Deciphering the genetic and metabolic basis of yeast aroma properties

Matthias Eder

THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE MONTPELLIER SUPAGRO

En Sciences pour l'Ingénieur – Biotechnologie et Microbiologie

École doctorale GAIA – Biodiversité, Agriculture, Alimentation, Environnement, Terre, Eau
Portée par l'Université de Montpellier

Unité Mixte de Recherches Sciences pour l'Œnologie

Deciphering the genetic and metabolic bases of yeast aroma properties

Présentée par Matthias EDER

Le 20.12.2017

Sous la direction de Sylvie DEQUIN

Devant le jury composé de

Elke NEVOIGT, Associate Professor, Jacobs University, Allemagne	Rapporteur
Ramon GONZALEZ, Research Professor, CSIC, Espagne	Rapporteur
Sophie LANDAUD, Professeur, AgroParisTech, Paris	Examinatrice
Bruno BLONDIN, Professeur, SupAgro, Montpellier	Président du jury, Examinateur
Sylvie DEQUIN, Directrice de recherche, INRA, Montpellier	Directrice de thèse
Jean-Luc LEGRAS, Ingénieur de recherche, INRA, Montpellier	Encadrant

Thèse confidentielle

UNIVERSITÉ DE MONTPELLIER

Montpellier SupAgro

Acknowledgments

This thesis was carried out in the mixed research unit Sciences for Oenology (UMR 1083 SPO) of INRA, SupAgro, Montpellier. The research project was part of the YEASTCELL initial training network (ITN) within the EU Marie Curie Actions.

First of all, I thank Jean-Marie Sablayrolles, Director of UMR SPO, for giving me the opportunity to join this institute. My biggest gratitude goes to my thesis supervisor Sylvie Dequin for her support and advice in the development of this work and also for her critical analyses and scientific exchanges. I am gratefully indebted for her very valuable contributions to every part of my scientific work. She constantly allowed this thesis to be my own, but steered it in the right direction whenever she thought it to be necessary.

I would also like to express my thankfulness to the members of my thesis jury, Elke Nevoigt, Ramon Gonzales, Sophie Landaud and Bruno Blondin, who agreed to evaluate my scientific work and to travel to the thesis defense close before Christmas.

I thank the organizers of YEASTCELL, John Morrissey and Francesca Doonan, for the countless efforts they put into setting up and managing the network.

In the course of the YEASTCELL ITN, two secondments were part of my thesis project. I would like to thank my supervisors during these secondments, Jean-Marc Daran at the group for Industrial Microbiology of the TU Delft, The Netherlands, and Jan Maarten Geertman with Niels Kuijpers at the Heineken Brewery in Zoeterwoude, The Netherlands.

Furthermore, I want to thank all PIs of YEASTCELL for their manifold contributions and all PhD students for their support and the great time we spent within and outside the ITN; it has been a matchless group!

Indeed, I also have to thank my former colleague Manuel Wittchen, who drew my attention to the YEASTCELL ITN and who encouraged me to apply at it. Without his intervention I would not have experienced this great adventure of the past years.

My appreciation goes to all members of UMR SPO, who helped and supported me in various ways during my work. Knowing that this list is incomplete, I want to thank in particular Audrey Bloem, Nicolas Bouvier, Pascale Brial, Claire Brice, Cécille Cadoux, Virginie Galeote, Stéphane Guezenec, Sandrine Mallet, Thibault Nidelet, Jessica Noble, Marc Perez, Christian Picou, Martine Pradal, Isabelle Sanchez and Peggy Rigou. I furthermore thank Virginie Bouckenooghe and Remy Schneider of Nyseos, Montpellier, for their scientific input to my thesis.

I have to address a big thank to Carole Camarasa and Jean-Luc Legras for their manifold support in conducting the experimental and scientific work. I must say that I especially learned to value the critical spirit of Jean-Luc. Without his passionate participation and input, this thesis would not have been successfully completed.

I also have to thank all PhD students of UMR SPO, Amandine, Camille, Elisa, Guillaume, Noemie, Souhir, Stéphanie, Thomas and Pauline. You were a great support and an excellent source of entertainment ;) I would like to thank Camille in particular for his willingness to help me with everything related to the pitfalls of the French language or the twisty French bureaucracy. Above all, my thank goes to my brother in spirit David (Dude …or Sweet? I don't know). All work and no play would have made Matthias a dull boy. Luckily, he avoided that in creative ways, including pizza, drinks and computer games.

Last but not least, I must express my very profound gratitude to my parents Edith and Johannes, and to my sister Johanna. Throughout all the years of my studies, they provided me with unfailing emotional (and financial) support and continuous encouragement. Above all, I thank my girlfriend and partner Anna. No matter what, she always made me know that I can count on her unconditional love and support. Her manifold help throughout the process of researching and writing this work made my thesis an accomplishment that could not have been possible without her. Thank you!

The research project behind this thesis received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007-2013/ under REA grant agreement n° 606795, project acronym: YEASTCELL.

Table of Content

Abbrevations

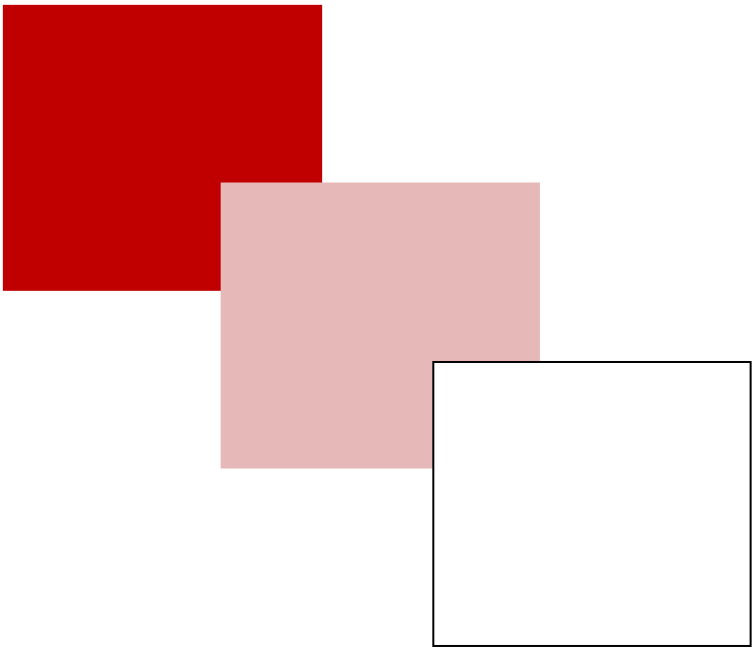
3MH	3-Mercaptohexanol
3MHA	3-Mercaptohexyl acetate
4MMP	4-Mercapto-4-methylpentan-2-one
ATP	Adenosine triphosphate
CBM	Constraint-based models
CCM	Central carbon metabolism
CO_2	Carbon dioxide
CoA	Coenzyme A
DMAPP	Dimethylallyl pyrophosphate
DMS	Dimethyl sulfide
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
FAS	Fatty acid synthetase
FBA	Flux balance analysis
FBPase	Fructose-1,6-bisphosphatase
FPP	Farnesyl pyrophosphate
f-QTL	Flux quantitative trait locus
GC-MS	Gas chromatography-mass spectrometry
GMO	Genetically modified organism
GPP	Geranyl pyrophosphate
hph^r	Hygromycin B resistance gene
H_2S	Hydrogen sulfide
HPLC	High-performance liquid chromatography
IPP	Isopentenyl pyrophosphate
LOD	Logarithm of odds
MFA	Metabolic flux analysis
NAD^+	Oxidized nicotinamid adenine dinucleotide
NADH	Reduced nicotinamid adenine dinucleotide
PCR	Polymerase chain reaction
PDH	Pyruvate dehydrogenase
pDMS	DMS potential
PPP	Pentose phosphate pathway
QTL	Quantitative trait locus
RHA	Reciprocal hemizygosity analysis
RNA	Ribonucleic acid
S. cerevisiae	*Saccharomyces cerevisiae*
SGD	*Saccharomyces* genome database
SIM	Selected ion monitoring
SM	Synthetic must
SMM	*S*-Methylmethionine
SNP	Single nucleotide polymorphisms
SPME	Solid phase micro extration
SO_2	Sulfur dioxide
TCA	Tricarboxylic acid
VA	Volatile acidity
YPD	Yeast extract peptone dextrose

List of Figures

List of Tables

List of Additional Files

Résumé Français



Les composés volatils jouent un rôle important dans la perception de la qualité du vin. Sans ces composés, le vin aurait peu de caractéristiques organoleptiques et ne serait pas apprécié de la même façon. Plus d'un millier de composés différents participent à l'arôme du vin, et interviennent dans des équilibres complexes pour participer aux principales perceptions gustatives et olfactives.

Les consommateurs partagent certaines attentes sur la flaveur du vin, mais leurs préférences peuvent changer au fil du temps. Récemment, le développement de vins fruités est devenu une priorité au niveau commercial. Dans un contexte visant à améliorer la santé publique, les vins allégés et réduits en alcool sont également plus populaires. Ceci est en contradiction avec les changements causés par le réchauffement climatique. La hausse des températures dans les régions viticoles entraîne une augmentation de la teneur en sucre des raisins et, par conséquent, des niveaux élevés d'alcool dans les vins finis. Le commerce du vin est un marché international important. Comme les vignes ne peuvent être cultivées que sous certaines latitudes, les régions productrices sont également impliquées dans l'export, dans un domaine devenant de plus en plus concurrentiel. Cela amène les producteurs de vin à constamment s'efforcer d'améliorer leurs produits afin de répondre aux attentes des consommateurs.

Outre les pratiques viticoles et de vinification, la levure a un impact significatif sur la qualité organoleptique des vins. Au cours de la fermentation, elle modifie et approfondit le profil aromatique en produisant des composés volatils à partir des nutriments présents dans le moût, en transformant les composés aromatiques apportés par les raisins, ou en libérant des molécules aromatiques à partir de molécules précurseurs. Parmi les composés d'arômes, les plus abondants, parmi ceux produits *de novo* par les levures, sont les alcools supérieurs, les acides gras à courte chaine, les acides gras à moyenne chaîne et leurs esters d'éthyle et d'acétate. En fonction du composé et de sa concentration, ces molécules peuvent produire des notes aromatiques agréables, comme le fruité ou fleuri par exemple, ou désagréables, telles que les odeurs de rance ou de solvant. Leur production est étroitement liée aux propriétés métaboliques des levures et est guidée à la fois par le métabolisme du carbone et de l'azote. Les alcools supérieurs et les acides à courte chaine peuvent dériver de la dégradation des acides aminés, mais aussi des intermédiaires du métabolisme carboné central. La production d'acides gras à moyenne chaîne repose sur l'acétyl-CoA qui est généré par le métabolisme carboné central. Les esters d'acétate et d'éthyle de tous ces composés sont formés par voie enzymatique, reposant sur des acétylations ou acylations. La levure a également un impact sur les arômes dérivés du raisin, soit par métabolisation des précurseurs aromatiques, soit par altération des composés d'arômes du raisin. Les terpénols, par exemple, sont synthétisés et apportés par les raisins. Selon le composé, ils produisent des arômes allant du pin à l'agrume et au fleuri. Des enzymes de levure catalysant des conversions entre différents terpénols pendant la fermentation ont été identifiées. De plus, à travers son activité fermentaire, la levure peut influencer la formation d'arômes post-fermentaires au cours du vieillissement du vin. Un exemple notable est le sulfure de diméthyle (DMS), qui peut conférer des notes de truffe et d'olives et a également un rôle d'exhausteur de l'arôme fruité du vin.

Les précurseurs du DMS (pDMS) sont présents dans les raisins et peuvent être métabolisés par les levures en DMS, ce dernier étant cependant éliminé en partie lors de la fermentation, du fait d'un entraînement par le CO_2 produit. Les précurseurs restant présents dans le vin jeune pourront être ensuite transformés chimiquement lors de la maturation.

La production et le métabolisme des composés aromatiques reposant sur les propriétés métaboliques de la levure, les bases génomiques sous-jacentes deviennent importantes à prendre en compte. L'évolution et la sélection humaine ont conduit à une population diversifiée de souches de levures *Saccharomyces cerevisiae* à forte variation génomique. Elucider les liens entre les variations génomiques et propriétés métaboliques est primordial pour pouvoir comprendre les mécanismes moléculaires responsables de l'orientation du métabolisme et pour pouvoir mieux exploiter, à terme, cette diversité au sein de l'espèce. Au final, cette connaissance peut permettre d'améliorer l'arôme et la qualité du vin et d'adapter les vins aux demandes des consommateurs.

Les bases génétiques des traits complexes, gouvernés par plusieurs gènes et donc quantitatifs, peuvent être explorées par des approches de cartographie QTL. Cette stratégie est basée sur la présence de marqueurs génétiques définis reposant sur des différences génotypiques, qui sont mis en lien avec des variations phénotypiques. Tout d'abord appliquée à la sélection végétale, cette approche a été adaptée à *S. cerevisiae* et est devenue un outil puissant pour élucider les bases génétiques de nombreux traits d'intérêt industriels, avec un nombre croissant d'études publiées durant les 10 dernières années y compris dans le domaine œnologique. Les approches de cartographie QTL sont continuellement améliorées afin d'augmenter la puissance des analyses et d'étendre la méthodologie à de nouveaux traits d'intérêt.

Le principal objectif de ce travail est d'identifier les bases génomiques et métaboliques de la formation d'arômes par *S. cerevisiae* pendant la fermentation du vin. Nous nous sommes intéressés aux esters, alcools supérieurs et acides organiques produits *de novo* par la levure, ainsi qu'aux terpénols et au DMS, formés à partir de précurseurs du moût de raisin. A cette fin, un croisement a été effectué entre deux souches de levures de vin sélectionnées en raison de leurs besoins différents en azote pendant la fermentation, paramètre considéré comme une indication de différences potentielles dans la formation des arômes. 130 ségrégants de génération F2 ont été obtenus et génotypés par séquençage complet de leur génome. Les données de génotypage ont été utilisées pour identifier l'ensemble des SNP différenciant les souches parentales et pour établir une carte de marqueurs sélectionnés. Les ségrégants ont été phénotypés individuellement pendant la fermentation du vin dans du moût de raisin synthétique, en utilisant des fermenteurs de 300 mL. Les concentrations en métabolites extracellulaires ont été mesurées en utilisant la chromatographie liquide à haute performance et la chromatographie en phase gazeuse associée à la spectrométrie de masse. Une analyse de liaison combinant les données de génotypage et de phénotypage a été réalisée pour identifier les variations alléliques ayant un impact sur la production d'arômes.

Dans la première partie de la thèse, les bases de l'analyse QTL ont été établies. Une cartographie simple-QTL a été réalisée pour trouver des loci dans le génome de la levure ayant une influence sur les paramètres généraux de fermentation, la production de métabolites extracellulaires principaux et la formation de métabolites volatils en fin de fermentation. L'utilisation d'un nombre relativement élevé de ségrégants (130) avec un taux de recombinaison augmenté (génération F2) a permis d'obtenir une puissance analytique élevée. De ce fait, nous avons pu appliquer plusieurs stratégies de cartographie QTL et notamment rechercher des interactions entre QTL. Ces analyses nous ont permis de détecter 65 QTL ayant une influence sur la formation de métabolites volatils et sur la production de métabolites extracellulaires et paramètres généraux de fermentation. Nous avons confirmé que la cartographie QTL multiple offre la possibilité de détecter des loci supplémentaires, en particulier mineurs. Des QTL interactifs ont pu être détectés pour trois caractères évalués : la formation de lactate d'éthyle, d'octanoate d'éthyle et de propanol. En validant les gènes candidats dans les QTL trouvés par analyse d'hemizygotie réciproque, nous avons démontré l'implication et l'impact de 13 variants alléliques au sein de 9 QTLs. Parmi ces allèles validés, 5 (*AGP1, ALP1, FAS1, ILV6* et *LEU9*) ont des rôles bien décrits dans les voies métaboliques conduisant à la formation d'arômes fermentaires de levure. Leur contribution à la production de volatils a été confirmée et des allèles expliquant une variation des caractères observés entre les souches parentales ont été identifiés et caractérisés. En outre, le caractère fructophile précédemment décrit de l'allèle *HXT3* de la souche MTF2621 a été confirmé dans cette étude. Pour les 7 autres gènes validés (*AGP2, IXR1, MAE1, NRG1, RGS2, RGT1* et *SIR2*), des contributions à la formation d'arômes fermentaires ont été révélées, celles-ci n'étant pas connues auparavant. En outre, une grande partie des gènes candidats validés sont des régulateurs transcriptionnels (*AGP2, IXR1, NRG1, RGS2, RGT1,* et *SIR2*), ce qui souligne le rôle de la régulation dans la formation des arômes fermentaires, assez peu étudié jusqu'à maintenant. Pour tous les gènes validés, des variants alléliques ayant des impacts différents ont été décrits. Des variants des transporteurs d'acides aminés Agp1 et Alp1, affectent la formation des produits de la voie d'Ehrlich par des différences d'expression ou d'affinité. L'implication du métabolisme de l'azote dans la formation d'arômes fermentaires est appuyée par l'identification de deux autres enzymes impliquées dans les voies de synthèse des acides aminés, Ilv6 et Leu9. Le fait que la production d'arômes fermentaires soit également liée au métabolisme central carboné est souligné par la validation des gènes cibles *MAE1* et *FAS1*. Mae1 catalyse la réaction du malate au pyruvate. Nos travaux suggèrent qu'un taux de conversion enzymatique plus élevé conféré par un allèle variant augmenterait la production de substances volatiles dérivées du pyruvate. Fas1 catalyse les étapes de la synthèse des acides gras à partir de malonyl-CoA, qui est formé par carboxylation de l'acétyl-CoA, et un allèle variant permettant une synthèse réduite des acides gras a été caractérisé.

La production d'arômes fermentaires de levure étant étroitement liée aux flux du métabolisme carboné central, une meilleure compréhension de l'impact des variations génomiques sur les flux intracellulaires est cruciale. Dans la deuxième partie de la thèse, nous avons cherché à évaluer la possibilité d'étendre la méthodologie de cartographie QTL à la

détection de loci influençant les flux métaboliques intracellulaires, appelés flux-QTL (f-QTL). Les concentrations en principaux métabolites extracellulaires mesurés pendant la phase de croissance exponentielle de la fermentation ont été utilisées pour estimer les flux intracellulaires du métabolisme central carboné de la levure en utilisant un modèle stœchiométrique à base de contraintes précédemment développé au laboratoire. Cela a conduit à l'intégration de traits quantifiables, par ailleurs indépendants. Les flux métaboliques estimés ont été utilisés comme données phénotypiques pour détecter les f-QTL. 4 QTLs ayant une influence sur 4 voies métaboliques principales et différents flux de transport de métabolites ont été détectés. Ces QTL n'ont pas pu être trouvés par analyse de liaison en considérant la concentration des métabolites extracellulaires seuls. Ceci démontre que des données phénotypiques modélisées, obtenues à partir d'un jeu de données expérimentales de taille limitée, peuvent permettre la détection de QTLs. La robustesse de l'approche a été confirmée par la validation de deux gènes cibles dans deux des QTL détectés, *PDB1* et *VID30*. Les variants alléliques de *PDB1* provoquent des différences dans les flux de synthèse du glycérol qui peuvent être liées aux déséquilibres redox engendrés par une modification de la conversion du pyruvate. Les variants de *VID30* ont une influence sur les flux de la glycolyse, de la synthèse de l'éthanol et du cycle de Krebs, pouvant être liée à des différences dans la dégradation des enzymes impliquées dans les réactions du métabolisme central du carbone. Ces résultats apportent un nouveau regard sur l'implication de certains gènes et de leurs variants sur le réseau métabolique, qui pourrait avoir des répercussions sur les propriétés aromatiques des levures. L'extension de l'approche QTL pour la détection des f-QTL pourrait être appliquée à l'élucidation d'autres traits, comme par exemple l'utilisation efficace de substrats alternatifs (glycérol, xylose …) par la levure ou la production de métabolites pour d'autres applications biotechnologiques.

Le troisième chapitre de la thèse applique la stratégie de cartographie QTL à la détection des régions génomiques impliquées dans la capacité de la levure à influencer le niveau de la S-méthylméthionine (SMM), le précurseur du DMS dérivé du raisin. A cette fin, la SMM a été ajoutée au moût synthétique et les 130 ségrégants ont été phénotypés pour leur capacité à éliminer la SMM du milieu en fin de fermentation. Ces résultats ont été utilisés comme données phénotypiques pour l'analyse QTL. Cela a conduit à la détection d'un QTL dans le génome de la levure ayant une forte influence sur le métabolisme de la SMM. Trois gènes candidats de cette région ont été évalués et nous avons pu démontrer que deux d'entre eux, *MMP1* et *YLL058W*, influencent le métabolisme de la SMM. Le gène du transporteur de la SMM, *MMP1*, a été identifié comme étant responsable de la majeure partie de la variation du phénotype. Ainsi, un allèle parental de *MMP1* exprime une capacité fortement diminuée à éliminer la SMM du milieu. En effectuant un échange allélique du gène entre les souches parentales à l'aide de la technique CRISPR-Cas9, nous avons identifié la mutation ponctuelle (SNP) à l'origine de ce caractère. Cette mutation génère un codon STOP dans le gène *MMP1*, ce qui conduit à l'expression d'un transporteur de la SMM tronqué, entrainant une meilleure conservation de la SMM qui est moins utilisé par la levure. En étudiant la distribution des variants de *MMP1* de 85 souches de levure d'origine différente, nous avons pu démontrer que

le variant Mmp1 codant pour une protéine tronquée est principalement trouvé dans les souches de levure de voile, indiquant ainsi un lien avec cette origine. La présence de ce SNP dans quelques souches de vin pourrait être attribuée à des transferts de gènes entre les souches de levure de voile et les levures de vin.

Dans le quatrième chapitre de la thèse, l'approche de cartographie QTL établie a été appliquée pour détecter les régions dans le génome qui influencent le métabolisme de la levure des terpénols dérivés du raisin. A cette fin, le géraniol, le précurseur de terpénols, a été ajouté au moût synthétique et tous les descendants ont été phénotypés pour leur capacité à transformer le géraniol pendant la fermentation œnologique. L'analyse de liaison a révélé 21 QTL dans le génome qui influencent la concentration de plusieurs terpénols en même temps ou la formation de composés uniques tels que le linalol ou l'α-terpinéol. Parmi les 21 QTL détectés, 13 régions correspondent à des QTL influençant d'autres traits déjà identifiés dans les chapitres précédents de la thèse. Ces QTL influencent les paramètres de fermentation ou la production des principaux métabolites, et également pour la plupart la formation d'arômes fermentaires. Cela indique une connexion potentielle entre le métabolisme des terpénols et la production de composés volatils, par exemple un mécanisme identique entraînant l'acétylation des alcools supérieurs formés et des terpénols présents.

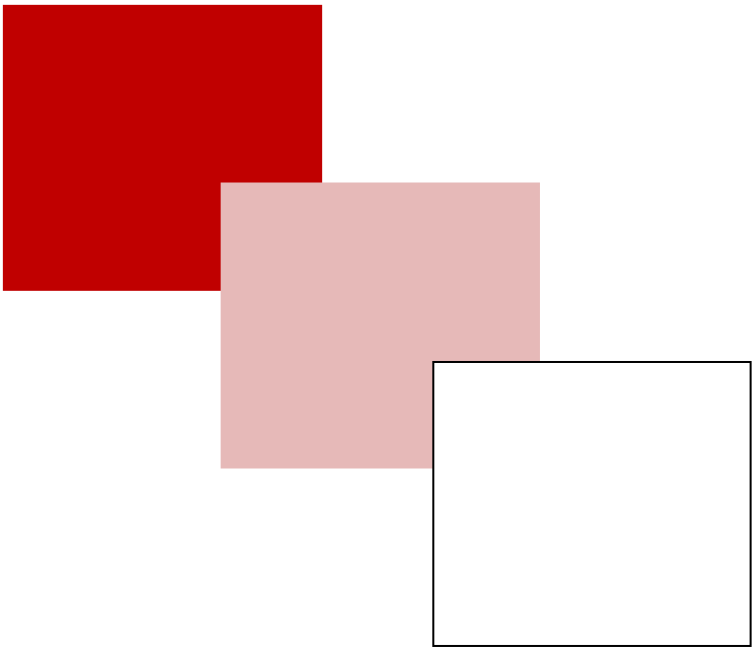
En conclusion, les résultats obtenus confirment et soulignent le rôle de la diversité génétique dans la formation des arômes fermentaires. La complexité des bases génomiques sous-jacentes est démontrée par le nombre élevé de QTL détectés et l'indication des interactions QTL. Nos travaux démontrent une fois de plus le lien étroit entre la formation d'arômes fermentaires et le métabolisme de l'azote et du carbone. En plus de confirmer l'intérêt de la cartographie QTL pour évaluer les traits œnologiques, les résultats de cette étude soulignent sa pertinence pour déchiffrer des traits complexes. En utilisant la cartographie QTL, il a été possible de valider des gènes cibles ayant une action globale sur des traits définis et d'identifier des gènes influençant les flux métaboliques (f-QTL).

Le secteur industriel perçoit parfois les universités et centres de recherche publics comme producteurs de connaissances théoriques peu pertinentes sur le plan économique. Les connaissances générées dans cette étude offrent des possibilités d'application à court terme. Bien que les allèles variants n'aient pas encore été introduits dans d'autres souches œnologiques ayant un fond génétique différent, l'identification d'allèles d'intérêt fournit clairement de nouvelles cibles pour construire des souches des levures de vin aux propriétés aromatiques améliorées pouvant être commercialisées. Cela peut inclure la construction de souches avec une production d'arômes fermentaires globalement plus élevée. Cependant, cette stratégie est susceptible de générer peu de valeur ajoutée, car des souches produisant différents niveaux d'arômes fermentaires existent déjà sur le marché. Une approche plus prometteuse serait de cibler certaines caractéristiques aromatiques. Par exemple, une surproduction d'acétate de 2-phényléthyle, qui apporte des notes florales et de rose aux boissons alcoolisées pourrait être envisagée. Des variants alléliques de Leu9 et Rgs2 affectant

la production d'acétate de 2-phényléthyle et l'impact de Agp1 et de Mae1 sur la formation de l'alcool supérieur correspondant, le 2-phényléthanol, ont pu être montrés. Une autre possibilité serait l'optimisation des souches pour le métabolisme des terpénols. Comme indiqué précédemment, les différents terpénols apportent différentes notes aromatiques au vin et sont notamment à la base de la typicité des muscats. Nous avons montré que les ségrégants différaient dans leur profil de terpénols et que des QTLs ayant un impact sur ce profil pouvaient être détectés. Cependant, il reste à valider les gènes candidats, ce qui permettra d'identifier les cibles d'amélioration des souches les plus pertinentes. Une troisième possibilité serait la sélection ou la construction de souches ayant des propriétés de préservation de la SMM. L'identification du variant tronqué de Mmp1 qui conduit à la préservation de la SMM dans le moût offre des perspectives prometteuses dans ce contexte. Ce variant pourrait être recherché dans des souches d'origine différente, permettant ainsi la sélection de souches d'intérêt. Cette information peut également conduire à l'insertion de l'allèle d'intérêt dans des souches déjà commercialisées par hybridation et sélection assistée par marqueur. Nous avons en outre montré que les variants alléliques de *PDB1* et *VID30* présentent des différences de distribution de flux entre la synthèse du glycérol et de l'éthanol. Bien que la variation causée soit relativement faible, cet allèle pourrait être utile, en appui à une autre stratégie, pour favoriser la surproduction de glycerol dans l'optique de diminuer le rendement en alcool.

Les gènes cibles et allèles les plus performants identifiés pourraient également être utilisées pour d'autres applications, incluant la synthèse de produits de chimie fine, par exemple d'alcools supérieurs ou d'acides organiques, pour la production de biocarburants ou de bioplastiques. Ces connaissances peuvent également s peuvent être utilisées pour caractériser d'autres variants alléliques dans des souches de levure d'origines diverses, et prédire ainsi leur phénotype.

General Introduction



Volatile aroma compounds play an outstanding role for the quality and perception of wine. They can emerge the liquid in the wine glass to become detectable by the human olfactory receptors. These volatile compounds give wine its organoleptic character and without it, wine would not be enjoyed in the same way it is.

Besides viticultural and winemaking practices, yeast, such as *Saccharomyces cerevisiae*, has a substantial influence on wine flavor. During fermentation, yeast influences the aromatic profile of wine by producing volatile compounds from nutrients in grape must, by the modification of aroma compound precursors found in the grapes or by releasing potent aroma molecules from odor-less grape precursor molecules. The production of fermentative aroma molecules is closely linked to yeast carbon and nitrogen metabolism. Yeast's ability to affects grape derived flavors, either by metabolism of odor-less precursors to aromas or by modification of grape aroma compounds, was also demonstrated to be dependent on yeast enzymatic properties. In addition, strain dependent influence of yeast on the development of post-fermentative aromas during wine aging was shown.

As the formation of numerous aroma compounds relies on yeast metabolic activities, it is therefore governed by the underlying genomic bases. Human use and selection of yeast for winemaking has led to a wide yeast population containing high genomic variation. In recent years, research has focused on the understanding of wine aroma formation as well as the genetic and phenotypic differences between yeast strains. However, information about the impact of genotypic diversity on phenotypic trait determination is still scarce. Understanding the links between genomic and metabolic properties and measuring the impact of genomic variation on metabolic traits is required for optimally exploiting yeast diversity in various applications.

The main aim of this work is to identify the genomic and metabolic bases for the formation of volatile aroma compounds by *S. cerevisiae* during wine fermentation. This includes the study of the *de novo* production of different aroma compounds such as esters, higher alcohols and organic acids as well as the investigation of compounds that are formed from precursors in the must like terpenols and dimethyl sulfide. The study will assess the impact of genomic variation among yeast strains on wine aroma trait determination to evaluate the hidden potential of genetic resources from divergent strains for the optimization of industrially used wine yeasts.

The results of the experimental work are divided in four chapters. Chapter one describes the assessment of two *S. cerevisiae* wine strains regarding their genomic difference and their ability to produce fermentative aromas such as higher alcohols, esters and acids. Furthermore, the strains were evaluated for general fermentation parameters and their production of main metabolites. Differences in fermentation parameters, main metabolite production and fermentative aroma formation were subsequently linked to genomic variations. Linkage analysis was extended to search for genomic regions that showed interaction among each other regarding their influence on fermentative aroma formation. Allelic gene variants

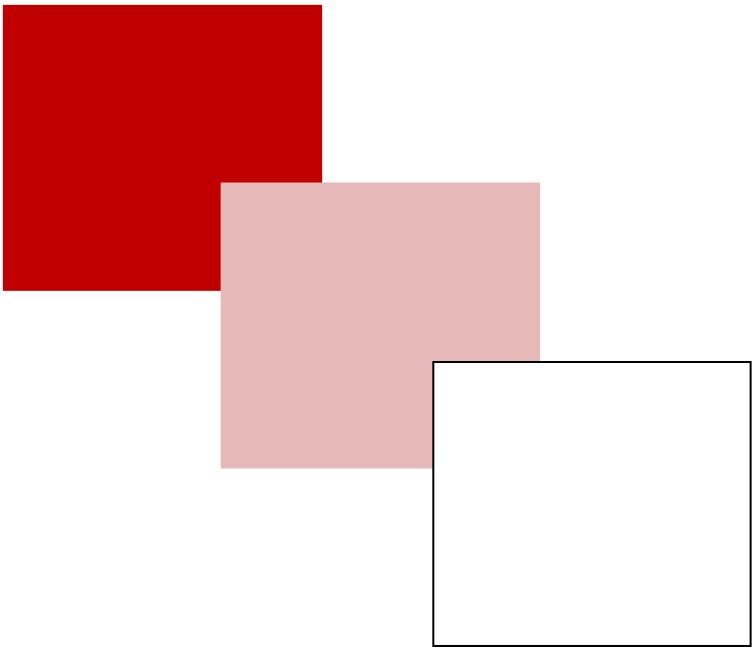
responsible for the phenotypic differences were identified and characterized within these regions.

As the production of yeast fermentative aroma is tightly associated to yeast metabolism, especially the central carbon metabolism, better understanding of the impact of genomic variation on intracellular fluxes is crucial. The second chapter of this thesis evaluates how genetic differences of two wine yeast strains can be associated to differences in intracellular carbon flux distributions. This was achieved by using a constraint-based model that takes determined concentrations of extracellular metabolites to estimate fluxes of the yeast central carbon metabolism. Subsequently, it was assessed if the obtained information can be used to also link differences in metabolic flux distributions to yeast genomic properties and to detect allelic variants that account for a diverse shaping of the metabolic profile.

The third and fourth chapter of the thesis extend the investigation to yeast's influence on the composition of grape derived aroma contributors. In the third chapter, the evaluated strains were characterized for their ability to preserve *S*-methylmethionine, an odor-less aroma precursor that is chemically transformed to the flavor compound dimethyl sulfide during wine maturation. Genomic bases behind strain differences were identified and linked to the strain's origin. In the fourth chapter, the strains were assessed for their different impact on terpenols, a class of potent grape-derived volatiles that can be altered by yeast metabolism. Phenotypic differences were again linked to the genomic properties of the strains and genomic regions behind terpenol metabolism were identified. These regions were compared to the detected genomic regions influencing other traits, such as fermentative aroma formation, in order to evaluate potential connections within the shaping of wine aromatic profile by yeast.

Chapter one was submitted to the journal "BMC Genomics" on 18.09.2017 and is currently under review. The chapters two and three are completed for submission. For chapter two, this process will be completed until the defense of the thesis. A patentability study is in progress on the results of chapter 3. Chapter four presents preliminary results on the genetic basis of the transformation of terpenols compounds during wine fermentation. Additional experimental work to validate the role of the identified alleles will be necessary before publication.

Introduction



1 The role of fermentation

Fermentation (from the Latin word *fervere*, which means “to boil”) plays a central role for a large share of foods that we consume every day. Fermentation is a natural process and plays a non-negligible role in nature. Usually, when a fruit is ripe it falls on the soil and the sugars that it contains are partly metabolized by microorganisms, particularly yeasts of different genera that are naturally present on the surface of the fruit. Alongside with the production of alcohol, fermentation leads to the production and release of volatile molecules that can be perceived as odors and flavors. In nature, these aromas produced by yeast may signal suited properties of the rotting fruit and may attract (or repel) insects and animals on the search for food. This can lead to the dispersal of both, yeast and fruit seeds.

1.1 Fermented Foods and beverages

It is estimated that a third of the food we eat is produced or refined by fermentation. Bread, cheese, yogurt, chocolate, vinegar, beer and wine are just examples of this non-exhaustive list of foods daily consumed worldwide. The use of fermentation by human is an inheritance. Archeology revealed that many civilizations used this process, from the Chinese (cabbage, tea) to the Mesopotamian, and Egyptian (bread, beer) and Romans (wine). But also the Aztec (cocoa) and Vikings (fermented milk) knew food preparation techniques that relied on fermentation.

Fermented foods and beverages have many advantages. The fermentation can allow the food to be more digestible. Contrary to apparent assumptions, the nutritional value is increased as vitamins or antioxidants are accumulated or release by the involved microorganisms (Steinkraus, 1994). Furthermore, the benefits of produced ethanol were probably one reason for early civilizations to develop fermentation techniques. Together with a reduction of pH during fermentation, ethanol promotes the conservation of food, thus making it safer to consume. In addition, it improves digestibility and on top of that acts as a euphorisant (Alba-Lois and Segal-Kischinevzky, 2010).

Fermented beverages play an important role in the diet of past and present mankind. Regardless the country of origin, their production process is fairly similar. A nutritious juice or liquid suspension is made from carbohydrate rich raw materials like cereals, fruits, milk or sap, which is then covered to protect it from oxidation. Subsequently, microorganisms present on the material or added carry out the fermentation. From this simple recipe, man was able to create a wide variety of alcoholic beverages. While bacteria and molds play the dominant role for the fermentation of foods, e.g. *Lactobacillus, Streptococcus* and *Penicillium* genera for the fermentation of dairy products, most fermented beverages are produced by yeast, particularly of the genus *Saccharomyces* and most often of the species *cerevisiae*, although bacteria might be additionally involved. In most cases, these yeasts are indigenously present on the raw

material, making an addition not necessarily required and leading to the involvement of a community of different species or strains in the fermentation.

1.2 History of fermented foods and beverages

It is estimated that the history of fermented foods and beverages began in Mesopotamia. Archaeologists have found indication of wine production in Iran and Egypt dating back to 6000 and 3000 BC (Cavalieri et al., 2003; McGovern et al., 1997) and more recently, an Armenian archaeological site has revealed the existence of a production unit of wine dating back to 4000 BC (Barnard et al., 2011). However, Chinese archaeological evidences suggest the production of fermented beverages as early as 7000 and 9000 BC (McGovern, 2009; McGovern et al., 2004). Investigations carried out show that sweetened rice mush was produced probably by mastication of rice. The juice could then have been fermented to an alcoholic beverage by yeasts. The more recent archaeological findings allow McGovern (2009) to push this theory further. According to him, the Neolithic revolution that began 11000 years ago would be a consequence of the use of alcoholic fermented foods, as consuming alcohol helped to survive in a hostile environment with few natural resources.

The production of fermented food and beverages began to take on bigger proportions in the early Middle Ages. At that time, small communities, usually monasteries, set up small scale production units for fermented foods and beverages such as cheese, bread, wine or beer. Towards the end of the 18th century, the industrial revolution led to the growth of larger units manufacturing these products. The progress of science, notably the work of Pasteur on yeasts in beer and wine, and the development of technologies such as artificial refrigeration and the steam engine prompted an enormous progress in the fermentation industry, especially the brewing industry, by achieving a better control over the fermentation process. On the other side, the development of these industries has made science particularly benefiting from the advances. The knowledge acquired on yeasts, for example, allowed later the use of particular and selected strains and resulted in considerable progress within this sector.

2 Yeast alcoholic fermentation

Alcoholic fermentation is a process of energy utilisation performed by yeast even under anaerobic conditions. The main reaction of this process is the bioconversion of sugars, such as glucose, fructose or galactose, into ethanol and carbon dioxide (CO_2):

$$C_6H_{12}O_6 \rightarrow 2\ CO_2 + 2\ C_2H_6O$$

One mole of sugar is converted to two moles of ethanol and two moles of CO_2. Two moles of adenosine triphosphate (ATP) are generated in the process, which the cell requires to maintain its metabolism (Figure 1).

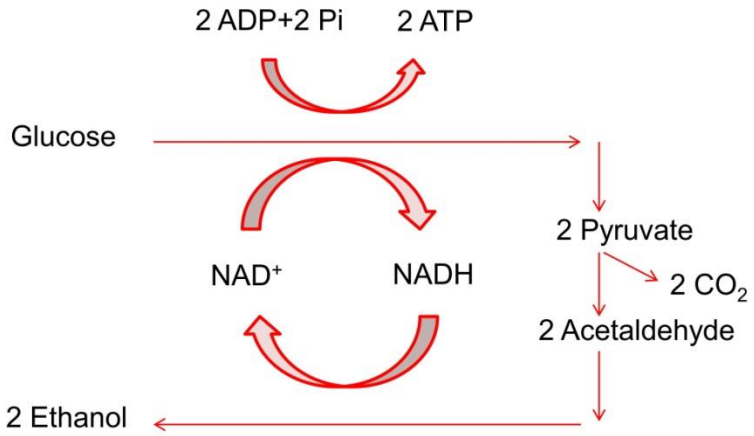


Figure 1: Simplified pathway of alcoholic fermentation by *Saccharomyces cerevisiae*.

Fermentation medium is an environment for many different competing species of microorganisms. The species with the best characteristics to survive and outcompete other opponents will proliferate and colonize the environment. The combination of the abilities of *Saccharomyces cerevisiae* to produce heat and to produce and tolerate ethanol in quantities of more than 5%, a toxic concentration for most other microorganisms, has enabled it to outgrow and eliminate competitors (Goddard, 2008). Other reasons include more efficient uptake of nitrogen sources or vitamins (Williams et al., 2015). Its genomic and metabolic properties are the basis behind the yeast's supremacy in fermentation conditions. During their evolution, *Saccharomyces* yeast and closely related *Saccharomycetaceae* have undergone a duplication of their whole genome (Conant and Wolfe, 2007). Together with additional genomic rearrangements, this has allowed *S. cerevisiae* to maintain a high metabolic activity even under anaerobiosis (Hagman et al., 2013). The drawback of alcoholic fermentation is a very low energy yield since only 2 molecules of ATP are produced per molecule of glucose, compared to 36 produced ATP molecules through respiration in the presence of oxygen. However, even under aerobic conditions *S. cerevisiae* and closely related species show the ability to produce and accumulate ethanol, which was termed Crabtree effect (De Deken, 1966; Postma et al., 1989). It has been suggested that overflow in sugar metabolism is the responsible mechanism behind the Crabtree effect and that this characteristic would have been acquired to enable rapid sugar consumption and to increase energy production rates. Further adaption would have led to the glucose-mediated repression of respiration in order to increase overflow metabolism and ethanol production (Hagman and Piškur, 2015). As *Saccharomyces* (and *Dekkera*) are also efficient in consuming ethanol, this strategy could provide the advantage to competitively dominate other yeasts by rapidly consuming sugars and producing alcohol, which is later consumed after establishing dominance: a make-accumulate-consume strategy (reviewed by Marsit and Dequin, 2015).

2.1 Oenological importance of alcoholic fermentation

Alcoholic fermentation is an essential step in the winemaking process, leading to the conversion of grape juice into wine. The sugars that are contained in grapes are converted into alcohol and at the same time many other biochemical, chemical and physicochemical processes take place. This leads to the production of other main metabolites such as glycerol and organic acids. In addition, several minor volatile compounds are formed, which are important for the sensory characteristics of wine and, in particular, give wine its vinous character (Romano et al., 2003). At the same time, grape derived aroma contributors get transformed by yeast's metabolism as well. Without the production of these substances, wine would have little organoleptic interest.

In general, grape must is characterized by an acidic pH (2.9 to 3.8) with high sugar content (150 to 260 g/L of an equimolar mix of glucose and fructose). It furthermore contains essential nutrients for yeast growth, such as nitrogen sources, lipids, vitamins and minerals. However, certain agents, such as sulfur dioxide, can act as growth inhibitors.

Unlike beer or sake, grape must does not undergo any heating stage and therefore contains a population of microorganisms deriving from grape berries and the extraction process. Grape flora is composed of different genera of yeast (*Saccharomyces, Kloeckera, Hanseniaspora, Rhodoturula, Candida, Debaryomyces, Brettanomyces,* ...), which may potentially contribute to the organoleptic characteristics of wine (Fleet, 1993; reviewed by Jolly et al., 2014). This biodiversity depends on several factors, such as the grape variety, the ripening state at harvest, antifungal treatments, the climatic conditions of the year, possible occurrence of fungal plagues and the general viticultural practices (Bokulich et al., 2014; Pretorius et al., 1999). Some of these yeast strains can enhance the aromatic composition of wine, while others have a deleterious effect on the organoleptic characteristics of the final product (i.e., *Dekkera bruxellensis*). The aromatic profile and the flavor of a wine are among its most important characteristics and in an increasingly competitive market, the modulation of fermentative aroma production to meet consumers' expectations has become one of the major challenges in winemaking.

Over the last 30 years the inoculation of must with selected yeasts, mainly the budding yeast *S. cerevisiae*, has become a widespread practice that concerns about 80% of all oenological fermentations today (Sablayrolles, 2009). The bulk implantation of one or more yeast strains selected for their technological properties such as fermentative capacity or for aromatic potential allows a better control of the fermentation process or the organoleptic characteristics of the wine.

The final flavor and aroma profile of wine is influenced by many viticultural effects and vinification methods used during the winemaking process, however, the impact given by the choice of the yeast strain plays a central role (Robinson et al., 2011). The various yeast strain marketed for oenology exhibit different abilities to produce or release aromas of different

kinds (Cadière et al., 2011; Camarasa et al., 2011; Patel and Shibamoto, 2003; Torija et al., 2003; Torrea et al., 2003; Vilanova et al., 2007).

Advances in scientific research have given winemakers the possibility to shape the aroma outcome of their wines in selecting commercially available wine yeasts by their biosynthesis of flavor active compounds and their ability to release or modify grape derived compounds (reviewed by Swiegers et al., 2005). However, a smaller share of wineries, especially traditional wine cellars, continue to use spontaneous alcoholic fermentation, as they find it grants their wines a greater complexity. Today, there is increasing interest on the contribution of non-*Saccharomyces* yeast species to wine sensorial attributes, either as wild microbiota or as mixed starters with *S. cerevisiae*.

Despite increased knowledge on the impact of *S. cerevisiae* metabolic properties on the synthesis of fermentative aroma, current information on the production and regulation of aroma biosynthesis is still incomplete. Many questions remain, notably about the influence of yeast genomic diversity on the winemaking process.

2.2 Fermentation process

During fermentation, the yeast is subjected to different stresses (high sugar content, acidic pH, nutrient limitation, increase in ethanol concentration, etc.), which require a continuous metabolic adaptation. Monitoring the alcoholic fermentation process is therefore a key step of winemaking. This involves keeping track of the consumption of sugars, which enables forecasting fermentation progression and anticipating possible problems, such as a sluggish or stuck fermentation. It can be achieved by different possible approaches:

- Monitoring the density of the fermentation medium, which correlates to the amount of present sugars (used in many cellars).
- Measuring the concentration of alcohol.
- Measuring the release of CO_2 (used in the laboratory and in an increasing number of cellars with on-line monitoring systems for fermentations, which allow precise monitoring and give access to the fermentation parameters).

Monitoring the yeast fermentation activity reveals three main phases (Figure 2).

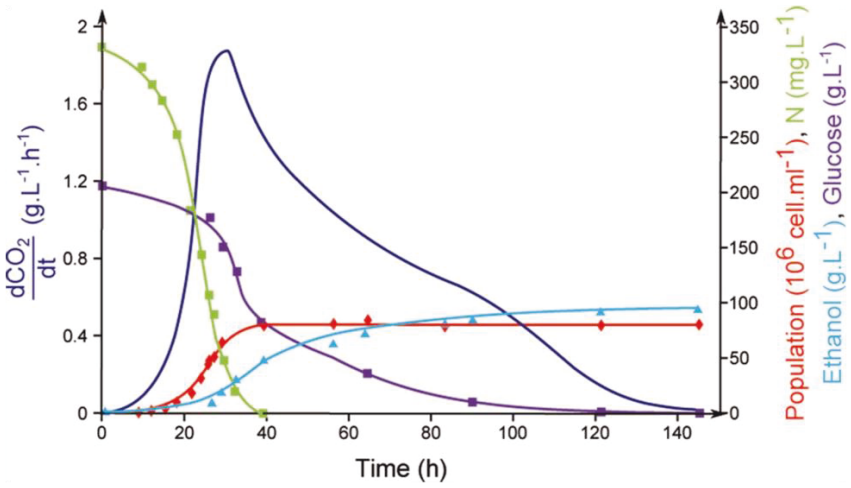


Figure 2: Representative development of main fermentation parameters during anaerobic wine fermentation by yeast. The different fermentation phases are connected to the availability of nutrients and the accumulation of metabolites. Dark blue: fermentation rate expressed as CO_2 release; light blue: ethanol concentration; red: yeast cell number; green: nitrogen concentration; purple: concentration of sugars; from Marsit and Dequin, (2015).

The **lag phase** corresponds to the adaptation of the yeast to a new environment (Pérez-Torrado et al., 2002). This phase can vary from a few hours to a few days, depending on the temperature, presence of inhibitors and the condition of the inoculated yeast cells (e.g. the preparation procedure of added active dry yeasts). During the lag phase, the yeast population does not increase, while the medium saturates with CO_2 (approximately 1.5 g/L). The metabolism of glycerol and ergosterol becomes important during the lag phase, as it leads to the accumulation of reserve molecules and cell protection factors. Ergosterol, for example, will later on protect the yeast from ethanol stress (Alexandre et al., 1994).

The **growth phase** or **exponential phase** starts when the cells are adapted to the new environment. It is characterized by an exponential cell multiplication, leading to a high release of CO_2 and a high fermentation rate. During this phase, the specific rate of CO_2 (dCO_2/dt) will reach a maximum value, corresponding to the maximum activity of the fermentation. Simultaneously, the transcriptomic activity remains stable during this period (Rossignol et al., 2003). The duration of the growth phase can last from 2 to 6 days, which is relatively short compared to the duration of the fermentation. It is highly influenced by the concentration of ammonia, amino acids and other nutrients, which are limiting factors (Ingledew and Kunkee, 1985) and by the presence of oxygen (Bisson, 1999).

The **stationary phase** in wine fermentation lasts between 3 and 20 days, during which the yeast does not multiply anymore, but retains fermentative activity. Most of the sugars (60-70%) are consumed during this phase. The cell viability remains generally higher than 80-90% and therefore the cell population maintains a maximum level, while the fermentation activity decreases progressively until most of the sugars are depleted. This decrease in activity is mainly due to increasing concentrations of ethanol and other substances produced during alcoholic fermentation, which are toxic for the cells. Furthermore, the nitrogen limitation

plays an essential inhibitory role on hexose transporters (Salmon, 1989). The stationary phase corresponds to a decrease in biosynthesis activities and a reorganization of the nitrogen metabolism. It is associated with important transcriptomic rearrangements, notably with the implementation of a stress response that is linked to the growth stop. Even though the stationary phase is considered a stress condition, the genes involved in glycolysis are maintained during this phase (Rossignol et al., 2003).

3 Wine aromas

Wine is a complex mixture and many molecules play a role in its organoleptic characteristics. Four senses are involved in defining the organoleptic quality of wine: sight, smell, taste and touch. Wine taste, perceived with the taste buds of the tongue, involves sweetness, acidity, bitterness, saltiness and the taste of umami. The mouth-feel of wine relates to the body and texture of the liquid and is influenced by factors such as the alcohol content or the presence of astringency. The general structure of a wine includes a wide range of perceptions, such as acidity, sweetness, occasional bitterness, tannins, alcohol content, palate weight and length, mouth-feel, the intensity of aroma and flavor and overall complexity (reviewed by Styger et al., 2011). These structural elements should be in balance and harmony. They are not assessed in isolation but in relationship to each other. Wine drinkers' senses are not uniformly sensitive to the subtle assortment of changing sensations. Some of the diversity in sensory perception and preferences for different wine styles among individuals and populations is cultural, some learned, some genetic (Pretorius and Høj, 2005). Furthermore, preferences are also influenced by factors such as gender and age (reviewed by Swiegers et al., 2005).

Wine aroma is related to the presence of compounds with low boiling points, which are, therefore, volatile. These compounds can escape the liquid in the wine glass to become detectable by the human nose (reviewed by Styger et al., 2011). While over 1000 aroma compounds were detected to be present in wine (Tao and Li, 2009), only a few compounds actually contribute to the sensory perception of wine flavor (reviewed by Polášková et al., 2008). Higher alcohols, acids and esters are quantitatively dominant in wine aroma and are important for the sensorial properties and the quality of wine (Stashenko et al., 1992).

In general, wine aroma can be divided into classes, depending on its origin. Varietal aroma is synthesized and contributed by the grapes, and is a distinction of grape variety. Wines made from specific grape varieties, like Gewürztraminer or Muscat, display distinctive aromas, which are induced by the corresponding variety (Duchêne et al., 2009; Guth, 1997). Prefermentative aroma originates during grape processing processes. Fermentative aroma is produced by yeast and bacteria during alcoholic and malolactic fermentations. Postfermentative aroma develops due to transformations that occur during conservation and aging of wine (Vilanova et al., 2010). Modern advances in scientific research give winemakers tools to shape their wines toward predetermined aroma outcomes. Today, wine yeast and

bacteria can be selected to optimally biosynthesize flavor-active compounds and to release or modify grape-derived flavor compounds without affecting the general fermentation performance. The most dominant aroma molecules are listed in Table 1, together with their perception threshold and concentration usually found in wine.

Table 1: Most dominant aroma and flavor compounds commonly found in wine, with the concentration range, their perception threshold and the associated scent; from Swiegers et al. (2005).

Compound	Concentration in wine (mg/L)	Perception threshold (mg/L)	Aroma description
Ethyl acetate	22.5–63.5	7.5*	VA, nail polish, fruity
2-phenylethyl acetate	0–18.5	0.25*	Flowery, rose, fruity
Isoamyl acetate	0.1–3.4	0.03*	Banana, pear
Isobutyl acetate	0.01–1.6	1.6****	Banana, fruity
Hexyl acetate	0–4.8	0.7**	Sweet, perfume
Ethyl butanoate	0.01–1.8	0.02*	Floral, fruity
Ethyl hexanoate	0.03–3.4	0.05*	Green apple
Ethyl octanoate	0.05–3.8	0.02*	Sweet soap
Ethyl decanoate	0–2.1	0.2*****	Floral, soap
Propanol	9.0–68	500**	Pungent, harsh
Butanol	0.5–8.5	150*	Fusel, spiritous
Isobutanol	9.0–174	40*	Fusel, spiritous
Isoamyl alcohol	6.0–490	30*	Harsh, nail polish
Hexanol	0.3–12.0	4**	Green, grass
2-phenylethyl alcohol	4.0–197	10*	Floral, rose
Acetic acid	100–1150	280*	VA, vinegar
Acetaldehyde	10–75	100**	Sherry, nutty, bruised apple
Diacetyl	<5	0.2** / 2.8***	Buttery
Glycerol	5–14 g/L	5.2 g/L**	Odourless (slightly sweet taste)
Linalool	0.0017–0.010	0.0015******/0.025*****	Rose
Geraniol	0.001–0.044	5******/30*	Rose-like
Citronellol	0.015–0.042	8******/100*	Citronella
2-acetyl-1-pyrroline	Trace	0.0001******	Mousy
2-acetyltetrahydropyridine	0.0048–0.1	0.0016******	Mousy
4-ethylphenol	0.012–6.5	0.14*/0.6***	Medicinal, barnyard
4-ethyl guaiacol	0.001–0.44	0.033*/0.11***	Phenolic, sweet
4-vinyl phenol	0.04–0.45	0.02******	phamaceutical
4-vinyl guaiacol	0.0014–0.71	10******	Clove-like, phenolic

* 10% ethanol, ** wine, *** red wine,**** beer, ***** synthetic wine, ****** water

3.1 Varietal aromas

Wines made from specific grape varieties typically display varietal character, e.g. distinctive aromas which evoke that variety. Varietal aroma contributors may exist as free forms or as odorless precursors in the grape berry, which are then released as odorous compounds during harvesting, vinification or ageing. These aromas are representative for the particular grape variety (Francis and Newton, 2005). The main compounds responsible for the intense aromas in Sauvignon Blanc wines have been assumed to be varietal thiols, while terpenols are most characteristic for Muscat varieties (Darriet et al., 1991). The majority of varietal aroma compounds are present in grapes as bound forms, e.g. glycosylated in the case of terpenols and cysteinylated or glutathionylated in the case of varietal thiols, which are then released during fermentation.

3.1.1 Terpenes

Terpenes constitute a large family of molecules and more than 4000 compounds have been identified. These compounds are polymers of isoprene units with five carbons (C5). Each terpene is classified according to the number of units: hemi- (C5), mono- (C10), sesqui- (C15), di- (C20), disesqui- (C25) and tri- (C30) terpenes.

40 and 30 different terpenes have been identified in grapes and wine, respectively. In wine, the most olfactory important terpenes are monoterpene alcohols (terpenols) (reviewed by Strauss et al., 1986). Geraniol, linalool, nerol and *α*-terpineol, but also *cis/trans* rose oxide, which is an oxidized form of citronellol (Luan et al., 2005), are characteristic aroma contributors for certain wines, such as Muscat or Gewürztraminer, and give them typical floral or citrusy notes (reviewed by Black et al., 2015; Ribéreau-Gayon et al., 1975; Simpson, 1979). Their biosynthesis in grapes involves the building block isopentenyl pyrophosphate (IPP), which is produced in the cytosol and in the chloroplast. This intermediate plays a central role in the production of many molecules in plants. It is known that IPP synthesis in the cytosol takes place via the mevalonate pathway (Wright, 1961), allowing the production of sesquiterpenes, sterols and ubiquinone, while the synthesized IPP in the chloroplast is produced by the 2-C-methyl-D-erythritol-4-phosphate pathway, allowing the production of monoterpenes, diterpenes, carotenoids and chlorophyll (Schwender et al., 1996). However, the precursor IPP can also pass from the chloroplast to the cytosol (Bick and Lange, 2003). The isopentenyl diphosphate isomerase produces an isomer of IPP, dimethylallyl pyrophosphate (DMAPP) (Banthorpe et al., 1983). Subsequently, geranyl pyrophosphate synthases allow the condensation of an IPP molecule on DMAPP to form geranyl pyrophosphate (GPP) (C10), or on GPP to form farnesyl pyrophosphate (FPP) (C15) (Figure 3). There are also other terpenes in wines such as limonene, *α*-pinene or myrcene, although it is not known whether they are produced in the vine by terpene synthases or if they are derived from a chemical conversion of present terpenols.

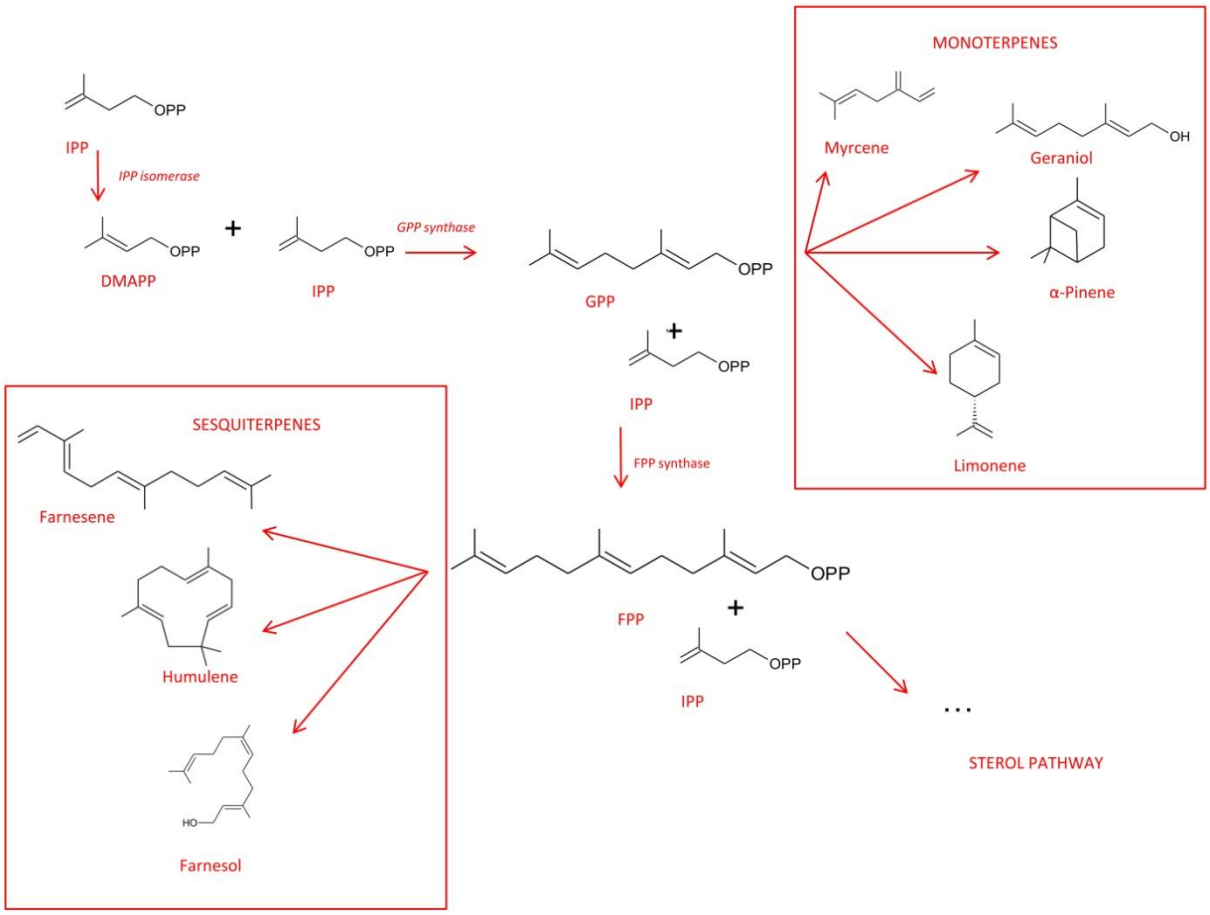


Figure 3: Synthesis of monoterpenes and sesquiterpenes (based on Cheng et al., 2007; McGarvey and Croteau, 1995; Nagegowda, 2010).

Although much of the variation in the level of terpenols in the finished wine is determined in the vineyard, yeast influences the levels of these compounds during fermentation (Furdíková et al., 2014; King and Dickinson, 2000). Yeast can catalyze the release of terpenols from glycosylated, non-volatile precursors (Voirin et al., 1990) and even *de novo* synthesize certain terpenols from geranyl pyrophosphate (Carrau et al., 2005). Koslitz et al. (2008) showed that *S. cerevisiae* produces both stereoisomers of rose oxide during fermentation. Furthermore, Steyer et al. (2012) reported several QTLs involved in the production of citronellol and *cis*-rose oxide by *S. cerevisiae* without identifying the underlying genes. Moreover, the study could identify the transcription factor Pdr8 to be responsible for the formation of nerolidol in wine fermentations. Later, Steyer et al. (2013) identified *OYE2* as the gene responsible for the reduction of geraniol into citronellol, and the involvement of *ATF1* in the acetylation of terpenols could be demonstrated.

The activated form of geraniol, geranyl pyrophosphate, is part of the sterol pathway (Figure 4). This led to the assumption that yeast could synthesize sterols from grape-derived terpenols as a reaction to anaerobiosis during alcoholic fermentation, which inhibits the oxygen dependent first steps of the sterol synthesis (Vaudano et al., 2004).

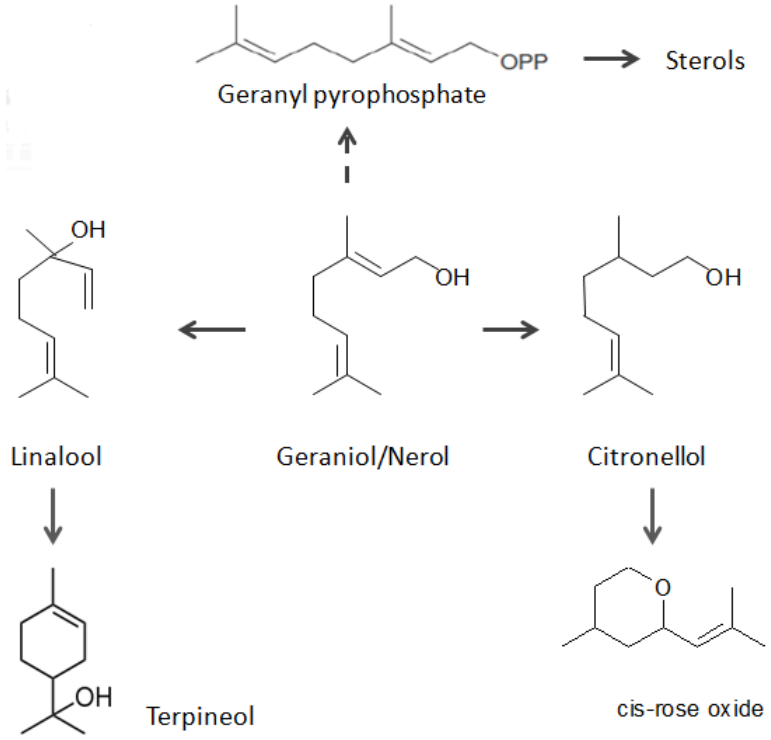


Figure 4: Schematic representation of the formation of terpenols during wine fermentation and the potential connection to yeast sterol pathway (based on King and Dickinson, 2000; Zea et al., 1995).

3.1.2 Volatile Phenols

Various phenolic substances contribute to taste, color and odor of wines (Dubois, 1983). From known classic wine components, vinylphenols (4-vinylguiacol, 4-vinylphenol) and ethylphenols (4-ethylguiacol, 4-ethylphenol) are the most significant phenolic aroma contributors (Etievant, 1981; Singleton and Esau, 1969). Their concentration can range from 0 to 6 mg/L. The derivatives 4-ethylguaiacol and 4-vinylguaiacol may add smoky, vanilla or clove-like odors to wine (Singleton and Esau, 1969). On the other hand, vinylphenols and in particular 4-ethylphenol may be responsible for phenolic off-flavors described as animal, horse sweat, stable, medical or "band aid". Their presence is considered undesirable since fruity nuances of the wine can be masked (Chatonnet et al., 1992; Dubois, 1983). This is even the case for concentrations lower than the perception threshold. It was demonstrated that while first the addition of 500 µg/L of 4-ethylphenol to a neutral wine resulted in medicinal, horse sweat or leather odors, the addition of 50 µg/L to the same wine already changed its sensory evaluation. However, this is not the case for 4-vinylguaiacol, which can enhance the varietal character of certain grape varieties, such as Gewürztraminer (Dubois, 1983; Grando et al., 1993). The described phenolic off-flavors most often develop in red wines during ageing, particularly when the wines are stored in old barrels and are seldom racked (Chatonnet et al., 1993, 1992).

Precursors of phenol compounds are produced by certain grape varieties and are often found in glycosylated forms (Singleton and Esau, 1969). Coumaric and ferulic acid are predominantly important as they can be transformed into volatiles by the cinnamate decarboxylase of *S. cerevisiae* during alcoholic fermentation (Chatonnet et al., 1993). As this enzyme is inhibited by catechin and catechin tannins, which are abundant in red wine, the levels of volatile phenols formed in red wines are in general lower than those found in white and rosé wines. The cinnamate decarboxylase of contaminating *Brettanomyces* or *Dekkera* species, however, is not inhibited by catechins. If present, these yeasts can reduce vinylphenols to ethylphenols, resulting in significant quantities of later compounds in red wines.

3.1.3 Volatile thiols

Sulfur-containing molecules are generally potent aroma contributors as their perception thresholds are low (Mestres et al., 2000), e.g., ranging from 0.8 ng/L to 60 ng/L for the volatile thiols 4-mercapto-4-methylpentan-2-one (4MMP), 3-mercaptohexanol (3MH) and its acetylated derivative 3-mercaptohexyl acetate (3MHA), which can impart pleasant smells of black currant/boxwood (4MMP), grapefruit (3MH) and passionfruit (3MHA) (Tominaga et al., 2006). 4MMP and 3MH are released from non-aromatic grape precursors by yeast during fermentation (Grant-Preece et al., 2010; Roland et al., 2010) and 3MHA is formed by acetylation of the released 3MH. While the production of 4MMP can differ considerably between different strains, the formation of 3MH and 3MHA is less varying and more dependent on grape must composition (Lee et al., 2008).

3.2 Fermentative aromas

During wine fermentation, the aromatic profile of must is changed and widened by the involved microorganisms (Figure 5). Besides altering varietal flavor compounds, yeast *de novo* synthesizes fermentative aromas during alcoholic fermentation. Higher alcohols and esters are hereby the most abundant volatiles synthesized by *S. cerevisiae* (reviewed by Cordente et al., 2012). In addition to environmental factors, such as nitrogen composition of the must or fermentation temperature, present yeast strains influence the nature and concentration of fermentative aromas (Molina et al., 2007). Gil et al., (1996) analyzed the aroma compounds of wines inoculated with pure and mixed cultures of apiculate and *S. cerevisiae* yeasts, showing that samples fermented with mixed cultures produce total higher amounts of alcohols and acids than wines produced with pure cultures of *S. cerevisiae*.

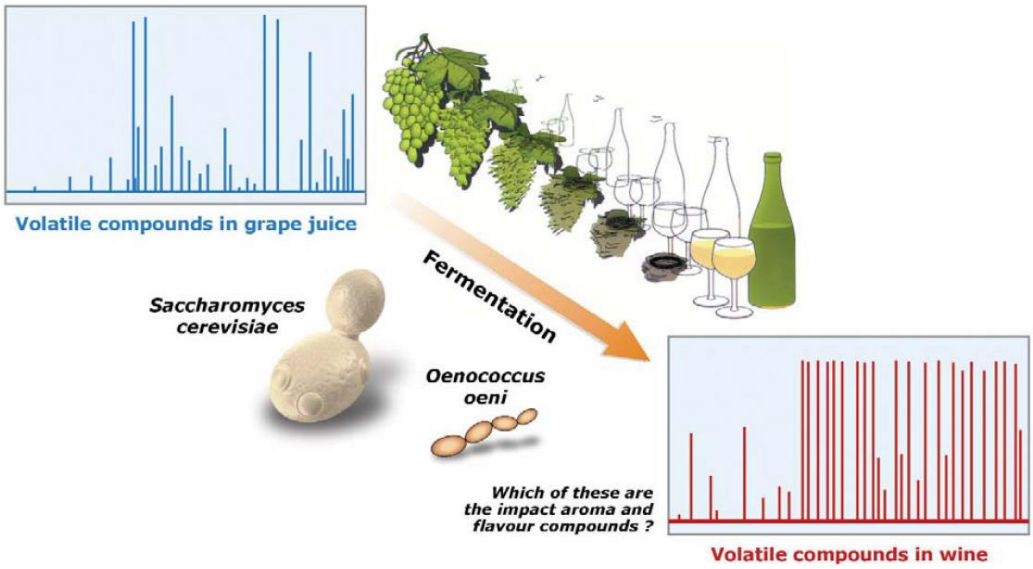


Figure 5: Schematic representation of the changes in aroma composition during wine fermentation, as pictured by chromatographic measurement. The yeast *S. cerevisiae* converts grape derived compounds and aroma precursors to volatile aromas; from Swiegers et al. (2005).

The production of fermentative aromas is closely linked to yeast central carbon and nitrogen metabolism (Figure 6) and is therefore dependent on the metabolic and enzymatic properties of the yeast (reviewed by Lambrechts and Pretorius, 2000).

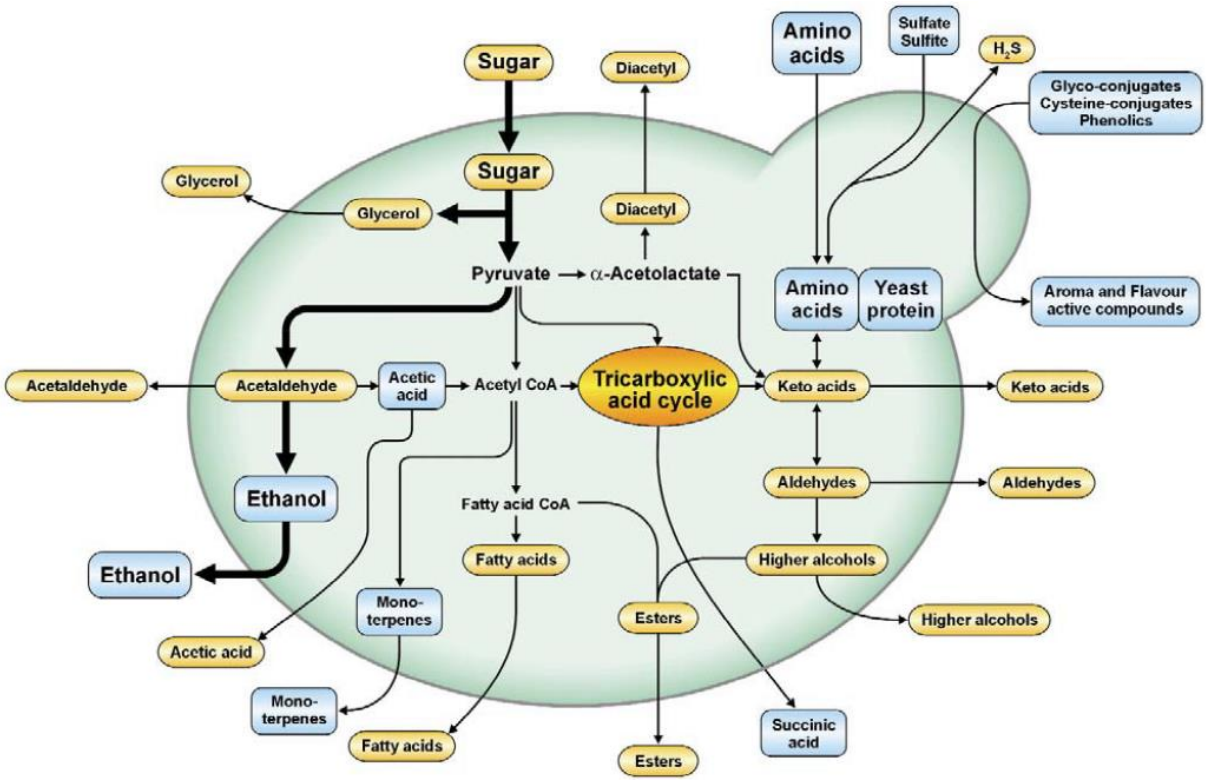


Figure 6: Simplified depiction of the derivation and synthesis of flavor-active compounds from carbon, nitrogen and sulfur metabolism of wine yeast; from Swiegers et al. (2005).

3.2.1 Higher alcohols

The presence of small amounts of higher alcohols, less than 300 mg/L, contributes to the fruitiness and the aromatic complexity of wines. 2-phenylethanol, for example, may bring floral notes to wine. If concentrations exceed 400 mg/L, however, these alcohols are considered deleterious, as they can impart strong and pungent odors (reviewed by Bell and Henschke, 2005). Higher alcohols in yeast are produced by the well-characterized Ehrlich pathway (reviewed by Hazelwood et al., 2008) (Figure 7).

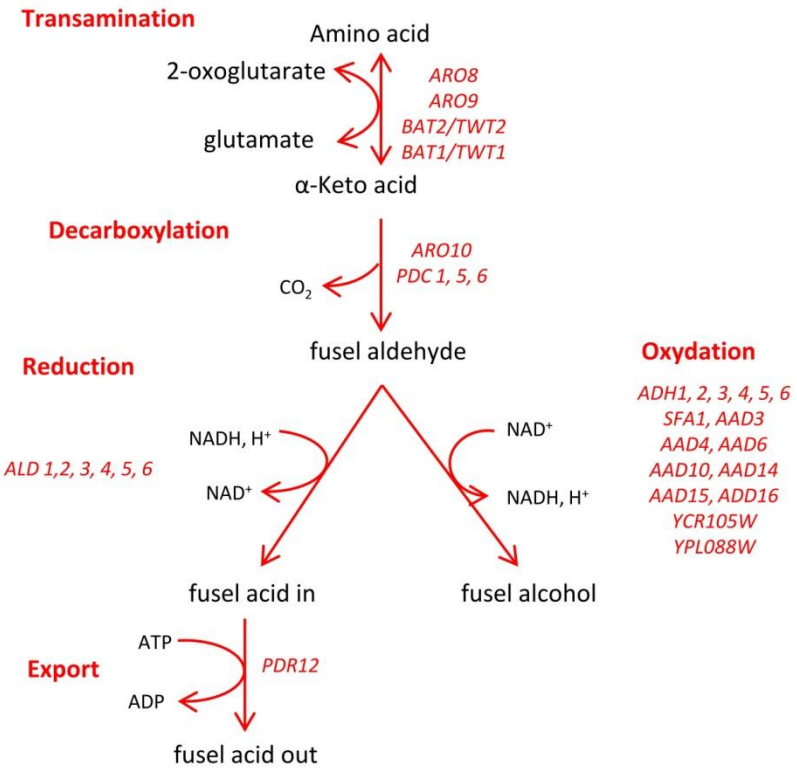


Figure 7: Production of higher alcohols in yeast (adapted from Hazelwood et al., 2008).

The production of higher alcohols is carried out by the decarboxylation and reduction of α-keto acids. These keto acids are either products of the central carbon metabolism, or they can be formed through transamination of different branched amino acids by Bat1 and Bat2 (Dickinson and Norte, 1993), respectively aromatic amino acids by Aro8 and Aro9. Other enzymes might be additionally involved in the transamination of branched amino acids (Eden et al., 2001). Overexpression of *BAT1* and *BAT2* in wine yeast resulted in increased levels of some higher alcohols, as well as in changes in the level of other aroma compounds in the finished wine (Lilly et al., 2006b). The production of higher alcohols is dependent on the availability of intracellular nitrogen. When the concentration of nitrogen is low, α-keto acids produced from the central carbon metabolism cannot be transaminated to form amino acids, and are metabolized into higher alcohols as a consequence (Moreno-Arribas and Polo, 2009). The decarboxylation of α-keto acids is carried out by the pyruvate decarboxylases Pdc1, Pdc5, Pdc6 and broad specificity decarboxylase Aro10. *THI3* is involved in leucine degradation, but it is assumed to encode a regulatory element rather than a catalytic enzyme (Hazelwood et al., 2008). The fusel aldehyde can then be reduced to a higher alcohol by a set of alcohol dehydrogenases (Adh1-Adh7) or by Sfa1. Alternatively, it can be oxidized to the corresponding fusel acid by a set of aldehyde dehydrogenases (Ald1-Ald6). The fate of the fusel aldehyde at this stage of conversion is dependent on the redox balance of the cell. Anaerobic, fermentative cultures favor the production of alcohols over acids to meet the demand for oxidizes nicotineamide adenine dinucleotide (NAD^+) (Hazelwood et al., 2008). The intermediates of the Ehrlich pathway are shown in Table 2, as well as their precursor amino acids.

Table 2: List of intermediates and products of the Ehrlich pathway, as well as their precursor amino acids.

Amino acid	α-keto acid	Fusel aldehyde	Fusel alcohol	Fusel acid
Leucine	α-ketoisocaproate	Isoamylaldehyde	Isoamyl alcohol	Isovalerate
Valine	α-ketoisovalerate	Isobutanal	Isobutanol	Isobutyrate
Isoleucine	α-ketomethylvalerate	Methylvaleraldehyde	Active amyl alchol	Methylvalerate
Phenylalanine	Phenylpyruvate	2-phenylacetaldehyde		2-phenylacetate
Tyrosine	*p*-hydroxy-phenylpyruvate	*p*-hydroxy-phenylacetaldehyde	*p*-hydroxy-phenylethanol	*p*-hydroxy-phenylacetate
Tryptophane	3-indole pyruvate	3-indole acetaldehyde	Tryptophol	
Methionine	α-keto-γ-(methylthio)butyrate	Methional	Methionol	

Even though the Ehrlich pathway is designated as the main synthesis route of higher alcohols (Hazelwood et al., 2008), most of these compounds can also be formed from glucose through the central carbon metabolism (reviewed by Bell and Henschke, 2005) (Figure 8). Indeed, the majority of the α-keto acids from which the higher alcohols are derived are metabolized from intermediates of the glycolysis or the tricarboxylic acid (TCA) cycle. For example, α-ketoisovalerate and α-ketoisocaproate, precursors for isobutanol and isoamyl alcohol, can be formed from pyruvate (Figure 8). Nevertheless, fusel alcohol production can not be seen as the result of amino acid imbalances, as recent analysis of the fate of amino acids during alcoholic fermentation revealed a complex of degradation and resynthesis of all amino acids (Crépin et al., 2017).

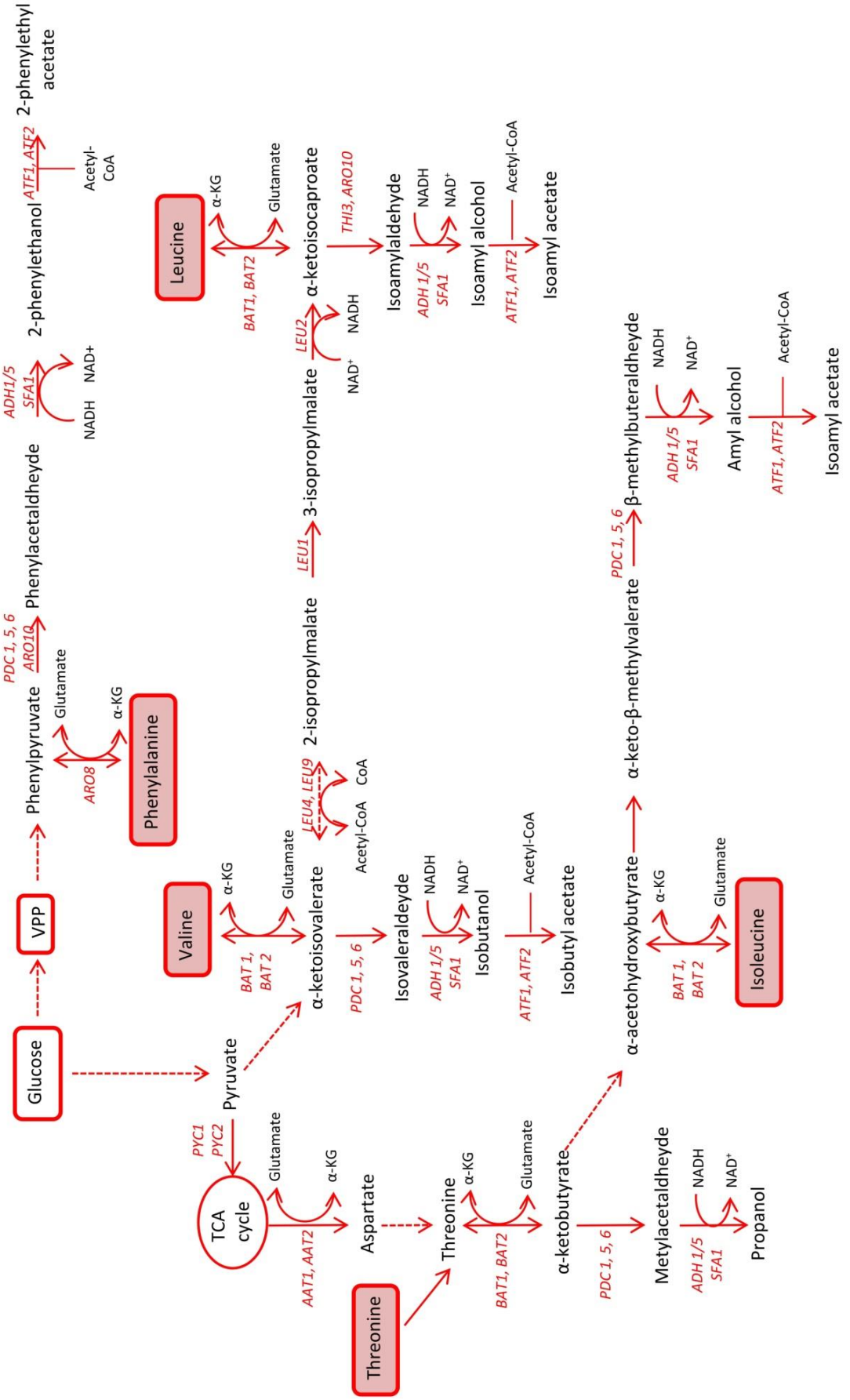


Figure 8: Synthesis of higher alcohols and acetate esters (adapted from Bell and Henschke, 2005).

