Identification of candidate biomarkers for quality assessment of hatchery-reared mussel larvae *via* GC/MS-based metabolomics

TIM YOUNG[1], ANDREA C. ALFARO[1*] & SILAS VILLAS-BÔAS[2]

[1]*Institute for Applied Ecology New Zealand, School of Applied Sciences, Auckland University of Technology, Auckland, New Zealand*

[2]*Centre for Microbial Innovation, School of Biological Sciences, The University of Auckland, Auckland, New Zealand*

[*] Corresponding author: Andrea C. Alfaro
Email: andrea.alfaro@aut.ac.nz

Identification of candidate biomarkers for quality assessment of hatchery-reared mussel larvae *via* GC/MS-based metabolomics

Abstract:

To ensure environmental and economic sustainability of future aquaculture growth, large-scale hatchery production of mollusc larvae is required. However, variation in larval quality currently limits potential maximum yields. Identification of biomarkers which reflect the immediate physiological condition of larvae during hatchery production could help monitor and determine causes of variation. Metabolomics is well-suited to this task due to its capacity for providing an instantaneous snapshot of the physiology of an organism through analysis of its metabolite profile. As a test, we applied GC/MS-based metabolomics for this purpose. Using a variety of univariate and multivariate feature selection methods, we identified four metabolite-metabolite ratios involving levels of succinate, glycine, alanine, pyroglutamate and myristic acid as candidate biomarkers for assessing mussel larval quality. These metabolites are known to have roles in energy metabolism, osmotic regulation, immune function and cell-cell communication. We anticipate that further investigation of these metabolites and their associated biochemical pathways will yield a more complete understanding of the factors responsible for larval production variability.

Keywords: aquaculture; bivalves; larval physiology; larval quality; metabolite profiles; metabolomics; mussel larvae; *Perna canaliculus*

Introduction:

Global mollusc production is rapidly increasing and is one of the largest aquaculture activities in the world. Most of the seed requirements are obtained from wild populations where seasonal and regional reliability of quality and quantity is inconsistent. Furthermore, these natural stocks are approaching, or have exceeded, maximum sustainable yields for some species (Helme et al. 2004). In recent years considerable attention has been placed in the production of hatchery-reared juveniles for natural stock enhancement and grow-out (O'Connor et al. 2012). However, a number of bottlenecks present problems for commercial viability, such as broodstock management and gamete quality, mass production and variable quality of microalgal diets, establishment of successful and reproducible larval rearing

practices, improvement of settlement rates and metamorphosis synchronisation and identification of causes and development of remedial strategies for poor larval health and growth.

In New Zealand, mussel (*Perna canaliculus*) exports represent the largest aquaculture sector by value and volume. Small-scale hatchery production of spat currently contributes only marginally towards industry's seed requirements. However, substantial research over the past decade has led to the on-going development of selective breeding lines (MacAvoy et al. 2008), establishment of a cryopreservation program (Paredes et al. 2012), optimisation of microalgae culture and larval rearing procedures (Ragg et al. 2010; Kaspar et al. 2014) and future strategies for extensive growth and up-scaling of hatchery facilities (Roberts 2013). The ability to provide consistency in larval quality and quantity is an important step towards reaching the commercial goal of successful large-scale production.

Larval quality is a term which refers to the physiological condition of larvae and is related to growth and survival during different developmental stages and under various environmental conditions (Racotta et al. 2003). For marine invertebrates, factors responsible for variation in larval quality may include parental genetics (Meyer & Manahan 2010), maternal size and life-history experience (Marshall & Keough 2004, 2006), level of maternal provisioning (Smith & Bolton 2007), sperm environment prior to fertilisation (Ritchie & Marshall 2013), sperm/egg ratios during fertilisation (Marshall & Keough 2003), embryo and larval rearing densities (Galley et al. 2010; Deng et al. 2013), availability of food (Aarab et al. 2013), presence of pathogens (Kesarcodi-Watson et al. 2009) and physical environmental parameters (Sánchez Lazo & Martínez-Pita 2012). However, even with attempts to control such factors within hatcheries, inter- and intra-cohort variation in larval quality is still commonly observed. The effects of this variation have far-reaching implications for mollusc aquaculture due to difficulties in predicting larval production volumes and potential carry-over effects on post-metamorphic survival, growth and reproduction (Emlet & Sadro 2006; Przeslawski & Webb 2009). Thus, continual monitoring of larval quality is critical to provide information for determining causation of poor production levels, which may lead to better management decisions. Identification of biomarkers which reflect the immediate physiological condition of larvae has the potential to provide industry with valuable tools to improve the commercial viability of hatchery operations.

