292–302.
10. Psychogios, N. *et al.* (2011) The human serum metabolome. *PLoS One* 6, e16957.
11. Olszewski, K.L. *et al.* (2010) Branched tricarboxylic acid metabolism in *Plasmodium falciparum*. *Nature* 466, 774–778.
12. Tielens, A.G. and van Hellemond, J.J. (2009) Surprising variety in energy metabolism within Trypanosomatidae. *Trends Parasitol.* 25, 482–490.
13. McConville, M.J. and Naderer, T. (2011) Metabolic pathways required for the intracellular survival of *Leishmania*. *Annu. Rev. Microbiol.* 6, 543–561.
14. Sernee, M.F. *et al.* (2006) *Leishmania* β-1,2-mannan is assembled on a mannose-cyclic phosphate primer. *Proc. Natl. Acad. Sci. USA* 103, 9458–9463.
15. Basant, A. *et al.* (2010) Alterations in urine, serum and brain metabolomic profiles exhibit sexual dimorphism during malaria disease progression. *Malar. J.* 9, 110.
16. Wang, Y. *et al.* (2008) Global metabolic responses of mice to *Trypanosoma brucei* infection. *Proc. Natl. Acad. Sci. USA* 105, 6127–6132.
17. Bakker, B.M. *et al.* (2010) The silicon trypanosome. *Parasitol.* 137, 1333–1341.
18. Zamboni, N. and Sauer, U. (2009) Novel biological insights through metabolomics and ^{13}C-flux analysis. *Curr. Opin. Microbiol.* 12, 553–558.
19. Amador-Noguez, D. *et al.* (2010) Systems-level metabolic flux profiling elucidates a complete, bifurcated tricarboxylic acid cycle in *Clostridium acetobutylicum*. *J. Bacteriol.* 192, 4452–4461.
20. Yuan, J. *et al.* (2008) Kinetic flux profiling for quantitation of cellular metabolic fluxes. *Nat. Protoc.* 3, 1328–1340.
21. Baran, R. *et al*. (2009) Mass spectrometry based metabolomics and enzymatic assays for functional genomics. *Curr. Opin. Microbiol.* 12, 547–552.
22. Eisenreich, W. *et al.* (2010) Carbon metabolism of intracellular bacterial pathogens and possible links to virulence. *Nat. Rev. Microbiol.* 8, 401–412.
23. Liu, W.T. *et al.* (2010) Imaging mass spectrometry of intraspecies metabolic exchange revealed the cannibalistic factors of *Bacillus subtilis*. *Proc. Natl. Acad. Sci. USA* 107, 16286–16290.

Biographies

Eleanor Saunders became interested in metabolomics and its application to *Leishmania* parasite during her PhD studies at the University of Melbourne. She is currently doing a postdoctorate and developing new approaches for measuring metabolic fluxes in the protozoan parasites.

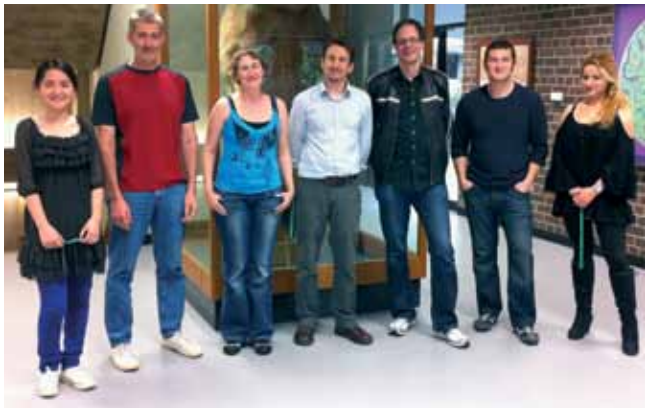
David de Souza has utilised metabolomics technologies to study a wide variety of microbial, animal and plant systems. He is completing his PhD studies at the University of Melbourne on parasite metabolomics as well as working as a research officer in the Bio21 node of Metabolomics Australia.

James MacRae completed his PhD studies at the University of Dundee, Scotland, studying the surface glycoproteins of the *Trypanosoma cruzi* parasite. He is currently utilising metabolomic approaches to identify potential drug targets in the apicomplexan parasites at the University of Melbourne.

Vladimir Likić is a bioinformatician and computational biologist with a focus on parasitic protozoa. He is a senior research fellow at the Bio21 Institute and head of Bioinformatics at Metabolomics Australia.

