Review

At the cutting-edge of grape and wine biotechnology

Anthony R. Borneman[1], Simon A. Schmidt[1], and Isak S. Pretorius[2]

[1] The Australian Wine Research Institute, PO Box 197, Adelaide, SA 5064, Australia
[2] University of South Australia, PO Box 2471, Adelaide, SA 5001, Australia

Wine is arguably the oldest biotechnological endeavor, with humans having been involved in wine production for at least 7000 years. Despite the artisan nature of its production, work by pioneering scientists such as Antoine-Laurent de Lavoisier and Louis Pasteur placed wine research in a prominent position for the application of cutting-edge biological and chemical sciences, a position it still holds to this day. Technologies such as whole-genome sequencing and systems biology are now revolutionizing winemaking by combining the ability to engineer phenotypes rationally, with a precise understanding of the genetic makeup and key phenotypic drivers of the key organisms that contribute to this age-old industry.

The interwoven histories of human civilization, winemaking, and science

Wine has been an integral part of human civilization for thousands of years, with the first signs of large-scale winemaking activities dating to at least 5000 BC [1] (Figure 1). Over the subsequent millennia, chance mating, traditional breeding, and selection strategies were applied to the wild precursors of modern grapevines to provide us with the array of *Vitis vinifera* cultivars that are used in wine production today [2]. However, it was not until the 19th and 20th centuries that the microorganisms that were responsible for the conversion of grape must into wine began to be formally classified (Figure 1). The yeast *Saccharomyces cerevisiae* was ultimately identified as the principal microorganism responsible for wine production, and the bacterium *Oenococcus oeni* was shown to be responsible for malolactic fermentation, a secondary fermentation that takes place in many wines (Figure 2) [3,4].

Ever since the historical work of Antoine-Laurent de Lavoisier defined the chemical reactions that underpin the fermentation of sugars into ethanol and carbon dioxide, and Louis Pasteur determined the biological basis of this reaction, wine has figured prominently as an applied platform for cutting-edge fundamental research. This rapid adoption of technology, although providing many benefits to the wine industry, including the use of sulfur as a preservative, grafted grapevines, analytical wine chemistry, and starter yeast and bacterial cultures, has also often created a divide between so called 'traditional' or 'Old World' winemakers and those of the 'New World' that seek to use technology to control the winemaking process. No issue has been more polarizing than the use of genetic modification (GM) of grapevines and wine yeast. Despite the power of this technology to significantly alter the characteristics of both the grapevine and wine microorganisms [5,6], GM technology remains all but unused due to bans on its application in wine throughout most major wine-producing countries.

Although the use of organisms produced by GM techniques in the food industry have been condemned by many, all but the most verdant supporters of 'traditional' winemaking practices support the application of classical agricultural improvement strategies, such as mutagenesis and breeding combined with phenotypic selection, to produce new grapevine clones or microbial strains. The development of modern molecular biology techniques, such as high-throughput genomics and systems biology, are now poised to revolutionize the winemaking process. Recent developments in modern molecular biology, such as high-throughput genomics and systems biology, are poised to revolutionize winemaking. By providing the means to rapidly characterize the scope of genetic diversity available to breeders and biotechnologists alike, these technologies provide the capability to fast-track identification of a phenotypic trait's genetic basis. Precisely defining the genetic sources of specific phenotypes will facilitate the systematic assembly of multiple desirable traits into grapevine and wine microorganisms. Improved genetic definition will increase precision and lead to efficiency gains in traditional development programs. Such opportunities may also

Glossary

AFLP: amplified fragment length polymorphism.
Linkage disequilibrium (LD): the non-random association of alleles between two loci or between a marker and a QTL within a population. Non-zero measures of LD indicate that the association is not random.
Linkage drag: during the introgression of desirable traits through crossing of two strains or varieties, genes other than that responsible for the selected trait may also be introduced due to tight linkage with the gene of interest. The introduction of these 'hitchhiker' genes may bring with them additional phenotypes, not always desirable.
Linkage group (LG): all genes situated on a single chromosome belong to a single LG. It is possible that a chromosome can appear as more than one LG when information about that chromosome is incomplete.
Quantitative trait loci (QTL): the identification of the causal genetic variation that underscores phenotypic characteristics often involves the mapping of QTLs. This process relies on crossing strains with contrasting phenotypes (e.g., a high and low producer) and analyzing the progeny to identify genetic markers that cosegregate with phenotypes of interest. This identifies genomic region(s) contributing to the trait while providing genetic markers that can be used to efficiently screen progeny.
RAPD: random amplification of polymorphic DNA.
SSR: simple sequence repeat.

Corresponding author: Pretorius, I.S. (sakkie.pretorius@unisa.edu.au).
Keywords: biotechnology; grapes; malolactic bacteria; wine yeast; winemaking.

  http://dx.doi.org/10.1016/j.tig.2012.10.014

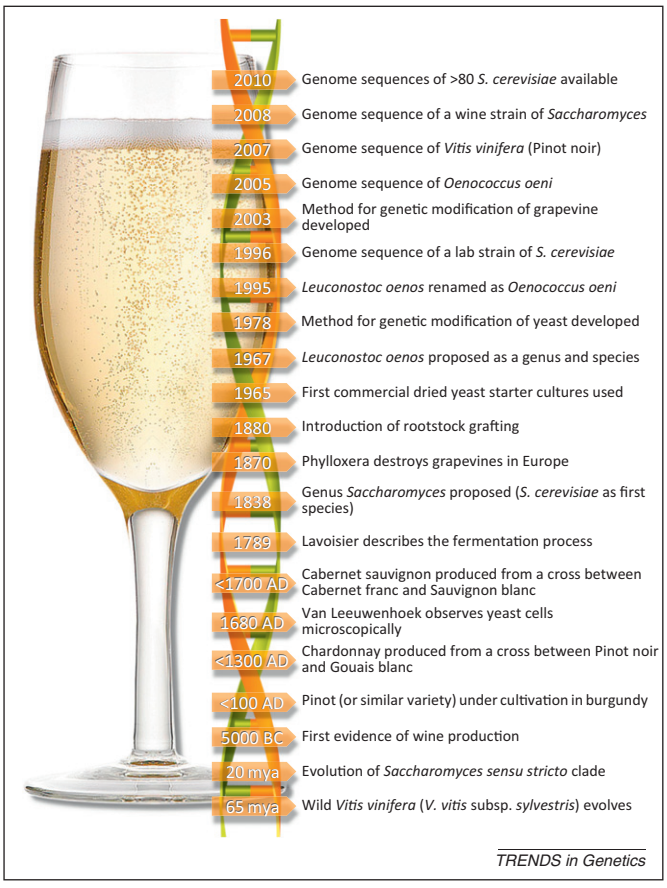

Figure 1. A timeline of significant events as winemaking transitioned through the ages of mystery to discovery and finally technology.

provide an area of compromise between traditionalists and technologists in the winemaking world.

