Australian Government
Wine Australia

CSIRO

Grapevine Germplasm Variety Identification



FINAL REPORT to WINE AUSTRALIA

Project Number: CSP1501

Principal Investigators: Mark Thomas, Mahmood Hassan, Ian Dry & Rob Walker

Research Organisation: CSIRO Agriculture & Food

Date: 9 June 2021

Title page: Grapevine germplasm variety identification

Authors:

Rob Walker, Mahmood Hassan, Ian Dry and Mark Thomas

CSIRO Agriculture and Food

Waite Campus

Locked Bag 2, Glen Osmond, SA 5064

Copyright:

© 2021 CSIRO. To the extent permitted by law, all rights are reserved and no part of this report covered by copyright may be reproduced or copied in any form or by any means except with the written permission of CSIRO.

Important Disclaimers:

CSIRO advises that the information contained in this report comprises general statements based on scientific research. The reader is advised and needs to be aware that such information may be incomplete or unable to be used in any specific situation. No reliance or actions must therefore be made on that information without seeking prior expert professional, scientific and technical advice. To the extent permitted by law, CSIRO (including its employees, affiliates and consultants) exclude all liability to any person for any consequences, including but not limited to all losses, damages, costs, expenses and any other compensation, arising directly or indirectly from using this report (in part or in whole) and any information or material contained in it.

Further, this study was made possible through a collaboration with the Spanish organisation Consejo Superior De Invetigaciones Cientificas (CSIC) through Instituto de Ciencias de la Vid y del Vino (ICVV), abbreviated ICVV-CSIC. Specifically, the work involved use of the ICVV-SNP database. Results based on access to the ICVV-SNP database are based on best available knowledge, and as such, the database may contain errors. ICVV-CSIC and CSIRO decline any responsibility for decisions or interpretations arising which could have been derived unknowingly from any inherent errors in the database.

Front cover: Photo of grapevine variety Rondinella

Contents

1 Abstract 1

2 Executive summary 2

3 Background 5

4 Project Aims and Performance targets 8

5 Methods 8

5.1 Plant Material 8

5.2 DNA extraction 8

5.3 SNP genotyping by Sequenonom iPLEX Chemistry Assay 8

5.4 Assigning variety names based on SNP analysis 9

5.5 SSR analysis 10

6 Results 10

6.1 Explanantion of naming outcomes from SNP analysis 10

6.2 Stage 1 - SNP analysis of Australian grapevine germplasm colletions 11

6.2.1 Genotypes assigned variety names 11

6.2.2 Genotypes assigned variety Prime names 12

6.2.2.1 General observations relating to genotypes assigned Outcome 1 variety Prime names 15

6.2.2.2 General observations relating to genotypes assigned Outcome 3 variety Prime names 16

6.2.3 Genotypes assigned Outcome 2/3 variety Provisional names 17

6.2.3.1 General observations relating to genotypes assigned Provisional names 19

6.2.4 Genotypes unable to be assigned a varietal name (Outcomes 4, 5 & 6) 19

6.3 Stage 2 - Verification of selected accessions to form the basis of an Australian national germplasm collection 20

6.3.1 Verification of assigned variety Prime names (Outcome 1) of selected accessions by repeated SNP analysis 20

6.3.2 Verification of assigned variety Prime and Provisional names (Outcomes 2 & 3) of selected accessions by SSR analysis 21

6.3.3 Verification of rootstock genotypes by SSR analysis 23

6.3.4 Mis-named accessions identified in the Australian germplasm collections 23

6.4 SNP analysis on DNA extracted from wood/bud samples 24

6.5 Previously unnamed varieties in historical collections assigned Prime names 26

7 Discussion……27
8 Outcome/Conclusions……30
9 Recommendations……31
Appendix 1: Communication……32
Appendix 2: Intellectual Property……33
Appendix 3: References……34
Appendix 4: List of Supplementary files……36
Appendix 5: Staff……37
Appendix 6: Acknowledgements……38

1 Abstract

The aim of this study was to validate the name and DNA profile of all accessions in the CSIRO and SARDI grapevine collections as well as unique accessions in four other Australian grapevine collections. Genotyping was carried out using a combination of DNA typing techniques over two stages and the results compared to databases of major European collections. In total, 1514 accessions were DNA typed, successfully confirming the identity of 363 unique wine, table/dried and multi-purpose grapevine varieties and 875 accessions currently within Australian germplasm collections. Some cases of misnaming were identified. The study also allowed variety Prime names to be assigned to valuable Australian historical material. The validated varieties form the nucleus for the establishment of the National Grapevine Collection and pave the way for the collections to be reopened to industry. The work also provides the foundation of a public register of planting material in Australia and has led to the creation of a database that could form the basis of a commercial DNA typing service for grapevine varieties.

2 Executive summary

Grapevine varieties in the major Australian grapevine germplasm collections have traditionally been identified and named by ampelography (visual inspection of the grapevine, particularly leaves). The Albariño/Savagnin blanc misidentification of 2008 highlighted the need to validate the names of varieties held in Australian collections and resulted in the closure of the CSIRO and SARDI collections, which hold up to 80% of the unique material in Australia. The primary objectives of this project were to validate the name and DNA profile of varieties using international references and to develop a single nucleotide polymorphism (SNP) profile database containing the validated varieties that could be used by an Australian service provider to offer an improved grapevine identification service to industry.

Initially, the project planned to work with three major European grapevine collections using a set of 48 SNP markers (Cabezas et al. 2011). Once an agreed list of varieties had been developed, DNA would be sampled from the varieties held in participating Australian collections and from reference vines of the same varieties held in the European collections. The DNA samples from the European collections would then be sent to Australia and along with the DNA samples from the Australian collections, all DNA samples would then be genotyped by the same provider using the same set of 48 SNP markers. However, the logistics and practicality of working with three European collection-holders proved difficult and following discussion with Wine Australia it was decided to work with one major European collection holder, the Spanish organisation, Consejo Superior De Investigaciones Cientificas (CSIC) through Instituto de Ciencias de la Vid y del Vino (ICVV).

In Stage 1 of the project a total of 1514 accessions were genotyped - 1037 from the CSIRO collection, 455 from the SARDI collection and a further 22 from OTHER collections including Adelaide Hills Vine Improvement, New South Wales Department of Primary Industries, West Australian Agriculture Authority and Yalumba Nursery. CSIRO arranged sampling of the accessions. DNA extraction and SNP genotyping was carried out by the Australian Genome Research Facility Limited (AGRF). The DNA profiles were then sent to the Spanish collaborator. The Spanish collaborator compared the SNP profiles of Australian accessions with accessions in their own collection and then nominated six different naming outcomes for each accession that was supplied:

- **Outcome 1**: The SNP pattern of the Australian sample fully matched a genotype in the ICVV-SNP database with a variety Prime name. The variety Prime name was assigned using upper case letters or, where the Prime name is a global name for a cluster of varieties derived from a single original embryo, using lower case letters.

- **Outcome 2**: The SNP pattern of the Australian sample fully matched a genotype in the ICVV-SNP database with a Provisional name. A Provisional name was assigned in ICVV-SNP database when there had been insufficient independent evidence to enable establishment of a variety Prime name for the genotype. Thus, the identification of the Australian accession was considered inconclusive.

- **Outcome 3**: The SNP pattern of the Australian sample almost fully matched a genotype in the ICVV-SNP database, so the name of the closest genotype was assigned. If the closest genotype was a Prime name, that Prime name was assigned. If the closest genotype was a Provisional

name, that Provisional name was assigned. In both cases the identification of the Australian accession was considered inconclusive.

- **Outcomes 4-6**: No name could be assigned because either the SNP pattern of the Australian sample did not match with any ICVV-SNP database genotype or matched a genotype in the database which matches multiple varieties or the SNP genotyping quality was too low.

Out of a total of 1514 accessions analysed, 1339 (88%) could be assigned a VIVC variety name (i.e. Outcomes 1-3) while 175 (12%) could not be assigned a variety name (Outcomes 4-6). Of the 1339 accessions that were assigned a name, 751 (56%) were assigned an Outcome 1 variety Prime name, 204 (15%) were assigned an Outcome 2 Provisional name and 384 (29%) were assigned an Outcome 3 Prime or Provisional name. It should be noted that the set of 48 SNPs that were used in this project was unable to distinguish clones and had insufficient resolution to reliably distinguish non-vinifera *Vitis* species and varieties with non-vinifera *Vitis* species in their pedigree. As most, if not all rootstocks used commercially in Australia, have American non-vinifera *Vitis* species in their background, it was not possible to assign names to these rootstocks using this SNP panel.

In Stage 2, a decision was made, following discussions with Wine Australia, to confirm the identity of a selection of the most important wine, table/dried and multi-purpose varieties from the germplasm collections that had been assigned Prime or Provisional names, which could form the basis of an Australian national germplasm collection. Leaf samples of these selected accessions were re-collected from the field in October 2020.

For accessions that had been assigned an Outcome 1 variety Prime name, SNP analysis was repeated, and the results analysed in-house using the SNP profile database that had been developed as part of Stage 1. In total, 583 accessions that had been assigned a variety Prime name from Stage 1, were re-sampled and re-analysed using this methodology and all were found to match the SNP profile obtained previously in Stage 1. This demonstrates the excellent reproducibility of the SNP DNA typing method.

For accessions that had been assigned Outcome 2 Provisional or Outcome 3 Prime or Provisional variety names, an alternative method of DNA typing based on simple sequence repeat (SSR) analysis was employed in an attempt to validate the variety names assigned by SNP analysis in Stage 1. Because of the much greater expense involved in SSR analysis, compared to SNP analysis, it was not feasible to analyse all 588 Outcome 2 and Outcome 3 accessions assigned a variety name by SSR analysis in Stage 1. Therefore, where there were multiple accessions with identical SNP patterns, a representative genotype was selected for SSR analysis. A total of 185 DNA samples of scion accessions were sent to the Institut Francais de la Vigne et du Vin (IFV) in France where they were subjected to SSR analysis. We also took the opportunity to send samples of the 20 most commonly used rootstocks for SSR analysis due to the issues observed with trying to genotype rootstock accessions using our 48 SNP panel.

From the 185 scion DNA samples analysed using SSRs, 150 (81%) were found to match samples in the IFV SSR database and were assigned a variety name. The remaining 35 samples did not match any SSR profiles in the IFV database and were classified as “unknowns”. Of the 150 that were assigned a variety name, 137 were found to match the Prime or Provisional name that had previously been assigned using SNP analysis in Stage 1. However, 12 of the 185 samples analysed by SSR analysis were assigned variety names that did not match the Outcome 2 or Outcome 3 Prime or Provisional variety names assigned using SNP analysis in Stage 1. In contrast to the results obtained from SNP analysis, SSR analysis confirmed the identity of the 19 of the 20 rootstocks analysed.

In a number of cases, the variety Prime name assigned by dual SNP analysis or combined SNP/SSR analysis differed from the name given to this accession in the Australian germplasm collection. These ranged from minor spelling differences to, in some cases, name differences. For many varieties, the name of the variety in the Australian collection was a recognized VIVC synonym of the assigned Prime name for that variety e.g., Shiraz (Australian collection) vs Syrah (VIVC) or Traminer (Australian collection) vs Savagnin (VIVC). Investigations showed that in the majority of these instances where cuttings had been distributed to industry, this was either under the correct variety name or the varieties are not being used for wine production in Australia.

