TÉCNICO LISBOA

UNIVERSIDADE DE LISBOA

INSTITUTO SUPERIOR TÉCNICO

Promoting the use of *Saccharomycodaceae* yeasts in winemaking by leveraging genome-wide analyses

Maria João Alves Tavares

Supervisor: **Doctor** Nuno Gonçalo Pereira Mira

Co-Supervisor: Doctor Ana Alexandra Mendes Ferreira

Thesis approved in public session to obtain the PhD Degree in Biotechnology and Biosciences

Jury final classification: Pass with Distinction

2024

UNIVERSIDADE DE LISBOA

INSTITUTO SUPERIOR TÉCNICO

Promoting the use of *Saccharomycodaceae* yeasts in winemaking by leveraging genome-wide analyses

Maria João Alves Tavares

Supervisor: **Doctor** Nuno Gonçalo Pereira Mira

Co-Supervisor: Doctor Ana Alexandra Mendes Ferreira

Thesis approved in public session to obtain the PhD Degree in Biotechnology and Biosciences

Jury final classification: Pass with Distinction

Jury

Chairperson: Doctor Arsénio do Carmo Sales Mendes Fialho, Instituto Superior Técnico, Universidade de Lisboa

Members of the Committee:

Doctor Timothy Alun Hogg, Escola Superior de Biotecnologia, Universidade Católica Portuguesa

Doctor Arsénio do Carmo Sales Mendes Fialho, Instituto Superior Técnico, Universidade de Lisboa

Doctor Miguel Nobre Parreira Cacho Teixeira, Instituto Superior Técnico, Universidade de Lisboa

Doctor Maria João Marques Ferreira Sousa Moreira, Escola de Ciências da Universidade do Minho

Doctor Nuno Gonçalo Pereira Mira, Instituto Superior Técnico, Universidade de Lisboa

Funding Institutions

FCT: Fundação para a Ciência e Tecnologia

2024

Promoting the use of *Saccharomycodaceae* yeasts in winemaking by leveraging genome-wide analyses

Abstract

Wine has been produced and consumed by humans for centuries. At the heart of winemaking is the activity of fermenting yeasts, particularly *Saccharomyces cerevisiae* that leads vinification. The growing trend of using starter-cultures of other yeasts alongside *S. cerevisiae* aims to improve wine properties and create distinguishable products. Yeasts from the *Saccharomycodacea* family, specifically the sister genera *Hanseniaspora* and *Saccharomycodes*, are considered promising in this context. However, their utilization is hampered by the poor knowledge of their physiology and interaction mechanisms with *S. cerevisiae*. This thesis addresses these gaps through a combination of knowledge-gathering and application.

The first genome sequence of a *Saccharomycodes ludwigii* wine strain is described in (**Chapter II**), facilitating comparative genomic analyses, both intra-species (**Chapter III**) and inter-species (**Chapter V**), against *Hanseniaspora*, providing insights into the physiology of *Sd. ludwigii* and the *Saccharomycodacea* family. Phenotypic profiling of Non-*Saccharomyces* isolates from grape musts in the Palmela winemaking area (Península de Setúbal region) enabled the identification of strain *Hanseniaspora uvarum* MJT198 (HuMJT198), based on its very high beta-glucosidase activity (**Chapter IV.A**). The combined use of HuMJT198 with *S. cerevisiae* produced wines with increased free terpenes, expanding the role of *H. uvarum* strains beyond modulation of secondary aroma. Exploring comparative genomics of HuMJT198 and other *Hanseniaspora* species beta-glucosidase-encoding genes were identified as contributors to this effect. Transcriptomic and metabolomic analyses revealed further interactions between HuMJT198 and *S. cerevisiae* at a molecular level (**Chapter IV.B**).

Altogether, the results described in this thesis supports the trend of exploring Non-*Saccharomyces* Yeasts as bio-flavourants in wine production but also other beverages, applying knowledge-driven and rational methods that improve the likelihood of success compared to the trial-and-error approaches currently in use.

Keywords: Non-*Saccharomyces*, *Saccharomycodaceae*, *Saccharomycodes ludwigii*, *Hanseniaspora uvarum*, winemaking

Promovendo o uso the leveduras *Saccharomycodaceae* na produção de vinho através da utilização de análises genómicas abrangentes

Resumo

A produção de vinho e o seu consumo acompanham o Ser Humano há vários séculos, com a atividade de leveduras fermentativas, como *Saccharomyces cerevisiae*, a liderar o processo de vinificação. A utilização de inóculos combinados de leveduras nativas e *S. cerevisiae*, é uma tendência crescente neste sector, ambicionando melhorar as propriedades do vinho e criando produtos distintos. Leveduras da família *Saccharomycodaceae*, particularmente dos géneros *Hanseniaspora* e *Saccharomycodes*, têm sido consideradas promissoras neste contexto. No entanto, o conhecimento limitado sobre a sua fisiologia e os mecanismos de interação com *S. cerevisiae*, restringe a sua utilização. Esta tese aborda estas lacunas, combinando investigação fundamental e aplicação prática.

No **Capítulo III**, é descrita a primeira sequência genómica de uma estirpe de *Saccharomycodes ludwigii* isolada de vinho, permitindo análises genómicas comparativas intraespécie (**Capítulo III**) e inter-espécies (**Capítulo V**), aprofundando a compreensão da fisiologia de *Sd. ludwigii* e da família *Saccharomycodaceae*. A fenotipagem de isolados Não-*Saccharomyces* obtidos de mostos da região vitivinícola da Península de Setúbal permitiu a identificação de estirpe *Hanseniaspora uvarum* (HuMJT198), pela sua elevada atividade de beta-glucosidase (**Capítulo IVA**). A utilização combinada de HuMJT198 com *S. cerevisiae* resultou em vinhos com maior concentração de terpenos livres, enaltecendo o papel das leveduras *Hanseniaspora* na modulação do aroma primário. A exploração genómica de HuMJT198 e outras espécies de *Hanseniaspora* identificou genes codificadores de beta-glucosidase como responsáveis por este efeito. Adicionalmente, análises transcriptómicas e metabolómicas revelaram interações moleculares entre HuMJT198 e *S. cerevisiae* (**Capítulo IVB**).

No geral, os resultados desta tese apoiam a exploração de leveduras Não-*Saccharomyces* como agentes bioaromatizantes, aplicando métodos baseados em conhecimento, com maior probabilidade de sucesso face à abordagem empírica atualmente utilizada.

Palavras-chave: Não-*Saccharomyces*, *Saccharomycodaceae*, *Saccharomycodes ludwigii*, *Hanseniaspora uvarum*, produção de vinho

“In the fields of observation chance favors only the prepared mind”

Louis Pasteur

“Enquanto houver estrada pra andar
A gente vai continuar
Enquanto houver estrada pra andar
Enquanto houver ventos e mar
A gente não vai parar
Enquanto houver ventos e mar”

Jorge Palma

Acknowledgments

It is through the hardworking hands, with skin hardened by calluses, and with nails dirty from the soil, of the one who has always been my grandfather, that I begin my thanks. In fact, if it weren't for him, this thesis or the work that gave rise to it would never have been produced. Slim and gentle, kind and serene, playful and adventurous, my uncle Virgílio Bacalhau (and, forgive me his grandchildren, my grandfather) swirled the glass of wine produced from his grapes, which grew in his land, for what seemed like an eternity. Like him, they faced many hardships, and they resisted those hardships, producing their fruits, which later became the wine my uncle sipped slowly, with the respect it deserved. And, like him, they had the ability to unite the large family that, at least once a year, armed themselves with pruning shears and buckets, rubber boots and berets, laughter, and good cheer. Thank you, dear uncle, for showing me this path, but also for so many other things that I hope I had the opportunity to tell you.

The doctoral journey is a marathon in which, often, the finish line seems to slip through one's fingers. It mirrors our weaknesses but also invites us to enhance our self-discipline and perseverance because motivation sometimes becomes scarce, even in the most enthusiastic individuals. And, definitively, it is a path no one walks alone. Just as any athlete has a good coach as part of the key to their success, a scientist also improves their qualities when their supervisor recognizes their characteristics and abilities, shaping their work and path to make them shine. Thank you, Nuno Mira, for providing a platform for my ideas and creating a supportive atmosphere where I could succeed but also learn from failures. Above all, I appreciate your prioritization of the individual over the scientist and your unwavering support that has guided me to this significant moment. I also appreciate my co-supervisor, Professor Alexandra Mendes-Ferreira, for welcoming me into her "home" and for sharing so much valuable knowledge with me. Thank you for giving me the opportunity to do the work that brought life and color to my thesis, the fermentations in natural must. And, of course, for introducing me to "covilhetes."

As a scientist, I have always been interested in seeing my research take shape and understanding its impact in the real world. I wondered how I could achieve this with yeast and wine: will it work? Are the questions I ask real? To answer these questions, I wandered around Palmela and its surroundings with determination, knocking on the doors of wineries and hearing many "no's." Until, without any delays or convincing, I heard a "yes." And that "yes" came from the mouth of Francisco Camolas, from Adega Camolas. I am grateful to him for the

opportunity given to this aspiring scientist to live in his winery, drink from his knowledge, and become intoxicated with the hustle and bustle of the grape harvest. I cherish the memories of stepping onto the grape crusher with a falcon tube in hand! Thank you also for providing me with what I thought was impossible: large-scale fermentations. I will never forget such generosity. Thanks to the oenologist António Sanches, whom I consider a friend, for opening the doors to his laboratory and sharing all his knowledge with me. Thanks to Bruno Correia, Mr. Raúl, and all the members of Adega Camolas who extended their hand to me.

I would also like to express my gratitude to all the collaborators whose contributions made much of the work presented in this thesis possible. Thank you to Professor Carlos Cordeiro, Doctor Marta Silva and their team from the Faculty of Sciences at the University of Lisbon for their support during the FT-ICR-MS experiments. I also extend my thanks to Professor Janja Trček from the University of Ljubljana for kindly providing us with the *Saccharomycodes ludwigii* BJK_5C strain.

As wine is not made solely from grapes, I am grateful for the educational and practical conditions provided to study my beloved yeasts. Thus, I also owe my learning and work to Professor Isabel Sá-Correia, the former director of the Biological Sciences Research Group (BSRG), who not only enabled my enrollment in the BIOTECnico doctoral program but also consistently supported my research. I appreciate her always relevant observations. Similarly, I extend my gratitude to Professor Arsénio Fialho, the current leader of BSRG, for the opportunity to work in a collaborative environment that fostered my growth as a scientist. To the Fundação para a Ciência e Tecnologia (FCT), I express my thanks for funding my doctoral scholarship PD/BD/142945/2018 under the BioTecnico doctoral program, as well as the funding for the SmartWine (PTDC/AGR-TEC/3315/2014) and SmartWine 2.0 (2022.06777.PTDC) projects that made the presented research possible. Funds received from FCT, I.P., in the scope of the project supporting the Research Unit Institute for Bioengineering and Biosciences (UIDB/04565/2020) and the project supporting Institute for Health and Bioeconomy, i4HB (LA/P/0140/2020) are also acknowledged.

I would also like to express my deep appreciation to all the colleagues from the NPM group, iBB-BSRG and WM&B- UTAD who crossed my path during these years and assisted me both in the laboratory and with their words of encouragement: Ana Vila-Santa, Bárbara Coelho, Diogo Santos, Isabel Seixas, Joana Pinheiro, João Pedro Silva, Mafalda Cavalheiro, Marcos Esteves, Marta Mota, Mónica Galocha, Nuno Pedro, Patrícia Lage, Sara Salazar. Your presence made my days much richer. I also thank Mónica Rato for all the logistical support

often needed. Thanks to Dona São for making me feel at home when I was away, and to Elsa, and Sara for being who they are.

Special thanks to Luís Valente and Raquel Sá-Leão for their invaluable support and understanding throughout this process.

I am also very grateful to the friends who never wavered and were a pillar in whom I confided, for both the venting of frustrations and celebrating achievements. Thank you, Inês and Pipa, who show me every day that distance is just a formality when friendship is already family. Thanks to Cuco, Raquel, Tomás, and André for your constant presence and for always asking about me, even when I sometimes hide away.

These acknowledgments wouldn't be complete without dedicating a paragraph to my family. Thank you, Francisco, for the immeasurable and unparalleled support during these years. Thank you for the friendship, love, and companionship, and for the scientific and philosophical discussions that often helped me navigate challenges. For challenging me and pushing me out of my comfort zone. For exploring the wineries of your hometown with me and joining in all the adventures. For bringing me to reason and offering me a helping hand. For letting me dream and keeping me grounded. Thank you for your generosity throughout this journey. Thank you Milinha and Aura for the logistical support. I also thank my parents, who see me complete another degree in this institute and share in my struggles and celebrations. Thank you for being the safe haven I can always turn to. Your unconditional support and love make me a better person and a more confident scientist. I am eternally grateful to you.

And it is with the small and fragile hands, with soft and delicate skin, and nails dirty from the soil of my daughter, that I conclude these acknowledgments. Thank you, dear daughter, for showing me the true path and accompanying me in this couple of years of life and thesis, which is as much yours as it is mine. This thesis that has as many of your tears as mine, as much of your joy as my enthusiasm, as much of your life as my persistence. I appreciate your company and presence, which have made my days infinitely happier. To you, Helena, I dedicate all of me. Thank you.

Agradecimentos

É pelas mãos trabalhadoras, com a pele enrijecida pelos calos, e com as unhas sujas de terra, daquele que foi sempre o meu avô, que começo os meus agradecimentos. Na verdade, se não fosse ele, esta tese ou o trabalho que lhe deu origem, nunca teriam sido produzidos. Delgado e franzino, bondoso e sereno, brincalhão e aventureiro, o meu tio Virgílio Bacalhau (e, perdoem-me os netos, o meu avô) balançava nos dedos, durante o que parecia ser uma eternidade, o copo de vinho produzido pelas suas uvas, que cresceram na sua terra. Tal como ele, enfrentaram muitas intempéries, e a essas intempéries resistiram, e produziram os seus frutos, que mais tarde originaram o vinho que o meu tio demoradamente bebericava, com o respeito que merecia. E, tal como ele, tiveram a capacidade de unir a numerosa família que, pelo menos uma vez no ano, se munia de tesouras de poda e baldes, galochas e boinas, de galhofa e boa disposição. Obrigada, querido tio, por me mostrares este caminho, mas também por tantas outras coisas que espero ter tido oportunidade de te dizer.

O doutoramento é uma maratona em que, muitas vezes, a fita da meta parece ir escapando por entre os dedos. É um espelho das nossas fraquezas, mas também um convite a aprimorarmos a nossa capacidade de autodisciplina e perseverança, pois a motivação por vezes escasseia, até na mais entusiasta das pessoas. E definitivamente é um caminho que ninguém faz sozinho. Como qualquer atleta tem num bom treinador uma parte da chave para o seu sucesso, também um cientista aprimora as suas qualidades quando o seu orientador reconhece nele as suas características e capacidades, modelando o seu trabalho e o percurso para o ver brilhar. Obrigada, Nuno Mira, por tantas vezes dares palco às minhas ideias, e por promoveres um ambiente seguro para acertar, mas também para muitas vezes falhar. E obrigada essencialmente por colocares a pessoa à frente do cientista, e por sempre me dares o apoio e o suporte para chegar até este momento. Agradeço também à minha co-orientadora Professora Alexandra Mendes-Ferreira por me ter recebido na sua "casa" e por comigo ter partilhado tanto e tão valioso conhecimento. Obrigada por me ter dado a oportunidade de ter feito o trabalho que deu vida e cor à minha tese, as fermentações em mosto natural. E por me ter apresentado aos covilhetes!

Enquanto cientista sempre me interessei por ver a minha pesquisa tomar forma, e conhecer o seu impacto no mundo real. Interessava-me saber como conseguiria fazê-lo com as leveduras e o vinho: será que funciona? Serão reais as perguntas que me coloco? Para responder a estas questões, percorri Palmela e arredores de punho fechado, batendo à porta de adegas e

ouvindo muitos “nãos”. Até que, sem qualquer delongas ou convencimentos, ouvi um “sim”. E esse “sim” veio da boca de Francisco Camolas, da Adega Camolas. A ele agradeço a oportunidade dada a esta cientista de viver na sua adega, beber do seu conhecimento, e ficar inebriada com a azáfama das vindimas. Guardo com muito carinho os saltos para o esmagador da uva, de tubo falcon em riste! Obrigada também por me proporcionar o que julgava ser impossível: as fermentações em larga escala. Nunca esquecerei tamanha generosidade. Obrigada ao enólogo António Sanches, em quem vejo um amigo, por ter aberto as portas ao seu laboratório e comigo ter partilhado todo o seu conhecimento. Obrigada ao Bruno Correia, ao senhor Raúl e a todos os membros da Adega Camolas que me estenderam a mão.

Gostaria também de agradecer a todos os colaboradores cujas contribuições tornaram possível muito do trabalho apresentado nesta tese. Obrigada ao Professor Carlos Cordeiro, à Doutora Marta Silva e à sua equipa pelo apoio durante os ensaios de FT-ICR-MS. De igual modo, agradeço à Professora Janja Trček da Universidade de Ljubljana por providenciar a estirpe *Saccharomycodes ludwigii* BJK_5C.

Como não só de uvas se faz vinho, agradeço as condições de cariz formativo e prático que me foram dadas para poder estudar as minhas queridas leveduras. Assim, devo também a minha aprendizagem e o meu trabalho à Professora Isabel Sá-Correia, antiga diretora do Biological Sciences Research Group (BSRG), que não só possibilitou a minha frequência no programa doutoral BIOTECnico, como sempre apoiou a minha pesquisa. Agradeço as sempre pertinentes observações. Do mesmo modo, agradeço também ao professor Arsénio Fialho, atual líder do BSRG, pela oportunidade de trabalhar num ambiente colaborativo e que promoveu a minha evolução enquanto cientista. À Fundação para a Ciência e Tecnologia (FCT) agradeço o financiamento da minha bolsa de doutoramento PD/BD/142945/2018 no âmbito do programa doutoral BioTecnico, bem como o financiamento dos projetos SmartWine (PTDC/AGR-TEC/3315/2014) e SmartWine 2.0 (2022.06777.PTDC) que possibilitaram a pesquisa aqui apresentada. Agradeço ainda à FCT, I.P., no âmbito do projeto que suportou a Unidade de Investigação Instituto de Bioengenharia e Biociências (UIDB/04565/2020) e o projeto que suportou o Instituto para a Saúde e Bioeconomia, (LA/P/0140/2020).

Gostaria também de deixar uma palavra de grande apreço a todos os colegas do grupo NPM, do iBB-BSRG e do WM&B- UTAD que cruzaram o meu caminho durante estes anos e me ajudaram quer no laboratório, quer com as suas palavras de encorajamento: Ana Vila-Santa, Bárbara Coelho, Diogo Santos, Isabel Seixas, Joana Pinheiro, João Pedro Silva, Mafalda Cavalheiro, Marcos Esteves, Marta Mota, Mónica Galocha, Nuno Pedro, Patrícia Lage, Sara

Salazar. A vossa presença tornou os meus dias muito mais ricos. Agradeço também à Mónica Rato por todo o apoio logístico tantas vezes necessários. Obrigada à Dona São por me fazer sentir em casa quando estava longe, e à Elsa e à Sara por serem quem são.

Obrigada ao Luís Valente e à Raquel Sá-Leão pela compreensão durante todo este processo.

Agradeço também aos amigos que nunca arredaram pé e que foram um pilar em quem confiei, tanto os desabafos, como as conquistas. Obrigada, Inês e Pipa, que todos os dias me mostram que a distância é apenas uma formalidade quando a amizade é já família. Obrigada, Cuco, Raquel, Tomás, e André pela presença constante, e por sempre perguntarem por mim, mesmo quando às vezes me escondo.

Estes agradecimentos não ficariam completos sem dedicar um parágrafo à minha família. Obrigada, Francisco, pelo apoio imensurável e inigualável durante estes anos. Obrigada pela amizade, amor, e companheirismo, e pelas discussões científicas e filosóficas que tantas vezes me ajudaram a sair de impasses. Por tantas vezes me desafiares e fazeres sair da zona de conforto. Por comigo percorreres as adegas da tua terra e por alinhares em todas as aventuras. Por me chamares à razão e por me dares a mão. Por me deixares sonhar e por me colares os pés na terra. Obrigada pela tua generosidade durante todo este caminho. Obrigada, Milinha e Aura, pelo tão importante apoio logístico. Agradeço também aos meus pais que me veem concluir mais um grau neste instituto, e que comigo sofrem e celebram. Obrigada por serem o porto de abrigo a quem posso sempre recorrer. O vosso apoio e amor incondicional fazem de mim uma pessoa melhor e uma cientista mais confiante. Sou-vos eternamente grata.

E é pelas mãos pequeninas e frágeis, de pele macia e fina, e (também) com as unhas sujas de terra da minha filha, que termino estes agradecimentos. Obrigada, querida filha, por me mostrares o verdadeiro caminho, e por me acompanhares nestes dois anos de vida e de tese, que é tão minha quanto tua. Esta tese que tem tantas das tuas lágrimas como das minhas, tanto da tua alegria como do meu entusiasmo, tanto da tua vida como da minha persistência. Agradeço a tua companhia e a tua presença que tornaram os meus dias infinitamente mais felizes. A ti, Helena, dedico toda de mim. Obrigada.

List of Acronyms

2-PA – 2-phenylethyl acetate

2-PE – 2-phenylethanol

AAT – Alcohol-O-acetyltransferase

ATF – Acetyltransferase

CCM – Carbon Central Metabolism

DAP – Di-Ammonium Diphosphate

DGE – Differential Gene Expression

DSBs – Double-stranded breaks

EGA – Esculin Glycerol Agar

ESI – Electrospray Ionization

FEL – Fast Evolving Lineage

FT-ICR-MS – Fourier Ion Cyclotron Resonance Mass Spectrometry

GlcNAc – N-Acetylglucosamine

GH – Glycoside Hydrolase

GJM – Grape Juice Medium (synthetic)

GO – Gene Ontology

GSEA – Gene Set Enrichment Analysis

IA – Isoamyl Alcohol

IAC – Isoamyl Acetate

ITS – Internal Transcribed Spacer

LSU – Large subunit

IPP – Isopentenyl Pyrophosphate

MLF – Malolactic Fermentation

MMB – Minimal Medium Broth

NES – Normalized Enrichment Score

NSY – Non-*Saccharomyces* Yeasts

PCA – Principal Components Analysis

PDC – Pyruvate decarboxylase

PFGE – Pulsed-Field Gel Electrophoresis

SEL – Slow Evolving Lineage

SO_2 – Sulfite Dioxide

SNP – Single Nucleotide Polymorphism

TDN – 1,1,6-trimethyl-1,2-dihydronaphtalene

TMM – Trimmed Mean of M values

uHRMS – Ultra-High Resolution Mass Spectrometry

UHPLC – Ultra-high Performance Liquid Chromatography

VFA – Volatile Fatty Acids

Table of Contents

List of Tables

Chapter I

Chapter II

Chapter III

Chapter IV.A

Chapter IV.B

List of Figures

Chapter I

Chapter II

Chapter III

Chapter IV.A

Chapter IV.B

Chapter V

Thesis Outline

The budding yeast *Saccharomyces cerevisiae,* the key species in vinification, is not very frequent in grape musts and is therefore usually inoculated in the form of commercial strains. Differently, other yeasts are naturally present in grapes and contribute for what is known as the "microbial *terroir*" which refers to the specific microbial communities, including bacteria, yeasts and fungi, that inhabit a particular vineyard or winemaking region, influencing the characteristics of the wine produced there.

The recent years of extensive advances in molecular microbiology approaches to understand how these microbial communities interact with each other showed that wine microbiota plays an important role in modulating its aroma and flavor. This thesis is focused on the study of yeasts belonging to the poorly characterized *Saccharomycodaceae* family, a sister family of *Saccharomycetaceae* (that includes *S. cerevisiae*) and that includes species belonging to the genera *Hanseniaspora* and *Saccharomycodes*, that are prevalent in grape musts, especially *Hanseniaspora* species. In the introductory chapter (**Chapter I**) an historical perspective on winemaking practices and wine consumption is provided, detailing its significance in human society and a description of its underneath microbial diversity. It is also detailed the impacts of Non-*Saccharomyces* Yeasts (or NSYs) in wine aroma (primary, secondary, tertiary) properties, along with a description of what are, currently the trends of their use as bio-flavourants.

Chapter II, the first chapter of results, describes the first whole-genome sequencing of a *Saccharomycodes ludwigii* strain, UTAD17 that has been isolated from a wine must in Portugal's Douro region. Notably resistant to sulfur dioxide, UTAD17 display distinct genomic features (especially compared to sister species of the *Hanseniaspora* genus) that may help to explain its tolerance to this chemical preservative. Additionally, a comparative genomic analysis encompassing a broader range of yeast species belonging to multiple genera (including *Candida*, *Pichia*, *Hanseniaspora* and *Saccharomyces*) revealed that UTAD17 encodes a few unique proteins not found in *Hanseniaspora* such as a set of beta-mannosyltransferases that are believed to significantly impact the structure of the cell wall. **Chapter III** continues the exploration of *Sd. ludwigii,* this time addressing a phenotypic and genomic characterization of *Sd. ludwigii* BJK_5C, a strain collected from an apple cider vinegar facility. Karyotyping for this strain, along with the previously used UTAD17 strain, revealed notable differences in genomic architecture of the species. Relevant genotype-phenotype assessments, based on

comparative phenotypic and genomic analyses of this strain along with UTAD17, were put forward concerning the different ability of the strains to grow on different carbon sources, their different levels of tolerance to sulfur dioxide or to acetic acid; and their different capabilities to ferment natural musts obtained from red and white grape varieties.

Chapter IV A focuses on the *Hanseniaspora* genus, addressing the potential role of these species in modulating primary wine aroma, something that has been little addressed since most literature focused the use of these species as modulators of secondary aroma. Extensive phenotyping of a collection of yeasts recovered from Moscatel Galego, a terpene-rich grape must, led to the identification of several high beta-glucosidase-producing strains, among which *H. uvarum* MJT198. This strain was further characterized from the genomic point of view and using extensive comparative genomic analysis with a number of publicly available *Hanseniaspora* genomes (herein annotated for the first time) two candidate beta-glucosidase-encoding genes are proposed to be underneath the observed phenotypes. Additionally, it is also described in this chapter the successful exploration of *H. uvarum* MJT198 as a co-adjuvant of *S. cerevisiae* in fermentations of Moscatel Galego natural must, resulting in wines with enhanced content of free terpenes.

In **Chapter IV B**, we continue the exploration of the results obtained in *H. uvarum* MJT198 natural must fermentations, performing an untargeted metabolomic analysis of the impact of this yeast in the supernatant of the fermentation and, necessarily, in the overall metabolome of the wine produced by the consortium, in comparison with the ones obtained with single-cultures of *H. uvarum* or *S. cerevisiae*. A molecular analysis, based on RNA-sequencing, was also undertaken to understand how *H. uvarum* genomic expression was impacted by the co-cultivation, an approach that is innovative since what has been performed in the literature in the field was much more focused on *S. cerevisiae* but not on the co-inoculated Non-*Saccharomyces* Yeasts (NSYs).

Chapter I. Introduction

Parts of the work described in this chapter are included in the following paper:

Tavares MJ, Santos D, Seixas I, Mendes Ferreira A, Mira NP, Wine aroma, a tale of molecules, enzymes and microbes: emphasis on the role of Yeasts. (submitted)

I.1 Tracing the Intertwined History of Wine and Humanity – A Journey of wine History

Winemaking has been linked to human civilization since ancient times. The origins of this practices date back to the Neolithic period, with the oldest evidence of wine production spanning from Georgia around 6000 BC (McGovern et al., 2017; Estreicher, 2017), Iran around 5000 BC (Estreicher, 2017; Voigt, 1983) and Armenia around 4100 BC (Estreicher, 2017), where the oldest-known winery was discovered (Keushguerian & Ghaplanyan, 2015). This early start indicates that the cultivation and fermentation of grapes began almost concurrently with the dawn of agriculture. Some historians suggest even that the cultivation of grapes for winemaking might have been one of the primary reasons for the development of agriculture (Bode, 1992). The ease of transport and storage, compared to other foods, is believed to turn wine into a valuable commodity for the producing communities, positively influencing agricultural practices and trade (Bode, 1992). The spread of winemaking coincided with the expansion of ancient civilizations (Bode, 1992). The Egyptians established a royal winemaking industry around 3000 BC (Verstrepen et al., 2006) after grape cultivation was introduced from the Levant, a region that embraced viticulture since the Bronze Age and maintained it until the decline of Byzantine rule (Robinson, 1994). This tradition, enriched over times, left an inheritance for the Greeks and Romans (Robinson, 1994). The ancient Greeks provided an indelible mark on oenology, nurturing unique grape varietals and inventing practices such as resin-lined wine vessels, an ancestor of the aromatic Retsina (Robinson, 1994). The Romans, inheriting this legacy, were instrumental in refining winemaking techniques, integrating wine into the daily fabric of their society (Robinson, 1994). Historically, wine has been more than just a beverage, holding significant cultural and religious importance. In fact, wine's role in religious ceremonies spans various cultures and times: ancient Greeks and Romans revered the wine gods, Dionysus and Bacchus, respectively (Varriano, 2010); it has been a part of Jewish rituals since Biblical times and is central to Christian communion practices (McGovern, 2013; Outreville, 2021). Even in Islamic history, despite prohibitions against alcohol, wine's distillation for medicinal purposes was (and is still) notable (Hayyan, 1935).

Wine also played crucial roles in shaping economies and societies, as previously said. In ancient Greece, wine began as a staple but grew into a significant economic driver, with Greek wine being widely known and exported (Bode, 1992; Li et al., 2018). The wine industry's impact is also evident in events like the French Revolution, where a wine tax contributed to the social upheaval leading to the fall of the monarchy (Plack, 2012). Additionally, in society, wine

has often been associated with celebration and social gatherings. It is a symbol of joy, prosperity, and is used in various ceremonies, from religious rituals to weddings and communal celebrations.

I.1.1The dynamic ecosystem of wine fermentation: exploring endogenous and inoculated yeasts

"Wine is the most healthful and hygienic of beverages."

Louis Pasteur

Despite its prevalence and ordinary consumption, wine is distinguished by various characteristics that render it unique and distinct across regions, vintages, and grape varietals. An experienced wine taster can discern over a hundred different aromas, contributing to the wine's complex bouquet (**Figure I. 1**). These aromas depend on factors ranging from the grape variety, which itself carries genetic information that can express a spectrum of scents from the subtle fragrance of fresh berries to the flowery undertones of roses; to the microbial *terroir*, where the unique yeast and bacteria in the vineyard environment influence fermentation and introduce distinctive flavors (*e.g.*, buttery notes or sharp, tangy accents). The fermentation conditions, such as temperature and duration, also play a critical role, with cooler fermentations preserving fresh fruit aromas while warmer conditions might encourage the development of spicier, more robust notes (Molina et al., 2007; Killian & Ough, 1979; Torija et al., 2003). Lastly, storage methods, including the type of vessel in which the wine is aged (be it stainless steel, which maintains a wine's crispness -Pichler et al., 2023; Montalvo et al., 2021-, or oak barrels, which can impart a vanilla richness -Echave et al., 2021; Issa-Issa et al., 2021-), along with the length of aging, further shapes the wine's profile, rounding out flavors and adding layers of complexity to the final product enjoyed by consumers worldwide (Styger et al., 2011).

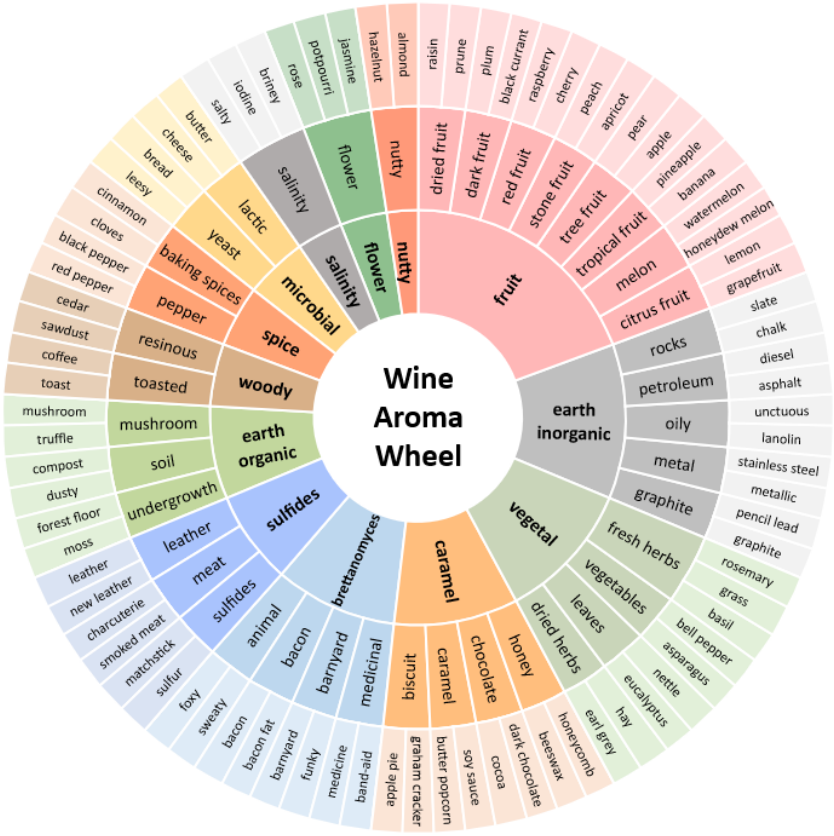


Figure I. 1- Wine Aroma Wheel. The wine aroma wheel is one of the tools used by wine tasters to characterize the sensory profile of a given wine during a wine tasting event (Schuring, 2019).

The aromatic complexity of wine is intricately linked to the dynamic interaction between its alcohol content and the many chemical compounds that contribute to its mouthfeel. The alcoholic content of this beverage results from the fermentation process facilitated by yeasts, such as *S. cerevisiae*, which convert natural grape sugars into ethanol and carbon dioxide. While modern understanding attributes the production of wine to specific yeast species, the historical context had winemaking surrounded by mysticism and mystery. In ancient times, the phenomenon of fermentation, transforming grape juice into a complex, aromatic, and alcoholic beverage, was often credited to spiritual or supernatural entities (Chambers & Pretorius, 2010). During the Middle Ages, empirical knowledge led winemakers to observe the froth on fermenting casks, using it to inoculate the next batch—a primitive form of yeast inoculation, marking a crucial development in winemaking (Unger, 2004).

The turning point came in 1680 when Dutch naturalist Anton van Leeuwenhoek became the first to directly observe yeast cells through his developed lenses (Chambers & Pretorius, 2010) (**Figure I. 2**). Two centuries later, Louis Pasteur postulated that these microscopic organisms were the key to fermentation, capable of imparting diverse flavors to wines from the same grapes (Chambers & Pretorius, 2010; Jolly et al., 2014). Emil Hansen's groundbreaking

work in 1890 led to the isolation of pure yeast strains. Dr. Hermann Müller-Thurgau then employed these strains as starter cultures for winemaking, marking a pivotal moment in yeast utilization (Verstrepen et al., 2006; Chambers & Pretorius, 2010; Nardi & Romano, 2023) (**Figure I. 2**). This newfound knowledge surrounding yeasts paved the way for the emergence of the first commercial *S. cerevisiae* cultures in 1965 (**Figure I. 2**). Winemakers embraced these innovative biological products to enhance predictability in fermentation speed and overall wine quality, especially in regions with limited winemaking traditions, thereby marking a significant stride in the evolution of winemaking practices (García et al., 2016).

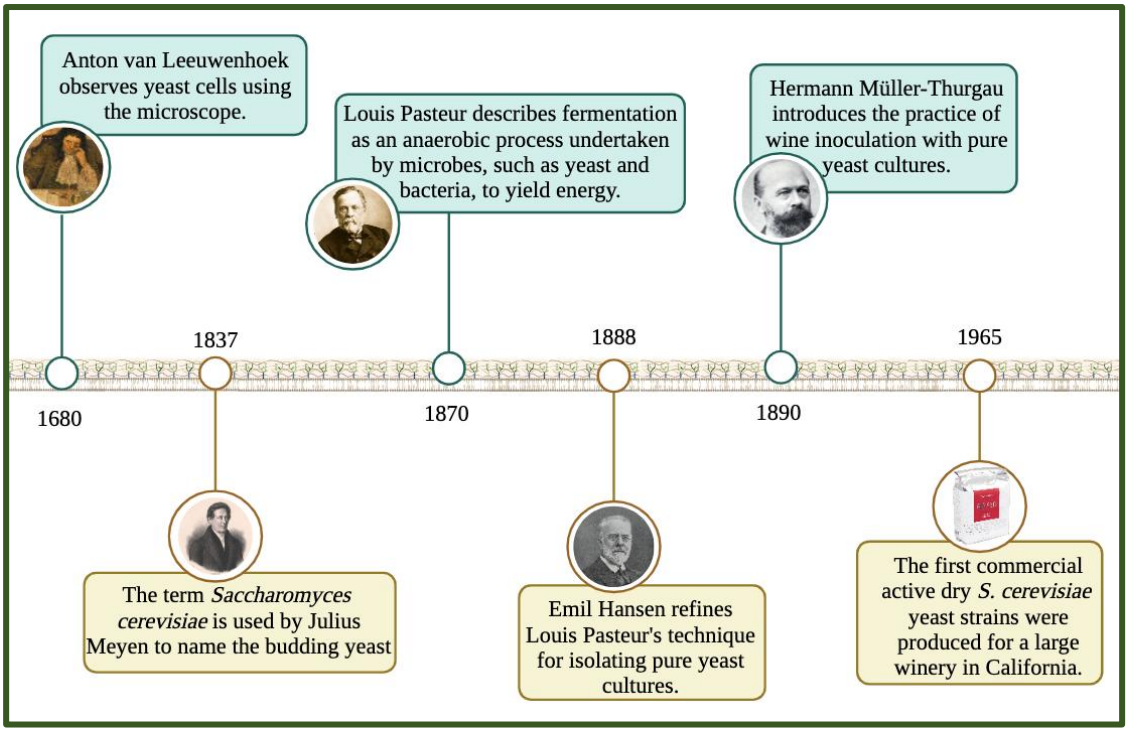


Figure I. 2– Historical milestones surrounding the knowledge of *S. cerevisiae* and yeast fermentation process. The first microscopic observation of *S. cerevisiae* cells was in 1680, by Anton van Leeuwenhoek. This milestone was the first step towards the creation of the first commercial dry inoculum for wine fermentation, that revolutionized winemaking. This illustration was specifically prepared for this thesis.

The evolution of genomics, catalyzed by the publication of the first *S. cerevisiae* genome in 1996, enabled the broadening of knowledge concerning the budding yeast. The expansion to wine strains unveiled the genomic nuances specific to the winemaking environment, laying the foundation for the study of other yeast genera (**Figure I. 3**). Although broadly used and commercialized in the winemaking context, the prevalence of wild *S. cerevisiae* strains in vineyards is scarce when comparing to other yeast genera (García et al., 2016; Cray et al., 2013; Boulton et al., 1999). In fact, there are several other wine-related yeast species that inhabit wineries and vineyards, that differently contribute to wine fermentation.

However, the high predictability of *S. cerevisiae* prompted the development of commercial *S. cerevisiae* starter cultures for wine production (*e.g.*, *S. cerevisiae* EC1118, *S. cerevisiae* QA23, *S. cerevisiae* VIN13) which made wine fermentations conducted with these strains much more reliable and faster, compared to spontaneous fermentations (**Figure I. 3**).

The vast majority of these other yeasts do not belong to the *Saccharomyces* genus and therefore they are known as Non-*Saccharomyces* Yeasts (or NSYs). Today, there are recognized 149 NSY genera which comprise nearly 1,500 species (Jolly et al., 2014; Kurtzman et al., 2011), 40 of these being commonly isolated from grape must (*e.g., Hanseniaspora uvarum, Torulaspora delbruecki, Metschnikowia pulcherrima*) (Jolly et al., 2014; Ciani et al., 2010). Historically, NSYs were deemed of importance and were mostly regarded as undesirable due to their capacity to cause spoilage (Jolly et al., 2014). This association resulted from NSYs being inefficient fermenters with a low tolerance for ethanol and thus these species were commonly isolated from contaminated musts, being linked to undesirable outcomes in wine quality, such as elevated volatile acidity levels (associated to acetic acid production), turbidity, or sedimentation (Jolly et al., 2014; Padilla et al., 2016). Nevertheless, ancient winemaking practices, which have long involved spontaneous fermentations, acknowledged the relevance of indigenous yeast presence in grape must to enhance particular characteristics in the resulting wines (Jolly et al., 2014). Therefore, despite posing a higher risk of spoilage and presenting challenges in the control of the winemaking process, wines fermented spontaneously exhibit heightened organoleptic properties, including superior mouthfeel and distinct aromas and flavors (Jolly et al., 2014; Fleet, 2003; Fleet, 2008). These contrasting features led to a scenario in which NSYs are friends, besides being foes.

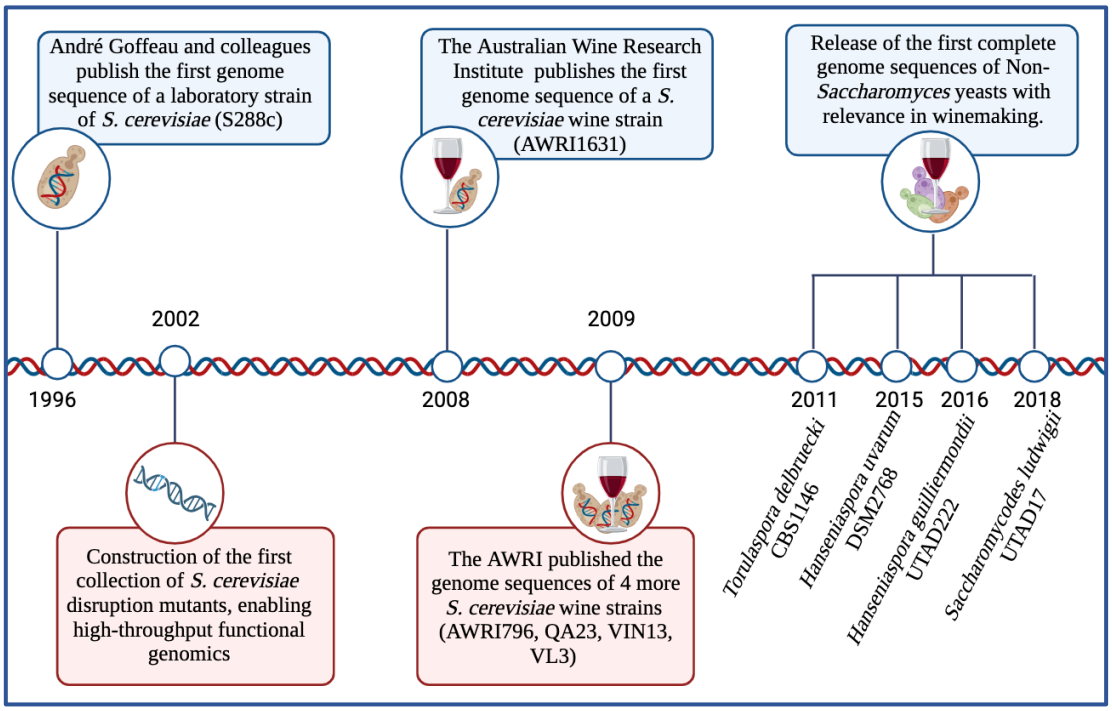


Figure I. 3– Milestones in yeasts' genomics, with interest in wine aroma. Starting with the first genome sequence of the laboratory strain *S. cerevisiae S288c*, to the senquencing and analysis of wine-related NSYs, such as *Torulaspora delbrueckii*, *Hanseniaspora uvarum*, *Hanseniaspora guilliermondii*, or *Saccharomycodes ludwigii*. This illustration was specifically prepared for this thesis.

I.1.2 The impact of Non- *Saccharomyces* Yeasts in winemaking

Wine-related NSYs include those found on grapes or in vineyards; in wines (whether healthy or spoiled); or associated with the winery/cellar structure and equipment or with the corks (Boulton, 1999). Insects have also been identified as important carriers of these yeasts into fermenting must (Jolly et al., 2014; Stefanini, 2018; Fogleman et al., 1982). Numerous external factors impact the presence and diversity of the wine-associated NSY community (Jolly et al., 2014), including the geographic location, climatic conditions, pesticide application, grape variety, the stage of ripening, grape health, harvesting methods, and specific weather conditions in each vintage have all been found to influence the yeast community associated to grape musts (Padilla et al., 2016). Hygiene practices, coupled with the consistent use of sulfur dioxide (SO_2) that is for long used to control microbial activity in musts, also affect the yeast-associated wine communities (Jolly et al., 2014) (**Figure I. 4**).

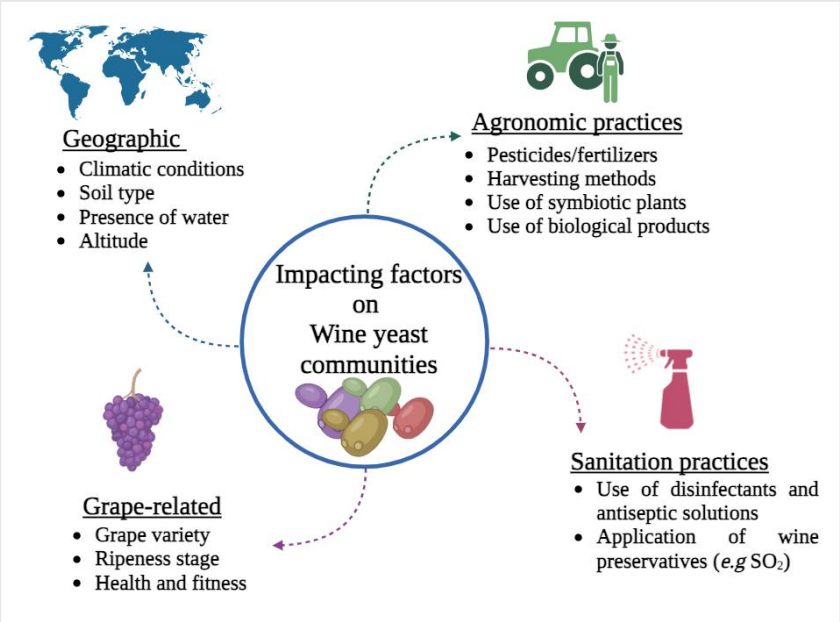


Figure I. 4– External factors impacting wine yeast communities. Numerous external factors impact the presence and diversity of the wine-associated NSY community, including greographic factors (*e.g.*, climate conditions, altitude, agronomic practices (*e.g.*, use of pesticides or fertilizers, harvesting methods), sanitation practices (*e.g.*, use of antiseptics and wine preservatives), or grape-related factors (*e.g.*, grape variety and ripeness).

Despite these variations, the NSYs species identified on grapes and in wines exhibit a notable similarity worldwide (Jolly et al., 2014). Specifically, in the initial days of spontaneous fermentation of grape must, the yeast population is predominantly composed of apiculate yeasts essentially belonging to the *Hanseniaspora/Kloeckera* (the teleomorphic and anamorphic state, respectively (Jolly et al., 2014) and *Candida* genera. This is followed by several species from the genera *Metschnikowia* and *Pichia*, and occasionally, species from *Brettanomyces*, *Kluyveromyces*, *Schizosaccharomyces*, *Torulaspora*, *Rhodotorula*, *Zygosaccharomyces*, and *Cryptococcus* genera (Jolly et al., 2014; Padilla et al., 2016; Fleet, 2003; Fleet, 2008; Lambrechts & Pretorius, 2000; Pretorius, 2000; Barata et al., 2000). Typically, these yeasts are divided into three main groups that try to cluster them according with their metabolic capacities: i) primarily aerobic yeasts; ii) apiculate yeasts with low fermentative activity; and iii) yeasts with fermentative metabolism (Jolly et al., 2014; Fleet, 1998; Combina et al., 2005) (**Table I. 1**).

Table I. 1- *Non-Saccharomyces* Yeasts found in wine. The abundance of NSYs in wine fermentation gradually changes according to the yeast's characteristics. Hence, in the fermentation onset, primarily aerobic yeast (like *Pichia* sp. or *Candida sp.*) dominate the yeast communities, being replaced by apiculate yeasts with low fermentative ability (like those from the *Hanseniaspora* genus). In the end of fermentation, there is a predominance of yeasts with fermentative metabolism, such as *Zygosaccharomyces bailii* (Jolly et al., 2014; Fleet, 2003; Fleet, 2008; Combina et al., 2005).

Non-*Saccharomyces* group	Fermentation stage	Genus	Species
Aerobic yeasts	Fermentation onset	*Pichia*	*P. kluyveri* *P. fermentans* *P. guilliermondii*
		Debaryomyces	*D. hansenii*
		Rhodotorula	*R. mucilaginosa*
		Candida	*C. stellata*
		Cryptococcus	*C. albidus* *C. laurentii*
Apiculate yeasts with low fermentative activity	Early stages of fermentation	*Hanseniaspora*	*H. uvarum* *H. guilliermondii* *H. occidentalis*
Yeasts with fermentative metabolism	End of fermentation	*Kluyveromyces*	*K. marxianus* *K. thermotolerans*
		Torulaspora	*T. delbruecki*
		Metschnikowia	*M. pulcherrima*
		Zygosaccharomyces	*Z. bailii*
		Saccharomycodes	*Sd. ludwigii*

During fermentation, with the increasing concentrations of ethanol, the abundance of each yeast varies sequentially (**Figure I. 5**). For instance, in fresh must and in the early stages of fermentation, there is a prevalence of aerobic and weak fermentative yeasts, such as *Pichia* or *Candida*; but these successively replaced by others that can endure progressively higher ethanol concentrations. Adding to the intoxicating effects of ethanol, the combined effects of decreased nutrient availability, SO_2, or oxygen deficiency will also cause impact in viability of the different yeast species. Interestingly, direct cell–to-cell contacts in the presence of high concentrations of viable *S. cerevisiae* yeasts has also been found to be an important mechanism by which the viability of NSYs can be reduced (Jolly et al., 2014; Nissen et al., 2003).

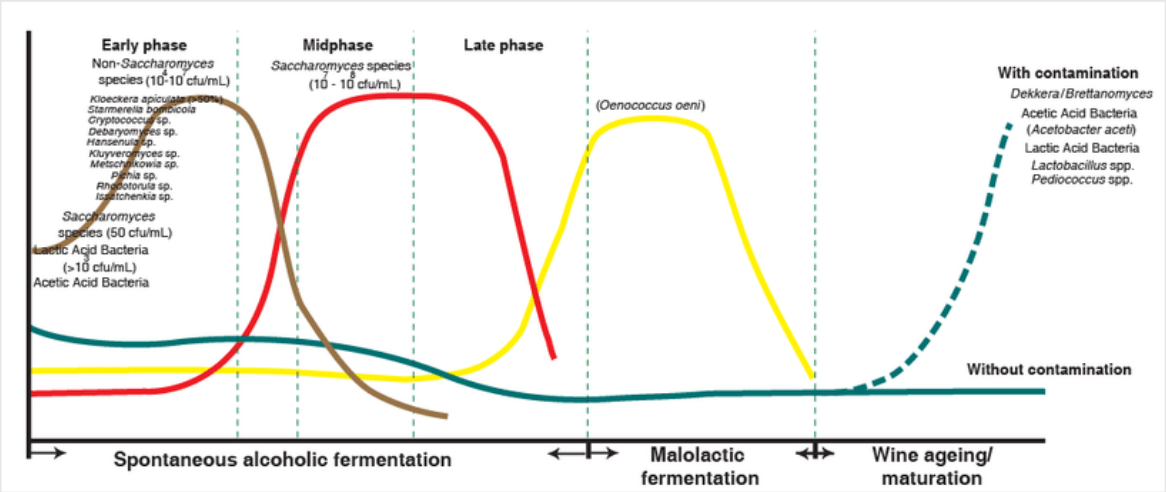


Figure I. 5– Microbial evolution during wine fermentation. In the early stages of fermentation, Non-*Saccharomyces* Yeasts dominate the wine must, due to low alcohol contents, being gradually replaced by *S. cerevisiae*. Adapted from Ghosh (2015).

I.1.3 NSYs and wine aroma - A chemical dance of grapes and microbes

The influence of NSYs on wine flavor is dependent on the concentration of metabolites they produce, which is intricately correlated to the activity of the producing species (Jolly et al., 2015). Environmental conditions in the must including high osmotic pressure, an equimolar mixture of glucose and fructose, the presence of SO_2, suboptimal growth temperature, increased alcohol concentrations, anaerobic conditions, and variable nutrient levels, all play crucial roles in shaping the survival and growth of different yeasts and, consequently, their impact on wine aroma. Wine aroma is a complex concept that usually combines the definitions of aroma, flavor and bouquet. While the aroma *per se* is defined as the mixture of odorous, volatile compounds; wine flavor is the overall sensory impression of both aroma and taste compounds, encompassing more complete factors such as acidity, alcoholic strength, astringency and bitterness (Lambrechts & Pretorius, 2000; Canon et al., 2022; Lin et al., 2019). Additionally, the wine bouquet is a broader concept, aggregating other factors that contribute to the overall sensorial perception of this beverage, such as grape variety, microbial terroir, fermentation conditions, microorganisms inoculated, aging or bottling. Hence, the definition of wine aroma bouquet encloses the overall sensory experience taken in all steps of wine production (Robinson et al., 2014; Ruiz et al., 2019). There are three main categories of wine aroma (**Figure I. 6**) varietal, or primary, that is the aroma fraction related to the grapes; secondary, dependent of the fermentation process; and tertiary, connected to the aromas developed during the aging process (Ruiz et al., 2019; Belda et al., 2017a).

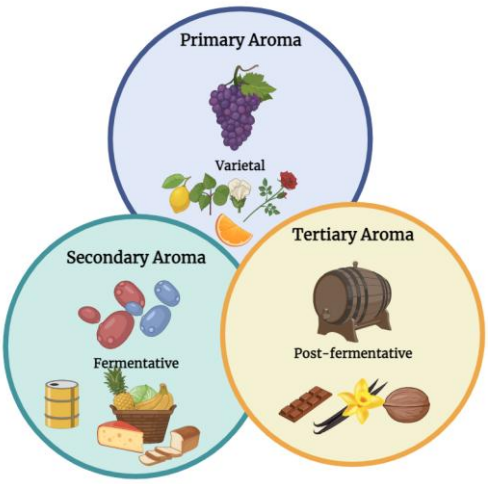


Figure I. 6- Wine Aroma. The primary aroma, or varietal, of wine originates in the grapes and is composed by free aromatic molecules and glycosylated precursors; The secondary aroma is the fermentative aroma that arises upon yeast activity; the tertiary aroma is the post-fermentative aroma, originated during wine aging. This image was especially prepared to be included in this thesis.

I.3.1 Primary Aroma

Wine primary aroma (also called varietal) is obtained from the principal flavor compounds in grape must (Ruiz et al., 2019) and originates from the secondary metabolites present in grapes. This aroma can be classified into several families related with specific descriptors, such as floral, green fruit, citrus fruit, stone fruit, tropical fruit, red fruit, black fruit, dry fruit, herbaceous, herbal, spices, and others (Ruiz et al., 2019). The volatile compounds that contribute to these varietal characteristics of wines are terpenes, methoxypyrazines, organic acids, tannins and precursors of aromatic aldehydes and volatile thiols (Styger et al., 2011; Chen et al., 2022). However, most of these compounds are present in wine in the form of water-soluble glycosides or conjugated with amino acids, an issue that hampers their aromatic potential (Styger et al., 2011; Belda et al., 2017a; Belda et al., 2017). For instance, although monoterpenes can be found as free volatile compounds, they are present at much higher concentrations in a non-odorous form, linked to sugar molecules (Caffrey & Ebeler, 2021). These compounds can, however, be released through the hydrolysis of the glycosidic precursor, achievable by microbe-mediated enzymatic degradation (Michlmayr et al., 2021; de Ovalle et al., 2021). Hence, although it is true that the primary aroma derives from grapes, it is also true that microbial activity (either endogenous, from the microbiome, or exogenously added) can modulate its perception through the secretion of glycosidases and peptidases that break the glycosidic bonds, releasing the terpenes into their volatile form (Michlmayr et al., 2021; González-Barreiro, 2015).

I.3.1.1Terpenes: Monoterpenes, sesquiterpenoids and norisoprenoids

Terpenes, or isoprenoids, are organic compounds consisting of isoprene (a five-carbon building block) and represent the largest class of natural products, encompassing different isomers and enantiomers (Buckle, 2015). Their distinctive carbon skeleton facilitates clustering in monoterpenes, sesquiterpenes, diterpenes, triterpenes, depending on whether they harbor two, three, four, or six isoprene units, respectively (Buckle, 2015; Aldred, 2009; Mosquera et al., 2021). In plants, these compounds are responsible for a wide range of basic cellular processes, such as respiration, photosynthesis, growth, development, reproduction, and adaptation to environmental conditions (Ninkuu et al., 2021; Singh & Sharma, 2015). Volatile terpenes are also believed to have a major contribution in plant defense, by serving as airborne signals that induce defense response in undamaged parts of the plant, and also by inducing responses in neighboring plants (Lin et al., 2019; Wang et al., 2019; Sharma et al., 2017). The most relevant functional present in terpenes are alcohols, but some terpenes contain ketones or exist in an oxidized form (being known in this case as terpenoids) harboring a basic isoprenoid structure coupled with an oxygen-containing ring (Jackson, 2008). In wine, monoterpenes (with a C10 skeleton) are largely dominant, but sesquiterpenes (C15 compounds) and C13-norisoprenoids can also be found in significant quantities (Slaghenaufi & Ugliano, 2018, Song et al., 2018). Both monoterpenes and sesquiterpenes derive from isopentenyl pyrophosphate (IPP) (**Figure I. 7**) (Wang et al., 2019). Usually located in the grape skin, monoterpenes are synthesized in the plastid through the 2C-methyl-D-erythritol-4-phosphate (DOX/MEP) biosynthetic pathway, while sesquiterpenes are synthesized in the cytosol following the mevalonate (MVA) pathway (**Figure I. 7**) (Wang et al., 2019; Carrau et al., 2008; Black et al., 2015). In both pathways, synthesis of the final terpene involves the enzyme terpene synthase (Carrau et al., 2008; Carrau et al., 2005).

Contrasting to monoterpenes and sesquiterpenes that have a dedicated biosynthetic pathway, C13-norisoprenoids result from the degradation of beta-carotenoids, which is an independent process that occurs in the chloroplast (**Figure I. 7**) (Black et al., 2015). As said above, although terpenes are typically volatile, a considerable portion of these compounds exist in the must complexed with glycosides, or as di- or triols, which are not aromatic (Caffrey & Ebeler, 2021; Michlmayr et al., 2012; Carrau et al., 2008; Martin et al., 2016; Zerbib et al., 2018; Mateo & Jiménez, 2000). Hence, recent studies have shown increasing interest in breaking these complexes, to release the terpenes and increase wine aromatic complexity.

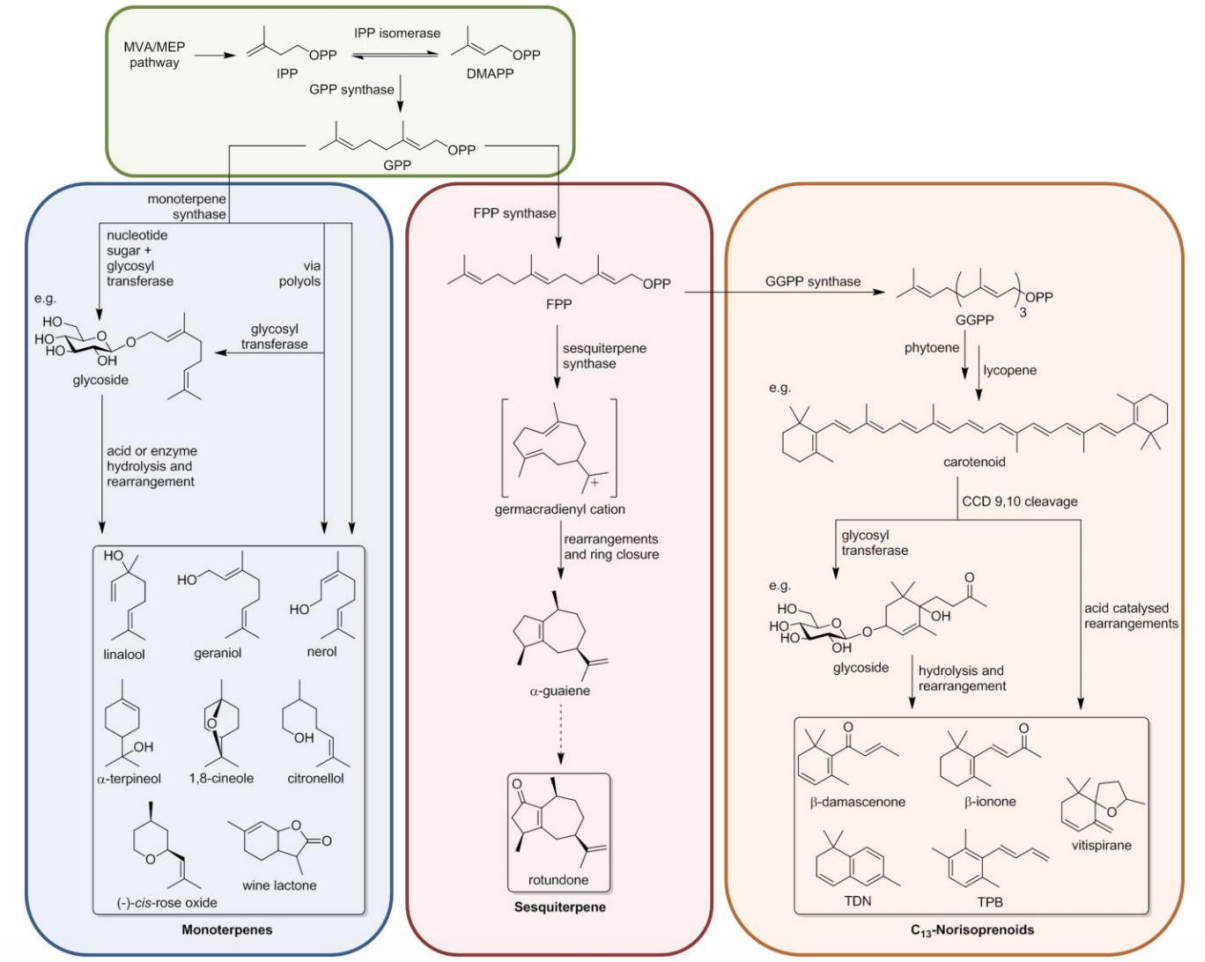


Figure I. 7 - Summary of the biosynthetic pathways of terpenes usually found in grapes. Monoterpenes and sesquiterpenes both derive from isopentenyl pyrophosphate (IPP), while monoterpenes are synthesized in the plastid following the mevalonate (MVA) pathway and sesquiterpenes are synthesized in the cytosol through the 2C-methyl-D-erythritol-4-phosphate (MEP) pathway. C13-norisoprenoids are produced through beta-carotene degradation, a process that occurs in the chloroplast. Adapted from Black et al., (2015).

Monoterpenes

Monoterpenes are cyclic molecules composed of two isoprene units, known to have a significant impact in modulation of wine aroma due to a potent aromatic footprint resulting from a very reduced sensorial threshold (Caffrey & Ebeler, 2021; Jackson, 2008; Flamini et al., 2018). In fact, more than 50 monoterpenic compounds have been isolated from wines including linalool (associated to citrus and flowery aroma), (E)-hotrienol (flowery), citronellol (green citrus), geraniol (roselike), nerol (rose-like), (-)-cis-rose oxide (geranium oil), and α-terpineol (floral, woody) (Ribéreau-Gayon et al., 2006) (**Figure I. 7; Figure I.9A**). Grape cultivars can be grouped according to their total free monoterpene concentration, being classified as neutral, if terpenes are present at very low concentrations (*e.g.*, Chardonnay); aromatic, if concentration of monoterpenes is between 1-4 mg/L (e.g., Riesling); and Muscat

if the concentration of monoterpenes is close to 6 mg/L (*e.g.*, Muscat blanc or Muscat of Alexandria) (Lin et al., 2019). Linalool and geraniol have the most importance in Muscat grapes (Mateo & Jiménez, 2000) being responsible for the floral, fruity and citrus aroma of their corresponding wines (Styger et al., 2011). Significantly, the interactions between combinations of monoterpenes and their metabolites results in a very characteristic spectrum of flavor intensities among Muscat varieties (Lin et al., 2019; González-Barreiro et al., 2015; Aldred, 2009). As stated before, glycosylated forms of monoterpenes are the more common form of these molecules (circa ~90%) (**Figure I. 8**) found in grape varieties such as Muscat, Sylvaner, Weisser, Riesling, and Gewürztraminer (Lin et al., 2019; Mateo & Jiménez, 2000; Ribéreau-Gayon, 2006). Terpenes can be found bound to disaccharides such as arabinoglycosides/vicianose (in which the terpene is bound to an arabinose-glucose saccharide), apioside/acuminose (in which the terpene is bound to an apiose-glucose saccharide), and rhamnoside (in which the terpene is bound to a rutinose-glucose saccharide) (Moreno & Peinado, 2012).

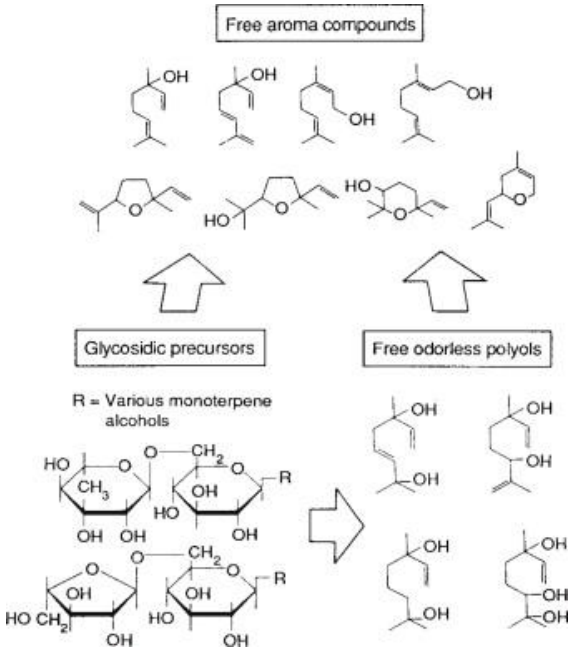


Figure I. 8- Glycosidic precursors and free odorless polyols are present in the grape, being a reserve of odorless precursors in the fruit. Adapted from Jackson et al., (2008).

Breaking of the glycosidic bond to increase the release of free monoterpenes, can be achieved enzymatically or by acidic hydrolysis. While acidic hydrolysis occurs very slowly during wine storage (although it can be accelerated by heat induction) (Sánchez-Acevedo, 2024), enzymatic activity is more frequent as the result of the secretion of beta-glucosidase hydrolytic enzymes by plant cells or by microorganisms present in the fermentation must

(Dziadas & Jeleń, 2016). The activity of beta-glucosidase enzymes is particularly relevant in white wines, due to their high levels of bound terpenes (Fernández-Pacheco, 2021), being less relevant in red cultivars as their activity can also result in loss of color due to anthocyanin breakage (Behrens, 2018; Monteiro et al., 2019). Despite this, the use of beta-glucosidases improved the organoleptic profile of wines obtained with Shiraz, a red grape variety (Jackson, 2008).

Sesquiterpenes

Sesquiterpenes are a terpene subclass with 15 carbon atoms and three isoprene units (**Figure I. 7**) (Li et al., 2020). These are widely used in the perfumery industry, due to their aromatic properties and abundance in plant oils (like the ones obtained from carrots or ginseng) (Slaghenaufi & Ugliano, 2018; Li et al., 2020). Based on the number of their carbon rings and chemical structure they can be classified as acyclic, monocyclic, bicyclic, tricyclic and tetracyclic, and can also be grouped according to the number of oxygen-containing groups (Li et al., 2020; Cincotta et al., 2015). In wine, several descriptors have been associated with sesquiterpenes, such as the "black pepper" aroma of Shiraz wines (Li et al., 2020; Wood et al., 2008) originated by rotundone, an oxygenated bicyclic sesquiterpenes (Li et al., 2020). Besides rotundone, there are other sesquiterpenes that can be identified in grapes and wine (97 compounds have been identified), such as alpha-langene, beta-caryophyllene, alpha-caryophyllene, and germacrene D, and oxygenated sesquiterpenes, including farnesol and nerolidol (**Figure I. 7, Figure I. 9B**) (Li et al., 2020). All sesquiterpenes identified in grapes are derived from farnesyl diphosphate that originates a farnesyl carbocation, that, depending on the downstream pathway can originate different molecules (Li et al., 2020; Könen and Wüst, 2019; Smit et al., 2019).

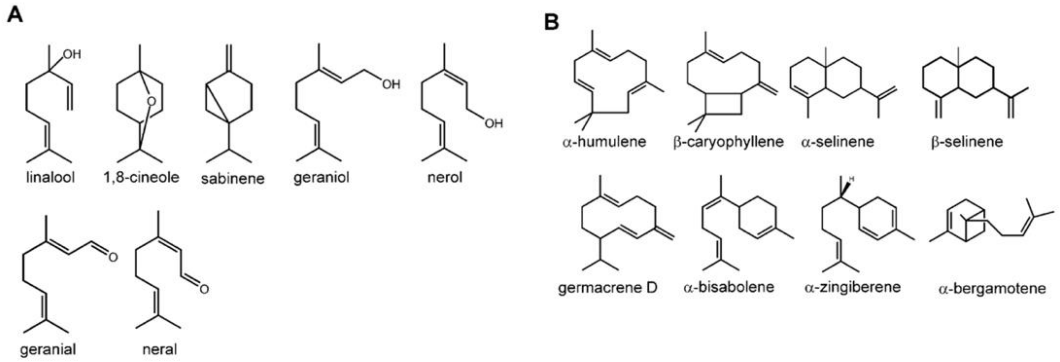


Figure I. 9– Examples of terpenes in the form of A) Monoterpenes, which are C10 compounds, and B) Sesquiterpenes, which are C15 compounds. In wine, the most common monoterpenes are linalool, geraniol and nerol, while germacrene D and ß-caryophyllene are sesquiterpenes also found in this environment. Adapted from Li et al., (2020).

C13-isoprenoids

Contrasting to monoterpenes and sesquiterpenes, whose biosynthesis originates in the common ancestor isopentenyl pyrophosphate, IPP, the C13-norisoprenoids detected in wine are produced from carotenoids (Black et al., 2015; Rodríguez-Bustamante & Sánchez, 2007; De Luca, 2011) via a three-step reaction, that starts with the oxidative cleavage (using molecular oxygen as a receptor for the electrons) of the carotenoid by a carotenoid cleavage dioxygenase, which results into C13 subunits. The resulting metabolites can be then further enzymatically transformed, through glycosyltransferases, in non-volatile C13-norisoprenoid glycoconjugates (Black et al., 2015; De Luca, 2011). An example of this reaction is the conversion of zeaxanthin or lutein into 3-hydroxy-beta-ionone and 4,9-dimethyldodeca-2,4,6,8,10-pentaene-1,12-dialdehyde, catalyzed by the enzyme carotenoid cleavage dioxygenase (De Luca, 2011) (**Figure I. 10**).

As is the case for the other terpenes, most terpenoids are also found in grapes in their glycosylated forms, usually more abundant (two to three times) in skin than in the pulp of the grapes (Black et al., 2015; De Luca, 2011). When free, C13-norisoprenoids provide aromas related with eucalyptus odor (produced by isomers of vitispirane) (Black et al., 2015; De Luca, 2011; Cheynier et al., 2010), tea (theaspirane) (Black et al., 2015; Cheynier et al., 2010), rose and exotic flowers (beta-damascenone) (Black et al., 2015; Cheynier et al., 2010), raspberry/violet (beta-ionone) (Black et al., 2015), but also kerosene (TDN - 1,1,6-trimethyl-1,2-dihydronaphtalene) (**Figure I. 10B**) (Black et al., 2015; Cheynier et al., 2010). Much of the differences observed in the typical aromas of a particular wine cultivars are attributed to C13-norisoprenoids, which is the case of TDN in Riesling wines that have characteristic petrol notes

(Black et al., 2015; Cheynier et al., 2010). Although the kerosene notes are well noticeable in these wines, TDN is not easily detected in grapes and must. Instead, it is much more frequent to find its precursors (like 2,6,10,10-tetramethyl-1-oxaspiro[4.5]deca-2,6-dien-8-diol), particularly abundant in grapes cultivated in sunny regions (Black et al., 2015). Besides TDN, the most important C13-norisoprenoids in wine are beta-damascenone and beta-ionone (Black et al., 2015; De Luca, 2011; Cheynier et al., 2010). Beta-damascenone adds the aroma of roses and honey, increases the aromas associated with ethyl esters and decreases methoxypyrazines-related aromas (Black et al., 2015). Since the concentration of this compound highly decreases with aging, its effects are more pronounced in young wines than in barrel-aged ones (Cheynier et al., 2010). Beta-ionone, attributes the aroma of violet and raspberry, and unlike beta-damascenone, its concentration is not reported to decrease over time (Black et al., 2015).

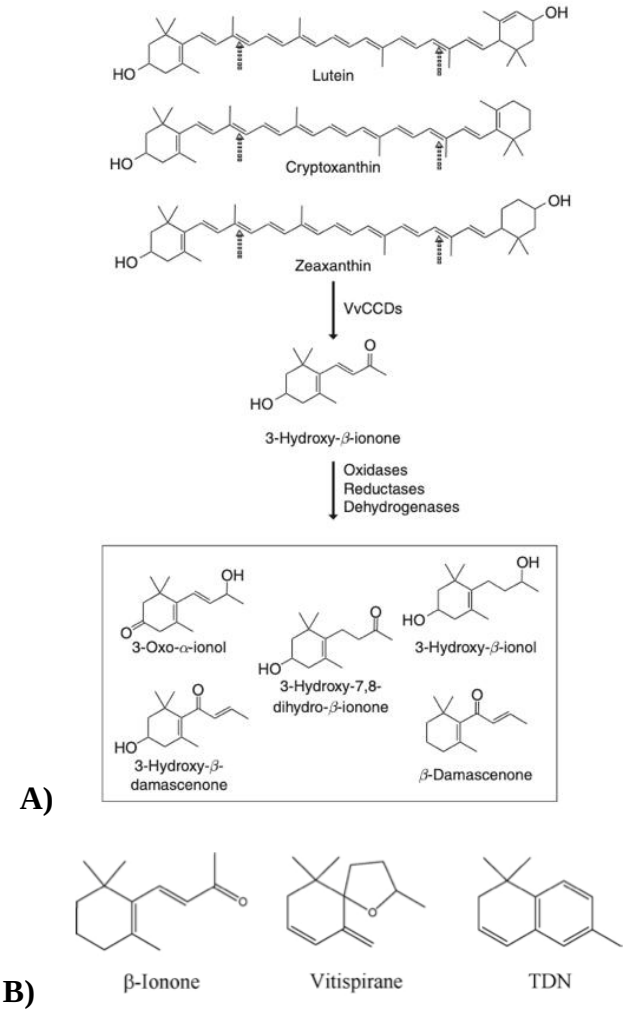
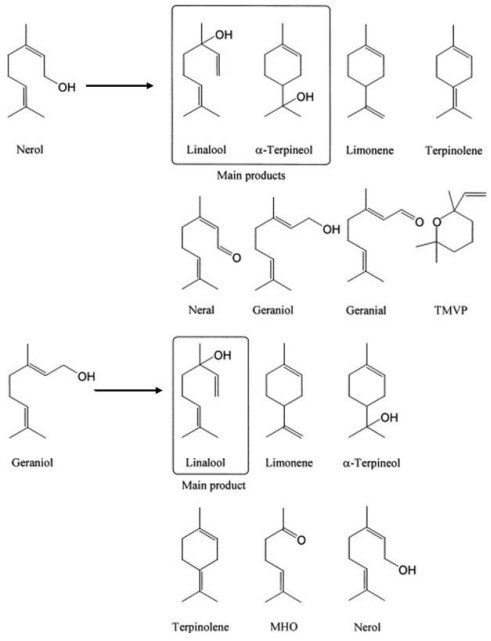


Figure I. 10 -C13-norisoprenoids. A) Conversion of lutein, cryptoxanthin or zeaxanthin into 3-hydroxy-β-ionone and 4,9-dimethyldodeca-2,4,6,8,10-pentaene-1,12-dialdehyde, catalyzed by the enzyme carotenoid cleavage dioxygenase (VcCCDs). B) Most common C13-norisoprenoids found in wine. The norisoprenoid β-ionone provides aromatic descriptors of violet and raspberry to wine, while β-damascenone adds the aroma of roses and honey. Contrasting with these pleasant aromas, TDN presence in wine results in a kerosene-like aroma. Adapted from De Luca et al., (2011) and Vinholes et al., (2009).

I.1.3.1.2 The evolution of terpene concentration and profile with wine aging

The concentration of terpenes has been reported to change along wine ageing (Jackson, 2008), which also impacts the type, and the proportions of terpenes present in aged wines. Although some increase in content is noticeable, because of the release of previously bound terpenes, the majority of the observed changes results in loss of terpene-related aromas (**Figure I. 11**) (Jackson, 2008). This occurs due to oxidations that may happen because of the higher temperatures and acidic conditions registered during ageing, resulting in a gradual substitution of monoterpene alcohols to monoterpene oxides, whose sensory thresholds are 10 times higher (Jackson, 2008). A paradigmatic example is linalool that can be oxidized, via an epoxide, to four oxides (cis- and trans- furan linalool oxide and cis/ trans pyran linalool oxide), all with higher sensorial threshold levels (Marais, 1983). These changes are more significant in grape varieties that depend largely on monoterpene alcohols to express their characteristic aroma and fragrance, such as Muscat (Jackson, 2008). The conversion of geraniol into alpha-terpineol and nerol (**Figure I. 11A**) is another modification registered to occur during ageing, which inevitable changes wine aroma from a more floral scent to a musty pine-like scent, associated to α-terpineol (**Figure I. 11A**). Other chemical modifications include the occurrence of cyclization, in which the terpene is transformed into a lactone; or conversion into a ketone, such as α- and β-ionone, or spiroethers, such as vitispirane (Jackson, 2008).



A)

B)

Compound	Olfatory perception threshold (! g/L)
" -terpineol	400
Citronellol	18
Geraniol	130
Linalool	50
Nerol	400

Figure I. 11- A) Terpene interconversion. During aging, the terpene profile of wines can be altered due to acid catalyzed reactions that can occur. For instance, during these reactions, geraniol is mainly converted into linalool, but also to α-terpineol and nerol. B) Sensory thresholds of the main monoterpenes found in wine. Adapted from Demyttenaere et al., (2000) and Cus et al., (2013).

I.1.3.2 Pyrazines and volatile thiols

Besides terpenes, methoxypyrazines are grape-derived compounds associated to the vegetal, green, and herbaceous aromas that characterize the Bordeaux cultivars, such as Sauvignon Blanc, Cabernet Sauvignon, Cabernet Franc, Merlot, or Carmenère. Other descriptors include bell pepper, vegetal and earthy (**Figure I. 12**). Methoxypyrazines are powerful odorants with very low perception thresholds (1-2 ng/L), with isobutyl-methoxypyrazine being the most relevant (5-30 ng/L) (Ruiz et al., 2019). Isopropyl-methoxypyrazine and sec-butyl-methoxypyrazyne can also be found, however at lower levels (Padilla et al., 2016; Allen & Lacey, 1998).

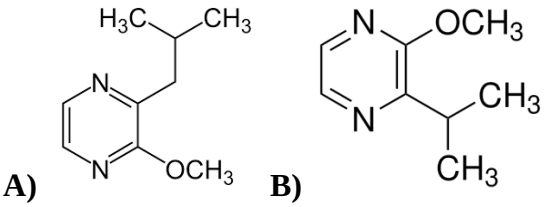


Figure I. 12–A) 3-isobutyl-2-methoxypyrazine (IBMP) and B) 2-isopropyl-3-methoxypyrazine (IPMP) are the most common methoxypyrazines found in wine. These compounds resemble the vegetal aromas (green bell pepper and grass, respectively) that are characteristic to some grape cultivars (*e.g.,* Cabernet Sauvignon).

Volatile thiols are organic sulfur derivates particularly relevant in some vine cultivars, such as Sauvignon Blanc (Padilla et al., 2016), but that have a very low sensorial threshold. Examples of volatile thiols are 4-mercapto-4-methylpentan-2-one, 3-mercaptohexan1-ol, and 3-mercaptohexyl acetate, that largely contribute to descriptors of passionfruit, grapefruit, gooseberry, blackcurrant, lychee, guava and box hedge (Swiegers & Pretorius, 2007; Swiegers & Pretorius, 2005). Volatile thiols are mainly found in their non-volatile form, by being conjugated to cysteine or glutathione moieties (**Figure I. 13B**). Enzymes, like carbon-sulfur-lyases release volatile thiols from their cysteinylated precursors (Cys-4MMP and Cys-3MH), having, like beta-glucosidases an oenological interest in enhancing wine varietal characteristics (Padilla et al., 2016; Tufariello et al., 2021).

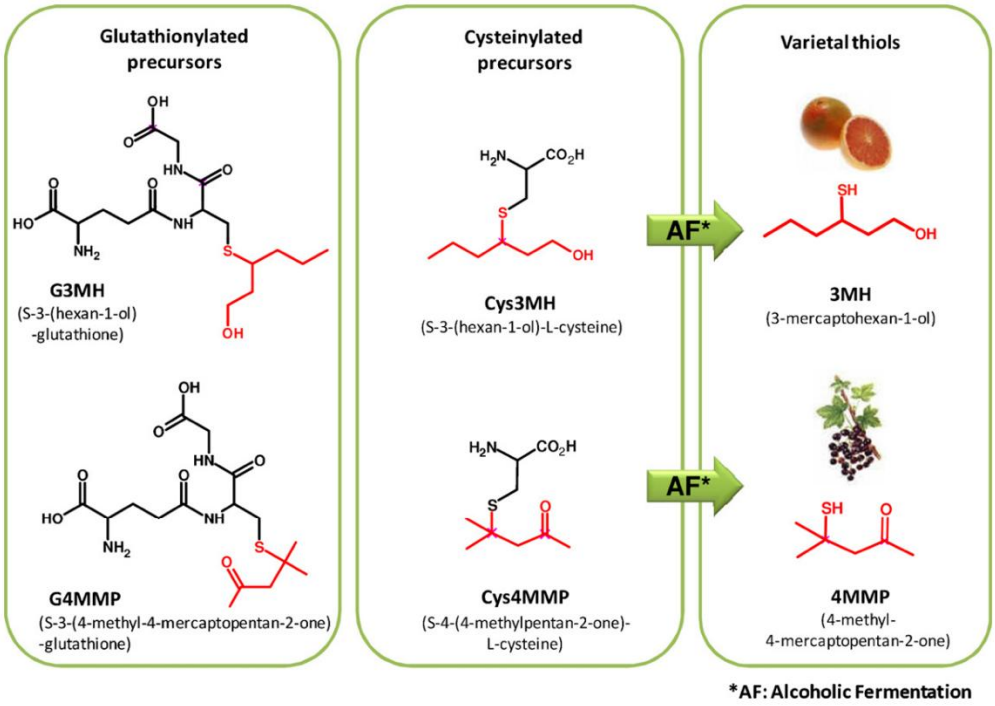


Figure I. 13- Volatile thiols are usually conjugated with glutathione or cysteine molecules, being non-volatile. Once released, their aromas resemble passionfruit and grapefruit, being very common in grape cultivars such as Sauvignon Blanc. Adapted from Roland et al., (2010).