Metabolomics is a non-targeted approach to comprehensively profile a broad range of metabolites within biological samples, such as cells, tissues and whole organisms. Presenting unique opportunities for novel biomarker discovery, metabolomics is one of the

latest fields gaining a stronghold in the various approaches towards molecular biology. Since metabolites are end products of gene and protein expression and are exceptionally sensitive to genetic and environmental perturbations, metabolomics offers a revolutionary framework for phenotyping organisms at the molecular level (Kuehnbaum & Britz-McKibbin 2013). Using such a hypothesis-free driven approach has yielded excellent information without *a priori* knowledge for understanding complex biological processes. Furthermore, this approach routinely demonstrates value in corroborating data from gene and protein expression studies, as well as providing additional and complimentary evidences (e.g., Wilmes et al. 2013). Examples of applications are widespread across diverse areas of research from human medicine (Jin et al. 2014) to deep-sea microbial ecology (Kimes et al. 2013).

While only limitedly applied to aquaculture research thus far, specific metabolomics-based investigations are starting to provide useful information for various sectors within the aquaculture industry. For example, identification of biomarkers and development of new techniques for monitoring health (Schock et al. 2012, 2013; Ji et al. 2013), growth (Almi 2012), nutritional condition (Cajka et al. 2013; Liu et al. 2013; Abro 2014; Wagner et al. 2014), culture environment (Savorani et al. 2010; Picone et al. 2011) and meat quality (Erikson et al. 2012) in fish, shellfish and crustaceans are now available. However, despite its wide applicability, the use of metabolomics in mollusc larval culture has not yet been realised. Here, we utilise GC/MS-based metabolomics to identify candidate biomarkers for assessing mussel larval quality using differential growth rates as a quality measure.

Methods:

Larval sampling and metabolite identification

Following standard industry protocols (Ragg et al. 2010), embryos produced from a selectively bred family line (F2) were reared to D-stage (2 days post-fertilisation [dpf]) in a 5000 L static tank and graded on an 80 µm sieve to retain healthy individuals from dead or undeveloped embryos and trocophores. These larvae were transferred to three 150 L conical flow-through tanks and fed *ad libitum* for two days. Umbo-stage larvae (4 dpf) were then separated on a mesh screen based on their relative sizes (>/<120 µm) to divide poor quality (slow growing) and high quality (fast growing) larvae. Six samples including pooled individuals (about 80,000) from each of the two size fractions and three replicate rearing tanks were snap frozen in liquid nitrogen and stored at -80°C until metabolite extraction. Using established methods (Villas-Bôas et al. 2011), metabolites were co-extracted with an internal standard (L-alanine-2,3,3,3-

*d*4) in cold MeOH:H_2O, derivatised by alkylation (MCF derivatives) and analysed *via* GC/MS. Raw spectral data were pre-processed according to Smart et al. (2010), using the R package Metab (Aggio et al. 2011) to automate AMDIS processing. Twenty nine metabolites from different chemical classes (e.g., amino acids, fatty acids, organic acids, cyclic alcohols, vitamins) were reliably identified by comparing retention times and m/z ratios against an in-house MS library.

Data processing

Metabolite peak intensities were normalized against the internal standard to compensate for potential technical variation (e.g., variable metabolite recoveries). Relative metabolite abundances ($\sum peaks_{i-j}/peak_i$) were calculated for each sample. Six metabolites with low abundance values (<0.1%) were manually filtered and removed as a data integrity precaution, and every possible remaining pair-wise metabolite level ratio (n = 253) was determined. All ratios were log-transformed to ensure independence of the metabolite order in which they were calculated. The dataset matrix was converted to a comma-separated value file and subjected to a variety of feature selection methods to identify potential biomarkers for discrimination of larval quality classes.