The single exception was the variety labelled as Petit Manseng in the CSIRO germplasm collection which was confirmed in 2019 to be MANSENG GROS BLANC (syn. Gros Manseng) and records indicated that this variety had been distributed from the CSIRO collection under the incorrect name. Furthermore, the National Vintage Survey reports that on average 44 tonnes per year of grapes labelled as Petit Manseng were crushed between 2015-19. As this has implications for the varietal labelling of the resultant wine, affected producers were notified by Wine Australia at the time through its regulatory function.

Discussions were also initiated with AGRF to investigate the possibility of them setting up a commercial service for grapevine DNA typing based on the SNP data collected from this project. To facilitate this process, we also carried out a pilot study looking at the feasibility of DNA extraction from buds and wood from canes, rather than leaf material, because leaf material may not always be available when a grape grower wishes to undertake DNA typing of germplasm in his vineyard. The results of the pilot study on 14 grapevine genotypes indicated that, although the failure rate for correctly reading individual SNPs on the panel was slightly higher with DNA extracted from buds and wood, than leaves, it did not significantly impact on varietal identification of the genotypes.

Finally, of special interest is the fact that this study has been able to assign variety Prime names to 11 accessions obtained from historical Australian collections and held in the CSIRO and SARDI collections for many years. Some of these accessions pre-date the phylloxera crisis in Europe in the late 1800s. These include six accessions from the Best's (Best Western, Victoria) collection, four accessions from the Coldstream (Victoria) collection and one accession from the Auldana (South Australia) collection.

Overall, the duplicate SNP/SNP analysis and the combined SNP/SSR analysis allowed assignment of a VIVC Prime/global Prime name to 362 unique genotypes and 875 accessions currently within Australian germplasm collections. These varieties form the nucleus for the establishment of the National Grapevine Collection (NGC) and a public register of planting material in Australia, and pave the way for the reopening of the collections to industry.

3 Background

Until the 1990s, ampelography was the only method available for identifying and naming grapevine varieties. Regular visits to Australia by leading international ampelographers and their inspections of accessions in the major grapevine germplasm collections, especially during the decades 1960s to 1990s, provided a level of confidence in the names of varieties that have underpinned the Australian wine industry expansion through the 1990s and consolidation over the last two decades.

Ground-breaking research by Thomas et al. (1994) led to the development of DNA profiling methods for grapevines. The initial methods were based on microsatellites or simple sequence repeats (SSR). The first DNA profiling service offered to the Australian industry was based on profiling of a cross-section of varieties in the CSIRO collection using SSR markers. The service was based on the identities of varieties held in the CSIRO collection at that time. During a visit to Australia in 2008, a visiting French ampelographer raised questions about the identity of what was believed to be a Spanish white wine variety. Subsequent checks including DNA profiling through an international service revealed that the variety was a French white wine variety. This highlighted the need to have varieties in Australian collections checked by DNA profiling against reference vines of varieties held in international collections. This project therefore aimed to develop a collaborative international effort to DNA profile varieties held in Australian germplasm collections to confirm identity.

DNA profiling has become an accepted method to identify grapevine varieties. Traditional ampelographic phenotyping methods nonetheless retain an important role, demonstrated by the Australian Plant Breeder's Rights (PBR) Office's reliance on traditional plant taxonomic methods for grapevine and other plant PBR. DNA profiling can identify naming errors and instances of synonyms (different names for the same variety) and homonyms (different varieties identified under the same name). The microsatellite (SSR) DNA marker type originally developed and described by Thomas et al. (1994) has been adopted world-wide and used successfully for many years in other countries to characterise variety collections. Advances in DNA technology have seen a new DNA marker type known as a SNP (single nucleotide polymorphism) emerge as a more suitable DNA marker for grapevine identification, as it has advantages over microsatellite markers (Cabezas et al. 2011). Rapid advances in SNP detection technology have reduced the SNP assay cost and the ease of transferability and comparison of DNA profiles between laboratories. In contrast, the SSR assay is based on DNA length differences, not sequence differences, and the technology used for SSR detection is basically unchanged from that used over 20 years ago.

Comparison of DNA profiles from different national grapevine collections has been shown to be a robust method to confirm grapevine identity. The Vitis International Variety Catalogue (VIVC), developed by the Julius Kuhn-Institut in Germany (Maul et al. 2021), has been a major advance assisting the standardisation of grapevine variety names. The Catalogue assigns a Prime name to each variety as well as known synonyms. The VIVC has linked over 1500 DNA profiles to these Prime names, providing a useful resource for comparing cultivar genetic differences. The concept of Prime names goes back to at least 1986, when records show that the Alcohol and Tobacco Tax and Trade Bureau unit within the United States Department of the Treasury defined a Prime name as the primary name approved for use on a wine label in the United States (Anon 2020). The Vitis International Variety Catalogue (Maul et al. 2021) defines Prime name as the name of the variety in its (probable) country of origin. Robinson et al. (2012), in their compilation of 1,368 grapevine varieties, apply Prime name to mean the country or region of origin of a grapevine variety.

Australia currently does not have an industry recognised national collection of grapevine varieties and clones. The two major collections that served as a resource for the Australian grape industries since the 1960s, the CSIRO and SARDI collections, were closed around 2009. Walker and Thomas (2009) addressed the future of grapevine germplasm collections in Australia and recommended a staged approach towards forming an Australian National Grapevine Collection including a DNA profiling study of available germplasm resources. The Australian Grape and Wine Authority, AGWA (now Wine Australia) took a lead role in attempting to plan a way forward for vine improvement in Australia and commissioned a number of projects over the past decade. These are summarised as follows:

- 'Grapevine Type Reference Pilot Study' (CSP 1102)
- 'Review of Grapevine Collections in Australia' (GWR 1112)
- 'A case for strategic national oversight of Australia's grapevine foundation assets – underpinning the competitiveness of the Australian wine sector' (WGG 1401)
- 'Australian Vine Quality Standards Scheme' (GWR 1002)
- 'The development of a national, grape-industry strategy for the implementation of a national standard for grapevine certification – Phase 1' (WAC 2101)
- 'National Germplasm Discussion Paper' (WAC 2009)
- 'National Grapevine Collection Coordinator' (FOV 2201)

The 'Grapevine Type Reference Pilot Study' trialled SNP marker technology and recommended that grapevine germplasm collections in Australia be DNA profiled to validate the identity of the varieties held in the collections. It was proposed that the validation process involve comparison of DNA profiles for accessions in Australian collections with DNA profiles of accessions in other international grapevine collections. The recommendation was accepted by Wine Australia which subsequently commissioned this project to address that need.

The study titled 'Review of Grapevine Collections in Australia' detailed the current state of both research and industry collections. It was found that the two main collections, those held by CSIRO and SARDI, hold up to 80% of unique material (McMichael 2013). The study noted that DNA profiling the major Australian collections would benefit both scientists and industry and should lead towards a pathway to reopen the collections to industry. Scientists would also benefit from the knowledge obtained in projects investigating grapevine improvement (rootstocks and scions), biosecurity, genetic diversity and adaptation to climate change.

The study, 'A case for strategic national oversight of Australia's grapevine foundation assets – underpinning the competitiveness of the Australian wine sector', identified a need for national oversight of Australia's grapevine foundation assets and industry access to the two largest collections in Australia maintained by CSIRO and SARDI.

The project 'Australian Vine Quality Standards Scheme – Vine Assurance' provided industry with a voluntary Australian Standard for Grapevine Propagation Material. An important component of the scheme is the use of DNA profiles to accurately identify grapevine varieties and to trace a variety from vine importation to nursery to vineyard to the winery for label integrity. In Australia, compliance with the Label Integrity Program is required for wine labelling and if a variety name is on the label an audit trail must be kept. The Australian Standard for Grapevine Propagation Material has not been adopted

widely by the sector but foreshadowed the need for a nuclear DNA certified collection, termed a G_0 collection, defined in the standard (Anon 2013) as "reference vine/s that have been authenticated by DNA testing".

The project 'The development of a national, grape-industry strategy for the implementation of a national standard for grapevine certification – Phase 1' aimed to lay the groundwork for a single, science-based quality assurance system to be adopted across the entire grapevine supply chain. It engaged all stakeholders to define the outcomes, develop a common language, define the quality assurance objectives and set out a roadmap for implementation in Phase 2.

The 'National Germplasm Discussion Paper' brought together previous work in the area (as above), reviewed the current state of play and proposed a model for an Australian National Grapevine Collection (NGC). It consists of three tiers of collections: a Vine Integrity Collection and two layers of DNA-tested, industry-accessible field collections which includes the CSIRO and SARDI accessions.

An implementation plan was developed to adopt the NGC model in a staged approach over three years. A 'National Grapevine Collection Coordinator' was appointed in January 2022 to manage the collection and its pillars, including a centrally maintained and publicly available national database of grapevine varieties and clones.

An updated DNA-based method and service for variety identification, based on a collection of varieties that have been DNA certified against reference varieties in a major international collection, would considerably reduce the risk to members of the Australian industry and to Australia's international reputation by minimising the chance of accidental mislabelling events. It would enable compilation of the public register of grapevine varieties which will underpin the National Grapevine Collection. An objective DNA-based method for variety identification would also assist those in the propagation sector who will adopt a national standard for grapevine certification as they could label their cuttings as DNA-certified.

This project aimed to address these needs by engaging and working collaboratively with the Spanish organisation, Consejo Superior De Investigaciones Cientificas (CSIC) through Instituto de Ciencias de la Vid y del Vino (ICVV) using a set of 48 SNPs (Cabezas et al. 2011) and the Institut Francais de la Vigne et du Vin (IFV) in France to access their SSR typing capability to genotype the accessions held by the major Australian germplasm collections.

4 Project Aims and Performance targets

The project aim was to DNA profile the varieties in the participating Australian collections to objectively identify each variety.

The specific objectives were to:

1. Confirm the variety name and DNA profile of Australian accessions through international collaborations.
2. Develop a database containing the validated varieties that can be used by an Australian service provider to offer a grapevine DNA identification service to industry.

5 Methods

5.1 Plant Material

Leaf samples were collected from growing shoot tips at the CSIRO Research Farm, Irymple, Victoria. Samples for Stage 1 SNP analysis were collected in November 2016, while samples for Stage 2 SNP and SSR analysis were collected in October 2020. Samples were also collected from the grapevine variety blocks at the South Australian Research and Development Institute (SARDI) Research Farm at Nuriootpa, South Australia. Plant material from the CSIRO and SARDI collections comprised 98.5% of all accessions sampled and genotyped. Plant material from accessions that were unique to three other collections, namely Adelaide Hills Vine Improvement Inc. (1 accession), New South Wales Department of Primary Industries (6 accessions), West Australian Agriculture Authority (6 accessions) and Yalumba Nursery (9 accessions) were also collected. Access to accessions from the non-CSIRO collections was enabled by Material Transfer Agreements between CSIRO and the respective collection owners.

