TÉCNICO LISBOA

Phenotypic analysis, genome sequencing and annotation of the non-conventional wine yeast *Saturnispora diversa*.

Sebastião Vaz Pardal da Silva Dias

Thesis to obtain the Master of Science Degree in

Biological Engineering

Supervisor:

Prof. Dr. Nuno Gonçalo Pereira Mira

Examination committee

President: Prof. Dr. Miguel Nobre Parreira Cacho Teixeira

Supervisor: Prof. Dr. Nuno Gonçalo Pereira Mira

Member of the committee: Prof. Dr. Ana Alexandra Mendes-Ferreira

December 2021

Preface

This document was written and made publicly available as an institutional academic requirement and as a part of the evaluation of the MSc thesis in Biological Engineering of the author at Instituto Superior Técnico. The work described herein was performed at the Institute for Bioengineering and Biosciences (IBB-IST) (Técnico,Lisbon, Portugal), during the period March-December 2021, under the supervision of Prof. Nuno Gonçalo Pereira Mira.

Declaration

I declare this document is an original work of my own authorship and that it fulfills all the requirements of the Code of Conduct and Good Practices of Universidade de Lisboa.

Acknowledgements

This work was only achievable due to the help I've received from various people, to whom I want to express my gratitude.

First and foremost, I would like to thank my supervisor, Prof. Dr. Nuno Gonçalo Pereira Mira, for all the guidance he gave me along the process of making this dissertation, for always being present and helping whenever I had doubts. Thank you for always challenging me to do better and for keeping me motivated at all times.

I also acknowledge funding by FCT (through contracts UIDB/04565/2020 and UIDP/04565/2020 to the Research Unit Institute for Bioengineering and Biosciences—iBB and the project LA/P/0140/2020 to the Associate Laboratory Institute for Health and Bioeconomy—i4HB).

My most sincere thanks to everyone in BSRG, Joana, Marta, Sara, Mónica and Nuno Pedro, for always being there to help me whenever I had questions or when things didn't go well. A special thanks to Maria João for her guidance throughout my work and for keeping me motivated, your help was pivotal for me to complete this dissertation.

Thank you, Professor Rogério, and Filipa, from Techlabs in FCUL, for your guidance as well in teaching me how to use PFGE equipment and the proper conditions to work on.

To my friends from the beginning, Guilherme, Barreto, Maria, Henrique, Diogo, Pedro, Fábrica, João and Manso, thank you for your constant support and friendship along all these years, you were very important for me to become the person I am today.

To the friends I made in Técnico, Rita, Laura, Glória, Narciso, Parada, Leonor, Cláudia, Daniela, José, Samuel and Ricardo, thank you for always being there in my journey throughout university, for helping in the bad moments and for always celebrating the achievements we made together.

Finally, I would like to express my incredible gratitude for the family I have, to my grandparents, to my uncles and cousins, you are my biggest supporters. To my brother, Francisco, thank you for always being there on my side, I'm incredibly lucky to call you my brother, and to my fathers, Patricia and Carlos, thank you for always believing in me, for the support you gave me in the decisions I've made along my life, you are all my pillars.

Abstract

Interest in the use of non-*Saccharomyces* strains as co-adjuvants in winemaking has emerged in the last few decades, given their capacity to produce aromatic molecules that *S. cerevisiae* does not produce. However, the lack of knowledge concerning the genetics and physiology of these species limits their more extensive utilization. The present work is focused on the genomic exploration of a *Saturnispora diversa* strain, MJT240, retrieved from grape must. Karyotyping of *S. diversa* MJT240, based on Pulse Field Gel Electrophoresis, showed the genomic DNA totals 9.68 Mb distributed along five chromosomes. The genomic sequence was obtained using Illumina and *de novo* assembled using SPAdes to yield 9.86 Mb, along 1150 contigs. Using manually curated *ab initio* gene detection, annotation of *S. diversa* MJT240 genome resulted in a predicted ORFeome of 5070 ORFs. Metabolic reconstruction revealed *S. diversa* is equipped with all genes for central carbon metabolic pathways, but does not contain key enzymes necessary for catabolism of galactose nor does it encode full biosynthetic pathways for the synthesis of thiamine and biotin. Enzymes necessary for the synthesis of the volatile aromatic compounds higher alcohols, ethyl esters, acetaldehydes were identified. Proteomic comparative analysis with species *Hanseniaspora guilliermondi, Torulaspora delbrueckii, Lachancea fermentati* and *Saccharomycodes ludwigii* revealed 1584 proteins for which no robust orthologue was identified, including five enzymes for the synthesis of glycosylphosphatidylinositol (GPI) proteins. Finally, the potential of *S. diversa* MJT240 to inhibit growth of the spoilage yeast *S. ludwigii* UTAD17 was herein studied using a co-culture setup, having resulted in *S. ludwigii* growth inhibition by 99.998%.

Keywords: *S. diversa* MJT240, non-*Saccharomyces* yeasts, genome annotation, metabolic reconstruction, proteome comparative analysis.

Resumo

O interesse na utilização de estirpes não-*Saccharomyces* como co-adjuvantes na vinificação surgiu nas últimas décadas, dada a sua capacidade de produzir moléculas aromáticas que *S. cerevisiae* não produz. Contudo, a falta de conhecimento sobre a genética e fisiologia destas espécies limita a sua utilização mais extensiva. O presente trabalho centra-se na exploração genómica de uma estirpe de *Saturnispora diversa*, MJT240, recuperada de mosto de uva. O cariótipo de *S. diversa* MJT240, baseado em Electroforese de Gel de Campo de Pulso, revelou um total de DNA genómico de 9,68 Mb distribuído ao longo de cinco cromossomas. A sequência genómica foi obtida utilizando Illumina e *de novo assembled* recorrendo ao algoritmo SPAdes para produzir 9,86 Mb, ao longo de 1150 *contigs*. Utilizando a detecção de genes *ab initio* com curação manual, a anotação do genoma *S. diversa* MJT240 resultou num ORFeoma previsto de 5070 ORFs. A reconstrução metabólica revelou que *S. diversa* está equipada com todos os genes para as vias metabólicas centrais do carbono, mas não contém enzimas chave necessárias para o catabolismo da galactose nem codifica as vias biossintéticas completas para a síntese de tiamina e biotina. Foram identificadas enzimas necessárias para a síntese dos compostos aromáticos voláteis álcoois superiores, ésteres etílicos, acetaldeídos. A análise comparativa proteómica com as espécies *Hanseniaspora guilliermondi, Torulaspora delbrueckii, Lachancea fermentati* e *Saccharomycodes ludwigii* revelou 1584 proteínas para as quais não foi identificado um ortólogo robusto, incluindo cinco enzimas para a síntese de proteínas de glicosilfosfatidilinositol (GPI). Finalmente, o potencial de *S. diversa* MJT240 para inibir o crescimento da levedura contaminante *S. ludwigii* UTAD17 foi aqui estudado em co-cultura, tendo resultado na inibição do crescimento de *S. ludwigii* em 99,998%.