Statistical analyses

All statistical analyses were performed using Metaboanalyst 2.0, a comprehensive web-based analytical pipeline for high-throughput metabolomics studies (Xia et al. 2012), and ROCCET, a web-based tool for common Receiver Operating Characteristic (ROC) curve analyses on metabolomics data sets (Xia et al. 2013). To visualise feature differences between larval quality classes, the top 50 metabolite ratios (ranked by their T-test statistics) were used to perform agglomerative hierarchical cluster analysis (distance = Pearson's correlation; aggregation = Ward's criterion). The constructed dendrogram is co-displayed with the results of heat map analysis to reveal metabolite ratio differences and assist conception of between-class/-feature clusters.

Four methods of feature reduction were used independently to minimise selection bias and provide a robust criteria for assisting candidate biomarker identification: 1) Volcano plot analysis was performed to simultaneously display fold-change and T-test statistics to maximise utilisation of statistical information from the data. Since the aim of this study is to identify low numbers of accurate biomarkers with the potential of commercial applicability, the T-test

significance level and fold-change cut-off thresholds were set at 0.01 and 3, respectively. 2) Supervised Projection to Latent Squares Discriminant Analysis (PLS-DA) was performed to help identify features which contribute most towards the class separation in the PLS-DA model. PLS-DA uses a multivariate regression technique to extract *via* linear combination of original variables the information that can predict the sample class membership. Leave One Out Cross Validation (LOOC) was used to assess the performance of the PLS-DA model. Feature selection was based on the Variable Importance in Projection (VIP) scores (a weighted sum of squares of the PLS loadings taking into account the amount of explained class-variation in each dimension). Features displaying VIP scores > 1.2 were considered as potential biomarkers of larval quality. 3) Significant Analysis of Metabolites (SAM) was performed to supply supportive evidence for the selection of features identified by PLS-DA. The advantages of SAM are that the analysis addresses the false discovery rate (FDR) when performing numerous tests on high-dimensional data and the use of permutation-based analysis avoids assumptions of metabolite independence by accounting for correlations. A delta value of 2.0 was selected after analysis of FDR *versus* the number of significant features as a function of various delta values. 4) Empirical Bayesian Analysis of Metabolites (EBAM) was used as a final independent feature selection method. EBAM is a variation of the SAM algorithm based on moderated *t*-statistics. The advantages of EBAM are that it also addresses FDR and is well-suited to analysis of data where biological replication is low.

To establish the best candidate biomarkers of larval quality, multivariate ROC analyses were performed (PLS-DA algorithm) using the common metabolite ratios identified from the four feature selection methods. ROC analyses which employ Monte-Carlo Cross Validation with multiple iterations are generally considered the method of choice for evaluating the performance of potential biomarkers. Using the ROC Explorer module to carry out repeated model testing, the two most robust predictive models were determined which incorporated the minimum number of the most stable features.

Results:

Unsupervised hierarchical clustering (Figure 1A) revealed that the metabolite ratio profiles of the samples (shown on the y-axis) can be distinguished clearly and robustly. Furthermore, the three biological replicates from each of the two larval quality classes are clearly clustered. Clustering of the features (shown on the x-axis) revealed several groups of

ratios that were either substantially higher or lower in the poor quality group compared with those in the high quality group.

The results of the hierarchical clustering were confirmed in the PLS-DA model (Figure 1B), with the first two components explaining 73.5% and 4.1% of the variation, respectively. The performance of the model was assessed *via* LOOCV, which showed a good level of predictive ability ($Q^2 = 0.80$; $R^2 = 0.95$). The metabolite ratios with the greatest influence on the clustering in the PLS-DA plot were identified by their first component VIP scores, resulting in selection of 47 features.

SAM analysis (Figure 1C) identified 25 ratios, with an FDR of 3.2%, that were significantly different between the larval quality classes. Of these features, 20 were lower in the poor quality group and five were higher. Using a variation of the SAM algorithm, EBAM analysis (Figure 1D) identified 26 ratios with an FDR of 4.2%, 20 of which were the same as those identified by SAM. Volcano plot analysis (Figure 1E) identified 29 ratios that satisfied the stringent criteria placed on feature selection (i.e., ratios having a fold-change of >3 and simultaneously having a T-test statistic of <0.01). The inclusion of this test removed potentially less biologically relevant features (i.e., those with small between-class differences) that may otherwise have been selected due to very low within-class variation. Of the metabolite ratios independently selected from the four methods, 19 ratios were commonly identified (Figure 1F), all of which were lower in the poor quality cohort.