Canes were collected in late March 2021 from the SARDI Nuriootpa vineyard at which point they were brown and lignified. Samples comprising 3-4 sections approximately 2mm in length were cut from each cane using sharp secateurs and placed into 2ml Eppendorf tubes in preparation for DNA extraction. Also 4-5 axillary buds were excised using a sharp scalpel and placed in 2 ml Eppendorf tubes. The secateurs and scalpel were wiped with 70% ethanol in between cutting canes of different varieties.

5.2 DNA extraction

DNA extraction of leaves and wood was carried out by the Australian Genome Research Facility Limited (AGRF).

5.3 SNP genotyping by Sequenonom iPLEX Chemistry Assay

A set of 48 SNP markers that were identified to have high discrimination power and are uniformly distributed across the grapevine (*Vitis vinifera* L.) genome (2-3 markers / chromosome) were proposed as a standard set for grapevine genotyping by Cabezas et al. (2011). The SNPs were included in Sequenom iPlex assay for genotyping carried out by the Australian Genome Research Facility Limited (AGRF).

All DNA samples (10 ng/µL) were transferred into 384 well PCR plates for genotyping. The genotyping analysis was performed as recommended by the manufacturer with reagents included in the iPLEX Gold SNP genotyping kit (Agena) and the software and equipment provided with the MassARRAY platform (Agena). Samples were amplified from a 5 µL final PCR volume composed of 1xPCR buffer, 2 mM $MgCl_2$, 500 µM deoxynucleotide triphosphates (dNTPs), 0.1 µM each PCR primer, 0.5 U of HotStar Taq enzyme and 1 µL DNA. The thermal cycling conditions consisted of a first denaturation step at 95°C for 2 min, followed by 45 cycles of denaturation at 95°C for 30 s, annealing at 56°C for 30 s and extension at 72°C for 1 min, with a final extension step at 72°C for 5 min. To neutralise unincorporated dNTPs, PCR products were treated with 0.5 U shrimp alkaline phosphatase by incubation at 37°C for 40 min, followed by enzyme inactivation by heating at 85°C for 5 min. The extension reaction was carried out in a final volume of 9 µL by adding 2 µL of an iPLEX Gold extension reaction cocktail to the purified PCR products, containing 0.222xiPLEX buffer, 1xiPLEX termination mix, 1xiPLEX enzyme and the SBE primer mix of extension primers. The iPLEX extension reaction was performed under the following thermal conditions: an initial denaturation step at 94°C for 30 s, followed by 40 cycles of a denaturation step at 94°C for 5 s, 5 cycles of annealing at 52°C for 5 s and extension at 80°C for 5 s and a final extension step at 72°C for 3 min. After desalting of the products by using SpectroCLEAN resins following the manufacturer's protocol, cleaned extension products were dispensed onto a 384 SpectroCHIP array using an RS1000 Nanodispenser, and finally, the array was introduced into a MassARRAY Compact mass spectrometer. Spectra were acquired using SpectroAcquire software, and data analysis, including automated allele calling, was done using MassARRAY Typer software, version 4.0.26.75.

5.4 Assigning variety names based on SNP analysis

SNP genotyping data was provided to Dr Javier Ibáñez, Instituto de Ciencias de la Vid y del Vino, La Rioja, Spain to compare to the DNA profiles on the SNP database maintained at the Spanish Instituto de Ciencias de la Vid y del Vino (ICVV). The Spanish grapevine SNP database was generated by SNP genotyping of their own grapevine germplasm collection, together with analysis of grapevine samples obtained from other grapevine germplasm collections throughout the world (including CSIRO), with the same set of 48 SNP markers.

The classification system used for genotype naming was designed by the ICVV in consultation with CSIRO. It involved six possible naming outcomes:

- **Outcome 1**: The SNP pattern of the Australian sample fully matched a genotype in the ICVV-SNP database with a variety Prime name. The variety Prime name was assigned using upper case letters or, where the Prime name is a global name for a cluster of varieties derived from a single original embryo, using lower case letters.

- **Outcome 2**: The SNP pattern of the Australian sample fully matched a genotype in the ICVV-SNP database with a Provisional name. A Provisional name was assigned in ICVV-SNP database when there had been insufficient independent evidence to enable establishment of a variety Prime name for the genotype. Thus, the identification of the Australian accession was considered inconclusive.

- **Outcome 3**: The SNP pattern of the Australian sample almost fully matched a genotype in the ICVV-SNP database so the name of the closest genotype was assigned. If the closest genotype was a Prime name, that Prime name was assigned. If the closest genotype was a Provisional

name, that Provisional name was assigned. In both cases the identification the Australian accession was considered inconclusive.

- **Outcome 4.** The SNP pattern of the Australian sample does not match with any ICVV-SNP database genotype. No name was assigned.
- **Outcome 5**. The SNP pattern of the Australian sample is identical or very close (1-2 mismatches) to genotypes existing in the database that cannot be used for identification because more than one variety matches to it. No name was assigned.
- **Outcome 6.** The SNP pattern of the Australian sample was of insufficient quality i.e. the number of SNPs obtained from the analysis was too low and needs to be repeated.

5.5 SSR analysis

Approximately 100 mg fresh leaf from the tip of growing shoots were collected on ice from CSIRO Irymple and SARDI Nuriootpa vineyards in late October 2020. The samples were then frozen in dry ice or liquid nitrogen. The frozen samples were submitted to AGRF (Waite Institute) for DNA extraction. Extracted DNA was normalized to a concentration of 20ng μl^{-1} and aliquoted into Axygen® 96-well Polypropylene PCR plates. The plates were heat sealed and sent to Institut Francais de la Vigne et du Vin (IFV) in France at ambient temperature by air freight where the SSR analysis was undertaken.

6 Results

6.1. Explanation of naming outcomes from SNP analysis

The ICVV-SNP database may contain errors. Results provided by the Spanish collaborator are provided to the best of its knowledge. Use of the results and decisions on classification of genotypes by CSIRO are also provided to the best of its knowledge and cannot be guaranteed to be free from error. ICVV-CSIC and CSIRO decline any responsibility derived from any such errors.

Full genotype matchings may occur between different varieties originated by somatic mutations. Since information about phenotypes is generally not available, it is not possible to distinguish between somatic variants. Hence Prime names (and VIVC number in the research database) refer to one variety of the group of somatic variants.

Prime names in upper case letters e.g., 'DOLCETTO' correspond to Prime names as occur in VIVC.

Prime names in lower case letters indicate a global name for a cluster of varieties derived from a single original embryo. For example, the global Prime name 'Pinot' denotes a genotype representing Pinot Noir, Pinot Blanc, Pinot Meunier, Pinot Gris etc.

A Prime name is usually assigned to a genotype in the ICVV-SNP database when there are at least two independent sources of evidence to support a variety identification. No absolute warranty is given by ICVV-CSIC that such evidence is correct nor independent and hence there is no absolute warranty assigned of any variety name.

When a genotype cannot be assigned a Prime name, normally a Provisional name is assigned (and presented in italics in Tables). Provisional names cannot be taken as the correct name for any genotype.

The set of 48 SNPs used in this study was designed from *Vitis vinifera* sequences and has low resolution for non-vinifera species. Consequently, different non-vinifera genotypes may have similar or fully identical genotypic profiles for this set of SNPs which makes it impossible to use for identification in these cases. That said, in some cases (mostly inter-specific hybrids with *Vitis vinifera*), a 48 SNP genotype was specific for a certain non-vinifera variety and thus it could be used for identification. It cannot be discounted that a similar variety not studied yet may have the same or a very close genotype.

6.2 Stage 1 - SNP analysis of Australian grapevine germplasm collections

A total of 1514 accessions were genotyped by AGRF using the 48 SNP panel. Of the total of 1514 accessions, 1037 (68.5%), 455 (31%) and 22 (1.5%) were from the CSIRO, SARDI and OTHER collections, respectively (Table 1). Variety names (Prime or Provisional) were assigned to 905 of 1037 (87.3%), 414 of 455 (91%) and 20 of 22 (90.9%) accessions from the respective collections. In total, across all collections, 1339 accessions (88.4%) were assigned names, with the remainder, 175 accessions (11.6%), unable to be assigned names.

Table 1. Total number of accessions genotyped from the various collections and number in each collection that were assigned a name (Prime or Provisional) or not assigned a name. OTHER = Adelaide Hills Vine Improvement, New South Wales Department of Primary Industries, West Australian Agriculture Authority and Yalumba Nursery.

Category	**Accessions analysed from the various collections**			
	CSIRO	**SARDI**	**OTHER**	**All**
Names assigned	905	414	20	1339
Names not assigned	132	41	2	175
Total	**1037**	**455**	**22**	**1514**

6.2.1 Genotypes assigned variety names

Accessions genotyped and assigned names were sub-divided into those with Prime names (Outcome 1), Provisional names (outcome 2) and Prime or Provisional names with outcome 3 (Table 2), according to the collections from which they originated. Of the 905 accessions in the CSIRO collection, 492 (54.4%) were assigned Prime name outcome 1, 153 (16.9%) were assigned Prime name outcome 3, 167 (18.5%) were assigned Provisional name outcome 2 and 93 (10.3%) were assigned Provisional name outcome 3. Of the 414 accessions in the SARDI collection, 247 (59.7%) were assigned Prime name outcome 1, 121 (29.2%) were assigned Prime name outcome 3, 34 (8.2%) were assigned Provisional name outcome 2 and 13 (3.1%) were assigned a Provisional name outcome 3. For the OTHER collections, 12 of 20 accessions (60%) were assigned Prime name outcome 1, 3 (15%) were assigned Prime name outcome 3, 4 (20%) were assigned Provisional name outcome 2 and 1 accession was assigned Provisional name outcome 3 (Table 3).

Full lists of accessions from the CSIRO, SARDI and OTHER collections that have been assigned variety Prime or Provisional names with outcomes 1, 2 or 3 are presented Supplementary Tables 1-3.

Table 2. Total number of accessions assigned variety Prime names or Provisional names with outcomes 1, 2 or 3 according to the collections from which they originated. OTHER = Adelaide Hills Vine Improvement, New South Wales Department of Primary Industries, West Australian Agriculture Authority and Yalumba Nursery.

Naming Outcome	**Number of Accessions / Naming outcome / Collection**			
	CSIRO	**SARDI**	**OTHER**	**All**
Prime (Outcome 1)	492	247	12	751
Provisional (Outcome 2)	167	34	4	204
Prime/Provisional (Outcome 3)	246[1]	134[1]	4[1]	384
Total	**905**	**414**	**20**	**1339**

[1]Prime names were 153, 121 and 3 for the CSIRO, SARDI and Other collections, respectively; Provisional names were 93, 13 and 1 for the CSIRO, SARDI and Other collections, respectively.

6.2.2 Genotypes assigned variety Prime names

The previous Tables provide information on accessions genotyped and accessions assigned variety Prime or Provisional names. Names in the ICVV database are formally genotypes, not varieties, because full phenotypic analysis of the genotypes in the ICVV collection is not available.

Where the DNA profile of a sample matches fully with a genotype in the ICVV-SNP database with a variety Prime name, that variety Prime name has been assigned. Variety Prime names in upper case letters usually correspond to Prime names as occur in VIVC. For many such varieties, i.e., that have been assigned variety Prime names, there are one or more accessions. Where there is just one accession of a variety, that accession is regarded as 'the variety' but where there are multiple accessions, the multiple accessions are clones or replicates of the variety.

