for molecular quality assessment of the fish *Sparus aurata*. *Nutrients* 3(2): 212–227.

Portella CAF, Smart KA, Tumanov S, Cook GM, Villas-Bôas SG (2014) Global metabolic response of *Enterococcus faecalis* to oxygen. *Journal of Bacteriology* 196(11): 2012–2022.

Qu JH, Cheng JH, Sun DW, Pu H, Wang QJ, Ma J (2015) Discrimination of shelled shrimp (*Metapenaeus ensis*) among fresh, frozen-thawed and cold-stored by hyperspectral imaging technique. *LWT-Food Science and Technology* 62(1): 202–209.

R Core Team (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (http://www.R-project.org).

Ramautar R, Somsen GW, de Jong GJ (2009) CE-MS in metabolomics. *Electrophoresis* 30(1): 276–291.

Ramautar R, Somsen GW, de Jong GJ 2015 CE-MS for metabolomics: Developments and applications in the period 2012–2014. *Electrophoresis* 36(1): 212–224.

Robles R, Lozano AB, Servilla A, Márquez L, Nuez-Ortin W, Moyano FJ (2013) Effect of partially protected butyrate used as feed additive on growth and intestinal metabolism in sea bream (*Sparus aurata*). *Fish Physiology and Biochemistry* 39(6): 1567–1580.

Rocha SM, Freitas R, Cardoso P, Santos M, Martins R, Figueira E (2013) Exploring the potentialities of comprehensive two-dimensional gas chromatography coupled to time of flight mass spectrometry to distinguish bivalve species: Comparison of two clam species (*Venerupis decussata* and *Venerupis philippinarum*). *Journal of Chromatography A* 1315: 152–161.

Römisch-Margl W, Prehn C, Bogumil R, Röhring C, Suhre K, Adamski J (2012) Procedure for tissue sample preparation and metabolite extraction for high-throughput targeted metabolomics. *Metabolomics* 8(1): 133–142.

Rubtsov DV, Jenkins H, Ludwig C, Easton J, Viant MR, Günther U, et al. (2007) Proposed reporting requirements for the description of NMR-based metabolomics experiments. *Metabolomics* 3: 223–229.

Samuelsson LM, Förlin L, Karlsson G, Adolfsson-Erici M, Larsson DJ (2006) Using NMR metabolomics to identify responses of an environmental estrogen in blood plasma of fish. *Aquatic Toxicology* 78(4): 341–349.

Sansone S-A, Schober D, Atherton HJ, Fiehn O, Jenkins H, Rocca-Serra P, et al. (2007) Metabolomics standards initiative—ontology working group: Work in progress. *Metabolomics* 3: 249–256.

Savorani F, Picone G, Badiani A, Fagioli P, Capozzi F, Engelsen SB (2010) Metabolic profiling and aquaculture differentiation of gilthead sea bream by ^{1}H NMR metabonomics. *Food Chemistry* 120(3): 907–914.

Savorani F, Rasmussen MA, Mikkelsen MS, Engelsen S B (2013) A primer to nutritional metabolomics by NMR spectroscopy and chemometrics. *Food Research International* 54(1): 1131–1145.

Schock TB, Duke J, Goodson A, Weldon D, Brunson J, Leffler JW, et al. (2013) Evaluation of Pacific white shrimp (*Litopenaeus vannamei*) health during a superintensive aquaculture growout using NMR-based metabolomics. *PLoS ONE* 8(3): e59521. doi:10.1371/journal.pone.0059521

Schock TB, Newton S, Brenkert K, Leffler J, Bearden DW (2012) An NMR-based metabolomic assessment of cultured cobia health in response to dietary manipulation. *Food Chemistry* 133(1): 90–101.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. (2003) Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Research* 13(11): 2498–2504.

Shao Y, Lia C, Chen X, Zhang P, Li Y, Li T, et al. (2015) Metabolomic responses of sea cucumber *Apostichopus japonicus* to thermal stresses. *Aquaculture* 435: 390–397.

Sheedy JR, Lachambre S, Gardner DK, Day RW ^{1}H-NMR metabolite profiling of abalone digestive gland in response to short-term starvation. *Aquaculture International* DOI: 10.1007/s10499-015-9941-4

Shi H, Ye T, Song B, Qi X, Chan Z (2015) Comparative physiological and metabolomic responses of four *Brachypodium distachyon* varieties contrasting in drought stress resistance. *Acta Physiologiae Plantarum* 37(6): 1–12.