Multivariate ROC analyses (Table 1) of the 19 selected features resulted in identification of four candidate biomarkers. Using balanced subsampling, two thirds of the samples were used to evaluate and rank the important features, which were then used to build classification models. These models were validated on the samples that were left out and the procedure was repeated multiple times to construct a series of most stable *n*-feature PLS-DA models. Although the proportions that the features were selected for in the iterated models differed (i.e., their rank frequencies), the four ratios identified were similarly incorporated in the 2- and 3-feature models. Candidate biomarker ratios and associated relative metabolite abundances for each larval quality class are presented in Table 2.

Discussion:

Of the initial 253 metabolite ratios analysed, we used four independent feature selection methods (Volcano plot, PLS-DA, SAM and EBAM) to identify candidate biomarkers for assessing mussel larval quality. These methods resulted in the selection of between 25 (SAM)

to 47 (PLS-DA) potentially important features. The 19 common features selected by the four tests were subjected to ROC-based analyses to reduce the number of final candidate biomarkers and assess their performance. Four metabolite ratios were identified as being good candidates for assessing larval quality: alanine/succinate, glycine/succinate, myristic acid/succinate and pyroglutamate/succinate. The use of ratios can have the advantage of relying on fluctuations of two metabolites rather than one, thereby substantially increasing biomarker sensitivity when these changes occur in opposite trajectories. Ratios of reaction-linked metabolites within a biochemical pathway can also help to identify functional changes not immediately apparent by revealing important associations at the enzyme level, which can be very useful in hypothesis-free studies. Each of the identified ratios were substantially lower in the poor quality larvae. Comparing this group with the faster-growing cohort, the trend is characterised by elevated levels of succinate and simultaneous reductions in relative metabolite abundances of alanine, glycine, pyroglutamate and myristic acid.

Succinate is a tricarboxylic acid (TCA) cycle intermediate and is common across all candidate biomarkers. Although many of the 19 potential biomarkers initially identified by the four feature selection methods did not contain succinate, incorporation of this metabolite into all ratios within the most stable 2- and 3-feature PLS-DA models highlights its significance in predicting larval quality. Furthermore, the approximate 10-fold difference in succinate abundance with concomitant ≤2-fold differences in the other four metabolites suggests that succinate is primarily responsible for the variation in biomarker ratios between larval quality classes. The accumulation of succinate in the poor quality larvae may indicate overall metabolic repression and lower energy production (Müller et al. 2012). Many marine molluscs have evolved well-adapted strategies for coping with dynamic environmental fluxes and stressful conditions, such as variations in salinity, pH and oxygen availability. For example, the capacity to easily convert between aerobic and anaerobic respiration under duress provides a mechanisms for molluscs to produce enough energy to preserve critical metabolic function without wasting energy on other non-critical processes, such as growth. Under aerobic conditions, succinate is readily oxidised to fumarate by succinate dehydrogenase. Interestingly, the fumarate/succinate ratio was identified as a potential biomarker candidate during our initial statistical screening in all four of the feature selection methods applied. As a classic example of a metabolite ratio with a direct substrate–product link, further integrated analysis of this reaction and its associated pathways at the transcript and/or protein level may provide important mechanistic insight into the physiological basis for intraspecific growth variation in mussel larvae.

Myristic acid is a C:14 fatty acid and is found in high concentrations (13-15% of lipids) in the microalgae diet fed to larvae during the rearing process (Ragg et al. 2010). The mean relative abundances of free myristic acid in the poor quality larvae were lower than in the high quality larvae (0.49 *versus* 0.65% respectively). This may reflect slight differences in lipid digestive abilities between the two classes. The fact that this discrepancy was small shows that the variation in the myristic/succinate ratios were dominated by large differences in succinate content, providing strong support for the potential importance of the TCA cycle intermediate as a marker of larval quality.

The FAA glycine was one of the most abundant metabolites found across all samples which is consistent with its role as a common osmolyte in marine bivalves (Kube et al. 2006). Our findings reveal that glycine content was significantly reduced in the poor quality cohort. Recent metabolomic and gene expression studies have demonstrated that glycine content is related to the health state of oysters and clams, with reductions of this amino acid occurring after infection with pathogens (Liu et al. 2013), exposure to arsenic (Ji et al. 2013) and under hypo-osmotic conditions (Meng et al. 2013). Levels of alanine (also serving osmolytic function) in the poor quality larvae mirrored the lower relative abundance of glycine. These observations may indicate the occurrence of stress-induced disruption in osmotic regulation.