Characterization of genetic diversity

There is a wide array of phenotypic variation present across different cultivars of *V. vinifera* and different strains of *S. cerevisiae* and *O. oeni*. These phenotypic differences have direct implications not only for wine quality but also for the efficiency of production of this industrialized product. However, the underlying genetic basis for the vast majority of these phenotypic differences remains unclear.

The genomic revolution that is being brought about by advances in next-generation sequencing is providing the means to answer these questions by making available the technology to map quickly and accurately the genetic diversity present within grapevine, wine yeast, and bacteria.

Owing to its position as one of the most intensively studied biological model systems, in 1996 the laboratory strain of *S. cerevisiae* (S288c) became the first eukaryote to have its genome fully sequenced [7]. Genomic data are now available for over 80 strains of *S. cerevisiae*, including six commercial wine yeasts and at least another dozen strains that were isolated either from wine, grapes, or vineyards [8–21]. The remaining sequences represent a variety of isolates from varied geographic and environmental sources, and strains from other yeast-based industries such as brewing, baking, and bioethanol production. This provides a rich dataset for assembling the *S. cerevisiae* phylogeny while also providing information linking genetic variation with specific environmental or industrial niches.

The first wine yeast genome sequence became available in 2008 [11]. Subsequent genomic analysis of *S. cerevisiae* wine yeasts have shown that wine strains form a tight geographic clade alongside environmental yeast isolates [11,21,22]. This is surprising given that some isolates originated outside of Europe (e.g., RM11-1a isolated from an American vineyard and DBVPG1106 isolated from Australian grapes). The close genotypic relationship is presumably due to the exportation of European wine yeast into wine-growing areas, either on grapevines or through the use of commercial inoculated yeast species at nearby wineries, although until large-scale haplotyping of these diploid strains can be reliably performed, admixture

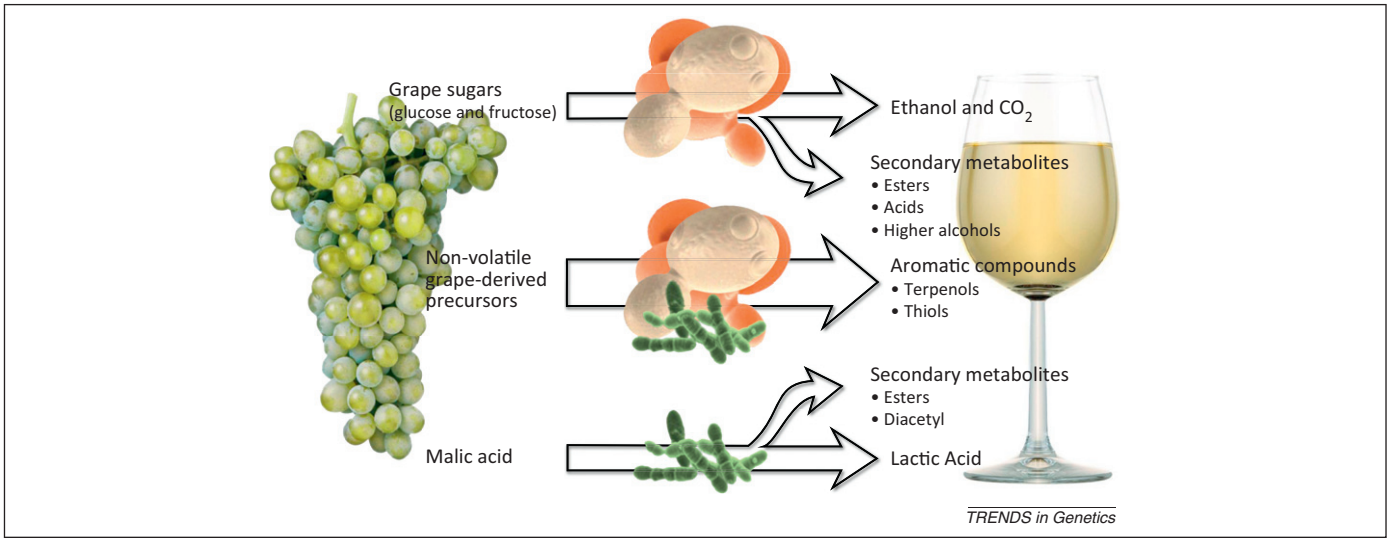

Figure 2. The impact of yeast and bacteria during wine fermentation. Wine contains a large number of volatile flavor and aroma compounds that are not present in grape juice and which are produced through the action of microorganisms. The primary role of the yeast *Saccharomyces cerevisiae* is to convert grape sugars into ethanol (top). However, during this process many secondary metabolites are also produced that impart distinct flavors and aromas to the final wine. The primary role of the bacterium *Oenococcus oeni* is to undertake malolactic fermentation in which wine is de-acidified through the conversion of malic acid to lactic acid (bottom). As seen for yeast, the growth of *O. oeni* also produces secondary metabolites that impact upon wine flavor. Both species also produce 'flavor-releasing enzymes' that release volatile aroma molecules from non-volatile precursors present in the grape (usually as sugar-bound glycosides). Adapted from [5].