Palavras-chave: *S. diversa* MJT240, leveduras não-*Saccharomyces*, anotação do genoma, reconstrução metabólica, análise comparativa.

Table of Contents

List of Figures

List of Tables

List of abbreviations

S. cerevisiae *Saccharomyces cerevisiae*

NSY Non-*Saccharomyces* Yeast

H. uvarum *Hanseniaspora uvarum*

H. guilliermondi *Hanseniaspora guilliermondi*

SO_2 Sulphur dioxide

S. ludwigii *Saccharomycodes ludwigii*

C. zemplinina *Candida zemplinina*

T. delbrueckii *Torulaspora delbrueckii*

C. pulcherrima *Candida pulcherrima*

HCA Hydroxycinnamic acids

B. bruxellensis *Brettanomyces bruxellensis*

DMDC Dimethyl dicarbonate

Z. bailii *Zygosaccharomyces bailii*

Blastn Blast nucleotide

rRNA ribossomal ribonucleic acid

YPD Yeast Peptone Dextrose

YCBAS Yeast Carbon Base Ammonium Sulphate

GJM Grape Juice Medium

WL Wallerstein Laboratory

OD Optic density

CFU Colony-forming Unit

Sd240 *Saturnispora diversa* MJT240

U17 *Saccharomyces ludwigii* UTAD17

PFGE Pulse Field Gel Electrophoresis

GC% Guanine-Cytosine Content

C. tropicallis *C. tropicallis*

CDS Coding sequence

E-value Expected value

PPP Pentose Phosphate Pathway

ORF Open Reading Frame

CoA Coenzyme A

PLP Pyridoxal-5-phosphate

RF Riboflavin

C. famata *Candida famata*

B. subtillis *Bacillus subtillis*

A. gossypii *Aspergillus gossypii*

GPI Glycosylphosphatidylinositol

Chapter 1 Introduction

1.1 Introduction to the theme of the thesis

Grape must is a low pH rich environment, enriched in sugars, where a wide variety of yeasts (among other microbes) thrive. These different yeast species can arise either from the surface of berries or from the equipment present in the cellar and the winery [1]. The microbes composition in grape must, generally defined as the must microbiome, depends on numerous factors including, climate conditions (temperature, humidity, etc), use of agrochemical treatments, geographical location or grapevine cultivar[1, 2]. The last of these is most evident in the comparison between white and red grapes, as, for example, red-pigmented basidiomycetous yeasts are more attracted to white grape varieties[3, 4]. Yeast's diversity on grapes was also found to depend on the water and nutrients available or the maturity stage of grapes[2, 3]. *Saccharomyces cerevisiae*, the leading species in vinification, isn't abundant in grape's must. In fact, the most abundant yeasts in the must are several non-*Saccharomyces* yeasts (NSY), especially in the early stages of grape processing and alcoholic fermentation [5], as seen in **Figure 1**.

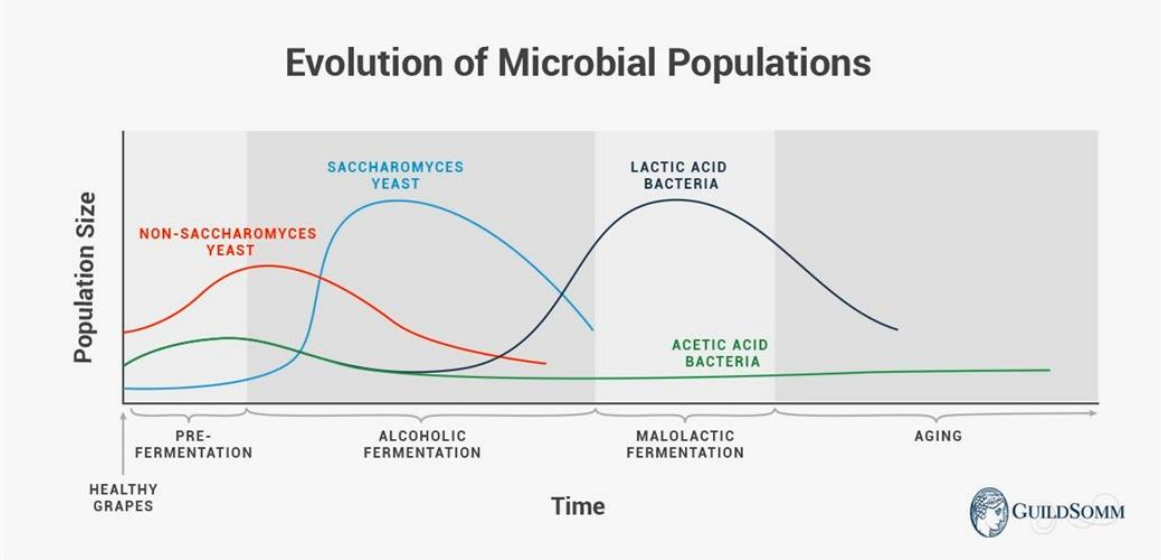


Figure 1- Evolution of microbial populations in the process of winemaking, from the harvesting of grapes to the end of fermentation and ageing. Non-S*accharomyces* are predominant in the early stages of grape processing and alcoholic fermentation, whereas in alcoholic fermentation *S. cerevisiae* is the leading yeast. Finally, in malolactic fermentation, lactic acid bacteria are predominant[6].