I.1.4 Microbiological effect on the final concentration of terpenes in wine

The concentration of free terpenes in wines, although intimately related to the grape variety, can also be modulated by microbial activity, especially by species encoding beta-glucosidases, capable of releasing bound terpenes into their volatile form. Microbial beta-glucosidases encoded by yeasts belonging to the *Saccharomyces* or non-*Saccharomyces* genera (*e.g.*, *Debaryomyces hansenii* - Rosi et al., 1994 -, *D. pseudopolymorphus* - Rosi et al., 1994 - or *H. uvarum* - Gaensly et al., 2015; Albertin et al., 2016) have gained attention for wine application due to their high activity under winemaking conditions. These enzymes show good activity under winemaking conditions including at low pHs, low temperatures or in the presence of high concentrations of sugar or ethanol (Manzanares et al., 2000; Zhang et al., 2020).

The disclosure of genomic sequences, together with more exhaustive phenotyping of wine NSYs species that are part of the must microbiome, has been revealing that several of these yeasts exhibit high/very high levels of beta-glucosidase activity such as the species belonging to *Candida*, *Pichia*, *Hanseniaspora* or *Metschnikowia* (especially *M. pulcherrima* although significant inter-strain variation has been observed-Morata et al., 2019; Huang et al., 2021-). *Starmerella stellata* (former *Candida stellata)* and *Trigonopsis vinaria* (former *Candida vinaria*) are other good examples of species described to secrete beta-glucosidases

with high activity (Huang et al., 2021; García et al., 2018). Among the species from the *Hanseniaspora* genus, *H. valbyensis*, *H. uvarum*, *H. guilliermondii*, *H. osmophila*, *H. occidentalis* and *H. vineae* were all reported to secrete beta-glucosidases, however, significant inter-strain variation has been observed (Huang et al., 2021; Fresni et al., 2021; López et al., 2015; Valera et al., 2021).

I.1.5 Secondary Aroma

Wine secondary aroma appears during the alcoholic fermentation of grape must and is originated from the activity of the wine microbiota, endogenous or added (Carpena et al., 2021). The aromas formed during this step can be classified as pre-fermentative (Ruiz et al., 2019; Carpena et al., 2021; Perestrelo et al., 2020), being related to the mechanical crushing of grapes (*e.g.*, C_6 alcohols and C_6 aldehydes, such as hexanol and hexanal, respectively (Perestrelo et al., 2020; Verzera et al., 2021); or as fermentative, originating during the alcoholic or malolactic fermentation process (*e.g.*, higher alcohols and esters) (Ruiz et al., 2019; Carpena et al., 2021; Perestrelo et al., 2020; Verzera et al., 2021). The concentration of metabolites produced during the fermentation process highly influences the wine secondary aroma, which is mainly composed by higher alcohols, volatile fatty acids, esters or aldehydes (Styger et al., 2011; Ruiz et al., 2019; Carpena et al., 2021). Due to the great diversity of the wine microbiota, the secondary aroma is the product of a countless number of interactions between microorganisms and their enzymatic and metabolic behavior (Fleet, 2003). In fact, several studies have already shown that the chemical composition of wine is greatly impacted by the interactions that are established between microbial species, especially between yeasts (Jolly et al., 2014; Roullier-Gall et al., 2020). For example, Barbosa and colleagues (2015) have demonstrated that yeast-yeast interactions contribute to aroma modulation, not only because of their individual contribution, but mostly due to changes in the genomic expression patterns of yeasts during the coexistence in wine fermentation. Hence, species and even, strain selection, is of high importance to obtain a wine with stylistic and distinguishable properties. The following sections detail the molecules that are more prominently linked to secondary wine aroma.

I.2.3.1 Higher alcohols

Higher alcohols, also known as fusel alcohols, are pivotal compounds in shaping the secondary aroma of wines (Carpena et al., 2021). While higher alcohols usually enhance wine

aroma, their impact is contingent upon concentration. Beyond an average threshold of 300 mg/L, higher alcohols can impart unpleasant aromas (Carpena et al., 2021). For instance, isoamyl alcohol and isobutanol provide desirable pepper notes in aged red wines at optimal concentrations but may yield negative descriptors like burnt toast and old wood (Carpena et al., 2021; De-La-Fuente-Blanco et al., 2016). Similarly, 2-phenylethanol contributes pleasant floral notes to wine aroma, but becomes pungent above 300 mg/L (Padilla et al., 2016; Carpena et al., 2021).

In yeasts, higher alcohol production occurs during fermentation from the decarboxylation of alpha-keto acids, a reaction included in the Ehrlich pathway (Hazelwood et al., 2008). This three-step pathway begins with the transamination of an aromatic amino acid to form an alpha-keto acid, which is then decarboxylated by pyruvate decarboxylase into an aldehyde. The aldehyde is finally reduced to the corresponding higher alcohol form (**Figure I. 14A**) (Hazelwood et al., 2008). Usually, the amino acid precursor determines which higher alcohol is formed: valine, leucine and isoleucine yield branched-chain aliphatic alcohols, namely isobutanol, amyl alcohol, and isoamyl alcohol, respectively. Conversely, phenylalanine, tyrosine, and tryptophan lead to the production of aromatic alcohols 2-phenylethanol, tyrosol, and tryptophol (Hazelwood et al.,2008, Holt et al., 2019). Additionally, other higher alcohols, like 2-methylbutanol-1, 3-or methyl-1-butanol-1, derived from isoleucine and leucine, respectively, are also formed during wine fermentation (Ravasio et al., 2014). However, although higher alcohol formation through the Ehrlich pathway is influenced by amino acids availability, it has been observed that there is often no direct link between higher alcohol production and its amino acid precursor present in the grape must. This is due to the intracellular alpha-keto acids pool, which can originate from sources other than amino acid metabolism (Crépin et al., 2017). In fact, in wine *S. cerevisiae* strains, it has been demonstrated that most alpha-keto acids (>90%) used to produce higher alcohols are derived from the Central Carbon Metabolism (CCM), through sugar catabolism, with a limited contribution from the amino acid metabolism (Crépin et al., 2017; Rollero et al., 2017). This could be explained as the Ehrlich pathway seems to function as a detoxification route for alpha-keto acids produced during sugar metabolism, as well as a regulatory mechanism for amino acid metabolism and catabolism (Romano et al., 2022).

The microbial metabolism of higher alcohols in wine plays a crucial role in aroma modulation. Studies have shown that the presence of NSYs in wines, either alone or alongside *S. cerevisiae*, can result in lower titers of higher alcohols (Jeromel et al., 2019). However,

certain yeast species, like *H. uvarum*, *C. zemplinina* and *P. anomala* are known for their higher alcohol production, while others like *H. guilliermondii*, *H. osmophila* and *P. membranifaciens* produce lower amounts (Carpena et al., 2021). While the conventional understanding of higher alcohol biosynthesis in yeasts revolves around the Ehrlich pathway, recent research by Seixas and colleagues (2023) has shed light into what appears to be a distinct pathway for higher alcohol production in *H. guilliermondii*. Unlike *S. cerevisiae*, which mostly utilizes keto-acids from the Krebs cycle and sugar catabolism as precursors (Crépin et al., 2017; Rollero et al., 2017; Romano et al., 2022), evidences point to the idea that *H. guilliermondii* uses amino acids assimilated from the growth medium (**Figure I. 14**). In particular, it was found that *S. cerevisiae* synthesizes higher alcohols at high carbon-nitrogen (C:N) ratios, while *H. guilliermondii* produces these compounds at low C:N ratios, presumably to efficiently utilize the carbon derived from assimilated amino acids (**Figure I. 14B**) (Seixas et al., 2023). Findings consistent with this model were also obtained for other *H. guilliermondii* strains (Viana et al., 2008) and for *Metschnikowia pulcherrima* (González et al., 2018) indicating that the observations are likely not specific to the strain utilized. This study marked an important advance on the comprehension of aroma modulation in NSYs, particularly concerning the metabolic preferences that lead to the production of higher alcohols and acetate esters, pointing that probably not all aspects of the physiology of NSYs cannot be extrapolated from the knowledge known in *S. cerevisiae*.

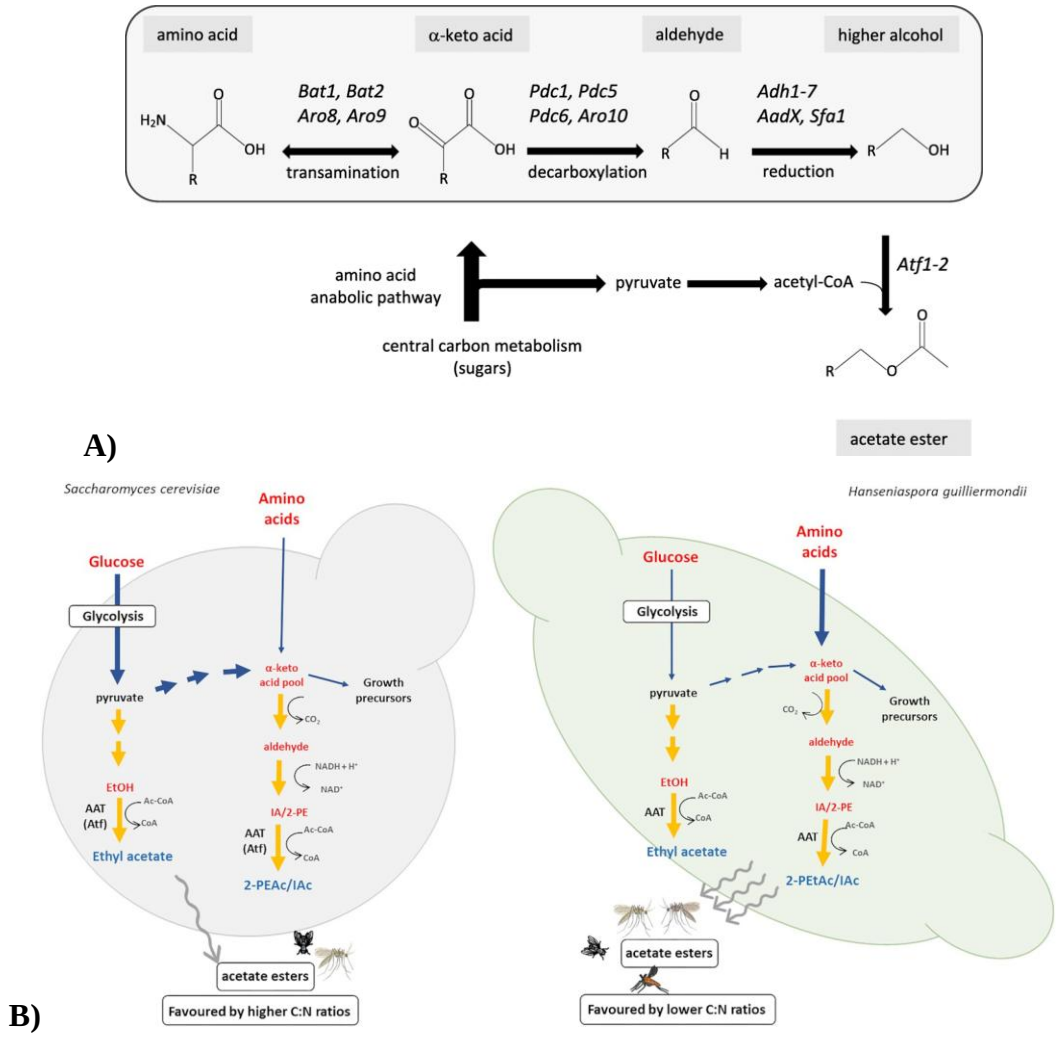


Figure I. 14- Higher alcohol and acetate ester production in yeasts**. A)** The Ehrlich pathway in *S. cerevisiae* for higher alcohol biosynthesis. The Ehrlich pathway is a three-step reaction that begins with a transamination of an aromatic amino acid into *α* -keto acids, which are then decarboxylated by the enzyme pyruvate decarboxylase (PDC) into aldehydes, and finally reduced to their higher alcohol form. Once produced, higher alcohols can be then esterified, by a condensation reaction with acetyl-CoA mediated by the enzyme acetyltransferase (ATF1-2), into their corresponding acetate esters. Adapted from Cordente et al., (2019). **B)** Schematic representation of higher alcohol/ acetate ester production in *S. cerevisiae* and in *H. guilliermondii*. In *S. cerevisiae* production of higher alcohols/ acetate esters is driven by keto-acids produced by the Krebs cycle, while in *H. guilliermondii* the precursors are the amino acids assimilated from the growth medium. IA: isoamylalcohol; 2-PE: 2-phenylethanol; 2-PEtAc: 2-phenylethyl acetate; IAc—isoamyl acetate. Adapted from Seixas et al., (2023).

I.2.3.2. Volatile fatty acids

Volatile fatty acids (VFA) are, in the wine context, mostly produced throughout the fermentation process and originate the corresponding ethyl esters, being categorized according to their chemical carbon structure (Carpena et al., 2021). These compounds, also known as low molecular weight organic acids, include a group of aliphatic monocarboxylic acids with two to six carbon atoms (C2 to C6): acetic (C2/HAc), propionic (C3/HPr), iso-butyric(iC4/iHBu), n-butyric (C4/HBu), iso-valeric (iC5/iHVa), n-valeric (C5/HVa), iso-caproic (iC6/HCa), and n-

caproic (iC6/HCa) acids (Agnihotri et al., 2022). In wines acetic acid is the most abundant volatile acid, being responsible for about 90% of the volatile fatty acid content, followed by propanoic acid and butanoic acid (Lambrechts & Pretorius, 2000). Usually, when in small amounts, VFAs contribute to the overall aroma of wine by adding a smokey, spicy or medicinal smell (Styger et al., 2011) and even to the perception of fresh fruit (Chambers & Pretorius, 2010, De-La-Fuente-Blanco et al., 2016). However, when in excess, VFAs usually result in defected wines by producing a sweaty, leathery, or barnyard aroma descriptors (Carpena et al., 2021; Agnihotri et al., 2022; Jackson, 2008).

The production of VFAs by NSYs is well documented. Species belonging to the *Hanseniaspora* and *Zygosaccharomyces* genera and also *Schizosaccharomyces pombe* have been described as producers of VFAs, especially acetic acid (Lambrechts & Pretorius, 2000). Additionally, wines produced with *H. vineae, H. uvarum, H. guilliermondii, C. stellata* or *C. zemplinina* were also reported to display higher concentrations of isobutyric acid (Carpena et al., 2021). In the specific case of *Hanseniaspora* species or genus, there have been reports indicating a positive correlation between the generation of acetic acid and acetate esters (Seixas et al., 2023; Rojas et al., 2001). This suggests that these species might possess a metabolic network that promotes a greater internal reservoir of acetate, thereby facilitating an increased supply of acetyl-CoA for the synthesis of acetate esters, in comparison to other yeast species (Seixas et al., 2023). Still, there is supported evidence of a great inter-species and inter-strain variability concerning this phenotype (Barbosa et al., 2015; Seixas et al., 2023; Lage et al., 2014), which likely results from not having been identified the genetic traits connected to the production of these molecules.

I.2.3.3. Esters

Esters constitute the predominant class of aroma compounds found in wine, with over 160 compounds identified (Padilla et al., 2016). Distinguished for enhancing wine aroma complexity with fruity descriptors, esters exert a highly positive influence, typically up to a threshold of 100 mg/L (Padilla et al., 2016; Lambrechts & Pretorius, 2000; Carpena et al., 2021). These compounds can be categorized into two main types: fatty acid ethyl esters, synthesized through the esterification of activated fatty acids (**Figure I. 15**) (Carpena et al., 2021); and acetate esters, derived from higher alcohols during the Ehrlich pathway, catalyzed by alcohol acetyltransferases (Carpena et al., 2021).

In wines, ethyl hexanoate, ethyl octanoate, and ethyl decanoate are the most abundant fatty acid ethyl esters, imparting fruity, sweet, and also floral notes (Carpena et al., 2021; Marullo et al., 2021). Meanwhile, common acetate esters include isobutyl acetate, amyl acetate, hexyl acetate, ethyl acetate, isoamyl acetate, and 2-phenylethyl acetate (2PA), contributing fruity and banana aromas (Carpena et al., 2021). Additionally, white wines often contain fatty acid ethyl esters, such as ethyl butanoate, caproate, caprylate, caprate and laureate (Carpena et al., 2021).

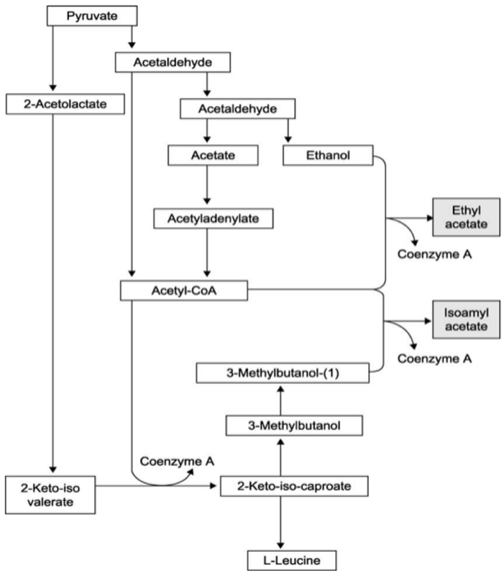


Figure I. 15- Esters in wine. A schematic representation of the biosynthetic pathway of ethyl acetate and isoamyl acetate in yeast. Adapted from Swiegers et al., (2007).

The ester profile of wines is greatly influenced by yeast activity during fermentation, underscoring the pivotal role of yeast strain selection in shaping the final product. While several *S. cerevisiae* strains are proficient in ester production, *Hanseniaspora, Candida* and *Pichia* have also been studied for their unique enzymatic mechanisms in this regard (Carpena et al.,2021; Nykanen & Nykanen, 1977; Martin et al., 2018; Hu et al., 2019). In *S. cerevisiae,* acetate ester biosynthesis is mediated by alcohol-O-acetyltransferases, (AATs), which conjugate an acetyl-CoA molecule to the corresponding higher alcohol (Seixas et al., 2023; Verstrepen et al., 2003). In *S. cerevisiae* three genes encoding AATs directly connected to acetate ester biosynthesis were identified: ScAtf1 and ScAtf2, which catalyze the majority of these reactions; and ScEat1, coding for an alternative enzyme, mostly involved in ethyl acetate biosynthesis (Seixas et al., 2023). The recent genomic analysis of *H. guilliermondii* strain UTAD222 (Seixas et al., 2019) revealed the presence of four genes suggested to be AATs.

Homologues to these genes were also found in *H. uvarum* and in *H. opuntiae* (Seixas et al., 2019, 2023). Notably, *Hanseniaspora* AATs bear a low similarity to those described in *S. cerevisiae*, although they still retain motifs that are essential for the activity of these enzymes, such as HXXXDG and WRLICLP, hypothesized as part of the active site of AATases (Seixas et al., 2019). These observations suggest the existence of a distinct family of alcohol acetyltransferase enzymes, responsible for acetate ester synthesis, whose properties, if different from the ones found in *S. cerevisiae* enzymes, can potentially explain the differences found in differences in acetate ester production observed among these species (Seixas et al.,2019). While some similarities between *S. cerevisiae* genes and those of some Non-*Saccharomyces* have been found, especially in *Lachancea* species, there is still a great discrepancy with other Non-*Saccharomyces* Yeasts from the wine environment (**Figure I. 16**).

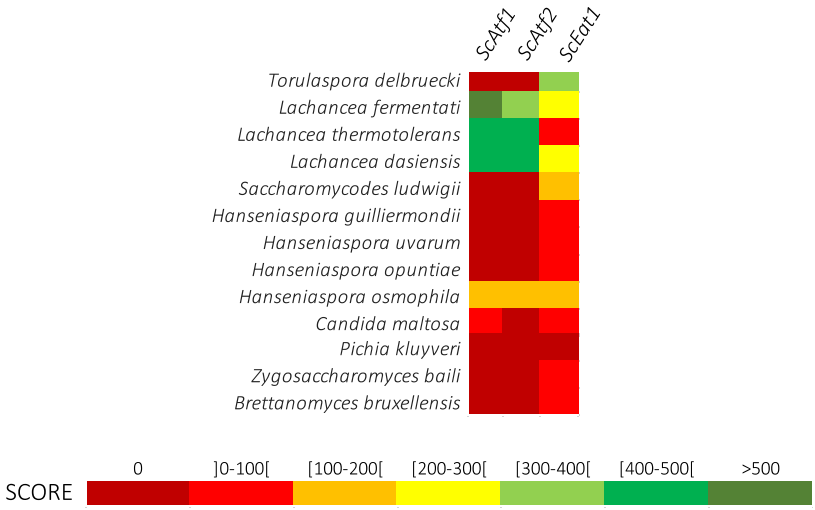


Figure I. 16– Heatmap based on NCBI's protein BLAST scoring. The scoring system used, BLOSUM62, is based on frequencies of amino acid substitutions in related proteins, which reflects the amino acid chemistry and protein structure. The heatmap herein presented ranges from deep red, which indicates the value 0, corresponding to no hit found, to deep green, with a score higher than 500.

I.2.3.4. Aldehydes

Aldehydes form a significant class of compounds contributing to the secondary aroma profile of wine (Padilla et al., 2016). Among them, acetaldehyde stands out as the most common, constituting over 90% of the total aldehyde content in wine (Lambrechts & Pretorius, 2000). With a low sensorial threshold (Romano et al., 1994), acetaldehyde imparts pleasant fruity aromas, reminiscent of apples, even at low concentrations. However, high levels can lead to undesirable characteristics resembling rotten apple (Styger et al., 2011; Liu & Pilone, 2000). Fortified wines exhibit varying levels of acetaldehyde, ranging from 175 mg/L in sherry wines to 300 mg/L in port wines; whereas concentrations in red or white wines typically range

between 30 mg/L and 80 mg/L, respectively (Romano et al., 1994; Liu & Pilone, 2000). Moreover, acetaldehyde serves as a precursor to other aroma-relevant compounds. For instance, diacetyl, acetoin and 2,3-butanediol derive from acetaldehyde. Diacetyl contributes nutty and toasty notes, acetoin imparts a buttery taste, while 2,3-butanediol does not have a strong odor (Styger et al., 2011; Cheynier et al., 2010). To produce diacetyl, acetaldehyde undergoes a condensation reaction. Then, diacetyl is rapidly reduced to form acetoin, which in turn can be reduced through the enzyme 2,3-butanediol dehydrogenase to produce 2,3-butanediol (Bae et al., 2016).

Acetaldehyde emerges from yeast activity during fermentation, being an intermediate compound in yeast metabolism, through the conversion of pyruvate via enzymes in the glycolytic pathway (Romano et al., 1994; García-Rios & Guillamón, 2019). Other aldehydes such as isobutyraldehyde (Atsumi et al., 2010) and phenylacetaldehyde (Vilela, 2020) are formed in the Ehrlich pathway and may accumulate in the fermenting must (Hazelwood et al., 2008). Besides *S. cerevisiae*, other NSYs, such as *Schizosaccharomyces pombe* or *Zygosaccharomyces bailii* have been described to produce acetaldehyde at considerable amounts, having a fundamental role in the aroma of fortified wines like port or sherry wines (Liu & Pilone, 2000). Acetic acid bacteria typically found in grape musts, such as those belonging to *Acetobacter, Gluconacetobacter* and *Komagateibacter* genera, are capable of producing acetaldehyde through the oxidation of ethanol (Cheynier et al., 2000; Liu & Pilone, 2000; Gomes et al., 2018), thus contributing to modulate its concentration in wines.

Besides microbial activity, the overall concentration of acetaldehyde can be influenced by other factors such as high temperatures or the presence of hydrogen peroxide (a byproduct of phenolic oxidation) that were both reported to elevate acetaldehyde levels (Liu & Pilone, 2000). Additionally, the addition of sulfur dioxide (SO_2) during winemaking is also known to increase the production of acetaldehyde in *S. cerevisiae* and in *Z. bailii* (Kuanyshev et al., 2017). At the wine pH (3-4), bisulfite (HSO_3^-) is the prevalent form obtained upon dissociation of sulfur dioxide. Acetaldehyde promptly binds to HSO3- and therefore the production of this carbohydrate is believed to represent an important mechanism by which yeasts try to counteract the deleterious effect of SO_2 (García-Rios & Guillamón, 2019; Casalone et al., 1992).. By adding SO_2 to wine, oenologists can, to a certain extent, neutralize the perceived aromatic character of acetaldehyde, rather than inhibiting acetaldehyde production (Liu & Pilone, 2000). The high reactivity of acetaldehyde also enables the formation of a wide range of flavor compounds (Styger et al., 2011). For instance, beneficial reactions between acetaldehyde and

flavonoids (through the formation of ethyl-bridges condensation products - Liu & Pilone, 2000) were reported to improve bouquets of red wines by increasing, not only its astringency, but also its color stability (Liu & Pilone, 2000).

I.2.4. Tertiary aroma

Throughout the aging process, the aroma profile of young wine, characterized by varietal nuances of the grape and the diverse attributes resulting from yeast fermentation, undergoes significant development, evolving into a more refined and complex bouquet (Styger et al., 2011; Slaghenaufi & Ugliano, 2018; Pereira et al., 2021). This transformation occurs as the wine interacts with French oak barrels, absorbing wood-derived volatiles, or through micro-oxygenation facilitated by cork stoppers, which enhances oxidation reactions. Consequently, the aroma developed during aging becomes one of the most defining features of the wine (Styger et al., 2011; Pereira et al., 2021). Typically, during aging, whether in barrels or in bottles, the fruity ester aromas, characteristic of young wines, are gradually replaced by warmer notes, such as wood, cocoa or chocolate, imparted by the oak (Pereira et al., 2021). Moreover, as amino acids and alcohols undergo oxidation, there is a shift in the balance of volatile compounds present in the wine, resulting in an increase in aldehydes (*e.g.*, acetaldehyde) and other oxidized compounds, such as ketones and organic acids (Styger et al., 2011; Slaghenaufi & Ugliano, 2018; G. Pereira et al., 2021). These compounds contribute to the formation or degradation of sulfur-containing volatile compounds. During aging in oak barrels, compounds such as cis- and trans-oak lactone are produced, as well as some volatile phenols such as eugenol, 4-methylguaiacol, and vanillin, introducing aromatic elements of wood, coconut (eugenol and 4-methylguaiacol), dark chocolate (4-methylguaiacol), and vanilla (vanillin) (Spillman et al., 2004).

I.2.4.1 Volatile phenols

Volatile phenols play a pivotal role in shaping the tertiary aroma of wine, emerging as the most influential volatile compounds in this developmental phase (Padilla et al., 2016; Swieger & Pretorius, 2005). Depending on their concentration, these compounds exhibit a wide range of aromatic description, spanning from the pleasant aromas of vanillin, methyl vanilla or homovanillyl alcohol, to undesirable off-flavors such as barnyard/stable (attributed to ethylphenols) or pharmaceutical-like scents, mostly in white wines (associated with vinylphenols - Padilla et al., 2016; Swiegers & Pretorius, 2005; Carpena et al., 2021). Among

the ethylphenols commonly encountered in wine, 4-ethylguaiacol and 4-ethylphenol, stand out, while notable exemples of vinylphenols include 4-vinylguaiacol or 4-vinylphenol (Swiegers & Pretorius, 2005).

The biosynthesis of vinylphenols, such as 4-vinylphenol and 4-vinylguaiacol, can occur during the alcoholic fermentation step or during aging in oak barrels (Padilla et al., 2016; Swiegers & Pretorius, 2005). Stereospecific decarboxylases present in *S. cerevisiae*, as well as in some NSYs (*Brettanomyces*/ *Dekkera* spp. are particularly known for this trait) facilitate the conversion of hydroxycinnamic acids (like p-coumaric and ferulic acid) into 4-vinylphenol and 4-vinylguaiacol, respectively (Nikfardjam et al., 2009). Moreover, lactic acid bacteria, including *Pediococcus* spp., also contribute to the synthesis of volatile phenols present in wine being capable of decarboxylating cinnamic acids into ethylphenols or converting p-coumaric and ferulic acids into 4-vinyl guaiacol and 4-vinylphenol (Swiegers & Pretorius, 2005).

I.2 The *Saccharomycodaceae* family as a case study of wine Non-*Saccharomyces* Yeasts

I.2.1 The *Saccharomycodaceae* family – A Taxonomic and Phylogenetic introduction

The *Saccharomycodaceae* family (Order: *Saccharomycocetales*, Class: *Saccharomycetes*, Division: *Ascomycota*, Kingdom: *Fungi*) was first described in 1954 by the Russian taxonomist Wladimir Kudrejawzew (Lodder et al., 1962) and is comprised by four genera of apiculate yeasts: *Hanseniaspora*, *Saccharomycodes*, *Nadsonia*, and *Wickerhamia* (van Wyk et al., 2023). Placed phylogenetically close to the family *Saccharomycetaceaee* (that, among other species, includes *Saccharomyces cerevisiae*), *Saccharomycodaceae* yeasts share the singular characteristic of cellular division through bipolar budding (Riley et al., 2016) (**Figure I. 17**).

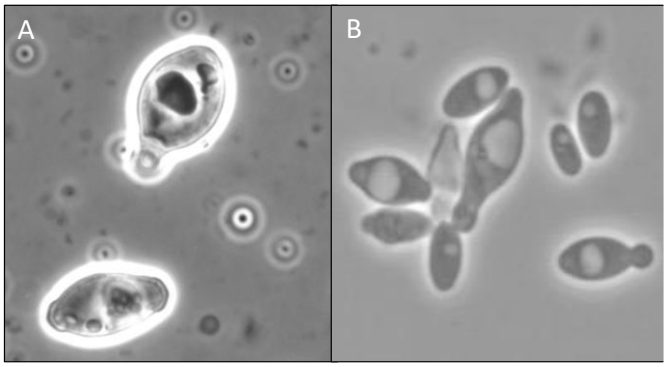


Figure I. 17– Optical Microscopy image of two *Saccharomycodaceae* species: A) *Saccharomycodes ludwigii* UTAD17 and B) *Hanseniaspora guilliermondii* UTAD222. Pictures taken with 400x resolution.

A preliminary study published in 1931 by Stelling-Decker positioned these yeasts in the tribe *Nadsonieae* of the subfamily *Saccharomycetoideae*. In this study, the authors emphasized the lemon-shaped vegetative cells which reproduced by bipolar budding (Spencer & Gorin, 1968; Stelling-Decker, 1931) (**Figure I. 18**).

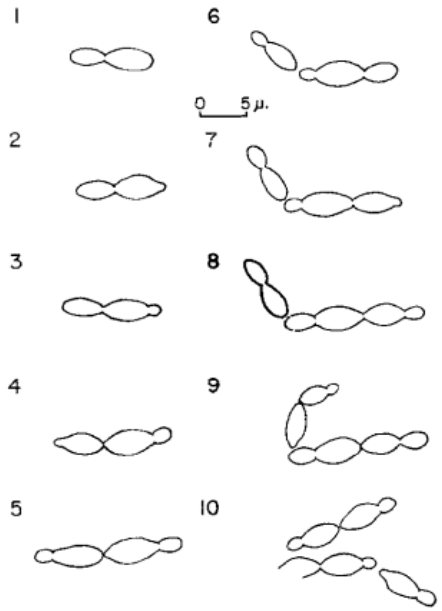
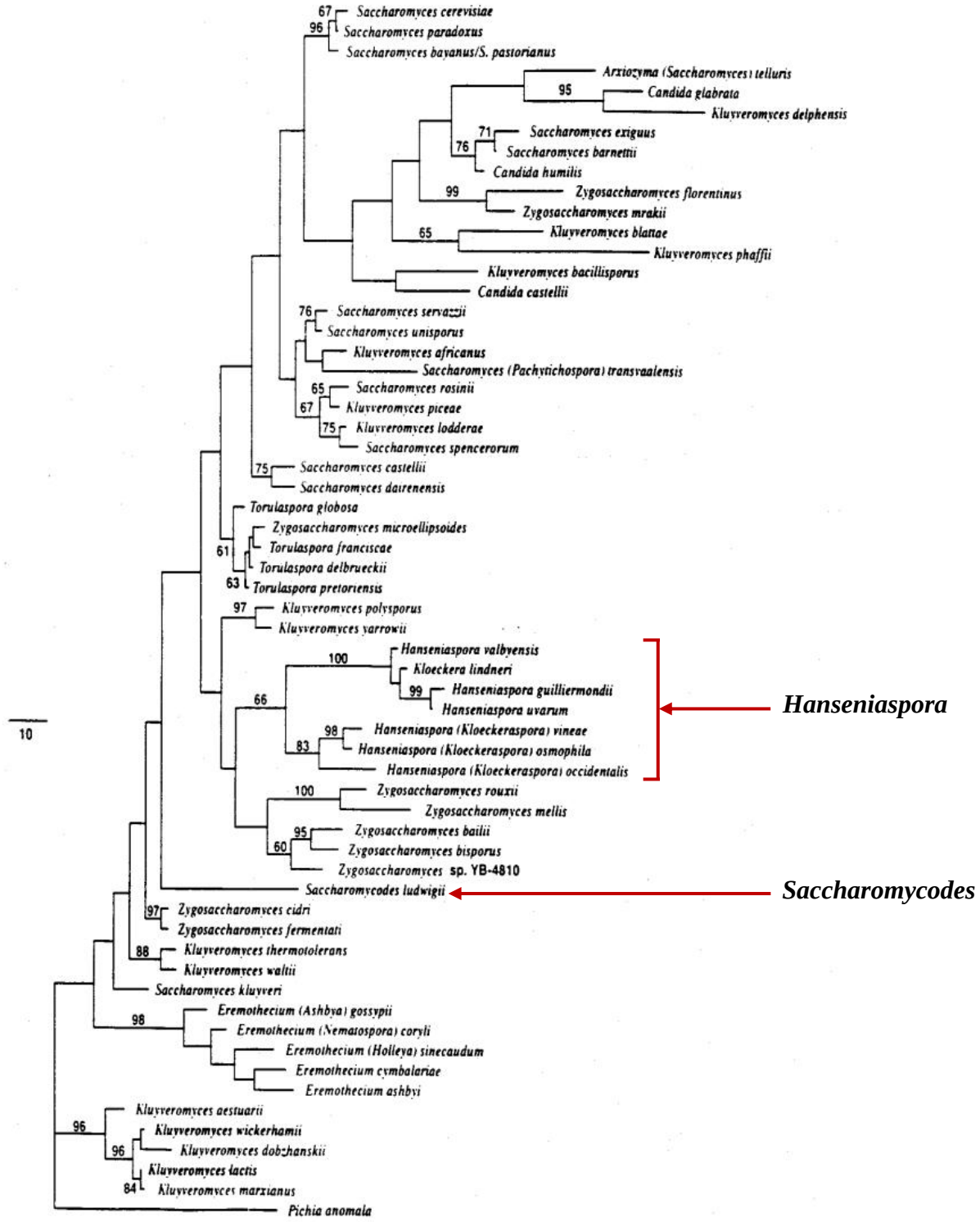


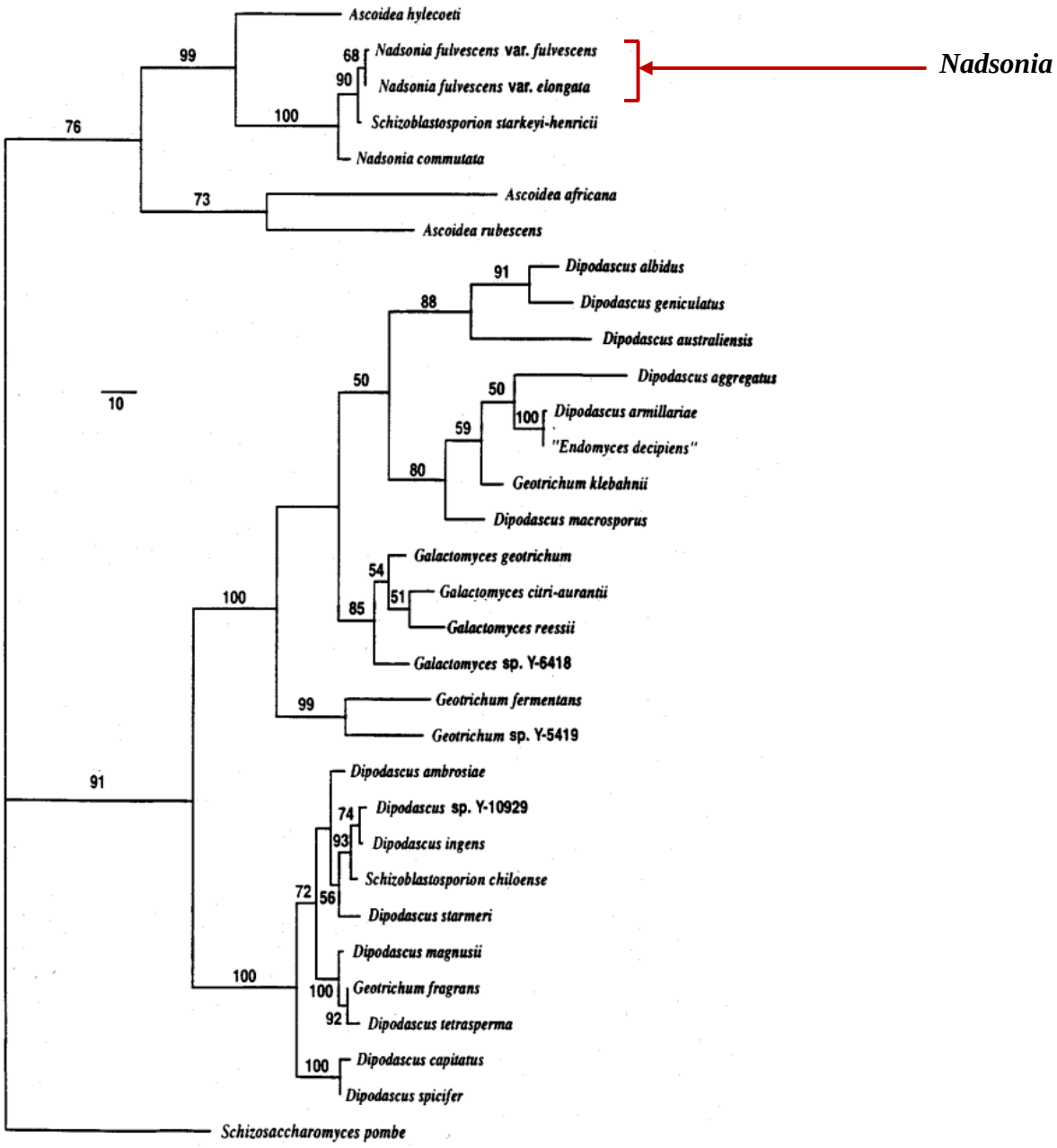
Figure I. 18– Drawings representing the vegetative reproduction process of *Hanseniaspora*. These drawings represent the division stages observed at 0, 30, 60, 85, 105, 125, 135, 145, 165 and 180 minutes from the bud stage to a separate cell in several *Hanseniaspora* species (*H. valbyensis, H. uvarum, K. apiculata*). Adapted from Miller & Phaff (1958).

However, while the genera *Hanseniaspora* and *Saccharomycodes* exhibit some degree of relatedness, *Nadsonia* and *Wickerhamia* are placed distant from one another as well as from the other genera of the *Saccharomycodaceae* family (Kurtzman, 2011; Kurtzman & Robnet, 1998;2003). A study conducted by Kurtzman and Robnett (1998) analysed the D1/D2 region of the large-subunit (LSU) of the ribosomal DNA to conclude that while *Nadsonia, Wickerhamia, Hanseniaspora,*and *Saccharomycodes* share the unique morphological property of bipolar budding, this trait is a weak predictor of phylogenetic relatedness (Kurtzman & Robnet, 1998). In particular, it was found that based on DNA sequence analyses, *Hanseniaspora* and *Saccharomycodes* are included in the *Saccharomyces* clade, *Nadsonia* is clustered in *Ascoidea,* and *Wickerhamia* included in *Debaryomyces* (**Figure I. 19**)(Kurtzman & Robnet, 1998).

A)


Saccharomyces cerevisiae
Saccharomyces paradoxus
Saccharomyces bayanus/S. pastorianus
Arxiozyma (Saccharomyces) telluris
Candida glabrata
Kluyveromyces delphensis
Saccharomyces exiguus
Saccharomyces barnettii
Candida humilis
Zygosaccharomyces florentinus
Zygosaccharomyces mrakii
Kluyveromyces blattae
Kluyveromyces phaffii
Kluyveromyces bacillisporus
Candida castellii
Saccharomyces servazzii
Saccharomyces unisporus
Kluyveromyces africanus
Saccharomyces (Pachytichospora) transvaalensis
Saccharomyces rosinii
Kluyveromyces piceae
Kluyveromyces lodderae
Saccharomyces spencerorum
Saccharomyces castellii
Saccharomyces dairenensis
Torulaspora globosa
Zygosaccharomyces microellipsoides
Torulaspora franciscae
Torulaspora delbrueckii
Torulaspora pretoriensis
Kluyveromyces polysporus
Kluyveromyces yarrowii
Hanseniaspora valbyensis
Kloeckera lindneri
Hanseniaspora guilliermondii
Hanseniaspora uvarum
Hanseniaspora (Kloeckeraspora) vineae
Hanseniaspora (Kloeckeraspora) osmophila
Hanseniaspora (Kloeckeraspora) occidentalis
Zygosaccharomyces rouxii
Zygosaccharomyces mellis
Zygosaccharomyces bailii
Zygosaccharomyces bisporus
Zygosaccharomyces sp. YB-4810
Saccharomycodes ludwigii
Zygosaccharomyces cidri
Zygosaccharomyces fermentati
Kluyveromyces thermotolerans
Kluyveromyces waltii
Saccharomyces kluyveri
Eremothecium (Ashbya) gossypii
Eremothecium (Nematospora) coryli
Eremothecium (Holleya) sinecaudum
Eremothecium cymbalariae
Eremothecium ashbyi
Kluyveromyces aestuarii
Kluyveromyces wickerhamii
Kluyveromyces dobzhanskii
Kluyveromyces lactis
Kluyveromyces marxianus
Pichia anomala
10
Hanseniaspora
Saccharomycodes

B)


Ascoidea hylecoeti
Nadsonia fulvescens var. fulvescens
Nadsonia fulvescens var. elongata
Schizoblastosporion starkeyi-henricii
Nadsonia commutata
Ascoidea africana
Ascoidea rubescens
Dipodascus albidus
Dipodascus geniculatus
Dipodascus australiensis
Dipodascus aggregatus
Dipodascus armillariae
"Endomyces decipiens"
Geotrichum klebahnii
Dipodascus macrosporus
Galactomyces geotrichum
Galactomyces citri-aurantii
Galactomyces reessii
Galactomyces sp. Y-6418
Geotrichum fermentans
Geotrichum sp. Y-5419
Dipodascus ambrosiae
Dipodascus sp. Y-10929
Dipodascus ingens
Schizoblastosporion chiloense
Dipodascus starmeri
Dipodascus magnusii
Geotrichum fragrans
Dipodascus tetrasperma
Dipodascus capitatus
Dipodascus spicifer
Schizosaccharomyces pombe
10
Nadsonia

C)

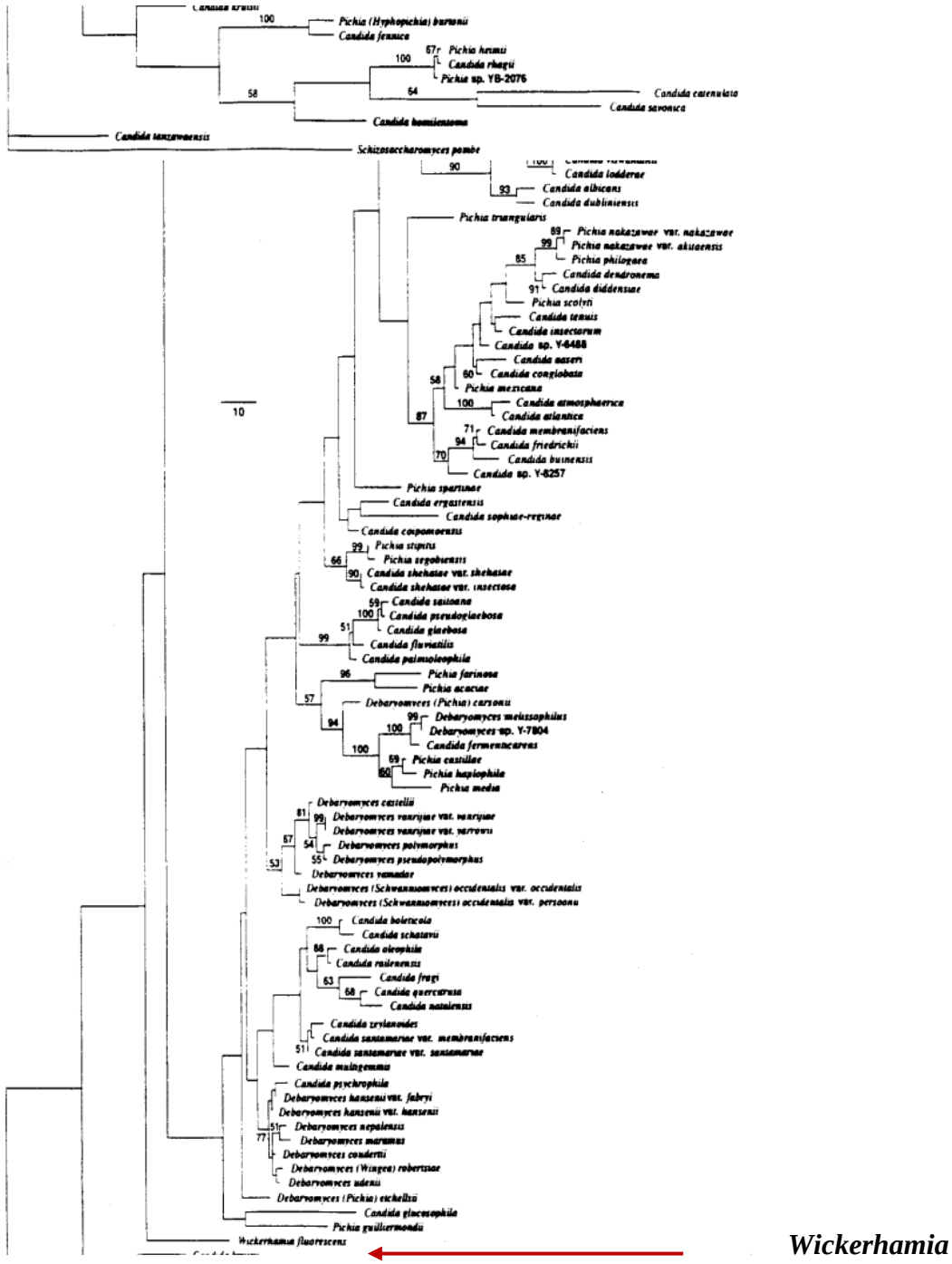


Figure I. 19 - Phylogenetic trees of *Saccharomycodaceae* family based on the D1/D2 domain of the large-subunit (LSU) ribosomal DNA. A) Phylogenetic tree of the *Saccharomyces* clade represented by 1 of 60 most parsimonious trees derived from maximum parsimony analysis of LSU domain D1/D2.; B) Phylogenetic tree of the *Ascoidea/Nadsonia/Dipodascus* clade represented by 1 of 2 most parsimonious trees derived from maximum parsimony analysis.; C) Phylogenetic tree of the *Debaryomyces/Lodderomyces* clade represented by 1 of 100 most parsimonious trees derived from maximum parsimony analysis. Adapted from Kurtzman and Robnett (1998).

Besides these genera, the *Saccharomycodaceae* family is composed by another genus that was defined as the imperfect form of *Hanseniaspora* (Miller & Phaff, 1958), the *Kloeckera* genus, that includes the asexual and anamorph form of *Hanseniaspora*. Within the *Kloeckera* genus, six species are recognized, as detailed in (**Table I. 2**). The introduction of molecular taxonomy enabled further associations between the species of *Hanseniaspora* and *Kloeckera*, first using both the GC content (Nakase & Komagata, 1970) and DNA–DNA relatedness (Meyer et el.,1978), and later, based on ITS sequences and protein-encoding genes (Kurtzman & Robnet, 2003; Yamada et al., 1992; Esteve-Zarzoso et al., 2001; Cadez et al., 2006).

Table I. 2- Relationships between species of the vegetative genus *Kloeckera* and *Hanseniaspora*, based on DNA composition and DNA-DNA relatedness, as proposed by Nakase & Komagata (1970) and Meyer et al (1978).

Kloeckera species	***Hanseniaspora species***
K. apiculata	*H. uvarum*
K.apis	*H. guilliermondii*
K. africana	*H. vineae*
K. corticis	*H. osmophila*
K. javanica	*H. occidentalis*
K. lindneri	*H. valbyensis*

Besides reproduction, one of the main phenotypic differences observed among *Saccharomycodaceae* yeasts concerned the mannan content of the cell polysaccharides. In specific, in 1968, Spencer and Gorin observed, using Nuclear Magnetic Resonance (NMR), that *Nadsonia elongata* and *N. fulvescens* both produced galactomannans, while the analysed species of the *Hanseniaspora* (*H. uvarum*, *H. valbyensis*, *H. osmophila*), *Kloeckera (K. africana*, *K. corticis* (previously known as *K. magna* (Cadez & Smith, 2011) and *K. javanica*) and *Saccharomycodes* (*Sd. ludwigii*) genera formed polysaccharides containing only mannose. Notably, *Sd. ludwigii* stood out in this study by having what appeared to be a more complex mixture of mannans (Spencer & Gorin, 1968) (**Figure I. 20**), which may be related to its described enhanced ability to excrete mannoproteins with a very high content of mannose (Domizio et al., 2011; Giovani et al., 2012). In this regard, the authors were also able to divide the *Hanseniaspora/Kloeckera* genera into three clusters based on their mannan content: one composed by *H. valbyensis* (anamorph *K. linderi*), *H. uvarum* anamorph *K. apiculata*), *K*; one composed by *H. osmophila* (anamorph *K. corticis*),and *H. vineae* (anamorph *K. africana)*; and

a third group containing a single species *H. occidentalis* (anamorph *K. javanica* - Spencer & Gorin, 1968).

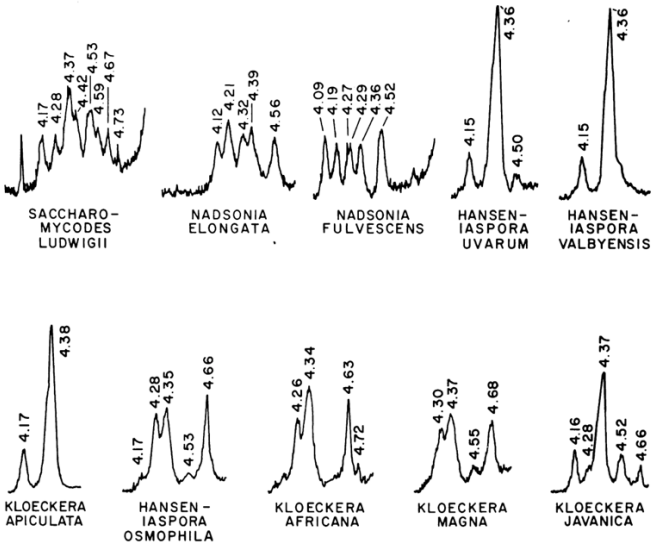


Figure I. 20- Nuclear Magnetic Resonance (NMR) spectra of the *Saccharomycodaceae* family. These spectra highlighted the yeast *Saccharomycodes ludwigii* as containing a more complex set of mannan polysaccharides, hinting for the presence of a mixture of mannans. Adapted from Spencer et al (1968).

The following sections provide further details about the biology and physiology of *Hanseniaspora* and *Saccharomycodes* species considering that within the *Saccharomycodaceae* family these are the species central to the development of this thesis.

I.2.1.1 Evolutionary and Genomic insights into the genus *Hanseniaspora*

Yeasts species within the *Hanseniaspora* genus are commonly found inhabiting various fruits, flowers, and barks (Saubin et al., 2020) being believed to be dispersed to these different environments (and to others, like vineyards) by insects (Seixas et al., 2019,2023, Kleman et al., 2022; Bianchi et al., 2020; Lam et al., 2015; Schubert et al., 2014; Giang et al., 2017). Among these insects, those from the *Drosophila* genus exhibit attraction to the volatiles produced by *Hanseniaspora* spp., including 2-phenylethyl acetate, isoamyl acetate, and ethyl acetate (Seixas et al., 2023; Kleman et al., 2022; Bianchi et al., 2020; Lam et al., 2015; Schubert et al., 2014; Giang et al., 2017). In 2024, there were 24 species of *Hanseniaspora* described, with the latest additions including *H. menglaensis* (2023) (Chen et al., 2023), *H. smithiae* (2021) (Cadez et al., 2021), *H. gamundiae* (2019) (Cadez et al., 2019), *H. terricola* (2019) (Liu et al.;2019) and *H. mollemarum* (2018) (Groenewald et al., 2018) (**Table I. 3**). However, it was not until 2014 that the first complete genome of a *Hanseniaspora* species was published,

specifically *H. vineae* (Giorello et al., 2014). Subsequently, others followed encompassing *H. guilliermondii* (Seixas et al., 2019), *H. opuntiae* (Sternes et al., 2016), *H. osmophila* (Sternes et al., 2016), *H. uvarum* (Sternes et al., 2016; Langenberg et al; 2017) and *H. valbyensis* (Riley et al., 2016). Using taxonomic markers, such as the D1/D2 domains of the 26S rRNA subunit of the Internal-Transcribed Spacer (ITS), it has been revealed that *Hanseniaspora* yeasts can be divided into two distinct lineages (Boekhout et al., 1994; Ouoba et al., 2015), the fast evolving lineage (or FEL) and a slow evolving lineage (or SEL). These findings were further supported by genomics-based analysis conducted by Steenwyck et al. (2019).

Table I. 3- *Hanseniaspora* species belonging to the Fast-Evolving lineage (FEL) and Slow-Evolving Lineage (SEL), and their corresponding isolation sources, and general genome information available (GC content, genome size, number of proteins)

	SPECIES	ISOLATION SOURCES	GC (%)	GENOME SIZE (Mbp)	GENBANK (EXAMPLE)	REF.
FEL	***H. clermontiae***	Stem rot, bark	37	8.7	GCA_030567235.1	Cadez et al., 2003
	H. guilliermondii	Grapes, wines, flowers, fruits	31	9	GCA_900119595.1	Seixas et al., 2019
	H. hatyaiensis	Rotten wood	36.5	9.6	GCA_030573555.1	Cadez et al., 2019
	H. jakobsenii	African palm wine	30	11	GCA_030573515.1	Ouoba et al., 2015
	H. lachancei	Fermenting agave juice	35	8.9	GCA_030567275.1	Cadez et al., 2003
	H. lindneri	Soil	36	10.8	GCA_019649525.1	Cadez et al., 2019
	H. menglaensis	Fruits, flowers	30.5	9.5	GCA_040256855.1	Chen et al., 2003; Ryan et al.,2024
	H. meyeri	Fruits	37	8.8	GCA_030370665.1	Cadez et al., 2003
	H. mollemarum	Garden soil	35	9.2	GCA_042466605.1	Groenewald et al., 2018; Steenwyk et al., 2019
	H. nectarophila	Nectar, flowers	34	10	GCA_030573495.1	Cadez et al., 2014; Badura et al., 2023
	H. opuntiae	Cactus, grape must	35	8.8	GCA_001749795.1	Cadez et al., 2003; Sternes et al., 2016
	H. pseudoguilliermondii	Orange juice concentrate	34.5	8.8	GCA_030573455.1	Cadez et al., 2006
	H. singularis	Flowers	26	8.9	GCA_030565715.1	Jindamorakot et al., 2009
	H. smithiae	stromata, bark, draught beer	n.a	n.a	n.a	Cadez et al., 2021
	H. taiwanica	soil	33	10	GCA_030580555.1	Cadez et al., 2019
	H. terricola	Flower, rotten wood	n.a	n.a	n.a	Liu et al., 2021

	H. thailandica	Insect frass	~34.9	10	GCA_030573475.1	Jindamorakot et al., 2009
	H. uvarum	Grapes, wines, flowers, fruits, soil, plants, insects, birds, mollusk and shrimps, honey by-products	32	9	GCA_037102615.1	Cadez et al., 2011; Sternes et al., 2016; Guaragnella et al.,2020
	H. valbyensis	Soil, Sap of tree, Beer, Tomato,Pulque Moss, balsamic vinegar	26.5	11.5	GCA_001664025.1	Riley et al., 2016
SEL	***H. gamundiae***	Fungal stromata from fermented beverage	37	10.2	GCA_003020785.1	Cadez et al., 2019
	H. osmophila	Grape, grape must	37	11.5	GCA_001747045.1	Granchi et al., 2002; Sternes et al., 2016
	H. occidentalis var citrica	Orange juice	35	11.3	GCA_030573535.1	Cadez et al., 2006; Steenwyk et al., 2019
	H. occidentalis var occidentalis	soil	34.5	11.5	GCA_030567355.1	Cadez et al., 2006; Esteves-Zarzoso et al.,2001
	H. vineae	Grapes, wines, fruits, soil	37.5	11.3	GCA_000585475.3	Esteves-Zarzoso et al., 2001; Giorello et al., 2014

The results of Steenwyk et al., (2019) suggested that the fast-evolving lineage (FEL) initiated its diversification approximately 87 million years ago and experienced a period of rapid evolution; while the slow-evolving lineage (SEL) began diversifying around 54 million years ago (Steenwyk et al., 2019) (**Figure I. 21**). While both lineages exhibited gene losses when compared to budding yeasts in the *Saccharomycotina* subphylum (*e.g., S. cerevisiae, Kluyveromyces marxianus, Cyberlindnera jadinii, Wickerhamomyces anomalus*), species within the FEL lineage displayed exceptionally high evolutionary rates, lower Guanine–Cytosine (GC) content, smaller genome sizes, and reduced gene numbers (Steenwyk et al., 2019). The genes lost by both lineages fall essentially into three categories: metabolism, DNA repair, and cell-cycle. Interestingly, the accelerated evolution observed in the FEL lineage was associated with an increased loss of genes related to DNA repair and cell-cycle regulation, that are believed to have facilitated genetic diversification (Steenwyk et al., 2019). In specific, it was found that FEL species, compared to SEL species, experienced increased mutational loads,

instabilities in homopolymers, and higher proportions of mutations linked to the common endogenously damaged base, 8-oxoguanine (Steenwyk et al., 2019).

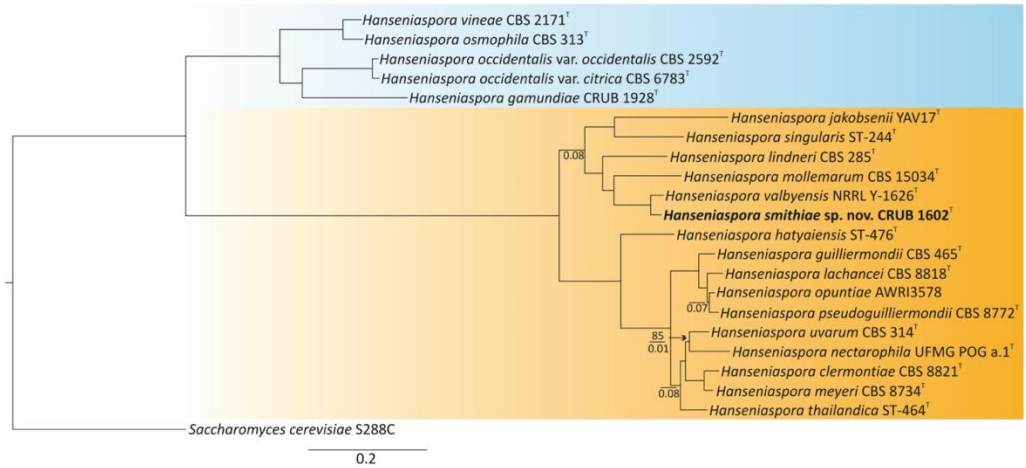


Figure I. 21– Phylogeny of the *Hanseniaspora* species, and the division between Fast-Evolving Lineage (FEL), represented inside the orange box, and Slow-Evolving Lineage (SEL), represented inside the blue box. Image adapted from Cadez et al., (2021)

Another consequence of FEL's loss of genes is ploidy variance (Steenwyk et al., 2019), although this hypermutation phenotype may have had opposite long-term effects, since the accumulation of deleterious mutations resulting from increased mutation rates may have ultimately slowed down sequence evolution (Ram & Hadany, 2012). Interestingly, this pattern of accelerated sequence evolution followed by a reduction in pace was also observed in SEL lineage of *Hanseniaspora*, though to a lower extent (Steenwyk et al., 2019).

This dynamic interplay between rapid evolution and subsequent deceleration suggests complex evolutionary trajectories within the *Hanseniaspora* genus, highlighting the delicate balance between adaptation and the accumulation of genetic load over time (Cadez et al., 2021). Still, the precise cause of this particular evolutionary strategy in *Hanseniaspora* remains elusive, although environmental factors may play a pivotal role in shaping those responses (Cadez et al., 2021). In fact, FEL's fast evolution, characterized by heightened mutation rates, may imply a more robust adaptive response to fluctuating and stressful environments, which can potentially increase the likelihood of beneficial mutations occurrence (Giraud et al., 2001; Healey et al., 2016). Given its frequent isolation from environments with abundant simple sugars, such as ripe fruit, *Hanseniaspora* may have evolved by selectively suppressing non-essential genes that do not contribute to its adaptation to these specific habitats (Cadez et al., 2021). Interestingly, the origin of grapes seems to coincide with the estimated origin time frame of this genus, suggesting a potential link between the evolutionary trigger of *Hanseniaspora* hypermutation and adaptation to the grapes' environment (Martin et al., 2018; Steenwyk et

al.,2019; Wikström et al.,2001). Phenotypically, the differential loss of genes experienced by the two lineages led to distinct characteristics and behaviors with SEL species being defined as the "fermentative clade" (encompassing, for example *H. vineae* and *H. osmophila* (van Wyk et al., 2022), while FEL lineage are designated as the "fruit clade" (encompassing *H. uvarum, H. guilliermondii* or *H. opuntiae* - van Wyk et al., 2022).

Genomic and Physiologic Insights into the *Hanseniaspora* genus

The disclosure and analysis of a *H. guilliermondii* genome (strain UTAD222) by Seixas and colleagues (2019) shed light on crucial aspects of the *Hanseniaspora* physiology. In fact, while possessing genes encoding enzymes for major pathways of central carbon metabolism, including glycolysis, TCA cycle and pentose phosphate pathway, the absence of a functional neoglucogenesis was revealed to be a metabolic characteristic across *Hanseniaspora* spp. (Seixas et al., 2019). This metabolic limitation was given by the absence of fructose-1,6-biphosphatase and phosphoenolpyruvate carboxykinase that are present in members of all the other families that compose the *Saccharomycetales* order (*e.g., Saccharomycetaceaee, Metschnikowiaceae* and *Pichiaceae*) (Seixas et al., 2019). Another important limitation found in *Hanseniaspora* metabolism concerned the absence of genes involved in biosynthesis of several essential vitamins and co-factors, such as thiamine, biotin, or pyridoxine (Seixas et al., 2019). In this work, the authors suggested that the deficiency in vitamin biosynthesis likely contributes significantly to the reduced fermentative capacity of *Hanseniaspora* (Seixas et al., 2019). This is particularly relevant because thiamine, in its active form as thiamine pyrophosphate (TPP), is an essential cofactor for pyruvate decarboxylase, playing a critical role in glycolysis and fermentation across a wide range of organisms (Hohmann & Meacock, 1998). In *S. cerevisiae*, for instance, thiamine availability directly influences glycolytic and fermentative fluxes by impacting pyruvate decarboxylase performance. When thiamine is limited, *S. cerevisiae* upregulates genes responsible for thiamine biosynthesis to maintain metabolic function (Seixas et al., 2019; Brion et al., 2014), and specific strains adapted to winemaking have evolved to enhance expression of these biosynthetic genes, thereby supporting fermentation under thiamine-limited conditions in grape musts (Seixas et al., 2019; Brion et al., 2014; Xu et al., 2012). Altogether, these findings suggest that *Hanseniaspora* species rely on environmental vitamins to survive, a hypothesis further supported by the strong up-regulation of *S. cerevisiae* genes involved in the biosynthesis of thiamine (ScTHI20 and ScTHI21), biotin (ScBIO3), and pyridoxin (ScSNO1) when in co-cultivation with *H.*

guilliermondii UTAD222 that is highly suggestive of nutrient competition (Barbosa et al., 2015; Seixas et al., 2019).

The genome analysis of strain *H. guilliermondii* UTAD222 also gave relevant insights into sugar uptake in *Hanseniaspora*. While *Hanseniaspora* species, and particularly *H. guilliermondii*, have been widely described as fructophilic (Ciani et al., 2010; Seixas et al., 2019; Cabral et al., 2015; Ciani et al.1999), possessing a particular affinity for fructose, no orthologue for fructose specific transporters (Ffz1 or Fsy1) could be found in the genome of *H. guilliermondii* UTAD222 (Seixas et al., 2019) nor in the genome of any other species of the *Hanseniaspora* genus, such as *H. uvarum, H. vineae* or *H. valbyensis* (Chen et al., 2023; Cadez et al., 2021; Cadez et al., 2019; Liu et al., 2019; Groenewald et al., 2018). Instead, the authors have disclosed two hexokinases in the genome of *H. guilliermondii* UTAD222 displaying a high level of similarity to those of *Z. bailii* strains, which have been demonstrated to favor fructose catabolism (Cordente et al., 2019; Giorello et al., 2014). While further studies on the *Hanseniaspora* hexose transporters are required, the phosphorylation of fructose mediated by hexokinases may be key to explain the fructophilic nature of *Hanseniaspora*, as described in wine *S. cerevisiae* strains (Cordente et al., 2019; Cadez et al., 2019; Groenewald et al., 2018). The analysis of the *H. guilliermondii "*transportome" led to the discovery that this species harbors multiple genes encoding sugar transporters, nearly half of which were hexose transporters, including high-affinity glucose and galactose permeases (Seixas et al., 2019). Most notably, the analysis of nitrogen transporters in *Hanseniaspora* revealed a preference for permeases with a broad substrate range over specific amino acid permeases, contrasting to *S. cerevisiae* that favors the use of specific amino acid permeases (Seixas et al., 2019). The absence of specificity in *H. guilliermondii's* permeases may, thus, represent an adaptive strategy to cope with the substantial variability and diversity of amino acids present in wine musts (Seixas et al., 2019).

I.2.1.2. The applications of *Hanseniaspora* in winemaking

The ability of the genus *Hanseniaspora* to modulate both primary and secondary wine aromas made it increasingly interesting and attractive for winemakers in pursuit for innovative products. Although commercial *Hanseniaspora* products are scarce, with only one commercially available starter culture (Fermivin VINEAE, Oenobrands, France), their abundance in the wine must, along with their potential to produce acetate esters and beta-glucosidase enzymes that, as described above, have an impact on the secondary and primary

wine aroma, turned interesting the utilization of *Hanseniaspora* species as co-adjuvants of *S. cerevisiae* (Martin et al., 2018; Steensels et al., 2014). The use of *Hanseniaspora* species in mixed fermentation with *S. cerevisiae* has been reported to produce wines with significantly different aroma profiles, usually with increased concentrations of higher alcohols, acetate esters and varietal aromas, namely isoamyl acetate and phenylethyl acetate, which are associated with increased fruity and floral characters of wines (Swiegers et al., 2005; Barbosa et al., 2015; Lage et al., 2014; Pietrafesa et al., 2020; Hu et al., 2018; Gallo et al., 2024). However, the formation of these aromas and by-products, as well as the population dynamics and fermentative ability of *S. cerevisiae* were found to be highly dependent on nutrient availability (Barbosa et al., 2015).

As previously described in this chapter, differential availability of nitrogen and sugars has been shown to exert opposite effects on acetate ester accumulation profiles between *S. cerevisiae* and *H. guilliermondii*. Specifically, a higher carbon-to-nitrogen (C/N) ratio leads to higher acetate ester production by *S. cerevisiae*, and a lower C/N ratio stimulates higher alcohol and acetate ester production in *H. guilliermondii* (Seixas et al., 2023). Additionally, transcriptomics studies on *H. guilliermondii* have revealed that it consumes less nitrogen in the presence of *S. cerevisiae*, due to competition for nutrients between the two species (Seixas et al., 2023). These findings support previous studies that showed that microbial interactions, influence *Hanseniaspora* behavior during fermentation (Wang et al., 2015; Harlé et al., 2020). Similarly, a study conducted by Barbosa and colleagues (2015) observed, for the first time, the effect of *H. guilliermondii* co-inoculation in the transcriptome of *S. cerevisiae* (Barbosa et al., 2015). In this work, the authors observed that the presence of *H. guilliermondii* led to 350 differently expressed *S. cerevisiae* genes, among which those related to vitamins biosynthesis were upregulated in mixed culture conditions, whereas genes involved in amino acids biosynthesis were enriched in single culture conditions (Barbosa et al., 2015). These findings, altogether with the observation of the decreased fermentative activity of the budding yeast in mixed fermentation (Barbosa et al., 2015), reinforce the need to study yeast physiology, their nutritional requirements and their interactions with other yeasts and the environment.

Hanseniaspora species are also notable for their enzymatic repertoire. Specifically, these yeasts produce beta-glucosidases, that besides contributing to terpene release can increase the antioxidant content in wines (van Wyk et al., 2023; Manzanares et al., 2000). For instance, compounds such as anthocyanins and the polyphenol resveratrol, a powerful antioxidant prevalent in red wines, are often bound to glycosylated precursors, inhibiting their beneficial

effects (Gaensly et al., 2015; Manzanares et al., 2011; Vidal et al., 2004). Furthermore, *Hanseniaspora* species have also been described as producers of beta-xylosidases, polygalacturonases and proteases, which are also highly desirable in winemaking, as they contribute to improving the wine's aromatic complexity and mouthfeel. Specifically, a study conducted by López et al. (2015) identified strains of *H. uvarum*, *H. osmophila* and *H. occidentalis* as having relevant beta-glucosidase and beta-xylosidase activities, as well as strains of *H. guilliermondii*, *H. vineae* and *H. occidentalis* as having high protease activity (López et al., 2015).

Relevant inter-species and inter-strain variability has been, however, observed among *Hanseniaspora* (Romano et al., 2003; Tristezza et al., 2016, Martin et al., 2018; López et al., 2015). Specifically, the *Hanseniaspora* species *H. osmophila*, *H. occidentalis* and *H. vineae* were reported to produce lower levels of acetaldehyde and ethyl acetate, when compared to *H. guilliermondii*, *H. meyeri*, *H. opuntiae*, *H. nectarophila* and *H. uvarum* (Badura et al., 2023). On the other hand, *H. uvarum* is typically a high beta-glucosidase producer (Hu et al., 2016; Zhang et al., 2021; Gao et al., 2022), as well as *H. vineae*, displaying higher activities compared to *H. osmophila* and even to *H. guilliermondii* (López et al., 2014; Martin et al., 2018). Despite these general trends, considerable variability also exists among strains within the same species, underscoring the importance of careful yeast selection to ensure desirable fermentation outcomes (Tristezza et al., 2016; López et al., 2015; Albertin et al., 2016). Moreover, the wine's phenotype can be influenced not only by strain selection but also by fermentation conditions. Factors such as the mode of inoculation, fermentation temperature or the presence or absence of aeration, significantly impact the aroma profile (Lage et al., 2014; Barbosa et al., 2022; Albergaria et al., 2003). Therefore, in addition to select the most suitable species or strain, a meticulous management and engineering of fermentation conditions are crucial to achieving optimal aroma attributes.

I.2.2. The *Saccharomycodes* Yeasts

First described in 1904 by E.C Hansen, *Saccharomycodes* yeasts are large, apiculate organisms that, like other yeasts from the *Saccharomycodaceae* family, divide asexually by bipolar budding (Boundy-Mills et al., 2011). *Sd. ludwigii* was initially isolated from deciduous trees in Europe (Phaff et al., 1998) and has since been found in soil, fruits and fruit juices, insects, and fermented beverages (Boundy-Mills et al., 2011). However, it is widely accepted

that this species originates from oak trees, as well as the slime fluxes and insects associated with oak trees (Phaff et al., 1998, Beech & Davenport, 1970, Lachance et al., 1995). The *Saccharomycodes* genus was first described in the 19th century by the German mycologist and botanist, Gustav Ludwig, with *Sd. ludwigii* as the only species (Ludwig, 1886). Later, *Sd. sinensis* was added to the genus based on sequence analysis of 18S ribosomal DNA (rDNA) (Yamada et al., 1992, Wang et al., 2015). However, subsequent works by Kurtzman and Robnett (1998), as well as Yamazaki et al. (2005) presented substantial evidence highlighting significant differences between *Sd. sinensis* and *Sd. ludwigii*, suggesting that they should not be classified within the same genus. Specifically, these authors demonstrated that the first type strain of *Sd. sinensis* (NRRL Y-12797*)* had a sequence of the D1/D2 domains of the 26S rDNA more identical to that of *Nadsonia fulvescens* var. *elongata* than to *Sd. ludwigii* (Kurtzman and Robnett, 1998, Boundy-Mills et al., 2011). Later, Yamazaki et al (2005), analyzed the D1/D2 sequences of another type strain of *S. sinensis* (IFO10111) and concluded that this species did not belong to any genera within *Saccharomycodaceae* (Wang et al., 2015). To correctly classify *Sd. sinensis*, a new genus within the *Saccharomycetaceaee* family, *Yueomyces*, was proposed in 2015, including *Yueomyces sinensis* (Wang et al., 2015). More recently, Wang et al. (2021) proposed a new species within the *Saccharomycodes* genus, *Sd. pseudoludwigii*, which was isolated from fruit and tree bark in China (Wang et al., 2021). These findings were based on the comparison of several gene sequences, namely the ribosomal genes 18S rRNA, D1/D2 domains of the 26S rRNA, and the ITS region (including the 5.8S rRNA), as well as the translation elongation factor EF-1, TEF1, and the RNA polymerase II gene, RPB2 from twelve different *Saccharomycodes* strains (nine *Sd. ludwigii* strains and three *Sd. pseudoludwigii*) (**Figure I. 22**). This study demonstrated that, while the three putative *Sd. pseudoludwigii* strains were close relatives to *Sd. ludwigii*, they exhibited significant dissimilarities. These differences included 4 substitutions and 1 insert/deletion in the D1/D2 domain, 12 substitutions and 9 inserts/deletions in the ITS regions, 39 synonymous and 4 non-synonymous substitutions in the RPB2 gene, and 18 synonymous and 3 non-synonymous substitutions in the TEF1 gene, between *Sd. pseudoludwigii* and *Sd. ludwigii* strains (Wang et al., 2021). Furthermore, the alignment of these gene sequences and the resultant multi-locus tree, clearly showed that the three *Sd. pseudoludwigii* strains formed a distinct clade from *Sd. ludwigii*, supporting the creation of a new species (**Figure I. 22**).

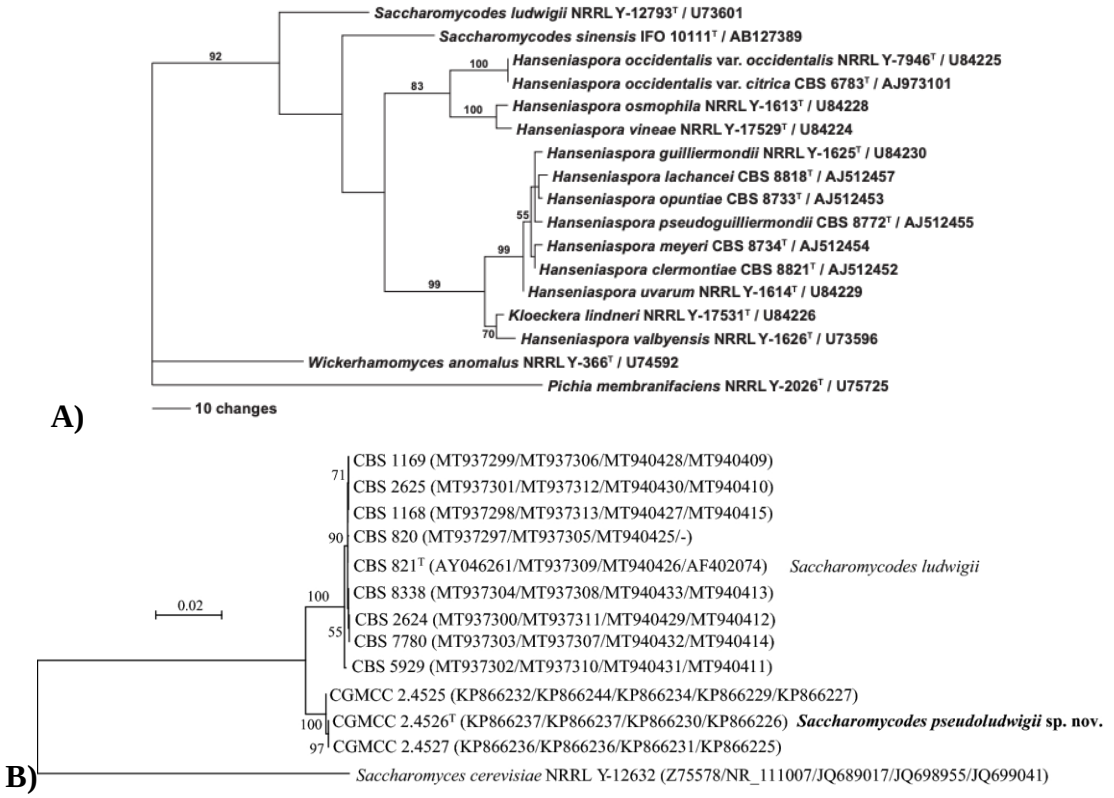


Figure I. 22– A) Phylogenetic placement of the *Saccharomycodes* and *Hanseniaspora* genera reveals that *Sd. ludwigii* is loosely related to *Hanseniaspora* and *Sd. sinensis. Pichia membranifaciens* was used as an outgroup. Adapted from Miller et al (2011). **B)** New species, *Sacccharomycodes pseudoludwigii* based on a multi-locus tree combining sequences of the18S rRNA, D1/D2 domains of the 26S rRNA, the ITS region (including the 5.8S rRNA), the *RPB2* and *TEF1* of nine *Sd. ludwigii* strains and three *Sd. pseudoludwigii* strains. Adapted from Wang et al., (2021).

Sd. ludwigii strains that had been subjected to karyotyping harbor seven chromosomes, with molecular sizes ranging from 0.8 to 2.3 Mbp, although strain-specific polymorphic chromosomal DNA has been observed (Yamazaki & Oshima, 1996). The first report of these findings was from a study conducted by Yamazaki and Oshima (1996) in which the authors performed the tetrad analysis of 24 genetic markers. These markers included the two mating type alleles (a/α), 21 auxotrophic mutant genes of various phenotypes, and two drug-resistant mutant genes of one *Sd. ludwigii* strain. Additionally, chromosomal electrophoresis through Pulsed-Field Gel Electrophoresis (PFGE) showed that the 24 genes analyzed were classified into seven linkage groups, with seven chromosomal bands also being observed in the PFGE gels (**Figure I. 23A**). However, strain-specific polymorphisms were noted in three pairs of homologous chromosomes in a diploid strain (**Figure I. 23A**). Specifically, these chromosomes appeared as doublets consisting of homologous chromosomes of slightly different sizes. To further investigate this hypothesis, the authors performed PFGE on four tetrad clones from a single ascus of the diploid strain, which resulted in a different band profile (**Figure I. 23B**).

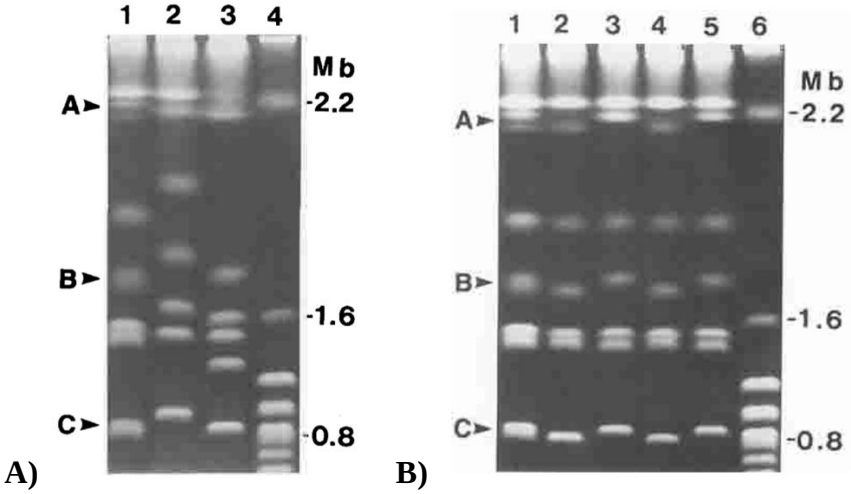


Figure I. 23– Electrophoretic karyotypes of *Sd. ludwigii* strains. A) Electrophoretic karyotypes of 3 *Sd. ludwigii* strains: Lane 1 consists of the diploid strain *Sd. ludwigii* 0-81; lane 2 and lane 3 consist of the haploid strains *Sd. ludwigii* IF0 1043 and *Sd. ludwigii* OUT 6282; lane 4 is *S. cerevisiae* YNN 295, used as a size marker. Arrowheads A, B and C indicate three putative doublet bands of chromosomes in diploid strain 0-81. B) Electrophoretic karyotypes of diploid strain 0-81 (lane 1) and its tetrad clones from a single ascus (lanes 2-5), and *S. cerevisiae* YNN 295 as size marker (lane 6). Arrowheads A, B and C indicate three putative doublet bands of chromosomes of strain 0-81. Adapted from Yamazaki and Oshima (1996).

