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?

The use of omics methodologies and systems biology-based approaches promises to transform industrial strain development. By integrating experimental and computational methodologies and creating mathematical models of biological systems, scientists will be better able to manipulate yeast to improve traits that are critical for industrial applications. This article describes the promise and some of the pitfalls of applying a systems biology approach in the development of novel wine yeast strains and highlights the feasibility of applying such approaches in Australia.

The promise

Systems biology seeks to study the relationships and interactions between all processes operating in a biological system and to integrate this information in order to understand how biological systems work. Systems biology involves simultaneous monitoring of numerous biological processes using omics technologies and high-powered computational approaches. A quantitative description of the biological system under study is required in order to explain and predict its behaviour[1-3]. Ideally, this description involves the formulation of mathematical models that can be used to simulate the biological system[1]. Predictions made by mathematical models can in turn be iteratively tested through additional experimentation, enabling fine-tuning of the model.

Industrial microbiology has benefited enormously from advances in technology. Omics approaches, combined with the capacity to store and interrogate massive datasets, are poised to drive the next major shift in microbial strain development, particularly if done in a systems biology framework. Over recent years laboratory strains of the yeast *Saccharomyces cerevisiae* have been the organisms of choice to establish methodologies for systems biology. Some of the advantages responsible for the position that this microorganism enjoys include: being easy to cultivate, having a small genome (6000 genes), being the first eukaryote to be sequenced[4], being easy to manipulate genetically[3,5], the availability of a vast array of molecular tools, the existence of knowledge-based resources such as the *Saccharomyces* Genome Database (SGD) and other databases containing massive datasets[6], having available a collection of single-deletion mutants[7], and the first comprehensive consensus metabolic network constructed for any organism[8,9].

With so many resources, information and data available, *S. cerevisiae* is well-suited for a quantitative description of its cellular processes and to map in detail all key interactions within the yeast cell[1,6,10-12].

The pitfalls

Although most industrial yeasts are the same species as the laboratory strains used to build the amazing resources available to the yeast scientific community, they are unique in many features, including robustness and 'productivity'; traits that are vital for industrial applications. For example, laboratory yeast strains are unable to successfully ferment grape juice to produce drinkable wine. There are also differences in the production of secondary metabolites, which are essential for the sensory properties of wines and beers. Consistent with these phenotypic variations, there is considerable variation at a genomic level between, for example, wine and laboratory strains of *S. cerevisiae*[13,14].

Because of these differences not all data and resources available for work on laboratory strains of *S. cerevisiae* are fully applicable to industrial strains. For example, commercially available *S. cerevisiae* DNA microarrays are designed using a laboratory yeast reference genome, thus the expression of some genes relevant to a particular industrial application might not be detected. Also, the *S. cerevisiae* deletion collection was developed in a laboratory strain[7] and although it can help to identify genes responsible for some industrially relevant traits in functional genomic screens, it would miss many additional genes that shape industry-relevant traits. From an industrial perspective, therefore, it is important in research and development, including strain development programs, to utilise industrial yeast strains.

Another potential pitfall for those wishing to adopt systems-based approaches to their research is the scale and associated costs. As mentioned above, systems biology typically involves the generation and integration of large datasets (for example, transcriptomics, proteomics, metabolomics), infrastructure to enable this, and the appropriate expertise over a range of disciplines. Much of this is beyond the reach of a single laboratory[1].

Systems biology in Australia

To explore the feasibility of applying systems biology approaches in Australia, Bioplatforms Australia (BPA) and The Australian Wine Research Institute (AWRI) formed The *BPA/AWRI Wine Yeast Systems Biology Consortium*. The Consortium** has embarked on a systems biology demonstration project that focuses on wine yeast fermentation. The aim of the project is to harness the expertise, infrastructure and technologies available across platforms of BPA (Genomics Australia, Proteomics Australia, Metabolomics Australia and Australian Bioinformatics Facility) to develop systems-based mathematical models of wine yeast metabolism, which will be used to develop predictive models to inform design of new strains with improved, wine-relevant, traits.

The first step in this project was defining the 'system'. This involved choosing a wine yeast strain and defining the growth and culture conditions. A primary consideration for choice of the wine yeast was the availability of genomic data and confidence in

its general winemaking capabilities, which, in this context, means being able to convert grape juice into wine. We used a haploid wine strain, AWRI1631, the genome of which was sequenced prior to commencement of the project[13]. A chemically defined 'white' grape juice (CDGJ) growth medium was developed to ensure reproducibility between experiments.