Simpson MJ, Bearden DW (2013) Environmental metabolomics: NMR techniques. *eMagRes* 2(4): DOI: 10.1002/9780470034590.emrstm1352

Smolinska A, Blanchet L, Buydens LM, Wijmenga SS (2012) NMR and pattern recognition methods in metabolomics: From data acquisition to biomarker discovery: A review. *Analytica Chimica Acta* 750: 82–97.

Søfteland L, Kirwan JA, Hori TS, Størseth TR, Sommer U, Berntssen MHG, et al. (2014) Toxicological effect of single contaminants and contaminant mixtures associated with plant ingredients in novel salmon feeds. *Food and Chemical Toxicology* 73: 157–174.

Song Q, Chen H, Li Y, Zhou H, Han Q, Diao X (2016) Toxicological effects of benzo (a) pyrene, DDT and their mixture on the green mussel *Perna viridis* revealed by proteomic and metabolomic approaches. *Chemosphere* 144: 214–224.

Southam AD, Easton JM, Stentiford GD, Ludwig C, Arvanitis TN, Viant MR (2008) Metabolic changes in flatfish hepatic tumours revealed by NMR-based metabolomics and metabolic correlation networks. *Journal of Proteome Research* 7(12): 5277–5285.

Spann N, Aldridge DC, Griffin JL, Jones OA (2011) Size-dependent effects of low level cadmium and zinc exposure on the metabolome of the Asian clam, *Corbicula fluminea*. *Aquatic Toxicology* 05(3): 589–599.

Srivastava V, Obudulu O, Bygdell J, Löfstedt T, Rydén P, Nilsson R, et al. (2013) OnPLS integration of transcriptomic, proteomic and metabolomic data shows multi-level oxidative stress responses in the cambium of transgenic hipI-superoxide dismutase Populus plants. *BMC Genomics* 14(1): 893.

Stentiford GD, Viant MR, Ward DG, Johnson PJ, Martin A, Wenbin W, et al. (2005) Liver tumors in wild flatfish: A histopathological, proteomic, and metabolomic study. *Omics: A Journal of Integrative Biology* 9(3): 281–299.

Steuer R (2006) Review on the analysis and interpretation of correlations in metabolomic data. *Briefings in Bioinformatics* 7(2): 151–158.

Steuer R, Mrogenthal K, Weckwerth W, Selbig J (2007) A gentle guide to the analysis of metabolomic data. In Weckworth W. (Ed.) *Metabolomics: Methods and Protocols* (pp 105–126), Humana Press Inc, New Jersey, US.

Su YB, Peng B, Han Y, Li H, Peng XX (2015) Fructose restores susceptibility of multidrug-resistant *Edwardsiella tarda* to kanamycin. *Journal of Proteome Research* 14(3): 1612–1620.

Sugimoto M, Kawakami M, Robert M, Soga T, Tomita M (2012) Bioinformatics tools for mass spectroscopy-based metabolomic data processing and analysis. *Current Bioinformatics* 7(1): 96–108.

Sumner LW, Amberg A, Barrett D, Beger R, Beale MH, Daykin C, et al. (2007) Proposed minimum reporting standards for chemical analysis. *Metabolomics* 3: 211–221.

Sun X, Weckwerth W (2012) COVAIN: A toolbox for uni-and multivariate statistics, time-series and correlation network analysis and inverse estimation of the differential Jacobian from metabolomics covariance data. *Metabolomics* 8(1), 81–93.

Szymańska E, Saccenti E, Smilde AK, Westerhuis JA (2012) Double-check: Validation of diagnostic statistics for PLS-DA models in metabolomics studies. *Metabolomics* 8(1): 3–16.

Taylor CF, Binz P-A, Aebersold R, Affolter M, Barkovich R, Deutsch EW (2008) Guidelines for reporting the use of mass spectrometry in proteomics. *Nature Biotechnology* 26: 860–861.

Tikunov AP, Johnson CB, Lee H, Stoskopf MK, Macdonald JM (2010) Metabolomic investigations of American oysters using ^{1}H-NMR spectroscopy. *Marine Drugs* 8(10): 2578–2596.

Tikunov AP, Stoskopf MK, Macdonald JM (2014) Fluxomics of the Eastern oyster for environmental stress studies. *Metabolites* 4(1): 53–70.

Tuffnail W, Mills GA, Cary P, Greenwood R (2009) An environmental ^{1}H NMR metabolomic study of the exposure of the marine mussel *Mytilus edulis* to atrazine, lindane, hypoxia and starvation. *Metabolomics* 5(1): 33–43.