Malcolm McConville has had a long-term interest in parasite metabolism. He is currently a NHMRC Principal Research Fellow and head of the Metabolomics Australia node at the Bio21 Institute, Melbourne University.

From omics to systems biology: Exploring the mystery box of microbial life



*Martin Ostrowski, Sasha Tetu, Karl Hassan, Anahit Penesyan, Kent Lim, Liam Elbourne, Liping Li, Deepa Varkey & Ian Paulsen**

Department of Chemistry and Biomolecular Sciences
Macquarie University, Sydney, NSW 2109

Microbial molecular biology has traditionally used very reductionist approaches; for example, find a gene of interest, clone it or knock it out and see if you can detect a phenotype. The genomics era has opened up the possibility of analysing microbes and communities at a systems level by combining high-throughput experimental data from genomic, transcriptomic, proteomic and phenomic techniques. This parallels earlier reductionist approaches by going from DNA to RNA to protein to phenotype, albeit on a global rather than individual gene scale.

In our group we are applying systems approaches at two levels; firstly we are looking at reconstructing the metabolism and physiology of pathogenic and environmental microbes. For example, we have been identifying the key drug resistance factors and elucidating the role of multidrug efflux pumps in the opportunistic pathogen *Acinetobacter baumannii*[1,2] and the regulation and biosynthesis of secondary metabolites in plant-associated pseudomonads[2]. Secondly, we are exploring the identity and function of populations within microbial communities to clarify their activity in global nutrient cycles. Following on from whole-community metagenomics projects, such as the Global Ocean Survey[3], we have been digging deeper into the functional diversity of marine cyanobacteria by combining

molecular ecology, flow sorting and metagenomic sequencing of *Synechococcus* populations from marine environments around the world. Thus 'omics' approaches allow us to delve into the inner workings of a cell, generate environmental hypotheses that can be tested at a systems level and extrapolate our knowledge across global ecosystems.

Moving from the small to the smaller: linking phenotypes with genes and genomes

One of the strengths of a multidisciplinary approach is the capacity to untangle complex networks of interactions at a regulatory or functional level in a bacterial cell. For example, we are interested in identifying the molecular basis of how certain plant commensal pseudomonads provide beneficial effects by suppressing plant pathogens. Mechanistically, pathogen suppression by biocontrol *Pseudomonas* spp. is facilitated largely through the production and secretion of bioactive secondary metabolites and exoenzymes. Our group is developing a more comprehensive appreciation of the genetic factors responsible for the production of these molecules and of their regulatory control. As such, we have adopted a multipronged omics-led approach, that examines biocontrol at four levels: genome, transcriptome, proteome and phenotype. The integration of data from these analyses can assist the identification of key biocontrol elements and establish factors influencing their expression. Ultimately, this information may be used to optimise conditions for the utilisation of existing biocontrol strains or tailor new strains to a particular application.

As a genus, *Pseudomonas* is able to facilitate suppression of a remarkably broad range of plant pathogens, from viruses to bacteria, fungi and insects, in an array of plant hosts and environments. Nonetheless, individual biocontrol *Pseudomonas* strains typically display niche biocontrol capacities, being optimally adapted to suppress only one or a few specific pathogens in particular plant systems. In an attempt to identify genetic differences between strains that provide the basis for diverse biocontrol abilities, we have sequenced a series of *P. chlororaphis*, *P. synxantha* and *P. fluorescens* isolates from different crop types, different areas of the plant (phyllosphere and rhizosphere), with different plant colonisation efficacies and that protect against different diseases. Multilocus and 16S rRNA gene phylogenetic analyses showed that these strains form a single, large clade, distinct from other well-characterised pseudomonads, such as *P. aeruginosa*, *P. syringae* and *P. putida*.

Despite their phylogenetic relatedness, the protein coding potential of the sequenced biocontrol *Pseudomonas* strains is highly diverse; as a group the strains share a core genome comprising only half of their predicted protein coding genes. The accessory biocontrol *Pseudomonas* genome (that is to say, the non-core genes) includes many novel putative biocontrol factors including secondary metabolite biosynthesis genes, protein secretion systems, insecticidal toxins and plant hormone biosynthesis genes. These genes are likely to define the differential biocontrol abilities of each strain and are usually clustered within genomic islands, that may have been acquired via horizontal gene transfer[4]. Nonetheless, due to the proteomic diversity between the strains, it is difficult to draw definitive conclusions about the precise factors responsible for each distinct phenotype, highlighting the need for additional functional analyses.