In the past, NSYs were generally considered deleterious and their presence by winemakers was feared due to the possibility of their activity depreciating wine properties like the aroma or the colour[7]. As more knowledge has been gathered concerning NSYs, it is becoming increasingly recognised that the use of NSYs as co-adjuvants of *S. cerevisiae* can be of interest to obtain wines with unique stylistic properties. This phenotype is attributed to NSYs as they are usually able to produce aroma molecules that are not produced (or that are produced in very

limited amounts) by *S. cerevisiae*[7]. Necessarily, the positive or negative influence of NSYs on wine flavour and other characteristics depends on the strain used, the metabolites these yeasts can produce, in what concentration and in the metabolic interactions established with *S. cerevisiae* and the remaining must microbiota[7, 8]. For instance, NSYs can promote the hydrolysis of several flavour compounds present in grapes through the excretion of β-glucosidases that can release aroma molecules that are attached to sugars, thereby enriching the sensory aromatic profile of wines. [7]

Generally speaking, there are three major groups of NSYs found in wine musts, especially in the early stages of the fermentation: apiculate yeasts with low fermentative activity like those belonging to the *Hanseniaspora* genus (like *H. uvarum* and *H. guilliermondii*); aerobic yeasts like those belonging to the *Pichia*, *Candida* or *Debranomyces* genera, and species with higher fermentative power like *Torulaspora delbruecki* or *Zygosaccharomyces bailli*. The majority of NSY species predominate in the early stages of the fermentation and their viability tends to decrease significantly as the fermentation progresses. The reasons underlying this decrease are complex and yet to be fully clarified (especially because many show a species or even strain dependence) but their reduced competitiveness due to limited metabolic capabilities, higher demand for oxygen that along vinification can be scarce and the increase in the concentrations of ethanol and of SO_2 (the preservative used by wine-makers to maintain microbiological stability) have been pointed as playing a role in this loss of viability[7]. The knowledge related to the physiology and biology of wine NSYs is still modest and this is illustrated, for example, by the number of genomic sequences that have been described and dully annotated for these species. In recent years, our team has been working in this field characterising, from the genomic point of view, members of the *Saccharomycodeacea* genus, which include not only the *Hanseniaspora* species but also those belonging to the *Saccharomycodes* genus[9, 10]. Among other aspects, the genomic analyses disclosed the lack of metabolic versatility of the *Hanseniaspora* species (due to impaired neoglucogenesis and glyoxylate cycle) or incomplete biosynthetic pathways for the synthesis of thiamine, a trait that may explain their reduced fermentation rates due to the essential role played by this vitamin in maintaining pyruvate decarboxylase activity[9]. In the case of *S. ludwigii*, the genomic analysis identified several genes that could be behind the extreme SO_2-tolerance phenotype for which this yeast species is known, as well as genes that render this species ability to assimilate N-acetylglucosamine, a trait uncommon in non-filamentous yeasts[10]. In this thesis, we progress the understanding of genomics in wine yeasts, this time focusing on an isolate retrieved during a survey of microbial profiling of wine musts produced in the Lisbon area, in 2019. This isolate, identified as belonging to the *Saturnispora diversa* species, is herein characterised and its genomic sequence obtained, annotated (with detailed manual curation) and preliminarily explored.

1.2 The world of non-*Saccharomyces* wine yeasts: a brief overview.

1.2.1 Potential of NSYs in winemaking

Nowadays, *Saccharomyces cerevisiae* commercial strains are used worldwide as starter cultures to undertake vinification, as they guarantee more predictable and reproducible fermentations (compared with spontaneous fermentations), consequently minimizing the risk of spoilage/contamination[11]. However, this practice also results in wines with reduced uniqueness and diversity in terms of aroma and flavour. In this context, the interest in using NSYs has grown in recent decades with a handful of strains already available in the market for commercial use. Besides this impact in aroma, NSYs have also been reported to produce wines with lower ethanol content, compared with *S. cerevisiae*, which can be an interesting trait in some consumer market segments[11, 12]. There are critical reasons for NSYs to be used together with *S. cerevisiae* and not as pure-cultures, including the possibility of producing higher concentrations of aroma molecules than those considered of interest (at a given threshold level the aroma content can be depreciative), high dependence of oxygen availability, limited tolerance to SO_2 and ethanol[13]. Furthermore, most NSYs also show low fermentative power and have slow fermentation kinetics, which can be a major drawback leading up to stuck fermentations[14]. In mixed fermentations with *S. cerevisiae*, either with co-inoculation or sequential inoculation, these negative features of NSY are mitigated as some of those traits are modulated by the established interaction with *S. cerevisiae*, therefore avoiding fermentation stagnation[15]. Since these yeasts have specific oenological characteristics that are not incorporated in *S. cerevisiae* commercial strains, controlled mixed cultures can have an additive impact on the wine quality, in terms of an aromatic and analytical profile. Several mixed fermentation studies have been performed in recent years to evaluate the potential to enhance given characteristics and/ or conditions of wine, such as the use of a *C. zemplinina* strains to increase the content of glycerol[16], mixed fermentations with *T. delbrueckii* strains to reduce acetic acid content and improve the overall quality of wine[17], among many other combinations. Nevertheless, most of the knowledge obtained concerning the impact of NSYs in fermentation is obtained at a lab-scale and a very limited number of studies have examined these yeasts' impact at large industrial scales[7, 11]. Furthermore, there is also a limited understanding on the dynamics of metabolic and physiological interactions between NSY and *S. cerevisiae* (and even with the remaining microbiota), such as the influence on growth kinetics upon mixed cultivation, the cell-to-cell contacts between both yeasts, the competition for space in medium, among others[8, 18]. Metabolic activities like the production of antimicrobial molecules/ toxic compounds, for instance, are other important aspects that require further investigations, as those may have an inhibitory or even a killing effect on NSYs. For instance, Branco et. al (2015)[19] demonstrated that *S. cerevisiae* strains produce antimicrobial peptides during alcoholic fermentation with *H. guilliermondi,* which have a negative impact on the latter by compromising the plasma membrane integrity and, thus, inducing a loss of culturability.

1.2.2 NSY communities in grape musts

As said before, the composition of yeast communities on grapes depends on multiple factors, including the stage of ripening. Most yeasts present on the grapes at harvest are washed into the must and may play an important role in fermentation. In the growth stages of grapes, the yeast flora present in their surface includes species of the *Basidiomycetes* phylum like *Rhodosporidium, Sporobolomyces, Cryptococcus or Rhodotorula*, possibly due to their higher capacity to grow in nutrient-limited environments[20]. These, however, aren't relevant to vinification as they are unable to ferment most juice sugars and survive in the winemaking process[20]. Species emerging in later stages of ripening, where nutrients leak from the surface, are generally *Ascomycetes* belonging to the *Metschnikowia, Candida* or *Hanseniaspora* genera[20]. In light of this profile of isolation, dependent on the ripening stage, the yeast species that are dominant in grape must and in the early stages of fermentation belong to the genera *Candida*, *Hanseniaspora*, *Metschnikowia*, *Lachancea, Pichia* and *Torulaspora*[21, 22]. It must be reinforced, however, that, as stated before, the must microbiome depends on a large multiplicity of factors and, therefore, the exact distribution of microbial communities, including yeast species, can present considerable differences. Pinto et. al (2015)[23] characterised and compared the yeast diversity in spontaneous fermentations across different wine regions of Portugal. The results confirmed considerable changes in biodiversity, unique for each region, with the genera *Hanseniaspora, Lachancea* and *Aureobasidium* being most predominant, as seen in **Figure 2**. As expected, isolation of *S. cerevisiae* was almost inexistent in the initial must (**Fig.2**). Compared to other regions in Europe[21, 24], there is a clear divergence in yeast community distributions, in which *H. uvarum* has the biggest prevalence, and species of the genera C*andida* and *Metschnikowia* have a higher frequency. Another difference noted by the authors was the lack of identification of yeasts belonging to the *Pichia* genus.