3.2.2 Esters

The presence of esters can have an important effect on the fruity and floral notes of the wine. The contribution of esters depends on the type of wine. It involves mechanisms of interactions and synergies (Escudero et al., 2007; Ferreira et al., 2000). The two main classes of esters important to wine are acetate esters, such as ethyl acetate, isoamyl acetate and phenylethyl acetate, and ethyl esters, like ethyl hexanoate and ethyl octanoate (reviewed by Styger et al., 2011).

Most ethyl esters contribute to the aroma of young wines and are the source of pleasant fruity notes, whereas among the acetate esters only amyl acetate, isoamyl acetate and 2-phenylethyl acetate contribute to floral and fruity notes. Certain esters, such as ethyl acetate, may provide solvent type aromas at high concentrations.

Acetate esters are produced from acetyl-coenzyme A (acetyl-CoA) and alcohol (ethanol or higher alcohols), by the alcohol acetyltransferases Atf1 and Atf2 (Figure 8). Atf1 is the more important of those two enzymes and is considered to be responsible for the synthesis of 80% of isoamyl acetate, 75% of phenylethyl acetate and about 40% of ethyl acetate. Numerous studies have shown that enzymatic activity is the most important factor in the production of these esters (Lilly et al., 2006a, 2000; Mason and Dufour, 2000; Verstrepen et al., 2003). In a yeast strain overexpressing *ATF1*, the production of isoamyl acetate is multiplied by more than factor 100, and the syntheses of ethyl acetate and phenylethyl acetate are increased between 10 and 200 times, while the overexpression of the esterase gene *IAH1* has a negative effect on ester production (Lilly et al., 2006a). Double deletion mutants of *ATF1* and *ATF2* still show considerable levels of some esters, suggesting the presence of other unknown ester-synthesizing enzymes (Verstrepen et al., 2003). In conclusion, the production of acetate esters depends on both, the availability of the two co-substrates (acetyl-CoA and higher alcohols) and the enzymatic activities of involved acetyltransferases and esterases.

Ethyl esters are the product of activated fatty acids (fatty acyl-CoA) and ethanol. The acyl-transferase genes *EHT1* and *EEB1* are responsible for both, synthesis and hydrolysis of ethyl esters (Figure 9) (Saerens et al., 2006). It has been shown that overexpression of these two genes does not lead to overproduction of ethyl esters, because these enzymes possess both synthetic activity and esterase activity. Eht1 catalyzes exclusively the synthesis of ethyl hexanoate from ethanol and hexanoyl-CoA, while Eeb1 is an ethanol acyltransferase responsible for most of the ethyl ester biosynthesis during fermentation. This enzyme also possesses esterase activity of short chain esters and may be involved in lipid metabolism (Saerens et al., 2008).

The most important factor influencing the production of ethyl esters is the availability of the two precursors, acyl-CoA and ethanol. In addition, the C8 ethyl esters and their precursor fatty acids are not entirely excreted in the medium. A certain share of the product is retained within the cell, about 60% for C8, 20% for 10 and almost 100% for C12 (Saerens et al., 2008). A recent

quantitative trait locus (QTL) mapping study by Steyer et al. (2012) found that allelic variants of *PLB2* show different influence on the levels of ethyl esters. Deletion of this gene resulted in a significant drop in the levels of ethyl octanoate and ethyl decanoate. The effect was stronger than for *EEB1* deletion mutants.

3.2.3 Fatty acids

Fatty acids have odors usually deemed unpleasant. On the other side, some authors have found positive correlations between fatty acid content and wine quality (San-Juan et al., 2011), however, it is supposed that the ethyl esters, directly derived from fatty acids, are responsible for these correlations (Maarse, 1991). Among the fatty acids, only octanoic acid could influence the aroma of wines when sole administered. However, if the fatty acids are evaluated as a mixture, their impact on the aroma increases. In certain wines, the perception threshold of 10 mg/L can be reached with a mixture of hexanoic, octanoic and decanoic acids.

Saturated fatty acids with an even number of carbon atoms, which are largely predominant, have two possible origins, the fatty acid synthesis (Figure 9) and the catabolism of long-chain fatty acid chains via β-oxidation (Schreier and Jennings, 1979). The anabolic pathway is the predominant pathway, whereas β-oxidation, which requires the presence of oxygen, is active only at the beginning of fermentation.

During their synthesis, fatty acids are formed successively by adding two carbon atoms from malonyl-CoA to an acyl-CoA, with acetyl-CoA being the first acyl-CoAs of this elongation cycle (Figure 9). The enzymatic reactions are carried out by the multi-enzymatic fatty acid synthase complex, which is composed of two subunits (Lynen et al., 1980). Acetyl-CoA is mainly formed by oxidative decarboxylation of pyruvate.

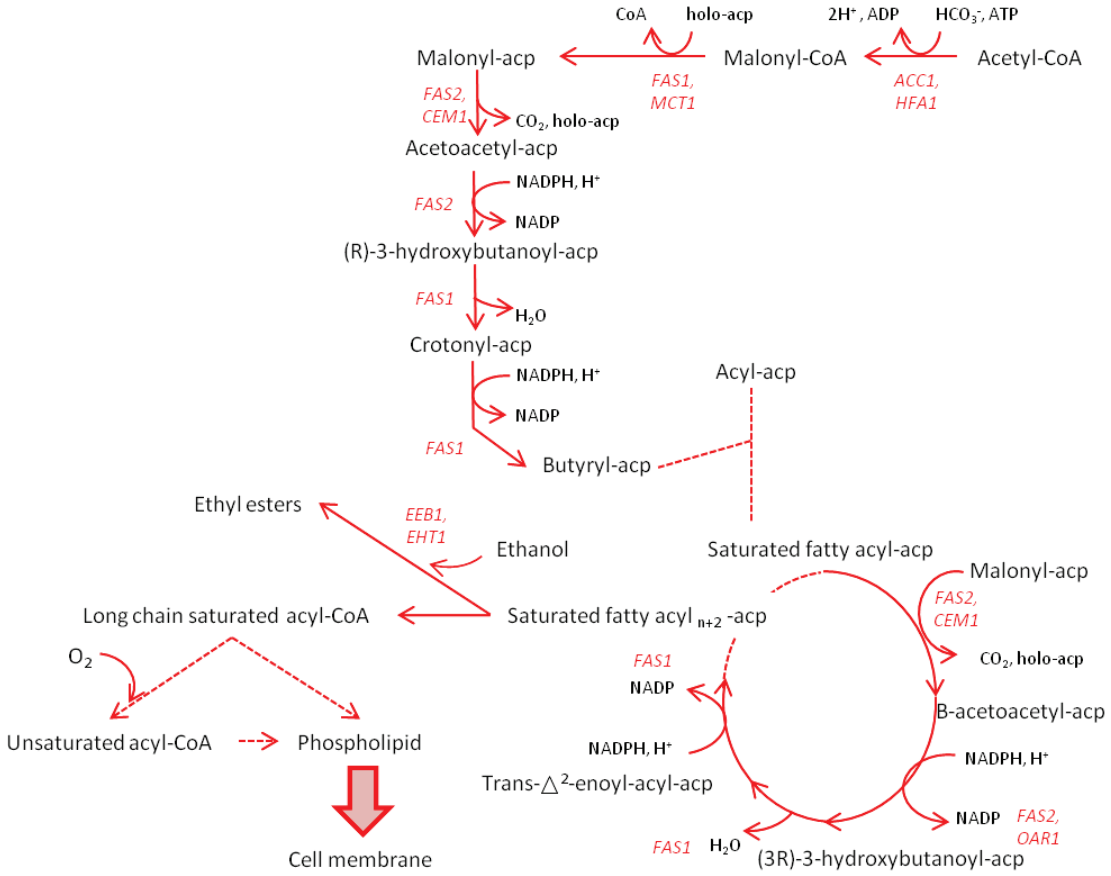


Figure 9: Fatty acid synthesis and their implication on ethyl esters formation (adapted from Saerens et al., 2010).

Fatty acid degradation by the β-oxidation pathway consists of a series of reactions, which release a molecule of acetyl-CoA from the fatty acid chain, resulting in a fatty acid reduced by two carbon atoms.

3.2.4 Volatile sulfur compounds

Sulfur dioxide (SO_2) and hydrogen sulfide (H_2S) are the most common sulfur-containing aroma compounds *de novo* generated by yeast from sulfur substrates during fermentation (Swiegers and Pretorius, 2007). Sulfur is present in grape juice as sulfate, which is contributed by grapes, and sulfite that is added by winemakers for microbial control. Besides prompting a health risk, their presence can impart off-flavors to wine (Mendes-Ferreira et al., 2009).

Other sulfur compounds with unpleasant smells may be produced during the catabolism of *S*-containing amino acids. Methionol, which has cauliflower and cooked cabbage aromas, is produced through the Ehrlich pathway, i.e., the transamination of methionine, decarboxylation of the resulting *α*-keto acid to methional and subsequent reduction (Perpète et al., 2005). Furthermore, methionine can be metabolized to methanethiol, which causes a

cooked cabbage flavor (Perpète et al., 2005). The yeast cell can then detoxify methanethiol to methyl thioacetate, which has a cooked vegetable smell (Rauhut, 2009).

3.3 Postfermentative aromas

Postfermentative aromas are formed during the maturation of wine due to chemical reactions. Varietal odorless precursors may undergo transformations, producing odor molecules. Fermentative aromas, such as esters, are hydrolysed, resulting in a modification of the fruity notes of the wine. In addition, the volatile substances of the wood diffuse into the wine during aging in oak barrels. These substances will be different depending on wood origin, manufacturing process and wine storage conditions, mainly wood composition and aging time (reviewed by Garde-Cerdán and Ancín-Azpilicueta, 2006).

Quantitatively, dimethyl sulfide (DMS) is the most important compound released during the maturation of wine, indicating the presence of DMS precursors in young wine. At low concentrations, it contributes towards the body of aged wines by accentuating berry fruit aromas (De Mora et al., 1987; Escudero et al., 2007; Lytra et al., 2016) and even notes of black olives and truffles (Segurel et al., 2004). DMS is known to confer pleasant notes to Cabernet-Sauvignon (De Mora et al., 1987) and Shiraz red wines (Segurel, 2005). Furthermore, different white wines with small additions of DMS were preferred over their non-treated counterparts (Spedding and Raut, 1982). At higher concentrations, the aroma of DMS is described as a fault, and can impart various vegetable flavors. DMS is a characteristic odor molecule found in several raw and processed vegetables (Buttery et al., 1971; Dignan and Wiley, 1976; Tulio et al., 2002; Ulrich et al., 2001). It is produced during wine alcoholic fermentation from dimethylsulfoxide (DMSO) or some sulfur amino acids by yeast metabolism (Bamforth and Anness, 1981; de Mora et al., 1986; Hansen, 1999).

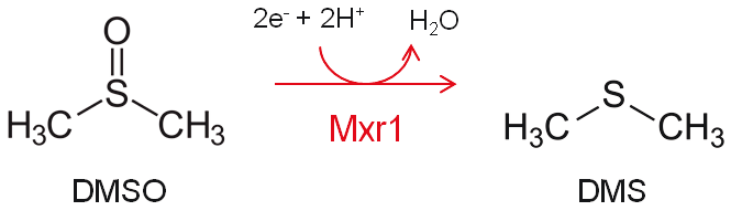


Figure 10: Formation of DMS from DMSO by the yeast methionine-*S*-sulfoxide reductase (Mxr1) during fermentation.

However, DMSO levels in must are low (Segurel, 2005) and the release of large amounts of CO_2 during alcoholic fermentation eliminates most of the DMS from the liquid, accounting for the low DMS concentrations observed just after the fermentation (Dagan, 2006). Loscos et al. (2008) revealed that *S*-methylmethionine (SMM), a methionine derivative, is the principal

precursor of DMS in grapes, accounting for more than 70% of the DMS potential (pDMS). Yeast expresses the high-affinity SMM transporter Mmp1 and potentially possesses other transport mechanisms (Rouillon et al., 1999). Intracellular SMM is metabolized to methionine by the methyl-transferase Mht1 (Thomas et al., 2000), therefore serving as an amino acid source (Figure 11).

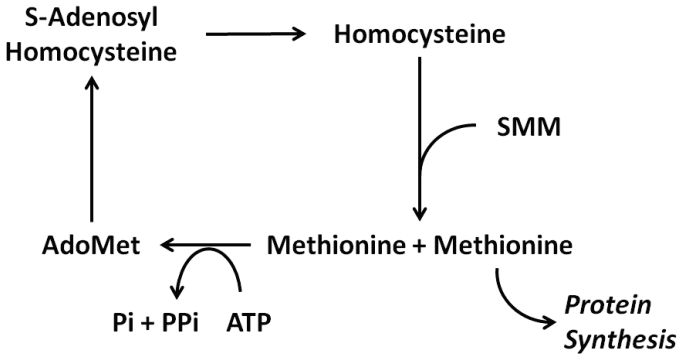


Figure 11: Metabolism of SMM by yeast. Enzymatic transfer of methyl-group from SMM to homocystein forms two molecules of methionine (adapted from Thomas et al., 2000).

Grape pDMS decreases during alcoholic fermentation, and only a small proportion of the pDMS is recovered in young wines. The recovery rate depends on numerous factors, such as the yeast strain, yeast assimilable nitrogen content in must and the general winemaking process (Dagan and Schneider, 2012). Nevertheless, the remaining pDMS leads to the release of DMS through degradation during wine aging, which is dependent on initial pDMS content and wine storage conditions, mostly storage temperature (Marais, 1979) (Figure 12).

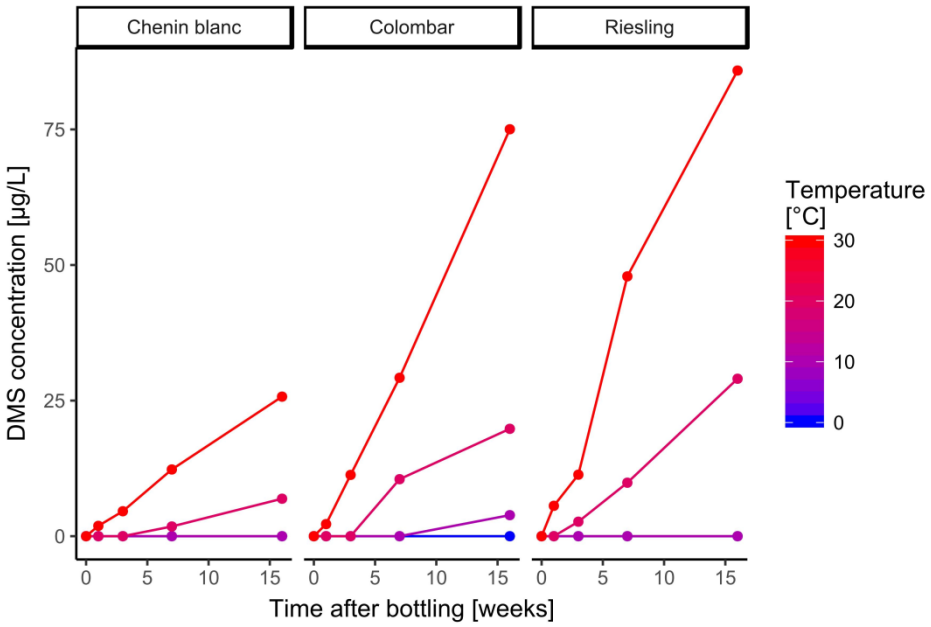


Figure 12: Development of DMS concentrations in wine stored over a 16-week period. Influence of wine grape variety (initial pDMS) and storage temperature on DMS formation (based on Marais, 1979).

In conclusion, the composition and level of listed aroma compounds influenced by yeast strongly determines the type and style of produced wine (Figure 13). Yeast strains producing characteristic flavor profiles were selected accordingly and are commercially available.

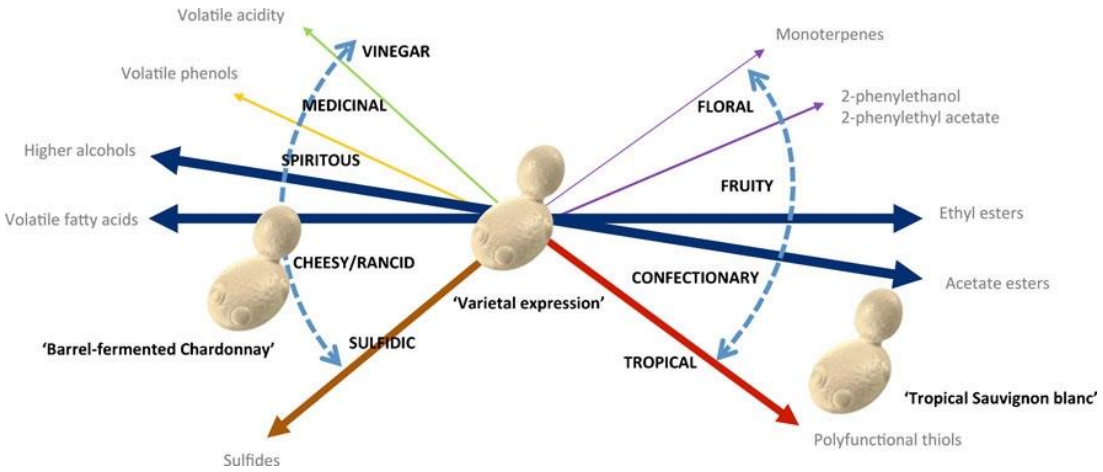


Figure 13: Potential range of wine flavor characteristics (dashed arrows) caused by the presence and level of representative aroma compounds (solid arrows, boldness correlating to magnitude of impact). Examples of desirable flavor phenotype represented by position of yeast; from Cordente et al. (2012).

3.4 Environmental factors affecting aroma formation

3.4.1 Temperature

Fermentation temperature is one of the main parameters affecting final contents of volatile compounds in wines. Temperature has a biological effect on yeast metabolism, as well as a physical effect on wine composition by causing evaporation of certain volatile compounds. The effect of temperature on the production of higher alcohols is complex. Beltran et al., (2008) showed that the total concentration of higher alcohols (isobutanol, isoamyl alcohol and 2-phenylethanol) increases with rising temperature, whereas Molina et al. (2007) only observed an increase in 2-phenylethanol, while other higher alcohols were decreased. When focusing on isobutanol and isoamyl alcohol, it was found that temperature has different effects on their synthesis kinetics. By using an online gas chromatograph (GC) system to determine the gas-liquid balance and kinetic parameters, Mouret et al. (2014) demonstrated that the total production of isobutanol increases with temperature, whereas production of isoamyl alcohol reaches its maximum value at 24 °C. Conversely, a negative effect of high temperatures (above 20 °C) on the final concentration of acetate and ethyl esters in wines is always observed, justifying the systematic use of low temperatures (between 15 and 20 °C) for the fermentation of white and rosé wines (Beltran et al., 2008; Molina et al., 2007; Mouret et al., 2014; Rollero et al., 2015).

3.4.2 Assimilable nitrogen content

Several studies have examined the impact of nitrogen supplementation of musts on the production of volatile compounds (Beltran et al., 2005; Jiménez-Martí et al., 2007; Mouret et al., 2014). The results obtained during these various studies are difficult to bring into line and are sometimes even contradictory, resulting from differences in experimental conditions and used strains. However, the synthesis of fermentative aromas is clearly influenced by the quantity and type of assimilable nitrogen, which is present in the musts or is added during the fermentation.

As higher alcohols derive partly from the catabolism of amino acids, their levels are influenced by the content of assimilable nitrogen. However, with the exception of propanol, this relationship is not monotonous. In fact, the addition of nitrogen to a deficient must (less than 60 mgN/L) leads to an increase in the production of higher alcohols, with a peak achieved with concentrations between 200 and 300 mg/L of assimilable nitrogen. Beyond these levels, the trend reverses and the production of higher alcohols decreases when the initial concentration of nitrogen further increases (Beltran et al., 2005; Jiménez-Martí et al., 2007).

The addition of nitrogen increases the synthesis of acetate esters as well (Beltran et al., 2005; Mouret et al., 2014), very likely because of a higher expression of genes encoding alcohol acetyltransferases that catalyze the conversion of higher alcohols to acetate esters

(Verstrepen et al., 2003). A similar observation was made for the production of ethyl esters, (Mouret et al., 2014).

3.4.3 Lipids

Sterols and fatty acids are not only essential components for maintaining the integrity of the yeast cell membrane, but they also act as precursors for the synthesis of certain aroma compounds (Drawert et al., 1966). When exogenous unsaturated fatty acids are available in abundance, the cells incorporate them into their membranes, which results in a significant reduction of *de novo* fatty acid synthesis. As a consequence, the production of aroma compounds, mostly fatty acid ethyl esters, is reduced (Yunoki et al., 2005). These observations are confirmed by Saerens et al., (2008), who showed that an addition of unsaturated fatty acids to the fermentation medium causes a reduction in the production of ethyl esters. The formation of ethyl decanoate is reduced by 50%, ethyl hexanoate by 33% and ethyl octanoate by 25%, which is sufficient to affect aromatic perception. The addition of unsaturated fatty acids decreases as well the production of acetate esters, e.g., isoamyl acetate is decreased by 32%. This is consistent with previous observations, which show that unsaturated fatty acids repress expression of *ATF1* (Fujii et al., 1997; Fujiwara et al., 1998).

3.4.4 Oxygen additions

In enological fermentation, oxygen availability is low. Some oxygen is available when grape must is brought into the fermentation vessel and in the case that oxygen is added during fermentation. This addition of oxygen, either from air or directly, is a popular practice that greatly reduces the risk of a sluggish or stuck fermentation (Blateyron and Sablayrolles, 2001). This technique does not affect total ester concentrations, but decreases the ratio of acetate esters to ethyl esters (Varela et al., 2012), which might be explained by the repression of *ATF1* by oxygen (Fujii et al., 1997). In agreement with this, Valero et al., (2002) observed a higher ratio of acetate esters to higher alcohols in a must without initial oxygenation.

3.4.5 Effect of other nutrients

In grape must, vitamins are generally present in sufficient quantities for a successful alcoholic fermentation, but their addition is beneficial for cell growth and may play a role in the production of aroma compounds (Hagen et al., 2008).

Minerals and metal ions are essential micronutrients that play important physiological roles during cell growth and alcoholic fermentation (Pereira, 1988), but these nutritional factors are often neglected regarding their contribution to the aroma of wines (Ibanez et al., 2008).

4 Other enological important metabolites

4.1 Glycerol

Excluding CO_2, glycerol is the second most produced metabolite during anaerobic fermentation. Its concentration ranges from 5 - 14 g/L in dry and semi-dry wines and can reach up to 25 g/L in botrytised wines, while white wines contain higher concentrations than red wines (Nieuwoudt et al., 2002). Glycerol can enhance mouthfeel and perceived sweetness of wine (Noble and Bursick, 1984), but the effect depends on wine type (higher effect in dry wines) and glycerol concentrations, with a sensory threshold of 5.2 g/L (Gawel et al., 2007). Glycerol plays essential roles to maintain the redox balance and to cope with osmotic stress (reviewed by Scanes et al., 1998). During fermentation, the production of glycerol allows to reoxidize the excess of reduced nicotinamide adenine nucleotide (NADH), which is generated during biomass formation and other NAD^+-dependent metabolic reactions. The production of glycerol involves two steps: the reduction of dihydroxyacetone phosphate to glycerol-3-phosphate, which is then dephosphorylated. Glycerol production is modulated by fermentation conditions, particularly heat, osmotic pressure, SO_2, nitrogen content and other factors affecting growth or physiological stress (Pigeau and Inglis, 2005).

In recent research, focus was placed on shifting yeast metabolism from ethanol towards glycerol production (reviewed by Kutyna et al., 2010; Tilloy et al., 2015) as an answer to higher alcohol contents in wine resulting from higher sugar content in grapes as a consequence of climate change (De Orduna, 2010). A wine yeast strain producing lower amounts of ethanol and high amounts of glycerol, 2,3-butanediol and succinate has been developed by evolutionary engineering (Tilloy et al., 2014) and has been commercialized.

4.2 Organic acids

Certain organic acids are derived from grapes, such as tartaric, malic and citric acid (Frayne, 1986), while others, such as succinic, acetic and pyruvic acid, are produced during fermentation in comparatively high concentrations (Whiting, 1976). They influence the overall acid-balance of wine and can therefore contribute to wine quality and organoleptic characteristics.

4.2.1 Acetic acid

Acetic acid is produced by yeast (and potentially acetic acid bacteria) as a byproduct during alcoholic fermentation (Malherbe et al., 2007). It is quantitatively and sensorial the most important volatile organic acid in wine and accounts for more than 90 % of volatile acidity and affects wine quality (Eglinton and Henschke, 1999). Acetic acid concentration depends on wine type and usually ranges from 0.2 - 0.8 g/L. In low concentrations, it provides warmth to

the palate but can provide a sourness taste and vinegary odor with exceeded levels (Lambrechts and Pretorius, 2000). At concentrations greater than 0.8 g/L, it is considered undesirable (Fleet, 1993). Acetic acid is produced as an intermediate of the pyruvate dehydrogenase (PDH) bypass, which converts pyruvate to acetyl-CoA. This pathway is the only source of cytosolic acetyl-CoA required for sterol and lipid synthesis (Pronk et al., 1996a, 1996b). Within the PDH bypass, acetic acid is formed from acetaldehyde by aldehyde dehydrogenases (Ald4, Ald5, Ald6) with Ald6 being the major isozyme (Saint-Prix et al., 2004). The production of acetic acid always increases rapidly at the beginning of fermentation, however, different yeast strains show variation (Shimazu and Watanabe, 1981).

4.2.2 Succinic acid

Succinic acid is by quantity the third most produced metabolite during wine alcoholic fermentation. It is present in concentrations of 0.6 - 1.2 g/L and contributes significantly to wine acidity. Succinic acid is derived from the TCA pathway. During anaerobiosis, this pathway functions as two branches (reductive and oxidative route) and succinic acid is mainly produced by the reductive branch during wine fermentation (Camarasa et al., 2003).

4.2.3 Pyruvic acid

Pyruvic acid is an important metabolic intermediate in alcoholic fermentation. It is for a large part decarboxylated to acetaldehyde by the pyruvate decarboxylase. Pyruvic acid has the ability to bind SO_2 (Rankine and Pocock, 1969), therefore influencing its germicidial power for which SO_2 is widely used in winemaking. The levels of pyruvic acid are initially low and increase with fermentation duration (Morata et al., 2003). Final levels of pyruvic acid are dependent on fermenting strain and pH of the must (Rankine and Pocock, 1969).

5 Deciphering complex traits

Phenotypic variations between individuals can be either characterized as qualitative or quantitative. Qualitative traits fit into discrete categories. They are shaped by a single locus or a small number of loci. In contrast, quantitative phenotypic traits present continuous ranges of variation and are typically controlled by several genomic regions, called quantitative trait loci or QTLs (Valdar et al., 2003). A QTL may contain a single gene or a group of linked genes, which influence the shape of the quantitative trait (Mackay, 2001).

The mapping of QTLs in the genome was first established in the twenties of the last century, with the study of Sax (1923) on the genetic analysis of the size of beans, associated with the color of pigmentation. This concept was then set up in plants (Lindstrom, 1924) and subsequently in other organisms, especially *S. cerevisiae* (Steinmetz et al., 2002; Deutschbauer

and Davis, 2005). As it is one of the main model organisms, the yeast has recently become an important model for elucidating the mechanisms that govern natural genetic and phenotypic variation. This success is partially due to its intrinsic biological features, such as the short sexual generation time, high meiotic recombination rate, and small genome size (Mancera et al., 2008). The development of DNA-microarray technology and second-generation sequencing methods enabled the fast and inexpensive genome-scale genotyping of yeast. This resulted in the generation of fully genotyped yeast crosses of homozygous strains that have been made available for QTL mapping (Cubillos et al., 2009).

5.1 Principle of QTL mapping

QTL mapping is performed with two strain lines of an organism capable of sexual reproduction. It consists in linking the genomic properties of a progeny of these strains, which are obtained by a genotyping step, to phenotypes of interest. The mapping can be divided into several steps, the crossing of two parental strains selected for their specific phenotypes, the sporulation of the cross, the genotyping of the daughter cells and the phenotyping of these strains for traits of interest. With this information, genomic regions that have an impact on the phenotypic traits can be detected and by dissecting these regions, alleles responsible for trait variation can be identified (Figure 14).

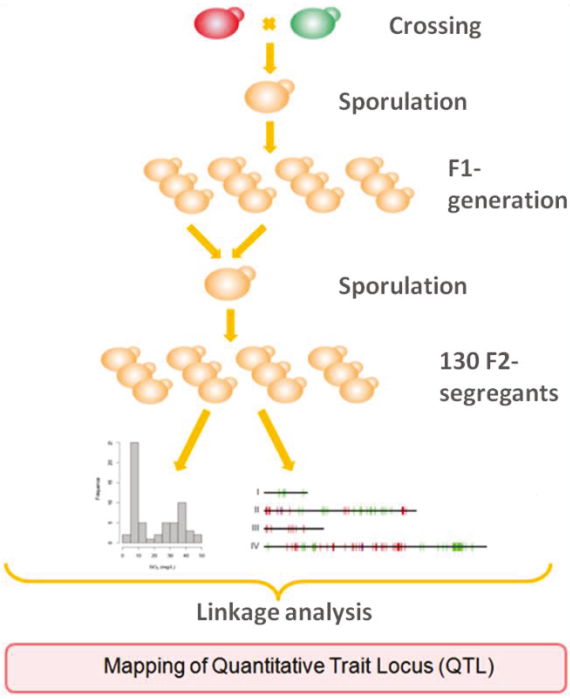


Figure 14 : Principle of QTL mapping used in this thesis. A yeast cross was sporulated to generate an F1-meiotic segregant population. Strains of this population were mated and sporulated again to obtain an F2-meiotic segregant population with increased recombination rate. 130 strains of this population were genotyped and phenotyped in order to perform linkage analysis on determined traits.

5.2 Establishment of a population of segregants

QTL mapping starts with the crossing of two parents, who may differ phenotypically in a trait of interest. However, this is not necessary as extreme differences in traits are often explained by major, but rare, loss of function alleles. Instead, crosses can be performed using the knowledge of population structure (Liti et al., 2017). After crossing, the sporulation of the diploid zygote leads to the production of segregants, the individuals of the F1 progeny, with genomic properties that are a mixture of the parent one's. The genetic diversity of the daughter generation is the result of random chromosomal recombination events during meiosis (~90 events per meiosis determined by Mancera et al., 2008). It is possible to further increase the recombination rate by crossing the F1 strains among each other to produce a second generation of segregants (F2) and eventually more (up to F13 for the production of a highly recombinant population by García-Ríos et al., 2017). This increases the resolution of the QTL mapping by shortening the genomic blocks that are inherited by each parent strain. The number of individuals constituting the evaluated population is another important criterion. A higher number of individuals increases the sensitivity of the analysis, as it strengthens the statistical power of the linkage analysis (Liti et al., 2017). Furthermore, a high number of segregants increases the probability of obtaining strains with extreme phenotypes, which can be advantageous for certain QTL mapping strategies, like bulk segregant analysis (Brauer et al., 2006).

5.3 Phenotyping of segregant population

During the phenotyping of the segregant population for the traits of interest, the reproducibility of the measurements is critical. If the selected traits are strongly influenced by environmental factors, or if the phenotypic determination is not reproducible enough, it can hinder the statistical analysis. The distribution of the phenotype gives an indication of the complexity of a trait. For example, the share of offspring individuals with the same character as the parent strains can give an idea about the number of alleles theoretically involved in the phenotypes (Lynch and Walsh, 1998). The distribution of the phenotype often follows one of the three main distribution laws, the Gaussian distribution, L-distribution or continuous distribution. The continuous and L-distribution correspond to the involvement of genes with major effects, while the Gaussian distribution is more characteristic for a phenotype resulting from the presence of multiple genes with minor effects (Griffiths, 2002).

5.4 Genetic markers and genotyping

The QTL mapping requires the use of genetic markers - polymorphic deoxyribonucleic acid (DNA) sequences, which determine the parental origin of a genomic region. The identification of markers and the assembly of a genetic map is therefore a key step in the search for QTLs. The most widely used genetic markers are single nucleotide polymorphisms (SNPs) and the developments of new genome sequencing technologies have open the way for high density

and even full genome genotypic maps, in which almost all the genetic variation between parents can be used. Two strategies can be chosen. For the first one, developed for plants by Michelmore et al. (1991) and called bulk segregant analysis, only those individuals that have the most extreme phenotypes are selected and genotyped as a group of individuals. For the second strategy, all individuals are genotyped, which gives the highest information on genetic diversity.

5.5 Linkage analysis

When the phenotyping and genotyping steps have been completed, linkage analysis consists in establishing a correlation between the presence of certain alleles and the evaluated quantitative phenotype. If the mean phenotype of segregants with allele A1 is different from those carrying allele A2 at locus A, this locus contains a QTL. Therefore, the simplest method is a variance analysis at each locus (Figure 15). The analytical power of this method is weakened by missing data and a low marker density in the region of the QTL of interest. A more complex methods is interval mapping method (Lander and Botstein, 1989), which is based on maximum likelihood parameter estimation and uses the information of markers flanking a position to determine the presence of QTLs along the genome and to evaluate a correlation between markers and the phenotype. Regression interval mapping approxes maximum likelihood interval mapping to reduce computation time (Haley and Knott, 1992).

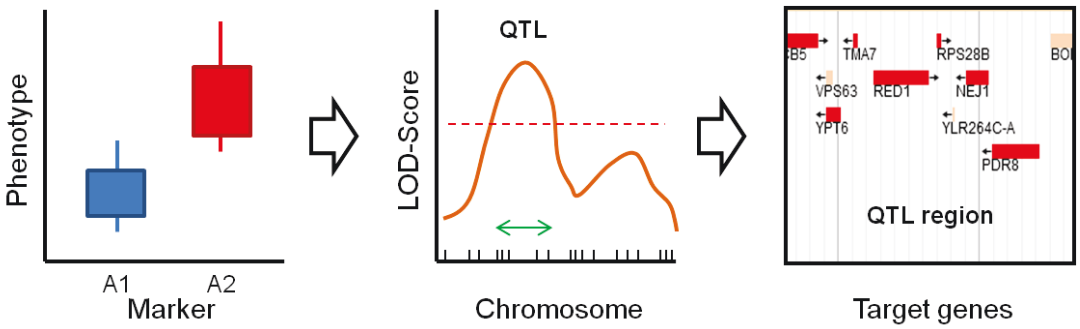


Figure 15: Simple depiction of linkage analysis. Phenotype distributions are evaluated at marker positions and a statistical analysis calculates the probability for each marker to have an influence on the trait, which therefore reveals the QTL positions with contained genes potentially involved.

The principle of multiple QTL research is to scan the genome for influences on the phenotype, but starting with a predefined assumption. For mapping double QTLs, an already detected single QTL is set as an explanatory variable in the model and a second QTL analysis is redone on the entire genome in the light of this additional variable. For detecting multiple QTLs, the number of QTLs set a priori as additional variables is not stated. Starting from a multiple

regression model, the genome is perused step by step, while only markers with an influence on the model are kept (Jansen, 1993; Zeng, 1993).

In general, the statistical power of the linkage analysis depends on the number of segregants and the quality of the marker map.

5.6 Dissecting a QTL and verifying candidate genes

The dissection of a QTL consists in reducing the size of the genetic region identified by the mapping step, and thus identifying the gene or genes present in this region, which have an influence on the examined trait. This step requires the use of databases that provide an annotation of the genes located in this region. The biological function of the genes is evaluated and genes with a potential link to the assessed trait are selected as target genes to verify the QTL. The complexity of this step is increased with the size of the identified region. The larger the genomic region, the more genes have potentially to be assessed. In a first step, the nucleotide sequence of each target gene is compared between the parental strains to determine polymorphisms, especially ones that result in a change of the protein's promoter sequence or amino acid chain. Subsequently, the impact of each target gene variant on the phenotype is tested, e.g. by reciprocal hemizygosity analysis (RHA) or allele swap (Figure 16) (Warringer et al., 2017). For RHA, a gene or allele is deleted in a heterozygous background, leading to hemizygosity of that sequence. By comparing two heterozygous strains with different hemizygous alleles, the impact of these alleles on the phenotype can be assessed. The impact of target gene variants on the assessed trait can also be validated by exchanging the corresponding DNA sequences between the parent cells. The strains constructed in this way are compared with the corresponding wild type parent strains to assess the change in phenotype caused by the swapped alleles.

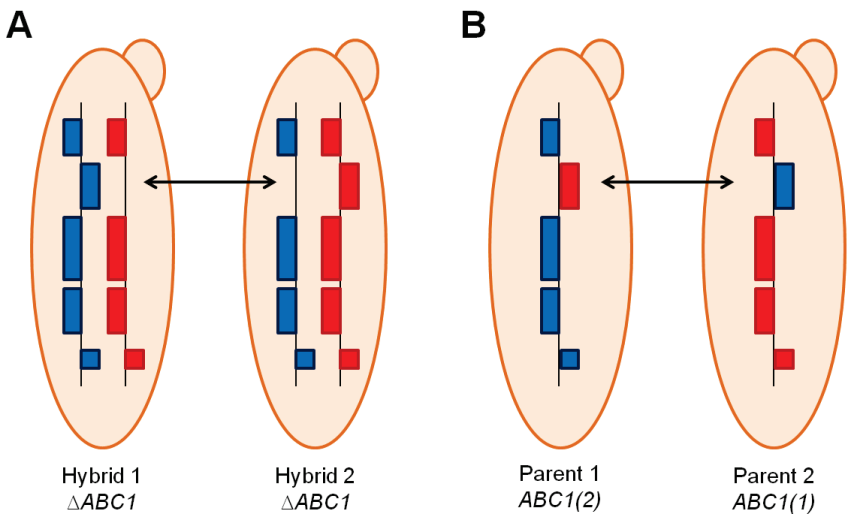


Figure 16: Methods of QTLs dissection used in this thesis. For RHA (**A**), both target gene alleles are separately deleted in the parental heterozygote. The resulting hemizygous strains are compared with each other to assess the impact of the remaining homozygous allele. For allele swap (**B**) the target gene alleles are exchanged between the parental strains and the resulting strains are compared with the undeleted parents to evaluate the resulting change in phenotype.

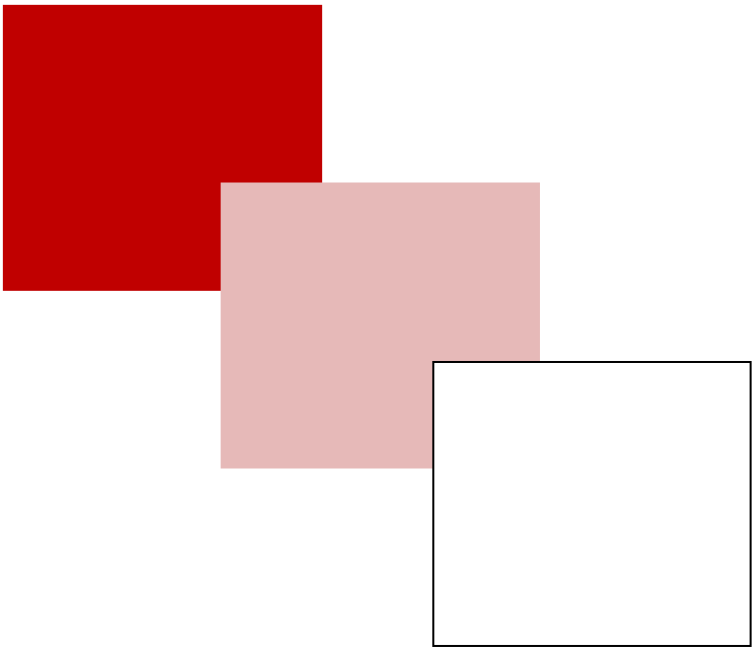
5.7 Industrial and enological applications of QTL mapping

Numerous QTL analyses have already been performed in yeast. This includes studies of genes involved in general yeast characteristics, e.g., high-temperature growth (Sinha et al., 2006), sporulation (Nogami et al., 2007) and stress tolerance (Brauer et al., 2006). Furthermore, an increasing number of QTL studies is carried out on industrial traits, such as traits important for, but not limited to, bio-ethanol production (Hu et al., 2007; Hubmann et al., 2013a, 2013b; Pais et al., 2013; Swinnen et al., 2016). In addition to that, the application of QTL mapping has shown to be a powerful tool for deciphering the genomic bases of enological traits. From its first use to map acetic acid production during wine fermentation (Marullo et al., 2007) it has recently led to a better understanding of the genomic bases influencing traits such as nitrogen source utilization (Brice et al., 2014a; Cubillos et al., 2017; Gutiérrez et al., 2013; Jara et al., 2014), sulfite production (Noble et al., 2015), the production of main metabolites as well as leaving of residual sugars (Salinas et al., 2012) and transcriptomic variation during fermentation (Ambroset et al., 2011; Brion et al., 2013). The formation of wine aroma is a quantitative and complex trait, as it is dependent on several genes, and recent QTL studies were also carried out to decipher the underlying genomic bases (Ambroset et al., 2011; Steyer et al., 2012).

The identification of QTLs and the characterization of alleles with different influence on the phenotype can also lead to the determination of molecular markers, which are then used for the selection or generation of new strains. As example, Ambroset et al. (2011), who used a cross between the lab strain S288C and a spore of the widely used winemaking strain EC1118, identified the gene *ABZ1* to be responsible for a QTL affecting fermentation rate. Using the same segregant population, Steyer et al. (2012) mapped QTLs for aroma production in

synthetic must and showed that the characterized variant of *ABZ1* also influences the production of 2-phenylethyl alcohol. Other QTL studies detected variants of *MET2* and *SKP2,* responsible for differences in the production of SO_2, H_2S and acetaldehyde (Noble et al., 2015). Salinas et al. (2012) used crosses from divergent yeast strains (wine, North American, West African and sake) and mapped several fermentation traits. They found variants affecting enological traits, such as acetic acid production (*ALD6*, subtelomeric region of chr II-R), residual sugar concentration (*MBR1*, *HAP4*, *YJR030C*), succinic acid production (*FLX1*, *MDH2*) and glycerol production (*YFL040W*, *GAT1*).

Materials and Methods



1 Media

1.1 Yeast culture medium

Yeast was cultured in Yeast extract Peptone Dextrose (YPD) medium at 28 °C under shaking. The composition of the medium is given in Table 3.

Table 3: Composition of YPD medium.

Compound	Amount per L
Yeast extract	10 g
Peptone	20 g
Glucose	20 g
Solid YPD media contained 1.5% agar	

Depending on the used resistance, selective YPD media contained 200 µg/mL geneticin (G418), 200 µg/mL nourseothricin (clonNAT) or 200 µg/mL hygromycin B respectively.

1.2 Micromanipulation media

The sporulation and dissection of heterozygous cells, which was carried out according to Codon et al., (1995), relies on the use of different media. The presporulation medium PRE5 allows the cells to reach the exponential phase. The SPO2 medium is a poor environment, leading to a state of nutritional deficiency, causing sporulation of the cells. The individual spores are deposited on YPD micromanipulation solid medium. The characteristic of this medium is its formulation with ultra-pure agar that is low in impurities, which could hinder the visualization of the asci. The media needed for cell sporulation are described in Table 4.

Table 4: Composition of yeast sporulation media.

Compound	Amount per L.
PRE5 medium	
Yeast extract	8 g
Peptone	3 g
Glucose	10 g
Potassium acetate	10 g
MiliQ water	to 1 L
SPO2 medium	
Potassium acetate	5 g
MiliQ water	to 1 L
Micromanipulation solid medium	
Yeast extract	10 g
Peptone	20 g
Glucose	20 g
Ultra pure agar	15 g
MiliQ water	to 1 L

1.3 Synthetic must

Fermentations were carried out in synthetic must (SM) that simulates standard grape juice (Bely et al., 1990). This medium is characterized by high equimolar concentrations of glucose and fructose, a limiting content of nitrogen and an acid pH of 3.3. Ammonia nitrogen and amino acid nitrogen are added as assimilable nitrogen sources. SM200 with 200 g/L of sugars, 200 mg/L of assimilable nitrogen and 5 mg/L of anaerobic factors (phytosterols) was used for this thesis. The composition of the medium and its components are given in Table 5 to Table 10. For mimicking grape derived aroma precursors, 2.0 mg/L geraniol and 500 µg/L *S*-methylmethionine (D/L-methionine methylsulfonium chloride) were added to the medium. SM was stored at -20 °C.

Table 5: Composition of the synthetic must SM200 used for fermentations.

Compound	Amount per L
Glucose	100 g
Fructose	100 g
Malic acid	6 g
Citric acid	6 g
Potassium phosphate	0.75 g
Potassium sulfate	0.5 g
Magnesium sulfate	0.25 g
Calcium chloride	0.155 g
Sodium Chloride	0.2 g
Ammonium chloride	0.22 g
Micronutrient stock solution	1 mL
Vitamins stock solution	10 mL
Anaerobic Factor stock Solution	0.33 mL
Amino acid stock solution	6.16 mL
Iron chloride stock solution	1 mL
The pH of the medium is adjusted to 3.3 with 10M Sodium hydroxide	

Table 6: Composition of micronutrient stock solution.

Compound	Amount per L
Manganese sulfate monohydrate	4 g
Zinc sulphate heptahydrate	4 g
Copper sulfate pentahydrate	1 g
Potassium iodide	1 g
Cobalt chloride hexahydrate	0.4 g
Boric acid	1 g
Ammonium heptamolybdate	1 g
The micronutrient stock solution is sterilized by 0.22 µm filtration and stored at 4 °C	

Table 7: Composition of the vitamin stock solution.

Compound	Amount per L
Myo-inositol	2 g
Calcium pantothenate	0.15 g
Thiamine hydrochloride	0.025 g
Nicotinic acid	0.2 g
Pyridoxine	0.025 g
Biotin	3 mL
The vitamin stock solution is stored at -20 °C	

Table 8: Composition of the anaerobic factors stock solution.

Compound	Amount per L
Phytosterols	1.5 mg
Tween 80	50 mL
Pure ethanol	100 mL
The anaerobic factors stock solution is stored at 4 °C	

Table 9: Composition of the iron chloride stock solution.

Compound	Amount per L
Iron (III) Chloride Hexahydrate	20 g
The iron chloride stock solution is stored at 4 °C	

Table 10: Composition of the amino acid stock solution.

Compound	Amount per L
Aspartic acid	3.4 g
Glutamic acid	9.4 g
Alanine	11.1 g
Arginine	28.6 g
Cysteine	1 g
Glutamine	38.6 g
Glycine	1.4 g
Histidine	2.5 g
Isoleucine	2.5 g
Leucine	3.7 g
Lysine	1.3 g
Methionine	2.4 g
Phenylalanine	2.9 g
Proline	46.8 g
Serine	6 g
Threonine	5.8 g
Tryptophan	13.7 g
Tyrosine	1.4 g
Valine	3.4 g
The amino acid stock solution is stored at -20 °C	

2 Yeast strains

The haploid *S. cerevisiae* strains MTF2621 (4CAR1 [Δ*HO*::*Neo*^{r}]) and MTF2622 (T73 [Δ*HO*::*Nat*^{r}]) were selected for the study according to their different need for nitrogen during wine fermentation. This was shown by using an approach based on the addition of nitrogen to keep the CO_2 production rate constant during nitrogen limitation (Brice et al., 2014b). The strain 4CAR1 belongs to the group of champagne strains, which originated through crossings between strains of the wine clade and the flor clade. The strain T73, however, belongs to the

phylogenetic clade of wine strains. For further analyses both strains were mated to form the parental heterozygote MTF2784.

3 Molecular biology techniques

3.1 DNA extraction

Yeast genomic DNA for sequencing was extracted using the MasterPure Yeast DNA Purification Kit (epicentre). To extract the DNA, 1.5 mL of an overnight yeast culture were transferred to a microcentrifuge tube and pelleted in a microcentrifuge for 5 min at 13000 rpm. The supernatant was removed and the pellet was resuspended in 300 µL of Yeast Cell Lysis Solution. Then, 1 µL of a 5 µg/µL RNase A solution was added to the tube, the suspension was vortexed and incubated at 65 °C for 15 min. The sample was placed on ice for 5 min, 150 µL of MPC Protein Precipitation Reagent were added and the mixture was vortexed for 10 sec. Cellular debris was pelleted by centrifugation in a microcentrifuge at 13000 rpm for 10 min. The supernatant was transferred to a clean microcentrifuge tube, 500 µL of isopropanol were added and the mixture was mixed by inversion. The DNA was pelleted by centrifugation in a microcentrifuge at 13000 rpm for 10 min and the supernatant was removed by pipetting. The pellet containing the DNA was washed with 0.5 mL of 70% ethanol. Then, the ethanol was carefully removed by pipetting and the microcentrifuge tube was left open at room temperature to evaporate any remaining ethanol. Finally, the DNA was suspended in 35 µL of TE buffer and stored at -20 °C.

Yeast genomic DNA for further analysis was extracted according to the protocol of Lõoke et al., (2011). One yeast colony was picked from an agar plate and suspended in 100 µL lysis buffer (200 mM LiOAC, 1% SDS). The solution was incubated for 15 min at 70 °C, 300 µL of 96% ethanol were added and the mixture was vortexed. The cell debris was spinned down in a microcentrifuge for 5 min at 13000 rpm and the pellet was subsequently washed with 70% ethanol. The pellet was spinned down in a microcentrifuge for 5 min at 13000 rpm, the ethanol was removed by pipetting and remaining ethanol was evaporated using a vacuum centrifuge. Then, the pellet was dissolved in 80 µL MilliQ water and the cell debris was spinned down in a microcentrifuge for 15 sec at 13000 rpm.

3.2 DNA amplification by polymerase chain reaction

Polymerase chain reaction (PCR) allows the amplification of a defined DNA strand, e.g. a genomic region. For this thesis, PCR was used for the construction of transformation cassettes and for the verification of performed transformations. The PCRs were carried out in a thermocycler with a heating lid (T3 thermocycler, Biometra or Mastercycler pro, Eppendorf). Two different polymerase enzymes were used for the PCR, the high fidelity Phusion

Polymerase (New England Biolabs) and the low fidelity Taq Polymerase (Thermo Scientific). The compositions of the PCR mixes are given in Table 11.

Table 11: Reaction mix for the PCR using Taq polymerase (left) and Phusion polymerase (right).

Compound	Quantity	Volume (µL)	Compound	Quantity	Volume (µL)
Buffer potassium chloride 10×	1/10	2.5	**Buffer HF 5×**	1/5	5
dNTPs 10 mM	200 µM	0.5	**dNTPs 10 mM**	200 µM	0.5
Primer fw (10 µM)	0.4 µM	1	**Primer fw (10 µM)**	0.5 µM	1.25
Primer rv (10 µM)	0.4 µM	1	**Primer rv (10 µM)**	0.5 µM	1.25
Taq polymerase	0.035 U/µL	0.25	**Phusion polymerase**	0.02 U/µL	0.25
DNA (25ng/µL)	50 ng	2	**DNA (25ng/µL)**	50 ng	2
Magnesium chloride 25 mM	2 mM	2.5			
MilliQ water		to 25	**MilliQ water**		to 25

The PCR programs were adapted to the chosen enzyme, to the hybridization temperature of the used primers and to the size of the fragment to be amplified. In general, the programs consist of an initial denaturation phase at 95 °C for 5 min, 30 - 35 cycles including a denaturation phase between 95 °C and 98 °C for 20 sec to 1 min, a hybridization phase between 52 °C and 70 °C for 20 to 30 sec and an elongation phase at 72 °C for 30 sec to 5 min. The PCR program ends with a terminal elongation phase at 72 °C for 5 to 10 min. All primers were synthesized by MWG (Eurofins MWG Operon).

3.3 DNA purification

DNA samples obtained by PCR were purified using the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel). One volume of sample was mixed with 2 volumes of Buffer NTI and 700 µL were loaded into a NucleoSpin Gel and PCR Clean-up Column placed into a collection tube (2 mL). The mixture was centrifuged in a microcentrifuge at 13000 rpm for 30 sec, the flow-through was discarded and the column was placed back into the collection tube. Then, 700 µL of Buffer NT3 were added to the column. The mixture was centrifuged at 13000 rpm for 30 sec, the flow-through was again discarded and the column was placed back into the collection tube. In order to remove Buffer NT3, the mixture was centrifuged at 13000 rpm for 1 min. Then, the column was placed into a new 1.5 mL microcentrifuge tube. After addition of 15-30 µL 70 °C warm Buffer NE, the tubes were incubated at room temperature for 1 min. Subsequently, the tubes were centrifuged at 13000 rpm for 1 min to eluate the DNA solution. The elution step was repeated to increase the DNA recovery rate.

3.4 Determination of DNA quantity and quality

After DNA extraction or PCR, the quality of the DNA was measured by spectrophotometry with a Nanodrop device (Nanodrop 1000, Thermo Scientific). The spectrophotometer gives certain wavelength signal ratios, which allows the detection of possible contaminations of RNA (260/280 nm higher than 2) and phenol (260/230 nm higher than 1.8).

DNA quantity was determined by fluorometric assay with the Quantus Fluorometer (Promega), using the QuantiFluor dsDNA system (Promega). First the QuantiFluor dsDNA Dye was diluted in 1:400 TE buffer to make the dye working solution. Then 1 µL of unknown DNA sample was added to 200 µL dye working solution into a 0.5 mL PCR tube. The samples were incubated at room temperature for 5 min and the fluorescence of the samples, indicating the DNA concentration, was measured.

3.5 Agarose gel electrophoresis

The validation of a PCR was performed by agarose gel electrophoresis, which allows the visualization of an amplification product and the determination of its size. The agarose gel consists of 0.8% agarose dissolved in 1 × TBE buffer (90 mM boric acid, 2 mM EDTA, 90 mM pH 8 Tris), with the addition of 0.2 µg/mL ethidium bromide. The PCR products were supplemented with a loading buffer, containing glycerol to increases the density of the sample and bromophenol blue to visualize its migration. A size marker (1 kb or 100 bp) was run along with each gel and allows the size determinations of the PCR products by comparing the migration distances. Gel electrophoresis was carried out in an electrophoresis device by submerging the gel in TBE buffer and applying a voltage of 120 V for 30 - 40 min. After migration, the ethidium bromide, which intercalates with the DNA, allowed the visualization of DNA by fluorescence under UV light. The gel was photographed subsequently (Kodak Gel Logic 100 Imaging system).

3.6 DNA sequencing

Genomic DNA samples were submitted to the sequencing platform Genotoul (Toulouse, France) and sequenced using the Illumina technology (paired end, 2 x 100 bp, HiSeq 2500) at a sequencing depth of 20-80 fold.

3.7 Yeast transformation

Yeast transformation was carried out according to the lithium acetate transformation method described by Gietz et al., (1995). Cells to transform were grown overnight in 5 mL of YPD. The optical density (OD) of the culture was determined and 50 mL of YPD were inoculated to an

OD of 0.5. The suspensions were grown until the exponential phase was reached, which corresponds to an OD of 2.0 – 2.2, to ensure maximal transformation efficiency. The cells were recovered by centrifugation (1960 × g for 5 min), washed with 40 mL of TRIS buffer (10 mM Tris HCl, pH 7.5), resuspended in 50 mL of lithium acetate buffer (0.1 M lithium acetate in 10 mM Tris HCl, pH 7.5) and incubated for 40 min at room temperature under gentle shaking. After centrifugation (1960 × g for 5 min), the cells were resuspended in 2.2 mL of lithium acetate buffer. In a 13 mL round cap tube, 100 µL of the cell suspension were mixed with 10 µL of salmon sperm DNA (denatured at 100 °C for 20 min and cooled in ice) and 1 – 3 µg of the transformation cassette. A second tube without the addition of transforming DNA was prepared as a negative control. After incubation for 10 min at room temperature, 300 µL PEG4000 solution (50 % w/v in TRIS buffer, sterile filtered) were added. The mixture was incubated for 10 min at room temperature. Subsequently a thermal shock was applied by submersion of the tubes in a water bath of 42 °C for 15 min. The thermal shock stimulates DNA uptake into the cells. Afterwards, the mixture was centrifuged (430 × g for 3 min), and the cell pallet was resuspended in 0.5 mL of YPD and incubated overnight without shaking. The cells were then centrifuged (1960 × g for 3 min), resuspended in TRIS buffer and spread on selective solid medium in two concentrations (1/10 and 9/10). The plates were incubated and grown colonies were isolated and verified by PCR.

4 Yeast sporulation and crossing

4.1 Microdissection of asci

The dissection and isolation of yeast asci was carried out using an MSM300 micromanipulator (Singer Instruments, UK). In order to obtain yeasts in the state of sporulation, a diploid yeast strain was transferred to 5 mL of YPD medium and grown overnight. Afterwards, the OD of the culture was determined and 25 mL of PRE5 medium were inoculated with the cells to an OD of 0.05. The culture was incubated for 16 h, the cells were washed with SPO2 medium and 10 mL of SPO2 medium were inoculated with the cells to an OD of 0.1 – 0.3. During incubation of the suspensions at room temperature and under gentle shaking, the rate of ascus formation was followed microscopically. When sporulation sufficiently took place (5 to 10 days), asci were treated with 0.2 mL of β-glucuronidase solution (15000 U/mL) for 15 to 30 min, resulting in a deterioration of the walls. 2 - 5 µL of cell suspension were spread on a seed strip at the side of the micromanipulation plate. The micromanipulation system was equipped with a needle that manually and individually retrieves the different spores of the same ascus and positions them on a grid, which served as a visual reference. Spore germination was stimulated by incubation for 48 to 72 h at 28 °C. The obtained colonies were subcultured onto solid YPD medium.

4.2 Yeast crossing and determination of mating types

Parental heterozygotes and F1 intercrosses were constructed by resuspending cell material of two haploid spores with different antibiotic resistances in 0.5 mL of YPD medium. The suspension was incubated for 4 – 8 h and 2 µL of the culture were dropped on double selective solid medium. Only mated heterozygotes with both antibiotic resistances were able to grow after incubation.

The mating type of haploid cells was also determined by yeast crossing. One spore of unknown mating type but known antibiotic resistance was resuspended in 0.5 mL of YPD medium, together with one of two reference strains with known mating type (a and α) and a different antibiotic resistance. The suspension was incubated for 4 – 8 h and 2 µL of the culture were dropped on double selective solid medium. Only spores with the opposite mating type of the used reference strain were able to mate and therefore grow after incubation.

5 Implementation of fermentations and monitoring of fermentation kinetics

The fermentations were carried out in glassware mini-fermenters of 300 mL, filled with 280 mL of SM200. The fermenters were equipped with water-filled airlocks to maintain anaerobiosis for the major part of the fermentation. Yeast strains were grown overnight in a shaking flask with 25 mL of YPD. The cell density was determined and the fermenters were inoculated to a cell density of 1.0×10^6 cells/mL. The fermentations were kept at a constant temperature of 24 °C under continuous magnetic stirring at 150 rpm. Fermentation progress was followed by the release of CO_2, which was determined by regularly measuring the weight loss, using a scale.

To facilitate the monitoring of the fermentation, a custom built robot (PlateButler) was used for parts of the experiments, which allowed the automatic weight measurement for up to 90 300-mL fermenters in 60 minute frequencies. The fermentation parameters were calculated and saved in a common data base.

Figure 17: Experimental layout of small scale wine fermentations with tracking of fermentation kinetics by hand (left) and using a custom-built robot (right).

6 Phenotyping

6.1 Determination of biomass

6.1.1 Cell population

Cell numbers and sizes were evaluated using an automated cell counter (Counter ZB-2, Beckman Coulter). Cell counting and size determination is based on the detection and measurement of changes in the electrical resistance produced by non-conductive particles. As a diluted suspension of cells in an electrolyte (Isoton solution, Beckman Coulter) is sucked through a small gap of 100 µm, the passage of each individual cell momentarily increases the electrical resistance between two electrodes, which are located on each side of the aperture. The volume of electrolyte displaced by each particle is measured in the form of a voltage pulse, whose height is proportional to the volume of the particle. The samples were diluted beforehand to remain within the linear range of the measurement between 20000 and 80000 cells/mL. Cell aggregates were disrupted by sonication, using an ultrasonic generator (Branson Sonifier, model 250). This measure allowed us to determine the number of cells per mL for a cell solution and the average cell volume (μm^3).

6.1.2 Dry matter

The dry weight was determined in duplicate by filtering 10 mL (V) of a cell culture through a nitrocellulose membrane with a porosity of 0.45 µm (Millipore, France) and known dry weight (P1). The membrane was rinsed twice with 10 mL of distilled water and dried for 48 h on a plate of known weight (P2) at 110 °C, in order to evaporate the water. Subsequently, the plate with the filter membrane was weighed (P3). The dry weight of the culture was calculated according to the formula:

$$\text{Dry weight (g/L)} = (P3\ (g) - (P2\ (g) + P1\ (g)) / V\ (L)$$

6.2 Determination of extracellular carbon compounds

The determination of remaining sugar substrates (glucose, fructose) and produced extracellular carbon metabolites (ethanol, glycerol, acetic acid, succinic acid, pyruvic acid, α-ketoglutaric acid) was carried out by high performance liquid chromatography (HPLC) (Camarasa et al., 2011). The flow rate of the HPLC (HPLC 1290 Infinity, Agilent Technologies, Santa Clara, California, USA) was set to 0.6 mL/min. Samples were separated by a pre-column and an ion-exclusion column (Phenomenex REZEX™ ROA-Organic Acid H+ (8%)), which was thermostatically controlled at 60 °C. The detection of the compounds was carried out using a refractometer in combination with a UV spectrometer at 210 nm. Chromatograms were processed by Agilent EZChrom software.

To prepare the samples, the cell culture was centrifuged for 5 min at 1360 × g. The supernatant was diluted 1/6 in the mobile phase, a degassed 0.005 N sulfuric acid solution. 25 µL of sample volume were injected automatically into the HPLC.

6.3 Determination of nitrogen compounds

Amino acids were derivatized with ninhydrin, which forms a purple compound detectable at 570 nm proportional to the concentration of the amino acid. Proline and hydroxiproline, which do not possess an amine function but an imine function, form a yellow compound detectable at 440 nm. The obtained spectra were analyzed with the Biochrom software.

After elimination of the cells by centrifugation, molecules with high molecular weights, like remaining proteins, were precipitated by adding 4 volumes of sample to one volume of 25% (w/v) sulfosalicylic acid, containing 2500 nmol/mL norleucine as internal standard. The mixture was incubated for 1 h at 4°C and centrifuged for 10 min at 4°C and 1960 × g. Samples were then filtered through a 0.22 µm Millipore nitrocellulose membrane. Amino acids were separated by liquid chromatography on an ion-exchange column (Ultrapac-8 lithium form;

Amersham Pharmacia Biotech) on Biochrom HPLC (Serlabo, France) and a temperature gradient from 32 °C to 75 °C, by using several buffers with a pH from 2.8 to 3.55 and a counterion concentration of lithium citrate from 200 to 1650 mM (Table 12, Serlabo, France). The column was then regenerated at 75 °C with a lithium hydroxide buffer of 300 mM.

Table 12: List of buffers used for the analysis of amino acids.

Buffer	Ion	Molarity (mM)	pH
Load	Lithium citrate	200	2.2
A	Lithium citrate	200	2.8
B	Lithium citrate	300	3
C	Lithium citrate	500	3.15
D	Lithium citrate	900	3.5
E	Lithium citrate	1650	3.55
F	Lithium hydroxide	300	

For quantification, a physiological standard was used. This standard contains one volume of acidic and neutral amino acids (A6407, Sigma-Aldrich, France), one volume of basic amino acids (A1585, Sigma-Aldrich, France), one volume of glutamine at 2500 nmol/mL and two volumes of lithium citrate. The standard was then processed as described above.

6.4 Determination of volatile compounds

6.4.1 Determination of fermentative aromas

Fermentative aroma samples were analyzed with a HP 6890 GC (Agilent Technologies, Santa Clara, California, USA) equipped with a CTC Combi PAL Autosampler Appl Microbiol Biotechnol AOC-5000 (Shimadzu, Columbia, USA) and coupled to a HP 5973 mass spectrometry (MS) detector (Agilent Technologies, Santa Clara, California, USA). The instrument was controlled and the data analyzed with the HP G1701DA ChemStation software (Agilent Technologies, Santa Clara, California, USA). The GC was fitted with a 60 m × 0.25 mm Phenomenex fused silica capillary column DB-WAX with 0.25 µm film thickness (Agilent Technologies, Santa Clara, California, USA). The carrier gas was helium, linear velocity was 36 cm/s and the flow rate 1.0 mL/min in constant flow mode. The initial oven temperature was 40 °C for 3 min. The temperature was increased by 4 °C/min until it reached 220 °C, and was held at this temperature for 20 min. The injector and the transfer line were held at 250 °C. The sample volume injected was 2 µL, and the splitter was opened after 30 sec, at 10:1. The focus liner (Agilent Technologies, Santa Clara, California, USA) was deactivated and tapered with glass wool (2–4mm). The mass spectrometer quadrupole temperature was set at 150 °C, the source was set at 230 °C, and the transfer line was held at 250 °C. For quantification, mass spectra were recorded in Selected Ion Monitoring (SIM) mode with positive ion electron impact at 70 eV. The ions monitored in SIM runs are shown in Table 13. For each compound, the ion marked

in bold was typically used for quantification, as it had the best signal-to-noise ratio and the least interference with other wine components. The other ions were used as qualifiers.

Table 13: Parameters for the determination of volatile fermentative aromas by GC-MS. Target ions for peak quantification in bold.

Compound	Retention time [min]	Qualifier ion m/z	Internal standard
Ethyl acetate	2.85	**61**, 70, 88	*d5*-ethyl butanoate
Ethyl propanoate	3.63	73, 75, **102**	*d5*-ethyl butanoate
Ethyl 2-methylpropanoate	3.77	71, 88, **116**	*d5*-ethyl butanoate
Propyl acetate	3.92	43, **61**, 73	*d5*-ethyl butanoate
2-Methylpropyl acetate	4.64	**56**, 73	*d5*-ethyl butanoate
Ethyl butanoate	5.15	71, **88**, 101	*d5*-ethyl butanoate
Propanol	5.2	31, 42, **59**	*d5*-ethyl butanoate
Ethyl 2-methylbutanoate	5.52	85, **102**, 115	*d5*-ethyl butanoate
Ethyl 3-methylbutanoate	5.95	85, **88**, 115	*d5*-ethyl butanoate
2-Methylpropranol	6.52	**43**, 74	*d5*-ethyl butanoate
2-Methylbutyl acetate	7.42	57, **72**, 74	*d5*-ethyl butanoate
3-Methylbutyl acetate	7.45	55, **70**, 87	*d5*-ethyl butanoate
Ethyl pentanoate	7.82	**85**, 88, 101	*d5*-ethyl butanoate
3-Methylbutanol	10.08	**55**, 57, 70	*d5*-ethyl hexanoate
Hexanol	10.59	**45**, 69, 87	*d5*-ethyl hexanoate
Ethyl hexanoate	10.96	88, 99, **115**	*d5*-ethyl hexanoate
2-Methylbutanol	11.48	**55**, 57, 70	*d5*-ethyl hexanoate
Hexyl acetate	12.25	56, 69, **84**	*d5*-ethyl hexanoate
Ethyl lactate	14.5	45, **75**	*d5*-ethyl octanoate
Ethyl octanoate	17.53	88, **101**	*d5*-ethyl octanoate
Propanoic acid	20.66	57, 73, **74**	*d5*-ethyl octanoate
2-Methylpropanoic acid	21.53	43, **73**, 88	*d7*-butyric acid
Butanoic acid	23.3	**60**, 73, 45	*d7*-butyric acid
Ethyl decanoate	23.72	**88**, 101	*d5*-ethyl decanoate
3-Methylbutanoic acid	24.49	157	*d7*-butyric acid
2-Methylbutanoic acid	24.51	**60**, 61, 87	*d7*-butyric acid
Diethyl succinate	24.79	73, **74**, 87	*d5*-ethyl decanoate
3-Methylthiopropanol	25.76	**101**, 129	*d5*-ethyl decanoate
Pentanoic acid	26.4	61, 73, **106**	*d7*-butyric acid
2-Phenylethyl acetate	28.47	60, **73**	*d5*-ethyl decanoate
Hexanoic acid	29.3	91, **104**	*d7*-butyric acid
Ethyl dodecanoate	29.4	**60**, 73, 87	*d5*-ethyl decanoate
2-Phenylethanol	30.9	**88**, 101	*d4*-phenylethanol
Octanoic acid	34.3	**91**, 92, 122	*d7*-butyric acid
Decanoic acid	36.84	**60**, 73, 101	*d7*-butyric acid
Dodecanoic acid	38.89	**60**, 73, 129	*d7*-butyric acid

The deuterated standard solution used for the quatification of fermentative aromas contained 100 µg/mL of ethyl-*d5* butanoate, 100 µg/mL of *d5*-ethyl hexanoate, 100 µg/mL of *d5*-ethyl octanoate, 100 µg/mL of *d7*-butyric acid, 100 µg/mL of *d5*-ethyl decanoate and 100 µg/mL of *d4*-phenyl ethanol. For the preperation of standard curves, a synthetic wine (12 % ethanol, 6

g/L malic acid, pH 3.3) was spiked with known concentrations of esters, higher alcohols and organic acids to measure. 10 µL of the deuterated standard solution were added to 5 mL of the synthetic wine in a 15 mL pyrex tube with a Teflon stopper (Supelco, Bellefonte, Pennsylvania, USA). Subsequently, 1 mL of dichloromethane was added and the mixture was gently shaken for 20 min on a shaking table. The samples were then centrifuged for 5 min at 1960 × g and 4 °C. The organic phase was recovered and the extraction was repeated once with the remaining aqueous phase. The resulting 2 mL of organic phase were dried with anhydrous sodium sulfate to remove all traces of the aqueous phase. Subsequently, the organic phase was transferred to a new vial and was concentrated through evaporation under nitrogen flow to a final volume of 0.5 mL. The concentrated sample was transferred to a concentrator insert (Supelco, Bellefonte, Pennsylvania, USA). New calibration standards were prepared with every set of fermentation samples to be measured. For the measurement of fermentation samples, the suspensions were centrifuged for 5 min at 1960 × g and 4 °C. The supernatants were treated as described above.

6.4.2 Determination of terpenols

For the analysis of terpenol composition, the samples were extracted by solid phase micro extraction (SPME) and measured via GC-MS (García et al., 1996). A Thermo Scientific Trace GC Ultra (Thermo Scientific, Waltham, Massachusetts, USA) equipped with a CTC Combi PAL Autosampler Appl Microbiol Biotechnol AOC-5000 (Shimadzu, Columbia, USA) and coupled to a Thermos Scientific ESQ™ single quadrupole MS detector (Thermo Scientific, Waltham, Massachusetts, USA) was used. The instrument was controlled and the data analyzed with the Thermo Xcalibur™ software (Thermo Scientific, Waltham, Massachusetts, USA). The GC was fitted with a 60 m × 0.25 mm Phenomenex fused silica capillary column DB-WAXTR with 0.25 µm film thickness (Agilent Technologies, Santa Clara, California, USA). The carrier gas was helium with a flow rate of 1.0 mL/min in constant flow mode. The injector was held at 270 °C, the source was set at 200 °C and the transfer line was held at 250 °C. The initial oven temperature was 40 °C PDMS/DVB/carboxen fiber was injected and the temperature was held for 5 min. Subsequently, the temperature was increased by 4 °C/min until it reached 180 °C and was then increased by 10 °C/min until it reached 240 °C. This temperature was held for 5 min. For quantification, mass spectra were recorded in Scan mode with positive ion electron impact at 70 eV. The ions monitored in Scan runs are shown in Table 14. For each compound, the ion marked in bold was typically used for quantification, as it had the best signal-to-noise ratio and the least interference with other compounds. The other ions were used as qualifiers. In the case of *cis*-rose oxide, geranyl acetate and citronellol, total ion current was used instead of qualifier ions.

Table 14: Parameters for the determination of terpenols by GC-MS. Target ions for peak quantification in bold.

Compound	Retention time [min]	Qualifier ion m/z	Internal standard
***cis*-rose oxide**	22.71	total ion current	*d5*-linalool
Linalool	29.13	55, **71**, 93	*d5*-linalool
Linallyl acetate	29.55	93, 69, **121**	*d5*-linalool
Citronellyl acetate	32.62	69, **81**, 95	*d5*-linalool
***α*-terpineol**	33.64	**59**, 93, 121	*d2*-geraniol
Neryl acetate	34.5	69, **93**, 121	*d2*-nerol
Geranyl acetate	35.28	total ion current	*d2*-geraniol
Citronellol	35.46	total ion current	*d5*-linalool
Nerol	36.45	41, 69, **93**	*d2*-nerol
Geraniol	37.87	41, 69, **139**	*d2*-geraniol

The deuterated standard solution used for the quantification of terpenols contained 8.35 µg/mL of *d2*-geraniol/*d2*-nerol and 0.835 µg/mL of *d5*-linalool. For the preperation of standard curves, a synthetic wine (12 % ethanol, 6 g/L malic acid, pH 3.3) was spiked with known concentrations of terpenols to measure. In a 20-mL glass flacon, 2.3 g of sodium chloride were dissolved in 7 mL of MilliQ water and the solution was kept on ice. 30 µL of internal deuterated terpenol standard and 1 mL of either calibration standard or sample were added, the tubes were closed and mixed well. The 2 cm long PDMS/DVB/carboxen fiber was inserted in the headspace of the vessel while agitating and incubating the sample for 30 min at 30 °C. Subsequently, the fiber was inserted in the GC-MS and terpenol content was determined as described above.

6.4.3 Determination of pDMS

For the measurement of remaining SMM, 10 mL of fermentation samples were centrifuged for 5 min at 1960 × g and 4 °C. The supernatants were kept at -20 °C and were submitted to Nyséos (Montpellier, France) for determination. The determination method first eliminated DMS already formed in the sample. Afterwards, the contained SMM was reacted to DMS by heat alkaline treatment and the concentration of produced DMS was determined by SPME and GC-MS measurement (Segurel et al., 2004).

7 Genotyping and QTL analyses

7.1 Calculation of heritability

The term heritability refers to the proportion of phenotypic variation within a population that is of genetic origin. In this thesis, heritability was calculated according to Brem et al., (2002). The total variation of a trait within a population of strains (var_{pop}) is

composed of the relative contributions of genetic differences (var_{gen}) and environmental influences (var_{env}).

$$var_{pop} = var_{gen} + var_{env}$$

Environmental influences var_{env} are estimated from the variance of a trait among the replicates of each parent of the study (where n is the number of replicates):

$$var_{env} = \frac{(n_{4CAR1} - 1)\, var_{4CAR1} + (n_{T73} - 1)\, var_{T73}}{n_{4CAR1} + n_{T73} - 2}$$

The population variance var_{pop} was calculated from the trait variance among the segregant population. Heritability was thus used to separate genetic and environmental components from phenotypic variation. Broad heritability (H2) was equal to the ratio of genetic variance to the total variance:

$$H2 = \frac{var_{pop} - var_{env}}{var_{pop}} \times 100$$

7.2 Genotyping and generation of marker map

Low-quality reads, obtained by the sequencing, were processed and filtered using the FASTX Toolkit v0.0.13.2 and TRIMMOMATIC v0.30 (Bolger et al., 2014), using a quality threshold of 20. First, reads were aligned to the S288C reference genome (release R64-1-1) using BWA v0.6.2 (Li and Durbin, 2009). Once the reads have been mapped, consensus genotype calling was performed using the tools available in the SAMtools package (Li et al., 2009). The initial variant set of 18155 biallelic variant positions was filtered for insuring a minimum spacing of 2.0 kb between SNPs. This resulted in a genotyping variant dataset of 3727 SNP markers.

7.3 Linkage analysis

The statistical analyses were performed using the programming language R v3.2.3 (www.r-project.org) with the program libraries R/qtl v1.40-8 and R/eqtl v1.1-7 (Broman et al., 2003). QTL mapping was performed with two different phenotype models, the normal model using Haley-Knott regression, and a non-parametric analysis. Two-dimensional, two-QTL scan was performed using the function *scantwo*. Multiple QTL mapping was performed with the function *stepwiseqtl*. Single QTL results were grouped as common QTL regions if their peaks were less than 10 cm apart. Proposed models of interaction were assessed with the function *fitqtl*.

7.4 Validation of QTL

7.4.1 Reciprocal hemizygosity analysis

QTLs were either evaluated as a whole, or single genes with a potential influence on the trait were chosen for evaluation. In both cases a disruption cassette containing the hygromycin B resistance gene (*hph*^{r}) from plasmid pAG32 (addgene) was amplified. The used primers del_[GENE]_fw/del_[GENE]_rv or del_[QTL]_fw/del_[QTL]_rv each contained an additional 40 bp overhang sequence, complementary to the upstream respectively downstream region of the target gene or region.

In the case of single gene assessment, these genes were deleted separately in both parent strains by homologous recombination with the disruption cassette. Transformed cells were selected on selective solid medium and correct gene deletion was verified by PCR, using the primer test_[GENE]_fw, which binds in the upstream region of the deleted gene, and the primer Hygro_rv, which binds within the deletion cassette. Deleted parent strains were subsequently mated with the opposite undeleted parent to form a heterozygote that was hemizygous for the target gene. The obtained strains were phenotyped (6) to assess the influence of the target gene respectively allelic variant on the trait.

In the case a QTL region is assessed as a whole, the disruption cassette was transformed into the parental heterozygote. Transformed cells were selected on selective solid medium and correct gene deletion was verified by PCR, using the primer test_[QTL]_fw, which binds in the upstream region of the deleted QTL, and the primer Hygro_rv, which binds within the deletion cassette. Deleted strains were further assessed by allelic PCR to determine the deleted variant. Hereby, the primer test_[QTL]_fw binds in the upstream region of the deleted QTL, and the primers tal_[GENE]_1/tal_[GENE]_2 bind at SNP positions of an allelic gene variant within the targeted region.

7.4.2 Allele swap

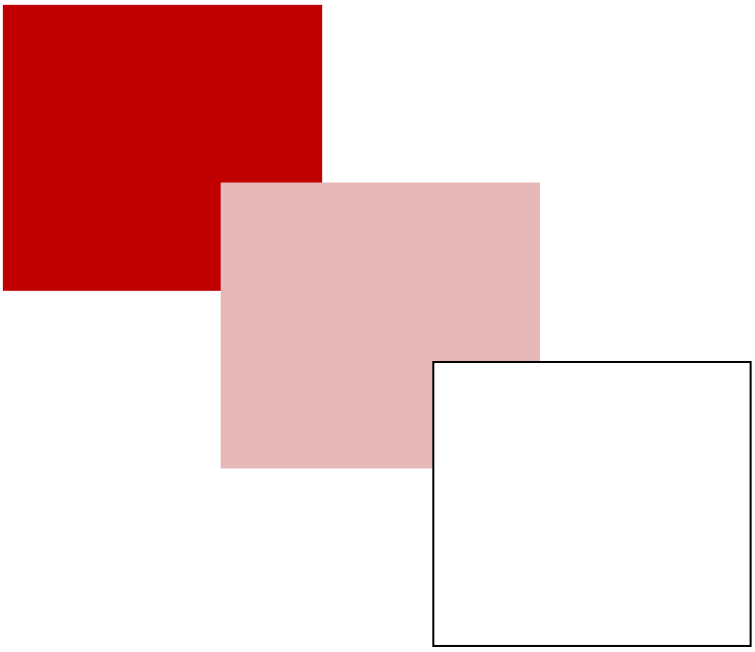
In the course of this thesis, allele swap was performed using the CRISPR/Cas9 toolbox developed by Mans et al. (2015). This approach has the advantage that genomic regions and even sole SNPs can be rapidly exchanged in a marker- and scarless way. The system consists in three components, the Cas9 enzyme, a spezific guide-ribonucleic acid (RNA) and a repair fragment. The guide RNA is complementary to a nucleotide sequence in the genome and binds to it. Cas9 recognizes the bound guide-RNA and performs a double-strand brake of the genomic DNA, which leads to DNA repair and alteration by the cell, using the repair fragment as a template.

The Cas9 expression cassette and the Hygromycin B resistance cassette were amplified by PCR from plasmids p414-TEF1p-cas9-CYC1t (DiCarlo et al., 2013) and pAG32 (addgene). The primers (Table 15) consisted of overlapping sequences for self-assembly of both cassettes and insertion of the resulting construct into the genomes of parent cells via homologous recombination replacing *GAL1*. The fragments for assembly of CRISPR transformation plasmids were amplified by PCR using template plasmids pROS11 – pROS17 (Mans et al., 2015) and primers containing allele specific guide sequences (Table 15). The plasmid fragments were transformed in *CAS9* carrying parental cells, together with a repair fragment mediating the allelic swap. The CRISPR plasmid was assembled by the cell, the guide RNA was expressed and mediates the double-strand break by Cas9. Positive sequence exchange mediated by the repair fragment was verified by allelic PCR, using a forward primer in the upstream region of the gene and one or more reverse primers on SNP positions within the gene. The obtained strains were phenotyped (6) to validate the impact of the allelic variants on the trait.

Table 15: Primers used for amplification of *CAS9* integration cassette (1678-5981) and CRISPR plasmids (6005 and CasGuide_GENE), according to the protocol of Mans et al. (2015).

Primer name	Nucleotide sequence (5‘ -> 3‘)
1678_GAL1DisB	AATGAGAAGTTGTTCTGAACAAAGTAAAAAAAAGAAGTATACTTACATAGGCCACTAGTGGATCTGTATAG
3093_tagA-pUG	ACTATATGTGAAGGCATGGCTATGGCACGGCAGACATTCCGCCAGATCATCAATAGGCACCTTCGTACGCTGCAGGTCGA
4653_A-CYC1t-rv	GTGCCTATTGATGATCTGGCGGAATGTCTGCCGTGCCATAGCCATGCCTTCACATATAGTCCGCAAATTAAAGCCTTCGAG
5981_Cas9_GAL1_fw	TTCACCGGTCGCGTTCCTGAAACGCAGATGTGCCTCGCGCCGCACACCGTATTACCGCCTTTGAGTG
6005_p426-CRISP_rv	GATCATTTATCTTTCACTGCGGAGAAG
CasGuide_GENE	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGATAAATGATC(20bp_homologous_sequence)GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAG

Results



<u>Chapter 1:</u> Establishment of QTL mapping strategy and QTL analysis of fermentative aroma formation and main metabolite production.

In the first part of this thesis the bases for the performed QTL analysis were established. Two wine yeast strains were chosen as they had previously been demonstrated to have different requirements for nitrogen during fermentation, which was seen as indication of different capacities to produce fermentative aromas. A population of 50 F2-segregants from a cross between these chosen strains was already available from previous studies in the working group. 80 F2-segregants were newly generated during this thesis by repeated crossing, sporulation and dissection. The strains were individually genotyped by whole genome sequencing and phenotyped during small scale (300 mL) wine fermentation in synthetic grape must by determining extracellular metabolites near the end of fermentation. The genotyping data was used to identify all SNPs between the parent cells and to establish a marker map of selected SNP identities for all segregant strains.

Subsequently, the phenotype data and marker map were used to perform linkage analysis in order to detect QTLs in yeast's genome that influence the production of main metabolites and formation of fermentative aroma molecules. Besides mapping single QTLs, several statistical approaches were applied to detect multiple QTLs and their interactions. A selection of detected QTLs was dissected by reciprocal hemizygosity analysis and allelic variants with different impact on aroma formation were described.