In addition to a range of putative biocontrol factors, each biocontrol strain encodes several hundred predicted regulatory genes. It is not clear what proportions of these genes are involved in regulating biocontrol phenotypes[5]. However, investigations from our and other groups have shown that the regulatory circuits coordinating the expression of biocontrol elements in *Pseudomonas* spp. are complex and intimately intertwined[6]. Furthermore, these circuits are responsive to a number of factors, including environmental conditions as well as the physiological state of the cell and can influence not only biocontrol, but a range of phenotypes including metabolism, motility and stress responses. Consequently we have conducted transcriptomic and proteomic experiments on a representative biocontrol strain *P. fluorescens* Pf-5. To date, we have applied these analyses to examine the consequences of deletion mutations to four regulators and several environmental stress conditions, such as micronutrient limitation, exposure to plant-derived toxins and heavy metals. The coupling of different high-throughput experimental approaches, as well as phenotypic studies has begun to reveal details of *P. flourescens* Pf-5 physiology; for example, the bioavailability of iron suppresses motility and antimicrobial biosynthesis but increases iron siderophore biosynthesis, iron uptake and metabolism. Expression analyses have also shown that antimicrobial secondary metabolite gene clusters and exoenzymes are generally under strong positive control by biocontrol regulators (Figure 1). Furthermore, the regulons of some loci, for example, the GacS/GacA two-component system (Figure 1), includes a range of biosynthetic gene clusters that were not previously recognised to function in biocontrol, highlighting this approach as a means of discovering novel potential biocontrol factors[2].

A primary factor limiting the use of biocontrol bacteria in agriculture is that many potentially effective strains do not readily colonise plant surfaces, such as leaves, blossoms, seeds and the rhizosphere, when applied artificially. The *Pseudomonas* isolates we have sequenced represent a range of colonisation abilities and include some strains displaying superior capacities to populate plant surfaces. A potential key factor in plant surface colonisation is the ability to utilise different plant-derived or environmental compounds as nutrient sources. To gain insight into the catabolic capabilities of the strains we have applied Biolog Phenotype Microarrays to simultaneous test hundreds of potential nutrient sources in a single high-throughput assay[7]. This work revealed that although there is significant diversity between the proteomes of the strains, they share very similar metabolic capabilities. This has led us to believe these organisms share

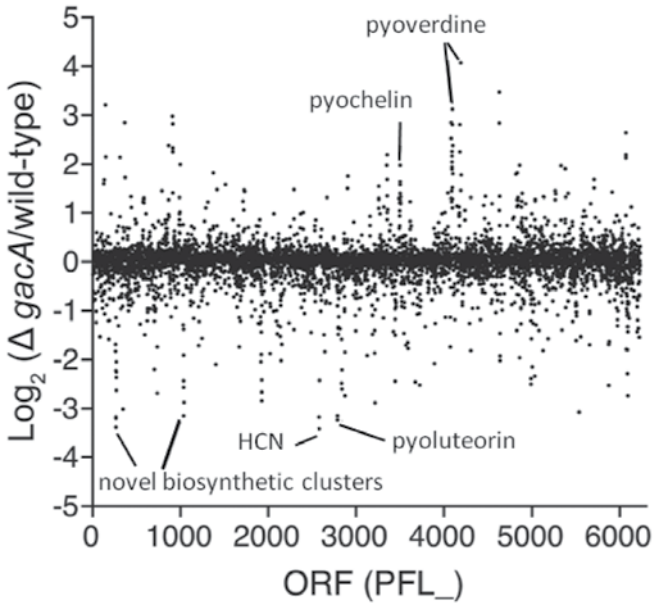


Figure 1. Differential gene transcription between wild-type *P. fluorescens* Pf-5 and a *gacA* mutant assessed in microarray experiments. Each point is one of the 6,147 annotated, protein-encoding genes in the Pf-5 genome, with the X axis showing gene order (the origin of replication at 0 and 6147), and the Y axis showing the $\log_2$ of transcript abundance of each gene in the *gacA* mutant relative to the wild-type strain Pf-5. Gene clusters whose members are regulated in a like manner by *gacA* are seen as spikes in the array data. The identities of highly modulated, well-characterised gene clusters are shown.

a common central metabolism, but that some strains differed in their ability to utilise specific, plant-derived compounds. To expand this analysis, we are currently building custom Biolog plates consisting of a wide range of compounds known to be produced by the respective plant hosts. To further explore the genetic factors required for plant colonisation, we are applying high-throughput transposon mutagenesis combined with next-generation sequencing to strains displaying superior colonisation potential. Mutant libraries are being passaged through selective plant surface environments to highlight genes essential for colonisation of these surfaces. Future studies may use the same mutant libraries to identify genes essential for suppression of particular plant pathogens.