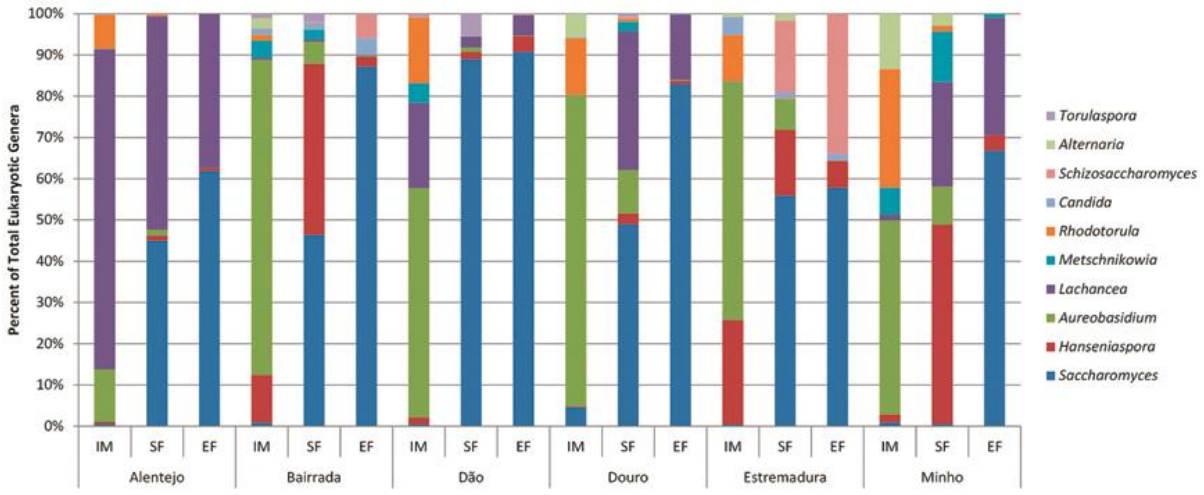


Figure 2- Eukaryotic microbial distribution, in cumulative frequency, over IM (initial Must), SF (start of fermentation) and EF (end of fermentation) in six different Portuguese wine regions: Alentejo, Bairrada, Dão, Douro, Estremadura and Minho.

1.2.3

1.2.4 Overview of wine NSYs

Among apiculate yeasts, species belonging to the *Hanseniaspora* genus, like *H. uvarum* and *H. guilliermondi* predominate in wine musts, usually representing 50-75% of the total yeast population[3]. *Hanseniaspora* species are weak fermenters and, as most NSYs, have a low tolerance to ethanol (that can range between 3 and 5% [25]). It is important to note that these values are highly variable between species and strains. For instance, *H. guilliermondi* strains with high tolerance to ethanol have been identified[26]. The more interesting attributes for wine production of *Hanseniaspora* strains/species include their capability of producing important volatile compounds, such as acetate and ethyl esters, that give the wine a positive fruity aroma, along with sulphur containing compounds and higher alcohols, considered to also have an additive impact to the aroma profile of wine[27]. *H. uvarum* was found to produce high quantities of isobutyl and isoamyl acetates, molecules that give the wine strawberry and banana aromas[25], while *H. guilliermondi* is reported to produce increased concentrations of acetate esters, like hexyl and ethyl acetate, and of 2- phenyl-ethyl acetate, which enhances flowery and fruity aromas in wine[28]. However, these species can also produce other molecules that may depreciate wine aroma such as 3-mercapto-1-propanol and trans-2-methyltetrahydrothiophen-3-ol, associated with a rancid/ sweaty aroma, being therefore essential to maintain a certain degree of control of their activity to ascertain wines with interesting properties. Both *H. uvarum* and *H. guilliermondi* preferably consume fructose over glucose, albeit the molecular reasons underlying this phenotype are not fully understood yet[9, 29]

Species from the genus *Candida* are also frequently found in grape musts with the more common species being *C. zemplinina* and *C. pulcherrima,* the anamorph form of *Metschnikowia pulcherrima*. Unlike other NSYs, *Candida zemplinina* (also described as *Starmerella bacillaris*), has been described to have a high tolerance to ethanol, superior to 10% (v/v), and may survive until the end of fermentation with a superior fermentative power[25, 30]. This species is known to produce low levels of acetaldehyde and acetic acid (which are deleterious to wine's quality) but it also produces high concentrations of glycerol, which enhances mouthfeel sensation[7, 30]. Also, *C. zemplinina* has been previously reported to favour fructose consumption in comparison with glucose and has a poor ethanol yield derived from sugar consumption, a trait considered useful to produce low-alcohol beverages[31]. *M. pulcherrima* is also abundant in the must, accounting for normally 5-20% of the overall yeast population in it[27]. Its relatively low fermentative power, compared to other NSYs, and low resistance to ethanol stress, ranging around 4% (v/v), are critical limitations to its usefulness in vinification[25]. Nevertheless, this species is known to contribute to the wine aroma profile due to its capability of producing β-glucosidases and β-xylosidades[32], enzymes that improve the production of the volatile aromatic compounds 2-phenyl ethanol and monoterpenes (nerols, geraniols and linalools) [32, 33]. However, some strains of this species have also been reported to produce "off-flavours" at relatively high concentrations [34]. Furthermore, this yeast is known to produce a red soluble pigment called pulcherrimin, in which its precursor pulcherriminic acid has antimicrobial activity towards other yeasts by reducing the iron available [35]. Consequently, it has been previously reported

that this compound may have an adverse effect on the growth of *S. cerevisiae* [7], albeit the results available in the literature suggest that the impact can be relatively minor[35].