Experimental design, the amount of replication, sampling regime, culture and sample volume, sample processing and sample analyses are influenced by the biological question and costs. In our case, because of the scale of the fermentations, culture volume was limited and, therefore, overall sample volume became a critical parameter in what was a time course experiment with multiple analyses to be performed at each time point. In order to decrease the number of samples for different analyses, we developed a method enabling the analysis of proteome and metabolome from the same sample.

Metabolite quantitative data can be highly variable depending on the sample treatment, extraction methods and analytical techniques employed[1]. Therefore, appropriate biological and technical replicates are key to obtaining the necessary statistical power and ensuring biologically meaningful results. We performed triplicate model batch wine fermentations that were sampled at five key time points, and for each of these: gene expression profiles were determined using RNAseq (this is a method of transcriptome analysis in which mRNAs are reverse transcribed into cDNAs which are then sequenced), proteomic analysis was performed using 2D-gels and iTRAQ (**I**sobaric **T**ags for **R**elative and **A**bsolute **Q**uantitation), and both targeted and non-targeted approaches were used for metabolomic profiling and metabolic footprinting (Figure 1). Metabolome samples were quenched and processed immediately after sampling.

For a comprehensive characterisation of metabolic networks and their functional operation, quantitative knowledge of intracellular fluxes is required. The fluxome, or the total set of fluxes in the metabolic network of a cell, represents integrative information on several cellular processes providing a true dynamic picture of the phenotype[2]. We plan to quantify intracellular fluxes by applying flux balancing on those five time points and also by using isotope-labelled precursors, specifically ^{13}C-labelled sugar, at exponential and stationary growth phases (Figure 1).

Several studies of wine fermentations have been made using two omics techniques, covering specifically transcriptome and proteome[15,16]; and transcriptome and metabolome[17]. At the outset we felt that metabolomic information needed to connect with, and be integrated with, transcriptional and proteomic analysis for a truly detailed description of wine fermentation. At the time of writing, transcriptional analyses appear to provide the most comprehensive view of an organism's response to its environment. However, it is not always the case that individual transcript levels correlate with the abundance of the corresponding protein. Thus, our view is that the different omics platforms are complementary and together provide the optimal level of detail for systems-level cellular modelling. With the support of BPA and our collaborators, all four omics techniques have been used to characterise quantitatively the wine yeast transcriptome, proteome, metabolome and fluxome.

Representative data collected to date are shown in Figure 2. Transcriptomic analysis identified 5556 genes, of these 2779 were significantly up- or down-regulated at least twofold during fermentation ($p<0.01$). Down-regulated genes included genes related to ribosome biogenesis, RNA processing, translation and metabolism of nitrogen compounds. This is consistent with the transition from exponential growth to fermentative metabolism. Up-regulated genes include those associated with oxidation-reduction processes and the generation of precursor metabolites and energy. Of the 1200 proteins identified (9% uncharacterised), the abundance of 400 proteins changed significantly over the fermentation time course. Cluster analysis revealed three major groups: proteins down-regulated relative to the first time point, including those involved in amino acid