The article was published the 01.03.2018 in *BMC Genomics* 19:166.

QTL mapping of volatile compound production in *Saccharomyces cerevisiae* during alcoholic fermentation

Matthias Eder[a], Isabelle Sanchez[a,b], Claire Brice[a], Carole Camarasa[a], Jean-Luc Legras[a], Sylvie Dequin[a,*]

[a]SPO, INRA, SupAgro, Université de Montpellier, F-34060 Montpellier, France

[b]MISTEA, INRA, SupAgro, F-34060 Montpellier, France

*Corresponding author. Mailing address: SPO, INRA, SupAgro, Université de Montpellier, F-34060 Montpellier, France;

Phone: +33 4 99 61 25 28; E-mail: sylvie.dequin@inra.fr

E-mail addresses:
Matthias Eder: matthias.eder@inra.fr
Isabelle Sanchez: isabelle.sanchez@inra.fr
Claire Brice: claire.bricecostal@gmail.com
Carole Camarasa: carole.camarasa@inra.fr
Jean-Luc Legras: jean-luc.legras@inra.fr
Sylvie Dequin: sylvie.dequin@inra.fr

Abstract

Background: The volatile metabolites produced by *Saccharomyces cerevisiae* during alcoholic fermentation, which are mainly esters, higher alcohols and organic acids, play a vital role in the quality and perception of fermented beverages, such as wine. Although the metabolic pathways and genes behind yeast fermentative aroma formation are well described, little is known about the genetic mechanisms underlying variations between strains in the production of these aroma compounds.

To increase our knowledge about the links between genetic variation and volatile production, we performed quantitative trait locus (QTL) mapping using 130 F2-meiotic segregants from two *S. cerevisiae* wine strains. The segregants were individually genotyped by next-generation sequencing and separately phenotyped during wine fermentation.

Results: Using different QTL mapping strategies, we were able to identify 65 QTLs in the genome, including 55 that influence the formation of 30 volatile secondary metabolites, 14 with an effect on sugar consumption and central carbon metabolite production, and 7 influencing fermentation parameters. For ethyl lactate, ethyl octanoate and propanol formation, we discovered 2 interacting QTLs each. Within 9 of the detected regions, we validated the contribution of 13 genes in the observed phenotypic variation by reciprocal hemizygosity analysis. These genes are involved in nitrogen uptake and metabolism (*AGP1, ALP1, ILV6, LEU9*), central carbon metabolism (*HXT3, MAE1*), fatty acid synthesis (*FAS1*) and regulation (*AGP2, IXR1, NRG1, RGS2, RGT1, SIR2*) and explain variations in the production of characteristic sensorial esters (e.g., 2-phenylethyl acetate, 2-metyhlpropyl acetate and ethyl hexanoate), higher alcohols and fatty acids.

Conclusions: The detection of QTLs and their interactions emphasizes the complexity of yeast fermentative aroma formation. The validation of underlying allelic variants increases knowledge about genetic variation impacting metabolic pathways that lead to the synthesis of sensorial important compounds. As a result, this work lays the foundation for tailoring *S. cerevisiae* strains with optimized volatile metabolite production for fermented beverages and other biotechnological applications.

Keywords: yeast, aroma compounds, metabolites, QTL mapping, fermentation

1 Background

The aroma of fermented beverages is the result of a complex blend of volatile compounds. In wine, these volatiles originate either directly from grape must or are produced *de novo* by yeast during alcoholic fermentation. Yeast utilizes the nutrients contained in grape must, which are mainly hexoses, nitrogen and lipid sources, for proliferation, whereas ethanol, CO_2 and various minor metabolites are produced as byproducts. Many of these metabolites are volatile with sensorial properties, which give wine its vinous character (Romano et al., 2003). Although the flavor and aroma profile of wine is influenced by vine environment and management techniques, the choice of yeast strain plays a central role (Robinson et al., 2011).

Higher alcohols and esters are the most abundant groups of fermentative aromas that are produced *de novo* in yeast metabolism (Cordente et al., 2012). Higher alcohols can impart a strong, pungent smell and taste when present in higher concentrations, but they result in a fruity character at low concentrations (Lambrechts and Pretorius, 2000). The formation of higher alcohols is carried out through decarboxylation and reduction of *α*-keto-acids, which derive either from central carbon metabolism or from the transamination of amino acids. Therefore, the synthesis of higher alcohols is linked to both carbon and nitrogen metabolism.

Acetate esters, which are produced by yeast from higher alcohols during fermentation, increase aroma complexity by imparting aromatic notes of pear, apple and banana to general fruitiness (Nykänen, 1986; Nykänen and Suomalainen, 1983). They are synthesized through acetyl transfer from acetyl-CoA to an alcohol by the acetyltransferases Atf1 and Atf2 (Mason and Dufour, 2000). Ethyl esters also contribute to global fruitiness perception and are synthesized through acyl transfer from an acyl-CoA to ethanol by the esterases Eeb1 and Eht1 (Saerens et al., 2006). The carboxylic acid molecules for ethyl ester synthesis predominantly originate from the degradation of *α*-keto-acids or fatty acid synthesis in lipid metabolism.

As a consequence, fermentative aroma can be seen as a complex mix of volatile compounds intimately associated with yeast metabolism. The diversity of yeast strains and variability in the regulation of yeast metabolism have a large impact on their production (Rossouw et al., 2012). Even though biochemical pathways have been established for most of these compounds and major genetic determinants have been identified, the genetic basis for the variation of volatile compound production between strains remains largely unknown.

The formation of several compounds important for wine aroma has been shown to be a quantitative and complex trait, as it is influenced by the contribution of multiple genes (Steyer et al., 2012). Quantitative trait locus (QTL) mapping has been demonstrated to be a powerful approach for deciphering the genetic bases of numerous complex traits (Deutschbauer and Davis, 2005; Steinmetz et al., 2002) and has been applied in several biotechnological applications. From its first use in enological studies to characterize allelic variants influencing acetic acid production (Marullo et al., 2007), it has been extended to decipher the genomic bases of fermentation parameters (Ambroset et al., 2011), the production of main metabolites

and residual sugar concentrations (Salinas et al., 2012), nitrogen utilization (Brice et al., 2014a), sulfite production (Noble et al., 2015) and secondary fermentation (Martí-raga et al., 2017). QTL mapping was also used for the detection of genomic regions influencing the production of volatile compounds by yeast during wine alcoholic fermentation (Steyer et al., 2012) using a population of 30 F1-segregants originating from a cross between an *S. cerevisiae* wine and a lab strain. One major QTL and seven minor genomic regions were found to influence the production of different volatile compounds despite high trait heritability. This result suggested that more analytical power is required in order to decipher the genetic bases of the production of volatiles during alcoholic fermentation. The sensitivity of QTL analysis and the ability to find loci with small contributions to phenotype variations can be increased by assessing a larger number of individuals (Bloom et al., 2015). Moreover, the resolution of the mapping can be improved and nearby QTLs can be unlinked by increasing the recombination rate of the segregants (García-Ríos et al., 2017). When multiple loci influence one trait, their contribution to trait variation can either be additive or interacting. Recent studies with a large yeast cross estimated that more than 40% of trait variations in a set of 20 traits could be explained by additive genetic effects, whereas pair-wise genetic interactions contributed to almost 10% of the phenotypic variance (Bloom et al., 2015). Multiple QTL mapping can not only detect linked QTLs but also provides more statistical power to find unlinked QTLs (Arends et al., 2010).

In this study, we addressed the complexity of the genetic basis underlying volatile metabolite production using a population of 130 F2-segregants obtained from a cross of two wine strains with different requirements for nitrogen (Brice et al., 2014b). In addition to performing a genome search for single QTLs, the large segregant population enabled us to increase the analytical strength by performing a search for multiple QTLs. As far as we know, this study is the first analysis of the interaction between loci influencing fermentative aroma formation. We identified a total of 65 QTLs in the genome that influence fermentation parameters and the production of metabolites, including 55 QTLs influencing the formation of 30 volatile secondary metabolites. For the production of ethyl lactate, ethyl octanoate and propanol, we could detect interacting QTLs. Finally, we experimentally validated the role of 13 genes in 9 of the identified genomic regions. These findings provide new information about the production of metabolites of interest due to their sensorial properties or other biotechnological value, such as medium chain fatty acids, fusel acids, higher alcohols and their esters. This opens new perspectives for engineering *S. cerevisiae* strains for broad biotechnological applications.

2 Materials and Methods

2.1 Media

Yeast was cultured at 28 °C in yeast extract peptone dextrose (YPD) media containing 10 g/L yeast extract, 20 g/L peptone and 20 g/L glucose. Solid YPD media contained 1.5% agar.

Selective YPD media containing 200 µg/mL geneticin (G418), 200 µg/mL nourseothricin (clonNAT) or 200 µg/mL hygromycin B were used.

Wine fermentations were carried out in synthetic must (SM) described by Bely et al. (1990). The medium contains glucose and fructose (each 100 g/L) and assimilable nitrogen (200 mg/L) in the form of ammonium and free amino acids, which mimics the nitrogen content of standard grape juice.

2.2 Yeast strains

The haploid *S. cerevisiae* strains MTF2621 (4CAR1 [Δ*HO*::*Neo*^{r}]) and MTF2622 (T73 [Δ*HO*::*Nat*^{r}]) exhibit different needs for nitrogen during wine fermentation, which may indicate different formations of aromas associated with nitrogen metabolism. The requirement was previously estimated using an approach based on the addition of nitrogen to keep the CO_2 production rate constant during limitation of this substrate (Brice et al., 2014b). The strain T73 belongs to the phylogenetic clade of wine strains, whereas strain 4CAR1 belongs to the group of champagne strains (Additional file 1), which originated through crossings between strains of the wine clade and the flor clade (Coi et al., 2016).

2.3 Generation of F2-segregants

Haploid spores of the strains MTF2621 and MTF2622 were mated to form a zygote, which was selected on YPD-agar containing G418 and clonNAT. The zygote was then sporulated in liquid sporulation media using the protocol of Codon et al. (1995). The resulting tetrads were dissected into single spores to obtain the F1-generation using a Singer MSM 400 workstation (Singer Instruments). In most cases, only one spore per tetrad was taken for further experiments to increase genomic independence among the spores. The antibiotic resistance of the obtained spores was determined by growth assay on YPD-agar plates containing G418 or clonNAT. Two spores with different antibiotic properties were mated, and the formed zygotes were subsequently selected on YPD-agar containing G418 and clonNAT. These zygotes were sporulated and dissected again. In total, 130 single spores from the F2-generation were used for this study.

2.4 Genotyping of strains

The genomic DNA of all 130 F2-segregants and both parent strains was isolated using the MasterPure™ Yeast DNA Purification Kit (epicentre) according to the protocol. The purity of the DNA was measured using a NanoDrop™ device (Thermo Fisher Scientific), and the concentration was determined by Qubit™ fluorometric quantification (Thermo Fisher

Scientific). The DNA samples were then used for sequencing using Illumina technology (HiSeq 2500, paired end, 2x100 bp, sequencing platform Genotoul in Toulouse, France) at a sequencing depth of 20- to 80-fold. For each library, low-quality reads were processed and filtered with the FASTX Toolkit v0.0.13.2 and TRIMMOMATIC v0.30 (Bolger et al., 2014) using a quality threshold of 20. First, reads were aligned to the S288C reference genome (release R64-1-1) using BWA v0.6.2 (Li and Durbin, 2009). Once the reads were mapped, consensus genotype calling was performed using the tools available in the SAMtools package (Li et al., 2009). The global set of variants obtained in VCF format contained 18155 biallelic variant positions with a genotyping quality greater than 100. The effect of SNPs on putative transcription factor binding sites was analyzed using YEASTRACT (release 2017; Teixeira et al., 2013). For the location of SNPs in annotated protein domains, information available in the *Saccharomyces* Genome Database (www.yeastgenome.org) was used. The initial variant set was filtered to ensure a minimum spacing of 2.0 kb between SNPs. This resulted in a genotyping variant dataset of 3727 SNP markers (Additional file 5). To increase the meaningfulness of the analysis, four strains with the most ambiguous markers were excluded. One strain was excluded because it was too close in genomic proximity to another segregant. This left a population of 125 F2-segregants for statistical analyses.

2.5 Phenotyping of strains

Segregants were fermented in duplicate with the parent strains as controls. The strains were grown overnight in 50 mL of YPD media. The cell density was determined using a Multisizer™ 3 Coulter Counter (Beckman Coulter). Sterilized 300-mL glassware mini fermenters were filled with 280 mL of SM200 and closed with an air lock. The fermenters were inoculated to a cell density of 1×10^6 cells/mL, weighed and left at 24 °C under stirring (300 rpm).

To determine the concentration of aroma compounds, a sample was taken when approx. 80% of the sugars were depleted. This corresponded to 67.9 – 75 g/L produced CO_2 and was determined by weighting the fermenters regularly to draw the weight decrease caused by the release of CO_2. Volatiles were extracted with dichloromethane according to the method described by Rollero et al. (2015). The concentrations of fermentative aromas were measured via GC/MS on full scan mode using a DB-WAX 60 m GC column. Thirty-four compounds were quantified using internal deuterated standards. In addition, the concentrations of extracellular metabolites after 80% of the fermentation were measured using HPLC (REZEX™ ROA-Organic Acid H+ (8%), 0.005 M H_2SO_4).

2.6 QTL mapping

The data obtained from phenotyping and genotyping were used to identify QTLs in the genome of yeast strains that influence the formation of volatile secondary metabolites during

wine fermentation. Furthermore, QTLs influencing fermentation parameters, substrate consumption and the production of extracellular main metabolites were examined. The statistical analyses were performed using the programming language R v3.2.3 (www.r-project.org) with the R/qtl v1.40-8 and R/eqtl v1.1-7 libraries (Broman et al., 2003). QTL mapping was performed with two different phenotype models, the normal model using Haley-Knott regression and a non-parametric analysis, resulting in logarithm of odds (LOD) scores for each marker and pseudo-marker every 2.5 cM (interval mapping method). An interval estimate of the location of each QTL was obtained as the 1-LOD support interval, the region in which the LOD score is within 1 unit of the peak LOD score. If the same locus was detected with both models, the results with the higher LOD score were selected. A two-dimensional, two-QTL scan was performed using the function *scantwo*. Multiple QTL mapping was performed twice with the function *stepwiseqtl*, once with strictly additive models and once with models that allowed for interactions. The limit of detected QTLs was set to 5. Newly detected QTL positions were counted when the LOD scores of models including these loci were higher than the added LOD score penalties of combining all loci of the respective model. For each method used, individual LOD score thresholds for a false discovery rate of 0.05 were determined with 1000 permutations. QTL mapping results for single traits were grouped as common QTL regions if their peaks were less than 10 cM apart. Proposed models of interaction were further assessed with the function *fitqtl*. The support of individual terms was evaluated by dropping each QTL from the proposed model, one at a time, and comparing the resulting models to the full model.

2.7 Reciprocal hemizygosity analysis

Validation of found QTLs was performed using reciprocal hemizygosity analysis (RHA) (Steinmetz et al., 2002). QTLs were either evaluated as a whole region or single genes with a potential influence on the trait were tested. For the deletion of selected regions, the parent strains were mated to form the heterozygote. Subsequently, one allele of the region was deleted randomly by homologous recombination with a disruption cassette containing the hygromycin B resistance gene (*hph*^{r}) that was obtained by PCR of the plasmid pAG32 (addgene) with the primers del_(QTL)_fw and del_(QTL)_rv (Additional file 2). Positive integration was selected by plating the transformed cells on YPD-agar plates containing hygromycin B. Correct deletion of the region was verified by PCR using primer test_(QTL)_fw that binds upstream of the deleted region and primer Hygro_rv that binds within the deletion cassette. The remaining allele of the QTL was identified by allelic PCR using primer test_(QTL)_fw that binds upstream of a selected gene in the hemizygous region and primers tal_(QTL)_1 or tal_(QTL)_2 that bind at a SNP position within the same gene.

For the deletion of single genes, the sequences were deleted in both parent strains by homologous recombination with a disruption cassette containing the hygromycin B resistance gene (*hph*^{r}) that was obtained by PCR of the plasmid pAG32 with the primers del_(GENE)_fw

and del_(GENE)_rv. Positive integration was selected by plating the transformed cells on YPD-agar plates containing hygromycin B. Correct deletion of the gene was verified by PCR using primer test_(GENE)_fw that binds upstream of the deleted gene and primer Hygro_rv that binds within the deletion cassette. Deleted parent strains were subsequently mated with the opposite undeleted parent to form a heterozygote that is hemizygous for the target gene.

Hemizygous constructions were phenotyped in triplicate (2.5). The significance of the influence of an allelic target region or gene variant on the trait was evaluated by student's t-test. If the impact of a variant on several traits was tested, p-values were not adjusted for multiple comparisons.

3 Results and Discussion

3.1 Phenotyping of strains

Using small scale fermenters, the F2-segregant population and both parental strains were phenotyped (Additional file 3 and Additional file 4) for the production of 43 extracellular metabolites that originate from nitrogen and central carbon metabolism (Figure 18).

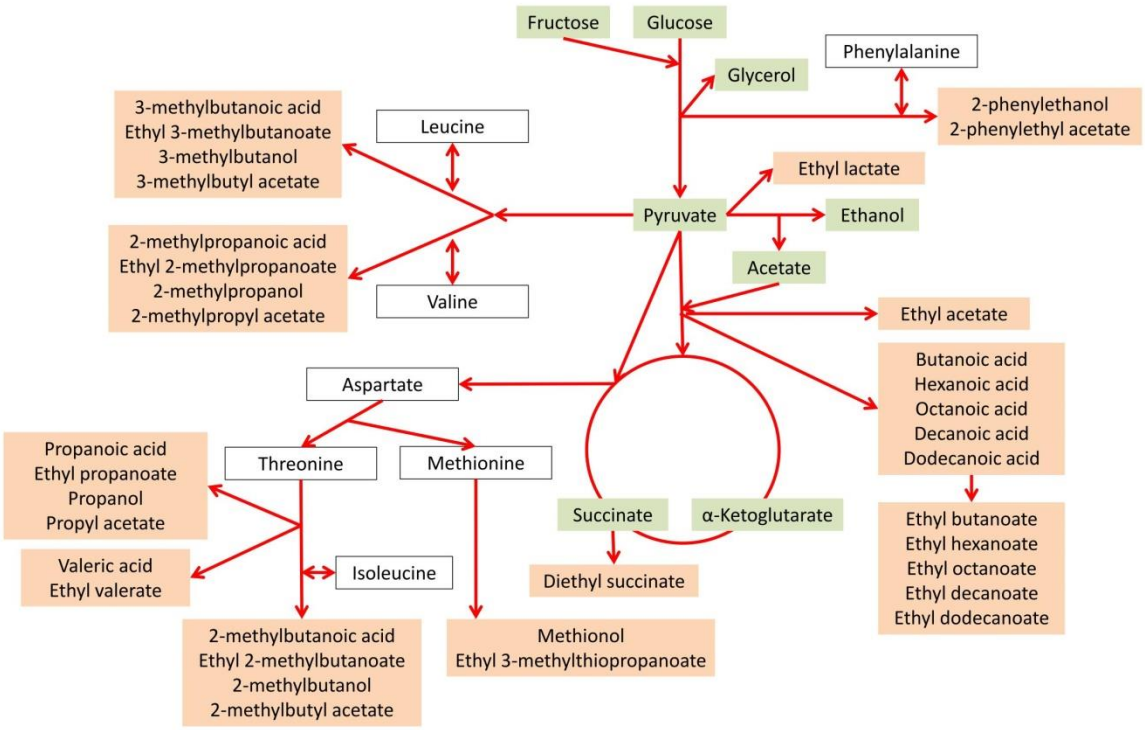


Figure 18: Simplified synthesis pathways of determined metabolites. Main and secondary metabolites determined in this study by HPLC (green) and GC-MS (red).

Most traits are normally distributed among the population, indicating that they are under polygenic control (Additional file 4). One exception is the ratio of glucose to fructose after 80% of the fermentation (G/F ratio), which shows a biphasic distribution, revealing the major influence of one locus for this trait. The phenotypes of the parental strains are located within the population of segregants, indicating transgression for most traits, which can be explained by the presence of alleles with opposite impacts on these traits in the parental genomes.

Heritability of the traits was calculated according to Brem et al. (2002) (Additional file 3). With a median of 70.09 and a maximum of 94.35, the determined heritability is high, indicating reproducible phenotyping and a strong genomic influence on trait variations. For the formation of 3-methylbutanol, decanoic acid, diethyl succinate, dodecanoic acid and ethyl dodecanoate, the heritability estimate is almost zero or negative, which might be associated with insufficient analytical reproducibility. We performed a principal component analysis (PCA) to reduce the complexity of the data set for the determined secondary metabolites and to estimate the potential common regulations (Figure 19).

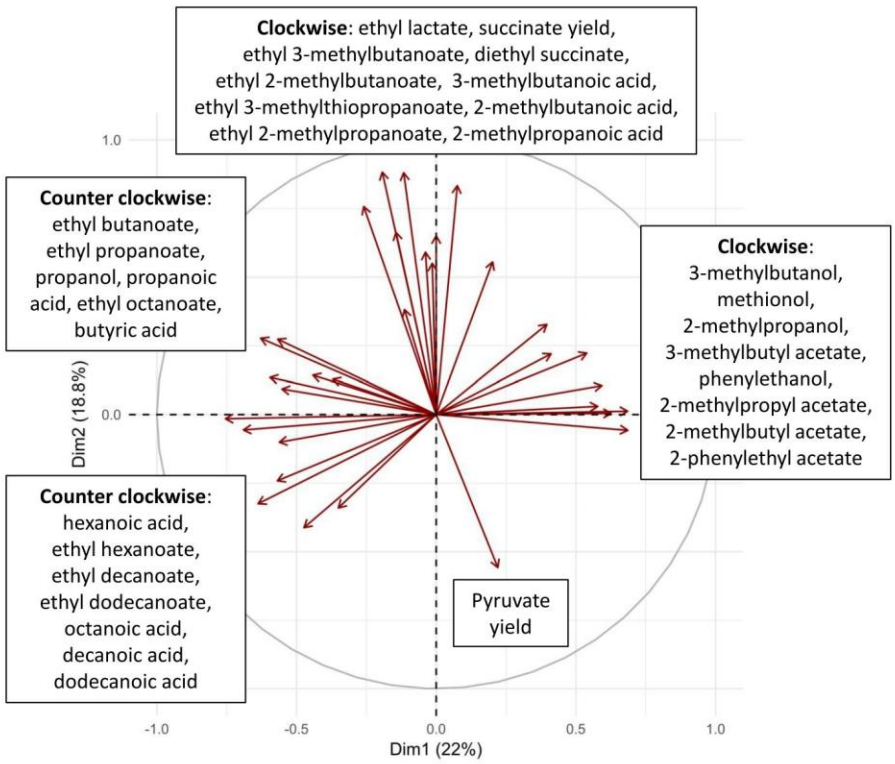


Figure 19: Principle component analysis. PCA for the formation of extracellular metabolites by *S. cerevisiae*. Traits that are less than 2% explained by the first two dimensions of the PCA were excluded (2-methylbutanol, acetate yield, alpha-ketoglutarate yield, ethanol yield, ethyl acetate, glycerol yield, propyl acetate, and valeric acid).

The first two dimensions of the PCA together explain 40.8% of global trait variance. It can be seen that several evaluated compounds are grouped according to their chemical family. The

production of all higher alcohols (except propanol and 2-methylbutanol) is correlated, together with the formation of their corresponding acetate esters. The synthesis of these molecules shares a common reaction step, i.e., the decarboxylation and reduction of α-keto-acids. Another linked group of volatiles is medium chain fatty acids with their ethyl esters. These compounds share a common pathway, namely, fatty acid synthesis. Interestingly, the formation of fusel acids is not correlated to the production of higher alcohols, although both compounds are metabolized from α-keto-acids. This suggests that the reduction or oxidation reactions, which lead to the formation of these compounds from fusel aldehydes, have a strong impact. The pyruvate yield is strongly negatively correlated to the production of ethyl lactate and loosely negatively correlated to the succinate yield and the formation of diethyl succinate. Pyruvate is a metabolic intermediate of both ethyl lactate and succinate formation (Figure 18).

3.2 Genome-wide identification of QTLs influencing fermentation parameters, main and secondary metabolite production

3.2.1 Simple QTL scan

The data obtained from phenotyping and the constructed marker map were used to perform a linkage analysis on 43 quantitative traits, including fermentation parameters, the production of main metabolites and the formation of volatile secondary metabolites. We were able to find a total of 32 QTLs influencing 32 traits (Table 16).

Table 16: QTLs detected with single QTL mapping. List of QTLs with an influence on fermentation parameters, the production of extracellular metabolites and volatile secondary metabolites that were detected with single QTL mapping. QTLs containing several single trait results with a peak distance < 10 cM are numbered in superscript.

	Trait	QTL name	Chromosome	Start position [bp]	End position [bp]	LOD Score
	Fermentation parameters					
[8]	CO_2 production rate at $t_{80\%}$	chr4@385.5	IV	1134839	1173812	4.77
[9]	CO_2 production rate at $t_{80\%}$	chr4@410.0	IV	1198692	1246959	3.69
	CO_2 production rate at $t_{80\%}$	chr10@241.5	X	717987	726938	4.67
[20]	Fermentation time $t_{80\%}$	chr13@7.9	XIII	20503	25723	4.86
[20]	CO_2 production rate at $t_{80\%}$	chr13@7.9	XIII	20503	25723	3.81
	Extracellular metabolites after 80% of the fermentation					
[2]	Pyruvate yield	chr2@172.5	II	507274	527387	3.54
[8]	G/F ratio	chr4@386.5	IV	1153678	1173812	10.49
[9]	G/F ratio	chr4@412.2	IV	1205742	1243242	5.49
[10]	Pyruvate yield	chr7@20.4	VII	56448	74414	6.15
[11]	Pyruvate yield	chr7@156.9	VII	463981	503880	3.75
[14]	Pyruvate yield	chr9@58.7	IX	173782	179168	4.00
[20]	Pyruvate yield	chr13@7.9	XIII	20503	25723	4.80

	Glycerol yield	chr13@19.3	XIII	52743	75040	4.82
	Volatile secondary metabolites after 80% of the fermentation					
[2]	Ethyl butanoate	chr2@166.4	II	488757	506771	3.96
	2-methylpropanoic acid	chr3@26.1	III	62518	111639	3.96
	Ethyl 2-methylbutanoate	chr4@71.3	IV	211091	234153	4.32
[3]	Ethyl 2-methylpropanoate	chr4@82.9	IV	216058	273735	3.96
[4]	Methionol	chr4@124.6	IV	365865	380035	4.37
[4]	3-methylbutyl acetate	chr4@133.6	IV	376106	407166	4.37
[4]	2-methylbutyl acetate	chr4@133.6	IV	397927	407166	5.86
[5]	2-phenylethyl acetate	chr4@161.9	IV	455335	505548	6.17
[5]	3-methylbutyl acetate	chr4@161.9	IV	455335	505548	4.57
[5]	2-methylbutyl acetate	chr4@161.9	IV	478242	505548	6.75
[6]	Ethyl lactate	chr4@175.0	IV	521776	527398	3.41
[6]	Propanoic acid	chr4@177.5	IV	524924	545742	4
[6]	Propanol	chr4@177.5	IV	527398	539089	5.19
[6]	Propyl acetate	chr4@179.4	IV	527398	560742	3.61
[10]	Diethyl succinate	chr7@15.0	VII	40689	56448	5.44
[10]	Ethyl 3-methylthiopropanoate	chr7@25.5	VII	50239	87729	4.47
[11]	Ethyl lactate	chr7@161.6	VII	458995	518880	3.97
	Dodecanoic acid	chr7@175.5	VII	494396	548880	4.76
[12]	Dodecanoic acid	chr7@195.5	VII	578880	601380	3.68
	2-phenylethyl acetate	chr7@294.6	VII	853536	885989	4.14
[17]	Octanoic acid	chr11@29.5	XI	77969	117578	3.99
[17]	Hexanoic acid	chr11@29.5	XI	82548	115238	6.94
[17]	Ethyl hexanoate	chr11@35.5	XI	97410	115238	5.43
[17]	Ethyl dodecanoate	chr11@41.8	XI	115238	145211	4.29
[17]	2-phenylethyl acetate	chr11@42.5	XI	125453	132044	4.29
[18]	2-methylpropyl acetate	chr11@123.8	XI	366406	391690	6.63
[18]	2-methylpropanol	chr11@127.6	XI	371345	400712	8.44
[18]	3-methylbutanol	chr11@132.7	XI	380437	403181	3.56
[18]	2-methylpropanoic acid	chr11@134.4	XI	400712	405331	4.64
	2-methylpropanol	chr11@158.4	XI	470852	477578	3.8
[19]	Propyl acetate	chr12@222.6	XII	662035	699182	4.83
[19]	Propanol	chr12@226.9	XII	662035	691268	4.08
	Valeric acid	chr13@102.0	XIII	296288	312983	5.87
[21]	Propanoic acid	chr14@41.9	XIV	119900	146614	4.67
[21]	Propanol	chr14@43.9	XIV	119900	146614	7.03
[21]	Propyl acetate	chr14@46.9	XIV	124114	150370	6.19
[21]	Valeric acid	chr14@48.9	XIV	125823	150370	7.72
	Propanoic acid	chr14@58.9	XIV	160420	185626	3.83
	Valeric acid	chr14@81.8	XIV	233520	259114	3.88
	Ethyl 3-methylbutanoate	chr15@77.3	XV	212898	239482	3.76
[22]	Ethyl decanoate	chr15@139.0	XV	409364	431700	3.97
[22]	Ethyl octanoate	chr15@142.3	XV	414810	438628	5.02
[23]	3-methylbutanoic acid	chr15@162.7	XV	485607	511993	4.29
[23]	2-phenylethyl acetate	chr15@176.5	XV	511993	545871	3.56
	Diethyl succinate	chr15@297.3	XV	879033	901450	4.46
	3-methylbutanoic acid	chr16@191.9	XVI	552371	593439	3.83

[25]	Diethyl succinate	chr16@303.9	XVI	899570	920003	4.08
[25]	Ethyl 3-methylbutanoate	chr16@303.9	XVI	904961	917224	3.81
[25]	Ethyl 2-methylbutanoate	chr16@304.1	XVI	904961	917224	3.79

The determined logarithm of odds (LOD) scores ranged from 3.41 to 10.49 with a median of 4.35. The highest LOD score was found for chr4@386.5, influencing the G/F ratio. The second highest LOD score of 8.44 was found for a QTL influencing the production of a volatile compound, namely, chr11@127.6 affecting the formation of 2-methylpropanol. Six major QTL regions were detected with LOD scores greater than six for at least one trait, which corresponds to an explained trait variation of more than 20% by these loci.

Globally, these QTLs were distributed over the whole genome, with exception of chromosomes I, V, VI and VIII. The size of the identified regions ranged from 5.2 kb to 65.7 kb with a median of 33.2 kb. The detected regions contained between 4 and 28 genes. Four QTLs were detected for both evaluated traits of the fermentation kinetics, the fermentation time ($t_{80\%}$) and CO_2 production rate at $t_{80\%}$. Eight QTLs were found for the concentration of extracellular main metabolites at $t_{80\%}$. These QTLs influenced three traits, which were glycerol yield, pyruvate yield and the G/F ratio. The most QTLs were detected for the formation of volatile secondary metabolites, namely, 28 QTLs influencing the formation of 27 volatiles. This included the production of characteristic sensorial compounds, such as 2-methylbutyl acetate, 3-methylbutanol and 3-methylbutyl acetate, and industrially relevant chemicals, such as higher alcohols and organic acids.

The detected QTLs were compared with loci found in QTL mapping studies of similar traits (Abt et al., 2016; Ambroset et al., 2011; Marullo et al., 2007; Salinas et al., 2012; Steyer et al., 2012). Only QTL chr7@161.6, which influences ethyl lactate formation, co-localizes with *PMA1*, a plasma membrane P2-type H+-ATPase that was shown by Abt et al. (2016) to be the responsible gene in a QTL affecting ethyl acetate production. More QTLs were in common with the findings of Rossouw et al. (2008), who used a comparative approach of combining transcriptomics and exo-metabolome analysis to predict candidate genes with a role in aroma profile modification. Several of the genes proposed by Rossouw et al. (2008), e.g., *ALP1, ILV6, LEU1, LEU2* and *LEU9*, were included in QTLs that we detected for the same or closely related traits.

3.2.2 Double and multiple QTL scan

To search for additional minor QTLs and to assess genetic interactions between our detected regions, we performed a two-dimensional, two-QTL scan and a multiple QTL search for all traits. The analyses confirmed 24 QTLs that were already found with the single QTL mapping and proposed 36 additional loci (Table 17).

Table 17: QTLs detected with double and multiple QTL mapping. QTLs with an influence on fermentation parameters and the production of extracellular metabolites and volatile secondary metabolites that were additionally detected with double and multiple QTL searches. QTLs containing several single trait results with a peak distance < 10 cM are numbered in superscript.

	Trait	QTL name	Chromosome	Peak [bp]	LOD Score
	Diethyl succinate	chr1@64.7	I	194100	4.23
[1]	Ethyl 2-methylbutanoate	chr2@116.2	II	348600	4.13
[1]	Hexanoic acid	chr2@122.1	II	366300	2.54
[2]	Ethyl lactate	chr2@181.0	II	543000	4.78
	Ethyl 2-methylpropanoate	chr3@78.8	III	236400	2.84
[3]	Pyruvate yield	chr4@91.9	IV	275730	2.39
[3]	Ethyl butanoate	chr4@92.0	IV	275985	3.63
[4]	Ethyl hexanoate	chr4@132.0	IV	396000	2.59
[5]	2-methylpropyl acetate	chr4@165.8	IV	497400	2.42
[6]	Diethyl succinate	chr4@175.0	IV	525000	2.99
[6]	Ethyl octanoate	chr4@175.0	IV	525000	2.14
	Dodecanoic acid	chr4@298.0	IV	894000	4.60
	Valeric acid	chr4@324.9	IV	974700	2.50
[7]	2-methylpropanol	chr4@348.9	IV	1046700	3.36
[7]	3-methylbutanol	chr4@348.9	IV	1046730	3.66
[9]	Pyruvate yield	chr4@411.9	IV	1235730	2.54
	G/F ratio	chr6@62.0	VI	186000	3.90
	Ethyl 3-methylbutanoate	chr6@85.2	VI	255600	4.14
[12]	Decanoic acid	chr7@198.0	VII	594000	3.65
[13]	2-methylpropyl acetate	chr9@16.9	IX	50700	2.09
[13]	Dodecanoic acid	chr9@17.6	IX	52800	3.20
	2-methylbutyl acetate	chr9@32.6	IX	97800	1.75
[14]	Dry weight	chr9@60.0	IX	180066	3.67
[14]	Ethyl propanoate	chr9@67.5	IX	202500	2.88
[14]	Propanol	chr9@67.5	IX	202500	1.85
	2-phenylethyl acetate	chr9@101.1	IX	303300	2.25
	Propyl acetate	chr10@27.2	X	81600	1.53
[15]	Ethyl hexanoate	chr10@61.3	X	183900	2.97
[15]	Hexanoic acid	chr10@61.3	X	183900	2.78
	Ethyl octanoate	chr10@88.5	X	265500	5.24
	Propanol	chr10@104.7	X	314100	2.08
[16]	2-methylpropanol	chr10@213.1	X	639300	2.29
[16]	2-methylpropyl acetate	chr10@221.7	X	665100	4.04
[16]	2-methylbutyl acetate	chr10@228.1	X	684300	2.34
[16]	3-methylbutyl acetate	chr10@228.8	X	686400	2.97
[17]	Decanoic acid	chr11@29.2	XI	87600	3.95
[17]	Ethyl decanoate	chr11@34.2	XI	102438	2.65
[17]	$t_{80\%}$	chr11@43.2	XI	129600	3.10
[17]	Ethyl octanoate	chr11@51.7	XI	155100	4.95
	Glycerol yield	chr11@70.0	XI	210000	2.72
	G/F ratio	chr11@85.6	XI	256800	2.65
	Ethyl propanoate	chr11@107.7	XI	323100	3.78
[18]	3-methylbutanoic acid	chr11@134.4	XI	403170	3.68
	2-methylpropyl acetate	chr12@98.3	XII	294900	3.15

	Ethyl 2-methylbutanoate	chr12@123.8	XII	371400	3.34
	Ethyl propanoate	chr12@139.6	XII	418800	3.67
	CO_2 production rate at $t_{80\%}$	chr12@257.3	XII	771900	2.47
	2-methylpropanol	chr12@317.1	XII	951300	3.08
	G/F ratio	chr13@164.5	XIII	493500	4.99
	Propyl acetate	chr13@248.3	XIII	744900	2.11
	2-phenylethyl acetate	chr13@304.0	XIII	912000	2.55
[21]	Dodecanoic acid	chr14@36.3	XIV	108900	4.69
	Ethyl 3-methylthiopropanoate	chr14@227.0	XIV	681000	3.42
[22]	2-phenylethanol	chr15@132.0	XV	396000	1.62
[23]	2-methylbutyl acetate	chr15@172.0	XV	516000	2.55
	Ethyl hexanoate	chr15@265.4	XV	796200	4.47
	Ethyl 3-methylbutanoate	chr15@304.4	XV	913200	2.77
	2-methylpropanoic acid	chr15@314.4	XV	943200	2.95
	2-phenylethyl acetate	chr15@352.0	XV	1056000	2.85
[24]	2-methylpropanol	chr16@13.8	XVI	41400	4.06
[24]	2-methylbutyl acetate	chr16@16.4	XVI	49200	2.78
[24]	G/F ratio	chr16@18.3	XVI	54900	2.75
	Valeric acid	chr16@116.5	XVI	349500	2.04
[25]	Ethyl lactate	chr16@304.0	XVI	912000	7.00

With the double QTL mapping, we found significant evidence for an interacting QTL pair at positions chr10@88.5 and chr11@51.7, influencing the formation of ethyl octanoate and accounting for 10.29% of trait variation. Another interacting QTL pair, chr2@181.0 and chr16@304.0, was found to influence the production of ethyl lactate, explaining 10.17% of trait variation. Multiple QTL mapping proposed interacting regions for a wide range of traits. However, their contributions to the respective phenotypes were low, with LOD scores of generally less than two. Due to penalization of the LOD score for more complex models of interaction, solely additive models were found to be more significant for all traits. This indicates that our number of segregants was still insufficient to achieve the statistical power required for the determination of QTL interactions. However, for the production of propanol the three involved QTLs, chr4@176.9, chr12@233.1 and chr14@45.8, could be detected with all three mapping strategies, giving strong evidence for their validity. Although models consisting of viewer QTLs were seen to be more probable by the multiple QTL mapping, the most likely model containing all three loci was an additive model with an interaction between chr4@176.9 and chr12@233.1. This indicates a remaining probability for the proposed interaction, which was calculated to potentially account for 5.58% of trait variation.

Combining the results from the single and multiple QTL mapping, each trait is influenced by a median of 3 QTLs, ranging from 1 to 7. The best explained trait is the G/F ratio with 6 detected loci accounting for 61.0% of determined trait variation. Regarding volatile formation, the difference in 2-methylpropanol production can be best elucidated, with 6 identified QTLs explaining 54.5% of trait variation. Several QTL display pleiotropic effects, as they influence many traits and can therefore be considered “hotspots”. The region affecting the most traits

is chrXI:77,969..155,100, which influences the fermentation time ($t_{80\%}$) and the formation of eight volatile compounds, namely, hexanoic acid, octanoic acid, decanoic acid, and their corresponding ethyl esters as well as ethyl dodecanoate and 2-phenylethyl acetate. Twenty-four other QTLs were found to influence more than one trait. This is often the case for the production of related compounds and indicates common regulation, which was already concluded by PCA. Examples for jointly influenced traits are sugar consumption and the CO_2 production rate, pyruvate yield and the formation of ethyl esters, the production of 2-methylpropanol and 3-methylbutanol, and the formation of several acetate esters.

3.3 Validation of genomic regions involved in metabolic traits

For the validation of single QTLs and the identification of impacting allelic variants within the corresponding region, 19 genes in 10 QTLs were further evaluated using RHA (Table 18). These target genes were chosen since they contained non-synonymous SNPs between the parent cells and were suspected to play a role in the detected traits, as their biologic functions were mostly connected to central carbon metabolism or nitrogen uptake and metabolism. In 9 QTLs, we could identify 13 genes that influence hexose transport and the formation of medium chain fatty acids, fusel acids, higher alcohols, and their corresponding esters (Table 18).

Table 18: Validated allelic variants in detected QTLs. Selected target genes for the verification of QTLs influencing fermentation kinetics, substrate consumption and the production of fermentative aromas; differences caused by the allelic gene variants regarding the influenced traits were detected by RHA and are given as the ratio of phenotype MTF2621 to phenotype MTF2622. (p-value: * ≤ 0.05, ** ≤ 0.01, *** ≤ 0.001)

QTL name	Trait	Evaluated genes	Different impact of allele on trait as MTF2621/MTF2622 [factor]
chr2@166.4	Ethyl butanoate	*AGP2*	0.79*** ethyl lactate
	Ethyl lactate		
	Pyruvate yield		
chr3@26.1	2-methylpropanoic acid	*AGP1*	1.26* 2-methylpropanoic acid
		ILV6	no effect
chr4@71.3	Ethyl 2-methylbutanoate	*YDL124W*	no effect
	Ethyl butanoate		
	Ethyl 2-methylpropanoate		
	Pyruvate yield		
chr4@133.6	2-methylbutyl acetate	*SIR2*	0.77* 3-methylbutyl acetate
	3-methylbutyl acetate		0.78** ethyl hexanoate
	Ethyl hexanoate		
	Methionol		
chr4@177.5	Ethyl lactate	*NRG1*	1.10* propanol
	Ethyl octanoate		
	Diethyl succinate		
	Propanoic acid		
	Propanol		

	Propyl acetate		
chr4@386.5	CO_2 production rate at $t_{80\%}$	*HXT3*	1.07* CO_2 production rate
	G/F ratio		1.86** G/F ratio
		HXT6	no effect
		HXT7	no effect
chr11@29.5	2-phenylethyl acetate	*ACP1*	no effect
	Ethyl decanoate	*FAS1*	0.81** ethyl hexanoate
	Ethyl dodecanoate		0.82** decanoic acid
	Ethyl hexanoate		0.84**hexanoic acid
	Ethyl octanoate		0.89** octanoic acid
	Decanoic acid	*FAT3*	no effect
	Hexanoic acid	*PXA2*	no effect
	Octanoic acid		
	$t_{80\%}$		
chr11@127.6	2-methylpropanoic acid	*IXR1*	1.14** 2-methylpropanol
	2-methylpropanol		1.16* 2-methylpropanoic acid
	2-methylpropyl acetate	*MAE1*	1.43** 2-methylpropanoic acid
	3-methylbutanoic acid		1.67*** 2-methylpropanol
	3-methylbutanol		1.27*** 3-methylbutanoic acid
			1.40* 3-methylbutanol
		RGT1	1.15*** 2-methylpropanol
chr12@226.9	Propanol	whole region	no effect
	Propyl acetate		
chr14@48.9	Dodecanoic acid	*ALP1*	0.90* dodecanoic acid
	Propanoic acid		1.07** propanol
	Propanol		1.26*** valeric acid
	Propyl acetate		
	Valeric acid		
chr15@176.5	2-methylbutyl acetate	*LEU9*	1.08* 2-phenylethyl acetate
	2-phenylethyl acetate	*RGS2*	1.21* 2-methylbutyl acetate
	3-methylbutanoic acid		0.83* 2-phenylethyl acetate

Table 19: Non synonymous SNPs between allelic variants. Differences in the amino acid (AA) sequence of the expressed protein resulting from non-synonymous SNPs between the allelic variants of the evaluated target genes. Comparison of the strains MTF2621 and MTF2622 with the *S. cerevisiae* reference strain S288C.

Gene	AA position	S288C	MTF2621	MTF2622
AGP1	7	P	P	L
	24	G	E	G
	142	N	S	N
	316	V	V	A
	370	F	L	F
	466	I	L	I
	540	L	I	L
	597	D	N	D
AGP2	256	H	Y	H
ALP1	126	V	V	A
FAS1	1504	V	A	V
	1715	V	A	V
	1970	V	I	V
ILV6	4	S	L	S
	56	A	P	A

IXR1	45	T	T	A
	65	Q	Q	-
	93	Y	Y	F
	104	-	ATTTTT	-
	291	M	M	L
	570	QQ	QQ	-
LEU9	76	D	D	H
	176	S	S	Y
MAE1	605	I	I	V
NRG1	129	P	H	P
	156	T	S	T
RGS2	99	Y	N	Y
RGT1	326	L	L	P
	717	I	I	V
	722	P	P	A
	729	S	S	N
SIR2	178	Q	Q	H
	201	S	G	S

3.3.1 Hexose transporter Hxt3 influences sugar utilization

Hexose transport is a limiting step for alcoholic fermentation speed (Elbing et al., 2004). QTL chr4@386.5, which influences the CO_2 production rate and G/F ratio, contains three hexose transporter genes, *HXT3*, *HXT6* and *HXT7*. We evaluated these genes individually by RHA. As the sequences of *HXT6* and *HXT7* are nearly identical, we assessed the effect of both genes together. Variation in *HXT3* was found as the sole effect influencing the CO_2 production rate and the G/F ratio (Additional file 7). The MTF2621 allele of the gene increased the CO_2 production rate by a factor of 1.07 and increased the G/F ratio by a factor of 1.86. An effect of this allelic variation on the production of determined volatiles could not be detected. A variant of *HXT3* has already been described in the literature by Guillaume et al. (2007) to have a higher affinity for fructose and was detected among flor strains (Coi et al., 2016). This variant originated through recombination between the orthologs *HXT1* and *HXT3*. Except for SNP T1411A, which results in amino acid change L471I, the MTF2621 allele of *HXT3* is identical to the variant described by Guillaume et al. (2007).

3.3.2 The formation of medium chain fatty acids and their ethyl esters is influenced by Agp2, Fas1 and Sir2

Ethyl esters of medium chain fatty acids provide floral and fruity notes to fermented beverages. In QTL chr4@133.6 and QTL hotspot chr11@29.5, which influence the formation of medium chain fatty acids and their ethyl esters, we identified *SIR2* and *FAS1* as causative

genes (Table 18). In chr2@166.4, a QTL impacting ethyl butanoate production with a lower significance (LOD 3.96), *AGP2* was found to modulate the formation of butanoic acid, the substrate for ethyl butanoate.

Fatty acids are synthesized from the repeated condensation of malonyl-CoA and acetyl-CoA, which is carried out by fatty acid synthetase (FAS). The FAS complex consists of the beta subunit Fas1 and the alpha subunit Fas2 (Kolodziej et al., 1996). Fas1, which was found to regulate the expression of *FAS2* (Wenz et al., 2001), possesses four independent enzymatic functions, i.e., acetyl transferase, enol reductase, dehydratase and malonyl/palmitoyl transferase (Schweizer et al., 1986). The parental allelic variants of *FAS1* differ in three non-synonymous SNPs (Table 19), of which one, SNP I1970V, lies in the malonyl-CoA-acyl carrier protein transacylase domain of the protein. The MTF2621 allele of Fas1 causes a significant decrease in the formation of hexanoic acid, ethyl hexanoate, valeric acid, octanoic acid, decanoic acid and dodecanoic acid by a factor of 0.78 – 0.89 (Figure 20). Therefore, we can suggest that the MTF2621 allele of *FAS1* is less active than the MTF2622 allele, and thus leads to a decreased synthesis rate of fatty acids.

The gene *SIR2* encodes an NAD^+-dependent deacetylase involved in chromatin silencing (Rine and Herskowitz, 1987). The allelic variants of *SIR2* differ in two non-synonymous SNPs (Table 19). The MTF2621 allele of the gene causes a decrease in the formation of hexanol, octanoic acid, decanoic acid, dodecanoic acid, ethyl butanoate and ethyl hexanoate up to a factor of 0.57 (Figure 20). In addition, the extracellular concentration of acetate was decreased by a factor of 0.8 (Additional file 7). Sir2 was found to influence the expression of the acetyl-CoA synthase *ACS2* (Lin et al., 2008), and a regulating function by Sir2 on the activity of acetyl-CoA synthase enzymes by deacetylation was proposed by Starai et al. (2003). Furthermore, Casatta et al. (2013) demonstrated that a null mutant of *SIR2* showed increased acetate metabolism and a lower excretion of acetate to the medium. Based on our observations, we can therefore suggest that the MTF2621 variant of Sir2 has lower deacetylase activity compared to the other variant, which results in decreased expression of *ACS2* and reduced activation of acetyl-CoA synthases. Consequently, this reduced activity leads to a lower availability of acetate and acetyl-CoA for fatty acid synthesis and elongation.

Other small but significant influences of *SIR2* can be seen in the formation of 2-methylpropanol, 3-methylbutanol and other degradation products of α-keto-acids, with the MTF2621 allele leading to a decrease of these compounds (Figure 21). Sir2 was found to modulate the expression of the amino acid permease *AGP1* (Liu et al., 2010). Altered expression of *AGP1* could impact the nitrogen assimilation and thus the formation of amino acid related fermentative aromas. Furthermore, as Sir2 is dependent on the cofactor NAD^+, its altered activity could influence redox homeostasis of the cell. Redox imbalances were reported to significantly affect the production of fermentative aromas by *S. cerevisiae* (Bloem et al., 2015).

AGP2 encodes a plasma membrane protein that is involved in the uptake of carnitine and polyamines (Schreve and Garrett, 2004; Van Roermund et al., 1999). Carnitine is important for intracellular acetyl transport between cellular compartments (Van Roermund et al., 1995), and the level of carnitine can therefore affect the availability of acetyl-CoA for the fatty acid synthesis. However, carnitine is not present in the synthetic medium used in this study. Agp2 positively regulates the expression of various proteins involved in substrate transport and other biological processes, and might also act as a sensor of environmental signals (Aouida et al., 2013). One non-synonymous SNP was found to distinguish the two parental variants (Table 19), and it is located in the extracellular region of the protein. The MTF2621 allele of *AGP2* causes an increase in the formation of butanoic acid, decanoic acid, dodecanoic acid and ethyl dodecanoate up to a factor of 1.63 (Figure 20). We suggest that the reported SNP in *AGP2* causes a higher formation rate of fatty acids for MTF2621, although the causative function remains unclear.

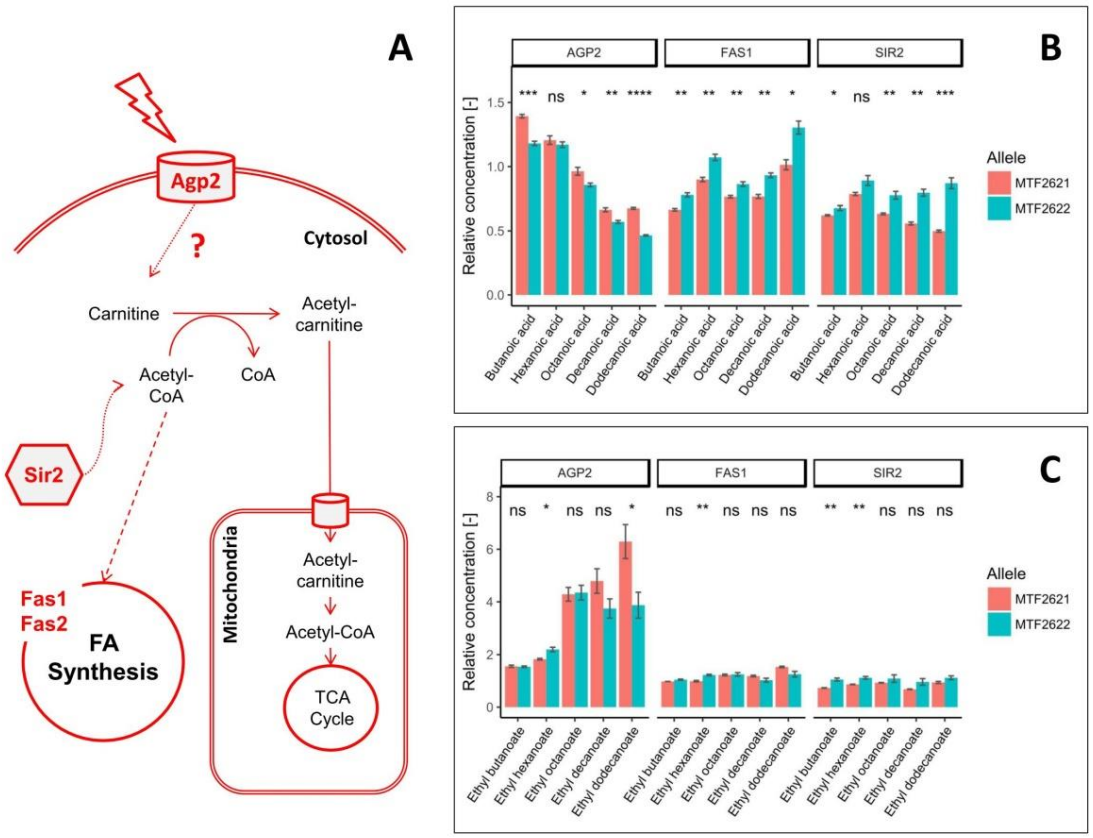


Figure 20: Effect of validated variants on medium chain fatty acid formation. Simplified pathway of fatty acid synthesis by the enzymes Fas1 and Fas2, which is dependent on intracellular acetyl transport (**A**). Allelic effect of the enzymes Agp2, Fas1 and Sir2 on the formation of fatty acids (**B**) and fatty acid ethyl esters (**C**) as determined by RHA. Concentrations are given in relation to the heterozygote of the parental strains MTF2621 and MTF2622. (p-value: * ≤ 0.05, ** ≤ 0.01, *** ≤ 0.001, **** ≤ 0.0001)

3.3.3 The formation of higher alcohols, fusel acids and their esters is influenced by Agp1, Ilv6, Mae1

Higher alcohols, fusel acids and especially their esters are essential fermentative aroma components that provide notes ranging from fruity to flowery. We identified *MAE1* in chr11@127.6, the QTL with the highest LOD score for volatile compounds, which influences the formation of five higher alcohols, fusel acids and acetate esters, (Table 18). The enzyme Mae1 catalyzes the oxidative decarboxylation of malate to pyruvate (Boles et al., 1998). Pyruvate is a precursor for the synthesis of the amino acids alanine, isoleucine, leucine and valine (Umbarger, 1978). An intermediate of valine and leucine biosynthesis, *α*–keto-isovalerate, can also be degraded to 2-methylpropanol and 2-methylpropanoic acid or to 3-methylbutanol and 3-methylbutanoic acid via *α*–keto-isocaproate (Figure 21). The allelic variants of *MAE1* differ in one non-synonymous SNP (Table 19), which is located in the NAD-binding domain of the protein. Furthermore, 5 SNPs in the 1000-bp upstream region of the gene affect predicted binding motifs for the proteins Azf1, Mot3, Rtg1, Rtg3, Stp1 and Stp2 (Additional file 6). The hemizygote carrying only the MTF2621 allele of *MAE1* shows increased formation of 2-methylpropanol, 3-methylbutanol, 2-methylpropanoic acid and 3-methylbutanoic acid by up to a factor of 1.67 (Figure 21). We can suggest that the MTF2621 allele of *MAE1* is superior to the MTF2622 allele and induces an increased flux of malate to pyruvate, leading to higher formation of *α*–keto-acids and their degradation products. This proposal is further supported by an observed increased formation of ethyl lactate (Additional file 7), which is also derived from pyruvate (Figure 18).

RHA detected several other minor influences of *MAE1* on traits that were not found by QTL mapping. The MTF2621 allele of the gene leads to a slightly higher production of 2-phenylethanol by a factor of 1.18 (Additional file 7). Mae1 was found to interact with Aro1 (Gavin et al., 2006), an enzyme catalyzing several steps of the chorismate pathway leading to the synthesis of aromatic amino acids, such as phenylalanine (Duncan et al., 1987). In addition, the MTF2621 allele of Mae1 leads to a decrease of several acetate esters and medium chain fatty acids up to a factor of 0.8 (Figure 21 and Additional file 7) and to an increase in the extracellular concentration of acetate by a factor of 1.28 (Additional file 7). These effects are consistent with the fact that Mae1 physically interacts with Acc1 (Gavin et al., 2006), an acetyl-CoA carboxylase that is involved in the regulation of acetyl-CoA and in the biosynthesis of medium and long chain fatty acids (Galdieri and Vancura, 2012; Mishina et al., 1980).

The gene *AGP1*, which encodes a low affinity amino acid permease for asparagine and glutamine (Schreve et al., 1998), was validated in QTL chr3@26.1 with an influence on the formation of 2-methylpropanoic acid. The two allelic variants of *AGP1* differ in 8 non-synonymous SNPs (Table 19), of which three lie in cytoplasmic domains of the protein and 5 in transmembrane domains. Another SNP is located in the 1000-bp upstream region of the gene, affecting the predicted binding motif for Ume6 (Additional file 6). The hemizygote carrying the MTF2621 allele of the gene shows a formation of 2-methylpropanol, 2-

methylpropyl acetate, 2-methylpropanoic acid and ethyl 2-methylpropanoate increased by a factor of 1.26 – 1.32 (Figure 21). Agp1 was found to transport valine to a lower extent (Schreve et al., 1998). Valine can be degraded to α-keto-isovalerate and then to 2-methylpropanol and 2-methylpropanoic acid by the Ehrlich pathway (Figure 21). We hypothesize that the reported SNPs lead to higher affinity of the MTF2621 allele of Agp1 for valine, causing a higher level of this amino acid in the cell. Another possible explanation is a different influence of the alleles on the transport of glutamine, which is important for the transamination of α-keto-acids in the cell. In this scenario, a reduced intracellular level of glutamine could lead to a decreased transamination of α-keto-isovalerate, which can therefore be degraded to 2-methylpropanol and 2-methylpropanoic acid. While this would also affect the transamination of other α-keto-acids and therefore the production of several higher alcohols or fusel acids, a significant, but small, influence of *AGP1* could only be additionally detected on the production of 2-phenylethanol.

In the same QTL (chr3@26.1), the variants of *ILV6* did not show significant differences in the formation of 2-methylpropanoic acid, but they did in the formation of the related higher alcohol 2-methylpropanol. Ilv6 is a regulatory subunit of the acetolactate synthase Ilv2, which catalyzes the first step of valine and leucine biosynthesis (Cullin et al., 1996). The allelic variants of *ILV6* differ in two non-synonymous SNPs (Table 19). SNP L4S lies in the N-terminal signal peptide domain of the protein, whereas SNP P56A is within a non-cytoplasmic domain. Another SNP is located in the 1000-bp upstream region and causes a loss of the predicted binding motifs for Msn2, Msn4, Nrg1 and Rph1 in strain MTF2621 (Additional file 6). The MTF2621 allele of *ILV6* leads to an increase in the formation of 2-methylpropanol and 3-methylbutanol by a factor of 1.24 (Figure 21). Therefore, we can hypothesize that the MTF2621 allele of *ILV6* stimulates a higher synthesis rate of acetolactate, which could result in higher synthesis of α–keto-isovalerate, including its degradation products.

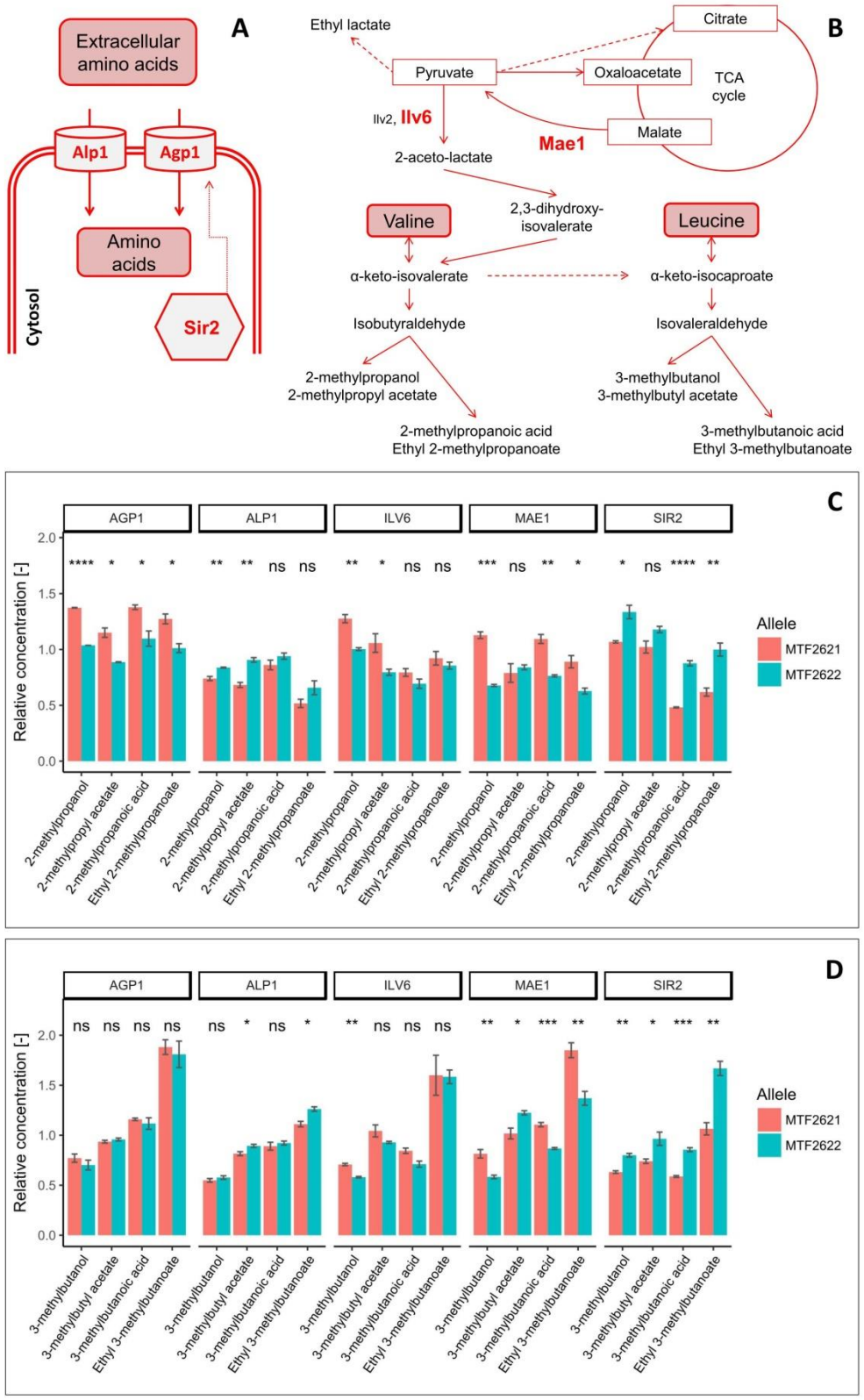


Figure 21: Effect of validated variants on of higher alcohol and fusel acid formation. Amino acids are transported into the cell by Agp1 and Alp1. The expression of *AGP1* is influenced by Sir2 (**A**). Simplified synthesis pathway of fermentative aromas connected to valine and leucine metabolism (**B**). Allelic effect of the involved enzymes Agp1, Alp1, Ilv6, Mae1 and Sir2 on the formation of volatiles deriving from *α*-keto-isovalerate (**C**) and *α*-keto-isocaproate (**D**) as determined by RHA. Concentrations are given in relation to the heterozygote of the parental strains MTF2621 and MTF2622. (p-value: * ≤ 0.05, ** ≤ 0.01, *** ≤ 0.001, **** ≤ 0.0001**)**

3.3.4 The formation of propanol is influenced by Alp1 and Nrg1

We assessed the three partly interacting genomic regions that were detected with the single, double and multiple QTL mapping to affect the production of propanol and related compounds (Table 18). *NRG1* and *ALP1* were validated in QTL chr4@177.5 and QTL chr14@43.9, respectively. As no clear candidate gene was identified in QTL chr12@226.9, we assessed the whole region by RHA; however, no significant impact could be detected for the production of propanol or propyl acetate. The QTL is likely to interact with chr4@177.5, and the negative validation of chr12@226.9 might indicate a possible epistatic interaction. Furthermore, chr12@226.9 is the weakest QTL of the three assessed loci, with an LOD score of 4.08 (Table 16), which could have hindered the validation.

Propanol and propanoic acid derive from the decarboxylation of *α*-keto-butyrate and the oxidation or reduction of the resulting propionaldehyde (Hazelwood et al., 2008). *α*-keto-butyrate is produced from the transamination of threonine, which is taken up from the medium as a nitrogen source or can be metabolized from pyruvate via aspartate through the amino acid pathway (Figure 22). It was shown, however, that the formation of propanol is mainly limited to the beginning of wine fermentation when nitrogen is present in the must and is dependent on the initial amount of available nitrogen (Mouret et al., 2014).

The protein Nrg1 is a transcriptional regulator of glucose repressed genes (Park et al., 1999; Zhou and Winston, 2001) and mediates a set of stress responsive genes (Vyas et al., 2005). The parental allelic variants of *NRG1* differ in two non-synonymous SNPs (Table 19), which are both located in the transcriptional repressor protein “yy” domain. The MTF2621 variant of the gene leads to an increase in propanol production by a factor of 1.10, whereas no significant effect could be detected in the formation of related compounds (Figure 22). The repressive function of Nrg1 is inhibited by Snf1; therefore, it is suspected to have a role in the response to nitrogen limitation (Kuchin et al., 2002). Furthermore, Nrg1 was found to influence the expression of *BAT1*, a mitochondrial aminotransferase involved in branched amino acid synthesis and Ehrlich pathway catabolism (Costanzo et al., 2010). With regard to this finding, we propose that the allelic variants of Nrg1 show a different response to nitrogen limitation, which affects the expression of *BAT1*, leading to a lower transamination rate of valine, leucine and isoleucine taken up from the medium. In this scenario, the overall availability of nitrogen for metabolism would be influenced, which therefore influences the synthesis of propanol.

The gene *ALP1* encodes a permease for cationic amino acids (Regenberg et al., 1999; Sychrova and Chevallier, 1994). The parental variants differ in one non-synonymous SNP (Table 19), which is located in a transmembrane domain. The MTF2621 variant increases the production of propanol by a factor of 1.07 (Figure 22). We can suggest that this variant of Alp1 leads to an increased uptake of amino acids from the medium at the beginning of the fermentation, which explains higher propanol formation. This hypothesis is supported by a significant decrease in fermentative aromas derived from *α*-keto-isovalerate and *α*-keto-isocaproate for the MTF2621 allele of the gene (Figure 21). The opposite correlation is reported in the

literature, in which a lower overall intracellular concentration of nitrogen leads to a higher level of fermentative aroma production due to a lower transamination rate of α-keto-acids derived from central carbon metabolism (Oshita et al., 1995; Vilanova et al., 2007).

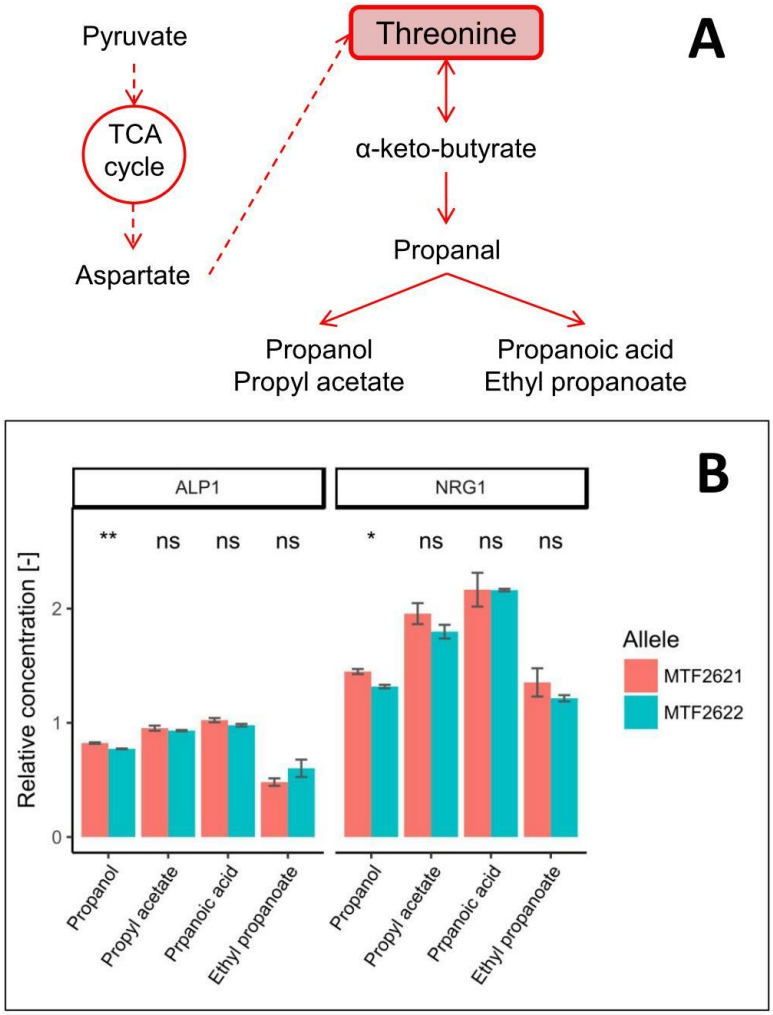


Figure 22: Effect of validated variants on propanol formation. Simplified synthesis pathway of fermentative aromas connected to threonine metabolism (**A**). Allelic effect of the involved enzymes Alp1 and Nrg1 on the formation of volatiles derived from α-keto-butyrate as determined by RHA (**B**). Concentrations are given in relation to the heterozygote of the parental strains MTF2621 and MTF2622. (p-value: * ≤ 0.05, ** ≤ 0.01).

4 Conclusions

In this study, we confirm the potential of QTL analysis for deciphering the impact of genetic variation on the production of volatile metabolites by *Saccharomyces cerevisiae* during alcoholic fermentation. We were able to enlarge the analytical power of the approach compared to previous studies by using a comparatively large number of 130 segregants originating from a cross of two wine strains and by increasing the recombination rate of the

segregants. This approach enabled us to perform single and multiple QTL mapping strategies, leading to the detection of 65 QTLs with an influence on the formation of volatile metabolites, the production of extracellular main metabolites and general fermentation parameters. Our results confirm that multiple QTL mapping offers the possibility to detect additional, particularly minor loci. We were furthermore able to detect interacting QTLs for three evaluated traits, i.e., the formation of ethyl lactate, ethyl octanoate and propanol. However, it could be seen that an even larger number of segregants is required for a thorough and significant determination of QTL interactions.

We validated 13 genes in 9 QTLs, and of these genes, five (*AGP1*, *ALP1*, *FAS1*, *ILV6* and *LEU9*) have well described roles in metabolic pathways leading to yeast fermentative aroma formation. We could confirm their contribution to volatile production and characterized allelic variants that explain variations in these traits between the parent strains. Furthermore, the previously described fructophilic character of the MTF2621 allele of *HXT3* was confirmed in this study. For the other 7 validated genes (*AGP2*, *IXR1*, *MAE1*, *NRG1*, *RGS2*, *RGT1* and *SIR2*), we revealed contributions to the formation of fermentative aromas that were not previously reported. The fact that 5 of the 12 validated genes involved in volatile formation have broad regulatory functions on gene expression reveals the significant role of gene regulation in fermentative aroma production. These results demonstrate that QTL mapping is an effective and advisable approach for detecting the impact of globally acting genes on individual traits.

In summary, our findings of QTLs, their interactions and underlying gene variants emphasize the complexity of yeast fermentative aroma formation and provide the most extensive analysis of the links between genetic variation and the fermentative production of sensorial important volatiles to date. The results of this study will lead to the improvement of commercial *S. cerevisiae* starter cultures for the production of fermented food and beverages by non-GMO methods, such as breeding via marker-assisted selection. As many of the described secondary metabolites are additionally used as biofuel additives or building blocks for chemical syntheses, improved knowledge about allelic variation may also open paths for improving strains in a wide range of biotechnological applications.

5 List of Abbreviations

FAS: Fatty acid synthetase; LOD: logarithm of odds; PCA: Principal component analysis; QTL: Quantitative trait locus; RHA: Reciprocal hemizygosity analysis; SM: Synthetic must; SNP: Single-nucleotide polymorphism; YPD: Yeast extract peptone dextrose.

6 Declarations

Ethics approval and consent to participate: Not applicable

Consent for publication: Not applicable

Availability of data and material: The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

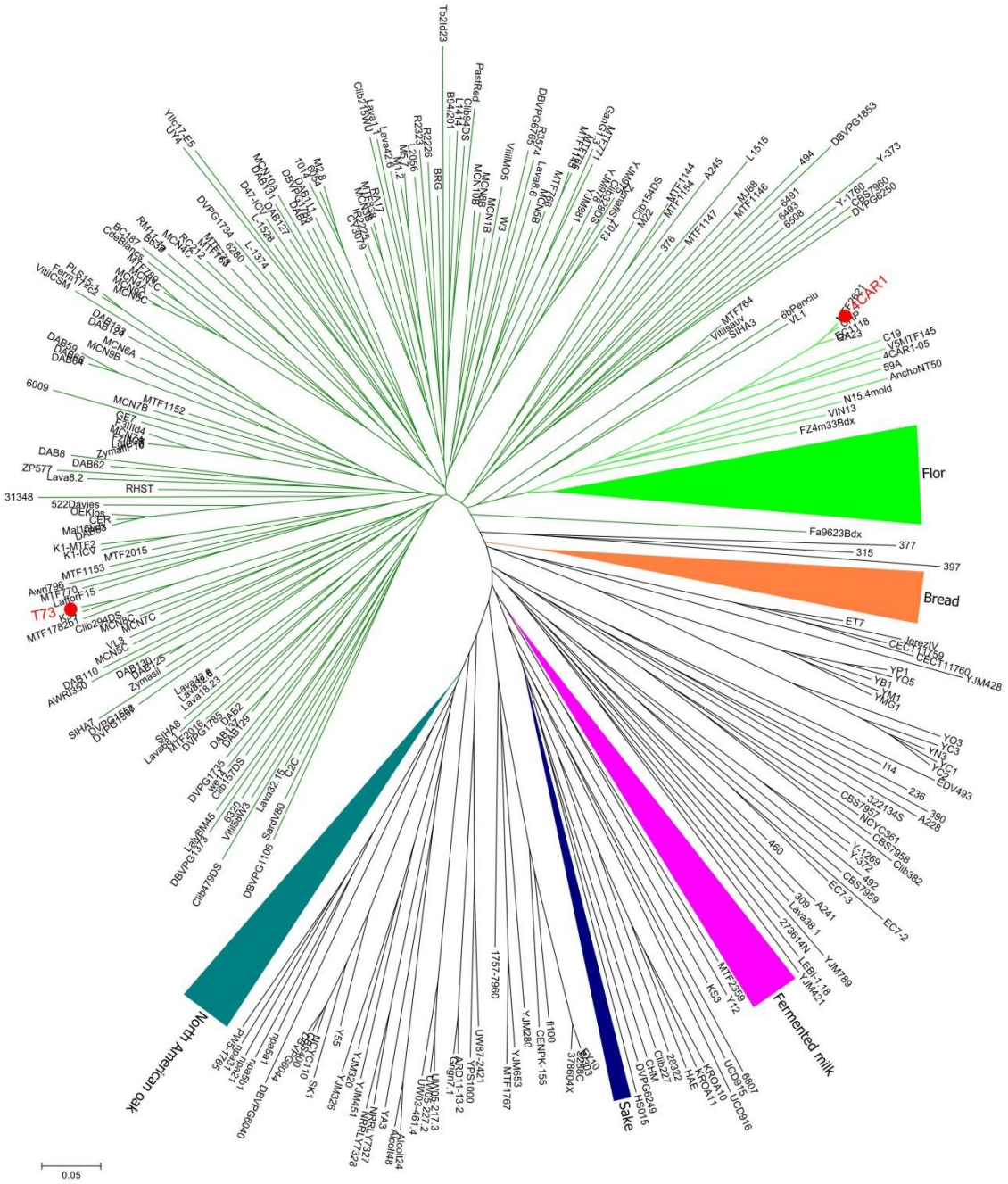
Competing interests: The authors declare that they have no competing interest.

Funding: This work was supported by the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007–2013/ under REA grant agreement n° 606795, project acronym: YEASTCELL.

Authors' contributions: ME conducted the experiments, contributed to the linkage analysis and wrote the manuscript. IS established and performed the linkage analysis. CB contributed to the selection of parent cells and generation of F2-segregants. CC contributed to the phenotyping of strains. JLL analyzed the genotyping data. SD conceived and planned the study, with contribution of CC and JLL. All authors revised the manuscript and read and approved the final manuscript.

Acknowledgements: We are grateful to Bruno Blondin for his contribution to the design of the QTL strategy. We thank Anna Hagstrom for her technical help in the construction of hemizygotes.

7 Supplementary Information



Additional file 1: Genomic background of parent strains. Location of the *S. cerevisiae* strains used in this study, MTF2621 (4CAR1) and MTF2622 (T73), within the genotypic subgroups of champagne strains (light green lines) and wine strains (dark green lines). Phylogenetic tree constructed with data from and as described by Legras et al. (2007).

Additional file 2: Table of primers used in this study.