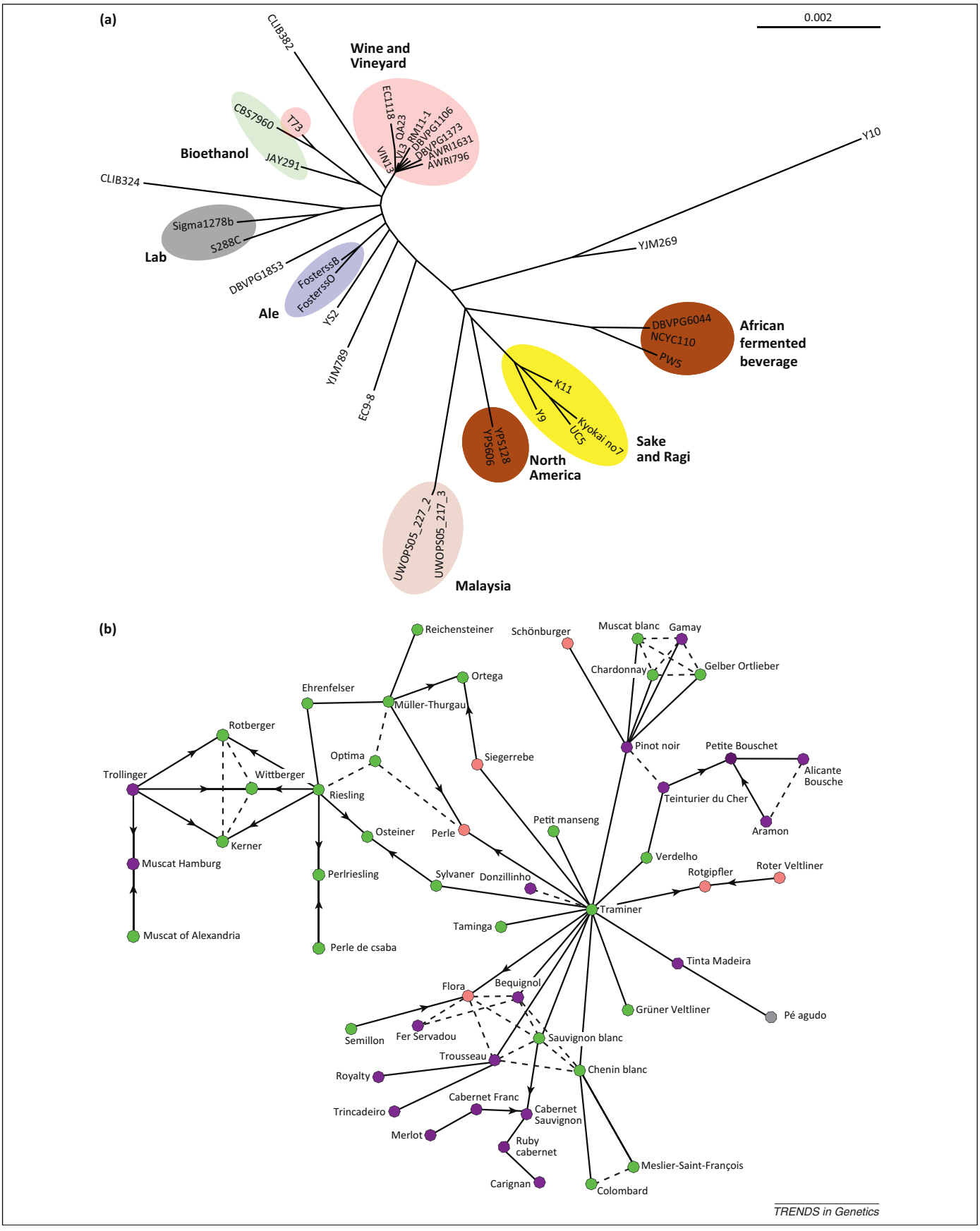

Figure 3. Genetic relationships between cultivars and strains of grapevines and yeast uncovered by genomic sciences. **(a)** A maximum-likelihood phylogeny for *Saccharomyces cerevisiae* based upon single-nucleotide polymorphism (SNP) variation present in whole-genome alignments. Color shading is used to highlight strains associated with specific industries or geographic locations. **(b)** A modified pedigree of 54 cultivars of *Vitis vinifera* as determined by [78] using whole-genome *Vitis* SNP genotyping arrays. Nodes are shaded according to the color of the grape berries produced by each cultivar (blanc, green; rosé, pink; noir, purple).

between these commercial strains and native yeast strains cannot be ruled out (Figure 3a). However, one of the most important findings that has emerged from these whole-genome comparisons is the differential presence of many genomic regions in specific subsets of *S. cerevisiae* strains. This is a situation that is common in many bacterial species, including *O. oeni* (see below), in which the species-wide, pan-genome is much larger than that found in any single representative [23].

The genomes of the commonly used wine yeast EC1118 and its relatives QA23 and VL3 were shown to contain at least one of two major telomeric insertions relative to the laboratory strain – a 40 kb insertion in chromosome VI and a 65 kb insertion in chromosome XV [11,18]. The insertion on chromosome XV contains at least three genes with potential to affect wine-relevant traits. One is a high-affinity fructose/H^+ symporter (a homolog of the *Saccharomyces pastorianus* gene *FSY1*) [24]; the other two comprise a pair of tandemly duplicated oligopeptide transporters, which potentially allow for a greater variety of small peptides to be used as nitrogen sources [18,25]. Interestingly a fourth open reading frame (ORF) from this genomic insertion was shown to encode a putative xylitol dehydrogenase that provides strains with the ability to utilize xylose as a carbon source, a phenotype of particular interest to the bioethanol industry [26].

The wine strain AWRI796 has a strain-specific ~45 kb insertion in the telomere of chromosome XV. This locus encodes up to 18 proteins, including three putative aryl-alcohol dehydrogenases, that are distinct from other members of the *S. cerevisiae* aryl-alcohol dehydrogenase family [11]. Given the expected role of these enzymes in converting ketones and aldehydes into corresponding aroma-active alcohols, these genes are candidates for shaping wine sensory attributes.

One of the most striking aspects of the wine yeast genome is the presence of a cluster of five genes that are postulated to have been horizontally transferred between *S. cerevisiae* and *Zygosaccharomyces* spp. [18]. Although present in many wine strains and at least one biofuel strain, the location, copy-number, and exact order of the genes within the cluster is both strain- and insertion site-dependent. The genesis of these various genomic insertions has been hypothesized to occur via the formation of a circular intermediate through an undetermined process that appears to be independent of classical recombinatorial or transposon-based duplication and insertion [11]. Interestingly, since the time of this discovery in yeast, similar multicopy insertions potentially involving circular intermediates have been found in the *Tilapia* (a freshwater fish) [27] and *Bos taurus* (bovine) [28] genomes, raising the possibility that a novel but widely conserved transposition mechanism may exist in eukaryotes.

The first genome sequence of *O. oeni*, strain PSU-1, was published in 2005 as part of a broad phylogenetic sequencing project focused on lactic acid bacteria. There are now an additional 13 strains of *O. oeni* for which whole-genome information is available [29–32]. The genome of *O. oeni* is compact (1.8 Mb), and individual strains encode between 1700 and 1900 proteins. However, as mentioned previously, like many bacterial species, this protein complement can vary by around 10% between strains. As a result, the pan-genome of *O. oeni* currently stands at around 3000 proteins, of which 1165 comprise the conserved core gene set found in all 14 strains [32].

The majority of differences in ORF content between strains of *O. oeni* are due to the differential presence of lysogenic bacteriophages [32]. However, the genome of the only other member of the *Oenococcus* genus, *Oenococcus kitaharae*, completely lacks lysogenic bacteriophage, presumably due to the presence of a functional CRISPR element pathway that is absent from *O. oeni* [33].