In addition, there are instances where genotypes have been assigned a variety Prime name but distinguished from the above by having the name presented in lower case letters. This indicates a global name for a cluster of varieties derived from a single original embryo. For example, Pinot is the global Prime name for the cluster of varieties including Pinot Noir, Pinot Gris and Pinot Meunier.

Table 3 shows the total number genotypes (as distinct from accessions shown in Table 2) in each collection that were assigned variety Prime names with one of three outcome numbers (defined in section 5.4).

Table 3. Total number of genotypes assigned a variety Prime name with outcome 1, 3 or mixed 1 & 3 (1/3) according to the collections from which they originated. OTHER = Adelaide Hills Vine Improvement, New South Wales Department of Primary Industries, West Australian Agriculture Authority and Yalumba Nursery.

Naming Outcome	Number of genotypes / Naming outcome (Prime names) / Collection			
	CSIRO	SARDI	OTHER	All
1	221	69	12	302
3	79	38	3	120
1/3	6	0	0	6
Total	**306**	**107**	**15**	**428**

Outcome 1 was the predominant outcome for all collections, comprising 221 of 306 (72.2%), 69 of 107 (64.5%) and 12 of 15 (80%) of assigned variety Prime names in the CSIRO, SARDI and OTHER collections, respectively. Outcome 3 comprised 79 of 306 (25.8%), 38 of 107 (35.5%) and 3 of 15 (20%) of assigned variety Prime names in the respective collections. Almost all genotypes in the CSIRO collection (221 + 79 = 300 of the total 306) had either an outcome 1 or 3. The remaining 6 genotypes, each with multiple accessions, had a mix of outcome 1 and 3. This category is shown in Table 4 under naming outcome 1/3. All 6 genotypes were from the CSIRO collection. They are:

- DANUGE, 3 accessions, 2 having variety Prime name outcome 1 and 1 having variety Prime name outcome 3.
- GAMAY NOIR, 16 accessions, 1 having variety Prime name outcome 1 and 15 having variety Prime name outcome 3. The collection name, Gamay, is a VIVC synonym for GAMAY NOIR. Gamay name is widely used and may or may not be replaced by GAMAY NOIR, although over time there may be an acknowledgement of GAMAY NOIR as the Prime name.
- TINTO CAO, 2 accessions, 1 having variety Prime name outcome 1 and 1 having variety Prime name outcome 3.
- VERDUZZO FRIULANO, 2 accessions, 1 having variety Prime name outcome 1 and 1 having variety Prime name outcome 3.
- Global Prime name Garnacha, 11 accessions, 1 having variety Prime name outcome 1 and 10 having variety Prime name outcome 3.
- Global Prime name Muscat of Alexandria, 26 accessions across the CSIRO and SARDI collections, with 24 having variety Prime name outcome 1 and 2 having variety Prime name outcome 3 (one in the CSIRO collection and one in the SARDI collection).

Reasons for assignment of different outcomes to accessions of the same genotype can be explained using Muscat of Alexandria as an example, where the 2 of 26 accessions that were assigned variety Prime name outcome 3 differed in one allele from the 24 (main genotype) accessions assigned variety Prime name outcome 1. The difference could be due to a mutation or it could be a technical error in genotyping. However, the 2 accessions assigned variety Prime name outcome 3 were included as Muscat of Alexandria because there were sufficient aligning SNPs and no non-matching alleles. 'Sufficient aligning SNPs' for a positive identification in this case means a set of 43 nuclear SNPs plus 5 chloroplast SNPs (which are much less informative) (Javier Ibáñez, personal communication).

As indicated in Table 2, a total of 1028 accessions were assigned variety Prime names (either outcomes 1 or 3), including 645, 368 and 15 accessions from the CSIRO, SARDI and OTHER collections, respectively. Then, as shown in Table 3, the actual number of genotypes assigned a variety Prime name (either outcomes 1 or 3) was 428, including 306, 107, and 15 from the CSIRO, SARDI and OTHER collections, respectively. The difference between the two sets of numbers is the number of clones or replicates of varieties where for each variety the accessions had the same genotype based on the set of 48 SNPs that were used. Instances of 'identical' genotypes designated as different clones may either have an observable phenotypic difference or differences which was/were unable to be differentiated with the set of 48 SNPs. In other cases, the accessions are true-to-type, or genotype replicates, since they are identical accessions of the same genotype propagated vegetatively, e.g. from hardwood cuttings, or, as in the CSIRO collection, varieties that were subjected to a virus elimination procedure, called Fragmented Shoot Apex Culture (FSAC) and propagated through tissue culture. Table 4 shows the number of clones or replicates of each genotype that were assigned a variety Prime name.

Table 4. Clones or replicates of genotypes that were assigned variety Prime names according to the collections from which they originated. OTHER = Adelaide Hills Vine Improvement, New South Wales Department of Primary Industries, West Australian Agriculture Authority and Yalumba Nursery.

Category	**Number per collection and total**			
	CSIRO	**SARDI**	**Other**	**All**
(a) Accessions with variety Prime names	643	368	15	1026
(b) Genotypes with variety Prime names	306	107	15	428
Clones or replicates (a minus b)	**337**	**261**	**0**	**598**

6.2.2.1 General observations relating to genotypes assigned Outcome 1 variety Prime names

General observations relating to genotypes that have been assigned variety Prime names with outcome 1:

- Accessions (total number shown within brackets) within the Australian collections with names Pinot Noir (39), Pinot Blanc (2), Pinot Gris (3), Pinot Precose (1), Pione (1) and Meunier (3) have been assigned the global Prime name Pinot, which, as previously explained, is a name for a cluster of varieties derived from a single original embryo. There are several other examples of global Prime names which are presented in the Tables and Appendices in lower case letters.
- Savagnin (1 accession) and Traminer (9 accessions) have been assigned the global Prime name Savagnin. Both Savagnin and Traminer are VIVC synonyms of Gewuerztraminer. Savagnin has green-white berries and Traminer has pink berries. Both names are likely to persist as common names. While Savagnin is a global Prime name, Savagnin Blanc and Savagnin Rose are variety Prime names in VIVC.
- All accessions in the collections with names involving variations of the name Sultana, Sultanina and Manik Chaman have been assigned the global Prime name Sultanina.
- All 10 accessions with the name Trebbiano have been assigned the variety Prime name TREBBIANO TOSCANO. Trebbiano is a VIVC synonym of TREBBIANO TOSCANO.
- All 5 accessions with the name Tinta Amarella have been assigned the variety Prime name TRINCADEIRA. Tinta Amarella is a VIVC synonym of TRINCADEIRA.
- All 5 accessions with the name Bastardo have been assigned the global Prime name Trousseau. Bastardo is a synonym of the VIVC variety Prime name TROUSSEAU NOIR.
- One accession with the name Grignolino and two accessions with the name Grey Riesling have been assigned the global Prime name Trousseau. Grey Riesling is a synonym of the VIVC Prime name TROUSSEAU GRIS.
- One accession of Rolle was assigned the variety Prime name VERMENTINO. Rolle is a VIVC synonym of VERMENTINO.
- All 12 accessions with the name Verdelho have been assigned the variety Prime name VERDELHO BRANCO. Verdelho is a VIVC synonym for VERDELHO BRANCO.
- Four accessions with the name Cornichon (one accession from the NSW DPI collection) or Purple Cornichon (3 accessions from the CSIRO collection) have been assigned the variety Prime name OLIVETTE NOIRE. Cornichon and Cornichon Violet are synonyms for OLIVETTE NOIRE.

There are several instances where the assigned variety Prime name outcome 1 is an international variant of the name commonly used in Australia. In terms of international naming convention, at least in a scientific framework, it is best to use the Prime names as they appear in VIVC. However, in Australia, the commonly used variant of the name will likely continue to be used. It is recommended that in official and scientific documents, both the variety Prime name and common name for the region or country be used. Examples are as follows:

- KOENIGIN DER WEINGAERTEN: German name for common name Queen of the Vineyard. Queen of the Vineyard is a VIVC synonym of KOENIGIN DER WEINGAERTEN

- KORINTHIAKI: Greek name for common name Zante Currant. Zante Currant is a VIVC synonym for KORINTHIAKI

- MUELLER THURGAU WEISS: German name for common name Muller Thurgau. Muller Thurgau is a VIVC synonym of MUELLER THURGAU WEISS WEINGAERTEN

- MUSCAT FLEUR D'ORANGER: French name for common name Orange Muscat. Orange Muscat is a VIVC synonym of MUSCAT FLEUR D'ORANGER

- SYRAH: French name for common name Shiraz

- VELTLINER GRUEN: VIVC assigns Italy as the country of origin of this variety. It is also known as Green (or White) Veltliner. Green Veltliner is a VIVC synonym for VELTLINER GRUEN

A total of 221, 69 and 12 (302) genotypes in the CSIRO, SARDI and OTHER collections, respectively, were assigned a variety Prime name with outcome 1 (Table 3). However, because some genotypes were present in more than one of the collections, this reduces to 236 unique genotypes across the three collections. Outcome 1 is a full match with a genotype in the ICVV-SNP database with a variety Prime name. Genotypes assigned a variety Prime name with outcome 1 (236), carry the most confidence with respect to identification, although as noted elsewhere in this report, no guarantees or warranties can be made.

6.2.2.1 General observations relating to genotypes assigned Outcome 3 variety Prime names

General observations relating to genotypes that almost fully match with a genotype in the ICVV-SNP database and assigned a variety Prime name with outcome 3 follow, noting that it is an inconclusive identification, since the assigned name is the Prime name of the closest genotype rather than of the same genotype:

- Two accessions in the CSIRO collection, one with collection name Baufrac and one with collection name Limberger, were assigned the variety Prime name BLAUFRAENKISCH. Limberger is a VIVC synonym for BLAUFRAENKISCH.
- All Malbec accessions have been assigned the variety Prime name COT. Malbec is a VIVC synonym for COT. The Malbec name is widely used in Australia and in other major wine regions, (for example, the Mendoza region of Argentina) and is therefore unlikely to be replaced by COT. The identification was subsequently confirmed by SSR analysis and variety Prime name outcome 1 assigned. Over time, there may be a gradual acceptance of COT as the Prime name with retention of Malbec as the common name.
- Six accessions named Mataro in the CSIRO collection have been assigned the variety Prime name MONASTRELL. Mataro is a VIVC synonym for MONASTRELL. While the identification is inconclusive, the Mataro name is widely used in Australia and in other major wine regions. If the identification was checked using SSRs and confirmed as MONASTRELL with variety Prime name outcome 1, the Mataro name is unlikely to be replaced by MONASTRELL, although over time there may be a gradual acceptance of MONASTRELL as the Prime name.
- A range of accessions with names including Brown Frontignan, Black Frontignan, Cape Currant, Grant's Currant, Muscat Blanc and Muscat Rouge have been assigned the variety Prime name Muscat a petits grains. Muscat Blanc and Muscat Rouge are VIVC synonyms for Muscat a petits grains.