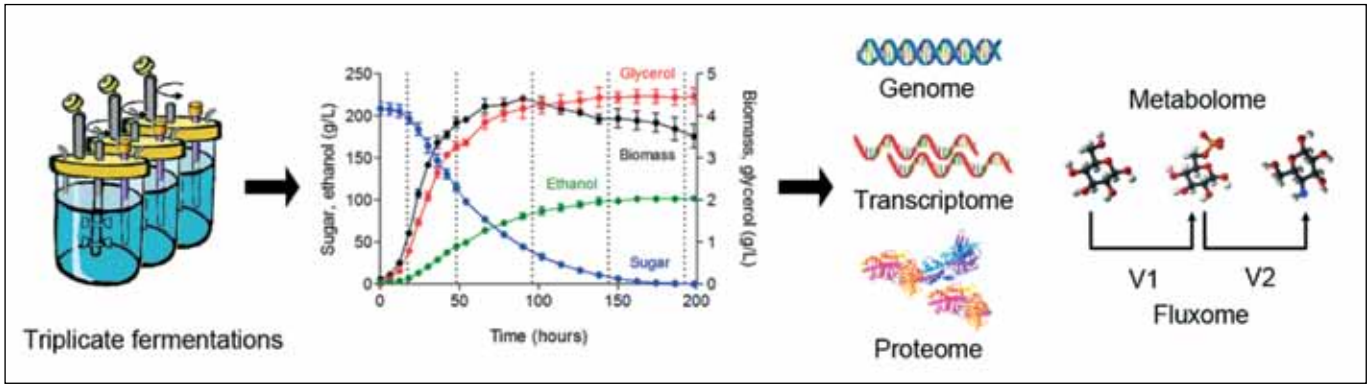

Figure 1. Wine yeast systems biology methodology. Triplicate model wine fermentations using AWRI1631 were sampled at five key time points (dotted lines). RNAseq was used to study the transcriptome, 2D-gels and iTRAQ were performed to explore the proteome, and both targeted and non-targeted approaches were used to study the metabolome. Metabolic flux balancing and isotope-labelled precursors will be applied to quantify the fluxome.