Tufi S, Lamoree MH, De Boer J, Leonards PEG (2015a) Cross-platform metabolic profiling: Application to the aquatic model organism *Lymnaea stagnalis. Analytical and Bioanalytical Chemistry* 407(7): 1901–1912.

Tufi S, Stel JM, De Boer J, Lamoree MH, Leonards PEG (2015b) Metabolomics to explore imidacloprid induced toxicity in the central nervous system of the freshwater snail *Lymnaea stagnalis*. *Environmental Science & Technology* DOI: 10.1021/acs.est.5b03282

van den Berg RA, Hoefsloot HC, Westerhuis JA, Smilde AK, van der Werf MJ (2006) Centering, scaling, and transformations: Improving the biological information content of metabolomics data. *BMC Genomics* 7(1): 142.

van der Werf MJ, Takors R, Smedsgaard J, Nielsen J, Ferenci T, Portais JC, et al. (2007) Standard reporting requirements for biological samples in metabolomics experiments: Microbial and in vitro biology experiments. *Metabolomics* 3: 189–194.

van Gulik WM, Canelas AB, Taymaz-Nikerel H, Douma RD, de Jonge LP, Heijnen JJ (2012) Fast sampling of the cellular metabolome. In *Microbial Systems Biology* (pp. 279–306). Humana Press.

Viant M, Sommer U (2013) Mass spectrometry based environmental metabolomics: A primer and review. *Metabolomics* 9: S144–S158.

Vidal NP, Goicoechea E, Manzanos MJ, Guillén MD (2014) Fourier transform infrared spectroscopy as a tool to study farmed and wild sea bass lipid composition. *Journal of the Science of Food and Agriculture* 94(7): 1340–1348.

Villas-Bôas SG, Smart KF, Sivakumaran S, Lane GA (2011) Alkylation or silylation for analysis of amino and non-amino organic acids by GC-MS? *Metabolites* 1(1): 3–20.

Vuckovic D (2012) Current trends and challenges in sample preparation for global metabolomics using liquid chromatography–mass spectrometry. *Analytical and Bioanalytical Chemistry* 403(6): 1523–1548.

Wagner L Trattner S, Pickova J, Gomez-Requeni P, Moazzami AA (2014) ^{1}H NMR-based metabolomics studies on the effect of sesamin in Atlantic salmon (*Salmo salar*). *Food Chemistry* 147: 98–105.

Wang Y, Kora G, Bowen BP, Pan C (2014) MIDAS: a database-searching algorithm for metabolite identification in metabolomics. *Analytical Chemistry* 86(19): 9496–9503.

Wei L, Wang Q, Wu H, Ji C, Zhao J (2015) Proteomic and metabolomic responses of Pacific oyster *Crassostrea gigas* to elevated pCO_2 exposure. *Journal of Proteomics* 112: 83–94.

Wei L, Wang Q, Wu H, Ji C, Zhao J (2015) Proteomic and metabolomic responses of Pacific oyster *Crassostrea gigas* to elevated *p*CO2 exposure. *Journal of Proteomics* 112: 83–94.

Wei X, Shi X, Kim S, Patrick JS, Binkley J, Kong M, et al. (2014) Data dependent peak model based spectrum deconvolution for analysis of high resolution LC-MS data. *Analytical Chemistry* 86(4): 2156–2165.

Wei X, Shi X, Kim S, Patrick JS, Binkley J, McClain C, et al. (2012) Data preprocessing method for liquid chromatography-mass spectrometry based metabolomics. Analytical Chemistry 84(18): 7963–7971.

Wei X, Shi X, Koo I, Kim S, Schmidt RH, Arteel GE, et al. (2013) MetPP: A computational platform for comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry-based metabolomics. *Bioinformatics* 29(14): 1786–1792.

Westerhuis JA, Hoefsloot HC, Smit S, Vis DJ, Smilde AK, van Velzen EJ, et al. (2008) Assessment of PLSDA cross validation. *Metabolomics* 4(1): 81–89.

Wojdyr M (2010) Fityk: A general-purpose peak fitting program. *Journal of Applied Crystallography* 43(5): 1126–1128.

Worley B, Powers R (2013) Multivariate analysis in metabolomics. *Current Metabolomics* 1(1): 92–107.

Wortberg F, Nardy E, Contzen M, Rau J (2012) Identification of *Yersinia ruckeri* from diseased salmonid fish by Fourier transform infrared spectroscopy. *Journal of Fish Diseases* 35(1): 1–10.