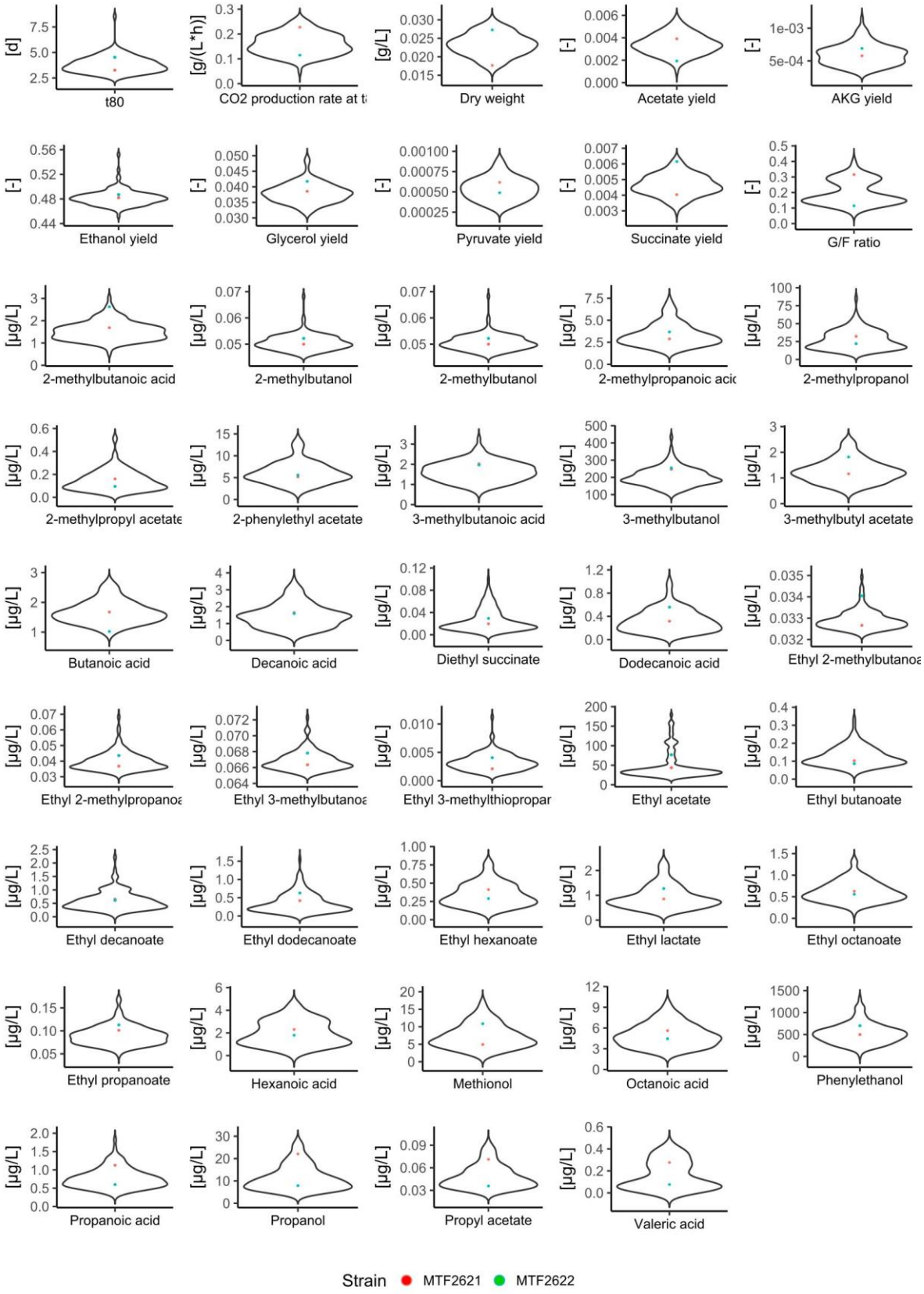
Primer name	Nucleotide sequence (5‘ -> 3‘)
del_ACP1_fw	ACAAACAACACAACTAACTCAATACAGCACCTTCCTTGCCTTCGTACGCTGCAGGTCGAC
del_ACP1_rv	GGGGTGACACGATACAATATAATAGAGCGGGGACGGACACGCATAGGCCACTAGTGGATCTG
del_AGP1_fw	AGAAGAAGCACGCTAATATAGACAAAGATAGCTTCGCACATTCGTACGCTGCAGGTCGAC
del_AGP1_rv	CAAAAATGAATAAATATAAAAGAAGTAAATGCTTTTTTTTGCATAGGCCACTAGTGGATCTG
del_AGP2_fw	AAGCTGCACTTACATTTTGCTTCCATAACTTTTGCCAAGCTTCGTACGCTGCAGGTCGAC
del_AGP2_rv	GCAGTCAATTTAAATTTGTGAATATAACGACATAATTGCAGCATAGGCCACTAGTGGATCTG
del_ALP1_fw	GTAGTGTTGCGATTATTGCCATGGATGAAACTGTGAACATTTCGTACGCTGCAGGTCGAC
del_ALP1_rv	GTGGTATGGAGTATTATTCTAAAATTATGAAAGGACATCCGCATAGGCCACTAGTGGATCTG
del_12@227_fw	TATTATTGTAATATGGGCGATGGTTAGGGTGACGGCGACTTTCGTACGCTGCAGGTCGAC
del_12@227_rv	GACGTGAAAAGCGGATCGTGTGTGTCTTGTATTTACGATGGCATAGGCCACTAGTGGATCTG
del_FAS1_fw	ATTTATTCGCCACACCTAACTGCTCTATTATTCGCTCATCTTCGTACGCTGCAGGTCGAC
del_FAS1_rv	AAGTTAAATATTTCTTACGGTTATATAATCACTTAAGAAAGCATAGGCCACTAGTGGATCTG
del_FAT3_fw	AAGGCGTTTGCTGCCTTAACCCAATTGATGGAAAATTCGGTTCGTACGCTGCAGGTCGAC
del_FAT3_rv	CAAGAAGGTCTGAGGGTTTTCTTGAGCCCAGGAAGTCAGCGCATAGGCCACTAGTGGATCTG
del_HXT3_fw	ATAGAATCACAAACAAAATTTACATCTGAGTTAAACAATCTTCGTACGCTGCAGGTCGAC
del_HXT3_rv	TAAAATACACTATTATTCAGCACTACGGTTTAGCGTGAAAGCATAGGCCACTAGTGGATCTG
del_HXT6_fw	GGCTTGCAGACAATGGAGAGCAAATGGGTATACAATATAGTTCGTACGCTGCAGGTCGAC
del_HXT6_rv	CAGAATTAGAGTGCATTTCAAATGCACAAATTAGAGCGTGGCATAGGCCACTAGTGGATCTG
del_HXT7_fw	CTTCACAATGTTCGAATCTATTCTTCATTTGCAGCTATTGTTCGTACGCTGCAGGTCGAC
del_HXT7_rv	GAGTACATTTCAAATGCACAAATTAGAGCGTGATCATGAAGCATAGGCCACTAGTGGATCTG
del_ILV6_fw	AATCTTTAGAACATCTGAGCTCACTAACCCAGTCTTTCTATTCGTACGCTGCAGGTCGAC
del_ILV6_rv	AGGAGAGTCCCGAGGGCGATCGCAAGGCCGAGAGACTAACGCATAGGCCACTAGTGGATCTG
del_IXR1_fw	CTGTGATATACGTACGACGCTAACAGTACCCACAACTGCATTCGTACGCTGCAGGTCGAC
del_IXR1_rv	TGGGATAATGTTACAGTGGAAAACTAAAGTTGTTTATTTGGCATAGGCCACTAGTGGATCTG
del_LEU9_fw	GGATAATACTATCAGCACATTATCATTTAGCCGCGTAGCCTTCGTACGCTGCAGGTCGAC
del_LEU9_rv	TATATATATAACATGAGTAATCATAAGCTACTCCTTTCTAGCATAGGCCACTAGTGGATCTG
del_MAE1_fw	AGTGCACATAAATACCAAGACAAAAGGTAGAAATACGGTTTTCGTACGCTGCAGGTCGAC
del_MAE1_rv	TTTTTTTTTTTAAGTGCAGGCGTTGGTTATGCTTCGTCTAGCATAGGCCACTAGTGGATCTG
del_NRG1_fw	CTCGACCAGCATATTACTACCCTTCGCAAACTTTCAGGCATTCGTACGCTGCAGGTCGAC
del_NRG1_rv	AGTAGTACTGCTAATGAGAAAAACACGGGTATACCGTCAAGCATAGGCCACTAGTGGATCTG
del_PXA2_fw	ATAATAATACAATTAAAAGTTACCGAAGAAAGATTTTATATTCGTACGCTGCAGGTCGAC
del_PXA2_rv	CAATTTATACATGATTTGGATTCTCCTTTGGCTATGTATGGCATAGGCCACTAGTGGATCTG
del_RGS2_fw	CCTTTGATACATAAAACGAAGAAAATTCAGCACATGCCAGTTCGTACGCTGCAGGTCGAC
del_RGS2_rv	TGAAGATATTTGTGTCTCCACAGATGATGAAGAGGCTATCGCATAGGCCACTAGTGGATCTG
del_RGT1_fw	GAAGCTGTACTCTTCTCTCAAACTCCAATATATTTCAAATTTCGTACGCTGCAGGTCGAC
del_RGT1_rv	GGAGAACCTGACCTACAGGAGAAGGGAGCATAGTTACCTGGCATAGGCCACTAGTGGATCTG
del_SIR2_fw	CATTCAAACCATTTTTCCCTCATCGGCACATTAAAGCTGGTTCGTACGCTGCAGGTCGAC
del_SIR2_rv	TATTAATTTGGCACTTTTAAATTATTAAATTGCCTTCTACGCATAGGCCACTAGTGGATCTG
del_YDL124W_fw	TAAACGGTTGTGTTACCCTAAAGAAACAGAGGTCAGTTAATTCGTACGCTGCAGGTCGAC
del_YDL124W_rv	ATACAAAACAAATATGACTCGTACATAAATTGTCCGGTATGCATAGGCCACTAGTGGATCTG
Hygro_rv	TGTTATGCGGCCATTGTC
tal_12@227_1	CGCCAGCATCAACATTAC
tal_12@227_2	CGCCGGCATCAACATTAC
test_ACP1_fw	GATAAGGCCGGTGCAACTTC
test_AGP1_fw	TTCTTCAGTGCCGCTTGAG
test_AGP2_fw	CGCCAAACCGGACCATTTAGATTC

test_ALP1_fw	CAGCTCCACTCCACATAAG
test_12@227_fw	ATATGGGCGATGGTTAGGG
test_FAS1_fw	ATTCGCCACACCTAACTG
test_FAT3_fw	TTTCTGCTTCGCTGTCTTCC
test_HXT3_fw	CGCGGAACATTCTAGCTCG
test_HXT6_fw	GATGTCTCGGATCTGTATGC
test_HXT7_fw	CCCACCATCTTTCGAGATCC
test_ILV6_fw	CAGACCGTCATGCAAGAATC
test_IXR1_fw	ATGGGTGGAACGGTTACTGAC
test_LEU9_fw	TTCTGGCTAGTTCCCTCGTTTC
test_MAE1_fw	CTTGGCTTCGACTCATCATC
test_NRG1_fw	CCTCAGGGAAAGCCAAGAAATG
test_PXA2_fw	AGATTAAGGGCGGCAATG
test_RGS2_fw	ACGCCGTTGCAGCTAAAGTC
test_RGT1_fw	GCTTATGAGCACCAGAGAC
test_SIR2_fw	GGCCTGCTATTTCTGTATCG
test_YDL124W_fw	TCCGCTTTCCAGTCGAACAC
test_YLR278C_fw	ACCTGAATCGCATCGTGTG

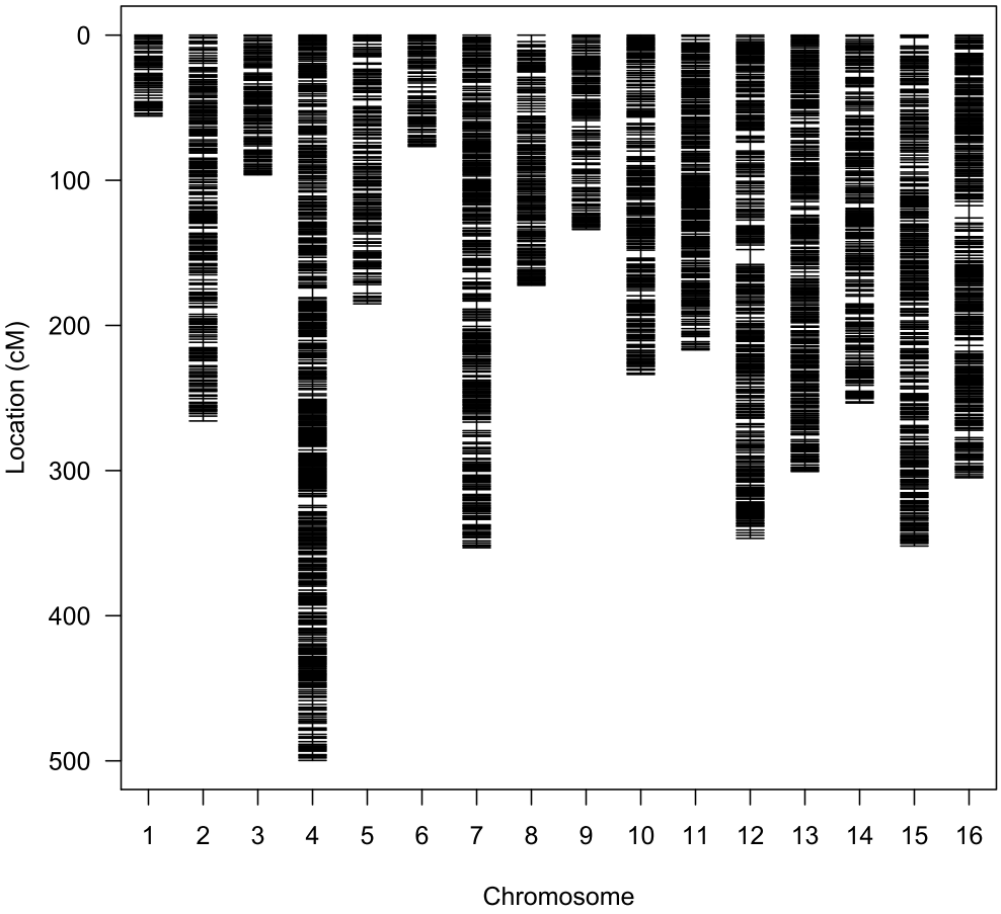
Additional file 3: Additional phenotypic information. Concentrations of determined secondary metabolites produced by the parental strains used in this study with trait variety among the segregant population given as Interquartile range (IQR) and heritability of evaluated traits.

Compound	MTF2621 [mg/L]	MTF2622 [mg/L]	Trait variety as IQR [mg/L]	Heritability
2-methylbutanoic acid	1.60 ± 0.19	2.65 ± 0.33	0,54	75,71
2-methylbutanol	0.05 ± 0.00	0.05 ± 0.00	0	58,14
2-methylbutyl acetate	0.26 ± 0.03	0.43 ± 0.09	0,14	92,87
2-methylpropanoic acid	2.80 ± 0.30	3.7 ± 0.51	1,61	94,35
2-methylpropanol	31.06 ± 4.63	21.6 ± 3.03	20,51	87,95
2-methylpropyl acetate	0.17 ± 0.02	0.10 ± 0.02	0,11	93,72
2-phenylethyl acetate	5.11 ± 1.15	5.65 ± 0.80	2,95	80,47
3-methylbutanoic acid	2.01 ± 0.23	1.98 ± 0.23	0,65	78,27
3-methylbutanol	234.41 ± 53.93	244.13 ± 46.85	60,79	-3,21
3-methylbutyl acetate	1.13 ± 0.11	1.93 ± 0.3	0,54	93,68
Butanoic acid	1.69 ± 0.36	0.95 ± 0.26	0,5	15,83
Decanoic acid	1.59 ± 0.90	1.66 ± 1.05	0,93	-56,77
Diethyl succinate	0.02 ± 0.02	0.03 ± 0.03	0,02	0,19
Dodecanoic acid	0.29 ± 0.20	0.62 ± 0.48	0,25	-4,48
Ethyl 3-methylthiopropanoate	$1.99 \pm 0.95 \times 10^{-3}$	$3.97 \pm 0.99 \times 10^{-3}$	0	69,38
Ethyl acetate	33.54 ± 16.01	74.83 ± 94.73	17,83	72,04
Ethyl butanoate	0.11 ± 0.02	0.09 ± 0.02	0,07	86,17
Ethyl decanoate	0.61 ± 0.29	0.55 ± 0.31	0,38	28,8
Ethyl dodecanoate	0.40 ± 0.27	0.6 ± 0.27	0,35	1,98
Ethyl hexanoate	0.43 ± 0.09	0.27 ± 0.05	0,24	69,56
Ethyl 2-methylpropanoate	0.04 ± 0.00	0.04 ± 0.01	0,01	88,05
Ethyl lactate	0.79 ± 0.36	1.21 ± 0.53	0,57	42,24
Ethyl octanoate	0.64 ± 0.19	0.53 ± 0.16	0,33	36,06
Ethyl propanoate	0.10 ± 0.01	0.11 ± 0.02	0,03	84,15

Ethyl 2-methylbutanoate	0.03 ± 0.00	0.03 ± 0.00	0	77,06
Ethyl 3-methylbutanoate	0.07 ± 0.00	0.07 ± 0.00	0	59,66
Hexanoic acid	2.34 ± 0.33	1.76 ± 0.29	1,8	91,33
Methionol	4.39 ± 1.96	10.53 ± 2.82	4,4	65,71
Octanoic acid	5.67 ± 1.33	4.48 ± 1.13	2,51	41,46
Phenylethanol	455.13 ± 105.33	711.57 ± 106.55	267,48	76,64
Propanoic acid	1.09 ± 0.25	0.53 ± 0.08	0,34	18,98
Propanol	21.41 ± 3.36	7.23 ± 0.96	6,81	63,61
Propyl acetate	0.07 ± 0.01	0.04 ± 0.00	0,02	70,62
Valeric acid	0.28 ± 0.03	0.06 ± 0.02	0,21	93,21



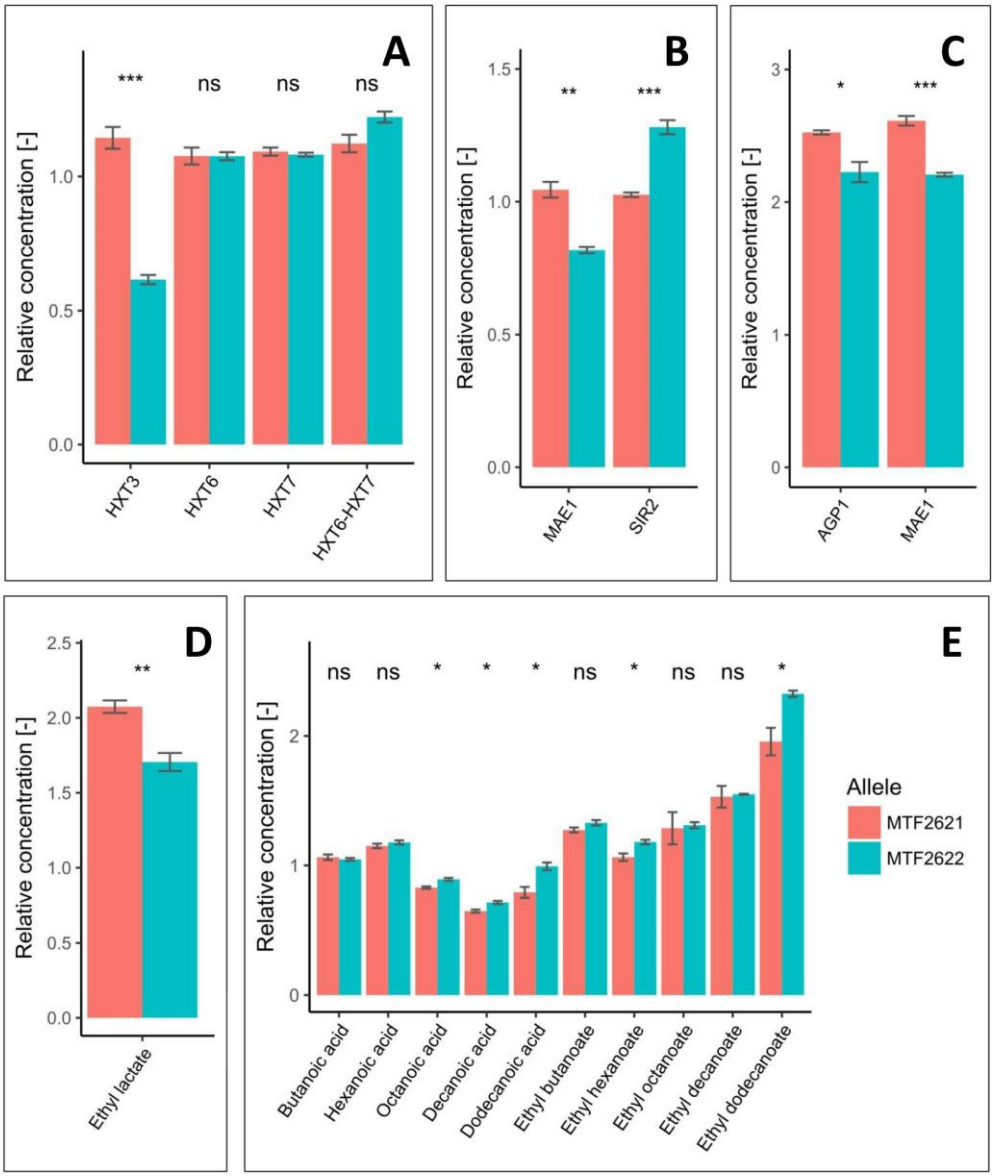
Additional file 4: Phenotype distributions among population. Distribution of evaluated traits for QTL mapping among all 130 F2-segregants of the study. The position of parental cells within the population is marked in red for MTF2621 and in green for MTF2622.



Additional file 5: Marker map. Graphic representation of marker positions that were used for linkage analysis.

Additional file 6: SNPs in predicted regulatory binding sites of validated genes. Detected SNPs in the 1000-bp upstream region of evaluated target genes that affect binding motifs for regulatory proteins as predicted with YEASTRACT (Teixeira et al., 2013). Comparison of the strains MTF2621 and MTF2622 with the *S. cerevisiae* reference strain S288C.

Gene	Nucleotide position	S288C	MTF2621	MTF2622	Affected binding motifs
AGP1	-523	A	G	A	Ume6
ILV6	-814	C	G	C	Msn2, Msn4, Nrg1, Rph1
LEU9	-709	T	T	C	Gcr1, Nrg1
MAE1	-989	T	C	T	Stp1, Stp2
	-707	A	A	G	Rtg1, Rtg3
	-694	G	G	A	Mot3
	-667	T	T	A	Azf1
	-351	C	C	G	Rtg1, Rtg3
RGT1	-640	A	G	A	Rtg1, Rtg3, Stb5
	-562	C	T	C	Ash1
	-438	C	A	C	Rtg1, Rtg3
	-241	A	C	T	Ste12



Additional file 7: Additionally detected allelic effects of the described enzymes as determined by RHA. Allelic effect of the sugar transporters Hxt3, Hxt6 and Hxt7 on the G/F ratio (**A**). Allelic effect of the enzymes Mae1 and Sir2 on the acetate yield (**B**). Allelic effect of the enzymes Agp1 and Mae1 on the production of 2-phenylethanol (**C**). Allelic effect of Mae1 on the formation of ethyl lactate (**D**) and fatty acids and fatty acid ethyl esters (**E**). Concentrations are given in relation to the heterozygote of the parental strains MTF2621 and MTF2622. (p-value: * ≤ 0.05, ** ≤ 0.01, *** ≤ 0.001)

<u>Chapter 2:</u> Development of a QTL mapping approach to detect loci that influence intracellular fluxes (f-QTLs) of the yeast central carbon metabolism.

The second part of the thesis assesses the possibility to extend QTL mapping to the detection of loci with an influence on intracellular metabolic fluxes (f-QTLs). In a wider view, another motivation behind the approach was the fact that the formation of fermentative aroma is connected to pathways of the yeast central carbon metabolism. The possibility to detect QTLs that influence flux distributions could therefore lead to the identification of allelic variants impacting aroma formation.

Extracellular metabolites were determined during the exponential phase of fermentation for all 130 F2-segregant strains. Intracellular fluxes were estimated by feeding the obtained metabolite concentrations into a constraint-based model of yeast central carbon metabolism previously developed in the lab. The estimated fluxes were then used as phenotype data for QTL mapping, relying on the marker map that was obtained during the first part of the project.

The chapter is composed as a research article, which will be submitted before the defense of this thesis.

QTL mapping of modeled metabolic fluxes reveals gene variants impacting yeast central carbon metabolism

Matthias Eder[a], Thibault Nidelet[a], Isabelle Sanchez[a,b], Carole Camarasa[a], Jean-Luc Legras[a], Sylvie Dequin[a,*]

[a]SPO, INRA, SupAgro, Université de Montpellier, F-34060 Montpellier, France

[b]MISTEA, INRA, SupAgro, F-34060 Montpellier, France

* Corresponding author. Mailing address: INRA, 2 Place Pierre Viala, 34060 Montpellier CEDEX 2, France; Phone: +33 4 99 61 25 28; E-mail: sylvie.dequin@inra.fr

E-mail addresses:
Matthias Eder: matthias.eder@inra.fr
Thibault Nidelet: thibault.nidelet@inra.fr
Isabelle Sanchez : isabelle.sanchez@inra.fr
Carole Camarasa: carole.camarasa@inra.fr
Jean-Luc Legras: jean-luc.legras@inra.fr
Sylvie Dequin: sylvie.dequin@inra.fr

Abstract

With the increasing availability of genomics data, elucidating the genotype-phenotype links has become a driving question today. We recently demonstrated great variations in metabolic fluxes of the central carbon metabolism between yeast strains of different origin. However, due to the complexity of flux regulatory mechanisms, we have a limited understanding of how fluxes are modulated. Here, we investigated the potential of the quantitative trait locus (QTL) mapping approach to elucidate the genetic variations responsible for differences in metabolic fluxes (f-QTL), using a population of 130 F2-segregants from a cross of two wine yeast strains. Intracellular metabolic fluxes were estimated by constraint-based modeling and used as quantitative phenotypes. Differences in fluxes were linked to genomic variations in the progeny population. Using this approach, we detected four main QTLs that influence metabolic pathways. The molecular dissection of these QTLs revealed the contribution of two allelic gene variants, *PDB1* and *VID30*, which have an influence on glycolysis, glycerol synthesis, ethanol synthesis, tricarboxylic acid cycle fluxes and transport and excretion of main metabolites.

Our study proves the feasibility of using model-estimated metabolic fluxes to decipher metabolic traits and increases the value of QTL mapping to elucidate the impact of genomic variation. We report the first elucidation of genetic determinants influencing metabolic fluxes. Deducing the mechanisms that control a metabolic network will allow the development of strains for producing food and beverages with optimized metabolite profiles using metabolic engineering or breeding strategies.

Keywords: f-QTL mapping, metabolic fluxes, central carbon metabolism, metabolic model, yeast

1 Introduction

Many phenotypic traits of the yeast *Saccharomyces cerevisiae* are dependent on the functional and regulatory properties of central carbon metabolism (CCM), and most of these traits are relevant for industrial processes such as the production of biofuels or the fermentation of foods and beverages. This includes fermentation rate, ethanol yield or the formation of other extracellular metabolites such as glycerol, acetic acid, succinic acid or pyruvic acid. Besides these main metabolites, some volatile aroma molecules are formed from intermediates of yeast CCM. Higher alcohols, fusel acids and their esters, e.g., originate from α-keto-acids, which can derive from pyruvate, phosphoenolpyruvate or oxaloacetate. Medium chain fatty acids and their esters are synthesized from acetyl-CoA. The composition and concentrations of main and volatile metabolites are an essential quality criterion for many fermented foods and beverages such as wine, where they play an important role for its sensorial perception.

Main and volatile metabolites are transformed from substrates of the fermentation medium by a large number of intracellular reactions. The formation of these metabolites is governed by metabolic fluxes, which refer to the turnover rate of substances through metabolic pathways. They reflect the integration of complex regulation on various biological levels, including genetic (transcription, translation, protein modifications, protein-protein interactions) and metabolic levels. Thus, they represent an interesting phenotypic trait, close to the cellular phenotype, which is of interest to study in order to better understand the genetic determinants contributing to the control of these fluxes.

While metabolite turnover rates are difficult to determine experimentally, they can be estimated by modeling with a limited set of data (reviewed by Österlund et al., 2012; Van Gulik and Heijnen, 1995). These constraint-based models (CBM) that formulate the metabolic network as a stoichiometry matrix exist in a wide range, from small models focusing on specific subtypes of cellular metabolism to genome-scale networks representing all metabolic reactions.

The first step of predicting fluxes from networks is to add constraints on input and output fluxes. Depending on network size and number of constraints, this approach, which is termed metabolic flux analysis (MFA), can already be sufficient to estimate fluxes. However, in most cases adding constraints on input and output data is not sufficient and there are two ways to deal with it, ^{13}C-MFA and flux balance analysis (FBA). The ^{13}C-MFA approach uses ^{13}C labeled substrates and tracking of ^{13}C across cellular metabolites generates information to also constraint intracellular fluxes in order to subsequently estimate them. The FBA applies the assumption that cellular functions of biochemical networks in a steady state are limited by physico-chemical constraints (reviewed by Palsson, 2000). Through linear optimization, FBA chooses the best fitting solution out of a narrowed solution frame defined by the stoichiometry matrix of the CBM (Varma and Palsson, 1994). The outcome of this flux

prediction depends on an applied objective function. Commonly used objective functions are the maximization of ATP production, minimalization of metabolic adjustment or, as in most cases, maximization of biomass production. The effectivity of these objective functions depends on experimental conditions, constraints and models.

A central issue for understanding cellular physiology is to understand the modulation of metabolic fluxes by genetic or environmental determinants and the application of CBM is a suited approach for that. In a previous study, our group used a combined ^{13}C-MFA/FBA approach to estimate intracellular fluxes of *S. cerevisiae* CCM in conditions of modified intracellular redox balance (Celton et al., 2012b). This model was used to assess the sensitivity of flux distribution to environmental conditions for *S. cerevisiae* wine yeast (Celton et al., 2012a). Another example for the application of FBA is the study of Quirós et al. (2013), who used a model developed by Vargas et al. (2011) to evaluate changes of yeast metabolism in high sugar must. In both studies, glycolytic fluxes were demonstrated to show the least variation, while fluxes of the pentose phosphate pathway (PPP) were highly variable. ^{13}C-MFA, on the other hand, was used to study network robustness (Blank et al., 2005) or the effects of deletion mutants (Velagapudi et al., 2007). The later study demonstrated interesting links between networks, e.g., a positive correlation between the PPP and biomass yield.

In recent years, research brought vast amounts of information about the genotypic and phenotypic diversity of *S. cerevisiae* (Fay and Benavides, 2005; Legras et al., 2007, 2005; Liti et al., 2009; Strope et al., 2015; Warringer et al., 2011). In these studies, the phenotypic diversity was mainly assessed by comparison of growth parameters in different media. However, several studies started to extend the characterization of diversity to a greater number of phenotypic traits. Spor et al. (2009) evaluated the diversity of six life-history traits and three metabolic traits among *S. cerevisiae* strains and could separate them into two groups of different life-history strategies. By performing wider phenotypic screening of 72 *S. cerevisiae* strains from different origins for seven life-history traits and eleven metabolic traits, Camarasa et al. (2011) showed that origin has a broad phenotypic impact.

In a previous study, our group assessed the diversity of flux distributions between *S. cerevisiae* strains from different origins (Nidelet et al., 2016). This was done by using the constraint-based model developed by Celton et al. (2012a) to estimate CCM flux distributions between 43 strains of different ecological origins that were grown under wine fermentation conditions. The study showed a contrasted image regarding flux variability with quasi-constancy of glycolysis and ethanol synthesis on the one hand, but high variation for other fluxes such as the PPP or acetaldehyde production on the other hand. In addition, the fluxes showed multimodal distributions that could be linked to ecological origin, which reveals an association between genetic origin and metabolic flux manifestation. Wine strains, e.g., exhibit relative increase of PPP and tricarboxylic acid cycle (TCA) cycle fluxes, while the production of acetaldehyde is strongly diminished. Flor strains, on the other hand, show a relative reduction in PPP and TCA cycle fluxes with a higher production of acetic acid (Nidelet et al., 2016).

Results gained from flux analysis have also been used for strain improvement by metabolic engineering based on flux predictions (Agren et al., 2013; Patil and Nielsen, 2005). Possible metabolic engineering strategies to increase ethanol yield where evaluated using CBM by Bro et al. (2006), which led to the development of a strain with ethanol yield increased by 3% and glycerol yield decreased by 40%. Further examples of genome scale model guided engineering strategy predictions include the optimization of yields of purine (Burgard and Maranas, 2003), succinic acid (Agren et al., 2013; Otero et al., 2013) or proline (Bundy et al., 2007). Therefore, more knowledge about the impact of genomic variation on metabolic flux distributions has potential for the selection or improvement of strains with diverse applications.

With numerous existing studies, QTL mapping has become an important approach to deeper understand genomic complexity of *S. cerevisiae* and to decipher the impact of genomic variation on yeast complex traits (Swinnen et al., 2012). This includes investigations of genetic determinants influencing enological important traits, which led to the discovery of allelic variants accounting for variations of these traits (Ambroset et al., 2011; Brice et al., 2014a; Eder et al., 2018; Martí-raga et al., 2017; Marullo et al., 2007; Noble et al., 2015; Salinas et al., 2012; Steyer et al., 2012). All these studies have in common that the assessed traits were straightforward to quantify. However, difficulties remain to detect QTLs for traits with small variations or more complex to determine, such as intracellular metabolic fluxes.

Thus, the possibility of using QTL mapping to decipher genomic variation impacting metabolic profiles rather than single metabolites would open ways to understand the mechanisms behind metabolic flux distributions and to engineer strains with superior metabolic properties for various applications.

To achieve this, we phenotyped 130 meiotic F2-segregants from a cross of two wine yeast strains for their production of extracellular main metabolites during exponential phase. We modeled intracellular fluxes of the yeast CCM by feeding these experimentally determined metabolite concentrations into a constraint-based model. Subsequently, we used these estimated fluxes as phenotypic data to perform QTL mapping on metabolic flux distributions. With this approach we were able to detect four QTLs with an influence on various metabolic fluxes. By performing reciprocal hemizygosity analysis (RHA), we confirmed the robustness of the method by validating the role of two genes, *PDB1* and *VID30*, within two QTLs. The allelic variants of these genes show a different effect on fluxes of the glycolysis, ethanol synthesis, glycerol synthesis, TCA cycle and the excretion of TCA cycle metabolites.

2 Materials and methods

2.1 Media

Yeast strains were cultured at 28 °C in yeast extract peptone dextrose (YPD) media, containing 10 g/L yeast extract, 20 g/L peptone and 20 g/L glucose. Solid YPD media contained 1.5% agar. Selective YPD media contained 200 µg/mL geneticin (G418), 200 µg/mL nourseothricin (clonNAT) or 200 µg/mL hygromycin B were used.

Wine fermentations were carried out in synthetic grape must (SM200) described by Bely et al. (1990). It contains 100 g/L of each sugar, glucose and fructose, and 200 mg/L of assimilable nitrogen. The amino acid composition hereby mimics the nitrogen content of standard grape juice.

2.2 Strain selection and generation

The haploid *S. cerevisiae* strains MTF2621 (4CAR1 [Δ*HO*::*Neo*^{r}]) and MTF2622 (T73 [Δ*HO*::*Nat*^{r}]) were selected for the study according to their different need for nitrogen during wine fermentation. The requirement of nitrogen was estimated by using an approach based on the addition of nitrogen to keep the CO_2 production rate constant during nitrogen limitation (Brice et al., 2014b). The strain T73 belongs to the phylogenetic clade of wine strains. The strain 4CAR1, however, belongs to the group of champagne strains, which originated through crossings between strains of the wine clade and the flor clade (Coi et al., 2016). From a crossing of both parent cells, 130 F2-segregants were generated as described in Eder et al. (2018).

2.3 Phenotyping of strains

Segregants were fermented in duplicates with the parent strains as controls. The strains were grown overnight in 50 mL YPD media. The cell density was determined using a Multisizer™ 3 Coulter Counter (Beckman Coulter). Sterilized 300-mL glassware mini fermenters were filled with 280 mL of SM200 and closed with an air lock. The fermenters were inoculated to a cell density of 1×10^6 cells/mL, weighed and left at 24 °C under stirring (300 rpm).

In order to determine the concentration of extracellular metabolites during the exponential phase, a sample was taken when approximately 10 g/L of CO_2 were produced. This was determined by weighting the fermenters regularly to draw the weight decrease caused by the release of CO_2. The dry weight was determined in duplicate by filtering 10 mL of cell suspension through a nitrocellulose membrane with a porosity of 0.45 µm (Millipore, France)

and known dry weight. The membrane was rinsed twice with 10 mL of distilled water, dried for 48 h and weighed to determine the dry biomass of the sample.

2.4 Modeling of metabolic fluxes

Extracellular metabolite concentrations (in mmol/mL) and dry mass weight (in g/L) obtained by the phenotyping of segregant strains were used to constrain a model of yeast central carbon metabolism previously developed (DynamoYeast, Celton et al., 2012a). The model covers 68 reactions (Additional file 10), 61 metabolites (Additional file 9) and distinguishes three compartments, the extracellular medium, cytoplasm and mitochondria. All predictions were performed using the programming language R v3.2.3 with the R/sybil v2.0.0 and R/sybilSBML v2.0.11 libraries (Gelius-Dietrich et al., 2013). Due to inexplicable block effects concerning the determination of succinate concentrations, succinate fluxes were constraint by setting a fixed range, corresponding to the maximum succinate flux variation between *S. cerevisiae* strains determined in our previous study (Nidelet et al., 2016). The error margin for the flux boundaries of the model was set to ±2.5%. We obtained the flux distribution throughout the metabolic network for each segregant by FBA with minimization of glucose input as objective function. For the modeling approach, fructose was hereby also considered as glucose, as this did not impact flux predictions. Finally, estimated fluxes were normalized to sugar uptake.

2.5 QTL mapping

QTL mapping of modeled metabolic fluxes was performed using a marker map that we obtained by whole genome sequencing of the individual segregant strains during a previous study (Eder et al., 2018). The statistical analyses were carried out using the programming language R v3.2.3 (www.r-project.org) with the R/qtl v1.40-8 and R/eqtl v1.1-7 libraries (Broman et al., 2003). QTL mapping was performed with two different phenotype models, the normal model using Haley-Knott regression and a non-parametric analysis, resulting in logarithm of odds (LOD) scores for each marker and pseudo-marker every 2.5 cM (interval mapping method). An interval estimate of the location of each QTL was obtained as the 1-LOD support interval, the region in which the LOD score is within 1 unit of the peak LOD score. If the same locus was detected with both models, the results with the higher LOD score were selected.

2.6 Reciprocal hemizygosity analysis

Molecular dissection of found QTLs was performed using RHA (Steinmetz et al., 2002; Warringer et al., 2017). Target genes in QTLs were chosen according to a biological function associated with metabolic processes and the gene's proximity to the determined QTL peak. The gene sequences were deleted in both parent strains by homologous recombination with a disruption cassette containing the hygromycin B resistance gene (*hph*^{r}). Positive integration was selected by plating the transformed cells on YPD-agar plates containing hygromycin B. Correct gene deletion was verified by PCR using a primer that binds in the upstream region of the deleted gene and a primer that binds within the deletion cassette. Deleted parent strains were subsequently mated with the opposite undeleted parent to form a heterozygote that is hemizygous for the target gene. The obtained strains were phenotyped in triplicate (2.3). Significant influence of the allelic gene variant on the trait was evaluated by student's t-test.

3 Results

3.1 Phenotyping of strains

Using the obtained metabolite data, metabolic fluxes of the CCM were predicted for all strains. As the single fluxes within main metabolic pathways strongly coincide with each other, a selection of fluxes representative for main metabolic pathways was made to facilitate the following analyses (Additional file 11). Principal component analysis (PCA) of these selected fluxes was performed to assess flux correlations and to evaluate the variation between parent and segregant strains (Figure 23).

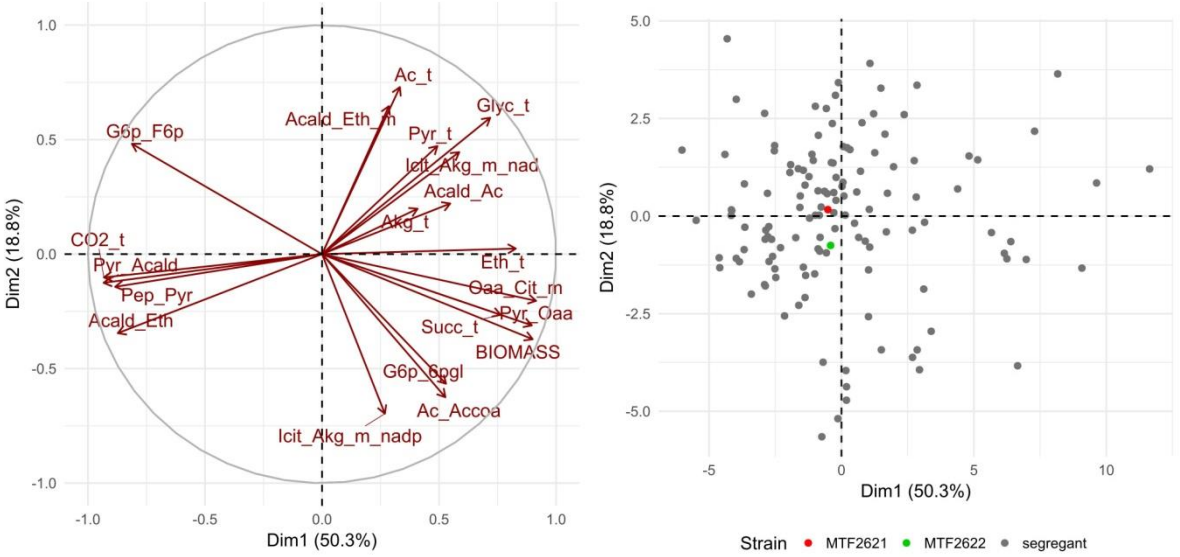


Figure 23: PCA of selected modeled fluxes (left) and variation among parents (red, green) and segregant strains (grey) (right).

With the first two dimensions explaining 69.1% of trait variation, the PCA of estimated fluxes well depicts the variation among strains. A positive correlation between fluxes of the PPP, TCA cycle oxidative branch and biomass formation can be seen. Together, these fluxes are negatively correlated to the upper glycolysis. Negative correlation can furthermore be found between fluxes of the lower glycolysis/ethanol synthesis and glycerol formation. It can be seen that the parent strains behave similar and show only small difference in their flux profile, while the segregant strains are more divergent. This is confirmed by the visualization of trait distributions (Additional file 12). The parental strains are located within the population of segregants for the majority of traits. To further assess variation among strains, coefficients of variation for estimated fluxes were separately calculated for parent and segregant strains (Figure 24).

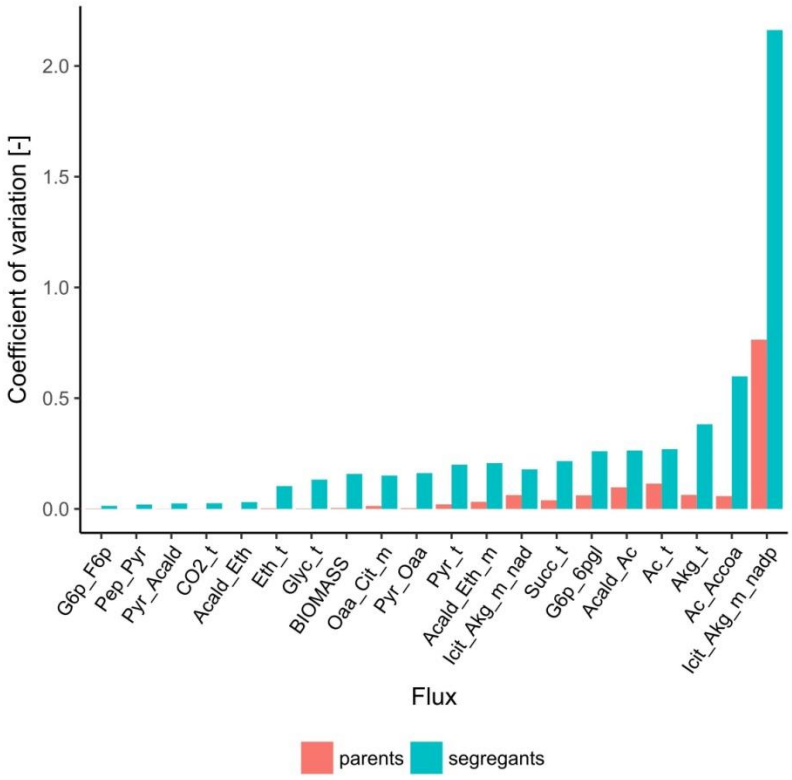


Figure 24: Coefficient of variation of selected representative fluxes among the parent strains of the study (blue) and among the resulting F2-segregants (red). Explanation of flux abbreviations is given in Additional file 9 and Additional file 10.

We show that the variation between segregant strains regarding the determined fluxes exceeds the variation between the parent strains, ranging from a factor of 2.4 for acetic acid production (Acald_Ac) to a factor of 268.8 for ethanol excretion (Eth_t). These results confirm the conclusion drawn by the PCA (Figure 23), explicitly that the parent strains do not show as much differences in flux distribution as the segregant population. However, differences in variation can be seen between single fluxes. To better visualize trait variation, the distributions for each flux were plotted around the mean value (Figure 25).

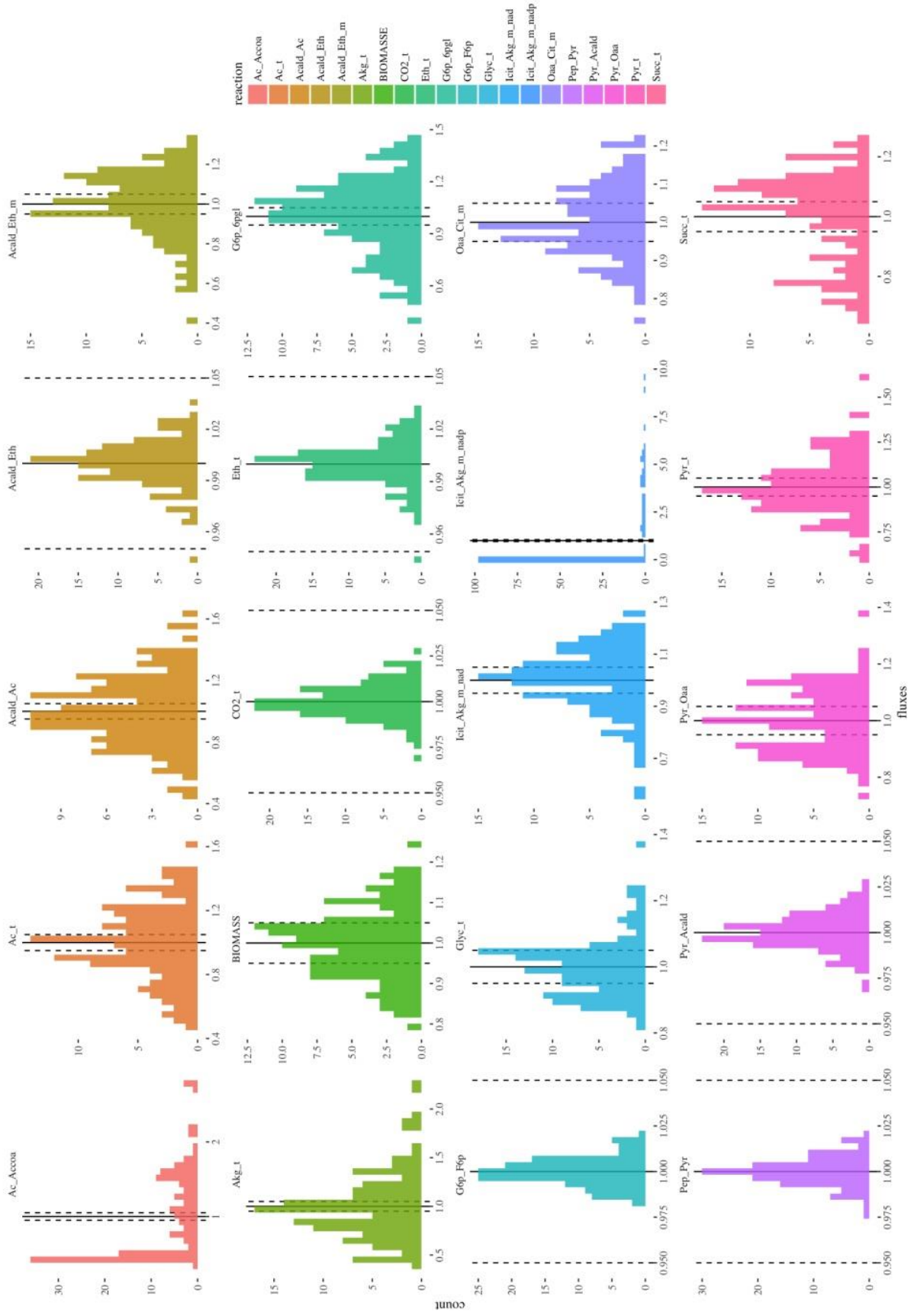


Figure 25: Variation of selected fluxes around the mean value among the segregant strains and ±5% interval indicated with dashed lines.

While fluxes of the glycolysis and ethanol synthesis only vary about ±2.5% around the mean value, fluxes of the PPP or the production of metabolites such as glycerol or acetic acid diverge up to 200% around the mean. Most fluxes are normally distributed, however, few outliers can be seen, particularly for glycerol, pyruvate and ethanol excretion. The distributions of three fluxes differ from a normal distribution, the synthesis of acetyl-CoA from acetate (Ac_Accoa), the NADP dependent mitochondrial flux from isocitrate to AKG (Icit_Akg_m_nadp) of the TCA cycle oxidative branch and the succinate excretion (Succ_t). In the case of Ac_Accoa, a subpopulation of segregant strains shows a strongly reduced flux towards acetyl-CoA. In the case of Icit_Akg_m_nadp, the analyses indicate that this flux is inactive in the majority of segregant strains, while this is not the case for the parental strains. In the case of Succ_t, two populations can be distinguished, which indicates a major influence of one allele on the trait. However, the flux of succinate excretion can be subject to greater inaccuracy since the flux was not constraint by experimental data but by an experimentally determined possible range (2.4).

3.2 Genome wide identification of QTLs influencing metabolic carbon fluxes

In a first step, QTL mapping was performed on obtained metabolite yields during exponential phase using the segregant marker map that we obtained from our previous study (Eder et al., 2018) (Table 20). This analysis led to the detection of 8 QTLs on 5 chromosomes influencing 7 traits. This included the majority of determined metabolite production yields as well as differences in sugar uptake (G/F ratio). The highest detected LOD-score was 4.71 for QTL chr4@152.6 influencing the succinate yield. Therefore, almost 16% of trait variation can be explained by the locus. The identification of two similar likely QTLs with an influence on succinate production does not confirm the previous assumption that one locus has a major impact on the trait.

Table 20: Detected QTLs influencing metabolite yields and ratio of remaining sugars (G/F ratio) during exponential phase.

Trait	QTL name	Chromosome	QTL start [bp]	QTL end [bp]	LOD
Succinate yield	chr4@122.8	IV	356071	380035	4.42
G/F ratio	chr4@125.4	IV	356071	400864	3.67
G/F ratio	chr4@139.3	IV	409433	448045	3.63
CO_2 yield	chr4@139.3	IV	410742	448045	4.05
Succinate yield	chr4@152.6	IV	426649	488205	4.71
G/F ratio	chr4@160.0	IV	448242	505548	3.76
CO_2 yield	chr4@160.0	IV	448242	505548	4.33
Glycerol yield	chr7@90.3	VII	252047	273771	3.76
AKG yield	chr10@242.3	X	717987	648141	4.38
Ethanol yield	chr13@208.9	XIII	622064	660267	4.22
G/F ratio	chr13@214.5	XIII	624189	648141	3.62
CO_2 yield	chr13@214.5	XIII	624189	648141	3.7
Acetate yield	chr13@237.7	XIII	710548	726277	3.84
AKG yield	chr15@31.8	XV	67745	111309	3.52

In a second step, QTL mapping was performed on estimated intracellular carbon fluxes (Table 21). A total of 4 QTLs influencing 7 traits could be detected on chromosomes II, V, VII and VIII. This included fluxes of the glycolysis/ethanol synthesis, glycerol synthesis, TCA cycle oxidative branch, biomass formation and metabolite transport/excretion. No QTLs could be detected for fluxes of the PPP, TCA cycle reductive branch and glutamate cycle, although these fluxes showed the most substantial variation among the segregant strains (Figure 24). The highest LOD-score of 4.63 was found for the influence of QTL chr7@18.0 on glycolysis and ethanol synthesis, meaning that 15.7% of the trait variation can be explained by the locus. The region of the QTL furthermore influences the most traits. Besides glycolysis and ethanol synthesis, fluxes of the biomass formation, TCA cycle oxidative branch and metabolite transport/excretion are affected.

Table 21: Detected QTLs influencing modeled metabolic fluxes.

Trait	QTL name	Chromosome	QTL start [bp]	QTL end [bp]	LOD
Glycerol synthesis	chr2@222.9	II	662795	701771	4.58
Malate transport	chr5@128.3	V	354177	400836	4.01
Glycolysis & ethanol synthesis	chr7@18.0	VII	40689	58851	4.63
Biomass	chr7@25.5	VII	52412	82449	3.73
TCA cycle oxidative branch	chr7@25.5	VII	52412	82449	4.05
Transport & excretion TCA cycle metabolites	chr7@25.5	VII	52412	82449	4.05
Ethanol transport	chr8@155.7	VIII	443664	483121	3.45

3.3 Validation of detected QTLs

The 2 QTLs with the highest LOD-score detected to influence estimated intracellular fluxes were selected for validation and candidate genes were chosen according to their distance to the QTL peak and biological function related to central carbon metabolism (Table 22). The

impact of these genes and their allelic variants was evaluated by RHA. The constructed strains were phenotyped for their formation of extracellular metabolites during exponential phase, intracellular fluxes were estimated using these metabolite concentrations and the differences between the alleles regarding flux distributions were assessed (Table 22).

Table 22: Validated allelic variants in detected QTLs influencing modeled metabolic fluxes. (p-value: * ≤ 0.05, ** ≤ 0.01)

QTL name	Trait	Evaluated genes	Different impact of allele on trait as MTF2621/MTF2622 [factor]
chr2@222.9	glycerol synthesis	*PDB1*	1.05* glycerol synthesis
chr7@25.5	biomass	*HAP2*	no effect
	glycolysis & ethanol synthesis	*VID30*	0.99* glycolysis & ethanol synthesis
	TCA oxidative branch		0.92* - 0.73** TCA cycle metabolite excretion
	TCA cycle metabolite transport & excretion		(0.90* TCA reductive branch)

Table 23: Differences in amino acids (AA) of validated gene variants caused by non-synonymous SNPs between parent strains. Comparison of SNP identity to *S. cerevisiae* type strain S288C.

Gene	Length in AA	AA position	S288C	MTF2621	MTF2622
PDB1	366	14	A	A	T
		26	-	A	-
		289	V	V	I
VID30	958	37	H	Y	H
		672	E	E	G
		882	I	V	I

In QTL chr2@222.9, detected to influence glycerol synthesis, *PDB1* was assessed by RHA and a significant influence of the allelic variants on the trait was detected. The allelic variants of the gene differ in three non-synonymous SNPs (Table 23). The MTF2621 allele of *PDB1* increases glycerol synthesis fluxes by 5% (Figure 26).

In region chrVII:40,689..82,449, detected to influence glycolysis, ethanol synthesis, biomass formation, TCA cycle fluxes and transport/excretion of TCA cycle metabolites, two genes were selected for validation, *HAP2* and *VID30*. While the variants of *HAP2* did not show significant difference regarding the metabolic fluxes, the contribution of *VID30* to the detected phenotype variations could be validated (Table 22). The allelic variants of *VID30* differ in three non-synonymous SNPs (Table 23). One SNP lies in the 1000-bp upstream region of the gene, however, no recognized binding site is affected according to the YEASTRACT database (Teixeira et al., 2013). The MTF2621 allele of Vid30 was found to decrease fluxes of glycolysis and ethanol synthesis by 1% (Figure 26). As the total variation regarding ethanol synthesis is only 6% among the segregant strains (Figure 25), this decrease by 1% was considered significant. In addition, the excretion of pyruvate, α-ketoglutarate and succinate was reduced

up to 27%. A significant influence of the alleles on the TCA cycle oxidative branch could not be detected. In contrast, the reductive branch of the TCA cycle was significantly affected, with the MTF2621 allele leading to a decrease in fluxes from malate to succinate by 10%.

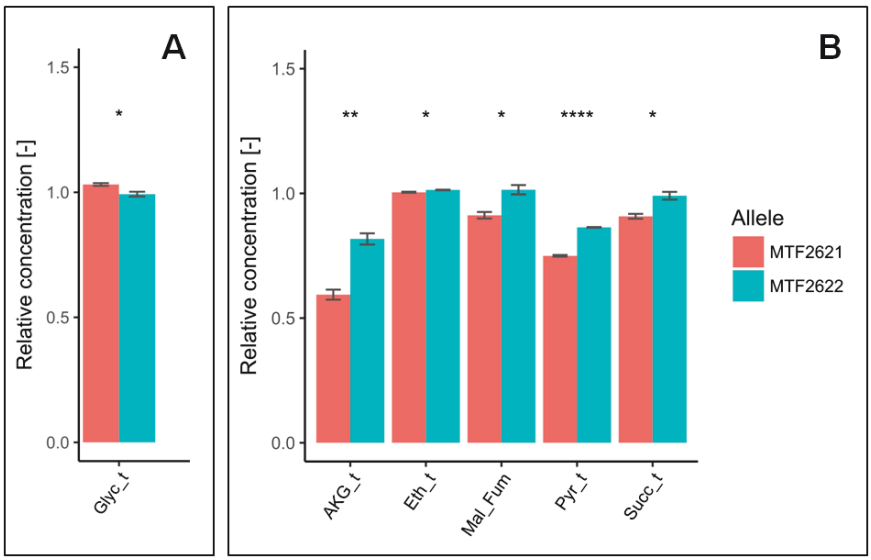


Figure 26: Allelic effect of *PDB1* (**A**) and *VID30* (**B**) on different estimated fluxes of yeast CCM. (p-value: * ≤ 0.05, ** ≤ 0.01, **** ≤ 0.0001)

3.4 Distribution of *PDB1* and *VID30* alleles in *S. cerevisiae* population

To visualize the natural variation of validated target genes within the *S. cerevisiae* population and to potentially link the variants to strain origins, phylogenetic trees were drawn using public available *PDB1* and *VID30* gene sequences (Figure 27 and Additional file 13).

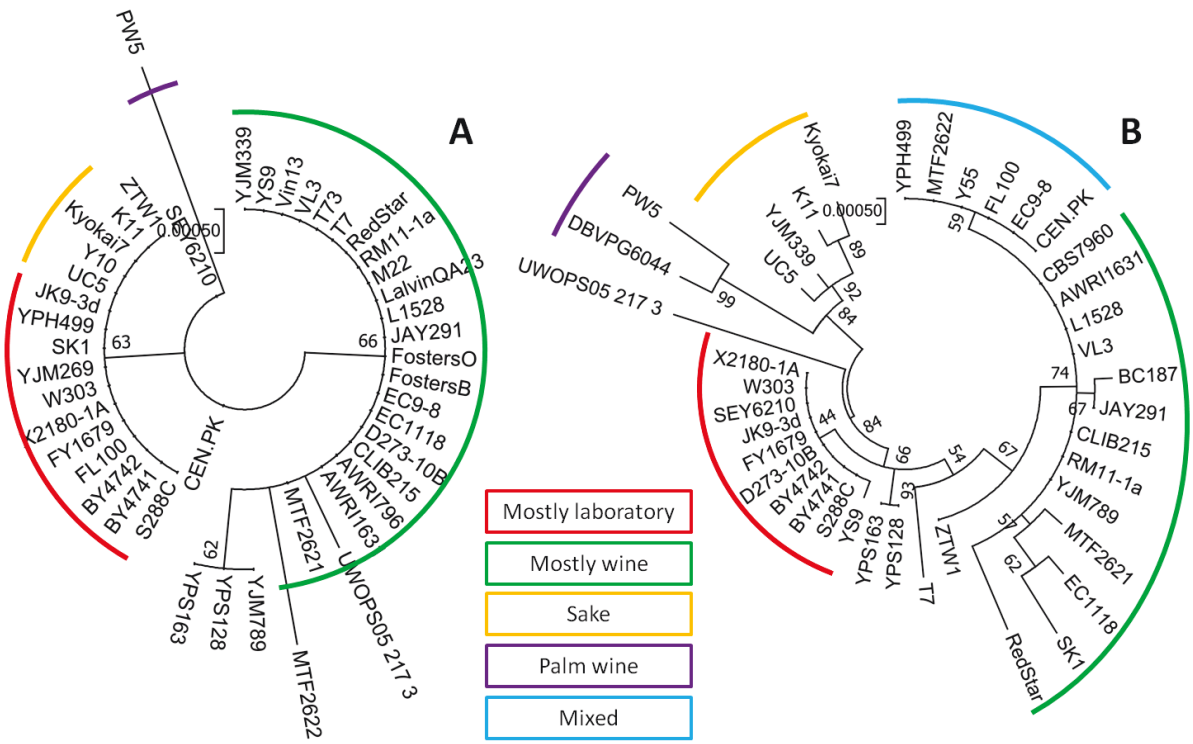


Figure 27: Phylogenetic trees of target genes (**A**) *PDB1* and (**B**) *VID30*, made from variant sequences of strains with different origin. Gene sequences were obtained from the *Saccharomyces* genome database (SGD). Maximum likelihood trees were constructed by bootstrap method with 200 replications using MEGA v7.0.26 (MEGA software).

Regarding *PDB1*, the allelic variants from yeast strains of different origin do not show much nucleotidic variation. Two main clusters can be seen, one consisting of mostly laboratory strains and the other of mostly wine strains. The allelic variants of the parental strains of this study are comparatively close. In contrast to this, the phylogenetic tree of *VID30* variants displays more variation between the strains. A probable explanation is the larger gene size. Several clusters can be distinguished, a laboratory strain cluster, a cluster consisting of African and sake strains and a wine strain cluster with a subcluster of mixed strains. The parental variants are more separated. While the MTF2621 allele is similar to the allele of strain EC1118, a genotypically close wine x flor strain, the MTF2622 allele is located within the mixed cluster.

4 Discussion

Many *S. cerevisiae* traits of interest for industrial applications are dependent on flux distributions within the CCM. We recently showed a pathway-dependent variability of flux distributions between *S. cerevisiae* strains, which was linked to the strain origin for some fluxes (Nidelet et al., 2016). These findings suggest the existence of a stock of genetic resources that can help to understand the genetic basis of flux distribution and to identify relevant targets for yeast strain improvement. In recent years, powerful methods such as QTL

mapping have been developed to link phenotypic and genomic variations. Our objective was to assess the potential of QTL mapping to detect genomic regions influencing metabolic fluxes (f-QTLs).

To this end, we used a population of 130 F2-segregants obtained from a cross of two wine yeast strains. Intracellular carbon fluxes were estimated for these strains using a constraint-based stoichiometric model of yeast CCM. Analysis of the flux deviations among the population of segregants indicated a positive correlation between fluxes of the PPP, TCA cycle oxidative branch and biomass formation, while these fluxes are negatively correlated to the upper glycolysis. Negative correlations were furthermore found between fluxes of the lower glycolysis/ethanol synthesis and glycerol formation (Figure 23). These observations are consistent with previous studies on *S. cerevisiae* strains (Heyland et al., 2009; Nidelet et al., 2016).

Although the parent strains do not display high variation for most modeled fluxes, a substantial variation among the segregants was observed (Figure 24). For some fluxes, for example PPP or main metabolite synthesis fluxes, the variation among segregants reached the variation among strains from different ecological origins determined by Nidelet et al. (2016). This emphasizes the complex nature of intracellular flux determination and indicates a rich genomic resource for metabolic profile optimization.

With 8 detected QTLs, the number of regions influencing metabolite yields during exponential phase was higher than the detected 4 QTLs influencing modeled metabolic fluxes. However, the regions from both QTL mappings differ from each other, showing that the modeling step was crucial for f-QTL detection.

All QTLs were compared to 8 loci detected by our previous study to influence extracellular metabolite production after 80% of fermentation using the same yeast cross (Eder et al., 2018). Only QTL chr7@18.0, which influences fluxes of the glycolysis, ethanol synthesis, biomass production, TCA cycle and transport/excretion of TCA cycle metabolites, was detected in our previous study to influence the pyruvate yield after 80% of fermentation. This indicates that the difference in flux distribution caused by QTL chr7@18.0 has a long-lasting effect on metabolite formation that can still be detected at the end of fermentation. However, regarding the concentrations of extracellular metabolites during exponential phase and after 80% of fermentation, all detected regions differ from each other and there are no common QTLs affecting metabolite production during both phases of fermentation. This indicates that different genomic regions could control metabolite production in growth phase and stationary phase, which actually corresponds to very different physiologic states of yeast during wine fermentation.

Within the two detected QTLs with the highest LOD-score, target genes with a potential role in CCM were identified and assessed by RHA. This study revealed the role of *PDB1* and *VID30* in the evaluated traits.

PDB1, shown to influence fluxes of the glycerol synthesis (Table 22, Figure 26), encodes the beta subunit of the pyruvate dehydrogenase (PDH), which is part of the large multienzyme PDH complex (Miran et al., 1993). Together with the other components dihydrolipoamide acetyltransferase and dihydrolipoamide dehydrogenase, the PDH complex converts pyruvate into acetyl-CoA (Pronk et al., 1996b). The allelic variants of the gene differ in three non-synonymous SNPs (Table 23) of which one, SNP V289I, lies in the pyruvate-ferredoxin oxidoreductase domain II of the protein. A possible explanation for the impact of Pdb1 variants on glycerol fluxes would be that the MTF2621 allele of Pdb1 shows an increased conversion rate of pyruvate to acetyl-CoA, which leads to a higher formation of the redox cofactor NADH. The resulting cofactor excess is then compensated through an increased glycerol synthesis, which maintains redox balance by NADH consumption (van Dijken and Scheffers, 1986). However, no significant difference in the estimated flux from pyruvate to acetyl-CoA (Pyr_Accoa_m) was detected by RHA for the alleles of *PDB1*.

The second validated gene, *VID30*, influences fluxes of the glycolysis/ethanol synthesis, TCA cycle reductive branch and excretion of TCA cycle metabolites (Table 22, Figure 26). Two functions of Vid30 could potentially account for these observed differences, the regulation of genes involved in glutamate/glutamine synthesis and the degradation of various metabolic enzymes.

The expression of *VID30* is repressed by ammonia and upregulated in response to low ammonia levels, a characteristic limitation during wine fermentation. Vid30 regulates various nitrogen catabolic genes, including *GDH1*, *GDH2*, *GDH3*, *GLN1* and *GLT1*. These genes express enzymes involved in the synthesis (and interconversion) of glutamate and glutamine from AKG and ammonia, therefore explaining the role of Vid30 for central carbon metabolism since AKG is part of fluxes of the TCA cycle oxidative branch. Gdh1, Gdh3 and Gln1 catalyze reactions from AKG to glutamine (Avendaño et al., 1997; Mitchell and Magasanik, 1984; Moye et al., 1985), whereas Gdh2 catalyzes the conversion of glutamate to AKG (Miller and Magasanik, 1990). Glt1 synthesizes glutamate from either AKG or glutamine (Filetici et al., 1996). In low ammonia environment, Vid30 behaves as a positive regulator for *GDH1*, *GDH3* and *GLT1*, which increases the flux from AKG to glutamate (van der Merwe et al., 2001). Since a decreased AKG production was detected for the MTF2621 allele of Vid30 (Figure 26), we suggest that this variant could stimulate an increased flux from AKG to glutamate through positive regulation of *GDH1*, *GDH3* and *GLT1*.

Another potential role of Vid30 in central carbon metabolism is its regulation of metabolic enzymes through degradation. When glucose-starved yeast is again transferred to glucose-rich medium, e.g., during inoculation, the metabolism increases the expression of glycolytic enzymes and simultaneously inactivates gluconeogenetic enzymes through catabolite inactivation. Vid30 possesses two functions in this process. It acts as a subunit of the glucose induced degradation (GID) protein complex that performs the ubiquitination of enzymes, which leads to their proteasome dependent inactivation (Hämmerle et al., 1998; Menssen et

al., 2012; Regelmann et al., 2003; Santt et al., 2008). Furthermore, Vid30 plays an important role for the formation of vesicles of the vacuole import and degradation pathway (Alibhoy et al., 2012), which carries out the degradation of enzymes expressed under growth on non-fermentable carbon sources (Huang and Chiang, 1997; Hung et al., 2004; Shieh and Chiang, 1998). Regulation peformed in this manner includes the turnover of hexose transporters Hxt3 and Hxt7 (Snowdon et al., 2007; Snowdon and Van der Merwe, 2012). Furthermore, various enzymes are regulated through degradation by Vid30 that catalyze gluconeogenesis reactions such as fructose-1,6-bisphosphatase, cytosolic malate dehydrogenase, isocitrate lyase and phosphoenolpyruvate carboxykinase (Carlson, 1999; Gancedo, 1998; Holzer, 1988; López-Boado et al., 1987). The reactions catalyzed by these enzymes strongly affect fluxes of the glycolysis and TCA cycle.

The allelic variants of *VID30* differ in three non-synonymous SNPs (Table 23) of which one, SNP V882I, lies in the CTLH/CRA domain of the protein, a protein-protein interaction domain also found in other components of the GID complex. We propose that the SNPs in the MTF2621 variant of Vid30 influence the protein's ability to inactivate hexose transporters and gluconeogenesis enzymes by degradation, hypothetically by an altered affinity to other components of the GID complex. This hypothesis is supported by the observed influence of the allelic variants on fluxes of the TCA cycle reductive branch (Figure 26), as the cytosolic malate dehydrogenase, which catalyzes the reaction from malate to oxaloacetate, is among the enzymes inactivated by Vid30 (Hung et al., 2004). Furthermore, the reported SNPs could affect the role of Vid30 in regulation of enzymes involved in the synthesis of glutamate from AKG. This hypothesis is supported by the detected significant influence of the Vid30 alleles on AKG formation (Figure 26). On the other hand, significant difference between the alleles in the flux from AKG to glutamate could not be detected by RHA. The difference in AKG formation could also be explained by the role of Vid30 for the degradation of isocitrate lyase. The enzyme catalyzes the reaction from isocitrate to succinate, which could influence AKG synthesis.

5 Conclusion

In this study we prove the feasibility of using modeled phenotypic data to detect regions in the genome with an influence on the underlying traits. We used extracellular main metabolites to estimate intracellular fluxes of the central carbon metabolism of *S. cerevisiae* using a constraint-based model. This led us to the integration of otherwise independent quantifiable traits. With this approach we detected 4 QTLs with an influence on 4 main metabolic pathways and various metabolite transport and excretion fluxes. These QTLs could not be found by linkage analysis considering extracellular metabolite concentration alone. As reported in the literature and seen in our study, variances of certain intracellular fluxes, such as the glycolysis, are in general low. Therefore, increased statistical power is needed for a more thorough determination of the impact of genomic variation on these fluxes. This could

be achieved by increasing the number of segregants or by performing multiple QTL mapping strategies, which has the potential to find QTLs with minor contributions.

The relevance of our approach was further confirmed by the validation of two target genes within found QTLs, *PDB1* and *VID30*. The allelic variants of *PDB1* cause differences in fluxes of the glycerol synthesis that we connected to redox imbalances as a result of altered pyruvate conversion. The variants of *VID30* impact fluxes of the glycolysis, ethanol synthesis and TCA cycle, which we propose to be caused by different regulation of enzymes catalyzing glutamate formation or a different catabolite induced degradation of enzymes involved in sugar uptake, gluconeogenesis and TCA cycle.

Compared to strains of other origins, the parental variants of the evaluated target genes are comparatively close. The characterization of more distant variants and the evaluation of their influence on intracellular flux distributions will increase knowledge about genetic resources that bear further potential to shape the metabolic profile of strains.

In summary, our findings of QTLs and allelic variants impacting metabolic fluxes increase knowledge about the links between genomic variation and yeast metabolic properties. The fact that we could demonstrate the applicability of QTL mapping on modeled phenotypic data, especially modeled metabolic fluxes, will open ways to improve strains not only for fermented beverages but for manifold purposes, e.g., the production of biofuels or other bulk and fine chemicals.

6 Supplementary Information

Additional file 8: List of primers used in this study with nucleotide sequence (5' -> 3')

Primer name	Nucleotide sequence (5' -> 3')
del_HAP2_fw	TGGAAGAGGAACAAGAACGCCATGTCAGCAGACGAAACGGTTCGTACGCTGCAGGTCGAC
del_HAP2_rv	TAAATAGGCCATATGGATACCATGTGGTATAAGAGGGCACGCATAGGCCACTAGTGGATCTG
del_PDB1_fw	CCTGTGTTTGTTCATTGATAATCGATCGCAGTTTAGTAAGTTCGTACGCTGCAGGTCGAC
del_PDB1_rv	ACTATTTCCGCGAAGAGGGTAGAAAGTGTAGGGTACAGGGGCATAGGCCACTAGTGGATCTG
del_VID30_fw	CGTTAAAGCCAAGCGTCGAATTTCAGCATAATTAAGAGGATTCGTACGCTGCAGGTCGAC
del_VID30_rv	ATGACTGATATCACATGGCTTTGTTGTTTGAAGGTGCTTGGCATAGGCCACTAGTGGATCTG
Hygro_rv	TGTTATGCGGCCATTGTC
test_HAP2_fw	CGTACAGCCATTGACCATAG
test_PDB1_fw	AATTCCGCCTCCCTCATAAC
test_VID30_fw	ACCTCTACTTCGACCATCAC

Additional file 9: Abbreviations for metabolites of the CCM and their compartmental localization.

Metabolite abbreviation	Compartment	Metabolite description
13dpg[c]	cytoplasm	3-Phospho-D-glyceroyl-phosphate
2pg[c]	cytoplasm	D-Glycerate-2-phosphate
3pg[c]	cytoplasm	3-Phospho-D-glycerate
6pgc[c]	cytoplasm	6-Phospho-D-gluconate
6pgl[c]	cytoplasm	6-phospho-D-glucono-1,5-lactone
ac[c]	cytoplasm	Acetate
ac[m]	mitochondria	Acetate
acald[c]	cytoplasm	Acetaldehyde
acald[m]	mitochondria	Acetaldehyde
accoa[c]	cytoplasm	Acetyl-CoA
accoa[m]	mitochondria	Acetyl-CoA
adp[c]	cytoplasm	ADP
adp[m]	mitochondria	ADP
akg[c]	cytoplasm	Alpha ketoglutarate
akg[m]	mitochondria	Alpha ketoglutarate
amp[m]	mitochondria	AMP
atp[c]	cytoplasm	ATP
atp[m]	mitochondria	ATP
cit[m]	mitochondria	Citrate
CO_2[c]	cytoplasm	CO_2
CO_2[m]	mitochondria	CO_2
coa[c]	cytoplasm	Coenzyme-A
coa[m]	mitochondria	Coenzyme-A
dhap[c]	cytoplasm	Dihydroxyacetone-phosphate
e4p[c]	cytoplasm	D-Erythrose-4-phosphate
etoh[c]	cytoplasm	Ethanol
etoh[m]	mitochondria	Ethanol
f6p[c]	cytoplasm	D-Fructose-6-phosphate
fdp[c]	cytoplasm	D-Fructose-1,6-bisphosphate

fum[c]	cytoplasm	Fumarate
fum[m]	mitochondria	Fumarate
g3p[c]	cytoplasm	Glyceraldehyde-3-phosphate
g6p[c]	cytoplasm	D-Glucose-6-phosphate
glc[c]	cytoplasm	D-Glucose
gln[c]	cytoplasm	L-Glutamine
glu[c]	cytoplasm	L-Glutamate
glu[m]	mitochondria	L-Glutamate
glyc[c]	cytoplasm	Glycerol
glyc3p[c]	cytoplasm	Glycerol-3-phosphate
icit[m]	mitochondria	Isocitrate
mal[c]	cytoplasm	L-Malate
mal[m]	mitochondria	L-Malate
nad[c]	cytoplasm	Nicotinamide-adenine-dinucleotide
nad[m]	mitochondria	Nicotinamide-adenine-dinucleotide
nadh[c]	cytoplasm	Nicotinamide-adenine-dinucleotide-reduced
nadh[m]	mitochondria	Nicotinamide-adenine-dinucleotide-reduced
nadp[c]	cytoplasm	Nicotinamide-adenine-dinucleotide-phosphate
nadp[m]	mitochondria	Nicotinamide-adenine-dinucleotide-phosphate
nadph[c]	cytoplasm	Nicotinamide-adenine-dinucleotide-phosphate-reduced
nadph[m]	mitochondria	Nicotinamide-adenine-dinucleotide-phosphate-reduced
oaa[c]	cytoplasm	Oxaloacetate
oaa[m]	mitochondria	Oxaloacetate
pyr[c]	cytoplasm	Pyruvate
pyr[m]	mitochondria	Pyruvate
r5p[c]	cytoplasm	alpha-D-Ribose-5-phosphate
ru5p[c]	cytoplasm	D-Ribulose-5-phosphate
s7p[c]	cytoplasm	Sedoheptulose-7-phosphate
succ[c]	cytoplasm	Succinate
succ[m]	mitochondria	Succinate
succoa[m]	mitochondria	Succinyl-CoA
xu5p[c]	cytoplasm	D-Xylulose-5-phosphate

Additional file 10: Estimated metabolic reactions with their corresponding compartment.

Flux abbrevation*	**Compartment**	**Reaction**
Glc_G6p	cytoplasm	glc[c] + atp[c] → g6p[c] + adp[c]
G6p_F6p	cytoplasm	g6p[c] ⇌ f6p[c]
F6p_Fdp	cytoplasm	f6p[c] + atp[c] → fdp[c] + adp[c]
Fdp_Dhap	cytoplasm	fdp[c] ⇌ dhap[c] + g3p[c]
Dhap_G3p	cytoplasm	dhap[c] ⇌ g3p[c]
G3p_13dpg	cytoplasm	g3p[c] + nad[c] ⇌ 13dpg[c] + nadh[c]
13dpg_3pg	cytoplasm	13dpg[c] + adp[c] ⇌ 3pg[c] + atp[c]
3pg_2pg	cytoplasm	3pg[c] ⇌ 2pg[c]
2pg_Pep	cytoplasm	2pg[c] ⇌ pep[c]
Pep_Pyr	cytoplasm	pep[c] + adp[c] → pyr[c] + atp[c]
G6p_6pgl	cytoplasm	g6p[c] + nadp[c] ⇌ 6pgl[c] + nadph[c]
6pgl_6pgc	cytoplasm	6pgl[c] → 6pgc[c]

6pgc_Ru5p	cytoplasm	6pgc[c] + nadp[c] → CO_2[c] + nadph[c] + ru5p[c]
Ru5p_Xu5p	cytoplasm	ru5p[c] ⇌ xu5p[c]
Ru5p_R5p	cytoplasm	ru5p[c] ⇌ r5p[c]
R5p_S7p	cytoplasm	r5p[c] + xu5p[c] ⇌ g3p[c] + s7p[c]
E4p_F6p	cytoplasm	e4p[c] + xu5p[c] ⇌ f6p[c] + g3p[c]
S7p_E4p	cytoplasm	g3p[c] + s7p[c] ⇌ e4p[c] + f6p[c]
Dhap_Glyc3p	cytoplasm	dhap[c] + nadh[c] → glyc3p[c] + nad[c]
Glyc3p_Glyc	cytoplasm	glyc3p[c] → glyc[c]
Pyr_Acald	cytoplasm	pyr[c] → acald[c] + CO_2[c]
Acald_Eth	cytoplasm	acald[c] + nadh[c] → etoh[c] + nad[c]
Acald_Ac	cytoplasm	acald[c] + nadp[c] → ac[c] + nadph[c]
Ac_Accoa	cytoplasm	ac[c] + 2 atp[c] → accoa[c] + 2 adp[c]
Pyr_Oaa	cytoplasm	pyr[c] + atp[c] + CO_2[c] → oaa[c] + adp[c]
Acald_Eth_m	mitochondria	acald[m] + nadh[m] ⇌ etoh[m] + nad[m]
Acald_Ac_m	mitochondria	acald[m] + nadp[m] → ac[m] + nadph[m]
Oaa_Mal	cytoplasm	oaa[c] + nadh[c] ⇌ mal[c] + nad[c]
Mal_Fum	cytoplasm	mal[c] → fum[c]
Fum_Succ	cytoplasm	fum[c] → succ[c]
Akg_Glu	cytoplasm	akg[c] + nadph[c] ⇌ glu[c] + nadp[c]
Glu_Akg_m	mitochondria	glu[m] + nad[m] → akg[m] + nadh[m]
Pyr_Accoa_m	mitochondria	pyr[m] + nad[m] → accoa[m] + nadh[m] + CO_2[m]
Oaa_Cit_m	mitochondria	accoa[m] + oaa[m] → cit[m]
Cit_Icit_m	mitochondria	cit[m] ⇌ icit[m]
Icit_Akg_m_nad	mitochondria	icit[m] + nad[m] → akg [m]+ CO_2[m] + nadh[m]
Icit_Akg_m_nadp	mitochondria	icit[m] + nadp[m] → akg[m] + CO_2[m] + nadph[m]
Akg_Succoa_m	mitochondria	akg[m] + nad[m] → succoa[m] + CO_2[m] + nadh[m]
Succoa_Succ_m	mitochondria	succoa[m] + adp[m] → succ[m] + atp[m]
Oaa_Mal_m	mitochondria	oaa[m] + nadh[m] → mal[m] + nad[m]
Mal_Fum_m	mitochondria	mal[m] → fum[m]
Fum_Succ_m	mitochondria	fum[m] → succ[m]
Mal_Pyr_m	mitochondria	mal[m] + nadp[m] → CO_2[m] + nadph[m] + pyr[m]
Acald_tm	transport	acald[c] ⇌ acald[m]
Succ_tm	transport	succ[c] + atp[c] → succ[m] + adp[c]
Mal_tm	transport	mal[c] + atp[c] → mal[m] + adp[c]
Mal_Succ_tm	transport	mal[c] + succ[m] ⇌ mal[m] + succ[c]
Pyr_tm	transport	pyr[c] + atp[c] → pyr[m] + adp[c]
Akg_tm	transport	akg[c] ⇌ akg[m]
Oaa_tm	transport	oaa[c] + atp[c] → oaa[m] + adp[c]
Eth_tm	transport	etoh[c] ⇌ etoh[m]
CO_2_tm	transport	CO_2[c] ⇌ CO_2[m]
Ac_tm	transport	ac[c] ⇌ ac[m]
Accoa_tm	transport	accoa[c] → accoa[m]
Glu_tm	transport	glu[c] + atp[c] → glu[m] + adp[c]
Glc_t	transport	glc[c] →
Eth_t	transport	etoh[c] →
Ac_t	transport	ac[c] →
Pyr_t	transport	pyr[c] →
Akg_t	transport	akg[c] →
Succ_t	transport	succ[c] →

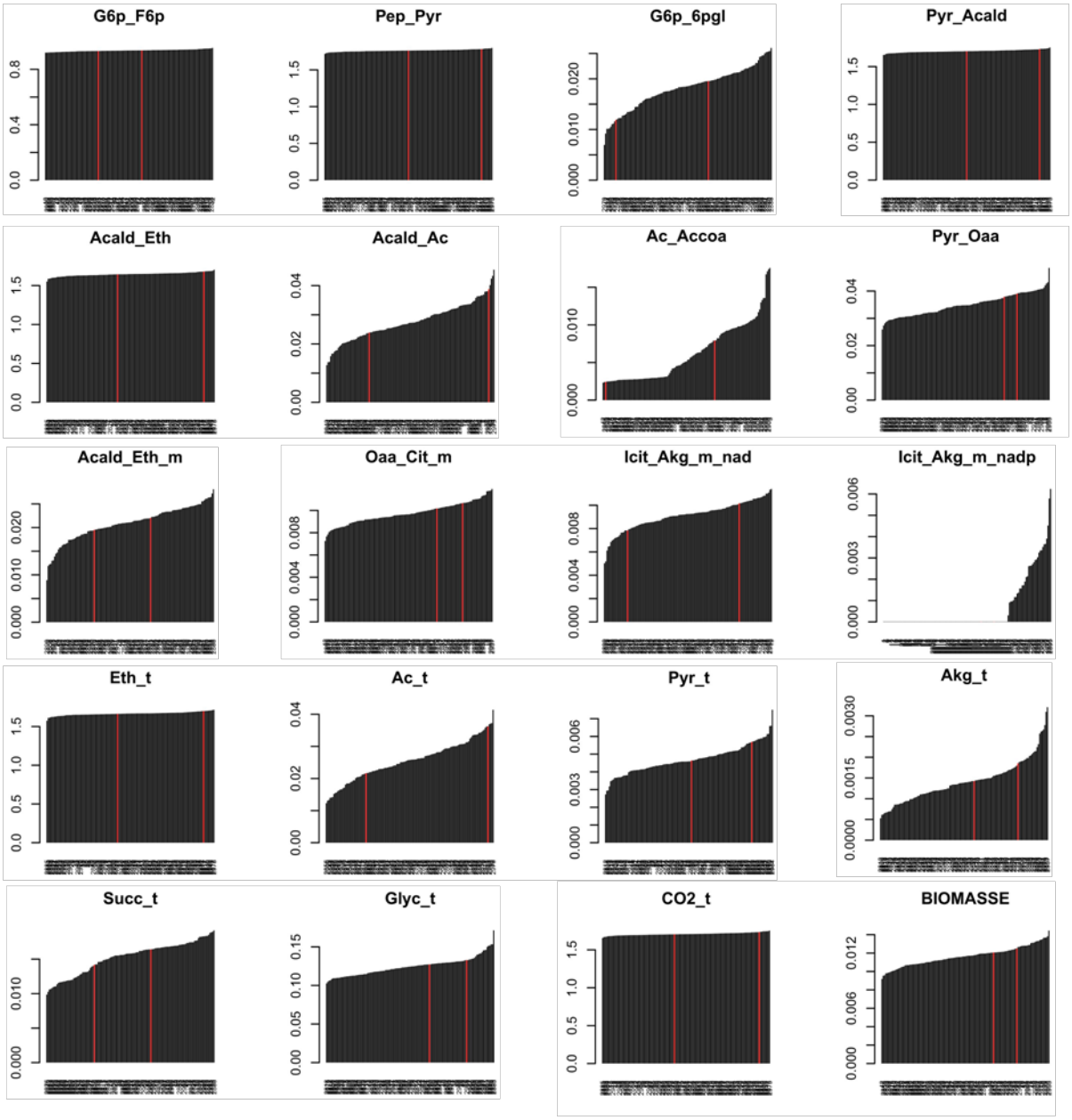
But_t	transport	but[c] →
Aceto_t	transport	aceto[c] →
Acald_t	transport	acald[c] →
Glyc_t	transport	glyc[c] →
CO_2_t	transport	CO_2[c] ⇌
ATP_Shuttle	cytoplasm	atp ⇌ adp
BIOMASS	biomass	3.96 g6p[c] + 0.258 r5p[c] + 0.129 e4p[c] + 0.116 g3p[c] + 0.303 3pg[c] + 0.232 pep[c] + 0.775 oaa[c] + 1.084 pyr[m] + 0 pyr[c] + 0.176 accoa[m] + 0.252 accoa[c] + 0.106 akg[m] + 0.366 akg[c] + 0 CO_2[c] + 0.136 glu[c] + 115 atp[c] + 0.106 atp[m] + 1.499 nad[c] + 0.176 nad[m] + 0.602 nadph[m] + 5.35 nadph[c] → 115 adp[c] + 0.106 adp[m] + 1.499 nadh[c] + 0.176 nadh[m] + 0.602 nadp[m] + 5.35 nadp[c]

* flux abbreviations are encoded as substrate and product connected with "_". For mitochondrial reactions we added "_m". Extracellular transport and mitochondrial transport reactions are marked with "_t" and a "_tm" respectively. Metabolite abbreviations can be found in Additional file 10.