Outside of integrated bacteriophage elements, the pan-genome of *O. oeni* appears to provide strain-specific differences in cell-wall exopolysaccharide content, sugar utilization, and transport across the plasma membrane [31,32,34]. There are also accounts of intrastrain variation in the presence of β-glycosidases [31,32]. These enzymes are important in liberating volatile aroma molecules from non-volatile sugar-bound precursors that are present in grape juice. β-Glycosidase enzyme variation between strains may be important in strain-specific flavor profiles that can be produced through the use of different strains of *O. oeni* for malolactic fermentation.

Given the much larger size of the *V. vinifera* genome (450 Mb) compared to both *S. cerevisiae* (12 Mb) and *O. oeni* (1.8 Mb), it is not surprising that genome sequencing data for this species have lagged behind that of its microbial counterparts. However, the sequencing of Pinot noir [35] and the Pinot-derived haploid PN40024 [36] in 2007 has revolutionized genetic studies of grapevine by providing linked genetic and physical maps that will ultimately be used to pinpoint genetic determinants of traits of viticultural and enological importance [37,38]. Analysis of these two genomes revealed the complex hexaploid genetic heritage of grapevine by identifying triploid paralogous regions within the genome, as well as the expansion of gene families with roles in wine quality metrics, such as anthocyanin (responsible for wine color) and monoterpene (an aroma compound important to particular wine styles) production, compared with other sequenced dicotyledonous species [39,40].

Of immediate biotechnological interest was the identification of 341 plant disease-resistance genes (R-genes) in the Pinot noir genome sequence. Several clusters of disease resistance genes were mapped to chromosomal regions where resistances to fungal diseases were previously assigned [41]. These clusters of disease resistance genes have been used to study grapevine genome evolution, an interesting example of the crossover between biotechnological imperatives and evolutionary understanding [42]. The triplet groupings of *V. vinifera* chromosomes and the stable gene order within R-gene clusters were used to identify putative component genomes of grapevine while also probing the underlying pathways of R-gene inheritance. This detailed knowledge of R-gene cluster structure and the relationships between them, delivered through whole-genome sequencing is laying the foundation for, and expanding knowledge of grapevine diversity, beyond that present in established germplasm collections (Box 1).

Box 1. Genomic diversity, R genes, and pathogen resistance

Expanding the understanding of grapevine genetic diversity beyond that which exists in established germplasm collections will contribute to the continued development of breeding lines containing stable and robust traits, especially for resistance to disease. Recent work characterizing resistance to the fungus *Plasmopara viticola* (downy mildew) is revealing in this context. *Rpv3*, the major component of defense in native North American grapevines, controls the ability to trigger a race-specific resistance response to downy mildew infection [95], and has been the basis of many breeding programs aimed at introducing downy mildew resistance into *Vitis vinifera* [96].

Rpv3 resides on the lower arm of chromosome 18 [linkage group (LG) 18] within 2 Mb, and tightly linked to, a second resistance locus, *Rpv2* [97]. Thirteen conserved haplotypes of *Rpv3* have been identified, with all haplotypes being present at a single locus [96,97]. This region contains more than 100 NBS-LRR (nucleotide binding site leucine-rich-repeat) R-genes, including some for powdery mildew [73], and exhibits a low recombination frequency. Tight linkage of the cluster has complicated positional cloning and potentially restricts the combining of different resistance alleles from within the linkage group.

Until recently, only one other downy resistance locus, *Rpv1* on LG 12, was known outside LG 18 [98]. As a result there has been a push to identify additional determinants of downy mildew resistance residing outside the *Rpv3* linkage group. Recent work with wild Asian cultivars has identified additional resistance loci, *Rpv10* on LG 9 [99] and *Rpv8* on LG 14 [100]. That this degree of diversity exists raises the prospect that stable and durable resistance to downy mildew may in fact be achievable within a single breeding line. Such findings are of particular importance in light of recent work describing the breakdown of *Rpv3*-based resistance by *P. viticola* in the field [101].

Linking genetic diversity with phenotypic variation

In commercial-scale agriculture, such as the modern-day wine sector, there is a need to relate genotypic variation with important phenotypes to efficiently exploit genetic resources in strain and clone development programs. Genotype-phenotype associations are increasingly being provided by integrated systems models, not only of individual organisms, but increasingly by a molecular-genetic grape-to-glass picture of winemaking.

Wine-omics and systems biology

Whole-genome methodologies such as transcriptomics, proteomics, and metabolomics (collectively termed 'omics techniques) are being applied in winemaking to define phenotypic variation at the molecular level and also to assign genetic contributors to variation.

The application of individual 'omic methodologies continues to provide insight into many areas of grapevine biology including vine development [43–50], disease resistance [51,52], and viticultural practices such as water and light deficit [53–55]. The existence of such diverse datasets has driven attempts at their integration [56]. Moreover, the range of different 'omics data types has led to the exploration of novel analytical methods, or the application of existing analytical methods in novel ways, in an attempt to extract biologically relevant information.

Multivariate analytical methods are increasingly being used to uncover relationships between complex 'omics datasets. An example of this approach is the identification of developmental stage- and process-specific biomarkers [57]. One study used a combination of principal component analysis and bidirectional orthogonal projections to latent structures (O2PLS-DS) on individual 'omics datasets to define developmental classes and class-specific variables. These class definitions and variable attributes were then used in subsequent rounds of two-way O2PLS comparisons of transcriptomic, proteomic, and metabolomic data to identify covariant structures between datasets. Systems-level representations of berry development and withering were built using these covariant structures, exposing the role of sphingolipid fatty acids during berry growth and the activation of specific stress-response genes during withering. For a summary of this and related multivariate approaches to data integration see [58].

Following the identification of key contributors to variation between treatments, as described above, some methods of analysis still require target identification. This is the case for some forms of metabolomic analysis. The use of untargeted metabolomics for the profiling of grapevine is increasingly being used as a discriminatory or phenotyping tool [43,45]. However, resources such as the tandem mass spectral database for phytochemicals promise to advance metabolite profiling beyond basic classification schemes to detailed phenotypic characterization, by providing the means to identify unknown metabolites unambiguously from high-quality profiling datasets [59].