Wu H, Ji C, Wei L, Zhao J, Lu H (2013a) Proteomic and metabolomic responses in hepatopancreas of *Mytilus galloprovincialis* challenged by *Micrococcus luteus* and *Vibrio anguillarum*. *Journal of Proteomics* 94: 54–67.

Wu H, Southam AD, Hines A, Viant MR (2008) High-throughput tissue extraction protocol for NMR-and MS-based metabolomics. *Analytical Biochemistry* 372(2): 204–212.

Wu H, Wang WX (2010) NMR-based metabolomic studies on the toxicological effects of cadmium and copper on green mussels *Perna viridis*. *Aquatic Toxicology* 100(4): 339–345.

Xi B, Gu H, Baniasadi H, Raftery D (2014) Statistical analysis and modeling of mass spectrometry-based metabolomics data. *Methods in Molecular Biology* 1198: 333–353.

Xia J, Broadhurst DI, Wilson M, Wishart DS (2013) Translational biomarker discovery in clinical metabolomics: An introductory tutorial. *Metabolomics* 9(2): 280–299.

Xia J, Sinelnikov I, Han B, Wishart DS (2015) MetaboAnalyst 3.0 - making metabolomics more meaningful. *Nucleic Acids Research* DOI: 10.1093/nar/gkv380

Xia J, Wishart DS (2010) MetPA: A web-based metabolomics tool for pathway analysis and visualization. *Bioinformatics* 26(18): 2342–2344.

Xiao JF, Zhou B, Ressom HW (2012) Metabolite identification and quantitation in LC-MS/MS-based metabolomics. *TrAC Trends in Analytical Chemistry* 32: 1-14.

Yan X, Xu J, Chen J, Chen D, Xu S, Luo Q, et al. (2012) Lipidomics focusing on serum polar lipids reveals species dependent stress resistance of fish under tropical storm. *Metabolomics* 8(2): 299–309.

Ye Y, Yang R, Lou Y, Chen J, Yan X, Tang H (2014) Effects of food processing on the nutrient composition of *Pyropia yezoensis* products revealed by NMR-based metabolomic analysis. *Journal of Food and Nutrition Research* 2(10): 749–756.

Yin P, Lehmann R, Xu G (2015) Effects of pre-analytical processes on blood samples used in metabolomics studies. *Analytical and Bioanalytical Chemistry* 407(17): 4879–4892.

Young T Alfaro AC, Villas-Bôas S (2015a) Metabolic profiling of mussel larvae: Effect of handling and culture conditions. *Aquaculture International* DOI: 10.1007/s10499-015-9945-0

Young T, Alfaro AC, Villas-Bôas S (2015b) Identification of candidate biomarkers for quality assessment of hatchery-reared mussel larvae via GC/MS-based metabolomics. *New Zealand Journal of Marine and Freshwater Research* 49(1): 87–95.

Yu WK, Shi YF, Fong CC, Chen Y, van de Merwe JP, Chan AKY, et al. (2013) Gender-specific transcriptional profiling of marine medaka (*Oryzias melastigma*) liver upon BDE-47 exposure. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics* 8(3): 255–262.Yu X, Niu X, Zhang X, Pei G, Liu J, Chen L, et al. (2015) Identification and mechanism analysis of chemical modulators enhancing astaxanthin accumulation in *Haematococcus pluvialis*. Algal Research 11: 284–293.

Zhao C, Gu D, Nambou K, Wei L, Chen J, Imanaka T, et al. (2015) Metabolome analysis and pathway abundance profiling of *Yarrowia lipolytica* cultivated on different carbon sources. *Journal of Biotechnology* 206: 42–51.

Zhao XL, Han Y, Ren ST, Ma YM, Li H, Peng XX (2015) L-proline increases survival of tilapias infected by *Streptococcus agalactiae* in higher water temperature. *Fish & Shellfish Immunology* 44(1): 33–42.

Zhou J, Chen B, Cai Z (2015) Metabolomics-based approach for assessing the toxicity mechanisms of dibutyl phthalate to abalone (*Haliotis diversicolor supertexta*). *Environmental Science and Pollution Research* 22(7): 5092–5099.

Zhu F, Jin M (2015) The effects of a thermophile metabolite, tryptophol, upon protecting shrimp against white spot syndrome virus. *Fish & Shellfish Immunology* 47(2): 777–781.

Zu X, Zeisei SH (2013) Spectral deconvolution for gas chromatography mass spectrometry-based metabolomics: Current status and future perspectives. *Computational and Structural Biotechnology Journal* 4(5):