In terms of its life-cycle, *Sd. ludwigii* was described in 1939 by Winge and Laustsen (1939) as a heterothallic species. Specifically, it was observed that after sporulation each cell becomes an ascus containing two pairs of spores, with one pair at each end (Yamazaki et al., 1976, Winge & Laustsen, 1939, Barnett, 2007). Typically, each pair of spores copulates upon germination, producing diploid cells and leaving no opportunity for haplophase multiplication (Winge and Laustsen, 1939, Yamazaki et al., 1976). However, with *Sd. ludwigii*, the authors were able to obtain haploid vegetative cells by single spore isolation (Winge & Laustsen, 1939). Interestingly, this work supports a theory first suggested by F. Ludwig in 1886, in which this author illustrated spore segregation inside the ascus, with the spores from each mating type adhering to each other and the resulting zygote producing normal vegetative cells (Ludwig, 1886) (**Figure I.24**).

Figure I. 24- Ascus of *Saccharomycodes ludwigii* showing paired ascospores. Illustration by Ludwig, published in 1886 and adapted from Barnett (2007).

In their publication, Winge and Laustsen described the behavior of *Saccharomycodes ludwigii's* tetrads. By dissecting several tetrads and applying different conditions to the spores, they observed various phenotypes (Winge & Laustsen, 1939). In one experiment, they allowed all four spores to germinate separately. In another, they separated one pair, allowing both spores to germinate independently while leaving the other pair to form a zygote (**Figure I. 25A**). They observed that in the separated pair, one spore continued to divide continuously, while the other ceased dividing after only a few divisions (**Figure I. 25A**). In homothallic yeasts, such as most natural isolates of *S. cerevisiae* (Katz Ezov et al., 2010), a haploid cell can switch its mating type through a gene-conversion process, allowing mating between mother and daughter cells (**Figure I. 25B**). This ability is not present in *Sd. ludwigii*. When haploid clones of *Sd. ludwigii* from the same mating type were kept apart from other colonies, they did not form diploids. This indicates that, unlike *S. cerevisiae*, which can exhibit both homothallic and heterothallic behaviors, *Sd. ludwigii* is exclusively heterothallic (Barnett, 2007).

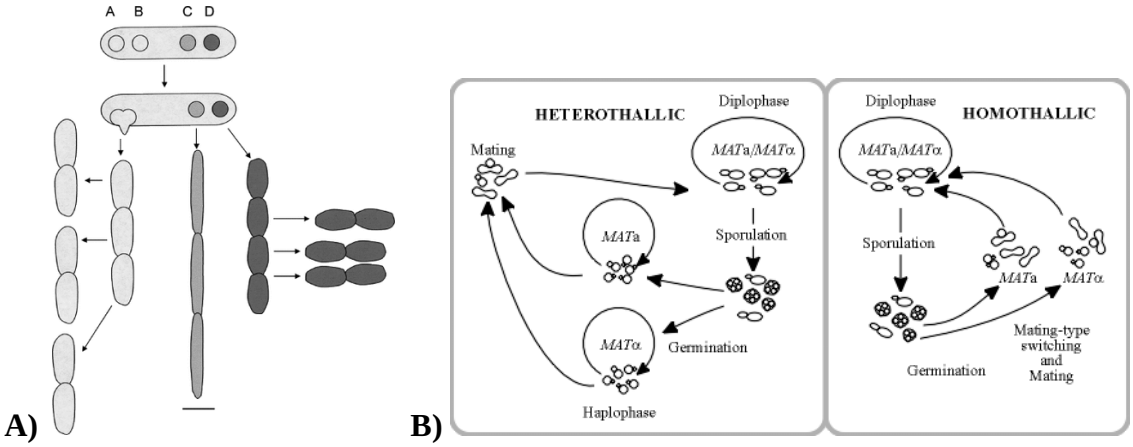


Figure I. 25– A) Winge and Laustsen model of *Sd. ludwigii's* tetrad. In this work, the authors found that ascospores A and B fused to form a diploid zygote, which buds and forms a growing colony of large cells. Ascospores C and D did not conjugate. Instead, they formed haploid daughter cells: those from ascospore C were elongate and stopped budding after two or three divisions, whereas the daughter cells from ascospore D were small and continued to grow to form a haploid colony (adapted from Barnett (2007). **B)** Differences in heterothallic and homothallic behaviour: in heterothallic yeasts, after sporulation, spores from different mating types, MAT-a and MAT-α, germinated maintaing haplophase; in homothallic yeasts, mother cells are able to switch mating-type and mate with daughter cells, becoming diploid. Adapted from Aksit (2012).

In their studies, Winge and Laustsen observed that *Sd. ludwigii* preferentially undergoes intra-tetrad mating (**Figure I. 26**), facilitated by strong interspore bridges that efficiently keep spores together in pairs of opposite mating types within meiotic tetrads (Winge & Laustsen, 1939, 1942, Papaioannou et al., 2021, Simmons & Ahearn, 1985). Recently, Miyakawa and colleagues have focused their work on understanding *Sd. ludwigii'*s interspore bridges and their role in this mating process (Miyakawa et al., 2012, 2016, 2020). Specifically, the authors

observed that the components of the interspore bridge were different from those of the spore wall, and that the lemon-shape morphology of *Sd. ludwigii*, as well as the polarity of the first and second meiotic divisions, were fundamental in producing a pair of spores of the opposite mating types (Miyakawa et al., 2012, 2016). In their pioneering work on this matter, Miyakawa's team isolated, for the first time, *Sd. ludwigii's* interspore bridges and observed that these structures detached from spore walls during germination and zygote formation (Miyakawa et al., 2020). A structural study revealed a different composition of these bridges compared to those in *S. cerevisiae*. Notably, while in *S. cerevisiae* spores can be separated with a glass needle, repeated sonication cycles could not break the bridges in *Sd. ludwigii* (Miyakawa et al., 2020). Furthermore, it was observed that these bridges do not contain proteins as major components, as no denaturation occurred following treatments with SDS and proteinase K (Miyakawa et al., 2020). These studies reinforce the unique and robust nature of *Sd. ludwigii's* interpsore bridges that tightly tether two spores of opposing mating types, facilitating the fusion of spores within the same tetrad, instead of dispersing and finding unrelated spores to mate with (Zakharov, 2023).

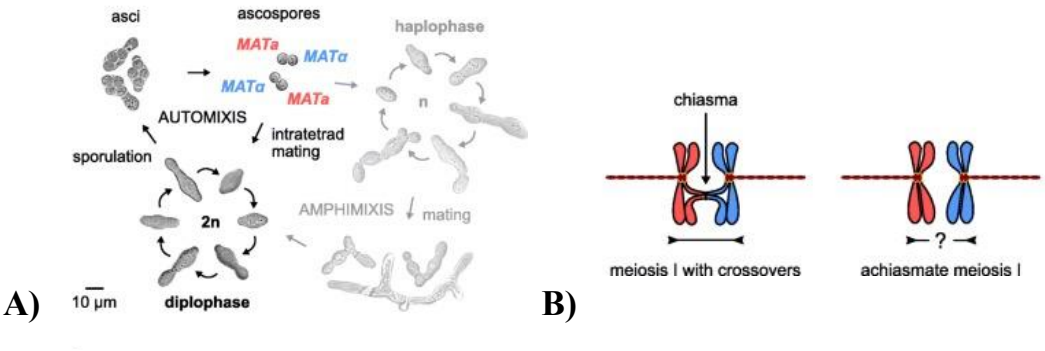


Figure I. 26– Life cycle of *Saccharomycodes ludwigii*. A) Brightfield images of strains *Sd. ludwigii* NBRC 1722 and diploid progenitor NBRC 1721 in their life cycle. B) Prophase I of meiosis with chiasma formation and crossover versus achiasmate meiosis observed in *Sd. ludwigii*. Adapted from Papaioannou et al., (2021).

In addition to intratetrad mating, Winge and Laustsen also noted an unusual segregation pattern in two pairs of marked genes, Nn and Ll (**Figure I. 27)** in *Sd. ludwigii*. The gene *N* was associated with the formation of normal cells, whereas the allele *n* resulted in a lethal, abnormal hyphae-like germination that persisted for only two or three generations. Additionally, the gene *L* produced long cells, while its allele *l* led to short cells (Winge & Laustsen, 1939, Barnett, 2007, Winge, 1946). These observations were later explained by Lindegren, who interpreted that *N* and *L* were on different chromosomes and positioned near the centromere, thereby preventing the occurrence of crossing over between them (Barnett, 2007, Lindegren, 1945).

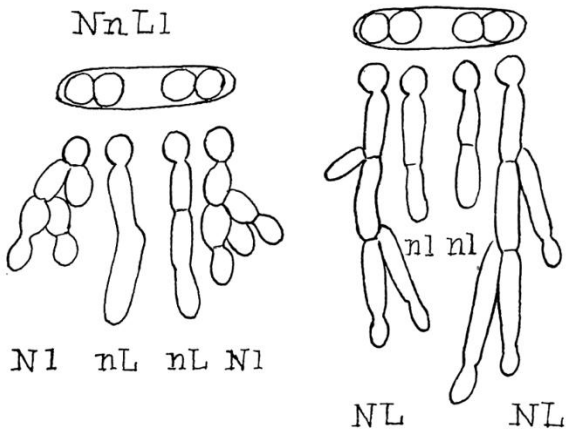


Figure I. 27– Two kinds of segregation in the asci of *Saccharomycodes ludwigii*, heterozygous for genes N (N,n) and L (L,l) as described by Winge and Laustsen (1942). The phenotypes observed for each allele, obtained from independent germination of each ascospore, were the following: N allowed normal growth; n was lethal and divided only a few times; L provided long cells, and l provided short cells. Adapted from Barnett (2007).

Recently, in 2021, Papaioannou and colleagues conducted an extensive study to understand the absence of meiotic crossing-over and the preference for intratetrad mating in *Sd. ludwigii*. By combining bioinformatics, molecular biology, and evolutionary approaches, they aimed to uncover the genetic and evolutionary reasons behind this unique meiotic behaviour (Papaioannou et al., 2021). The researchers first hypothesized that significant sequence variation or structural differences between homologue chromosomes might explain the lack of crossing-over. However, genome sequencing of two parental haploid *Sd. ludwigii* strains showed high collinearity, ruling out major differences between homologues. Additionally, by analyzing non-random chromosomal distribution of genetic markers, the authors observed that all chromosomes were covered by these markers, indicating that these genes captured most of the existing crossing-over, and that no significant chromosomal structural differences could be observed (Papaioannou et al., 2021). Then, to investigate the genetic causes for the absence of crossing-over in *Sd. ludwigii*, the authors studied the meiotic machinery of this yeast. Interestingly, they observed that, with the exception of gene *MER1*, which is absent in *Sd. ludwigii* and in *S. cerevisiae* is responsible for major defects in meiotic recombination, *Sd. ludwigii* retains the essential machinery for meiotic recombination (Papaioannou et al., 2021). Supporting these findings was the observation that some of the meiotic genes that were missing in *Sd. ludwigii* were also absent in other members of the *Saccharomycetaceaee* (*e.g., Lachancea kluyvery*), which are capable of performing meiotic crossing-over (Papaioannou et al., 2021). Interestingly, similar to *S. cerevisiae*, key meiotic recombination components are also required in *Sd. ludwigii's* meiosis. Specifically, the

deletion of the meiotic recombination protein SPO11 in *Sd. ludwigii* led to reduced sporulation and spore viability, highlighting its role in generating meiotic DNA double-strand breaks (DSBs). In addition, immunostaining of the recombinase protein RAD51 demonstrated its specific localization during meiosis and its dependence on SPO11. The presence of discrete foci in wild-type strains contrasted with their absence in *Δrad51* or *Δspo11*mutants, indicating the necessity of these proteins for normal meiotic function. Additionally, the deletion of the DMC1 gene, another meiosis-specific recombinase, resulted in elongated, filamentous RAD51 foci, similar to patterns observed in *S. cerevisiae*, indicating its role in DSB repair and in interhomolog interactions (Papaioannou et al., 2021). Furthermore, the Single Nucleotide Polymorphism (SNP) analysis of *Sd. ludwigii*'s meiosis performed between a haploid strain and a geographically distant strain revealed low sequence divergence, and crosses between these strains frequently produced tetrads with four viable spores (Papaioannou et al., 2021). Sequencing of all spores from five full tetrads of this cross, followed by analysis of the SNP segregation patterns, revealed a complete absence of meiotic crossing-overs. Additionally, the authors extended this analysis to two tetrads from another diploid strain and confirmed the extreme rarity of meiotic crossing-over in *Sd. ludwigii*. Additionally, they investigated the existence of crossing-over in telomeric regions of chromosomes. Their findings revealed that the initiation of recombination in *Sd. ludwigii* was not biased towards telomeres or adjacent regions. Instead, the extreme suppression of crossing-over was consistent across the entire genome (Papaioannou et al., 2021).

Finally, the authors searched for evolutionary signs of homologue interactions in *Sd. ludwigii*, by comparing divergent strains. They sequenced ten haploid and diploid strains, finding variable sequence divergence and heterozygosity levels in diploids. By performing a phylogenetic analysis based on genome-wide SNP content, they identified a major cluster of seven strains and four more divergent ones. Comparison between the dendrograms of individual chromosomes revealed chromosome-dependent topologies and distances for certain strains, with chromosomes showing unstable topologies having significantly different SNP densities from their genomic averages. This suggested that the absence or rarity of meiotic recombination over time has allowed chromosomes to accumulate SNPs independently. Supporting this hypothesis, similar analysis in *Lachancea kluyveri*, a related yeast capable of meiotic recombination, showed very stable topologies with all chromosomes matching the whole-genome topology (Papaioannou et al., 2021).

Overall, these findings suggested that the genetic diversity within populations and the level of heterozygosity in individuals are significantly influenced by mating strategies. Fertilization within the meiotic tetrad offers *Sd. ludwigii* two significant biological advantages over random spore pairing: firstly, individual spores are unlikely to find a suitable mate during fungal cell dispersal, thus maintaining spore connection after meiosis allows them to fuse in pairs, preserving the diploid stage; secondly, intratetrad mating preserves heterozygosity in large genome sections, enhancing the development of improved biological features in daughter cells (Zakharov, 2023). Specifically, the centromeric regions of all chromosomes remain heterozygous, provided that the mating type locus is closely linked to the centromere of its chromosome (Zakharov, 2023) (**Figure I. 28**). Therefore, according to these authors, *Sd. ludwigii* may have evolved to perform achiasmate meiosis to ensure a successful adaptation to its environment, by thriving with minimal levels of rare meiotic recombination (Papaioannou et al., 2021).

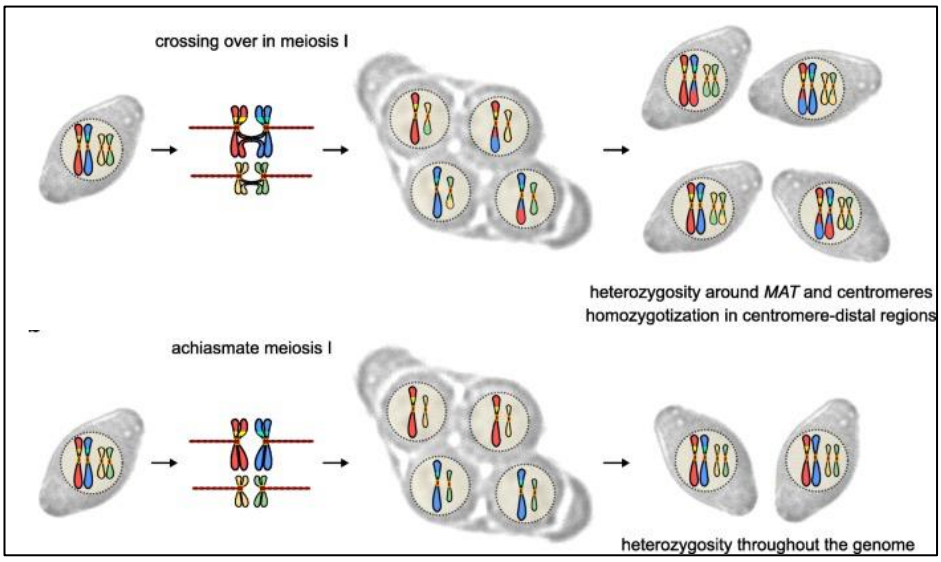


Figure I. 28– Preservation of heterozygosity in *Sd. ludwigii* through high-rates of intratetrad mating coupled with suppression of meiotic recombination. A) Intratetrad mating with crossing-over in meiosis I results in heterozygosity around the mating-type locus (MAT) and centromeres, but in homozygotization in centromere-distal regions. B) Intratetrad mating with achiasmate meiosis (without crossing-over) is the preferred mating strategy in *Sd. ludwigii* and results in heterozygosity throughout the genome. Adapted from Papaioannou et al., (2021).

I.2.2.1. The role of *Saccharomycodes* in winemaking, and additional applications

In the winemaking context, *Sd. ludwigii* is usually regarded as a spoilage agent, being usually detected in the last stages of fermentation, or even, during storage, in bottled wines or in badly sanitized corks, usually resulting in cloudiness and turbidity phenotypes (Vejarano, 2018, Esteves et al., 2019, Loureiro & Malfeito-Ferreira, 2003, Thomas, 1993). Due to its

unique high tolerance to sulfur dioxide (Vejarano, 2018, Stratford et al., 1987, Domizio et al., 2011), along with its capacity to ferment and endure ethanol concentrations up to 12% (Ciani & Maccarelli, 1997), this species is difficult to eradicate from wineries and cellar equipment. For this reason, it is often coined as the "winemakers nightmare" (Vejarano, 2018). Aside from negatively impacting the visual quality of wines, *Sd. ludwigii* is also able to modify its aroma and flavor, resulting in a completely different product from an organoleptic point of view (Vejarano, 2018, Esteves et al., 2019, Romano et al., 1999). Aromas associated to *Sd. ludwigii* include an increased production of ethyl acetate, acetaldehyde, acetoin, as well as in the higher alcohols isobutanol, amyl alcohol, and isoamyl alcohol (Domizio et al., 2011, Vejarano, 2019, Esteves et al., 2019, Ciani & Maccarelli, 1997, Romano et al., 1999, Granchi et al., 2002), which can confer negative undertones to the wine upon exceeding their respective thresholds of perception.

A much more consensual application of *Sd. ludwigii* is related to its use in the production of low-alcohol beers or with differentiated flavor (De Francesco et al., 2015, Sileoni et al., 2023, Adamenko et al., 2020, Jackowski et al., 2023). In fact, a wide number of reports have stated that the use of *Sd. ludwigii* in beer production leads to a desirable aroma, characterized by a high ester content (De Francesco et al., 2015), low concentration of the undesirable diacetyl (De Francesco et al., 2015), and high antioxidant potential (higher polyphenol content) (Adamenko et al., 2020). The good results obtained with this yeast for beer production have led to the development of a commercial product, Fermentum Mobile FM58, (Fermentum Mobile, Gdańsk, Poland) (Jackowsky et al., 2023). Additionally, *Sd. ludwigii* has been appointed to produce other fermented beverages, such as cider (Estela-Escalante et al., 2011), or fruit juices (Romano et al., 1999). Besides the beverage industry, a recent study conducted by Pilap et al., (2022) has revealed a novel potential for the second-generation of bioethanol production using an *Sd. ludwigii* strain, due to its thermotolerant nature and high acetic acid resistance.

While OMICS-related studies are more abundant for *Hanseniaspora*, there is still a significant knowledge gap regarding *Saccharomycodes*. This work aims to address this gap by exploring wine *Saccharomycodacea* yeasts using comparative OMICs, aiming to enhance their application in various biotechnological aspects.

Chapter II.

Genome Sequencing, annotation and exploration of the SO2-tolerant non-conventional yeast *Saccharomycodes ludwigii*

Part of the results shown in this chapter were published in:

Tavares MJ, Güldener U, Esteves M, Mendes-Faia A, Mendes-Ferreira A, Mira NP, "Genome Sequence of the Wine Yeast *Saccharomycodes ludwigii* UTAD17". Microbiol Resour Announc, 7 (2018). doi: /10.1128/mra.01195-18

Tavares MJ, Güldener U, Mendes-Ferreira A and Mira NP, "Genome sequencing, annotation and exploration of the SO_2-tolerant non-conventional yeast *Saccharomycodes ludwigii*", BMC Genomics, 22, 131(2021). doi: /10.1186/s12864-021-07438-z

II.1 Abstract

Saccharomycodes ludwigii belongs to the poorly characterized *Saccharomycodaceae* family and is known for its ability to spoil wines, primarily due to its high tolerance to sulfur dioxide (SO_2). To enhance our understanding of *Sd. ludwigii*, our group initially performed Illumina sequencing, resulting in a draft genome consisting of over 1,000 contigs. Subsequently, we re-sequenced the *Sd. ludwigii* UTAD17 genome using PacBio technology, which produced 20 contigs assembled into a 13 Mb genome, representing 95% of the DNA content of this strain, as estimated by karyotying. Annotation of the assembled UTAD17 genome predicted 4,644 protein-encoding genes. Comparative analysis of the predicted *Sd. ludwigii* ORFeome with other *Saccharomycodaceae* species led to the identification of 213 unique proteins. These included six enzymes involved in N-acetylglucosamine catabolism, four cell wall beta-mannosyltransferases, several flocculins and three acetoin reductases. Unlike its sister species in the *Hanseniaspora* genus, *Sd. ludwigii* possesses functional pathways for gluconeogenesis, the glyoxylate cycle and thiamine biosynthesis. Additionally, *Sd. ludwigii's* genome revealed four efflux pumps similar to the Ssu1 sulfite exporter, and robust orthologues for 65% of the SO_2-tolerance genes found in *S. cerevisiae*, suggesting a genetic basis for its high SO_2 tolerance.

This work provides the first genome-wide picture of a *Sd. ludwigii* strain, advancing our understanding of the physiology and genetics of this species and the *Saccharomycodaceae* family. The release of this genomic sequence and the insights gained from it can guide the development of improved wine preservation strategies to counteract spoilage prompted by *Sd. ludwigii*. Furthermore, it will accelerate the exploration of this species as a cell factory, particularly in the production of fermented beverages, where the use of Non-*Saccharomyces* species, including of spoilage species, is gaining popularity.

Keywords: *Saccharomycodes ludwigii, Saccharomycodaceae,* Non-*Saccharomyces* wine yeast, Sulfur resistance, Genome sequencing

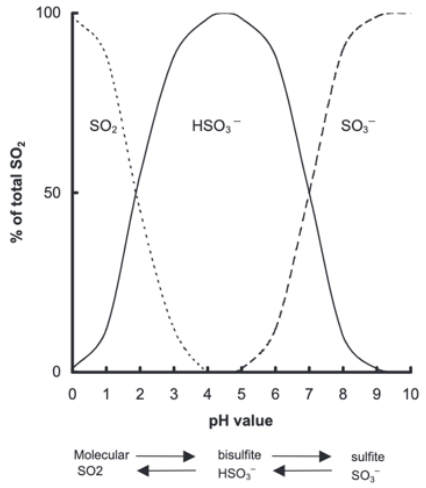
II.2 Introduction

Saccharomycodes ludwigii is a budding yeast in the *Saccharomycodaceae* family, known for its large-apiculate morphology, and notorious for spoiling wines due to its high tolerance to sulfur dioxide (SO_2) (Vejarano et al., 2018). While the *Saccharomycodaceae* family also includes the *Hanseniaspora* genus - species like *H. guilliermondii* and *H. uvarum*, which positively impact wine aroma (Boundy-Mills et al., 2011; Martin et al., 2018; Lage et al., 2014) - *Sd. ludwigii* is predominantly a spoilage species. *Sd. ludwigii* strains are frequently isolated from sulfite-preserved grape musts, end-of-vinification samples, and during storage (Boundy-Mills et al., 2011; Romano et al., 1999; Granchi et al., 2002; Boulton et al., 1999). Several sources have been suggested to serve as reservoirs of *Sd. ludwigii*, including grape surfaces (Combina et al., 2005; Barata et al., 2012), non-sanitized corks (Combina et al., 2005; Vejarano et al., 2018; Loureiro & Malfeito-Ferreira, 2003), and cellar equipment (Vejarano, 2018, Loureiro & Malfeito-Ferreira, 2003; Stringini et al., 2009), complicating spoilage control. Additionally, the yeast has been identified in plant fluids (Combina et al., 2005; Stringini et al., 2009) and the intestinal microbiota of vineyard insects (Fogleman et al., 1982; Stefanini, 2018), suggesting possible transportation vectors. Spoilage effects associated to *Sd. ludwigii* include off-flavors like acetoin, ethyl acetate, acetaldehyde, and acetic acid (Romano et al., 1999; Vejarano, 2018; Loureiro & Malfeito-Ferreira, 2003), as well as increased sediment and cloudiness in wines (Loureiro & Malfeito-Ferreira, 2003; Fleet, 2003). Beyond wines, *Sd. ludwigii* has been found in spoiled carbonated beverages (Hutzler et al., 2012), fermented fruit juices (Erkmen, O. & Bozoglu, 2016; Roller, S. & Covill, 1999), and high-ethanol drinks like mezcal and tequila (Lachance, 1995).

The spoilage ability of *Sd. ludwigii* is largely due to its high SO_2 tolerance, which is used by winemakers as an antimicrobial agent, with its effectiveness dependent on the concentration of undissociated SO_2, predominant at low pH levels (Stratford et al., 1987; Stratford & Rose, 1986). At wine pH (3-4), bisulfite is the most abundant form, and once inside the cell, it accumulates due to its inability to cross the plasma membrane. After crossing the microbial plasma membrane by simple diffusion, the lipophilic molecular SO_2 dissociates in the near-neutral cytosol resulting in the release of protons and of bisulfite which, due to its negative charge, cannot cross the plasma membrane and accumulates internally (**Figure II. 1A**) (Stratford et al., 1987; Stratford & Rose, 1986). Notably, studies with *Sd. ludwigii* (at pH 4.0) revealed that the accumulation of SO_2 inside the cells was significantly lower than the one registered for *S. cerevisiae* (Stratford et al., 1987) that is much more susceptible to SO_2. That

different accumulation was hypothesized (but not experimentally demonstrated) to result from the different lipid composition of the plasma membrane of these two yeasts that may result in different permeabilities to SO_2 (Stratford et al., 1987). In the presence of SO_2 *Sd. ludwigii* cells excrete high amounts of the SO_2-sequestering molecule acetaldehyde, however this response does not seem to account for the enhanced tolerance of this species since similar excretion rates were observed in susceptible *S. cerevisae* strains (Stratford et al., 1987).

A)



B)

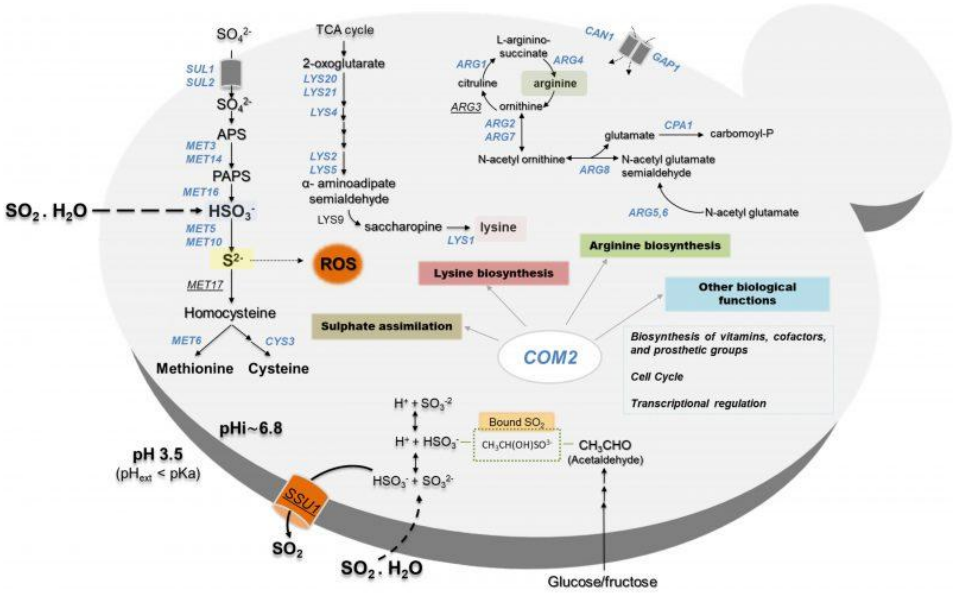


Figure II. 1– A) Distribution of sulfite, bisulfite, and molecular SO_2 as a function of pH in aqueous solution. Adapted from Jarvis (2014 B) Schematic representation of *S. cerevisae* response to SO_2. Adapted from Lage et al., (2019).

To counter-act the deleterious effect of intracellular accumulation of SO_2, *S. cerevisiae* relies on the activity of the sulfite plasma membrane transporter Ssu1, which is believed to

promote the extrusion of metabisulfite (Lage et al., 2019; Avram & Bakalinsky, 1997) (**Figure II. 1B**). The high tolerance to SO_2 of *Brettanomyces bruxellensis*, another relevant wine spoilage species, was also associated to the activity of Ssu1 (Varela et al., 2019), however, in *Sd. ludwigii* no such similar transporter has been described until thus far. In fact, the molecular traits underlying the high tolerance to SO_2 of *Sd. ludwigii* remain largely uncharacterized.

To fill this gap and further understand *Sd. ludwigii*, our group isolated and sequenced the first publicly genome of a *Sd. ludwigii* strain, UTAD17, isolated from a wine must obtained from the demarcated Douro region, in Portugal. This sequencing project was performed using Illumina and PacBio, resulting in 20 contigs and a predicted ORFeome of 4,528 canonical protein-coding genes, comparable to other *Saccharomycodaceae* members like *H. osmophila*, which has 4,657 predicted proteins (Sternes et al., 2016).

This chapter describes the information extracted from the genomic sequence of the UTAD17 strain shedding light into the biology and physiology of the *Sd. ludwigii* species with emphasis on the "SO_2 tolerance" phenotype. Not only this is expected to contribute for the design of better preservation strategies by the wine industry to circumvent spoilage caused by *Sd. ludwigii*, but this is also expected to accelerate the exploration of this species (and specially of this strain) in production of fermented beverages and in other biotechnological applications. In fact, there is a growing interest of using Non-conventional yeast species, including those previously considered spoilage organisms, to enhance the aroma profile of fermented beverages, and this portfolio of new potentially interesting species includes *Sd. ludwigii* (Holt et al., 2018; Steensels et al., 2015; Domizio et al., 2014).

II.3 Methods

II.3.1 Strains and Media

The strains used in this chapter were *Saccharomycodes ludwigii* UTAD17 and *Candida albicans* SC5314. UTAD17 was isolated from a wine must obtained from the Douro demarcated region harvested from the experimental vineyard of UTAD (with approved use for research), in Portugal. *C. albicans* SC5314 is a reference strain of *C. albicans.*

Cells were grown in rich growth media Yeast Peptone Dextrose (YPD) or in Minimal Medium (MM). YPD contains, per liter, 20 g glucose (Labchem), 10 g yeast extract (Gibco) and 20 g Peptone (HiMedia Laboratories). MM contains, per liter, 20 g glucose (Labchem), 2.65 g $(NH_4)_2SO_4$ (Merck Millipore) and 1.7 g of yeast nitrogen base without amino acids and without ammonium sulphate (Difco). YPD and MM medium were sterilized by autoclaving for 15 minutes at 121ºC and 1 atm.

II.3.2 Pulsed-field gel electrophoresis (PFGE)

Separation of *Sd. ludwigii* UTAD17 chromosomal DNA was carried out as described by Sipiczki et al. (2003) and as modified by El Hage & Houseley (2013). Briefly, yeast chromosomes were separated in 1% agarose gels in 0.5 x TBE buffer (0.13 M tris-base (pH 7.6), 45 mM Boric Acid, 2.5 mM EDTA) cooled at 12 °C in a BioRad CHEF-DRIII electrophoresis apparatus (Bio-Rad, Hercules, CA, USA). Electrophoresis was conducted at 3 V/cm for 36 h with a 200–300 s ramping switch interval and for 60 h with a 300–600 s ramping switch interval. The CHEF-DNA size markers used to calculate the molecular sizes of UTAD17 chromosomal bands were *Hansenula wingei*, for chromosome bands ranging from 1.05 to 3.13 Mbp and *S. cerevisiae* (for chromosomes below 1.05 Mbp). The molecular sizes for *Sd. ludwigii* UTAD17 chromosomes were then calculated through a calibration curve (band distance vs molecular size) making use of ImageJ software.

II.3.3 Genome sequencing, assembly and annotation of *Sd. ludwigii* UTAD17

Genomic DNA extraction of *Sd. ludwigii* UTAD17, as well as subsequent sequencing and assembly was performed by two different companies: Illumina sequencing was performed by STABVIDA (Caparica, Setúbal, Portugal), and PacBio sequencing was conducted by CD Genomics (Shirley, New York, United States). Briefly, genomic DNA of *Sd.*

ludwigii UTAD17 was extracted using Quiagen Magattract HMW kit according to the manufacturer's instructions. The DNA quality was evaluated using a Qubit fluorometer and a Fragment Analyzer™ Automated CE System combined High Sensitivity Large Fragment 50Kb Analysis Kit.

For Illumina paired-end sequencing, the DNA libraries of *Sd. ludwigii* UTAD17 were prepared using the ThruPLEX DNA-seq kit to produce inserts in the range of 250-300 bp. Library amplification was performed on the cluster generation station of the GAIIx and using the Illumina MiSeq cluster generation kit. To obtain the paired-end reads, primers were designed to hybridized with Illumina specific adaptors resulting in reading of each end as a separate run. The sequencing reaction was run for 100 cycles (tagging, imaging and cleavage of one terminal base at a time), and four images of each tile on the chip were taken in different wavelengths for exciting each base-specific fluorophore.

For PacBio, qualified genomic DNA was fragmented using Covaris g-TUBE devices and were subsequently repaired by treating the sample with a DNA-damage repair mix. Following DNA-damage repair, blunt ends were created on each end and then hairpin adapters incorporating a unique barcode were ligated to each blunt end. The SMRTbell DNA template libraries were selected using a bluepippin system targeting a fragment size > 10 kb. Library quality was analyzed by Qubit and average fragment size was estimated using an Agilent 2100 Bioanalyzer. We used Sequel Sequencing kit 2.1 to sequence the library in PacBio Sequel platform. For the bioinformatics analysis, we first demultiplexed the PacBio subread file with Lima package, after which the demultiplexed bam file was converted to FASTA format using SAMtools FASTA. Flye was used to assemble the FASTA file with "--plasmids --iterations 2 --asm-coverage 120" parameters. The completeness of the genomics data was assessed using BUSCO11 (version 4.1.2, run in mode genome and proteome with the lineage dataset: saccharomycetes_odb10). 91% was obtained when the software was run in the genome mode and 96% when run in the proteome mode. The results obtained with BUSCO allowed us to estimate the degree of reads contamination as being below 1%.

To further curate the PacBio-assembled contigs and avoid sequencing mistakes, the reads obtained from the Illumina sequencing were mapped in the 20 contigs obtained (resulting in more than 95% mapping). The annotation of the 20 curated PacBio contigs was performed in the Geneious software framework (version 2019.2.3). Afterwards the ab initio gene detection algorithm Augusts (trained in *S. cerevisiae*, *S. pombe* and *A. nidulans*) was used to identify putative CDSs in the sequence of the contigs. Whenever the predicted gene models coincided

with CDSs previously described as belonging to the *Sd. ludwigii* UTAD17 ORFeome these were automatically validated. BLASTP analysis using the UNIPROT database as a target was used to curate and modify the gene models predicted in silico. Those gene models having an identified hit at UNIPROT were considered valid while those that didn't comply with this criterion were considered hypothetical. To obtain further information concerning the annotation, including functional categorization, the OmicsBox (version 1.1.164) framework was used.

II.3.4 Metabolic reconstruction and comparative proteomic analysis of *Sd. ludwigii* UTAD17 ORFeome and other yeast species

Metabolic reconstruction of *Sd. ludwigii* UTAD17 was performed making use of KEGG BlastKoala annotation tool (Kanehisa et al., 2016) using as a query dataset the 5,008 genes predicted in the *in silico* annotation, choosing Fungi as the taxonomic group and enabling Koala to search against the family_eukaryotes.pep KEGG database. To further improve this functional annotation the eggNOG-mapper (Huerta-Cepas et al., 2017; Jensen et al., 2008; Huerta-Cepas et al., 2019) set at the default parameters was also used. For the comparative analysis of the *Sd. ludwigii* UTAD17 ORFeome with the proteomes of *S. cerevisiae* EC1118, *H. guilliermondii* UTAD222, *H. uvarum* AWRI 3580, *H. osmophila* AWRI 3579, *T. delbrueckii* CBS 1146, and *L. fermentati* CBS 6772 pairwise BLASTp analyses were performed using the sets of proteins publicly available at UNIPROT for each strain. Two proteins from the different yeast species under analysis were considered highly similar whenever identity associated with the pairwise alignments was above 50% had an associated *e*-value below e^{-50}. Whenever protein pairwise alignments resulted in identities between 30 and 50% with an associated e-value below e^{-20}, the corresponding proteins were considered similar. In all the other cases the protein pairs were considered dissimilar. In order to assess genetic relatedness between the different strains used in the comparative proteomic analysis the IST sequence of these strains (and also of others belonging to the same species and available at NCBI) was aligned using MUSCLE (Edgar, 2004) and used for phylogenetic distance analysis on MEGA X (Tamura et al., 2021). For this, the maximum likelihood method was used and general time reversible model chosen based on Neighbour-Join and BioNJ algorithms applied to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach. The phylogenetic analysis was performed using default parameters, and a bootstrap method analysis with 250 replications. The ITS sequence of *Schizosaccharomyces pombe* 972 h was used as an outgroup.

II.3.5 Morphological studies of *Sd. ludwigii* in the presence of N-Acetylglucosamine

Sd. ludwigii UTAD17 and *C. albicans* SC5314 cells were overnight grown in liquid YPD medium, at 30ºC with orbital agitation of 250 rpm. Following this period, cells were inoculated into MM with an optical density of OD_{600}=0.1 and incubated in the same conditions. After reaching the exponential state (~3h), cells were harvested and inoculated at OD_{600}=0.1 into different culture medium: i) MM; ii) MM + 50mM N-Acetylglucosamine, and incubated at two different temperatures: i) 30ºC and ii) 37 ºC with orbital agitation (250 rpm), for 18 hours. After this period, cells were observed under an optical microscope, with 100x magnification.

II.4 Results and Discussion

II.4.1 Overview on the genomic sequence of *Sd. ludwigii* UTAD17 and on the corresponding functional annotation

In order to have a suitable portrait of the genomic architecture of the *Sd. ludwigii* UTAD17, strain karyotyping based on PFGE was performed (**Figure II. 2**). The results obtained revealed seven clearly separated chromosomal bands, ranging from 0.9 Mbp to 2.9 Mb, totaling 13.75 Mbp (**Figure II. 2**).

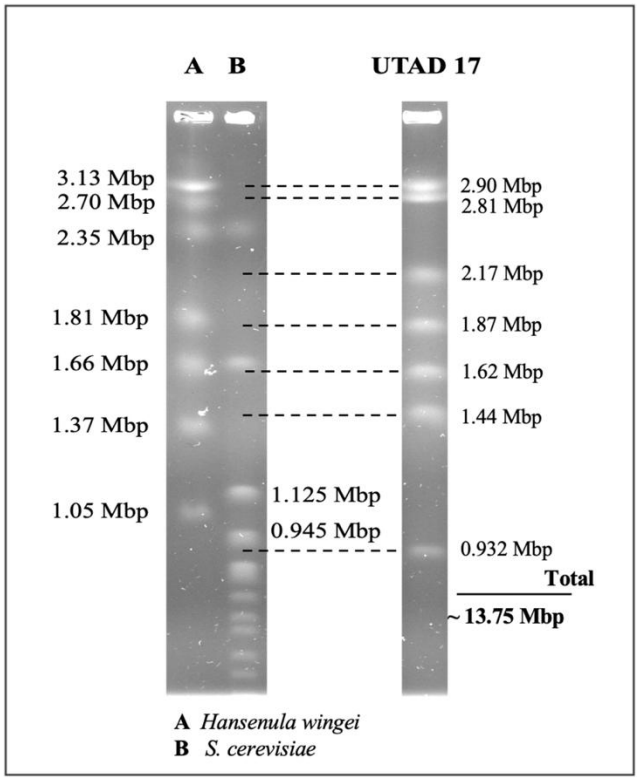


Figure II. 2- Karyotyping of *Saccharomycodes ludwigii* UTAD17, based on PFGE. Total DNA of *Sd. ludwigii* UTAD17 was separated by PFGE, as detailed in materials and methods. In the end of the run 7 clearly separated bands, presumed to correspond to the 7 chromosomes of *Sd. ludwigii* UTAD17, were obtained. Molecular sizes of these chromosomes was estimated based on the migration pattern obtained for the chromosomal bands from *Hansenula wiingei* (lane A) and *Saccharomyces cerevisiae* BY4741 that were used as markers (lane B).

This number of chromosomes and their size range is consistent with what was previously described for other *Sd. ludwigii* strains (Yamazaki & Oshima, 1996) and is also in line with what is reported for other members of the *Saccharomycodaceae* family (Esteve-Zarzoso et al., 2001; Cadez et al., 2002; Seixas et al., 2019). Sequencing with Illumina generated 20,333,547 reads of 250 bp on average and were assembled *de novo* into 1,360 contigs. These contigs were obtained after a filter to have a coverage above 300× and a size above 1,000 nucleotides, and resulted in N50 length of 17,540 bp. The sum of the assembled

contigs totaled 10,785,241 bp (**Table II. 1**). Sequencing with PacBio generated 585,118 reads (with a 445.3 coverage) which were *de novo* assembled into 20 contigs and an assembled genome of 12,999,941 bp (**Table II. 1).** This assembled genome size corresponds to approximately 95% of the estimated size this strain, based on the karyotyping shown above in **Figure II. 2**.

Table II. 1 – Sequencing results of *Sd. ludwigii* UTAD17 genome generated by Illumina MiSeq and PacBio.

	UTAD17 Illumina sequencing summary	**UTAD17 PacBio sequencing summary**
Number of reads	20,333,547	585,118
Number of contigs	1,360	20
N50	17,540 bp	1.48 Mbp
Assembled genome	10,785,241 bp	12,999,941 bp

Using the gathered genomic information from *Sd. ludwigii* UTAD17, *in silico* annotation was performed exploring results provided by different algorithms used for *ab initio* gene detection, afterwards subjected to an exhaustive manual curation. Using this approach 5,033 protein-encoding genes (CDS) were predicted in the genome of *Sd. ludwigii* UTAD17, out of which 4,644 are believed to encode canonical protein-encoding genes and 389 were considered putative genes since upon BLAST against the UNIPROT database no hit was found (details are provided in **Appendix Table II.1**). The putative CDSs were distributed throughout 17 of the 20 assembled contigs with genes not being detected only in contigs 14, 16 and 19 (**Appendix Table II.2**). Contigs 14 and 19 share high similarity (above 95% at the nucleotide level) with described mitochondrial DNA from other *Sd. ludwigii* strains, for which we anticipate these correspond to portions of UTAD17 mitochondrial DNA.

To get a more functional view of the *Sd. ludwigii* UTAD17 ORFeome all the predicted proteins were organized into biological functions using for that the eggNOG-mapper, a tool that enables functional annotation using COG categories (Huerta-Cepas et al., 2017; Jensen et al., 2008; Huerta-Cepas et al., 2019) (**Figure II. 3**). The highest number of proteins for which it was possible to assign a biological function were clustered in the “Intracellular traficking”, “Transcription”, “Translation” and “Post-translational modification” classes (**Figure II. 3**) which is consistent with the distribution obtained for *Hanseniaspora* species and for *S. cerevisiae* (**Figure II. 3**). The number of *S. cerevisiae* genes clustered in 12 of the 21 functional COG classes surpassed those of *Sd. ludwigii* UTAD17 by approximately 20% (details provided

in **Appendix Table II.3**, an observation that is consistent with the later species being pre-whole genome duplication like the other species of the *Saccharomycodaceae* family (Boundy-Mills et al., 2011; Seixas et al., 2019; Wolfe et al., 2015). Indeed, further mining of *Sd. ludwigii* UTAD17 genome revealed traits found in pre-whole genome duplication species such as disassembly of the genes necessary for allantoin metabolism, absence of galactose catabolism genes and the lack of a functional pathway for de novo nicotinic acid biosynthesis (Wolfe et al., 2015). Furthermore, out of the 555 ohnologue pairs identified in *S. cerevisiae* (Byrne & Wolfe, 2005) we could identify homologues for 517 in the genome of *Sd. ludwigii* UTAD17, with 512 of these existing in single-copy (that is, the two ohnologues were similar to the same *Sd. ludwigii* UTAD17 protein) (**Appendix Table II.4**).

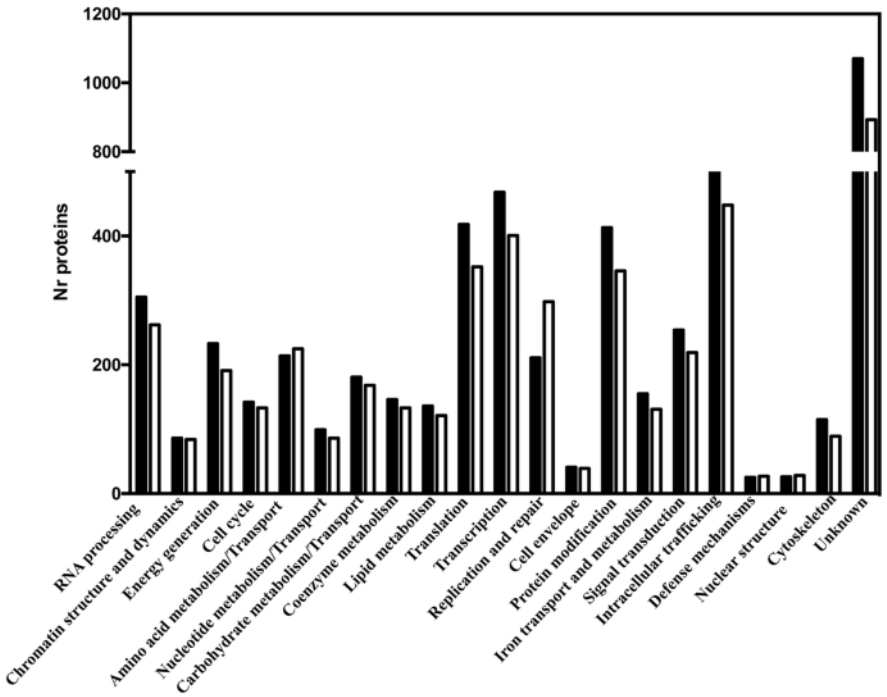


Figure II. 3- Functional categorization of the predicted ORFeome of *Sd. ludwigii* UTAD17. After annotation of the assembled genomic sequence, the validated gene models were clustered according with the biological function they are predicted to be involved in (using COG functional categories) using the eggNOG-mapper tool (black bars). As a comparison, the distribution of the *S. cerevisiae* proteome is also shown (white bars).

II.4.2 Comparative analysis of the predicted proteomes of *S. ludiwgii* with members of the *Saccharomycetaceaee* and *Saccharomycodaceae* families.

The get further hints into the physiology of *Sd. ludwigii* the predicted ORFeome of the UTAD17 strain was compared with the one predicted for *H. guilliermondii*, *H. uvarum* and *H. osmophila*, these representing three species of the *Saccharomycodaceae* family with an available annotated genomic sequence. Three *Saccharomycetaceae* species with relevance in the wine environment were also included in this comparative analysis: *Lachancea fermentati*, *Torulaspora delbrueckii* and the *S. cerevisiae* wine strain EC1118 (**Figure II.4**).

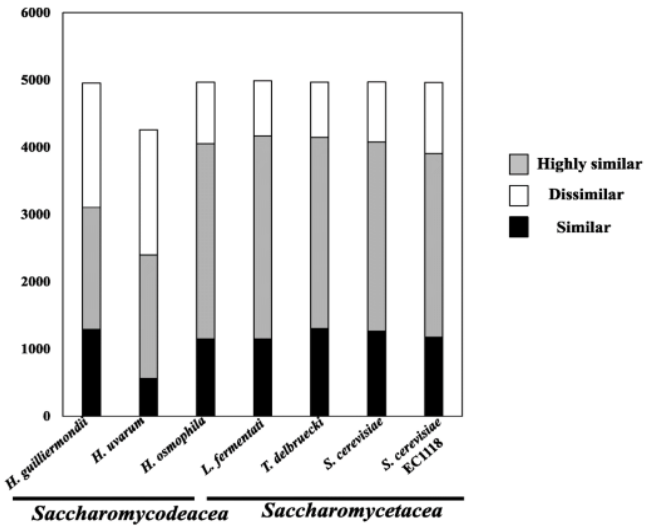


Figure II. 4– Comparative analysis of the predicted proteome of the *Saccharomycodaceae* species *Sd. ludwigii*, *H. guilliermondii*, *H. uvarum* and *H. osmophila*. The ORFeome predicted for *Sd. ludwigii* UTAD17 strain was compared with the one of the *Hanseniaspora* species that also belong to the *Saccharomycodaceae* family using pairwise BlastP alignments. Three species belonging to the *Saccharomycetaceae* family with relevance in the wine environment, *S. cerevisiae*, *L. fermentati* and *T. delbrueckii* were also included in this comparative analysis. The graph shows the number of *Sd. ludwigii* proteins highly similar (e-value below or equal to e^{-20} and identity above 50%), similar (e-value below or equal to e^{-20} and identity between 30 and 50%) or dissimilar (e-value above e^{-20}) from those found in the other yeast species considered.

The *Sd. ludwigii* UTAD17 ORFeome showed the highest degree of similarity with *L. fermentati*, *T. delbrueckii* and *H. osmophila*, while similarity with the predicted proteomes of *H. uvarum* and *H. guilliermondii* was considerably smaller (**Figure II. 4**). This observation was surprising but is in line with the results obtained by phylogenetic analysis of the the ITS sequence of the strains used in this comparative proteomic analysis that also shows a higher divergence of *H. guilliermondii* and *H. uvarum* species within the *Saccharomycodaceae* family (**Appendix Figure II.1**). *H. osmophila* was described to have phenotypic traits similar to those exhibited by *Sd. ludwigii*, including the ability to survive in high sugar grape musts or reasonable fermentative capacity (Granchi et al., 2002; Viana et al., 2008) two traits not observed for *H. uvarum* or *H. guilliermondii*. Similarly, *L. fermentati*, formerly described as *Zygosaccharomyces fermentati* (Kurtzman & Robnet, 2003), also shares phenotypic traits with *Sd. ludwigii* including tolerance to SO_2 and ethanol and the ability to grow on grape musts or wines with high residual sugar content (Bellut et al., 2020). Thus, it is possible that the observed higher similarity of the *Sd. ludwigii* with *H. osmophila*, *L. fermentati* and *T. delbrueckii* proteomes can result from the evolution of similar adaptive responses to the challenging environment of wine musts, not reflecting their phylogenetic relatedness. In this context, it is intriguing why *H. guilliermondii* and *H. uvarum* are apparently so divergent considering they

are also present in grape musts. To capture more specific features of the *Sd. ludwigii* species, the proteins considered dissimilar from those found in the set of yeasts used for the comparative proteomic analysis were compared resulting in the Venn plot depicted in **Figure II. 5**.

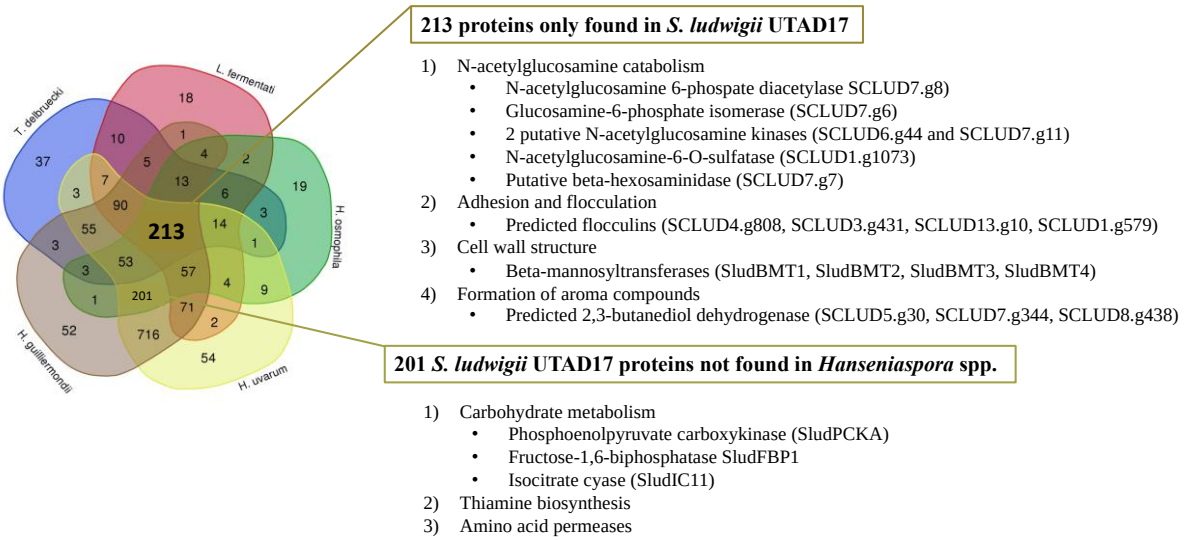


Figure II. 5- The *Sd. ludwigii* UTAD17 proteins found to be dissimilar from those found in the other yeast species were compared and the results are shown in the Venn plot. In the picture are highlighted the 213 proteins that were unique of *Sd. ludwigii*, as no robust homologue could be found in any of the other yeast species considered. Additionally, 201 proteins of *Sd. ludwigii* that were found in the *Saccharomycetaceae* species but not in the other *Saccharomycodaceae* species. Some of the functions represented in these two protein datasets are highlighted in this picture.

This analysis identified 213 proteins that were only found in *Sd. ludwigii* (detailed in **Appendix Table II.5**). This set included six enzymes required for catabolism of N-acetylglucosamine (GlcNAc) into fructose 6-phosphate including a N-acetylglucosamine-6-phosphate deacetylase (SCLUD7.g8), a glucosamine-6-phosphate isomerase (SCLUD7.g6) and two putative N-acetylglucosamine kinases (SCLUD6.g44 and SCLUD7.g11) (**Figure II. *5***, **Figure II. 7** and **Appendix Table II.5**). A predicted N-acetylglucosamine permease (SCLUD1.g377) was also identified in the genome of *Sd. ludwigii* UTAD17, however, this was also present in the genome of the other four yeast species considered. The set of *Sd. ludwigii* specific proteins also included a protein weakly similar to a described bacterial N-acetylglucosamine-6-O-sulfatase (SCLUD1.g1073) and a putative beta-hexosaminidase (SCLUD7.g7), these two enzymes being required for catabolism of polysaccharides harboring GlcNAc like heparin sulphate (**Figure II. *5*** and **Appendix Table II.5**). In yeasts GlcNAc metabolism has been essentially described in dimorphic species like *Candida albicans* or *Yarrowia lypolytica*, where it serves as a potent inducer of morphological transition (Novotný et al., 1994). Recently, the ability of *Scheffersomyces stipitis* to consume GlcNAc was

described enlarging the panoply of GlcNAc consuming yeasts to non-dimorphic species (Passoth et al., 1996).

Aiming to uncover why GlcNAc catabolism is present in *Sd. ludwigii*, but not in the other *Saccharomycodaceae* species, we grew UTAD17 in minimal medium supplemented with 50 mM of GlcNAc. This assay aimed to determine if, similarly to yeasts like *C. albicans*, GlcNAc could induce hyphae formation in *Sd. ludwigii* UTAD17. We first tested *Sd. ludwigii's* filamentation capacity at 30ºC, using *C. albicans* SC5314 as a positive control due to its well-established role of GlcNAc in inducing filamentation in this species (Naseem et al., 2011; Su et al., 2018; Rao et al., 2021). At this temperature, no hyphae formation was observed in UTAD17, although hyphae structures began to appear in *C. albicans* (**Figure II. 6**). Increasing temperature to 37ºC, led to *Sd. ludwigii* UTAD17 cells becoming elongated and developing what looked like hyphae-like structures, suggesting a positive relationship between GlcNAc catabolism and dimorphic growth in *Sd. ludwigii* (**Figure II. 6**). In *C. albicans*, GlcNAc catabolism is closely associated with pathogenic and commensal relationships, as GlcNAc serves as a major carbon source that supports *C. albican's* growth in humans and other mammals (Liu et al., 2013; Yang et al., 2023). It also functions as a powerful inducer of morphological transitions such as hyphae development, being intimately linked to the yeast's virulence and infection of host niches (Liu et al., 2013; Yang et al., 2023). No reports of *Sd. ludwigii* associated to human infections has been described, although it is frequently isolated from the gut of insects (Fogleman et al., 1982; Stefanini, 2018), where the ability to degrade this carbon source could provide an advantage for effective colonization of this host. Additionally, the ability of *Sd. ludwigii* to use GlcNAc as a carbon source will likely provide an important advantage in the competitive environment of wine musts (Liu et al., 2013; Yang et al., 2023).

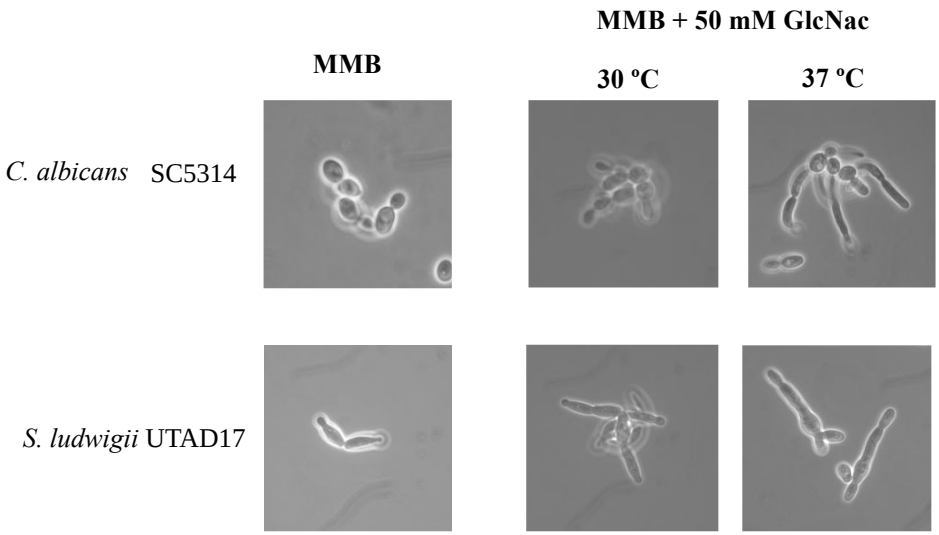


Figure II. 6– N-acetylglucosamine promotes hyphae formation in *Sd. ludwigii* UTAD17. Like in *C. albicans*, in *Sd. ludwigii* UTAD17 is also possible to observe the formation of hyphae-like structures, at 37ºC, suggesting that GlcNAc is capable of inducing dimorphic growth in this species

A set of proteins with a predicted function in adhesion and flocculation also emerged among the set of *Sd. ludwigii*-specific proteins (**Figure II. 5**). The ability of *Sd. ludwigii* to cause cloudiness in bottled wines has been described as well as its ability to grow on biofilms or to flocculate even when growing in synthetic growth medium (Loureiro & Malfeito-Ferreira, 2003). Further investigations should focus on what could be the role played by these flocullins/adhesins in the aggregation and ability of *Sd. ludwigii* to form biofilms considering that they are considerably different from the flocullins/adhesins found in the closely related yeast species. A particularly interesting aspect will be to investigate whether these adhesins mediate *Sd. ludwigii* adherence to the abiotic surfaces of cellars or to cellar equipment.

II.4.3 Metabolic reconstruction of *Sd. ludwigii* UTAD17

To reconstruct the *Sd. ludwigii* metabolic network, the ORFeome predicted for this strain was used as an input for the BlastKOALA tool (Kanehisa et al., 2016) resulting in the schematic representation shown in **Figure II. 7** (the corresponding functional distribution is shown in **Appendix Figure II.2** while in **Appendix Table II.6** are provided further details about the genes clustered in each of the metabolic pathways). This analysis shows that *Sd. ludwigii* UTAD17 is equipped with all the genes of the main pathways of central metabolism including the pentose phosphate pathway, glycolysis, gluconeogenesis, Krebs cycle and oxidative phosphorylation, besides the already discussed capacity to use GlcNAc (**Figure II. 7**; the identity of enzymes associated to the different enzymatic steps shown in the metabolic

map are provided in **Appendix Table II.6**). The fact that *Sd. ludwigii* UTAD17 is equipped with neoglucogenic enzymes, an isocitrate lyase and all the enzymes required for biosynthesis of thiamine are marked differences from what is observed in *Hanseniaspora* species (Seixas et al., 2019) (**Figure II. 7**; **Figure II. 5** and **Appendix Table II.6**). Considering the critical role of thiamine in driving fermentation, the fact that *Sd. ludwigii* cells are able to biosynthesize it can be responsible for the higher fermentative capacity of these cells, compared with *Hanseniaspora* spp. that are auxotrophic for this vitamine (Seixas et al., 2019; Steenwyk et al., 2019). A closer look into the genes involved in thiamine biosynthesis in *Sd. ludwigii* UTAD17 revealed that this yeast encodes 10 enzymes required for conversion of histidine and pyridoxal-phosphate into the thiamine precursor hydroxymethylpyrimidine diphosphate (HMP-P), three enzymes for conversion of HMP-P into HMP-PP and four predicted thiamine transporters (**Figure II. 7**). This is interesting since in *L. fermentati* and in *T. delbrueckii* we could only identify one enzyme for each of the different enzymatic steps required for biosynthesis of 3-HMP-PP, similar to what is reported for *Kluveromyces lactis*, *K. thermotolerans* or *Saccharomyces kluyveri* (Wightman & Meacock, 2003) (**Figure II. 7** and **Appendix Table II.6**). In fact, until thus far the expansion of enzymes involved in synthesis of 3-HMP-PP has been described as a specific feature of the *Saccharomyces* sensu strictu species that harbours 3 enzymes for the synthesis of 3-HMP-P (Thi5, Thi11, Thi12 and Thi13) and two for the synthesis of 3-HMP-PP (Wightman & Meacock, 2003). The amplification of only the 3-HMP-P branch, but not the HET branch (which provides the other precursor for thiamine biosynthesis; **Figure II. 7**) is noteworthy (**Figure II. 7**). 3-HMP-P has only been described as an intermediate of thiamine biosynthesis and, thus, it is not clear the outcome that might be obtained by *Sd. ludwigii* cells with the expansion of enzymes involved in synthesis of this precursor. In the case of *S. cerevisiae* the expansion of genes producing 3-HMP-P was proposed to assure proper channelling of pyridoxine to thiamine biosynthesis avoiding depletion in biosynthesis of amino acids (Wightman & Meacock, 2003). Genes required for catabolism of lactate, mannose, sucrose, raffinose and starch were also found in the genome of UTAD17 (**Figure II. 7**, **Appendix Table II.6**) consistent with the demonstrated ability of *Sd. ludwigii* UTAD17 and other strains of this species to grow on these sources (results not shown). As said above, *Sd. ludwigii* UTAD17 strain is not equipped with genes allowing catabolism of galactose and we could also not detect genes for catabolism of lactose, this observation being in line with the reported inability of this species to grow on these carbon sources (Boundy-Mills et al., 2011).

Concerning nitrogen metabolism, all the genes required for synthesis of proteogenic amino acids, synthesis and degradation of GABA and for conversion of amino acids into higher alcohols through the Ehrlich pathway were found in the *Sd. ludwigii* UTAD17 ORFeome (**Appendix Table II.6** ; **Appendix Table II.7** and **Figure II. 7**). No genes encoding enzymes for the synthesis of spermidine, spermine or putrescine, or for biosynthesis of methionine through the salvaging pathway (the main source of precursors for the biosynthesis of polyamines) were found in the genome of UTAD17, similar to what was observed for *Hanseniaspora* species (Seixas et al., 2019; Steenwyk et al., 2019). Although this observation is intriguing, considering that polyamines, and specially spermidine, plays a detrimental role in mediating growth in *S. cerevisiae* (Chattopadhyay et al., 2003), it is in line with previous reports of the inability of the UTAD17 strain, and of *Sd. ludwigii* species in general, to produce biogenic amines (which are produced from polyamines) (Esteves et al., 2019; Ivit et al., 2018). Another noticeable difference between *Sd. ludwigii* UTAD17 and the *Hanseniaspora* species was the observation that *Sd. ludwigii* is equipped with specific permeases for methionine, GABA, histidine, proline, glutamine, lysine, arginine, choline, isoleucine/valine/isoleucine, besides encoding five putative general amino acid permeases while *Hanseniaspora* encodes only two specific amino acid permeases but thirteen general amino acid permeases (**Figure II. 7**and **Appendix Table II.6**) (Seixas et al., 2019).

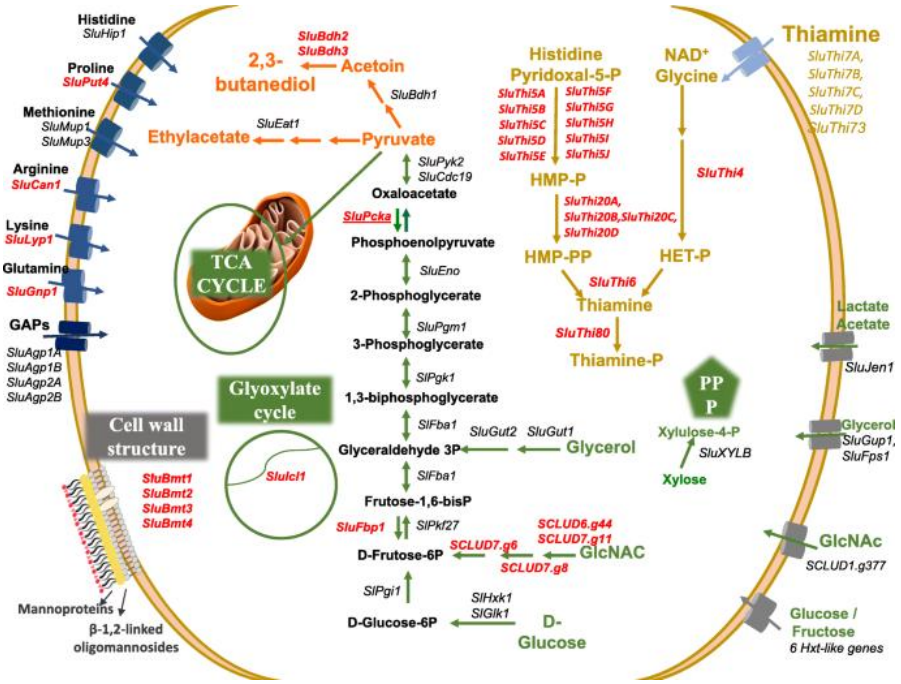


Figure II. 7- Schematic overview on the central carbon and nitrogen metabolic networks of *Sd. ludwigii* UTAD17. The predicted ORFeome of *Sd. ludwigii* was used as an input in the metabolic networks reconstruction tools eggNOG-mapper and KEEG Koala to gather a schematic representation of the metabolic pathways linked to central carbon and nitrogen metabolism active in *Sd. ludwigii* UTAD17. The picture schematically represents some of the active pathways identified in this *in silico* analysis, emphasizing in red proteins that were found in *Sd. ludwigii* but in other *Saccharomycodaceae*. This schematic representation is original and was specifically prepared by the authors to be presented in Tavares et al., (2021).

II.4.4 The predicted FLAVOROMA genes in *Sd. ludwigii* UTAD17

One of the aspects for which *Sd. ludwigii* is considered to have strong biotechnological potential is its use in tailored flavour-fermented beverages (Romano et al., 1999; Esteves et al., 2019; Ivit et al., 2018; Holt et al., 2018; Michel et al., 2016; De Francesco et al., 2015). In this context, we searched the UTAD17 ORFeome for genes predicted to be involved in formation of volatile aroma compounds, with the more relevant aspects of this analysis being summarized in the metabolic map shown in **Figure II. 7** and further detailed in **Appendix Table II.7**. *Sd. ludwigii* UTAD17 is equipped with enzymes for biosynthesis of ethyl esters, namely ethyl acetate (**Figure II. 7** and **Appendix Table II.7**). In specific, we could identify two alcohol acetyl-transferases in the genome of the UTAD17 strain, SCLUD4.g700, moderately similar to *S. cerevisiae* Eat1 and SCLUD6.g215, weakly similar to the *Kluveromyces lactis* KlEat1 and *Wickerhamomyces anomalus* WaEat1 (**Figure II. 7** and **Appendix Table II.7**). Eat1 from *W. anomalus* and *K. lactis* were recently described as part of a novel family of alcohol acetyltransferases essentially responsible for ethyl acetate production (Kruis et al., 2017). No orthologues of ScAtf1 or ScAtf2, two other alcohol acetyltransferases that in *S. cerevisiae* also contribute for synthesis of ethyl acetate, were identified in the ORFeome of *Sd. ludwigii* UTAD17. Similarly, no orthologues for ScAtf1 or ScAtf2 were found in the genome of *Hanseniaspora* species whose ability to produce ethyl esters (a trait for which these species are particularly known for –*e.g*., *H. guilliermondii* - Lage et al., 2014), was hypothesized to result from the activity of a specific set of putative set alcohol acetyl transferases only found in the genomes of these species (Seixas et al., 2019) and whose functional characterization is being pursued in our laboratory. Strikingly, although the UTAD17 strain is equipped with enzymes leading to synthesis of ethyl esters and *Sd. ludwigii* is known for high capability to produce ethyl esters (Romano et al., 1999; Michel et al., 2016; Domizio et al., 2011; Ciani & Maccarelli et al., 1997), during fermentation of natural grape juice the production of these volatiles by UTAD17 cells was almost negligible, even below the one exhibited by *S. cerevisiae* (Esteves et al., 2019). Further investigations will have to be performed to better understand this observation specially focusing whether this trait is specific of the UTAD17 strain (which by some reason could have the activity of alcohol acetyl-transferase enzymes impaired) or whether this resulted from the composition of the grape juice used in the fermentations that could be less favourable for production of ethyl esters by *Sd. ludwigii* cells. *Sd. ludwigii* is also known for its ability to produce 2,3-butanediol and acetoin (Romano et al., 1999; De Francesco et al., 2015; Romano et al., 1998). Consistently, three predicted 2,3-butanediol dehydrogenases

(SCLUD5.g30, SCLUD7.g344 and SCLUD8.g438) are found in the genome of the UTAD17 strain (**Figure II. 7** and **Appendix Table II.7**). Enzymes for the production of higher alcohols, isoamyl alcohols and a putative beta-glucosidase were also detected in the genome of the UTAD17 strain (**Appendix Table II.7)**.

The retention of aromatic compounds in wines (namely fruity esters) has been linked to increased content of mannoproteins, which have also been found to increase mouthfeel, provide protection against protein and tartaric instability and reduced astringency (Caridi et al., 2006; Domizio et al., 2014). *Sd. ludwigii* is known for its enhanced ability to excrete mannoproteins (Esteves et al., 2019; Domizio et al., 2011; Giovani et al., 2012). Further characterization of these mannoproteins released by *Sd. ludwigii* showed a very high content (above 90%) of mannose suggesting that hyper-mannosylation of these proteins released from the cell wall occurs (Giovani et al., 2012). Notably, mining of the *Sd. ludwigii* UTAD17 genome allowed us to identify five putative beta-mannosyltransferases (SludBMT1, SludBMT2, SludBMT3 and SludBMT4) with strong homology to beta-mannosyltransferases described in *Candida albicans*, *C. glabrata* or *Pichia pastoris* (Mille et al., 2008; Krainer et al., 2013) (**Appendix Table II.8** and **Figure II. 7**). In those species, these enzymes mediate the incorporation of beta-1,2-linked oligomannosides in the cell wall (a unique feature since in *S. cerevisiae* these are α-1,4-linked) and to promote hyper-mannosylation of secreted proteins (Mille et al., 2008; Krainer et al., 2013). No orthologues of these enzymes could be found in *H. uvarum* or in *H. guilliermondii*, while *H. osmophila* appears to encode only one of these mannosyltransferases (**Appendix Table II.8**). The presence of these beta-mannosyltransferases in the genome of *Sd. ludwigii* UTAD17 strongly suggests that the composition of the cell wall should be quite different from the one found in other *Saccharomycodaceae* and in *S. cerevisiae*, which can affect the function of this structure as a selective barrier (for example, against SO_2, as it will be discussed below).

II.4.5 Elucidating *Sd. ludwigii* stress responses relevant in the context of wine fermentation: emphasis on tolerance to sulfur dioxide

The fact that contamination by *Sd. ludwigii* is observed either in sulfited grape-musts or in stabilized wines (Romano et al., 1999; Barata et al., 2010; Loureiro & Malfeito-Ferreira, 2003) indicates that this yeast is equipped with means to survive the harsh environment of vinification which include, among others, the high concentration of sugars present in the beginning of the fermentation or the high concentrations of ethanol obtained in the end, besides

the presence of inhibitory concentrations of SO_2. Although not much is known concerning how wine Non-*Saccharomyces* Yeasts respond to environmental stress, a lot of knowledge was gathered on the field in *S. cerevisiae* including the genes necessary for tolerance to ethanol (van Voorst et al., 2006; Yoshikawa et al., 2009), to high concentrations of glucose (Teixeira et al., 2010) or even those required for this species to thrive in oenologically relevant conditions (the so-called fermentome - Walker et al., 2014). We have searched the genome of UTAD17 for orthologues of these sets of stress-tolerance genes described in *S. cerevisiae* and found that *Sd. ludwigii* UTAD17 harbours around 40–45% of those genes, as detailed in **Appendix Table II.9**. This percentage is considerably higher than the one found in *H. uvarum* or *H. guilliiermondii* that only harboured 33% of the ethanol-tolerance genes and 35% of the "fermentome" genes (Seixas et al., 2019). A closer comparison revealed that *Sd. ludwigii* UTAD17 encodes all the "Sc fermentome"-genes identified in *Hanseniaspora* and 37 additional others that were not found in *Hanseniaspora* (**Appendix Table II.9**). *Sd. ludwigii* UTAD17 was also found to encode 226 *S. cerevisiae* ethanol-resistance genes that could not be identified in its sister species *Hanseniaspora* (details in **Appendix Table II.9**). Although these numbers need to be taken carefully because the proteins might have suffered divergent evolutive paths in *S. cerevisiae* and in the other yeasts and therefore the absence of an homologue does not necessarily mean that a functional orthologue is absent, the higher presence of stress genes in *Sd. ludwigii* UTAD17 is consistent with the described increased resilience of this species to wine-related stresses, compared with *Hanseniaspora* species. From the analysis performed it stood out that a large cohort of peroxisomal genes and mitochondrial proteins involved in translation and in the respiratory chain are present both in *S. cerevisiae* and in *Sd. ludwigii* but are absent from *Hanseniaspora* genomes (**Appendix Table II.6**). Although the molecular mechanism by which the peroxisomal function contributes for tolerance to ethanol in *S. cerevisiae* could not be uncovered until thus far, it was clear that when cells are challenged with toxic concentrations of ethanol a proliferation of these organelles occurs (Teixeira et al., 2009), which is also consistent with the protective effect exerted by PEX genes (Yoshikawa et al., 2009; Teixeira et al., 2009) and with their reported up-regulation under ethanol stress (Stanley er al., 2010). Further investigations will be required to confirm whether or not peroxisomal function plays a role in mediating tolerance to ethanol in *Sd. ludwigii.*

Sd. ludwigii UTAD17 does not encode a clear orthologue for the transcription factors Msn2 and Msn4 (Gasch et al., 2000), responsible for the control of the environmental stress

response in *S. cerevisiae* and that also play a role in response of this species to vinification conditions (Cardona et al., 2007; Marks et al., 2008). Closer mining of the *Sd. ludwigii* genome allowed us to identify one protein (SCLUD3.g330) that shows similarity at the level of the C-terminal domain (67% identity) to the C-terminus of ScMsn2 and ScMsn4, the region that comprises the DNA binding domain of these regulators (Görner et al., 2002) (**Appendix Figure II.3**). Due to evolution, transcription factors tend to conserve their homology across more distant species essentially at the DNA binding domain, while transactivation domains are largely more variable. In *Hanseniaspora* species no orthologue of ScMsn2/ScMsn4 (Seixas et al., 2019) nor this was also found in the *Saccharomycetaceaee* species *S. pombe* or *S. kluyverii* (Brion et al., 2016). This does not mean that *Hanseniaspora* or *Sd. ludwigii* do not mount an environmental stress response with some similarities to the one described in *S. cerevisiae* since such a response was described to occur in *S. pombe* or *S. kluyverii* (Brion et al., 2016), albeit the absence of a Msn2/Msn4 clear orthologue.

Tolerance to SO_2 in *S. cerevisiae* has been largely attributed to the activity of the sulfite export pump Ssu1 (Lage et al., 2019; Avram & Bakalinsky, 1997; Nadai et al., 2016), which was also found to influence tolerance to this preservative in the more tolerant strain *Brettanomyces bruxellensis* (Varela et al., 2019). Notably, the genome of *Sd. ludwigii* UTAD17 was found to encode four genes with a strong similarity (above 45% at the amino acid level) with the efflux pump ScSsu1: SCLUD1.g608, SCLUD1.g608b, SCLUD1.g612, SCLUD1.g612b) (**Appendix Figure II.4**). These four genes are arranged in tandem and appear as two duplicated pairs with SCLUD1.g608 and SCLUD1.g612 showing higher similarity to ScSsu1 (47.5% identity) and SCLUD1.g608b, SCLUD1.g612b showing a lower similarity due to a premature STOP codon that renders the proteins shorter at the C-terminal region (**Figure II. 5**and **Appendix Figure II.4**). This is very interesting since it is the first time that a candidate sulfite export system is described in *Sd. ludwigii*. Our preliminary results from a transcriptomic analysis undertaken in SO_2-challenged *Sd. ludwigii* cells confirms that the four *SSU1* genes are transcribed (albeit the shorter versions at a considerably lower extent than the larger ones) and their transcription is augmented upon exposure to SO_2 (results not shown). Further studies will be required to investigate the individual role of these four "Ssu1-like" pumps in determining tolerance to SO_2 in *Sd. ludwigii*. More SO_2-tolerant *B. bruxellensis* strains were found to encode alleles with higher activity of SSU1 than those encoded by less susceptible strains, this being attributed to the existence of point mutations that result in increased activity of the pumps (Varela et al., 2019). In this sense, it will also be interesting to investigate if a similar trait is

observed in the case of the *SSU1* genes encoded by *Sd. ludwigii* considering that a strong inter-strain variability concerning tolerance to SO_2 has also been described (Vejarano et al., 2018). In *S. cerevisiae* transcription of *SSU1* is largely dependent of the transcriptional activator Fzf1 (Avram et al., 1999) but we could not identify an orthologue for this regulator in the genome of *Sd. ludwigii* (nor in *B. bruxellensis*) suggesting that in this species this sulfite efflux pump could be under the control of a different regulatory circuit.

Recently our group has performed a genome-wide phenotypic screening that identified around 200 genes required for tolerance to SO_2 in *S. cerevisiae* expanding the set of resistance determinants to this preservative well beyond Ssu1 (Lage et al., 2019). Among these newly identified SO_2-resistance genes was the Com2 transcription factor, an orphan homologue of Msn2, which was identified as being critical not only for tolerance but also for the reprogramming of *S. cerevisiae* transcriptome in response to SO_2 stress (Lage et al., 2019). Although the SCLUD3.g330 regulator discussed above shows more homology with ScMsn2 than with ScCom2 (**Appendix Figure II.3**), it still remains the question of whether or not this regulator could mediate response and/or tolerance of *Sd. ludwigii* to SO_2. Around 65% of the other SO_2-resistance genes identified in *S. cerevisiae* had a robust orthologue in the ORFeome of *Sd. ludwigii* UTAD17 including the genes mediating the sulfur assimilation pathway (*e.g.*, *MET14*, *MET16* and *MET3*) or the genes involved in biosynthesis of lysine and arginine (**Figure II. 8** and **Appendix Table II.10**).

Another aspect that might also influence the extreme tolerance to SO_2 of *Sd. ludwigii* is the different structure of the cell wall which, as discussed above, is likely to be enriched in β-1,2-mannosides due to the presence of β-mannosyltransferases. This modification could result in a lower permeability of *Sd. ludwigii* cell wall to SO_2 and thus explain the observed reduced diffusion rate into the inside of these cells, compared to the one observed for *S. cerevisiae* (Stratford et al., 1987; Stratford & Rose, 1986). Since ethanol tolerance has also been demonstrated to depend on diffusion across the cell wall, it is possible that this anticipated difference in the cell wall of *Sd. ludwigii* can also contribute for its higher tolerance to ethanol, especially when compared to its closely related *Hanseniaspora* species. In **Figure II. 8** we have schematically represented the features that might influence tolerance to SO_2 in *Sd. ludwigii*, as uncovered by the herein described genomic analysis. What can be the individual contribution of these different players for the overall phenotype exhibited by this species and how they might determine intra-strain variability, is something that will need future studies focused on the herein uncovered candidates.