Torulaspora delbrueckii is another NSY with oenological potential that is frequently isolated from wine musts[36]. Compared to *H. uvarum or C. zemplinina*, has a higher fermentation rate and is considered a low ethanol producer. However, it is less tolerant to lower oxygen levels during fermentation [37]. This yeast has been reported to produce low levels of volatile acidity, acetoin, hydrogen sulfide and acetaldehyde, associated with off-flavours, both when used alone or in combination with *S. cerevisiae* [17, 38]. Concerning the chemical composition and aroma profile of the wine, *T. delbrueckii* is reported to over-produce succinic acid and in some strains also linalool and 3-ethoxy-1-propanol, two volatile compounds derived from monoterpene alcohols [7, 34] that are associated with citrus fruit descriptors [34].

1.2.5 NSYs: friends, but also foes.

Despite what has been discussed above concerning the generally positive influence of NSYs in the grape must, some species are particularly feared due to their track record as spoilage yeasts, either because they over-produce compounds that can reduce organoleptic properties of wine or because they augment ethanol concentrations, cause sediments or cloudiness along with wine maturation [39]. The most commonly described species associated with wine spoilage are *Brettanomyces/ Dekkera bruxellensis*, *Zygosaccharomyces bailli* and *Saccharomycodes ludwigii. Brettanomyces bruxellensis* has been considered an undesirable contaminant of wine for many years, mainly due to its ability to convert hydroxycinnamic acids (HCAs) into vinyl phenols and ethyl phenols, two molecules associated with "horse sweat" trace aroma [39, 40]. This yeast can stay viable for long periods and, consequently, proliferate during wine ageing and in wine bottles, thereby being able to trigger contamination even after long periods since fermentation has ceased. It is, currently, one of the main concerns in winemaking, and an effective method to control its activity has not been fully developed[41]. *B. bruxellensis* proliferation is normally controlled with the addition of SO_2, however, it has been demonstrated that some strains may endure extremely high concentrations of this chemical, in some cases above the legally allowed limit[41, 42]. Also, Capozzi et. al (2016)[43] demonstrated that SO_2 may induce the formation of viable but non-culturable cells of *B. bruxellensis*, which may lead to new contamination of the yeast in the final stages of winemaking, after fermentation. The use of killer toxins to prevent the proliferation of the yeast are being studied and may represent a very useful tool for practical use in the future, due to their selectivity and stability[42]. *Saccharomycodes ludwigii* is another spoilage yeast that, like *B. bruxellensis*, shows extreme tolerance to SO_2 [10, 44]. The principal deleterious effects described in association with *S. ludwigii* contamination of wines are associated with growth in the final stages of winemaking, including in bottled wines, where it can lead to turbidity and sedimentation, and even re-fermentation of the wine[45]. Compared to *B. bruxellensis*, the frequency of S. ludwigii isolation is considerably smaller [46]. Some of the possible treatments to control *S. ludwigii* are the use of dimethyl dicarbonate (DMDC), as a replacement of SO_2. The use of killer toxins, produced by other yeasts may also be another useful tool for control of *S. ludwigii*[44]. *Zygosaccharomyces bailii* is also considered a problematic spoilage wine yeast as it is very tolerant to harsh/stress conditions such as low pH,

organic acids and preservatives, and due to its high osmotolerance[40]. The effects in wine contaminated by this species include cloudiness, visible sediment formation in dry wines and refermentation in sweet wines. Also, this yeast can produce undesirable odour active metabolites to the wine aromatic profile. [45, 47] The methods of control more effective for *Z. bailli* inhibition include the use of DMDC or chitosan [47].

Albeit being considered spoilage yeasts due to the above-presented features, all three yeasts have been recognised to have an interesting potential as bio-flavouring agents[7, 44, 48]. For instance, the use of *S. ludwigii strains* in mixed fermentations with *S. cerevisiae* resulted in a positive increase, compared to pure cultures of *S. cerevisiae,* of the volatile aromatic compounds isoamyl acetate, glycerol and 2-phenyl ethanol, as well as in lower volatile acidity, all traits associated to improved wine quality[44, 49]. In the trials performed by Vejarano et. al (2018)[44], another interesting feature found in mixed fermentations undertaken with *S. cerevisiae* and *S. ludwigii* is the reduction of the ethanol content in 1.7% (v/v), compared to pure fermentation of *S. cerevisiae*. *B. bruxellensis* also has a biotechnological application in lambic beer fermentation due to the production and liberation of volatile compounds that enhance the floral/ fruity traits of beers through the over-production of higher alcohols and several esters (ethyl lactate, ethyl acetate, ethyl caprate and ethyl caprylate)[48].

1.3 Relevant traits about the physiology of wine NSYs in the context of winemaking disclosed by genomic analyses.

Studies regarding the use of NSY in wine fermentations show us that certain specific traits enable them to be (or not to be) suitable for use in alcohol fermentation, including their ability to adapt to the stressful environment of wine must or to produce certain metabolites that are important for wine aroma complexity. These phenotypic traits are, surely, the result of the interplay of different molecular mechanisms, however, how those mechanisms function and inter-connect is not completely understood, something that can't be easily perceived from the available literature, which focuses more on the result of the application of the yeasts (rather than elaborating on the reasons underlying it). The use of comparative and functional genomics is expected to shed light on this by disclosing, to a first extent, the portrait of the ORFeomes of different species, that can, in a first attempt, help to understand what their potential is concerning a determined phenotype. The disclosure of the genomic sequence, accompanied by its annotation, is also an important step to undertake analysis at a subsequent cellular level (such as the transcriptome or the proteome) that can provide better insights into the regulation of the different molecular mechanisms governing phenotypes of interest in NSYs, as also on how they can inter-connect and shape the resulting metabolic interactions. Seixas et al. (2019) [9] performed an extensive analysis of the genomic sequence of *H. guilliermondii* that involved, among other aspects, the metabolic reconstruction and comparative analysis with the genomes of *S. cerevisiae* and other wine NSYs. That analysis showed that *H. guilliermondii* is not equipped with a full gluconeogenesis nor glyoxylate cycle, and also determined the absence of genes encoding for enzymes necessary for catabolism of galactose, xylose or lactate. These genes were also absent from *H. opuntiae* and *H. uvarum*, two species closely related to *H. guilliermondii*. Altogether, these observations sustain the idea that the reduced competitiveness of *Hanseniaspora* species in wine fermentation can also come from their higher dependence on the availability of glucose and fructose. The