Given its pioneering position in genomics, *S. cerevisiae* has also been at the forefront of transcriptomic work, and this remains the major 'omics platform used for the study of yeast biology [60,61]. Several transcriptomic studies that have examined the impact of various wine-relevant environmental perturbations on gene expression, including the rehydration of dried active wine yeast [62,63] and the fermentation process either in isolation [64] or co-inoculated with *O. oeni* [65]. However, data from a comparative standpoint are limited to a group of studies that compared the transcriptomic response of five different commercial wine strains with phenotypic data and metabolomic analysis to ultimately link differences in aroma production to variation within transcription factor networks [66,67]. In addition, a recent study applied a systems-biology framework of transcriptomics, proteomics, and flux analysis to investigate the impact of NADP availability on ethanol production in a genetically-modified low ethanol-producing strain [68].

Gene mapping

The extensive development of genetic resources including AFLP, SSR, and RAPD markers (see Glossary) for grapevine genetics in the era before the availability of the Pinot noir genome sequence meant that the majority of traits of viticultural and enological interest were first described as quantitative trait loci (QTL). However, the genomic intervals that could be defined for these traits, even when combining multiple marker types, was often limited to regions of hundreds of kb or more [69]. In addition, high percentages of repetitive sequence elements, variable recombination frequencies, and large clusters of homologous sequences made defining precise genetic determinants of these traits difficult. As a result, genetic determinants are known for a relatively small number of traits such as color

and aroma [70,71], dwarfism [72], flower hermaphroditism [73], and plant pathogen resistance [74,75].

Markers, predominantly SSRs, derived from classical genetic maps used to map the genes discussed above are still used to define the limits of genetic variation within existing germplasm collections [76,77]. However, the availability of next-generation sequencing and whole-genome sequence information is leading to the development of single-nucleotide polymorphism (SNP)-based genotyping arrays that are revealing the level of genetic diversity at the disposal of breeders. Using SNP genotyping arrays, it was shown that 75% of the US Department of Agriculture grape germplasm collection was related to at least one other cultivar by a first-degree relationship [78] (Figure 3b). Furthermore, it was shown that the majority of the genetic diversity present in domesticated grapevines is the result of spontaneous somatic mutation as opposed to sexually derived, segregating polymorphisms.

Although low-density SNP-based arrays allowed the rapid assessment of genetic diversity, they have also exposed the limitation of this technology in the assessment of high-diversity species via genome-wide association studies (GWAS). Owing to the rapid decay of linkage disequilibrium in grapevine, these low-density genotyping platforms, although providing orders of magnitude more markers than traditional methods, still fall short of providing significant linkage to many phenotypic characteristics [79]. However, the rapid development of next-generation sequencing technology means that genotyping-by-sequencing is now a viable option for GWAS because it effectively provides an infinitely dense SNP map tailored to the population under investigation, thereby leading to 'whole-genome-assisted' rather than simply marker-assisted breeding techniques [80,81].

In *S. cerevisiae*, the availability of molecular techniques such as efficient transformation allow gene mapping by molecular genetic means, including mutant isolation and gene complementation, rather than via techniques such as QTL analysis. However, because desirable industrial yeast traits may be polygenic, methods such as QTL mapping have received increasing attention in recent years. Despite the later application of QTL mapping in yeast relative to grapevine, the early development of genomic resources has seen the rapid adoption of high-density SNP genotyping microarrays [82] and, recently, genotyping-by-sequencing [83] being used routinely for QTL analysis in this species, further accelerating the potential for this classical technique to be used in a powerful whole-genome sequencing-based workflow.

Given the benefits of using whole-genome data for QTL mapping, several studies have applied this technique in industrial strains of *S. cerevisiae*. These identified the loci responsible for several brewing characteristics, including key flavor metabolite production, in saké yeast [84], and loci responsible for ethanol resistance [19] and xylose utilization [26] for application in the bioethanol industry. However, despite the promise of this technique, only two studies have focused on wine-relevant traits, through mapping the loci responsible for influencing acetic acid production [85] and differences in fermentation performance between wine and laboratory strains [86].

The future of winemaking – is there a place for GM?

Moving forward, the winemaking industry faces challenges both from environmental change and from consumer demands. The application of genomics and genome-assisted breeding will provide the means to address many issues regarding the tailoring of wine flavor and the production of industrial microorganisms with increased robustness and fermentation performance. However, there are future challenges to the wine industry that may exceed the capacity of even these technology-assisted classical techniques to address efficiently. In these situations a combination of GM with non-GM techniques will need to be deployed within a systems-biology framework. In fact, it may be in addressing such vital needs of the industry, while providing a direct measurable benefit to the consumer, that GM strains may obtain the level of industry backing that is required to address the concerns of the general public.

The issue of climate change, combined with consumer preferences for full-flavored ripe berry characteristics, has meant that sugar levels in the mature grape berries, and therefore alcohol concentrations in finished wines, have been steadily increasing in recent years [87]. The push to produce full-flavored wines that are lower in alcohol has therefore received considerable attention, primarily from yeast researchers who have looked to divert carbon flux away from ethanol [87]. However, it may be that the search for a low-ethanol wine yeast strain is one area that may not be readily achievable by a non-GM strategy, even with the application of systems biology. In this case, researchers are effectively trying to unravel millions of years of evolution that has made *S. cerevisiae* a highly efficient and reliable producer of ethanol with very little variation in the amount of ethanol that is produced by different *S. cerevisiae* strains [88]. As such, although there has been one study in which a non-GM, adaptive evolution, approach was used to reduce ethanol with minor success [89], GM options in this area have produced far more significant reductions [87].

Likewise, one of the challenges to viticulture is the consumer and economic push to limit the use of pesticides in the vineyard where, particularly in areas with relatively cool, wet summers, such as large parts of Europe, large fungicide inputs are required. In addition to representing a significant economic burden to producers, especially in situations where fungicide resistance is emerging [90], these significant and repeated chemical additions pose potential environmental and public health concerns. Although the introgression of fungicide-resistance alleles into susceptible grapevine clones is benefiting from genome-assisted breeding (Box 1), complications with classical grapevine breeding, such as inbreeding repression during back-crossing and negative trait introduction through linkage drag, contribute to difficulties in genetic trait maintenance and recovery [91–93].

Transgenic approaches offer the possibility of circumventing these difficulties by avoiding traditional breeding and backcrossing altogether, while also bridging mating barriers to introduce novel resistance mechanisms to the species. This is exemplified by the recent success in engineering resistance to *Xylella fastidiosa*, a bacterial pathogen that causes Pierce disease in grapes. Through the

introduction of a protein chimera of human neutrophil elastase and cecropin B into Thompson seedless table grapes, resistance to this bacterium was established. This represents a phenotype that is lacking in natural cultivars of *V. vinifera*, and which would therefore be impossible to introduce by classical selection and breeding [94].