Pyroglutamate is the cyclic lactam of glutamic acid and, despite its existence being known for over a century, its functions remain poorly understood. When protein-bound, the inclusion of a terminal pyroglutamate residue stabilises proteins by making them resistant to degradation by amino peptidases (Kumar & Bachhawat 2012). In its free form as a cellular metabolite, pyroglutamate is an intermediate in glutathionine degradation, can act as an osmoregulator and is also readily interconverted to glutamate, which itself serves various biological roles (Kumar & Bachhawat 2012). For example, glutamate is a molluscan neurotransmitter (Hatakeyma et al. 2010), a common osmolyte in marine bivalves (Kube et al. 2006) and the precursor to proline and alanine (Wang et al. 2012). Thus, variations in pyroglutamate levels between larval quality classes may indicate differential immune function and capacities for protein stabilisation, osmoregulation and cell-cell communication.

In conclusion, the results of this study demonstrates the potential application of metabolomic-based approaches in aquaculture research to: 1) classify mollusc larvae based on their quality, 2) construct prediction models for larval quality assessment, and 3) identify biochemical pathways which may be under differential regulation to reveal important mechanistic insights for future investigation. Indeed, we anticipate that further analysis of the single metabolites and their ratios will reveal additional information, and when integrated with

gene and protein expression data could provide new avenues for selective breeding programs to consistently yield high quality larvae. Supplementary experiments incorporating metabolomics-based approaches to investigate other measures of larval quality (e.g., health after immunological challenge and nutritional condition in response to diet manipulation) have the potential to offer industry with a suite of biomarkers for monitoring the physiological state of larvae throughout the rearing process. Such a tool kit could help identify causes of larval batch failures and develop remedial strategies to enhance overall larval production within hatcheries.

Acknowledgments:

We are thankful to Rodney Roberts, Dan McCall, Nicola Hawes and the Shellfish Production and Technology New Zealand Ltd (SPATnz) for assisting with larval production and to Margarita Markovskaya for assistance with the sample processing at the Metabolomics Lab, University of Auckland. This project was partially funded by a Faculty Research Grant to A. C. Alfaro. We are grateful for the fruitful discussions with the Aquaculture Biotechnology Group, which improved the manuscript.

REFERENCES

Aarab L, Pérez-Camacho A, Viera-Toledo MP, Courtois de Viçose G, Fernández-Palacios H, Molina L 2013. Embryonic development and influence of egg density on early veliger larvae and effects of dietary microalgae on growth of brown mussel *Perna perna* (L. 1758) larvae under laboratory conditions. Aquaculture International 21: 1065–1076.

Abro R 2014. Digestion and metabolism of carbohydrates in fish. Unpublished PhD dissertation, Swedish University of Agricultural Sciences, Uppsala, Sweden. 58 p.

Aggio R, Villas-Bôas SG, Ruggiero K 2011. Metab: An R package for high-throughput analysis of metabolomics data generated by GC-MS. Bioinformatics 27: 2316–2318.

Almli M 2012. Effects of different live feed on larval growth and development in ballan wrasse (*Labrus bergylta* Ascanius, 1767): A metabolomics study. Unpublished MSc thesis, Norwegian University of Science and Technology, Trondheim, Norway. 59 p.

Cajka T, Danhelova H, Vavrecka A 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.

Deng Y, Fu S, Liang F, Xie S 2013. Effects of stocking density, diet, and water exchange on growth and survival of pearl oyster *Pinctada maxima* larvae. Aquaculture International 21: 1185–1194.

Emlet RB, Sadro SS 2006. Linking stages of life history: How larval quality translates into juvenile performance for an intertidal barnacle (*Balanus glandula*). Integrative and Comparative Biology 46: 334–346.

Erikson U, Standal IB, Aursand IG, Veliyulin E, Aursand M 2012. Use of NMR in fish processing optimization: A review of recent progress. Magnetic Resonance in Chemistry 50: 471–80.

Galley TH, Batista FM, Braithwaite R, King J, Beaumont AR 2010. Optimisation of larval culture of the mussel *Mytilus edulis* (L.). Aquaculture International 18: 315–325.

Hatakeyama D, Mita K, Kobayashi S et al. 2010. Glutamate transporters in the central nervous system of a pond snail. Journal of Neuroscience Research 88: 1374–1386.