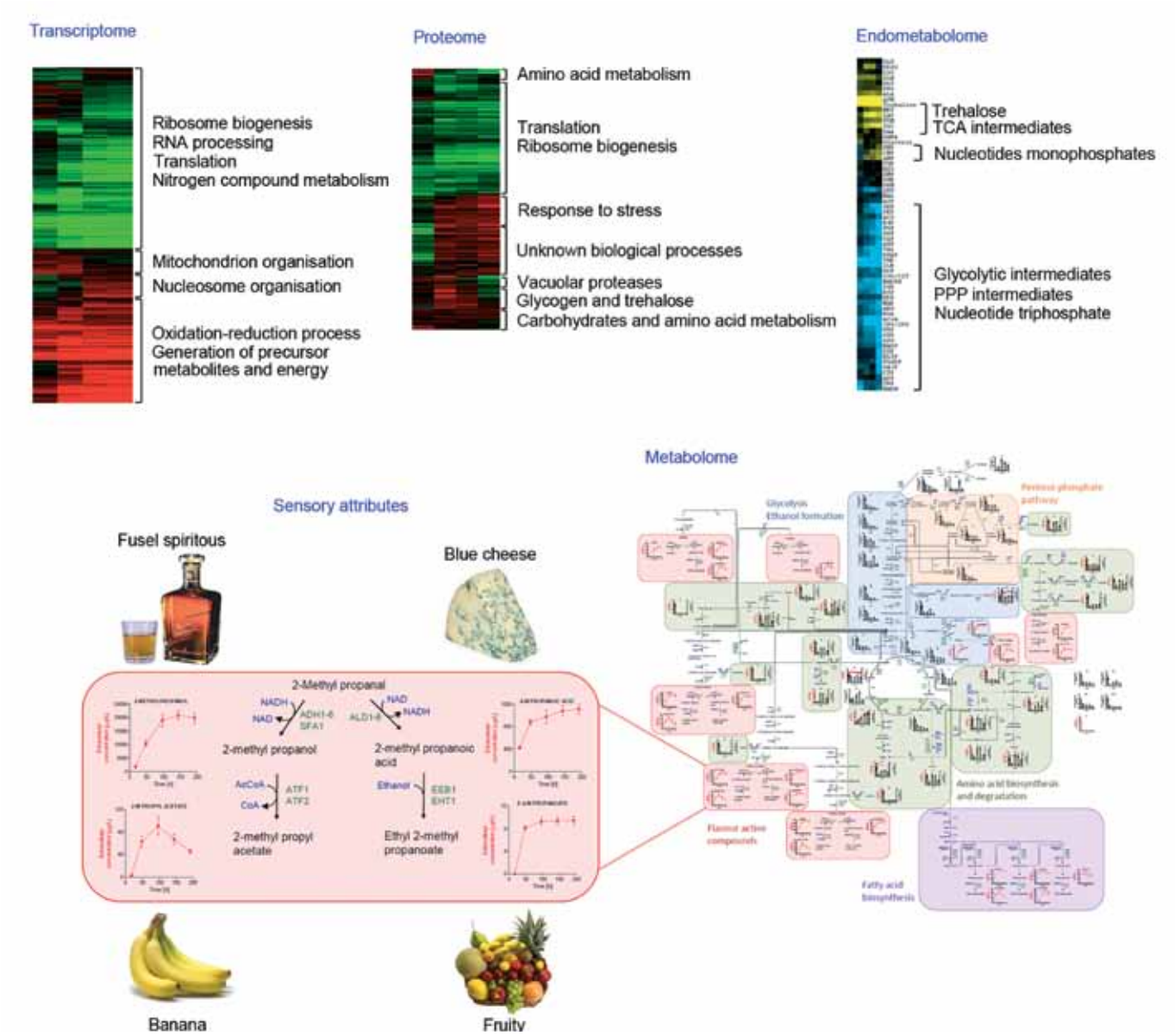

Figure 2. Hierarchical clustering and heat maps for transcriptome, proteome and endometabolome. Metabolite concentrations were used to build a metabolic map, which includes central carbon and nitrogen metabolism, fatty acid biosynthesis and the formation of flavour-active compounds. Many metabolites relate to sensory attributes we can smell in wine.

metabolism and transport, translation and ribosome biogenesis; proteins up-regulated relative to the first time point, including proteins related to stress response and several proteins with unknown biological functions; and proteins up-regulated after 48 hours (second time point), these were associated with vacuolar processes, related to glycogen and trehalose metabolism, and carbohydrate and amino acid metabolic processes. Such a profile is consistent with recycling of cellular components. Quantification of 115 metabolites was achieved using a targeted approach, while non-targeted analysis revealed 340 potential metabolites. Cluster analysis of the exometabolome showed two groups of metabolites: glucose, fructose, ammonia and amino acids decreasing during fermentation and esters, higher alcohols, medium-chain fatty acids (MCFAs) and glycerol increasing. Cluster analysis of the endometabolome revealed, as one might predict, that trehalose and tricarboxylic acid (TCA) cycle intermediates increased, nucleotide monophosphates increased moderately, while glycolytic intermediates, intermediates of the pentose phosphate pathway and nucleotide triphosphates decreased. Metabolite concentrations have been used to build a metabolic map covering central carbon metabolism, nitrogen and amino acid metabolism, fatty acid biosynthesis and the formation of flavour-active compounds. Flavour-active compounds are often associated with sensory attributes and, therefore, give valuable information about one of the key system phenotypes, wine aroma.

We are currently working to integrate individual data from the different platforms to build a picture of our system and identify gaps that need to be filled. With this information as a foundation

we are embarking on the development of predictive mathematical models, enabling informed development of improved, novel wine yeast strains.

Acknowledgements

The AWRI, a member of the Wine Innovation Cluster in Adelaide, is supported by Australia's grape growers and winemakers through their investment body the *Grape and Wine Research Development Corporation* with matching funds from the Australian Government. Systems biology research at the AWRI is performed using resources provided as part of the *National Collaborative Research Infrastructure Strategy*, an initiative of the Australian Government, in addition to funds from the Government of South Australia.

** The Australian Wine Yeast Systems Biology Consortium:

Adelson DL (University of Adelaide), Baker MS (Macquarie University), Barrero R (ABF), Bellgard M (ABF), Borneman AR (AWRI), Chambers PJ (AWRI), Forrest S (GA), Hack J (MA), Hoffman P (PA), Janitz K (GA), Khan A (PA), Kroemer JO (MA), Li S (UNSW), Mercurio M (MA), Moolhuijzen P (ABF), Roessner U (MA), Schmidt SA (AWRI), Varela C (AWRI) and Wilkins MR (UNSW).

ABF, Australian Bioinformatics Facility; AWRI, Australian Wine Research Institute; GA, Genomics Australia; MA, Metabolomics Australia; PA, Proteomics Australia; UNSW, University of New South Wales.