Concluding remarks

New genomic technologies are providing the methodologies for the rapid identification of the genetic loci that shape agronomically and industrially important traits. This information will provide precise markers for genome-assisted breeding programs and enable the development of new grapevine cultivars, wine yeasts, and malolactic bacteria. These strategies, combined with the possibility of applying GM techniques under the rigorously controlled circumstances provided by a systems-biology framework, will offer improved performance from the vine to the glass and a means of tailoring wine sensory properties to meet consumer demand.

Acknowledgments

The Australian Wine Research Institute (AWRI), a member of the Wine Innovation Cluster situated at the Waite research precinct in Adelaide, is supported by Australia's grapegrowers and winemakers through their investment body the Grape and Wine Research Development Corporation, with matching funding from the Australian Government. Systems-biology research at the AWRI uses resources provided as part of the National Collaborative Research Infrastructure Strategy (NCRIS), an initiative of the Australian Government, in addition to funds from the South Australian State Government. AWRI collaborating partners within this NCRIS-funded initiative – overseen by Bioplatforms Australia – are Genomics Australia, Proteomics Australia, Metabolomics Australia (of which the Microbial Metabolomics unit is housed at the AWRI) and Bioinformatics Australia. The authors would also like to thank Dr Paul Chambers for critical comments on this manuscript.

References

1 McGovern, P.E. *et al.* (1996) Neolithic resinated wine. *Nature* 381, 480–481
2 This, P. *et al.* (2006) Historical origins and genetic diversity of wine grapes. *Trends Genet.* 22, 511–519
3 Barnett, J.A. (1992) The taxonomy of the genus *Saccharomyces meyen ex reess*: a short review for non-taxonomists. *Yeast* 8, 1–23
4 Garvie, E.I. (1967) *Leuconostoc oenos* sp. nov. *J. Gen. Microbiol.* 48, 431–438
5 Pretorius, I.S. *et al.* (2012) The winemaker's bug: from ancient wisdom to opening new vistas with frontier yeast science. *Bioeng. Bugs* 3, 147–156
6 Vivier, M.A. and Pretorius, I.S. (2002) Genetically tailored grapevines for the wine industry. *Trends Biotechnol.* 20, 472–478
7 Goffeau, A. *et al.* (1996) Life with 6000 genes. *Science* 274, 546–567
8 Akao, T. *et al.* (2011) Whole-genome sequencing of sake yeast *Saccharomyces cerevisiae* Kyokai no. 7. *DNA Res.* 18, 423–434
9 Argueso, J.L. *et al.* (2009) Genome structure of a *Saccharomyces cerevisiae* strain widely used in bioethanol production. *Genome Res.* 19, 2258–2270
10 Babrzadeh, F. *et al.* (2012) Whole-genome sequencing of the efficient industrial fuel-ethanol fermentative *Saccharomyces cerevisiae* strain CAT-1. *Mol. Genet. Genomics* 287, 485–494
11 Borneman, A.R. *et al.* (2011) Whole-genome comparison reveals novel genetic elements that characterize the genome of industrial strains of *Saccharomyces cerevisiae*. *PLoS Genet.* 7, e1001287
12 Borneman, A.R. *et al.* (2012) The genome sequence of the wine yeast VIN7 reveals an allotriploid hybrid genome with *Saccharomyces cerevisiae* and *Saccharomyces kudriavzevii* origins. *FEMS Yeast Res.* 12, 88–96
13 Borneman, A.R. *et al.* (2008) Comparative genome analysis of a *Saccharomyces cerevisiae* wine strain. *FEMS Yeast Res.* 8, 1185–1195
14 Doniger, S.W. *et al.* (2008) A catalog of neutral and deleterious polymorphism in yeast. *PLoS Genet.* 4, e1000183
15 Dowell, R.D. *et al.* (2010) Genotype to phenotype: a complex problem. *Science* 328, 469
16 Halfmann, R. *et al.* (2012) Prions are a common mechanism for phenotypic inheritance in wild yeasts. *Nature* 482, 363–368
17 Nijkamp, J.F. *et al.* (2012) *De novo* sequencing, assembly and analysis of the genome of the laboratory strain *Saccharomyces cerevisiae* CEN.PK113-7D, a model for modern industrial biotechnology. *Microb. Cell Fact.* 11, 36
18 Novo, M. *et al.* (2009) Eukaryote-to-eukaryote gene transfer events revealed by the genome sequence of the wine yeast *Saccharomyces cerevisiae* EC1118. *Proc. Natl. Acad. Sci. U.S.A.* 106, 16333–16338
19 Swinnen, S. *et al.* (2012) Identification of novel causative genes determining the complex trait of high ethanol tolerance in yeast using pooled-segregant whole-genome sequence analysis. *Genome Res.* 22, 975–984
20 Wei, W. *et al.* (2007) Genome sequencing and comparative analysis of *Saccharomyces cerevisiae* strain YJM789. *Proc. Natl. Acad. Sci. U.S.A.* 104, 12825–12830
21 Liti, G. *et al.* (2009) Population genomics of domestic and wild yeasts. *Nature* 458, 337–341
22 Dunn, B. *et al.* (2012) Analysis of the *Saccharomyces cerevisiae* pan-genome reveals a pool of copy number variants distributed in diverse yeast strains from differing industrial environments. *Genome Res.* 22, 908–924
23 Tettelin, H. *et al.* (2008) Comparative genomics: the bacterial pan-genome. *Curr. Opin. Microbiol.* 11, 472–477
24 Galeote, V. *et al.* (2010) FSY1, a horizontally transferred gene in the *Saccharomyces cerevisiae* EC1118 wine yeast strain, encodes a high-affinity fructose/H^+ symporter. *Microbiology* 156, 3754–3761
25 Damon, C. *et al.* (2011) A novel fungal family of oligopeptide transporters identified by functional metatranscriptomics of soil eukaryotes. *ISME J.* 5, 1871–1880
26 Wenger, J.W. *et al.* (2010) Bulk segregant analysis by high-throughput sequencing reveals a novel xylose utilization gene from *Saccharomyces cerevisiae*. *PLoS Genet.* 6, e1000942
27 Fujimura, K. *et al.* (2011) Circular DNA intermediate in the duplication of *Nile tilapia vasa* genes. *PLoS ONE* 6, e29477
28 Durkin, K. *et al.* (2012) Serial translocation by means of circular intermediates underlies colour sidedness in cattle. *Nature* 482, 81–84
29 Makarova, K. *et al.* (2006) Comparative genomics of the lactic acid bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 103, 15611–15616
30 Mills, D. *et al.* (2005) Genomic analysis of PSU-1 and its relevance to winemaking. *FEMS Microbiol. Rev.* 29, 465–475
31 Borneman, A.R. *et al.* (2010) Genotypic diversity in *Oenococcus oeni* by high-density microarray comparative genome hybridization and whole genome sequencing. *Appl. Microbiol. Biotechnol.* 86, 681–691
32 Borneman, A.R. *et al.* (2012) Comparative analysis of the *Oenococcus oeni* pan genome reveals genetic diversity in industrially-relevant pathways. *BMC Genomics* 13, 373
33 Borneman, A.R. *et al.* (2012) Functional divergence in the genus *Oenococcus* as predicted by genome sequencing of the newly-described species, *Oenococcus kitaharae*. *PLoS ONE* 7, e29626
34 Dimopoulou, M. *et al.* (2012) Exploration of phenomena contributing to the diversity of *Oenococcus oeni* exopolysaccharides. *Int. J. Food Microbiol.* 153, 114–122
35 Velasco, R. *et al.* (2007) A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. *PLoS ONE* 2, e1326
36 Jaillon, O. *et al.* (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* 449, 463–467
37 Moroldo, M. *et al.* (2008) A physical map of the heterozygous grapevine 'Cabernet Sauvignon' allows mapping candidate genes for disease resistance. *BMC Plant Biol.* 8, 66
38 Scalabrin, S. *et al.* (2010) Physical mapping in highly heterozygous genomes: a physical contig map of the Pinot Noir grapevine cultivar. *BMC Genomics* 11, 204
39 Falginella, L. *et al.* (2010) Expansion and subfunctionalisation of flavonoid 3′,5′-hydroxylases in the grapevine lineage. *BMC Genomics* 11, 562
40 Martin, D.M. *et al.* (2010) Functional annotation, genome organization and phylogeny of the grapevine (*Vitis vinifera*)