While we have just begun to comprehend the physiological basis of biocontrol in plant commensal pseduomonads, our combined 'omics' approach has identified a number of key biocontrol factors that may have gone undetected using classical gene-specific analyses.

Moving from the small to the big: from genes and genomes to geosciences

On a global scale, marine microbes play a key role in controlling nutrient cycles that have a profound affect on our biosphere. They produce and catabolise a major portion of organic matter and should be a key component of ecosystem and climate models. Because small bacteria do not sink out of the euphotic zone (the illuminated portion of the ocean) they form the basis of a tightly coupled loop where photosynthetically fixed organic matter is recycled. Less abundant, but larger cyanobacteria, such as *Synechococcus*, and eukaryotic algae also make a significant contribution to carbon fixation, even in nutrient poor environments, despite only accounting for a small fraction of cell numbers and DNA[8].

Over the past decade we have been exploring the functional and genomic diversity of *Synechococcus* spp. isolated from distinct habitats around the world's oceans[9,10]. This work has revealed the spatial structuring of globally distributed primary producers into locally adapted 'ecotypes'. As for biocontrol pseudomonads, a high degree of genomic variation underpins their adaptation to a wide range of ecological niches while the strong pressure of natural selection eliminates redundant genes[11]. Genome comparisons layered onto ecological data have, therefore, enabled us to define sets of genes and associated physiological traits that explain the adaptation of individual lineages to specific niches. Comparison of model coastal (strain CC9311) and oceanic (strain WH8102) *Synechococcus* isolates highlight interesting examples, such as a higher number of genes important for trace metal homeostasis in CC9311, and a wider range of phosphorus uptake and scavenging genes in WH8102[12]. The genome-encoded regulatory capacity of each strain reinforces the distinction between the coastal and oceanic lifestyle. For example, the coastal strains have lost their two-component sensory and regulatory system for phosphate, *PhoBR,* while the oceanic strain WH8102, which is the dominant strain in phosphorus-limited regions of the ocean, encode multiple periplasmic binding proteins for PO_4^{3-}, express a diverse range of Zn^{2+}-dependent genes for scavenging phosphorus from organic sources and encode an additional novel phosphorus regulator.

By constructing targeted gene knockouts in both phosphorus regulatory systems (PhoBR and PtrA) we have been able to document the phosphorus regulon in this model oceanic isolate and characterise a number of genes of unknown function[13], including several unusual phosphatases. We also demonstrate the interplay of the two regulators in a signal cascade that grades the cellular response to the extent and duration of phosphorus-limitation[14]. It is interesting to note that many genes that have been recently acquired by lateral gene transfer, including a large cluster of genes involved in cell-surface modification (Figure 2), have been recruited into the P-regulon in WH8102, a result that highlights the importance of genome plasticity and environmental selection in niche adaptation.

While we have delved deeply into the gene activity of a few selected strains, recent estimates suggest that there are numerous abundant lineages for which there are no cultured representatives[15]. Community metagenomics has the capacity to survey microbial diversity and illuminate the biochemical functioning of relatively simple systems, yet whole community approaches suffer from diminishing returns when the ecosystem

diversity is large. This is particularly evident in oceanic systems that are numerically dominated by small bacteria such as *Prochlorococcus* and *Pelgibacter*. As a consequence of this disconnection between genome copy numbers, biomass and activity, whole community sequencing efforts are literally just scraping the surface of the functional diversity of organisms that are influential in a range of climate-critical nutrient cycles.

In order to overcome this limitation we have been utilising fluorescence activated cell sorting (FACS) to fast-track the capture and isolation of *Synechococcus* strains that are indigenous in Australian coastal waters. The power of fluorescent sorting has also enabled us to fractionate different functional groups from seawater samples for targeted genomics or 'population – genomics' in order to characterise environmentally selected sets of genes at contrasting sites[9,16]. By sifting through the unique features of genetic repertoires of *Synechococcus* populations characteristic of distinct habitats along the coast, we hope to identify the forces shaping environmental communities. Targeted genomics provides us with a novel perspective to understand additional dimensions of biotic and abiotic pressures that are difficult to quantify with traditional oceanographic measurements. Because of their ubiquitous distribution and key role at the base of the marine food chain, *Synechococcus* could serve as powerful indicators of the status of ecosystems that could help us monitor the impacts of climate change along Australia's coasts.