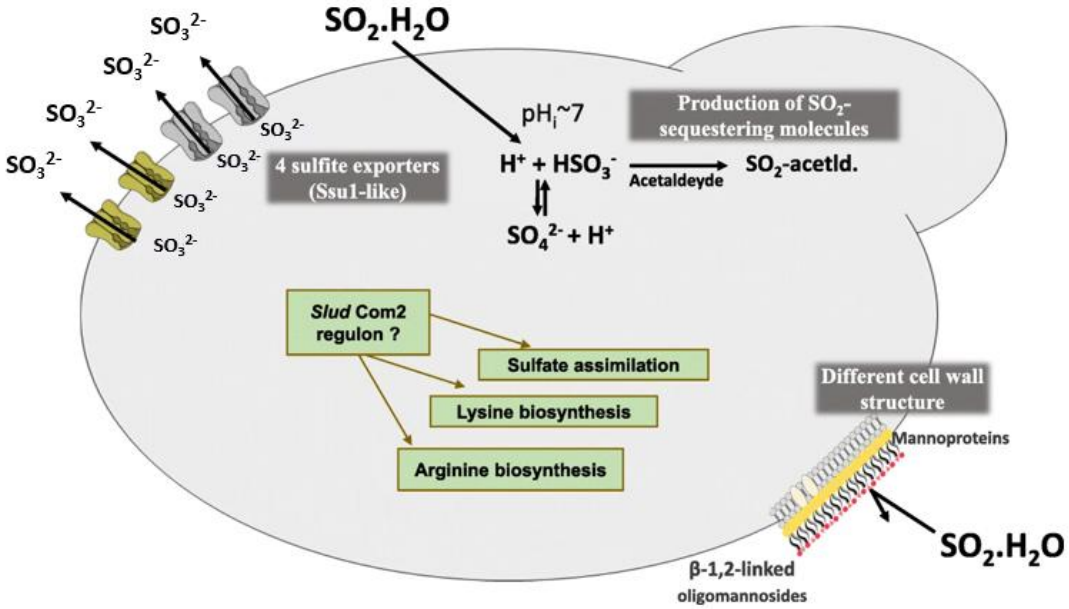


Figure II. 8- Mechanisms suggested to contribute for tolerance to SO_2 in *Sd. ludwigii*, as suggested by mining the genome of the UTAD17 strain. The candidate players that might contribute for the enhanced tolerance to SO_2 exhibited by *Sd. ludwigii* cells, as suggested by mining of the genome of the UTAD17 strain, were selected and are herein highlighted. Besides the four predicted sulfite exporters with similarity with the sulfite export pump Ssu1 from *S. cerevisiae*, orthologues for genes that been found to mediate tolerance to SO_2 in *S. cerevisiae* such as genes involved in biosynthesis of lysine and arginine or the genes involved in the sulfate assimilation pathway, are also indicated. The eventual involvement of a putative Com2-regulatory pathway in *Sd. ludwigii* is also hypothesized, based on the existence of a transcription factor with some degree of similarity to this crucial SO_2-determinant in *S. cerevisiae* (see details for further discussion in the text). It is also hypothesized whether the presumed different structure of the *Sd. ludwigii* cell wall, resulting from this species harbouring a set of mannoproteins and a different structure of the β-glucan (compared to the one exhibited by in other *Saccharomycodaceae* species and by *Saccharomycetaceae*) can contribute for the reported reduced diffusion of SO_2 into the inside of *Sd. ludwigii* cells (Stratford et al., 1987). This schematic representation was adapted from Tavares et al., (2021).

II.5 Conclusions

In this work we have deepened the genomic sequence and annotation of the wine spoilage species *Sd. ludwigii* UTAD17 shedding light into relevant aspects of the biology and physiology of this species such as its high resilience to the wine preservative SO_2 or its resilience to thrive in the challenging environment of the wine must. Compared to its sister *Hanseniaspora* species, that also belong to the *Saccharomycodaceae* family, significant differences were observed including functional neoglucogenesis, glyoxylate and thiamine pathways or a different cell wall structure. We have also unravelled aspects that might render a suitable exploration of *Sd. ludwigii* as an interesting microbial cell factory either as a co-adjuvant in production of fermentation beverages, or in production of aroma compounds (*e.g.*, ethyl acetate) or of biofuels (*e.g.*, production of bioethanol using the N-GlcNAc enriched renewable source chitin).

Chapter III.

Comparative genomics of *Saccharomycodes ludwigii* strains unveils relevant aspects of the physiology of this species, and adaptive traits underlying tolerance to sulfur dioxide

Part of the results shown in this chapter was published in:

Esteves M, Barbosa C, Vasconcelos I, Tavares MJ, Mendes-Faia A, Mira NP and Mendes-Ferreira A. "Characterizing the potential of the non-conventional yeast *Saccharomycodes ludwigii* UTAD17 in winemaking", Microorganisms, 7(11), 478 (2019). doi: 10.3390/microorganisms7110478

Tavares MJ, Esteves M, Coelho B, Butler G, Sousa MJ, Mendes-Ferreira A and Mira NP. "Comparative genomics of *Saccharomycodes ludwigii* strains unveils relevant aspects of the physiology of this species and adaptive traits underlying tolerance to sulfite dioxide.", Microbial Genomics (manuscript in preparation).

III.1 Abstract

Sd. ludwigii is a poorly-studied member of the *Saccharomycodaceae* family and studies addressing the establishment of relevant genotype-phenotype associations are very scarce. In this chapter, the genomic sequence of the UTAD17 strain, described in the previous chapter, is leveraged to examine such type of studies using a comparative analysis between UTAD17 and strain BJK_5C, isolated from an apple cider vinegar facility in Slovenia. Significant phenotypic differences were observed between UTAD17 and BJK_5C concerning their different levels of tolerance to SO_2 or to acetic acid; pattern of sugar consumption and aroma formation under fermentation. Results obtained concerning the performance of these strains while fermenting natural grape musts are also described, pointing to a possible application of these strains as adjuncts of *S. cerevisiae*. To correlate these differences with genotypic features, the genome of BJK_5C was obtained and compared to the one of UTAD17. The results obtained concerning this inter-strain *Sd. ludwigii* comparative genomic and phenotypic analyses are herein discussed.

Keywords: *Saccharomycodaceae*, *Saccharomycodes ludwigii*, genomics of wine spoilage yeast, wine fermentation, sulfur dioxide

III.2 Introduction

Saccharomycodes ludwigii is a non-*Saccharomyces* yeast from the *Saccharomycodacea* family that can be found in wine musts and bottled wines (Yamazaki & Oshima, 1996; Vejarano et al., 2018). However, besides this niche, *Sd. ludwigii* strains have been isolated from other instances such as fruit juices and their fermented products (Pilap et al., 2022; Roller & Covill, 1999; Wang et al., 2021), tequila and mezcal (Lachance, 1995), soil samples, insect guts (Lachance et al., 1995), and tree secretions (Stringini et al., 2009; Boundy-Mills et al., 2011). Such diversity is suggestive of a relatively high adaptability of the species to different environments. The recent description of a *Sd. ludwigii* strain capable of enduring high temperatures (capable of growing in 43ºC) and exhibiting multistress tolerance against several stress agents (such as acetic acid, furfural, 5-hydroxymethyl furfural and ethanol) further supports this idea (Pilap et al., 2022). Recently, a novel species, closely related to *Sd. ludwigii*, was proposed for the first time (Wang et al., 2021). This species, termed *Sd. pseudoludwigii*, was isolated from fruit and tree bark in China and exhibited relevant differences in several genomic markers (such as considerable differences in the sequences of the ITS region, and *RPB1* and *TEF1* genes), compared to *Sd. ludwigii* (Wang et al., 2021).

Following the previous chapter, in which the genome of a *Sd. ludwigii* strain (UTAD17), isolated from grape must, was thoroughly examined, in this work the focus lies in observing the main phenotypic and genomic differences observed when using a *Sd. ludwigii* strain (BJK_5C) isolated from an apple cider vinegar facility. Specifically, this work will compare the fermentative performance of the BJK_5C strain using a red and white natural must, comparing the results to what had been also performed for the UTAD17 strain (Esteves et al., 2019). The relevance of this experiment emerges after several studies have shown that, depending on the strain, *Sd. ludwigii* can be used an adjunct of *S. cerevisiae*, in the fermentation of grape musts, resulting in wines with increased organoleptic properties, such as differential aroma profiles and mouthfeel (Domizio et al., 2014; Domizio et al., 2011a; Domizio et al., 2011b).

In a previous work, the strain *Sd. ludwigii* UTAD17 was used as an adjunct of *S. cerevisiae* to ferment a Touriga Nacional red grape must (Esteves et al., 2019). In this study, it was found that all the glucose and fructose available in the must were consumed in a timely manner (5 days when *Sd. ludwigii* and *S. cerevisiae* were co-inoculated, and 8 days when the inoculation was sequential - *Sd. ludwigii* UTAD17 inoculated first and, after 72h, *S. cerevisiae*) (Ciani & Maccarelli, 1997). The analysis of the fermented wines obtained with *Sd. ludwigii* UTAD17

and *S. cerevisiae* showed low levels of acetic acid and ethyl acetate, which contrasts with the results obtained in similar studies conducted with other *Sd. ludwigii* strains that reported, in all cases, high levels of these metabolites (Domizio et al., 2011; Romano et al., 1999; Ciani & Maccarelli, 1997). In fact, the wines obtained by *Sd. ludwigii* UTAD17 in single culture produced lower levels of ethyl acetate than those with *S. cerevisiae*, whether in single inoculation or in sequential co-inoculation (Esteves et al., 2019). Additionally, the wines obtained with *Sd. ludwigii* were also found to have low concentrations of secondary aroma compounds, pointing to a possible use of this strains in the production of wines with a greater expression of their varietal character (Esteves et al., 2019).

In this Chapter, we have undertaken fermentations of two natural musts (Touriga Nacional, a red grape variety, and Sauvignon Blanc, a white grape variety) using two different *Sd. ludwigii* strains: UTAD17 and BJK_5C. The results obtained differed, at some point, between the strains, confirming, that the background of the *Sd. ludwigii* strain used in starter cultures is relevant for the outcome of the obtained product. To get clarifying insights into this aspect, the genome sequence of BJK_5C strain was also herein obtained and discussed, in comparison with the findings obtained before for UTAD17.

III.3 Methods and Materials

III.3.1 Strains and growth media

The experiments conducted in this study used two *Sd. ludwigii* strains: BJK_5C, isolated from an apple cider vinegar facility in Slovenia and kindly provided to us by Professor Janja Trcek (University of Maribor, Slovenia), and the UTAD17 strain, described in the previous chapter (Tavares et al., 2021). Strains were preserved at -80ºC in YPD (containing 20 g/L glucose, 10g/L peptone, 5 g/L yeast extract), and routinely maintained at 4ºC also in solid YPD plates.

For the different experiments, both strains were cultivated in minimal medium (MMB) [containing 20 g/L glucose, 1.7 g/L nitrogen base without amino acids, and 1.67 g/L ammonium sulphate] or in synthetic grape juice medium (GJM). GJM, designed to mimic natural grape must, was prepared with a slight modification of the formulation originally described by Henschke and Jiranek (1993), as previously published by Seixas et al (2023). In specific, equimolar concentrations of glucose and fructose (200 g/L) were used as carbon source. Nitrogen was supplied as a mixture of di-ammonium phosphate (DAP) and amino acids (from a stock solution containing: 1g/L alanine, 9.1 g/L arginine·HCl, 1.7 g/L asparagine·H_2O, 3.5 g/L aspartic acid, 6.22 g/L glutamic acid·HCl, 2 g/L glutamine, 0.74 g/L glycine·HCl, 2.04 g/L histidine·HCl H_2O, 2 g/L isoleucine, 3 g/L leucine, 3.12 lysine·HCl, 1.5 g/L phenylalanine, 5 g/L proline, 4 g/L serine, 3.5 g/L threonine, 1 g/L tryptophan, 0.2 g/L tyrosine, 2 g/L valine, and 1.5 g/L methionine).

III.3.2 Growth of *Sd. ludwigii* strains in different carbon sources and under different relevant environmental stresses.

The capacity of *Sd. ludwigii* UTAD17 and BJK_5C strains to grow in different carbon sources and to endure different stress conditions was assessed based on spot assays. To perform this phenotypic assay, the two *Sd. ludwigii* strains were grown overnight in 25 mL liquid minimal medium MMB, at 30ºC with orbital agitation of 250 rpm, in 50mL shake flasks. On the next day, a new suspension was prepared in fresh medium (aiming for an initial OD_{600nm} of 0.1) and the cells incubated at 30ºC and 250 rpm until exponential phase was reached (after 3 hours). At this point, a cell suspension having an OD_{600nm} of 0.05 was prepared in water as well and two subsequent dilutions (1:5 and 1:25). Of these cell suspensions, 4 µL were applied as spots onto MMB medium having 2% of glucose (used as a control) or fructose, saccharose,

galactose, maltose, mannose, mannitol, rhamnose, sorbitol, xylose, glycerol or D-L-lactic acid. The media and the organic acids concentrations were adjusted to pH 3.5. Cell growth was monitored for 72h. This same experimental setup was used to monitor growth of UTAD17 and BJK_5C in MMB medium supplemented with ethanol (2% to 6%), acetic acid (10 mM to 60 mM, final pH 3.5) and SO_2 (from 2 to 6 mM). SO_2 was included in media by incorporating equal amounts of freshly prepared stock solutions of potassium metabisulfite (Merck) in water, with pH adjusted to 3.5, to obtain the desired final concentrations.

III.3.3 Karyotyping based on pulsed-field gel electrophoresis (PFGE)

PFGE was used to separate the chromosomal DNA bands of *Sd. ludwigii* BJK_5C in similar terms to what had been done before for UTAD17. The experimental setting described by Sipiczki et al. (2003), modified by Hage & Houseley (2013) was used. In brief terms, the yeast chromosomal bands were separated in a BioRad CHEF-DRIII electrophoresis apparatus (Bio-Rad, Hercules, CA, USA) using a 1% agarose gel in 0.5x TBE buffer at 12ºC. Electrophoresis ran at 3 V/cm for 36h with a 200-300s ramping switch interval and for 60h with a 300-600s ramping switch interval. The size comparison of the obtained bands and the correspondent calculation of their molecular sizes was made possible by using CHEF-DNA size markers of *Hansenula wingei* (with chromosome bands ranging from 1.05 to 3.13 Mbp) and of *S. cerevisiae* (using chromosomes with a molecular weight below 1.05 Mbp). The molecular sizes of the chromosomal bands obtained in the PFGE gel of *Sd. ludwigii* BJK_5C were calculated using a calibration curve (band distance vs molecular size of chromossomes obtained for *H. wingei* and *S. cerevisiae*) using the ImageJ software.

III.3.4 Quantification of *Sd. ludwigii* UTAD 17 and *Sd. ludwigii* BJK_5C total genomic DNA and ploidy by flow cytometry

Total genomic DNA content from *Sd. ludwigii* strains UTAD17 and BJK_5C was performed using a SYBR Green-based protocol, as described before (Mira et al., 2014). Briefly, cells were grown in YPD medium at 26ºC until mid-exponential phase (OD_{600nm} ~ 0.1), harvested by centrifugation, washed twice with H_2O and fixed overnight in 500mL of 70% ethanol (vol/vol). After the fixing step, cells were collected by centrifugation, washed with 50 mM of sodium citrate buffer (pH 7.5), re-suspended in 750 μL of 50mM sodium citrate buffer (pH 7.5) supplemented with 1mg of RNAse, and incubated at 50ºC for 1h. Afterwards, 1 mg

of proteinase K was added to the cell suspension and the mixture was left at 50ºC for another hour. Later, cells were stained using 20 μL of SYBR Green I working solution (500-fold dilution of the commercial solution), and samples were sonicated at low power before being analyzed in an Epicsw XLTM (Beckman Coulter) flow cytometer equipped with an argon ion laser emitting a 488 nm beam at 15 mW. The green fluorescence was collected through a 488 nm blocking filter, a 550 nm long-pass dichroic and a 525 nm bandpass. The mean fluorescent intensities obtained for *S. cerevisiae* BY4741 and BY4743 were used to build a calibration curve, from which it was possible to estimate the size of the genomes of the *Sd. ludwigii* strains BJK_5C and UTAD17.

III.3.5 Grape juice preparation, fermentation trials and analytics.

Natural grape juice was obtained by crushing grapes of the *Vitis vinifera* L. cv. Touriga Nacional (red) and *Vitis vinifera* L. cv. Sauvignon Blanc (white). After homogenization, the juice was clarified, by centrifugation at 12,734× g for 10 min, to carefully separate the solid fraction. A sample of the grape-juice was collected at this point for routine analysis. After pasteurization at 70 °C for 10 min, the grape-juice was immediately cooled on ice. For each *Sd. ludwigii* strain (UTAD17 and BJK_5C), the inoculum was prepared by separately pre-growing the yeast cells in 50 mL-flasks, containing 25 mL of synthetic grape-juice medium (GJM). Nitrogen was added up to 267 mg YAN/L, supplied as di-ammonium phosphate (DAP). The flasks were incubated overnight at 25 °C in an orbital shaker set at 150 rpm min^{-1}. Strains *Sd. ludwigii* UTAD17 and BJK_5C were inoculated individually in grape-juice with an initial cellular concentration of 10^6 CFU/mL. Fermentations trials were conducted by inoculating single cultures of *Sd. ludwigii* strains UTAD17 and BJK_5C to each grape must. Single culture fermentations were carried out in triplicates using a previously described system (Mendes-Ferreira et al., 2009) consisting of 100 mL flasks filled to 2/3 of their volume (80 mL) and fitted with a side-arm port sealed with a rubber septum for anaerobic sampling. Two flasks containing not inoculated grape-must were used as control. The flasks containing Touriga Nacional grape must were maintained at 25 °C under static conditions, while Sauvignon Blanc fermentations were carried out at 15 °C, also under static conditions. Fermentations were monitored daily by weight loss (as an estimation of CO_2 production) and were allowed to proceed until no further weight loss was observed. For the assessment of growth parameters and analytical determinations, aseptic sampling was performed using a syringe-type system.

After fermentation, the wines were centrifuged for 10 min at 5500 rpm, to remove yeast cells and were kept at −20 °C until the analytical determinations were performed.

The amount of glucose, fructose, acetic acid, as well as Yeast Assimilable Nitrogen (YAN) (comprising primary amino nitrogen – PAN - and ammonium) were enzymatically determined using a Y15 autoanalyzer (Biosystems S.A, Barcelona, Spain). Total SO_2, pH, and titratable acidity were determined according to the standard methods compiled in the Compendium of International Methods of Analysis of Musts and Wines (Paris, 2020). Ethanol and glycerol concentrations were determined in a high-performance liquid chromatography system (HPLC Flexar, PerkinElmer, Shelton, Connecticut, EE. UU) equipped with the ion exclusion cation exchange column Aminex HPX-87H (Bio-Rad Laboratories, Hercules, CA, USA) and refractive index detector. Samples were eluted with sulfuric acid (0.005 M) at 60 °C and a 0.6 mL/min flow rate. Samples were previously filtered through a membrane (Millipore, 0.22 μm pore size) before an injection of 6 μL. The components were identified through their relative retention times, compared to the respective standards, using the Perkin Elmer Chromera Software.

Aliphatic higher alcohols (1-propanol, 1-butanol, 2-methyl-1-butanol and 3-methyl-1-butanol), acetaldehyde, and ethyl acetate present in differently obtained wines were analyzed as described Moreira et al. (2011), using a Hewlett-Packard 5890 (Hewlett-Packard, Palo Alto, CA 94304, USA) gas chromatograph equipped with a flame ionization detector (GC-FID) and connected to a H.P. 3396 Integrator. Fifty microliters of 4-methyl-2- pentanol at 10 g L^{-1} were added to 5 mL of wine as the internal standard. The sample (1 μL) was injected (split, 1: 30) into a CP-WAX 57 CB column (Chrompack) of 50 m × 0.25 mm and 0.2 μm phase thickness. The program temperature varied from 40 °C (10 min) to 80 °C (10 min) at 3 °C min−1 and from 80 °C to 200 °C (4 min) at 15 °C min^{-1}. Injector and detector temperatures were set at 220 °C. Carrier gas was H_2 at 1–2 mL min^{-1}. The determination of 2-phenylethanol, acetates of higher alcohols (isoamyl acetate, 2-phenylethyl acetate) and ethyl esters of fatty acids (ethyl butanoate, ethyl hexanoate and ethyl octanoate), volatile fatty acids (butyric, isobutyric, isovaleric acids) and free fatty acids (hexanoic, octanoic and decanoic acids) was performed in a Hewlett Packard 5890 gas chromatograph, equipped with a flame ionization detector. For this purpose, 50 mL of wine, with 4-decanol at 1.5 mg/L as the internal standard, was extracted successively with 4, 2, and 2 mL of ether–hexane (1:1 v:v−1) for 5 min. The organic phase (1 μL) was injected (splitless) into a BP21 (SGE) column of 50 m × 0.22 mm and 0.25 μm phase thickness. The temperature program was 40 °C (1 min) to 220 °C (15 min), at 2 °C·min−1.

Injector and detector temperatures were set at 220 °C. The carrier gas used was H2 at 1–2 mL min^{-1}.

III.3.6 Genome sequencing, assembly and annotation of *Sd. ludwigii* BJK_5C

To obtain genomic DNA of *Sd. ludwigii* BJK_5C, cells were grown overnight in YPD at 28ºC, 250rpm, and DNA extraction was performed as described previously (Salazar et al., 2018). Sequencing of the genomic DNA obtained was performed by Illumina MiSeq and MinION, enabling the obtention of long reads together with shorter, more accurate, reads. For the Illumina-based sequencing, the DNA libraries were prepared using the ThruPLEX DNA-seq kit, and paired-end sequencing of the generated DNA fragments was performed on a MiSeq platform with two sequencing rounds, generating 250 bp paired-end reads. The quality of the Illumina data was assessed using Fastqc. Reads were trimmed with Skewer version 0.2.2, choosing the paired-end mode, and a minimum read length after trimming of 50 bases. For the MinION, sequencing libraries were prepared by barcoding with the Rapid Barcoding Kit (SQK-RBK004) from Oxford Nanopore Technologies (ONT). The prepared libraries were loaded onto an ONT MinION flow cell (FLO-MIN 106) and sequenced for 50h. Basecalling was performed using the ONT Guppy software, version 3.2.4+d9ed22f with the following parameters; "--input_path fast5 --save_path fastq --flowcell FLO-MIN106 --kit SQK-RBK004 --verbose_logs --cpu_threads_per_caller 5 --num_callers 7. Basecalled data were demultiplexed using qcat version 1.1.0 with the following parameters; "--fastq fastq/all_multiplexed_reads.fastq --barcode_dir demultiplex_qcat --detect-middle --min-read-length 1 --trim --kit RBK004 --epi2me, and basecalled data were demultiplexed using qcat version 1.1.0. Read quality was checked with NanoPlot version 1.23.1, and potential contaminant reads were identified by command-line BLASTN from the BLAST+ package version 2.2.31 and reads with Q < 7 and length < 1 kilobase (Kb) were remived with NanoFilt version 2.3.0. Filtered MinION reads were assembled using Canu version 1.8 using recommended parameters for haplotype separation and Illumina reads were used to polish the assembly with Pilon version 1.2.3. Assembly statistics were checked using QUAST version Genome annotation of the 8 contigs obtained was performed using the WebAugustus tool within the framework of Geneious software (version 2021.2.2) using as training datasets the genomes of *S. cerevisiae, S. pombe* and *Sd. ludwigii* UTAD17 to identify putative gene models. BLASTp analysis against the UNIPROT database was used in the manual curation of the predicted gene models.

III.3.7 Comparative genomic analysis of *Sd. ludwigii* UTAD17 and BJK_5C.

A low and high-resolution approach was used for the comparative analysis of the genomes of *Sd. ludwigii* BJK_5C and UTAD17. The low-resolution analysis was performed by comparing, in MAUVE, the 8 contigs obtained after the mixed assembly of the obtained BJK_5C reads with the 20 contigs that had been obtained for the UTAD17 strain. The high-resolution analysis was focused on the identification of single nucleotide polymorphisms (SNP) in the homologous gene alleles encoded by the two strains. For this, the Illumina BJK_5C reads were mapped against the UTAD17 annotated genome using the tools available at CLC Genomics Workbench. To make the SNP call detection, a variant detection using probabilistic and quality-based variant detection were used. Only those SNPs with coverage > 20%, forward/reverse balance > 0.4 and frequency > 70 were considered. The predicted set of BJK_5C proteins (the BJK_5C ORFeome) was compared with the UTAD17 proteome established in the previous chapter using BLASTp. Proteins were considered highly similar in between the two strains if the BLAST associated to the pairwise alignment was above 50%, with an associated e-value below 1E-20. If the pairwise alignment identity fell between 30 and 50%, and the associated e-value below 1E-20E, we considered these proteins to be similar. In every other case, the proteins pairs were considered dissimilar. To get further insights into biological function of the identified BJK_5C proteins, functional analysis was performed using BLASTKoala (Kanehisa et al., 2016).

III.4 Results and Discussion

III.4.1 *Sd. ludwigii* BJK_5C and *Sd. ludwigii* UTAD17 show different phenotypic traits concerning tolerance to environmental stressors and utilization of carbon sources.

To see whether the different isolation environment from where the *Sd. ludwigii* strains UTAD17 (retrieved from wine must) and BJK_5C (retrieved from a cider plant) were recovered impacted the physiology of these strains, their growth was examined using different carbohydrates as carbon sources including glucose, fructose, glycerol, DL-lactic acid, and maltose. In the presence of glucose or fructose, growth of the two strains was identical, however, in all other cases *Sd. ludwigii* BJK_5C exhibited a more robust growth than the UTAD17 strain (**Figure III. 1A**). The difference is particularly noticeable when the cells were cultivated in media having lactate, maltose, glycerol and galactose as carbon sources that only resulted in an almost negligible growth of UTAD17 cells, while BJK_5C cells exhibited a clear growth under these conditions. (**Figure III. 1A**). We have also compared the ability of these strains to tolerate environmental stresses relevant in the context of their ecological niches including tolerance to ethanol and SO_2 (relevant in winemaking) and high concentrations of acetic acid, relevant in the context of production of apple cider vinegar. For this analysis, the strains were always cultivated in a media containing glucose as the carbon source. The results obtained, shown in **Figure III. 1B**, demonstrate a clear higher tolerance of UTAD17 cells towards SO_2-induced stress, while BJK_5C cells were clearly more tolerant to EtOH and to acetic acid (**Figure III. 1B**.). This pattern is, somehow consistent with the expected adaptive responses evolved by these cells to thrive in the environments where they were isolated from, since high SO_2 concentrations are much more prevalent in the wine environment, where UTAD17 originates, while acetic acid is extremely prominent in the context of vinegar production (acetic acid concentration in vinegar varies from 4 to 8% - Stratford, 1999), where BJK_5C was isolated. *Sd. ludwigii* BJK_5C was also shown to tolerate higher ethanol concentrations than UTAD17. This observation may be due to a differential metabolization of ethanol between the two strains. Additionally, strain BJK_5C may possess a higher tolerance to stress conditions, due to the constant exposure to an extreme environment, similarly to what was registered in *S. cerevisiae* (Teixeira et al., 2009; Hirasawa et al., 2007; Yoshikawa et al., 2009; Guaragnella et al., 2021). Further studies are, thus, needed to better understand these mechanisms to proper leverage them for potential biotechnological applications.

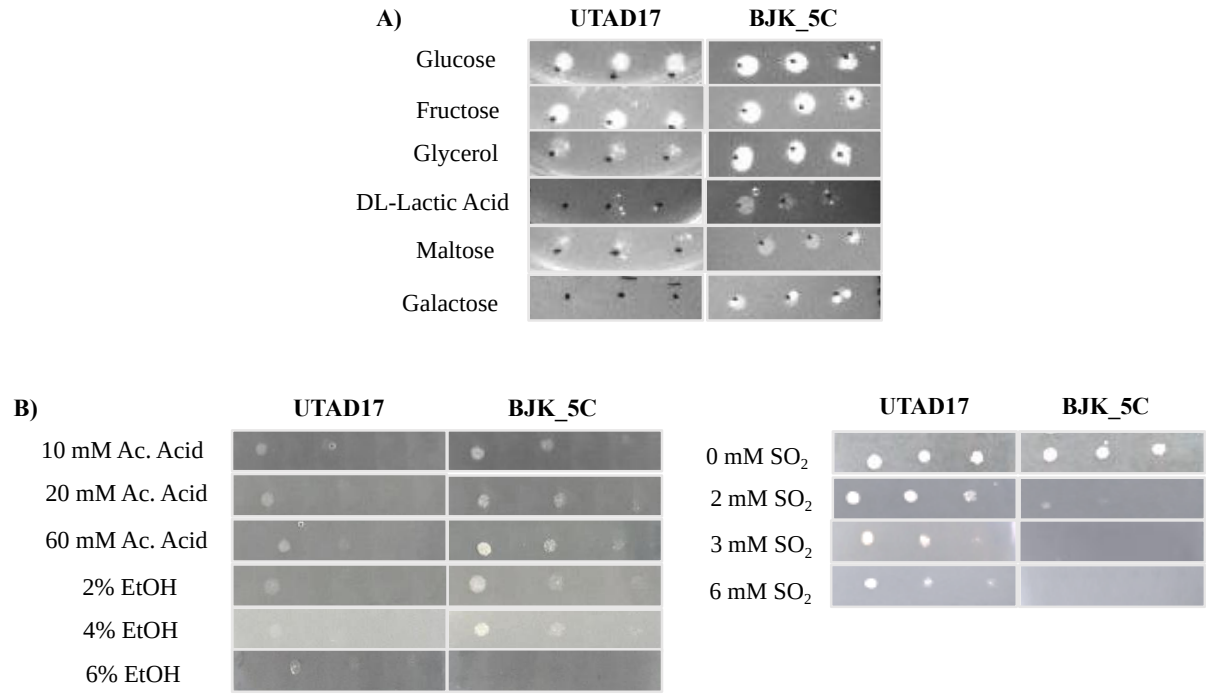


Figure III. 1- Phenotypic characterization of strains *Sd. ludwigii* BJK_5C and *Sd. ludwigii* UTAD17 by growing the cells in spot assays in media containing A) different carbon sources; B) increasing concentrations of acetic acid and ethanol; C) increasing concentrations of potassium metabisulfite after 72h.

III.4.2 Strains *Sd. ludwigii* BJK_5C and UTAD17 exhibit different fermentation profiles of natural grape must, resulting in wines with different aromatic profiles

Prior studies describe the utilization of *Sd. ludwigii* strains as adjuncts of *S. cerevisiae* for the production of beers (De Francesco et al., 2015; Jackowski et al., 2023). In all these cases had been used *Sd. ludwigii* strains with origin in grape musts or in fermented beverages. This was also the case of the UTAD17 strain that had been previously used, together with *S. cerevisiae* QA23 commercial strain, to ferment red grape musts (Esteves et al., 2019). In this study we decided to examine the performance of the BJK_5C strain in fermentation of two natural grape musts (red and white) considering its origin in a cider plant. As a control and point of comparison we used the UTAD17 strain. The vinifications were performed in red and in white grape musts, from Touriga Nacional and Sauvignon Blanc grape varieties. The results obtained concerning the fermentation parameters are shown in **Figure III. 2**. Growth dynamics and fermentation profiles of the two strains in Touriga Nacional (**Figure III. 2B**) were similar (fermentation was complete in both cases after approximately 384 hours) but in Sauvignon Blanc a clear difference was observed, with the UTAD17 cells fermenting the must at a much higher rate (**Figure III. 2A**) *Sd. ludwigii* UTAD17 completed the fermentation after ~384

hours, while *Sd. ludwigii* BJK_5C cells only completed the fermentation after 816 hours. One possibility that may be investigated in the future as a determining factor of the slower fermentation rate of BJK_5C cells is this strain's inferior ability to ferment at lower temperatures, since fermentations of white grape musts were conducted at 15ºC, while red grape musts were fermented at 25ºC.

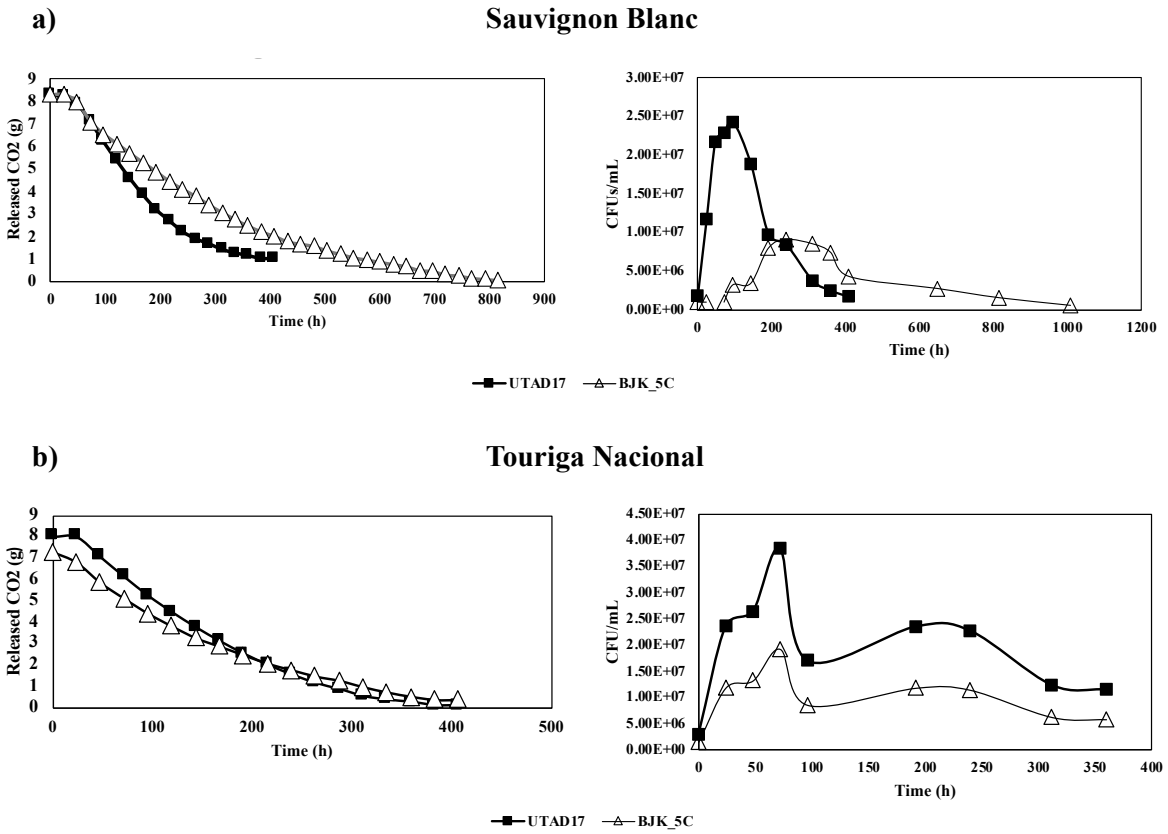


Figure III. 2– Means of the fermentation profiles obtained with *Sd. ludwigii* UTAD 17 and *Sd. ludwigii* BJK_5C that followed the weight loss and corresponding release of CO_2 in A) Vitis vinifera L. cv. Sauvignon Blanc (white grape variety) and B) Vitis vinifera L. cv. Touriga Nacional (red grape variety).

The red and white wines produced by the two *Sd. ludwigii* strains exhibited very similar properties in terms of ethanol, acetic acid and glycerol production (**Table III. 2**). In particular, both strains generated wines with approximately the same amount of ethanol (~10%), despite Sauvignon Blanc having more total sugars than Touriga Nacional (**Table III. 1**; **Table III. 2**), with UTAD17 showing a slightly higher concentration in both cases. Red wines had higher concentrations of acetic acid, compared to the white wine obtained for both strains, with those obtained by UTAD17 having slightly higher concentrations. In contrast, glycerol production was marginally higher in red wines produced by BJK_5C cells (**Table III. 2**).

Table III. 1 – Physicochemical properties of the initial Sauvigon Blanc and Touriga Nacional grape musts.

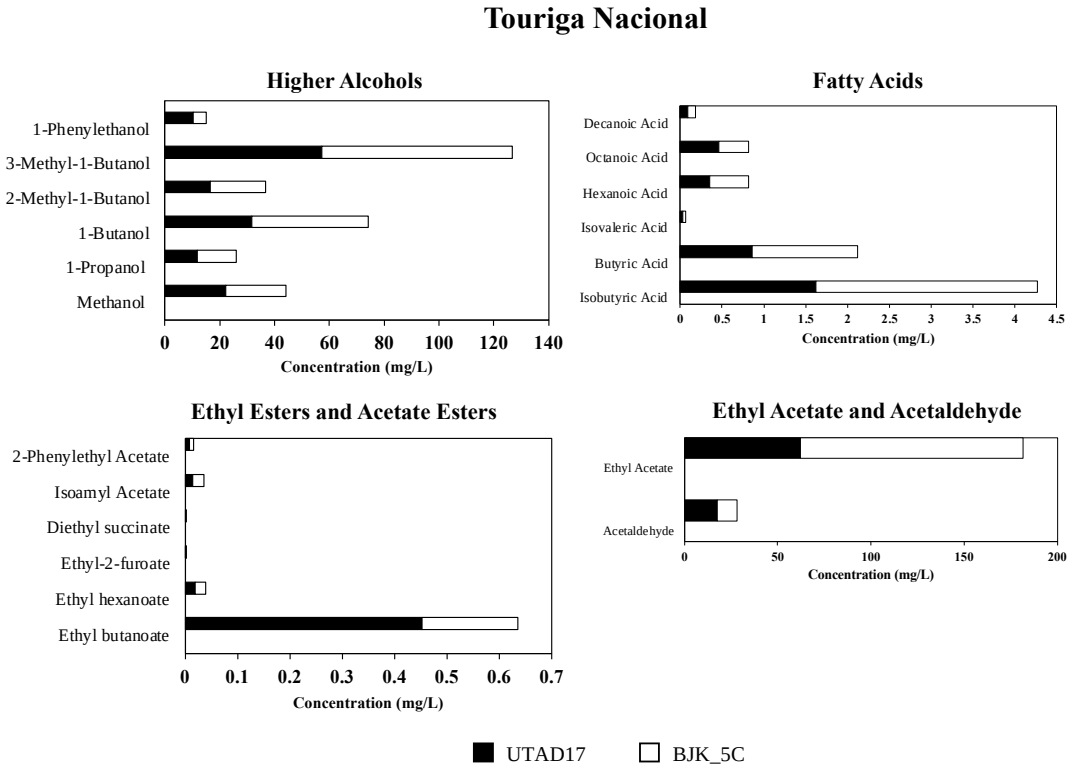
Parameter	Sauvignon Blanc	Touriga Nacional
Total Sugars (g/L)	197.04	185.87
Glucose (g/L)	101.08	95.42
Fructose	95.96	90.45
YAN (mg/L)	220.6	189.38
Ammonia (mg/L)	80	109

Table III. 2 - Concentrations of ethanol, acetic acid and glycerol of the wines obtained from the fermentation of the grape varieties Touriga Nacional (red grape variety) and Sauvignon Blanc (white grape variety) with the *Sd. ludwigii* strains UTAD17 and BJK_5C.

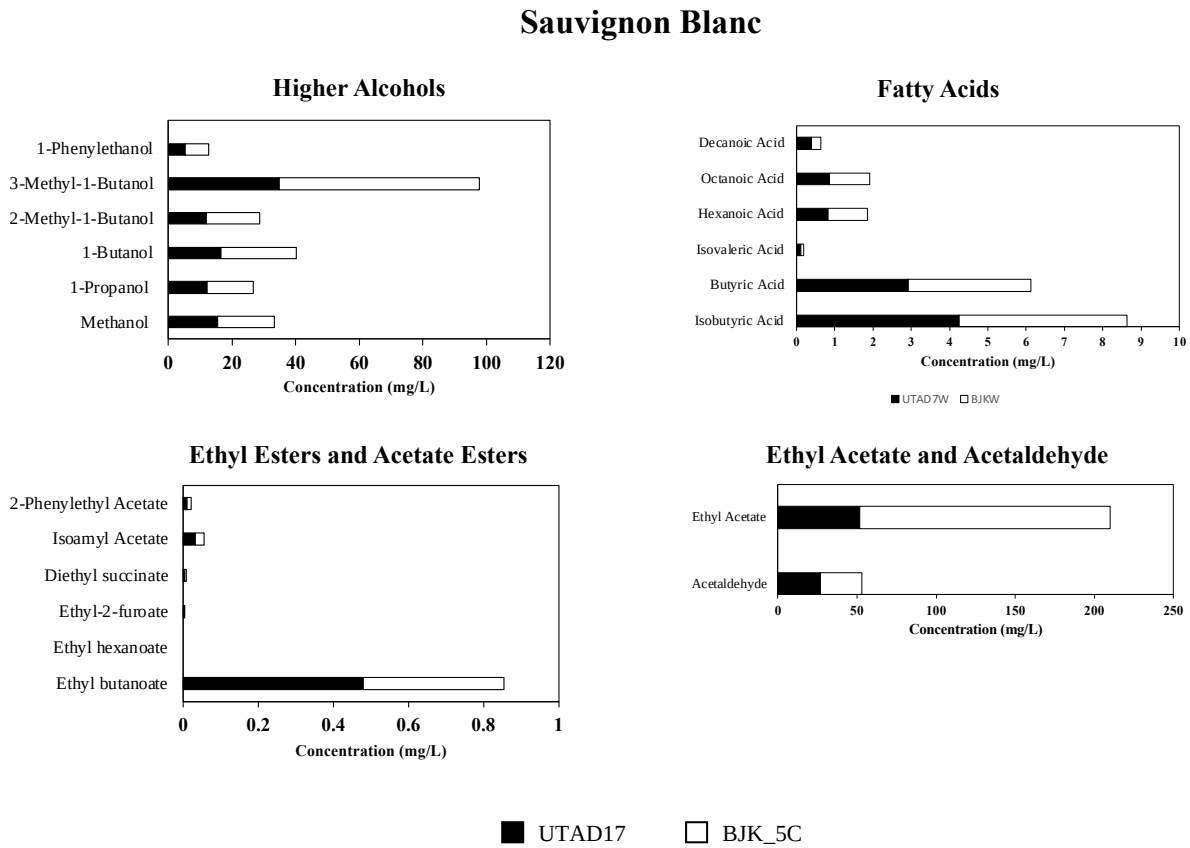
	Touriga Nacional			Sauvignon Blanc		
Sd. ludwigii strain	**Ethanol (%)**	**Acetic acid (%)**	**Glycerol (g/L)**	**Ethanol (%)**	**Acetic acid (%)**	**Glycerol (g/L)**
UTAD17	10.4	0.438	6.1	10.7	0.306	5.4
BJK_5C	10.2	0.435	7.3	10.2	0.213	6.7

Concerning the aroma compounds profile, shown in **Figure III. 3**, it was evident a higher concentration of higher alcohols, ethyl acetate and short-chain fatty acids (butyric and isobutyric acids) in both wines, compared to the amount of acetate esters (with the exception of ethyl butanoate), acetaldehyde and intermediate chain fatty acids (decanoic acid, octanoic acid) (**Figure III. 3**). In general, the volatiles detected in the red and white wines were similar, with the exception of the fatty acids that were present in considerably higher amounts (~2-fold) in the white wines. Notably, the effect of the strain of *Sd. ludwigii* used to produce the wines was much more evident than the type of must used (that is, we found much more differences in the wines produced by UTAD17 or BJK_5C, than those obtained when comparing the white/red wine fermented by the same strain, **Figure III. 3**). In general, wines produced by strain BJK_5C had higher amounts of volatiles, compared to wines produced by strain UTAD17. This difference was more noticeable for ethyl acetate that achieved a concentration ~2-fold higher in the red wine and a 3-fold increase in the white wine (**Figure III. 3)**. This high production of ethyl acetate by BJK_5C is, in fact, consistent with results from studies with other *Sd. ludwigii* strains under different conditions (Estela-Escalante, 2011; Estela-Escalante, 2018; Romano et al., 1999; Vejarano, 2018) thus suggesting that the low production prompted by UTAD17 cells is a strain-specific phenotype.

A)


Touriga Nacional
Higher Alcohols
1-Phenylethanol
3-Methyl-1-Butanol
2-Methyl-1-Butanol
1-Butanol
1-Propanol
Methanol
0 20 40 60 80 100 120 140
Concentration (mg/L)
Fatty Acids
Decanoic Acid
Octanoic Acid
Hexanoic Acid
Isovaleric Acid
Butyric Acid
Isobutyric Acid
0 0.5 1 1.5 2 2.5 3 3.5 4 4.5
Concentration (mg/L)
Ethyl Esters and Acetate Esters
2-Phenylethyl Acetate
Isoamyl Acetate
Diethyl succinate
Ethyl-2-furoate
Ethyl hexanoate
Ethyl butanoate
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7
Concentration (mg/L)
Ethyl Acetate and Acetaldehyde
Ethyl Acetate
Acetaldehyde
0 50 100 150 200
Concentration (mg/L)
UTAD17
BJK_5C

B)


Sauvignon Blanc
Higher Alcohols
1-Phenylethanol
3-Methyl-1-Butanol
2-Methyl-1-Butanol
1-Butanol
1-Propanol
Methanol
0 20 40 60 80 100 120
Concentration (mg/L)
Fatty Acids
Decanoic Acid
Octanoic Acid
Hexanoic Acid
Isovaleric Acid
Butyric Acid
Isobutyric Acid
0 1 2 3 4 5 6 7 8 9 10
Concentration (mg/L)
UTAD7W
BJKW
Ethyl Esters and Acetate Esters
2-Phenylethyl Acetate
Isoamyl Acetate
Diethyl succinate
Ethyl-2-furoate
Ethyl hexanoate
Ethyl butanoate
0 0.2 0.4 0.6 0.8 1
Concentration (mg/L)
Ethyl Acetate and Acetaldehyde
Ethyl Acetate
Acetaldehyde
0 50 100 150 200 250
Concentration (mg/L)
UTAD17
BJK_5C

C)

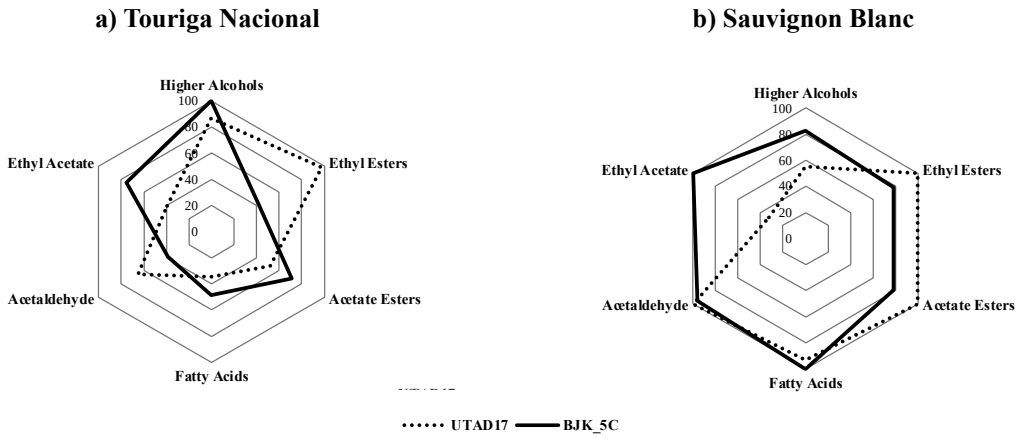


Figure III. 3- Normalized aroma compounds of wines obtained from A) the red grape variety, Touriga Nacional, and the B) white grape variety, Sauvignon Blanc, fermented with *Sd. ludwigii* strains UTAD17 (black filling) and BJK_5C (white filling). C) Radar charts of the obtained wines, summarizing the main differences found between *Sd. ludwigii* strains BJK_5C (complete line) and UTAD17 (dashed line). Generally, wines obtained with the strain BJK_5C showed an overproduction of ethyl acetate, while those obtained with UTAD17 revealed more ethyl esters. The effects of both strains were more pronounced in the white wine than in the red wine. However, strain UTAD17 produced wines with a more neutral yeast contribution to the overall aroma.

III.4.3 Genome sequencing unveils differences in architecture and in the ORFeomes of *Sd. ludwigii* UTAD17 and BJK_5C strains.

In the previous chapter, karyotyping of the genome of *Sd. ludwigii* UTAD17 strain revealed that this strain is equipped with 7 chromosomes and a genome totaling ~13.75 Mb. This number of chromosomes was in line with what has been reported for other strains of this species (Yamazaki & Oshima, 1996). In this context, we have used PFGE to obtain the number, and size, of chromosomes for BJK_5C, having obtained 9 chromosomal bands (with sizes ranging from 0.929 Mbp and 3.13 Mbp) as detailed in (**Figure III. 4**). The total genome size is circa 17.75 Mbp. Compared to the genome of UTAD17, two extra chromosomal bands and approximately more 4 Mbp were obtained in the PFGE of the BJK_5C strain. We wondered whether these differences can reflect a different ploidy of the UTAD17 and BJK_5C strains, that may not be easily perceived by the PFGE, as sister chromosomes (with the same molecular weight) will appear under a single band, eventually more intense. Indeed, when comparing the PFGE gel obtained for the two *Sd. ludwigii* strains (**Figure III. 4**), it is possible to observe that two of the chromosomal bands (2.87 Mbp and 1.48 Mbp) present in the karyotype of BJK_5C appear to have a lower intensity than the others, suggesting that these bands may represent

homologous chromosomes with different sizes (or, alternatively, pieces of chromosomes that became fragmented).

The existence of small chromosomes in other yeasts has been reported, including in *C. glabrata* and *S. cerevisiae*, often leading to environmental adaptations (Ahmad et al., 2013). In a specific study undertaken on *C. glabrata*, the authors proposed two different mechanisms explaining the occurrence of these small chromosomes (Ahmad et al., 2013): i) involving segmental duplication covering the centromeric region, resulting in duplicated genes and mitotic instability; ii) involving a translocation event where a larger chromosome arm moved to another chromosome, leaving the centromere with a shorter arm. This led to genes being present in only one copy within the genome, resulting in mitotic stability (Ahmad et al., 2013). Although the existence of small chromosomes has not been reported in *Sd. ludwigii*, the unusual cell division observed in this species, characterized by achiasmate meiosis, may increase its meiotic complexity, potentially leading to the development of diverse structures and phenotypes (Papaioannou et al., 2021).

To further explore this issue, we estimated the genome sizes of BJK_5C and of UTAD17 strains by flow cytometry, using the established genome sizes of *S. cerevisiae* haploid (BY4741) and diploid (BY4743) strains as references (**Appendix Table III.1**). The results obtained indicate that the fluorescence intensities of the *Sd. ludwigii* cells in G0 and G1 phases are quite different, suggesting different ploidy levels, with BJK_5C likely being triploid, and UTAD17 diploid. Polyploid states in yeasts have been linked to higher genomic complexity, resulting in increased stress tolerance and evolutionary adaptation (Mozzachiodi et al., 2022; Selmecki et al., 2015), but also sterility and reduced genetic diversity (Mozzachiodi et al., 2022; Ezov et al., 2006). For instance, a study by Ezov et al. (2006) on wild triploid *S. cerevisiae* strains revealed that these yeasts performed clonal reproduction more frequently than sexual reproduction, leading to a complete homogenization of their ribosomal DNA (Ezov et al., 2006). Furthermore, the domestication of *S. cerevisiae* has resulted in complex polyploid genomes, such as triploid and tetraploid forms, in industrial strains (Mozzachiodi et al., 2022; Ezov et al., 2006). Being isolated from an industrial apple cider vinegar plant, it is possible that BJK_5C cells have been exposed to various industrial stresses that may have boosted alterations in genomic architecture and evolution towards higher ploidy, an aspect that will require further elucidation.

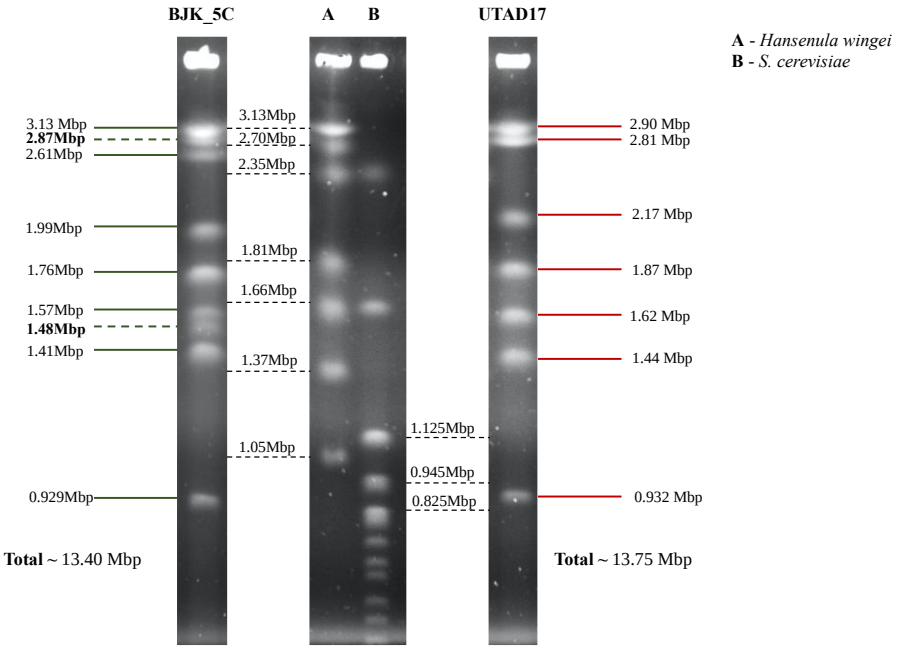


Figure III. 4- Pulsed-field gel electrophoresis of *Sd. ludwigii* strains BJK_5C and UTAD17. In this experiment *Sd. ludwigii* BJK_5C chromosomal bands are separated by electrophoresis into 9 visible bands which can be compared with the 7 bands of *Sd. ludwigii* UTAD17. *S. cerevisiae* and *H. wingei* are used as weight markers for lower and higher molecular weights respectively. The total sum of the 9 separated bands yields 17.75 Mbp. Two extra chromosomal bands were observed in BJK_5C and are represented by dashed lines.

To expand our knowledge about genomics of the *Sd. ludwigii* strains we decided to obtain the genome sequence of the BJK_5C strain, using a hybrid assembly of Illumina MiSeq and MinION reads. With this mixed approach we obtained a reduced number of contigs after the assembly, while obtaining high-quality reads suitable for a fine-tuned SNPs analysis. After assembly, the reads were assembled into 8 contigs, with sizes ranging from 3.045 Mbp to 86.26 Kbp and an average GC content of 30.7%, as detailed in **Table III. 3.** As expected, the GC content obtained for *Sd. ludwigii* BJK_5C was identical to the one reported for the UTAD17 strain (**Table III. 3**). The assembled genome size for BJK_5C was 12.25 Mbp, which differs from the estimated size of 17.75 Mbp obtained by PFGE by ~5Mbp.

Table III. 3– Genome general assembly and annotation statistics of *Sd. ludwigii* BJK_5C and *Sd. ludwigii* UTAD17

Genome assembly statistics	***Sd. ludwigii* BJK_5C**	***S.ludwigii* UTAD17**
Number of contigs	8	20
N50 (mbp)	1.82	1.48
Assembly size (mbp)	12.25	13
Number of predicted genes	4,901	5,008
Average GC content (%)	30.7	31

A proposed correspondence between the obtained contigs and the chromosomal bands observed in the PFGE, based on genomic size, can be found in **Table III. 4**. In this correspondence, the two additional BJK_5C chromosomal bands identified in the PFGE were left out. Additionally, a smaller contig, identified as the mitochondrial contig, during sequencing (Scaffold 8) was also excluded from the correspondence, due to its reduced size.

Table III. 4– Proposed correspondence between contigs obtained through genome assembly and Pulsed-Field Gel Electrophoresis (PFGE).

ASSEMBLY			**PFGE**		
Scaffold	**Size (Mbp)**		**Chromosomes**	**Size (Mbp)**	
1	3.05	→	1	3.13	3.13
			2	**2.87**	
2	2.18	→	3	2.61	2.61
3	1.82	→	4	1.99	1.99
4	1.56	→	5	1.76	1.76
5	1.43	→	6	1.57	1.57
			7	**1.48**	
6	1.32	→	8	1.41	1.41
7	0.818	→	9	0.929	0.929
8	0.086				
-	-				
Sum	12.25		Sum (Mbp)	17.749	13.40

We undertook a whole alignment (using MAUVE) of BJK_5C and UTAD17 assembled contigs resulting in the chromosomal map that is shown in **Figure III. 5**. Following this analysis, we observed that most of the contigs from UTAD17 could align with those obtained for BJK_5C (**Figure III. 5**). The whole-contig alignment of both *Sd. ludwigii* strains' also revealed what appears to be several genomic arrangements including inversions (verified in the Scaffolds 4 and 7 from BJK_5C, comparing to UTAD17's contigs SCLUD5 and SCLUD7); as well as regions that are present in the genome of BJK_5C's but without homology in UTAD17 (see scaffolds 1, 2, 4 and 6 (**Figure III. 5**). One observation that caught our eye was the existence of several UTAD17 contigs that are very similar to each other (*e.g.*, contig 19 is 99% identical to contig 12), or similar parts of other contigs (*e.g.*, contig 15 appears to be identical to contig 1 and contig 8, with more than 99% identity) (**Figure III. 5B**). The existence of these seemingly duplication of contigs may result from sequencing errors but they can also be attributable to some putative loss of chromosomal homozygosity that increased genomic

dissimilarities and prevented the overlapping of sequences during assembly. This is an aspect that will have to be studied in the future in order to get a full picture of what is the BJK_5C genomic architecture. To understand whether the observed differences in the whole-genome comparisons of BJK_5C and UTAD17, we aligned the contigs obtained for BJK_5C with those published for the strain *Sd. ludwigii* NBRC1722 (Papaioannou et al., 2021). The results showed a very high similarity in the chromosomal arrangement of both strains, with the only differences being the absence of match of the predicted mitochondrial contig in BJK_5C (0.086 Mbp) (**Appendix Figure III.1**). These findings suggest that the genomic sequence and chromosomal arrangement of UTAD17 is the different player, which may result from an adaptation to the wine environment. However, further genomic analyses focused on other *Sd. ludwigii* strains are needed to help us answer this question.

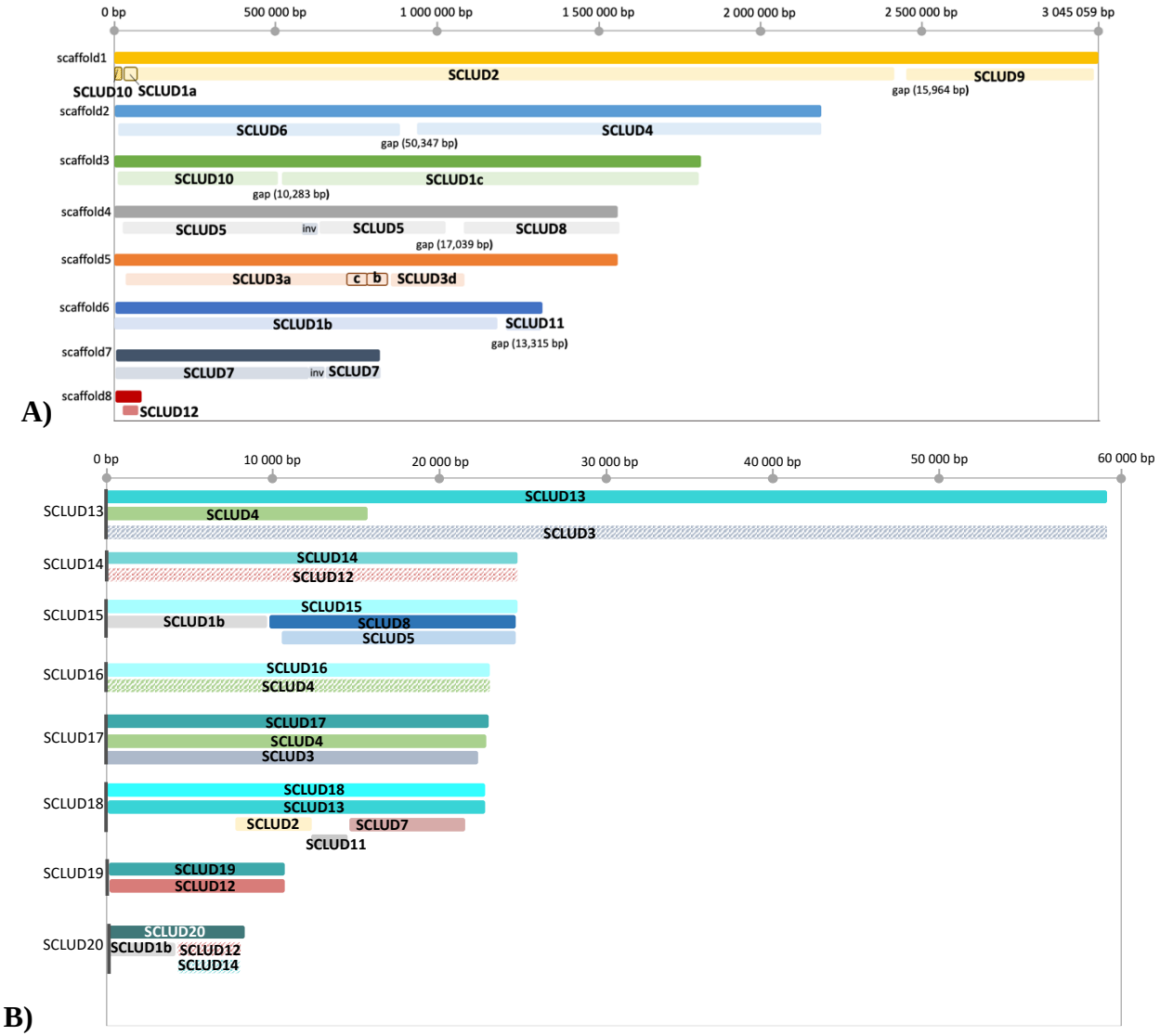


Figure III. 5- A) Whole-genome alignment of *Sd. ludwigii* BJK_5C contigs (Scaffold 1 to 8) and *Sd. ludwigii* UTAD17 contigs (SCLUD1-SCLUD20). This alignment facilitated a better understanding of the genomic similarities and differences between the two *Sd. ludwigii* strains. The analysis revealed that the genetic information in contig 1 of UTAD17 is arranged differently in BJK_5C, as it is divided into three different contigs. Additionally, seven contigs of UTAD17 could not be rearranged with BJK_5C genome despite having a complete BLASTn match against it. B) Genomic sequence alignment of the seven UTAD17 contigs that could not be aligned with BJK_5C with the genome of UTAD17 reveals that they completely (full colouring), or partially (dashed colouring), align with other UTAD17 contigs, which may have interfered in the whole-genome alignment between the two strains.

III.4.4 Comparison of the ORFeomes of BJK_5C and UTAD17

Having the genomic sequence of the BJK_5C strain available, we undertook *in silico* annotation of its genome using the predicted ORFeome of the UTAD17 strain as reference. After manually curating the results suggested by the automatic gene finders, a total of 4,901 putative ORFs were predicted to be encoded by BJK_5C cells, which is in line with the number predicted for UTAD17 cells (**Table III. 3**). We started exploring the biological information emerging from this comparative analysis by focusing on aspects that were identified as characteristic of the *Sd. ludwigii* species while analysing the UTAD17 genome (described in chapter 1). These include the high number of genes involved in thiamine biosynthesis , which were not found in other yeasts of the *Saccharomycodaceae* family (*e.g*., *THI5*, encoding the first enzyme of thiamine biosynthesis, was present in 10 copies in the genome of UTAD17-Tavares et al., 2021); the set of beta-mannosyltransferases similar to those found in *C. albicans* or *Pichia pastoris*, but absent in other species more closely related to *Sd. ludwigii*; the genes involved in the catabolism of N-acetylglucosamine. All the genes necessary for thiamine biosynthesis were present in the genome of the BJK_5C strain, although at a lower copy number than in UTAD17 (Tavares et al., 2021). Specifically, while UTAD17 encoded 10 and 6 copies of the *THI13* and *THI20/21* genes; BJK_5C, encoded 5 and 2 copies of these genes. Like UTAD17, BJK_5C encodes four thiamine transporters (**Figure III. 6**). The genome of *Sd. ludwigii* NBRC1722 also revealed a similar multiplication of genes linked to thiamine biosynthesis (**Appendix Table III.2**) suggesting that this is a characteristic of the *Sd. ludwigii* species. While the physiological impact of this amplification remains unclear, it represents an interesting aspect for further studies.

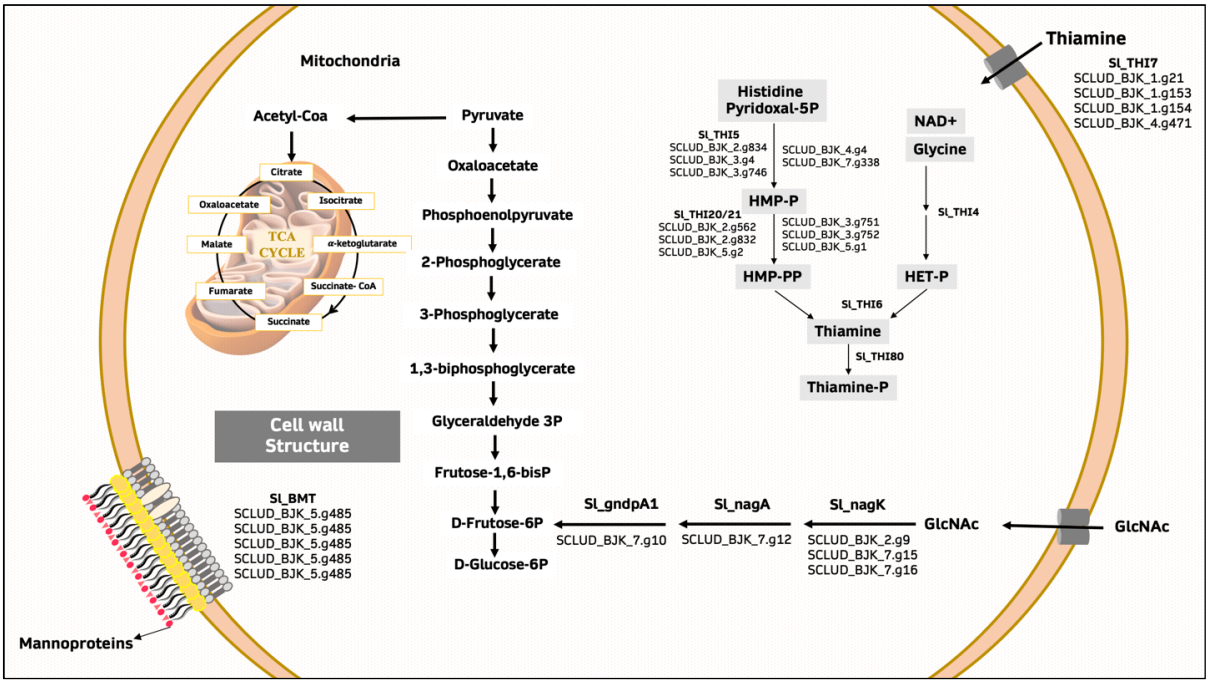


Figure III. 6– Metabolic aspects that were found to be specific to the genomes of *Sd. ludwigii* UTAD17, including the presence of genes involved in thiamine biosynthesis, in the catabolism of N-acetylglucosamine, and specific beta-mannosyltransferases., could also be found in the genome of strain BJK_5C.

We could also identify in the genome of the BJK_5C and of the NBRC172 strains orthologues for the beta-manosyltransferases that had been previously identified in the genome of UTAD17 (**Figure III. 6, Appendix Table III.2**). Interestingly, in all these strains (see **Appendix Table III.2**), these five predicted beta-mannosyltransferases genes appear arranged in tandem in the genome. Similarly, the genes required for the catabolism of N-acetylglucosamine were also detected in the genomes of BJK_5C and NBRC1722 (**Figure III. 6, Appendix Table III.2**). Notably, N-acetylglucosamine kinase was found in three copies in BJK_5C, while UTAD17 and in NBRC1722 only have two copies (**Appendix Table III.2**). The presence of N-acetylglucosamine induced morphological changes in UTAD17 cells (**Chapter II**). To determine if this same behavior could also be observed for other *Sd. ludwigii* strains, we conducted the same experiment with strain BJK_5C. The results confirmed that, at 37ºC, GlcNAc triggered formation of hyphae in BJK_5C cells, as it had been observed for the UTAD17 strain (**Figure III. 7**).

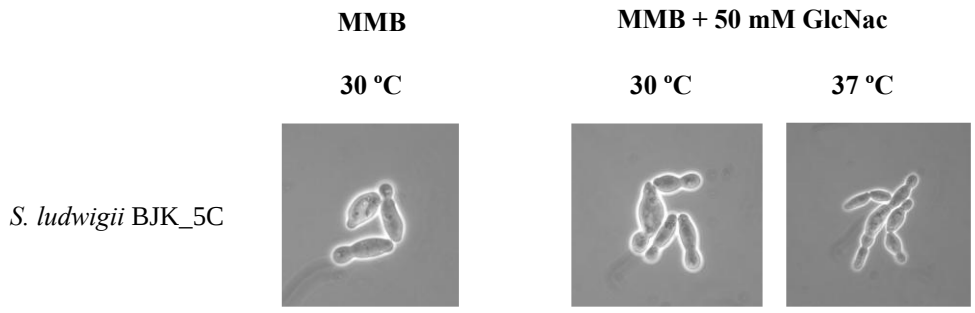
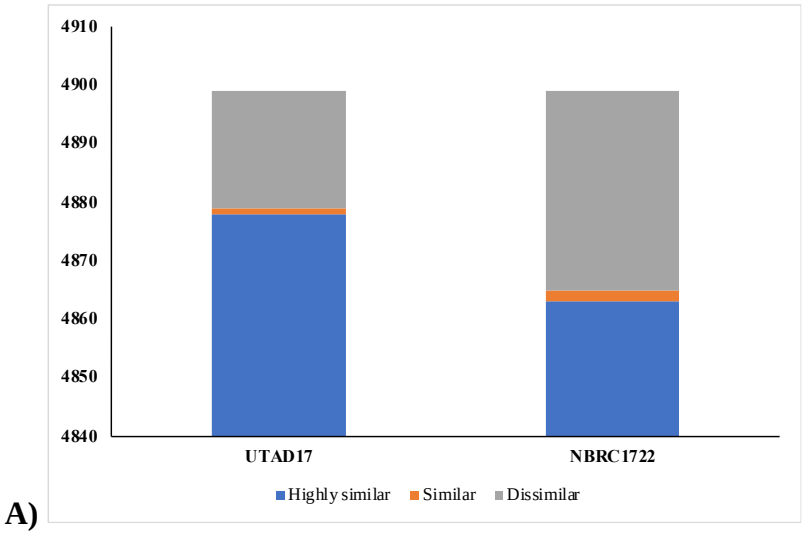


Figure III. 7– Induction of hyphae-like structures on *Sd. ludwigii* BJK_5C at 37ºC, caused by N-acetylglucosamine.

To pinpoint more general differences in the proteomes of BJK_5C and UTAD17 strains, we conducted a pairwise BLASTP analysis that allowed us to categorize the proteins into three groups: highly similar (e-value ≤ 1E-20 and identity> 50%); similar (e-value ≤ 1E-20 and identity between 30 % and 50%); and dissimilar (e-value > 1E-20 or no BLASTp hit) (**Figure III. 8A**). Most proteins predicted for BJK_5C (4,879) were similar to those predicted for UTAD17, with only 20 proteins being identified as dissimilar (**Figure III. 8A**). Among these 20 was included a predicted mobile viral element (SCLUD_BJK_1.g875), a carbohydrate-binding domain (SCLUD_BJK_2.g794), a Pac2 protein involved in reproduction (SCLUD_BJK_7.g170), among others (**Table III. 5**).

To further explore proteome variation across *Sd. ludwigii* strains, we conducted a similar ORFeome analysis for the NBRC1722 laboratory strain (Papaioannou et al., 2021). Of the proteins predicted to be encoded by the NBRC1722, 4,863 were highly similar to those encoded by BJK_5C, 2 were considered similar and 34 were dissimilar (**Figure III. 8A**). Among the dissimilar proteins, 8 were common to both UTAD17 and NBRC1722 (**Table III. 5**). For functional insights into these proteins, we performed an INTERPRO search combined with UNIPROT. This identified two proteins with defined functions: a urea active transporter (SCLUD_BJK_5.g556) and a heat shock protein (SCLUD_BJK_6.g9). Additionally, we identified one protein with a Zinc/Cysteine DNA binding domain, typically involved in carbohydrate catabolism such as galactose and maltose (SCLUD_BJK_4.g165); a cell wall protein (SCLUD_BJK_5.g153); and a protein involved in preribossome assembly or transport (SCLUD_BJK_7.g305). The remaining two proteins (SCLUD_BJK_1.g167 and SCLUD_BJK_3.g86) could not be assigned a function (**Figure III. 8B**). The complete list of dissimilar proteins is available in **Table III. 5**.



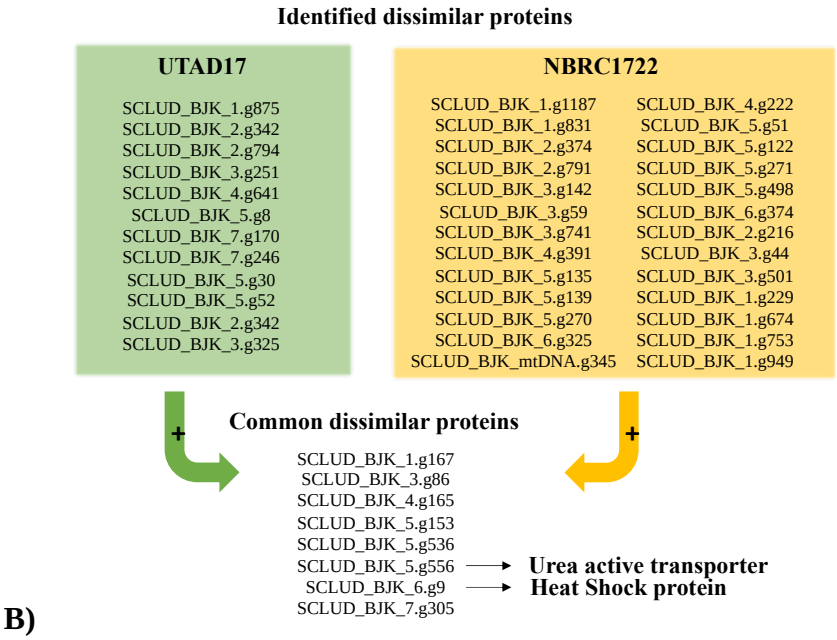


Figure III. 8- Results of the BLASTp search between the proteomes of *Sd. ludwigii* strain BJK_5C and strains UTAD17 and NBRC1722. A) From the selected ORFs that constitute the proteome of BJK_5C, 4,879 protein-encoding genes were considered highly similar, 20 were dissimilar and 1 was considered similar to proteins from UTAD17. The BLASTp search against NBRC1722 revealed 4,863 highly similar proteins, 34 dissimilar and 2 similar. B) A functional analysis revealed that, from the 5 proteins that did not have a hit against both *Sd. ludwigii* proteomes, only 2 had an assigned function (Urea active transporter and heat shock protein).

The comparative analysis of the *Sd. ludwigii* BJK_5C proteome with those of two different *Sd. ludwigii* strains, UTAD17 and NBRC1722, revealed several key insights. As expected, BJK_5C showed a higher similarity to UTAD17, another natural isolate, than to NBRC1722, a laboratory strain. Despite this, our analysis identified 8 unique proteins in BJK_5C that were dissimilar from both UTAD17 and NBRC1722 proteomes (**Table III. 5**). Although these sequences provided limited functional information, —with only 2 having assigned functions—they may still provide valuable insights into the genetic adaptation of

BJK_5C to its industrial environment. Furthermore, we could also identify dissimilar proteins in the proteomes of UTAD17 or NBRC1722. The differences observed can, however, result from issues occurred during sequencing or through different gene-finding algorithms that may have prevented gene detection.

Table III. 5– Complete list of dissimilar proteins between the proteomes of *Sd. ludwigii* BJK_5C and UTAD17 and NBRC1722, including their probable function and associated GO terms.

***Sd. ludwigii* BJK_5C proteins without homologue in the ORFeome of *Sd. ludwigii* UTAD17 and *Sd. ludwigii* NBRC1722**				
PROTEIN	**PROBABLE FUNCTION**	**INTERPRO FAMILY**	**INTERPRO DOMAIN**	**GO TERMS**
SCLUD_BJK_1.g167	X	X	X	X
SCLUD_BJK_3.g86	X	X	X	X
SCLUD_BJK_4.g165	The N-terminal region of a number of fungal transcriptional regulatory proteins contains a Cys-rich motif that is involved in zinc-dependent binding of DNA. The region forms a binuclear Zn cluster, in which two Zn atoms are bound by six Cys residues. A wide range of proteins are known to contain this domain. These include the proteins involved in arginine, proline, pyrimidine, quinate, maltose and galactose metabolism, amide and GABA catabolism, leucine biosynthesis, amongst others	X	Zn(2)Cys(6) fungal-type DNA-binding domain (IPR001138)	regulation of DNA-templated transcription (GO:0006355); zinc ion binding (GO:0008270); DNA-binding transcription factor activity, RNA polymerase II-specific (GO:0000981)
SCLUD_BJK_5.g153	Cell wall protein; Fungal-specific domain that contains eight cysteines and is found in some proteins with proposed roles in fungal pathogenesis	X	Extracellular membrane protein, CFEM domain (IPR008427)	X
SCLUD_BJK_5.g536	X	X	X	X
SCLUD_BJK_5.g556	Urea active transporter	Sodium/solute symporter (IPR001734); Urea active transporter (IPR031155)	X	transmembrane transport (GO:0055085); urea transmembrane transport (GO:0071918); transmembrane transporter activity (GO:0022857); urea transmembrane transporter activity (GO:0015204); membrane (GO:0016020)
SCLUD_BJK_6.g9	Heat shock protein 78, mitochondrial	X	X	X

SCLUD_BJK_7.g305	They are nucleolar proteins that contain a central domain consisting of ten repeats of acidic serine clusters alternating with lysine-, alanine- and proline-rich basic stretches; may be involved in preribosome assembly or transport	X	Srp40, C-terminal (IPR007718)	X

***Sd. ludwigii* BJK_5C proteins without homologue in the ORFeome of *Sd. ludwigii* UTAD17**

PROTEIN	PROBABLE FUNCTION	INTERPRO FAMILY	INTERPRO DOMAIN	GO TERMS
SCLUD_BJK_1.g875	Indicative of a mobile element such as a retrotransposon or retrovirus; retroviral insertion elements and lies just upstream of the integrase region on the polyproteins	Protein of unknown function DUF5314 (IPR035179)	Reverse transcriptase, RNA-dependent DNA polymerase (IPR013103); Integrase, catalytic core (IPR001584); GAG-pre-integrase domain (IPR025724)	DNA integration (GO:0015074); nucleic acid binding (GO:0003676)
SCLUD_BJK_2.g342	Microtubule-binding subcomplex of the outer kinetochore that is essential for proper chromosome segregation; mediates the formation and maintenance of bipolar kinetochore-microtubule attachments by forming closed rings around spindle microtubules and establishing interactions with proteins from the central kinetochore	DASH complex subunit Spc19 (IPR013251)	X	Attachment of spindle microtubules to kinetochore (GO:0008608); DASH complex (GO:0042729); spindle microtubule (GO:0005876)
SCLUD_BJK_2.g794	Putative carbohydrate binding domain. The domain contains up to eight conserved cysteine residues that may be involved in disulphide bridges	X	Carbohydrate-binding WSC (IPR002889)	X
SCLUD_BJK_3.g251	Function as an upstream regulator of the Fab1 lipid kinase pathway. The Fab1 lipid pathway is important for correct regulation of membrane trafficking events.	Vacuolar segregation subunit 7 (IPR024260)	X	X
SCLUD_BJK_4.g641	X	X	X	X
SCLUD_BJK_5.g8	X	X	X	X

SCLUD_BJK_7.g170	Promotes the onset of gluconate uptake upon glucose starvation. The Pac2 protein controls the onset of sexual development, by inhibiting the expression of ste11, in a pathway that is independent of the cAMP cascade	Gti1/Pac2 family (IPR018608)	X	X
SCLUD_BJK_7.g246	X	X	X	X
SCLUD_BJK_5.g30	U3 small nucleolar RNA-associated protein 15 - Involved in nucleolar processing of pre-18S ribosomal RNA. Required for optimal pre-ribosomal RNA transcription by RNA polymerase I together with a subset of U3 proteins required for transcription (t-UTPs).	X	X	X
SCLUD_BJK_5.g52	X	X	X	X
SCLUD_BJK_2.g342	DASH complex subunit SPC19; Microtubule-binding subcomplex of the outer kinetochore that is essential for proper chromosome segregation; mediates the formation and maintenance of bipolar kinetochore-microtubule attachments by forming closed rings around spindle microtubules and establishing interactions with proteins from the central kinetochore	DASH complex subunit Spc19 (IPR013251)	X	attachment of spindle microtubules to kinetochore (GO:0008608); spindle microtubule (GO:0005876); DASH complex (GO:0042729)
SCLUD_BJK_3.g325	X	X	X	X

***Sd. ludwigii* BJK_5C proteins without homologue in the ORFeome of *Sd. ludwigii* NBRC1722**

PROTEIN	PROBABLE FUNCTION	INTERPRO FAMILY	INTERPRO DOMAIN	GO TERMS
SCLUD_BJK_1.g1187	Major adhesin for controlling filamentous growth, mat, and biofilm formation of Baker's yeast	X	Flocculin 11 domain (IPR018789)	X

SCLUD_BJK_1.g831	Involved in the RNA-processing pathways; involved in numerous processes associated with RNA processing and gene expression regulation; important modulators of RNA biogenesis and function	Small nuclear ribonucleoprotein Sm D2 (IPR027248)	Sm domain, eukaryotic/archaea-type (IPR001163); Sm domain (IPR047575)	RNA splicing (GO:0008380); RNA binding (GO:0003723); small nuclear ribonucleoprotein complex (GO:0030532)
SCLUD_BJK_2.g374	May be involved in protein-protein interactions as well	X	High mobility group box domain (IPR009071)	X
SCLUD_BJK_2.g791	Involved in sporulation and affects the sphingolipid composition of the plasma membrane	Membrane protein SUR7/Rim9-like, fungi (IPR009571)	Ubiquitin-like domain (IPR000626); Ubiquitin domain (IPR019956)	protein binding (GO:0005515)
SCLUD_BJK_3.g142	Ribosome biogenesis protein	X	RNA recognition motif domain (IPR000504)	nucleic acid binding (GO:0003676); RNA binding (GO:0003723)
SCLUD_BJK_3.g59	Involved in sporulation and affects the sphingolipid composition of the plasma membrane	Membrane protein SUR7/Rim9-like, fungi (IPR009571)		plasma membrane (GO:0005886)
SCLUD_BJK_3.g741	X	X	X	X
SCLUD_BJK_4.g391	Mitochondrial intermembrane chaperone that participates in the import and insertion of multi-pass transmembrane proteins into the mitochondrial inner membrane. Also required for the transfer of beta-barrel precursors from the TOM complex to the sorting and assembly machinery (SAM complex) of the outer membrane. Acts as a chaperone-like protein that protects the hydrophobic precursors from aggregation and guide them through the mitochondrial intermembrane space.	X	Tim10-like (IPR004217)	X
SCLUD_BJK_5.g135	E3 ubiquitin-protein ligase that mediates monoubiquitination of histone H2B to form H2BK123ub1 in association with the E2 enzyme RAD6/UBC2. H2BK123ub1 gives a specific tag for epigenetic transcriptional activation, elongation by RNA polymerase II, telomeric silencing, and is also a prerequisite for H3K4me and H3K79me formation. It thereby plays a central role in histone	E3 ubiquitin ligase Bre1 (IPR013956)	Zinc finger, RING-type (IPR001841)	Histone monoubiquitination (GO:0010390); ubiquitin-protein transferase activity (GO:0004842)

	code and gene regulation. Also modulates the formation of double-strand breaks during meiosis.			
SCLUD_BJK_5.g139	Carbohydrate-binding domain found in fungal adhesins (also referred to as agglutinins or flocculins)	X	PA14/GLEYA domain (IPR037524); GLEYA adhesin domain (IPR018871)	flocculation (GO:0000128)
SCLUD_BJK_5.g270	Transcription activator involved in the regulation of genes expressed in response to environmental changes. When overexpressed it activates transcription of the multidrug resistance ABC transporter PDR5, thus conferring resistance to the fungicide fluconazole (FCZ) and cycloheximide. When overexpressed, it also confers, independent of PDR5, increased resistance to 4-nitroquinoline-N-oxide (4-NQO)	X	Basic-leucine zipper domain (IPR004827)	regulation of DNA-templated transcription (GO:0006355); DNA-binding transcription factor activity (GO:0003700)
SCLUD_BJK_6.g325	X	X	X	X
SCLUD_BJK_mtDNA.g345	Cytochrome c oxidase subunit 2	Cytochrome c/quinol oxidase subunit II (IPR045187)	Cytochrome c oxidase subunit II-like C-terminal (IPR002429)	X
SCLUD_BJK_4.g222	60S ribosomal protein L12-B	Ribosomal protein L11/L12 (IPR000911)	Ribosomal protein L11, N-terminal (IPR020784); Ribosomal protein L11, C-terminal (IPR020783)	translation (GO:0006412); structural constituent of ribosome (GO:0003735); ribosome (GO:0005840)
SCLUD_BJK_5.g51	Putative nitroreductase	X	X	X
SCLUD_BJK_5.g122	X	X	X	X
SCLUD_BJK_5.g271	Catalyses the posttranslational formation of 4-hydroxyproline in -xaa-pro-gly-sequences in collagens and other proteins. Prokaryotic enzymes might catalyse hydroxylation of antibiotic peptides. These are 2-oxoglutarate-dependent dioxygenases, requiring 2-oxoglutarate and dioxygen as cosubstrates and ferrous iron as a cofactor	Prolyl 4-hydroxylase (IPR045054)	Prolyl 4-hydroxylase, alpha subunit (IPR006620); Prolyl 4-hydroxylase alpha subunit, Fe(2+) 2OG dioxygenase domain (IPR044862); Oxoglutarate/iron-dependent dioxygenase (IPR005123)	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen (GO:0016705); iron ion binding (GO:0005506); L-ascorbic acid binding (GO:0031418)
SCLUD_BJK_5.g498	60S ribosomal protein L37-A	Ribosomal protein L37e (IPR001569)		translation (GO:0006412); structural constituent of ribosome (GO:0003735); ribosome (GO:0005840)

SCLUD_BJK_6.g374	X	X	X	X
SCLUD_BJK_2.g216	X	X	X	X
SCLUD_BJK_3.g44	Zinc metabolism: altering membrane sterol content or by directly altering cellular zinc levels.	X	X	X
SCLUD_BJK_3.g501	X	X	X	X
SCLUD_BJK_1.g229	DNA damage repair and genome maintenance	Centromere protein X (IPR018552)		kinetochore assembly (GO:0051382); DNA repair (GO:0006281); protein heterodimerization activity (GO:0046982)
SCLUD_BJK_1.g674	X	X	X	X
SCLUD_BJK_1.g753	Transcription regulation		SWIB/MDM2 domain (IPR003121); SWIB domain (IPR019835)	
SCLUD_BJK_1.g949	Uncharacterized protein YGL108C	X	X	X

To complement the comparative analysis of the *Sd. ludwigii* ORFeomes, a high-resolution analysis, based on identification of Single Nucleotide Polymorphisms (SNPs), was performed. For this, the Illumina reads obtained from the BJK_5C genome were mapped to the genome sequence of the UTAD17 strain (as detailed in materials and methods) resulting in the identification of 51,505 SNPs; 23,439 of which were in coding sequences; and 9,888 corresponding to non-synonymous SNPs. After filtering these SNPs to match the designated quality parameters, the number of non-synonymous SNPs that resulted in amino acid changes was 3,872. In **Figure III. 9**, we show the distribution of these non-synonymous SNPs in the protein pairs, being particular noticeable the existence of a higher degree of changes in genes predicted to encode flocculins (*e.g.*, SCLUD10.g193, and SCLUD7.g3, each harbouring 16 SNPs, compared to the alleles encoded by the UTAD17 strain), SCLUD1.g770, an orthologue of *S. cerevisiae* nucleoporin NUC159; and SCLUD5.g229, an orthologue of *S. cerevisiae* signalling mucin HKR1 (**Figure III. 9**). For future works, it would be interesting to study if these gene modifications directly impact, or not, the function of their corresponding proteins and, consequently, modify the phenotype. In *S. cerevisiae*, adhesins of the Flo family are involved in cell flocculation and play an important role in formation of filaments and biofilms, allowing the transition from a planktonic lifestyle to more structured and multicellular structures (Willaert et al., 2021). Mucin HXK1 is also involved in the activation of a pathway leading to filamentous growth, regulating, in *S. cerevisiae*, both cell wall beta-glucan synthesis and budding pattern in *S. cerevisiae* (Yabe et al., 1996). Interestingly, a study performed by Coi et al (2017) with *S. cerevisiae* wine strains also revealed a high number of mutations in the gene *FLO11* (encoding a flocculin), compared to the corresponding alleles encoded by flor strains. Further studies will be required to dissect what are the functional outcomes (if any) of the different alleles of flocculins encoded by BJK_5C and by UTAD17, for example, in the ability of the cells to form biofilms, a trait that may be relevant in the context of *Sd. ludwigii* spoilage capacity that has been linked to the formation of cloudiness (corresponding to cellular aggregates or biofilms). It is important to denote, however, that the sub-telomeric location of these genes (strongly suggested by their location on the contig ends, SCLUD1.g3 and SCLUD10.g193), potentiates a high level of mutations in these proteins, since chromosome ends are more prone to elevated rates of meiotic recombination, sister chromatid exchange, DNA transfer and also to sequencing mistakes and assembly issues (Luo et al., 2011).