- Accessions (with number of accessions between brackets) named Boal (2), Farana (2) and Bobal (1) have been assigned the variety Prime name PLANTA FINA. Farana is a VIVC synonym for PLANTA FINA.
- Three accessions in the CSIRO collection named Touriga have been assigned the variety Prime name TOURIGA NACIONAL. Touriga is a VIVC synonym for TOURIGA NACIONAL.
- All accessions of Riesling have been assigned the variety Prime name RIESLING WEISS. Riesling is a VIVC synonym for RIESLING WEISS.
- As for some genotypes assigned variety Prime name outcome 1, there are examples of genotypes assigned variety Prime name outcome 3 where the assigned name is the European name rather than the name commonly known for the variety in Australia. An example is Garnacha, which is the Spanish name, while Grenache is the common name used in Australia.

6.2.3 Genotypes assigned Outcome 2/3 variety Provisional names

Where a genotype cannot be assigned a variety Prime name, a Provisional name is normally assigned. A Provisional name was assigned in the ICVV-SNP database when there had been insufficient independent evidence to enable establishment of a variety Prime name for the genotype. Thus, importantly, Provisional names cannot be taken as the right names for any genotype.

Full lists of varieties in alphabetical order from the CSIRO, SARDI and OTHER collections that have been assigned Provisional names with outcomes 2 or 3 are presented in Supplementary Tables 1-3.

Table 5 shows the total number genotypes in each collection that were assigned Provisional names outcomes 2 and 3 (both defined in section 5.4). Outcome 2, involving a full match with a genotype in the ICVV-SNP database with a Provisional name was the predominant outcome for all collections, comprising 132 of 206 (64%), 23 of 36 (63.9%) and 4 of 5 (80%) Provisional names in the CSIRO, SARDI and OTHER collections, respectively. Outcome 3, where the sample DNA profile almost fully matched with a genotype in the ICVV-SNP database resulting in the name of the closest genotype being assigned, in this case a Provisional name, included 74 of 206 (35.9%), 12 of 36 (33.3%) and 1 of 5 (20%) of genotypes in the respective collections.

For 246 of 247 genotypes with multiple accessions, all accessions of a specific genotype had either an outcome 2 or 3 (Table 5). However, for the genotype assigned the Provisional name Tinto in the SARDI collection, with 2 accessions, the accession with collection name Americano was assigned outcome 2 and the accession with collection name Dolcetto was assigned outcome 3. This example is shown in Table 5 under naming outcome 2/3.

Of the 247 genotypes assigned Provisional names, 206 (83%), 36 (15%) and 5 (2%) were from the CSIRO, SARDI and OTHER collections, respectively (Table 5).

Table 5. Total number of genotypes assigned Provisional names with outcomes 2 or 3 or mixed outcomes 2 & 3 according to the collections from which they originated. OTHER = Adelaide Hills Vine Improvement, New South Wales Department of Primary Industries, West Australian Agriculture Authority and Yalumba Nursery.

Naming Outcome	Number of genotypes / Naming outcome (Provisional names) / Collection			
	CSIRO	SARDI	Other	All
2	132	23	4	159
3	74	12	1	87
2/3	0	1	0	1
Total	**206**	**36**	**5**	**247**

As previously discussed, the difference between the number of accessions assigned a Provisional name and the number of genotypes assigned a Provisional name, for each naming outcome, is the number of clones or replicates of genotypes where for each genotype the accessions had the same DNA profile based on the set of 48 SNPs that were used. These 'identical' genotypes, in some cases, are different clones that may have an observable phenotypic difference or differences which was/were unable to be differentiated with the set of 48 SNPs. Alternatively, they may be identical (true-to-type) accessions of the same genotype propagated vegetatively, e.g. from hardwood cuttings, or, as in the CSIRO collection, varieties that were subjected to a virus elimination procedure, called Fragmented Shoot Apex Culture (FSAC) and propagated through tissue culture.

Table 6 shows the number of clones or replicates of each genotype in each collection that has been assigned a Provisional name. In total, of 314 accessions with Provisional names, there were 247 genotypes and 67 clones or replicates.

Table 6. Clones or replicates of genotypes that were assigned Provisional names according to the collections from which they originated. OTHER = Adelaide Hills Vine Improvement, New South Wales Department of Primary Industries, West Australian Agriculture Authority and Yalumba Nursery.

Category	Number per collection and total			
	CSIRO	SARDI	Other	All
(a) Accessions with Provisional names	262	47	5	314
(b) Genotypes with Provisional names	206	36	5	247
Clones or replicates (a minus b)	56	11	0	67

6.2.3.1 General observations relating to genotypes assigned Provisional names

General observations relating to genotypes that have been assigned Provisional names follow:

- Seven accessions with the name Doradillo were assigned the Provisional name *Mantuo De Pilas* (outcome 2). Doradillo and Mantuo Pilas are VIVC synonyms of Mantuo De Pilas.
- A range of CSIRO named varieties, for example *M51-18, M13-01, M44-14, M48-42, Merbein 5489, Merbein 5512 and Merbein 6262* have been assigned Provisional names, which is because of their novelty and hence not currently recorded in VIVC, which may change in the future.
- There are several instances of genotypes with Provisional names where there is a number in brackets after the name. Examples are *Baco* (3), *Bonarda* (2), *Corvina Veronese* (2), *Montils* (2) and *Santa Paula* (2). This means there is more than one genotype in the ICVV collection with the same Provisional name. However, the number between brackets does not refer to the total number of genotypes with the same Provisional name, rather it is used to distinguish each of those genotypes. For example, for the variety Corvina Veronese, there is an accession from Italy, an accession from Argentina and an accession from the CSIRO collection which was provided in a 2009 study. To explain further, GEN_SNP_1343 (from Italy) is provisionally named *Corvina Veronese (1)*, GEN_SNP_1130 (from CSIRO in 2009) is provisionally named *Corvina Veronese (2)* and GEN_SNP_2531 (from Argentina) is provisionally named *Corvena Veronese (3).* The accession from the CSIRO collection in 2019, with collection name Corvena Veronese, matches to GEN_SNP_1130, so it has been assigned Provisional name *Corvina Veronese (2).*
- Further emphasising the important point that 'Provisional names cannot be taken as the right names for any genotype', there are instances where Provisional names have been assigned outcome 3 but where the reference genotype in the ICVV database is based on single samples provided by CSIRO as part of a previous DNA profiling project with ICVV. Corvina Veronese is one such example. This is an instance where the sample genotype is matching against 'itself' and therefore remains inconclusive. Some CSIRO bred varieties that are not in the public domain are in this category. However, for CSIRO bred varieties, where detailed ampelographic information is available and used for the purposes of obtaining Plant Breeders Rights, the naming can be considered conclusive.

6.2.4 Genotypes unable to be assigned a varietal name (Outcomes 4, 5 & 6)

Of the 1514 accessions analysed, 175 (11.6%) could not be assigned either a Prime or Provisional name. This category includes 132 of 1037 (12.7%) of CSIRO accessions, 41 of 455 (9%) of SARDI accessions, and 2 of 22 (9.1%) of accessions from the OTHER collections.

It was noteworthy that of the 175 accessions which could not be assigned names, 139 (79.4%) were, to the best of our knowledge, genotypes with one or more non-vinifera species in their pedigree or background, many of which are rootstocks. This highlights the point made earlier that it may not be possible to use the set of 48 SNPs for identification of such genotypes.

6.3 Stage 2 - Verification of selected accessions to form the basis of an Australian national germplasm collection

In Stage 2, a decision was made, following discussions with Wine Australia, to confirm the identity of a selection of the most important wine, table/dried and multi-purpose grape varieties from the germplasm collections that had been assigned Prime or Provisional names, which could form the basis of an Australian national germplasm collection.

Clonal accessions that that were either replicated within individual collections or duplicated between collections were identified and narrowed down to a single representative clone. Leaf samples of these selected accessions were collected from the germplasm collections in October 2020 and sent to AGRF for DNA extraction.

6.3.1 Verification of assigned variety Prime names (Outcome 1) of selected accessions by duplicate SNP analysis

For accessions that had been assigned an Outcome 1 variety Prime name, SNP analysis was repeated by AGRF, using the same SNP panel as in Stage 1, and the SNP results were compared to our own in-house SNP profile database that had been developed as part of Stage 1. In total, 583 accessions (Table 7) were resampled and re-analysed using this methodology. The results presented in Supplementary Table 4 show that all of the accessions sampled gave a SNP profile that matched that obtained by the SNP analysis in Stage 1. This demonstrates the excellent reproducibility of the SNP-based DNA typing method.

Table 7. Number of genotypes and accessions and in the different collections with verified Prime names based on duplicated SNP analysis.

Category	Number per collection				
	CSIRO only	**SARDI only**	**CSIRO & SARDI**	**YALUMBA only**	**Total**
Accessions with verified variety Prime names	243	6	331	3	583
Genotypes with verified variety Prime names	176	5	49	3	233

A total of 233 unique wine, table/dried and multi-purpose grape varieties genotypes were successfully verified by this duplicate SNP analysis. Of these, 204 are wine grape varieties, 28 multi-purpose wine grape varieties and 25 table/dried grape varieties. Over 75% of these genotypes are maintained in the CSIRO collection, 21% maintained in both the CSIRO and SARDI collections and the remainder unique to the SARDI or Yalumba genotype collections.

When the clones associated with the verified genotypes are also taken into account, the number of unique accessions verified by SNP analysis increases to 583. However, it should be remembered that

the set of 48 SNPs that were used in this project cannot be used to distinguish clones. Thus, the true number of unique clones of each verified genotype within the germplasm collections remains to be confirmed by other techniques such as next-generation sequencing.

6.3.2 Verification of selected accessions assigned Prime/Provisional names (Outcomes 2 & 3) by SSR analysis

For accessions that had been assigned Outcome 2 or Outcome 3 Prime or Provisional variety names in Stage 1, an alternative method of DNA typing based on simple sequence repeat (SSR) analysis was employed to verify the SNP analysis. However, because of the much greater expense involved in SSR analysis, compared to SNP analysis, it was not feasible to undertake SSR analysis on all 588 Outcome 2 and Outcome 3 accessions assigned a variety name by SNP analysis in Stage 1. Therefore, a different sampling strategy was devised.

Where there was only a single representative of a unique variety in the collection, as identified by a unique SNP profile, this accession was selected for SSR analysis. Similarly, where there were multiple accessions (clones) of a unique variety which all had identical SNP profiles, a single representative accession was chosen for SSR analysis. Finally, where more than one clone of a single variety was found to have a minor difference in SNP profile, one representative of each SNP profile pattern was selected for SSR analysis. We reasoned that the SSR variety name assigned to one representative of a group of clones, with an identical SNP pattern, could be assigned to all members of that group.

A total of 184 scion accessions that had been assigned Outcome 2 or Outcome 3 Prime or Provisional variety names by Stage 1 SNP analysis, were selected for SSR analysis. Leaf material was sampled from the vineyard, DNA extracted by AGRF and the DNA samples sent to the Institut Francais de la Vigne et du Vin (IFV) in France for SSR analysis. The results of this scion analysis are shown in Supplementary Table 5.