Conclusions

These types of multidisciplinary approaches allow us to start to pry open the mystery box of microbial life and explore microbial systems from a single cell to a biosphere level. The *Synechococcus* and *Pseudomonas* work currently under way in our laboratory provides us with excellent models to explore the links between the fundamental units of selection: genes and genomes, with the ecology and activity of microorganisms over a range of different spatial scales.

References

1. Eijkelkamp, B.A. *et al.* (2011) Development of a high-throughput cloning strategy for characterization of *Acinetobacter baumannii* drug transporter proteins. *J. Molec. Microbiol. Biotechnol.* In Press.
2. Hassan, K.A. *et al.* (2011) Roles of DHA2 family transporters in drug resistance and iron homeostasis in *Acinetobacter* spp. *J. Molec. Microbiol. Biotechnol.* 20, 116–124.
3. Rusch, D.B. *et al.* (2007) The Sorcerer II Global Ocean Sampling Expedition: Northwest Atlantic through Eastern Tropical Pacific. *PLoS Biol* 5, e77. DOI:10.1371/journal.pbio.0050077
4. Paulsen, I.T. *et al.* (2005) Complete genome sequence of the plant commensal *Pseudomonas fluorescens* Pf-5: insights into the biological control of plant disease. *Nature Biotechnol.* 23, 873–878.
5. Loper, J.E. *et al.* (2007) The genomic sequence of *Pseudomonas fluorescens* Pf-5: Insights into Biological Control. *Phytopathology* 97, 233–238.
6. Haas, D. and Keel, C. (2003) Regulation of antibiotic production in root-colonizing *Pseudomonas* spp. and relevance for biological control of plant disease. *Ann. Rev. Phytopathol.* 41, 117–153.
7. Bochner, B.R. (2009) Global phenotypic characterization of bacteria. *FEMS Microbiol. Rev.* 33, 191–205.
8. Jardillier, L. *et al.* (2010) Significant CO_2 fixation by small prymnesiophytes in the subtropical and tropical northeast Atlantic Ocean. *ISME J.* 4, 1180–1192.
9. Palenik, B. *et al.* (2009) Coastal *Synechococcus* metagenome reveals major roles for horizontal gene transfer and plasmids in population diversity. *Environ. Microbiol.* 11, 349–359.
10. Scanlan, D.J. *et al.* (2009) Ecological genomics of marine picocyanobacteria. *Microbiol. Mol. Biol Rev* 73, 249–299.
11. Dufresne, A. *et al.* (2008) Unraveling the genomic mosaic of a ubiquitous genus of marine cyanobacteria. *Genome Biol.* 9, R90.
12. Palenik, B. *et al.* (2006) Genome sequence of *Synechococcus* CC9311: Insights into adaptation to a coastal environment. *Proc. Natl. Acad. Sci. USA* 103, 13555–13559.
13. Tetu, S.G. *et al.* (2009). Microarray analysis of phosphate regulation in the marine cyanobacterium *Synechococcus* sp WH8102. *ISME J.* 3, 835–849.