The establishment of what is the outcome of the SNPs found across the *Sd. ludwigii* ORFeome is difficult, except for proteins that might become truncated due to frameshifts or

premature STOP codons. Thus, we focused attention on these and found 13 genes exhibiting frame-shifting mutations leading to premature truncations in BJK_5C: two flocculins (SCLUD7.g3 and SCLUD17.g359), a pheromone-processing carboxypeptidase (SCLUD1.g368), a mannosyl-oligosaccharide glucosidase (SCLUD4.g637), an RNA polymerase-associated protein LEO1 (SCLUD4.g637), a vacuolar protein sorting-associated protein 5 (SCLUD.g380), and a Nucleosome assembly protein NAP1 (SCLUD7.g139) could also be identified (**Table III. 6**). Although SNP identification in polyploids can be more challenging than in haploids or diploids, due to the need to distinguish homoeologous SNPs (polymorphic positions occurring across subgenomes within and among individuals) from allelic SNPs (polymorphic positions occurring within a single subgenome among individuals), the complete genome sequencing performed ensures the obtention of reads from all the copies of each gene. Taking a closer look into the SNPs herein identified, it can be useful to determine their real phenotype in the cells, especially in those that appear completely truncated (*e.g.*, SCLUD3.g14 that is truncated at 11 amino acids).

Table III. 6– Single Nucleotide polymorphisms (SNPs) exhibiting frame-shifts identified in the gene sequences encoded by *Sd. ludwigii* BJK_5C when compared to strain UTAD17. *n.a* – Function not available

UTAD17 Contig	ORF name	Probable function	Non-synonymous SNPs in BKJ_5C compared to UTAD17
SCLUD1	SCLUD1.g268	n.a	Glu481fs
SCLUD1	SCLUD1.g368	Pheromone-processing carboxypeptidase KEX1	Asn689fs
SCLUD2	SCLUD2.g650	Probable endonuclease	Asn636fs
SCLUD2	SCLUD2.g669	Uncharacterized transporter SLY41	Arg174fs
SCLUD3	SCLUD3.g14	Hypothetical protein	Phe11fs
SCLUD3	SCLUD3.g16	Hypothetical protein	Pro76fs
SCLUD4	SCLUD4.g637	Mannosyl-oligosaccharide glucosidase	Asn343fs
SCLUD5	SCLUD5.g307	RNA polymerase-associated protein LEO1	Glu144fs
SCLUD5	SCLUD5.g380	Vacuolar protein sorting-associated protein 5	Gln292fs
SCLUD5	SCLUD5.g445	Hypothetical protein	Lys203fs
SCLUD7	SCLUD7.g3	Flocculation protein FLO1	Phe3017fs
SCLUD7	SCLUD7.g139	Nucleosome assembly protein NAP1	Ser439fs
SCLUD17	SCLUD17.g359	Flocculation protein FLO1	Thr31fs

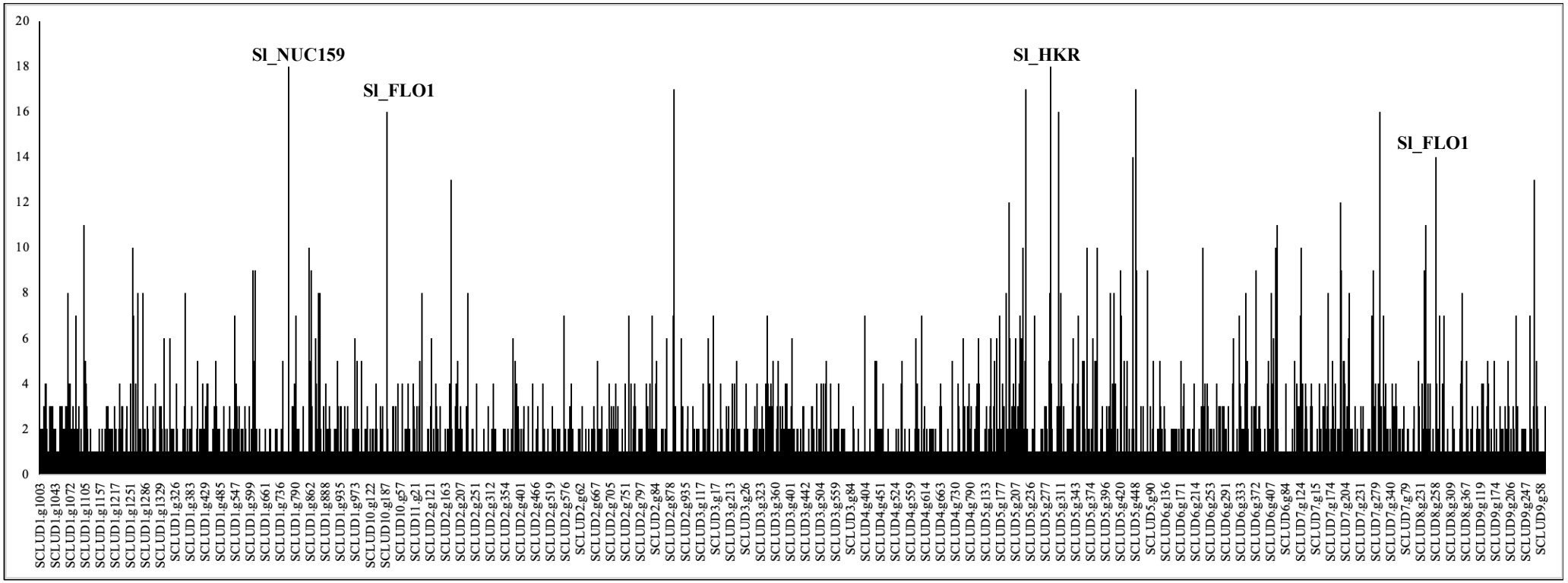


Figure III. 9- Non-synonymous Single Nucleotide Polymorphisms (SNPs) between *Sd. ludwigii* BJK_5C and *Sd. ludwigii* UTAD17, being identified in bold the genes which accumulated a higher number of SNPs (orthologues to ScFLO1, ScNUC159 and ScHKR).

III.4.5 Genotype-Phenotype-Ecotype associations in the genome of *Sd. ludwigii* BJK_5C

In this section we investigated whether certain differential phenotypes that were observed for UTAD17 and BJK_5C strains could be attributed to differences in proteins linked to these phenotypes. To recapitulate the main findings that had been described above for the two strains a schematic representation was built (**Figure III. 10**).

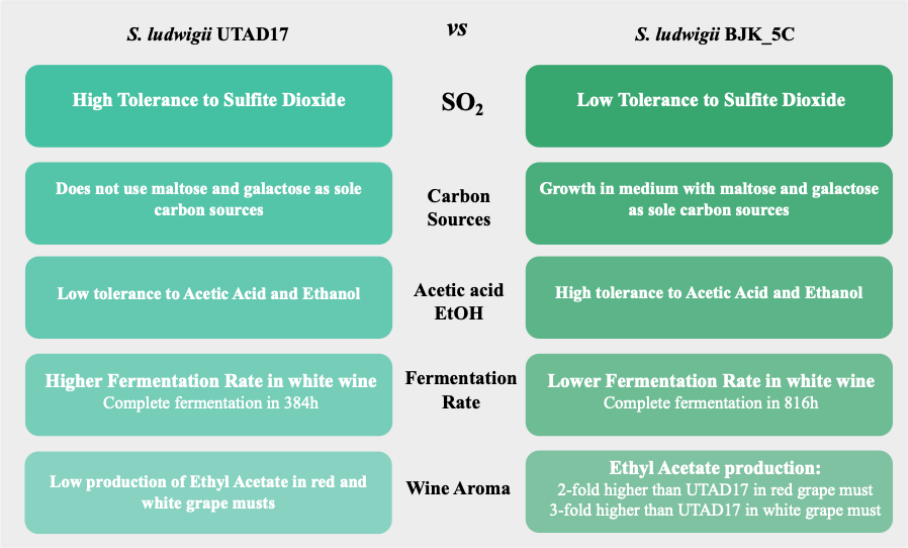


Figure III. 10– Differences in *Sd. ludwigii* phenotypes: strain UTAD17 showed higher tolerance to sulfur dioxide than BJK_5C. In opposition, BJK_5C demonstrated a higher tolerance than UTAD17 to increased concentrations of acetic acid and ethanol. Additionally, wines fermented with both strains resulted in different outcomes in terms of fermentation rates and aroma properties: UTAD17 took half of the time of BJK_5C in completely fermenting white wine. The fermentation produced by BJK_5C resulted in unpleasant concentrations of ethyl acetate, while those obtained by UTAD17 were more neutral.

III.4.5.1 Maltose and Galactose as sole carbon sources

The phenotypic analysis undertaken showed that *Sd. ludwigii* BJK_5C cells are able to use maltose and galactose as sole carbon sources, while strain UTAD17 does not grow on these sources (**Figure III. 1**). In yeasts, maltose assimilation usually requires a maltose permease (MALT), a maltase to undertake cleavage of the two glucose monomers (MALS) and a transcriptional activator for these genes (MALR) (Wanke et al., 1997; Novak et al., 2004). We could find orthologues for these genes in both BJK_5C and UTAD17 cells (

Table III. *7*), but only BJK_5C could grow on maltose as a sole carbon source (**Figure III. 1**). These findings suggest that the utilization of maltose by these yeasts might be differently regulated in both strains. However, additional experimental data, such as experiments of direct consumption coupled with transcriptional studies of the *MALT* (maltose transporter) and *MALS* (maltase) genes (that are repressed in the presence of glucose and induced in the presence of maltose) (Novak et al., 2004) should be conducted to draw proper conclusions on this aspect. As for galactose utilization, it follows the Leloir pathway that converts galactose into glucose-6-phosphate, that is subsequently mobilized via glycolysis (Sellick et al., 2008; Holden et al., 2003). The enzymes required for this are galactose mutarotase (GAL10), galactokinase (GAL1), galactose-1-phosphate uridyltransferase (GAL7), UDP-galactose-4-epimerase (GAL10), and phosphoglucomutase (PGM1/PGM2) (Sellick et al., 2008; Holden et al., 2003). Both BJK_5C and UTAD17 strains encode homologues for 5 of these six enzymes, but it was not possible to detect in any of them the GAL7 gene (**Figure III. 11,**

Table III. *7*). Thus, while this absence is consistent with the lack of strain UTAD17 to consume galactose, it is contradictory with the capability of strain BJK_5C to do so. One possible explanation for this phenotype is the potential presence of a GAL7-like gene in the genome of BJK_5C, enabling its growth.

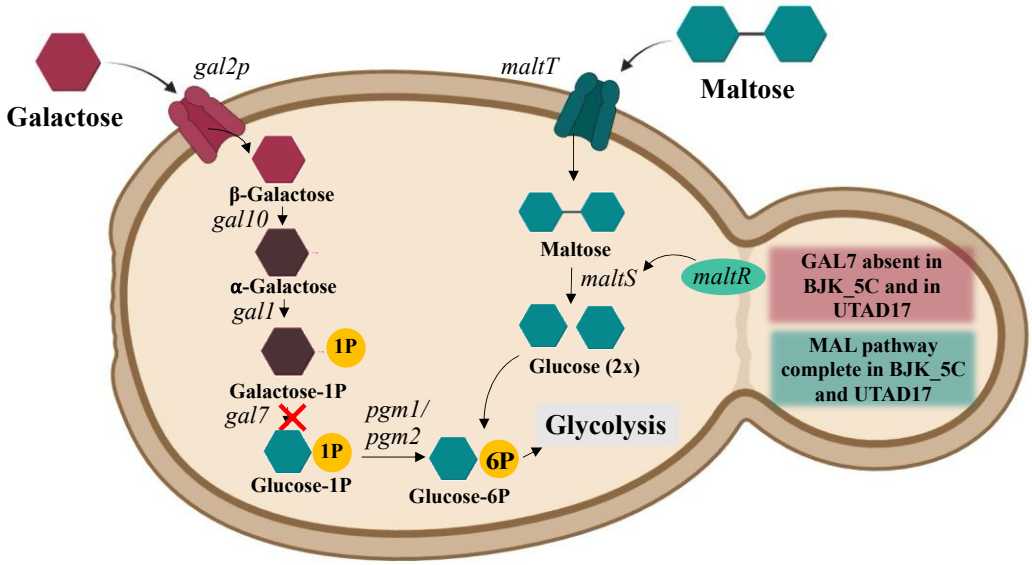


Figure III. 11– Maltose and Galactose degradation pathways in yeast. The comparative genomic analysis of both *Sd. ludwigii* strains, BJK_5C and UTAD17, revealed that both strains are equipped with the same genomic machinery for the degradation of galactose and maltose. While maltose degradation pathway appears to be complete in both strains, the galactose degradation pathway is missing the GAL7 gene, encoding a galactose-1-phosphate uridyltransferase responsible for converting the galactose-1P into glucose-1P.

Table III. 7– Genes involved in the degradation of maltose and galactose and their presence/absence in the genomes of *Sd. ludwigii* strains BJK_5C and UTAD17. These genes were double-confirmed using BLASTp against UNIPROT and BLAST Koala (KO number). *n.a* – not available

Pathway	Gene	KO	Description	*Sd. ludwigii* BJK_5C gene	*Sd. ludwigii* UTAD17 gene
Maltose degradation	*malt/malt61*	*n.a*	Maltose permease	SCLUD_BJK_7.g13	SCLUD7.g9
	mals/mal12	K00700	1,4-alpha-glucan-branching enzyme	SCLUD_BJK_4.g418	SCLUD5.g33
	malr/mal3	K10436	Microtubule integrity protein	SCLUD_BJK_7.g218	SCLUD7.g219
Galactose degradation (Leloir pathway)	*gal10*	K07748	Mutarotase	SCLUD_BJK_2.g661	SCLUD4.g645
	gal1	K00869	Galactokinase	SCLUD_BJK_2.g42	SCLUD6.g78
	gal7	K00965	Galactose-1P-Uridyltransferase	n.a	n.a
	gal10	K07748	UDP-glucose-4-epimerase	SCLUD_BJK_2.g661	SCLUD4.g645
	pgm1/pgm2	K01835	Phosphoglucomutase	SCLUD_BJK_5.g355 SCLUD_BJK_1.g1183	SCLUD3.g215 SCLUD9.g215

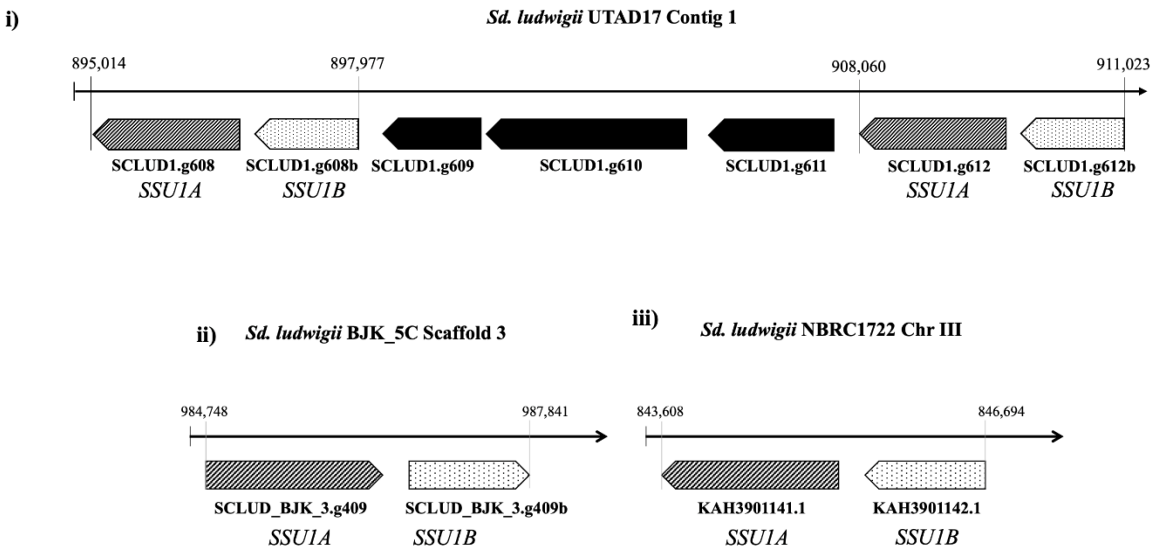
III.4.5.2 Tolerance to acetic acid, ethanol and SO_2

As demonstrated in the first section of this chapter, *Sd. ludwigii* BJK_5C cells grew considerably better in the presence of stressful concentrations of acetic acid or ethanol, compared to UTAD17. The *Sd. ludwigii* genes that determine tolerance to these stressors are not known. Interestingly, the orthologues of *S. cerevisiae* genes that are involved in acetic acid and ethanol tolerance, such as aldehyde and alcohol dehydrogenase (Trivic et al., 2000; Remize et al., 2000; Chakraboti et al., 2009), were found to be present in both *Sd. ludwigii* strains (**Appendix Table III.4**). Since no genomic information regarding *S.ludwigii*'s tolerance mechanisms to these stresses is currently available, it would be valuable to investigate whether these orthologues play a similar role in response to acetic acid and ethanol exposure. Additionally, examining the composition of the cell membrane and cell wall could be crucial, as these structures may influence tolerance to changing environmental conditions. External factors, such as extreme pH, can induce structural modifications that enhance resistance to environmental stress (Snoek et al., 2016). For instance, changes in unsaturated fatty acid composition or ergosterol content of the membrane may contribute to increased ethanol tolerance, as ethanol directly affects membrane fluidity and structure, thereby affecting membrane proteins (Snoek et al., 2016).

Unlike what was observed to occur under stress caused by ethanol or acetic acid, *Sd. ludwigii* UTAD17 cells were more tolerant to SO_2, compared to BJK_5C. Such phenotypic observation may have something to do with the lower concentrations of sulfite that are used in cider vinegar production (between 10 and 50 ppm - Almeida et al., 2018), while in wines this may reach 350 ppm (Pisoschi & Pop, 2018). In **Chapter II** we reported the existence of two genes in the UTAD17 genome encoding for predicted sulfite efflux pumps (*SSU1A* and *SSU1B*). A closer look into the genomic arrangement of these genes showed us that they are, in fact, duplicated with the two copies being arranged in tandem. In the genome of BJK_5C we could also identify two *SSU1* genes (*SSU1A* and *SSU1B*) arranged in tandem (**Figure III. 12**). However, in this case, we could not find a duplication of this locus. Thus, it is possible to hypothesize that the expression levels of *SSU1* genes in the BJK_5C are lower than those achieved in UTAD17 cells, which may have consequences in tolerance to SO_2. Further comparison of the Ssu1A protein sequences showed that the BJK_5C, (SCLUD_BJK_3.g409) allele is predicted to be slightly larger than its UTAD17 counterpart, with 34 more amino acids. On the other hand, the *SSU1B* gene (SCLUD_BJK_3.g409b) is the same length as its counterpart in UTAD17. The *SSU1A* gene locus in NBRC1722 genome has a structure similar

to the observed for BJK_5C, with only two *SSU1* genes (one *SSU1A* and one *SSU1B*), arranged in tandem, being both the same size as those in BJK_5C (**Figure III. 12**).

A)



B)

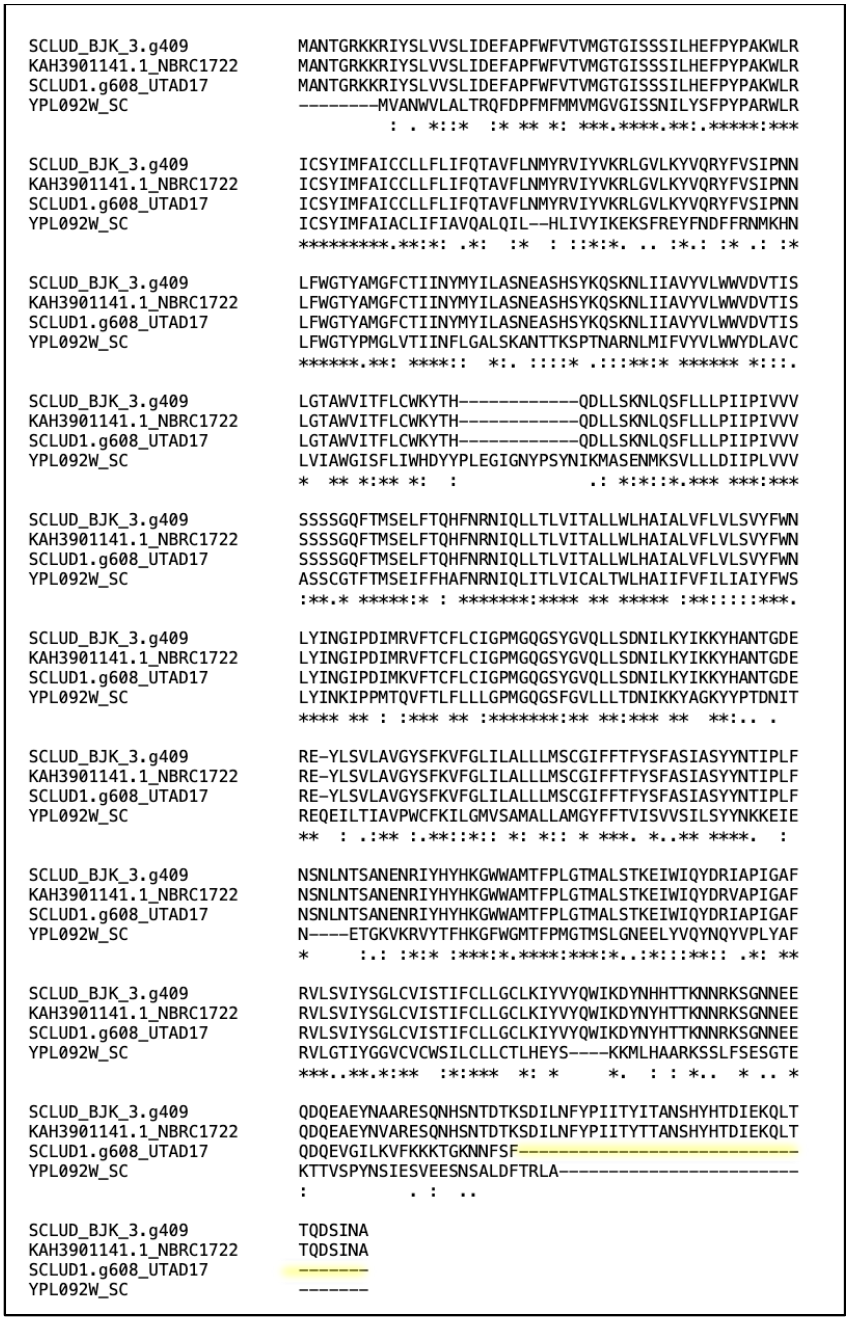


Figure III. 12 – The sulfite efflux pumps of *Sd. ludwigii*. A) Genomic location of the *SSU1* genes in three *Sd. ludwigii* strains UTAD17 (i), BJK_5C (ii) and NBRC1722 (iii) shows a duplication of these genes in UTAD17. B) ClustalW alignment of the predicted Ssu1 sequences in UTAD17, BJK_5C, NBRC1722 and *S. cerevisiae*

The duplication of the region encompassing the two *SSU1* genes in UTAD17 prompted us to investigate the surrounding genomic sequences further. When comparing the genomes of UTAD17, BJK_5C and NBRC1722, we found a common structure, involving the predicted genes for the *S. cerevisiae* orthologues *DYN1* (coding for a dynein heavy chain), the two *SSU1* genes (A and B), and *LCB3* (coding for a dihydrosphingosine-1-phosphate phosphatase). Interestingly, in UTAD17 we have also discovered three additional genes (SCLUD1.g607, SCLUD1.g610, SCLUD1.g611) that have been previously annotated but did not have any hits

against the UNIPROT database and were absent from the corresponding region in the other two *Sd. ludwigii* strains. These genes represent three copies of the same putative gene (designated by *GENEX*), although SCLUD1.g607 encodes only half of the sequence coded by SCLUD1.g607 and SCLUD1.g611 (**Appendix Table III.5**).

We propose a structure for the gene cluster encompassing the sulfite efflux pump in *Sd. ludwigii*, as presented in **Figure III. 13**. Observing this figure, it appears that in UTAD17, a duplication event of the genomic region including *SSU1A*, *SSU1B* and *LCB3* must have occurred, resulting in the existence of two combinations of these genes downstream from *DYN1* (**Figure III. 13**). Moreover, the absence of the *GENEX* in BJK_5C and NBRC1722, coupled with the complete lack of BLASTx matches against the NCBI and UNIPROT databases, suggests that these sequences may result from sequencing errors or might be genomic artifacts specific to the UTAD17 genome.

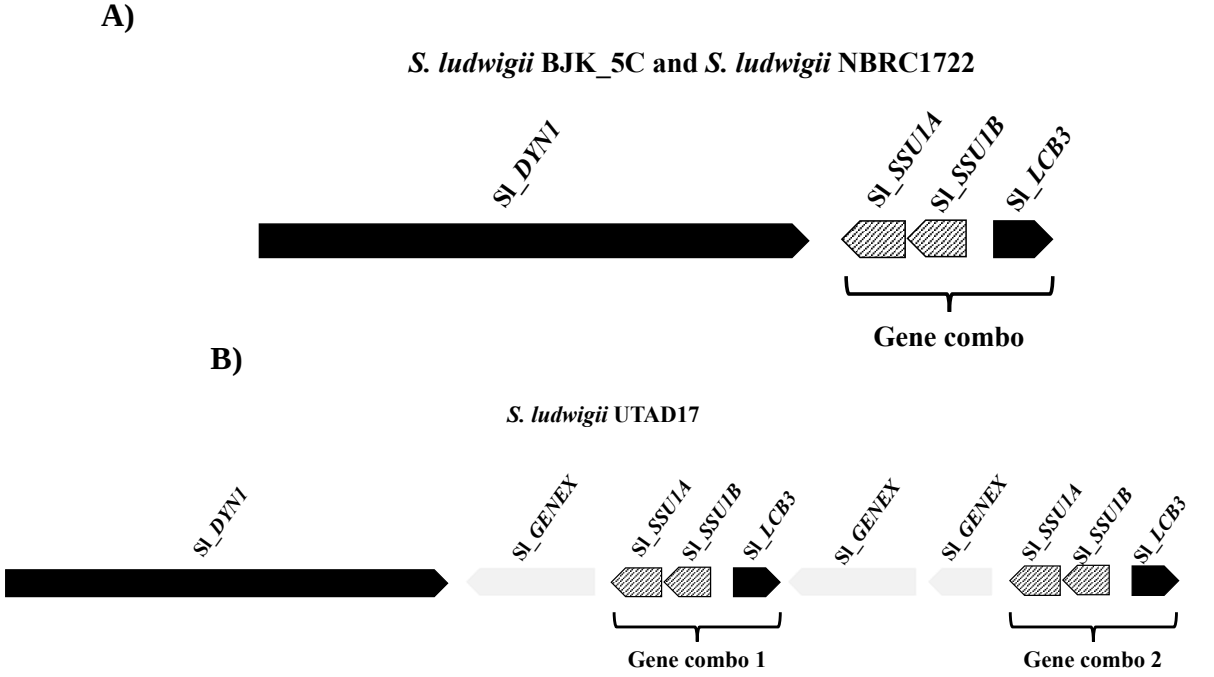


Figure III. 13- Proposed structure for the gene cluster encompassing the sulfite efflux pump in *Sd. ludwigii* . A) Gene cluster of *Sd. ludwigii* BJK_5C and NBRC1722 presents only one Gene combo downstream from gene *DYN1* (*SSU1A*, *SSU1B*, and *LCB3*). B) Gene cluster of *Sd. ludwigii* UTAD17 presents a duplication of the gene combo downstream from gene *DYN1*.

III.4.5.3 Enzymes involved in the synthesis of aroma volatiles

Vinification of Touriga Nacional and Sauvignon Blanc grape musts by BJK_5C or UTAD17 cells resulted in a distinct aroma bouquet, with emphasis on the considerably higher production of ethyl acetate by BJK_5C cells, a trait that characterizes other *Sd. ludwigii* strains but that was not observed for UTAD 17 (**Figure III. 3)**. Ethyl acetate results from the activity of alcohol acetyltransferases (AATs) that use acetyl-CoA and ethanol, resulting in release of free CoA (Bohnenkamp et al., 2020). Two types of ethyl acetate-producing AAT enzymes have been described in *S. cerevisiae*: Sc*EAT1,* the main contributor for ethyl acetate production (Kruis et al., 2017; Kruis et al., 2020), and *ScATF* (Lilly et al., 2000; Fukuda et al., 1998), and *ScIMO32* (Kruis et al., 2017). The genomes of *Sd. ludwigii* strains BJK_5C and UTAD17 encode three genes predicted to be alcohol acetyltransferases (**Table III. 8**). Two of these are similar to those described in *S. cerevisiae*: one analogous to *ScEAT1,* and the other to a probable alcohol acetyltransferase encoded by *ScIMO32* (Kruis et al., 2017). The third protein is similar to an ethanol acetyltransferase described in *K. marxianus*, which is interesting since in *K. marxianus* is able to produce substantial amounts of ethyl acetate (Kruis et al., 2018). No differences in these three enzymes were found between strains UTAD17 and BJK_5C (**Appendix Table III.3**). Therefore, we hypothesize that the robust difference found concerning ethyl acetate production between these two *Sd. ludwigii* strains can result from differences in gene expression, eventually combined with differences in the internal concentration of the metabolites required for its synthesis (**Figure III. 14**). Further studies will be required to further investigate this matter.

Table III. 8- Alcohol acetyltransferase enzymes from *Sd. ludwigii*. Both *Sd. ludwigii* strains (UTAD17 and BJK_5C) have 3 genes encoding Alcohol acetyltransferases, with similar corresponding proteins. The genes related to BJK_5C are in bold.

AAT from *Sd. ludwigii*	UNIPROT HIT	E-value	Query coverage (%)	Identity (%)
SCLUD_BJK_2.g179	EAT1_KLUMD Ethanol acetyltransferase 1 (*Kluyveromyces marxianus*)	4.89E-145	80.20	60.2
SCLUD6.g215				
SCLUD_BJK_2.g717	EAT1_YEAST Ethanol acetyltransferase 1 (*Saccharomyces cerevisiae*)	5.12E-42	99.29	35.7
SCLUD4.g700				
SCLUD_BJK_1.g697	IMO32_YEAST Probable alcohol acetyltransferase (*Saccharomyces cerevisiae*)	1.14E-83	82.32	43.8
SCLUD2.g693				

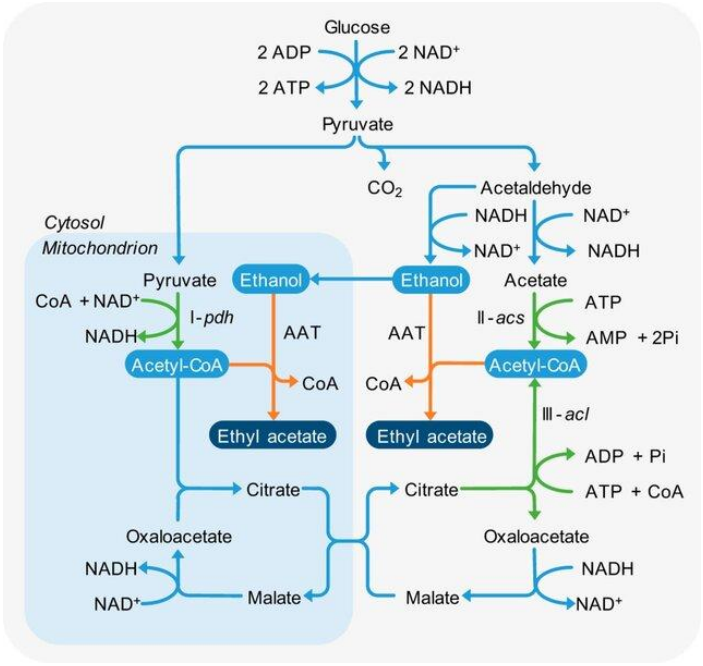


Figure III. 14 Potential pathways of ethyl acetate production in yeast, through the activity of alcohol acetyltransferase (AAT) enzymes. The reactions catalyzed by AATs are indicated in orange, while the the three reactions forming acetyl-CoA during glucose catabolism are shown in green. Reaction I, pyruvate dehydrogenase (pdh); reaction II, acetyl-CoA synthetase (acs); and reaction III, ATP citrate lyase (acl). Adapted from Kruis et al., (2018).

III.5 Discussion

Sd. ludwigii is widely known for its high spoilage capacity in stored wines, being able to endure high concentrations of sulfur dioxide and often resulting in the excessive production of ethyl acetate and acetaldehyde (Vejarano et al., 2018; Esteves et al., 2019; Tavares et al., 2021). The sequencing and analysis of the first genome of a grape-must-native *Sd. ludwigii* strain (UTAD17), as detailed in **Chapter II**, allowed us to point out important traits that had not been described for this species (and being more interesting because they were not found in the sister *Hanseniaspora* species) including the presence of enzymes required for the catabolism of N-acetylglucosamine (and the consequent observations in morphology prompted by this carbohydrate); a set of five beta-mannosyltransferases and an expansion of genes required for thiamine biosynthesis and of the predicted sulfite efflux pump *SSU1* (Tavares et al., 2021).

In this chapter we moved this along by focusing the genome of other *Sd. ludwigii* strains, BJK_5C, isolated from an industrial vinegar production environment, and the laboratorial strain NBRC1722 (meiotic descendant of wild-type isolate NBRC 1721, isolated from wine must, in Japan - Papaioannou et al., 2021), enabling inter-strain comparisons within the *Sd. ludwigii* species. In all three *Sd. ludwigii* strains, the *SSU1* genes are flanked by *DYN1*, encoding a dynein, and *LCB3*, encoding a dihydrosphingosine 1-phosphate phosphatase. However, while in UTAD17 the region encompassing the two *SSU1* genes and *LCB3* was found to be duplicated (with the duplicates arranged in tandem), in BJK_5C and NBRC1722, only one copy of the *SSU1* locus was observed. These results strongly point to the occurrence of local duplication event in strain UTAD17, being tempting to speculate whether that has been prompted to occur considering the wide application of sulfite in wine-making that may have forced the strains to evolve efficient adaptive responses. Indeed, in *S. cerevisiae*, such adaptive responses have also been described to occur involving *SSU1*, being considered one of the most effective resistance mechanisms in *S. cerevisiae* (Zara & Nardi, 2021; Park & Bakalinsky, 2000; Avram & Bakalinsky, 1997), as it is essential for efficient sulfite efflux. High polymorphism in SSU1 gene and chromosomal rearrangements, such as translocations between chromosomes VIII and XVI, led to an increase *SSU1* expression, providing an adaptive advantage in sulfite-rich wine environments (Yuasa et al., 2004; Pérez-Ortín et al., 2002). This highlights the selective pressure sulfites impose, driving genetic adaptations in *S. cerevisiae* strains from the wine ecosystem (Zara & Nardi, 2021). While such genomic adaptations are less frequently documented in NSYs, the duplication of the *SSU1* gene in *Sd. ludwigii* marks

the first noted adaptation of this nature, indicating potential similar responses in other NSYs within winemaking contexts.

Interestingly, gene duplication has also apparently occurred in the genes involved in thiamine biosynthesis: *THI13*, which initiates the thiamine biosynthetic pathway, is present in 10 copies in UTAD17 and in 5 copies in BJK_5C. In eukaryotes, gene duplication is recognized as an important source of new genes and a significant foundation for genetic variation (Ames et al., 2010; Conant & Wolfe, 2008). However, it is not always possible to pinpoint the direct phenotypic consequences of these events (Ames et al., 2010). Still, several studies in yeasts suggested that gene duplication events often led to functional improvement, as recognized across several species (Ames et al., 2010). For example, *S. cerevisiae* and *C. glabrata* share a common ancestor that underwent a whole-genome duplication event approximately 100 million years ago (Wolfe & Shields, 1997; Kellis et al., 2004). This event, followed by gene loss and specialization, resulted in significant evolutionary and functional innovation in both species (Ames et al., 2010; Poláková et al., 2009). Duplications of large genomic segments are also common among yeasts, for example, in *C. glabrata*, segmental duplications have led to the creation of entirely new chromosomes (Poláková et al., 2009). Importantly, the new genes generated by these events are often found in tandem sets (Dujon et al., 2004), as it had been observed for the *SSU1* and thiamine genes. These gene-duplication events are lineage-specific and cause variability among strains of the same species (Ames et al., 2010), usually in an environment-dependent manner (Ames et al., 2010). The obtained results suggest that duplication of *SSU*'s and of the thiamine genes may have occurred while *Sd. ludwigii* UTAD17 strain adapted to the challenging environment of grape musts, although this is still a hypothesis that lacks some confirmation. Although it had not been yet demonstrated that the Ssu1 efflux pumps from *Sd. ludwigii* also play a role in tolerance of this species to SO_2, as it has been already extensively shown for its *S. cerevisiae* counter-partners, the observation that UTAD17 is considerably more tolerant to this stress than BJK_5C is suggestive of that.

Another indicative of what can be *Sd. ludwigii*'s adaptation to different environments are different ploidies observed of UTAD17 (diploid) and BJK_5C (triploid), along with the distinct chromosomal architecture. Specifically, the differences in the number of chromosomal bands between these strains (7 for UTAD17 and 9 for BJK_5C) is suggestive of aneuploidy in *Sd. ludwigii* BJK_5C, which is characterized by abnormal chromosome numbers, that occurred after an error in cell division (Gilchrist & Stelkens, 2019). In yeasts, aneuploidy often leads to an increased cell doubling time (Gilchrist & Stelkens, 2019), but some studies also found that

in some situations it enhances cell proliferation (Gilchrist & Stelkens, 2019; Zhu et al., 2012; Sheltzer & Amon, 2011). Recent research has linked the occurrence of aneuploidy to adaptation, proposing that this chromosomal imbalance may serve as a natural, although transient, pathway to adaptive evolution in microbial populations (Smukowski et al., 2017; Chang et al., 2013; Hose et al., 2015; Scope et al., 2021). For instance, studies on *S. cerevisiae* (Chang et al., 2013; Gorter et al., 2017; Pavelka et al., 2010; Selmecki et al., 2015), *C. albicans* (Yang et al., 2021) or *Cryptococcus neoformans* (Gerstein et al., 2015) demonstrated that aneuploidy can increase stress resistance. In these cases, the loss or gain of extra chromosomes likely has a direct impact on gene dosage, causing higher phenotypic diversity (Gilchrist & Stelkens, 2019). In a recent study by Scopel et al. (2021), genetic variation in aneuploidy prevalence and tolerance was examined across *S. cerevisiae* species. The authors found that karyotype imbalances were more frequent as ploidy levels increased, which was consistent with previous conclusions (Gilchrist & Stelkens, 2019; Scopel et al., 2021; Zhu et al., 2016). Additionally, the study noted that varying industrial conditions could lead to the selection of locally adapted strains, with differences in chromosomal amplifications observed between strains from different industrial origins (Large et al., 2020). Moreover, it was found that some yeast clades display varying capacities to manage the stress associated with aneuploidy (Scopel et al., 2021). For example, strains from the Sake environment exhibited an inherent tolerance to karyotype imbalance, resulting in a higher aneuploidy rate and unstable karyotypes (Scopel et al., 2021). In particular, one study showed that when generating meiotic products from a sake strain (TCR7) (Kadowaki et al., 2017), a large fraction of aneuploid spores, absent in the parental strain, could be observed, which was consistent with a meiotic segregation defect (Kadowaki et al., 2017). Additional research on strain K9 revealed variability in chromosomal duplication, with different chromosomal amplifications appearing in replicate genomic analyses (Gasch et al., 2016). These findings suggested that specific environments can promote the emergence of strains with unstable karyotypes, by fostering higher rates of segregation defects and an increasing tolerance to aneuploidy. In *Sd. ludwigii*, cell division is characterized by a high rate of intra-tetrad mating and an unusual meiotic process, in which homologue interactions are suppressed during prophase I (Papaioannou et al., 2021; Winge, 1946; Simmons & Ahearn, 1985). A recent comparative study of divergent *Sd. ludwigii* strains from international origins mostly identified haploid and diploid strains, but also two aneuploid strains (2n +1 trisomies) (Papaioannou et al., 2021). This study confirmed the total absence of crossing-over, consistent with previous research on *Sd. ludwigii* meiosis (Yamazaki & Oshima,

1996; Papaioannou et al., 2021; Yamazaki et al., 1976). Additionally, the authors demonstrated that *Sd. ludwigii* undergoes frequent intratetrad mating, which preserves high levels of heterozygosity in parts of the genome, while efficiently purging deleterious mutations and enhancing fitness (Papaioannou et al., 2021; Zakharov, 2023). This strategy also ensures that a full diploid chromosome is restored, even in the case of individual aneuploid spores (Papaioannou et al., 2021). Therefore, the adaptation of *Sd. ludwigii* to its environment may be linked to its unique genetic mechanisms, including aneuploidy and its distinctive meiotic process. *Sd. ludwigii*'s ability to tolerate and thrive with chromosomal imbalances, as seen in BJK_5C, along with meiotic strategies like intratetrad mating, highlights its evolutionary flexibility. Therefore, these mechanisms can not only enhance stress resistance and phenotypic diversity, but also provide a pathway for rapid evolution and driving evolutionary innovation in yeast.

In addition to exploring the genomic traits of *Sd. ludwigii*, this work also assessed the utilization of strains BJK_5C and UTAD17 in wine production. Both strains were used in the vinification of two natural wine musts: one from a red grape variety (Touriga Nacional) and another from a white grape variety (Sauvignon Blanc). Aroma analysis of the resulting wines revealed consistent differences between the strains across both grape varieties. Notably, BJK_5C produced significantly higher levels of ethyl acetate compared to UTAD17 (both in red and in white wines), with concentrations exceeding the acceptable sensory threshold for these compounds (AWRI, 2024). Based on the data obtained here, it was not possible to determine the reasons for this increased ethyl acetate production in BJK_5C. However, it was clear that BJK_5C may not be an ideal choice for winemaking, in contrast to UTAD17, which shows potential for producing varietal wines.

Overall, this chapter broadened our view into *Sd. ludwigii's* genomics and physiology by highlighting its adaptive strategies that have allowed it to thrive in diverse environments, but also exploring its potential role in the winemaking context.

Chapter IV.A.

Exploration of a *Hanseniaspora uvarum* high beta-glucosidase-producing strain in winemaking: molecular analyses envisaging identification of causative genes and usefulness as a starter-culture

Part of the results shown in this chapter are described in:

Tavares MJ, Santos D, Seixas I, Mendes Ferreira A, Mira NP, Wine aroma, a tale of molecules, enzymes and microbes: emphasis on the role of Yeasts (submitted);

Tavares MJ, Ferreira, V., Sanches, A., Camolas, F., Ferreira A, Mira NP, Exploration of a *Hanseniaspora uvarum* high beta-glucosidase-producing strain in winemaking: molecular analyses envisaging identification of causative genes and usefulness as a starter-culture (submitted)

IV.A.1 Abstract

Varietal aroma in wine, also known as primary aroma, is determined by the chemical compounds specific of each grape variety, such as monoterpenes, that contribute to the floral notes commonly found in wines. However, most monoterpenes present in the grape juice are bound to sugars through glycosidic forms, turning them non-volatile therefore incapable of contributing to the wine aroma. One way to release monoterpenes is through the activity of beta-glucosidases produced by the microbiota present in the grape must. Following this line, in this work, a panoply of ~ 100 Non-*Saccharomyces* Yeast isolates were collected from Moscatel Galego grapes, a variety known to have a high content in terpenes. These isolates were surveyed for their beta-glucosidase activity potential leading to the identification of one *Hanseniaspora uvarum* strain (MJT198), as a particularly high producer. To understand the molecular basis of this phenotype the genome of this strain was obtained and compared, in a systematic manner, with the genome of many other Non-*Saccharomyces* Yeasts, resulting in the identification of two putative genes that are believed to encode beta-glucosidases and thus be responsible for the observed phenotype. Considering that a very high strain-to-strain variability has been observed concerning the beta-glucosidase production potential across wine yeast strains, the identification of the key players that might be involved in this phenotype is key to assure a stable selection of strains with higher potential to be used in the industry. In this work we also show that the use of *H. uvarum* MJT198 as an adjunct of *S. cerevisiae* results in wines with higher concentration of terpenes, thus indicating its usefulness as a modulator of primary wine aroma

Keywords: Beta-glucosidase, *Hanseniaspora*, comparative genomics, wine fermentation, natural must

IV.A.2 Introduction

As detailed in the introductory chapter of this thesis, wine aroma is usually divided into three main categories: the primary aroma, comprising the compounds associated to the grape varieties used to produce the wine (Zalacain et al., 2007; Ruiz et al., 2019; Rapp & Mandery, 1986); the secondary aroma, stemming from molecules directly produced during the fermentation process by microorganisms present during the vinification (Denat et al., 2021; Belda et al., 2017a; Belda et al., 2017b); and the tertiary aroma, developed during wine ageing and dependent of storage conditions. Wine primary aroma is also called varietal aroma and its impact includes the contributions of many different molecules including terpenes, methoxypyrazines, organic acids, tannins, the precursors of aromatic aldehydes, and volatile thiols (Zalacain et al., 2007; Chen et al., 2022; Styger et al., 2011). Most of these compounds are present in the grape must in the form of water-soluble glycosides or conjugated with amino acids, which turns them non-odorous (Zalacain et al., 2007; Caffrey & Ebeler, 2021). Necessarily, primary wine aromas are highly dependent on the variety of grape used, with some grapes being already more aromatic (*e.g.,* Muscat and Riesling - Zalacain et al., 2007; Loscos et al., 2009; Pedroza et al., 2010) than others (*e.g.,* Chardonnay - Zalacain et al., 2007; Loscos et al., 2009). However, this does not necessarily imply that different grape varieties have different total amounts of those more aromatic molecules since it may simply be that some grape varieties have these compounds in their non-odorous form, bound to sugars. Muscat grapes (Wilson et al., 1986; Sun et al., 2020; Hjelmeland & Ebeler, 2015) are known for their recognizably high content in monoterpenes that upon hydrolysis, occurring naturally during wine aging or assisted by enzymes (like beta-glucosidases or peptidases) produced by indigenous species of the must microbiota (Zalacain et al., 2007; Wilson et al., 1986; Sun et al., 2020; Hjelmeland & Ebeler, 2015; de Ovalle et al., 2021; de Ovalle et al., 2018; Michlmayr et al., 2012), can release their aroma. This chapter focuses on the influence that Non-*Saccharomyces* Yeasts (NSYs) may have in the production of beta-glucosidases and with that in modulating wine primary aroma. This trait is especially relevant as beta-glucosidase activity is rarely found in *S. cerevisiae* strains (Zhang et al., 2021). Extensive work has been done concerning the effect of NSYs as modulators of secondary wine aroma, however, much less knowledge has been generated on their impact in primary wine aroma. Beta-glucosidases (beta-D-glucopyranoside glucohydrolases, E.C. 3.2.1.21) belong to the Glucoside Hydrolases (GH) superfamily that encompasses the GH1, GH3, GH5, GH9, and GH30 families (Zhang et al., 2020). The diversity of families, possible substrates, catalytic mechanisms, and enzyme

locations (intracellular, or extracellular) complicates the identification of what can be the enzymes that can specifically impact wine aroma (Rudakiya et al., 2019). Consequently, several studies have demonstrated a very high strain-to-strain variability in the beta-glucosidase production potential of different strains, highlighting the complexity of enzyme activity among NSYs in winemaking (Villena et al., 2007). However, the molecular reasons behind such variability remain elusive. In particular, the specific genes encoding beta-glucosidases have yet to be fully identified, despite the availability of many genomic sequences of relevant wine NSYs (Zhang et al., 2021). Shedding light into this aspect was another objective of the work pursued in this chapter emphasizing for that matter in the species of the *Hanseniaspora* genus considering their high abundance in grapes and recognized capability of many of its species (*e.g., H. uvarum, H. guilliermondii*) to produce beta-glucosidases (Gallo et al., 2023; Zhang et al., 2020; Aguiar-Cervera et al., 2024; Gallo et al., 2024; Wang et al., 2015; López et al., 2016; Tristezza et al., 2016; Hu et al., 2016; Moreira et al., 2011). Conducting a variety of genomic analyses helped us to pinpoint the causative beta-glucosidase genes with impact in wine aroma. It was also investigated the impact of using a high beta-glucosidase producing strain as a co-adjuvant with *S. cerevisiae* in natural must fermentation, in order to assess whether its high beta-glucosidase activity registered *in vitro*, correlated with an ability to increase free terpene levels in wines and, with that, enhance primary wine aroma.

IV.A.3 Methods and Materials

IV.A.3.1 Isolation of indigenous Non-*Saccharomyces* Yeast strains from Moscatel Galego must

The strains used in this work were isolated from a Moscatel Galego grape must during the 2019 harvest, in a local winery in the Setúbal Peninsula, in Portugal. Must samples were collected during grape crushing in a Destemmer-Crusher and used to prepare serial dilutions (using as solvent a solution of 0.9% NaCl) from 10^{-1} to 10^{-6}. The solutions prepared were plated onto the surface of YPD solid medium (20g/L glucose, 10 g/L peptone, 5 g/L yeast extract, 20g/L agar) and random colonies were selected to grow in WL Nutrient agar medium, that contained 4 g/L yeast extract (Gibco, ThermoFisher Scientific, USA), 5 g/L pancreatic digest of casein (Difco, Fisher Scientific, UK) , 50 g/L glucose (LabChem, USA), 0.55g/L KH_2PO_4 (Merck, Germany), 425 mg/L KCl (Merck, Germany), 125 mg/L $CaCl_2$ (Merck, Germany), 125 mg/L $MgSO_4$ (PanReac, Spain), 2.5 mg/L $FeCl_3$ (ChemLab, Belgium), 2.5 mg/L $MnSO_4$, (Sigma-Aldrich, USA), 22 mg/L bromocresol green (PanReac, Spain), and 20g/L agar (LabChem, USA) and incubated at 30ºC for 48h until colony development. Isolated colonies with distinct morphologies were selected, re-streaked on the surface of yeast extract peptone dextrose (YPD) solid medium (20g/L glucose, 10 g/L peptone, 5 g/L yeast extract, 20g/L agar), and incubated at 30ºC. Finally, the isolated colonies were added to 1 mL liquid YPD medium to which was added 1mL glycerol (Scharlau, Spain) to assure successful cryopreservation at -80ºC until further use.

IV.A.3.2 Screening of beta-glucosidase-activity in the cohort of yeast isolates

Screening of isolates for extracellular beta-glucosidase activity was performed in esculin medium using a moderate modification of the method described by Pérez et al (2011). Briefly, strains were grown overnight in liquid YPD medium. Afterwards, optical density (OD_{600}) was measured, and the cultures were diluted with saline solution 0.9% NaCl until OD_{600}=0.04. Of that dilution, 1uL of each culture was inoculated onto Esculin Glycerol Agar (EGA) plates containing 1 g/L esculin (Sigma-Aldrich, USA), 0.3 g/L ferric chloride (ChemLab, Belgium), 1 g/L casein hydrolysate (Difco,Fisher Scientific, UK), 25 g/L yeast extract (Gibco, ThermoFisher Scientific, USA), 8ml/L glycerol (Scharlau, Spain), and 20 g/L agar, pH 6.0 (LabChem, USA), and incubated at 30ºC. The plates were observed each 30 minutes after inoculation, with the appearance of black spots indicating the presence of extracellular beta-glucosidase activity. Isolates that demonstrated activity were clustered,

qualitatively, in 4 different categories: i) “very high activity” if the black spot was only visible 60 minutes after inoculation; ii) “high activity” if the black spot was only visible between 60 minutes and 8h after inoculation; iii) “moderate activity” if the black spot was only visible between 8 and 24 h; iv) X (denoting no activity) if no visible black spot could be detected after 24h.

IV.A.3.3 Molecular identification of yeast isolates.

From the obtained Non-*Saccharomyces* isolates, 24 were selected based on their beta-glucosidase production potential for molecular identification based on Sanger sequencing of the conserved Internal Transcribed Spacer (ITS) region. For this identification, the genomic DNA of the selected isolates was obtained by growing the cells overnight in YPD liquid medium, at 30ºC with 250 rpm agitation. On the next day, cells were centrifuged, resuspended in lysis buffer (50mM Tris, 50mM EDTA pH 8.0, 250mM NaCl, 0.3% SDS) and vortexed in the presence of 100 µL of glass beads. This step was followed by an incubation step at 65ºC, for 1h, followed by an incubation on ice for 2 minutes. Then, DNA was precipitated using 20µL 3M sodium acetate and 2 volumes (400 µL) 100% ethanol, followed by overnight incubation at -20ºC. Afterwards, DNA was washed with 200µL of 70% ethanol and resuspended in 50 µL ddH_2O. The ITS region was amplified by PCR using the universal primers ITS1 (5'-TCCGTAGGTGAACCTGCGG-3’) and ITS4 (5’-TCCTCCGCTTATTGATATGC-3’) in a reaction mixture that included: 1U Horse-Power™ Taq DNA Polymerase enzyme (Canvax, Valladolid, Spain), 50 ng of the recovered genomic DNA, 1.2 pmol/ µL of MgCl2 (Canvax, Valladolid, Spain), 1.6 pmol/ µL of each primer and 0.016 pmol/ µL of dNTPs. For the PCR amplification the following conditions were used: initial denaturation at 94ºC for 1 minute, 30 cycles of denaturation at 94ºC for 30 seconds, annealing at 55ºC for 1 minute, and extension at 72ºC for 40 seconds. A final extension of 72ºC for 10 minutes was performed. To confirm the existence of a single band, the resulting PCR product was subjected to an electrophoresis run in an 1.5% agarose gel in TAE buffer (Tris-acetate-EDTA) pH 8.0 for 45 minutes at 100V.

IV.A.3.4 Genomic sequencing of strain *H. uvarum* MJT198

H. uvarum MJT198 was sequenced in an Illumina MiSeq platform in CD genomics (Shirley, NY, USA). The DNA libraries were prepared using the ThruPLEX DNA-seq kit and paired-end sequencing of the generated DNA fragments was performed on a MiSeq platform.

After two sequencing rounds, reads of 250 bp on average were obtained and *de novo* assembled using CLC Genomics Workbench (Qiagen, Germany) assembly tools with default parameters.

IV.A.3.5 *In silico* analysis and characterization of the beta-glucosidase-encoding genes among the *Hanseniaspora* genus

Thirty one genomes belonging to different *Hanseniaspora* species/strains deposited at NCBI assembly database were used in a standardized annotation procedure using WebAugustus (Hoff & Stanke, 2013; Stanke et al., 2006) (trained on annotated genome of *S. cerevisiae* S288c). With the results of that automatic genome annotation an in-house database containing the obtained proteomes for the 33 strains examined was created (**Appendix Table IV.A.1**). To identify among those proteomes the beta-glucosidase genes, the dbCAN Meta Server was used to identify which proteins belonged to the Glycoside Hydrolase family (Yin et al., 2012), enabling the identification of proteins from the GH3 family (putative beta-glucosidase sequences). Then, the putative beta-glucosidase sequences discovered were submitted to an INTERPRO (Blum et al., 2021; Jones et al., 2014) search, in which it was possible to obtain further functional information of the proteins, such as families, domains and relevant sites, as is the case of signal-P, whose existence hints for a possible extracellular protein. Phylogenetic analysis of the proteins found to encode different beta-glucosidases was performed by whole-sequence alignment, using MUSCLE (Edgar, 2004), followed by a phylogenetic analysis in MEGA (Tamura et al., 2021) and in the interactive tool for phylogeny analysis, iTOL (Maximum-Likelihood Tree) (Letunic & Bork, 2021).

IV.A.3.6 Fermentations in Moscatel Galego natural must

Fermentations were carried out using natural grape must obtained from Moscatel Galego grapes, harvested in 2019, at a local winery in the Setúbal Peninsula, Portugal. The grape juice was obtained by pressing the grapes in a sterile bag, followed by a gentle sulfiting, and then stored at -20ºC. Before storage, aliquots of the must were taken for the quantification of glucose, glucose-fructose, primary amino nitrogen (PAN), yeast assimilable nitrogen (YAN), ammonia, pH, total acidity, free SO_2, total SO_2, copper, malic acid, iron, calcium, magnesium, potassium, and Brixº. These measurements were conducted using a Y15 enzymatic autoanalyzer (Biosystems, Barcelona, Spain) with corresponding kits provided by the manufacturer. Before the fermentations, the must was gradually thawed at 4ºC and

distributed into six 250 mL flasks (labeled F1 through F6), each filled to two-thirds of their volume. Flasks F4-F6 were inoculated solely with *S. cerevisiae* QA23, while flasks F1-F3 underwent sequential inoculation: first with *H. uvarum* MJT198, followed by *S. cerevisiae* QA23 72 hours later. Starter cultures of *H. uvarum* MJT198 and *S. cerevisiae* QA23 were prepared by cultivating the cells overnight at 25ºC and 150 rpm in 100 mL Erlenmeyer flasks containing 50 mL of a synthetic Grape Juice Medium (GJM) (Henscke & Jiranek, 1993). A cell density of $1x10^6$ CFU/mL was used to inoculate both strains. To enable anaerobic and aseptic sampling, each flask was equipped with a sidearm port sealed with a rubber septum and incubated in a static growth chamber at 13ºC. Yeast growth was monitored through periodic optical density measurements (OD_{600}) and colony-forming unit (CFUs) counts on solid agar plates of YPD (in fermentations containing only *S. cerevisiae* QA23 or *H. uvarum* MJT198) or WL (for mixed fermentations with both *S. cerevisiae* QA23 and *H. uvarum* MJT198). Fermentation progression was tracked by daily measurement of weight loss of the flasks, which estimated the production of CO_2. Fermentation was deemed complete when no further weight loss was observed. Aseptic sampling of the fermentation broth was performed periodically, using a syringe system. These samples were used for the determination of major and minor volatiles (as detailed below), as well as for the quantification of glucose, fructose, organic acids and ethanol, performed by HPLC, using a refractive index (RI) detector, following separation on an Aminex HPX-87H column, eluted at room temperature with 0.005 M H_2SO_4 at a flow rate of 0.6 mL/min, for 30 minutes.

IV.A.3.7 Quantification of major and minor volatiles

The determination of major (volatiles usually present in wines at levels above 0.2 mg/L) and minor (volatiles usually present in samples between 0.1 and 200 μg/L) volatiles present in the different samples taken along the single and co-inoculated fermentations of the must was performed, as a paid service, with the Laboratory for Aroma Analysis and Enology (LAAE), from the University of Zaragoza, Spain. In total, major and minor volatile compounds (**Appendix Table IV.A.2**) were quantified in 15 samples: three replicas of the initial grape must, three replicas of the two sets of fermentations taken at 72h and three replicas of the two sets of fermentations taken in the end of the fermentation. The samples with an expected high concentration of sugars (initial grape must and samples took at 72h) were subjected to a procedure that aimed to eliminate residual sugars and acids, while recovering the Polyphenol and Aromatic Fraction (PAF), as described by Alegre et al., (2020). In brief, PAF extraction

was performed with C18 resin packed in 6 mL SPE cartridges, previously conditioned with methanol and MilliQ water containing 2 % (v/v) of ethanol. After charging the resin with the samples, a washing step with MiliQ water at pH 3.5 was done, followed by a drying step under vacuum. The aroma precursor fractions were recovered by elution with 20 mL of ethanol. Prior quantification of the major and minor volatiles, samples were subjected to an accelerated aging via acid hydrolysis at 50ºC, in anoxia for two weeks. For this, the samples were placed into a free O_2 chamber Jacomex (Dagneux, France), as developed by Vela et al., (2017) and adapted for fermented media by Denat et al., (2017) and Oliveira & Ferreira (2019). Samples were incubated at 50ºC for 2 weeks. Following the aging step, higher alcohols and their acetates, volatile fatty acids and their ethyl esters, branched fatty acids, acetoin, diacetyl, and acetaldehyde were quantified by the GCFID analysis of a dichloromethane micro-extract. For this, the samples were diluted in an ammonium sulfate solution and placed horizontally for agitation during 90 minutes. Then, extraction was performed with dichloromethane. Samples were centrifuged (2500 rpm, 10 minutes), and the organic phase recovered with a syringe. The minor aroma compounds terpenes, norisoprenoids, vanillin derivatives, volatile phenols branched ethyl esters, were analyzed by GC-MS after a solid phase extraction (SPE) step. For this, the samples were supplemented with appropriate amounts of 2-octanol, 3- octanone, 3,4-dimethylphenol that were used as internal standards. The samples were passed through the SPE cartridge (ISOLUTE® ENV+ SPE, Biotage, Sweden), previously conditioned with dichloromethane, methanol and 12 % (v/v) of a hydroalcoholic solution. The resin was then washed with a solution of 30 % (v/v) of methanol and 1 % (m/v) of sodium bicarbonate, and dried under vacuum. Elution was performed with dichloromethane containing 5 % (v/v) of methanol. The concentrations of both major and minor volatiles were calculated with the relative response factor obtained from the analysis of a calibrated synthetic wine.

IV.A.4 Results and Discussion

IV.A.4.1 Screening for beta-glucosidase-producing yeast isolates from Moscatel Galego highlights the species *H. uvarum* as a top-producer

The isolation of Non-*Saccharomyces* yeast isolates from a Moscatel Galego must, (obtained from a local winery in the Setúbal region) led to the recovery of 105 isolates. These isolates were subsequently profiled for their beta-glucosidase activity using a solid esculin medium, as detailed in the Materials and Methods section. Moscatel grape varieties are known for their high terpene content (Sánchez et al., 2007), which was hypothesized to favor the isolation of yeast strains with enhanced beta-glucosidase activity. Such screening identified 11 isolates with very high beta-glucosidase activity, indicated by the rapid appearance of a black halo in the esculin-containing medium within the first 60 minutes after inoculation, while 32 strains exhibited high beta-glucosidase activity, with the black halo appearing between 1- and 8-hours post- inoculation (**Appendix Figure IV.A.1**). 27 strains exhibited a moderate activity (black halo formation observed after overnight growth) and 35 showed no beta-glucosidase activity (no black halo formed even after overnight incubation) (**Figure IV.A. 1**).

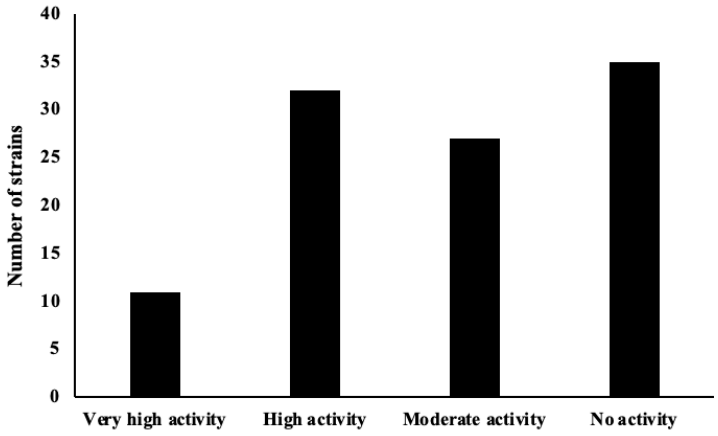


Figure IV.A. 1 - Beta-glucosidase activity on esculin medium from Moscatel Galego isolates. From 105 isolates, 11 demonstrated a very high activity on esculin, 32 had high activity, 27 demonstrated moderate activity, and 35 had no beta-glucosidase activity on esculin.

To identity the species to which belong the high beta-glucosidase producing isolates, molecular identification, based on sequencing the ITS region, was used. Additionally, we randomly selected for molecular identification 5 isolates with high activity, 5 with moderate activity and 3 with no activity observed (**Table IV.A. 1**). All the 11 isolates showing very high beta-glucosidase activity belonged to the *H. uvarum* species, aligning with other studies that

have also identified this species as having high activity (Hu et al., 2016; Hu et al., 2018; Zhang et al.,2023). The 5 selected "high-activity" were also identified as belonging to the *Hanseniaspora* genus, including *H. uvarum* (3 isolates), *H. opuntiae* (1 isolate) and *H. vineae* (1 isolate) (**Table IV.A. 1**). The 3 isolates without beta-glucosidase activity were identified as *Lanchancea thermotolerans* and *Saturnispora diversa* (**Table IV.A. 1**). Searches for carbohydrate-active enzymes in the genomes of two publicly available *L. thermotolerans* strains (CBS6340 and NRRL Y-8284) revealed the absence of proteins predicted to belong to the GH3 family, the most common family of beta-glucosidases found in yeasts (Zhang et al., 2021). This absence of GH3 may align with our phenotypic observation of limited beta-glucosidase activity, as GH3 enzymes typically contribute to such activity in other yeast species (Huang et al., 2021). Instead, proteins from the GH5 family were identified in both strains: three GH5 proteins in *L. thermotolerans* CBS6340 and two in *L. thermotolerans* NRRL Y-8284 (results not shown).

Our screening and subsequent identification highlighted *Hanseniaspora* species as prominent producers of beta-glucosidases with high activity, though some inter-strain variability was observed. In contrast, beta-glucosidase activity in *L. thermotolerans* varied across strains: one isolate exhibited moderate activity, while two others showed no detectable activity. Further research is, thus, necessary as this variation may indicate that GH5 beta-glucosidases, if present in these *L. thermotolerans* isolates, have a different substrate specificity compared to GH3 enzymes, typically associated with strong beta-glucosidase activity on substrates like esculin. A recent work by Zhang et al., (2021) that compiled the existing studies on beta-glucosidase activity in different wine yeasts also denoted this significant inter-strain variability, among yeast strains from the same species. Such differences were observed under different experimental conditions (*e.g.*, different pH and temperatures) and using different substrates (which may indicate that different yeast species and strains have beta-glucosidases with distinct substrate preferences). Thus, more than inter-species variability, the display of beta-glucosidase activity is also very dependent of the strain, and the substrate used.

Table IV.A. 1 - Classification on esculin and molecular identification (ITS), based on BLASTn on NCBI, of selected 24 strains isolated from Moscatel Galego must. In bold, strain *H. uvarum* MJT198, was the selected strain for continuing the study on beta-glucosidase production.

Beta-glucosidase activity	Isolate	BLAST results		
		Species Identification	Identity (%)	E-value
Very high activity	MJT164	*H. uvarum*	99.82	0.0
	MJT166		99.65	0.0
	MJT180		100	0.0
	MJT183		99.66	0.0
	MJT198		99.83	0.0
	MJT207		100.0	0.0
	MJT208		99.49	0.0
	MJT227		98.08	0.0
	MJT236		99.83	0.0
	MJT254		100.0	0.0
	MJT260		100.0	0.0
High activity	MJT175	*H. opuntiae*	94.63	0.0
	MJT182	*H. uvarum*	99.66	0.0
	MJT189	*H. uvarum*	99.83	0.0
	MJT218	*H. vineae*	99.31	0.0
	MJT242	*H. uvarum*	100.0	0.0
Moderate activity	MJT165	*H. uvarum*	94.51	0.0
	MJT176	*H. uvarum*	99.66	0.0
	MJT185	*H. uvarum*	99.66	0.0
	MJT211	*L. thermotolerans*	99.83	0.0
	MJT241	*H. vineae*	99.15	0.0
No activity	MJT162	*L. thermotolerans*	99.66	0.0
	MJT240	*S. diversa*	99.82	0.0
	MJT250	*L. thermotolerans*	99.15	0.0

IV.A.4.2 Genomic analysis of Non-*Saccharomyces* Yeasts of the *Hanseniaspora* genus enables the identification of beta-glucosidase-encoding genes

This study and others (such as those gathered in Zhang et al., 2021) highlighted the high prevalence of yeast species within the *Hanseniaspora* genus as exhibiting very high beta-glucosidase activity (despite the observed strain-to-strain variability). Despite this, the genes that are responsible for such activity remain to be characterized, a knowledge that may also help us to shed light into the strain-to-strain variability since different strains may expressed more or less active alleles, pending alterations occurring in their coding sequences. To shed light into this aspect, the predicted ORFeomes of 33 different *Hanseniaspora* strains (including of the herein identified high beta-glucosidase producing isolate *H. uvarum* MJT198) (**Figure IV.A. 2**), were inspected. Particularly, for 33 of these strains only the genomic sequences of 13 species were deposited at NCBI and they had to be subjected to annotation (using the Augustus gene finder trained in *S. cerevisiae* S288c). To identify beta-glucosidase sequences in the generated ORFeomes, we employed the dbCAN Meta Server (Yin et al., 2012), an

informatic tool that detects proteins belonging to the glycoside hydrolase (GH) family. This approach led to the identification of three predicted GH3 family enzymes in the genome of *H. guilliermondii* UTAD222, consistent with what has been the manually curated annotation undertaken previously by our group (Seixas et al., 2019). All the examined *Hanseniaspora* ORFeomes included GH3 enzymes with varying levels of sequence homology, which clustered into four distinct groups. The most prevalent group, named *BGLU1*, was found in all *Hanseniaspora* species examined with the exception of *H. gamundiae*, *H. osmophila*, *H. vineae*, and *H. occidentalis* (**Figure IV.A. 2**; **Figure IV.A. 3**). Although *H. gamundiae* CRUB1928 contained a GH3 enzyme, this sequence resembled an importin alpha-re-exporter, which is distinct from any other predicted BGLU proteins. The three other beta-glucosidase groups identified were designated BGLU2, BGLU3, and BGLU4 (**Figure IV.A. 2**; **Figure IV.A. 3**), with BGLU2 appearing to be phylogenetically closer to BGLU1 but only being found in 11 different species, including *H. guilliermondii, H. pseudoguilliermondii, H. opuntiae, H. thailandica, H. osmophila, H. meyeri, H. lachancei, H. uvarum, H. hatyaiensis, H. clermontiae* and *H. lindneri*. The third group, BGLU3, was only found in *H. vineae, H. osmophila* and *H. occidentalis* (**Figure IV.A. 2**; **Figure IV.A. 3**). BGLU4 group included the most different sequences of beta-glucosidases, being also of interest the observation that these proteins appear to harbor a signal peptide that might suggest potential extracellular location. This gene was present in 9 out of 20 *Hanseniaspora* species examined, including *H. guilliermondii*, *H. opuntiae* and *H. singularis*.

The observed variations in beta-glucosidase sequences across *Hanseniaspora* species may reflect evolutionary adaptations, possibly influencing carbon source scavenging (Zhang et al., 2021), since beta-glucosidases play essential roles in releasing aromatic compounds during winemaking and in producing monomeric sugars from complex oligosaccharides, facilitating yeast consumption of these sugars (Zhang et al., 2021). Another interesting finding was that the species that possess the *BGLU3* genes are all part of the slow-evolutionary lineage (SEL) within *Hanseniaspora*, characterized by a slower evolutionary rate compared to the fast-evolutionary lineage (FEL) (Steenwyk et al., 2019), which includes species such as *H. uvarum* and *H. guilliermondii*. This observation supports the hypothesis that beta-glucosidases may play an evolutionary role; however, additional studies on these genes and are required to fully clarify their specific function in releasing bound substrates.

Hanseniaspora strain	BGLU1	BGLU2	BGLU3	BGLU4
H. clermontiae NRRL-Y27515	1	2	0	1
H. gamundiae CRUB 1928	1	0	0	0
H. guilliermondii NRRL-Y1625	1	1	0	1
H. guilliermondii UTAD222	1	1	0	1
H. hatyaiensis ST476	1	1	0	0
H. jakobsenii ZIM 2603	1	1	0	1
H. lachancei NRRL- Y27514	1	1	0	1
H. lindneri CBS 285	1	1	0	0
H. meyeri NRRL-Y27513	1	1	0	1
H. mollemarum CBS 15034	2	0	0	0
H. nectarophila CBS 13383	1	0	0	1
H. occidentalis var. citrica CBS 6783	0	0	1	0
H. occidentalis var. occidentalis NRRL-Y7946	0	0	3	0
H. opuntiae AWRI 3578	1	1	0	1
H. osmophila AWRI3579	0	0	3	0
H. osmophila NRRL-Y1613	0	0	3	0
H. pseudoguilliermondii ZIM 213	1	1	0	1
H. singularis ST244	1	0	0	1
H. smithiae CRUB 1602	1	0	0	1
H. thailandica ZIM 2325	1	1	0	1
H. uvarum 349	1	1	0	0
H. uvarum AWRI 3580	1	1	0	0
H. uvarum AWRI 3581	1	1	0	0
H. uvarum B079	1	2	0	0
H. uvarum CBA 6001	1	1	0	0
H. uvarum DSM 2768	1	1	0	0
H. uvarum H2	1	1	0	0
H. uvarum H20	1	1	0	0
H. uvarum H4	1	1	0	0
H. uvarum MJT198	1	1	0	0
H. uvarum NRRL-Y1614	1	1	0	0
H. valbyensis NRRL-Y1626	1	0	0	1
H. vineae T0205AF	1	0	1	0

Figure IV.A. 2 - Presence and absence of beta-glucosidase encoding genes across the *Hanseniaspora* genus. In this genus, it was possible to identify four beta-glucosidase-encoding genes. Overall, the most prominent gene is *BGLU1*, which is present in all tested strains, apart from *H. gamundiae* CRUB1928.

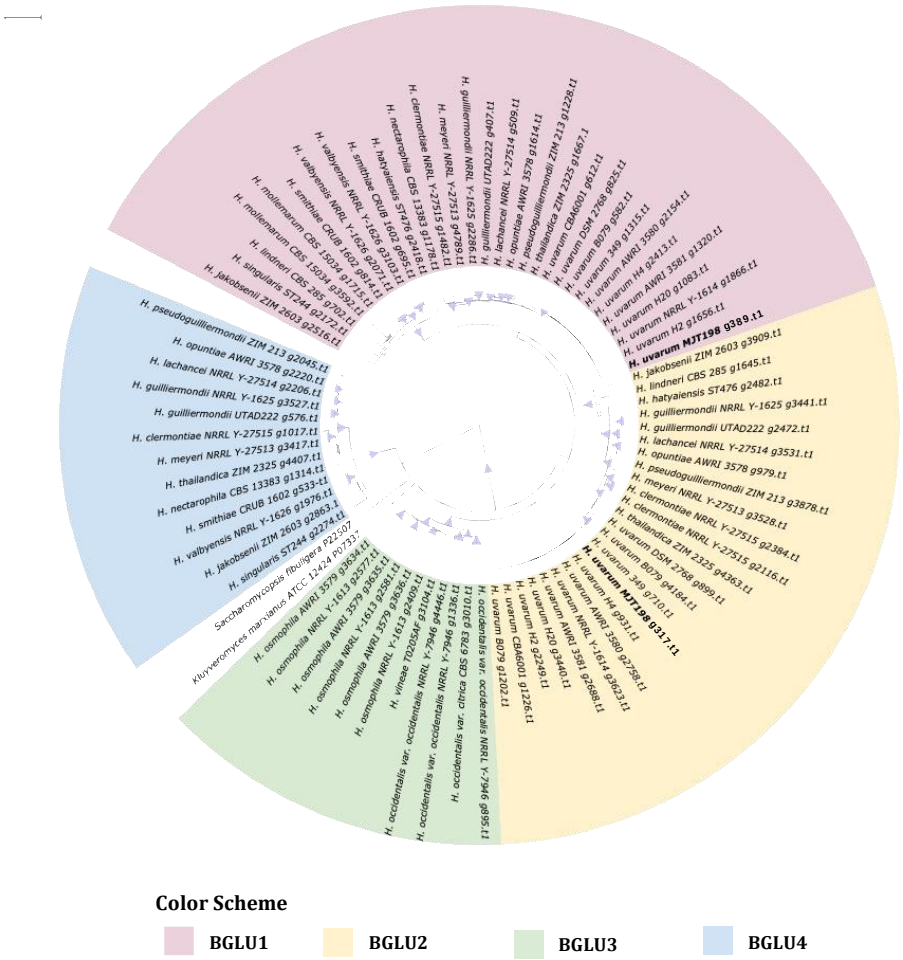


Figure IV.A. 3 - Phylogenetic display of the beta-glucosidase sequences identified in the ORFeomes of the 31 strains of the *Hanseniaspora* genus. In this phylogenetic tree it possible to observe four different types of beta-glucosidases: *BGLU1* (red); *BGLU2* (yellow); *BGLU3* (green); *BGLU4* (blue).

Confronting the results of the *in silico* analysis with those of the phenotypic screening for beta-glucosidase activities, it was found that in the inspected ORFeomes belonging to *H. uvarum* (the species that in our screening comprised the isolates with highest activity) do not appear to encode enzymes of the *BGLU3* or *BGLU4* genes. These observations turn tempting to say that the beta-glucosidases with highest capability to recognize esculin are not enzymes belonging to these two groups but rather to the BGLU1 and BGLU2 group. Thus, the evidence obtained seem to suggest that the genes encoding beta-glucosidases with impact in esculin degradation should be homologous to those encoded by genes HU_g389.t1 and HU_g317.t1 found in the genome of *H. uvarum* MJT198 (**Figure IV.A. 3**).