Additional file 11: Used selection of 20 fluxes representative for main metabolic pathways.

Flux abbrevation*	**Pathway**	**Reaction**
G6p_F6p	Upper glycolysis	g6p[c] ⇌ f6p[c]
Pep_Pyr	Lower glycolysis	pep[c] + adp[c] → pyr[c] + atp[c]
G6p_6pgl	PPP	g6p[c] + nadp[c] ⇌ 6pgl[c] + nadph[c]
Pyr_Acald	Ethanol synthesis	pyr[c] → acald[c] + CO_2[c]
Acald_Eth	Ethanol synthesis	acald[c] + nadh[c] → etoh[c] + nad[c]
Acald_Ac	Acetate metabolism	acald[c] + nadp[c] → ac[c] + nadph[c]
Ac_Accoa	Ac-CoA metabolism	ac[c] + 2 atp[c] → accoa[c] + 2 adp[c]
Pyr_Oaa	TCA reductive branch	pyr[c] + atp[c] + CO_2[c] → oaa[c] + adp[c]
Acald_Eth_m	Ethanol synthesis	acald[m] + nadh[m] ⇌ etoh[m] + nad[m]
Oaa_Cit_m	TCA oxidative branch	accoa[m] + oaa[m] → cit[m]
Icit_Akg_m_nad	TCA oxidative branch	icit[m] + nad[m] → akg [m]+ CO_2[m] + nadh[m]
Icit_Akg_m_nadp	TCA oxidative branch	icit[m] + nadp[m] → akg[m] + CO_2[m] + nadph[m]
Eth_t	Ethanol excretion	etoh[c] →
Ac_t	Acetate excretion	ac[c] →
Pyr_t	Pyruvate excretion	pyr[c] →
Akg_t	AKG excretion	akg[c] →
Succ_t	Succinate synthesis	succ[c] →
Glyc_t	Glycerol synthesis	glyc[c] →
CO_2_t	CO_2 synthesis	CO_2[c] ⇌
BIOMASS	Biomass formation	3.96 g6p[c] + 0.258 r5p[c] + 0.129 e4p[c] + 0.116 g3p[c] + 0.303 3pg[c] + 0.232 pep[c] + 0.775 oaa[c] + 1.084 pyr[m] + 0 pyr[c] + 0.176 accoa[m] + 0.252 accoa[c] + 0.106 akg[m] + 0.366 akg[c] + 0 CO_2[c] + 0.136 glu[c] + 115 atp[c] + 0.106 atp[m] + 1.499 nad[c] + 0.176 nad[m] + 0.602 nadph[m] + 5.35 nadph[c] → 115 adp[c] + 0.106 adp[m] + 1.499 nadh[c] + 0.176 nadh[m] + 0.602 nadp[m] + 5.35 nadp[c]

* flux abbreviations are encoded as substrate and product connected with "_". For mitochondrial reactions we added "_m". Extracellular transport and mitochondrial transport reactions are marked with "_t" and a "_tm" respectively. Metabolite abbreviations can be found in Additional file 10.



Additional file 12: Distributions of selected estimated CCM fluxes between the segregant strains and position of parent strains within the population (red lines).

Additional file 13: List of strains used for the phylogenetic analysis of target genes *PDB1* and *VID30*. Genomic sequences and description of strain origin were obtained from the *Saccharomyces* genome database (SGD).

Strain	Origin
DBVPG6044	african
PW5	african
SK1	african
Y55	african
Y10	asian
FostersB	beer
FostersO	beer

CLIB215	bread
YS9	bread
YJM339	clinical
YJM789	clinical
ZTW1	ethanol
BY4741	laboratory
BY4742	laboratory
CEN.PK	laboratory
D273-10B	laboratory
FL100	laboratory
FY1679	laboratory
JK9-3d	laboratory
S288C	laboratory
SEY6210	laboratory
W303	laboratory
X2180-1A	laboratory
YPH499	laboratory
T7	oak NA
YPS128	oak NA
YPS163	oak NA
CBS7960	rum
JAY291	rum
K11	sake
Kyokai7	sake
UC5	sake
EC9-8	soil
UWOPS05_217_3	soil
AWRI1631	wine
AWRI796	wine
BC187	wine
L1528	wine
LalvinQA23	wine
M22	wine
RedStar	wine
RM11-1a	wine
T73	wine
VL3	wine
YJM269	wine
EC1118	wine x flor
Vin13	wine x flor

<u>Chapter 3:</u> Determination of genomic regions with an influence on yeast's characteristic to metabolize the grape-derived aroma precursor *S*-methylmethionine

Dimethylsulfide (DMS) is a potent aroma enhancer and contributor in wine. However, yeast can metabolize DMS precursors during alcoholic fermentation and therefore decreases the potential level of the volatile in wine. This chapter of the thesis applies the established QTL mapping strategy to the detection of genomic regions behind the yeast's ability to influence the level of the grape-derived DMS precursor *S*-methylmethionine (SMM). For this purpose, SMM was added to the synthetic must and all 130 F2-segregants were phenotyped for their ability to leave SMM in the medium at the end of fermentation. The results were used as phenotype data for QTL mapping, using the marker map that was obtained during the first part of the project.

During this thesis project, a secondment was done at the group for Industrial Microbiology of the TU Delft, The Netherlands, under supervision of Dr. Jean-Marc Daran. During this secondment, the CRISPR/Cas9 genome editing system was established in the yeast strains used in this study.

The chapter is composed as a research article for submission.

Genomic bases for the metabolism of the DMS precursor *S*-methylmethionine by *Saccharomyces cerevisiae*

Matthias Eder[a], Isabelle Sanchez[a,b], Carole Camarasa[a], Jean-Marc Daran[c], Jean-Luc Legras[a], Sylvie Dequin[a,*]

[a]SPO, INRA, SupAgro, Université de Montpellier, F-34060 Montpellier, France

[b]MISTEA, INRA, SupAgro, F-34060 Montpellier, France

[c]Department of Biotechnology, Delft University of Technology, NL-2629 HZ Delft, The Netherlands

* Corresponding author. Mailing address: INRA, 2 Place Pierre Viala, 34060 Montpellier CEDEX 2, France; Phone: +33 4 99 61 25 28; E-mail: sylvie.dequin@inra.fr

E-mail addresses:
Matthias Eder: matthias.eder@inra.fr
Isabelle Sanchez: isabelle.sanchez@inra.fr
Carole Camarasa: carole.camarasa@inra.fr
Jean-Marc Daran: j.g.daran@tudelft.nl
Jean-Luc Legras: jean-luc.legras@inra.fr
Sylvie Dequin: sylvie.dequin@inra.fr

Abstract

Dimethyl sulfide (DMS) is a well described sulfur containing flavor active volatile that enhances general fruity aroma and imparts aromatic notes in wine. The most important precursor of DMS is *S*-methylmethionine (SMM), which is synthesized by grapes and can be metabolised by the yeast *S. cerevisiae* during wine fermentation. Precursor molecules left after fermentation are chemically converted to DMS during wine maturation, meaning that the level of DMS in wine is determined by the amount of remaining precursors at bottling.

To elucidate SMM metabolism in yeast we performed a quantitative trait locus (QTL) mapping approach using a population of 130 F2-segregants obtained from a cross between two wine yeast strains and we detected one major QTL explaining almost 30% of the trait variation. Within the QTL, the gene *YLL058W* and SMM transporter gene *MMP1* were found to have an influence on SMM metabolism, from which *MMP1* has the bigger impact. We identified and characterized a variant coding for a truncated transporter with superior SMM preserving attributes. A population analysis with 85 yeast strains from different origins revealed a significant association of the variant to flor strains and minor occurrence in cheese and wine strains.

These results will help in the selection and improvement of *S. cerevisiae* strains for the production of wine and other fermented foods containing DMS such as cheese or beer.

Keywords: QTL mapping, DMS, SMM metabolism, varietal aroma, yeast

1 Introduction

The perception of wine is tightly linked to its aroma, a complex blend of compounds with low boiling points, which are therefore volatile. These molecules can emerge the liquid in the wine glass to become detectable by the human nose (Swiegers et al., 2005). Over 1000 aroma compounds have been found and described in wine (Tao and Li, 2009). Depending on their origin these compounds can be divided into classes. Varietal aromas are contributed by the grapes and are distinctive for each grape variety. Fermentative aromas are produced by yeast and bacteria during alcoholic and malolactic fermentation. Post-fermentative aromas develop during conservation and aging of wine due to occurring transformations. However, these classes are not strictly separated. Varietal aroma contributors, for example, may exist as odorless precursors in the grape berry, which are then transformed to odorous compounds during harvesting, winemaking or ageing.

Dimethyl sulfide (DMS), a well described sulfur-containing odorant (Anness and Bamforth, 1982), is one of the most important aroma compounds released from grape derived precursors during the maturation of wine (Simpson, 1979). The olfactory perception threshold of DMS was described to be 27 µg/L in red wines (Beloqui et al., 1996) and 25 µg/L in white wines (Spedding and Raut, 1982). DMS is known to confer pleasant notes to Cabernet-Sauvignon (De Mora et al., 1987) and Shiraz red wines (Segurel, 2005). Furthermore, different white wines with small additions of DMS were preferred over their non-treated counterparts (Spedding and Raut, 1982). In low concentrations, DMS enhances the berry fruit aroma of wines (De Mora et al., 1987; Escudero et al., 2007; Lytra et al., 2016) and was reported to impart additional truffle and black olive notes (Segurel et al., 2004). In fact, DMS is a characteristic aroma molecule of black and summer truffles (Culleré et al., 2010). At higher concentrations and in certain wines, the presence of DMS is perceived less pleasant (Spedding and Raut, 1982). The compound can impart vegetal or molasses notes (Mestres et al., 2000), and DMS was described as a characteristic flavor compound in many raw and processed vegetables, such as tomatoes, asparagus, broccoli or cooked corn (Buttery et al., 1971; Dignan and Wiley, 1976; Tulio et al., 2002; Ulrich et al., 2001).

Different grape derived precursor molecules are able to form DMS in wine and therefore determine the DMS potential (pDMS). Besides from dimethyl sulfoxide (DMSO), yeast can generate DMS from the sulfur amino acids cysteine, cystine and glutathione (Anness and Bamforth, 1982; de Mora et al., 1986). With a vapor pressure of 53 kPa at 20 °C DMS is relatively volatile (Lestremau et al., 2003), and molecules produced during fermentation are mostly driven off by CO_2, which results in low DMS concentrations observed just after the fermentation (Dagan, 2006).

The main precursor of DMS, however, was identified to be *S*-methylmethionine (SMM), an amino acid derivative accounting for more than 70% of formed DMS (Loscos et al., 2008). The synthesis of SMM is only reported in plants, where it is synthesized from methionine and AdoMet (Mudd and Datko, 1990). SMM plays an important role for the phloem sulfur

transport (Bourgis et al., 1999) and in addition acts as a precursor for the biosynthesis of dimethylsulfoniopropionate, an osmoprotectant (Trossat et al., 1998).

The uptake of SMM into the cell is achieved by the high affinity SMM permease, encoded by gene *MMP1,* and by another low affinity transporter system that could also account for the low affinity uptake of *S*-adenosyl methionine (AdoMet), an intermediate of methionine metabolism (Rouillon et al., 1999). Inside the cell, SMM is metabolized, together with homocystein, into two molecules of methionine by the SMM-homocystein methyl-transferase Mht1 (Figure 28) (Thomas et al., 2000). SMM can therefore be used as a sulfur source by the yeast (Rouillon et al., 1999).

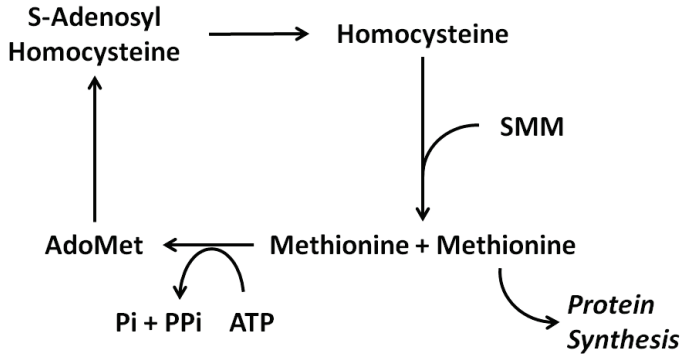


Figure 28: Metabolism of SMM by *S. cerevisiae*. Two molecules of methionine are synthesized from SMM and homocysteine; One methionine molecule is recycled to homocysteine, the other methionine molecule can be used for protein synthesis; adapted from Thomas et al. (2000).

The genes *MMP1* and *MHT1* are clustered in the genome of *S. cerevisiae* and oriented in opposite direction, being only separated by a *cis*-acting regulator sequence for the MET gene network (Thomas et al., 2000), a set of genes implicated in the steps of methionine biosynthesis (reviewed by Thomas and Surdin-Kerjan, 1997). *MHT1* is regulated along with other MET genes and both genes, *MMP1* and *MHT1*, are repressed by methionine and require the Met4p activator for activation (Thomas et al., 2000).

The presence of DMS in wine mainly depends on its release from conserved precursor molecules through a chemical process during wine aging (Segurel et al., 2005). Only a small proportion of pDMS is recovered in young wines. The recovery rate depends on numerous factors, such as the yeast strain, yeast assimilable nitrogen content in must and the general winemaking process (Dagan and Schneider, 2012). Nevertheless, the remaining pDMS leads to the release of DMS during maturation, and DMS levels far exceeding the olfactory threshold were detected in older vintages (Dagan, 2006; de Mora et al., 1993). Although the release of DMS is dependent on the wine storage conditions, mostly storage temperature, the final level is determined by the pDMS content at bottling (Fedrizzi et al., 2007; Marais, 1979).

As the decrease of SMM is a metabolic process and dependent on several genes, the genetic properties of yeast are expected to account for trait variation, and differences between yeast strains in pDMS preservation could be observed (Dagan and Schneider, 2012). One approach to link genetic to phenotypic variation is QTL mapping, which has already been successfully applied in yeast to decipher enological important traits (Ambroset et al., 2011; Brice et al., 2014a; Martí-raga et al., 2017; Marullo et al., 2007; Noble et al., 2015; Salinas et al., 2012; Steyer et al., 2012). In addition to investigate fermentative aroma formation, Steyer et al. (2012) used a QTL mapping approach to assess the role of yeast metabolism on the transformation of grape derived aroma molecules during wine fermentation, and they found 5 genomic regions influencing the alteration of grape terpenols. However, the impact of yeast metabolism on the degradation of pDMS has never been investigated until now.

The aim of this study was to understand the impact of genetic variation on the metabolism of pDMS. We performed a QTL mapping approach with a set of 130 F2-segregants previously used to detect genomic regions influencing fermentative aroma formation (Eder et al., 2018). We detected a locus in the yeast genome with a dominant effect on SMM preservation and could validate two contained genes within the QTL, *YLL058W* and *MMP1*, to have an impact on the trait. The SMM-transporter gene *MMP1* showed to have the major impact on SMM metabolism and an Mmp1 variant with superior preservation properties was identified and characterized. The detection of *MMP1* and *YLL058W* variants will help to select commercially available yeast starter cultures for the production of more aromatic wines and other fermented foods, or to improve commercial *S. cerevisiae* strains through non-genetically modified organisms (GMO) methods, like breeding via marker-assisted selection.

2 Materials and methods

2.1 Media

Yeast strains were cultured at 28 °C in yeast extract peptone dextrose (YPD) media, containing 10 g/L yeast extract, 20 g/L peptone and 20 g/L glucose. Solid YPD media contained 1.5% agar. Selective YPD media contained 200 µg/mL geneticin (G418), 200 µg/mL nourseothricin (clonNAT) or 200 µg/mL hygromycin B were used.

Wine fermentations were carried out in synthetic grape must (MS200) described by Bely et al. (1990). It contains 100 g/L of each sugar, glucose and fructose, and 200 mg/L of assimilable nitrogen. The amino acid composition hereby mimics the nitrogen content of standard grape juice.

2.2 Strain selection, segregant generation and genotyping

The haploid *S. cerevisiae* strains MTF2621 (4CAR1 [Δ*HO*::*Neo*^{r}]) and MTF2622 (T73 [Δ*HO*::*Nat*^{r}]) were selected for the study according to their different need for nitrogen during wine fermentation. The requirement of nitrogen was estimated by using an approach based on the addition of nitrogen to keep the CO_2 production rate constant during nitrogen limitation (Brice et al., 2014b). The strain T73 belongs to the phylogenetic clade of wine strains. The strain 4CAR1, however, belongs to the group of champagne strains, which originated through crossings between strains of the wine clade and the flor clade (Coi et al., 2016). For this study, we used a population of 130 segregants of the F2 generation previously obtained from a cross between these two strains (Eder et al., 2018). The strains had been genotyped by whole genome sequencing in order to generate a marker map for QTL mapping.

2.3 Phenotyping of strains

Segregants were fermented in duplicates with the parent strains as controls. Strains were precultured overnight in 50 mL YPD media and cell density was determined using a Multisizer™ 3 Coulter Counter (Beckman Coulter). Sterilized 300-mL glassware mini fermenters were filled with 280 mL of SM200, supplemented with 500 µg/L *S*-methylmethionine chloride and closed with an air lock. Fermenters were inoculated to a cell density of 1×10^6 cells/mL and incubated at 24 °C under stirring (300 rpm). Flasks were weighted twice daily to determine CO_2 production, which directly reflects sugar consumption. Samples were taken when approx. 80% of the sugars were depleted in order to determine the concentration of *S*-methylmethionine near the end of the fermentation. These samples were then analyzed for the quantity of remaining pDMS by solid phase micro extraction and GC-MS (Nyseos in Montpellier, France). pDMS concentrations are given as equivalent of DMS formation.

2.4 QTL mapping

The phenotyping and genotyping data were used to identify QTLs in the yeast genome that influence pDMS metabolism during wine fermentation. The statistical analyses were carried out using the programming language R v3.2.3 (www.r-project.org) with the R/qtl v1.40-8 and R/eqtl v1.1-7 libraries (Broman et al., 2003). QTL mapping was performed with two different phenotype models, the normal model using Haley-Knott regression and a non-parametric analysis, resulting in logarithm of odds (LOD) scores for each marker and pseudo-marker every 2.5 cM (interval mapping method). An interval estimate of the location of each QTL was obtained as the 1-LOD support interval, the region in which the LOD score is within 1 unit of the peak LOD score. If the same locus was detected with both models, the results with the higher LOD score were selected.

2.5 Reciprocal hemizygosity analysis

Molecular dissection of QTLs was performed using reciprocal hemizygosity analysis (RHA) (Steinmetz et al., 2002; Warringer et al., 2017). The gene sequences of selected target genes were deleted in both parent strains by homologous recombination with a disruption cassette containing the hygromycin B resistance gene (*hph*^{r}). Positive integration was selected by plating the transformed cells on YPD-agar plates containing hygromycin B. Correct gene deletion was verified by PCR using the primer test_[GENE]_fw which binds in the upstream region of the deleted gene and the primer Hygro_rv which binds within the deletion cassette. Deleted parent strains were subsequently mated with the opposite undeleted parent to form a heterozygote that is hemizygous for the target gene. The obtained strains were phenotyped in triplicate (2.3). Significance of the impact of allelic variation on the trait was evaluated by student's t-test.

2.6 Allelic swap

The impact of the allelic variants on the phenotype was validated by exchanging the corresponding sequences between the parent cells, using the CRISPR/Cas9 toolbox developed by Mans et al. (2015). This approach has the advantage that genes and even sole SNPs can be rapidly exchanged in a marker- and scarless way. The *CAS9* expression cassette was transformed into both parent strains via homologous recombination replacing *GAL1*. For the allele swap, the allelic gene variants were amplified with the Phusion DNA polymerase according to the protocol (Thermo Fisher Scientific). An allele specific guide-RNA was designed and transformed into the parent cells where it induced the Cas9 mediated double-strand break within the variant sequence, while the corresponding allele of the other parent was provided as repair fragment. Positive sequence exchange was verified by allelic PCR, using a forward primer in the upstream region of the gene and one or more reverse primer on SNP positions within the gene. The obtained strains were phenotyped in triplicate (2.3) to validate the impact of the allelic variants on the trait.

3 Results and Discussion

3.1 Phenotyping of strains

First, the ability of both parental strains to metabolize SMM, the precursor of DMS, was compared. The content of pDMS in synthetic must before fermentation was determined at 282.2 ± 48.0 µg/L (as equivalent of DMS formation). The two parent strains consumed a major fraction of pDMS after 80% of the fermentation but differ significantly in their ability to reduce the precursor. While the residual concentration of pDMS was 73.04 ± 18.30 µg/L for strain MTF2621, almost no precursor was left by strain MTF2622 (6.54 ± 6.95 µg/L pDMS).

As DMS is mainly originating from amino acid related compounds, it was necessary to assess whether the consumption of pDMS is connected to the consumption of certain amino acids. This could indicate a common regulation or the involvement of an amino acid transporter in pDMS uptake. For both parent strains the uptake of nitrogen sources was compared to the decrease of pDMS in the medium during the first 28 hours of the fermentation (Figure 29).

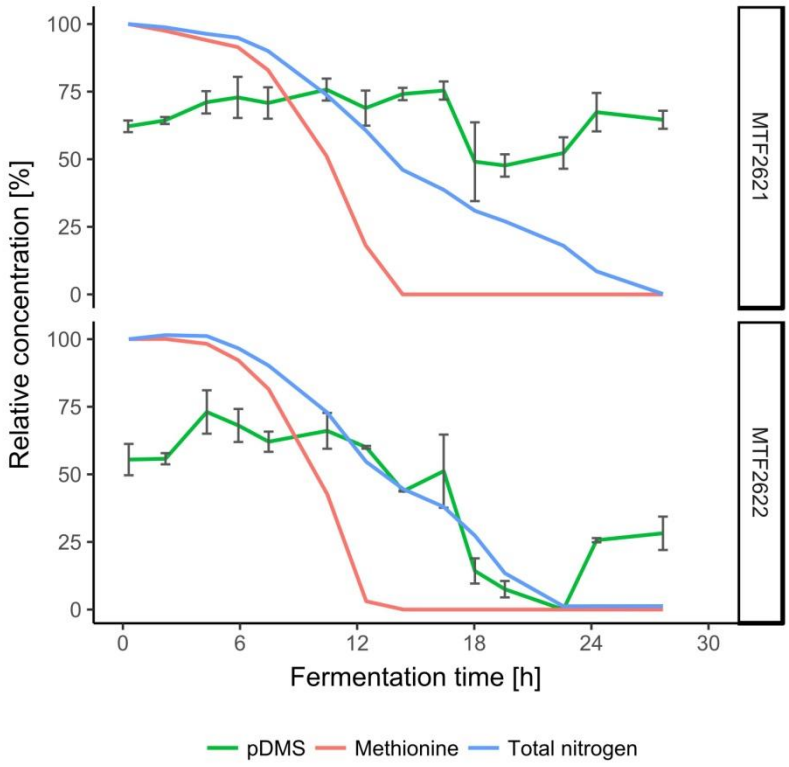


Figure 29: Determined concentrations of pDMS over the first 28 hours of the fermentation for the parental strains MTF2621 (top) and MTF2622 (bottom). Values for methionine and the amount of total available nitrogen are shown for comparison. Concentrations are given in relation to the starting concentration.

Interestingly, the parental strains show a different behavior in their consumption of nitrogen sources. Strain MTF2621 consumed all available nitrogen within 28 h, while MTF2621 achieved this within 23 h. A rapid decrease in the pDMS concentration could be observed for both strains in the beginning of the fermentation. In the first sample, taken after 20 min of the fermentation, the pDMS concentration was measured to be roughly 60% of the concentration determined in the initial must. A rapid decrease of pDMS after inoculation has been previously reported (Dagan and Schneider, 2012). For MTF2622 a second decrease of pDMS in the medium could be observed after 14 h. As the decrease occurs approximately with the depletion of methionine from the medium, it suggests that pDMS uptake is repressed by methionine. Strain MTF2622 had completely taken up all available pDMS after 23 h, which correlates to the depletion of all available nitrogen sources from the medium. However, the

concentration of pDMS in the medium subsequently increased again and reached almost 30% of the initial concentration at the end of the measurement. This indicates that pDMS was actively taken up and that it was very likely not metabolized completely and released to the medium at the end of the fermentation. For MTF2621 a minor decrease to 50% of the initial pDMS concentration could be observed between 16.0 h and 24.0 h of the fermentation. In this case the decrease of pDMS did not correlate with the depletion of methionine. In addition, the concentration of pDMS in the medium rose again before the depletion of all nitrogen sources and reached its initial level at the end of the follow-up. This suggests that strains MTF2622 and MTF2622 differ in their ability to take up pDMS.

3.2 Genome wide identification of QTLs influencing pDMS metabolism

In order to discover the genetic basis behind the different abilities to metabolize pDMS, we performed a QTL analysis. 130 F2-segregants from a cross between the two parent strains were phenotyped for their ability to consume pDMS. The distribution of pDMS content at the end of fermentation between the strains revealed two populations (Figure 30).

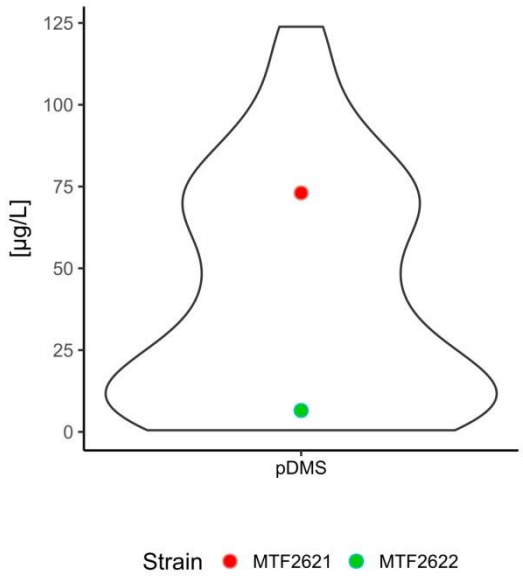


Figure 30: Distribution of measured pDMS concentrations after 80% of the fermentation (as equivalent of DMS concentration) among the population of segregants, together with the concentration mean and interquartile range (boxplot).

A large share of 45.4% of the segregants had metabolized 90-100% of the pDMS present in the medium while a smaller fraction of the strains had left approximately 70.0 µg/L pDMS. This matches the determined concentrations for the parental strains. For the segregant with the highest preservation of pDMS, a concentration of 123.9 µg/L could be measured, which still corresponds to a pDMS loss of 56.1% compared to the unfermented medium. These results indicate the major influence of one locus on the trait.

We used the segregant marker map that we obtained from our previous study (Eder et al., 2018) to perform a linkage analysis of the segregants' ability to metabolize pDMS. One major QTL on chromosome XII was detected (Figure 31), which matches the previous observation that the trait is mainly influences by one allele (Figure 30).

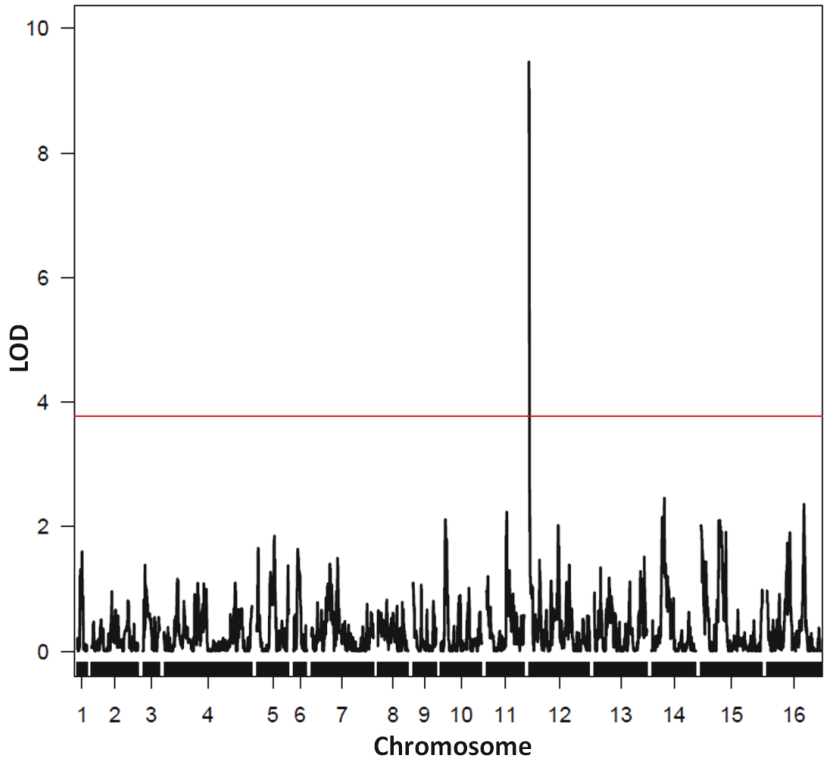


Figure 31: Detected QTL on chromosome XII with an influence on the metabolism of pDMS.

The detected QTL had an LOD-score of 9.46, meaning that 29.4% of the trait variation can be explained by the locus. The detected region chrXII:17,035..24,535 has a size of 7.5 kB and contains four genes: *MHT1, MMP1, GTT2, and YLL058W*. Interestingly, *MHT1* and *MMP1* are described to be involved in pDMS metabolism. *MHT1* is the only gene that does not contain non-synonymous SNPs between the parent strains. Two SNPs are located in the 1000-bp upstream region of the gene at positions -878 and -848, however, they are already located in *MMP1*.

The high impact of the detected region chrXII:17,035..24,535 could have masked the detection of additional QTLs with low influence on the trait. For this reason, QTL mapping was repeated with the marker chr12_20766 as covariate. No additional QTL was found, although a region on chromosome XVI around 656700 bp was close to being significant with an LOD of 3.59. The region was closer observed, but no promising target gene with a potential function for the trait was identified.

3.3 Variant of *S*-methylmethionine transporter Mmp1 causes pDMS preservation

By performing RHA we evaluated the allelic effects of the three candidate genes within the QTL region that contained non-synonymous SNPs, *MMP1*, *GTT2* and *YLL058W*. While no difference could be seen for the allelic variants of *GTT2*, the other two candidate genes showed significant difference between the alleles (Figure 32). A smaller effect can be seen for the allelic variants of *YLL058W*, the MTF2621 allele causes the hemizygote to leave 2.2 times more pDMS at the end of fermentation. The variants differ in six non-synonymous SNPs within the coding region (Table 24), of which three lie in the pyridoxal phosphate-dependent transferase domain of the protein. This domain is shared with several proteins involved in amino acid metabolic processes (*Saccharomyces* genome database). *YLL058W* encodes a protein of unknown function that shows similarity to the cystathionine gamma-synthase Str2 (Giaever et al., 2002). Str2 converts cysteine into cystathionine and is involved in the regulation of sulfur assimilation genes (Hansen and Johannesen, 2000).

The parental variants of *MMP1* showed to have a bigger impact on the metabolism of pDMS. The MTF2621 allele causes an increased pDMS preservation by factor 5.64 in comparison to the MTF2622 allele. The variants differ in one non-synonymous SNPs within the coding region (A536G), which causes the introduction of a STOP codon at position 179 of the 583 AA long polypeptide chain in strain MTF2622. This position is located within the amino acid permease sequence of the protein. Another SNP was found in the 1000bp upstream region of the gene although no recognized transcription factor binding site was affected according to the YEASTRACT database (Teixeira et al., 2013).

Table 24: Differences in peptide chains of validated gene variants caused by non-synonymous SNPs between parent strains. Comparison of SNP identity to *S. cerevisiae* type strain S288C.

Gene	AA position	S288C	MTF2621	MTF2622
MMP1	179	W	W	STOP
YLL058W	98	K	K	E
	107	T	T	R
	131	A	A	T
	207	G	G	D
	213	P	P	S
	252	S	S	L

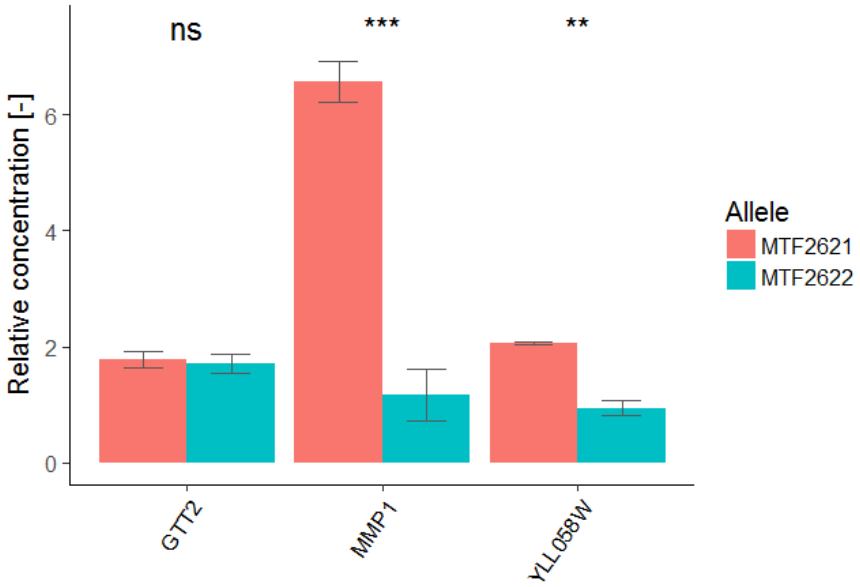


Figure 32: Allelic impact of the genes *GTT2*, *MMP1* and *YLL058W* on the concentration of preserved pDMS after 80% of the fermentation, assessed by RHA. Concentrations are given in relation to the undeleted parental heterozygote. p-value: ns > 0.05, ** ≤ 0.01, *** ≤ 0.001.

To assess whether the non-synonymous SNP in *MMP1* is responsible for the increased preservation of pDMS, we exchanged nucleotide 536 between the parent strains. This allelic swap resulted in strain MTF2621 *MMP1*(2622) with an eliminated STOP codon in *MMP1* and strain MTF2622 *MMP1*(2621) with an introduced STOP codon in *MMP1*. The elimination of the STOP codon completely inversed the phenotype of MTF2621 to MTF2622 (Figure 33). Strain MTF2621 *MMP1*(2622) had consumed the same share of pDMS after 80% of the fermentation as MTF2622 and was not anymore able to preserve SMM. A similar result could be seen for the opposite allelic swap: the introduction of the STOP codon within *MMP1* of MTF2622 almost completely inversed the phenotype of MTF2622 to MTF2621. However, the concentration of preserved pDMS by MTF2622 *MMP1*(2621) was slightly lower than for strain MTF2621, although the difference was not significant. This could be an indication that other alleles, like *YLL058W*, or environmental factors, like fermentation time, have an influence on the trait.

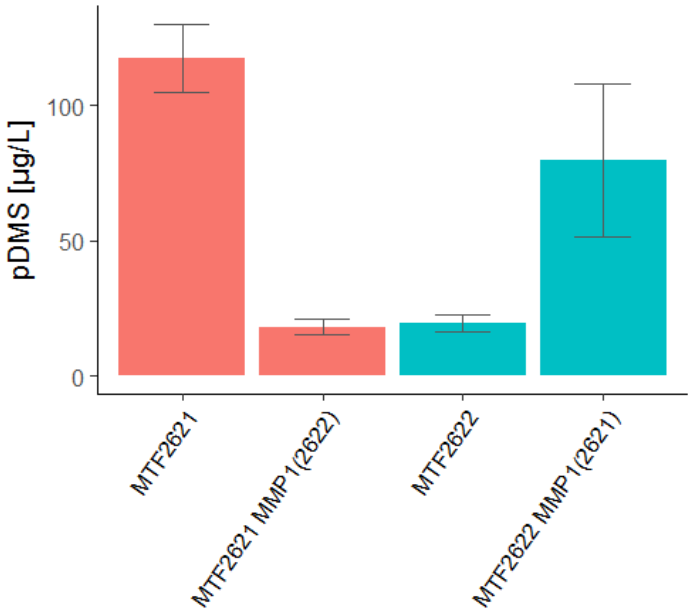


Figure 33: Effect on the concentration of residual pDMS after 80% of the fermentation **for the a**llelic swap of *MMP1* (nucleotide 536) between the parent cells.

During the fermentation, the concentration of pDMS decreased for all observed segregants, even for strains containing the truncated variant of Mmp1. This could indicate the existence of other transport mechanisms, either active through transporter proteins or passive through permeases. The MTF2622 strain showed marked uptake of pDMS at the beginning of the fermentation, although it contains the truncated Mmp1 transporter (Figure 29). For both strains a release of pDMS could be seen after an initial uptake, indicating diffusion of pDMS with the concentration gradient. Other possibilities for the observed precursor loss are yeast-independent influences, like degradation or evaporation.

To assess the occurrence of the truncated Mmp1 variant within *S. cerevisiae* strains of different origin, a blast of the nucleotide sequence was performed against 85 yeast genomes recently sequenced in our group (Coi et al., 2017; Legras et al., submitted) or available in the *Saccharomyces* genome database (SGD) (Table 25, Additional file 15).

Table 25: Distribution of the truncated SMM transporter Mmp1, resulting from SNP G536A in the gene, between 85 assessed *S. cerevisiae* strains of different origin.

Strain origin	Full lenght Mmp1	Truncated Mmp1
African	4	0
Beer	3	0
Bread	2	1
Cheese	3	3
Clinical	1	0
Flor	0	8
Laboratory	6	0
Oak Mediterranean	3	0
Oak North American	7	1
Others	8	0
Rum	8	2
Wine	19	4
Wine x flor	2	0
Total	66	19

SNP G536A, resulting in the STOP codon at amino acid position 179 within *MMP1*, was found in 19 out of 85 strains. All 8 assessed flor strains contain the truncated Mmp1 variant, whereas 83% of wine strains possess the full-length variant. Apart from flor strains, considerable occurrence of the truncated Mmp1 variant can only be seen in cheese strains, with three of the six assessed strains containing the variant.

The significant concentration of the truncated Mmp1 allele in flor strains indicates a relation of this mutation and the origin. To further evaluate this potential link, a phylogenetic analysis of all complete *MMP1* sequences (79) from the assessed strains was performed (Figure 34).

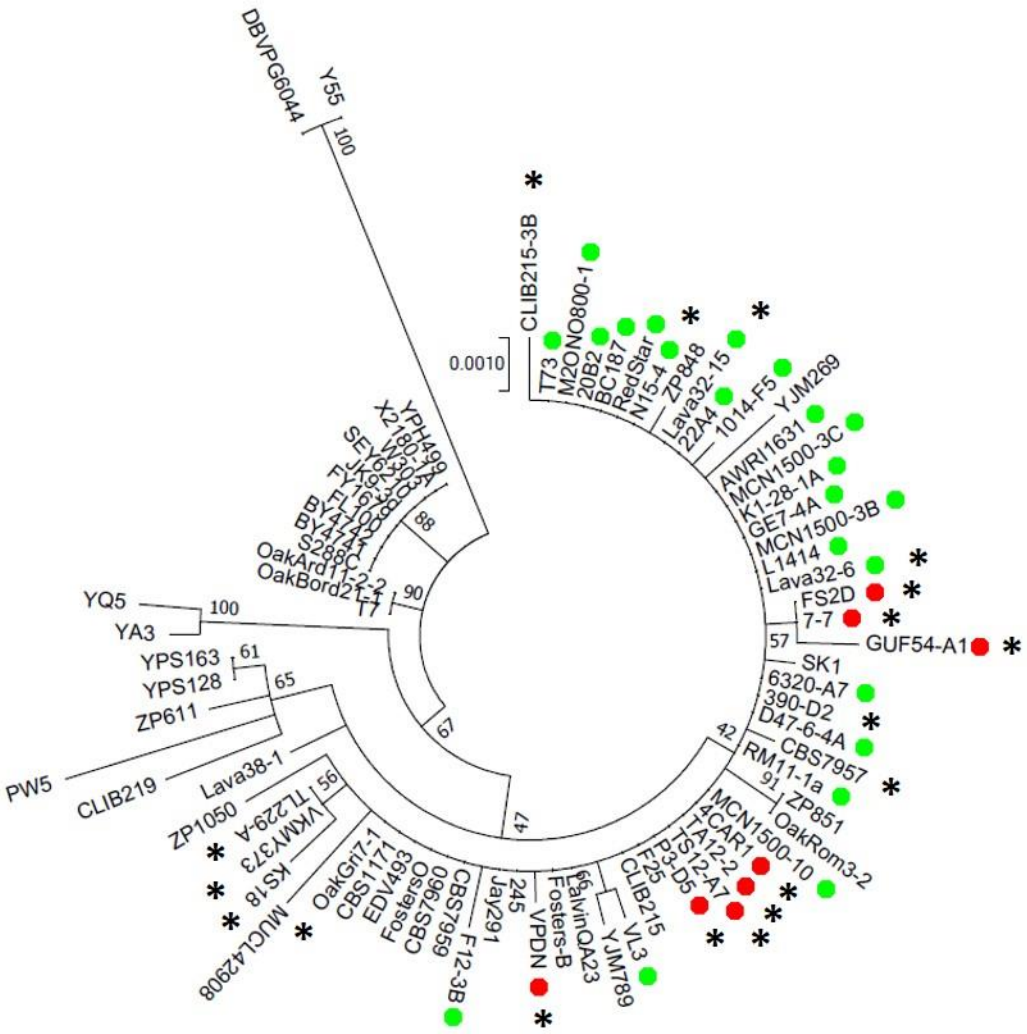


Figure 34: Phylogenetic analysis of *MMP1* alleles from 79 strains of different origin. Wine strains are marked green and flor strains are marked red. Strains with truncated Mmp1 transporter are indicated with asterisks. Maximum likelihood tree was constructed by bootstrap method with 200 replications using MEGA v7.0.26 (MEGA software).

Regarding the sequence of *MMP1,* no clustering of flor alleles can be seen in the pylogenetic analysis and, except for SNP G536A, wine and flor alleles share high global similarity. This and the fact that all assessed flor strains contain the truncated Mmp1 allele makes us presume that the SNP arose in flor strains. In this scenario, the occurrence of truncated Mmp1 alleles in wine strains (and others) could have been caused by gene flow from flor to wine yeast strains, an event recently described by Coi et al. (2017). The disproportionately high presence of the truncated Mmp1 variant in cheese strains could indicate human selection resulting from desired flavor characteristics of elevated DMS levels in cheese.

4 Conclusion

In this study we confirmed the potential of QTL analysis to decipher enological important traits of *Saccharomyces cerevisiae*, in particular the ability to metabolize grape derived aroma precursors during wine fermentation. By phenotyping and genotyping a set of 130 F2-segregants deriving from a cross between two strains isolated from wine, we were able to detect a genomic region with a strong influence on the metabolism of SMM, the main precursor for DMS.

Two genes in the QTL, *MMP1* and *YLL058W*, were found to influence SMM metabolism and the SMM transporter gene *MMP1* was identified to be responsible for the major part of trait variation. The MTF2621 allele of *MMP1* expresses a strongly decreased ability to take up SMM from the medium and by performing allelic swap of the gene between the parent strains of the study, we identified SNP 536 G -> A to be causative for this behavior. The nucleotide change leads to the introduction of a STOP codon at position 179 of the 583 AA long polypeptide chain, which results in the expression of a truncated SMM transporter Mmp1.

By comparing *MMP1* variants from 85 yeast strains of different origin, we could demonstrate that the truncated Mmp1 variant predominantly occurs in flor strains, therefore indicating a link to this origin. Rare findings of the described SNP in wine strains could be attributed to gene flow from flor to wine yeast strains.

These findings offer new perspectives for the management of pDMS in grape must. The introduction of the truncated allele of *MMP1* into industrial strains by non-GMO methods, like breeding via marker-assisted selection, will generate superior strains for enhancing the preservation of SMM during wine fermentation and will therefore lead to a more fruity and aromatic wine. As DMS is a flavor active volatile present in other fermented foods, like beer (Anness and Bamforth, 1982) and cheese (Carbonell et al., 2002), the discovered genetic variation may therefore be used to improve yeast starter cultures for a wide range of fermented products.

5 Supplementary information

Additional file 14: List of primers used in this study with nucleotide sequence (5' -> 3').

Primer name	Nucleotide sequence (5' -> 3')
1678_GAL1DisB	AATGAGAAGTTGTTCTGAACAAAGTAAAAAAAAGAAGTATACTTACATAGGCCACTAGTGGATCTGTATAG
3093_tagA-pUG	ACTATATGTGAAGGCATGGCTATGGCACGGCAGACATTCCGCCAGATCATCAATAGGCACCTTCGTACGCTGCAGGTCGA
4653_A-CYC1t-rv	GTGCCTATTGATGATCTGGCGGAATGTCTGCCGTGCCATAGCCATGCCTTCACATATAGTCCGCAAATTAAAGCCTTCGAG
5981_Cas9_GAL1_fw	TTCACCGGTCGCGTTCCTGAAACGCAGATGTGCCTCGCGCCGCACACCGTATTACCGCCTTTGAGTG
6005_p426-CRISP_rv	GATCATTTATCTTTCACTGCGGAGAAG
6209_Cas9ORF_rv	CTGCGCCGTGCTGTTCTTTTGAG
962_GAL1_fw	TACGGATTAGAAGCCGCCGAGC
CasGuide_MMP1_1	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGATAAATGATCCGATGGTGCCTAATAAGAGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAG
CasGuide_MMP1_2	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGATAAATGATCCGATGGTGCCCAATAAGAGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAG
CasRep_MMP1_fw	GGTAGGATGTTCGCTTACCATC
CasRep_MMP1_rv	CCACCACCTGCAATGAGAAC
del_GTT2_fw	AGAAACTTGGCGCTCTATATAAAGTACCTACAAAGGATACTTCGTACGCTGCAGGTCGAC
del_GTT2_rv	GTTTCCACGTACCACGCAAAACTTGTCTCAAGTAGCCACTGCATAGGCCACTAGTGGATCTG
del_MMP1_fw	TCAAGAGTTCTCCATCAGGATACCATGAAAGATACCCGAATTCGTACGCTGCAGGTCGAC
del_MMP1_rv	ATTATTATTAATTAATATTCAAATGAAGTACTTCACTAAGGCATAGGCCACTAGTGGATCTG
del_YLL058W_fw	AACCACCACTGAACAACTATTAATCATCGTGGAAATTATCTTCGTACGCTGCAGGTCGAC
del_YLL058W_rv	AAAAAATTCTTGATTTTTGATATTCTACTCATGATCTGCTGCATAGGCCACTAGTGGATCTG
Hygro_rv	TGTTATGCGGCCATTGTC
tal_MMP1_1	CCGATGGTGCCTAATAAGAG
tal_MMP1_2	CCGATGGTGCCCAATAAGAG
test_GTT2_fw	TCGATGGATGCCGTATCAC
test_MMP1_fw	CTACCACCTTTCGCTCTTG
test_YLL058W_fw	ACACCAGGATGACTAACC

Additional file 15: Occurrence of truncated SMM transporter Mmp1, caused by SNP G536A within the gene, in 85 strains of different origin. Genome sequences were obtained from data bases Evolya, Genowine and the *Saccharomyces* genome database (SGD).

Strain	Origin	Mmp1 variant	Source
DBVPG6044	african	full	SGD
PW5	african	full	SGD
SK1	african	full	SGD
Y55	african	full	SGD
YQ5	atypical bread	full	Evolya
YA3	atypical bread (sake like)	full	Evolya
NRRY1791	atypical cheese (sake like)	full	Evolya
CLIB219	atypical grape	full	Evolya
Lava38-1	atypical grape	full	Evolya

CBS1171	beer	full	Evolya
ForstersB	beer	full	SGD
ForstersO	beer	full	SGD
6464	bread	full	Evolya
CLIB215	bread	full	Evolya
CLIB215-3B	bread	truncated	Evolya
KS11	cheese	full	Evolya
KS18	cheese	truncated	Evolya
MUCL42908	cheese	full	Evolya
NRRLY1545	cheese	full	Evolya
TL229-A	cheese	truncated	Evolya
VKMY373	cheese	truncated	Evolya
YJM789	clinical	full	SGD
7-7	flor	truncated	Genowine
F25	flor	truncated	Genowine
FS2D	flor	truncated	Genowine
GUF54-A1	flor	truncated	Genowine
P3-D5	flor	truncated	Genowine
TA12-2	flor	truncated	Genowine
TS12-A7	flor	truncated	Genowine
VPDN	flor	truncated	Evolya
BY4741	laboratory	full	SGD
BY4742	laboratory	full	SGD
FL100	laboratory	full	SGD
FY1679	laboratory	full	SGD
JK9-3D	laboratory	full	SGD
SEY6210	laboratory	full	SGD
W303	laboratory	full	SGD
X2180-1A	laboratory	full	SGD
YPH499	laboratory	full	SGD
OakRom3-2	oak MED	full	Evolya
ZP848	oak MED	full	Evolya
ZP851	oak MED	full	Evolya
OakArd11-2-2	oak NA	full	Evolya
OakBord21-1	oak NA	full	Evolya
OakGri7-1	oak NA	truncated	Evolya
T7	oak NA	full	SGD
YPS128	oak NA	full	SGD
YPS163	oak NA	full	SGD
ZP1050	oak NA	full	Evolya
ZP611	oak NA	full	Evolya
245	rum	full	Evolya
309	rum	full	Evolya
376	rum	full	Evolya
390-D2	rum	truncated	Evolya
460	rum	full	Evolya

CBS7957	rum	truncated	Evolya
CBS7959	rum	full	Evolya
CBS7960	rum	full	SGD
EDV493	rum	full	Evolya
JAY291	rum	full	SGD
1014-F5	wine	full	Genowine
20B2	wine	truncated	Genowine
22A4	wine	full	Genowine
6320-A7	wine	full	Genowine
AWRI1631	wine	full	SGD
AWRI796	wine	full	SGD
BC187	wine	full	SGD
D47-6-4A	wine	full	Genowine
F12-3B	wine	full	Genowine
GE7-4A	wine	full	Genowine
K1-28-1A	wine	full	Genowine
L1414	wine	full	Genowine
Lava32-15	wine	truncated	Evolya
Lava32-6	wine	truncated	Evolya
M2ONO800-1	wine	full	Evolya
MCN1500-10	wine	full	Evolya
MCN1500-3B	wine	full	Evolya
MCN1500-3C	wine	full	Evolya
N15-4	wine	truncated	Evolya
REDSTAR	wine	full	SGD
RM11-1a	wine	full	SGD
VL3	wine	full	SGD
YJM269	wine	full	SGD
QA23	wine x flor	full	SGD
Vin13	wine x flor	full	SGD

<u>Chapter 4:</u> Preliminary results about the genetic bases of terpenol transformation by *S. cerevisiae* during alcoholic fermentation

This chapter describes additional results of this thesis that were obtained during experimental work on terpenol transformation during wine fermentation, but that could not be completed for publishing as a research article.

Terpenols represent a group of compounds that are important for the bouquet of certain wine styles. They are produced in grape vine and are introduced to must by the grapes. However, terpenols are transformed during yeast alcoholic fermentation and a strain dependency of this conversion was observed.

The population of 130 F2-segregants described in chapter 1 was phenotyped for the strains' ability to metabolize the terpenol precursor geraniol, which was added to the synthetic must. Samples were taken near the end of fermentation and terpenol content was determined by SPME-GC-MS. The results were used as phenotype data for QTL mapping, using the marker map that was obtained during the first part of the project.

This approach was successful and allowed to identify 21 QTLs with an influence on terpenol composition as well as several potential target genes within the QTL regions. The role of these allelic variants remains to be confirmed.

1 Introduction

Terpenols are volatile compounds that are produced by algae and higher plants from the common precursor geranyl pyrophosphate (reviewed by Jüttner, 1995, and by Nagegowda, 2010, respectively). Terpenols such as geraniol, nerol, linalool, citronellol, α-terpineol and cis-rose oxide are potent aroma compounds with a rose-flowery or citrusy scent. Geraniol can also be produced by some *S. cerevisiae* strains in generally very low amounts (Carrau et al., 2005), although some mutant strains have gained the ability to produce significantly elevated levels of geraniol and linalool (Javelot et al., 1991).

Muscat and Gewürztraminer wines contain up to 1.5 and 4 mg/L of linalool and geraniol respectively, in contrast to other wines where the contents are 10 to 20 fold lower (reviewed by Black et al., 2015; Ribéreau-Gayon et al., 1975; Simpson, 1979; Steyer et al., 2013). In addition, these grape varieties present a characteristic proportion between linalol and geraniol. While high concentrations of free linalool are found in muscat grapes, high amounts of glycosidically bound geraniol are present in gewürztraminer grapes (Duchêne et al., 2009; Ribéreau-Gayon et al., 1975).

Although the vast majority of terpenols found in wine are produced by grape vine, accumulated in the grape berry during its development and introduced to must by the grapes, they are modified during wine fermentation by *S. cerevisiae* and other yeasts (Gamero et al., 2011; King and Dickinson, 2000). It could be demonstrated that the shape of these conversions is strain dependent, indicating an impact of yeast metabolic properties (Furdíková et al., 2014). Some steps in the transformation of terpenols are mainly driven by simple chemical reactions such isomerization reactions at acidic pH (Simpson and Miller, 1983). This is the case for the conversion of geraniol to linalool (Zea et al., 1995), although yeast might influence the conversion rate directly through additional enzymatic catalyzation or indirectly through affecting environmental conditions such as pH. From linalool, α-terpineol can be formed during fermentation (King and Dickinson, 2000). The reduction of geraniol to citronellol was already elucidated to be enzymatically catalyzed and was reported to be a three step reaction involving two alcohol dehydrogenases and the oxidoreductase Oye2 (Steyer et al., 2013). Some of the citronellol produced during fermentation is the basis for the synthesis of the highly potent aroma contributor *cis*-rose oxide (Koslitz et al., 2008). Different yeast strains present varying abilities to produce this compound (Furdíková et al., 2014). In addition, terpenols are primary alcohols and can therefore be acetylated like other aliphatic alcohols such as propanol or hexanol. Atf1 has been shown to be the main acetyltransferase involved in terpenol acetylation (Steyer et al., 2013). Several studies report a loss in total terpenol concentrations during alcoholic fermentation (King and Dickinson, 2000; Soares et al., 2015; Vaudano et al., 2004). Possible explanations include evaporation or the adsorption to yeast cell wall structures (Bishop et al., 1998). As the precursor of geraniol is geranyl pyrophosphate, an intermediate of the sterol pathway, this also led to the assumption that yeast can incorporate geraniol into sterol synthesis, especially during alcoholic fermentation under

oxygen limitation (Vaudano et al., 2004). However, this hypothesis was never confirmed and sterol auxotrophic mutants with limitations in pathway steps before Erg20 were unable to use geraniol for growth. These aspects were not assessed in this study.

The conversion from geraniol to other terpenols during fermentation is presented in Figure 35.

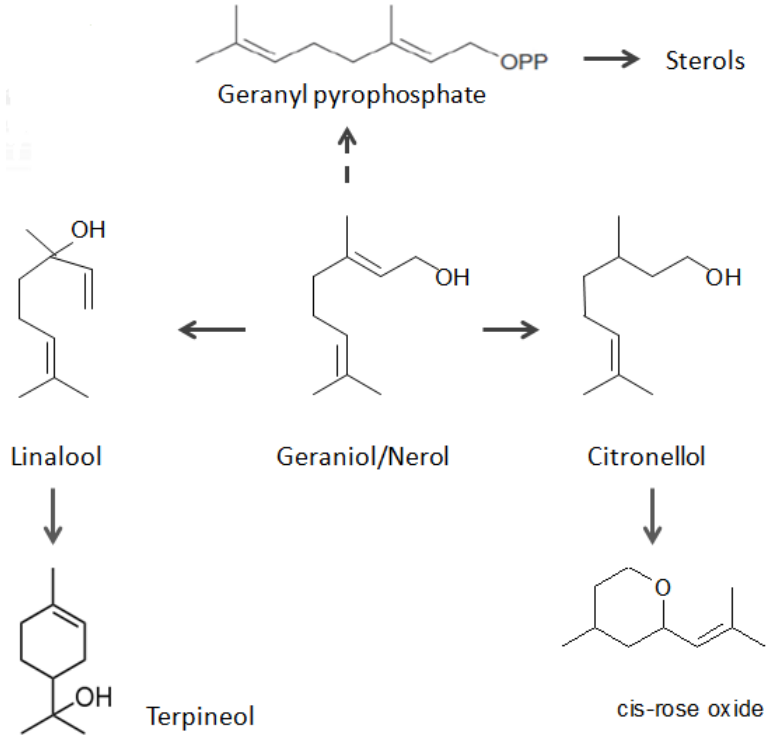


Figure 35: Schematic representation of the formation of terpenols during wine fermentation and the potential connection to yeast sterol pathway (based on King and Dickinson, 2000; Zea et al., 1995).

The enzymatic bases and principles of several reactions of terpenol transformation are still not yet understood. This includes the conversion of linalool to α-terpineol or of citronellol to *cis*-rose oxide. In a similar manner, no clear explanation has been given to the loss of overall terpenols during alcoholic fermentation. As terpenol compounds have a high impact on wine aroma, we have evaluated potential differences in the production of terpenols among a population of segregant strains to gain insights into the metabolism of these compounds during alcoholic fermentation. For this purpose, we have used the same experimental design as described above in order to correlate differences in terpenol formation to genetic variation among strains.

2 Materials and Methods

2.1 Strain selection, segregant generation and genotyping

This experiment is associated to the previous experiment described in chapter 1. The evaluated strains consisted in the haploid *S. cerevisiae* strains MTF2621 (4CAR1 [Δ*HO*::*Neo*^{r}]) and MTF2622 (T73 [Δ*HO*::*Nat*^{r}]). From a cross between these parents, a population of 130 segregants of the F2 generation was obtained. The strains had been genotyped by whole genome sequencing in order to generate a marker map for QTL mapping (Chapter 1).

2.2 Phenotyping of segregant population

The segregant strains were phenotyped during wine fermentation in 280 mL of synthetic grape must spiked with 2.0 mg/L geraniol. Samples were taken after 80% of fermentation, which was determined by following the weight loss caused by the production of CO_2.

For the analysis of terpenol composition, the samples were extracted by solid phase micro extraction (SPME) and measured via GC-MS (García et al., 1996). In a 20-mL glass flacon, 2.3 g of sodium chloride were dissolved in 7 mL of MilliQ water and the solution was kept on ice. 30 µL of internal deuterated terpenol standard and 1 mL of sample were added, the tubes were closed and mixed well. A PDMS/DVB/carboxen fiber was inserted in the headspace of the vessel while agitating and incubating the sample for 30 min at 30 °C. Subsequently, the fiber was inserted in the GC-MS and terpenol content was determined.

2.3 QTL mapping

The phenotyping and genotyping data were used to identify QTLs in the yeast genome that influence terpenol transformation during wine fermentation. The statistical analyses were carried out using the programming language R v3.2.3 (www.r-project.org) with the R/qtl v1.40-8 and R/eqtl v1.1-7 libraries (Broman et al., 2003). QTL mapping was performed with two different phenotype models, the normal model using Haley-Knott regression and a non-parametric analysis, resulting in logarithm of odds (LOD) scores for each marker and pseudo-marker every 2.5 cM (interval mapping method). An interval estimate of the location of each QTL was obtained as the 1-LOD support interval, the region in which the LOD score is within 1 unit of the peak LOD score. If the same locus was detected with both models, the results with the higher LOD score were selected.

3 Results and Discussion

In order to obtain precise quantitative data on the conversion of geraniol to other terpenols for the QTL mapping, we first considered potential biases related to different fermentative behaviors of the segregants. Indeed, as higher fermentation rates lead to a more vigorous production and release of CO_2, one possible cause for terpenol loss during fermentation could be the stripping of compounds from the medium by CO_2. To evaluate the significance of this factor and its potential impact on QTL mapping, the influence of different CO_2 production rates was evaluated for two segregant strains with different fermentation kinetics in synthetic must containing 10 mg/L of geraniol (Table 26).

Table 26: Difference of terpenol formation for two segregant strains with extreme fermentation kinetics. Samples were taken after 80% of fermentation, which corresponded to 73h for the fast fermenting strain and to 261h for the slow fermenting strain.

	Concentration [µmol/L]		
Compound	**Must**	**Slow fermenter**	**Fast fermenter**
***cis*-rose oxide**	0.13	0.12 ± 0.0	0.12 ± 0.01
Citronellol	0.33	10.58 ± 0.43	8.38 ± 0.50
Citronellyl acetate	0.07	1.38 ± 0.31	2.97 ± 0.12
Geraniol	61.55	7.94 ± 0.78	5.58 ± 0.14
Geranyl acetate	0.03	0.61 ± 0.07	1.69 ± 0.06
Linalool	0.32	0.86 ± 0.06	0.66 ± 0.04
Nerol	0.51	0.70 ± 0.01	0.68 ± 0.05
Neryl acetate	0.02	0.04 ± 0.01	0.09 ± 0.01
α-terpineol	0.00	0.01 ± 0.0	0.01 ± 0.0
Total	63.07	22.24 ± 1.68	20.18 ± 0.32
	100%	35.27 ± 2.67%	32.00 ± 0.50%

A decrease in geraniol concentration from 10.0 mg/L at the beginning to less than 2.0 mg/L at the end of fermentation can be observed for both strains. Total terpenol concentrations also declined during fermentation. This could either result from a loss of terpenols through evaporation, adsorption to cellular structures or conversion of terpenols into compounds not recorded by the phenotyping. Although differences exist between the strains regarding terpenol composition after fermentation, no significant difference in total terpenol content can be observed. This is in agreement with former observations, which reported low stripping of geraniol during fermentation (Ferreira et al., 1996; Steyer et al., 2013). This indicates that fermentation kinetics should not affect the search for QTLs influencing terpenol metabolism.

Subsequently, the phenotyping of our 130 F2-segregant population was carried out for the content of terpenols after 80% of fermentation. PCA was performed to show the correlations between dfferent terpenols compounds (Figure 36). The first two dimensions of the PCA explain 62.1% of variation between the fermentations.

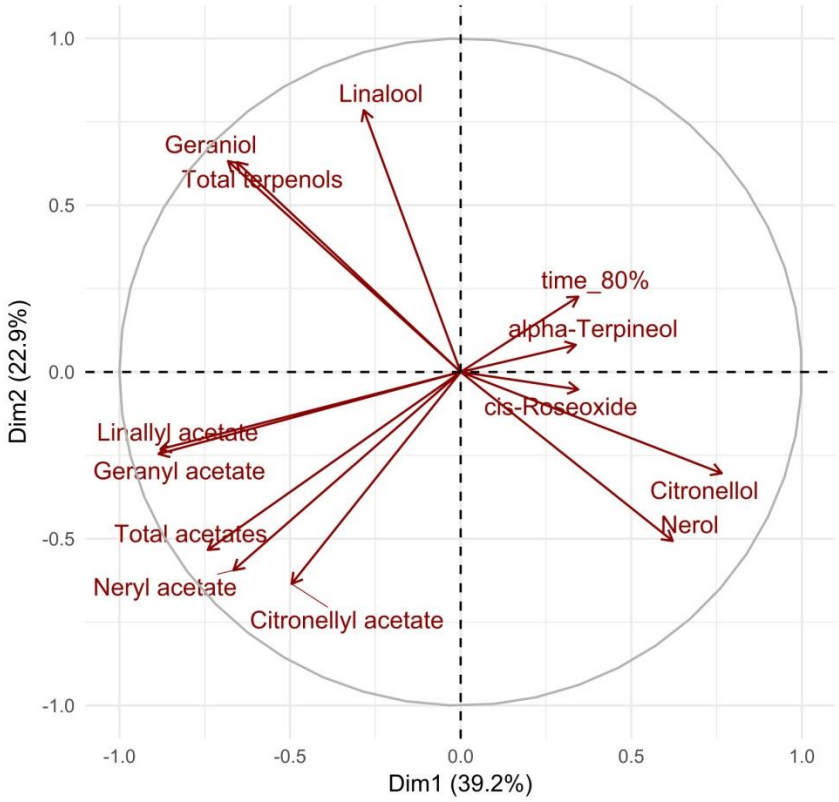


Figure 36: PCA of fermentation time (80%) and determined terpenol concentrations at this time point.

Terpenol acetates are correlated with each other while they are negatively correlated to fermentation time. However, the fermentation time is only little explained by the PCA, as is the formation of *cis*-rose oxide and α-terpineol. The concentrations of geraniol and linalool are correlated. This can be explained by the chemical conversion of geraniol to linalool (Zea et al., 1995), which decreases with diminished geraniol concentrations. Geraniol is negatively correlated to nerol and citronellol formation. This was expected as citronellol was already shown to be synthesized from geraniol by enzymatic reactions (Steyer et al., 2013). Nerol is the isomer of geraniol and the negative correlation between both suggests an influence of yeast on the isomerization reaction. Finally, the concentration of *cis*-rose oxide is correlated to the citronellol content, which is in agreement with previous reports (Furdíková et al., 2014; Steyer et al., 2013).

The phenotypic data set was then used to detect QTLs in the yeast genome that influence terpenol metabolism (Table 27). This approach revealed as many as 21 QTLs distributed over seven chromosomes that influence 8 traits.

Table 27: Detected QTLs that influence the metabolism of terpenols by *S. cerevisiae*.

Trait	QTL name	Chromosome	QTL start [bp]	QTL end [bp]	LOD-Score
Linalool	chr2@172.5	II	493459	527387	3.97
α-Terpineol	chr4@141.1	IV	410742	426649	4.00
α-Terpineol	chr4@160.0	IV	448242	512368	3.62
Terpenol acetates	chr4@379.2	IV	1132305	1141711	4.19
Linallyl acetate	chr7@16.8	VII	32927	56448	4.13
Neryl acetate	chr7@16.8	VII	47026	56448	5.36
Linalool	chr7@198.0	VII	589048	608880	4.75
Neryl acetate	chr7@294.6	VII	867733	896533	4.81
Terpenol acetates	chr7@340.5	VII	1013880	1025828	6.63
α-Terpineol	chr8@134.5	VIII	386376	412965	3.55
Linallyl acetate	chr13@7.9	XIII	20503	25723	3.50
Linalool	chr13@7.9	XIII	20503	30565	4.15
Geraniol	chr13@53.3	XIII	130890	190371	3.79
Total terpenols	chr13@69.8	XIII	193006	217967	3.93
Geraniol	chr14@36.3	XIV	78386	119900	3.76
Total terpenols	chr14@36.3	XIV	78386	119900	4.65
Geraniol	chr14@65.6	XIV	194351	201973	3.98
Linalool	chr14@65.6	XIV	194351	201973	5.54
Total terpenols	chr14@78.6	XIV	204725	247403	4.22
Citronellyl acetate	chr14@102.4	XIV	290197	314885	4.30
α-Terpineol	chr15@27.9	XV	67745	93027	3.96
Citronellyl acetate	chr15@299.1	XV	873008	904240	3.73
Terpenol acetates	chr15@336.2	XV	1004482	1011768	3.77
Terpenol acetates	chr16@286.3	XVI	840264	876279	3.85
Terpenol acetates	chr16@299.9	XVI	896990	904961	4.64
Linalool	chr16@300.7	XVI	896990	917224	4.01

Traits influenced by these detected QTLs include the total amount of terpenols and terpenol acetates as well as concentrations of single compounds. The fact that several loci could be detected for the formation of linalool and α-terpineol indicates an influence of yeast on the corresponding reactions. However, no QTLs with an impact on citronellol or *cis*-rose oxide concentrations were found. The highest determined logarithm of odds (LOD)-score is 6.63 for QTL chr7@340.5 influencing terpenol acetate formation, meaning that up to 21 % of trait variation can be explained by the locus. Most QTLs (10) were detected to influence the formation of terpenol acetates. Two mechanisms may explain differences in acetate ester production, namely a different amount of terpenol substrates or a different general acetylation activity of the metabolism. As most QTLs for terpenol acetate formation showed no influence on single terpenol concentrations, it is therefore likely that these loci impact the acetylation reaction.

The detected regions were compared to four QTLs reported by Steyer et al. (2012) to influence the formation of citronellol and *cis*-rose oxide, but no overlap was found for any of these regions. This could be due to the different origin of strains. Steyer et al. (2012) used a cross

from one *S. cerevisiae* wine and one lab strain to map QTLs for terpenol metabolism, therefore, the more divergent genomic background could have led to the discovery of different regions.

The QTLs detected in this study were furthermore compared with regions that influence the formation of main and other volatile metabolites, as well as estimated metabolic fluxes and fermentation parameters that were detected during the first parts of this project (Table 28).

Table 28: QTLs for terpenol metabolism overlapping with regions detected by previous experiments to influence other traits.

QTL name	Influenced terpenol traits	Influenced other traits
chr2@172.5	linalool	ethyl butanoate
		ethyl hexanoate
		pyruvate yield
chr4@141.9	α-terpineol	ethyl hexanoate
		higher alcohol acetates (2x)
chr4@161.9	α-terpineol	higher alcohol acetates (4x)
chr4@379.2	terpenol acetates	CO_2 production rate t_{80}
		ratio glucose/fructose
chr7@26.5	terpenol acetates	pyruvate yield
		glycolysis and ethanol synthesis
		biomass
		TCA cycle oxidative
		TCA cycle metabolites
		ethyl esters (2x)
chr7@198.0	linalool	decanoic acid
		dodecanoic acid
chr7@294.6	neryl acetate	2-phenylethyl acetate
chr13@7.9	linalool	CO_2 production rate t_{80}
	linallyl acetate	pyruvate yield
		t_{80}
chr14@36.3	geraniol	dodecanoic acid
	total terpenols	propanoic acid
		propanol
		propyl acetate
		valeric acid
chr14@65.6	geraniol	propanoic acid
	linalool	
chr14@78.6	total terpenols	valeric acid
chr15@299.1	citronellyl acetate	ethyl esters (2x)
chr16@299.9	linalool	ethyl esters (4x)
	terpenol acetates	

Out of the 21 QTLs detected for terpenol related compounds, 13 regions had been found in previous chapters of this thesis to influence other traits such as the formation of medium chain fatty acids, fatty acid ethyl esters or propanol and related compounds. In two cases, QTLs affecting the formation of terpenol acetates also influence the CO_2 production rate after 80% of fermentation, a measure of fermentation kinetic. This confirms deductions drawn by the

PCA, namely a negative correlation of terpenol acetate formation and fermentation time (Figure 36).

In five QTLs detected to influence both terpenol metabolism and fermentative aroma formation, target genes were previously validated for their influence on fermentative aroma formation (chapter 2). In QTL chr2@172.5, membrane sensor protein Agp2 was validated influencing several medium chain fatty acids and their ethyl esters. In QTL chr4@141.9, histone deacetylase Sir2 was identified to also have an influence on the formation of fatty acids and fatty acid ethyl esters. Hexose transporter Hxt3 was validated in QTL chr4@379.2 to influence CO_2 production rate and sugar metabolism. In QTL chr7@26.5, Vid30, which is involved in vacuolar import and protein ubiquitylation, has been shown to impact metabolic fluxes. In QTL chr14@36.3, amino acid transporter Alp1 was identified to influence the production of propanol, dodecanoic acid and valeric acid. However, it is necessary to assess whether the validated genes are also responsible for variations of terpenol content in wine.

Candidate genes with a potential role in terpenol metabolism were also searched for in remaining QTLs with LOD-scores higher than 4.42. This corresponds to an explained variation of at least 15% by these loci (Table 29).

Table 29: Suggestions for target genes to validate in detected QTLs with LOD-scores of more than 4.42.

QTL name	Trait	Taget gene	Function
chr7@16.8	linallyl acetate neryl acetate	*HAP2*	Subunit of DNA-binding complex involved in activation and repression of respiratory gene expression
chr7@198.0	linalool	*no clear candidate*	
chr7@294.6	neryl acetate	*PDX1*	Part of pyruvate dehydrogenase complex
		TDH3	Glyceraldehyde-3-phosphate dehydrogenase
chr7@340.5	terpenol acetates	*SAY1*	Sterol deacetylase
chr14@36.3	geraniol total terpenols	*ERG24*	C-14 sterol reductase
chr14@65.6	geraniol linalool	*ZWF1*	Glucose-6-phosphate dehydrogenase
chr16@299.9	linalool	*GDB1*	Glycogen debranching enzyme
	terpenol acetates	*HDA3*	Subunit of histone deacetylase

In 6 of the evaluated loci, 8 genes with biological functions potentially involved in terpenol transformation could be detected. This includes enzymes with roles in yeast central carbon metabolism (CCM), as the production of terpenol acetates relies on the availability of acetyl-CoA, an intermediate of CCM. In the same manner, deacetylases could have an influence on terpenol acetate formation. Furthermore, dehydrogenases were chosen as candidate genes since the reaction from geraniol to citronellol was already proposed to be partly dependent on alcohol dehydrogenases (Steyer et al., 2013). In addition, the chosen candidate genes contain non-synonymous SNPs (*GDB1*, *HAP2, PDX1, SAY1, ZWF1*) or SNPs in the upstream

region (*ERG24, HDA3, TDH3*) between the parental variants, making these alleles potential targets for evaluation.

4 Conclusion

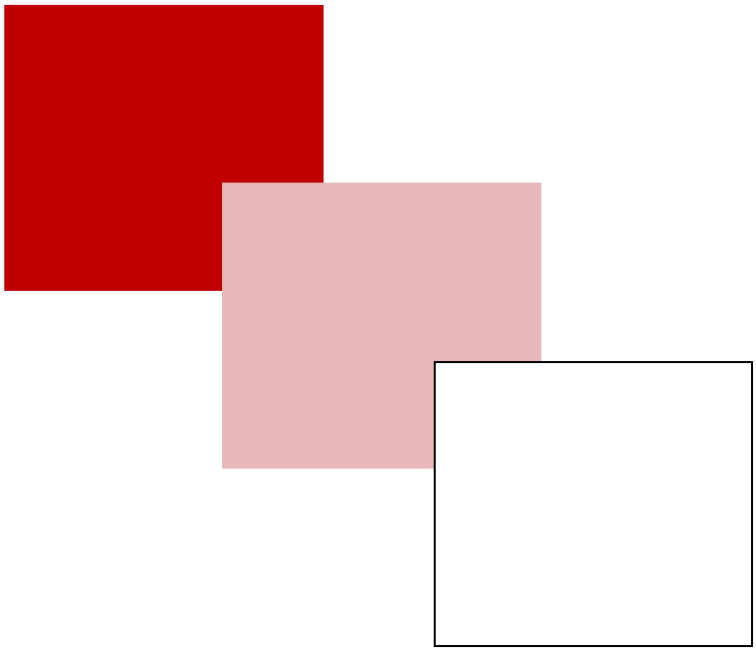
In this chapter, we used the QTL mapping approach established in the first part of the thesis to evaluate the ability of *S. cerevisiae* to transform terpenol compounds during alcoholic fermentation. We detected 21 QTLs affecting the total content of terpenol or concentrations of particular compounds. This indicates that yeast genomic properties have an influence on the shaping of wine terpenol profile. The QTLs could be separated in two groups, loci affecting the concentrations of terpenol compounds and loci with an influence on various terpenol acetates. We conclude that most QTLs found to influence terpenol acetates rather affect yeast's general acetylation capability since those regions did not influence the concentrations of primary terpenols.

Several QTLs were detected with an impact on single terpenols, particularly on linalool and α-terpineol. This demonstrates that the formation of these compounds is influenced by yeast. Although no QTL was detected to affect nerol concentration, a determined negative correlation between the isomers geraniol and nerol indicates an influence of yeast on the isomerization step.

Most of the detected QTLs were already found in previous experiments to influence the production of fermentative aromas or metabolic fluxes (chapters 1 and 2). Thus we can conclude that terpenol transformation during wine fermentation is very likely connected to yeast metabolism and as consequence to fermentative aroma formation. In some of these regions, target genes were already demonstrated to account for the influence on fermentative aromas or metabolic flux distribution. It has to be evaluated if these genes are also responsible for terpenol transformation and which mechanism is behind this process.

In several newly detected QTLs we identified target genes that have a probable function in terpenol transformation and contain non-synonymous SNPs or SNPs in the upstream region between the parent strains. The evaluation of these candidates using reciprocal hemizygosity analysis or allele swap is still pending. Assessing these genes will lead to a better understanding of the genomic bases and metabolic processes behind grape terpenol transformation by yeast during wine fermentation. The identification of allelic variants that account for differences in terpenol composition provides potential to shape the aromatic profile of yeast starter cultures for wine production. As hop-derived terpenols are also important contributors to beer flavor and biotransformation of terpenols was shown during ale and lager beer fermentation (King and Dickinson, 2003; Takoi et al., 2010a, 2010b), this knowledge can furthermore be applied to improve commercially used beer yeast strains.

Concluding Remarks



Wine market is a big and global business, driven partially by consumers' expectations for wine flavor that is susceptible to change. The actual market attitude is a general priority for fruity wines and the background of health concerns should make light and alcohol reduced wines more popular. This is in conflict with a general increase in grape ripeness that is caused by global climate change and leads to elevated alcohol contents in finished wine. As vines can only be grown in limited latitudes, wine producing regions are generally also wine exporters and a competition between these countries for market share exists. This brings wine producers to constantly strive for improving their product in order to meet the customers' expectations in the context of a changing environment.

Besides viticultural and winemaking practices, yeast has significant impact on wine flavor. During fermentation, it changes and deepens the aromatic profile by producing volatile compounds from nutrients in the must, by transforming aroma compounds that are contributed by the grapes or by releasing potent aroma molecules from odor-less, grape-derived precursor molecules. Among the most abundant aroma compounds *de novo* produced by yeast are higher alcohols, fusel acids, medium chain fatty acids and their corresponding ethyl and acetate esters. Depending on compound and concentration, these molecules can provide pleasant flavors, such as fruity or flowery, and unpleasant odors, such as rancid or solvent. Their production is closely linked to yeast metabolic properties and is driven by both, the carbon and nitrogen metabolism. Higher alcohols and fusel acids can be derived from the degradation of absorbed amino acids but also from intermediates of the CCM. The production of medium chain fatty acids relies on acetyl-CoA that emerges from the CCM. Acetate and ethyl esters of all listed compounds are formed enzymatically by acetylation or acylation. Yeast also affects grape derived flavors, either by metabolism of odor-less precursors to aromas or by alteration of grape aroma compounds. Terpenols, for example, are synthesized and contributed by grapes. Depending on the compound, they provide flavors ranging from piney to citrusy and flowery. Yeast enzymes were identified that catalyze conversions between different terpenols during fermentation. Furthermore, the influence of yeast on wine composition has an effect on the formation of post-fermentative aromas during wine aging. The most notable example is DMS, which increases general fruitiness and can impart notes of truffle and olives. The precursors of DMS are present in grapes and yeast can metabolize them to DMS, which is, however, driven off by the CO_2 produced during fermentation. Only precursor molecules left in young wine are then transformed chemically during maturation.

As the shape of described aroma compounds relies on yeast metabolic properties, it is therefore governed by the underlying genomic bases. Human use and selection of yeast for winemaking has led to a wide yeast population containing high genomic variation. Understanding the links between genomic and metabolic properties, but also the impact of genomic variation on metabolic traits is a requirement for the exploitation of yeast diversity. In the end, this has the potential to improve wine aroma and quality and to adapt to consumers' demands.

The genomic basis of complex traits, meaning traits that are driven by several genes and are therefore quantitative, can be fathomed by QTL mapping. Different strategies were developed, which have in common that the presence of defined genetic markers, i.e., genotype differences, is connected to phenotypic variations. Starting from plant breeding, this approach was adapted to *S. cerevisiae* and has become a powerful tool for enological studies with numerous analyses published. QTL mapping is continuously further improved to enhance the outcomes of the analyses and to expand the methodology to new traits of interest.

Referring to the presented background, future research on wine strain development will pursue the following key points: (i) more profound determination and assessment of wine sensory characteristics; (ii) evaluation of consumers' preferences for aroma and flavor to predict wine choice behavior in key markets; (iii) development and application of methodologies to identify genomic and metabolic factors that influence both, the *de novo* production of volatile compounds and the release or transformation of grape flavor precursors by yeast; (iv) population analyses and characterization of superior alleles in terms of aroma formation and (v) improvement of existing commercially used wine strains by non-GMO methods to further shape and improve their aromatic properties in order to meet consumers' demands.

The research outlined in this thesis contributes to points (iii) and (iv). QTL mapping was applied to detect regions in the genome of *S. cerevisiae* that influence the production of fermentative aromas, the alteration of grape aroma compounds and the potential for post-fermentative aroma formation during wine maturation. The use of a comparatively high number of segregants with increased recombination rate led to an elevated analytic strength of the analysis. As result, several QTL mapping strategies could be applied, including the search for interacting QTLs using multiple QTL mapping. The outcome was the detection of over 80 QTLs, including loci with minor effects, and the identification of interacting QTLs for several traits. By validating candidate genes within found QTLs, the involvement and impact of several allelic variants of these genes was demonstrated. The connection of some genes to the observed traits has already been described before (*AGP1*, *ALP1*, *ILV6*, *LEU9*), but some genes could be newly connected to influenced traits (*AGP2*, *IXR1*, *MAE1*, *NRG1*, *RGS2*, *RGT1*, *SIR2*). In all cases, allelic variants with different impact were described.

Variants of the amino acid transporters Agp1 and Alp1 affect the formation of Ehrlich pathway products through proposed differences in expression or affinity. The involvement of nitrogen metabolism in the formation of fermentative aroma was further demonstrated by the identification of two other enzymes involved in amino acid synthesis pathways, Ilv6 and Leu9. The fact that fermentative aroma production is also linked to the CCM was underlined with the validation of the target genes *MAE1* and *FAS1*. Mae1 catalyzes the reaction of malate to pyruvate and a higher conversion rate of this enzyme was proposed to increase the production of volatiles derived from pyruvate. Fas1 catalyzes reaction steps of the fatty acid synthesis starting from malonyl-CoA, which is formed by carboxylation of acetyl-CoA. In addition to that,

the fact that a large share of validated candidate genes are transcriptional regulators (*AGP2, IXR1, NRG1, RGS2, RGT1, SIR2*) emphasizes the role of regulation for fermentative aroma formation, a relation that has been less investigated yet. Candidate genes could also be validated that influence pDMS preservation. In particular, the validation of SMM transporter *MMP1* and identification of a truncated allele with superior pDMS preserving properties confirm both, the influence of yeast on grape-derived aroma precursors and the suitability of QTL mapping for elucidating this context.

The analytical methodology was further developed by extending the QTL mapping approach to the identification of genomic regions influencing modeled metabolic fluxes. This demonstrated that modeled phenotypic data, which was obtained with limited experimental data, could be used to detect QTLs with an influence on these modeled traits. The robustness of the approach was further proven by validation of two target genes, *PDB1* and *VID30*, whose allelic variants differently influence glycolysis, ethanol and glycerol synthesis, fluxes of the TCA cycle and various metabolite transport and excretion fluxes. Better understanding of the impact of genomic variation on intracellular fluxes is crucial as the production of yeast fermentative aroma is closely linked to fluxes of the CCM. In addition, the approach also has the potential to gain further knowledge about yeast substrate utilization or metabolite production for other biotechnological applications.