pathways leading up to biosynthesis of thiamine or biotin were also determined to be incomplete in *Hanseniaspora* species and no orthologue for the sulphite efflux pump Ssu1, an essential determinant of tolerance to SO_2, could be identified. These genomic traits are in line with the reduced fermentation rates exhibited by *Hanseniaspora* species and with their generally high susceptibility to SO_2. Tavares et. al (2021)[10] have also analysed the genomic sequence of the *S. ludwigii* strain, which resulted in the uncovering of several genes that are candidates to mediate tolerance to SO_2, with emphasis on the identification of four Ssu1-like proteins. The multiplication of this potent Ssu1-tolerance determinant can be behind the extremely high tolerance phenotype exhibited by *S. ludwigii* cells to SO_2. Notably, *S. ludwigii* genomic analysis also disclosed a complete gluconeogenesis and glyoxylate cycle, as well as complete pathways for the synthesis of biotin and thiamine, unlike what was observed in the sister *Hanseniaspora* species. Similarly using functional and comparative genomics, Junior et al. 2018 [50] performed whole-genome analysis of *C. zemplinina* (*S. bacillaris*) strains and have uncovered the presence of genes encoding enzymes involved in the synthesis of higher alcohols, consistent with the reported capability of this species in over-producing these molecules.

1.4 The *Saturnispora diversa* species.

This thesis is focused on an isolate (named MJT240) obtained in 2019, retrieved from grape must obtained from Muscat grapes harvested in the Lisbon area. After isolation, the isolate was identified (based on sequencing of the highly conserved D1/D2 region) as *Saturnispora diversa*. Formerly, *S. diversa* was classified as *Candida diversa*[51, 52], with a very limited amount of information being published about it. Despite that, description of *Saturnispora diversa* isolates from grape musts was already performed in the region of Ontario or Slovenia and Assam tea plants in Thailand [53-55]. Description of three *S. diversa* isolates (named 11.77, 11.79 and 11.86) from spontaneous fermenting musts obtained with Touriga Nacional grapes was also described in a study undertaken in Portugal, that examined the impact of using these isolates in producing wines [56]. To get a glimpse of how these isolates compared to ours, it is shown in **Fig.3** a nucleotide alignment of the D1/D2 sequences of the different isolates. The resulting alignment showed a very high identity between these different isolates. BLASTN analysis of the 16S RNA sequence at the NCBI database also revealed an extreme similarity to sequences attributed to other *S. diversa* strains (shown in **Appendix A**). Altogether, these observations sustained our confidence in the identification of the isolates as belonging to the *S. diversa* species.

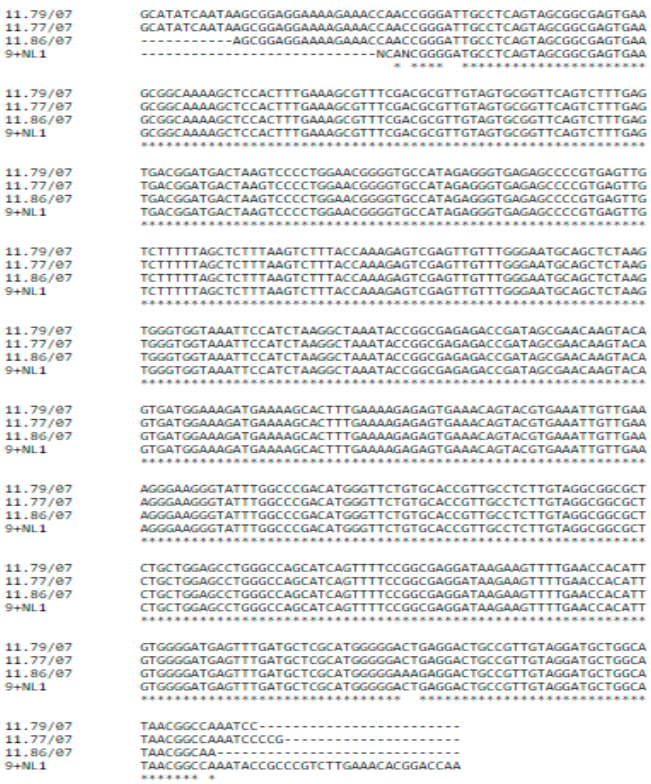


Figure 3- Multiple sequence alignment, in ClustalW interface, of rRNA sequences of isolates 11.86, 11.79, 11.77 and 9+NL1 (SdMJT40). The result shows a very high identity in all isolates, suggesting these are identified as belonging to *S. diversa* species.

The first description of *Saturnispora diversa* was made in 1960 by Ohara et al (1960)[57], reporting it as *Candida fimetaria var. diversa*[51]. Interestingly, this was also a strain retrieved from a grape must. In 1970, *C. fimetaria* was transferred to *C. lambica* and, due to physiological differences, *C. diversa* was elevated to the rank of species[51, 58]. Later on, Kurtzman et. al (1998) [59] determined, through phylogenetic analysis of the D1/D2 large sequence, that this yeast belongs to the *Saturnispora* clade. However, as the yeast is anamorphic and, according to the International Code of Botanical Nomenclature (Vienna Code), it was required that anamorphic and teleomorphic species must be assigned to different genera, the species must remain allocated to the *Candida* genus[52, 60]. With the introduction of the Melbourne Code, International Code of Nomenclature for Algae, fungi and plants (adopted in July 2011), it is now permitted to assign to the same genus related species, either teleomorphic or anamorphic. Therefore, many anamorphic species that were shown to be members of teleomorphic clades were assigned to the correspondent genera. Thus, *Candida diversa* is now assigned to the *Saturnispora* clade[52]. Also, other species were shown to be phylogenetically separated from their known genera and were, therefore, assigned to the same genus[52].