With this *in silico* identification, the essential future step to leverage this research is to confirm the expression of these genes across different *H. uvarum* strains and to assess whether variations in expression levels or sequence alterations may explain the differential capacity for

beta-glucosidase activity against esculin. Furthermore, it is also crucial to consider the representativeness of esculin as a substrate in this context, given that monoterpenes may differ in their structural relevance to winemaking applications. While esculin serves as an important substrate, and identifying beta-glucosidases that degrade it is valuable, understanding the extent to which these findings apply to monoterpene metabolism in wine will be essential for a more comprehensive interpretation.

IV.A.4.3 Effect of using *H. uvarum* MJT198 as an adjunct of *S. cerevisiae* in fermentation of Moscatel Galego wine

The high beta-glucosidase activity exhibited by strain *H. uvarum* MJT198 in the phenotypic assay conducted using esculin as a substrate, prompted us to evaluate whether that capacity could also be detectable when this strain is used as an adjunct of *S. cerevisiae* in fermentation of Moscatel Galego grape musts. To test that, an experimental setting was designed, involving a sequential inoculation of the two yeasts, with *H. uvarum* being inoculated in the grape must and *S. cerevisiae* after 4 days of fermentation (**Figure IV.A. 4**).

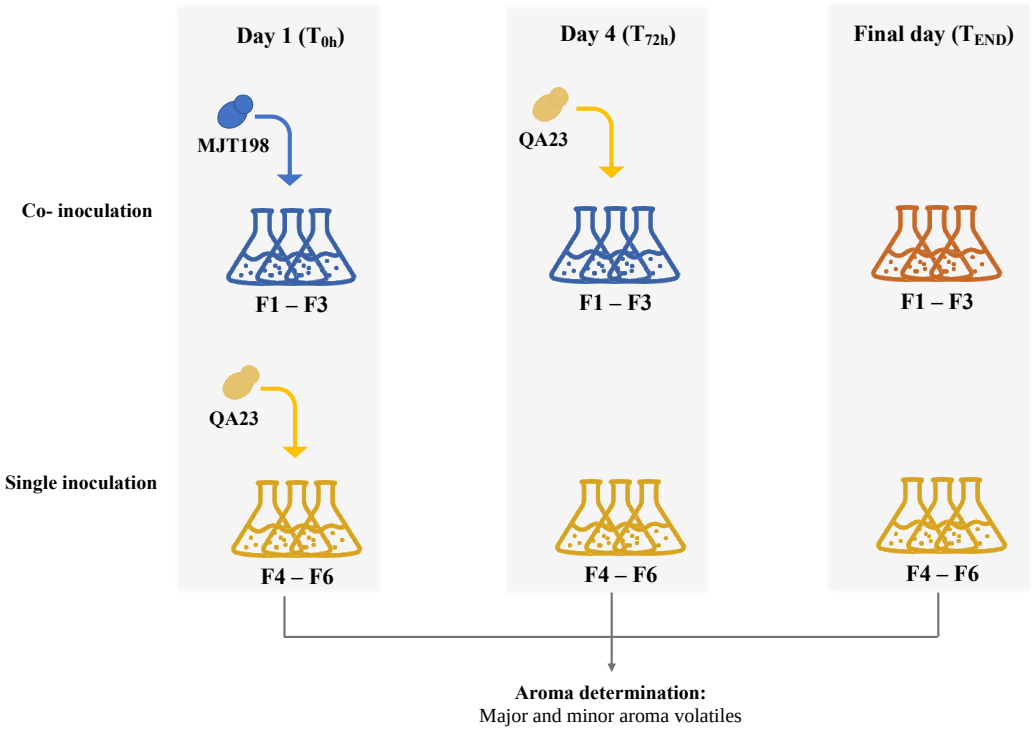


Figure IV.A. 4 - Inoculation and sampling strategies for the fermentations conducted on natural must with *H.uvarum* MJT198 and *S. cerevisiae* QA23. These fermentations consisted of two different inoculation strategies: a sequential inoculation which started with *H. uvarum* MJT198 followed by an addition of *S. cerevisiae* QA23 at 72h; and a single culture (control) with *S. cerevisiae*. For both inoculation strategies, three samples were taken for aroma determination: the initial timepoint, taken immediately after inoculation (T_{0h}), at 72h after inoculation (T_{72h}), and at the end of fermentation (T_{end}).

Chemical analysis of the grape juice obtained immediately before the inoculation of the yeasts revealed a total sugar content of 150.61 g/L, approximately half of glucose and half of fructose (**Table IV.A. 2**); a total acidity was 3.2 g/L and a pH of 3.99; and a yeast assimilable nitrogen (YAN) content of 272 mg/L, distributed between 177 mg/L of primary amino nitrogen (PAN) and 116 mg/L of ammonia. The concentration of initial free sulfite in suspension was 49 mg/L, being the total SO_2 of 85 mg/L.

Table IV.A. 2 - Physico-chemical characterization of the initial grape juice, before fermentation. In this characterization, are presented values of total sugars, glucose, malic acid, total acidiy, yeast assimilable nitrogen (YAN), primary assimilable nitrogen (PAN), ammonia and sulfite.

Parameter	Concentration
Total sugar (Glucose + Fructose)	150.61 g/L
Glucose	75.56 g/L
Malic acid	2.4 g/L
Total acidity	3.2 g/L
YAN	272 mg/L
PAN	177 mg/L
Ammonia	116 mg/L
Free SO_2	49 mg/L
Total SO_2	85 mg/L
pH	3.99
Brix (º)	17.5
Theoretical EtOH = (Brix x 0.6) x(90-100%)	9.45-10.5%
ºBrix/Total acidity (optimal value 3.1)	4.48
ºBrix/pH^2 [200-270]	278.6
ºBrix/pH [85-95]	69.8

Dynamics of growth of the two yeasts as well as the fermentation profiles are shown in **Figure IV.A. 5 .**

Carbon Dioxide (CO_2) release was calculated taking into consideration the theoretical sugar concentration, the molecular weights of CO_2 and glucose and the fermentation volume, being given by the equation below:

$$CO2\ release\ (g) = \frac{Theoretical\ total\ sugar\ concentration\ \left(\frac{g}{L}\right) \times 2 \times MW\ CO2\ \left(\frac{g}{mol}\right)}{MW\ glucose\ \left(\frac{g}{mol}\right)} \times Fermentation\ Volume\ (L) \quad \textbf{(1)}$$

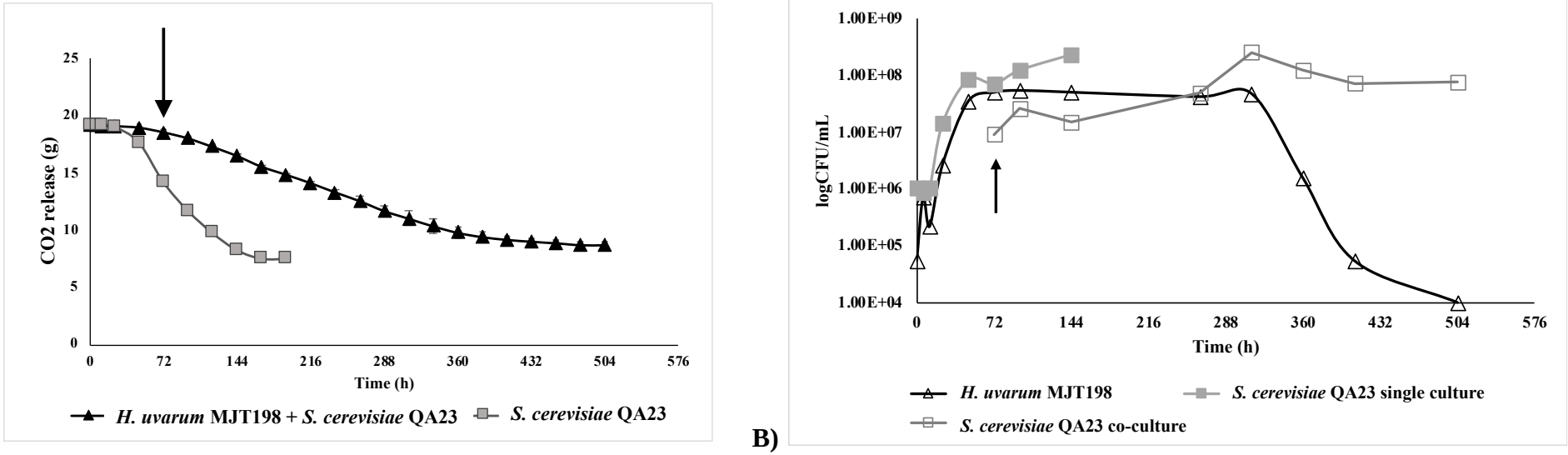


Figure IV.A. 5 - Fermentation monitoring. **A)** Fermentation profiles showing the production and release of carbon dioxide over time. The averages and standard deviations among replicates are displayed for i) sequential inoculation of *H. uvarum* MJT198 and *S. cerevisiae* QA23 (black triangles); and ii) single inoculation of *S. cerevisiae* QA23 (gray squares). The sequential inoculation took 504h to complete, while the single inoculation only took 192h. **B)** Yeast cell counts for the sequential inoculation with *H. uvarum* MJT198, represented by open black triangles and *S. cerevisiae* QA23 by open gray squares; yeast cell counts for the single inoculation with *S. cerevisiae* QA23 are represented by filled gray squares. In both images, the addition of *S. cerevisiae* at 72h is indicated by a black arrow.

Both the co-inoculated and the single *S. cerevisiae* fermentations were successfully completed, although with substantial differences in the kinetics, which aligns with findings from other studies performed using co-cultures of *S. cerevisiae* and *H. uvarum* (Wang et al., 2015; Tristezza et al., 2016). Specifically, the fermentations with *S. cerevisiae* QA23 were completed within 192 hours, while the co-inoculated sequential fermentations took 504 hours to finish (**Figure IV.A. *5***). These differences in fermentation time reflect on one side, the limited fermentative capacity of *H. uvarum* in the beginning, but also the competitive nature of the co-culture since the fermentation of *S. cerevisiae* is also slower compared to what is observed when these cells are cultivated alone.

H. uvarum MJT198 maintained relatively constant viability for 12 days, until 288h into fermentation (**Figure IV.A. *5***). The fermented must obtained after 72h of inoculation with *H. uvarum* showed a higher consumption of fructose (15% consumed) than of glucose (2% consumed), which is line with the described fructophilic behavior of this species (Ciani & Fatichenti, 1999; De Benedictis et al., 2011; Seixas et al., 2018). At this same time point, the fermentations conducted by *S. cerevisiae* QA23 showed a consumption of 37% glucose and 29% fructose, which is indicative of a co-consumption of both sugars. Although *S. cerevisiae* is highly glucophilic, preferentially fermenting glucose while suppressing the mechanisms for utilizing other sugars in its presence (Barbosa et al., 2014), strain QA23 strain appears to have adapted its sugar consumption pattern in a way that allows co-consumption, this observation also being reported in other studies (Barbosa et al., 2014; Contreras et al., 2023). Indeed, a study conducted by Barbosa and colleagues also describes the co-consumption of glucose and fructose by QA23 cells, also adding that the fast metabolization of these sugars by these cells was highly dependent on nitrogen consumption (Barbosa et al., 2014). The relatively low sugar consumption during the early stages of fermentation is reflected in the modest ethanol titers observed at 72h, with 0.5% being present in the musts fermented by *H. uvarum* MJT198 and 2.35% in those fermented by *S. cerevisiae* QA23 (**Table IV.A. 3**). By the end of the sequential fermentation undertaken by both *H. uvarum* and *S. cerevisiae*, the ethanol concentration was 9.18% (**Table IV.A. 3**) which was, interestingly, 0.78% higher than that in the fermentation carried out with *S. cerevisiae* alone. This suggests that the presence of *H. uvarum* may have contributed for this slightly enhanced ethanol production. A similar outcome in terms of ethanol production was reported by Cervera et al., (2024), where a co-fermentation of beer wort with *H. vineae* and *S. cerevisiae* (inoculated at equal proportions (50:50) also resulted in 1.05% more ethanol comparing to the fermentations undertaken with *S. cerevisiae* alone. The

increase in proportion of *H. vineae* cells (90:10), ethanol production decreased by 2%, comparing to *S. cerevisiae* alone, indicating that the ratio of yeast species in fermentation can have a significant influence in their interaction and in the overall outcome of ethanol.

Table IV.A. 3 - Average values for concentration of total sugars (glucose + fructose), glucose, fructose and ethanol at 72h and at the end of fermentation for i) sequential inoculation of *H. uvarum* MJT198 and *S. cerevisiae* QA23 and ii) the single inoculation of *S. cerevisiae* QA23. Additionally, the values for the initial wine must are also represented.

Concentration (g/L)	**Initial must**	**Sequential fermentation**		***S. cerevisiae* QA23**	
	T_{0h}	**T_{72h} (with only Hu)**	**T_{end} (Hu+Sc)**	**T_{72h}**	**T_{end}**
Total sugars (glucose + fructose)	150.6	138.08 ± 2.4	0.08 ± 0.06	101.03±5.11	1.52±0.54
Glucose	75.56	74.13 ± 1.63	0.01 ± 0	47.64±1.83	0.06 ± 0.03
Fructose	75.05	63.95 ± 0.90	0.07 ± 0.06	53.39±3.43	1.46±0.512
Ethanol (%)	0	0.493±0.115	9.18 ± 0.72	2.35±0.122	8.4±0.385

Considering the crucial role that aging of wines in determine wine aroma profile (impacts acid hydrolysis, chemical rearrangements of unstable molecules and concentration of terpenes – Waterhouse et al., 2016; Denat et al., 2021; Simpson & Miller, 1983 –) we chose to analyze whether the presence of *H. uvarum* MJT198 affected or not the primary aroma using a procedure that, in the lab, "recreates" the aged wine. Quantification of major (usually present in wines at levels above 0.2 mg/L) and minor (usually present in samples between 0.1 and 200 μg/L) volatiles showed that the presence of *H. uvarum* had a significant impact in the overall profile of aroma molecules, as reflected in the PCA analysis shown in **Figure IV.A. 6**. This analysis clearly reveals the distinct impact of *H. uvarum* with the samples obtained from wines fermented with *H. uvarum* MJT198 being located in the first quadrant, while those from wines fermented solely with *S. cerevisiae* were located in the fourth quadrant **Figure IV.A. 6**). Samples taken at 72 hours of fermentation clustered closely with the initial wine must, probably due to the early stage of the fermentation. A closer look into the PCA shows that the biplot analysis revealed that ethyl acetate and isoamyl alcohol were the most relevant compounds driving the differentiation of clusters B (wines using *S. cerevisiae* and *H. uvarum*) and C (wines obtained using only *S. cerevisiae*), respectively (**Figure IV.A. 6**). This finding is particularly significant as *H. uvarum* inoculation has been linked to an increase in ethyl acetate levels (Moreira et al., 2000; Li et al., 2000; Johnson et al., 2020) and *S. cerevisiae* QA23 alone has been described as a good producer of isoamyl alcohol, outperforming other commercial *S. cerevisiae* strains, like CY3079, VL3 and EC1118 (Seixas et al., 2023; Oberholster et al., 2018).

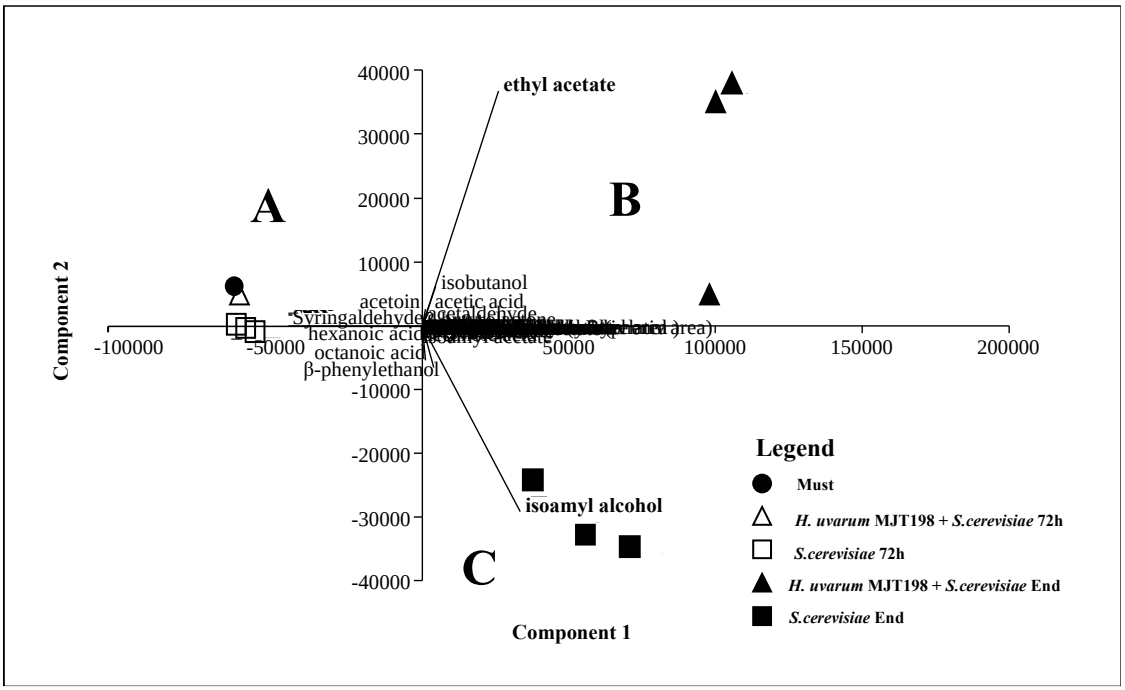


Figure IV.A. 6 - Principal Components Analysis (PCA) biplot showing sample clustering based on aroma production. This analysis enabled the creation of three different sample clusters: A) Cluster formed by the initial wine must samples (filled dots) and the samples took at 72h (*H. uvarum_S. cerevisiae* T72h-triangles, and *S. cerevisiae* T72h – squares); B) Cluster formed by the final samples of the young wine produced by the consortium of *H. uvarum* MJT198 and *S. cerevisiae* QA23 (filled triangles); C) Cluster formed by the final samples of the young wine produced by the single inoculation of *S. cerevisiae* QA23 (filled squares).

The quantification of the major and minor aroma compounds (shown in **Figure IV.A. 7**) of the fermented musts after 72 hours of inoculation revealed that those fermented by *S. cerevisiae* QA23 in single culture exhibited higher levels of beta-phenylethanol, isoamyl alcohol, butyric, octanoic, hexanoic and decanoic acids; beta-phenylethyl acetate, ethyl octanoate and ethyl hexanoate (**Figure IV.A. 7I**). Higher levels of beta-citronelol, cis/trans rose oxide, 4-vinylguaiacol, trans-isoeugenol, ethyl-2-methylbutyrate, ethyl isovalerate and acetovanillone were also observed in the musts fermented by *S. cerevisiae*, compared to those fermented only by *H. uvarum* MJT198 (at 72h only this yeast was present since *S. cerevisiae* was inoculated at this time only). Differently, the musts fermented exclusively by *H. uvarum* MJT198 exhibited higher levels of 1-hexanol and methionol; of the monoterpenols linalool, geraniol and to a lesser extent, nerol (**Figure IV.A. 7I**). Overall, the determination of the major and minor volatiles revealed that *S. cerevisiae* had a more pronounced effect on the secondary aroma metabolites of wine at this stage, which was also reflected in the greater number of overall metabolites detected. This outcome was, somehow, expected since the production of these metabolites is associated with increased metabolic activity, which is higher in *S. cerevisiae* than in *H. uvarum*. In contrast, the higher metabolites linked to primary aroma (such

as monoterpenols) detected in the musts fermented by *H. uvarum* MJT198 suggest that this yeast's enzymatic activity (such as beta-glucosidase) may play a crucial role in the early stages of fermentation helping to release aromatic compounds from glycosidic precursors.

Comparing to the wines produced by *S. cerevisiae* QA23, those obtained using the consortium of *H. uvarum* MJT198 and *S. cerevisiae* QA23 exhibited significantly higher levels of major volatiles such as 1-hexanol and methionol, isobutanol, isobutyric acid, ethyl propanoate, acetaldehyde, and ethyl acetate (**Figure IV.A. 7II**), as well 4-vinylphenol, gamma-butyrolactone, beta-citronellol or beta-damascenone. In the other direction, the wines produced by *S. cerevisiae* QA23 exhibited higher levels of octanoic, butyric, and decanoic acids; of beta-phenylethyl acetate, isoamyl and hexyl acetate, ethyl decanoate, ethyl butyrate, and of 1-butanol; and of the minor volatiles ethyl isovalerate and gamma-octalactone (**Figure IV.A. 7II**). Consistent with the results obtained at 72 hours, the wines obtained by fermentation with *H. uvarum* were richer in terms of minor aroma volatiles, while those obtained with *S. cerevisiae* QA23 alone showed a strong influence on the production of volatile fatty acids. The higher levels of beta-phenylethyl acetate and isoamyl acetate obtained in wines fermented with *S. cerevisiae* alone were, at some extent, surprising since the use of *H. uvarum* as starter cultures usually leads to wines with increased concentrations of acetate esters (Zhang et al., 2023; Mestre et al., 2019; Hu et al., 2018; Hall et al., 2017; Bordet et al., 2020).

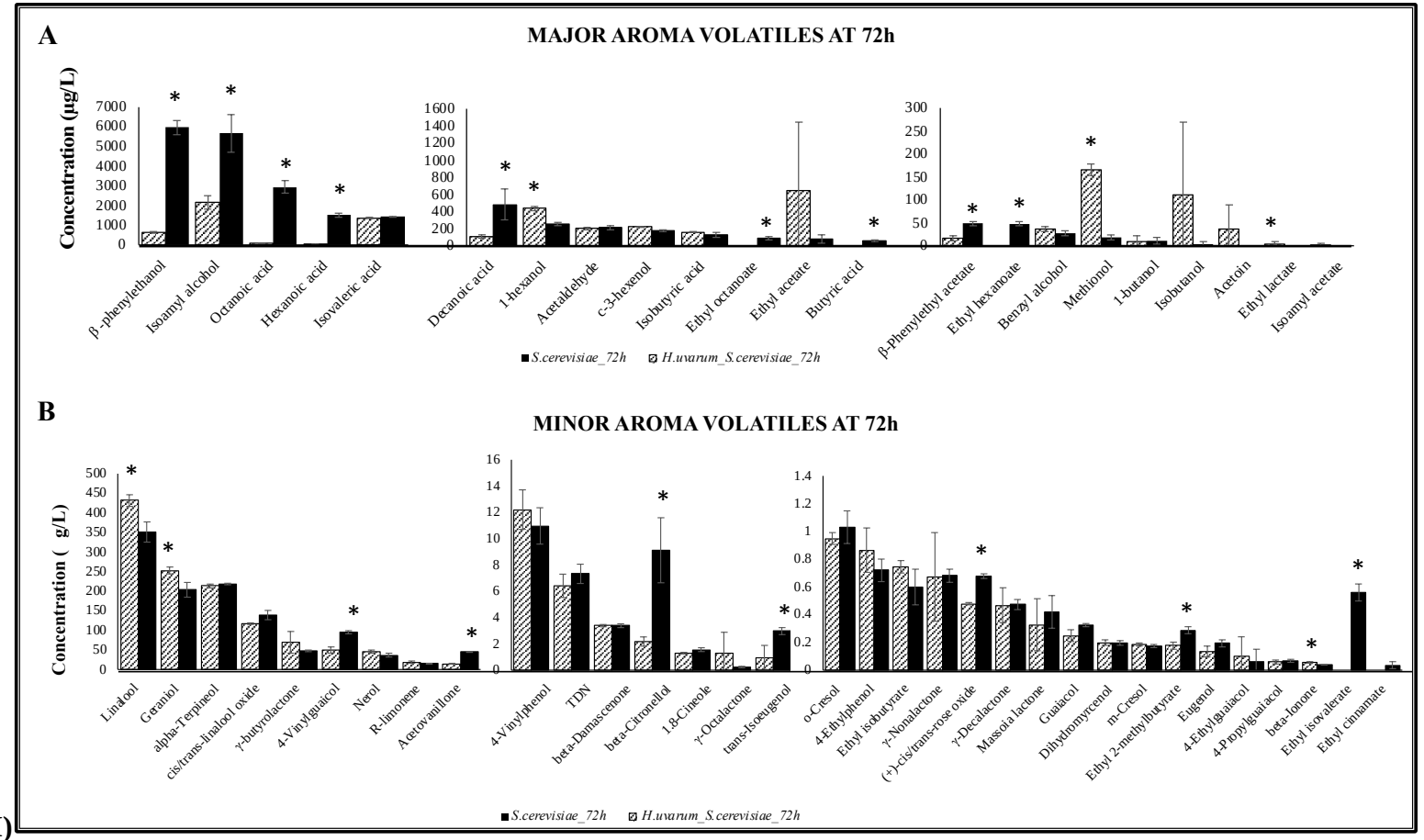
A study performed by Tristezza et al. (2016) that also explored a *H. uvarum* strain in sequential and mixed cultures with *S. cerevisiae* also resulted in wines with reduced content of beta-phenylethyl acetate in the wines fermented with *H. uvarum,* whether in simultaneous or sequential inoculation with *S. cerevisiae*, compared to those fermented with *S. cerevisiae* alone. These findings seem to suggest that inter-strain variability may exist, and some *H. uvarum* strains may be more prone to produce beta-phenylethyl acetate than others. As observed in the herein described fermentations, other studies also exploring *H. uvarum* described a positive impact in terpene concentrations of using this species as a starter culture (Hu et al., 2016; López et al., 2016; López et al., 2015). Notably, the strain *S. cerevisiae* QA23 herein used is also described as having beta-glucosidase activity (Lallemand, 2020), however, the wines produced through co-inoculation with *H. uvarum* exhibited higher levels of nearly all terpenes tested (though not all increases were statistically significant), compared to wines fermented solely with *S. cerevisiae*.

The webcharts represented in **Figure IV.A. 8** summarize the influence of the inoculation strategy on the wine major and minor volatiles. For instance, at 72h, the must

fermented with *S. cerevisiae* QA23 contained higher levels of minor aroma volatiles compared to that of *H. uvarum* (**Figure IV.A. 8, frame i**), while by the end of fermentation (**Figure IV.A. 8, frame iii**), the scenario had significantly changed: the presence of strain *H. uvarum* MJT198 also impacted led to an increase in every minor volatile category, including ethyl esters, monoterpenols, cinnamates and lactones (**Figure IV.A. 8, frame iii**). In contrast, the single fermentation with *S. cerevisiae* QA23 resulted in a higher concentration of major aroma volatiles, except for carbonyl compounds and alcohols that were higher in the wines obtained with the consortium (**Figure IV.A. 8, frame iv**).

Overall, these findings may suggest that the use of *H. uvarum* enhances the overall release of terpenes, probably because the enzymes of this non-*Saccharomyces* yeast have distinct properties than those described in *S. cerevisiae* (*e.g.*, different substrate specificities). This enzymatic diversity between the two species may allow them to complement each other's metabolic capabilities, leading to a broader release of aromatic compounds related to wine primary aroma.

I)


A
MAJOR AROMA VOLATILES AT 72h
Concentration (µg/L)
β-phenylethanol
Isoamyl alcohol
Octanoic acid
Hexanoic acid
Isovaleric acid
Decanoic acid
1-hexanol
Acetaldehyde
c-3-hexenol
Isobutyric acid
Ethyl octanoate
Ethyl acetate
Butyric acid
β-Phenylethyl acetate
Ethyl hexanoate
Benzyl alcohol
Methionol
1-butanol
Isobutanol
Acetoin
Ethyl lactate
Isoamyl acetate
S.cerevisiae_72h
H.uvarum_S.cerevisiae_72h
B
MINOR AROMA VOLATILES AT 72h
Concentration (g/L)
Linalool
Geraniol
alpha-Terpineol
cis/trans-linalool oxide
γ-butyrolactone
4-Vinylguaicol
Nerol
R-limonene
Acetovanillone
4-Vinylphenol
TDN
beta-Damascenone
beta-Citronellol
1,8-Cineole
γ-Octalactone
trans-Isoeugenol
o-Cresol
4-Ethylphenol
Ethyl isobutyrate
γ-Nonalactone
(+)-cis/trans-rose oxide
γ-Decalactone
Massoia lactone
Guaiacol
Dihydromyrcenol
m-Cresol
Ethyl 2-methylbutyrate
Eugenol
4-Ethylguaiacol
4-Propylguaiacol
beta-Ionone
Ethyl isovalerate
Ethyl cinnamate
S.cerevisiae_72h
H.uvarum_S.cerevisiae_72h

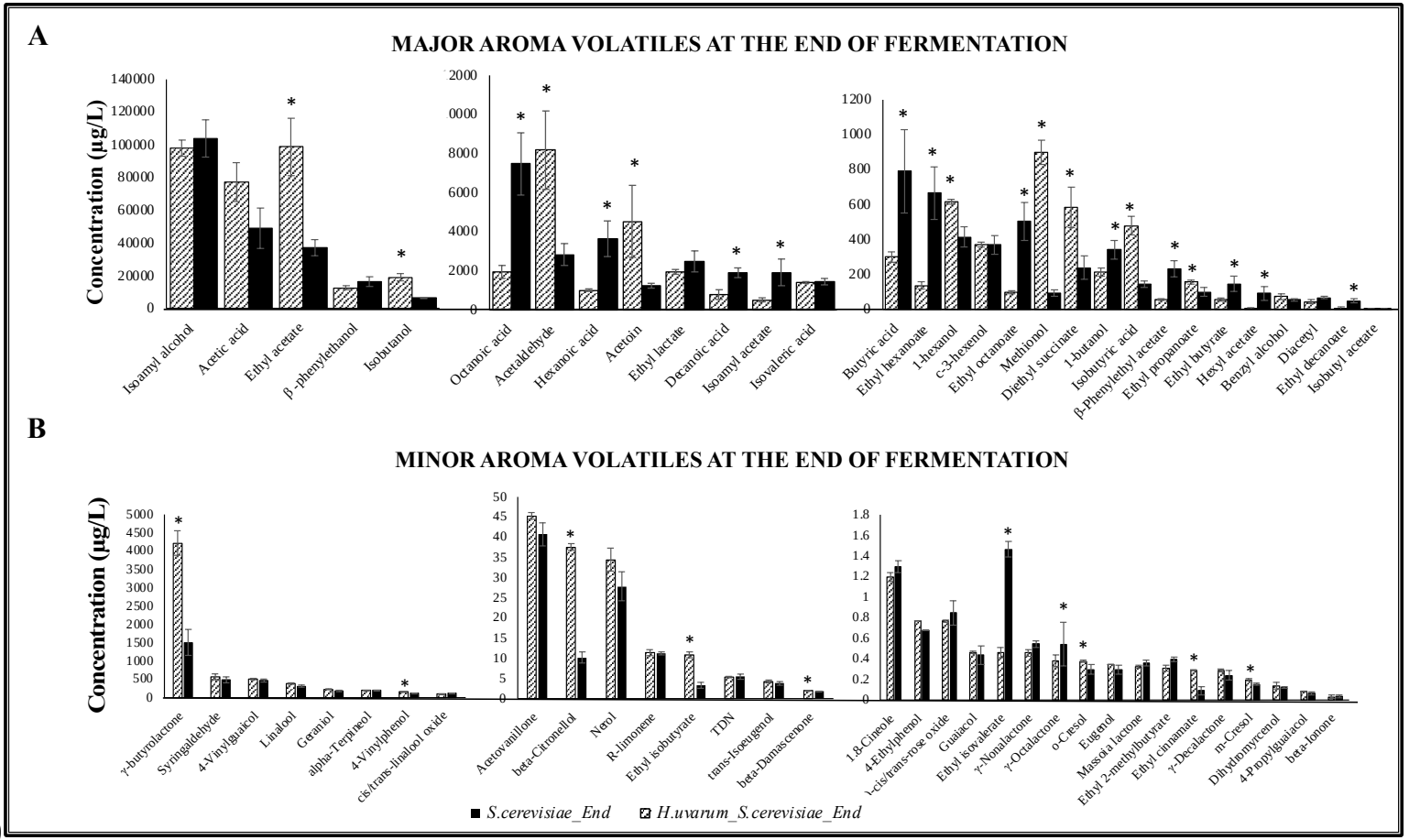


Figure IV.A. 7 - Major and minor aroma volatiles by *H. uvarum* MJT198 and *S. cerevisiae* QA23. I) Volatiles produced at 72h; II) Volatiles produced at the end of fermentation. Statistically significant differences (p-value < 0.05) are represented by *.

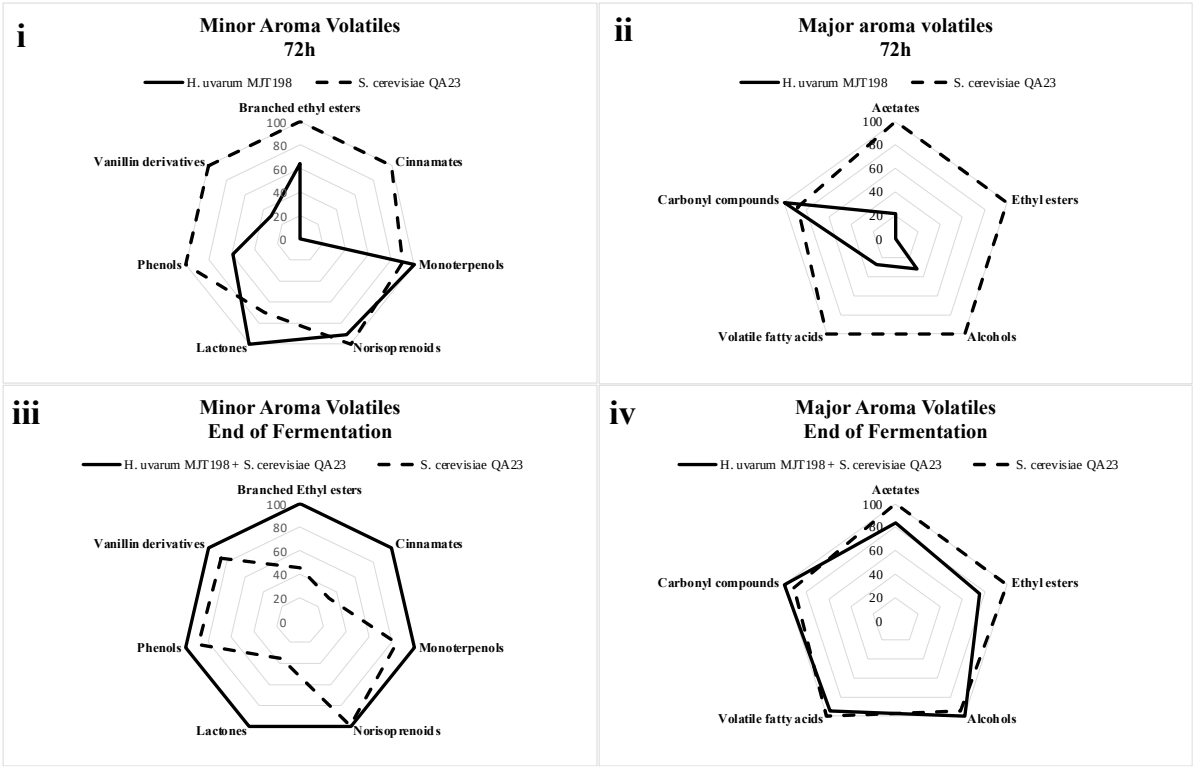


Figure IV.A. 8 - Webcharts showing the differences in volatile wine aroma compounds in the wines obtained with *S. cerevisiae* QA23 (complete line) and with the co-inoculation of *H. uvarum* MJT198 and *S. cerevisiae* QA23.

Chapter IV. B.

OMICs-analysis of mixed *H. uvarum*-*S. cerevisiae* fermentation: the perspective of the Non-*Saccharomyces* species

IV.B.1 Abstract

The complex dynamics among yeast species during fermentation plays a pivotal role in shaping the aromatic characteristics of wines. Gradually shifting from the conventional approach of using *Saccharomyces cerevisiae* as primary wine starter culture, there is a growing exploration of Non-*Saccharomyces* Yeasts (NSYs) due to their potential to diversify wine aromas. *Hanseniaspora uvarum* is one of such NSYs whose utilization as an adjunct of *S. cerevisiae* is being considered due to its aromatic properties. In this chapter, the changes in the gene expression of *H. uvarum* while in co-culture with *S. cerevisiae* are disclosed, an approach that differs from most work undertaken in the literature that focused on the budding yeast but not the NSY. Also, an untargeted metabolomics analysis (based on Fourier Transform Ion Cyclotron Resonance Mass Spectrometry, FT-ICR-MS) was used to understand the overall impact of the presence of *H. uvarum* in the body of metabolites accumulated in the fermented wine, allowing us to take a clearer (and broader) picture of the bio-flavoring potential of this species.

Keywords: Metabolomics, FT-ICR-MS, untargeted wine metabolome, non-volatile, transcriptomics

IV.B.2 Introduction

The integration of "OMICs" technologies as a mean to unveil the dynamics of microbial interactions in the context of winemaking has gained considerable attention and unveiled important aspects that show that the interaction is more than the sum of the partners involved (Tzachristas et al., 2021; Roullier-Gall et al., 2020). In particular, the use of genomics and transcriptomics, coupled with phenotypic data, was found to offer an integrated perspective into yeast responses to the environment during fermentation (Tzachristas et al., 2021; Roullier-Gall et al., 2020; Lloyd et al., 2015). Additionally, metabolomics, defined as the study of all metabolites in a biological system under specific physiological conditions (Clish et al., 2015), is a valuable tool to comprehend the microbial interactome during fermentation and to obtain metabolic footprints of yeast strains in various fermentation conditions (Lloyd et al., 2015).

Despite several studies have addressed the metabolomic analysis of wine fermentations, these mainly focused the changes on the volatile wine metabolome, traditionally employing methodologies (like gas chromatography coupled with mass spectrometry (GC-MS) (Rossi et al., 2023), that are oriented towards a certain panoply of metabolites that are selected, *a priori*, by the user as targets. More recently, the wine non-volatile metabolome has been studied using other methodologies, such as Ultra-high resolution mass spectrometry (uHRMS), by Fourier transform ion-cyclotron mass spectrometry (FT-ICR-MS) or ultra-high performance liquid chromatography-electrospray tandem mass spectrometry (UHPLC-MS) (Rocchetti et al., 2018; Bordet et al., 2023; Bordet et al., 2021; Wang et al., 2024; Roullier-Gall et al., 2020; Roullier-Gall et al., 2014), allowing a more detailed view regarding yeast interactions during wine fermentations. In this study we have explored FT-ICR-MS and for those reasons this technique is detailed below.

FT-ICR-MS is an ultra-high-resolution technique that allows an unmatched mass accuracy thus making it a powerful tool for an untargeted metabolomics analysis (Xie et al., 2022). The mode of action of a FT-ICR-MS is considerably distinct from the other MS methods, since it is the only technique where the ions are not detected by hitting a detector, but only passing near the detection plates (de Hoffman et al., 2007) (**Figure IV.B. 1**) .

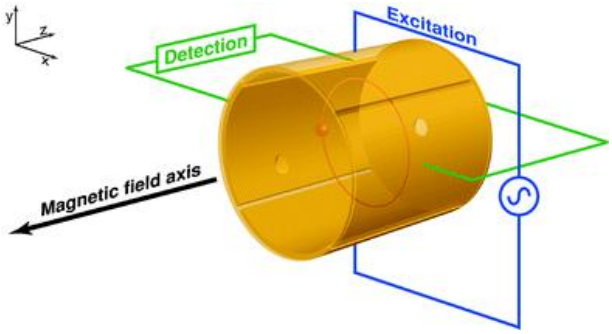


Figure IV.B. 1 - Graphical representation of a cylindrical FT-ICR analyzer cell. In this representation, the magnetic field is located in the z-axis, while the excitation plates are located along the y-axis and the detection plates along the x-axis. The trapping electrodes are located at each end of the cell, and the orbiting ions are represented by the red circle. From Barrow et al., 2004

The masses are determined by the ion cyclotron resonance frequency that each ion produced as it rotates in the magnetic field, in opposition to other techniques where mass is known in space and in time (de Hoffman et al., 2007; Comisarow et al., 1974; Barrow et al., 2005). This causes the ions to be detected simultaneously during the detection interval increasing the signal-to-noise ratio, and consequently, resolution de Hoffman et al., 2007; Comisarow et al., 1974; Barrow et al., 2005). Prior to entering the detector, ions are generated externally by an ion source and then injected into the analyzer cell which is located within a strong magnetic field de Hoffman et al., 2007; Comisarow et al., 1974; Barrow et al., 2005). Electrospray ionization (ESI) is one of the most used ionization techniques, especially coupled to FT-ICR-MS, being extremely fit for the analysis of thermally labile large supramolecules (Banerjee & Mazumdar, 2012). Being classified as a soft ionization technique, this technique achieves little to no fragmentation of the molecules, enabling the ionization of intact chemical species by multiple charging. These characteristics make ESI+ suitable for the determination of large, biologically important macromolecules, like proteins and nucleic acids and their consequent mass spectrometric analysis and structural characterization (**Figure IV.B. 2**).

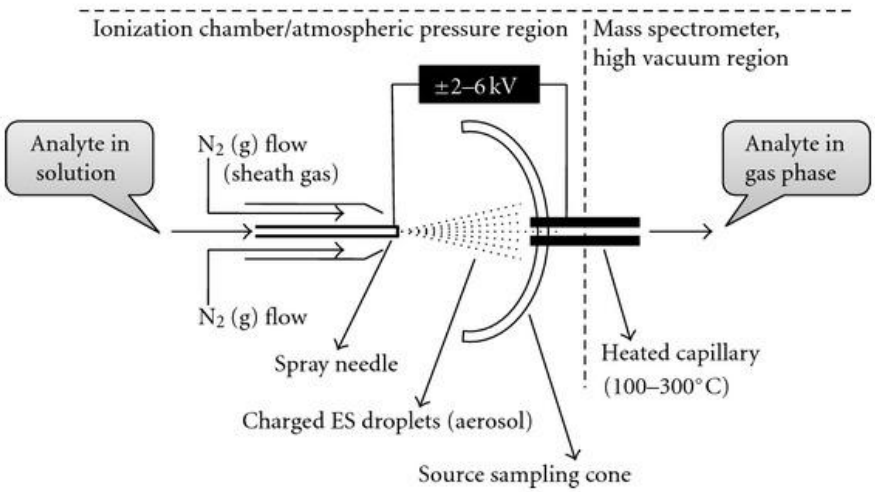


Figure IV.B. 2 - Functioning of ESI. I) The analyte in a diluted solution is injected into the system by a needle; ii) a very high voltage (2-6 kV) is applied to the tip of the metal capillary, causing the dispersion of the sample into an aerosol of charged droplets; iii) The charged analytes are released from the droplets and some pass through the orifice of a heated capillary into the analyzer of the mass spectrometer. The coaxial flow of dry N2 around the capillary is determinant for an effective nebulization. From Banerjee & Mazumbdar, 2012.

There have been several works exploring FT-ICR-MS, and other metabolomics techniques in the study of wine metabolomics. For example, a recent study conducted by Roullier-Gall and colleagues (2020), employed FT-ICR-MS to explore the chemical non-volatile metabolome from pure and mixed culture fermentations of Sauvignon Blanc with *Lachancea thermotolerans*, *Starmerella bacillaris*, *Metschnikowia pulcherrima*, and *S. cerevisiae*. The non-volatile metabolomic analysis confirmed that mixed fermentations lead to distinct wine metabolite compositions (Roullier-Gall et al., 2020). Wines produced from sequential fermentations using co-cultures differed from those conducted in single fermentation, suggesting non-neutral yeast interactions during the winemaking process (Roullier-Gall et al., 2020). Additionally, in a recent work, Cabernet Sauvignon wines fermented by *H. uvarum* in the presence and absence of *S. cerevisiae* were analyzed using UHPLC-MS (Wang et al., 2024). The non-volatile metabolomics analysis showed specific metabolomic signatures of the different wine fermentations performed using single strains (i.e. *H. uvarum*) or consortia (*i.e. H. uvarum* + *S. cerevisiae*). In particular, it was found that the fermentations conducted using *H. uvarum* strains led to depletion of polypeptides, amino acids, and nucleotides, compared to the fermentations conducted only with *S. cerevisiae*. The presence of *H. uvarum* strains also altered the polyphenolic composition ratios and decreased indole derivatives in the fermented wine (Wang et al., 2024). Ultimately, the obtained data indicates that metabolomics is a valuable tool for deepening our understanding of the influence yeast strains have on the final wine. By extending beyond the analysis of volatile compounds, which alone cannot fully

capture the complexity perceived by a wine taster, metabolomics offers a more comprehensive view over the wine's characteristics.

In the preceding chapter of this thesis, strain *H. uvarum* MJT198, isolated from a terpene-rich Muscat wine must and presenting very high beta-glucosidase activity, was shown to modify primary wine aroma when used in co-culture with *S. cerevisiae* in Moscatel Galego wine fermentations. This positive effect of *H. uvarum* MJT198 was consistent with the identification in its genome of a set of beta-glucosidase encoding genes. In this chapter, we present a detailed metabolomic (focused on non-volatiles and using ultra-high-resolution mass spectometry based on FT-ICR-MS coupled with electrospray ionization in the positive mode) and transcriptomic analysis of those fermentations of Moscatel Galego must undertaken by *H. uvarum* MJT198 and *S. cerevisiae* QA23. It is the goal of these OMICS analyses to provide a broader understanding of the interactions established between the two yeast species during wine fermentation, as well as the role of the NSY in modifying the wine metabolome beyond the "expected players".

IV.B.3 Methods and Materials

IV.B.3.1 Fermentations in natural must with *H. uvarum* MJT198 and *S. cerevisiae* QA23

H. uvarum MJT198, isolated from Moscatel Galego must harvested in 2019 at a local winery in the Setúbal Peninsula, Portugal, and the commercial wine yeast strain *S. cerevisiae* QA23 (Proenol S.A, Vila Nova de Gaia, Portugal) were used. Both strains were preserved in YPD medium at -80ºC until further use. The fermentations of Moscatel Galego must using *H. uvarum* MJT198 and *S. cerevisiae* QA23 were those described in the previous **Chapter IV.A.**

IV.B.3.2 Metabolomic profiling using FT-ICR-MS.

To analyze the metabolome of the different samples taken along the fermentation (F1, F2 and F3 (resulting from the sequential inoculation of *H. uvarum* MJT198 and *S. cerevisiae* QA23), and F4, F5 and F6 (resulting from the *S. cerevisisae* QA23 control) of the Moscatel Galego must, FT-ICR-MS was used. For this, the samples taken from the fermentation were centrifuged (10 minutes at 13500 rpm, 4ºC), diluted 1:100 in a solution of methanol and water in a 1:1 ratio, containing 1 µL of formic acid (Sigma Aldrich, MS grade). 1 mg/mL leucine-enkephalin (YGGFL, Sigma Aldrich) was added to each sample to serve as the internal standard (molecular mass of [M+H]+ = 556.276575 Th), considered for analysis by electrospray inonization in positive mode (ESI+). Afterwards, the samples were injected into a Solarix XR 7T FTICR-MS (Fourier-Transform Ion Cyclotron mass spectrometer, Brüker Daltonics) using direct infusion at a flow rate of 240 µL/h and employing electrospray ionization in positive mode. Mass spectra were acquired in absorption mode, accumulating 300 scans of time-domain transient signals in 4M-point time-domain data sets. The mass spectra were recorded in the range of 100 to 1000 m/z. The injection system was cleaned with methanol (LC-MS grade, Merck) between samples. The obtained spectra were visualized and analyzed using Compass DataAnalysis software V 5.0 (Brüker Daltonics) to: i) determine the m/z values (obtained using FTMS parameters without isotopic deconvolution) of all observed peaks; ii) identify possible compounds, through the Smart Formula addon and resorting to an in-house database of metabolites encompassing the human metabolome database (HMDB) and the yeast metabolome database (YMDB).

IV.B.3.3 Total RNA extraction and sequencing

Total RNA was isolated from the samples obtained from the cell pellets recovered along the fermentation using the RiboPure Yeast Kit (Thermofisher Scientific, Life Technologies, Carlsbad, CA, USA) according with the manufacturer's instructions. While we extracted RNA from three biological replicates, due to technical issues with the samples, reliable transcriptomic data could only be obtained for two out of the three biological replicates considered for each time point.

The extracted RNA was treated with DNAse I (DNA-free™ Kit Ambion, Life Technologies) and the yield and purity of the obtained RNA analysed using an Agilent 2100 Bioanalyzer. The QuantSeq libraries were prepared using Lexogen's QuantSeq 3′ mRNA-Seq Library Prep Kit for Illumina, according to the manufacturer's instructions. The QuantSeq libraries were sequenced using the Illumina NextSeq 2000 to produce 120 bp single-end reads for each sample. Library preparation and sequencing was done at the Genomics Unity of the Gulbenkian Science Institute (Oeiras, Portugal). The obtained reads from the sequenced were processed using the OmicsBox software (Biobam, v3.12). Reads were trimmed using the Trimmomatic (Bolger et al., 2014) package, using a quality cut-off of Q20; and afterwards aligned to the genomes of *H. uvarum* MJT198 and *S. cerevisiae* QA23 using STAR (Spliced Transcripts Alignment to a Reference) (Dobin et al., 2013). Pairwise differential gene expression (DGE) analysis was performed to identify differentially expressed genes using the software package edgeR (empirical analysis of DGE in R, v3.28.0) (Robinson et al., 2010). The normalization method chosen was TMM (Trimmed Mean of M values) and the statistical test used to assess statistical significance of the differences found was GLM (Likelihood Ratio Test). Genes were considered to be differentially expressed in a given pairwise comparison when the associated fold-change was above 2 and the corresponding p-value inferior to 0.05.

Results and Discussion

IV.B.4.1 Non-volatile Metabolomics enabled the metabolic fingerprint of mixed fermentations using *H. uvarum* MJT198 and *S. cerevisiae* QA23

To broaden the study of the impact of *H. uvarum* in the body of metabolites present in the fermentation broth of natural Muscat grapes, in this chapter it was performed an untargeted metabolomic analysis (based on FT-ICR-MS) of the supernatants obtained in the fermentations described in **Chapter IV.A** corresponding to: i) fermentations of Muscat grape must carried out by *S. cerevisiae* at 72 hours and at the end of fermentation; ii) fermentations of Moscatel Galego grape must carried out by *H. uvarum* MJT198 at 72 hours, and by *H. uvarum* MJT198 and *S. cerevisiae* at the end of fermentation; iii) supernatants obtained immediately after inoculation of the two yeasts in the grape must. An overview of the experimental setting is provided in **Figure IV.B. 3.**

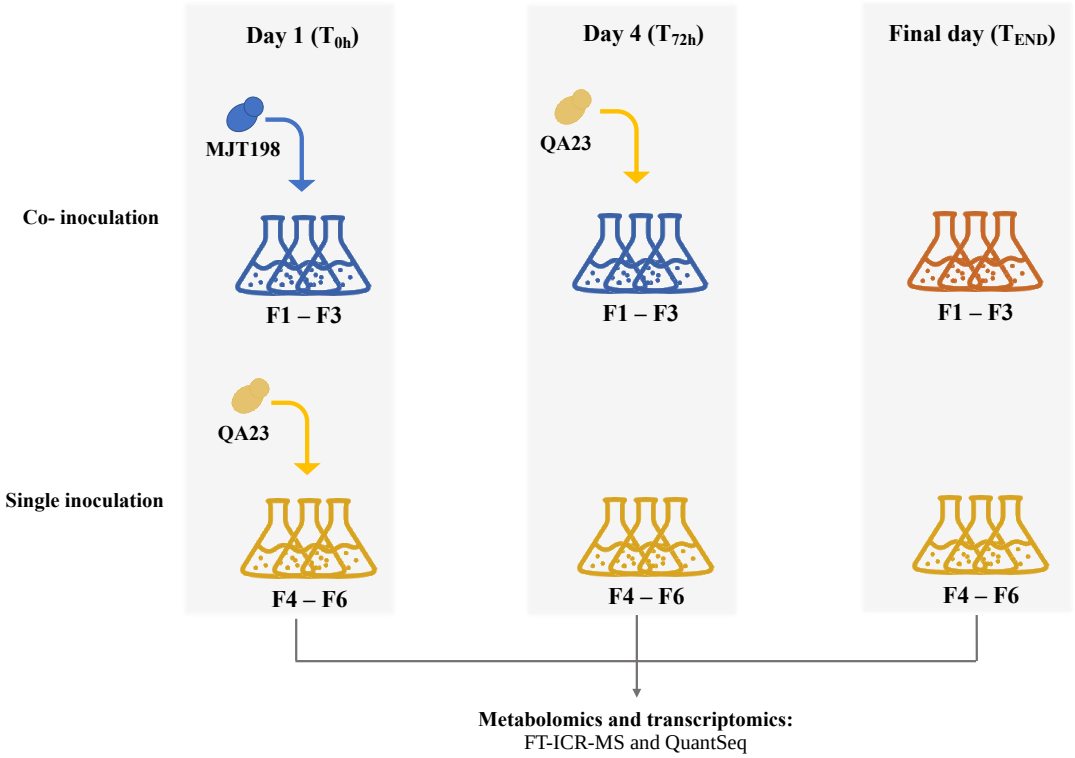
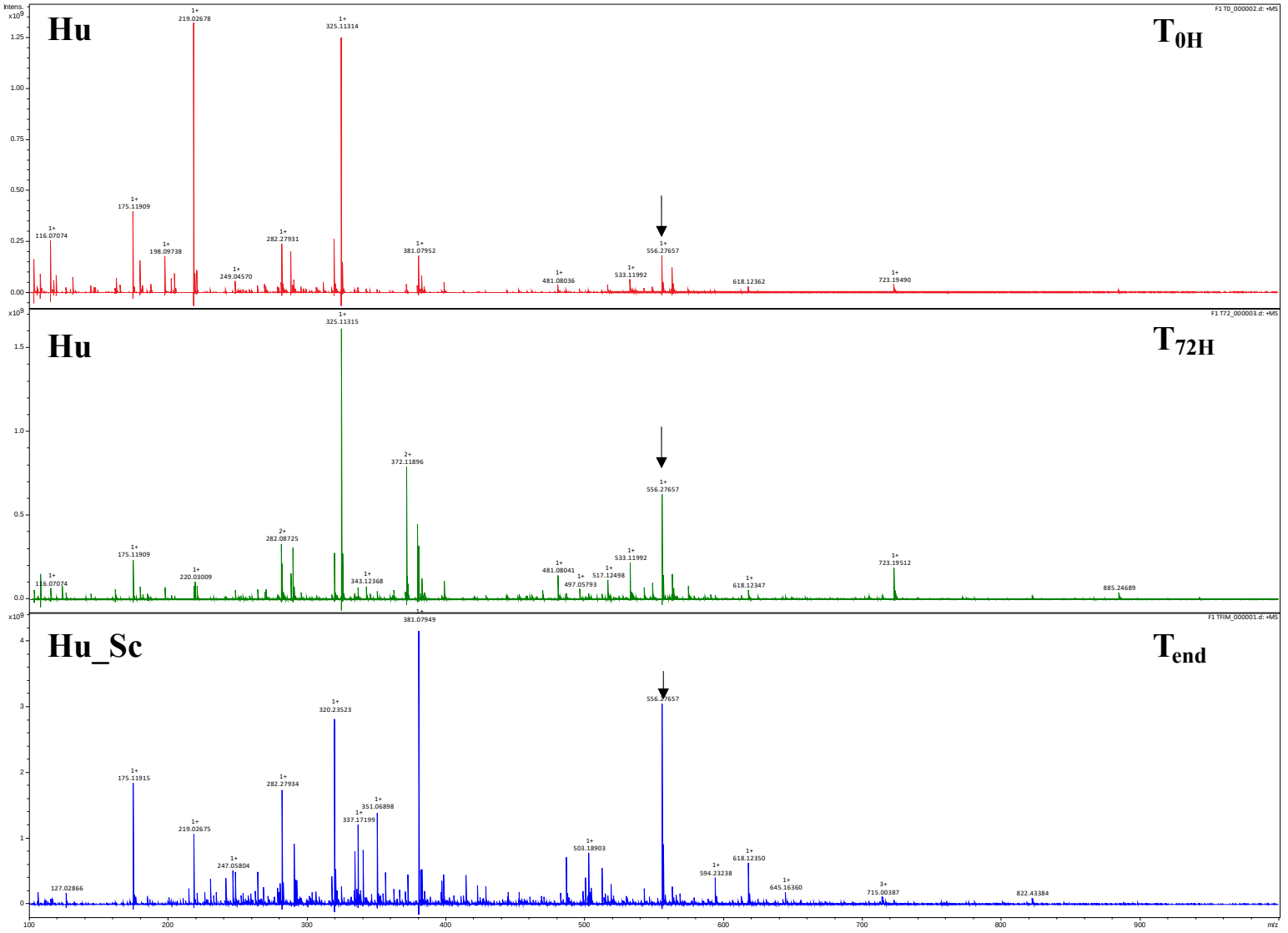


Figure IV.B. 3– Experimental setup for the analysis of the non-volatile metabolome and transcriptome during fermentations in Moscatel galego must. The samples were obtained following the sequential inoculation of *H. uvarum* MJT198 followed by *S. cerevisiae* QA23, and single inoculation of *S. cerevisiae*, at different timepoints.

The experimental setup and subsequent analysis performed represented the first exploration of *H. uvarum's* influence on the wine metabolome, coupled with transcriptomics analysis, in co-culture with *S. cerevisiae*, using FT-ICR-MS. In **Figure IV.B. 4** it is shown a comparison of the raw FT-ICR-MS spectra of the wines obtained after fermentations undertaken by *S. cerevisiae* QA23, alone or together with *H. uvarum* (after 72h and in the end of the fermentations), in comparison with the spectra obtained with the grape must immediately after inoculation of the two yeasts. Such results render clear the impact of the yeasts in the fermentation metabolome, since the profile of the spectra obtained at 72h and, more prominently, at the end of the fermentation, were much different from those of the initial grape must specially in what concerns the number of peaks.

Overall, FT-ICR-MS identified 2,468 peaks with different m/z for the different samples tested (**Appendix Table VI.B.1**). To improve the reliability of the comparisons of the different FT-ICR-MS spectra, a normalization step was performed: each peak's intensity was divided by the intensity of the internal standard (leucine-enkephalin) to mitigate variance arising from sample preparation and instrument response, and afterwards logarithmically transformed.



A)

B)

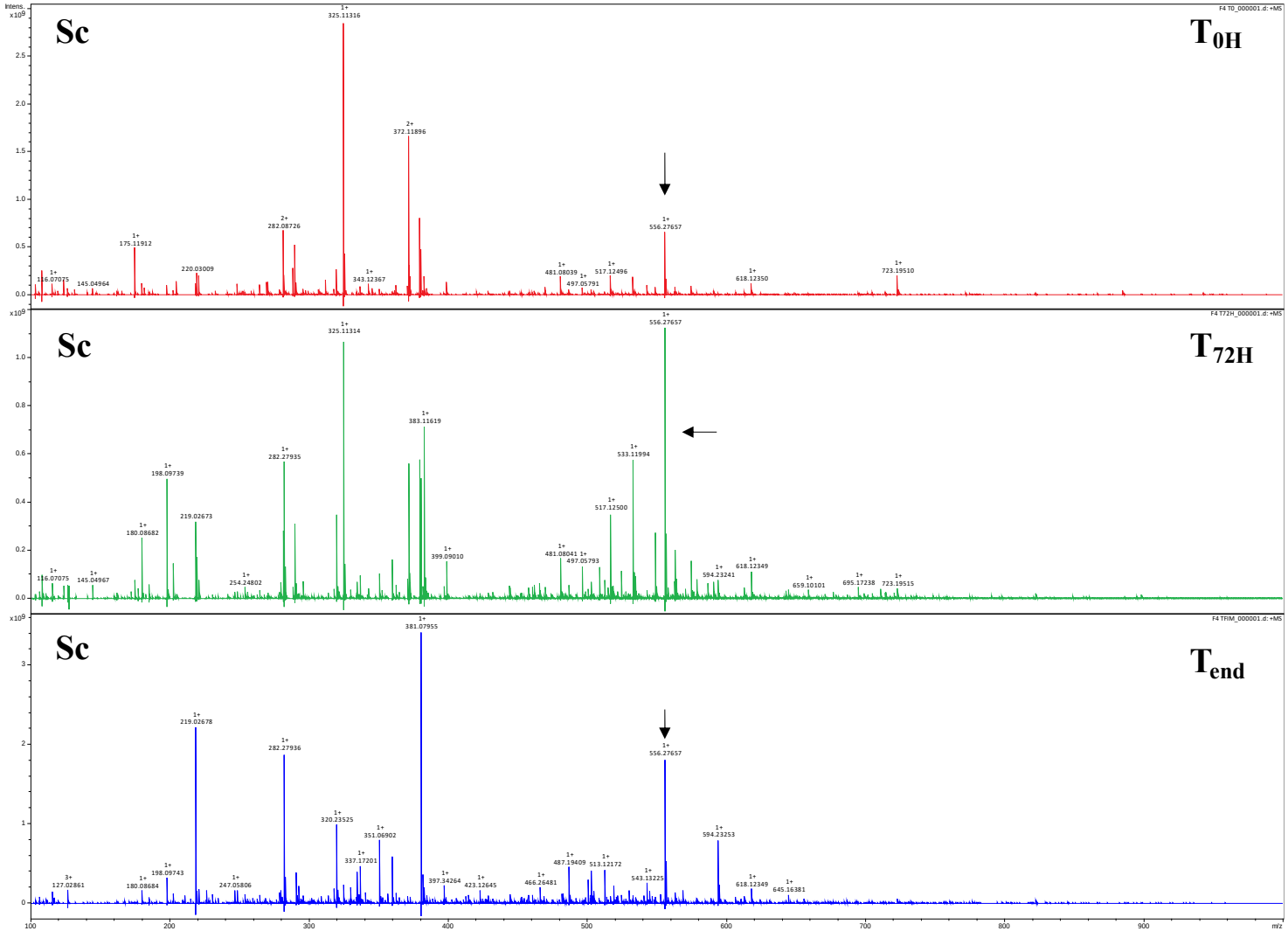


Figure IV.B. 4– Visual spectra obtained from FT-ICR-MS at T0h, T72h, and Tend, for A) The sequential inoculation strategy of *H. uvarum* MJT198 followed by *S. cerevisiae* QA23, and B) The single inoculation of *S. cerevisiae*. The peak corresponding to the internal standard Leucine-enkephalin is highlighted with an arrow and can be found at m/z 556.276575.

Principal component analysis (PCA) of normalized data (**Figure IV.B. 5**) led to the identification of four clusters: cluster A) - corresponding to the initial samples (T_{0h}) and samples obtained after 72 hours of fermentation with *H. uvarum*; cluster B) - corresponding to the samples obtained from the broth fermented only by *S. cerevisiae* QA23 after 72 hours; cluster C) - corresponding to the samples obtained at the end of the mixed fermentation by *H. uvarum* MJT198 and *S. cerevisiae* QA23 (T_{end}); cluster D) - corresponding to the samples acquired at the end of fermentation by *S. cerevisiae* QA23 (T_{end}) (**Figure IV.B. 5**). This separation of the data was interesting as it was consistent with the low fermentative activity of *H. uvarum* (thus resulting in less changes in the fermentation broth compared to the initial must), with higher fermentative activity of *S. cerevisiae* that necessarily changes the fermentation broth to something different from the initial must and also with the impact of using one or the two yeasts in the inoculum.

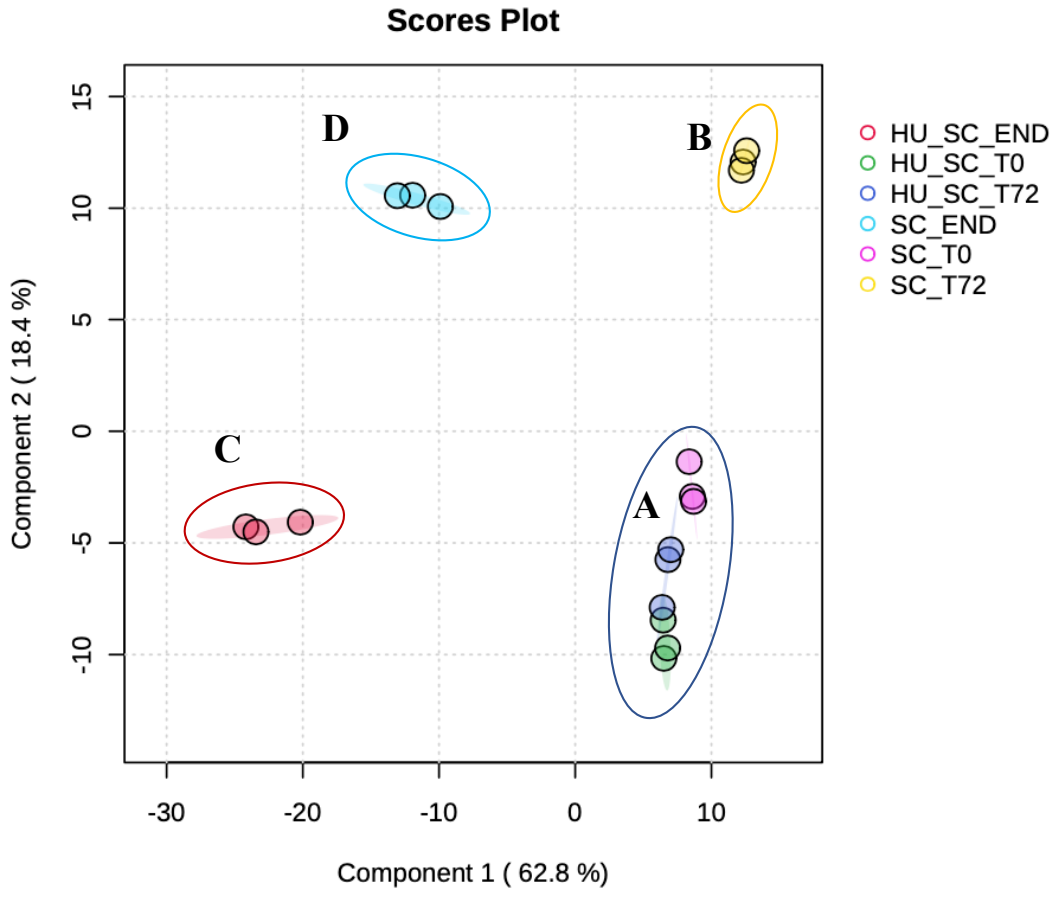


Figure IV.B. 5 - PCA biplot of the data obtained by FT-ICR-MS. In this chart it is possible to find four cluster of samples: A) composed by the samples took at T_{0h} for every sample, and T_{72h} for the sequential inoculation; B) composed by the samples of the single inoculation of *S. cerevisiae* at T72h; C) composed by the samples of the sequential inoculation of *H. uvarum* MJT198 and *S.cerevisiae* QA23 at the end of fermentation; D) composed by the samples of the single inoculation with *S. cerevisiae* at the end of fermentation. PCA was performed using the statistical analysis of the MetaboAnalyst web-based platform (Xia et al. 2009).

Out of the 2,468 peaks analysed, only 1,956 (79%) could be assigned a molecular formula using the SmartFormula add-on in MetaboScape, and of those, only 693 (28%) were successfully matched to an identified molecular name. The low identification rate reflects a common challenge in metabolomics, and the need for continuously updated and expanded metabolomic databases, particularly for yeast and plant metabolites. It is also important to note that the adducts and mass tolerance selected for peak annotation, greatly influenced the identification assigned to each peak, resulting in several peaks having the same annotation. To unequivocally confirm each peak identity, other technologies, such as GC-MS, could be used.

To obtain a baseline of the grape must composition, we analyzed the spectra that were obtained immediately after the inoculation with *H. uvarum* (samples HU_T0h) or with *S. cerevisiae* (samples SC_T0h) in order to identify the peaks common to the two. From this analysis, a total of 1,508 common peaks were identified in the several analyzed spectra this being believed to correspond to the components of the initial wine must (**Figure IV.B. 6)**. Among the identified compounds in these peaks assigned to the grape must were L-Arginine, L-Aspartic Acid, L-Glutamic Acid, L-Glutamine and L-Isoleucin; complex glucosides like 2,3-butanediol glucoside and Cyanidin 3-glucoside were found; along with plant-derived compounds like gibberellin (a key phytohormone) and epicatechin (a flavonol) (**Figure IV.B. 6)**. Interestingly, the presence of chitin was also detected, which can be a reflection of the presence of yeasts (that have chitin on their cell walls) or even small insects that could have been present on the grape surface during crushing (**Figure IV.B. 6)**.

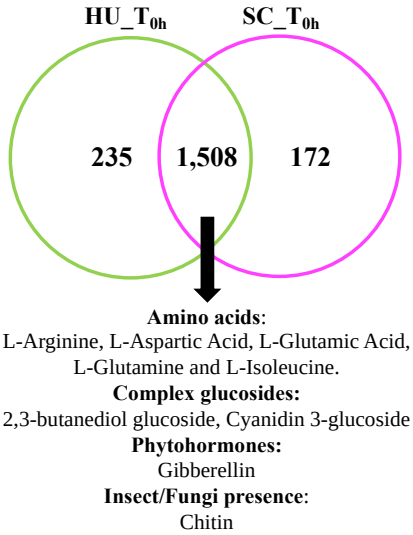


Figure IV.B. 6 - Specific and common peaks identified in samples taken at the fermentation onset, after inoculation with *H. uvarum* MJT198 (HU) and *S. cerevisiae* QA23 (SC). Some common peaks were regarded as belonging to the wine must.

To better understand the major differences observed 72 hours after inoculation of the yeasts, the peaks (representing metabolites) identified in the spectra of samples taken after 72 hours of *H. uvarum* fermentation were compared with those obtained immediately after the inoculation of the yeasts in the grape must (sample 0h). This analysis identified 220 peaks (corresponding to different metabolites) that were only present after the 72h of *H. uvarum* activity, while 1,568 peaks were found at time 0h and after 72h (**Figure IV.B. 7A**). The same comparison made with the spectra obtained from the musts fermented by *S. cerevisiae*, showed 354 peaks that were not initially present and 1,090 in common with those initially present (**Figure IV.B. 7B**).

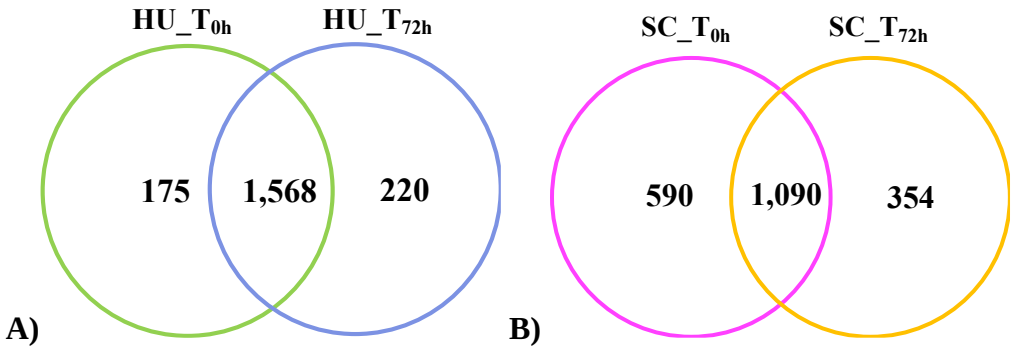


Figure IV.B. 7– Venn diagrams depicting the evolution in the number of peaks from the fermentation onset to 72 hours obtained for A) the sequential inoculation with *H. uvarum* MJT198 (HU) and B) the single inoculation with *S. cerevisiae* (SC).

When we compared the spectra of the fermentations obtained with *H. uvarum* or with *S. cerevisiae* at the 72h of fermentation, we found 1,602 peaks in common, while 506 were only found in the *H. uvarum* samples (140 identified) and 382 (79 identified) in the *S. cerevisiae* fermented samples **Figure IV.B. 8**). Among the "*H. uvarum*-exclusive" peaks, we found L-proline, L-arginine, L-isoleucine, D-glucose, glycogen, glucosamine, several arabinose derivatives (beta-D-Xylopyranosyl-(1->5)-alpha-L-arabinofuranosyl-(1->3)-L-arabinose and b-D-Xylopyranosyl-(1->4)-a-L-rhamnopyranosyl-(1->2)-L-arabinose), methoxypyrazines (such as 2-(methylthio)pyrazine), these being compounds typically found in green plant tissues, including grape berries, and associated with green pepper flavor (Reynolds, 2022). Additionally, several weak acids were also identified in these samples, as detailed (**Appendix Table IV.B.2**):

- (1xi,3xi)-1,2,3,4-Tetrahydro-1-methyl-beta-carboline-3-carboxylic acid;
- 3-(6-hydroxy-7-methoxy-2H-1,3-benzodioxol-5-yl)propanoic acid;
- 3,4,5-trihydroxy-6-{[(2E)-2-methyl-3-phenylprop-2-en-1-yl]oxy}oxane-2-carboxylic acid; 4-O-alpha-D-Galactopyranuronosyl-D-galacturonic acid;

- 3,4,5-trihydroxy-6-{3-hydroxy-4-[5-hydroxy-3-(3-hydroxy-3-methylbutyl)-8,8-dimethyl-4-oxo-4H,8H-pyrano[2,3-f]chromen-2-yl]phenoxy}oxane-2-carboxylic acid;
- ent-Epiafzelechin(2a->7,4a->8)epiafzelechin 3-(4-hydroxybenzoic acid);
- 2-({[3,4-dihydroxy-5-(3,4,5-trihydroxybenzoyloxy)phenyl](hydroxy)methylidene}amino)acetic acid
- Palmitic acid;
- Nicotinic acid

The strong acid [3-(4-methoxyphenyl)propoxy]sulfonic acid was also identified in the *H.uvarum* exclusive peaks.

In the samples obtained from the fermentations undertaken only with *S. cerevisiae* at 72 hours, we found the amino acids L-aspartic acid, L-phenylalanine, L-threonine and L-tyrosine (**Figure IV.B. 8**). Notably, no simple or complex sugars were detected in *S. cerevisiae* samples which may suggest a more advanced fermentation stage, compared to the *H. uvarum* samples at this timepoint where some of such sugars were identified. Additionally, weak acids were also prevalent in the *S. cerevisiae* peaks at this stage, which correlates with the results from the previous Chapter **(Appendix Table IV.B.2):**

- (S)-3-[(Cyanophenylmethyl)amino]-3-oxopropanoic acid
- 2-({[3,4-dihydroxy-5-(3,4,5-trihydroxybenzoyloxy)phenyl](hydroxy)methylidene}amino)acetic acid
- 2-hydroxy-3-(sulfooxy)butanedioic acid
- 3-Deoxy-D-glycero-D-galacto-2-nonulosonic acid
- 3-Hydroxysebacic acid
- 3,4,8,9,10-pentahydroxy-6-oxo-6H-benzo[c]chromene-1-carboxylic acid
- Citric acid
- Suberic acid

The analyses of the spectra obtained from the final wines revealed 979 peaks in common in the wines obtained with the consortia of yeasts or with *S. cerevisiae* alone, while 136 peaks were only found in the spectra obtained from the consortium (57 identified), and 247 only found in the fermentations carried out only with *S. cerevisiae* QA23 (**Figure IV.B. 8**) (83 identified). The peaks found only in the fermentations obtained with the consortia of yeasts included suberic acid, (±)-3-hydroxynonanoic acid, L-arginine, glycogen and glutathione (**Figure IV.B. 8**) Notably, the peaks found only in the fermentations undertaken by *S. cerevisiae* QA23 included several glucosides, such as linalool oxide, D 3-[apiosyl-(1->6)-glucoside], 7-hydroxy-2-methyl-4-oxo-4H-1-benzopyran-5-carboxylic acid 7-glucoside, and (S)-3-octanol glucoside. The fact that these molecules are present in the spectra obtained *with S. cerevisiae* but not with the consortia of yeasts may result from the absence in *S. cerevisiae*

of beta-glucosidase(s) that may drive the degradation of these molecules. The esters (S)-3-Mercaptohexyl pentanoate and 6-beta-Hydroxy-mometasone furoate, L-proline and L-tryptophan were also only identified in the samples fermented by the consortium (**Figure IV.B. *8***). The complete list of compounds and their annotation can be found in **Appendix Table IV.B.1,** while the peaks corresponding to each timepoint can be found in **Appendix Table IV.B.2.**

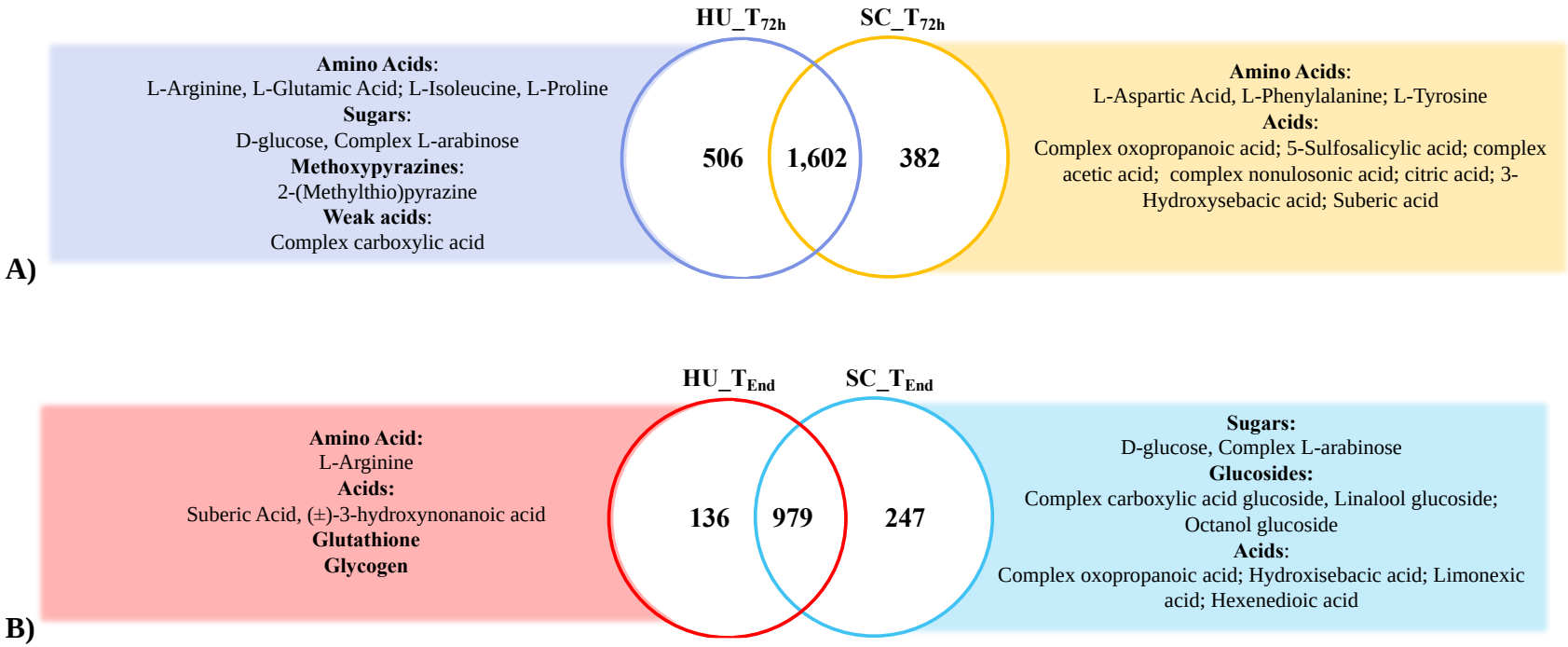
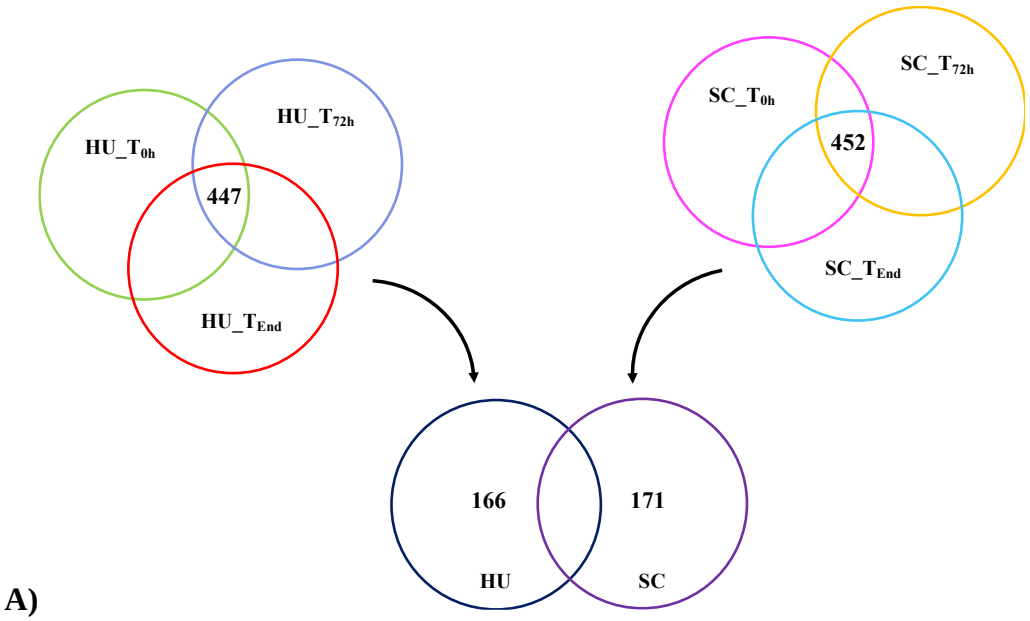


Figure IV.B. 8– Venn charts displaying specific and common peaks identified in the samples fermented by the sequential inoculation with *H. uvarum* MJT198 (HU) and the single inoculation with *S. cerevisiae* QA23 (SC) at A) 72 hours after inoculation, and B) the end of fermentation.

To further analyze the metabolomic fingerprints associated with *H. uvarum* inoculation in the must, an additional Venn diagram was constructed in which the peaks found in the spectra obtained only with *H. uvarum* (at 0h, 72h or in the end) were analyzed to find those in common that could not be found in any of the spectra obtained with *S. cerevisiae* alone (at 0h, 72h or in the end). Such analysis, shown in **Figure IV.B. 9**, allowed us to obtain more specific signatures of *H. uvarum*. With this analysis we could identify 166 peaks (presumed metabolites) linked to the presence of *H. uvarum*, while 171 were only associated with *S. cerevisiae* (**Figure IV.B. 9***)*. Among those exclusive to *H. uvarum*, it was possible to find L-arginine, palmitic acid, suberic acid, limocitrol and glycogen. In contrast, the fermentations carried out exclusively by *S. cerevisiae* revealed the presence of L-proline, glucosamine, and the monoterpene glucoside linalool-glucoside.