terpene synthase gene family based on genome assembly, FLcDNA cloning, and enzyme assays. *BMC Plant Biol.* 10, 226
41 Di Gaspero, G. *et al.* (2007) Linkage maps of grapevine displaying the chromosomal locations of 420 microsatellite markers and 82 markers for R-gene candidates. *Theor. Appl. Genet.* 114, 1249–1263
42 Malacarne, G. *et al.* (2012) Deconstruction of the (paleo)polyploid grapevine genome based on the analysis of transposition events involving NBS resistance genes. *PLoS ONE* 7, e29762
43 Ali, K. *et al.* (2011) Monitoring biochemical changes during grape berry development in Portuguese cultivars by NMR spectroscopy. *Food Chem.* 124, 1760–1769
44 Fernandez, L. *et al.* (2010) Transposon-induced gene activation as a mechanism generating cluster shape somatic variation in grapevine. *Plant J.* 61, 545–557
45 Toffali, K. *et al.* (2011) Novel aspects of grape berry ripening and post-harvest withering revealed by untargeted LC-ESI-MS metabolomics analysis. *Metabolomics* 7, 424–436
46 Mica, E. *et al.* (2009) High throughput approaches reveal splicing of primary microRNA transcripts and tissue specific expression of mature microRNAs in *Vitis vinifera*. *BMC Genomics* 10, 558
47 Lijavetzky, D. *et al.* (2012) Berry flesh and skin ripening features in *Vitis vinifera* as assessed by transcriptional profiling. *PLoS ONE* 7, e39547
48 Guillaumie, S. *et al.* (2011) Transcriptional analysis of late ripening stages of grapevine berry. *BMC Plant Biol.* 11, 165
49 Fortes, A.M. *et al.* (2011) Transcript and metabolite analysis in *Trincadeira cultivar* reveals novel information regarding the dynamics of grape ripening. *BMC Plant Biol.* 11, 149
50 Fasoli, M. *et al.* (2012) The grapevine expression alas reveals a deep transcriptome shift driving the entire plant into a maturation program. *Plant Cell* 24, 3489–3505
51 Milli, A. *et al.* (2012) Proteomic analysis of the compatible interaction between *Vitis vinifera* and *Plasmopara viticola*. *J. Proteomics* 75, 1284–1302
52 Polesani, M. *et al.* (2010) General and species-specific transcriptional responses to downy mildew infection in a susceptible (*Vitis vinifera*) and a resistant (*V. riparia*) grapevine species. *BMC Genomics* 11, 117
53 Grimplet, J. *et al.* (2009) Proteomic and selected metabolite analysis of grape berry tissues under well-watered and water-deficit stress conditions. *Proteomics* 9, 2503–2528
54 Sreekantan, L. *et al.* (2010) Differential floral development and gene expression in grapevines during long and short photoperiods suggests a role for floral genes in dormancy transitioning. *Plant Mol. Biol.* 73, 191–205
55 Tillett, R.L. *et al.* (2011) Identification of tissue-specific, abiotic stress-responsive gene expression patterns in wine grape (*Vitis vinifera* L.) based on curation and mining of large-scale EST data sets. *BMC Plant Biol.* 11, 86
56 Grimplet, J. *et al.* (2009) VitisNet: 'Omics' integration through grapevine molecular networks. *PLoS ONE* 4, e8365
57 Zamboni, A. *et al.* (2010) Identification of putative stage-specific grapevine berry biomarkers and omics data integration into networks. *Plant Physiol.* 154, 1439–1459
58 Richards, S.E. *et al.* (2010) Intra- and inter-omic fusion of metabolic profiling data in a systems biology framework. *Chemometr. Intell. Lab. Syst.* 104, 121–131
59 Sawada, Y. *et al.* (2012) RIKEN tandem mass spectral database (ReSpect) for phytochemicals: a plant-specific MS/MS-based data resource and database. *Phytochemistry* 82, 38–45
60 DeRisi, J.L. *et al.* (1997) Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278, 680–686
61 Nagalakshmi, U. *et al.* (2008) The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* 320, 1344–1349
62 Rossignol, T. *et al.* (2006) Analysis of the genomic response of a wine yeast to rehydration and inoculation. *Appl. Microbiol. Biotechnol.* 71, 699–712
63 Winter, G. *et al.* (2011) Effects of rehydration nutrients on H2S metabolism and formation of volatile sulfur compounds by the wine yeast VL3. *AMB Express* 1, 36
64 Marks, V.D. *et al.* (2008) Dynamics of the yeast transcriptome during wine fermentation reveals a novel fermentation stress response. *FEMS Yeast Res.* 8, 35–52
65 Rossouw, D. *et al.* (2012) The impact of co-inoculation with *Oenococcus oeni* on the trancriptome of *Saccharomyces cerevisiae* and on the flavour-active metabolite profiles during fermentation in synthetic must. *Food Microbiol.* 29, 121–131
66 Rossouw, D. *et al.* (2012) Transcriptional regulation and the diversification of metabolism in wine yeast strains. *Genetics* 190, 251–261
67 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
68 Celton, M. *et al.* (2012) A comparative transcriptomic, fluxomic and metabolomic analysis of the response of *Saccharomyces cerevisiae* to increases in NADPH oxidation. *BMC Genomics* 13, 317
69 Fischer, B.M. *et al.* (2004) Quantitative trait locus analysis of fungal disease resistance factors on a molecular map of grapevine. *Theor. Appl. Genet.* 108, 501–515
70 Walker, A.R. *et al.* (2007) White grapes arose through the mutation of two similar and adjacent regulatory genes. *Plant J.* 49, 772–785
71 Battilana, J. *et al.* (2011) Functional effect of grapevine 1-deoxy-D-xylulose 5-phosphate synthase substitution K284N on Muscat flavour formation. *J. Exp. Bot.* 62, 5497–5508
72 Boss, P.K. and Thomas, M.R. (2002) Association of dwarfism and floral induction with a grape 'green revolution' mutation. *Nature* 416, 847–850
73 Riaz, S. *et al.* (2006) Refined mapping of the Pierce's disease resistance locus, PdR1, and Sex on an extended genetic map of *Vitis rupestris* x *V. arizonica*. *Theor. Appl. Genet.* 113, 1317–1329
74 Hwang, C-F. *et al.* (2010) Cloning and characterization of XiR1, a locus responsible for dagger nematode resistance in grape. *Theor. Appl. Genet.* 121, 789–799
75 Barker, C.L. *et al.* (2005) Genetic and physical mapping of the grapevine powdery mildew resistance gene, *Run1*, using a bacterial artificial chromosome library. *Theor. Appl. Genet.* 111, 370–377
76 Laucou, V. *et al.* (2011) High throughput analysis of grape genetic diversity as a tool for germplasm collection management. *Theor. Appl. Genet.* 122, 1233–1245
77 Cipriani, G. *et al.* (2010) The SSR-based molecular profile of 1005 grapevine (*Vitis vinifera L.*) accessions uncovers new synonymy and parentages, and reveals a large admixture amongst varieties of different geographic origin. *Theor. Appl. Genet.* 121, 1569–1585
78 Myles, S. *et al.* (2011) Genetic structure and domestication history of the grape. *Proc. Natl. Acad. Sci. U.S.A.* 108, 3530–3535
79 Myles, S. *et al.* (2010) Rapid genomic characterization of the genus *Vitis*. *PLoS ONE* 5, e8219
80 Elshire, R.J. *et al.* (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* 6, e19379
81 Riedelsheimer, C. *et al.* (2012) Genome-wide association mapping of leaf metabolic profiles for dissecting complex traits in maize. *Proc. Natl. Acad. Sci. U.S.A.* 109, 8872–8877
82 Winzeler, E.A. *et al.* (1998) Direct allelic variation scanning of the yeast genome. *Science* 281, 1194–1197
83 Ehrenreich, I.M. *et al.* (2010) Dissection of genetically complex traits with extremely large pools of yeast segregants. *Nature* 464, 1039–1042
84 Katou, T. *et al.* (2009) QTL mapping of sake brewing characteristics of yeast. *J. Biosci. Bioeng.* 107, 383–393
85 Marullo, P. *et al.* (2007) Single QTL mapping and nucleotide-level resolution of a physiologic trait in wine *Saccharomyces cerevisiae* strains. *FEMS Yeast Res.* 7, 941–952
86 Ambroset, C. *et al.* (2011) Deciphering the molecular basis of wine yeast fermentation traits using a combined genetic and genomic approach. *G3 (Bethesda)* 1, 263–281
87 Kutyna, D.R. *et al.* (2010) Microbiological approaches to lowering ethanol concentration in wine. *Trends Food Sci. Technol.* 21, 293–302
88 Varela, C. *et al.* (2008) Taking control of alcohol. *Aust. N. Z. Wine Ind. J.* 23, 41–43
89 Kutyna, D.R. *et al.* (2012) Adaptive evolution of *Saccharomyces cerevisiae* to generate strains with enhanced glycerol production. *Appl. Microbiol. Biotechnol.* 93, 1175–1184
90 Gadoury, D.M. *et al.* (2012) Grapevine powdery mildew (*Erysiphe necator*): a fascinating system for the study of the biology, ecology and epidemiology of an obligate biotroph. *Mol. Plant Pathol.* 13, 1–16