Helm MM, Bourne N, Lovatelli A 2004. Hatchery culture of bivalves. FAO fisheries technical paper 471. Rome, Italy. 177 p.

316 Ji C, Wu H, Liu X, Zhao J, Yu J, Yin X 2013. The influence of salinity on toxicological effects of arsenic in digestive gland of clam
317 *Ruditapes philippinarum* using metabolomics. Chinese journal of oceanology and limnology 31: 345–352.
318 Ji C, Wu H, Wei L, Zhao J, Wang Q, Lu H 2013. Responses of *Mytilus galloprovincialis* to bacterial challenges by metabolomics and
319 proteomics. Fish & Shellfish Immunology 35: 489–498.
320 Jin X, Yun SJ, Jeong P, Kim IY, Kim WJ, Park S 2014. Diagnosis of bladder cancer and prediction of survival by urinary
321 metabolomics. Oncotarget 5: 1635–1645.
322 Kaspar HF, Keys EF, King N, Smith KF, Kesarcodi-Watson A, Miller MR 2014. Continuous production of *Chaetoceros calcitrans* in a
323 system suitable for commercial hatcheries. Aquaculture 420–421: 1–9.
324 Kesarcodi-Watson A, Kaspar H, Lategan MJ, Gibson L 2009. Two pathogens of Greenshell™ mussel larvae, *Perna canaliculus*: *Vibrio*
325 *splendidus* and a *V. coralliilyticus/neptunius*-like isolate. Journal of Fish Disease 32: 499–507.
326 Kimes NE, Callaghan AV, Aktas DF et al. 2013. Metagenomic analysis and metabolite profiling of deep–sea sediments from the Gulf of
327 Mexico following the Deepwater Horizon oil spill. Frontiers in Microbiology 4: 50.
328 Kube S, Gerber A, Jansen JM, Schiedek D 2006. Patterns of organic osmolytes in two marine bivalves, *Macoma balthica*, and *Mytilus* spp.,
329 along their European distribution. Marine Biology 149: 1387–1396.
330 Kuehnbaum NL, Britz-McKibbin P 2013. New advances in separation science for metabolomics: Resolving chemical diversity in a post-
331 genomic era. Chemical Reviews 113: 2437–2468.
332 Liu X, Ji C, Zhao J, Wu H 2013. Differential metabolic responses of clam *Ruditapes philippinarum* to *Vibrio anguillarum* and *Vibrio*
333 *splendidus* challenges. Fish & Shellfish Immunology 35: 2001–2007.
334 MacAvoy ES, Wood AR, Gardner JPA 2008. Development and evaluation of microsatellite markers for identification of individual
335 Greenshell™ mussels (*Perna canaliculus*) in a selective breeding programme. Aquaculture 274: 41–48.
336 Marshall DJ, Keough MJ 2003. Sources of variation in larval quality for free-spawning marine invertebrates: Egg size and the local sperm
337 environment. Invertebrate Reproduction and Development 44: 63–70.
338 Marshall DJ. Keough MJ 2004. When the going gets rough: Effect of maternal size manipulation on larval quality. Marine Ecology Progress
339 Series 272: 301–305
340 Marshall DJ, Keough MJ 2006. Complex life cycles and offspring provisioning in marine invertebrates. Integrative & Comparative Biology
341 46: 643–651.
342 Meng J, Zhu Q, Zhang L, Li C, Li L, She Z, Huang B, Zhang G 2013. Genome and transcriptome analyses provide insight into the
343 euryhaline adaptation mechanism of *Crassostrea gigas*. PLoS One 8: e58563.
344 Meyer E, Manahan DT 2010. Gene expression profiling of genetically determined growth variation in bivalve larvae (*Crassostrea gigas*).
345 The Journal of Experimental Biology 213: 749–758.
346 O'Connor W, Dove M, O'Connor S, Luu LT, Xan L, Giang CT 2012. Building bivalve hatchery production capacity in Vietnam and
347 Australia. Final report [FR2012-25] for project FIS/2005/114). Canberra, Australian Centre for International Agricultural Research.
348 Paredes E, Adams SL, Tervit HR et al. 2012. Cryopreservation of Greenshell™ mussel (*Perna canaliculus*) trochophore larvae. Cryobiology
349 65: 256–262.
350 Picone G, Engelsen S.B, Savorani F, Testi S, Badiani A, Capozzi F 2011. Metabolomics as a powerful tool for molecular quality
351 assessment of the fish *Sparus aurata*. Nutrients 3: 212–227.
352 Przeslawski R, Webb AR 2009. Natural variation in larval size and developmental rate of the Northern quahog *Mercenaria mercenaria* and
353 associated effects on larval and juvenile fitness. Journal of Shellfish Research 28:505–510.
354 Racotta IS, Palacios E, Ibarra AM 2003. Shrimp larval quality in relation to broodstock condition. Aquaculture 227:107–130.
355 Ragg NLC, King N, Watts E, Morrish J 2010. Optimising the delivery of the key dietary diatom *Chaetoceros calcitrans* to intensively
356 cultured Greenshell™ mussel larvae, *Perna canaliculus*. Aquaculture 306: 270–280.
357 Ritchie H, Marshall DJ 2013. Fertilisation is not a new beginning: Sperm environment affects offspring developmental success. Journal of
358 Experimental Biology 216: 3104–3109.
359 Roberts R 2013. SPATnz. Seafood New Zealand 21(2): 20–21.
360 Sánchez-Lazo C, Martínez-Pita I 2012. Effect of temperature on survival, growth and development of *Mytilus galloprovincialis* larvae.
361 Aquaculture Research 43: 1127–1133.
362 Savorani F, Picone G, Badiani A, Fagioli P, Capozzi F, Engelsen SB 2010. Metabolic profiling and aquaculture differentiation of gilthead
363 sea bream by 1H NMR metabonomics. Food Chemistry 120: 907–914.
364 Schock TB, Newton S, Brenkert K, Leffler J, Bearden DW 2012 An NMR-based metabolomic assessment of cultured cobia health in
365 response to dietary manipulation. Food Chemistry 133: 90–101.