The detection of linalool-glucoside in the *S.cerevisiae* fermentations, but not in those involving *H. uvarum* is interesting as it may suggest that the action of beta-glucosidase(s) may have enhanced hydrolysis of the glucoside, releasing the bound linalool. In line with this hypothesis, the analysis of the volatile metabolome in the previous chapter highlighted higher levels of linalool concentration (431 μg/L compared to 351 μg/L) at 72 hours in the fermentations by *H. uvarum*, compared to those by *S. cerevisiae*. At the end of the fermentation, the levels of free linalool in the wines were also higher in the wines obtained with the consortia than with those obtained with *S. cerevisiae* alone (377 μg/L compared to 307 μg/L). The higher concentration of amino acids in the wines produced by *S. cerevisiae* is in line with the findings reported in Wang et al (2004), who also reported lower levels of L-lysine, L-arginine, L-histidine, L-leucine and D-proline in wines produced by *H. uvarum* compared to those produced by *S. cerevisiae*. Consistent with our findings, Wang et al. (2024) also obtained higher concentrations of suberic and palmitic acid in the *H. uvarum*- fermented wines, alongside a lower prevalence of fatty acids compared to *S. cerevisiae* wines. The low fatty acid content in *H. uvarum* wines correlates with a decreased ethyl ester production observed in the volatile metabolome. This is supported by literature, that emphasizes that ethyl ester concentrations are closely tied to fatty acid production (Hu et al., 2019; Hu et al., 2018).



B)

Exclusive to *H. uvarum*	Exclusive to *S. cerevisiae*
N-Acetylornithine	L-Proline
Hexanethioic acid S-propyl ester	L-Tryptophan
L-Arginine	Benzyl chloride
Palmitic acid	3-Hexenedioic acid
Suberic Acid	Glucosamine
Arginyl-Glutamine	Complex oxopropanoic acid
Limocitrol	(S)-3-Octanol glucoside
Glycogen	Linalool glucoside

Figure IV.B. 9– Venn diagrams that lead to the identification of the peaks exclusive to the presence of *H. uvarum* in the fermentation, and of *S. cerevisiae* in single inoculation. The upper left Venn diagram disclosed 447 peaks as common to the inoculation of *H. uvarum* in all timepoints. The upper right Venn diagram identified 452 peaks as common to all timepoints for the single inoculation with *S. cerevisiae*. The bottom Venn diagram is the resulting diagram that enabled the identification of 166 peaks exclusive to the presence of *H. uvarum* in the fermentation, and of 171 peaks exclusive to *S. cerevisiae.* B) Identification of some peaks exclusive to *H. uvarum* fermentations (left) and *S. cerevisiae* fermentations (right).

The analysis of the different peaks across the timepoints enabled the generation of distinct barcodes, representing the presence of *H. uvarum* in natural must fermentation and, for comparative purposes, the single inoculation of *S. cerevisiae* (**Figure IV.B. 10**). Given the high precision of FT-ICR-MS, this graphical representation of *H. uvarum* presence may offer a promising diagnostic tool for identifying this yeast in the fermentation medium. Moreover, if extended to other Non-*Saccharomyces* Yeasts, this approach could also serve as a metabolomic identity card. By tracing the unique footprints of different species, this barcode may enable the

detection of potential contaminants, potentiating early intervention and more refined control over wine quality.

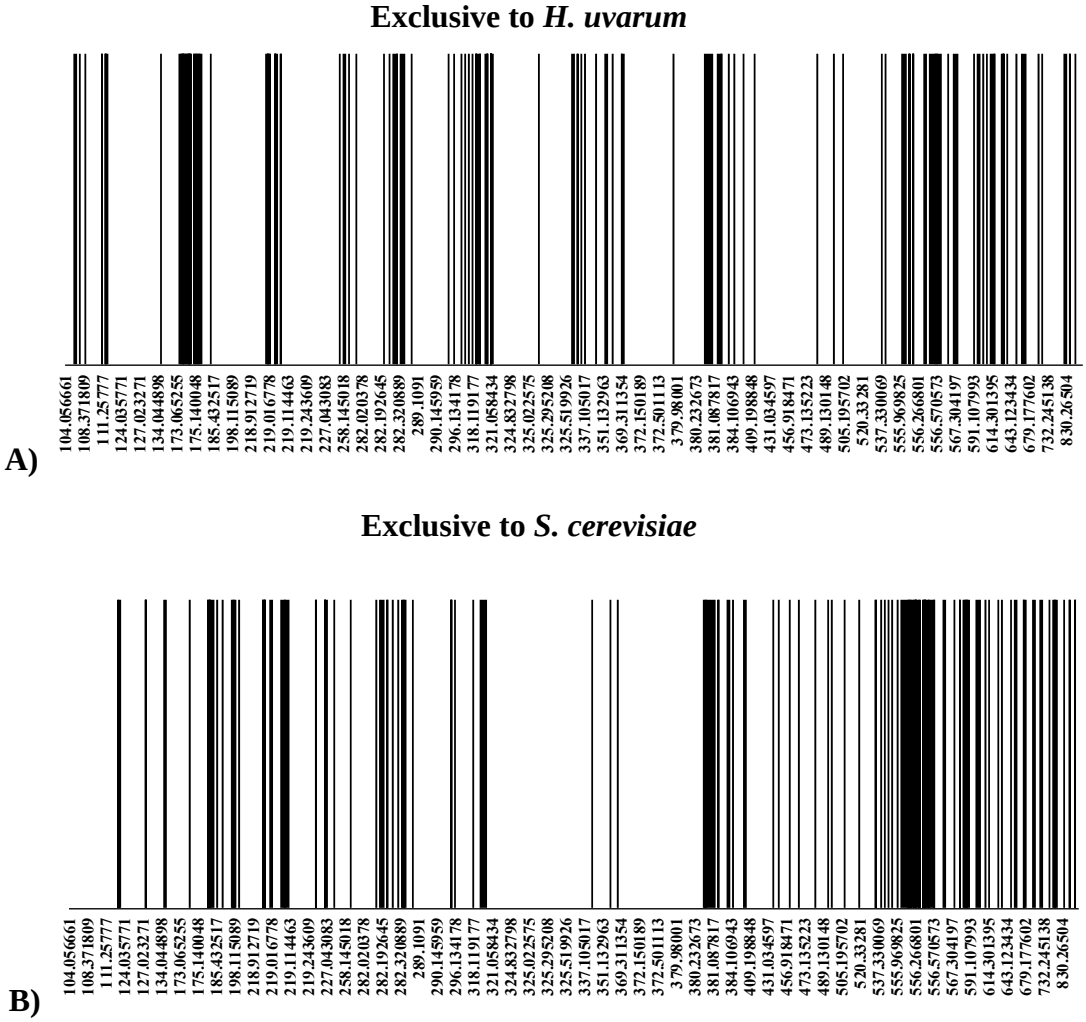


Figure IV.B. 10– Barcodes obtained through metabolomic profiling of the fermentation of natural must with A) *H. uvarum* and *S. cerevisiae* in sequential inoculation; B) single inoculation with *S. cerevisiae*. These barcodes indicate presence (1) or absence (0) of m/z peaks.

The results shown below provide a qualitative perspective of the analysis of the non-volatile metabolome focused on the presence or absence of peaks and, consequently, of metabolites. After that, we focused a quantitative perspective, taking into consideration how normalized peak intensity changed in the different samples. To get the more reliable results, a statistical analysis was performed by comparing peak evolution through time for each inoculation strategy. By comparing the normalized intensity of peaks present in the spectra obtained in the 72 hours and in the initial timepoint of the fermentations carried out with *H. uvarum* MJT198, 309 peaks exhibited decreased intensity, while 389 peaks exhibited increased

intensity, as shown in the volcano plot in **Figure IV.B. 11** and **Table IV.B 1.** Among the peaks with decreased intensity after 72h of fermentation with *H. uvarum* are included several plant-related compounds such as the flavonoid 6-hydroxykaempferol 3,6-diglucoside 7-glucuronide, the kavalactone 5,6-Dihydro-11-methoxyyangonin, the catechin epicatechin-(2 beta->7,4beta->6)-catechin, N-acetylornithine, and the phenylethanoid acteoside (**Table IV.B 1**). Conversely, among the compounds that increased concentration after 72 hours of fermentation by *H. uvarum*, we found the quercetin glucoside quercetin 3-O-glucuronide, glutaminylisoleucine, nicotinic acid, and L-Glutamic acid (**Table IV.B 1**). The same analysis made with the wines obtained in end of fermentation and the 72h sample in the sequential fermentations performed with *H. uvarum* and afterwards with *S. cerevisiae*, we found 277 peaks that decreased intensity (*e.g.*, quercetin 3-O-glucuronide, glutathione, 3'-deaminofusarochromanone, N-hydroxyneosaxitoxin, and trans-o-Coumaric acid 2-glucoside (**Table IV.B 1**) while 476 increased concentrations (*e.g.*, the alkaloid 8-methyldihydrochelerythrine, or the hydroxycoumarin liqcoumarin (**Table IV.B 1**).

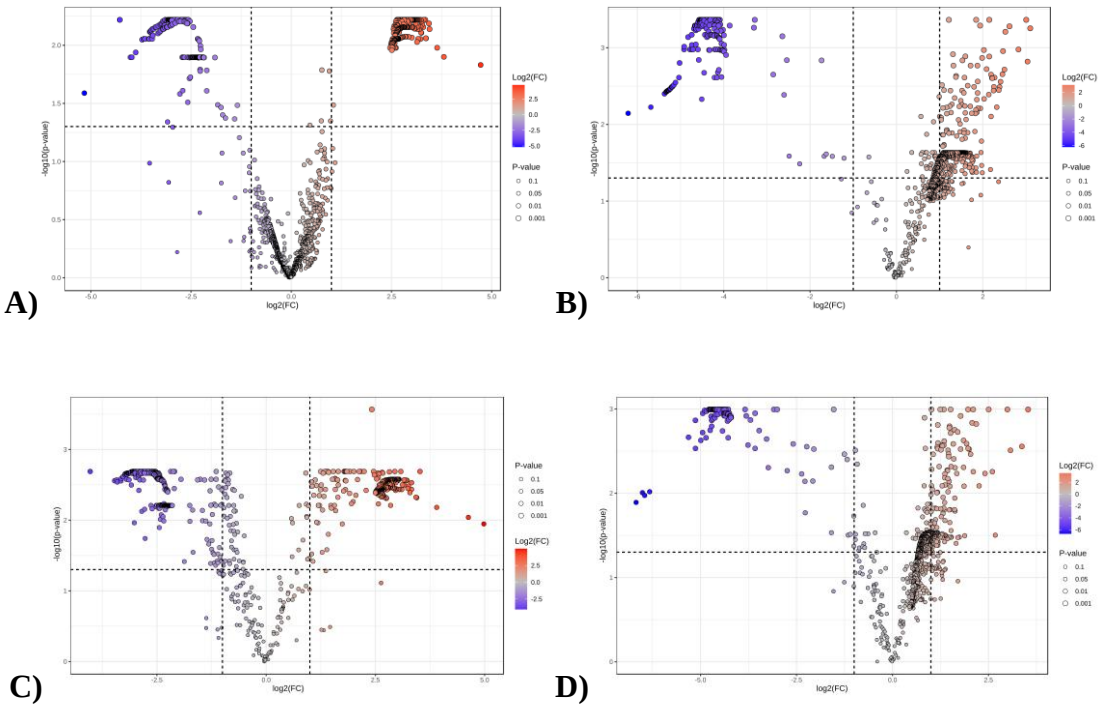


Figure IV.B. 11– Volcano plots indicating the peaks with significant intensities variations in the fermentations with A) *H. uvarum* MJT198 between the timepoint at T_{72h} and T_{0h}; B) *H. uvarum* MJT198 and *S. cerevisiae* QA23 between the timepoint at T_{End} and T_{72h}; C) *S. cerevisiae* QA23 between the timepoint at T_{72h} and T_{0h}; B) *S. cerevisiae* QA23 between the timepoint at T_{End} and T_{72h}. Red represents peaks with increasing intensity, while blue represents decreasing intensity.

When the same analysis was performed with the supernatants obtained from the fermentations carried out with *S. cerevisiae* alone, we found 311 peaks with significantly

higher concentrations after 72h (comparing with the initial time-point) including ergosterol and dimethylglycine (**Figure IV.B. 11** and **Table IV.B 1**). Conversely, 453 exhibited decreased concentrations including linalool glucoside, glutathione and the di-peptide asparaginyl-Histidine (**Figure IV.B. 11** and **Table IV.B 1**). Comparing the wines obtained with *S. cerevisiae* alone with the timepoint at 72 hours, 145 peaks decreased concentration (corresponding, for example, to N-succinyl-2-amino-6-ketopimelate, nicotinic acid, or the benzopyran mollicelin A), while only 14 exhibited a decreased concentration (*e.g.*, melibiose, galactose-beta-1,4-xylose, epicatechin or benzyl chloride (**Figure IV.B. 11** and **Table IV.B 1**). Interestingly, melibiose is a disaccharide formed by an alpha-1,6 linkage between galactose and glucose, is typically formed after the breakdown of raffinose by hydrolysis, which is usually present in plant seeds, leaves, stems and roots, such as those from grapevine (Moreno & Peinado, 2012). The majority of *Saccharomyces* species are not able to metabolize this sugar, including *S. cerevisiae*, which may explain its enhanced presence by the end of fermentation (Tosi et al., 2009).

Table IV.B 1– Peaks with significantly different intensities between timepoints in the sequential and single inoculation.

		Peaks with increasing intensity		Peaks with decreasing intensity	
Inoculation	**Timepoint**	**Number**	**Top examples**	**Number**	**Top examples**
***H. uvarum* MJT198 + *S. cerevisiae* QA23**	T_{72} vs T_0	309	Quercetin 3-O-glucuronide [4,5-dihydroxy-2-(5-hydroxy-7-methoxy-4-oxo-2-phenyl-4H-chromen-8-yl)oxan-3-yl]oxidanesulfonic acid Glutaminylisoleucine Nicotinic acid Carbinoxamine Demethylated antipyrine L-Glutamic acid	389	6-Hydroxykaempferol 3,6-diglucoside 7-glucuronide 5,6-Dihydro-11-methoxyyangonin Epicatechin-(2beta->7,4beta->6)-catechin N-Acetylornithine Acteoside
	T_{end} vs T_{72}	476	8-Methyldihydrochelerythrine 5-Methoxydimethyltryptamine 3,4,5-trihydroxy-6-{[7-hydroxy-4-oxo-2-phenyl-8-(3,4,5-trihydroxyoxan-2-yl)-4H-chromen-5-yl]oxy}oxane-2-carboxylic acid Trimethaphan (S)-Multifidol 2-[apiosyl-(1->6)-glucoside] Liqcoumarin	277	Quercetin 3-O-glucuronide Glutathione 3'-Deaminofusarochromanone N'-Hydroxyneosaxitoxin trans-o-Coumaric acid 2-glucoside
***S. cerevisiae* QA23**	T_{72} vs T_0	311	Nicotinic acid Ergosterol 3,4,5-trihydroxy-6-{[(2E)-2-methyl-3-phenylprop-2-en-1-yl]oxy}oxane-2-carboxylic acid Dimethylglycine 6-Phenyl-3-hexen-2-one	453	Linalool oxide D 3-[apiosyl-(1->6)-glucoside] Asparaginyl-Histidine 2-O-alpha-D-Galactopyranosyl-1-deoxynojirimycin Glutathione Phlorin
	T_{end} vs T_{72}	14	Benzyl chloride 4,8-dihydroxy-2H-furo[2,3-h]chromen-2-one Galactose-beta-1,4-xylose Epicatechin Melibiose	145	N-Succinyl-2-amino-6-ketopimelate Dinoseb acetate Nicotinic acid [4,5-dihydroxy-2-(5-hydroxy-7-methoxy-4-oxo-2-phenyl-4H-chromen-8-yl)oxan-3-yl]oxidanesulfonic acid Mollicellin A

The overall comparison of peak intensity trend with time between the two inoculation strategies is depicted in **Figure IV.B. 12**. A total 309 peaks exhibited a significant intensity decrease between the end of fermentation and the initial timepoint in the sequential inoculation with *H. uvarum* MJT198 and *S. cerevisiae* QA23, compared to 1,013 peaks in the single inoculation with *S. cerevisiae* (**Figure IV.B. 12)**. Of these, 250 peaks showed the same decreasing trend in both inoculation strategies, while 763 decreased exclusively in the *S. cerevisiae* inoculations, and 59 exclusively in the sequential inoculation (**Figure IV.B. 12)**. Conversely, 605 peaks significantly increased in intensity over time in the sequential inoculation, while 840 peaks increased in the single *S. cerevisiae* inoculation. Among these, 522 peaks were common to both fermentation strategies, with 318 unique to *S. cerevisiae* single inoculation and 83 exclusive to the sequential inoculations. Out of the peaks with contradicting tendencies between inoculation strategies, it is possible to identify Limonexic acid, whose intensity increases in *S. cerevisiae* but decreases in the sequential inoculation; and the compounds 8-Hydroxypinoresinol 8-glucoside and 5-Methoxydimethyltryptamine, whose intensity increases in the sequential inoculation but decreases in the *S. cerevisiae* fermentations (**Table IV.B 2).**

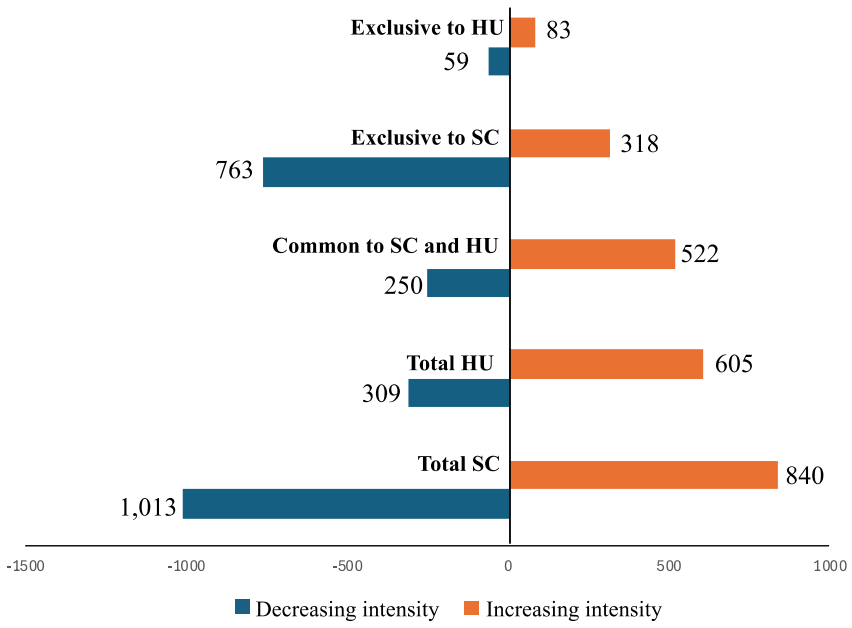


Figure IV.B. 12– Overall comparison of the evolution of peaks between the end and the beginning of fermentation, being represented the peaks exclusive to the sequential inoculation with *H. uvarum* (HU), to the single inoculation with *S. cerevisiae* (SC), and the total peaks analysed.

Table IV.B 2– Peak intensities with contradicting tendencies in both inoculation strategies, when comparing the end of fermentation with the beginning, and their putative identification.

INCREASING IN SC AND DECREASING IN HU_SC

m/z	Name	Fold-change > |2|	
		SC	HU_SC
314.108359	Limonexic acid	1.5302	-5.1845
198.097393	n.a	1.3841	-4.081
219.087947	n.a	1.4679	-4.2437
219.101846	n.a	1.5014	-4.6365
218.977629	n.a	1.4747	-4.1786
218.951419	n.a	1.5864	-4.5599
219.101442	Metanephrine	1.4695	-4.5179
218.964213	N-Hexadecanoylpyrrolidine	1.3693	-4.4529
219.024983	n.a	1.2105	-5.7235
219.002109	n.a	1.3845	-3.8238

INCREASING IN HU_SC AND DECREASING IN SC

m/z	Name	Fold-change > |2|	
		SC	HU_SC
175.140367	n.a	-4.058	1.087
381.115404	n.a	-4.0097	2.1224
175.140048	n.a	-3.9642	1.2178
175.090787	n.a	-3.9132	1.0098
175.148235	n.a	-3.9357	1.011
320.259271	n.a	-4.3495	1.1835
265.016061	n.a	-3.7314	1.1069
550.747742	n.a	-3.6981	1.5063
317.182144	8-Hydroxypinoresinol 8-glucoside	-3.6931	1.4003
261.144692	5-Methoxydimethyltryptamine	-3.6809	1.6748

IV.B.4.2 Transcriptomics analysis of the *H. uvarum*-*S. cerevisiae* co-culture during fermentation during fermentation of Muscat grape must

To provide further information about the interactions established between *H. uvarum* and *S. cerevisiae* during the fermentation of Moscatel Galego must, cells were collected along the different time-points, as detailed in **Figure IV.B. 3**. From a molecular perspective, interactions established in yeast consortia used as starter cultures have been studied from the perspective of *S. cerevisiae*, while less has been described concerning how the co-cultivation affects genomic expression of the Non-*Saccharomyces* Yeasts used, specially using natural grape musts. Gene expression was monitored at three timepoints that were analyzed: fermentation onset, after 72 hours of inoculation and by the end of fermentation. After RNA-sequencing an average of 15.1M reads, per replica, per condition were obtained, as detailed in **Table IV.B 3.** These reads were mapped to the genome of *H. uvarum* MJT198 resulting in alignment percentages above 70% from the samples obtained at 0h and at 72h which this yeast was present alone (**Table IV.B 3**). It is possible that the remaining 30% of RNA sequenced may correspond to RNA from other sources, in particular, originating from the grape cells. In the case of the samples obtained at the end of the fermentation, where *S. cerevisiae* was already present, the percentage of unique mapping to the genome of *H. uvarum* was lower, ranging from 37% to 50% (**Table IV.B 3**). The reduction in the number of reads mapping to the genome of *H. uvarum* in the end of the co-inoculated fermentation was expected on one side due to the presence of *S. cerevisiae* cells and also from the reduced viability of *H. uvarum* to the ethanol present in the broth that may necessarily affect mRNA production.

Table IV.B 3-Number of reads per replicate and percentage mapped to the genome of *H. uvarum* MJT198.

Samples	Number of reads	%Reads mapped to *H. uvarum* MJT198
F1_t_{0h}	17,946,080	70.99
F2_ t_{0h}	14,029,779	75.38
F1_ t_{72h}	16,815,924	75.47
F2_t_{72h}	15,877,663	76.78
F1_t_{end}	11,762,643	35.41
F2_t_{end}	13,638,204	47.89

To confirm whether the reads that mapped to the genome of *H. uvarum* corresponded to genomic regions specifically assigned to this genome (not matching to *S. cerevisiae*) we mapped them to the genome of *S. cerevisiae* QA23 as well. The results obtained showed that

94.34% of these reads also align with *S. cerevisiae* genome likely corresponding to genes whose sequences are identical in between the two yeasts. The reads that could be specifically mapped to the *H. uvarum* were of only circa 5%, which would translate in a very low vertical coverage, preventing us from an accurate quantification of genomic expression. Therefore, it was decided to proceed with the analysis of results obtained only when *H. uvarum* was growing alone in the Moscatel Galego grape must (that is, the 72h samples) (**Appendix Table IV.B.4**). Although this is not as informative as we had designed the experiment to be, it still is relevant since it is the first OMICS analysis performed for this species when fermenting a natural grape must.

After 72h of fermentation of the Moscatel Galego grape must, we found that the expression of 373 *H. uvarum* genes was reduced, compared to the expression levels they exhibited immediately after inoculation in this must. In contrast, 160 genes showed an increased expression (**Figure IV.B. 13**). A closer analysis of the data revealed 25 genes that were only expressed at 72 hours but not at the start of fermentation, and conversely, 180 genes were expressed only at the beginning but not after 72 hours. The *H. uvarum* genes that we found to be expressed only after 72 hours of must fermentation were orthologues of the ribosomal genes *RRP1*, *RL16A*, and *MRT4*, alongside genes encoding mitochondrial proteins like Mix23, Idh2, and Syim. On the other hand, the set of genes expressed only at the fermentation onset but not at 72 hours included also ribosomal proteins Rm22, Rm02, and Rs23, a sulfiredoxin (Srx1), and several transporters (*e.g.*, Vacuolar transporter, *VTC2;* probable transporter *MCH4*, biotin transporter; carboxylic acid transporter). The full list of genes uniquely expressed or whose expression was no longer detected at 72 hours can be found in **Appendix Table IV.B.4**, while the complete list of upregulated and downregulated genes can be found in **Appendix Table IV.B.5.**

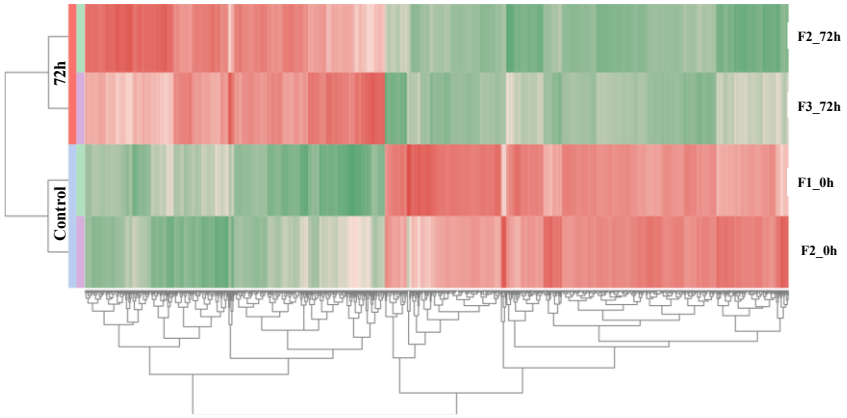
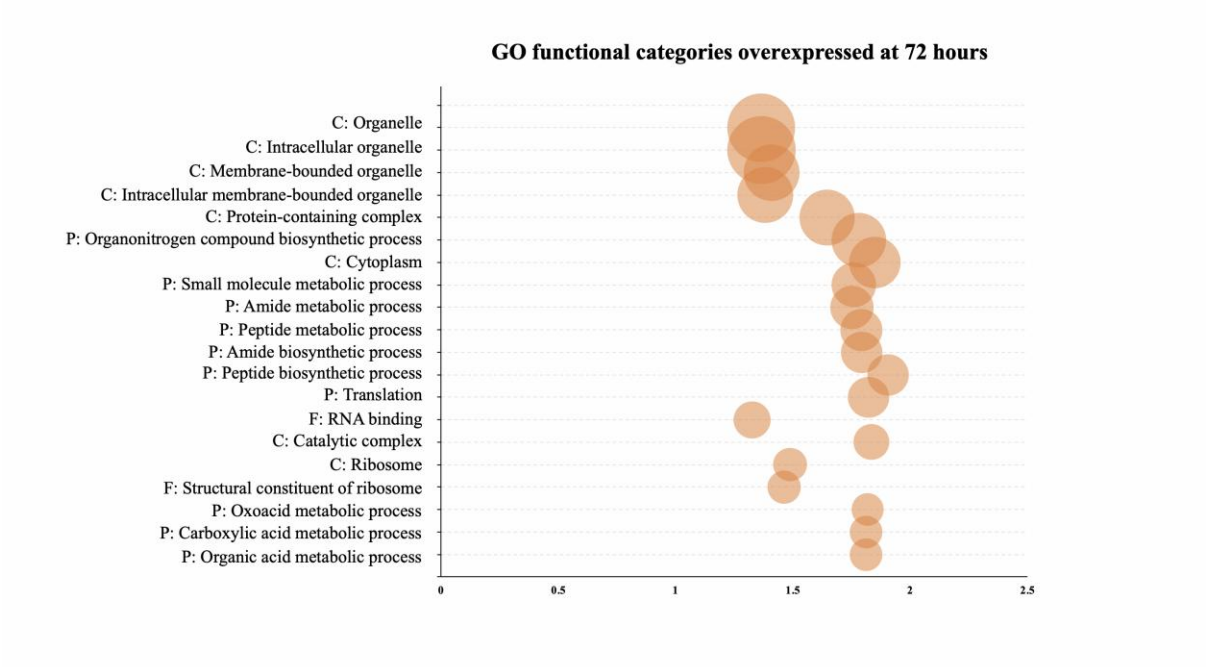


Figure IV.B. 13– Heatmap demonstrating the *H. uvarum* MJT198 genes that were found to be up-regulated (represented in red) and down-regulated (represented in green) at 72h after the beginning of fermentation, when comparing to the initial fermentation timepoint. In total, 160 genes were found to be up-regulated, while 373 were down-regulated.

To uncover the biological significance of the differentially expressed genes identified in this study, Gene Ontology (GO) Enrichment Analysis and Gene Set Enrichment Analysis (GSEA) were employed. These tools allowed us to predict the major gene pathways significantly altered 72 hours after the onset of grape must fermentation with *H. uvarum* MJT198 (**Figure IV.B. 14**). GO enrichment and GSEA facilitated the classification of the overexpressed genes into 112 GO categories. Among these, 60 belonged to Biological Processes (P), 34 to Cellular Component (C), and 18 to Molecular Function (F). In contrast, the under-expressed genes were organized into only 16 GO categories, with 8 assigned to Biological Processes, 7 to Molecular Function, and just 1 to Cellular Component (**Figure IV.B. *15***). This discrepancy in GO categorization, despite having a higher number of under-expressed genes, may suggest that the overexpressed genes are more functionally diverse and involved in a wider range of biological pathways.

Figure IV.B. 14 illustrates the size of the gene sets associated with each GO functional category, while **Figure IV.B. 15** presents the distribution of the most significant GO functional categories, ranked by their Normalized Enrichment Score (NES). The NES accounts for the gene set size (i.e., the number of genes assigned with each category), allowing for meaningful comparisons between categories. The complete distribution of GO categories can be found in **Appendix Table IV.B.6**. From **Figure IV.B. 14** it is possible to observe that the largest gene set among the overexpressed genes are those related to organelle compounds. In contrast, for downregulated genes, the largest sets are associated with DNA binding and DNA metabolic processes.

A)



B)

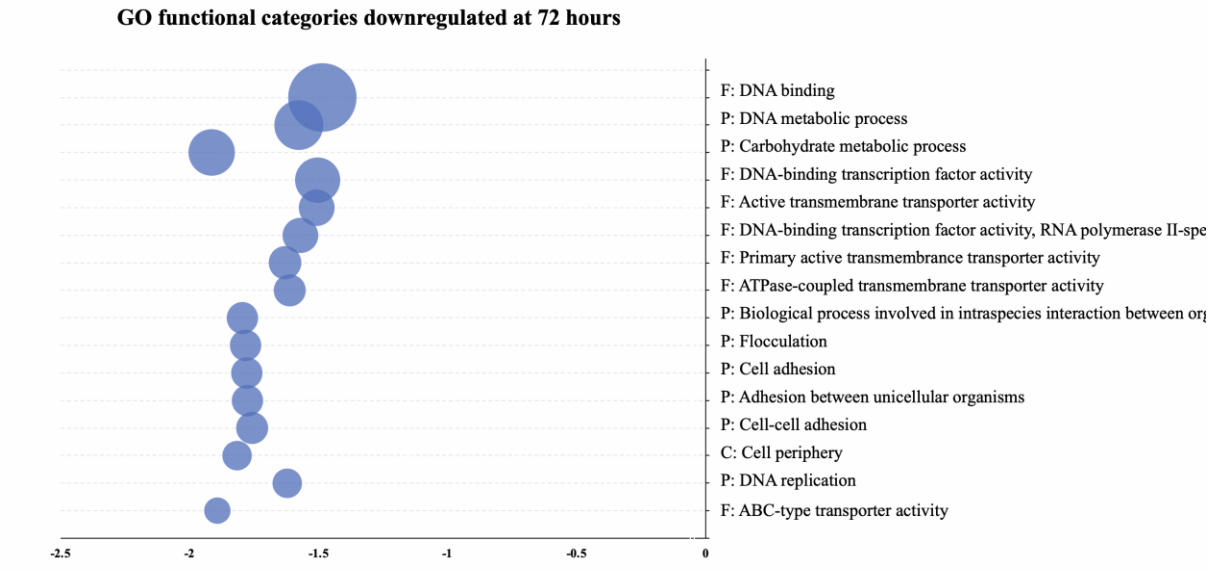


Figure IV.B. 14 – GO functional categories distribution according to the gene set size that originated them. A) GO categories upregulated at 72 hours and B) GO categories downregulated at 72 hours.

However, after data normalization, we observe that the most significant categories of downregulated genes at 72 hours include those related to carbohydrate metabolic processes,

ABC-type transporter activity, as well as those related to flocculation, cell adherence and cell-cell interaction (**Figure IV.B. 15**). At this timepoint, the abundance of carbohydrates sources was substantial (0.4 % ethanol produced, and 137 g/L of total sugars still available– see **Chapter IV.A**). This high sugar availability likely reduced the need for enhanced carbohydrate metabolic machinery, compared to the cells that had been immediately inoculated in the must and that came from a pre-inoculum where sugar has probably been exhausted. At the same time, the downregulation of genes associated with flocculation, cell-adherence, and cell-cell interactions supports this theory. *H. uvarum* cells at T_{0h} appeared to be in a more stressful state than at 72 hours, characterized by a higher biomass degree (from the pre-inoculum), since typically flocculation is induced as a response to adverse environmental conditions, such as nutrient scarcity (Soares, 2011). Regarding the categories of upregulated genes, it was noticeable the high number of categories related to RNA metabolism (ncRNA, tRNA), which may reflect an extended adaptive response of *H. uvarum* MJT198 at the level of regulating gene and genomic expression. Additionally, several categories related to amino acid metabolism are also overrepresented in this analysis, which likely reflects the specific preference of amino acids as nitrogen source associated to *Hanseniaspora* (Seixas et al., 2023).

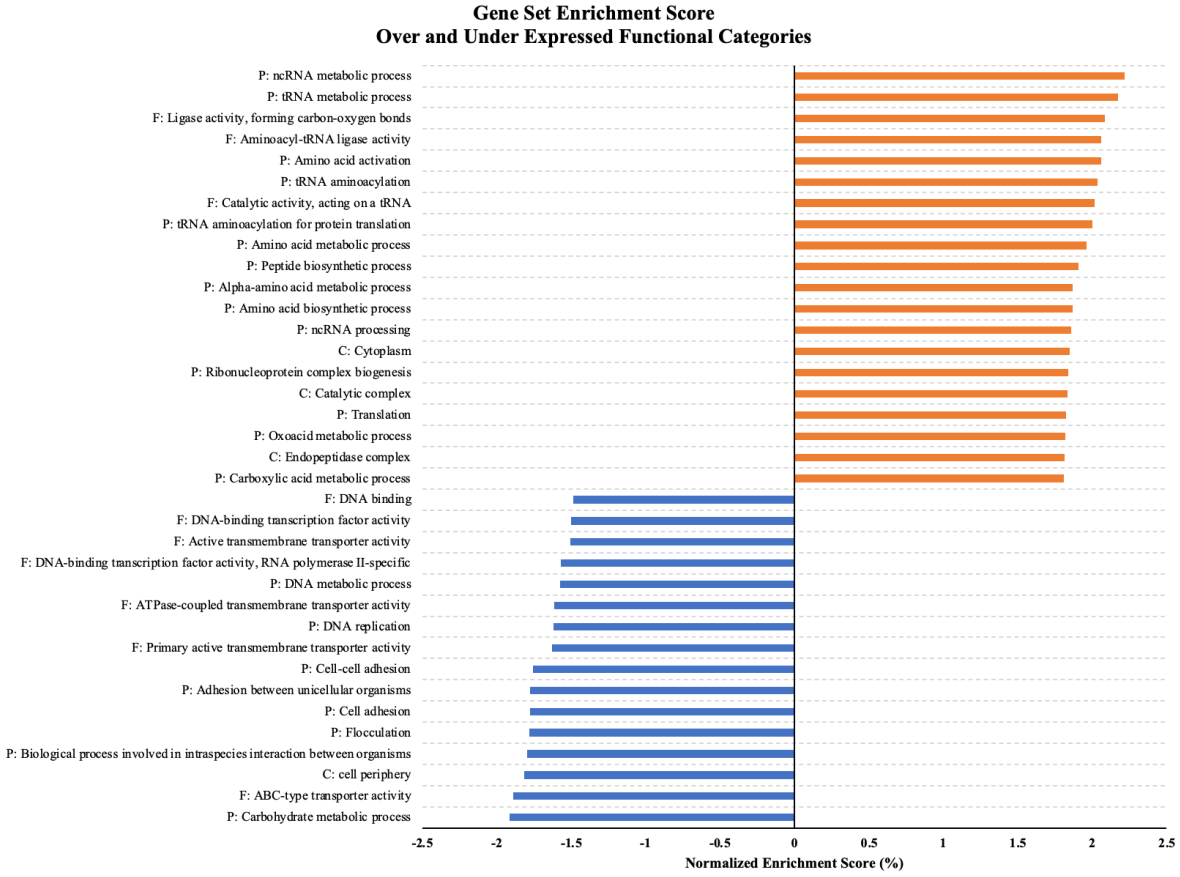


Figure IV.B. 15– Gene Set Enrichment Score Analysis (GSEA) with the most significant functional categories over and underexpressed at 72 hours.

Alterations in expression of genes involved in formation of aroma compounds.

Considering that one of the major goals of this analysis was the establishment of relevant links between the presence of *H. uvarum* and changes in the aroma profile, we have examined in closer detail the changes in expression of those genes encoding enzymes with functions related with the formation of such molecules. We started by the 4 genes encoding predicted alcohol acetyl transferases (AAT): HU_g1549.t1, HU_g2292.t1, HU_g2974.t1 identified based on their homology with the recently characterized orthologues in *H. guilliermondii* and in *H. uvarum* AWRI3580. Although they were all found to expressed, their transcription either identical or slightly diminished after 72h of fermentation in the Moscatel Galego grape must, comparing to the initial values (**Table IV.B 4).** The regulatory mechanisms that govern expression of these genes in *H. uvarum* had not been established, although it has recently been shown that in *H. guilliermondii* their expression is conditions where nitrogen is more abundant than carbon (low C:N ratio) (Seixas et al., 2023). Compared to the initial timepoint, where cells derived from a pre-inoculum, that likely faced some level of carbon source exhaustion (leaving nitrogen in excess), the C:N ratio at 72 hours is expected to be higher. If *H. uvarum's* genes are regulated similarly to *H. guilliermondii's*, this means that at 72 hours, the expression of the genes coding for alcohol acetyltransferase will be lower than at T_{0h}, which, contrasting with gene Hu_g2292.t1, is the tendency observed. Therefore, similar to the study conducted on *H. guilliermondii* (Seixas et al., 2019), it will be valuable in future research to investigate these genes further, examining their regulatory differences and potential impacts on acetate ester formation, under varying C:N ratios.

Table IV.B 4– Genes coding for alcohol acetyltransferases in *H. uvarum* MJT198 and their expression at 72 hours.

GENE	FOLD-CHANGE	LOG_2FC	EXPRESSION
Hu_g1549.t1	0.76	-0.40	decreased
Hu_g2292.t1	1.09	0.13	increased
Hu_g2974.t1	0.52	-0.94	decreased

At 72 hours, the concentration of the higher alcohols methionol and isobutanol was significantly increased in the musts fermented by *H. uvarum* MJT198, comparing to those fermented with *S. cerevisiae*. The expression of genes encoding enzymes of the Ehrlich pathway (*Hu_ADH1*, coding for alcohol dehydrogenases, and the 4 copies of *Hu_ARO9*, encoding the aromatic amino acid aminotransferase responsible for the initial step of the

pathway), was lower in the fermenting *H. uvarum* cells than in the cells that were immediately inoculated in the must (*ADH1* by 4.5-fold (Hu_g4911.t1) and 28 (Hu_g603.t1); and the *ARO9* (Hu_g3487.t1, Hu_g3702.t1, Hu_g1212.t1) by 4,8,9 and 19 fold, respectively) (**Appendix Table IV.B.5**). This discrepancy in gene expression contrasts with the increased concentration of the higher alcohols observed, however, it is important to denote that besides gene expression, the production of higher alcohols is governed by other factors such as availability of precursors and of carbon sources. Notably, the deletion of *ARO8* and *ARO9* in *S. cerevisiae* has been demonstrated to result in increased methionol levels (Deed et al., 2019), which has been explained by the activity of additional transaminases that may compensate for the lack of these genes, or an alternative pathway that may convert methanethiol to methionol.

Turning the attention to minor volatiles, a significant prevalence of the monoterpenes linalool, geraniol, and nerol was observed in the samples obtained after 72h of fermentation with *H. uvarum*, a trend that persisted until the end of fermentation. Interestingly the two genes encoding beta-glucosidase in *H. uvarum* MJT198 (Hu_g317.t1 and Hu_g389.t1) exhibited divergent transcriptional behaviors. While the gene corresponding to *Hu_BGLU1* (Hu_g389.t1) displayed a negative fold-change of 6.49, *Hu_BGLU2* (Hu_g317.t1) was transcribed with a positive fold-change of 2.97 (**Appendix Table IV.B.5**). It will be interesting to further examine what could be the effect of these changes in expression in the release of the terpenic fractions present in the sugar

Chapter V.

Final Discussion

Overview

Wine production has traditionally relied on *Saccharomyces cerevisiae* as the primary yeast species responsible for alcoholic fermentation (Padilla et al., 2016; Fleet, 2008; Barnett, 2007). While this yeast is effective in ensuring consistent fermentation, its dominance and consequent limited diversity of yeast species used in winemaking, limits the stylistic variety in the final product. In recent years, however, Non-*Saccharomyces* Yeasts (NSYs) have emerged in the winemaking market as a promising solution to diversify wine profiles. By their unique enzymatic activities, these yeasts can enhance the complexity of wine by generating a broader array of aromas and flavors (Gschaedler, 2017; Capece et al., 2018; Capece & Romano, 2019) (**Figure V. 1**). The introduction of novel metabolic pathways into winemaking has the potential to enrich the organoleptic characteristics of wines, creating more diverse and distinctive styles. However, their use remains limited due to factors such as their slower fermentation rates, low tolerance to ethanol and different susceptibility levels to preservatives like sulfur dioxide, which makes their application challenging in large-scale winemaking.

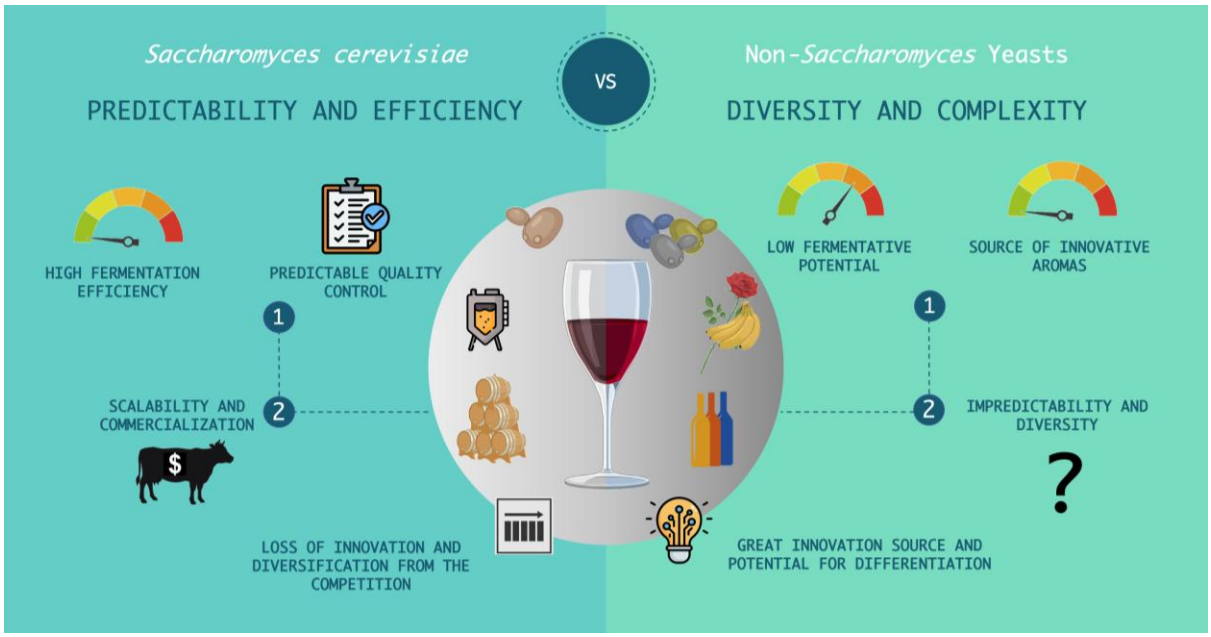


Figure V. 1– *S. cerevisiae* vs Non-*Saccharomyces* Yeasts as winemaking products. While *S. cerevisiae* has long been the traditional choice in winemaking due to its high fermentative efficiency and predictability, the future of winemaking may lie on NSYs, which offer the potential for greater aroma differentiation and complexity.

This thesis focused on the exploration of two NSYs from the *Saccharomycodaceae* family, which, despite their phylogenetic proximity, are associated with distinct roles in winemaking. On one hand, *Saccharomycodes ludwigii* is often regarded as a spoilage yeast, due to its high resistance to sulfur dioxide, making it challenging to eliminate from wineries

and their equipment (Vejarano, 2018; Pilap et al., 2022; Romano et al., 1999; Tavares et al., 2021; Startford et al., 1987; Esteves et al., 2019; Jackowski et al., 2023). Its presence is usually connected to cloudiness and sedimentation in bottled wines, significantly compromising the commercial viability of wine (Vejarano, 2018; Pilap et al., 2022; Romano et al., 1999; Tavares et al., 2021; Startford et al., 1987; Esteves et al., 2019; Jackowski et al., 2023) (**Figure V. 2**). On the other hand, *Hanseniaspora uvarum* is recognised for its potential to enhance wine aroma, mostly through its high enzymatic activities, particularly beta-glucosidase, which increases the concentration of free volatiles compounds, such as monoterpenes and norisoprenoids that are key to wine's primary aroma (Martin et al., 2018; Moreira et al., 2011; van Wyk et al., 2023; Moreira et al., 2008; Wang et al., 2015; Kleman et al., 2022; Albertin et al., 2016; Hu et al., 2019; Hu et al., 2018; Zhang et al., 2023; Guaragnella et al., 2020; Tristezza et al., 2016; Du Plessis et al., 2019; Langenberg et al., 2017). In addition, *H. uvarum* is also associated to the increased production of acetate esters, such as isoamyl acetate and 2-phenylethyl acetate (Moreira et al., 2008; Gao et al., 2022), contributing to more pronounced fruity flavors, likely driven by distinct acetyltransferase enzymes that differ from those of *S. cerevisiae* (Martin et al., 2018; Moreira et al., 2011; van Wyk et al., 2023; Moreira et al., 2008; Wang et al., 2015; Kleman et al., 2022; Albertin et al., 2016; Hu et al., 2019; Zhang et al., 2023; Guaragnella et al., 2020; Seixas et al., 2023). However, the excessive production of ethyl acetate by *Hanseniaspora* species remains a concern (Hu et al., 2018), as it can lead to undesirable aromas in wines (when present in concentrations above 150-200 mg/L, ethyl acetate can lead to solvent-like or nail polish aromas - Li et al., 2020) (**Figure V. 2**). While over the years, research has been more focused on *H. uvarum* than in *Sd. ludwigii*, there are still considerable gaps in the genomic and physiological characteristics of both species. Hence, this work sought to address these gaps by exploring the genomic mechanisms underlying their dual roles in winemaking, and to uncover specific traits that could be harnessed for future applications.

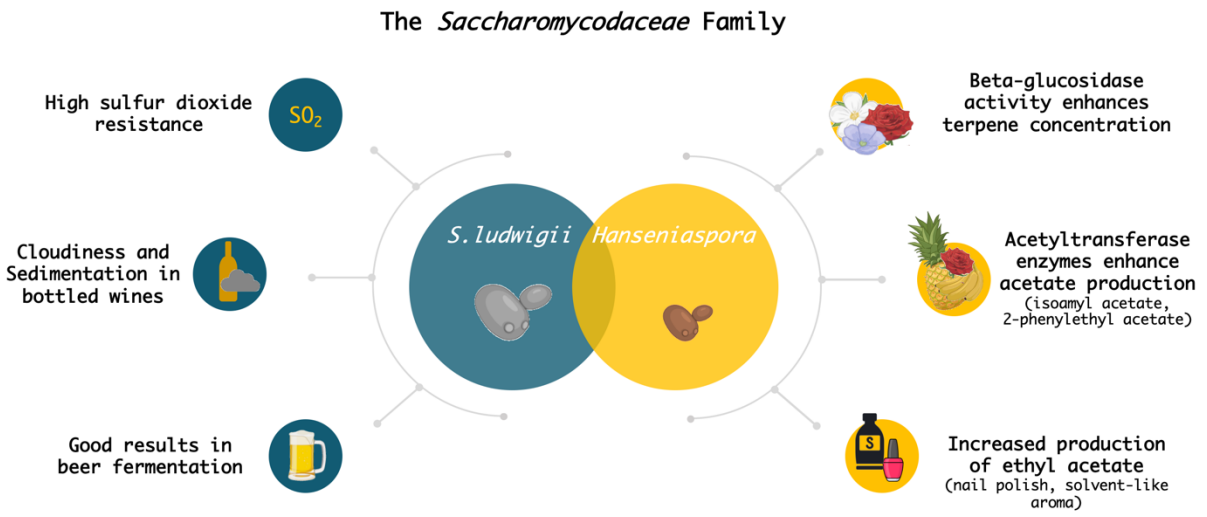


Figure V. 2– *Saccharomycodes* vs *Hanseniaspora*: existing information in the literature places *Sd. ludwigii* as a spoilage yeast in winemaking, yet, positive outcomes have been registered in the fermentation of other beverages, such as beer. On another hand, the *Hanseniaspora* genus has been regarded as having potential as bioflavourant in winemaking, due to its powerful enzymatic activity. Nonetheless, the usually high ethyl acetate concentrations associated to *Hanseniaspora* limit the application of these yeasts.

Insights from *Sd. ludwigii* UTAD17

Beginning with the exploration of the lesser-studied species, *Saccharomycodes ludwigii*, we aimed to uncover the primary factors that contributed to its reputation as the "winery nightmare". Was *Sd. ludwigii* truly the villain, or was it merely misunderstood? Could there be hidden potential within this yeast for bio-flavouring, or was it simply a problem that should be avoided at all costs? And perhaps more importantly, was strain specificity a key factor that could lead to different outcomes in winemaking? To address these questions, we undertook a comprehensive genomic analysis of the wine strain *Sd. ludwigii* UTAD17 (**Chapter II**). Not only this strain possessed four genes encoding sulfite efflux pumps, that may have granted it a significant advantage in resisting sulfur dioxide, but its genome was also equipped with intriguing genes that provided important clues for its high endurance during wine fermentation[13]. For instance, the existence of a set of beta-mannosyltransferase enzymes that could only find orthologues in very distantly related species, such as *Candida albicans* or *Pichia pastoris*, hinted for a unique cell wall mannosylation (Tavares et al., 2021; Mille et al., 2008; Krainer et al., 2013). In *Candida* species, this cell wall structure is typically linked to increased pathogenicity, improving their adherence to host cells (Mille et al., 2008). Yet, *Sd. ludwigii* is not known to be pathogenic in humans. So why would this yeast benefit from such a high mannan content in its cell wall? Previous studies on *Sd. ludwigii* had already demonstrated its unusually high mannan content (Spencer & Gorin, 1968), which could

significantly reduce the diffusion of free sulfite into the cells, enhancing its resistance to this preservative. Another possibility lies in its common isolation niche, the insect gut (Fogleman et al., 1982; Stefanini, 2018). A more complex cell wall may improve its ability to colonize this niche, ensuring its survival through better adhesion to the host and dispersion. Aligning with this, the genome of *Sd. ludwigii* UTAD17 disclosed the genes involved in the catabolism of N-acetylglucosamine (GlcNAc) (Tavares et al., 2021). Being the monomer of chitin, *Sd. ludwigii's* capacity to metabolize GlcNAc can be a further hint into its ability to colonize insects. In the fermentation environment, this trait could also offer a competitive advantage by allowing the yeast to use an alternative carbon source, clearly distinguishing it from other yeast species that solely rely on traditional fermentable sugars. In addition, another interesting finding in the genome of UTAD17 was the presence of multiple copies of genes involved in the thiamine biosynthetic pathway. Unlikely closely related yeast species, which lack this gene set, *Sd. ludwigii's* ability to synthesize thiamine likely enhances its fermentative capacity (thiamine is the co-factor of several key enzymes), further differentiating it from the *Hanseniaspora* genus. This suggests that *Sd. ludwigii* not only thrives in challenging environments (such as high ethanol and SO_2 concentrations) but also possesses unique genetic adaptations that could be harnessed for fermentation processes (**Figure V. 3**).

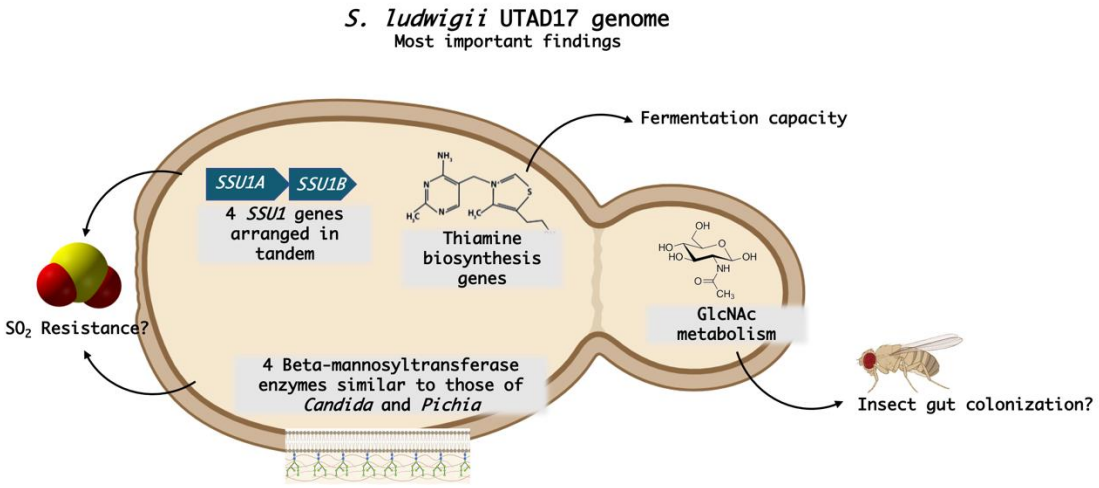


Figure V. 3– Most important findings of the analysis of the *Sd. ludwigii* UTAD17 genome. Among the most distinctive features are the existence of four sulfite efflux pumps genes (SSU1), genes responsible for thiamine biosynthesis, four beta-mannosyltransferase enzymes, and genes responsible for metabolizing N-acetylglucosamine (GlcNAc)

Sd. ludwigii vs *Sd. ludwigii* – strain specificity in *Saccharomycodes*

While the genomic exploration of *Sd. ludwigii* UTAD17 provided valuable insights into the unique characteristics of this wine strain, the lack of additional genomes for comparison, at the time, limited our ability to determine whether these traits were common across other *Sd. ludwigii* strains or specific to UTAD17. Hence, the opportunity to explore another strain, *Sd. ludwigii* BJK_5C, isolated from an apple cider vinegar facility, allowed us to deeply investigate strain-specific variability within this species (**Chapter III**). The obtained results showed that strain BJK_5C exhibited two additional chromosomal bands and increased ploidy when compared to UTAD17 (which displayed the expected seven bands as reported in the literature - Yamazaki & Oshima, 1996), leading us to explore the complexities of *Sd. ludwigii* reproduction. In yeasts, polyploidy is often linked to greater genomic complexity, which in turn enhances tolerance to stress and promotes evolutionary adaptation (Selmecki et al., 2015; Mozzachiodi et al., 2022; Byrne et al., 2005). This was likely the case for BJK_5C, which was subjected to the stressful conditions of a vinegar production industrial facility.

These differences in the karyotypes were also reflected in the phenotypic traits of the strains (**Figure V. 4**). Indeed, while BJK_5C demonstrated growth on a variety of carbon sources and resistance in the presence of high acetic acid concentrations, it was surprisingly vulnerable to sulfur dioxide. Conversely, UTAD17 thrived in sulfur dioxide but displayed less flexibility in utilizing different carbon sources. The genomic differences between these two strains further supported these phenotypic observations (**Figure V. 4**). In fact, BJK_5C harboured half of the copies of sulfite efflux pumps found in UTAD17 and fewer copies of the thiamine biosynthesis genes, probably affecting its overall fermentation performance. This hypothesis was confirmed through fermentation experiments, where both strains performed similarly on the fermentation of a red grape variety, but BJK_5C showed a significant delay in white wine fermentation, possibly due to its increased susceptibility to the lower temperatures typically used in white wine production. Going deeper into the analysis, we found that strain specificity in *Sd. ludwigii* also extended to aroma production, with the wines fermented by each strain exhibiting significantly different aroma profiles. Indeed, BJK_5C produced three times more ethyl acetate than UTAD17, resulting in ethyl acetate concentrations that were considered defective. In contrast, UTAD17 modulated both major and minor aroma volatiles in a way that enhanced the expression of the grape variety, making it more suitable for varietal wine production, as previously reported (Esteves et al., 2019). These findings highlight the importance of understanding strain-specificity within *Sd. ludwigii*, and also of the epigenetic

factors modulating gene expression, as these peculiarities can greatly influence winemaking outcomes. However, the potential of these yeasts as bio-flavoring agents is highly dependent on the fermentation matrix (since different precursors result in distinct aroma profiles), and also on the type of beverage being produced, each demanding different flavor compounds. For instance, *Sd. ludwigii* has gained interest in beer production due to the unique specificities of the brewing process, illustrating how a species considered problematic in one context may offer advantages in another (Romano et al., 1999; Jackowski et al., 2023; De Francesco et al., 2015; Sileoni et al., 2023). Additionally, the complex enzymatic activity of NSYs and the increasing knowledge on this matter, opens opportunities for their exploration in the production of enzymes with significant biotechnological potential.

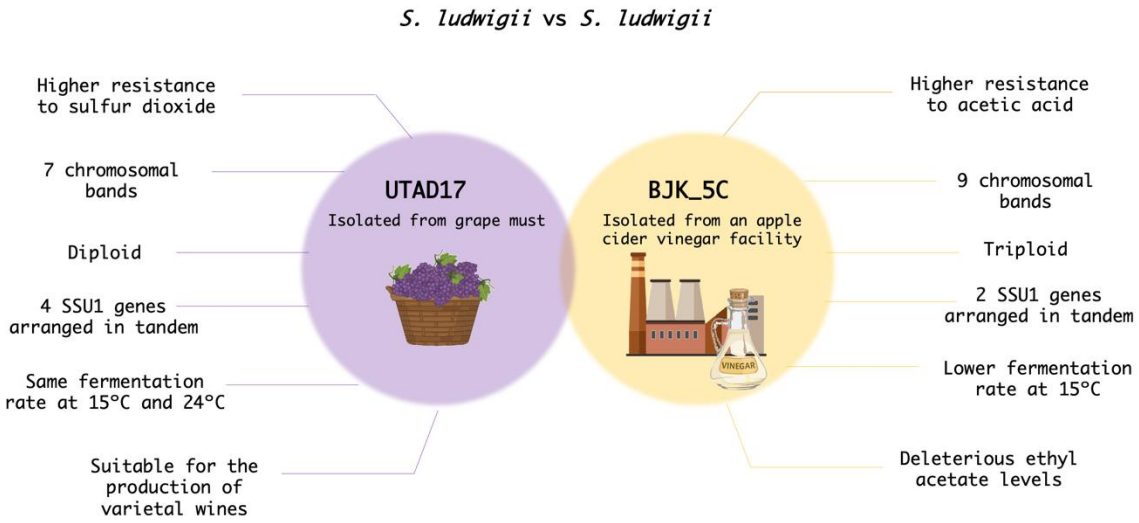


Figure V. 4– Main findings obtained from the comparison of *Sd. ludwigii* strains UTAD17 and BJK_5C, and their correlation to their isolation environment. Since UTAD17 was isolated from the wine environment, it showed an increased resistance to sulfur dioxide, a higher number of sulfite-efflux pump genes, and a higher fermentation rate at different temperatures. Conversely, strain BJK_5C, isolated from an apple cider vinegar facility, showed higher resistance to acetic acid, but less resistance to sulfur dioxide and less copies of the sulfite efflux pump genes. It also showed decreased fermentation capacity at lower temperatures.

Beta-glucosidase in the *Hanseniaspora* and their role in wine primary aroma

In **Chapter VI.A** we embarked on the search for beta-glucosidase-producing yeasts, focusing on strains isolated from Moscatel Galego must, which is known for its high concentration of natural glucosides. Beta-glucosidases play a critical role in wine aroma enhancement as they release aromatic volatiles that are bound as glycosylated compounds, such as terpenes and norisoprenoids, which are present in grape must but often unavailable in terms of aroma (Rosi et al., 1994; Zhang et al., 2021). Unsurprisingly, *H. uvarum* natural isolates emerged as the most promising candidates for beta-glucosidase production. Although some research has been done with this species, there had not been a thorough genomic investigation for the molecular players that were behind this important phenotype. Our work identified four

different beta-glucosidase genes (*BGLU1-4*) across the *Hanseniaspora* genus, with *H. uvarum* specifically carrying two of them (*BGLU1* and *BGLU2*). Interestingly, this genomic analysis also uncovered two additional beta-glucosidase genes (*BGLU3* and *BGLU4*), that were identified in a small subset of *Hanseniaspora* species. One of these, *BGLU3*, was found in *Hanseniaspora* species (*e.g., H. osmophila*, *H. vineae*) belonging to the slow-evolving lineage (SEL) of *Hanseniaspora* (Steenwyk et al., 2019; Cadez et al., 2021), which demonstrate relatively high fermentation rates (Granchi et al., 2002; Viana et al., 2009; Badura et al., 2023), and increased similarity to the genomes of *Sd. ludwigii* UTAD17 (Tavares et al., 2021; Granchi et al., 2002). The *BGLU4* gene was only detected in a small subset of *Hanseniaspora* FEL species and was the only encoded protein that contained a signal peptide, suggesting a potential extracellular location (**Figure V. 5**).

Ultimately, this research opened a promising new pathway for investigating *BGLU* genes and their corresponding enzymes. While this thesis did not go into those details, it raised several important questions for future research. What substrates trigger the expression of these genes? How are they influenced by fermentation conditions such as pH, sugar, and ethanol levels? Are these enzymes suitable for wine production, and can they resist to the fermentation turmoil? Most importantly, can we evolve strains to produce more resilient enzymes? Alternatively, could we isolate, purify, and produce them independently while preserving their properties? How do they compare to commercially available enzymes? To generate impactful, biotechnologically-relevant work, the future of *Hanseniaspora* beta-glucosidase research should focus on answering these questions.

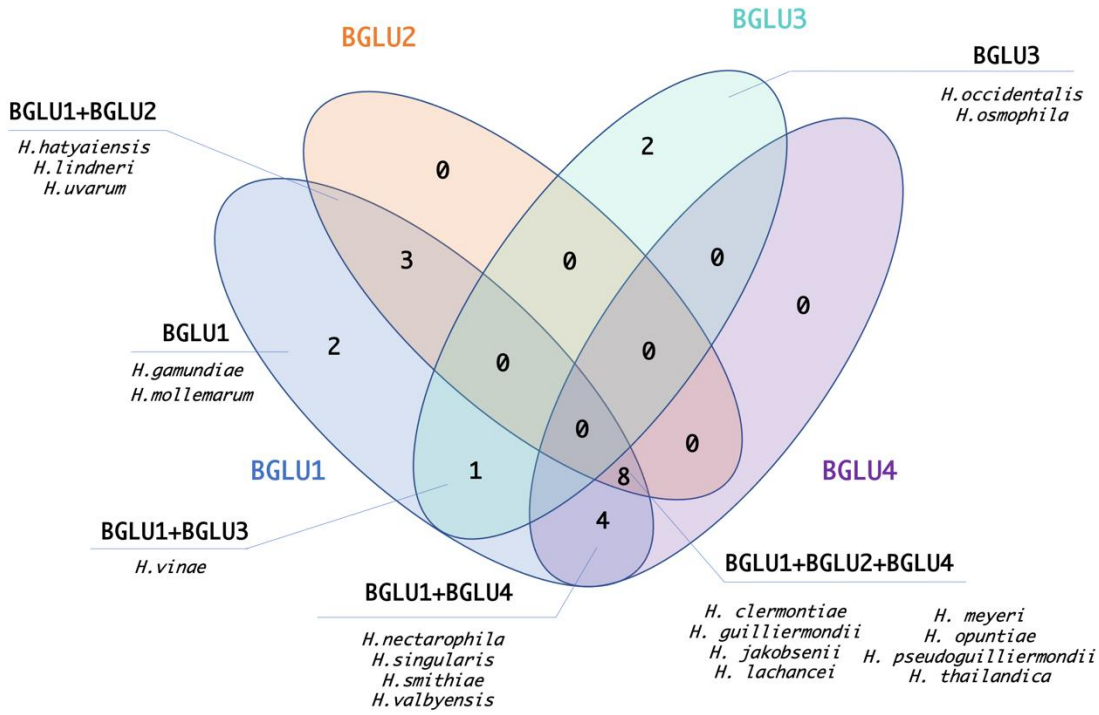


Figure V. 5- Venn diagram picturing the *Hanseniaspora* species analyzed in Chapter V and their corresponding beta-glucosidase genes.

In **Chapter IV.A** we also studied the impact of *H. uvarum and S. cerevisiae* co-inoculation in Moscatel Galego wine fermentation and aroma development. Could the beta-glucosidase producing *H. uvarum* modulate wine aroma development? Having identified a promising *H. uvarum* strain (MJT198), we conducted fermentations in natural Moscatel Galego must in sequential inoculation with *S. cerevisiae*. The obtained results demonstrated that the presence of strain MJT198 influenced primary wine aroma (such as those dependent on beta-glucosidase activity), increasing the concentration of free terpenes when compared to wines fermented solely with *S. cerevisiae*. While the results clearly demonstrated the impact of *H. uvarum* MJT198 in the modulation of wine aroma development, not much was understood about the molecular mechanisms involved. Consequently, we decided to study the metabolomic and transcriptomic signatures of the *H. uvarum and S. cerevisiae* fermentations.

In **Chapter IV. B** we analyzed the non-volatile metabolome of the wines produced in **Chapter IV.A**, following an untargeted metabolomics approach with FT-ICR-MS. This highly sensitive technique revealed profound changes in the non-volatile profile of the wines fermented with *H. uvarum and S. cerevisiae* when compared to those fermented solely by *S. cerevisiae*. Several unique metabolites were detected in wines inoculated with *H. uvarum,* indicating that this NSY plays a significant role in the modulation of the wine chemical profile. However, despite the clear modifications introduced by the *H. uvarum* strain to the wine matrix,

a major limitation of the untargeted metabolomic analysis was related to difficulties in the unequivocal identification of the specific metabolites, posing a bottleneck in our analysis. For this technique to be more broadly and successfully employed, more comprehensive and robust databases are essential to extract more meaningful and insightful information (**Figure V. 6**)

The subsequent transcriptomic analysis also revealed valuable information regarding the *H. uvarum* molecular mechanisms involved in wine modulation, such as the upregulation of *BGLU2* and downregulation of *BGLU1* at 72 hours. Still, the transcriptomic analysis was not without its limitations. Despite drawing important conclusions about the *H. uvarum* genes involved in wine aroma expression, the lack of at least three biological replicates inhibited an effective statistical analysis and diminished the overall impact of the experiment and the strength of our findings. Additionally, at the end of fermentation, we were unable to analyze the *H. uvarum* gene expression, due to the dominance of *S. cerevisiae* in the culture medium. Although it would be interesting to assess this gene set, we found that it was not a significant setback, as the analysis at 72 hours captured *H. uvarum's* precise role in modulating both major and minor volatile aroma compounds. The results obtained in this chapter not only reinforced findings from the volatile-targeted metabolome but also uncovered key genes that may play a role in shaping wine aroma. An interesting future approach to this experiment could be to track the gene expression of *H. uvarum's* and *S. cerevisiae*, both in single and co-culture, throughout the fermentation process. This would allow for the identification of the putative yeast-yeast interaction mechanisms that may be influencing the production of major and minor aroma volatiles.

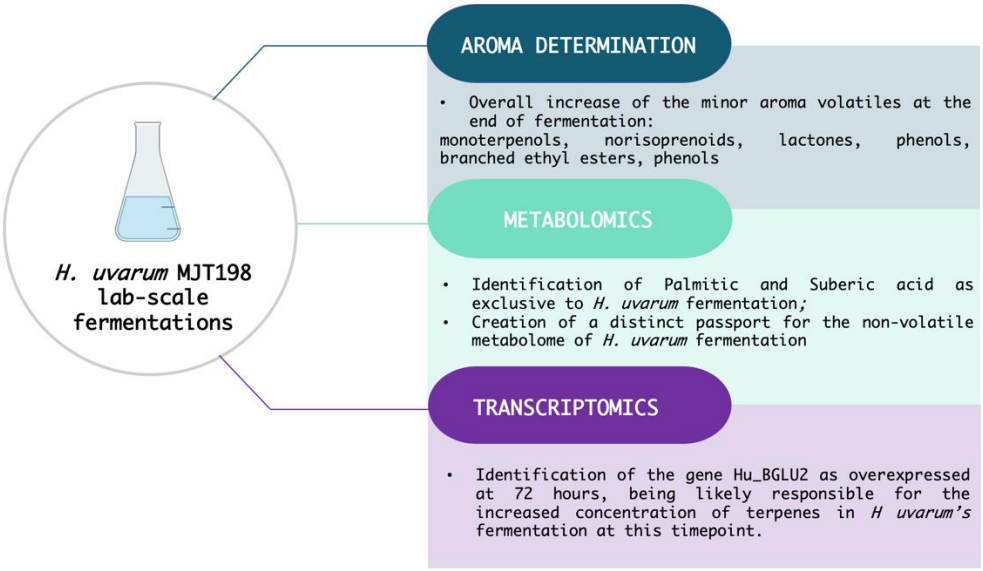


Figure V. 6– Summary of the findings extracted from the lab fermentation trials with *H. uvarum* MJT198 in Moscatel Galego wine must, and consequent volatile and non-volatile metabolomics analysis and transcriptomic analysis.

Conclusions

Overall, this thesis significantly advanced our understanding of strain-specific genomic adaptations in *Sd. ludwigii* and deepened our knowledge of *H. uvarum's role* in winemaking. Through comprehensive genomic and phenotypic analyses, we identified probable key mechanisms behind *Sd. ludwigii*'s persistence in challenging fermentation environments, such as multiple sulfite efflux pumps, thiamine biosynthesis and unique cell wall mannosylation. Additionally, the discovery of *Sd. ludwigii*'s strain-specific traits, given by the differing ploidy levels and varying abilities to utilize alternative carbon sources, suggested that there is more to *Sd. ludwigii* than its traditional association with spoilage, opening new doors for other applications in fermentation. On the other hand, *H. uvarum* was shown to enhance wine aroma beyond the typical secondary aroma modulation. Through the expression of beta-glucosidase enzymes, *H. uvarum* released volatile aromatic compounds and modulated wine primary aroma, contributing to a more complex sensory profile. However, additional studies are necessary to avoid off-flavors.

Despite the valuable insights gained, a significant path lies ahead. To fully harness the potential of NSYs, further exploration of their behavior in diverse grape musts is needed. For instance, after confirming *H. uvarum's* MJT198 ability to increase free terpene concentrations in natural must fermentations, it would be particularly useful to test this strain in grape varieties with low terpene content. This would allow us to explore the full potential of *H. uvarum* in

modulating aroma profiles in varieties that are naturally poor in aromatic compounds, thereby expanding its application in winemaking. As the wine industry adapts to changing climatic conditions, it is crucial not only to study these strains more thoroughly and systematically, identifying their key enzymes and how they respond to different fermentation conditions, but also to test them in diverse musts with varying characteristics. Matching the right yeast strain to the right grape variety will help maximize a grape's potential, modulating wine's aroma to truly capture its essence and its nature (**Figure V.7**).

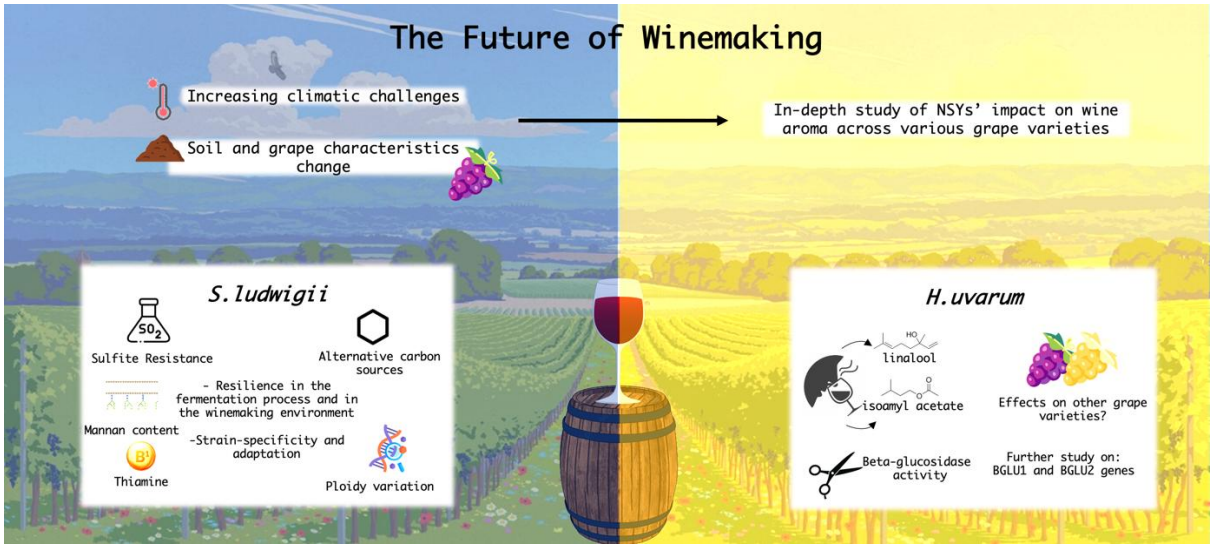


Figure V. 7- The Future of Winemaking is expected to feature the challenges of climate change and focus on more comprehensive studies of alternative technologies, with novel yeasts expectedly playing a key role in wine production.

References

Aa, E., Townsend, J. P., Adams, R. I., Nielsen, K. M. & Taylor, J. W. Population structure and gene evolution in *Saccharomyces cerevisiae*. FEMS Yeast Res 6, 702–715 (2006).

Adamenko, K., Kawa-Rygielska, J. & Kucharska, A. Z. Characteristics of Cornelian cherry sour non-alcoholic beers brewed with the special yeast *Saccharomycodes ludwigii*. Food Chem 312, 125968 (2020).

Agnihotri, S. et al. A Glimpse of the World of Volatile Fatty Acids Production and Application: A review. Bioengineered 13, 1249–1275 (2022).

Ahmad, K. M. et al. Small chromosomes among Danish *Candida glabrata* isolates originated through different mechanisms. Antonie Van Leeuwenhoek 104, 111–122 (2013).

Aksit, A. Molecular Characterization of Ethanol Resistance In *Saccharomyces cerevisiae*. (University of Groningen, Groningen, 2012).

Albergaria, H. & Arneborg, N. Dominance of *Saccharomyces cerevisiae* in alcoholic fermentation processes: role of physiological fitness and microbial interactions. Appl Microbiol Biotechnol 100, 2035–2046 (2016).

Albergaria, H., Torrao, A., Hogg, T. & Girio, F. Physiological behaviour of in aerobic glucose-limited continuous cultures. FEMS Yeast Res 3, 211–216 (2003).

Albertin, W. et al. *Hanseniaspora uvarum* from Winemaking Environments Show Spatial and Temporal Genetic Clustering. Front Microbiol 6, 1569 (2016).

Aldred, E. Terpenes. Pharmacology 167–174 (2009) doi:10.1016/B978-0-443-06898-0.00022-0.

Allen, M. S. & Lacey, M. J. Methoxypyrazines of Grapes and Wines. ACS Symposium Series 714, 31–38 (1998).

Ames, R. M. et al. Gene Duplication and Environmental Adaptation within Yeast Populations. Genome Biol Evol 2, 591–601 (2010).

Atsumi, S. et al. Engineering the isobutanol biosynthetic pathway in *Escherichia coli* by comparison of three aldehyde reductase/alcohol dehydrogenase genes. Appl Microbiol Biotechnol 85, 651–657 (2010).

Australian Wine Research Institute. Wine flavours, faults and taints. (2024).

Avram, D. & Bakalinsky, A. T. *SSU1* encodes a plasma membrane protein with a central role in a network of proteins conferring sulfite tolerance in *Saccharomyces cerevisiae*. J Bacteriol 179, 5971–5974 (1997).

Avram, D., Leid, M. & Bakalinsky, A. T. Fzf1p of *Saccharomyces* cerevisiae is a positive regulator of *SSU1* transcription and its first zinc finger region is required for DNA binding. Yeast 15, 473–480 (1999).

Badura, J. et al. Aroma Profiles of *Vitis vinifera* L. cv. Gewürztraminer Must Fermented with Co-Cultures of *Saccharomyces cerevisiae* and Seven *Hanseniaspora* spp. Fermentation 9, 109 (2023).

Bae, S.-J., Kim, S. & Hahn, J.-S. Efficient production of acetoin in *Saccharomyces cerevisiae* by disruption of 2,3-butanediol dehydrogenase and expression of NADH oxidase. Sci Rep 6, 27667 (2016).

Banerjee, S. & Mazumdar, S. Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. Int J Anal Chem 2012, 1–40 (2012).

Barata, A., Malfeito-Ferreira, M. & Loureiro, V. The microbial ecology of wine grape berries. Int J Food Microbiol 153, 243–259 (2012).

Barbosa, C., Mendes-Faia, A., Lage, P., Mira, N. P. & Mendes-Ferreira, A. Genomic expression program of *Saccharomyces cerevisiae* along a mixed-culture wine fermentation with *Hanseniaspora guilliermondii*. Microb Cell Fact 14, 1–17 (2015).

Barbosa, C., Ramalhosa, E., Vasconcelos, I., Reis, M. & Mendes-Ferreira, A. Machine Learning Techniques Disclose the Combined Effect of Fermentation Conditions on Yeast Mixed-Culture Dynamics and Wine Quality. Microorganisms 10, 107 (2022).

Barbosa, R. et al. Multiple Rounds of Artificial Selection Promote Microbe Secondary Domestication—The Case of Cachaça Yeasts. Genome Biol Evol 10, 1939–1955 (2018).

Barnett, J. A history of research on yeasts 10: foundations on yeasts genetics. Yeast 24, 799–845 (2007).

Barrow, M. P., Burkitt, W. I. & Derrick, P. J. Principles of Fourier transform ion cyclotron resonance mass spectrometry and its application in structural biology. Analyst vol. 130 18–28 Preprint at https://doi.org/10.1039/b403880k (2005).

Barton, A. B., Pekosz, M. R., Kurvathi, R. S. & Kaback, D. B. Meiotic Recombination at the Ends of Chromosomes in *Saccharomyces cerevisiae*. Genetics 179, 1221–1235 (2008).

Bataillon, M., Rico, A., Sablayrolles, J.-M., Salmon, J.-M. & Barre, P. Early thiamin assimilation by yeasts under enological conditions: Impact on alcoholic fermentation kinetics. J Ferment Bioeng 82, 145–150 (1996).

Beech, F. & Davenport, R. The role of yeasts in cider-making. Agricultural and Food Sciences, Biology (1970).

Behrens, C. J., Krahe, N. K., Linke, D. & Berger, R. G. BadGluc, a β-glucosidase from *Bjerkandera adusta* with anthocyanase properties. Bioprocess Biosyst Eng 41, 1391–1401 (2018).

Belda, I. et al. Influence of *Torulaspora delbrueckii* in varietal thiol (3-SH and 4-MSP) release in wine sequential fermentations. Int J Food Microbiol 257, 183–191 (2017).

Belda, I. et al. Microbial contribution to Wine aroma and its intended use for Wine quality improvement. Molecules vol. 22 Preprint at https://doi.org/10.3390/molecules22020189 (2017).

Bellut, K., Krogerus, K. & Arendt, E. K. *Lachancea fermentati* Strains Isolated From Kombucha: Fundamental Insights, and Practical Application in Low Alcohol Beer Brewing. Front Microbiol 11, 764 (2020).

Bianchi, F. et al. Persistence of a Yeast-Based (*Hanseniaspora uvarum*) Attract-and-Kill Formulation against *Drosophila suzukii* on Grape Leaves. Insects 11, 810 (2020).

Black, C. A., Parker, M., Siebert, T. E., Capone, D. L. & Francis, I. L. Terpenoids and their role in wine flavour: recent advances. Aust J Grape Wine Res 21, 582–600 (2015).

Bode, W. K. H. The Ancient History of the Making and Development of Wine. International Journal of Wine Marketing 4, 36–43 (1992).

Boekhout, T., Kurtzman, C. P., O'Donnell, K. & Smith, M. T. Phylogeny of the Yeast Genera *Hanseniaspora* (Anamorph *Kloeckera*), *Dekkera* (Anamorph *Brettanomyces*), and *Eeniella* as Inferred from Partial 26S Ribosomal DNA Nucleotide Sequences. Int J Syst Bacteriol 44, 781–786 (1994).

Bohnenkamp, A. C. et al. Multilevel optimisation of anaerobic ethyl acetate production in engineered *Escherichia coli*. Biotechnol Biofuels 13, 65 (2020).

Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120 (2014).

Bordet, F. et al. Different Wines from Different Yeasts? "*Saccharomyces cerevisiae* Intraspecies Differentiation by Metabolomic Signature and Sensory Patterns in Wine". Microorganisms 9, 2327 (2021).

Bordet, F. et al. Expanding the diversity of Chardonnay aroma through the metabolic interactions of *Saccharomyces cerevisiae* cocultures. Front Microbiol 13, (2023).

Boulton, R. B., Singleton, V. L., Bisson, L. F. & Kunkee, R. E. Microbiological Spoilage of Wine and its Control. in Principles and Practices of Winemaking 352–381 (Springer US, Boston, MA, 1999). doi:10.1007/978-1-4757-6255-6_9.

Boulton, R. B., Singleton, V. L., Bisson, L. F. & Kunkee, R. E. Yeast and Biochemistry of Ethanol Fermentation. in Principles and Practices of Winemaking 102–192 (Springer US, Boston, MA, 1999). doi:10.1007/978-1-4757-6255-6_4.