91 Chaïb, J. *et al.* (2010) The grape microvine – a model system for rapid forward and reverse genetics of grapevines. *Plant J.* 62, 1083–1092
92 Bouquet, A. (1981) Recurrent selection in grapevine breeding. *Agronomie* 1, 65–73
93 Di Gaspero, G. and Cattonaro, F. (2010) Application of genomics to grapevine improvement. *Aust. J. Grape Wine Res.* 16, 122–130
94 Dandekar, A.M. *et al.* (2012) An engineered innate immune defense protects grapevines from Pierce disease. *Proc. Natl. Acad. Sci. U.S.A.* 109, 3721–3725
95 Casagrande, K. *et al.* (2011) Defence responses in Rpv3-dependent resistance to grapevine downy mildew. *Planta* 234, 1097–1109
96 Di Gaspero, G. *et al.* (2012) Selective sweep at the Rpv3 locus during grapevine breeding for downy mildew resistance. *Theor. Appl. Genet.* 124, 277–286
97 Bellin, D. *et al.* (2009) Resistance to *Plasmopara viticola* in grapevine 'Bianca' is controlled by a major dominant gene causing localised necrosis at the infection site. *Theor. Appl. Genet.* 120, 163–176
98 Merdinoglu, D. *et al.* (2003) Genetic analysis of downy mildew resistance derived from *Muscadinia rotundifolia*. *Acta Horticult.* 1, 451–456
99 Schwander, F. *et al.* (2012) Rpv10: a new locus from the Asian *Vitis* gene pool for pyramiding downy mildew resistance loci in grapevine. *Theor. Appl. Genet.* 124, 163–176
100 Blasi, P. *et al.* (2011) Construction of a reference linkage map of *Vitis amurensis* and genetic mapping of Rpv8, a locus conferring resistance to grapevine downy mildew. *Theor. Appl. Genet.* 123, 43–53
101 Peressotti, E. *et al.* (2010) Breakdown of resistance to grapevine downy mildew upon limited deployment of a resistant variety. *BMC Plant Biol.* 10, 147