All together, these results confirm and emphasize the role of genetic diversity for fermentative aroma formation. The complexity of the underlying genomic bases is demonstrated by the high number of detected QTLs and the indication of QTL interactions. The study demonstrated once more the tight link of fermentative aroma formation to nitrogen and carbon metabolism. Besides confirming the value of QTL mapping for assessing enological traits, its utility for deciphering complex traits in general was demonstrated. Globally acting target genes could be validated for their impact on single traits and genes with influence on metabolic fluxes could be identified by flux QTL mapping.

In comparison to this study, previous QTL mapping studies on yeast aroma formation used a cross between an *S. cerevisiae* lab and wine strain. This had the advantage that the strains were more different from a genomic point of view. Greater differences between allelic gene variants might have helped to detect their general impact on the observed traits as they can result in more diverse phenotypes. On the other hand, the use of *S. cerevisiae* lab strains to investigate winemaking traits might generally bring the disadvantage to map SNPs that correspond to deleterious mutations accumulated in lab strains. In this study, two wine strains were used for the QTL mapping. Choosing this approach, alleles with an actual application for winemaking could be detected and were characterized. In addition to that, generating the F2-generation increased the recombination rate of the segregants and a comparatively large number of strains was assessed. These measures led to the detection of QTLs with minor influence and QTL interactions.

One strain of this study, 4CAR1, is a hybrid between a wine and a flor strains and is used for champagne making. Some phenotypic traits could be connected to this genomic origin such as the sugar uptake, for which the causative allelic variant of hexose transporter *HXT3* was found to be shared with other flor strains. In addition, an allelic variant of pDMS transporter *MMP1* was characterized for its ability to preserve pDMS in must and was connected to the flor origin. However, compared to yeast strains with other applications such as the production of beer or sake, the parental strains are close from a genomic point of view. Nevertheless, variation could be seen for most traits and the cross of both parents generated an even greater variation among the segregant population. This did not only concern the aroma production, but the metabolic profile in general. It indicates the presence of a rich genetic resource for strain improvement by in-crossing of superior alleles. In addition, if these close strains already have the potential to generate great variety, the use of more distant yeasts such as beer, rum or sake strains offers even wider possibilities for breeding programs.

A general evaluation of the allele frequency among the yeast population can be made for the three validated genes with performed sequence alignments. The alleles of SMM transporter gene *MMP1* are highly similar among wine and flor strains, which includes both parental alleles. However, the occurrence of the SNP causative for superior SMM preservation is rather rare and mainly connected to flor origin. Concerning pyruvate dehydrogenase subunit gene *PDB1*, besides an insertion of three nucleotides that is only found in three other strains (soil, wine and wine x flor origin), the MTF2621 allele is common among wine strains and is also found in strains of other niches such as beer or bread. Surprisingly, the *PDB1* allele of wine strain MTF2622 is more different to other assessed wine strains. The allele of palm wine strain PW5 is the most divergent in comparison to all other assessed strains. For the gene of GID complex component Vid30, both parental alleles differ considerably from the majority of evaluated wine strains. The MTF2621 allele is located between most wine alleles and the genotypically similar wine x flor strain EC1118. The *VID30* allele of MTF2622 shares similarity with a cluster of several strains with different origins, namely laboratory, African and soil. *VID30* alleles of palm wine and sake strains are the most divergent to those of wine strains.

Universities and research institutes often face the prejudice that they produce theoretical and basic knowledge with little industrial or economic relevance. The results obtained by this study prove otherwise. Although the findings were not used to construct new wine yeast strains with novel aromatic properties for commercialization, the identification of superior allelic gene variants provides targets for future strain improvements. This could include the construction of strains with higher general production of fermentative aromas. However, this broad increase is likely to generate little added value, as a variety of strains producing different levels of fermentative aromas already exist. A more promising approach is the targeted shape of certain aroma characteristics. This could include an overproduction of single fermentative compounds, e.g., 2-phenylethyl acetate, which contributes floral and rose notes to alcoholic beverages. Allelic variants of Leu9 and Rgs2 were demonstrated to affect 2-phenylethyl acetate production and an impact of Agp1 and Mae1 on the formation of the underlying higher

alcohol 2-phenylethanol could be shown. Another possibility would be the optimization of strains regarding their terpenol metabolism. As pointed out, different terpenol compounds contribute different flavors to wine. It was shown that the segregant strains differed in their terpenol profile and QTLs with an impact on this profile could be detected. However, target genes remain to be validated, which would facilitate the process of strain improvement if performed in future studies. A third possibility would be the selection or construction of strains with pDMS conserving properties. This is facilitated by the identification of the truncated Mmp1 variant that leads to the conservation of pDMS in the must. This variant could be identified in various strains of different origin, therefore enabling the selection of a suited strain. In addition, the information can also lead to the insertion of the allele in existent industrial strains by non-GMO methods. It was furthermore shown that the detected allelic variants of *PDB1* and *VID30* show differences in flux distribution between the glycerol and ethanol synthesis. Although the caused variation is relatively small, the redistribution of carbon to glycerol production plays an important role for the construction of wine yeast strains with reduced alcohol production.

The obtained information about target genes and superior alleles can also be used for other purposes. This includes the biosynthesis of bulk and fine chemicals, e.g., higher alcohols or organic acids, for applications such as biofuel or bioplastic production. In addition, knowledge about all presented target genes can be used to characterize other allelic variants in yeast strains from various origins regarding their impact on the described traits.

The presented results give answers to the majority of issues drawn in the objectives of the thesis. However, they also led to the rise of new questions:

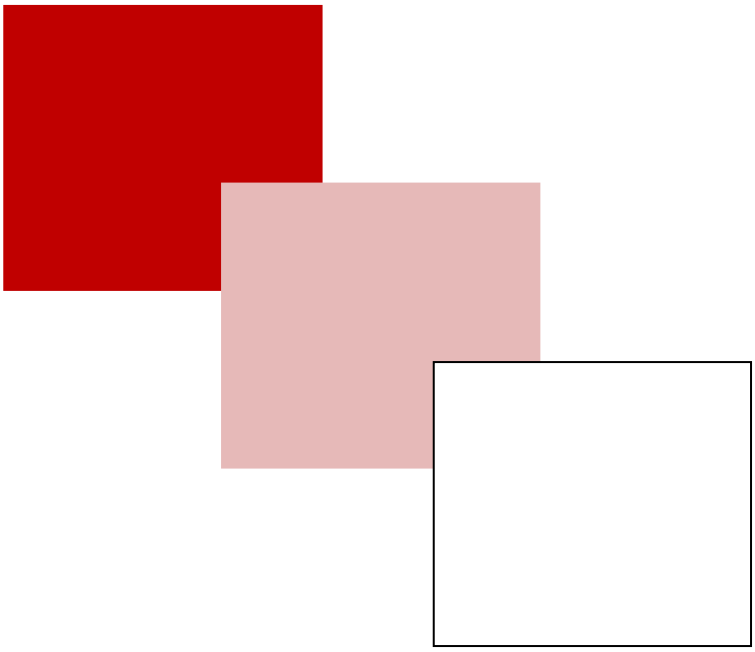
- What proportion does the contribution of QTL interactions have for fermentative aroma formation and how important is their global influence?
  To fully answer this question, a much larger population of segregants is needed to significantly detect QTLs and QTL interactions with small impacts.
- Which characteristics do different allelic variants of identified target genes show and can superior variants be found in strains from other niches?
- Why does pDMS concentration rapidly decrease at the beginning of fermentation?
  This can be assessed by the inoculation with inactivated yeast and the subsequent determination of pDMS. An equal decrease in pDMS concentrations would indicate the physical adsorption of pDMS to cellular components.
- Which other genes are involved in pDMS consumption?
  One possibility to determine additionally involved genes would be repeated QTL mapping with newly generated or existing crosses of yeast strains that both contain an identical *MMP1* allele.
- Why do overall terpenol levels decrease during fermentation? Are further compounds produced that were not determined during this study? Do these compounds have organoleptic properties for wine?

In the same manner as for pDMS, the use of inactivated cells can also indicate physical adsorption of terpenols to components of the cell.

- Which allelic variants are responsible for terpenol transformation in detected QTLs?
- By which means do conversions between terpenols occur during fermentation?
  If target genes for terpenols are validated, this question can already be answered by considering the biological function of these genes.
- Can grape-derived terpenols be incorporated in the sterol pathway of yeast and therefore support cell proliferation? This hypothesis could be evaluated by fermenting yeast strains in synthetic medium spiked with ^{13}C-labeled geraniol. By extracting and determining synthesized sterols after fermentation, the presence of ^{13}C-labeled sterols would affirm the question.

It is expected that these questions are pursued in future studies for which this thesis lays the foundation. Furthermore, the generated segregant population with associated genotype information can be used to decipher the bases for additional traits.

Bibliography



Abt, T. Den, Souffriau, B., Foulquié-Moreno, M.R., Duitama, J., Thevelein, J.M., 2016. Genomic saturation mutagenesis and polygenic analysis identify novel yeast genes affecting ethyl acetate production, a non-selectable polygenic trait. Microb. Cell 3, 159–175.

Agren, R., Otero, J.M., Nielsen, J., 2013. Genome-scale modeling enables metabolic engineering of Saccharomyces cerevisiae for succinic acid production. J. Ind. Microbiol. Biotechnol. 40, 735–747.

Alba-Lois, L., Segal-Kischinevzky, C., 2010. Beer & wine makers. Nature.

Alexandre, H., Rousseaux, I., Charpentier, C., 1994. Ethanol adaptation mechanisms in Saccharomyces cerevisiae. Biotechnol. Appl. Biochem. 20, 173–183.

Alibhoy, A.A., Giardina, B.J., Dunton, D.D., Chiang, H.-L., 2012. Vid30 is required for the association of Vid vesicles and actin patches in the vacuole import and degradation pathway. Autophagy 8, 29–46.

Ambroset, C., Petit, M., Brion, C., Sanchez, I., Delobel, P., Guérin, C., Chiapello, H., Nicolas, P., Bigey, F., Dequin, S., Blondin, B., 2011. Deciphering the molecular basis of wine yeast fermentation traits using a combined genetic and genomic approach. G3 Genes, Genomes, Genet. 1, 263–281.

Anness, B.J., Bamforth, C.W., 1982. Dimethyl sulphide - a review. J. Inst. Brew. 88, 244–252.

Aouida, M., Texeira, M.R., Thevelein, J.M., Poulin, R., Ramotar, D., 2013. Agp2, a member of the yeast amino acid permease family, positively regulates polyamine transport at the transcriptional level. PLoS One 8, e65717.

Arends, D., Prins, P., Jansen, R.C., Broman, K.W., 2010. R/qtl: High-throughput multiple QTL mapping. Bioinformatics 26, 2990–2992.

Avendaño, A., Deluna, A., Olivera, H., Valenzuela, L., Gonzalez, A., 1997. GDH3 encodes a glutamate dehydrogenase isozyme, a previously unrecognized route for glutamate biosynthesis in Saccharomyces cerevisiae. J. Bacteriol. 179, 5594–5597.

Bamforth, C.W., Anness, B.J., 1981. The role of dimethyl sulphoxide reductase in the formation of dimethyl sulphide during fermentations. J. Inst. Brew. 87, 30–34.

Banthorpe, D. V, Long, D.R.S., Pink, C.R., 1983. Biosynthesis of geraniol and related monoterpenes in Pelargonium graveolens. Phytochemistry 22, 2459–2463.

Barnard, H., Dooley, A.N., Areshian, G., Gasparyan, B., Faull, K.F., 2011. Chemical evidence for wine production around 4000 BCE in the Late Chalcolithic Near Eastern highlands. J. Archaeol. Sci. 38, 977–984.

Bell, S., Henschke, P.A., 2005. Implications of nitrogen nutrition for grapes, fermentation and wine. Aust. J. Grape Wine Res. 11, 242–295.

Beloqui, A.A., Kotseridis, Y., Bertrand, A., 1996. Determination of the content of dimethyl sulphide in some red wines. OENO One 30, 167–170.

Beltran, G., Esteve-Zarzoso, B., Rozès, N., Mas, A., Guillamón, J.M., 2005. Influence of the timing of nitrogen additions during synthetic grape must fermentations on fermentation kinetics and nitrogen consumption. J. Agric. Food Chem. 53, 996–1002.

Beltran, G., Novo, M., Guillamón, J.M., Mas, A., Rozès, N., 2008. Effect of fermentation temperature and culture media on the yeast lipid composition and wine volatile compounds. Int. J. Food Microbiol. 121, 169–177.

Bely, M., Sablayrolles, J.-M., Barre, P., 1990. Automatic detection of assimilable nitrogen deficiencies during alcoholic fermentation in oenological conditions. J. Ferment. Bioeng. 70, 246–252.

Bick, J.A., Lange, B.M., 2003. Metabolic cross talk between cytosolic and plastidial pathways of isoprenoid biosynthesis: unidirectional transport of intermediates across the

chloroplast envelope membrane. Arch. Biochem. Biophys. 415, 146–154.
Bishop, J.R., Nelson, G., Lamb, J., 1998. Microencapsulation in yeast cells. J. Microencapsul. 15, 761–773.
Bisson, L.F., 1999. Stuck and sluggish fermentations. Am. J. Enol. Vitic. 50, 107–119.
Black, C.A., Parker, M., Siebert, T.E., Capone, D.L., Francis, I.L., 2015. Terpenoids and their role in wine flavour: recent advances. Aust. J. grape wine Res. 21, 582–600.
Blank, L.M., Lehmbeck, F., Sauer, U., 2005. Metabolic-flux and network analysis in fourteen hemiascomycetous yeasts. FEMS Yeast Res. 5, 545–558.
Blateyron, L., Sablayrolles, J.M., 2001. Stuck and slow fermentations in enology: statistical study of causes and effectiveness of combined additions of oxygen and diammonium phosphate. J. Biosci. Bioeng. 91, 184–189.
Bloem, A., Sanchez, I., Dequin, S., Camarasa, C., 2015. Metabolic impact of redox cofactor perturbations on the formation of aroma compounds in Saccharomyces cerevisiae. Appl. Environ. Microbiol. AEM-02429.
Bloom, J.S., Kotenko, I., Sadhu, M.J., Treusch, S., Albert, F.W., Kruglyak, L., 2015. Genetic interactions contribute less than additive effects to quantitative trait variation in yeast. Nat. Commun. 6, 8712.
Bokulich, N.A., Thorngate, J.H., Richardson, P.M., Mills, D.A., 2014. Microbial biogeography of wine grapes is conditioned by cultivar, vintage, and climate. Proc. Natl. Acad. Sci. 111, E139–E148.
Boles, E., de Jong-Gubbels, P., Pronk, J.T., 1998. Identification and Characterization of MAE1, the Saccharomyces cerevisiae Structural Gene Encoding Mitochondrial Malic Enzyme. J. Bacteriol. 180, 2875–2882.
Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120.
Bourgis, F., Roje, S., Nuccio, M.L., Fisher, D.B., Tarczynski, M.C., Li, C., Herschbach, C., Rennenberg, H., Pimenta, M.J., Shen, T.-L., 1999. S-methylmethionine plays a major role in phloem sulfur transport and is synthesized by a novel type of methyltransferase. Plant Cell 11, 1485–1497.
Brauer, M.J., Christianson, C.M., Pai, D.A., Dunham, M.J., 2006. Mapping novel traits by array-assisted bulk segregant analysis in Saccharomyces cerevisiae. Genetics 173, 1813–1816.
Brem, R.B., Yvert, G., Clinton, R., Kruglyak, L., 2002. Genetic dissection of transcriptional regulation in budding yeast. Science (80-.). 296, 752–755.
Brice, C., Sanchez, I., Bigey, F., Legras, J.-L., Blondin, B., 2014a. A genetic approach of wine yeast fermentation capacity in nitrogen-starvation reveals the key role of nitrogen signaling. BMC Genomics 15, 495.
Brice, C., Sanchez, I., Tesnière, C., Blondin, B., 2014b. Assessing the mechanisms responsible for differences between nitrogen requirements of Saccharomyces cerevisiae wine yeasts in alcoholic fermentation. Appl. Environ. Microbiol. 80, 1330–1339.
Brion, C., Ambroset, C., Sanchez, I., Legras, J.-L., Blondin, B., 2013. Differential adaptation to multi-stressed conditions of wine fermentation revealed by variations in yeast regulatory networks. BMC Genomics 14, 681.
Bro, C., Regenberg, B., Förster, J., Nielsen, J., 2006. In silico aided metabolic engineering of Saccharomyces cerevisiae for improved bioethanol production. Metab. Eng. 8, 102–111.
Broman, K.W., Wu, H., Sen, Ś., Churchill, G.A., 2003. R/qtl: QTL mapping in experimental crosses. Bioinformatics 19, 889.
Bundy, J.G., Papp, B., Harmston, R., Browne, R.A., Clayson, E.M., Burton, N., Reece, R.J., Oliver,

S.G., Brindle, K.M., 2007. Evaluation of predicted network modules in yeast metabolism using NMR-based metabolite profiling. Genome Res. 17, 510–519.

Burgard, A.P., Maranas, C.D., 2003. Optimization-based framework for inferring and testing hypothesized metabolic objective functions. Biotechnol. Bioeng. 82, 670–677.

Buttery, R.G., Seifert, R.M., Guadagni, D.G., Ling, L.C., 1971. Characterization of additional volatile components of tomato. J. Agric. Food Chem. 19, 524–529.

Cadière, A., Ortiz-Julien, A., Camarasa, C., Dequin, S., 2011. Evolutionary engineered Saccharomyces cerevisiae wine yeast strains with increased in vivo flux through the pentose phosphate pathway. Metab. Eng. 13, 263–271.

Camarasa, C., Grivet, J.-P., Dequin, S., 2003. Investigation by 13C-NMR and tricarboxylic acid (TCA) deletion mutant analysis of pathways for succinate formation in Saccharomyces cerevisiae during anaerobic fermentation. Microbiology 149, 2669–2678.

Camarasa, C., Sanchez, I., Brial, P., Bigey, F., Dequin, S., 2011. Phenotypic landscape of Saccharomyces cerevisiae during wine fermentation: evidence for origin-dependent metabolic traits. PLoS One 6, e25147.

Carbonell, M., Nuñez, M., Fernández-García, E., 2002. Seasonal variation of volatile compounds in ewe raw milk La Serena cheese. Lait 82, 699–711.

Carlson, M., 1999. Glucose repression in yeast. Curr. Opin. Microbiol. 2, 202–207.

Carrau, F.M., Medina, K., Boido, E., Farina, L., Gaggero, C., Dellacassa, E., Versini, G., Henschke, P. a., 2005. De novo synthesis of monoterpenes by Saccharomyces cerevisiae wine yeasts. FEMS Microbiol. Lett. 243, 107–115.

Casatta, N., Porro, A., Orlandi, I., Brambilla, L., Vai, M., 2013. Lack of Sir2 increases acetate consumption and decreases extracellular pro-aging factors. Biochim. Biophys. Acta (BBA)-Molecular Cell Res. 1833, 593–601.

Cavalieri, D., McGovern, P.E., Hartl, D.L., Mortimer, R., Polsinelli, M., 2003. Evidence for S. cerevisiae fermentation in ancient wine. J. Mol. Evol. 57, S226–S232.

Celton, M., Goelzer, A., Camarasa, C., Fromion, V., Dequin, S., 2012a. A constraint-based model analysis of the metabolic consequences of increased NADPH oxidation in Saccharomyces cerevisiae. Metab. Eng. 14, 366–379.

Celton, M., Sanchez, I., Goelzer, A., Fromion, V., Camarasa, C., Dequin, S., 2012b. A comparative transcriptomic, fluxomic and metabolomic analysis of the response of Saccharomyces cerevisiae to increases in NADPH oxidation. BMC Genomics 13, 317.

Chatonnet, P., Dubourdie, D., Boidron, J., Pons, M., 1992. The origin of ethylphenols in wines. J. Sci. Food Agric. 60, 165–178.

Chatonnet, P., Dubourdieu, D., Boidron, J., Lavigne, V., 1993. Synthesis of volatile phenols by Saccharomyces cerevisiae in wines. J. Sci. Food Agric. 62, 191–202.

Cheng, A., Lou, Y., Mao, Y., Lu, S., Wang, L., Chen, X., 2007. Plant terpenoids: biosynthesis and ecological functions. J. Integr. Plant Biol. 49, 179–186.

Codon, A.C., Gasent-Ramirez, J.M., Benitez, T., 1995. Factors which affect the frequency of sporulation and tetrad formation in Saccharomyces cerevisiae baker's yeasts. Appl. Environ. Microbiol. 61, 630–638.

Coi, A.L., Bigey, F., Mallet, S., Marsit, S., Zara, G., Gladieux, P., Galeote, V., Budroni, M., Dequin, S., Legras, J.L., 2017. Genomic signatures of adaptation to wine biological ageing conditions in biofilm-forming flor yeasts. Mol. Ecol. 26, 2150–2166.

Coi, A.L., Legras, J.-L., Zara, G., Dequin, S., Budroni, M., 2016. A set of haploid strains available for genetic studies of Saccharomyces cerevisiae flor yeasts. FEMS Yeast Res. 16, fow066.

Cordente, A.G., Curtin, C.D., Varela, C., Pretorius, I.S., 2012. Flavor-active wine yeasts. Appl.

Microbiol. Biotechnol. 96, 601–618.
Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E.D., Sevier, C.S., Ding, H., Koh, J.L.Y., Toufighi, K., Mostafavi, S., 2010. The genetic landscape of a cell. Science (80-.). 327, 425–431.
Crépin, L., Truong, N.M., Bloem, A., Sanchez, I., Dequin, S., Camarasa, C., 2017. Management of Multiple Nitrogen Sources during Wine Fermentation by Saccharomyces cerevisiae. Appl. Environ. Microbiol. 83, e02617-16.
Cubillos, F.A., Brice, C., Molinet, J., Tisné, S., Abarca, V., Tapia, S.M., Oporto, C., García, V., Liti, G., Martínez, C., 2017. Identification of nitrogen consumption genetic variants in yeast through QTL mapping and Bulk segregant RNA-seq analyses. G3 Genes, Genomes, Genet. 7, 1693–1705.
Cubillos, F.A., Louis, E.J., Liti, G., 2009. Generation of a large set of genetically tractable haploid and diploid Saccharomyces strains. FEMS Yeast Res. 9, 1217–1225.
Culleré, L., Ferreira, V., Chevret, B., Venturini, M.E., Sánchez-Gimeno, A.C., Blanco, D., 2010. Characterisation of aroma active compounds in black truffles (Tuber melanosporum) and summer truffles (Tuber aestivum) by gas chromatography–olfactometry. Food Chem. 122, 300–306.
Cullin, C., Baudin-Baillieu, A., Guillemet, E., Ozier-Kalageropoulos, O., Ozier-Kalogeropoulos, O., 1996. Functional analysis of YCL09C: evidence for a role as the regulatory subunit of acetolactate synthase. Yeast 12, 1511–1518.
Dagan, L., 2006. Potentiel aromatique des raisins de Vitis Vinifera L. cv. Petit Manseng Gros Manseng. Contrib. al'arôme des vins pays des Côtes Gascogne., Agro M.
Dagan, L., Schneider, R., 2012. Le sulfure de diméthyle: quels moyens pour gérer ses teneurs dans les vins en bouteille. L'arômes des Vins. Toulouse/Zaragoza 7–10.
Darriet, P., Lavigne-Cruège, V., Boidron, J.-N., Dubourdieu, D., 1991. Caractérisation de l'arôme variétal des vins de Sauvignon par couplage chromatographie en phase gazeuse-odométrie. OENO One 25, 167–174.
De Deken, R.H., 1966. The Crabtree effect: a regulatory system in yeast. Microbiology 44, 149–156.
de Mora, S.J., Eschenbruch, R., Knowles, S.J., Spedding, D.J., 1986. The formation of dimethyl sulphide during fermentation using a wine yeast. Food Microbiol. 3, 27–32.
De Mora, S.J., Knowles, S.J., Eschenbruch, R., Torrey, W.J., 1987. Dimethyl sulphide in some Australian red wines. Vitis.
De Orduna, R.M., 2010. Climate change associated effects on grape and wine quality and production. Food Res. Int. 43, 1844–1855.
Deutschbauer, A.M., Davis, R.W., 2005. Quantitative trait loci mapped to single-nucleotide resolution in yeast. Nat. Genet. 37, 1333–1340.
DiCarlo, J.E., Norville, J.E., Mali, P., Rios, X., Aach, J., Church, G.M., 2013. Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. Nucleic Acids Res. 41, 4336–4343.
Dickinson, J.R., Norte, V., 1993. A study of branched-chain amino acid aminotransferase and isolation of mutations affecting the catabolism of branched-chain amino acids in Saccharomyces cerevisiae. FEBS Lett. 326, 29–32.
Dignan, D.M., Wiley, R.C., 1976. DMS levels in the aroma of cooked frozen sweet corn as affected by cultivar, maturity, blanching and packaging. J. Food Sci. 41, 346–348.
Drawert, F., Heimann, W., Emberger, R., Tressl, R., 1966. Über die Biogenese von Aromastoffen bei Pflanzen und Früchten, II. Enzymatische Bildung von Hexen-(2)-al-(1),

Hexanal und deren Vorstufen. European J. Org. Chem. 694, 200–208.
Dubois, P., 1983. Volatile phenols in wines. Flavour Distill. beverages. London, UK Soc. Chem. Ind. 110–119.
Duchêne, E., Legras, J.L., Karst, F., Merdinoglu, D., Claudel, P., Jaegli, N., Pelsy, F., 2009. Variation of linalool and geraniol content within two pairs of aromatic and non-aromatic grapevine clones. Aust. J. Grape Wine Res. 15, 120–130.
Duncan, K., Edwards, R.M., Coggins, J.R., 1987. The pentafunctional arom enzyme of Saccharomyces cerevisiae is a mosaic of monofunctional domains. Biochem. J. 246, 375–386.
Dunn, B., Sherlock, G., 2008. Reconstruction of the genome origins and evolution of the hybrid lager yeast Saccharomyces pastorianus. Genome Res. 18, 1610–1623.
Eden, A., Van Nedervelde, L., Drukker, M., Benvenisty, N., Debourg, A., 2001. Involvement of branched-chain amino acid aminotransferases in the production of fusel alcohols during fermentation in yeast. Appl. Microbiol. Biotechnol. 55, 296–300.
Eder, M., Sanchez, I., Brice, C., Camarasa, C., Legras, J.-L.J., Dequin, S., 2018. QTL mapping of volatile compound production in Saccharomyces cerevisiae during alcoholic fermentation. BMC Genomics 19, 166.
Eglinton, J.M., Henschke, P.A., 1999. The occurrence of volatile acidity in Australian wines. Aust. Grapegrow. Winemak. 7–14.
Elbing, K., Larsson, C., Bill, R.M., Albers, E., Snoep, J.L., Boles, E., Hohmann, S., Gustafsson, L., 2004. Role of hexose transport in control of glycolytic flux in Saccharomyces cerevisiae. Appl. Environ. Microbiol. 70, 5323–5330.
Escudero, A., Campo, E., Fariña, L., Cacho, J., Ferreira, V., 2007. Analytical characterization of the aroma of five premium red wines. Insights into the role of odor families and the concept of fruitiness of wines. J. Agric. Food Chem. 55, 4501–4510.
Etievant, P.X., 1981. Volatile phenol determination in wine. J. Agric. Food Chem. 29, 65–67.
Fay, J.C., Benavides, J.A., 2005. Evidence for domesticated and wild populations of Saccharomyces cerevisiae. PLoS Genet. 1, e5.
Fedrizzi, B., Magno, F., Badocco, D., Nicolini, G., Versini, G., 2007. Aging effects and grape variety dependence on the content of sulfur volatiles in wine. J. Agric. Food Chem. 55, 10880–10887.
Ferreira, V., Lopez, R., Cacho, J.F., 2000. Quantitative determination of the odorants of young red wines from different grape varieties. J. Sci. Food Agric. 80, 1659–1667.
Ferreira, V., Peña, C., Escudero, A., Cacho, J., 1996. Losses of volatile compounds during fermentation. Zeitschrift für Leb. Und-forsch. A 202, 318–323.
Filetici, P., Martegani, M.P., Valenzuela, L., González, A., Ballario, P., 1996. Sequence of the GLT1 gene from Saccharomyces cerevisiae reveals the domain structure of yeast glutamate synthase. Yeast 12, 1359–1366.
Fleet, G.H., 1993. Wine microbiology and biotechnology. CRC Press.
Francis, I.L., Newton, J.L., 2005. Determining wine aroma from compositional data. Aust. J. Grape Wine Res. 11, 114–126.
Frayne, R.F., 1986. Direct analysis of the major organic components in grape must and wine using high performance liquid chromatography. Am. J. Enol. Vitic. 37, 281–287.
Fujii, T., Kobayashi, O., Yoshimoto, H., Furukawa, S., Tamai, Y., 1997. Effect of aeration and unsaturated fatty acids on expression of the Saccharomyces cerevisiae alcohol acetyltransferase gene. Appl. Environ. Microbiol. 63, 910–915.
Fujiwara, D., Yoshimoto, H., Sone, H., Harashima, S., Tamai, Y., 1998. Transcriptional co-

regulation of Saccharomyces cerevisiae alcohol acetyltransferase gene, ATF1 and Δ-9 fatty acid desaturase gene, OLE1 by unsaturated fatty acids. Yeast 14, 711–721.
Furdíková, K., Makyšová, K., Ďurčanská, K., Špánik, I., Malík, F., 2014. Influence of yeast strain on aromatic profile of Gewürztraminer wine. LWT-Food Sci. Technol. 59, 256–262.
Galdieri, L., Vancura, A., 2012. Acetyl-CoA carboxylase regulates global histone acetylation. J. Biol. Chem. 287, 23865–23876.
Gamero, A., Manzanares, P., Querol, A., Belloch, C., 2011. Monoterpene alcohols release and bioconversion by Saccharomyces species and hybrids. Int. J. Food Microbiol. 145, 92–97.
Gancedo, J.M., 1998. Yeast carbon catabolite repression. Microbiol. Mol. Biol. Rev. 62, 334–361.
García-Ríos, E., Morard, M., Parts, L., Liti, G., Guillamón, J.M., 2017. The genetic architecture of low-temperature adaptation in the wine yeast Saccharomyces cerevisiae. BMC Genomics 18, 159.
García, D.D. la C., Magnaghi, S., Reichenbächer, M., Danzer, K., 1996. Systematic optimization of the analysis of wine bouquet components by solid-phase microextraction. J. Sep. Sci. 19, 257–262.
Garde-Cerdán, T., Ancín-Azpilicueta, C., 2006. Review of quality factors on wine ageing in oak barrels. Trends food Sci. Technol. 17, 438–447.
Gavin, A.-C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L.J., Bastuck, S., Dümpelfeld, B., 2006. Proteome survey reveals modularity of the yeast cell machinery. Nature 440, 631–636.
Gawel, R., Sluyter, S.A.N., Waters, E.J., 2007. The effects of ethanol and glycerol on the body and other sensory characteristics of Riesling wines. Aust. J. grape wine Res. 13, 38–45.
Gelius-Dietrich, G., Desouki, A.A., Fritzemeier, C.J., Lercher, M.J., 2013. sybil–Efficient constraint-based modelling in R. BMC Syst. Biol. 7, 125.
Gietz, R.D., Schiestl, R.H., Willems, A.R., Woods, R.A., 1995. Studies on the transformation of intact yeast cells by the LiAc/ss-DNA/PEG procedure. Yeast 11, 355–360.
Gil, J. V, Mateo, J.J., Jiménez, M., Pastor, A., Huerta, T., 1996. Aroma compounds in wine as influenced by apiculate yeasts. J. Food Sci. 61, 1247–1250.
Goddard, M.R., 2008. Quantifying the complexities of Saccharomyces cerevisiae's ecosystem engineering via fermentation. Ecology 89, 2077–2082.
Grando, M.S., Versini, G., Nicolini, G., 1993. Selective use of wine yeast strains having different volatile phenols production. Vitis 32, 43–50.
Grant-Preece, P.A., Pardon, K.H., Capone, D.L., Cordente, A.G., Sefton, M.A., Jeffery, D.W., Elsey, G.M., 2010. Synthesis of wine thiol conjugates and labeled analogues: fermentation of the glutathione conjugate of 3-mercaptohexan-1-ol yields the corresponding cysteine conjugate and free thiol. J. Agric. Food Chem. 58, 1383–1389.
Griffiths, A.J.F., 2002. Modern genetic analysis: integrating genes and genomes. Macmillan.
Guillaume, C., Delobel, P., Sablayrolles, J.M., Blondin, B., 2007. Molecular basis of fructose utilization by the wine yeast Saccharomyces cerevisiae: A mutated HXT3 allele enhances fructose fermentation. Appl. Environ. Microbiol. 73, 2432–2439.
Guth, H., 1997. Identification of character impact odorants of different white wine varieties. J. Agric. Food Chem. 45, 3022–3026.
Gutiérrez, A., Beltran, G., Warringer, J., Guillamón, J.M., 2013. Genetic basis of variations in nitrogen source utilization in four wine commercial yeast strains. PLoS One 8, e67166.
Hagen, K.M., Keller, M., Edwards, C.G., 2008. Survey of biotin, pantothenic acid, and assimilable nitrogen in winegrapes from the Pacific Northwest. Am. J. Enol. Vitic. 59, 432–

436.
Hagman, A., Piškur, J., 2015. A study on the fundamental mechanism and the evolutionary driving forces behind aerobic fermentation in yeast. PLoS One 10, e0116942.
Hagman, A., Säll, T., Compagno, C., Piskur, J., 2013. Yeast "make-accumulate-consume" life strategy evolved as a multi-step process that predates the whole genome duplication. PLoS One 8, e68734.
Haley, C.S., Knott, S.A., 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity (Edinb). 69, 315–324.
Hämmerle, M., Bauer, J., Rose, M., Szallies, A., Thumm, M., Düsterhus, S., Mecke, D., Entian, K.-D., Wolf, D.H., 1998. Proteins of newly isolated mutants and the amino-terminal proline are essential for ubiquitin-proteasome-catalyzed catabolite degradation of fructose-1, 6-bisphosphatase of Saccharomyces cerevisiae. J. Biol. Chem. 273, 25000–25005.
Hansen, J., 1999. Inactivation of MXR1 abolishes formation of dimethyl sulfide from dimethyl sulfoxide in Saccharomyces cerevisiae. Appl. Environ. Microbiol. 65, 3915–3919.
Hazelwood, L.A., Daran, J., Maris, A.J.A. Van, Pronk, J.T., Dickinson, J.R., van Maris, A.J.A., Pronk, J.T., Dickinson, J.R., 2008. The Ehrlich Pathway for Fusel Alcohol Production: a Century of Research on. Society 74, 2259–2266.
Heyland, J., Fu, J., Blank, L.M., 2009. Correlation between TCA cycle flux and glucose uptake rate during respiro-fermentative growth of Saccharomyces cerevisiae. Microbiology 155, 3827–3837.
Holzer, H., 1988. Proteolytic catabolite inactivation in Saccharomyces cerevisiae. Revis. sobre Biol. Cel. RBC 21, 305–319.
Hu, X.H., Wang, M.H., Tan, T., Li, J.R., Yang, H., Leach, L., Zhang, R.M., Luo, Z.W., 2007. Genetic dissection of ethanol tolerance in the budding yeast Saccharomyces cerevisiae. Genetics 175, 1479–1487.
Huang, P.-H., Chiang, H.-L., 1997. Identification of novel vesicles in the cytosol to vacuole protein degradation pathway. J. Cell Biol. 136, 803–810.
Hubmann, G., Foulquié-Moreno, M.R., Nevoigt, E., Duitama, J., Meurens, N., Pais, T.M., Mathé, L., Saerens, S., Nguyen, H.T.T., Swinnen, S., 2013a. Quantitative trait analysis of yeast biodiversity yields novel gene tools for metabolic engineering. Metab. Eng. 17, 68–81.
Hubmann, G., Mathé, L., Foulquié-Moreno, M.R., Duitama, J., Nevoigt, E., Thevelein, J.M., 2013b. Identification of multiple interacting alleles conferring low glycerol and high ethanol yield in Saccharomyces cerevisiae ethanolic fermentation. Biotechnol. Biofuels 6, 87.
Hung, G.-C., Brown, C.R., Wolfe, A.B., Liu, J., Chiang, H.-L., 2004. Degradation of the gluconeogenic enzymes fructose-1, 6-bisphosphatase and malate dehydrogenase is mediated by distinct proteolytic pathways and signaling events. J. Biol. Chem. 279, 49138–49150.
Ibanez, J.G., Carreon-Alvarez, A., Barcena-Soto, M., Casillas, N., 2008. Metals in alcoholic beverages: A review of sources, effects, concentrations, removal, speciation, and analysis. J. food Compos. Anal. 21, 672–683.
Ingledew, W.M., Kunkee, R.E., 1985. Factors influencing sluggish fermentations of grape juice. Am. J. Enol. Vitic. 36, 65–76.
Jansen, R.C., 1993. Interval mapping of multiple quantitative trait loci. Genetics 135, 205–211.
Jara, M., Cubillos, F.A., García, V., Salinas, F., Aguilera, O., Liti, G., Martínez, C., 2014. Mapping genetic variants underlying differences in the central nitrogen metabolism in fermenter

yeasts. PLoS One 9, e86533.
Javelot, C., Girard, P., Colonna-Ceccaldi, B., Vladescu, B., 1991. Introduction of terpene-producing ability in a wine strain of Saccharomyces cerevisiae. J. Biotechnol. 21, 239–251.
Jiménez-Martí, E., Aranda, A., Mendes-Ferreira, A., Mendes-Faia, A., lí del Olmo, M., 2007. The nature of the nitrogen source added to nitrogen depleted vinifications conducted by a Saccharomyces cerevisiae strain in synthetic must affects gene expression and the levels of several volatile compounds. Antonie Van Leeuwenhoek 92, 61–75.
Jolly, N.P., Varela, C., Pretorius, I.S., 2014. Not your ordinary yeast: non-Saccharomyces yeasts in wine production uncovered. FEMS Yeast Res. 14, 215–237.
Jüttner, F., 1995. Physiology and biochemistry of odorous compounds from freshwater cyanobacteria and algae. Water Sci. Technol. 31, 69–78.
King, A., Dickinson, J.R., 2003. Biotransformation of hop aroma terpenoids by ale and lager yeasts. FEMS Yeast Res. 3, 53–62.
King, A., Dickinson, J.R., 2000. Biotransformation of monoterpene alcohols by Saccharomyces cerevisiae, Torulaspora delbrueckii and Kluyveromyces lactis. Yeast 16, 499–506.
Kolodziej, S.J., Penczek, P. a., Schroeter, J.P., Stoops, J.K., Kolodziej, S.J., Schroeter, J.P., Stoops, J.K., Penczek, P. a., Schroeter, J.P., Stoops, J.K., 1996. Structure-Function Relationships of the Saccharomyces cerevisiae Fatty Acid Synthase Three-Dimensional Structure. J. Biol. Chem. 271, 28422–28429.
Koslitz, S., Renaud, L., Kohler, M., Wüst, M., 2008. Stereoselective formation of the varietal aroma compound rose oxide during alcoholic fermentation. J. Agric. Food Chem. 56, 1371–1375.
Krogerus, K., Magalhães, F., Vidgren, V., Gibson, B., 2015. New lager yeast strains generated by interspecific hybridization. J. Ind. Microbiol. Biotechnol. 42, 769–778.
Kuchin, S., Vyas, V.K., Carlson, M., 2002. Snf1 protein kinase and the repressors Nrg1 and Nrg2 regulate FLO11, haploid invasive growth, and diploid pseudohyphal differentiation. Mol. Cell. Biol. 22, 3994–4000.
Kutyna, D.R., Varela, C., Henschke, P.A., Chambers, P.J., Stanley, G.A., 2010. Microbiological approaches to lowering ethanol concentration in wine. Trends Food Sci. Technol. 21, 293–302.
Lambrechts, M.G., Pretorius, I.S., 2000. Yeast and its Importance to Wine Aroma - A Review. South African J. Enol. Vitic. 21, 97–129.
Lander, E.S., Botstein, D., 1989. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121, 185–199.
Lee, P., Shooter, D., Eschenbruch, R., 1993. The analysis and importance of dimethylsulfoxide in wine. Am. J. Enol. Vitic. 44, 327–332.
Lee, S.A., Rick, F.E., Dobson, J., Reeves, M., Clark, H., Thomson, M., Gardner, R.C., 2008. Grape juice is the major influence on volatile thiol aromas in Sauvignon Blanc. Aust. New Zeal. Grapegrow. Winemak. 533, 78–86.
Legras, J.-L., Ruh, O., Merdinoglu, D., Karst, F., 2005. Selection of hypervariable microsatellite loci for the characterization of Saccharomyces cerevisiae strains. Int. J. Food Microbiol. 102, 73–83.
Legras, J., Merdinoglu, D., Cornuet, J., Karst, F., 2007. Bread, beer and wine: Saccharomyces cerevisiae diversity reflects human history. Mol. Ecol. 16, 2091–2102.
Lestremau, F., Andersson, F.A.T., Desauziers, V., Fanlo, J.-L., 2003. Evaluation of solid-phase microextraction for time-weighted average sampling of volatile sulfur compounds at ppb

concentrations. Anal. Chem. 75, 2626–2632.
Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25, 1754–1760.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078–2079.
Libkind, D., Hittinger, C.T., Valério, E., Gonçalves, C., Dover, J., Johnston, M., Gonçalves, P., Sampaio, J.P., 2011. Microbe domestication and the identification of the wild genetic stock of lager-brewing yeast. Proc. Natl. Acad. Sci. 108, 14539–14544.
Lilly, M., Bauer, F.F., Lambrechts, M.G., Swiegers, J.H., Cozzolino, D., Pretorius, I.S., 2006a. The effect of increased yeast alcohol acetyltransferase and esterase activity on the flavour profiles of wine and distillates. Yeast 23, 641–659.
Lilly, M., Bauer, F.F., Styger, G., Lambrechts, M.G., Pretorius, I.S., 2006b. The effect of increased branched-chain amino acid transaminase activity in yeast on the production of higher alcohols and on the flavour profiles of wine and distillates. FEMS Yeast Res. 6, 726–743.
Lilly, M., Lambrechts, M.G., Pretorius, I.S., 2000. Effect of increased yeast alcohol acetyltransferase activity on flavor profiles of wine and distillates. Appl. Environ. Microbiol. 66, 744–753.
Lin, Y., Qi, Y., Lu, J., Pan, X., Yuan, D.S., Zhao, Y., Bader, J.S., Boeke, J.D., 2008. A comprehensive synthetic genetic interaction network governing yeast histone acetylation and deacetylation. Genes Dev. 22, 2062–2074.
Lindstrom, E.W., 1924. A genetic linkage between size and color factors in the tomato. Science (80-.). 60, 182–183.
Liti, G., Carter, D.M., Moses, A.M., Warringer, J., Parts, L., James, S.A., Davey, R.P., Roberts, I.N., Burt, A., Koufopanou, V., Tsai, I.J., Bergman, C.M., Bensasson, D., O'Kelly, M.J.T., van Oudenaarden, A., Barton, D.B.H., Bailes, E., Nguyen, A.N., Jones, M., Quail, M. a, Goodhead, I., Sims, S., Smith, F., Blomberg, A., Durbin, R., Louis, E.J., 2009. Population genomics of domestic and wild yeasts. Nature 458, 337.
Liti, G., Warringer, J., Blomberg, A., 2017. Mapping quantitative trait loci in yeast. Cold Spring Harb. Protoc. 2017, pdb-prot089060.
Liu, B., Larsson, L., Caballero, A., Hao, X., Öling, D., Grantham, J., Nyström, T., 2010. The polarisome is required for segregation and retrograde transport of protein aggregates. Cell 140, 257–267.
Lõoke, M., Kristjuhan, K., Kristjuhan, A., 2011. Extraction of genomic DNA from yeasts for PCR-based applications. Biotechniques 50, 325.
López-Boado, Y.S., Herrero, P., Gascon, S., Moreno, F., 1987. Catabolite inactivation of isocitrate lyase from Saccharomyces cerevisiae. Arch. Microbiol. 147, 231–234.
Loscos, N., Ségurel, M., Dagan, L., Sommerer, N., Marlin, T., Baumes, R., 2008. Identification of S-methylmethionine in Petit Manseng grapes as dimethyl sulphide precursor in wine. Anal. Chim. Acta 621, 24–29.
Luan, F., Mosandl, A., Münch, A., Wüst, M., 2005. Metabolism of geraniol in grape berry mesocarp of Vitis vinifera L. cv. Scheurebe: demonstration of stereoselective reduction, E/Z-isomerization, oxidation and glycosylation. Phytochemistry 66, 295–303.
Lynch, M., Walsh, B., 1998. Genetics and analysis of quantitative traits. Sinauer Sunderland, MA.
Lynen, F., Engeser, H., Foerster, E., Fox, J.L., Hess, S., Kresze, G., Schmitt, T., Schreckenbach,

T., Siess, E., Wieland, F., 1980. On the structure of fatty acid synthetase of yeast. FEBS J. 112, 431–442.

Lytra, G., Tempere, S., Marchand, S., de Revel, G., Barbe, J.-C., 2016. How do esters and dimethyl sulphide concentrations affect fruity aroma perception of red wine? Demonstration by dynamic sensory profile evaluation. Food Chem. 194, 196–200.

Maarse, H., 1991. Volatile compounds in foods and beverages. CRC press.

Mackay, T.F.C., 2001. The genetic architecture of quantitative traits. Annu. Rev. Genet. 35, 303–339.

Magalhães, F., Krogerus, K., Castillo, S., Ortiz-Julien, A., Dequin, S., Gibson, B., 2017a. Exploring the potential of Saccharomyces eubayanus as a parent for new interspecies hybrid strains in winemaking. FEMS Yeast Res. 17, fox049.

Magalhães, F., Krogerus, K., Vidgren, V., Sandell, M., Gibson, B., 2017b. Improved cider fermentation performance and quality with newly generated Saccharomyces cerevisiae× Saccharomyces eubayanus hybrids. J. Ind. Microbiol. Biotechnol. 1–11.

Malherbe, S., Bauer, F.F., Du Toit, M., 2007. Understanding problem fermentations–A review.

Mancera, E., Bourgon, R., Brozzi, A., Huber, W., Steinmetz, L.M., 2008. High-resolution mapping of meiotic crossovers and noncrossovers in yeast. Nature 454, 479.

Mans, R., van Rossum, H.M., Wijsman, M., Backx, A., Kuijpers, N.G.A., van den Broek, M., Daran-Lapujade, P., Pronk, J.T., van Maris, A.J.A., Daran, J.-M.G., 2015. CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in Saccharomyces cerevisiae. FEMS Yeast Res. 15.

Marais, J., 1979. Effect of storage time and temperature on the formation of dimethyl sulphide and on white wine quality". Vitis 18, 254–260.

Marsit, S., Dequin, S., 2015. Diversity and adaptive evolution of Saccharomyces wine yeast: a review. FEMS Yeast Res. 15.

Martí-raga, M., Mas, A., Beltran, G., 2017. Genetic causes of phenotypic adaptation to the second fermentation sparkling wines in. G3 Genes| Genomes| Genet. 7, 399.

Marullo, P., Aigle, M., Bely, M., Masneuf-Pomarède, I., Durrens, P., Dubourdieu, D., Yvert, G., Masneuf-Pomarede, I., Durrens, P., Dubourdieu, D., Yvert, G., Masneuf-Pomarède, I., Durrens, P., Dubourdieu, D., Yvert, G., Masneuf-Pomarede, I., Durrens, P., Dubourdieu, D., Yvert, G., Masneuf-Pomarède, I., Durrens, P., Dubourdieu, D., Yvert, G., 2007. Single QTL mapping and nucleotide-level resolution of a physiologic trait in wine Saccharomyces cerevisiae strains. FEMS Yeast Res. 7, 941–952.

Mason, A.B., Dufour, J.P., 2000. Alcohol acetyltransferases and the significance of ester synthesis in yeast. Yeast 16, 1287–1298.

McGarvey, D.J., Croteau, R., 1995. Terpenoid metabolism. Plant Cell 7, 1015.

McGovern, P.E., 2009. Uncorking the past. Quest Wine, Beer, Other Alcohol. Beverages, Berkeley.

McGovern, P.E., Hartung, U., Badler, V.R., Glusker, D.L., Exner, L.J., 1997. The beginnings of winemaking and viniculture in the ancient Near East and Egypt. Expedition 39, 3–21.

McGovern, P.E., Zhang, J., Tang, J., Zhang, Z., Hall, G.R., Moreau, R.A., Nuñez, A., Butrym, E.D., Richards, M.P., Wang, C., 2004. Fermented beverages of pre-and proto-historic China. Proc. Natl. Acad. Sci. U. S. A. 101, 17593–17598.

Mendes-Ferreira, A., Barbosa, C., Falco, V., Leão, C., Mendes-Faia, A., 2009. The production of hydrogen sulphide and other aroma compounds by wine strains of Saccharomyces cerevisiae in synthetic media with different nitrogen concentrations. J. Ind. Microbiol. Biotechnol. 36, 571–583.

Menssen, R., Schweiggert, J., Schreiner, J., Kušević, D., Reuther, J., Braun, B., Wolf, D.H., 2012. Exploring the topology of the Gid complex, the E3 ubiquitin ligase involved in catabolite-induced degradation of gluconeogenic enzymes. J. Biol. Chem. 287, 25602–25614.

Mestres, M., Busto, O., Guasch, J., 2000. Analysis of organic sulfur compounds in wine aroma. J. Chromatogr. A 881, 569–581.

Michelmore, R.W., Paran, I., Kesseli, R. V, 1991. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc. Natl. Acad. Sci. 88, 9828–9832.

Miller, S.M., Magasanik, B., 1990. Role of NAD-linked glutamate dehydrogenase in nitrogen metabolism in Saccharomyces cerevisiae. J. Bacteriol. 172, 4927–4935.

Miran, S.G., Lawson, J.E., Reed, L.J., 1993. Characterization of PDH beta 1, the structural gene for the pyruvate dehydrogenase beta subunit from Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. 90, 1252–1256.

Mishina, M., Rogguenkamp, R., Schweizer, E., 1980. Yeast Mutants Defective in Acetyl-Coenzyme A Carboxylase and Biotin: Apocarboxylase Ligase. FEBS J. 111, 79–87.

Mitchell, A.P., Magasanik, B., 1984. Biochemical and physiological aspects of glutamine synthetase inactivation in Saccharomyces cerevisiae. J. Biol. Chem. 259, 12054–12062.

Molina, A.M., Swiegers, J.H., Varela, C., Pretorius, I.S., Agosin, E., 2007. Influence of wine fermentation temperature on the synthesis of yeast-derived volatile aroma compounds. Appl. Microbiol. Biotechnol. 77, 675–687.

Morata, A., Gómez-Cordovés, M.C., Colomo, B., Suárez, J.A., 2003. Pyruvic acid and acetaldehyde production by different strains of Saccharomyces cerevisiae: relationship with vitisin A and B formation in red wines. J. Agric. Food Chem. 51, 7402–7409.

Moreno-Arribas, M.V., Polo, M.C., 2009. Wine chemistry and biochemistry. Springer.

Mouret, J.-R., Camarasa, C., Angenieux, M., Aguera, E., Perez, M., Farines, V., Sablayrolles, J.M., 2014. Kinetic analysis and gas–liquid balances of the production of fermentative aromas during winemaking fermentations: effect of assimilable nitrogen and temperature. Food Res. Int. 62, 1–10.

Moye, W.S., Amuro, N., Rao, J.K., Zalkin, H., 1985. Nucleotide sequence of yeast GDH1 encoding nicotinamide adenine dinucleotide phosphate-dependent glutamate dehydrogenase. J. Biol. Chem. 260, 8502–8508.

Mudd, S.H., Datko, a H., 1990. The S-Methylmethionine Cycle in Lemna paucicostata. Plant Physiol. 93, 623–630.

Nagegowda, D.A., 2010. Plant volatile terpenoid metabolism: biosynthetic genes, transcriptional regulation and subcellular compartmentation. FEBS Lett. 584, 2965–2973.

Nidelet, T., Brial, P., Camarasa, C., Dequin, S., 2016. Diversity of flux distribution in central carbon metabolism of S. cerevisiae strains from diverse environments. Microb. Cell Fact. 15, 58.

Nieuwoudt, H.H., Prior, B.A., Pretorius, I.S., Bauer, F.F., 2002. Glycerol in South African table wines: an assessment of its relationship to wine quality.

Noble, A.C., Bursick, G.F., 1984. The contribution of glycerol to perceived viscosity and sweetness in white wine. Am. J. Enol. Vitic. 35, 110–112.

Noble, J., Sanchez, I., Blondin, B., 2015. Identification of new Saccharomyces cerevisiae variants of the MET2 and SKP2 genes controlling the sulfur assimilation pathway and the production of undesirable sulfur compounds during alcoholic fermentation. Microb. Cell Fact. 14, 68.

Nogami, S., Ohya, Y., Yvert, G., 2007. Genetic complexity and quantitative trait loci mapping of yeast morphological traits. PLoS Genet. 3, e31.

Nykänen, L., 1986. Formation and occurrence of flavor compounds in wine and distilled alcoholic beverages. Am. J. Enol. Vitic. 37, 84–96.

Nykänen, L., Suomalainen, H., 1983. Aroma of beer, wine and distilled alcoholic beverages. Springer Science & Business Media.

Oshita, K., Kubota, M., Uchida, M., Ono, M., 1995. Clarification of the relationship between fusel alcohol formation and amino acid assimilation by brewing yeast using 13C-labeled amino acid, in: Proceedings of Congress-European Brewery Convention. Oxford University Press, p. 387.

Österlund, T., Nookaew, I., Nielsen, J., 2012. Fifteen years of large scale metabolic modeling of yeast: Developments and impacts. Biotechnol. Adv. 30, 979–988.

Otero, J.M., Cimini, D., Patil, K.R., Poulsen, S.G., Olsson, L., Nielsen, J., 2013. Industrial systems biology of Saccharomyces cerevisiae enables novel succinic acid cell factory. PLoS One 8, e54144.

Pais, T.M., Foulquié-Moreno, M.R., Hubmann, G., Duitama, J., Swinnen, S., Goovaerts, A., Yang, Y., Dumortier, F., Thevelein, J.M., 2013. Comparative Polygenic Analysis of Maximal Ethanol Accumulation Capacity and Tolerance to High Ethanol Levels of Cell Proliferation in Yeast. PLoS Genet. 9.

Palsson, B., 2000. The challenges of in silico biology. Nat. Biotechnol. 18, 1147.

Park, S.H., Koh, S.S., Chun, J.H., Hwang, H.J., Kang, H.S., 1999. Nrg1 Is a Transcriptional Repressor for Glucose Repression of STA1 Gene Expression in Saccharomyces cerevisiae. Mol. Cell. Biol. 19, 2044–2050.

Patel, S., Shibamoto, T., 2003. Effect of 20 different yeast strains on the production of volatile components in Symphony wine. J. Food Compos. Anal. 16, 469–476.

Patil, K.R., Nielsen, J., 2005. Uncovering transcriptional regulation of metabolism by using metabolic network topology. Proc. Natl. Acad. Sci. U. S. A. 102, 2685–2689.

Pereira, C.F., 1988. The importance of metallic elements in wine. A literature survey. Zeitschrift für Leb. Und-forsch. A 186, 295–300.

Pérez-Torrado, R., Carrasco, P., Aranda, A., Gimeno-Alcañiz, J., Pérez-Ortín, J.E., Matallana, E., del Olmo, M., 2002. Study of the first hours of microvinification by the use of osmotic stress-response genes as probes. Syst. Appl. Microbiol. 25, 153–161.

Perpète, P., Duthoit, O., De Maeyer, S., Imray, L., Lawton, A.I., Stavropoulos, K.E., Gitonga, V.W., Hewlins, M.J.E., Richard Dickinson, J., 2005. Methionine catabolism in Saccharomyces cerevisiae. FEMS Yeast Res. 6, 48–56.

Pigeau, G.M., Inglis, D.L., 2005. Upregulation of ALD3 and GPD1 in Saccharomyces cerevisiae during Icewine fermentation. J. Appl. Microbiol. 99, 112–125.

Polášková, P., Herszage, J., Ebeler, S.E., 2008. Wine flavor: chemistry in a glass. Chem. Soc. Rev. 37, 2478–2489.

Postma, E., Verduyn, C., Scheffers, W.A., Van Dijken, J.P., 1989. Enzymic analysis of the crabtree effect in glucose-limited chemostat cultures of Saccharomyces cerevisiae. Appl. Environ. Microbiol. 55, 468–477.

Pretorius, I.S., Høj, P.B., 2005. Grape and wine biotechnology: challenges, opportunities and potential benefits. Aust. J. grape wine Res. 11, 83–108.

Pretorius, I.S., Van der Westhuizen, T.J., Augustyn, O.P.H., 1999. Yeast biodiversity in vineyards and wineries and its importance to the South African wine industry. A review. South African J. Enol. Vitic. 20, 61–70.

Pronk, J.T., Flikweert, M.T., van der Zanden, L., Janssen, W.M., Steensma, H.Y., van Dijken, J.P., 1996a. Pyruvate decarboxylase: an indispensable enzyme for growth of Saccharomyces cerevisiae on glucose. Yeast, 12 (3), 247-257.

Pronk, J.T., Yde Steensma, H., van Dijken, J.P., 1996b. Pyruvate metabolism in Saccharomyces cerevisiae. Yeast 12, 1607–1633.

Quirós, M., Martínez-Moreno, R., Albiol, J., Morales, P., Vázquez-Lima, F., Barreiro-Vázquez, A., Ferrer, P., Gonzalez, R., 2013. Metabolic flux analysis during the exponential growth phase of Saccharomyces cerevisiae in wine fermentations. PLoS One 8, e71909.

Rankine, B.C., Pocock, K.F., 1969. Influence of yeast strain on binding of sulphur dioxide in wines, and on its formation during fermentation. J. Sci. Food Agric. 20, 104–109.

Rauhut, D., 2009. Usage and formation of sulphur compounds, in: Biology of Microorganisms on Grapes, in Must and in Wine. Springer, pp. 181–207.

Regelmann, J., Schüle, T., Josupeit, F.S., Horak, J., Rose, M., Entian, K.-D., Thumm, M., Wolf, D.H., 2003. Catabolite degradation of fructose-1, 6-bisphosphatase in the yeast Saccharomyces cerevisiae: a genome-wide screen identifies eight novel GID genes and indicates the existence of two degradation pathways. Mol. Biol. Cell 14, 1652–1663.

Regenberg, B., Düring-Olsen, L., Kielland-Brandt, M.C., Holmberg, S., 1999. Substrate specificity and gene expression of the amino-acid permeases in Saccharomyces cerevisiae. Curr. Genet. 36, 317–328.

Ribéreau-Gayon, P., Boidron, J.N., Terrier, A., 1975. Aroma of Muscat grape varieties. J. Agric. Food Chem. 23, 1042–1047.

Rine, J., Herskowitz, I., 1987. Four genes responsible for a position effect on expression from HML and HMR in Saccharomyces cerevisiae. Genetics 116, 9–22.

Robinson, A.L., Boss, P.K., Heymann, H., Solomon, P.S., Trengove, R.D., 2011. Influence of Yeast Strain, Canopy Management, and Site on the Volatile Composition and Sensory Attributes of Cabernet Sauvignon Wines from Western Australia. J. Agric. Food Chem. 59, 3273–3284.

Roland, A., Schneider, R., Le Guernevé, C., Razungles, A., Cavelier, F., 2010. Identification and quantification by LC–MS/MS of a new precursor of 3-mercaptohexan-1-ol (3MH) using stable isotope dilution assay: Elements for understanding the 3MH production in wine. Food Chem. 121, 847–855.

Rollero, S., Bloem, A., Camarasa, C., Sanchez, I., Ortiz-Julien, A., Sablayrolles, J.-M., Dequin, S., Mouret, J.-R., 2015. Combined effects of nutrients and temperature on the production of fermentative aromas by Saccharomyces cerevisiae during wine fermentation. Appl. Microbiol. Biotechnol. 99, 2291.

Romano, P., Fiore, C., Paraggio, M., Caruso, M., Capece, A., 2003. Function of yeast species and strains in wine flavour. Int. J. Food Microbiol. 86, 169–180.

Rossignol, T., Dulau, L., Julien, A., Blondin, B., 2003. Genome-wide monitoring of wine yeast gene expression during alcoholic fermentation. Yeast 20, 1369–1385.

Rossouw, D., Jacobson, D., Bauer, F.F., 2012. Transcriptional Regulation and the Diversification of Metabolism in Wine Yeast Strains. Genetics 190, 251 LP-261.

Rossouw, D., Næs, T., Bauer, F.F., 2008. Linking gene regulation and the exo-metabolome: a comparative transcriptomics approach to identify genes that impact on the production of volatile aroma compounds in yeast. BMC Genomics 9, 530.

Rouillon, A., Surdin-kerjan, Y., Thomas, D., 1999. Transport of Sulfonium Compounds 274, 28096–28105.

Sablayrolles, J.M., 2009. Control of alcoholic fermentation in winemaking: Current situation

and prospect. Food Res. Int. 42, 418–424.
Saerens, S.M.G., Delvaux, F., Verstrepen, K.J., Van Dijck, P., Thevelein, J.M., Delvaux, F.R., 2008. Parameters affecting ethyl ester production by Saccharomyces cerevisiae during fermentation. Appl. Environ. Microbiol. 74, 454–461.
Saerens, S.M.G., Delvaux, F.R., Verstrepen, K.J., Thevelein, J.M., 2010. Production and biological function of volatile esters in Saccharomyces cerevisiae. Microb. Biotechnol. 3, 165–177.
Saerens, S.M.G., Verstrepen, K.J., Van Laere, S.D.M., Voet, A.R.D., Van Dijck, P., Delvaux, F.R., Thevelein, J.M., 2006. The Saccharomyces cerevisiae EHT1 and EEB1 genes encode novel enzymes with medium-chain fatty acid ethyl ester synthesis and hydrolysis capacity. J. Biol. Chem. 281, 4446–4456.
Saint-Prix, F., Bönquist, L., Dequin, S., 2004. Functional analysis of the ALD gene family of Saccharomyces cerevisiae during anaerobic growth on glucose: the NADP+-dependent Ald6p and Ald5p isoforms play a major role in acetate formation. Microbiology 150, 2209–2220.
Salinas, F., Cubillos, F. a., Soto, D., Garcia, V., Bergström, A., Warringer, J., Ganga, M.A., Louis, E.J., Liti, G., Martinez, C., 2012. The genetic basis of natural variation in oenological traits in Saccharomyces cerevisiae. PLoS One 7, e49640.
Salmon, J.M., 1989. Effect of sugar transport inactivation in Saccharomyces cerevisiae on sluggish and stuck enological fermentations. Appl. Environ. Microbiol. 55, 953–958.
San-Juan, F., Ferreira, V., Cacho, J., Escudero, A., 2011. Quality and aromatic sensory descriptors (mainly fresh and dry fruit character) of Spanish red wines can be predicted from their aroma-active chemical composition. J. Agric. Food Chem. 59, 7916–7924.
Santt, O., Pfirrmann, T., Braun, B., Juretschke, J., Kimmig, P., Scheel, H., Hofmann, K., Thumm, M., Wolf, D.H., 2008. The yeast GID complex, a novel ubiquitin ligase (E3) involved in the regulation of carbohydrate metabolism. Mol. Biol. Cell 19, 3323–3333.
Sax, K., 1923. The association of size differences with seed-coat pattern and pigmentation in Phaseolus vulgaris. Genetics 8, 552.
Scanes, K.T., Hohrnann, S., Prior, B.A., 1998. Glycerol production by the yeast Saccharomyces cerevisiae and its relevance to wine: a review. South African J. Enol. Vitic. 19, 17–24.
Schreier, P., Jennings, W.G., 1979. Flavor composition of wines: a review. Crit. Rev. Food Sci. Nutr. 12, 59–111.
Schreve, J.L., Garrett, J.M., 2004. Yeast Agp2p and Agp3p function as amino acid permeases in poor nutrient conditions. Biochem. Biophys. Res. Commun. 313, 745–751.
Schreve, J.L., Sin, J.K., Garrett, J.M., 1998. The Saccharomyces cerevisiae YCC5 (YCL025c) gene encodes an amino acid permease, Agp1, which transports asparagine and glutamine. J. Bacteriol. 180, 2556–2559.
Schweizer, M., Roberts, L.M., Höltke, H.-J., Takabayashi, K., Höllerer, E., Hoffmann, B., Müller, G., Köttig, H., Schweizer, E., 1986. The pentafunctional FAS1 gene of yeast: its nucleotide sequence and order of the catalytic domains. Mol. Gen. Genet. MGG 203, 479–486.
Schwender, J., Seemann, M., Lichtenthaler, H.K., Rohmer, M., 1996. Biosynthesis of isoprenoids (carotenoids, sterols, prenyl side-chains of chlorophylls and plastoquinone) via a novel pyruvate/glyceraldehyde 3-phosphate non-mevalonate pathway in the green alga Scenedesmus obliquus. Biochem. J. 316, 73–80.
Segurel, M., 2005. Contribution des précurseurs glycosidiques et du sulfure de diméthyle des baies de Vitis vinifera L. Cv Grenache noir et Syrah à l'arôme des vins de la Vallée du Rhône.

Segurel, M. a., Razungles, A.J., Riou, C., Salles, M., Baumes, R.L., 2004. Contribution of dimethyl sulfide to the aroma of syrah and grenache noir wines and estimation of its potential in grapes of these varieties. J. Agric. Food Chem. 52, 7084–7093.

Segurel, M. a., Razungles, A.J., Riou, C., Trigueiro, M.G.L., Baumes, R.L., 2005. Ability of possible DMS precursors to release DMS during wine aging and in the conditions of heat-alkaline treatment. J. Agric. Food Chem. 53, 2637–2645.

Shieh, H.-L., Chiang, H.-L., 1998. In vitro reconstitution of glucose-induced targeting of fructose-1, 6-bisphosphatase into the vacuole in semi-intact yeast cells. J. Biol. Chem. 273, 3381–3387.

Shimazu, Y., Watanabe, M., 1981. Effects of yeast strains and environmental conditions on formation of organic acids in must during fermentation. J. Ferment. Technol. 59, 27–32.

Simpson, R.F., 1979. Aroma composition of bottle aged white wine. Vitis-Berichte ueber Rebenforsch. mit Dokumentation der Weinbauforsch. (Germany, FR).

Simpson, R.F., Miller, G.C., 1983. Aroma composition of aged Riesling wine. Vitis 22, 51–63.

Singleton, V.L., Esau, P., 1969. Phenolic substances in grapes and wine, and their significance. Adv. Food Res. Suppl. 1, 1–261.

Sinha, H., Nicholson, B.P., Steinmetz, L.M., McCusker, J.H., 2006. Complex genetic interactions in a quantitative trait locus. PLoS Genet. 2, e13.

Snowdon, C., Hlynialuk, C., Van Der Merwe, G., 2007. Components of the Vid30c are needed for the rapamycin-induced degradation of the high-affinity hexose transporter Hxt7p in Saccharomyces cerevisiae. FEMS Yeast Res. 8, 204–216.

Snowdon, C., Van der Merwe, G., 2012. Regulation of Hxt3 and Hxt7 turnover converges on the Vid30 complex and requires inactivation of the Ras/cAMP/PKA pathway in Saccharomyces cerevisiae. PLoS One 7, e50458.

Soares, R.D., Welke, J.E., Nicolli, K.P., Zanus, M., Caram??o, E.B., Manfroi, V., Zini, C.A., 2015. Monitoring the evolution of volatile compounds using gas chromatography during the stages of production of Moscatel sparkling wine. Food Chem. 183, 291–304.

Spedding, D.J., Raut, P., 1982. The influence of dimethyl sulphide and carbon disulphide in the bouquet of wines. Vitis 21, 240–246.

Spor, A., Nidelet, T., Simon, J., Bourgais, A., de Vienne, D., Sicard, D., 2009. Niche-driven evolution of metabolic and life-history strategies in natural and domesticated populations of Saccharomyces cerevisiae. BMC Evol. Biol. 9, 296.

Starai, V.J., Takahashi, H., Boeke, J.D., Escalante-Semerena, J.C., 2003. Short-chain fatty acid activation by acyl-coenzyme A synthetases requires SIR2 protein function in Salmonella enterica and Saccharomyces cerevisiae. Genetics 163, 545–555.

Stashenko, H., Macku, C., Shibamoto, T., 1992. Monitoring volatile chemicals formed from must during yeast fermentation. J. Agric. Food Chem. 40, 2257–2259.

Steinkraus, K.H., 1994. Nutritional significance of fermented foods. Food Res. Int. 27, 259–267.

Steinmetz, L.M., Sinha, H., Richards, D.R., Spiegelman, J.I., Oefner, P.J., McCusker, J.H., Davis, R.W., 2002. Dissecting the architecture of a quantitative trait locus in yeast. Nature 416, 326–330.

Steyer, D., Ambroset, C., Brion, C., Claudel, P., Delobel, P., Sanchez, I., Erny, C., Blondin, B., Karst, F., Legras, J.-L., 2012. QTL mapping of the production of wine aroma compounds by yeast. BMC Genomics 13, 573.

Steyer, D., Erny, C., Claudel, P., Riveill, G., Karst, F., Legras, J.-L.L., 2013. Genetic analysis of geraniol metabolism during fermentation. Food Microbiol. 33, 228–234.

Strauss, C.R., Wilson, B., Gooley, P.R., Williams, P.J., 1986. Role of monoterpenes in grape and

wine flavor. ACS Publications.
Strope, P.K., Skelly, D.A., Kozmin, S.G., Mahadevan, G., Stone, E.A., Magwene, P.M., Dietrich, F.S., McCusker, J.H., 2015. The 100-genomes strains, an S. cerevisiae resource that illuminates its natural phenotypic and genotypic variation and emergence as an opportunistic pathogen. Genome Res. 25, 762–774.
Styger, G., Prior, B., Bauer, F.F., 2011. Wine flavor and aroma. J. Ind. Microbiol. Biotechnol. 38, 1145.
Swiegers, J.H., Bartowsky, E.J., Henschke, P. a, Pretorius, I.S., 2005. Yeast and bacterial modulation of wine aroma and flavour. Aust. J. grape wine Res. 11, 139–173.
Swiegers, J.H., Pretorius, I.S., 2007. Modulation of volatile sulfur compounds by wine yeast. Appl. Microbiol. Biotechnol. 74, 954–960.
Swinnen, S., Ho, P.-W., Klein, M., Nevoigt, E., 2016. Genetic determinants for enhanced glycerol growth of Saccharomyces cerevisiae. Metab. Eng. 36, 68–79.
Swinnen, S., Thevelein, J.M., Nevoigt, E., 2012. Genetic mapping of quantitative phenotypic traits in Saccharomyces cerevisiae. FEMS Yeast Res. 12, 215–227.
Sychrova, H., Chevallier, M.R., 1994. Yeast sequencing reports. APL1, a yeast gene encoding a putative permease for basic amino acids. Yeast 10, 653–657.
Takoi, K., Itoga, Y., Koie, K., Kosugi, T., Katayama, Y., Nakayama, Y., Watari, J., 2010a. The Contribution of Geraniol Metabolism to the Citrus Flavour of Beer : Synergy of Geraniol and β-Citronellol Under Coexistence with Excess Linalool. J. Inst. Brew. 116, 251–260.
Takoi, K., Koie, K., Itoga, Y., Katayama, Y., Shimase, M., Nakayama, Y., Watari, J., 2010b. Biotransformation of hop-derived monoterpene alcohols by lager yeast and their contribution to the flavor of hopped beer. J. Agric. Food Chem. 58, 5050–5058.
Tao, Y.-S., Li, H., 2009. Active volatiles of cabernet sauvignon wine from Changli County. Health (Irvine. Calif). 1, 176.
Teixeira, M.C., Monteiro, P.T., Guerreiro, J.F., Gonçalves, J.P., Mira, N.P., dos Santos, S.C., Cabrito, T.R., Palma, M., Costa, C., Francisco, A.P., 2013. The YEASTRACT database: an upgraded information system for the analysis of gene and genomic transcription regulation in Saccharomyces cerevisiae. Nucleic Acids Res. 42, D161–D166.
Thomas, D., Becker, A., Surdin-Kerjan, Y., 2000. Reverse methionine biosynthesis from S-adenosylmethionine in eukaryotic cells. J. Biol. Chem. 275, 40718–40724.
Thomas, D., Surdin-Kerjan, Y., 1997. Metabolism of sulfur amino acids in Saccharomyces cerevisiae. Microbiol. Mol. Biol. Rev. 61, 503–532.
Tilloy, V., Cadière, A., Ehsani, M., Dequin, S., 2015. Reducing alcohol levels in wines through rational and evolutionary engineering of Saccharomyces cerevisiae. Int. J. Food Microbiol. 213, 49–58.
Tilloy, V., Ortiz-Julien, A., Dequin, S., 2014. Reduction of ethanol yield and improvement of glycerol formation by adaptive evolution of the wine yeast Saccharomyces cerevisiae under hyperosmotic conditions. Appl. Environ. Microbiol. 80, 2623–2632.
Tominaga, T., Niclass, Y., Frérot, E., Dubourdieu, D., 2006. Stereoisomeric distribution of 3-mercaptohexan-1-ol and 3-mercaptohexyl acetate in dry and sweet white wines made from Vitis vinifera (Var. Sauvignon Blanc and Semillon). J. Agric. Food Chem. 54, 7251–7255.
Torija, M.J., Beltran, G., Novo, M., Poblet, M., Guillamón, J.M., Mas, A., Rozes, N., 2003. Effects of fermentation temperature and Saccharomyces species on the cell fatty acid composition and presence of volatile compounds in wine. Int. J. Food Microbiol. 85, 127–136.

Torrea, D., Fraile, P., Garde, T., Ancí, C., 2003. Production of volatile compounds in the fermentation of chardonnay musts inoculated with two strains of Saccharomyces cerevisiae with different nitrogen demands. Food Control 14, 565–571.

Trossat, C., Rathinasabapathi, B., Weretilnyk, E.A., Shen, T.-L., Huang, Z.-H., Gage, D.A., Hanson, A.D., 1998. Salinity promotes accumulation of 3-dimethylsulfoniopropionate and its precursor S-methylmethionine in chloroplasts. Plant Physiol. 116, 165–171.

Tulio, A.Z., Yamanaka, H., Ueda, Y., Imahori, Y., 2002. Formation of methanethiol and dimethyl disulfide in crushed tissues of broccoli florets and their inhibition by freeze− thawing. J. Agric. Food Chem. 50, 1502–1507.

Ulrich, D., Hoberg, E., Bittner, T., Engewald, W., Meilchen, K., 2001. Contribution of volatile compounds to the flavor of cooked asparagus. Eur. Food Res. Technol. 213, 200–204.

Umbarger, H.E., 1978. Amino acid biosynthesis and its regulation. Annu. Rev. Biochem. 47, 533–606.

Valdar, W.S.J., Flint, J., Mott, R., 2003. QTL fine-mapping with recombinant-inbred heterogeneous stocks and in vitro heterogeneous stocks. Mamm. genome 14, 830–838.

Valero, E., Moyano, L., Millan, M.C., Medina, M., Ortega, J.M., 2002. Higher alcohols and esters production by Saccharomyces cerevisiae. Influence of the initial oxygenation of the grape must. Food Chem. 78, 57–61.

van der Merwe, G.K., Cooper, T.G., van Vuuren, H.J.J., 2001. Ammonia regulates VID30 expression and Vid30p function shifts nitrogen metabolism toward glutamate formation especially when Saccharomyces cerevisiae is grown in low concentrations of ammonia. J. Biol. Chem. 276, 28659–28666.

van Dijken, J.P., Scheffers, W.A., 1986. Redox balances in the metabolism of sugars by yeasts. FEMS Microbiol. Lett. 32, 199–224.

Van Gulik, W.M., Heijnen, J.J., 1995. A metabolic network stoichiometry analysis of microbial growth and product formation. Biotechnol. Bioeng. 48, 681–698.

Van Roermund, C.W., Elgersma, Y., Singh, N., Wanders, R.J., Tabak, H.F., 1995. The membrane of peroxisomes in Saccharomyces cerevisiae is impermeable to NAD (H) and acetyl-CoA under in vivo conditions. EMBO J. 14, 3480.

Van Roermund, C.W.T., Hettema, E.H., Van Den Berg, M., Tabak, H.F., Wanders, R.J.A., 1999. Molecular characterization of carnitine-dependent transport of acetyl-CoA from peroxisomes to mitochondria in Saccharomyces cerevisiae and identification of a plasma membrane carnitine transporter, Agp2p. EMBO J. 18, 5843–5852.

Varela, C., Torrea, D., Schmidt, S.A., Ancin-Azpilicueta, C., Henschke, P.A., 2012. Effect of oxygen and lipid supplementation on the volatile composition of chemically defined medium and Chardonnay wine fermented with Saccharomyces cerevisiae. Food Chem. 135, 2863–2871.

Vargas, F.A., Pizarro, F., Pérez-Correa, J.R., Agosin, E., 2011. Expanding a dynamic flux balance model of yeast fermentation to genome-scale. BMC Syst. Biol. 5, 75.

Varma, A., Palsson, B.O., 1994. Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type Escherichia coli W3110. Appl. Environ. Microbiol. 60, 3724–3731.

Vaudano, E., Moruno, E.G., Di Stefano, R., 2004. Modulation of Geraniol Metabolism During Alcohol Fermentation. J. Inst. Brew. 110, 213–219.

Velagapudi, V.R., Wittmann, C., Schneider, K., Heinzle, E., 2007. Metabolic flux screening of Saccharomyces cerevisiae single knockout strains on glucose and galactose supports elucidation of gene function. J. Biotechnol. 132, 395–404.

Verstrepen, K.J., Van Laere, S.D.M., Vanderhaegen, B.M.P., Derdelinckx, G., Dufour, J.-P., Pretorius, I.S., Winderickx, J., Thevelein, J.M., Delvaux, F.R., 2003. Expression levels of the yeast alcohol acetyltransferase genes ATF1, Lg-ATF1, and ATF2 control the formation of a broad range of volatile esters. Appl. Environ. Microbiol. 69, 5228–5237.

Vilanova, M., Genisheva, Z., Masa, A., Oliveira, J.M., 2010. Correlation between volatile composition and sensory properties in Spanish Albariño wines. Microchem. J. 95, 240–246.

Vilanova, M., Ugliano, M., Varela, C., Siebert, T., Pretorius, I.S., Henschke, P.A., 2007. Assimilable nitrogen utilisation and production of volatile and non-volatile compounds in chemically defined medium by Saccharomyces cerevisiae wine yeasts. Appl. Microbiol. Biotechnol. 77, 145–157.

Voirin, S.G., Baumes, R.L., Bitteur, S.M., Gunata, Z.Y., Bayonove, C.L., 1990. Novel monoterpene disaccharide glycosides of Vitis vinifera grapes. J. Agric. Food Chem. 38, 1373–1378.

Vyas, V.K., Berkey, C.D., Miyao, T., Carlson, M., 2005. Repressors Nrg1 and Nrg2 regulate a set of stress-responsive genes in Saccharomyces cerevisiae. Eukaryot. Cell 4, 1882–1891.

Warringer, J., Liti, G., Blomberg, A., 2017. Yeast reciprocal hemizygosity to confirm the causality of a quantitative trait loci-associated gene. Cold Spring Harb. Protoc. 2017, pdb-prot089078.

Warringer, J., Zörgö, E., Cubillos, F.A., Zia, A., Gjuvsland, A., Simpson, J.T., Forsmark, A., Durbin, R., Omholt, S.W., Louis, E.J., 2011. Trait variation in yeast is defined by population history. PLoS Genet. 7, e1002111.

Wenz, P., Schwank, S., Hoja, U., Schüller, H.-J., 2001. A downstream regulatory element located within the coding sequence mediates autoregulated expression of the yeast fatty acid synthase gene FAS2 by the FAS1 gene product. Nucleic Acids Res. 29, 4625–4632.

Whiting, G.C., 1976. Organic acid metabolism of yeasts during fermentation of alcoholic beverages—a review. J. Inst. Brew. 82, 84–92.

Williams, K.M., Liu, P., Fay, J.C., 2015. Evolution of ecological dominance of yeast species in high-sugar environments. Evolution (N. Y). 69, 2079–2093.

Wright, L.D., 1961. Biosynthesis of isoprenoid compounds. Annu. Rev. Biochem. 30, 525–548.

Yunoki, K., Yasui, Y., Hirose, S., Ohnishi, M., 2005. Fatty acids in must prepared from 11 grapes grown in Japan: comparison with wine and effect on fatty acid ethyl ester formation. Lipids 40, 361–367.

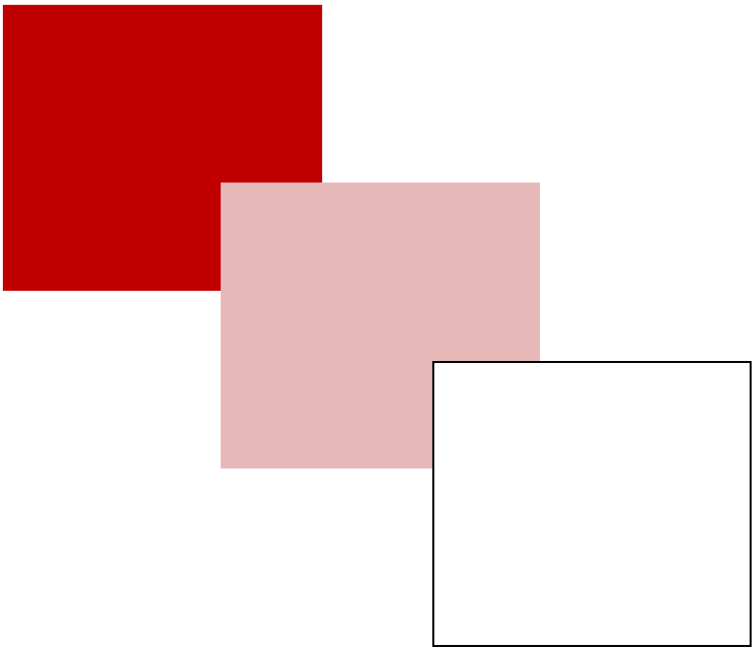
Zastrow, C.R., Hollatz, C., De Araujo, P.S., Stambuk, B.U., 2001. Maltotriose fermentation by Saccharomyces cerevisiae. J. Ind. Microbiol. Biotechnol. 27, 34–38.

Zea, L., Moreno, J., Ortega, J.M., Medina, M., 1995. Content of free terpenic compounds in cells and musts during vinification with three Saccharomyces cerevisiae races. J. Agric. Food Chem. 43, 1110–1114.

Zeng, Z.-B., 1993. Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. Proc. Natl. Acad. Sci. 90, 10972–10976.

Zhou, H., Winston, F., 2001. NRG1 is required for glucose repression of the SUC2 and GAL genes of Saccharomyces cerevisiae. BMC Genet. 2, 5.