From the 185 scion DNA samples analysed using SSRs, 150 (81%) were found to match samples in the IFV SSR database and could be assigned a variety name (Table 8). The remaining 35 samples did not match any SSR profiles in the IFV database and were classified as ‘unknowns’. Of the 150 that were assigned a variety name by SSR analysis, 137 were found to match the Prime or Provisional name that had previously been assigned using SNP analysis in Stage 1, and one matched a previous SSR analysis carried out on this accession. Another 12 accessions that were assigned variety names by SSR analysis, did not match the Outcome 2 Provisional or Outcome 3 Prime or Provisional variety names assigned using SNP analysis in Stage 1.

Table 8. Number of accessions assigned variety names based on SSR analysis

Stage 1 SNP Outcome	Number of Accessions			
	SSR assigned name = SNP assigned name	**SSR assigned name ≠ SNP assigned name**	**Unknown**	**Total**
Outcome 2 - Provisional variety name (*p, 2*)	39	6	22	67
Outcome 3 - Prime variety name (P, 3)	75	1	1	77
Outcome 3 - Provisional variety name (*p,* 3)	23	5	12	40
Total	**137**	**12**	**35**	**184[1]**

[1]*The actual number of scion DNA samples analysed by SSR analysis was 185. Of these 184 were Outcome 2/3 genotypes based on Stage 1 SNP analysis and one was a sample of a variety that had previously been confirmed by ADH using SSR analysis to be CASTETS.*

A total of 130 unique wine, table/dried and multi-purpose grape varieties genotypes were successfully verified by this combined SNP/SSR analysis. Of these, 93 are wine grape varieties, 27 multi-purpose wine grape varieties and 10 table grape varieties. Over 66% of these genotypes are maintained in the CSIRO collection, 19% maintained in both the CSIRO and SARDI collections and the remainder unique to the SARDI or Adelaide Hills Vine Improvement genotype collections (Table 9).

When the clones associated with these verified genotypes are also taken into account, the number of accessions verified by combined SNP/SSR analysis increases to 292. However, as highlighted above, the true number of unique clones of each verified genotype within the germplasm collections remains to be confirmed by other techniques such as next-generation sequencing.

Table 9. Number of accessions and genotypes in the different collections with verified Prime names based on combined SNP and SSR analysis.

Category	Number per collection				
	CSIRO only	**SARDI only**	**CSIRO & SARDI**	**ADH only**	**Total**
Genotypes with verified variety Prime names	86	18	25	1	130
Accessions with verified variety Prime names	102	21	168	1	292

6.3.3 Verification of rootstock genotypes by SSR analysis

As highlighted in section 6.2.4, the 48 SNPs used for SNP analysis were not useful for identifying genotypes with one or more non-vinifera species in their pedigree or background. As a result, it was not possible to genotype the rootstock accessions analysed in Stage 1. We therefore took the opportunity to send DNA samples of 20 of the most important rootstocks, in addition to the 185 scion accessions, to IFV for SSR analysis.

In contrast to results obtained by SNP analysis, SSR analysis confirmed the identity of the 19/20 rootstocks analysed (Supplementary Table 6). The only rootstock that could not be identified by SSR analysis was K51-32.

6.3.4 Mis-named accessions identified in the Australian germplasm collections

Supplementary Tables 4 and 5 list accessions in the CSIRO and SARDI collections which show significant differences between the varietal name assigned in the Australian germplasm collections and the Prime or global Prime name assigned based on the SNP/SNP (Supplementary Table 4) or SNP/SSR (Supplementary Table 5) analysis. All collection names were checked to see if they were recognised synonyms of the variety Prime or global Prime names in the VIVC database names. If they were not accepted synonyms they were listed as mis-named accessions.

A total of 34 accessions are highlighted in Supplementary Tables 4 and 5 as being mis-named in the existing collections; 29 in the CSIRO collection and 5 in the SARDI collection.

In order to assess the potential impact on the Australian wine industry a number of parameters were assessed. Firstly, the 8 mis-named accessions that are only used for table or dried grape production were excluded. Next the records of distribution of canes from both the CSIRO and SARDI collections were examined to determine if any of the 26 remaining mis-named accessions had been distributed from the collections. Our investigations revealed that none of the mis-named accessions had been distributed from the SARDI collection, but 11 had been distributed from the CSIRO Merbein collection prior to its closure in 2008 (Supplementary Table 7).

One possible explanation for the mis-naming of at least some of CSIRO accessions now present in the Irymple collection, is mis-labelling that occurred during the transfer of the collection from Merbein to Irymple. Fortunately, it was discovered that 9 of these 11 mis-named accessions had been previously sampled and genotyped in 2008, prior to the move, as part of a SNP pilot study carried out by CSIRO. However, at that time, many of the genotypes could not be confirmed because of the limited number of SNP profiles in the Spanish ICVV database. We therefore re-analysed the 2008 SNP profile data for these 9 samples, with the more extensive SNP database generated from this project. This analysis confirmed the identity of the DURIF, SAPERAVI, Barbera (BARBERA NERA), Malbec (COT), Crujidera (MARUFO), Red Veltliner (VELTLINER ROT) and REICHENSTEINER accessions in the Merbein collection, prior to their transfer to Irymple. In contrast, the accession named Lignan Blanc was confirmed to be AGOSTENGA and the accession named Siderites was confirmed to be RODITIS, based on the 2008 sampling.

No sampling of Riesling was carried out in 2008 for the SNP pilot study. However, CSIRO's distribution records indicated that Riesling 110-18 had been provided to the Victorian and Murray Valley Vine Improvement Association (VAMVVIA) and enquires revealed that this material was still present in the VAMVVIA collection. This material was subsequently sampled in 2021 and SNP genotyped and confirmed to be Riesling (RIESLING WEISS) (Supplementary Table 7).

In summary, this analysis confirmed that 8 out of the 11 mis-named accessions currently in the CSIRO Irymple collection (Supplementary Table 7), were the result of mis-labelling that occurred during the move of the CSIRO germplasm collection from Merbein to Irymple. Thus, any material of these 8 accessions that was distributed prior to 2008 was correctly named. However, it would appear that the accessions labelled Lignan Blanc and Siderites were misnamed in the original Merbein collection.

6.4 SNP analysis on DNA extracted from wood/bud samples

Although the best quality and yield of DNA for SNP analysis is generally obtained from leaf samples collected from growing shoot tips, such material may not be available to a grower when they decide they wish to undertake SNP genotyping on grapevines in their vineyard. For a commercial SNP-based service to be viable, it should be capable of undertaking SNP analysis on DNA extracted from a range of different grapevine samples at different times of the year.

We undertook a pilot study to examine the feasibility of using DNA extracted from dormant buds and wood from canes in comparison to leaf material. Material was collected from 14 different grapevine genotypes for which we already had SNP profiles. The results shown in Table 10 indicate that the failure rate for calling individual SNPs of the 48 SNP panel was often higher with DNA extracted from buds and wood compared to leaves. This probably reflects the poorer quality and/or yield of the DNA extracted from these tissue types compared to leaves. As a result, the amount of product produced in the Sequenom iPlex assay is not sufficient to allow detection of a SNP product ('fail') or results in an incorrect reading of the SNP output by the mass spectrometer ('mis-read').

Despite this, the slightly higher SNP failure rate observed with DNA extracted from buds or wood had no impact on our ability to assign a varietal name to the genotypes because only a minimum 36 SNPs are required to accurately assign a genotype (Javier Ibáñez, personal communication).

Table 10. Comparison of SSR analysis results obtained using DNA extracted from leaf, bud and cane wood samples. The same panel comprising 48 SNPs was used as described previously.

Varieties sampled	Source Material	Fail	Misread
SYRAH	Bud	0	1
	Wood	2	0
	Fresh leaf	0	0
ARNEIS	Bud	3	0
	Wood	1	0
	Fresh leaf	1	0
KORITHI ASPRO	Bud	1	0
	Wood	1	0
	Fresh leaf	0	0
DOLCETTO	Bud	0	0
	Wood	0	0
	Fresh leaf	0	0
Pedro Ximenes	Bud	0	0
	Wood	1	0
	Fresh leaf	0	0
Valensi	Bud	3	0
	Wood	0	2
	Fresh leaf	0	0
MUSCADELLE	Bud	1	0
	Wood	0	0
	Fresh leaf	0	0
Chardonnay	Bud	0	0
	Wood	0	0
	Fresh leaf	0	0
Muscat a petits grains	Bud	1	0
	Wood	0	1
	Fresh leaf	0	0
SAUVIGNON BLANC	Bud	3	0
	Wood	2	0
	Fresh leaf	0	0
COLOMBARD	Bud	0	1
	Wood	1	0
	Fresh leaf	0	0
Clairette	Bud	0	0
	Wood	0	0
	Fresh leaf	0	0
Villard Blanc	Bud	1	0
	Wood	1	0
	Fresh leaf	0	0
MANTUO DE PILAS	Bud	1	0
	Wood	4	0
	Fresh leaf	0	0

6.5 Previously unnamed genotypes in historical collections assigned Prime names

It is interesting to note that this DNA typing project has also enabled us to assign VIVC Prime names to un-named genotypes in the CSIRO and SARDI germplasm collections that had previously been collected from historical collections (Table 11; Supplementary Tables 4 & 5).

Table 11. List of previously un-named genotypes in historical collections that have been assigned VIVC Prime names.

VIVC Prime Name	Name in Collection
ASPIRAN NOIR	Best's R1V12
ASPIRAN NOIR	Best's R2V40
GALEGO DOURADO	Best's R15V92 ex Great Western
Palomino	Best's R5V104
TRINCADEIRA	Best's R3V54
Tibouren	Best's R3V47
BOURBOULENC	Coldstream R1V13
BOURBOULENC	Coldstream R2V10
BOURBOULENC	Coldstream R6V11
Chasselas	Coldstream R10V10
LISTAN PRIETO	Auldana No 3

Genotypes from the historical collections include the names 'Bests', 'Coldstream, Graytown' and Auldana, along with identification, for example, row and vine numbers, accession number or letter, that reflect their origins.

Joseph Best was the original planter of the Seppelt's Vineyard at Great Western, located approximately 219 km west of Melbourne (Evans 1973). After successive changes of ownership through the late 1800s into the early 1900s, considerable expansion of vineyards took place in the Great Western region between the 1940s and 1960s (Evans 1973).

The younger brother of Joseph Best, Henry Best, established a wide range of grapevine varieties at a vineyard on the banks of the Concongella River, close to the Great Western Township, in the 1860s (Evans 1973).

Establishment of these vineyards in the mid-1800s was prior to the outbreak of phylloxera in Europe in the late 1800s during which some *Vitis vinifera* varieties were lost in Europe.

Further collections of grapevine varieties were maintained at Coldstream, around 40 km north-east of Melbourne and at Graytown, a former gold mining town in central Victoria, about 117 km north of Melbourne. Much less information is available about the Coldstream and Graytown plantings.

Dr Alan Antcliff, the scientist appointed to lead grapevine breeding at CSIRO in the 1960's through to 1980's, coordinated the procurement of grapevine accessions from sources within Australia and from

major grapevine collections internationally. This included sourcing accessions from the Great Western, Coldstream and Graytown regions of Victoria. The procurement program aimed to expand the genetic resources available for breeding. The accessions were identified by collection name, plus a letter or number code specific to the collection. The accessions were established in CSIRO's Merbein collection mostly during the 1960s and 70s.