Schock TB, Duke J, Goodson A 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.

Smart K.F., Aggio R.B., Van Houtte J.R., Villas-Bôas S.G. 2010. Analytical platform for metabolome analysis of microbial cells using MCF derivatization followed by gas chromatography-mass spectrometry. Nature Protocols 5: 1709–1729.

Smith BL, Bolton TF 2007. Consequences of natural variation in maternal investment and larval nutrition on larval growth of a marine polychaete. Marine and Freshwater Research, 58: 1002–1007.

Steuer R, Kurths J, Fiehn O, Weckwerth W 2003. Interpreting correlations in metabolic networks. Biochemical Society Transactions 31: 1476–1478.

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: 3–20.

Wagner L, Trattner S, Pickova J, Gómez-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, Li E, Yu N et al. 2012. Characterization and expression of glutamate dehydrogenase in response to acute salinity stress in the Chinese mitten crab, *Eriocheir sinensis*. PLoSOne 7(5): e37316.

Wilmes A, Limonciel A, Aschauer L et al. 2013. Application of integrated transcriptomic, proteomic and metabolomic profiling for the delineation of mechanisms of drug induced cell stress. Journal of Proteomics 79: 180–194.

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

FIGURES

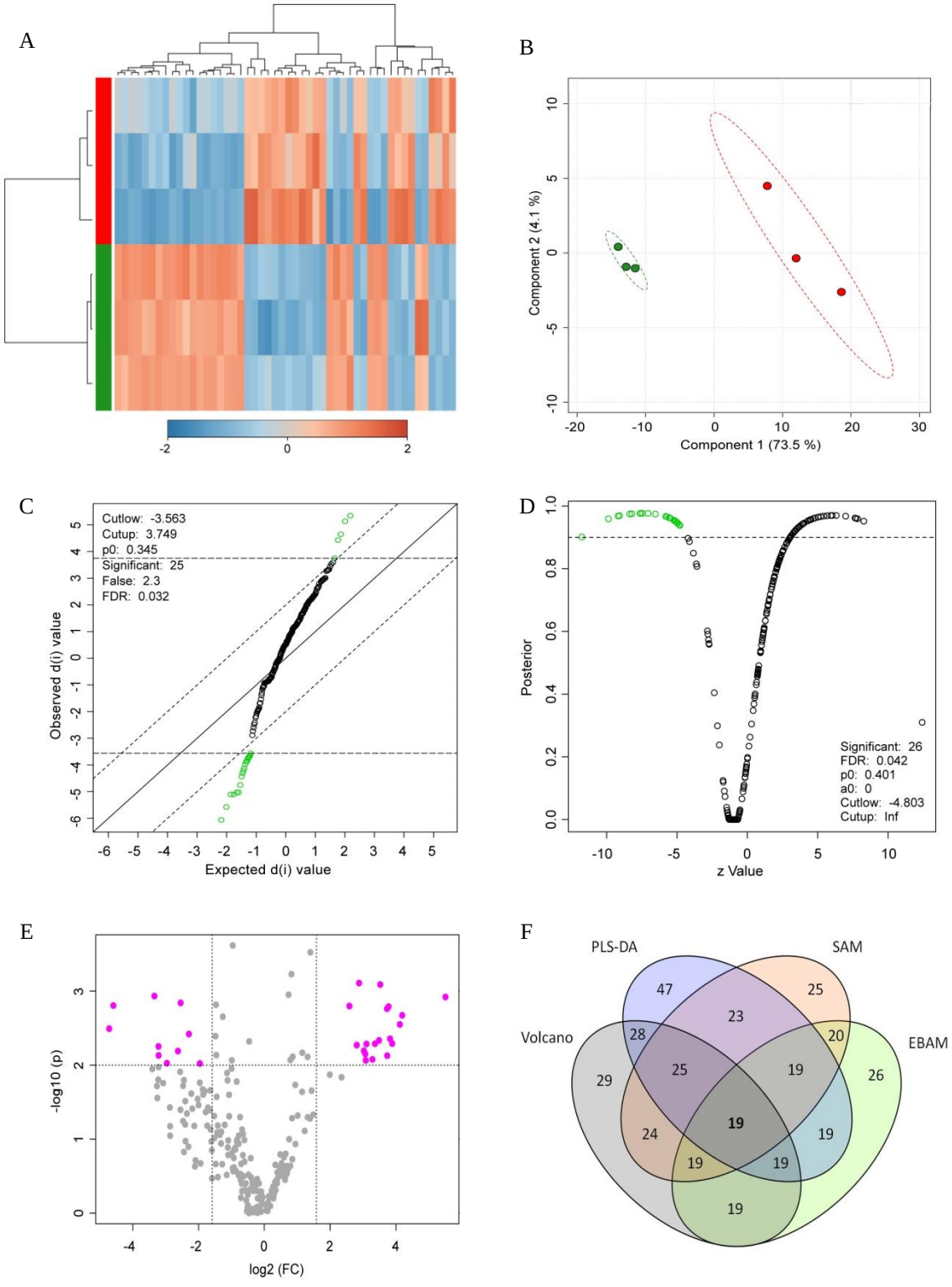
A
-2
0
2
B
Component 2 (4.1 %)
Component 1 (73.5 %)
C
Cutlow: -3.563
Cutup: 3.749
p0: 0.345
Significant: 25
False: 2.3
FDR: 0.032