The *Saturnispora* clade was originated to accommodate Saturn-spore species that were previously assigned to the genera *Pichia* and *Williopsis*[61]. From the phylogenetical analysis of partial sequences of subunit rRNAs, it was possible to determine that the *Williopsis* species were very separated from each other and, consequently, should not be assigned to a new genus. However, from the same rRNA analysis, it was determined that all *Pichia* species were very similar and showed a great genetic discrepancy to the *Williopsis* species despite having very

similar phenotypic characteristics. Therefore, *Pichia* species were assigned to a single new genus, the *Saturnispora,* inserted in the family *Pichiaceae*[61]. The assigned species were *Saturnispora dispora, Saturnispora satoi, Saturnispora zaruensi* and *Saturnispora ahearnii* [61, 62]*.* Nowadays, the genus is well-supported by phylogenetic analyses established from the analysis of multiple genes sequences of the conserved translation elongation factor-1α gene, besides D1/D2 [63]. As of 2019, there are a total of 20 species inserted in the *Saturnispora* genus, 11 of which are teleomorphic [63, 64]: *Saturnispora besseyi, Saturnispora ahearnii, Saturnispora dispora, Saturnispora satoi, Saturnispora zaruensi, Saturnispora gosingensis, Saturnispora hagleri, Saturnispora mendoncae, Saturnispora quitensis, Saturnispora bothae and Saturnispora serradocipensis* – and 9 are anamorphic species [52, 60, 65, 66] - *Saturnispora diversa, Saturnispora halmiae, Saturnispora sanitii, Saturnispora sekii, Saturnispora siamensis, Saturnispora silvae* and *Saturnispora suwanaritii, Saturnispora mangrovi and Saturnispora kantuleensis.*

Saturnispora cells normally exhibit a spherical/ ovoid shape and the formation of pseudohyphae can be observed[65]. Cells reproduce asexually through multilateral budding on a narrow base and are described as not being able to consume D- xylose nor nitrate while glucose is fermented[65]. Up to now, *S. diversa* is the only species from the *Saturnispora* clade found in grape must or damaged/ rotten grapes[67]. Other *Saturnispora* species have been found in a wide variety of rich-specie sources and habitats around the world, as seen in **Table 1**.

Table 1- Description of species *S. mangrovi, S. quitensis, S. kantuleensis, S. hagleri, S. halmiae, S. siamensis* and *S. sylvae*, included in the same genus as *S. diversa*, and their correspondent isolation locals.

Specie	Isolation local
Saturnispora mangrovi	Syhat mangroves in Saudi Arabia[66]
Saturnispora quitensis	Maquipucuna cloud forest reserve in Ecuador[63]
Saturnispora kantuleensis	Peat swamps in Thailand[68]
Saturnispora hagleri	Drosophila flies in Atlantic rainforest in Brazil[69]
Saturnispora halmiae	Ghanaian cocoa fermentations[70]
Saturnispora siamensis	Rotten wood, mangroves in Thailand[71]
Saturnispora silvae	Plant phytotelmata waters in Brazil[72]

Saturnispora diversa is described to assimilate citric and succinic acid, D-mannitol, D-glucitol, ethanol, and to undertake glucose fermentation[51]. The yeast has also been described to assimilate glycerol, xylitol and D-xylose[73]. Its growth is inhibited in agar containing trehalose, lactose, raffinose, rhamnose, methanol, erythritol, inulin, β-Glucosides, melibiose, nitrite, nitrate[51]. When cultivated in a rich medium (YPD medium), cells present an ovoid to short cylindrical form and are distributed singly or in short chains. However, in different medium conditions - on yeast carbon base ammonium sulphate (YCBAS) agar, at 18ºC and after 14 days - cells change their morphology to a poorly developed pseudohyphae form, consisting of branched chains of cylindrical cells, as seen in **Figure 4** [51]. There is little information available about this yeast since it is difficult to evaluate the

occurrence of *S. diversa* due to its similarity based on physiology to other yeasts, particularly *S. silvae*. Nevertheless, the most useful characteristics to identify a microorganism as belonging to *S. diversa* species is their ability to assimilate citric acid and D-mannitol, their lack of growth in medium with β-Glucosides or medium containing trehalose as carbon source[51]. Also, their ability to ferment glucose is useful to identify *S. diversa* species[51].

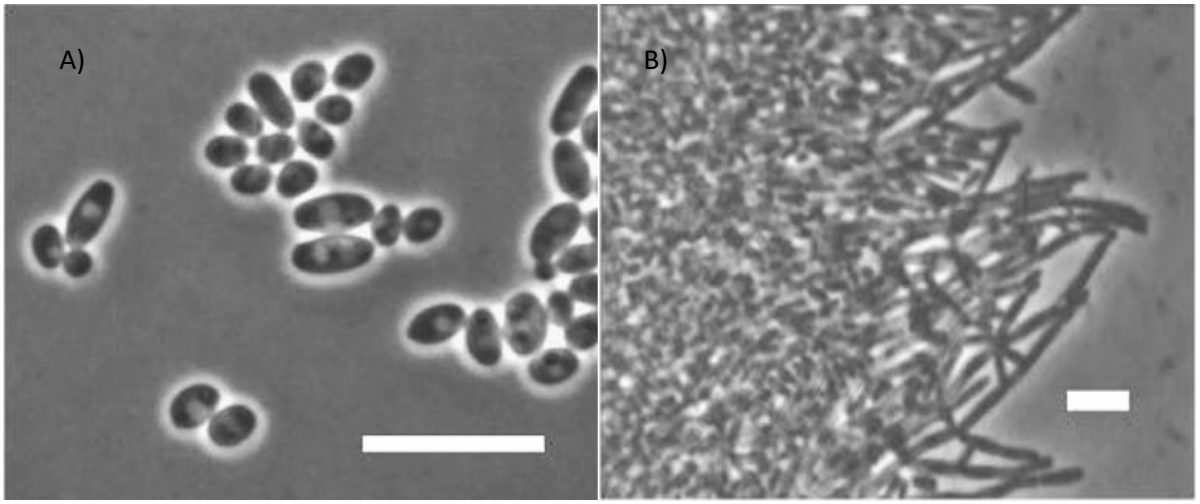


Figure 4 - Microscopic visualization of *S. diversa*: A) Cells grown in rich medium (YPD broth) and ambient temperature present an ovoid to cylindrical form and are distributed single; and B) Cells grown in YCBAS agar, 18ºC present pseudohyphae form and are aggregated.

Albeit *S. diversa* has not been extensively characterised, there are several yeast characterisation studies in which *S. diversa* (formerly *C. diversa*) was reported. In this context, *S. diversa* was isolated from soils of Assam tea plants and in peat swamps of Thailand, and there have been studies to assess the role played by yeasts in promoting plant growth. In these studies, the production of the extracellular enzymes was evaluated and it was determined that *S. diversa* produces extracellular lipases, with an enzyme activity Index of 1.20 and 2.4, in the soils of Assam tea plants and peat swamps, respectively [55, 74]. Differently, no extracellular exoglucanase, xylanase, pectinase, amylase and proteases were produced by *S. diversa*[55, 74]. Interestingly, ammonia production[55] and indole-3-acetic acid production (IAA)[74], a hormone of the class auxin responsible for cellular mechanisms related to the growth of plants, were detected in supernatants obtained by *S. diversa* cultures in concentrations up to 41.6 mg/L, which compared to the other identified yeasts was particularly high[74]. The same isolates obtained from peat swamps in Thailand showed the capability to produce ethanol from glucose in concentrations reaching 50g/L, whereas when using xylose as carbon source no ethanol production was detected[74]. Lentz et al. 2014[73] studied the potential of using different yeasts in the fermentation of beers and determined that growth of *S. diversa* isolates was inhibited when concentrations of ethanol were above 8%, thus showing a good ethanol tolerance compared to other yeasts. In the same study, it was also observed the acidophilic behaviour of *S. diversa,* exhibiting growth in a wide range of pH, from 5 to 2.4.