Similarly, SARDI procured a limited number of accessions from what was called the 'Auldana collection' in South Australia. Auldana is now a suburb of Adelaide not far from Penfolds Estate Cellar Door at Magill. The suburb stands on the site of the once famous Auldana vineyard established by Patrick Auld in 1847. An accession with the collection name Auldana No. 3 has been assigned variety Prime Name with outcome 1 of Listan Prieto.

It is interesting also to note that three Best's accessions in the CSIRO collection genotyped the same as an accession with the Provisional Name *Graytown R3#2,* demonstrating interchange of accessions between historical collections. All three Best's accessions were assigned the Provisional Name *Graytown R3#2,* however, a true identification of the variety is yet to be determined and these accessions do not appear in Table 11.

7 Discussion

Genotyping by DNA profiling is used by all major international grapevine collection holders for the characterisation of their germplasm collections (Meredith et al. 1995; Lopez et al. 1999; Borrego et al. 2001; Dettweiler and Eibach 2003; This et al. 2004; Santiago et al. 2005; Frei et al. 2006; Moreira et al. 2014; Maul et al. 2015). Prior to the development of SSR markers for grapevine variety identification (Thomas et al. 1994), the more traditional science of ampelography was relied upon for germplasm characterisation. Ampelography is still used, increasingly in parallel with DNA profiling, particularly where there is a need to resolve questions about grapevine identity raised by DNA profiling results or to resolve clonal differences. The combination of DNA profiling and ampelography has been used where required in this study.

Microsatellite (SSR) markers have proven very useful for genetic identification and parentage analysis because of their relative advantage such as co-dominance, multi-allelism and high levels of polymorphism (Cabezas et al. 2011). However, there are also a number of disadvantages, the most important of which is allele binning whereby raw allele lengths must be converted into allele classes expressed by integer numbers. Alternatives to SSR markers have been investigated. This includes single nucleotide polymorphisms (SNPs). A set of 48 stable SNP markers has been selected with high discrimination power and proposed for use in grapevine genotyping (Cabezas et al. 2011). We used this set of 48 SNPs to genotype accessions in major Australian grapevine collections. Then in collaboration with Dr Javier Ibanez (ICVV), the DNA profiles of accessions in the Australian collections were compared with DNA profiles in the ICVV collection which had been genotyped with the same set of 48 SNPs. Where a match or close match occurred between the Australian sample DNA profile and the DNA profile of a genotype in the ICVV collection, a naming outcome was assigned. However, it needs to be remembered that certain variety names assigned to some SNP profiles within the ICVV database must be considered Provisional only because they have not been verified on multiple independent samples. This is the reason that this project originally set out with the aim of comparing our SNP data with the SNP databases of multiple international collaborators. However, as this could

not be organized in the time frame of the project, we relied on the ICVV SNP database for assigning varietal names.

The SNP analysis in Stage 1 of 1514 accessions collected from the two major and four of the minor grapevine germplasm collections in Australia identified 751 accessions that showed a perfect match to a SNP profile within the Spanish ICVV database with a variety Prime name (Outcome 1). This enabled us to assign a VIVC Prime or global Prime variety name to these 751 accessions in the list of Australian germplasm accessions. Subsequent re-sampling and re-analysis by SNP genotyping of a subset of 583 unique wine, table/dried and multi-purpose grape varieties selected from these 751 accessions confirmed the identity of all 583 accessions comprising 233 unique genotypes. This demonstrates the excellent reproducibility of the SNP DNA genotyping method.

The SNP analysis in Stage 1 also identified 588 accessions that could not be assigned a variety Prime name because they were either (a) a perfect match to genotypes within the ICVV database with a Provisional name (Outcome 2) or (b) almost fully matched with profiles of genotypes within the ICVV database with a variety Prime or Provisional name (Outcome 3). As this group included accessions of a number of important wine grape varieties such as Riesling (RIESLING WEISS), Malbec (COT), and Grenache (**Garnacha**), SSR analysis was carried out on these accessions in an attempt to verify their identity. Re-sampling and SSR analysis of 184 representative accessions of this group resulted in the verification of a further 130 unique genotypes and 292 accessions.

When the outputs from both the duplicate SNP/SNP analysis and the combined SSR/SNP analysis are considered together, this study has been able to assign a VIVC Prime/global Prime name to 362 unique genotypes and 875 accessions currently within these six Australian germplasm collections which could form the basis of a future Australian National Grapevine Germplasm Collection.

In addition, this study has resulted in the development of a SNP profile database that could be used by a commercial DNA typing service to identity 336 unique grapevine genotypes with a high degree of reliability and reproducibility. Initial testing of DNA extracted from dormant buds and cane wood also suggests that it will be feasible for grape growers to submit vine samples at any time of the year and not be restricted to having to wait until spring to collect leaves from growing shoots.

While SNP genotyping was shown to be highly successful with identifying *V. vinifera* scions there remain some limitations. The first is that the 48 SNP panel used in this study was ineffective at identifying genotypes with non-vinifera Vitis species in their pedigree. Of the 175 accessions that could not be assigned varietal names in the Stage 1 SNP analysis (Table 1), 139 (79.4%) had a non-vinifera background. Many of these accessions were rootstocks which were subsequently confirmed by SSR analysis (Supplementary Table 6). It would appear that at least for the present, SSRs will remain the technique of choice for the identification of genotypes containing significant amounts of non-vinifera *Vitis* species.

The other major limitation of the 48 SNP marker set used in this study is its inability to be able to differentiate clones of varieties. Most notable was the inability of the marker set to distinguish colour variants and potentially also variants in seediness, e.g. seeded and seedless accessions of Sultanina, and leaf hairiness, which differs between three accessions of LUGLIENGA BIANCA. The colour locus is well separated from the loci for the 48 SNPs, so it is not surprising that colour variants were not identified. This is best demonstrated by the need to use the global Prime name 'Pinot' to indicate a cluster of varieties derived from a single original embryo. SNP analysis using the 48 SNP panel was

unable to differentiate between Pinot Noir, Pinot Gris, Pinot Blanc and Pinot Meunier. Another example is the necessity to use the global Prime name Sultanina to cover a wide range of clonal variants of Sultana. Clonal differentiation requires the collection of much more genome sequencing information as was used to differentiate Chardonnay clones (Roach et al. 2018). In fact, with the ongoing development of more rapid and cheaper sequencing technologies, it is possible that SNP genotyping will be replaced by whole genome sequencing for both grapevine variety and clonal identification in the not-to-distant future.

Of the 875 accessions assigned a Prime or global Prime name based on either SNP/SNP or SNP/SSR DNA typing, 34 were found to have a collection name that was significantly different to the assigned name and was not a recognized VIVC synonym of the Prime or global Prime name (Supplementary Table 7). Of these 34, 29 are present in the CSIRO Irymple collection and 5 in the SARDI Nuriootpa collection.

Three accessions in the CSIRO germplasm collection were confirmed by SNP analysis to be mis-named in the original Merbein collection (Supplementary Table 4) and confirmed to have been distributed from the collection to outside sources. These are Lignan Blanc (correct variety name AGOSTENGA), Siderites (correct variety name RODITIS) and Petit Manseng (correct variety name MANSENG GROS BLANC syn. Gros Manseng). However, a search of the National Vintage Survey for the years 2015 to 2019 found no mention of the varieties Lignan Blanc or Siderites indicating that it is not currently being used for wine grape production in Australia.

In contrast, the National Vintage Survey indicated that an average of 44 tonnes per year of the variety Petit Manseng was reported for the period 2015-2019. If the vine material currently being used for production of Petit Manseng grapes in Australia was originally derived from the CSIRO germplasm collection, our data would indicate that these grapes are actually of the variety MANSENG GROS BLANC (syn. Gros Manseng) and not Petit Manseng. In 2020, Wine Australia identified the potentially impacted parties through the export and the vintage survey data and contacted them directly about this issue.

The Australian Wine Industry has a label integrity program and the results of this study have potential consequences for the program. Some 23% of the genotypes identified within the CSIRO and SARDI collections are labelled in the collections with synonyms of the recognized VIVC Prime name. Therefore, a revision of the list of names and synonyms that are officially recognised by the Australian Wine Industry should occur. Furthermore, it is important that there is consistency in the naming of grapevine varieties between various official registers of grapevine varieties in Australia and internationally. The Grapevine Variety Selector Tool should be updated for consistency between names used in the Selector Tool and assigned variety Prime names from this study.

8 Outcomes/Conclusions

This project has successfully demonstrated the reliability and reproducibly of SNP genotyping for the identification of wine, table and dried grape varieties. The SNP-based technique can be used successfully on DNA extracted from leaves, cane wood or dormant buds. Furthermore, the analysis can be undertaken by commercial companies within Australia, meaning that samples do not have to be sent to facilities overseas. However, the current 48 SNP panel is not suitable for reliably DNA typing genotypes with non-vinifera *Vitis* species in their pedigree e.g. rootstocks.

A survey of the two major and four of the minor Australian grapevine germplasm collections has identified 362 unique wine, table/dried and multi-purpose grapevine varieties and 875 accessions that could form the basis of the National Grapevine Collection (NGC). These accessions have been confirmed by the collection of two independent samples and DNA typing by either duplicate SNP analysis or a combination of SNP and SSR analysis. Of the 875 accessions, 871 are held in either the CSIRO Irymple collection (63%) or the SARDI Nuriootpa collection (37%). The 362 unique genotypes have been assigned Prime or global Prime names based on the Vitis International Variety Catalogue (VIVC) which, in a number of cases, differs from the varietal name currently used by the Australian wine industry

A SNP database containing the SNP profiles of all 362 unique genotypes has been assembled that could form the basis of a commercial SNP-based DNA typing service which is much more extensive than the SSR-based DNA typing service currently available in Australia. However, the 48 SNP panel used in this study was not able to differentiate between clones of individual genotypes.

An outcome of potential general interest has been the assignment of variety Prime names to 11 accessions from historical Australian collections that were held in the CSIRO and SARDI collections for many years. Some of these accessions pre-date the phylloxera crisis in Europe in the late 1800's. These include 5 accessions from the Best's (Best Western, Victoria) collection, 5 accessions from the Coldstream (Victorian) collection and 1 accession from the Auldana (South Australia) collection.

The results of this project pave the way for the reopening of the CSIRO and SARDI germplasm collections, which have been closed to industry since 2009. Varieties held in these collections form an integral part of the NGC which is currently being established, and access to material will be made possible under the auspices of the NGC.

7 Recommendations

1. A revised Australian grapevine DNA genotyping service should be considered based on the 48 SNP profiles of the 362 unique varieties assigned Prime or global Prime names in this study.
2. The list of names and synonyms that are officially recognised by the Australian Wine Industry should be revised for consistency in naming of grapevine varieties between various official registers of grapevine varieties in Australia and internationally
3. The officially-recognised names should be incorporated into a national public register of planting material in Australia, being developed as part of the National Grapevine Collection.
4. Variety names in the Grapevine Variety Selector Tool, currently being developed by Wine Australia, should cross-reference for consistency, at least for varieties assigned Prime names as listed in Supplementary Tables 4 and 5.
5. An industry-led strategy should be developed to deal with instances where accessions in germplasm collections have been found to have variety names that are different to those assigned by DNA typing in this study and are not a recognised synonym of the assigned name.
6. The results of this project pave the way for the reopening of the CSIRO and SARDI germplasm collections, which have been closed to industry since 2009. Varieties held in these collections have the potential to add significant benefits to the sector and include varieties with immediate commercial potential, heritage/pre-phylloxera selections and climate change adaption varieties. These varieties will form an integral part of the National Grapevine Collection (NGC) which is currently being established, and access to material will be made possible under the auspices of the NGC.