Observed d(i) value
Expected d(i) value
D
Posterior
z Value
Significant: 26
FDR: 0.042
p0: 0.401
a0: 0
Cutlow: -4.803
Cutup: Inf
E
-log10 (p)
log2 (FC)
F
PLS-DA
SAM
Volcano
EBAM
47
25
23
28
20
29
25
19
26
19
24
19
19
19
19

Figure 1. Hierarchical clustering and heat map (A) of poor (■) and high (■) quality larvae based on the top 50 features ranked by their T-test statistic (samples on the y-axis, features on the x-axis). PLS-DA scores plot (B) of poor (●) and high (●) quality larvae. SAM plot (C) of significantly different features (○) between quality classes (upper right = ratios that were higher in poor quality group; lower left = ratios that were lower). EBAM plot (D) of significantly different features (○) between quality classes (all lower in the poor quality group). Volcano plot (E) of features (●) with a between-class fold-change >3 and a T-test statistic <0.01 (upper left = ratios that were higher in the poor quality group; upper right = ratios that were lower). Venn diagram (E) displays the counts of commonly identified ratios by the four independent feature selection methods.

Table 1. Degree of importance and rank frequency of candidate biomarkers in models for larval class separation.

Candidate biomarker	Importance in models	2-feature model rank frequency	3-feature model rank frequency
Pyroglutamate/Succinate	1.39	0.68	0.90
Glycine/Succinate	1.28	0.54	0.92
Alanine/Succinate	1.27	0.38	0.78
Myristic acid/Succinate	1.16	0.40	0.40