Annex



Additional experiments

During the experimental work of this thesis, additional results were obtained related to yeast alcoholic fermentation or fermentative aroma. As the haploid parent strains of this study showed spontaneous mating type switch and self-diploidization, the influence of differences in ploidy on several wine fermentation traits was assessed for these strains. Furthermore, a secondment was done in the course of the thesis at the Heineken brewery, The Netherlands, under supervision of Dr. Jan Maarten Geertman. This opportunity was used to perform a study in cooperation with Frederico Magalhães, PhD student at VTT Finnland, to evaluate the suitability of newly generated *Saccharomyces pastorianus* strains for beer brewing.

1 Effect of yeast ploidy on wine fermentation traits

Saccharomyces cerevisiae can survive and grow in two different genomic states, haploid and diploid. While the majority of isolated *S. cerevisiae* wine strains is diploid (Legras et al., 2007), research on wine yeast phenotypes is mainly performed on haploid spores. This raises the question of the impact of wine yeast ploidy on aroma formation and fermentation kinetics. To evaluate this context, diploid strains of parental spores that had formed through self-diploidization after spontaneous mating type switch were isolated. This behavior could be observed in both parent strains, although the *HO* gene had been deleted in beforehand. Ploidy of strains was verified by mating type PCR on the MAT-locus. Haploid and diploid strains were phenotyped in triplicate during small scale wine fermentation in 280 mL of synthetic grape must. After 80% of fermentation, samples were taken and the concentrations of extracellular main metabolites and volatile metabolites were determined by HPLC and GC-MS.

Principle component analyses (PCAs) were performed on fermentation parameters, sugar consumption and main metabolite yields (Figure 37), as well as on the formation of volatile metabolites (Figure 38).

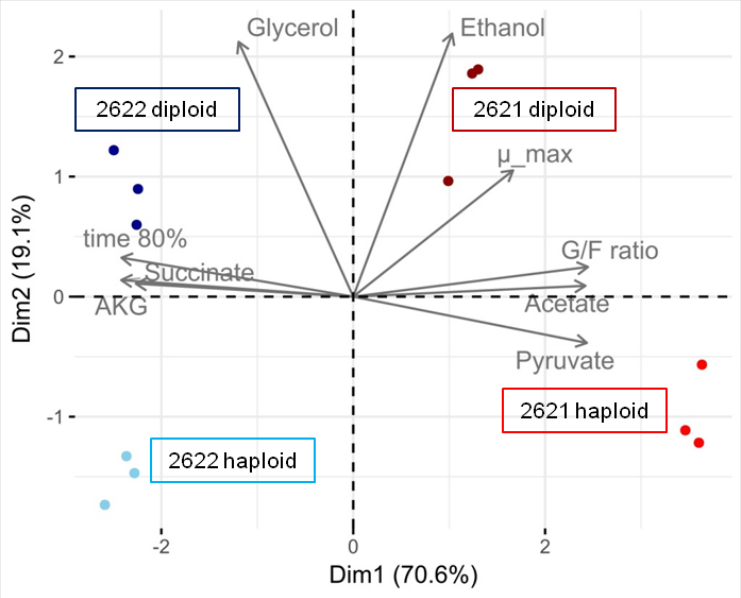


Figure 37: PCA of extracellular main metabolite yields, fermentation parameters and ratio of left sugars after 80% of fermentation, dependent on strain and ploidy.

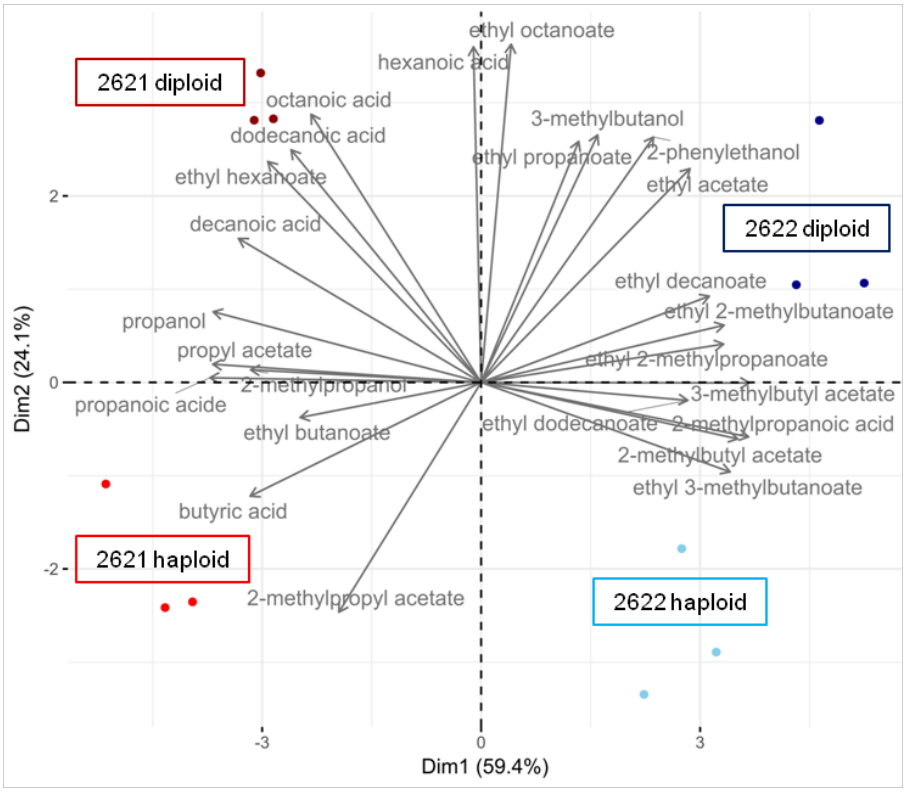


Figure 38: PCA of volatile compound formation (selection) after 80% of fermentation, dependent on strain and ploidy.

With the first two dimensions explaining 89.7% of variance for main metabolite production and 83.5% of variance for fermentative aroma formation, both PCAs well describe the differences between strains. Although differences between both parent strains have a

stronger impact on the determined traits, a distinction between haploid and diploid cells can be made. This notably concerns the yields of glycerol and ethanol, the maximum fermentation rate (μ_max), the formation of some higher alcohols and the formation of medium chain fatty acid and derived volatiles.

The following table breaks down the changes in all observed traits when haploid parental strains are compared to their diploid counterpart (Table 30).

Table 30: Comparison of haploid and diploid parental strains in terms of fermentation parameters, main metabolite yields, ratio of remaining glucose/fructose (G/F ratio) and formation of volatile metabolites after 80% of fermentation.

Phenotype	Trait	haploid/diploid MTF2621 [%]	haploid/diploid MTF2622 [%]
Fermentation parameters	μ max	3.9	-9.1
	$t_{\mu\ max}$	0.9	5.0
	$t_{80\%}$	-16.6	-10.7
Main metabolite yields	Acetate	12.7	37.5
	AKG	-33.0	15.2
	Ethanol	-0.7	-0.6
	Glycerol	-8.0	-4.3
	Pyruvate	34.3	0.8
	Succinate	-13.0	-9.5
Sugar consumption	G/F ratio	13.0	12.3
Acetate esters	2-methylbutyl acetate	15.2	-7.6
	2-methylpropyl acetate	59.0	21.2
	3-methylbutyl acetate	-3.9	-14.9
	2-phenylethyl acetate	2.3	-20.4
	ethyl acetate	-25.1	-25.2
	propyl acetate	25.2	-0.6
Ethyl esters	diethyl succinate	-16.1	-23.3
	ethyl 2-methylbutanoate	-31.8	-26.3
	ethyl 2-methylpropanoate	-11.5	-26.4
	ethyl 3-methylbutanoate	189.1	-17.0
	ethyl 3-methylthiopropanoate	-0.8	4.8
	ethyl butanoate	4.8	8.4
	ethyl decanoate	-4.4	-24.3
	ethyl dodecanoate	10.8	-11.9
	ethyl hexanoate	-18.6	-20.7
	ethyl lactate	17.9	-11.6
	ethyl octanoate	-23.3	-25.6
	ethyl propanoate	-23.2	-37.6
higher alcohols	2-methylbutanol	55.8	-29.2
	2-methylpropanol	3.3	0.8
	2-phenylethanol	-16.4	-20.4
	3-methylbutanol	-9.0	-3.2
	hexanol	-3.4	3.7
	methionol	11.2	-17.7
	propanol	4.1	-4.9

fusel acids	2-methylbutanoic acid	-28.8	-32.1
	3-methylbutanoic acid	-24.3	-18.2
	2-methylpropanoic acid	0.7	-4.0
	propanoic acid	13.5	9.8
	valeric acid	3.2	-9.8
Medium chain fatty acids	butanoic acid	36.3	9.4
	hexanoic acid	-24.0	-15.4
	octanoic acid	-15.3	-13.7
	decanoic acid	-4.4	-17.1
	dodecanoic acid	-15.0	-32.1

It can be seen that most traits are strongly influenced by the ploidy of the strains. Fermentation time, succinate yield, ethanol yield and the formation of ethyl acetate and most medium chain fatty acids are reduced, while the production of some desirable flavor compounds such as 2-methylpropyl acetate and ethyl butanoate is increased for fermentations of haploid parents. Therefore, it seems to be of advantage to use haploid strains for wine fermentation. However, for both parent strains only 62% of all observed traits follow the same pattern by changes in ploidy, which indicates significant strain dependence. Nevertheless, it is evident that ploidy influences the strain's behavior in terms of fermentation kinetics and metabolite formation, which has to be taken in consideration when evaluating the characteristics of wine strains.

2 Evaluation of newly generated *Saccharomyces eubayanus* x *Saccharomyces cerevisiae* (wine) hybrids for the fermentation of beer

2.1 Introduction

Lager strains (*S. pastorianus*) occurred probably during beer brewing through hybridization of an *S. cerevisiae* ale strain and another parental yeast species that was long time unknown (Dunn and Sherlock, 2008). Only recently, the second parent has been discovered and characterized as *S. eubayanus* (Libkind et al., 2011). The hybridizytion of new *S. pastorianus* strains is currently being performed by many research groups to extend the selection of lager beer strains for brewing.

The use of non-ale *S. cerevisiae* parent strains, like wine strains, could provide new aromatic characteristics to *S. pastorianus* beer strains. In addition, this approach can furthermore be used to create cold tolerant strains for wine fermentation, since the use of lower fermentation temperatures preserves volatiles produced by the yeast or present in the must (Magalhães et al., 2017a).

In this experiment, it was evaluated if an *S. cerevisiae* wine strain parent can impart new aromatic properties to an *S. pastorianus* hybrid strain used for the production of beer. Hybrid strains were generated by mating of the *S. eubayanus* isolate C902 and a natural lysine

auxotrophic mutant of the *S. cerevisiae* wine strain 59A, a haploid derivate of the widely used strain EC1118 (Krogerus et al., 2015). Six *S. pastorianus* hybrids were selected during fermentation at 37 °C in the absence of lysine supplementation. These strains have already been assessed in cider (Magalhães et al., 2017b) and wine making conditions (Magalhães et al., 2017a).

2.2 Materials and Methods

The assessment of the generated hybrids in beer making conditions was carried out at the Heineken brewery in Zoeterwoude, The Netherlands. The strains were fermented anaerobically at 22 °C in 400 mL of 12 °P Heineken wort with the addition of 9.2 µM ZnSO4 and 0.5 mL/L antifoam emulsion. The wort was inoculated to a cell density of 1.0×10^7 cells/mL and the fermentation was followed with the Alcoholic Fermentation Monitor (AFM, Applikon Biotechnology) by measuring the exhaust of produced CO_2. Samples of 40 mL were taken after 18 h, 40 h and at the end of fermentation. The samples were submitted to the Heineken Sample Service Center and were analyzed for residual sugars, fermentative aromas, sulfur dioxide and 4-vinylguaiacol.

2.3 Results and Discussion

Both parents and the hybrid strains were fermented in 12°P Heineken wort at 22 °C. The corresponding fermentation kinetics are presented in Figure 39 and Figure 40.

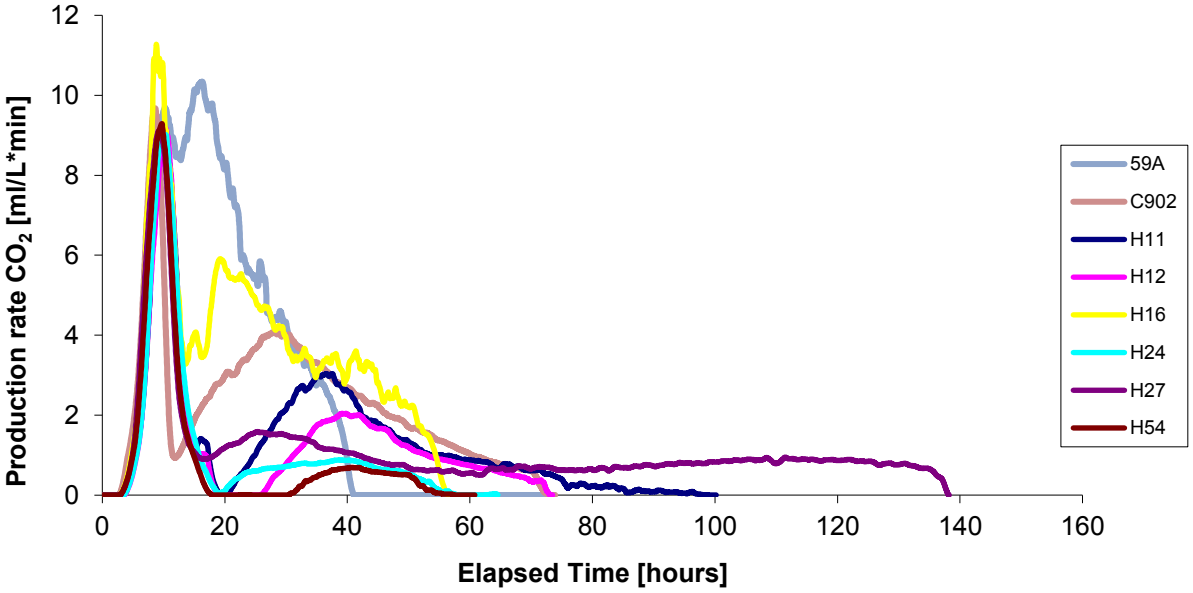


Figure 39: CO_2 production rate of the fermented parent strains C902 and 59A as well as the hybrids H11 – H54 in 12 °P Heineken wort at 22 °C.

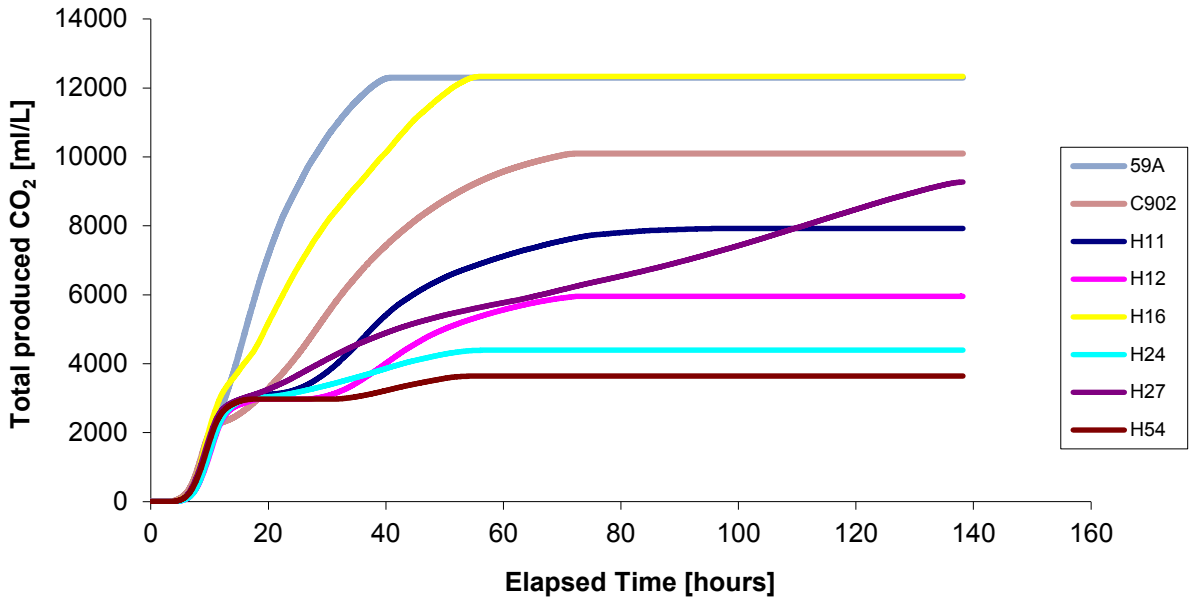


Figure 40: Total CO_2 production of the fermented parent strains C902 and 59A as well as the hybrids H11 – H54 in 12 °P Heineken wort at 22 °C.

10 h after the beginning of fermentation, a drop in CO_2 production rate can be observed for all strains. Four of the hybrid strains (H11, H12, H24 and H54) even show a second lag phase up to 14 h. The *S. eubayanus* parent C902, as well as all hybrid strains except H16, did not produce the same total amount of CO_2 as the *S. cerevisiae* parent 59A and hybrid H16. This indicates differences in the total consumption of present sugars (maltose, glucose, maltotriose, fructose, sucrose). Residual sugars after the end of fermentation were determined for both parents and the hybrid strains (Table 31).

Table 31: Concentration of residual sugars [g/L] after the fermentation of the hybrid strains in 12 °P Heineken wort at 22 °C.

Strain	Glucose	Fructose	Sucrose	Maltose	Maltotriose	Total
Wort	1.23	0.30	0.28	5.22	1.42	8.44
59A	0.00	0.01	0.00	0.00	1.47	1.53
C902	0.00	0.01	0.00	0.00	1.44	1.50
H11	0.00	0.00	0.00	0.35	1.36	1.71
H12	0.00	0.00	0.00	1.77	1.44	3.21
H16	0.00	0.00	0.00	0.00	1.33	1.33
H24	0.00	0.00	0.00	3.72	1.38	5.10
H27	0.00	0.00	0.00	0.00	1.40	1.40
H54	0.00	0.00	0.00	4.00	1.46	5.46

It can be seen that all strains fully metabolize the sugars glucose, fructose and sucrose. Differences exist between the strains in the metabolism of maltose, which is only fully consumed by the parent strains and the hybrids H16 and H27. This is contradictory to observable differences between the fermentation curves in Figure 40, where a lower CO_2 production can be seen for parent C902 and H27, indicating that less sugars were consumed. A possible explanation could be the continuation of fermentation after the samples were taken, although samples were kept cold before submission to the Sample Service Center. Another explanation could be that not all fermenters were airtight sealed. Maltotriose is not consumed significantly by any of the strains. In some cases, the concentration of maltotriose seems higher than the initial concentration in wort. This was caused by inoculation with maltotriose rich preculture. The disablilty to consume maltotriose must not be a fault for beer brewing strains. Maltotriose does not have high sweetening power and a large share of *S. cerevisiae* ale strains does not fully metabolize the sugar (Zastrow et al., 2001). However, the incomplete consumption of maltose, as it was the case for several of the hybrid strains (Table 31), is more problematic as it leads to sweeter and less alcoholic beers that are more prone to contamination.

The strains show a sequential sugar metabolism, which could be demonstrated by measuring the sugar concentrations in samples taken during the fermentation (data not shown). Glucose is consumed first, then maltose. The switch between both sugar metabolisms causes the observed decline in fermentation rate after 10 h of fermentation, which is strain dependent (Figure 39 and Figure 40).

End point samples of the hybrid fermentations were submitted to the Heineken Sample Service Center and the concentrations of main fermentative aromas were determined. Unfortunately, no samples of the parent strain fermentations were submitted for aroma compound determination. The concentrations of volatile metabolites are given in Table 32.

Table 32: Concentrations of fermentative aromas [mg/L] produced by the hybrid strains H11 – H54 during fermentation in 12 °P Heineken wort at 22 °C.

Compound	H11	H12	H16	H24	H27	H54
Acetaldehyde	3.37	1.66	0.68	2.34	2.42	2.07
DMS [µg/L]	19.33	10.93	13.08	15.40	14.40	15.60
Acetone	0.52	0.36	0.68	0.68	0.57	0.67
Ethyl formiate	0.09	0.06	0.07	0.05	0.09	0.05
Ethyl acetate	18.65	14.98	23.31	6.50	30.07	5.35
Methanol	2.56	2.26	2.53	2.19	2.37	2.18
Ethyl propionate	0.07	0.04	0.14	0.02	0.09	0.02
Propanol	30.57	19.94	27.67	12.13	26.86	11.58
Isobutanol	34.48	31.61	43.67	13.08	39.47	10.23
Isoamyl acetate	1.82	1.52	4.06	0.95	3.58	0.80
Amyl alcohols	98.46	82.20	141.10	56.12	100.46	53.16
Ethyl capronate	0.12	0.11	0.49	0.09	0.39	0.07
Total higher alc.	163.60	133.70	212.50	81.30	166.90	75.00
4-vinylguaiacol	2.04	2.14	2.08	1.94	2.07	2.03

The hybrid strains show considerable variation regarding the production of fermentative aromas. The variation ranges from factor 1.1 for the production of 4-vinylguaiacol to factor 7.0 for the formation of the esters ethyl propionate and ethyl capronate. The production of 4-vinylguaiacol was therefore uniform among the strains. With more than factor 5.0, the variation for acetate ester formation (ethyl acetate and isoamyl acetate) was also high. The formation of higher alcohols varied by factor 2.8 between the hybrid strains.

The olfactory properties of the hybrid strains were rudimentary assessed by smelling 20 mL of the fermentation supernatants in a 50 mL Falcon tube. The detected odors are listed in Table 33.

Table 33: Aroma characteristics of the hybrid strain fermentations gained by the olfactory analysis.

Strain	Odor
H11	fruity
H12	fresh, fruitier than H11, slightly sweet
H16	unpleasant
H24	sweet, candy like, sligh solvent odor, stronger perception of 4-vinylguaiacol
H27	odd, sulfuric, worty, unpleasant
H54	similar to H24, also stronger perception of 4-vinylguaiacol

The perceived fruity aroma produced by strains H12 and H11 can be explained with higher levels of produced esters. The unpleasant aroma of strain H27 could occur due to the higher levels of produced higher alcohols (Table 32). The stronger perception of 4-vinylguaiacol, a

clove-like aroma, for strains H24 and H54 can not be explained with the measured levels of the compound, which were similar among all strains (Table 32). It is more likely that the aroma is perceived stronger since less other fermentative aromas were produced due to the unfinished fermentation (Table 32). The aroma of the majority of fermented beer samples was considered pleasant.

Based on the olfactory analysis, it was decided to ferment strain H12 in a 200-L scale in the Heineken pilot brewery. The resulting beer was filtered, filled in bottles and offered to participants of the 33rd International Specialized Symposium on Yeast (ISSY33 in Cork, Ireland). The beer was well perceived and attributed commercial potential. The taste and flavor perception was generally described as medium-bodied, round, slightly sweet and fruity.

2.4 Conclusion

In conclusion, the novel hybridization of *S. pastorianus* can generate great variety between the resulting hybrid strains. The use of non-ale *S. cerevisiae* strains as parent additionally increases the potential for obtaining strains with new attributes. The resulting strains do not only provide new characteristics for the fermentation of lager beer, but also extend the properties of strains used for the production of other fermented foods, like wine or cider.

More studies are necessary to fully assess the potential of the evaluated hybrids. This includes the fermentation at lower temperatures and the comparison with existing commercialized strains in order to evaluate the novelty of obtained flavor profiles. Nevertheless, it was shown that the generation of new *S. pastorianus* hybrids has potential to add more variety to lager beer strains.

Contribution to other publications

Publication 1

Explorer et exploiter la biodiversité des levures œnologiques : comment optimiser la qualité organoleptique des vins ?

Matthias Eder[a], Jessica Noble[a], Stéphanie Rollero[a,b], Damien Steyer[c], Jean-Luc Legras[a], Bruno Blondin[a], Sylvie Dequin[a]

[a] INRA, Montpellier Supagro, Université de Montpellier, UMR1083 Science pour l'œnologie, 2 place Viala, F-34060 Montpellier

[b] Lallemand, SAS, 19 rue des Briquetiers, F-31700 Blagnac

[c] Twistaroma, F-68021 Colmar

Corresponding author

Dr. Sylvie Dequin; INRA UMR1083 SPO, 2 place Viala, 34060 Montpellier CEDEX 2, France; Phone: +33 4 99 61 25 28; sylvie.dequin@supagro.inra.fr

Innovations Agronomiques 44 (2015), 55-67

Explorer et exploiter la biodiversité des levures œnologiques : comment optimiser la qualité organoleptique des vins?

Eder M.[1,2,3], Noble J.[1,2,3,4], Rollero S.[1,2,3,4], Steyer D.[5], Legras J-L.[1,2,3], Blondin B.[1,2,3], Dequin S.[1,2,3]

[1] INRA, UMR1083 Sciences pour l'œnologie, 2 Place Viala, F- 34060 Montpellier

[2] SupAgro, UMR1083 Sciences pour l'œnologie, 2 Place Viala, F- 34060 Montpellier

[3] Université de Montpellier, UMR1083 Sciences pour l'œnologie, 2 Place Viala, F- 34060 Montpellier

[4] Lallemand, SAS, 19 rue des Briquetiers, F- 31700 Blagnac

[5] Twistaroma, F-68021 Colmar

Correspondance : sylvie.dequin@supagro.inra.fr

Résumé

Aujourd'hui, l'industrie du vin doit faire face à de nombreux défis, liés à l'évolution des pratiques, des modes de consommation, au changement climatique et à la concurrence internationale. Un enjeu important de la filière est le développement de levures présentant des propriétés nouvelles, adaptées à ce contexte. La fermentation alcoolique, bien que faisant appel à une flore microbienne diverse présente dans le moût, est réalisée majoritairement par l'espèce *Saccharomyces cerevisiae*. C'est donc sur cette espèce qu'ont porté la majorité des efforts de recherche. Après la sélection de souches issues de la diversité naturelle, ces dix dernières années ont vu l'essor de techniques de génétique et d'hybridation permettant de combiner les propriétés intéressantes de plusieurs individus, ou la sélection de nouvelles variations en utilisant des approches d'évolution dirigée. Ces approches ont conduit à l'obtention de levures industrielles aux propriétés organoleptiques supérieures. Une autre voie consiste à combiner les propriétés d'espèces *Saccharomyces* par hybridation, ou encore à utiliser des espèces non *Saccharomyces* co-inoculées avec *S. cerevisiae* afin d'obtenir des profils aromatiques plus complexes. Nous présentons ici les avancées récentes dans l'amélioration des levures *S. cerevisiae* et la sélection de levures non conventionnelles pour optimiser la qualité organoleptique des vins.

Mots-clés : Fermentation œnologique, levures, arômes, sulfites, alcool, évolution dirigée, hybridation

Abstract: Exploration and exploitation of the biodiversity of wine yeast: how to optimize the organoleptic quality of wines?

Today, the wine industry is facing many challenges related to the development of new practices, changes in wine consumption, climate change and international competition. An important issue is the development of yeasts with new properties, adapted to this context. Wine fermentation, although involving a diverse microbial flora present in the must, is mainly carried out by the species *Saccharomyces cerevisiae*. Thus, this species has been the focus of most research efforts. After selecting strains from natural diversity, the last ten years have seen the rise of genetic and hybridization techniques to combine interesting properties of several individuals, and of evolutionary engineering to select new variations. These approaches resulted in the development of several industrial yeasts with superior organoleptic properties. Another strategy is to combine the properties of different *Saccharomyces* species by hybridization, or to use non-*Saccharomyces* species co-inoculated with *S. cerevisiae* in order to obtain more complex flavor profiles. We present recent advances in the improvement of *S. cerevisiae* and in the selection of non-conventional yeasts to optimize the organoleptic quality of the wines.

Keywords: Wine fermentation, yeast, aromas, sulfite, alcohol, adaptive evolution, hybridization

Introduction

Le vin a une longue et riche histoire, datant de plusieurs milliers d'années, étroitement liée à l'histoire de l'agriculture. Des études archéologiques et historiques permettent de penser que le vin est apparu durant la période néolithique (8500–4000 av. J.-C.). La sédentarisation des hommes durant cette période a favorisé, de manière générale, le développement de la transformation et du stockage des aliments (Cavalieri et al., 2003). L'une des premières preuves de la production de vin remonte à 5400-5000 av. J.-C, avec la découverte de la présence de tartrate de calcium dans des jarres datées de cette époque, sur le site néolithique de Hajji Firuz en Iran (McGovern et al., 1996). De là, la vinification s'est étendue à l'Egypte et à la Mésopotamie (~ 3500–3000 av. J.-C.) pour arriver en Crête (~2200 av. J.-C.) avant d'atteindre Europe et, à partir de là, le ''nouveau monde'' (McGovern, 2003).

La fermentation alcoolique est non seulement une méthode efficace pour la préservation de la qualité et la sécurité des boissons et des aliments, mais le vin est aussi une drogue et un remède largement utilisés dans l'antiquité, en raison de ses propriétés analgésiques, désinfectantes et de conservation. Au fil du temps, le vin a influencé la géographie, l'économie, l'archéologie, l'histoire, les mythologies et les religions, les arts et les traditions, le droit et la médecine. Aujourd'hui, cette boisson a une place unique dans la plupart des sociétés, avec une forte valeur économique et culturelle.

Ce n'est qu'en 1860 que Louis Pasteur a découvert l'origine de la fermentation et l'implication de la levure (Pasteur, 1860). Au début des années 1880, Emile Christian Hansen, du laboratoire Carlsberg au Danemark, a développé la première culture pure de levure et la première inoculation d'un moût de raisin avec une culture de levure pure a été réalisée par Müller-Thurgau en 1890. Cette pratique a commencé à être utilisée en œnologie dans les années 1970. Depuis, elle s'est généralisée et actuellement, la majorité des vins sont élaborés à l'aide de levures sélectionnées appartenant principalement à l'espèce *Saccharomyces cerevisiae*. Ces pratiques ont permis un meilleur contrôle et fiabilité du processus de fermentation, en limitant les risques d'altérations microbiologiques, et ont largement contribué à l'amélioration de la qualité des vins au cours des dernières décennies.

Aujourd'hui, la filière œnologique doit faire face à de nombreux défis, liés à l'évolution des pratiques, des modes de consommation, au changement climatique et à la concurrence internationale. Ce contexte se traduit par une dépendance accrue à l'égard de l'innovation technologique. La qualité sensorielle est devenue un critère essentiel pour le consommateur, et ce pour de nombreux produits fermentés (Hugenholtz, 2013). Le vin n'échappe pas à cette tendance et l'obtention de signatures uniques permettant de distinguer un vin des autres sur le marché est recherchée.

L'arôme du vin se compose d'une grande variété de composés volatils, certains d'entre eux provenant des raisins (arômes variétaux), de produits secondaires synthétisés lors de la fermentation (arômes fermentaires) et générés lors du vieillissement du vin (arômes post-fermentaires).

L'utilisation et le développement de souches et espèces de levures présentant des propriétés nouvelles, adaptées à ce contexte, représentent de ce fait un enjeu important. *S. cerevisiae* reste la principale levure sur laquelle ont été concentrés la plupart des efforts de recherche d'avant-garde en génétique, génomique, physiologie pour l'amélioration des levures de vin. Ces connaissances ont permis de réelles possibilités d'innovation, basées sur une meilleure exploitation de la diversité de cette espèce ou sur le développement de variants présentant de nouvelles capacités. Des stratégies non OGM ont été privilégiées ces dix dernières années, reposant principalement sur (i) des travaux de génétique quantitative qui visent à identifier les bases génétiques des propriétés des souches, de façon à développer de manière rationnelle des levures présentant une combinaison de traits optimaux par hybridation, (ii) des approches d'évolution dirigée, qui sont des stratégies alternatives aux approches basées sur le génie génétique permettant de générer de la diversité. Par ailleurs, les avancées en génomique ont permis ces dernières années l'identification parmi les levures œnologiques de plusieurs hybrides inter-espèces *Saccharomyces*, qui présentent un certain nombre d'avantages par rapport à *S. cerevisiae*. Enfin, les espèces non-*Saccharomyces* présentes dans le moût peuvent contribuer aux

caractéristiques organoleptiques des vins, que ce soit de manière positive ou négative (Fleet, 1993 ; Jolly et al., 2014). Actuellement, une tendance forte est d'exploiter cette diversité d'espèces de façon contrôlée, en utilisant des inoculum mixtes non *Saccharomyces*/S. *cerevisiae*.

Dans cette revue, nous insisterons plus particulièrement sur les développements récents de levures *S. cerevisiae* en lien avec l'amélioration des qualités organoleptiques du vin, et nous discuterons plus brièvement l'intérêt de levures moins conventionnelles, comme les hybrides *Saccharomyces* inter-espèces et des espèces non *Saccharomyces*.

1. Evolution de la flore pendant la fermentation œnologique

La fermentation des moûts de raisins peut se produire spontanément par l'activité de divers micro-organismes naturellement présents sur les baies de raisin. Plus de 40 espèces de levures ont été identifiées à partir de moût de raisins (Fleet, 1993 ; Jolly et al., 2014). Les genres les plus fréquents sont *Hanseniaspora* (*Kloeckera*), *Candida*, *Pichia*, *Rhodotorula*, *Debaryomyces*, *Metschnikowia*, *Kluyveromyces*, *Schizosaccharomyces*, *Torulaspora*, *Zygosaccharomyces* et *Dekkera*. Ces espèces sont majoritaires lors des étapes préfermentaires de l'élaboration des vins et en début de fermentation. Une succession séquentielle de ces espèces est observée au cours de la première phase de la fermentation spontanée. La plupart de ces espèces disparaissent rapidement, bien que certaines puissent persister plus longtemps (Fleet, 1993). Ce phénomène peut s'expliquer par leur faible capacité fermentaire, leur faible tolérance à l'anaérobiose et aux concentrations élevées en SO_2 et en éthanol.

La fermentation spontanée est une pratique pouvant conduire à l'obtention de profils aromatiques plus complexes, en raison du répertoire d'espèces et de souches mises en jeu dans le moût. Cependant, le microbiote naturel, qui provient de la flore présente sur les baies et dans les caves, est inconnu et variable d'une année sur l'autre, ce qui en fait une pratique imprévisible et donc risquée. Même dans le cas des fermentations spontanées, la prédominance de *S. cerevisiae* (inoculée ou indigène) dans la fermentation est attendue et désirée. Cette espèce est responsable de la dégradation de la majeure partie des sucres en éthanol et CO_2. Une des caractéristiques les plus remarquables de *S. cerevisiae* et d'autres espèces étroitement apparentées est leur capacité à produire et accumuler l'éthanol, appelé effet Crabtree, même en présence d'oxygène. C'est la combinaison de plusieurs traits «gagnants», comme la dégradation rapide des sucres, la production et l'accumulation d'éthanol associée à la production de chaleur, la tolérance à ce composé et aux fortes températures, et la capacité de croissance en anaérobiose qui fait de *S. cerevisiae* la levure œnologique par excellence.

2. La levure *S. cerevisiae* : rôle et diversité intra espèce

La fermentation œnologique expose les levures à une variété de stress : stress osmotique dû à la forte concentration en sucres du moût de raisin (180-260 g/L), niveaux élevés de sulfite, anaérobiose, stress acide, limitation en éléments nutritifs (azote, lipides et vitamines) et toxicité de l'éthanol. Une fermentation de vin typique comporte une phase de latence, qui dure plusieurs heures, une courte phase de croissance de 24 à 36 heures, suivie d'une phase stationnaire, pendant laquelle la majeure partie du sucre (entre 50 et 80%) est fermentée. Durant cette phase, l'activité des levures diminue constamment, bien que les niveaux de viabilité restent élevés (en général plus de 90%), jusqu'à ce que le sucre soit épuisé.

Dans ce contexte, la levure *S. cerevisiae* joue un rôle majeur. Elle est en effet non seulement responsable de la dégradation des sucres en éthanol et CO_2, mais produit de nombreux sous-produits de fermentation.

Les arômes produits par *S. cerevisiae* comprennent entre autres des aldéhydes, des alcools supérieurs, des acides gras à moyenne et longue chaîne, des acides gras, des esters d'éthyle et des esters d'acétate ainsi que des composés soufrés (Swiegers et al., 2005). Ces composés volatils sont issus de la dégradation des sucres, des acides aminés et des acides gras, et sont formés à travers un processus métabolique complexe et dynamique pendant la fermentation. De nombreux facteurs contrôlent leur production : le cépage, la nature des précurseurs (principalement les acides aminés), la disponibilité en micronutriments et en azote, les traitements du moût et les conditions de fermentation, y compris la température, ainsi que la souche de levure. Ces composés fermentaires jouent un rôle clé dans la qualité sensorielle et la typicité des vins. Par exemple, certains alcools supérieurs, les esters d'éthyle et des esters d'acétate génèrent des notes fruitées et florales (Swiegers et al., 2005). L'acide acétique est un composé clé dont la production doit rester faible, et dont le niveau de production varie très fortement d'une souche à l'autre (Camarasa et al., 2011 ; Marullo et al., 2006 ; Salinas et al., 2012). Des composés cétoniques comme l'acétaldéhyde peuvent également avoir un impact organoleptique en combinant le SO_2. Les composés soufrés comme les sulfites (SO_2) ou le sulfure d'hydrogène (H_2S) affectent négativement les qualités sensorielles d'un vin. La comparaison de 72 souches *S. cerevisiae* d'origines diverses dans des conditions de fermentation œnologique a révélé des variations importantes non seulement dans les propriétés fermentaires mais aussi dans la production de métabolites, en particulier l'acide acétique et les esters (Camarasa et al., 2011). Une autre étude, basée sur un nombre restreint de souches et l'analyse sensorielle, a suggéré que les levures de vin produisent plus d'arômes fruités du vin que les souches provenant d'autres origines (Hyma et al., 2011).

En plus de cette production *de novo* de métabolites secondaires, les levures peuvent également transformer des précurseurs du raisin pour libérer des arômes variétaux (Swiegers et al., 2005 ; Swiegers et Pretorius, 2005). Les levures *Saccharomyces* peuvent par exemple former des thiols comme le 4MMP (4-méthyl-4-mercapto-pentan-2-one), le 3MH (3-mercapto-hexan-1-ol) et son acétate (A3MH) à partir de différents précurseurs. Ces composés, caractéristiques des vins de Sauvignon blanc, apportent des notes aromatiques de fruit de la passion, pamplemousse, citron et buis (Tominaga et al., 2000, Aznar et al., 2001). L'effet de la souche de levure sur la révélation des thiols volatils peut varier d'un facteur 20 (Dubourdieu et al., 2006 ; Swiegers et Pretorius, 2005). D'autres composés, comme les terpénols sont des composés très importants de l'arôme des cépages musqués (muscat, gewurztraminer…). Lors de la fermentation, la levure *S. cerevisiae* agit sur les composés terpéniques en hydrolysant les glycosides, et en isomérisant ou en réduisant les alcools terpéniques libérés comme le géraniol pour donner du citronellol ou du géranyl diol et donner du cis rose oxyde, l'un des composés d'arôme clé du gewurztraminer (Koslitz et al., 2008). Enfin, la souche de levure peut également avoir un impact sur la formation d'autres composés, comme par exemple le sulfure de diméthyle (DMS), un composé soufré léger qui peut contribuer à l'arôme, positivement ou négativement selon sa concentration et la typologie du vin. Différents travaux ont montré que ce composé est un exhausteur des arômes fruités (Segurel et al., 2004 ; Dagan et Schneider, 2012). La S-méthylméthionine (SMM) représente l'essentiel des précurseurs de DMS des raisins et du moût, mais une part importante de ce potentiel aromatique est perdue au cours de la vinification. Les levures ont différentes capacités d'assimilation de ce précurseur pendant la fermentation (Dagan et Schneider, 2012), ce qui ouvre des perspectives pour développer des souches préservant ce potentiel.

Les stratégies mises en œuvre pour développer de nouvelles souches *S. cerevisiae* plus performantes sur le plan organoleptique ainsi que plusieurs exemples illustrant ces recherches seront détaillées ci-après.

3. Amélioration des souches : nouvelles cibles technologiques

3.1 Stratégies non OGM : cartographie de QTL et évolution dirigée

3.1.1 Cartographie de QTL

Les variations du phénotype (par exemple la production de quantité différente d'un métabolite) sont la conséquence de mécanismes génétiques complexes. On définit un QTL (Quantitative Trait Locus) comme une région du génome responsable de ce caractère quantitatif. Le principe de recherche de QTL consiste à établir un lien statistique entre l'hérédité d'un marqueur génétique (le plus souvent un polymorphisme de séquence) et la valeur d'un caractère quantitatif. La recherche de ce lien s'effectue par des approches de cartographie génétique et phénotypiques, qui sont résumées sur la Figure 1. Différents approches de cartographie de QTL ont été menées pour rechercher des gènes impliqués dans des caractères technologiques, comme la production de sulfite ou la production de composés volatils aromatiques.

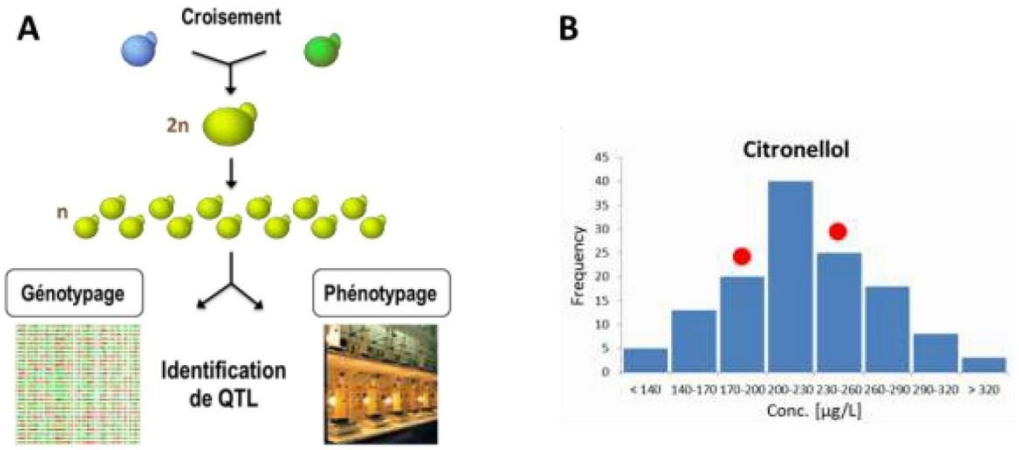


Figure 1 : Démarche de recherche de QTL. (A) Deux souches parentales haploïdes ayant des valeurs phénotypiques différentes sont croisées, générant un individu diploïde hétérozygote. La sporulation de cet individu permet d'obtenir une population de descendants haploïdes ayant subi plusieurs évènements de recombinaison lors de la méiose et possédant une distribution aléatoire des différents loci parentaux. Ces descendants sont génotypés (actuellement par séquençage complet de leur génome) afin de localiser ces marqueurs, et phénotypés pour les traits quantitatifs d'intérêt. L'analyse statistique de liaison est la dernière étape qui permet la détection des QTL. Un QTL est identifié lorsqu'il existe une corrélation entre un marqueur et le paramètre quantitatif étudié. (B) Fréquence de la distribution d'un trait (citronellol) dans la population de descendants. La production des 2 parents est indiquée par un rond rouge.

3.1.2 Evolution dirigée

Les approches d'évolution expérimentale, ou évolution adaptative, sont basées sur la culture à long-terme d'un organisme en présence d'une pression de sélection (Figure 2). Ceci favorise l'émergence de variations génétiques, qui peut être suivie par l'évolution adaptative de la population de levure et par la sélection de variants présentant un phénotype souhaité. La caractérisation des souches évoluée et ancestrale par séquençage de leur génome et différentes analyses de génomique fonctionnelle (comme l'étude de l'expression du génome par analyse transcriptomique) est généralement utilisée pour essayer d'identifier les mutations impliquées, ce qui reste un objectif parfois difficile à atteindre. Une fois

identifiées, les mutations peuvent alors être transférées à une autre souche par hybridation. Les approches d'évolution ont été appliquées récemment au développement de souches aromatiques et de souches à faible rendement en alcool. Elles ont pour avantage de pouvoir être utilisées sans connaissance génétique préalable, et de pouvoir être appliquées à n'importe quelle espèce.

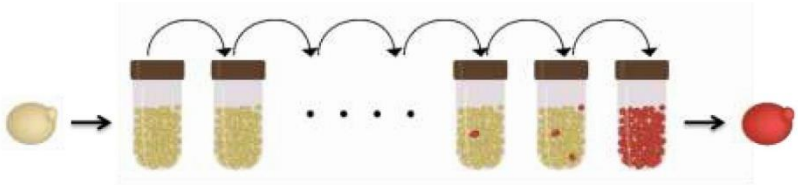


Figure 2 : Principe de l'évolution adaptative. Une culture prolongée est réalisée afin de favoriser l'accumulation de mutations spontanées. L'utilisation de conditions sélectives (présence d'un stress, croissance en présence d'un substrat mal assimilé etc…) permet de sélectionner les souches ayant acquis une mutation qui confère un bénéfice dans ces conditions.

3.2 Exemples de développement de souches S. cerevisiae

3.2.1 Levure œnologique faible productrice de sulfites

Les levures œnologiques forment divers composés soufrés en fermentation qui peuvent avoir des impacts négatifs sur la qualité organoleptique, la qualité sanitaire ou la maîtrise des procédés. C'est le cas des sulfites et du H_2S qui sont parfois formés en quantités excessives lors des fermentations alcooliques. Les niveaux de sulfites des vins sont règlementés pour des raisons sanitaires ce qui suppose la maîtrise de leur formation au cours de l'élaboration des vins. Les sulfites sont aussi des inhibiteurs des bactéries lactiques qui réalisent la fermentation malolactique ce qui suppose de maîtriser leur teneur transitoire au cours des procédés.

Il est donc important de maîtriser la formation de ces composés. Des travaux récents de génétique des levures ont permis d'identifier les bases génétiques des variations de formation de composés soufrés. Une démarche de génétique quantitative et de recherche de QTL (quantitative trait locus) a été menée dans deux souches aux comportements contrastés, ce qui a permis d'identifier deux gènes candidats impliqués dans le métabolisme du soufre (*SKP2* et *MET2*) pour lesquels l'impact des variations alléliques sur la formation de SO_2, de H_2S et d'acétaldéhyde a été démontré.

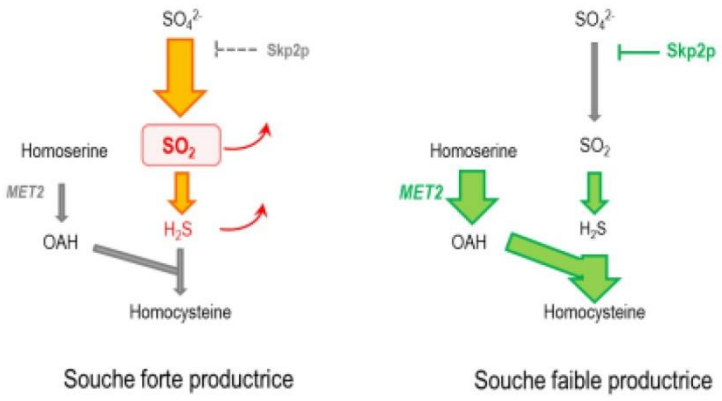


Figure 3 : Allèles impliqués dans la formation de SO_2, H_2S et acétaldéhyde

Ceci a permis d'expliquer les différences de comportement des souches. En effet, la combinaison des allèles favorables permet, à la fois, de limiter la capacité d'incorporation du sulfate très précocement dans le métabolisme et d'augmenter l'utilisation des formes sulfites et H_2S par l'activité de la voie d'apport des squelettes azotés conduisant à la cystéine. Une fois ces gènes/allèles identifiés, il a été possible de les introduire à façon dans d'autres souches de levures, par des approches d'hybridation, ce qui a permis de générer de nouvelles souches faibles productrices de SO_2, H_2S et acétaldéhyde (Noble et al., 2015).

3.2.2 Levures œnologiques à caractère aromatique

Thiols : Des avancées récentes ont permis d'identifier un des gènes clés impliqués dans la formation d'un des thiols du Sauvignon Blanc. En utilisant une approche de cartographie de QTL, il a été possible d'identifier un allèle du gène *IRC7*, qui code pour une cystéine desulfydrase impliquée dans la forte production de 4MMP. La plupart des souches de *S. cerevisiae*, y compris la souche de laboratoire de référence, ont une délétion de 38 pb (paires de bases) qui inactive ce gène. La surexpression de l'allèle long identifié dans une levure de vin a permis d'augmenter la production de 4MMP dans un vin Sauvignon Blanc, qui est alors passé d'un niveau indétectable (<10 ng/L) à des concentrations de 1000 ng/L, ainsi que celles de 3MH et A3MH (Roncoroni et al., 2011). L'identification de ce gène a permis par la suite le développement de souches fortes productrices de thiols par des approches d'hybridation (Dufour et al., 2013)

Terpénols : Les terpénols dont fait partie le géraniol sont des composés très importants de l'arôme des cépages musqués (muscat, gewurztraminer…). Différentes approches de recherche de QTL ont permis d'identifier des régions du génome impliquées dans la production de rose oxyde caractéristique du gewurztraminer, à partir du géraniol à l'odeur de rose (Steyer et al., 2012), mais le gène impliqué dans ces différences n'a pas encore été identifié. Ces études ont également éclairé les étapes de la transformation du géraniol. Il est transformé en un autre composé aromatique aux notes citronnées (citronellol) sous l'action du gène *OYE2* et est également transformé en ester durant la fermentation sous l'action du gène *ATF1*, ce pool d'ester de géraniol pouvant ensuite être réhydrolysé par la suite durant la conservation du vin (Steyer et al., 2013). Un des enjeux actuels, poursuivis dans le cadre du projet européen YeastCell (ITN-2013-606795) dont l'INRA est partenaire, est d'identifier d'autres allèles potentiellement impliqués dans la formation de ces composés.

Esters : Récemment une levure aromatique surproductrice d'esters a été obtenue en utilisant une approche d'évolution adaptative. Une souche de levure œnologique commerciale a été cultivée pendant 70 générations sur un milieu contenant du gluconate, un sucre mal assimilé par *S. cerevisiae* et métabolisé au niveau de la voie des pentoses phosphate (VPP). Cette approche a permis de sélectionner un variant présentant une meilleure croissance sur ce substrat (Cadiere et al., 2011 ; Cadiere et al., 2012). Lors de la fermentation du vin, la souche évoluée se caractérise par une forte production d'esters d'acétate, qui participent au caractère fruité du vin, et par une diminution de la production d'acide acétique, composé indésirable à forte concentration. Ces modifications sont dues à un métabolisme fortement remanié, notamment à une amplification de la voie des pentoses phosphates, une synthèse accrue de lipides et des modifications du métabolisme azoté et secondaire. Des vinifications à l'échelle pilote ainsi que des essais en cave ont confirmé le potentiel de cette souche pour la production de vins aromatiques et fruités. Cette souche est commercialisée depuis 2012.

Les effets de différents facteurs environnementaux (azote, phytostérols et température) sur la production des arômes de la souche évoluée et de la souche ancestrale ont été évalués. Une surproduction systématique des alcools supérieurs (à l'exception du propanol) et des esters d'acétate par la souche évolué a été observée quelles que soient les conditions de fermentation (Figure 4).

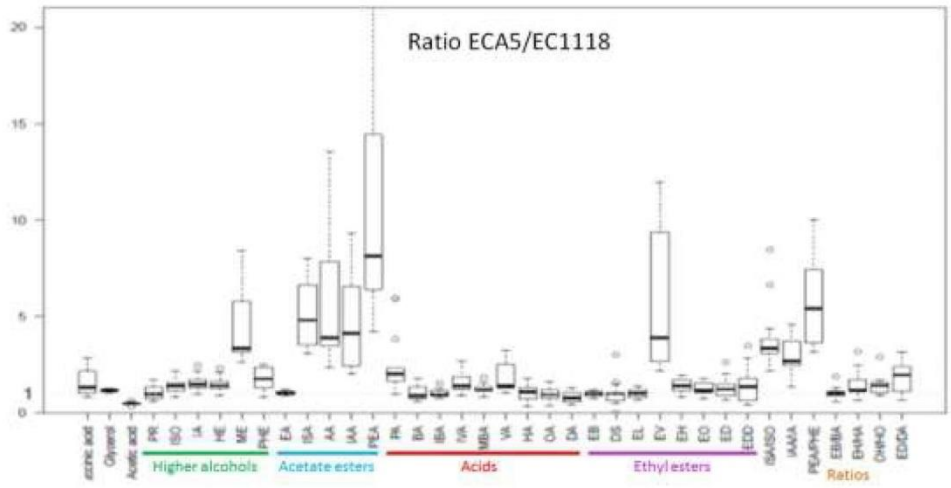


Figure 4 : Caractéristiques aromatiques de la souche évoluée Affinity™ ECA5. Ratios entre les concentrations finales des arômes fermentaires produites par Affinity ™ ECA5 (souche évoluée) et Lalvin EC1118® (souche ancestrale).

PR: propanol; ISO: isobutanol; IA: alcool isoamylique; HE: hexanol; ME: méthionol; PHE: 2-phényléthanol; EA: acétate d'éthyle; ISA: acétate d'isobutyle; AA: acétate d'amyle; IAA: acétate d'isoamyle; PEA: 2-phenylethylacetate; PA: acide propanoïque; BA: acide butanoïque; IBA: acide isobutanoïque; IVA: acide isovalérique; MBA: acide 2-méthylbutanoïque; VA: acide valérique; HA: acide hexanoïque; OA: acide octanoïque; DA: acide décanoïque; EB: butanoate d'éthyle; DS: succinate de diéthyle; EL: lactate d'éthyle; EV: valérate d'éthyle; EH: hexanoate d'éthyle; EO: octanoate d'éthyle; ED: décanoate d'éthyle; JED: dodécanoate d'éthyle

Le séquençage complet du génome de plusieurs souches évoluées ainsi obtenues a récemment permis d'identifier la mutation responsable de cette surproduction d'esters, ce qui offre la possibilité de transférer ce caractère à d'autres souches œnologiques. Par ailleurs, d'autres travaux basés sur la recherche de QTL sont actuellement développés par l'INRA dans le cadre du projet européen YeastCell (ITN-2013-606795), afin d'identifier des gènes et allèles impliqués dans la formation des alcools supérieurs et esters.

3.2.3 Souche œnologique faible productrice d'éthanol

L'évolution des pratiques vinicoles, la sélection de cépages à fort rendement en sucres et le changement climatique ont conduit ces 30 dernières années à une augmentation de la teneur en alcool des vins de l'ordre de 2% v/v. Cette augmentation a des conséquences à plusieurs niveaux : (i) sur la qualité sensorielle des vins et le déroulement des fermentations, (ii) économique, ainsi (iii) qu'en termes de santé publique. C'est avec l'objectif de réorienter le métabolisme des levures vers la formation de sous-produits ayant un impact favorable sur la qualité du vin, comme le glycérol, au détriment de l'éthanol, que des travaux récents ont été conduits visant à sélectionner des souches par adaptation sur des milieux à forte pression osmotique.

L'obtention de souches à faible rendement en alcool repose sur la modification de leur métabolisme afin de détourner une partie des sucres vers d'autres sous-produits que l'éthanol. Une démarche d'évolution expérimentale a été conduite sur une souche œnologique commerciale, consistant à maintenir en culture les levures pendant plusieurs centaines de générations dans un milieu salin contrôlé (en présence de KCl) imposant un stress osmotique afin de favoriser l'apparition et la sélection de mutations naturelles orientant leur métabolisme vers une surproduction de glycérol. Il a ainsi été possible de dévier le métabolisme des sucres vers la formation de glycérol (qui confère au vin de la rondeur, du moelleux) et de 2,3-butanediol (neutre d'un point de vue sensoriel), sans accumulation de

métabolites indésirables. Une approche génétique basée sur des croisements de descendants méiotiques de la levure évoluée a permis d'obtenir une souche présentant une déviation métabolique plus importante. Une souche finale sélectionnée a été évaluée à l'échelle pilote sur moûts de raisin : les vins obtenus présentent une diminution de la teneur en alcool de l'ordre de 1 % v/v et contiennent très peu d'acidité volatile. Leur acidité totale est par contre plus élevée, ce qui présente un fort intérêt puisque l'augmentation de la teneur en alcool des vins est le plus souvent associée à des pH trop élevés (Tilloy et al., 2014). Ces recherches s'inscrivent dans l'effort actuel des chercheurs de l'INRA au sein du programme LACCAVE (Impacts et adaptations à long terme de la filière viti-vinicole au Changement Climatique) pour mieux cerner les impacts du changement climatique et étudier les stratégies d'adaptation de la filière viticole.

4. Autres espèces *Saccharomyces* et hybrides

En dehors de *S. cerevisiae*, une autre espèce du clade *Saccharomyces*, *Saccharomyces uvarum*, est également capable d'achever la fermentation alcoolique du jus de raisin. On retrouve cette espèce également associée à la fermentation du cidre (Nguyen et al., 2000). Il s'agit d'une espèce plus cryo-tolérante que *S. cerevisiae* (Belloch et al., 2008), qui confère au vin des notes florales particulières liées notamment à la production de phényl-2-éthanol et de son acétate. Elle produit également peu d'acide acétique et peut produire des terpènes (Masneuf-Pomarede et al., 2010 ; Gamero et al., 2011).

Au cours des dernières décennies, un nombre croissant d'hybrides naturels interspécifiques entre deux ou plusieurs espèces de levure *Saccharomyces* a été identifié dans les procédés fermentaires conduisant à des boissons fermentés. L'exemple le plus connu est la levure de brasserie *S. pastorianus*, qui est un hybride entre *S. cerevisiae* et *S. eubayanus* (Libkind et al., 2011). La caractérisation moléculaire de levures de vin et de cidre a également révélé de nombreux hybrides formés indépendamment entre *S. cerevisiae* / *S. kudriawzevii* (González et al., 2006 ; Erny et al., 2012),. *S. cerevisiae / S uvarum* (Masneuf et al., 1998 ; Masneuf et al., 2002 ; Sipiczki, 2008) ou entre *S. cerevisiae / S. kudriawzevii / S. uvarum* (González et al., 2006).

L'hybridation interspécifique, en fournissant de nouvelles combinaisons de gènes, peut conférer de nouveaux avantages par rapport aux espèces parentales. Ainsi, plusieurs travaux ont révélé une plus grande tolérance des hybrides aux stress ou aux basses températures par rapport à *S. cerevisiae* lors de la fermentation du vin (Belloch et al., 2008 ; González et al., 2007). Par exemple, *S. uvarum* et *S. kudriavzevii* sont mieux adaptées à la croissance à basse température que *S. cerevisiae*, tandis que *S. cerevisiae* est plus tolérante à l'alcool. Ces hybrides peuvent être particulièrement avantageux lors de la fermentation des vins blancs, qui a lieu à basse température (généralement à 10-15° C) afin de réduire la perte de composés volatils aromatiques. Certains hybrides sont également décrits comme forts producteurs d'alcools supérieurs, d'esters et de thiols (González et al., 2007 ; Dubourdieu et al., 2006 ; Swiegers et al., 2009). Les caractéristiques aromatiques de souches de différentes espèces et d'hybrides interspécifiques ont été comparées lors de vinifications sur Gewurztraminer (Figure 5). Cette analyse, bien que portant sur un nombre restreint d'individus de chaque espèce, a permis de différencier les origines génétiques sur la base de la production de certaines familles d'arômes : par exemple, les hybrides *S. cerevisiae* x *S. kudriavzevii* se caractérisent par une plus forte production de terpénols.

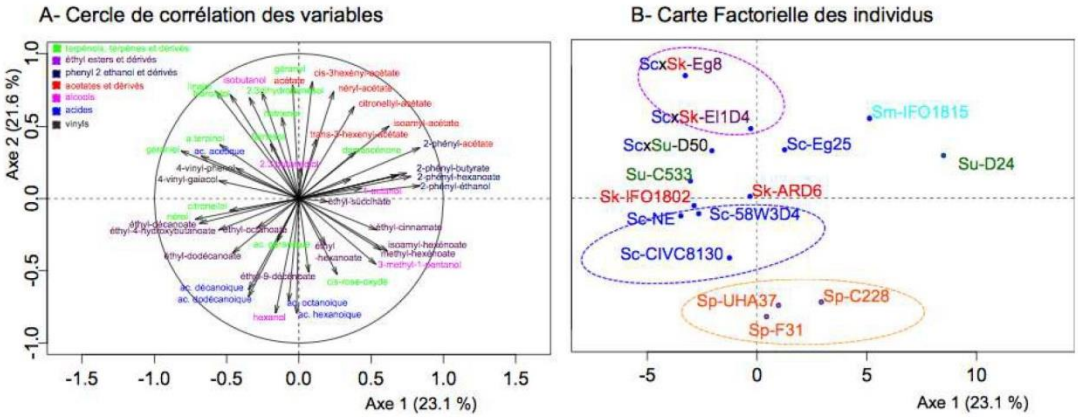


Figure 5 : Analyse en composantes principales (ACP) des concentrations en composés volatils de vins de Gewurztraminer (2008) produits par différentes souches et espèces. Fermentations réalisées à l'échelle de 2 litres en triplicat. A : Cercle de corrélation des variables. B : Carte factorielle des individus. Sp-F31, Sp-UHA37, Sp-C228 : *S. paradoxus* ; Su-D24 : *S. uvarum* ; Sc x Sk-Eg8, Sc x Sk-El1D4 (hybrides *S. cerevisiae* x *S. kudriavzevii*), Sc-CIVC8130, Sc-Eg25, Sc-58W3D4 : *S. cerevisiae* ; Sk-Ifo1815 : *S. mikatae,* Sk-Ifo1802 : *S. kudriavzevii* (ici probablement dominé par *S. cerevisiae, et Sc-NE : non ensemencé*). D'après Steyer et Legras (données non publiées).

5. Espèces non *Saccharomyces*

L'intérêt œnologique des levures non-*Saccharomyces* a connu un regain d'intérêt ces dernières années. Si leur impact a longtemps été jugé comme négatif, plusieurs études récentes ont dévoilé certaines aptitudes technologiques de ces levures, notamment pour leur production d'arômes. D'une manière générale, l'addition des levures non-*Saccharomyces*, en combinaison avec *S. cerevisiae*, pourrait permettre d'accroitre la complexité du produit final, tout en évitant les risques liés à la fermentation spontanée. Leur contribution à la production d'arômes fermentaires et leur capacité à excréter des activités enzymatiques d'intérêt sont autant de potentialités qui ont été décrites dans de nombreuses publications (Jolly et al., 2014 ; Sadoudi et al., 2012 ; Dashko et al., 2015 ; Renault et al., 2015). Cela n'exclue pas la production parfois élevée de certains composés indésirables ; comme chez *S. cerevisiae*, il existe une grande diversité phénotypique au sein de chaque espèce, ce qui nécessite de cribler un grand nombre de souches pour obtenir la bonne combinaison de propriétés d'intérêt. Une autre difficulté est l'existence d'interactions complexes et en grande partie inconnues entre levures. Par exemple, des effets synergiques ou au contraire négatifs sur la production de composés aromatiques ont été observés lors de co-cultures de certaines espèces avec *S. cerevisiae* (Sadoudi et al., 2012). Par conséquent, la qualité du vin qui résulte de fermentation après ensemencement avec un mélange de levures est assez imprévisible. Quelques couples mixtes sont actuellement commercialisés, offrant de nouveaux outils pour moduler l'équilibre organoleptique des vins.

Conclusions et perspectives

Plus de 800 composés ont été identifiés comme participant au profil aromatique des vins (Mendes-Pinto, 2009). Certains de ces composés possèdent une contribution individuelle significative, tandis que d'autres vont agir en synergie ou de façon antagoniste. Les arômes peuvent être issus directement du raisin, produits par le métabolisme des micro-organismes, notamment la levure, ou développés au cours du vieillissement du vin (Francis et Newton, 2005).

Aujourd'hui, les propriétés aromatiques et en particulier le caractère fruité des vins sont de plus en plus recherchés. La souche de levure, la composition du moût et le mode de conduite de la fermentation ont un impact majeur sur la production d'arômes au cours de cette étape du procédé de vinification. De ce fait, la production d'arômes fermentaires peut être modulée en jouant sur la souche de levure ou sur les conditions de fermentation (températures, nutriments). Les avancées récentes en matière de développement de nouvelles souches produisant plus d'arômes fermentaires ou capables de convertir de manière plus efficace les précurseurs du moût en arômes variétaux, offrent des perspectives considérables pour orienter le profil aromatique des vins. Ces approches ont déjà abouti à la mise sur le marché de nouvelles souches. Leur poursuite devrait permettre d'identifier les bases génétiques des différences de production d'un plus grand nombre d'arômes, offrant de nouvelles perspectives d'amélioration.

D'autre part, l'utilisation des levures non-*Saccharomyces* a connu un regain d'intérêt ces dernières années. Certaines souches appartenant à ces espèces présentent en effet des potentialités intéressantes, notamment aromatiques. Il existe de ce fait une tendance récente et avérée pour développer des fermentations en culture mixte, combinant une espèce « non conventionnelle » avec une *S. cerevisiae*. Si l'utilisation de ces levains mixtes peut constituer de nouveaux outils pour moduler l'équilibre organoleptique des vins, leur maitrise nécessite une meilleure connaissance, jusqu'alors fragmentaire, du métabolisme des levures « non conventionnelles » ainsi que des mécanismes d'interactions avec *S. cerevisiae*.

Références bibliographiques

Aznar M., Lopez R., Cacho J.F., Ferreira V., 2001. Identification and quantification of impact odorants of aged red wines from Rioja. GC-olfactometry, quantitative GC-MS, and odor evaluation of HPLC fractions. Journal of Agricultural and Food Chemistry 49, 2924-2929

Belloch C., Orlic S., Barrio E., Querol A., 2008. Fermentative stress adaptation of hybrids within the *Saccharomyces* sensu stricto complex. International Journal of Food Microbiology 122, 188-195.

Cadiere A., Camarasa C., Julien A., Dequin S., 2011. Evolutionary engineered *Saccharomyces cerevisiae* wine yeast strains with increased in vivo flux through the pentose phosphate pathway. Metabolic Engineering 13, 263-271.

Cadière A., Aguera E., Caillé S., Ortiz-Julien A., Dequin S., 2012. Pilot-scale evaluation the enological traits of a novel, aromatic wine yeast strain obtained by adaptive evolution. Food Microbiology 32, 332-337.

Camarasa C., Sanchez I., Brial P., Bigey F., Dequin S., 2011. Phenotypic landscape of *Saccharomyces cerevisiae* during wine fermentation: evidence for origin-dependent metabolic traits. PLoS One 6, e25147.

Cavalieri D., McGovern P.E., Hartl D.L., Mortimer R., Polsinelli M., 2003. Evidence for *S. cerevisiae* Fermentation in Ancient Wine. Journal of Molecular Evolution 57, 226–232.

Dagan L., Schneider R., 2012. Le sulfure de diméthyle : Quels moyens pour gérer ses teneurs dans les vins en bouteille ? Actes du Colloque Arômes des vins, Toulouse, pp7-10.

Dashko S., Zhou N., Tinta T., Sivilotti P., Lemut M.S., Trost K., Gamero A., Boekhout T., Butinar L., Vrhovsek U., Piskur J., 2015. Use of non-conventional yeast improves the wine aroma profile of Ribolla Gialla. Journal of Industrial Microbiology and Biotechnology, Apr 23. DOI 10.1007/s10295-015-1620-y

Dubourdieu D., Tominaga T., Masneuf I., Peyrot des Gachons C., Murat M.L., 2006. The Role of Yeasts in Grape Flavor Development during Fermentation: The Example of Sauvignon blanc. American Journal of Enology and Viticulture 57, 81-88.

Dufour M., Zimmer A., Thibon C., Marullo P., 2013. Enhancement of volatile thiol release of *Saccharomyces cerevisiae* strains using molecular breeding. Applied Microbiology and Biotechnology 97, 5893-5905.

Erny C., Raoult P., Alais A., Butterlin G., Delobel P., Matei-Radoi F., Casaregola S., Legras J.L., 2012. Ecological success of a group of *Saccharomyces cerevisiae/Saccharomyces kudriavzevii* hybrids in the northern European wine-making environment. Applied and Environmental Microbiology 78, 3256-65.

Fleet G.H., 1993. The microorganisms of winemaking – isolation, enumeration and identification. Wine Microbiology and Biotechnology, 1–25. In Fleet GH (ed.), Harwood Academic Publishers, Switzerland.

Francis I.L., Newton J.L., 2005. Determining wine aroma from compositional data. Australian Journal of Grape and Wine Research 11, 114-126.

Gamero A., Manzanares P., Querol A., Belloch C., 2011. Monoterpene alcohols release and bioconversion by *Saccharomyces* species and hybrids. International Journal of Food Microbiology 145, 92-97.

González S.S., Barrio E., Gafner J., Querol A., 2006. Natural hybrids from *Saccharomyces cerevisiae*, *Saccharomyces bayanus* and *Saccharomyces kudriavzevii* in wine fermentations. FEMS Yeast Research 6, 1221–1234.

González S.S., Gallo L., Climent M.D., Barrio E., Querol A., 2007. Enological characterization of natural hybrids from *Saccharomyces cerevisiae* and *S. kudriavzevii*. International Journal of Food Microbiology 116, 11–18.

Hugenholtz J., 2013. Traditional biotechnology for new foods and beverages. Current Opinion in Biotechnology 24, 155-159.

Hyma K.E., Saerens S.M., Verstrepen K.J., Fay J.C., 2011. Divergence in wine characteristics produced by wild and domesticated strains of *Saccharomyces cerevisiae*. FEMS Yeast Research 11, 540-51.

Jolly N.P., Varela C., Pretorius I.S., 2014. Not your ordinary yeast: non-*Saccharomyces* yeasts in wine production uncovered. FEMS Yeast Research 14, 215–237.

Koslitz S., Renaud L., Kohler M., Wust M., 2008. Stereoselective formation of the varietal aroma compound rose oxide during alcoholic fermentation. Journal of Agricultural and Food Chemistry 56,1371-5.

Libkind D., Hittinger C.T., Valério E., Goncalves C., Dover J., Johnston M., Goncalves P., Sampaio J.P., 2011. Microbe domestication and the identification of the wild genetic stock of lager-brewing yeast. Proceedings of the National Academy of Sciences 108, 14539-14544.

Masneuf I., Hansen J., Groth C. Piskur J., 1998. New Hybrids between *Saccharomyces* Sensu Stricto Yeast Species Found among Wine and Cider Production Strains. Applied and Environmental Microbiology 64, 3887–3892.

Masneuf I., Murat M., Naumov G., Tominaga T., Dubourdieu D., 2002. Hybrids *Saccharomyces cerevisiae & Saccharomyces bayanus var. uvarum* having a high liberating ability of some sulfur varietal aromas of *Vitis vinifera* sauvignon blanc wines. Journal International des Sciences de la Vigne et du Vin 36, 205–212.

Masneuf-Pomarede I., Bely M., Marullo P., Lonvaud-Funel A., Dubourdieu D., 2010. Reassessment of phenotypic traits for *Saccharomyces bayanus var. uvarum* wine yeast strains. International Journal of Food Microbiology 139, 79-86.

Marullo P., Bely M., Masneuf-Pomarede I., Pons M., Aigle M., Dubourdieu D., 2006. Breeding strategies for combining fermentative qualities and reducing off-flavor production in a wine yeast model. FEMS Yeast Research 6, 268-279.

McGovern P.E., Glusker D.L., Exner L.J., Voigt M.M., 1996. Neolithic resonated wine. Nature 381, 480-481.

McGovern P.E., 2003. Ancient Wine: The Search for the Origins of Viniculture. Princeton Univ. Press, Princeton.

Mendes-Pinto M.M., 2009. Carotenoid breakdown products - the norisoprenoids - in wine aroma. Archives of Biochemistry and Biophysics 483, 236-45.

Nguyen H.V., Lepingle A., Gaillardin C.A., 2000. Molecular typing demonstrates homogeneity of *Saccharomyces uvarum* strains and reveals the existence of hybrids between *S. uvarum* and *S.*

cerevisiae, including the *S. bayanus* type strain CBS 380. Systematic and Applied Microbiology 23, 71-85.

Noble J., Sanchez I., Blondin B., 2015. Identification of new *Saccharomyces cerevisiae* variants of the *MET2* and *SKP2* genes controlling the sulfur assimilation pathway and the production of undesirable sulfur compounds during alcoholic fermentation. Microbial Cell Factory 14, 68.

Pasteur, L., 1860. Mémoire sur la fermentation alcoolique. Ann Chim Phys 58: 323-426.

Renault P., Coulon J., de Revel G., Barbe J.C., Bely M., 2015. Increase of fruity aroma during mixed *T. delbrueckii/S. cerevisiae* wine fermentation is linked to specific esters enhancement. International Journal of Food Microbiology 207, 40-48.

Roncoroni M., Santiago M., Hooks D.O., Moroney S., Harsch M.J., Lee S.A., Richards K.D., Nicolau L., Gardner R.C., 2011. The yeast *IRC7* gene encodes a β-lyase responsible for production of the varietal thiol 4-mercapto-4-methylpentan-2-one in wine. Food Microbiology 28, 926-35.

Sadoudi M., Tourdot-Maréchal R., Rousseaux S., Steyer D., Gallardo-Chacón J.J., Ballester J., Vichi S., Guérin-Schneider R., Caixach J., Alexandre H., 2012. Yeast-yeast interactions revealed by aromatic profile analysis of Sauvignon Blanc wine fermented by single or co-culture of non-Saccharomyces and Saccharomyces yeasts. Food Microbiology 32, 243-53.

Salinas F., Cubillos F.A., Soto D., Garcia V., Bergstrom A., Warringer J., Ganga A., Louis E., Liti G., Martinez C., 2012. The genetic basis of natural variation in oenological traits in *Saccharomyces cerevisiae*. PLoS One 7: e49640.

Segurel M.A., Razungles A.J., Riou C., Salles M., Baumes R.L., 2004. Contribution of dimethyl sulfide to the aroma of Syrah and Grenache Noir wines and estimation of its potential in grapes of these varieties. Journal of Agricultural and Food Chemistry 52, 7084-7093.

Sipiczki M., 2008. Interspecies hybridization and recombination in *Saccharomyces* wine yeasts. FEMS Yeast Research 8, 996–1007.

Steyer D., Ambroset C., Brion C., Claudel P., Delobel P., Sanchez I., Erny C., Blondin B., Karst F., Legras J-L., 2012. QTL mapping of the production of wine aroma compounds by yeast. BMC Genomics 13, 573.

Steyer D., Erny C., Claudel P., Riveill G., Karst F., Legras J-L., 2013. Genetic analysis of geraniol metabolism during fermentation. Food Microbiology 33, 228–34.

Swiegers J.H., Bartowsky E.J., Henschke P.A., Pretorius I.S., 2005 Yeast and bacterial modulation of wine aroma and flavor. Australian Journal of Grape and Wine Research 11, 139-173.

Swiegers J.H., Pretorius I.S., 2005. Yeast modulation of wine flavor. Advances in Applied Microbiology 57, 131-17

Swiegers J.H., Kievit R.L., Siebert T., Lattey K.A., Bramley B.R., Francis I.L., King E.S., Pretorius I.S., 2009. The influence of yeast on the aroma of Sauvignon Blanc wine. Food Microbiology 26, 204-211.

Tilloy V., Ortiz-Julien A., Dequin S., 2014. Reduction of ethanol yield and improvement of glycerol formation by adaptive evolution of the wine yeast *Saccharomyces cerevisiae* under hyperosmotic conditions. Applied and Environmental Microbiology 80, 2623-32.

Tominaga T., Baltenweck-Guyot R., Peyrot Des Gachons C., Dubourdieu D., 2000. Contribution of Volatile Thiols to the Aromas of White Wines Made From Several *Vitis vinifera* Grape Varieties. American Journal of Enology and Viticulture 51, 178-181.

Publication 2

The contribution to this publication consisted in the qualitative analysis of high resolution mass spectrometry data and its graphic representation.

Targeted filtering reduces the complexity of UHPLC-Orbitrap-HRMS data to decipher polyphenol polymerization

Anna Vallverdú-Queralt[a*], Emmanuelle Meudec[a], **Matthias Eder**[a], Rosa M. Lamuela-Raventos[b,c], Nicolas Sommerer[a] and Véronique Cheynier[a]

[a] SPO, INRA, Montpellier Supagro, Université de Montpellier, 2, place Viala, 34060 Montpellier, France.

[b] Nutrition and Food Science Department, XaRTA, INSA, Pharmacy School, Av Joan XXIII s/n, University of Barcelona, Barcelona, Spain

[c] CIBER Fisiopatología de la Obesidad y la Nutrición (CIBEROBN), Madrid, Spain

Corresponding author

Dr. Anna Vallverdu-Queralt; SPO, INRA, Montpellier Supagro, Université de Montpellier, 2, place Viala, 34060 Montpellier, France; Phone: +33 499612656; avallverdu@ub.edu

Food Chemistry 227 (2017) 255–263

Contents lists available at ScienceDirect

Food Chemistry

journal homepage: www.elsevier.com/locate/foodchem

Targeted filtering reduces the complexity of UHPLC-Orbitrap-HRMS data to decipher polyphenol polymerization

Anna Vallverdú-Queralt [a,*], Emmanuelle Meudec [a], Matthias Eder [a], Rosa M. Lamuela-Raventos [b,c], Nicolas Sommerer [a], Véronique Cheynier [a]

[a] SPO, INRA, Montpellier Supagro, Université de Montpellier, 2, place Viala, 34060 Montpellier, France
[b] Nutrition and Food Science Department, XaRTA, INSA, Pharmacy School, Av Joan XXIII s/n, University of Barcelona, Barcelona, Spain
[c] CIBER Fisiopatología de la Obesidad y la Nutrición (CIBEROBN), Madrid, Spain

ARTICLE INFO

Article history:
Received 4 August 2016
Received in revised form 20 December 2016
Accepted 21 January 2017
Available online 22 January 2017

Keywords:
Petroleomics-inspired strategy
Van Krevelen diagrams
Kendrick mass defect plots
Epicatechin
Orbitrap-HRMS

ABSTRACT

UHPLC-LTQ-Orbitrap-high resolution mass spectrometry (HRMS) was applied to investigate complex polymeric polyphenols, before and after acid-catalysed depolymerisation in the presence of a nucleophile (phloroglucinol). Reaction products of (−)-epicatechin with acetaldehyde formed in model solution were selected for a proof-of concept experiment. The complexity of the UHPLC-HRMS dataset obtained after 4 h incubation was reduced with petroleomics-inspired strategies using Van Krevelen diagrams and modified Kendrick mass defect filtering targeting ethyl-epicatechin ($C_{17}H_{16}O_6$) units. Combining these approaches with mass fragmentation and phloroglucinolysis allowed us to describe reaction of epicatechin and acetaldehyde. More than 65 compounds were found, including the homogeneous bridged derivatives (up to the undecamer), vinyl and ethanol adducts, and xanthene and xanthylium salt derivatives which were identified for the first time.

1. Introduction

Plant polyphenols are structurally diverse and complex (Andersen & Markham, 2006; Cheynier, 2005). The presence of several building blocks and several linkage types and positions in polymeric structures leads to hundreds of potential compounds, including various reaction products formed in processed foods and beverages, such as tea or wine (Tarascou et al., 2010). The main challenge lies in the detection of thousands of known and unknown components in a wide range of chemistries, molecular masses, dynamic concentration range, and MS responses, ideally in a single analysis.

Polyphenol composition is usually analyzed by high performance chromatography (HPLC). Equipment development has enabled improvement of method resolution, sensitivity, and selectivity and given access to analysis of long lists of compounds, including some present in trace amounts. Sensitive and selective methods based on Ultra High Performance Liquid Chromatography coupled to triple-quadrupole Mass Spectrometry (UHPLC-QqQ-MS) enable quantification of large numbers of polyphenols in wine (Lambert et al., 2015). However, such targeted approaches need to be completed by untargeted strategies to fully describe the polyphenol composition of complex matrices such as plant-based foods and beverages (Hall, 2006; Vallverdu-Queralt, Medina-Remon, Casals-Ribes, Amat, & Lamuela-Raventos, 2011, Vallverdú-Queralt et al., 2013). High resolution mass spectrometry (HRMS) has emerged as a powerful tool for metabolomics, i.e. extensive description of all low molecular weight metabolites in a given sample. HRMS provides exact mass measurements thus leading to elemental composition assignment of molecules which is an essential step for compound identification (Gougeon et al., 2009). HRMS analysis of wines has revealed markers or signatures of the geographical origin of the wine and oak cooperage (Gougeon et al., 2009) and of the grape variety (Delcambre & Saucier, 2012). Numerous signals could be attributed to elemental compositions but only a few corresponding to structures reported in data bases or described earlier in the literature were identified. Comparison of storage conditions, including oxygen exposure (Arapitsas et al., 2012) and cellar or house storage (Arapitsas, Speri, Angeli, Perenzoni, & Mattivi, 2014) using HRMS enabled detection and tentative identification of new compounds arising from reactions of phenolic compounds in wine. Nevertheless, these molecules represent only the emerged part of the iceberg while the immerged part is composed of complex polyphenols, including polymers.

* Corresponding author at: SPO, INRA, Montpellier Supagro, Université de Montpellier, 2, place Viala, 34060 Montpellier, France.
E-mail address: avallverdu@ub.edu (A. Vallverdú-Queralt).

http://dx.doi.org/10.1016/j.foodchem.2017.01.106
0308-8146/

Mass defect filtering techniques inspired from Kendrick mass plot (Kendrick, 1963) have been proposed to simplify interpretation of electrospray ionization HRMS data obtained from complex mixtures of phenolic compounds such as thearubigins (Kuhnert, 2010), or lignin degradation products (Dier, Egele, Fossog, Hempelmann, & Dietrich, 2016). These studies provided new information on the sample composition and cascades of reactions involved in the studied oxidative polymerization and degradation processes. However, only low molecular weight compounds ($m/z < 1000$) were detected. As larger species (up to 2100) have been detected by MALDI-TOF MS in the thearubigin samples (Kuhnert, 2010), this limit is likely due to the domination of electrospray ionization spectra by lower molecular weight compounds which is well established for other phenolic polymers such as tannins (Mané et al., 2007; Mouls, Mazauric, Sommerer, Fulcrand, & Mazerolles, 2011; Yanagida, Kanda, Shoji, Ohnishi-Kameyama, & Nagata, 1999). Introducing a LC separation step before the HRMS analysis is expected to limit the in source competition and ionization suppression responsible for this phenomenon. Moreover HRMS analysis is typically performed using Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometry, as it offers better mass resolution than any other mass analyzer. However, the duty cycle of a typical high resolution FT-ICR experiment is too long for routine analysis with optimized UHPLC narrow chromatographic peaks. Moreover, the high costs associated with FT-ICR limit its use to a small number of research laboratories. LTQ-Orbitrap, presenting significantly lower costs relative to typical FT-ICR instruments, and thus more accessible, provides mass resolution inferior to the modern FT-ICR instrument but still sufficient to assign molecular formulas to individual components present in mixtures.