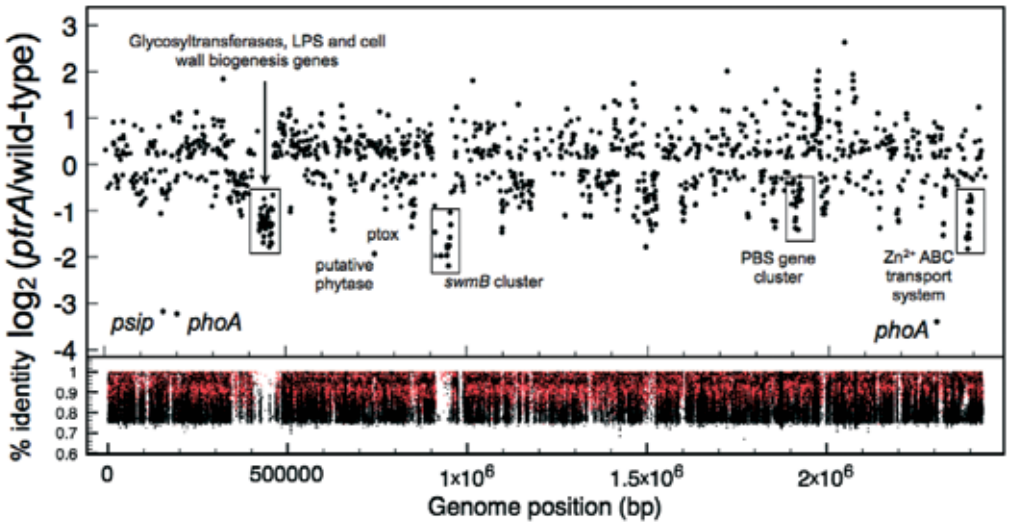


Figure 2. Microarray analyses of regulatory mutants show the recruitment of laterally acquired genes into the phosphorus regulon in the genome of *Synechococcus* strain WH8102 (top panel). Tiling of environmental metagenomic sequences against the WH8102 genome highlight genomic islands that are absent in indigenous *Synechococcus* populations (bottom panel). Of interest is a large island containing cell-surface modification genes that are under the influence of a novel phosphorus regulator, ptrA. Modifying the cell surface may be a key defence against predators, such as grazers and phage.

14. Ostrowski, M. *et al.* (2010) PtrA is required for coordinate regulation of gene expression during phosphate stress in a marine *Synechococcus*. *ISME J.* 4, 908–921.

15. Mazard, S. *et al.* (2011a) Multi-locus sequence analysis, taxonomic resolution and biogeography of marine *Synechococcus*. *Environ. Microbiol.* DOI: 10.1111/j.1462-2920.2011.02514.x

16. Mazard S. *et al.* (2011b) A targeted metagenomic approach to determine the 'population genome' of marine *Synechococcus*. In: *Handbook of Molecular Microbial Ecology II: Metagenomics in Different Habitats*. (FJ de Bruijn, ed).

Biographies

Ian Paulsen is Professor of Genomics at Macquarie University and is an ISI Highly Cited Researcher. As a former faculty member at The Institute for Genomic Research, he has led the sequencing of many microbial genomes. His research interests are focused on using genome sequencing, metagenomics and functional genomics to understand how lateral gene transfer in bacteria enables them to adapt to different environmental niches.

Other members of the Paulsen group at Macquarie University: **Martin Ostrowski** is a postdoctoral researcher interested in marine microbial ecology; **Sasha Tetu** is a postdoctoral fellow working on metagenomics of unusual microbial communities; **Karl Hassan** is an ARC Postdoctoral Fellow who is using functional genomics to characterise biocontrol pseudomonads; **Anahit Penesyan** is an ARC SuperScience Fellow working on cell-cell communication in *Pseudomonas*; **Kent Lim** is a PhD student finishing up his thesis on functional genomics of biocontrol mediated by *Pseudomonas fluorescens*; **Liam Elbourne** is a postdoctoral bioinformatics scientist working on more projects than he can possibly count; **Liping Li** and **Deepa Varkey** have just completed honours and master degrees, respectively.

* The photograph includes, from left to right: Ms Liping Li, Dr Liam Elbourne, Dr Sasha Tetu, Dr Martin Ostrowski, Prof Ian Paulsen, Dr Karl Hassan, Dr Anahit Penesyan. Ms Deepa Varkey is not in the photograph.

Systems biology: a new paradigm for industrial yeast strain development




Cristian Varela[1], Simon A Schmidt[1], Anthony R Borneman[1], Jens O Krömer[2], Alamgir Khan[3], Paul J Chambers[1] & The Australian Wine Yeast Systems Biology Consortium[4]*

[1] The Australian Wine Research Institute, PO Box 197, Glen Osmond, Adelaide, SA 5064
[2] Australian Institute for Bioengineering and Nanotechnology (AIBN), University of Queensland, St Lucia, QLD 4072
[3] Australian Proteome Analysis Facility (APAF), Macquarie University, Sydney, NSW 2109
[4] Genomics Australia, Proteomics Australia, Metabolomics Australia, Australian Bioinformatics Facility and The Australian Wine Research Institute[**]
* Corresponding author: Cristian.Varela@awri.com.au

One of the key challenges for industrial yeast strain development is to obtain a thorough understanding of the biology of yeast and to apply this knowledge to develop novel strains with improved features. The detailed study of individual biological components and the use of metabolic engineering have benefited the development of industrial strains enormously; however, such approaches have failed to describe yeast behaviour in the detail required to reveal the complex interactions operating within such biological systems. How can we accurately describe the biological processes and the interactions that occur during fermentation or cell growth?