References

1. Canelas, A.B. *et al.* (2010) Integrated multilaboratory systems biology reveals differences in protein metabolism between two reference yeast strains. *Nat. Commun.* 1, 145.
2. Cascante, M. and Marin, S. (2008) Metabolomics and fluxomics approaches. *Essays in Biochemistry: Systems Biology* 45, 67–81.
3. Snyder, M. and Gallagher, J.E.G. (2009) Systems biology from a yeast omics perspective. *FEBS Letters* 583, 3895–3899.
4. Goffeau, A. *et al.* (1996) Life with 6000 genes. *Science* 274, 546, 563–567.
5. Hohmann, S. (2005) The Yeast Systems Biology Network: mating communities. *Curr. Op. Biotechnol.* 16, 356–360.
6. Petranovic, D. and Nielsen, J. (2008) Can yeast systems biology contribute to the understanding of human disease? *Trends Biotechnol.* 26, 584–590.
7. Winzeler, E.A. *et al.* (1999) Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* 285, 901–906
8. Dobson, P.D. *et al.* (2010) Further developments towards a genome-scale metabolic model of yeast. *BMC Systems Biol.* 4, 145.
9. Herrgard, M.J. *et al.* (2008) A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology. *Nature Biotechnol.* 26, 1155–1160.
10. Borneman, A.R. *et al.* (2007) Yeast systems biology: modelling the winemaker's art. *Trends Biotechnol.* 25, 349–355.
11. Kresnowati, M.T.A.P. *et al.* (2006) When transcriptome meets metabolome: fast cellular responses of yeast to sudden relief of glucose limitation. *Mol. Syst. Biol.* 2, 49.
12. Walther, D. *et al.* (2010) Metabolic pathway relationships revealed by an integrative analysis of the transcriptional and metabolic temperature stress-response dynamics in yeast. *OMICS* 14, 261–274.
13. Borneman, A.R. *et al.* (2008) Comparative genome analysis of a *Saccharomyces cerevisiae* wine strain. *FEMS Yeast Research* 8, 1185–1195.
14. Borneman, A.R. *et al.* (2011) Whole-genome comparison reveals novel genetic elements that characterize the genome of industrial strains of *Saccharomyces cerevisiae*. *Plos Genetics* 7, e1001287.
15. Rossignol, T. *et al.* (2009) The proteome of a wine yeast strain during fermentation, correlation with the transcriptome. *J. Appl. Microbiol.* 107, 47–55.
16. Zuzuarregui, A. *et al.* (2006) Transcriptomic and proteomic approach for understanding the molecular basis of adaptation of *Saccharomyces cerevisiae* to wine fermentation. *Appl. Environ. Microbiol.* 72, 836–847.
17. Rossouw, D. *et al.* (2008) Linking gene regulation and the exo-metabolome: A comparative transcriptomics approach to identify genes that impact on the production of volatile aroma compounds in yeast. *BMC Genomics* 9, 530.

Biographies

Dr Cristian Varela is a senior research scientist at the AWRI, leading the Low-Ethanol Project. He completed his undergraduate degree in Biochemistry at the Catholic University of Chile where he also completed a Master of Biochemistry and a PhD in Chemical Engineering and Bioprocesses. Cristian joined the AWRI in 2004 and his main research interests are to understand yeast metabolic networks and the interaction between gene expression, protein levels and metabolite concentrations, that is, systems biology.

Dr Simon Schmidt is a research scientist at the Australian Wine Research Institute (AWRI). He earned his Bachelor of Science in Biological Science from Flinders University, followed by an honours degrees studying cellular differentiation in *Dictyostelium discoidium* and subsequently worked on signal transduction pathways controlling intercellular communication via gap junctions. Simon completed his PhD studying signalling mechanisms leading to disease resistance in plants and was the first to develop a recombinant expression and purification method for the 'nucleotide binding site-leucine rich repeat' class of plant disease resistance receptor proteins. Since moving to the AWRI, Simon's key interests have evolved to include the relationship between nutrient availability and yeast fermentation performance and developing a systems-level understanding of *Saccharomyces cerevisiae* fermentative metabolism.

Dr Alamgir Khan obtained his PhD from Macquarie University studying biochemical interactions of nematophagous fungi and plant parasitic nematodes eliciting molecular mechanisms of infection for biocontrol. Alamgir has over 13 years' experience in proteomics, primarily in the technology development, sample preparation, two-dimensional gel-based separation of proteins and bead-based immunodetection of low abundance proteins in biological samples for biomarker discovery. Alamgir joined APAF in 2005, following five years of industry experience gained at Proteome Systems Ltd.

Dr Paul Chambers is Research Manager at The Australian Wine Research Institute where he oversees research projects that cover a range of aspects of improving wine yeast performance and generating novel wine yeast strains with improved winemaking characteristics.