Boundy-Mills, K., Stratford, M. & Miller, M. W. *Saccharomycodes* E.C. in The Yeasts vol. 2 747–750 (Elsevier, 2011).

Brion, C., Ambroset, C., Delobel, P., Sanchez, I. & Blondin, B. Deciphering regulatory variation of *THI* genes in alcoholic fermentation indicate an impact of Thi3p on *PDC1* expression. BMC Genomics 15, 1085 (2014).

Brion, C., Pflieger, D., Souali-Crespo, S., Friedrich, A. & Schacherer, J. Differences in environmental stress response among yeasts is consistent with species-specific lifestyles. Mol Biol Cell 27, 1694–1705 (2016).

Brown, C. A., Murray, A. W. & Verstrepen, K. J. Rapid Expansion and Functional Divergence of Subtelomeric Gene Families in Yeasts. Current Biology 20, 895–903 (2010).

Buckle, J. Basic Plant Taxonomy, Basic Essential Oil Chemistry, Extraction, Biosynthesis, and Analysis. in Clinical Aromatherapy 37–72 (Elsevier, 2015). doi:10.1016/B978-0-7020-5440-2.00003-6.

Byrne, K. P. & Wolfe, K. H. The Yeast Gene Order Browser: Combining curated homology and syntenic context reveals gene fate in polyploid species. Genome Res 15, 1456–1461 (2005).

Cabral, S., Prista, C., Loureiro-Dias, M. C. & Leandro, M. J. Occurrence of *FFZ* genes in yeasts and correlation with fructophilic behaviour. Microbiology (N Y) 161, 2008–2018 (2015).

Cadez, N. & Smith, M. Th. *Kloeckera*. in The Yeasts 1287–1289 (Elsevier, 2011).

Cadez, N. et al. *Hanseniaspora smithiae* sp. nov., a Novel Apiculate Yeast Species From Patagonian Forests That Lacks the Typical Genomic Domestication Signatures for Fermentative Environments. Front Microbiol 12, (2021).

Cadez, N., Bellora, N., Ulloa, R., Hittinger, C. T. & Libkind, D. Genomic content of a novel yeast species *Hanseniaspora gamundiae* sp. nov. from fungal stromata (*Cyttaria*) associated with a unique fermented beverage in Andean Patagonia, Argentina. PLoS One 14, e0210792 (2019).

Cadez, N., Raspor, P. & Smith, M. Th. Phylogenetic placement of *Hanseniaspora–Kloeckera* species using multigene sequence analysis with taxonomic implications: descriptions of *Hanseniaspora pseudoguilliermondii* sp. nov. and *Hanseniaspora occidentalis* var. citrica var. nov. Int J Syst Evol Microbiol 56, 1157–1165 (2006).

Cadez, N., Raspor, P., Cock, A. W. A. M., Boekhout, T. & Smith, M. T. Molecular identification and genetic diversity within species of the genera *Hanseniaspora* and *Kloeckera*. FEMS Yeast Res 1, 279–289 (2002).

Caffrey, A. & Ebeler, S. E. The Occurrence of Glycosylated Aroma Precursors in *Vitis vinifera* Fruit and *Humulus lupulus* Hop Cones and Their Roles in Wine and Beer Volatile Aroma Production. Foods 2021, Vol. 10, Page 935 10, 935 (2021).

Canon, F., Caillé, S., Sarni-Manchado, P. & Cheynier, V. Wine taste and mouthfeel. Managing Wine Quality 41–95 (2022)

Capece, A. & Romano, P. Yeasts and Their Metabolic Impact on Wine Flavour. in Yeasts in the Production of Wine 43–80 (Springer New York, New York, NY, 2019).

Capece, A., Romaniello, R., Siesto, G. & Romano, P. Conventional and Non-Conventional Yeasts in Beer Production. Fermentation 4, 38 (2018).

Cardona, F., Carrasco, P., Pérez-Ortín, J. E., del Olmo, M. lí & Aranda, A. A novel approach for the improvement of stress resistance in wine yeasts. Int J Food Microbiol 114, 83–91 (2007).

Caridi, A. Enological functions of parietal yeast mannoproteins. Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology 89, 417–422 (2006).

Carpena, M. et al. Secondary Aroma: Influence of Wine Microorganisms in Their Aroma Profile. Foods 10, (2021).

Carrau, F. M. et al. De novo synthesis of monoterpenes by *Saccharomyces cerevisiae* wine yeasts. FEMS Microbiol Lett 243, 107–115 (2005).

Carrau, F. M., Boido, E. & Dellacassa, E. Terpenoids in Grapes and Wines: Origin and Micrometabolism during the Vinification Process.

Casalone, E. et al. Mechanism of resistance to sulphite in *Saccharomyces cerevisiae*. Curr Genet 22, 435–440 (1992).

Chambers, P. J. & Pretorius, I. S. Fermenting knowledge: the history of winemaking, science and yeast research. EMBO Rep 11, 914–920 (2010).

Chang, S.-L., Lai, H.-Y., Tung, S.-Y. & Leu, J.-Y. Dynamic Large-Scale Chromosomal Rearrangements Fuel Rapid Adaptation in Yeast Populations. PLoS Genet 9, e1003232 (2013).

Chattopadhyay, M. K., Tabor, C. W. & Tabor, H. Spermidine but not spermine is essential for hypusine biosynthesis and growth in *Saccharomyces cerevisiae*: Spermine is converted to spermidine in vivo by the FMS1-amine oxidase. Proc Natl Acad Sci U S A 100, 13869–13874 (2003).

Chen, D. et al. Global Transcriptional Responses of Fission Yeast to Environmental Stress. Mol Biol Cell 14, 214–229 (2003).

Chen, K. & Li, J. A glance into the aroma of white wine. White Wine Technology 313–326 (2022)

Chen, X., Qiao, Y.-Z. & Hui, F.-L. *Hanseniaspora menglaensis* f.a., sp. nov., a novel apiculate yeast species isolated from rotting wood. Int J Syst Evol Microbiol 73, (2023).

Cheynier, V., Schneider, R., Salmon, J. M. & Fulcrand, H. Chemistry of Wine. Comprehensive Natural Products II: Chemistry and Biology 3, 1119–1172 (2010).

Ciani, M. & Fatichenti, F. Selective sugar consumption by apiculate yeasts. Lett Appl Microbiol 28, 203–206 (1999).

Ciani, M. & Maccarelli, F. Oenological properties of Non-*Saccharomyces* yeasts associated with wine-making. World J Microbiol Biotechnol 14, 199–203 (1997).

Ciani, M., Comitini, F., Mannazzu, I. & Domizio, P. Controlled mixed culture fermentation: a new perspective on the use of Non-*Saccharomyces* yeasts in winemaking. FEMS Yeast Res 10, 123–133 (2010).

Cincotta, F., Verzera, A., Tripodi, G. & Condurso, C. Determination of Sesquiterpenes in Wines by HS-SPME Coupled with GC-MS. Chromatography 2015, Vol. 2, Pages 410-421 2, 410–421 (2015).

Clish, C. B. Metabolomics: an emerging but powerful tool for precision medicine. Molecular Case Studies 1, a000588 (2015).

Coi, A. L. et al. Genomic signatures of adaptation to wine biological ageing conditions in biofilm-forming flor yeasts. Mol Ecol 26, 2150–2166 (2017).

Combina, M. et al. Yeasts associated to Malbec grape berries from Mendoza, Argentina. J Appl Microbiol 98, 1055–1061 (2005).

Comisarow, M. B. & Marshall, A. G. Fourier transform ion cyclotron resonance spectroscopy. Chem Phys Lett 25, 282–283 (1974).

Conant, G. C. & Wolfe, K. H. Turning a hobby into a job: How duplicated genes find new functions. Nat Rev Genet 9, 938–950 (2008).

Cordente, A. G., Schmidt, S., Beltran, G., Torija, M. J. & Curtin, C. D. Harnessing yeast metabolism of aromatic amino acids for fermented beverage bioflavouring and bioproduction. Appl Microbiol Biotechnol 103, 4325–4336 (2019).

Cray, J. A. et al. The biology of habitat dominance; can microbes behave as weeds? Microb Biotechnol 6, 453–492 (2013).

Cus, F. & Jenko, M. The influence of yeast strains on the composition and sensory quality of Gewurztraminer wine. Food Technol Biotechnol 51, 547–554 (2013).

De Francesco, G., Turchetti, B., Sileoni, V., Marconi, O. & Perretti, G. Screening of new strains of *Saccharomycodes ludwigii* and *Zygosaccharomyces rouxii* to produce low-alcohol beer. Journal of the Institute of Brewing 121, 113–121 (2015).

de Hoffmann, E. & Stroobant, V. Mass Spectrometry: Principles and Applications, 3rd Edition. Wiley, John Wiley & Sons, Ltd (2007).

De Luca, V. Wines. in Comprehensive Biotechnology vol. 4 241–255 (Elsevier, 2011).

de Ovalle, S., Brena, B. & González-Pombo, P. Influence of beta glucosidases from native yeast on the aroma of Muscat and Tannat wines. Food Chem 346, 128899 (2021).

De-La-Fuente-Blanco, A., Sáenz-Navajas, M. P. & Ferreira, V. On the effects of higher alcohols on red wine aroma. Food Chem 210, 107–114 (2016).

Deed, R. C., Hou, R., Kinzurik, M. I., Gardner, R. C. & Fedrizzi, B. The role of yeast *ARO8*, *ARO9* and *ARO10* genes in the biosynthesis of 3-(methylthio)-1-propanol from L -methionine during fermentation in synthetic grape medium. FEMS Yeast Res 19, (2019).

Demyttenaere, J. C. R., Del Carmen Herrera, M. & De Kimpe, N. Biotransformation of geraniol, nerol and citral by sporulated surface cultures of *Aspergillus niger* and *Penicillium* sp. Phytochemistry 55, 363–373 (2000).

Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21 (2013).

Domizio, P. et al. Outlining a future for Non-*Saccharomyces* yeasts: Selection of putative spoilage wine strains to be used in association with *Saccharomyces cerevisiae* for grape juice fermentation. Int J Food Microbiol 147, 170–180 (2011).

Domizio, P. et al. Potential spoilage Non-*Saccharomyces* yeasts in mixed cultures with *Saccharomyces cerevisiae*. Ann Microbiol 61, 137–144 (2011).

Domizio, P., Liu, Y., Bisson, L. F. & Barile, D. Use of non-*Saccharomyces* wine yeasts as novel sources of mannoproteins in wine. Food Microbiol 43, 5–15 (2014).

Du Plessis, H. et al. Modulation of Wine Flavor using *Hanseniaspora uvarum* in Combination with Different *Saccharomyces cerevisiae*, Lactic Acid Bacteria Strains and Malolactic Fermentation Strategies. Fermentation 5, 64 (2019).

Dujon, B. et al. Genome evolution in yeasts. Nature 430, 35–44 (2004).

Dziadas, M. & Jeleń, H. H. Comparison of enzymatic and acid hydrolysis of bound flavor compounds in model system and grapes. Food Chem 190, 412–418 (2016).

Echave, J., Barral, M., Fraga-Corral, M., Prieto, M. A. & Simal-Gandara, J. Bottle Aging and Storage of Wines: A Review. Molecules 26, 713 (2021).

Edgar, R. C. MUSCLE: A multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5, 1–19 (2004).

Erkmen, O. & Bozoglu, F. Food Microbiology: Principles into Practice. (Wiley, 2016).

Estela-Escalante, W. et al. Actividad Fermentativa de *Saccharomycodes ludwigii* y Evaluación de la Síntesis de Compuestos de Importancia Sensorial durante la Fermentacion de Jugo de Manzana. Rev. Espec. Cienc. Quim. Biol. 14, 12–23 (2011).

Esteve-Zarzoso, B., Peris-Torán, M. J., Ramón, D. & Querol, A. Molecular characterisation of *Hanseniaspora* species. Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology 80, 85–92 (2001).

Esteves, M. et al. Characterizing the potential of the non-conventional yeast *Saccharomycodes ludwigii* UTAD17 in winemaking. Microorganisms 7, (2019).

Estreicher, S. K. The beginning of wine and viticulture. Physica status solidi c 14, (2017).

Ezov, T. K. et al. Molecular-Genetic Biodiversity in a Natural Population of the Yeast *Saccharomyces cerevisiae* From "Evolution Canyon": Microsatellite Polymorphism, Ploidy and Controversial Sexual Status. Genetics 174, 1455–1468 (2006).

F, V., Gil, J., Genoves, S., Valles, S. & Manzanares, P. Rational selection of non-*Saccharomyces* wine yeasts for mixed starters based on ester formation and enological traits. Food Microbiol 25, 778–785 (2008).

Fay, J. C. & Benavides, J. A. Evidence for Domesticated and Wild Populations of *Saccharomyces cerevisiae*. PLoS Genet 1, e5 (2005).

Fernández-Pacheco, P., García-Béjar, B., Briones Pérez, A. & Arévalo-Villena, M. Free and Immobilised β-Glucosidases in Oenology: Biotechnological Characterisation and Its Effect on Enhancement of Wine Aroma. Front Microbiol 12, 2327 (2021).

Fidalgo, M., Barrales, R. R., Ibeas, J. I. & Jimenez, J. Adaptive evolution by mutations in the *FLO11* gene. Proceedings of the National Academy of Sciences 103, 11228–11233 (2006).

Flamini, R. et al. Combining liquid chromatography and tandem mass spectrometry approaches to the study of monoterpene glycosides (aroma precursors) in wine grape. J Mass Spectrom 53, 792–800 (2018).

Fleet, G. H. The microbiology of alcoholic beverages. in Microbiology of Fermented Foods 217–262 (Springer US, Boston, MA, 1998). doi:10.1007/978-1-4613-0309-1_9.

Fleet, G. H. Wine yeasts for the future. in FEMS Yeast Research vol. 8 979–995 (2008).

Fleet, G. Yeast interactions and wine flavour. Int J Food Microbiol 86, 11–22 (2003).

Fogleman, J. C., Starmer, W. T. & Heed, W. B. Comparisons of yeast florae from natural substrates and larval guts of southwestern Drosophila. Oecologia 52, 187–191 (1982).

Fresno, J. M. Del et al. The Impact of *Hanseniaspora vineae* Fermentation and Ageing on Lees on the Terpenic Aromatic Profile of White Wines of the Albillo Variety. International Journal of Molecular Sciences 2021, Vol. 22, Page 2195 22, 2195 (2021).

Fukuda, K. et al. Balance of Activities of Alcohol Acetyltransferase and Esterase in *Saccharomyces cerevisiae* Is Important for Production of Isoamyl Acetate. Appl Environ Microbiol 64, 4076–4078 (1998).

G. Pereira, A. et al. Management of Wine Aroma Compounds: Principal Basis and Future Perspectives. in Chemistry and Biochemistry of Winemaking, Wine Stabilization and Aging (IntechOpen, 2021). doi:10.5772/intechopen.92973.

Gaensly, F., Agustini, B. C., da Silva, G. A., Picheth, G. & Bonfim, T. M. B. Autochthonous yeasts with β-glucosidase activity increase resveratrol concentration during the alcoholic fermentation of *Vitis labrusca* grape must. J Funct Foods 19, 288–295 (2015).

Gallo, A. et al. Aroma Features of *Hanseniaspora vineae* Hv205 Wines in Sequential and Co-Inoculation Strategies. Fermentation 10, 191 (2024).

Gallone, B. et al. Domestication and Divergence of *Saccharomyces cerevisiae* Beer Yeasts. Cell 166, 1397-1410.e16 (2016).

Gao, P. et al. Indigenous Non-*Saccharomyces* Yeasts With β-Glucosidase Activity in Sequential Fermentation With *Saccharomyces cerevisiae*: A Strategy to Improve the Volatile Composition and Sensory Characteristics of Wines. Front Microbiol 13, (2022).

García-Ríos, E. & Guillamón, J. M. Mechanisms of Yeast Adaptation to Wine Fermentations. in 37–59 (2019). doi:10.1007/978-3-030-13035-0_2.

García-Ríos, E. & Guillamón, J. M. Sulfur dioxide resistance in *Saccharomyces cerevisiae*: beyond *SSU1*. Microbial Cell 6, 527–530 (2019).

García, M., Esteve-Zarzoso, B. & Arroyo, T. Non-*Saccharomyces* Yeasts: Biotechnological Role for Wine Production. in Grape and Wine Biotechnology (InTech, 2016). doi:10.5772/64957.

García, M., Esteve-Zarzoso, B., Cabellos, J. M. & Arroyo, T. Advances in the Study of *Candida stellata*. Fermentation 2018, Vol. 4, Page 74 4, 74 (2018).

Gasch, A. P. & Werner-Washburne, M. The genomics of yeast responses to environmental stress and starvation. Funct Integr Genomics 2, 181–192 (2002).

Gasch, A. P. Comparative genomics of the environmental stress response in ascomycete fungi. Yeast 24, 961–976 (2007).

Gasch, A. P. et al. Further support for aneuploidy tolerance in wild yeast and effects of dosage compensation on gene copy-number evolution. Elife 5, (2016).

Gasch, A. P. et al. Genomic Expression Programs in the Response of Yeast Cells to Environmental Changes. Mol Biol Cell 11, 4241–4257 (2000).

Gerstein, A. C. et al. Polyploid Titan Cells Produce Haploid and Aneuploid Progeny To Promote Stress Adaptation. mBio 6, (2015).

Gerstein, A. C. et al. Too Much of a Good Thing: The Unique and Repeated Paths Toward Copper Adaptation. Genetics 199, 555–571 (2015).

Ghosh, S. Metagenomic screening of cell wall hydrolases, their anti-fungal activities and potential role in wine fermentation. PhD Dissertation (2015). Available at: https://core.ac.uk/download/pdf/37438565.pdf

Giang, T., He, J., Belaidi, S. & Scholz, H. Key Odorants Regulate Food Attraction in *Drosophila melanogaster*. Front Behav Neurosci 11, (2017).

Gilchrist, C. & Stelkens, R. Aneuploidy in yeast: Segregation error or adaptation mechanism? Yeast 36, 525–539 (2019).

Giorello, F. et al. Genomic and Transcriptomic Basis of *Hanseniaspora vineae's* Impact on Flavor Diversity and Wine Quality. Appl Environ Microbiol 85, (2019).

Giorello, F. M. et al. Genome Sequence of the Native Apiculate Wine Yeast *Hanseniaspora vineae* T02/19AF. Genome Announc 2, (2014).

Giovani, G., Rosi, I. & Bertuccioli, M. Quantification and characterization of cell wall polysaccharides released by non-*Saccharomyces* yeast strains during alcoholic fermentation. Int J Food Microbiol 160, 113–118 (2012).

Giraud, A. et al. Costs and Benefits of High Mutation Rates: Adaptive Evolution of Bacteria in the Mouse Gut. Science (1979) 291, 2606–2608 (2001).

Gomes, R. J., Borges, M. de F., Rosa, M. de F., Castro-Gómez, R. J. H. & Spinosa, W. A. Acetic Acid Bacteria in the Food Industry: Systematics, Characteristics and Applications. Food Technol Biotechnol 56, (2018).

González-Barreiro, C., Rial-Otero, R., Cancho-Grande, B. & Simal-Gándara, J. Wine Aroma Compounds in Grapes: A Critical Review. Crit Rev Food Sci Nutr 55, 202–218 (2015).

González, B. et al. The production of aromatic alcohols in non-*Saccharomyces* wine yeast is modulated by nutrient availability. Food Microbiol 74, 64–74 (2018).

Görner, W. et al. Acute glucose starvation activates the nuclear localization signal of a stress-specific yeast transcription factor. EMBO Journal 21, 135–144 (2002).

Gorter, F. A. et al. Genomics of Adaptation Depends on the Rate of Environmental Change in Experimental Yeast Populations. Mol Biol Evol 34, 2613–2626 (2017).

Granchi, L., Ganucci, D., Messini, A. & Vincenzini, M. Oenological properties of *Hanseniaspora osmophila* and *Kloeckera corticis* from wines produced by spontaneous fermentations of normal and dried grapes. FEMS Yeast Res 2, 403–407 (2002).

Groenewald, M. et al. Diversity of yeast species from Dutch garden soil and the description of six novel Ascomycetes. FEMS Yeast Res (2018).

Gschaedler, A. Contribution of non-conventional yeasts in alcoholic beverages. Current Opinion in Food Science vol. 13 73–77 Preprint (2017).

Guaragnella, N. & Bettiga, M. Acetic acid stress in budding yeast: From molecular mechanisms to applications. Yeast 38, 391–400 (2021).

Guaragnella, N. et al. Genome Sequencing and Comparative Analysis of Three *Hanseniaspora uvarum* Indigenous Wine Strains Reveal Remarkable Biotechnological Potential. Front Microbiol 10, (2020).

Guillamón, J. M. & Barrio, E. Genetic Polymorphism in Wine Yeasts: Mechanisms and Methods for Its Detection. Front Microbiol 8, (2017).

Guillaume, C., Delobel, P., Sablayrolles, J.-M. & Blondin, B. Molecular Basis of Fructose Utilization by the Wine Yeast *Saccharomyces cerevisiae* : a Mutated *HXT3* Allele Enhances Fructose Fermentation. Appl Environ Microbiol 73, 2432–2439 (2007).

Hage, A. El & Houseley, J. Resolution of budding yeast chromosomes using pulsed-field gel electrophoresis. Methods in Molecular Biology 1054, 195–207 (2013).

Hakes, L., Pinney, J. W., Lovell, S. C., Oliver, S. G. & Robertson, D. L. All duplicates are not equal: the difference between small-scale and genome duplication. Genome Biol 8, R209 (2007).

Harrison, M.-C. et al. Machine learning enables identification of an alternative yeast galactose utilization pathway. Proceedings of the National Academy of Sciences 121, (2024).

Hayyan, J. ibn. Kitab Ikhraj Ma Fi Al-Quwwa Ila al-Fi`l, Part of Mukhtarat Rasa'il Jabir Ibn Hayyan. (1935).

Hazelwood, L. A., Daran, J. M., Van Maris, A. J. A., Pronk, J. T. & Dickinson, J. R. The Ehrlich Pathway for Fusel Alcohol Production: a Century of Research on *Saccharomyces cerevisiae* Metabolism. Appl Environ Microbiol 74, 2259 (2008).

Healey, K. R. et al. Prevalent mutator genotype identified in fungal pathogen *Candida glabrata* promotes multi-drug resistance. Nat Commun 7, 11128 (2016).

Henscke, P. A. & Jiranek, V. Yeasts-metabolism of nitrogen compounds in Wine Microbiology and Biotechnology. Harwood Academic Publishers 77–164 (1993).

Hirasawa, T. et al. Identification of target genes conferring ethanol stress tolerance to *Saccharomyces cerevisiae* based on DNA microarray data analysis. J Biotechnol 131, 34–44 (2007).

Hittinger, C. T. et al. Remarkably ancient balanced polymorphisms in a multi-locus gene network. Nature 464, 54–58 (2010).

Hittinger, C. T., Rokas, A. & Carroll, S. B. Parallel inactivation of multiple *GAL* pathway genes and ecological diversification in yeasts. Proceedings of the National Academy of Sciences 101, 14144–14149 (2004).

Holden, H. M., Rayment, I. & Thoden, J. B. Structure and Function of Enzymes of the Leloir Pathway for Galactose Metabolism. Journal of Biological Chemistry 278, 43885–43888 (2003).

Holt, S., Miks, M. H., De Carvalho, B. T., Foulquié-Moreno, M. R. & Thevelein, J. M. The molecular biology of fruity and floral aromas in beer and other alcoholic beverages. FEMS Microbiology Reviews vol. 43 193–222 (2019).

Holt, S., Mukherjee, V., Lievens, B., Verstrepen, K. J. & Thevelein, J. M. Bioflavoring by non-conventional yeasts in sequential beer fermentations. Food Microbiol 72, 55–66 (2018).

Horowitz, H., Thorburn, P. & Haber, J. E. Rearrangements of highly polymorphic regions near telomeres of *Saccharomyces cerevisiae*. Mol Cell Biol 4, 2509–2517 (1984).

Hose, J. et al. Dosage compensation can buffer copy-number variation in wild yeast. Elife 4, (2015).

Hu, K. et al. Enhancing wine ester biosynthesis in mixed *Hanseniaspora uvarum*/*Saccharomyces cerevisiae* fermentation by nitrogen nutrient addition. Food Research International 123, 559–566 (2019).

Hu, K. et al. Potential of Glycosidase from Non-*Saccharomyces* Isolates for Enhancement of Wine Aroma. J Food Sci 81, M935–M943 (2016).

Hu, K., Jin, G. J., Xu, Y. H. & Tao, Y. S. Wine aroma response to different participation of selected *Hanseniaspora uvarum* in mixed fermentation with *Saccharomyces cerevisiae*. Food Research International 108, 119–127 (2018).

Hu, K., Jin, G.-J., Mei, W.-C., Li, T. & Tao, Y.-S. Increase of medium-chain fatty acid ethyl ester content in mixed *H. uvarum*/*S. cerevisiae* fermentation leads to wine fruity aroma enhancement. Food Chem 239, 495–501 (2018).

Hu, K., Jin, G.-J., Xu, Y.-H. & Tao, Y.-S. Wine aroma response to different participation of selected *Hanseniaspora uvarum* in mixed fermentation with *Saccharomyces cerevisiae*. Food Research International 108, 119–127 (2018).

Huang, R. et al. Characterization of the β-Glucosidase activity in indigenous yeast isolated from wine regions in China. J Food Sci 86, 2327–2345 (2021).

Huerta-Cepas, J. et al. EggNOG 5.0: A hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res 47, D309–D314 (2019).

Huerta-Cepas, J. et al. Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. Mol Biol Evol 34, 2115–2122 (2017).

Hutzler, M., Riedl, R., Koob, J. & Jacob, F. Fermentation and spoilage yeasts and their relevance for the beverage industry - A review. BrewingScience 65, 33–52 (2012).

International Organisation of Vine and Wine. Compendium of International Methods of Wine and Must Analysis. vol. 1 (Paris, 2020).

Issa-Issa, H. et al. Effect of Aging Vessel (Clay-Tinaja versus Oak Barrel) on the Volatile Composition, Descriptive Sensory Profile, and Consumer Acceptance of Red Wine. Beverages 7, 35 (2021).

Ivit, N. N. et al. Making natural sparkling wines with non-*Saccharomyces* yeasts. European Food Research and Technology 244, 925–935 (2018).

Jackowski, M. et al. Comparison of Two Commercially Available Strains, *Saccharomycodes ludwigii* and *Torulaspora delbrueckii*, for the Production of Low-Alcohol Beer. Beverages 9, 66 (2023).

Jackson, R. S. Chemical Constituents of Grapes and Wine. Wine Science 270–331 (2008).

Jackson, R. Wine Science: Principles and Applications. (Elsevier, 2008).

Jarvis, B. Cider (Cyder; Hard Cider). in Encyclopedia of Food Microbiology 437–443 (Elsevier, 2014).

Jensen, L. J. et al. eggNOG: Automated construction and annotation of orthologous groups of genes. Nucleic Acids Res 36, (2008).

Jeromel, A., Korenika, A. M. J. & Tomaz, I. An influence of different yeast species on wine aroma composition. Fermented Beverages: Volume 5. The Science of Beverages 171–285 (2019).

John Varriano. Wine: A Cultural History. (Reaktion Books, 2010).

Jolly, N. P., Varela, C. & Pretorius, I. S. Not your ordinary yeast: Non-*Saccharomyces* yeasts in wine production uncovered. FEMS Yeast Research vol. 14 215–237 (2014).

Kadowaki, M. et al. Chromosomal Aneuploidy Improves the Brewing Characteristics of Sake Yeast. Appl Environ Microbiol 83, (2017).

Kanehisa, M., Sato, Y. & Morishima, K. BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. J Mol Biol 428, 726–731 (2016).

Katz Ezov, T. et al. Heterothallism in *Saccharomyces cerevisiae* isolates from nature: effect of HO locus on the mode of reproduction. Mol Ecol 19, 121–131 (2010).

Kellis, M., Birren, B. W. & Lander, E. S. Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. Nature 428, 617–624 (2004).

Keushguerian, V. & Ghaplanyan, I. Carving out a new market niche: Historic world of wines. BIO Web Conf 5, 03002 (2015).

Killian, E. & Ough, C. S. Fermentation Esters — Formation and Retention as Affected by Fermentation Temperature. Am J Enol Vitic 30, 301–305 (1979).

Kleman, I., Rehermann, G., Kwadha, C. A., Witzgall, P. & Becher, P. G. *Hanseniaspora uvarum* Attracts *Drosophila suzukii* (Diptera: Drosophilidae) With High Specificity. J Econ Entomol 115, 999–1007 (2022).

Könen, P. P. & Wüst, M. Analysis of sesquiterpene hydrocarbons in grape berry exocarp (*Vitis vinifera* L.) using in vivo-labeling and comprehensive two-dimensional gas chromatography–mass spectrometry (GC×GC–MS). Beilstein Journal of Organic Chemistry 15, 1945 (2019).

Krainer, F. W. et al. Knockout of an endogenous mannosyltransferase increases the homogeneity of glycoproteins produced in *Pichia pastoris*. Sci Rep 3, 1–13 (2013).

Kruis, A. J. et al. Alcohol Acetyltransferase Eat1 Is Located in Yeast Mitochondria. Appl Environ Microbiol 84, (2018).

Kruis, A. J. et al. Ethyl acetate production by the elusive alcohol acetyltransferase from yeast. Metab Eng 41, 92–101 (2017).

Kruis, A. J. et al. From Eat to trEat: engineering the mitochondrial Eat1 enzyme for enhanced ethyl acetate production in *Escherichia coli*. Biotechnol Biofuels 13, 76 (2020).

Kuanyshev, N., Adamo, G. M., Porro, D. & Branduardi, P. The spoilage yeast *Zygosaccharomyces bailii* : Foe or friend? Yeast 34, 359–370 (2017).

Kurtzman, C. & Robnett, C. Phylogenetic relationships among yeasts of the '*Saccharomyces* complex' determined from multigene sequence analyses. FEMS Yeast Res 3, 417–432 (2003).

Kurtzman, C. P. & Robnett, C. J. Identification and phylogeny of ascomycetous yeasts from analysis of nuclear large subunit (26S) ribosomal DNA partial sequences. Antonie Van Leeuwenhoek 4, 331–371 (1998).

Kurtzman, C. P., Fell, J. W. & Boekhout, T. The Yeasts, a Taxonomic Study - Volume 1. (Elsevier, 2011).

Lachance, M. A. Yeast communities in a natural tequila fermentation. Antonie Van Leeuwenhoek 68, 151–160 (1995).

Lachance, M.-A., Gilbert, D. G. & Starmer, W. T. Yeast communities associated with *Drosophila* species and related flies in an eastern oak-pine forest: A comparison with western communities. J Ind Microbiol 14, 484–494 (1995).

Lage, P. et al. *H. guilliermondii* impacts growth kinetics and metabolic activity of *S. cerevisiae*: the role of initial nitrogen concentration. Int J Food Microbiol 172, 62–9 (2014).

Lage, P., Sampaio-Marques, B., Ludovico, P., Mira, N. P. & Mendes-Ferreira, A. Transcriptomic and chemogenomic analyses unveil the essential role of Com2-regulon in response and tolerance of *Saccharomyces cerevisiae* to stress induced by sulfur dioxide. Microbial Cell 6, 509–523 (2019).

Lam, S. S. T. H. & Howell, K. S. *Drosophila* -associated yeast species in vineyard ecosystems. FEMS Microbiol Lett 362, fnv170 (2015).

Lambrechts, M. G. & Pretorius, I. S. Yeast and its Importance to Wine Aroma - A Review. South African Journal of Enology and Viticulture 21, 97–129 (2000).

Langenberg, A.-K. et al. Glycolytic Functions Are Conserved in the Genome of the Wine Yeast *Hanseniaspora uvarum*, and Pyruvate Kinase Limits Its Capacity for Alcoholic Fermentation. Appl Environ Microbiol 83, (2017).

Large, C., Hanson, N. & Tsouris, A. Genomic stability and adaptation of beer brewing yeasts during serial repitching in the brewery. (2020).

Legras, J.-L. et al. Adaptation of *S. cerevisiae* to Fermented Food Environments Reveals Remarkable Genome Plasticity and the Footprints of Domestication. Mol Biol Evol 35, 1712–1727 (2018).

Legras, J., Merdinoglu, D., Cornuet, J. & Karst, F. Bread, beer and wine: *Saccharomyces cerevisiae* diversity reflects human history. Mol Ecol 16, 2091–2102 (2007).

Li, H. et al. The worlds of wine: Old, new and ancient. Wine Economics and Policy 7, 178–182 (2018).

Li, J. et al. Next-generation metabolic engineering of non-conventional microbial cell factories for carboxylic acid platform chemicals. Biotechnol Adv 43, 107605 (2020).

Li, Y.-Q., Hu, K., Xu, Y.-H., Mei, W.-C. & Tao, Y.-S. Biomass suppression of *Hanseniaspora uvarum* by killer *Saccharomyces cerevisiae* highly increased fruity esters in mixed culture fermentation. LWT 132, 109839 (2020).

Li, Z., Howell, K., Fang, Z. & Zhang, P. Sesquiterpenes in grapes and wines: Occurrence, biosynthesis, functionality, and influence of winemaking processes. Compr Rev Food Sci Food Saf 19, 247–281 (2020).

Lilly, M., Lambrechts, M. G. & Pretorius, I. S. Effect of Increased Yeast Alcohol Acetyltransferase Activity on Flavor Profiles of Wine and Distillates. Appl Environ Microbiol 66, 744–753 (2000).

Lin, J., Massonnet, M. & Cantu, D. The genetic basis of grape and wine aroma. Hortic Res 6, 81 (2019).

Lindegren, C. Yeast genetics: Life cycles, cytology, hybridization, vitamin synthesis, and adaptative enzymes. Bacteriol Rev 9, (1945).

Liti, G. et al. Population genomics of domestic and wild yeasts. Nature 458, 337–341 (2009).

Liu, L. et al. Microbial production of glucosamine and N-acetylglucosamine: advances and perspectives. Appl Microbiol Biotechnol 97, 6149–6158 (2013).

Liu, S. Q. & Pilone, G. J. An overview of formation and roles of acetaldehyde in winemaking with emphasis on microbiological implications. Int J Food Sci Technol 35, 49–61 (2000).

Liu, Z. et al. *Hanseniaspora terricola* sp. nov., an ascomycetous yeast isolated from Tibet. Int J Syst Evol Microbiol 71, (2019).

Lloyd, N., Johnson, D. L. & Herderich, M. J. Metabolomics approaches for resolving and harnessing chemical diversity in grapes, yeast and wine. Aust J Grape Wine Res 21, 723–740 (2015).

Lodder, J. & Kreger-Van Rij, N. J. W. Wladimir I. Kudrjawzew, Die Systematik der Hefen. Übersetzt von R. Wittwer; in deutscher Sprache herausgegeben von R. Dickscheit. Berlin 1960: Akademie-Verlag. 324 S., 11 Abb., 80 Tafeln, 5 Tab. DM 43. —. Z Allg Mikrobiol 2, 77–80 (1962).

López, S., Mateo, J. & Maicas, S. Screening of *Hanseniaspora* Strains for the Production of Enzymes with Potential Interest for Winemaking. Fermentation 2, 1 (2015).

Löser, C., Urit, T. & Bley, T. Perspectives for the biotechnological production of ethyl acetate by yeasts. Appl Microbiol Biotechnol 98, 5397–5415 (2014).

Loureiro, V. & Malfeito-Ferreira, M. Spoilage yeasts in the wine industry. International Journal of Food Microbiology vol. 86 23–50

Ludwig, F. Ueber Alkoholgahrung und Schleimfluss lebender Baume und deren Urheber. Berichte der Deutschen Botanischen Gesellschaf 4, XVII–XXVII (1886).

Luo, Y. et al. Diverse mutational mechanisms cause pathogenic subtelomeric rearrangements. Hum Mol Genet 20, 3769–3778 (2011).

Macías, L. G. et al. Convergent adaptation of *Saccharomyces uvarum* to sulfite, an antimicrobial preservative widely used in human-driven fermentations. PLoS Genet 17, e1009872 (2021).

Manzanares, P., Rojas, V., Genovés, S. & Vallés, S. A preliminary search for anthocyanin-β-D-glucosidase activity in non-*Saccharomyces* wine yeasts. Int J Food Sci Technol 35, 95–103 (2000).

Manzanares, P., Vallés, S. & Viana, F. Non-*Saccharomyces* Yeasts in the Winemaking Process. in Molecular Wine Microbiology (2011). doi:10.1016/B978-0-12-375021-1.10004-9.

Marais, J. Terpenes in the Aroma of Grapes and Wines: A Review. South African Journal of Enology and Viticulture 4, 49–58 (1983).

Marks, V. D. et al. Dynamics of the yeast transcriptome during wine fermentation reveals a novel fermentation stress response. FEMS Yeast Res 8, 35–52 (2008).

Marsit, S. et al. Evolutionary Advantage Conferred by an Eukaryote-to-Eukaryote Gene Transfer Event in Wine Yeasts. Mol Biol Evol 32, 1695–1707 (2015).

Martin, V. et al. De Novo Synthesis of Benzenoid Compounds by the Yeast *Hanseniaspora vineae* Increases the Flavor Diversity of Wines. J Agric Food Chem 64, 4574–4583 (2016).

Martin, V., Jose Valera, M., Medina, K., Boido, E. & Carrau, F. Oenological impact of the *Hanseniaspora*/*Kloeckera* yeast genus on wines — A review. Fermentation vol. 4 (2018).

Marullo, P. et al. Metabolic, organoleptic and transcriptomic impact of *Saccharomyces cerevisiae* genes involved in the biosynthesis of linear and substituted esters. Int J Mol Sci 22, (2021).

Mateo, J. J. & Jiménez, M. Monoterpenes in grape juice and wines. J Chromatogr A 881, 557–567 (2000).

McCubrey, J. A. et al. Roles of the Raf/MEK/ERK pathway in cell growth, malignant transformation and drug resistance. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research 1773, 1263–1284 (2007).

McGovern, P. E. Ancient Wine. (Princeton University Press, Princeton, 2013). doi:10.1515/9781400849536.

McGovern, P. et al. Early Neolithic wine of Georgia in the South Caucasus. Proceedings of the National Academy of Sciences 114, (2017).

Mendes-Ferreira, A., Barbosa, C., Falco, V., Leão, C. & Mendes-Faia, A. 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 (2009).

Meyer, S., Smith, M. & Simione, F. P. Systematics of *Hanseniaspora* Zikes and *Kloeckera* Janke. Antonie Van Leeuwenhoek 44, 79–96 (1978).

Michel, M. et al. Review: Pure non- *Saccharomyces* starter cultures for beer fermentation with a focus on secondary metabolites and practical applications. Journal of the Institute of Brewing 122, 569–587 (2016).

Michlmayr, H. et al. Release of wine monoterpenes from natural precursors by glycosidases from *Oenococcus oeni*. Food Chem 135–334, 80 (2012).

Mille, C. et al. Identification of a new family of genes involved in β-1,2- mannosylation of glycans in *Pichia pastoris* and *Candida albicans*. Journal of Biological Chemistry 283, 9724–9736 (2008).

Miller, M. W. & Phaff, H. J. A comparative study of the apiculate yeasts. Mycopathol Mycol Appl 10, 113–141 (1958).

Mira, N. P. et al. The Genome Sequence of the Highly Acetic Acid-Tolerant *Zygosaccharomyces bailii*-Derived Interspecies Hybrid Strain ISA1307, Isolated From a Sparkling Wine Plant. doi:10.1093/dnares/dst058.

Miyakawa, I., Matsuo, E., Yagi, R., Inai, T. & Morifuku, Y. Isolation of Interspore Bridges from the Budding Yeast *Saccharomycodes ludwigii*. Cytologia (Tokyo) 85, 307–312 (2020).

Miyakawa, I., Nakahara, A. & Ito, K. Morphology of mitochondrial nucleoids, mitochondria, and nuclei during meiosis and sporulation of the yeast *Saccharomycodes ludwigii*. J Gen Appl Microbiol 58, 43–51 (2012).

Miyakawa, I., Yagi, R., Sugihara, K. & Inai, T. Sporulation Culture of Spheroplasts Disturbs the Formation of Normal Spore Pairs in the Yeast *Saccharomycodes ludwigii*. Cytologia (Tokyo) 81, 47–52 (2016).

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

Montalvo, F. F., García-Alcaraz, J. L., Cámara, E. M., Jiménez-Macías, E. & Blanco-Fernández, J. Environmental impact of wine fermentation in steel and concrete tanks. J Clean Prod 278, 123602 (2021).

Monteiro, L. M. O. et al. Efficient hydrolysis of wine and grape juice anthocyanins by *Malbranchea pulchella* β-glucosidase immobilized on MANAE-agarose and ConA-Sepharose supports. Int J Biol Macromol 136, 1133–1141 (2019).

Morano, K. A., Grant, C. M. & Moye-Rowley, W. S. The Response to Heat Shock and Oxidative Stress in *Saccharomyces cerevisiae*. Genetics 190, 1157–1195 (2012).

Morata, A. et al. Applications of *Metschnikowia pulcherrima* in Wine Biotechnology. Fermentation 2019, Vol. 5, Page 63 5, 63 (2019).

Moreira, N. et al. Volatile compounds contribution of *Hanseniaspora guilliermondii* and *Hanseniaspora uvarum* during red wine vinifications. Food Control 22, 662–667 (2011).

Moreira, N., de Pinho, P. G., Santos, C. & Vasconcelos, I. Relationship between nitrogen content in grapes and volatiles, namely heavy sulphur compounds, in wines. Food Chem 126, 1599–1607 (2011).

Moreira, N., Mendes, F., Guedes de Pinho, P., Hogg, T. & Vasconcelos, I. Heavy sulphur compounds, higher alcohols and esters production profile of *Hanseniaspora uvarum* and *Hanseniaspora guilliermondii* grown as pure and mixed cultures in grape must. Int J Food Microbiol 124, 231–238 (2008).

Moreno, J. & Peinado, R. Must Aromas. in Enological Chemistry 23–39 (Elsevier, 2012).

Moreno, J. & Peinado, R. Sugars. in Enological Chemistry 77–93 (Elsevier, Córdoba, 2012).

Mosquera, Marta. E. G. et al. Terpenes and Terpenoids: Building Blocks to Produce Biopolymers. Sustainable Chemistry 2021, Vol. 2, Pages 467-492 2, 467–492 (2021).

Mozzachiodi, S., Krogerus, K., Gibson, B., Nicolas, A. & Liti, G. Unlocking the functional potential of polyploid yeasts. Nat Commun 13, 2580 (2022).

Murakami, H., Mu, X. & Keeney, S. How do small chromosomes know they are small? Maximizing meiotic break formation on the shortest yeast chromosomes. Curr Genet 67, 431–437 (2021).

Nadai, C., Treu, L., Campanaro, S., Giacomini, A. & Corich, V. Different mechanisms of resistance modulate sulfite tolerance in wine yeasts. Appl Microbiol Biotechnol 100, 797–813 (2016).

Nakase, T. & Komagata, K. Significance of DNA base composition in the classification of yeast genera *Hanseniaspora* and *Kloeckera*. J. Gen. Appl. Microbiol 16, 241–250 (1970).

Nardi, T. & Romano, P. The changing role of Women in Food Microbiology: the case history of wine microbiologists in Italy. Front Microbiol 14, (2023).

Nikfardjam, M., May, B. & Tschiersch, C. Analysis of 4-vinylphenol and 4-vinylguaiacol in wines from the Wurttemberg region (Germany). Mitteilungen Klosterneuburg 84–89 (2009).

Ninkuu, V. et al. Biochemistry of Terpenes and Recent Advances in Plant Protection. Int J Mol Sci 22, 5710 (2021).

Nissen, P., Nielsen, D. & Arneborg, N. Viable *Saccharomyces* cerevisiae cells at high concentrations cause early growth arrest of non-*Saccharomyces* yeasts in mixed cultures by a cell–cell contact-mediated mechanism. Yeast 20, 331–341 (2003).

Novak, S., Zechner-Krpan, V. & Mari, V. Maltose Transport and Metabolism in *S. cerevisiae*. Food Technol. Biotechnol 42, 213–218 (2004).

Novo, M. et al. Eukaryote-to-eukaryote gene transfer events revealed by the genome sequence of the wine yeast *Saccharomyces cerevisiae* EC1118. Proceedings of the National Academy of Sciences 106, 16333–16338 (2009).

Novotný, Doležalová, L. & Lieblová, J. Dimorphic growth and lipase production in lipolytic yeasts. Folia Microbiologica: Official Journal of the Institute of Microbiology, Academy of Sciences of the Czech Republic 39, 71–73 (1994).

Nykänen, L. & Nykänen, I. Production of esters by different yeast strains in sugar fermentations. Journal of the Institute of Brewing 83, 30–31 (1977).

Ouoba, L. I. I. et al. *Hanseniaspora jakobsenii* sp. nov., a yeast isolated from Bandji, a traditional palm wine of *Borassus akeassii*. Int J Syst Evol Microbiol 65, 3576–3579 (2015).

Outreville, J.-F. Wine Consumption and Religions: A Research Note. Beverages 7, 70 (2021).

Padilla, B., Gil, J. V. & Manzanares, P. Past and future of non-*Saccharomyces* yeasts: From spoilage microorganisms to biotechnological tools for improving wine aroma complexity. Front Microbiol 7, 1–20 (2016).

Papaioannou, I. A. et al. Sex without crossing over in the yeast *Saccharomycodes ludwigii*. Genome Biol 22, 1–28 (2021).

Park, H. & Bakalinsky, A. T. *SSU1* mediates sulphite efflux in *Saccharomyces cerevisiae*. Yeast 16, 881–888 (2000).

Passoth, V., Zimmermann, M. & Klinner, U. Peculiarities of the regulation of fermentation and respiration in the crabtree-negative, xylose-fermenting yeast *Pichia stipitis*. Applied Biochemistry and Biotechnology - Part A Enzyme Engineering and Biotechnology 57–58, 201–212 (1996).

Pavelka, N. et al. Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast. Nature 468, 321–325 (2010).

Perestrelo, R., Silva, C., Gonçalves, C., Castillo, M. & Câmara, J. S. An Approach of the Madeira Wine Chemistry. Beverages 2020, Vol. 6, Page 12 6, 12 (2020).

Pérez-Ortín, J. E., Querol, A., Puig, S. & Barrio, E. Molecular Characterization of a Chromosomal Rearrangement Involved in the Adaptive Evolution of Yeast Strains. Genome Res 12, 1533–1539 (2002).

Peter, J. et al. Genome evolution across 1,011 *Saccharomyces* cerevisiae isolates. Nature 556, 339–344 (2018).

Phaff, H. J., Miller, M. W. & Em, M. The Life of Yeasts. (Harvard University Press: Cambridge, UK, 1998).

Pichler, A. et al. Aroma Profile of Merlot Red Wine Stored in Stainless-Steel Tanks and Wooden Barrels with Different Toasting Methods. Foods 13, 45 (2023).

Pietrafesa, A., Capece, A., Pietrafesa, R., Bely, M. & Romano, P. *Saccharomyces cerevisiae* and *Hanseniaspora uvarum* mixed starter cultures: Influence of microbial/physical interactions on wine characteristics. Yeast 37, 609–621 (2020).

Pilap, W., Thanonkeo, S., Klanrit, P. & Thanonkeo, P. The potential of multistress tolerant yeast, *Saccharomycodes ludwigii*, for second-generation bioethanol production. Sci Rep 12, 22062 (2022).

Plack, N. Liberty, Equality and Taxation: Wine in the French Revolution. Soc Hist Alcohol Drugs 26, 5–22 (2012).

Poláková, S. et al. Formation of new chromosomes as a virulence mechanism in yeast *Candida glabrata*. Proceedings of the National Academy of Sciences 106, 2688–2693 (2009).

Pontes, A., Hutzler, M., Brito, P. H. & Sampaio, J. P. Revisiting the Taxonomic Synonyms and Populations of *Saccharomyces cerevisiae*—Phylogeny, Phenotypes, Ecology and Domestication. Microorganisms 8, 903 (2020).

Prestianni, R. et al. Use of sequentially inoculation of *Saccharomyces cerevisiae* and *Hanseniaspora uvarum* strains isolated from honey by-products to improve and stabilize the quality of mead produced in Sicily. Food Microbiol 107, 104064 (2022).

Pretorius, I. S. Tailoring wine yeast for the new millennium: novel approaches to the ancient art of winemaking. Yeast 16, 675–729 (2000).

Pryde, F. E., Gorham, H. C. & Louis, E. J. Chromosome ends: all the same under their caps. Curr Opin Genet Dev 7, 822–828 (1997).

Ram, Y. & Hadany, L. The evolution of stress-induced hypermutation in asexual populations. Evolution (N Y) 66, 2315–2328 (2012).

Ravasio, D., Wendland, J. & Walther, A. Major contribution of the Ehrlich pathway for 2-phenylethanol/rose flavor production in *Ashbya gossypii*. FEMS Yeast Res. 14, 833–844 (2014).

Reynolds, A. G. 11 - Viticultural and vineyard management practices and their effects on grape and wine quality. in Woodhead Publishing Series in Food Science, Technology and Nutrition, Managing Wine Quality 443–539 (Woodhead Publishing, 2022).

Ribéreau-Gayon, P., Glories, Y., Maujean, A. & Dubourdieu, D. Varietal Aroma. Handbook of Enology 205–230 (2006).

Riley, R. et al. Comparative genomics of biotechnologically important yeasts. Proceedings of the National Academy of Sciences 113, 9882–9887 (2016).

Robinson, A. L. et al. Origins of Grape and Wine Aroma. Part 1. Chemical Components and Viticultural Impacts. Am J Enol Vitic 65, 1–24 (2014).

Robinson, J. The Oxford Companion to Wine. (Oxford University Press, 1994).

Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR : a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140 (2010).

Rocchetti, G., Gatti, M., Bavaresco, L. & Lucini, L. Untargeted metabolomics to investigate the phenolic composition of Chardonnay wines from different origins. Journal of Food Composition and Analysis 71, 87–93 (2018).

Rodríguez-Bustamante, E. & Sánchez, S. Microbial production of C13-norisoprenoids and other aroma compounds via carotenoid cleavage. Crit Rev Microbiol 33, 211–230 (2007).

Rojas, V., Gil, J. V, Piñaga, F. & Manzanares, P. Studies on acetate ester production by non-*Saccharomyces* wine yeasts. Int J Food Microbiol 70, 283–289 (2001).

Roland, A. Influence des phénomènes d'oxydation lors de l'élaboration des moûts sur la qualité aromatique des vins de Melon B. et de Sauvignon Blanc en Val de Loire. undefined (2010).

Roller, S. & Covill, N. The antifungal properties of chitosan in laboratory media and apple juice. Int J Food Microbiol 47, 67–77 (1999).

Romano, P., Brandolini, V., Ansaloni, C. & Menziani, E. The production of 2,3-butanediol as a differentiating character in wine yeasts. World J Microbiol Biotechnol 14, 649–653 (1998).

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

Romano, P., Marchese, R., Laurita, C., Saleano, G. & Turbanti, L. Biotechnological suitability of *Saccharomycodes ludwigii* for fermented beverages. World J Microbiol Biotechnol 15, 451–454 (1999).

Romano, P., Suzzi, G., Turbanti, L. & Polsinelli, M. Acetaldehyde production in *Saccharomyces cerevisiae* wine yeasts. FEMS Microbiol Lett 118, 213–218 (1994).

Rosi, I., Vinella, M. & Domizio, P. Characterization of Beta-Glucosidase Activity in Yeasts of Enological Origin. Journal of Applied Bacteriology 77, 519–527 (1994).

Rossi, L., Foschi, M., Biancolillo, A., Maggi, M. A. & D'Archivio, A. A. Optimization of HS-SPME-GC/MS Analysis of Wine Volatiles Supported by Chemometrics for the Aroma Profiling of Trebbiano d'Abruzzo and Pecorino White Wines Produced in Abruzzo (Italy). Molecules 28, 1534 (2023).

Roullier-Gall, C., David, V., Hemmler, D., Schmitt-Kopplin, P. & Alexandre, H. Exploring yeast interactions through metabolic profiling. Sci Rep 10, 6073 (2020).

Roullier-Gall, C., Witting, M., Gougeon, R. D. & Schmitt-Kopplin, P. High precision mass measurements for wine metabolomics. Front Chem 2, (2014).

Ruiz, J. et al. Effects on varietal aromas during wine making: a review of the impact of varietal aromas on the flavor of wine. Appl Microbiol Biotechnol 103, 7425–7450 (2019).

Ryan, A. P. et al. Genome Analysis of a Newly Discovered Yeast Species, *Hanseniaspora menglaensis*. Journal of Fungi 10, 180 (2024).

Ryan, O. et al. Global Gene Deletion Analysis Exploring Yeast Filamentous Growth. Science (1979) 337, 1353–1356 (2012).

Saerens, S. M. G. et al. Parameters Affecting Ethyl Ester Production by *Saccharomyces cerevisiae* during Fermentation. Appl Environ Microbiol 74, 454–461 (2008).

Salazar, S. B. et al. Comparative genomic and transcriptomic analyses unveil novel features of azole resistance and adaptation to the human host in *Candida glabrata*. FEMS Yeast Res 18, (2018).

Sánchez-Acevedo, E., Lopez, R. & Ferreira, V. Kinetics of aroma formation from grape-derived precursors: Temperature effects and predictive potential. Food Chem 438, 137935 (2024).

Saubin, M. et al. Investigation of Genetic Relationships Between *Hanseniaspora* Species Found in Grape Musts Revealed Interspecific Hybrids With Dynamic Genome Structures. Front Microbiol 10, (2020).

Schubert, M., Hansson, B. S. & Sachse, S. The banana code—natural blend processing in the olfactory circuitry of *Drosophila melanogaster*. Front Physiol 5, (2014).

Schuring, R. Wine Emmbeddings and a Wine Recommender. Towards Data Science (2019).

Scopel, E. F. C., Hose, J., Bensasson, D. & Gasch, A. P. Genetic variation in aneuploidy prevalence and tolerance across *Saccharomyces* cerevisiae lineages. Genetics 217, (2021).

Seixas, I. et al. Genome sequence of the non-conventional wine yeast *Hanseniaspora guilliermondii* UTAD222 unveils relevant traits of this species and of the *Hanseniaspora* genus in the context of wine fermentation. DNA Research 26, 67–83 (2019).

Seixas, I., Santos, D., Vasconcelos, I., Mira, N. P. & Mendes-Ferreira, A. Insights into the transcriptional regulation of poorly characterized alcohol acetyltransferase-encoding genes (HgAATs) shed light into the production of acetate esters in the wine yeast *Hanseniaspora guilliermondii*. FEMS Yeast Res 23, (2023).

Sellick, C. A., Campbell, R. N. & Reece, R. J. Chapter 3 Galactose Metabolism in Yeast-Structure and Regulation of the Leloir Pathway Enzymes and the Genes Encoding Them. Int Rev Cell Mol Biol 269, 111–150 (2008).

Selmecki, A. M. et al. Polyploidy can drive rapid adaptation in yeast. Nature 519, 349–352 (2015).

Sharma, E., Anand, G. & Kapoor, R. Terpenoids in plant and arbuscular mycorrhiza-reinforced defence against herbivorous insects. Ann Bot 119, 791–801 (2017).

Sheltzer, J. M. & Amon, A. The aneuploidy paradox: costs and benefits of an incorrect karyotype. Trends in Genetics 27, 446–453 (2011).

Sheltzer, J. M. et al. Aneuploidy Drives Genomic Instability in Yeast. Science (1979) 333, 1026–1030 (2011).

Shen, X.-X. et al. Tempo and Mode of Genome Evolution in the Budding Yeast Subphylum. Cell 175, 1533-1545.e20 (2018).

Sileoni, V., Maranghi, S., De Francesco, G., Perretti, G. & Marconi, O. Flavour Stability of a Cold-Stored Unpasteurized Low-Alcohol Beer Produced by *Saccharomycodes ludwigii*. Food Bioproc Tech 16, 2471–2482 (2023).

Simmons, R. B. & Ahearn, D. G. Ascospore Ornamentation in *Saccharomycodes ludwigii*. Mycologia 77, 660–662 (1985).

Singh, B. & Sharma, R. A. Plant terpenes: defense responses, phylogenetic analysis, regulation and clinical applications. 3 Biotech 5, 129–151 (2015).

Sipiczki, M., Miklos, I., Leveleki, L. & Antunovics, Z. Genetic and Chromosomal Stability of Wine Yeasts. in Food Microbiology Protocols 273–281 (Humana Press, New Jersey, 2003).

Slaghenaufi, D. & Ugliano, M. Norisoprenoids, sesquiterpenes and terpenoids content of Valpolicella wines during aging: Investigating aroma potential in relationship to evolution of tobacco and balsamic aroma in aged wine. Front Chem 6, 66 (2018).

Smit, S. J., Vivier, M. A. & Young, P. R. Linking terpene synthases to sesquiterpene metabolism in grapevine flowers. Front Plant Sci 10, 177 (2019).

Smukowski Heil, C. S. et al. Loss of Heterozygosity Drives Adaptation in Hybrid Yeast. Mol Biol Evol 34, 1596–1612 (2017).

Snoek, T., Verstrepen, K. J. & Voordeckers, K. How do yeast cells become tolerant to high ethanol concentrations? Curr Genet 62, 475–480 (2016).

Soares, E. V. Flocculation in *Saccharomyces* cerevisiae: a review. J Appl Microbiol 110, 1–18 (2011).

Song, M., Fuentes, C., Loos, A. & Tomasino, E. Free Monoterpene Isomer Profiles of *Vitis Vinifera* L. cv. White Wines. Foods 7, (2018).

Spencer, J. F. & Gorin, P. A. Mannose-containing polysaccharides of the apiculate yeasts *Nadsonia*, *Hanseniaspora*, *Kloeckera*, and *Sacharomycodes*, and their use as an aid in classification. J Bacteriol 96, 180–3 (1968).

Spillman, P. J., Sefton, M. A. & Gawel, R. The contribution of volatile compounds derived during oak barrel maturation to the aroma of a Chardonnay and Cabernet Sauvignon wine. Aust J Grape Wine Res 10, 227–235 (2004).

Stanley, D., Bandara, A., Fraser, S., Chambers, P. J. & Stanley, G. A. The ethanol stress response and ethanol tolerance of *Saccharomyces* cerevisiae. J Appl Microbiol 109, 13–24 (2010).

Steensels, J. et al. Brettanomyces yeasts - From spoilage organisms to valuable contributors to industrial fermentations. International Journal of Food Microbiology vol. 206 24–38 (2015).

Steensels, J. et al. Improving industrial yeast strains: Exploiting natural and artificial diversity. FEMS Microbiol Rev 38, 947–995 (2014).

Steenwyk, J. L. et al. Extensive loss of cell-cycle and DNA repair genes in an ancient lineage of bipolar budding yeasts. PLoS Biol 17, e3000255 (2019).

Stefanini, I. Yeast-insect associations: It takes guts. Yeast 35, 315–330 (2018).

Stelling-Decker, N. M. Die Hefesammlung Des , , Centraal-Bureau Voor Schimmelcultures". (Amsterdam, 1931).

Sternes, P. R., Lee, D., Kutyna, D. R. & Borneman, A. R. Genome Sequences of Three Species of *Hanseniaspora* Isolated from Spontaneous Wine Fermentations. Genome Announc 4, e01287-16 (2016).

Stratford, M. & Rose, A. H. Transport of sulphur dioxide by *Saccharomyces* cerevisiae. J Gen Microbiol 132, 1–6 (1986).

Stratford, M. PRESERVATIVES | Traditional Preservatives – Organic Acids. in Encyclopedia of Food Microbiology 1729–1737 (Elsevier, 1999). doi:10.1006/rwfm.1999.2035.

STRATFORD, M., MORGAN, P. & ROSE, A. H. Sulphur Dioxide Resistance in *Saccharomyces cerevisiae* and *Saccharomycodes ludwigii*. Microbiology (N Y) 133, 2173–2179 (1987).

Stringini, M., Comitini, F., Taccari, M. & Ciani, M. Yeast diversity during tapping and fermentation of palm wine from Cameroon. Food Microbiol 26, 415–420 (2009).

Styger, G., Prior, B. & Bauer, F. F. Wine flavor and aroma. J Ind Microbiol Biotechnol 38, 1145–1159 (2011).

Swiegers, J. H. & Pretorius, I. S. Modulation of volatile sulfur compounds by wine yeast. Appl Microbiol Biotechnol 74, 954–960 (2007).

Swiegers, J. H. & Pretorius, I. S. Yeast modulation of wine flavor. Advances in Applied Microbiology vol. 57 131–175(2005).

Tamura, K., Stecher, G. & Kumar, S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Mol Biol Evol 38, 3022–3027 (2021).

Tavares, M. J. et al. Genome Sequence of the Wine Yeast *Saccharomycodes ludwigii* UTAD17. Microbiol Resour Announc 7, (2018).

Tavares, M. J., Güldener, U., Mendes-Ferreira, A. & Mira, N. P. Genome sequencing, annotation and exploration of the SO2-tolerant non-conventional yeast *Saccharomycodes ludwigii*. BMC Genomics 22, 1–15 (2021).

Teixeira, M. C., Raposo, L. R., Mira, N. P., Lourenço, A. B. & Sá-Correia, I. Genome-wide identification of *Saccharomyces cerevisiae* genes required for maximal tolerance to ethanol. Appl Environ Microbiol 75, 5761–5772 (2009).

Teixeira, M. C., Raposo, L. R., Palma, M. & Sá-Correia, I. Identification of genes required for maximal tolerance to high-glucose concentrations, as those present in industrial alcoholic fermentation media, through a chemogenomics approach. OMICS 14, 201–210 (2010).

Thomas, D. S. Yeasts as Spoilage Organisms in Beverages. The Yeasts (Academic Press, New York, USA, 1993).

Torija, M. J. et al. 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 (2003).

Torres, E. M. et al. Effects of Aneuploidy on Cellular Physiology and Cell Division in Haploid Yeast. Science (1979) 317, 916–924 (2007).

Tosi, E., Azzolini, M., Guzzo, F. & Zapparoli, G. Evidence of different fermentation behaviours of two indigenous strains of *Saccharomyces cerevisiae* and *Saccharomyces uvarum* isolated from Amarone wine. J Appl Microbiol 107, (2009).

Tristezza, M. et al. The Oenological Potential of *Hanseniaspora uvarum* in Simultaneous and Sequential Co-fermentation with *Saccharomyces cerevisiae* for Industrial Wine Production. Front Microbiol 7, 670 (2016).

Tufariello, M. et al. Influence of Non-*Saccharomyces* on Wine Chemistry: A Focus on Aroma-Related Compounds. Molecules 26, (2021).

Tyczewska, A. & Grzywacz, K. tRNA-derived fragments as new players in regulatory processes in yeast. Yeast 40, 283–289 (2023).

Tzachristas, A., Dasenaki, M. E., Aalizadeh, R., Thomaidis, N. S. & Proestos, C. Development of a Wine Metabolomics Approach for the Authenticity Assessment of Selected Greek Red Wines. Molecules 26, 2837 (2021).

Unger, R. W. Beer in the Middle Ages and Renaissance. (2004).

Vale-Silva, L., Beaudoing, E., Tran, V. D. T. & Sanglard, D. Comparative Genomics of Two Sequential *Candida glabrata* Clinical Isolates. G3 Genes|Genomes|Genetics 7, 2413–2426 (2017).

Valera, M. J., Boido, E., Dellacassa, E. & Carrau, F. Comparison of the Glycolytic and Alcoholic Fermentation Pathways of *Hanseniaspora vineae* with *Saccharomyces cerevisiae* Wine Yeasts. Fermentation 6, 78 (2020).

Valera, M. J., Olivera, V., Boido, E., Dellacassa, E. & Carrau, F. Wine Aroma Characterization of the Two Main Fermentation Yeast Species of the Apiculate Genus *Hanseniaspora*. Fermentation 2021, Vol. 7, Page 162 7, 162 (2021).

van Voorst, F., Houghton-Larsen, J., Jønson, L., Kielland-Brandt, M. C. & Brandt, A. Genome-wide identification of genes required for growth of *Saccharomyces cerevisiae* under ethanol stress. Yeast 23, 351–359 (2006).

van Wyk, N., Badura, J., von Wallbrunn, C. & Pretorius, I. S. Exploring future applications of the apiculate yeast *Hanseniaspora*. Crit Rev Biotechnol 1–20 (2023)

Varela, C., Bartel, C., Roach, M., Borneman, A. & Curtin, C. *Brettanomyces bruxellensis SSU1* Haplotypes Confer Different Levels of Sulfite Tolerance When Expressed in a *Saccharomyces cerevisiae SSU1* Null Mutant. Appl Environ Microbiol 85, (2019).

Vejarano, R. *Saccharomycodes ludwigii*, control and potential uses in winemaking processes. Fermentation, 4 (2018).

Venkatesh, A., Murray, A. L., Coughlan, A. Y. & Wolfe, K. H. Giant GAL gene clusters for the melibiose-galactose pathway in *Torulaspora*. Yeast 38, 117–126 (2021).

Verstrepen, K. et al. The alcohol acetyl transferase gene is a target of the cAMP/PKA and FGM nutrient-signalling pathways. FEMS Yeast Res 4, 285–296 (2003).

Verstrepen, K. J., Chambers, P. J. & Pretorius, I. S. The Development of Superior Yeast Strains for the Food and Beverage Industries: Challenges, Opportunities and Potential Benefits. in Yeasts in Food and Beverages 399–444 (Springer Berlin Heidelberg, Berlin, Heidelberg).

Verzera, A. et al. Varietal Aromas of Fortified Wines from Different Moscato Var. (*Vitis vinifera* L.) under the Same Pedoclimatic Conditions. Foods 10, 2549 (2021).

Viana, F., Gil, J. V., Genovés, S., Vallés, S. & Manzanares, P. Rational selection of non-*Saccharomyces* wine yeasts for mixed starters based on ester formation and enological traits. Food Microbiol 25, 778–785 (2008).

Viana, F., Gil, J. V., Vallés, S. & Manzanares, P. Increasing the levels of 2-phenylethyl acetate in wine through the use of a mixed culture of *Hanseniaspora osmophila* and *Saccharomyces cerevisiae*. Int J Food Microbiol 135, 68–74 (2009).

Vidal, S. et al. The mouth-feel properties of polysaccharides and anthocyanins in a wine like medium. Food Chem 85, 519–525 (2004).

Vilela, A. Modulating wine pleasantness throughout wine-yeast co-inoculation or sequential inoculation. Fermentation 6, (2020).

Vinholes, J., Coimbra, M. A. & Rocha, S. M. Rapid tool for assessment of C13 norisoprenoids in wines. J Chromatogr A 1216, 8398–8403 (2009).

Voigt, M. M. Hasanlu, Volume I: Hajji Firuz Tepe, Iran--The Neolithic Settlement (University Museum Monograph). (1983).

Walker, M. E. et al. Genome-wide identification of the Fermentome; genes required for successful and timely completion of wine-like fermentation by *Saccharomyces cerevisiae*. BMC Genomics 15, 552 (2014).

Wang, C., Mas, A. & Esteve-Zarzoso, B. Interaction between *Hanseniaspora uvarum* and *Saccharomyces cerevisiae* during alcoholic fermentation. Int J Food Microbiol 206, 67–74 (2015).

Wang, J. et al. Metabolomics and flavor diversity in Cabernet Sauvignon wines fermented by various origins of *Hanseniaspora uvarum* in the presence and absence of *Saccharomyces cerevisiae*. LWT 203, 116396 (2024).

Wang, L., Groenewald, M., Wang, Q.-M. & Boekhout, T. Reclassification of *Saccharomycodes sinensis*, Proposal of *Yueomyces sinensis* gen. nov., comb. nov. within *Saccharomycetaceaee* (Saccharomycetales, Saccharomycotina). PLoS One 10, e0136987 (2015).

Wang, M.-M. et al. Intraspecific nucleotide divergence in *Saccharomycodes ludwigii*, and proposal of *Saccharomycodes pseudoludwigii* sp. nov, a new apiculate yeast isolated from China. Antonie Van Leeuwenhoek 114, 553–559 (2021).

Wang, Q., Quan, S. & Xiao, H. Towards efficient terpenoid biosynthesis: manipulating IPP and DMAPP supply. Bioresources and Bioprocessing 2019 6:1 6, 1–13 (2019).

Wanke, V., Vavassori, M., Thevelein, J. M., Tortora, P. & Vanoni, M. Regulation of maltose utilization in *Saccharomyces cerevisiae* by genes of the RAS/protein kinase A pathway. FEBS Lett 402, 251–255 (1997).

Wightman, R. & Meacock, P. A. The THI5 gene family of *Saccharomyces cerevisiae*: Distribution of homologues among the hemiascomycetes and functional redundancy in the aerobic biosynthesis of thiamin from pyridoxine. Microbiology vol. 149 1447–1460 (2003).

Wikström, N., Savolainen, V. & Chase, M. W. Evolution of the angiosperms: calibrating the family tree. Proc R Soc Lond B Biol Sci 268, 2211–2220 (2001).

Willaert, R. G., Kayacan, Y. & Devreese, B. The Flo Adhesin Family. Pathogens 10, 1397 (2021).

Winge, O. & Laustsen, O. On two types of spore germination, and on genetic segregations in *Saccharomyces*, demonstrated through single-spore cultures. Cr Trav Lab Carlsberg Ser Physiol 22, 99–117 (1942).

Winge, O. & Laustsen, O. *Saccharomycodes ludwigii* Hansen, a balanced heterozygote. Comptes Rendus Lab Carlsberg Ser Physiol. 22, 357–379 (1939).

Winge, O. The segregation in the ascus of *Saccharomycodes ludwigii*. Cr Trav Lab Carlsberg Ser Physiol 24, 223–226 (1946).

Wolfe, K. H. & Shields, D. C. Molecular evidence for an ancient duplication of the entire yeast genome. Nature 387, 708–713 (1997).

Wolfe, K. H. et al. Clade- and species-specific features of genome evolution in the *Saccharomycetaceaee*. FEMS Yeast Research, 15 (2015).

Wood, C. et al. From Wine to Pepper: Rotundone, an Obscure Sesquiterpene, Is a Potent Spicy Aroma Compound. J Agric Food Chem 56, 3738–3744 (2008).

Xia, J., Psychogios, N., Young, N., Wishart, D.S., MetaboAnalyst: a webserver for metabolomic data analysis and interpretation. Nucl. Acids Res. 37, W652-660 (2009)

Xie, Y. R., Castro, D. C., Rubakhin, S. S., Sweedler, J. V. & Lam, F. Enhancing the Throughput of FT Mass Spectrometry Imaging Using Joint Compressed Sensing and Subspace Modeling. Anal Chem 94, 5335–5343 (2022).

Yabe, T. et al. HKR1 encodes a cell surface protein that regulates both cell wall beta-glucan synthesis and budding pattern in the yeast *Saccharomyces cerevisiae*. J Bacteriol 178, 477–483 (1996).

Yamada, Y., Maeda, K. & Banno, I. The phylogenetic relationships of the Q6-equipped species in the teleomorphic apiculate yeast genera *Hanseniaspora*, *Nadsonia* and *Saccharomycodes* based on the partial sequences of 18S and 26S ribosommal ribonucleic acids. Journal of General Applied Microbiology 38, 585–596 (1992).

Yamazaki, T. & Oshima, Y. *Saccharomycodes ludwigii* has seven chromosomes. Yeast 12, 237–240 (1996).

Yamazaki, T., Nakagawa, Y., Hayakawa, M. & Iimura, Y. Phylogenetic position of the yeast strain *Saccharomycodes sinsensis* IFO10111T. J Gen Appl Microbiol 51, 35–39 (2005).

Yamazaki, T., Ohara, Y. & Oshima, Y. Rare occurrence of tetratype tetrads in *Saccharomycodes ludwigii*. J Bacteriol 125, 461–466 (1976).

Yang, D., Zhang, M., Su, C., Dong, B. & Lu, Y. *Candida albicans* exploits N-acetylglucosamine as a gut signal to establish the balance between commensalism and pathogenesis. Nat Commun 14, 3796 (2023).

Yang, F. et al. The fitness costs and benefits of trisomy of each *Candida albicans* chromosome. Genetics 218, (2021).

Yoshikawa, K. et al. Comprehensive phenotypic analysis for identification of genes affecting growth under ethanol stress in *Saccharomyces cerevisiae*. FEMS Yeast Res 9, 32–44 (2009).

Yuasa, N., Nakagawa, Y., Hayakawa, M. & Iimura, Y. Distribution of the sulfite resistance gene *SSU1*-R and the variation in its promoter region in wine yeasts. J Biosci Bioeng 98, 394–397 (2004).

Zakharov, I. A. Intratetrad mating as the driving force behind the formation of sex chromosomes in fungi. Trends in Genetics and Evolution 6, (2023).

Zara, G. & Nardi, T. Yeast Metabolism and Its Exploitation in Emerging Winemaking Trends: From Sulfite Tolerance to Sulfite Reduction. Fermentation 7, 57 (2021).

Zerbib, M., Cazals, G., Ducasse, M. A., Enjalbal, C. & Saucier, C. Evolution of flavanol glycosides during red grape fermentation. Molecules 23, 3300 (2018).

Zhang, B.-Q., Shen, J.-Y., Duan, C.-Q. & Yan, G.-L. Use of Indigenous *Hanseniaspora vineae* and *Metschnikowia pulcherrima* Co-fermentation With *Saccharomyces cerevisiae* to Improve the Aroma Diversity of Vidal Blanc Icewine. Front Microbiol 9, (2018).

Zhang, P., Zhang, R., Sirisena, S., Gan, R. & Fang, Z. Beta-glucosidase activity of wine yeasts and its impacts on wine volatiles and phenolics: A mini-review. Food Microbiol 100, 103859 (2021).

Zhang, Z. et al. Wine aroma modification by *Hanseniaspora uvarum*: A multiple-step strategy for screening potential mixed starters. Food Chem X 20, 100930 (2023).

Zhu, J., Pavelka, N., Bradford, W. D., Rancati, G. & Li, R. Karyotypic Determinants of Chromosome Instability in Aneuploid Budding Yeast. PLoS Genet 8, e1002719 (2012).

Zhu, Y. O., Sherlock, G. & Petrov, D. A. Whole Genome Analysis of 132 Clinical *Saccharomyces cerevisiae* Strains Reveals Extensive Ploidy Variation. G3 Genes|Genomes|Genetics 6, 2421–2434 (2016).

Zhuang, S. et al. Production of medium-chain volatile flavour esters in *Pichia pastoris* whole-cell biocatalysts with extracellular expression of *Saccharomyces* cerevisiae acyl-CoA:ethanol O-acyltransferase Eht1 or Eeb1. Springerplus 4, 467 (2015)

Appendix

Hyperlinks for the Appendix Tables and Figures

Appendix Tables

Chapter II

Appendix Table II.1 - List of *Sd. ludwigii* UTAD17 genes uncovered in the PacBio assembly but not identified in the Illumina assembly . Genes considered hypothetical, based on the lack of a hit upon BLAST against. the UNIPROT database, are highlighted in grey.

Appendix Table II.2 - Number of CDSs identified in the 20 contigs obtained upon assembly of the PacBio reads

Appendix Table II.3 - Number of genes classified in the different functional COG categories in *S. cerevisiae* and in *Sd. ludwigii* UTAD17

Appendix Table II.4- Orthologues found in the genome of *Sd. ludwigii* UTAD17 of the ohnologues pairs identified in the genome of *S. cerevisiae* (as indicated by Byrne and Wolfe, Genome Res.15: 1456, 2005).

Appendix Table II.5 - *Sd. ludwigii* UTAD17 genes for which we could not find orthologues in the genome of the *Saccharomycodaceae* species *Hanseniaspora guilliermondii* and *Hanseniaspora uvarum* nor in the genome of the *Saccharomycetaceae* species *Lachancea fermentati*, *T. delbruecki* and *Saccharomyces cerevisiae*. Genes highlighted in grey correspond to those whose best hit found in the UNIPROT database was identified based on lower degrees of homology, as assessed by an e-value parameter below e-20

Appendix Table II.6 - Functional annotation of the main *Sd. ludwigii* UTAD17 carbon and nitrogen metabolic pathways, also being indicated the corresponding orthologues (it is provided the UNIPROT accession code for each orthologue) present in the *Saccharomycodaceae* species *H. guilliermondii* and *H. uvarum* and in the *Saccharomycetaceae* species *S. cerevisiae* EC1118 and S288c, *L. fermentati* and *T. delbruecki*.

Appendix Table II.7- List of *Sd. ludwigii* UTAD17 genes predicted to be involved in pathways leading to the formation of aroma compounds. The closest orthologue identified in *S. cerevisiae* are also indicated

Appendix Table II.8 - Genes coding for beta-mannosyltransferases found in the genome of *Sd. ludwigii* UTAD17 and their orthologues in other *Saccharomycodaceae* and *Saccharomycetaceae* species used in the comparative proteomic analysis, and in *C. albicans* and *P. pastoris*.

Appendix Table II.9 - Identification in the predicted *Sd. ludwigii* UTAD17 ORFeome of orthologues for the genes described to mediate *S. cerevisiae* tolerance to high concentrations of ethanol, glucose and to vinification conditions (also known as the fermentome). The datasets of the *S. cerevisiae* genes were retrieved from Teixeira et al., (2009), Teixeira et al., (2010) and Walker et al., (2014). The genes shaded in grey correspond to genes identified in *S. cerevisiae* but for which we could not identify an orthologue in *Sd. ludwigii* UTAD17.

Appendix Table II.10 - Identification in the predicted *Sd. ludwigii* UTAD17 ORFeome of orthologues for the genes described to mediate *S. cerevisiae* tolerance to SO_2 at low pH, according with the data from Lage et al., (2019).

Chapter III

Appendix Table III.1- Ploidy estimation, through Flow Cytometry fluorescence measurements of meiosis, of *Sd. ludwigii* UTAD17 and BJK_5C. The determination was performed using *S. cerevisiae* strains BY4741 and BY4743 as haploid and diploid controls, respectively.

Appendix Table III.2- Genes specific of *Sd. ludwigii* UTAD17, and their homologues in strains UTAD17, BJK_5C and NBRC1722. A) Genes involved in thiamine biosynthesis; B) Genes involved in the catabolism of N-acetylglucosamine; C) Beta-mannosyltransferase.

Appendix Table III.3- Non-synonymous SNPs found between the genomes of *Sd. ludwigii* strains BJK_5C and UTAD17.

Appendix Table III.4 - Genes involved in the resistance to weak acids and ethanol in *Sd. ludwigii* strains BJK_5C and UTAD17.

Appendix Table III.5 - Sulfite efflux genes in *Sd.ludwigii*. **A**) BLASTp search results between the SSU1p sequences of strain BJK_5C and those of strains UTAD17 and NBRC1722. **B)** The *SSU1* "operon", with the corresponding locations within the genome, gene sizes and putative description of the resulting protein.

Chapter IV A

Appendix Table IV.A.1 – List of *Hanseniaspora* strains used in this study and their corresponding NCBI accession number.

Appendix Table IV.A.2 –Major and Minor aroma volatiles analyzed for the wines fermented by *H. uvarum* MJT198 in sequential inoculation with *S. cerevisiae* QA23 ; and for the wines fermented solely with *S. cerevisiae* QA23, at 72 hours after inoculation and at the end of fermentation.

Chapter IV. B

Appendix Table IV.B.1 - FT-ICR-MS raw data featuring the m/z peaks and their assigned number and molecular mass, and their corresponding intensities across three timepoints (T0h, T72h, and end of fermentation) for the sequential inoculation (HU_SC) and single inoculation (SC). The table also features the ions considered for molecular identification, the molecular formula obtained through the SmartFormula add-on, and the annotations obtained through the human metabolome database (HMDB) and yeast metabolome database (YMDB).

Appendix Table IV.B.2 - Peaks exclusive to *H.uvarum* (green) and *S. cerevisiae* (blue) at 72 hours and at the end of fermentation.

Appendix Table IV.B.3 - Resulting of read mapping against the *S. cerevisiae* QA23 genome at the end of fermentation from the fermentations conducted by *H. uvarum* MJT198 and *S. cerevisiae* QA23.

Appendix Table IV.B.4 - Genes with A) Expression Initiated at 72 Hours and B) Expression Lost at 72 Hours and their identification.

Appendix Table IV.B.5 - Up and down-regulated *H. uvarum* MJT198 genes between the timepoint at 72 hours and the initial timepoint.

Appendix Table IV.B.6 - Over and under-represented GO categories at 72 hours comparing to the beginning of fermentation, after Gene Expression Statistical Analysis (GSEA), grouped by category size and by their Normalized Enrichment Score (NES)

Appendix Figures

Chapter II

Appendix Figure II.1 - Dendogram of the alignment obtained for the Internal transcribed spacer (ITS) of different yeast strains belonging to the *Saccharomycodaceae* and *Saccharomycetaceae* families, including the strains that were used for the comparative proteomic analysis (underlined in red). The ITS sequences of the strains were retrieved directly from the NCBI database or in some cases from available genomic sequences.

Appendix Figure II.2 - Functional analysis of *Sd. ludwigii* UTAD17 as indicated by the metabolic reconstruction tool BLASTkoala. The number of *Sd. ludwigii* UTAD17 genes clustered in each of the functional categories is indicated inside the pie chart

Appendix Figure II.3 - Comparison of the amino acid sequence of the transcription factors ScCom2, ScMsn2 and SlMsn2 (ORF *SCLUD3.g330*). Using the available amino acid sequences for the different proteins a multiBLASTP analysis was undertaken using "Geneious Multiple Sequence Alignment ClustalW". Conserved residues are indicated in black boxes, being clear the highest degree of similarity at the C-terminal domain of the proteins which are known to harbor the DNA-binding domain of ScMsn2.

Appendix Figure II.4 - Amino acid sequence alignment of the sulfite efflux pumps *Ssu1* of *S. cerevisiae* and the four predicted orthologues in *Sd. ludwigii* UTAD17. This alignment was performed with "Geneious Multiple Sequence Alignment ClustalW" using the amino acid sequences ScSsu1, SCLUD1.g608, SCLUD1.g612, SCLUD1.g608b, SCLUD1.g612b. The red arrow indicates the premature STOP codon found in the predicted sequences of two of the *Sd. ludwigii* UTAD17 orthologues.

Chapter III

Appendix Figure III.1 - Whole-genome alignment using MAUVE of the *Sd. ludwigii* genomes of strains BJK_5C (used as reference) and NBRC1722.

Chapter IV A

Appendix Figure IV.A.1 – Phenotypic assessment in esculin medium of beta-glucosidase activity in natural yeast isolates of Moscatel Galego. Enzymatic activity is represented by a black halo around the yeast colony.