The aim of the work is to develop an original strategy combining non targeted UHPLC-LTQ-Orbitrap-HRMS analysis to unravel the composition of complex polyphenol mixtures. The products arising from reactions of epicatechin with acetaldehyde have been selected as a model system to generate tannin-like complexity for this proof of concept study. Such reactions are known to occur in wine (Saucier, Little, & Glories, 1997) but also during fruit maturation (Tanaka, Takahashi, Kouno, & Nonaka, 1994; Tarascou et al., 2011), resulting in color and taste changes. Studies in model solutions have demonstrated condensation of acetaldehyde with flavonoids such as catechin and epicatechin and identified the first products of these reactions, *i.e.* flavanol-acetaldehyde adducts and methyl-methine linked flavanol dimers and oligomers (often called ethyl-linked oligomers), (Es-Safi, Fulcrand, Cheynier, & Moutounet, 1999a; Fulcrand, Doco, Es-Safi, Cheynier, & Moutounet, 1996; Saucier et al., 1997). These products are highly unstable and rapidly proceed to a wide variety of products covering a large range of molecular weights through polymerization, cleavage, and recombination reactions. Recent studies have established additional structures arising from dimerization of vinylcatechin (Cruz et al., 2009), which can result from cleavage of ethyl-linked flavanol oligomers yields intermediate vinylflavanol adducts (Cruz, Borges, Silva, Mateus, & De Freitas, 2008; Es-Safi, Fulcrand, Cheynier, & Moutounet, 1999b). However, description of the reaction medium is still incomplete. In particular, the intense orange color developing during the reaction has remained unexplained although it likely contributes to the oxidative browning of white wines, correlated with the wine flavanol content (Cheynier, Rigaud, Souquet, Barillere, & Moutounet, 1989; Simpson, 1982).

We demonstrate a successful application of UHPLC-LTQ-Orbitrap-HRMS, assisted with petroleomics-inspired interpretation of the data, to resolve and identify the most abundant molecular formulas in the positive ion electrospray ionization mass spectrum of acetaldehyde-mediated epicatechin condensation reactions. Moreover, application of the same approach after phloroglucinolysis (acid-catalysed depolymerization in the presence of a nucleophilic agent) has provided additional structural information on polymer building blocks, as shown earlier for the ethyl-linked oligomers (Es-Safi et al., 1999a; Tanaka et al., 1994).

2. Materials and methods

2.1. Reagents and materials

Methanol (CH_3OH), ethanol (CH_3CH_2OH), formic acid (HCO_2H), ascorbic acid, sodium hydroxide (NaOH), hydrochloric acid (HCl) and acetic acid (CH_3COOH) were obtained from Prolabo (Fontenay S/Bois, France). (−)-Epicatechin, ammonium formate (NH_4HCO_2) and phloroglucinol were purchased from Sigma (St. Louis, MO). Acetaldehyde was obtained from Merck (Darmstadt, Germany). Deionized water was purified with a Milli-Q water system (Millipore, Bedford, MA) prior to use.

2.2. Reactions

Model solutions have been prepared as described by Es-Safi et al. (1999a). Briefly, an acidic solution was prepared with 17 μL of CH_3COOH and 50 μL of CH_3CH_2OH in 373 μL of H_2O, giving a pH value of 2.2. Six milligrams of (−)-epicatechin (40 mmol/L) and 60 μL of acetaldehyde (2135 mmol/L) were then added. The reactions were monitored by UHPLC coupled with an Orbitrap-HRMS detector.

2.3. Phloroglucinolysis

In order to gain a first insight into the nature of the acetaldehyde-induced cross-linking, the polymers were degraded by acid-catalysed depolymerisation in the presence of phloroglucinol as described by Fournand et al. (1999). Two aliquots of the previous extract (500 μL each) described in 2.2 were taken to dryness with a centrifugal solvent evaporator (Genevac, IPSWICH, UK) operated at 35 °C and 10 mbar. The aliquots were then dissolved in 250 μL of CH_3OH–HCl (0.2 mol/L) containing phloroglucinol (50 g/L) and L-ascorbic acid (10 g/L). After heating in a water bath at 50 °C for 20 min, the phloroglucinolysis reaction was stopped by adding an equal volume of NH_4HCO_2 (250 mmol/L) and the reaction medium was analyzed by UHPLC-Orbitrap-HRMS.

2.4. Isolation of the xanthylium salts by solid phase extraction (SPE)

The xanthylium salts were obtained by upscaling the model solution reaction (2 mL total volume) and incubating for 4 h. They were isolated by SPE performed with MCX cartridges (Waters Sep Pak®Vac 1 cc mixed-mode cation-exchange and reversed-phase solvent–30 mg). After conditioning with 2 mL of CH_3OH and 2 mL of H_2O at 2% HCO_2H, the cartridges were loaded with the samples (2 mL). A washing step was carried out with H_2O at 2% HCO_2H. Xanthylium salts were eluted and recovered with CH_3OH:NH_4OH (98:2 V/V). The elution fraction was then evaporated and reconstituted with H_2O adjusted at pH 1.5 with HCO_2H, before injection in the UHPLC-Orbitrap-HRMS.

2.5. Instrumentation UHPLC-LTQ-Orbitrap-HRMS and flow injection analysis (FIA)-LTQ-Orbitrap-HRMS

The analyses were carried out in an Accela UHPLC system from Thermo Fisher Scientific (San Jose, CA, USA) consisting of an autosampler, a quaternary pump, a vacuum degasser, and a thermostated column compartment.

For UHPLC analysis, chromatographic separation was performed on an Acquity BEH C_{18} column (150 mm length, 1 mm internal diameter, 1.7 μm particle size; Waters) at 35 °C. The mobile phase consisted of H_2O/HCO_2H (99/1, V/V) (solvent A) and CH_3OH/HCO_2H (99/1, V/V) (solvent B). The flow rate was 0.08 mL/min and the injection volume 0.5 μL. The elution program was as follows: isocratic with 2% B (1 min), 2–30% B (1–10 min), isocratic with 30% B (10–12 min), 30–75% B (12–25 min), 75–90% B (25–30 min), and isocratic with 90% B (30–35 min).

For FIA analysis (i.e. LC injection without column separation which allows very short duty cycle and sample injection every four minutes), a sample volume of 0.5 μL was injected using the Accela autoinjector set at 4 °C and was delivered with an isocratic mobile phase consisting of H_2O/MeOH (50:50, V/V) with 1% HCO_2H at a flow rate of 0.08 mL/min during 4 min.

The mass spectrometer was a linear ion trap-Orbitrap (LTQ-Orbitrap Velos) from Thermo Fisher Scientific (San Jose, CA, USA) equipped with an electrospray ionization (ESI) source. The specific conditions were adapted from a previous method reported by Vallverdú-Queralt et al. (2015)). Instrument control and data acquisition were performed with Xcalibur 3.0. software (Thermo Fisher Scientific). An external calibration for mass accuracy was carried out the day before the analysis according to the manufacturer's guidelines.

In order to treat the LC–MS data, a single mass spectrum was reconstructed from the chromatogram by extracting mass signals from the chromatogram in 3 segments of 10 min each. The chromatography was used only as a mean to reduce ionization competition in the ion source for an effective mass spectrometric detection.

2.6. UHPLC-HRMS metabolomics

LTQ-Orbitrap-HRMS data were treated with R software to compress large datasets into more discernable ones. The following chemical constraints were applied: composition was restricted to C, H, and O with double bond equivalents (DBE) between 7 and 65; O/C ratio $\leqslant 1$; 2.0 > H/C ratio > 0.5 (Tobias & Fiehn, 2007). The peak table output of R was then used for statistical analysis. The m/z signals were exported to Excel spreadsheets and experimental accurate masses were converted to a new mass scale inspired from the "Kendrick" mass scale (based on the 12C atomic mass as exactly 12 a.m.u.) and representing the epicatechin-ethyl units ($C_{17}H_{16}O_6$).

$$\textbf{Kendrick mass} = \text{experimental accurate mass} \times (316/316.0941)$$

where 316 u is the nominal mass of epicatechin-ethyl units ($C_{17}H_{16}O_6$) and 316.0941 u represents the exact mass of this repeating unit. Kendrick mass defects (KMD) were calculated as follows:

$$\text{KMD} = (\text{nominal mass} - \text{Kendrick mass})$$

The Kendrick graphic displayed nominal mass vs KMD.

The 2D dimensional van Krevelen Diagram provides a qualitative description of the data, displaying the hydrogen/carbon (H/C) vs. oxygen/carbon (O/C) ratios of the elemental formulas. 3D Van Krevelen diagram displaying the hydrogen/carbon (H/C) vs. oxygen/carbon (O/C) vs. molecular weight were also constructed.

Metabolite identification was performed by comparing retention times and accurate mass spectra for the parent ion mass and two fragment ions to those of the standard, when available. Tentative annotation of the chromatographic peaks was made according to the measured accurate masses and the isotopic patterns. Under these conditions, all proposed molecular formulas were estimated with mass errors below 2 milli mass units (mmu). The MS/MS spectra were compared with those of candidate compounds found in previous literature.

3. Results and discussion

Mass spectrometer resolution and mass accuracy are the primary considerations for determining whether an instrument suits its intended purpose. To obtain the widest coverage of molecules with the best resolution and mass accuracy, the classical ionization and transmission parameters were optimized: source voltage, sheath gas, auxiliary gas, sweep gas and capillary temperature. The epicatechin-acetaldehyde model solution obtained after 4 h of incubation was injected onto the UHPLC-LTQ-Orbitrap-HRMS under different conditions and HRMS in the positive ion mode and negative ion mode were recorded. Both modes yielded similar results but some signals were detected only in the positive ion mode which was thus selected for further analysis.

Restrictions for chemical compositions of measured m/z values were set with relative abundances >0.08% (corresponding to S/N > 3) in the m/z ranging from 250 to 2000 applying the restriction rules described in the Experimental section.

FIA was compared to UHPLC. Although FIA-MS provided fast analysis and method simplicity, only low molecular weight compounds (m/z < 1000) were detected, in comparison to UHPLC which enabled peak detection up to 2000. Therefore, introducing a UHPLC separation step before the HRMS analysis limited the in source competition and ionization suppression.

3.1. Metabolomics approach to decrease the complexity of UHPLC-HRMS data

The epicatechin-acetaldehyde model solutions were analyzed at time 0 and after 2, 4, and 24 h of incubation in the positive ion mode over the mass range m/z 220 to 2000.

Application of the above mentioned restriction rules led to selection of 69 peaks which can be unambiguously attributed to elemental CHO compositions below 2 mmu tolerance (see below for structural interpretation Table 1). Each peak corresponds to a series of isomers that are not differentiated by the applied approach. The majority of the compounds were detected as singly charged ions. However, compounds with molecular weights >2000 were detected as doubly or triply charged ions. Three replicate measurements were carried out after 4 h of incubation; relative standard deviations (RSD) of intensities <22.5% for lower intensity peaks, RSD <10% for peaks with higher intensity and RSD <3.6% for mmu error were obtained (Table 1), suggesting that the method is robust and reproducible for experimental mass measurements.

Annotations were first based on the reaction products described in the literature (Fig. 1). Condensation of epicatechin with acetaldehyde has been reported to give rise to ethanol-flavanol adducts and methyl-methine linked flavanol dimers (called ethyl-linked dimers), that polymerize further to ethyl-linked oligomers as both positions 6 and 8 of the flavanol are reactive (Fulcrand et al., 1996). These compounds can also yield vinyl-flavanol adducts (Cruz et al., 2009; Es-Safi et al., 1999a), which can cyclize to dihydrofuran-flavanol, be reduced to ethyl-flavanol, or dimerize (Cruz et al., 2009). Moreover, similar products formed by reaction of flavanols with other aldehydes proceed to xanthene and xanthylium salts (Es-Safi, Le Guernevé, Fulcrand, Cheynier, & Moutounet, 2000) but these molecules have not been reported after reaction with acetaldehyde. Ethyl bridges are susceptible to acid catalysed cleavage in the presence of a nucleophilic agent (Es-Safi et al., 1999a), yielding mono or di-substituted adducts, as illustrated in the case of phloroglucinolysis (Figure S1).

Table 1
Chemical compounds arising from acetaldehyde and epicatechin polymerization reactions.

Compound	Experimental mass	Mean Mmu error	RSD of Mmu error	Molecular formula	Mean intensity	RSD of mean intensity
1 Epicatechin	291.0863	0.2	1.1	C15H15O6	1.48E+06	2.9
2 Epicatechin-vinyl adduct	317.1020	0.4	0.9	C17H17O6	9.98E+06	3.5
3 Epicatechin-ethyl adduct	319.1181	1.2	0.4	C17H19O6	1.07E+04	19.4
4 Epicatechin-ethanol adduct	335.1125	0.5	1.1	C17H19O7	8.13E+06	5.7
5 Epicatechin-vinyl-vinyl adduct	343.1178	0.3	1.2	C19H19O6	8.95E+06	5.9
6 Epicatechin-vinyl-ethyl adduct	345.1338	0.9	1.5	C19H21O6	1.70E+04	20.3
7 Epicatechin-ethyl-ethyl adduct	347.1494	1.5	1.3	C19H23O6	1.07E+04	22.5
8 Epicatechin-vinyl-ethanol adduct	361.1281	0.6	1.4	C19H21O7	1.67E+05	8.4
9 Epicatechin-vinyl-vinyl-vinyl adduct	369.1332	0.6	2.5	C21H21O6	9.09E+05	5.3
10 Epicatechin-ethanol-ethanol adduct	379.1374	1.8	0.8	C19H23O8	1.01E+06	5.8
11 Epicatechin-vinyl-ethanol-ethanol adduct	405.1541	0.8	1.1	C21H25O8	1.22E+05	12.3
12 Epicatechin-ethanol-ethanol-ethanol adduct	423.1639	1.6	1.9	C21H27O9	1.57E+05	11.6
13 Xanthylium salt	587.1541	1.2	1.4	C32H27O11	2.68E+05	10.3
14 Xanthene	589.1701	0.8	1.8	C32H29O11	2.16E+06	7.1
15 Hydrated Xanthylium salt	605.1643	1.6	2.4	C32H29O12	3.83E+06	4.7
16 Epicatechin-ethyl-epicatechin	607.1810	0.6	1.0	C32H31O12	1.34E+07	2.1
17 Epicatechin-ethyl-epicatechin-vinyl adduct	633.1963	0.9	1.1	C34H33O12	9.34E+06	4.5
18 Epicatechin-ethyl-epicatechin-ethyl adduct	635.2128	1.3	1.2	C34H35O12	1.15E+04	15.1
19 Epicatechin-ethyl-epicatechin-ethanol adduct	651.2072	0.5	0.9	C34H35O13	3.81E+06	4.1
20 Epicatechin-ethyl-epicatechin-vinyl-vinyl adduct	659.2122	0.7	1.9	C36H35O12	6.32E+05	7.8
21 Epicatechin-ethyl-epicatechin-vinyl-ethyl adduct	661.2285	1.5	2.5	C36H37O12	1.25E+04	17.9
22 Epicatechin-ethyl-epicatechin-ethyl-ethyl adduct	663.2441	1.4	3.6	C36H39O12	1.47E+04	20.5
23 Unknown 1	665.1866	0.8	1.9	C34H33O14	1.01E+05	12.4
24 Epicatechin-ethyl-epicatechin-vinyl-ethanol adduct	677.2221	1.2	1.0	C36H37O13	6.62E+05	9.7
25 Epicatechin-ethyl-epicatechin-vinyl-vinyl-vinyl adduct	685.2276	0.9	2.5	C38H37O12	1.23E+06	5.7
26 Epicatechin-ethyl-epicatechin-ethanol-ethanol adduct	695.2334	0.6	2.6	C36H39O14	9.34E+05	10.1
27 Unknown 2	707.1970	0.9	1.7	C36H35O15	1.24E+05	14.5
28 Epicatechin-ethyl-epicatechin-vinyl-ethanol-ethanol adduct	721.2488	0.8	0.9	C38H41O14	4.53E+05	12.6
29 Epicatechin-ethyl-epicatechin-ethanol-ethanol-ethanol adduct	739.2592	1.0	1.8	C38H43O15	9.41E+04	15.9
30 Epicatechin-ethyl-xanthylium salt	903.2494	0.7	0.9	C49H43O17	1.29E+05	12.6
31 Epicatechin-ethyl-xanthene	905.2645	1.2	1.4	C49H45O17	9.59E+05	9.3
32 Epicatechin-ethyl-hydrated Xanthylium salt	921.2591	1.4	1.5	C49H45O18	1.01E+06	7.4
33 (Epicatechin-ethyl)$_2$-epicatechin	923.2756	0.6	2.1	C49H47O18	9.38E+06	5.4
34 (Epicatechin-ethyl)$_2$-epicatechin-vinyl adduct	949.2912	0.7	2.3	C51H49O18	4.57E+06	5.2
35 (Epicatechin-ethyl)$_2$-epicatechin-ethyl adduct	951.3075	1.3	2.7	C51H51O18	9.95E+03	20.3
36 (Epicatechin-ethyl)$_2$-epicatechin-ethanol adduct	967.3019	0.6	1.9	C51H51O19	1.81E+06	4.6
37 (Epicatechin-ethyl)$_2$-epicatechin-vinyl-vinyl adduct	975.3061	1.4	1.4	C53H51O18	4.32E+05	11.7
38 (Epicatechin-ethyl)$_2$-epicatechin-vinyl-ethyl adduct	977.3231	1.3	1.6	C53H53O18	1.05E+04	16.8
39 (Epicatechin-ethyl)$_2$-epicatechin-ethyl-ethyl adduct	979.3380	1.5	2.1	C53H55O18	1.27E+04	19.4
40 Unknown 3	981.2805	0.8	2.4	C51H49O20	1.13E+05	16.3
41 (Epicatechin-ethyl)$_2$-epicatechin-vinyl-ethanol adduct	993.3170	1.1	2.1	C53H53O19	3.65E+05	12.4
42 (Epicatechin-ethyl)$_2$-epicatechin-ethanol-ethanol adduct	1011.3281	0.7	0.8	C53H55O20	9.56E+05	10.3
43 (Epicatechin-ethyl)$_2$-epicatechinethanol-ethanol-ethanol adduct	1055.3543	0.5	1.3	C55H59O21	1.34E+05	11.6
44 Xanthene-ethyl-xanthene	1203.3488	1.0	2.5	C66H59O22	9.52E+03	19.8
45 Xanthylium salt-ethyl-hydrated xanthylium salt	1217.3269	2.0	2.4	C66H57O23	1.11E+04	17.7
46 Xanthene-ethyl-hydrated xanthylium salt	1219.3439	0.8	1.8	C66H59O23	9.24E+03	16.8
47 (Epicatechin-ethyl)$_2$-xanthylium salt	1219.3441	0.6	1.6	C66H59O23	9.68E+04	18.1
48 (Epicatechin-ethyl)$_2$-xanthene	1221.3591	1.2	3.4	C66H61O23	7.59E+05	14.5
49 Hydrated xanthylium salt-ethyl-hydrated xanthylium salt	1235.3385	1.1	3.5	C66H59O24	1.01E+04	17.6
50 (Epicatechin-ethyl)$_2$-hydrated xanthylium salt	1237.3541	1.2	2.1	C66H61O24	8.12E+05	14.3
51 (Epicatechin-ethyl)$_3$-epicatechin	1239.3703	0.5	2.4	C66H63O24	7.66E+06	5.7
52 (Epicatechin-ethyl)$_3$-epicatechin-vinyl adduct	1265.3861	0.5	2.8	C68H65O24	3.79E+06	6.1
53 (Epicatechin-ethyl)$_3$-epicatechin-ethanol adduct	1283.3965	0.7	2.1	C68H67O25	1.57E+06	4.3
54 (Epicatechin-ethyl)$_3$-epicatechin-vinyl-vinyl adduct	1291.4012	1.0	1.3	C70H67O24	3.77E+05	10.2
55 (Epicatechin-ethyl)$_3$-epicatechin-ethanol-ethanol adduct	1327.4228	0.6	0.9	C70H71O26	7.33E+05	9.2
56 (Epicatechin-ethyl)$_3$-xanthylium salt	1535.4380	1.4	0.6	C83H75O29	2.37E+04	15.7
57 (Epicatechin-ethyl)$_3$-xanthene	1537.4535	1.5	1.5	C83H77O29	4.57E+05	12.4
58 (Epicatechin-ethyl)$_3$-hydrated xanthylium salt	1553.4481	1.8	0.8	C83H77O30	5.12E+05	10.1
59 (Epicatechin-ethyl)$_4$-epicatechin	1555.4645	1.1	0.9	C83H79O30	5.68E+06	4.9
60 (Epicatechin-ethyl)$_4$-epicatechin-vinyl adduct	1581.4807	0.5	2.5	C85H81O30	1.60E+06	5.8
61 (Epicatechin-ethyl)$_4$-ethanol adduct	1599.4912	0.6	3.3	C85H83O31	9.78E+05	7.9
62 (Epicatechin-ethyl)$_4$-hydrated xanthylium salt	1869.5432	1.4	3.5	C100H93O36	2.42E+05	12.8
63 (Epicatechin-ethyl)$_5$-epicatechin	1871.5598	0.5	3.1	C100H95O36	3.50E+06	6.1
64 (Epicatechin-ethyl)$_5$-epicatechin-vinyl adduct	1897.5751	0.8	2.6	C102H97O36	3.97E+05	10.2

Table 1 (*continued*)

Compound	Experimental mass	Mean Mmu error	RSD of Mmu error	Molecular formula	Mean intensity	RSD of mean intensity
65 (Epicatechin-ethyl)$_6$-epicatechin	1094.3311*	0.6	1.1	C117H111O42	2.16E+06	7.3
66 (Epicatechin-ethyl)$_7$-epicatechin	1252.3785*	0.5	1.8	C134H127O48	9.59E+05	8.4
67 (Epicatechin-ethyl)$_8$-epicatechin	1410.4239*	1.5	1.7	C151H143O54	7.69E+05	8.1
68 (Epicatechin-ethyl)$_9$-epicatechin	1045.9795**	1.3	2.8	C168H159O60	3.91E+05	10.2
69 (Epicatechin-ethyl)$_{10}$-epicatechin	1151.3443**	1.6	1.9	C185H175O66	1.29E+05	11.5

* RSD: relative standard deviation detected as doubly charged ions.
** Detected as triply charged ions.

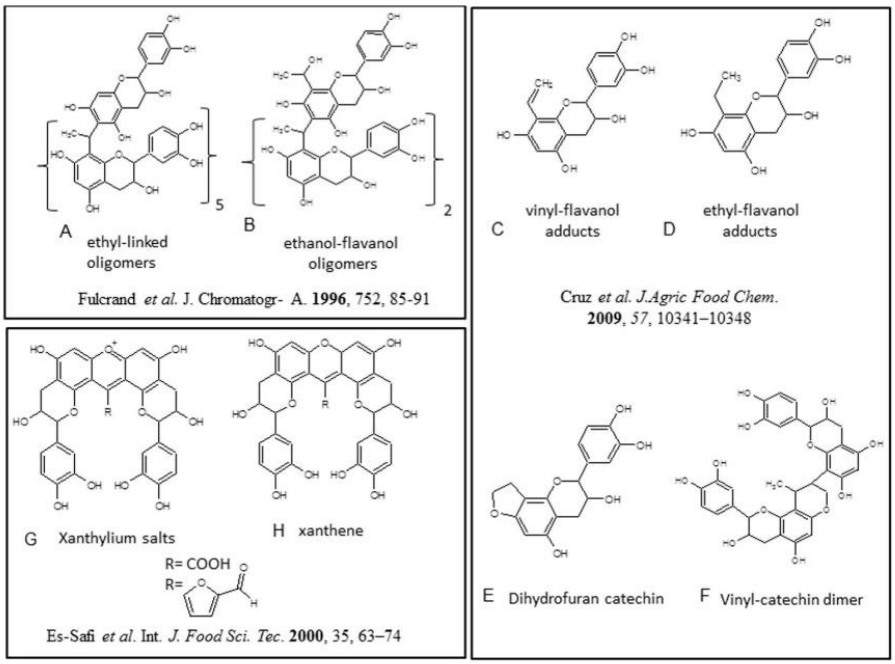


Fig. 1. Epicatechin-acetaldehyde condensation products, described in the literature.

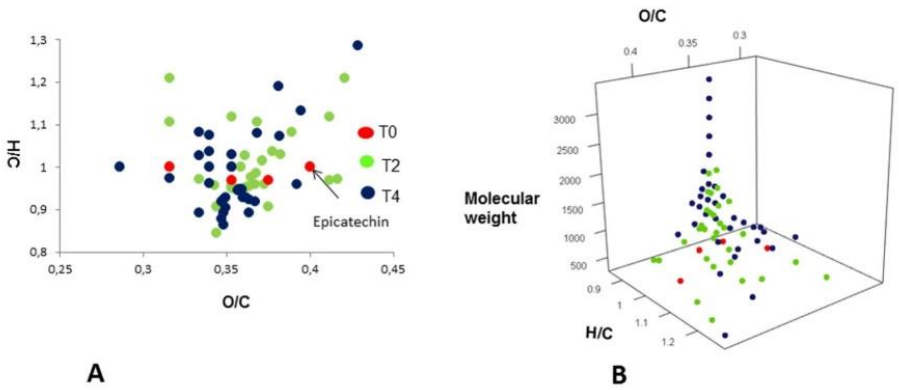


Fig. 2. The van Krevelen diagram performed just after addition of acetaldehyde, and after 2 h and 4 h of incubation. 2D (A), 3D (B).

The van Krevelen diagram performed just after addition of acetaldehyde and after 2 h and 4 h of incubation (Fig. 2A) showed that some reaction products formed immediately (red dots), but the medium was much more complex after 2 h (green dots) and even more after 4 h (blue dots). No additional compound was detected after longer incubation. The 3D van Krevelen diagram

(Fig. 2B), in which the third dimension presents the molecular weight, indicates that most of the increased complexity is due to further polymerisation towards higher polymers. This 3D van Krevelen representation of the reaction complexity displays a fractal pattern limited by compound ionization, HRMS dynamic range and resolving power. Thus, similar polymerization reactions can be observed and assigned at higher masses, as far as the mass spectrometer is able to ionize, transmit, separate and detect those larger molecules.

3.2. Mass defect filtering for detection and sorting of product series

Based on the assumption that most compounds contained epicatechin-ethyl units (316 a.m.u), a Kendrick plot with this mass increment was created (Fig. 3). Several series of compounds of the same class and type but different number of epicatechin-ethyl bridge units lining up on parallel horizontal lines were thus distinguished. Different classes of compounds are now readily identified, because the Kendrick mass defect for compounds of a given class is displaced vertically from those of other classes.

3.2.1. The homogeneous bridged derivatives of epicatechin

The most abundant series corresponded to the homogeneous bridged derivatives of epicatechin (n = 2 to 11 epicatechin units, Fig. 1A, compounds **1**, **16**, **33**, **51**, **59**, **63** and **65**–**69**) linked by ethyl bridges. The dimer, trimer, tetramer, pentamer, and hexamer were detected as singly charged ions at *m*/*z* 607.1810, 923.2756, 1239.3703, 1555.4645, and 1871.5598, respectively. The hexamer was also detected as doubly charged ion at 936.2839. Similarly, the heptamer, octamer, and nonamer were detected as doubly charged ions at 1094.3311 Th, 1252.3785 Th, and 1410.4239 Th, respectively. Lastly, the decamer and undecamer were detected as triply charged ions at 1045.9795 Th and 1151.3443 Th. The dimers were formed instantly after addition of acetaldehyde while trimers, tetramers, and pentamers were detected after 2 h of incubation and all others only after 4 h. Confirmation of the postulated structures was provided by fragmentation spectra. The characteristic fragment ions of these compounds were detected at *m*/*z* 317.1011 (epicatechin linked to an ethyl bridge), *m*/*z* 291.0845 (epicatechin), and ions corresponding to successive losses of ethyl-epicatechin units (−316.0934). The detection of oligomers beyond the hexamer species, reported here for the first time, is enabled by the high resolution and dynamic range offered by the HRMS equipment.

3.2.2. The ethanol-epicatechin adducts

Another important series corresponded to ethanol adducts on epicatechin monomer and ethyl-linked epicatechin dimer, trimer, tetramer, and pentamer detected at *m*/*z* 335.1125, 651.2072, 967.3019, 1283.3965, and 1599.4912 (Figure S2, compounds **4**, **19**, **36**, **53** and **61**). All of them yielded fragment ions attributed to $[M+H-18]^+$ (loss of water), $[M+H-44]^+$ (loss of CH_3-COH), $[M+H-152]^+$ (corresponding to retro-Diels Alder (RDA) fragmentation in the C-ring involving 1,3 scission), and $[M+H-290]^+$ (loss of an epicatechin unit), in agreement with the postulated structures. Fulcrand et al. (1996) had already shown the presence of ethanol adducts on monomer, dimer, and trimer, supporting the mechanism postulated by Timberlake and Bridle (1976), but it is the first time that adducts on tetramer and pentamer are detected, again due to the higher resolution and dynamic range of the HRMS detection.

Two other series corresponding to addition of two and three ethanol groups on epicatechin and ethyl-linked epicatechin oligomers were also visible on the Kendrick diagram established with the 316 mass increment. Thus the signals at *m*/*z* 379.1374, 695.2334, 1011.3281, and 1327.4228 (**10**, **26**, **42** and **55**) with fragment ions corresponding to $[M+H-18]^+$, $[M+H-18-18]^+$, $[M+H-44]^+$, $[M+H-44-44]^+$, $[M+H-152]^+$, and $[M+H-290]^+$ were attributed to two ethanol adducts on the monomer, dimer, trimer, and tetramer. Similarly, peaks detected at *m*/*z* 423.1639, 739.2592, and 1055.3543 (compounds **12**, **29** and **43**) appearing after 4 h of incubation, correspond to three ethanol adducts on the monomer, dimer, and trimer. As only two substitution positions are available on the molecules, this implies that at least two ethanol units are linked together (Figure S2). These motives may come from acetaldehyde oligomerization reactions.

3.2.3. The vinyl-epicatechin adducts

Ions detected at *m*/*z* 317.1020, 633.1963, 949.2912, 1265.3861, 1581.4807, and 1897.5751 are tentatively attributed to the vinyl adducts on the monomer, and ethyl-linked dimer, trimer, tetramer, pentamer, and hexamer (Figure S3, compounds **2**, **17**, **34**, **52**, **60** and **64**). The hexamer was also detected as doubly charged ion at 949.2916 Th. The ions at *m*/*z* 317.1020 and 633.1963 may also correspond respectively to the dihydrofuran formed by cyclisation of the vinyl-catechin and to the vinylcatechin dimer described earlier (Cruz et al., 2009). A vinylcatechin trimer structure has also been postulated for the ion detected at *m*/*z* 949.2916 and the other ions could similarly correspond to the equivalent tetramers, pentamers,

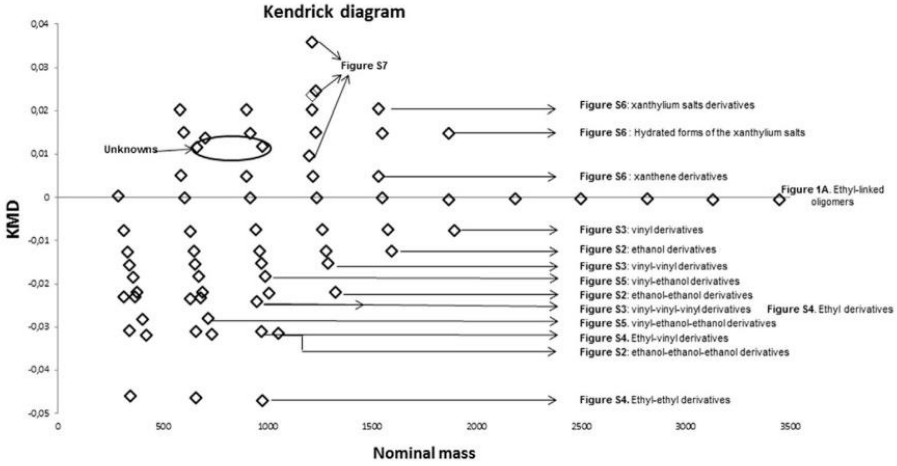


Fig. 3. Epicatechin-acetaldehyde condensation products plotted on a Kendrick diagram.

and hexamers. These ions yielded fragment ions $[M+H-18]^+$, $[M+H-26]^+$, $[M+H-152]^+$, $[M+H-316]^+$, and epicatechin (*m/z* 291.0845). Although the presence of the dihydrofuran derivatives cannot be ruled out and most of the fragments are consistent with either structure, the loss of 26 mass units is specific of vinyl derivatives. A similar pattern was shown for two vinyl adducts on the monomer, dimer, trimer, and tetramer at *m/z* 343.1178, 659.2122, 975.3061, and 1291.4012 (Figure S3, compounds **5**, **20**, **37** and **54**). The mass fragments of these ions were $[M+H-26]^+$, $[M+H-26-26]^+$, $[M+H-152]^+$, $[M+H-290]^+$, $[M+H-316]^+$, and *m/z* 291.0845. Similarly, three vinyl adducts on the monomer and dimer were found at *m/z* 369.1332 and 685.2276 (compounds **9** and **25**), after 4 h of incubation. As explained above for ethanol-ethanol-ethanol derivatives, the presence of three vinyl groups implies that at least two of them are linked together. These structures may arise from addition of acetaldehyde on a vinyl-flavanol unit.

Moreover, ethyl adducts on the monomer, and ethyl-linked dimer, trimer were also found at *m/z* 319.1181, 345.1338, 347.1494, 635.2128, 661.2285, 663.2441, 951.3075, 977.3231, 979.3380 (Figure S4, compounds **3**, **6**, **7**, **18**, **21**, **22**, **35**, **38** and **39**). Cruz et al. (2009) have already reported the presence of a *m/z* 319, which is consistent with the structure of a 8-(or 6-)-ethylcatechin. However, the other ethyl-adducts have never been reported before. Ethyl-derivatives come from the reduction of the corresponding vinyl-derivatives.

3.2.4. The vinyl-ethanol-epicatechin adducts

The vinyl-ethanol adduct intermediates were also detected on the monomer, dimer, and trimer, at *m/z* 361.1281, 677.2221, and 993.3170 (Figure S5, compounds **8**, **24** and **41**). The mass fragments of these ions were $[M+H-18]^+$, $[M+H-26]^+$, $[M+H-44]^+$, $[M+H-152]^+$ and the epicatechin ion at *m/z* 291.0845. A similar pattern was shown for ethanol-ethanol-vinyl adducts on the monomer and dimer appearing after 4 h of incubation and detected at *m/z* 405.1541 and *m/z* 721.2488 (compounds **11** and **28**). The mass fragments of these ions were the loss of $[M+H-26]^+$ and the ion at *m/z* 291.0845.

3.2.5. Xanthene and xanthylium derivatives

Three additional series corresponding to compounds unreported until now were detected in the Kendrick diagram (Fig. 3) at *m/z* 589.1701, 905.2645, 1221.3591, and 1537.4535 (compounds **14**, **31**, **48** and **57**), at *m/z* 587.1541, 903.2494, 1219.3441 and 1535.4380 (compounds **13**, **30**, **47** and **56**) at *m/z* 605.1643, 921.2591, 1237.3541, 1553.4481, and 1869.5432 (compounds **15**, **32**, **50**, **58** and **62**). This last ion was also detected as doubly charged ion at 935.2760 Th. The first two series are 18 and 20 mass units lower than the ethyl-bridged oligomers, suggesting that they result from these compounds, respectively by dehydration and by dehydration followed by an oxidation step. The loss of a water molecule may be achieved between the two 7-OH groups, giving a xanthene which oxidizes to a xanthylium salt, as described earlier for condensation products of flavan-3-ols with other aldehydes (Es-Safi et al., 2000). The xanthylium salts may be responsible for the brown color of the solutions, which is so far unexplained. The fragment ions of the compounds detected at *m/z* 589.1701, and 587.1541 could be attributed to the loss of water $[M+H-18]^+$, the loss of a hydroxyvinylphenol group ($C_8H_8O_3$) obtained by RDA fission $[M+H-152]^+$ and the loss of a methyl group $[M+H-14]^+$, in agreement with the proposed structures (Figure S6). The xanthylium salts were not detected in the negative ion mode, confirming the postulated cation structure. The third series was attributed to the hydrated form of the xanthylium salts (*m/z* 605.1643, 921.2591, 1237.3541, 1553.4481, and 1869.5432), on the basis of its fragmentation pattern (loss of water and RDA fragmentation).

Four other compounds, also unreported until now, were also found at *m/z* 1203.3488, 1217.3269, 1219.3439 and 1235.3385, and were tentatively attributed to xanthene-ethyl-xanthene, xanthylium-ethyl-hydrated xanthylium, xanthene-ethyl-hydrated xanthylium and hydrated xanthylium-ethyl-hydrated xanthylium, respectively (Figure S7, compounds **44**, **45**, **46** and **49**).

In order to corroborate the identity of the xanthylium salts, a SPE was carried out after 4 h of incubation. With a MCX phase, only positively charged species are properly retained. The eluted phase was evaporated and reconstituted with H_2O adjusted at pH 1.5 with HCO_2H. This solution had a brown color and contained the ion detected at *m/z* 587.1541 with the same fragmentation pattern as described above. That allowed us to corroborate the hypothesis of xanthylium salts structures.

*3.2.6. Unknown compounds (**23**, **27** and **40**)*

After 4 h of incubation, three other compounds were detected at *m/z* 665.1863, 707.1970, and 981.2805. These correspond to molecular formulas $C_{34}H_{33}O_{14}$, $C_{36}H_{35}O_{15}$, and $C_{51}H_{49}O_{20}$, respectively. Compounds **23** and **40** have a mass increment of 316 on the Kendrick diagram (Fig. 3) and differ from structures $C_{34}H_{33}O_{12}$ (**17**) and $C_{51}H_{49}O_{18}$ (**34**), i.e. (epicatechin-ethyl)$_n$-epicatechin vinyl adducts, with n = 1 and n = 2, respectively, by the presence of two additional oxygen atoms. Compound **27** contains three additional oxygen atoms compared to structure $C_{36}H_{35}O_{12}$ (**20**), epicatechin-ethyl-epicatechin-vinyl-vinyl adduct. These structural modifications suggest that vinyl units are needed to form these new compounds.

3.3. Data analysis after depolymerization using phloroglucinolysis

All these analyses were supported by the phloroglucinolysis experiments as illustrated by comparison of the van Krevelen diagram before (Fig. 4A) and after (Fig. 4B) phloroglucinolysis. The composition appeared much simpler after phloroglucinolysis (light blue dots disappeared and the intensity of dark blue dots increased in respect with the same sample before phloroglucinolysis), but new peaks, corresponding to the phloroglucinol adducts, were detected (red dots), which shifted, in general, to molecules with higher oxygen content.

After the phloroglucinolysis experiments, the compounds **1**–**15** (Table 1) were detected with higher intensity than before the phloroglucinolysis experiments (Fig. 4) while all larger molecular weight species were no longer detected. Additional products could be attributed to the derivatives resulting from addition of phloroglucinol onto the ethyl carbocation formed after acid catalysed cleavage of the ethyl bond. All these compounds show a characteristic fragment ion corresponding to the loss of the phloroglucinol group $[M+H-126]^+$. Thus, phloroglucinol-ethyl-phloroglucinol was detected at *m/z* 279.0866, as reported earlier (Drinkine, Lopes, Kennedy, Teissedre, & Saucier, 2007). This molecule may result from cleavage of both flavanol-ethyl bonds (Figure S1, 1 + 2) or form directly from acetaldehyde. The phloroglucinol-ethyl-epicatechin and the phloroglucinol-ethyl-epicatechin-ethyl-phloroglucinol were detected at *m/z* 443.1333 and 595.1815, with a fragment ion at 317.1021 and 469.1491, respectively. The latter arises from phloroglucinolysis of trimers or higher polymers. The mass signal detected at *m/z* 469.1491, with a fragment ion at 343.1179 may correspond to the phloroglucinol–ethyl adducts of vinyl-epicatechin or dihydrofuran-epicatechin or arise from phloroglucinolysis of the vinylcatechin dimer described by Cruz et al. (2009). Other mass signals could be attributed to the phloroglucinol-ethyl derivatives of epicatechin (*m/z* 471.1655), ethanol-epicatechin (*m/z* 487.1598), vinyl-vinyl-epicatechin (*m/z* 495.1649), vinyl-ethyl-epicatechin (*m/z* 497.1811), ethyl-ethyl-epicatechin (*m/z* 499.1968), vinyl-ethanol-epicatechin (*m/z*

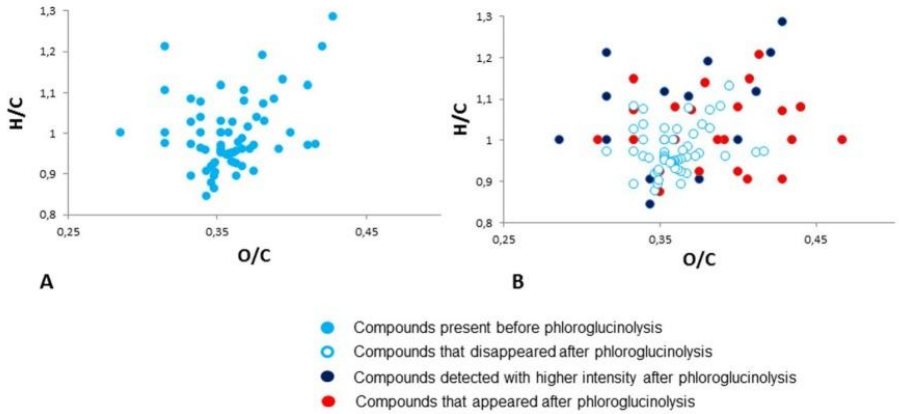


Fig. 4. The van Krevelen diagram before (A) and after (B) phloroglucinolysis.

513.1755), vinyl-vinyl-vinyl-epicatechin (*m*/*z* 521.1806), ethanol-ethanol-epicatechin (*m*/*z* 531.1860), ethanol-ethanol-ethanol-epicatechin (*m*/*z* 575.2123). The presence of several vinyl and/or ethanol in the structures detected after phloroglucinolysis allowed us to confirm the structures presented in Figures S2–7. In addition, new species detected at *m*/*z* 741.2177, *m*/*z* 739.2021, and *m*/*z* 757.2126 could be attributed to the corresponding phloroglucinol–ethyl adducts of compounds **13**–**15**, further supporting the postulated structures (compounds **30**–**32**, **44**–**50**, **56**–**58** and **62**).

3.4. General cascade reaction mechanism based on petroleomics-inspired data interpretation

With all this data, a cascade reaction mechanism, involving successive polymerization, oxidation, reduction, cyclisation, dehydration, and hydration steps, for the formation of epicatechin-acetaldehyde oligomers was postulated (Fig. 5).

Each polymerisation step involves addition of a flavanol-derivative (through its nucleophilic C6 or C8 position) onto a protonated acetaldehyde, generating an ethanol adduct intermediate. This epicatechin-(ethyl-epicatechin)$_n$-ethanol intermediate can be reduced to the corresponding vinyl (which can isomerize to the furan derivatives or polymerize) (Cruz et al., 2009), and then to the ethyl derivatives. They can also undergo protonation and react with another flavanol A-ring to form epicatechin-(ethyl-epicatechin)$_{n+1}$.

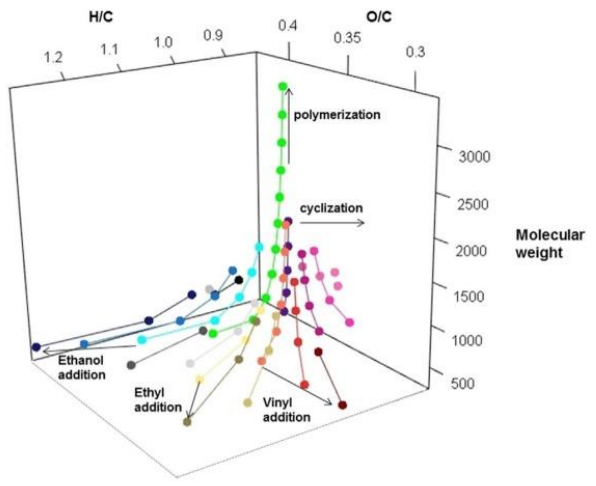


Fig. 5. Cascade reaction mechanism plotted on a 3D Van Krevelen diagram. Labels are provided in supplementary data Figure S8.

In addition, two adjacent flavanol units in ethyl-linked oligomers can cyclize to form xanthenes that oxidize to xanthylium salts which can yield the corresponding hydrated forms.

Due to acetaldehyde oligomerization reactions, ethanol-ethanol-epicatechin adducts and ethanol-ethanol-ethanol-epicatechin adducts can also be formed. Vinyl-vinyl epicatechin adducts and vinyl-vinyl-vinyl-epicatechin adducts result from the addition of acetaldehyde on a vinyl-flavanol unit or from reduction of the ethanol adducts. These vinyl-vinyl-flavanol polymers can be reduced further to ethyl-vinyl-flavanol polymers or ethyl-ethyl-flavanol polymers. All these molecules can be involved in the reaction cascade through the residual free C6 or C8 of the flavanol moieties.

In conclusion, UHPLC-LTQ-Orbitrap-HRMS, which is now a "standard" instrument in metabolomics laboratories, provides access to a large data set but needs targeted filtering to reduce the complexity of polyphenol polymerization spectra. Combining petroleomics-inspired interpretation of UHPLC-HRMS data, in particular using the Kendrick-inspired mass defect filtering approach targeting the epicatechin-ethyl units ($C_{17}H_{16}O_6$) and Van Krevelen diagrams, with mass fragmentation and phloroglucinolysis allowed us to describe the products resulting from reaction of epicatechin and acetaldehyde.

This approach gave access to larger molecular weight ethyl-linked epicatechin oligomers from the heptamer to the undecamer, so far unreported. Several xanthylium salts and xanthenes were reported for the first time. Consequently, in this work, a structural model for acetaldehyde-mediated reaction cascades involving formation of ethanol adducts, vinyl adducts, ethyl bridges, loss of water molecules to form xanthylium salts was proposed. Although the reaction conditions are not similar to those found in wine (only epicatechin, large excess of acetaldehyde), this study paves the way towards detection of more complex molecules formed in wine through similar mechanisms.

Authors contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

Acknowledgements

We thank the scientific platform of Barcelona for the technical support. Anna Vallverdú-Queralt is grateful to the Alfonso Martín Escudero Foundation for the postdoctoral fellowship for carrying out research abroad. Financial support from the Instituto de Salud Carlos III, ISCIII (CIBEROBN), GIS IBiSA (Infrastructures en Biologie Santé et Agronomie), Région Languedoc Roussillon, and INRA CNOC for funding of the UHPLC-MS equipment are also acknowledged.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.foodchem.2017.01.106.

References

Andersen, Ø. M., & Markham, K. R. (Eds.). (2006). *Flavonoids: Chemistry, biochemistry and applications* (pp. 1–34). Boca Raton, New York: CRC Press.

Arapitsas, A., Scholz, M., Vrhovsek, U., Di Blasi, S., Bartolini, A., Masuero, D., et al. (2012). A metabolomic approach to the study of wine micro-oxygenation. *PLoS One, 7*, e37783–e37788.

Arapitsas, P., Speri, G., Angeli, A., Perenzoni, D., & Mattivi, F. (2014). The influence of storage on the "chemical age" of red wines. *Metabolomics, 10*, 816–832.

Cheynier, V. (2005). Polyphenols in foods are more complex than often thought. *The American Journal of Clinical Nutrition, 81*, 223S–229S.

Cheynier, V., Rigaud, J., Souquet, J.-M., Barillere, J.-M., & Moutounet, M. (1989). Effect of pomace contact and hyperoxidation on the phenolic composition and quality of Grenache and Chardonnay wines. *American Journal of Enology and Viticulture, 40*, 36–42.

Cruz, L., Bras, R. F., Teixeira, N., Fernandes, A., Mateus, N., Ramos, J., et al. (2009). Synthesis and structural characterization of two diasteroisomers of vinylcatechin dimers. *Journal of Agricultural and Food Chemistry, 57*, 10341–10348.

Cruz, L., Borges, E., Silva, A. M. S., Mateus, N., & De Freitas, V. (2008). Synthesis of a new (+)-catechin-derived compound: 8-vinylcatechin. *Letters in Organic Chemistry, 5*, 530–536.

Delcambre, A., & Saucier, C. (2012). Identification of new flavan-3-ol monoglycosides by UHPLC-ESI-Q-TOF in grapes and wine. *Journal of Mass Spectrometry, 47*, 727–736.

Dier, T., Egele, K., Fossog, V., Hempelmann, R., & Dietrich, A. (2016). Enhanced mass defect filtering to simplify and classify complex mixtures of lignin degradation products. *Analytical Chemistry, 88*, 1328–1335.

Drinkine, J., Lopes, P., Kennedy, J., Teissedre, P. L., & Saucier, C. (2007). Ethylidene-bridged flavan-3-ols in red wine and correlation with wine age. *Journal of Agricultural and Food Chemistry, 55*, 1109–1116.

Es-Safi, N.-E., Fulcrand, H., Cheynier, V., & Moutounet, M. (1999a). Competition between (+)-catechin and (−)-epicatechin in acetaldehyde- induced polymerization of flavanols. *Journal of Agricultural and Food Chemistry, 47*, 2088–2095.

Es-Safi, N. E., Fulcrand, H., Cheynier, V., & Moutounet, M. (1999b). Studies on the acetaldehyde-induced condensation of (−)-epicatechin and malvidin 3-*O*-glucoside in a model solution system. *Journal of Agricultural and Food Chemistry, 47*, 2096–2102.

Es-Safi, N. E., Le Guernevé, C., Fulcrand, E., Cheynier, V., & Moutounet, M. (2000). Xanthylium salts involved in wine color evolution. *International Journal of Food Science & Technology, 35*, 63–74.

Fournand, D., Vicens, A., Sidhoum, L., Souquet, J. M., Moutounet, M., & Cheynier, V. (1999). Accumulation and extractability of grape skin tannins and anthocyanins at different advanced physiological stages. *Journal of Agricultural and Food Chemistry, 54*, 7331–7338.

Fulcrand, H., Doco, T., Es-Safi, E. N., Cheynier, V., & Moutounet, M. J. (1996). Study of the acetaldehyde induced polymerisation of flavan-3-ols by liquid chromatography-ion spray mass spectrometry. *Journal Chromatography A, 752*, 85–91.

Gougeon, R. D., Lucio, M., Frommberger, M., Peyron, D., Chassagne, D., Alexandre, H., et al. (2009). The chemodiversity of wines can reveal a metabologeography expression of cooperage oak wood. *Proceedings of the National Academy of Sciences, 106*, 9174–9179.

Hall, R. D. (2006). Plant metabolomics: from holistic hope, to hype, to hot topic. *New Phytologist, 169*, 453–468.

Kendrick, E. (1963). A mass scale based on CH2 = 14.0000 for high resolution mass spectrometry of organic compounds. *Analytical Chemistry, 35*, 2146–2154.

Kuhnert, N. (2010). Unraveling the structure of the black tea thearubigins. *Archives of Biochemistry and Biophysics, 501*, 37–51.

Lambert, M., Meudec, E., Verbaere, A., Mazerolles, G., Wirth, J., Masson, G., et al. (2015). A high-throughput UHPLC-QqQ-MS method for polyphenol profiling in Rosé Wines. *Molecules, 2015*(20), 7890–7914.

Mané, C., Sommerer, N., Yalcin, T., Cheynier, V., Cole, R. B., & Fulcrand, H. (2007). Assessment of the molecular weight distribution of tannin fractions through MALDI-TOF MS analysis of protein−tannin complexes. *Analytical Chemistry, 79*, 2239–2248.

Mouls, L., Mazauric, J.-P., Sommerer, N., Fulcrand, H., & Mazerolles, G. (2011). Comprehensive study of condensed tannins by ESI mass spectrometry: average degree of polymerisation and polymer distribution determination from mass spectra. *Analytical and Bioanalytical Chemistry, 400*, 613–623.

Saucier, C., Little, D., & Glories, Y. (1997). First evidence of acetaldehyde-flavanol condensation Products in red wine. *American Journal Of Enology And Viticulture, 48*, 370–373.

Simpson, R. F. (1982). Factors affecting oxidative browning of white wine. *Vitis, 21*, 233–239.

Tanaka, T., Takahashi, R., Kouno, I., & Nonaka, G.-I. J. (1994). Chemical evidence for the de-astringency (insolubilization of tannins) of persimmon fruit. *Journal of the Chemical Society, Perkin, 20*, 3013–3022.

Tarascou, I., Souquet, J. M., Mazauric, J. P., Carrillo, S., Coq, S., Canon, F., et al. (2010). The hidden face of food phenolic composition. *Archives of Biochemistry and Biophysics, 501*, 16–22.

Tarascou, I., Mazauric, J.-P., Meudec, E., Souquet, J.-M., Cunningham, D., Nojeim, S., et al. (2011). Characterisation of genuine and derived cranberry proanthocyanidins by LC-ESI-MS. *Food Chemistry, 128*, 802–810.

Timberlake, C. F., & Bridle, P. (1976). Interactions between anthocyanins, phenolic compounds and acetaldehyde and their significance in red wines. *American Journal of Enology and Viticulture, 27*, 97–105.

Tobias, K., & Fiehn, O. (2007). Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. *BMC Bioinformatics, 8*, 105.

Vallverdu-Queralt, A., Medina-Remon, A., Casals-Ribes, I., Amat, M., & Lamuela-Raventos, R. M. (2011). A metabolomic approach differentiates between conventional and organic ketchups. *Journal of Agricultural and Food Chemistry, 59*, 11703–11710.

Vallverdú-Queralt, A., Oms-Oliu, G., Odriozola-Serrano, I., Lamuela-Raventós, R. M., Martín-Belloso, O., & Elez-Martínez, P. (2013). Metabolite profiling of phenolic and carotenoid contents in tomatoes after moderate-intensity pulsed electric field treatments. *Food Chemistry, 136*, 199–205.

Vallverdú-Queralt, A., Boix, N., Piqué, E., Gómez-Catalan, J., Medina-Remon, A., Sasot, G., et al. (2015). Identification of phenolic compounds in red wine extract samples and zebrafish embryos by HPLC-ESI-LTQ-Orbitrap-MS. *Food Chemistry, 181*, 146–151.

Yanagida, A., Kanda, T., Shoji, T., Ohnishi-Kameyama, M., & Nagata, N. (1999). Fractionation of apple procyanidins by size-exclusion chromatography. *Journal Chromatography A, 855*, 181–190.

Publication 3

The contribution to this publication consisted in the qualitative analysis of high resolution mass spectrometry data.

The hidden face of wine polyphenol polymerization highlighted by high resolution mass spectrometry

Anna Vallverdú-Queralt[a*], Emmanuelle Meudec[a], **Matthias Eder**[a], Rosa M. Lamuela-Raventos[b,c], Nicolas Sommerer[a] and Véronique Cheynier[a]

[a] SPO, INRA, Montpellier Supagro, Université de Montpellier, 2, place Viala, 34060 Montpellier, France.

[b] Nutrition and Food Science Department, XaRTA, INSA, Pharmacy School, Av Joan XXIII s/n, University of Barcelona, Barcelona, Spain

[c] CIBER Fisiopatología de la Obesidad y la Nutrición (CIBEROBN), Madrid, Spain

Corresponding author

Dr. Anna Vallverdu-Queralt; SPO, INRA, Montpellier Supagro, Université de Montpellier, 2, place Viala, 34060 Montpellier, France; Phone: +33 499612656; avallverdu@ub.edu

ChemPubSoc Europe

DOI: 10.1002/open.201700044

Open Access ChemistryOPEN Communications

The Hidden Face of Wine Polyphenol Polymerization Highlighted by High-Resolution Mass Spectrometry

Anna Vallverdú-Queralt,*[a] Emmanuelle Meudec,[a] Matthias Eder,[a] Rosa M. Lamuela-Raventos,[b, c] Nicolas Sommerer,[a] and Véronique Cheynier[a]

Polyphenols, including tannins and red anthocyanin pigments, are responsible for the color, taste, and beneficial health properties of plant-derived foods and beverages, especially in red wines. Known compounds represent only the emerged part of the "wine polyphenol iceberg". It is believed that the immersed part results from complex cascades of reactions involving grape polyphenols and yeast metabolites. We used a non-targeted strategy based on high-resolution mass spectrometry and Kendrick mass defect plots to explore this hypothesis. Reactions of acetaldehyde, epicatechin, and malvidin-3-*O*-glucoside, representing yeast metabolites, tannins, and anthocyanins, respectively, were selected for a proof-of-concept experiment. A series of compounds including expected and so-far-unknown structures were detected. Random polymerization involving both the original substrates and intermediate products resulting from cascade reactions was demonstrated.

Polyphenols attract considerable interest because of their ubiquitous occurrence within the plant kingdom and their numerous important properties, related to their high structural diversity.[1] Owing to the acidic character of their hydroxyl groups and the nucleophilic properties of the phenolic rings, these molecules are highly reactive and undergo various types of reactions in the course of food processing and storage.[2] The major polyphenols in fruits and especially in grape are anthocyanins, the pigments of red and dark cultivars, and flavan-3-ols, including monomers (e.g. epicatechin) and proanthocyanidin oligomers and polymers, commonly called tannins,[1,2] which confer astringency.[3] Their reactions, further increasing structural diversity, are responsible for the color and taste changes occurring during wine making and ageing, and may influence the health benefits of wine.[4] A number of reaction products have been unraveled in red wine.[5–7] In particular, oxygen exposure has been reported to promote the accumulation of products derived from reaction of anthocyanins and/or flavan-3-ols with acetaldehyde, a yeast metabolite and oxidation product of ethanol present in a large amount in wine.[8,9] Investigations in model solutions have shown that reactions of the red anthocyanin pigments with acetaldehyde yield both purple pigments[10] and orange pyranoanthocyanin derivatives.[11] However, known polyphenols, including tannins, red anthocyanin pigments, and reaction products identified so far, represent the emerged part of the "polyphenol iceberg", explaining only a small proportion of the observed color. Similarly, oxidative browning of white wines is related to their flavan-3-ol content[12,13] and might result from flavan-3-ol reactions with acetaldehyde, but the structure of the brown pigments is still unknown. It is believed that the immersed part of the polyphenol iceberg results from random complex cascades of reactions involving grape polyphenols and other wine components such as yeast metabolites.

The purpose of the present work was to explore this hypothesis and shed light on the immersed part of the polyphenol iceberg. To this aim, non-targeted metabolomics approaches based on high-resolution mass spectrometry (HRMS) and petroleomics-derived data interpretation strategies, namely Kendrick mass defect plots, were implemented.[14] A simple model consisting of (−)-epicatechin (Ec) and malvidin-3-*O*-glucoside (Mv3G), representative of the two major families of grape polyphenols, and acetaldehyde, was selected to perform a proof-of-concept experiment. Using the restriction rules described in Supporting Information (Materials and Methods), 160 mass signals were retained and unambiguously attributed to elemental CHO compositions below 3 mmu tolerance.

In the selected model system, Ec and Mv3G are the two initial flavonoid building blocks. Mv3G exists as several forms in equilibrium.[13] In acidic aqueous solutions (pH <2), the red flavylium cation ($Mv3G_F$) is predominant. In mildly acidic solutions, such as wine, it undergoes nucleophilic attack of water at position 2 (or 4) of the pyrylium nucleus, leading to the colorless hemiketal ($Mv3G_{OH}$), and deprotonation, giving rise to blue–purple quinonoid bases ($Mv3G_B$).[15] This leads to four initial building blocks, as depicted in Figure 1.

Ec and Mv3G possess nucleophilic carbons in the 6 and 8 positions, which can react with protonated acetaldehyde to

[a] Dr. A. Vallverdú-Queralt, E. Meudec, M. Eder, Dr. N. Sommerer, Dr. V. Cheynier
Department Sciences pour l'œnologie
Institution INRA, UMR1083
2 Place Pierre Viala, Montpellier 34000 (France)
E-mail: avallverdu@ub.edu

[b] Dr. R. M. Lamuela-Raventos
Nutrition and Food Science Department
University of Barcelona
Av Joan XXIII s/n, 08007 Barcelona (Spain)

[c] Dr. R. M. Lamuela-Raventos
Instituto de Salud Carlos III, ISCIII (CIBEROBN)
C/ Sinesio Delgado, 4, 28029 Madrid (Spain)

Supporting Information and the ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/open.201700044.

ChemPubSoc Europe

Open Access ChemistryOPEN Communications

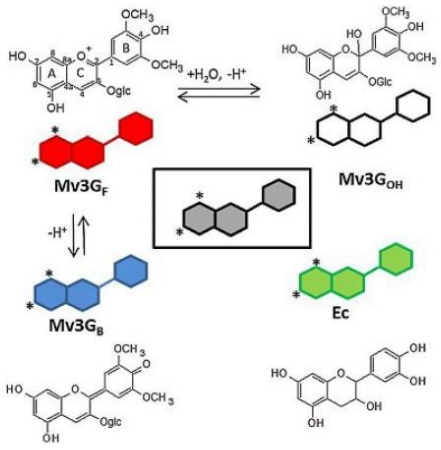


Figure 1. Initial building blocks of acetaldehyde mediated reaction. *Nucleophilic activity.

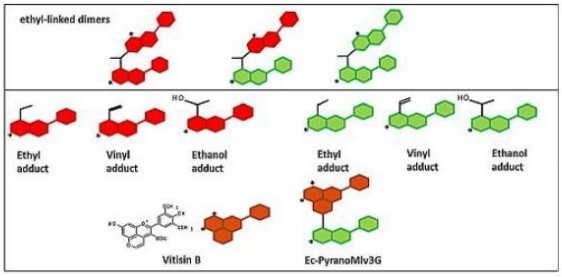


Figure 2. Other known building blocks arising from Mv3G$_F$ and Ec. *Nucleophilic activity

yield ethanol–flavonoid adducts and methyl–methine-linked flavonoid dimers (called ethyl-linked dimers) (see Figure 2).[16,17] By dehydration and cleavage of the methyl-methine bridge, respectively, both generate vinyl-flavonoid intermediates and ethyl-flavonoids.[18] The flavylium cation can also undergo cycloaddition at the C-4 and 5-OH positions with vinylphenol structures[19] or with carbonyl compounds (with dehydration),[20] followed by an oxidation step to yield an additional pyrane ring. PyranoMv3G (vitisin B)[21] and Ec-pyranoMv3G,[22] resulting from reaction of Mv3G with acetaldehyde and with vinyl-Ec, respectively, have been detected in wine. All of these reaction products show nucleophilic activity in C-6 and C-8 positions and, thus, can be considered as additional building blocks for the acetaldehyde condensation (Figure 2).

To investigate step-growth condensation, a Kendrick plot was created with a mass increment corresponding to Ec-ethyl units (316 amu) based on the assumption that most compounds contained sequences of this structural motif. Each horizontal line corresponds to a series of compounds differing by the number of Ec-ethyl units, starting from a given base unit, as illustrated in Figure 3 for Ec and Mv3G.

Increases corresponding to Mv3G-ethyl units (518 amu) are also displayed, showing that Ec and Mv3G compete in random polymerization to yield a large number of Ec, Mv3G, and mixed Ec–Mv3G ethyl-bridged oligomers, containing up to 6, 4, and 10 units, respectively (Table S1, new compounds in bold). It should be emphasized that each signal detected represents several isomers with different random sequences and different linkage positions and conformations (6-6, 6–8, R and S, 8-8). Moreover, Mv3G was found in its different forms in the homogenous Mv3G polymers, as described earlier,[17] whereas the flavylium form was predominant in mixed derivatives, suggesting that they are less susceptible to hydration than Mv3G, as shown earlier for the dimers.[12] Note that the quinonoidal base form is detected as its protonated adduct and, thus, in

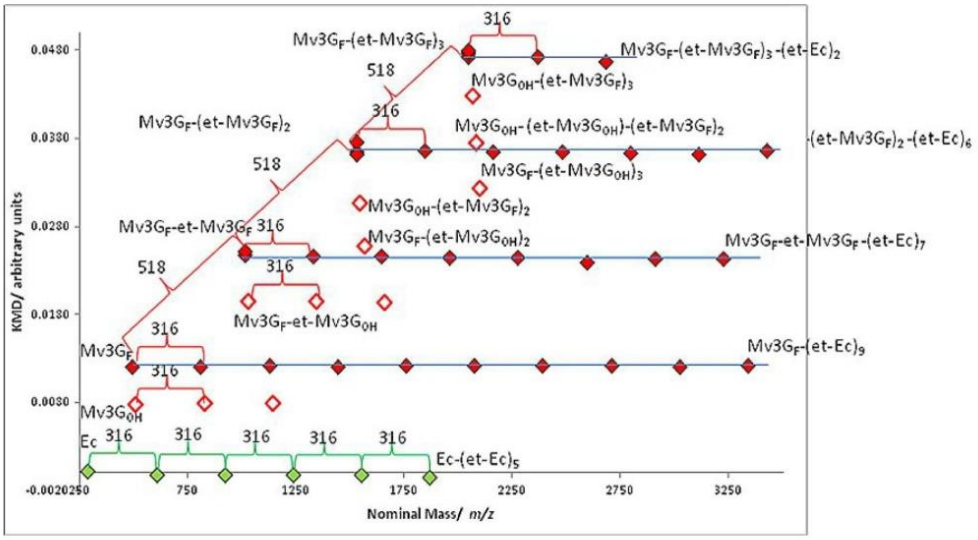


Figure 3. Diagram of Kendrick for Ec and Mv3G.

ChemPubSoc Europe
Open Access ChemistryOPEN Communications

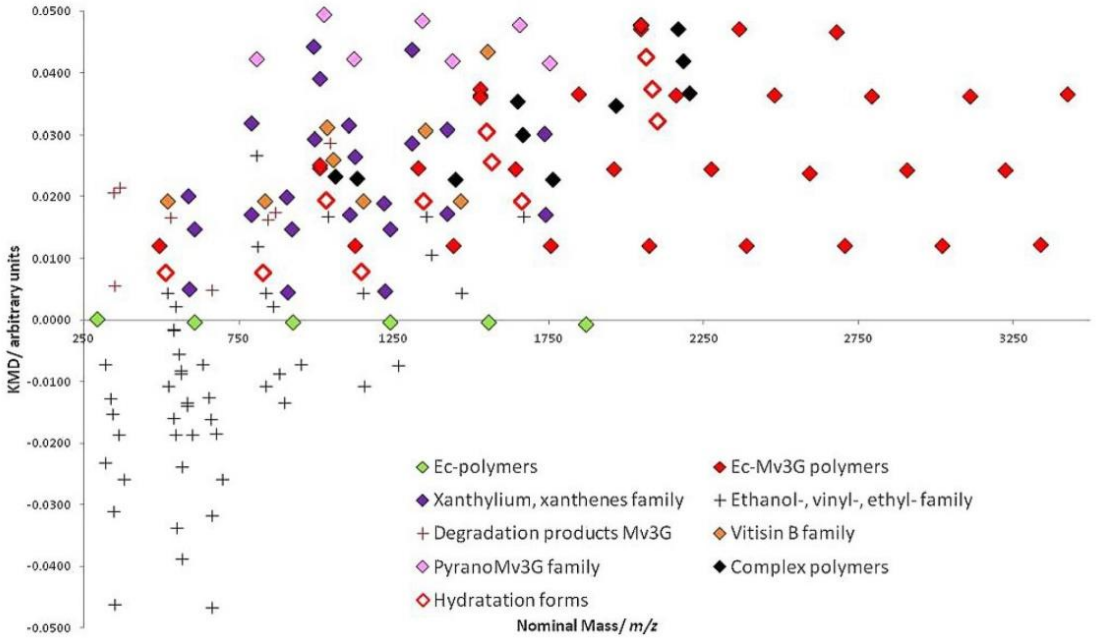


Figure 4. Diagram of Kendrick for all molecules.

some structures, cannot be distinguished from the flavylium form (Tables S1-S7).

The Kendrick representation enabled the detection of several base units, in addition to the initial building blocks (i.e. Ec and Mv3G), each giving rise to a polymer series formed through successive additions of ethyl-Ec and/or ethyl-Mv3G (Figure 4, Tables S2-S7).

These additional building blocks included the known ethyl, vinyl, and ethanol adducts of Ec and Mv3G (Table S2), pyranoMv3G (vitisin B) (Table S**3**) and Ec-pyranoMv3G (Table S4), but also so far unreported compounds (Tables S1–S7). Among the latter, Mv3G–pyranoMv3G, arising from the cycloaddition of vinyl-Mv3G onto another Mv3G molecule, and its hydrated form were detected (Figure 5, Table S8).

Other new series corresponded to derivatives of xanthene, xanthylium, and hydrated xanthylium forms (Figure 5, Table S8). Two adjacent Ec and/or Mv3G units in ethyl-linked oligomers can cyclize to form xanthenes that oxidize to xanthylium salts, which are in equilibrium with their hydrated form (Table S5). These reactions have been described earlier for condensation products of flavan-3-ols and/or anthocyanins with other aldehydes[23–24] or with Ec and acetaldehyde,[26] but this is the first report of such molecules arising from the reaction of Ec and/or Mv3G with acetaldehyde.

Each of these building blocks can also combine randomly in the polymerization process, as illustrated by the detection of complex molecules such as vitisin B–ethyl–vitisin B (M^{2+} detected at $m/z = 530.1424$), vitisin B–ethyl–Ec-pyranoMv3G (M^{2+} detected at $m/z = 674.1741$), vitisin B–ethyl–Mv3G–pyranoMv3G (M^{3+} detected at $m/z = 517.1346$), and vitisin B–ethyl–

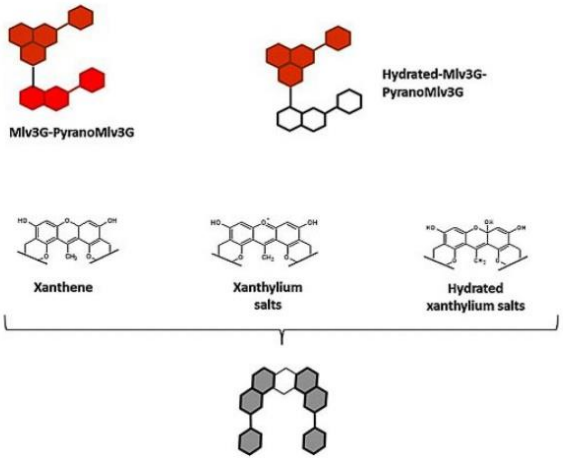


Figure 5. More complex building blocks.

xanthene (M^{+} detected at $m/z = 1131.3138$), all of which had not been reported (Table S6).

Other compounds detected on the diagram of Kendrick were degradation products of Mv3G, namely malvone, malvone aglycone, and 3-hydroxyphenylacetyl glucoside,[27] that also serve as building blocks because their ethyl-flavonoid derivatives were detected, as well as *cis*- and *trans*-anthocyanone A (8-β-D-glucopyranosyl-2,4-dihydroxy-6-oxo-cyclohexa-2,4-dienylacetic acid) (Table S**7**).

ChemPubSoc Europe

Open Access ChemistryOPEN Communications

In summary, we demonstrated that, starting from a very simple solution, reaction cascades involving acetaldehyde–flavonoid condensation, dehydration of flavonoid–ethyl–flavonoid units, cycloaddition (yielding pyranoanthocyanins), oxidation (of xanthene to xanthylium), and hydration (of flavylium and xanthylium salts) yield an extremely complex composition, including new compounds. We believe that similar random cascade mechanisms involving many more building blocks and reactions generate the composition of complex real systems such as wines (or other processed food and non-food products). HRMS data analysis associated with a Kendrick mass defect strategy can be successfully applied to describe these systems, detect additional products, building blocks, and reactions, and, using appropriate chemometrics, establish relationships between composition and quality.

Experimental Section

See the Supporting Information.

Acknowledgements

We thank the scientific platform of CCIT UB for the technical support. Funding for this work was provided by the Instituto de Salud Carlos III, ISCIII (CIBEROBN), the Quality Group from Generalitat de Catalunya (GC) 2014 SGR 773, and by INRA Departement CEPIA.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: acetaldehyde-mediated condensation · flavonoids · Kendrick mass defect plots · mass spectrometry · polymerization

[1] S. Quideau, D. Deffieux, C. Douat-Casassus, L. Pouységu, *Angew. Chem. Int. Ed.* **2011**, *50*, 586 – 621; *Angew. Chem.* **2011**, *123*, 610 – 646.
[2] I. Tarascou, J. M. Souquet, J. P. Mazauric, S. Carrillo, S. Coq, F. Canon, H. Fulcrand, V. Cheynier, *Arch. Biochem. Biophys.* **2010**, *501*, 16 – 22.
[3] F. Canon, A. R. Milosavljevič, G. Van Der Rest, M. Réfrégiers, L. Nahon, P. Sarni-Manchado, V. Cheynier, A. Giuliani, *Angew. Chem. Int. Ed.* **2013**, *52*, 8377 – 8381; *Angew. Chem.* **2013**, *125*, 8535 – 8539.
[4] A. Crozier, I. Jaganath, M. N. Clifford, *Nat. Prod. Rep.* **2009**, *26*, 1001 – 1043.
[5] N. Mateus, A. Silva, C. Santo sBuelga, J. C. Rivas-Gonzalo, V. De Freitas, *J. Agric. Food Chem.* **2002**, 50, 2110 – 2116.
[6] V. De Freitas, N. Mateus, *Environ. Chem. Lett.* **2006**, *4*, 175 – 183.
[7] V. Cheynier, M. Duenas-Paton, E. Salas, C. Maury, J. M. Souquet, P. Sarni-Manchado, H. Fulcrand, *Am. J. Enol. Vitic.* **2006**, *57*, 298 – 305.
[8] J. Wirth, C. Morel-Salmi, J. M. Souquet, H. Fulcrand, V. Cheynier, *Food Chem.* **2010**, *123*, 107 – 116.
[9] A. Arapitsas, M. Scholz, U. Vrhovsek, S. Di Blasi, A. Bartolini, D. Masuero, D. Perenzoni, A. Rigo, F. Mattivi, *PLoS One* **2012**, *7*, e37783 – e37788.
[10] M. Duenas, H. Fulcrand, V. Cheynier, *Anal. Chim. Acta* **2006**, *563*, 15 – 25.
[11] V. De Freitas, N. Mateus, *Anal. Bioanal. Chem.* **2011**, *401*, 1463 – 1477.
[12] V. Cheynier, J. Rigaud, J.-M. Souquet, J.-M. Barillere, M. Moutounet, *Am. J. Enol. Vitic.* **1989**, *40*, 36 – 42.
[13] R. F. Simpson, *Vitis* **1982**, *21*, 233 – 239.
[14] T. Dier, K. Egele, V. Fossog, R. Hempelmann, A. Dietrich, *Anal. Chem.* **2016**, *88*, 1328 – 1335.
[15] A. Vallverdú-Queralt, M. Biler, E. Meudec, C. L Guernevé, A. Vernhet, J.-P. Mazauric, J.-L. Legras, M. Loonis, P. Trouillas, V Cheynier, O. Dangles, Int. J. Mol. Sci. **2016**, *17*, 1842.
[16] N.-E. Es-Safi, H. Fulcrand, V. Cheynier, M. Moutounet, *J. Agric. Food Chem.* **1999**, *47*, 2088 – 2095.
[17] V. Atanasova, H. Fulcrand, C. Le Guerneve, V. Cheynier, M. Moutounet, *Tetrahedron Lett.* **2002**, *43*, 6151 – 6153.
[18] L. Cruz, R. F. Bras, N. Teixeira, A. Fernandes, N. Mateus, J. Ramos, J. Rodriguez-Borges, V. De Freitas, *J. Agric. Food Chem.* **2009**, *57*, 10341 – 10348.
[19] H. Fulcrand, P. J. Cameira dos Santos, P. Sarni-Manchado, V. Cheynier, J. J. Favre-Bonvin, *Chem. Soc. Perkin. Trans.* **1996**, *1*, 735 – 739.
[20] H. Fulcrand, C. Benabdeljalil, J. Rigaud, V. Cheynier, M. Moutounet, *Phytochemistry* **1998**, *7*, 1401 – 1407.
[21] J. Bakker, P. Bridle, T. Honda, H. Kuwano, N. Saito, N. Terahara, C. Timberlake, *Phytochemistry* **1997**, *7*, 1375 – 1382.
[22] L. Cruz, N. Teixeira, A. M. Silva, N. Mateus, J. Borges, V. de Freitas, *J. Agric. Food Chem.* **2008**, *56*, 10980 – 10987.
[23] N. E. Es-Safi, V. Cheynier, M. Moutounet, *J. Agric. Food Chem.* **2000**, *48*, 5946 – 5954.
[24] N. E. Es-Safi, C. Le Guernevé, V. Cheynier, M. Moutounet, *J. Agric. Food Chem.* **2000**, *48*, 4233 – 4240.
[25] N. E. Es-Safi, V. Cheynier, M. Moutounet, *J. Agric. Food Chem.* **2002**, *25*, 5586 – 5595.
[26] A. Vallverdú-Queralt, E. Meudec, M. Eder, R. M. Lamuela-Raventos, N. Sommerer, V. Cheynier, *Food Chem.* **2017**, *227*, 255 – 263.
[27] A. Vallverdú-Queralt, E. Meudec, N. Ferreira, N. Sommerer, O. Dangles, V. Cheynier, C. Le Guerneve, *Food Chem.* **2016**, *199*, 902 – 910.

Received: March 2, 2017
Version of record online May 15, 2017

Oral presentations at international scientific meetings

Matthias Eder, Isabelle Sanchez, Peggy Rigou, Thibault Nidelet, Carole Camarasa, Jean-Luc Legras, Sylvie Dequin (2014). Deciphering the Genetic and Metabolic Bases of Yeast Aroma Properties. 31st International Symposium on Yeast (ISSY31), Vipavia, Slovenia.

Matthias Eder, Isabelle Sanchez, Peggy Rigou, Thibault Nidelet, Carole Camarasa, Jean-Luc Legras, Sylvie Dequin (2016). Deciphering the Genetic and Metabolic Bases of Yeast Aroma Properties. 12th Levures, Modèles et Outils meeting (LMO12), Bruxelles, Belgium.

Matthias Eder, Isabelle Sanchez, Peggy Rigou, Thibault Nidelet, Carole Camarasa, Jean-Luc Legras, Sylvie Dequin (2016). Deciphering the Genetic and Metabolic Bases of Yeast Aroma Properties. 6th Conference on Physiology of Yeast and Filamentous Fungi (PYFF6), Lisbon, Portugal.

Poster presentations at international scientific meetings

Matthias Eder, Isabelle Sanchez, Peggy Rigou, Thibault Nidelet, Carole Camarasa, Jean-Luc Legras, Sylvie Dequin (2015). Deciphering the Genetic and Metabolic Bases of Yeast Aroma Properties. 32nd International Symposium on Yeast (ISSY32), Perugia, Italy.
(The poster was honored with a best poster price award by the organizing committee of the meeting).

Matthias Eder, Isabelle Sanchez, Peggy Rigou, Thibault Nidelet, Carole Camarasa, Jean-Luc Legras, Sylvie Dequin (2015). Deciphering the Genetic and Metabolic Bases of Yeast Aroma Properties. EMBO conference “Exploring the genomic complexity and diversity of eukaryotes”, Sant Feliu de Guixols, Spain.

Matthias Eder, Isabelle Sanchez, Peggy Rigou, Thibault Nidelet, Carole Camarasa, Jean-Luc Legras, Sylvie Dequin (2017). Deciphering the Genetic and Metabolic Bases of Yeast Aroma Properties. 33rd International Symposium on Yeast (ISSY33), Cork, Ireland.
(The poster was honored with a best poster price award by the organizing committee of the meeting).

Résumé : La levure *Saccharomyces cerevisiae* joue un rôle essentiel dans la production de composés aromatiques, tels que les esters, les alcools supérieurs et les acides organiques, ainsi que dans la transformation de précurseurs d'arômes du raisin pendant la fermentation du vin. Afin d'identifier les bases génomiques et métaboliques de ces propriétés, un croisement a été réalisé entre deux souches de levures de vin, sélectionnées pour leurs besoins en azote différents lors de la fermentation. 130 ségrégants de génération F2 ont été génotypés par séquençage complet du génome et individuellement phénotypés pendant la fermentation en mesurant les métabolites extracellulaires par HPLC et GC-MS. Les flux métaboliques intracellulaires ont été estimés à l'aide d'un modèle à base de contraintes. Une analyse QTL (quantitative trait locus) a été utilisée pour identifier les allèles influençant les variations d'arômes et de flux métaboliques. Plus de 80 QTL expliquant la variation de 59 caractères quantitatifs ont été détectés. Ces caractères comprennent des paramètres fermentaires, de consommation de substrat, la production de principaux métabolites et d'arômes fermentaires, ainsi que le métabolisme de composés aromatiques du raisin comme de terpénols ou le sulphur de diméthyle. L'intérêt de la cartographie QTL pour identifier les déterminants génétiques de variations de flux intracellulaires (f-QTLs) a par ailleurs été démontrée. Les QTL détectés ont été disséqués et des gènes dont les allèles contribuent spécifiquement aux variations phénotypiques ont été identifiés. Ces résultats soulignent la complexité génomique et métabolique de la synthèse et de la transformation d'arômes par la levure. L'identification de ces déterminants génétiques permet de mieux comprendre les liens entre variation génétique des levures et traits technologiques et fournit une base précieuse pour le développement de souches optimisées par des stratégies génétiques de croisement assisté par marqueurs.

Titre : Décrypter les bases génétiques et métaboliques des propriétés aromatiques de la levure

Mots clés : Fermentation œnologique, levure de vin industrielle, production d'arômes, arômes fermentaires, terpénols, sulfure de diméthyle, flux métaboliques

Abstract: The yeast *Saccharomyces cerevisiae* plays a vital role in the production of aroma compounds, such as esters, higher alcohols and organic acids, and the conversion of grape-derived aroma precursors during wine fermentation. To identify the genomic and metabolic bases for these processes, a cross was performed between two wine yeast strains selected because of their different nitrogen requirement during fermentation. 130 F2-segregants were genotyped by whole genome sequencing and individually phenotyped during wine fermentation by measuring extracellular metabolites using HPLC and GC-MS. Intracellular metabolic fluxes were estimated by constraint-based modeling. Quantitative trait locus (QTL) mapping was used to identify allelic variants influencing variations in the aroma profile and metabolic fluxes. More than 80 QTLs explaining variation in 59 quantitative traits were detected. These traits consisted of general fermentation parameters, substrate consumption, the production of main metabolites and fermentative aromas and the metabolism of grape aroma compounds such as terpenols or dimethyl sulfide. The applicability of QTL mapping to detect regions influencing intracellular fluxes (f-QTLs) was furthermore demonstrated. Found QTLs were dissected and genes with allele specific contributions to the phenotype were identified. These results emphasize the genomic and metabolic complexity of yeast aroma formation. In addition, the identification of genetic determinants increases knowledge about the links between genetic variation and industrial traits and provides a valuable foundation for the development of optimized strains by marker-assisted selection breeding strategies.

Title: Deciphering the genetic and metabolic bases of yeast aroma properties

Keywords: Wine fermentation, industrial wine yeast, aroma production, fermentative aromas, terpenols, dimethyl sulfide, metabolic fluxes