Appendix 1: Communication

Publications

No scientific publications were completed during the course of the project

Industry articles

The following article was published soon after the commencement of the project:

- 'Another big step on the road to a national germplasm collection', written by Nick Carne and published in R&D@Work. https://www.wineaustralia.com/news/articles/another-big-step-on-the-road-to-a-national-germpla

Media exposure

The topic of grapevine variety identification is an industry-sensitive topic and it was felt best to have any media related activities at the conclusion of the project after results were summarised, reviewed and presented to Wine Australia.

However, during the undertaking of this project, an article appeared on social media suggesting that growers who had planted what they thought was Petit Manseng was actually Gros Manseng. This information came to light after an ampelographic inspection of germplasm in Monash collection carried out in 2019. The Stage 1 SNP analysis that we had independently undertaken confirmed this mis-naming (Supplementary Table 1). Wine Australia subsequently published a story about the mis-identification on their website: https://www.wineaustralia.com/news/articles/grape-variety-misidentified on 30th April 2020. This attracted some media interest and ABC radio asked to interview Ian Dry about the issue. Following the interview the ABC published an article about the mis-naming issue on their website: https://www.abc.net.au/news/2020-05-06/decades-old-petit-manseng-mix-up-confirmed-by-csiro/12209306 on 6th May 2020.

Appendix 2: Intellectual Property

A Register of Background Intellectual Property was developed and submitted to Wine Australia at the time of project commencement.

The primer sequences for the SNP markers are confidential.

Material Transfer Agreements were prepared with Australian grapevine collection holders willing to submit accessions to the project. These include:

- Adelaide Hills Vine Improvement
- West Australian Agriculture Authority
- New South Wales Depart of Primary Industries
- Yalumba Nursery

A Technical Support Contract was entered into between Consejo Superior De Investigaciones Cientificas (CFIC) and CSIRO to enable use of the ICVV-SNP database for variety identification of Australian grapevine germplasm. This included two variations to extend the end date of the contract.

Genotyping was done by the Australian Genome Research Facility Limited under a contract Research Services to CSIRO Agreement with CSIRO.

Appendix 3: References

Anon (2013) Australian Standard. Grapevine propagation material. AS 5588-2013. Standards Australia Limited, Sydney, New South Wales, Australia.

Anon (2020) Grapes. Foundation Plant Services Fact Sheet. University of California, Davis. College of Agriculture and Environmental Sciences.

Cabezas, J.A., Ibanez, J., Lijavetzky, D., Valez, D., Bravo, G., Rodriguez, V., Carreno, I., Jermakow, A.M., Carreno, J., Ruiz-Garcia, L., Thomas, M.R. and Martinez-Zapater, J.M. (2011) A 48 SNP set for grapevine cultivar identification. BMC Plant Biology 11:153

Borrego, J., Rodriguez, I., de Andres, M.T., Martin, J., Chavez, J., Cabello, F. and Ibanez, J. (2001) Characterisation of the most important Spanish grape varieties through isozyme and microsatellite analysis. Acta Horticulturae 546, 371-375.

Evans, L. (1973) Australia and New Zealand Complete Book of Wine. Pual Hamlyn Pty Ltd, Dee Why West, New South Wales, Australia, 2099.

Frei, A., Porret, N.A., Frey, J.E. and Gafner, J. (2006) Identification and characterisation of Swiss grapevine cultivars using microsatellite markers. Mitteilungen Klosterneuburg 56(5-6), 147-156.

Lopes, M.S., Sefe, K.M. and Dias, E.E. (1999) The use of microsatellites for germplasm management in a Portugese grapevine collection. Theoretical and Applied Genetics 99, 733-739.

Maul, E., Ganesch, A., Röckel, F. and Töpfer, R. (2021) Vitis International Variety Catalogue. www.vivc.de [accessed 26/10/2022].

Maul, E., Topfer, R., Carka, F., Cornea, V., Cerespan, M., Dallakyan, M., De Andres Dominguez, T., De Lorenzis, G., Dejeu, L., Goryslavets, S., Grando, S., Hovannisyan, N., Hudcovicova, M., Hvarleva, T., Ibanez, J., Kiss, E., Kocsis, L., Lacombe, T., Laucou, V., Maghradze, D., Maletic, E., Melyan, G., Mihaljevic, M.Z., Munoz-Organero, G., Musayev, M., Nebish, A., Popescu, C.F., Regner, F., Risovanna, V., Ruisa, S., Salimov, V., Savin, G., Schneider, A., Stajner, N., Ujmajuridze, L. and Failla, O. (2015) Identification and characterisation of grapevine genetic resources maintained in Eastern European collections. Vitis 54, 5-12.

McMichael, P. (2013) Summary of a review of grapevine germplasm collections in Australia. Final Report GWR1112 to Wine Australia by Scholfield Robinson Horticultural Services Pty Ltd.

Meredith, P.C., Dangl, S.G. and Bowers, E.J. (1996) Clarifying the identity of some Californian wine grapes by DNA profiling. American Journal of Enology and Viticulture 49(1), 65-68.

Moreira, F.M., Lorenzi, S., Laera, M., Sodro, M., Zulini, L., Stefanini, M. and Grando, M.S. (2014) Molecular characterisation of the grapevine germplasm collection held at the Fondazione Edmund Mach. Proceedings Xth International Conference on Grapevine Breeding and Genetics. Eds. B.I. Reisch and J. Londo. Acta Horticulturae 1046, 673-676.

Roach, M.J., Johnson, D.L., Bohlmann, J., van Vuuren, H.J.J., Jones, S.J.M., Pretorius, I.S., Schmidt, S.A. and Borneman, A. R. (2018) Population sequencing reveals clonal diversity and ancestral inbreeding in the grapevine cultivar Chardonnay. PLoS Genet 14, e1007807.

Robinson, J., Harding, J. and Vouillamoz, J. (2012) Wine Grapes. A complete guide to 1,368 vine varieties, including their origins and flavours. Penguin Books Ltd, 80 Strand, London WC2R ORI, England.

Santiago, J.L., Boso, S., Martin, J.P., Ortiz, J.M. and Martinez, M.C. (2005) Identification of grapevine cultivars (Vitis vinifera L.) from northwestern Spain using microsatellite markers and ampelometric methods. Vitis 44(2), 67-72.

This, P., Jung, A., Boccacci, P., Borrego, J., Botta, J., Constantini, L., Crespan, M., Dangl, G.S., Eisenheld, C., Ferreira-Monteiro, F., Grando, M.S., Ibanez, J., Lacombe, T., Laucou, V., Magalhaes, R., Meredith, C.P., Milani, N., Peterlunger, E., Regner, F., Zulini, L. and Mayl, E. (2004) Development of a standard set of microsatellite reference alleles for identification of grape cultivars. Theoretical and Applied Genetics 109, 1448-1458.

Thomas, M.R., Cain, P. and Scott, N.S. (1994) DNA typing of grapevines: A universal methodology and database for describing cultivars and evaluating genetic relatedness. Plant Molecular Biology 25, 939-949.

Walker, R.R. and Thomas, M.R. (2009) Evolution of grapevine variety collections in Australia. The Australian and New Zealand Grapegrower and Winemaker, Annual Technical Issue 2009, p. 78-80.

Appendix 4: List of Supplementary files

Supplementary Table 1: CSIRO accessions assigned VIVC Prime or Provisional names based on Stage 1 SNP analysis.

Supplementary Table 2: SARDI accessions assigned VIVC Prime or Provisional names based on Stage 1 SNP analysis.

Supplementary Table 3: OTHER collections accessions assigned VIVC Prime or Provisional names based on Stage 1 SNP analysis.

Supplementary Table 4: Accessions assigned VIVC Prime/global Prime names based on duplicate SNP analysis.

Supplementary Table 5: Accessions assigned VIVC Prime/global Prime names based on combined SNP & SSR analysis.

Supplementary Table 6: Verification of rootstock genotypes by SSR analysis.

Supplementary Table 7: Mis-named winegrape and multi-purpose varieties in Australian germplasm collections.

Appendix 5: Staff

CSIRO

Mark Thomas was the project leader. He developed the project proposal and liaised with potential international collaborators regarding availability of international reference collections. Mark was involved in the work to develop the set of 48 SNP markers that was used for the study. He was responsible for the design of the project and arranged genotyping with an appropriate service provider.

Mahmood Hassan was responsible for all project activity including collection of samples for DNA profiling, sending samples to the service provider for genotyping, and along with Mark Thomas, sending DNA profiles to the international collaborator for checking against the international reference collection. Mahmood also took the lead role in assessing the results supplied by the international collaborator. This included careful and thorough checking of 48 SNP profiles for accessions and assisting Rob Walker in development of summary lists for inclusion in the Final Report.

Rob Walker has been involved with CSIRO's grapevine collection since the 1980s. He has worked at the interface with industry on matters relating to the collection since that time. During the 1990s he developed more formal arrangements for access by industry to the collection which were used until its closure in 2009. He wrote a Discussion Paper on the future of grapevine germplasm collections in Australia which has contributed to CSIRO and industry strategy on collections since that time. He assisted with arrangements to establish the project and was a primary author of the Final Report to Wine Australia.

Ian Dry joined the project on the early retirement of Dr Mark Thomas prior to the completion of the project. He was involved in the planning of Stage 2 of the project and liaising with IFV to organize the SSR analysis. He wrote the sections relating to Stage 2 of the project in the Final report.

ICVV-CSIC

Dr Javier Ibáñez made a substantial contribution to the project. He is widely acknowledged for his contributions towards building the ICVV-SNP database, a non-public database made from the analysis of more than 10,000 grapevine samples sourced from around the world. The database contains 2800 unique *Vitis vinifera* genotypes with DNA profiles for 48 SNPs. Dr Ibáñez compared DNA profiles supplied by CSIRO for 48 SNPs against profiles for varieties in the ICVV-SNP database using the same 48 SNPs and assigned Prime, Provisional names or no names with one of six naming outcomes for each genotyped accession.

Appendix 6: Acknowledgements

CSIRO

Peter Clingeleffer made a very significant contribution by applying valuable skills in ampelography to address questions about accession identity. He also provided very helpful advice and discussion during the process of preparing the Final Report.

Hilary Davis assisted Mahmood Hassan in sampling accessions in the CSIRO collection.

Darren Jones assisted with searching relevant literature.

Australian Genome Resource Facility Limited (AGRF)

We thank David Hawkes and Nicole Burtt (Supervisor of the AGRF Adelaide node) for facilitating genotyping services that were an important part of the project.

Other collection holders

We thank the following people for kindly facilitating supply of leaf material for DNA extraction and genotyping from varieties that are unique to their respective collections: Nick Dry, Yalumba Nursery and Oxford Landing Estates; Prue Henschke, Adelaide Hills Vine Improvement Inc.; Diana Fisher, Western Australian Agriculture Authority; and Darren Fahey, New South Wales Department of Primary Industry.