The genetic basis of acid composition in developing berries of the cultivated grapevine *Vitis vinifera*

Emily Grace Higginson

School of Agriculture, Food and Wine

University of Adelaide

December 2015

Table of Contents

Abstract

Grapevines contain many different organic acids and the two most abundant are tartaric acid and malic acid. Malic acid and tartaric acid both increase in concentration up until veraison then after veraison malic acid is broken down as sugar increases. Malic acid has been studied in a variety of fruits for it is a very common acid. However tartaric acid is an uncommon primary acid in fruits and very little is known about its synthesis in grapevine. However, tartaric acid is important in providing a low pH which is important for the prevention of microbial spoilage during the winemaking process. A high pH of juice means that more tartaric acid will need to be added in the winery increasing the cost to wine makers. By discovering more about the genes involved in the synthesis of malic and tartaric acid and the breakdown of malic acid this knowledge could be used to breed vines with higher acid concentrations.

L-Idonate dehydrogenase (L-IDH) is one of only two genes known to participate in the tartaric acid synthesis pathway. Since its initial characterisation two more isoforms have been annotated in the grapevine genome based on sequence similarity. The characterisation of these isoforms was undertaken using a variety of techniques including expression of the proteins in *E. coli* and *in vitro* protein activity assays and also *in planta* expression in the microvine with the creation of transgenic microvines.

To try and discover regions of the genome that might be involved in acid metabolism in grapevine berries, malic and tartaric acid concentrations were measured from four progeny populations. The individuals of these populations were then sequenced using a genotyping

by sequencing method to find SNPs markers for a Genome Wide Association Study (GWAS). This GWAS was then verified with genetic mapping and QTL analysis.

During the process of measuring acid from these progeny populations the question of variability in acid concentration between berries from the same vine arose. A preliminary study into this variability was conducted to determine the variability of malic and tartaric acid in berries both within a bunch and between bunches from the same vine. This data was then used to predict the error in sampling subsets of berries of different sizes.

Tartaric acid concentration in tissues other than the berry was also explored. Acid concentration was measured in several tissues including root, shoots and leaves. It was found that tartaric acid was present in these tissues with varying concentrations. Tartaric acid concentration in leaves was then studied further try see if there was a link between the age of the leaf and tartaric acid concentration and also between leaf tartaric acid concentration and berry tartaric acid concentration. There was found to be no link between the two in these preliminary studies.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name for an other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

I consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying subject to the provisions of the Copyright Act 1968.

The Author acknowledges that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has be granted by the University to restrict access for a period of time.

Acknowledgements

I would like to thank the members, past and present, of the Thomas Laboratory and the Ford Laboratory, especially, Dr Mahmood Hassan, Pat Corena, Don McKenzie, Dr. Crystal Sweetman and Yong Jia and also everyone at CSIRO agriculture Flagship at Waite Campus for their help.

I would like to thank Dr. Natoyia Lloyd from the Australian Wine Research Institute for conducting the UHPLC-MS/MS to measure acid concentration in berry samples and Dr. Olena Kravchuk for providing statistical analysis to determine the error of sampling subsets of berries.

I would like to thank Angela Lush from Lush Logic for help with spelling and editing of the thesis manuscript.

I would like to thank the Australian Grape and Wine Authority for providing me with a top-up and laboratory materials scholarship, the University of Adelaide for providing me with the C.J Everard Scholarship and CSIRO Agriculture Flagship for providing the space and materials so I could conduct my research.

I would like to thank my family and my partner Eric for providing me with support during my PhD.

Finally I would like to thank my Supervisors Dr Mark Thomas and Associate Professor Christopher Ford for their guidance and help with my research and thesis writing.

List of Abbreviations

AGRF	Australian Genome Research Facility
AWRI	Australian Wine Research Institute
EST	Expressed Sequence Tag
FPLC	Fast Protein Liquid Chromatography
GA	Gibberellins
GBS	Genotyping By Sequencing
GLM	General Linear Model
GUI	Graphical User Interface
GWAS	Genome Wide Association Study
HPLC	High Performance Liquid Chromatography
ICP	Inductively Coupled Plasma Mass Spectrometry
IPTG	Isopropyl β-D-1-thiogalactopyranoside
5KGA	5-keto-D-gluconic acid
L-IDH	L-idonate Dehydrogenase
LOD	Log Of Odds
MLM	Mixed Linear Models
MS	Malate Synthase
NAD-cyMDH	cytoplasmic NAD dependent malate dehydrogenase
NAD-mMDH	Mitochondrial NAD Dependent Malate Dehydrogenase
NADP-ME	NADP dependent malic enzyme
NCBI	The National Center for Biotechnology Information
NGS	Next Generation Sequencing
NMR	Nuclear Magnetic Resonance
PCA	Principal Component Analysis
PEP	Phosphoenolpyruvate
PEPC	Phosphoenolpyruvate Carboxylase
PEPCK	Phosphoenolpyruvate Carboxykinase
QTL	Quantitative Trait Loci
qRT-PCR	Quantitative Reverse Transcription PCR
RT-PCR	Reverse Transcription PCR
SARDI	South Australian Research and Development Institute
SDS-PAGE	Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis
SE	Standard Error
SNP	Single Nucleotide Polymorphism

TCA	Tricarboxylic Acid
UHPLC-MS/MS	Ultra-High Performance Liquid Chromatography Mass Spectrometry