The *S. diversa* strains 11.77, 11.79 and 11.86, mentioned above as being retrieved from spontaneous fermenting musts of Touriga Nacional variety, were tested for their oenological potential as starter-cultures[56]. Interestingly, *S. diversa* emerged, together with *S. bacilaris*, as having good oenological potential, with the resulting wines having intense and diverse aroma characteristics and high balance. Some of the parameters of the wine after fermentation that were considered relevant in the wines produced with *S. diversa* as a starter culture, were the low alcohol content (7.8% (v/v), compared with 12,79% (v/v) obtained using *S. cerevisiae*), the high glycerol concentration (13.11g/ L) and fermentation purity [volatile acidity (g/l)/ ethanol % (v/v)]. The positive aroma descriptors identified in the wines produced by *S. diversa* were the nutty (volatile compounds benzaldehyde, 2,4-octadienal, etc[75]) and dried fruits descriptors (esters, isopentyl alcohol[76]). Concerning the mouthfeel descriptors, softness and sweetness were higher in the wines produced with *S. diversa*, compared with those obtained with *S. cerevisiae*. Characteristics and parameters of the wine fermented by *S. diversa,* compared to the reference (*S. cerevisiae*), are presented in **Table 2**.

Table 2- Parameters relative to the result of fermentation of Touriga Nacional musts, using as starting culture the yeasts *S. diversa* and *S. cerevisiae* (reference), respectively.

Parameters	***Saturnispora diversa***	**Reference**
Alcohol content (% v/v)	7.80	12.79
Fructose ($g.L^{-1}$)	70.98	70.98
Glucose ($g.L^{-1}$)	22.17	0.25
Volatile acidity ($g.L^{-1}$)	1.08	0.78
Glycerol ($g.L^{-1}$)	13.11	4.17
Fermentation rate ($g.L^{-1}$ of CO_2/day)	9.4	22.2
Fermentation purity (Volatile acidity / ethanol % (v/v)	0.14	0.06
pH	3.41	3.47

In another study, conducted by Buzzini et al. 2003[77], the volatile organic compounds (VOC) produced by different yeast strains retrieved from rain forests in Brazil was studied, including in that set strains identified as *Saturnispora diversa*. Gas chromatography-mass spectroscopy (GC-MS) analysis led to the conclusion that *S. diversa* cells can produce amyl and isoamyl alcohol, associated with fusel aromas[78] and isobutyl, isoamyl and 2-methyl butyl acetate, natural esters associated with fruity odour/flavour. From the esters, isoamyl acetate has a higher impact on wine aroma [79].

An interesting trait that has been linked to *Saturnispora diversa* is its antagonistic behaviour, that is, its ability to inhibit growth, activity and/ or the reproduction of phytopathogens [80-83]. Specifically, it was observed that *S.*

diversa shifts from single cells to a pseudohyphal morphology when cultivated on 0.3% agar YPD, exhibiting considerable biocontrol performance against *Botrytis cinerea* in post-harvested fruits (apple and kiwi). A noticeable aspect was that the biocontrol activity was more pronounced at 4ºC in treatment for apples, with a disease incidence decrease of more than 80%[80]. In a subsequent study [82], it was demonstrated that the use of *S. diversa* together with harpin, an acidic and heat-stable protein that acts as a hypersensitive response elicitor, provided superior control of post-harvested kiwifruit. This was partly due to an increased level of activity of SOD, a defence-related enzyme that plays a role in detoxifying reactive oxygen species, thus inhibiting the oxidative effect caused by fungal pathogens. More recently, in a study performed by Raspor et. al (2010) [83], a biocontrol assay on grape berries assessed that *S. diversa*, in its usual form, had a biocontrol activity of 32.2% against *B. cinerea,* when grown in NYDA medium.

Chapter 2 Materials and methods

2.1 Yeast Strains and Culture media

The yeast *S. diversa* MJT240, a strain retrieved from grape must obtain from Muscat grapes harvested in the Lisbon area, and the yeast *S. ludwigii* UTAD17 [84], an autochthonous Douro Wine Region strain, were both isolated in our laboratory and used in this study. Yeasts were routinely maintained at 4 °C on Yeast Peptone Dextrose agar plates (YPD) which contains, per litre, 20 g of glucose, 10 g of peptone, 5 g of yeast extract, and 20 g of agar from stocks stored at −80 °C. Prior to use, the yeasts were transferred to a new slant of YPD and incubated for 24–48 h at 30 °C, unless otherwise stated. Strains were also cultivated in liquid YPD medium (whose composition is identical to the one described above without agar), in selective minimal medium plates (20g/L of carbon source, 20g/L of agar, 2,67g/L of ammonium sulphate, 1.7g/L of YNB w/out amino acids nor ammonium sulphate) and in YPD plates supplemented with 6 mM SO_2. For that, a stock solution of SO_2 0.1M, previously sterilized by filtration, was used. The two strains were also cultivated in synthetic grape juice medium (GJM), from a known receipt [85], which contains, per litre, 100g of glucose, 100 g of fructose, 3g of malic acid, 0.2g of citric acid, 1.14g of K2HPO4, 1.23g of MgSO4.7H2O, 0.44g of CaCl2.2H2O, 0.504g of di-ammonium potassium and 5g of tartrate potassium, 100 mL of vitamin solution 20x (composition provided in **Table 3** below), 1mL of mineral solution 10000x (composition provided in **Table 3** below), 0.5 mL of ergosterol solution 10x and 22.12mL of amino acid solution 25x (composition provided in **Table 4** below). Tartrate potassium was the last compound added to the medium, considering its low solubility. After the mixture of all reagents, the pH of the GJM medium was adjusted to 3.5 (using HCl 1M), and the GJM medium was filtered and stored at 4ºC until further use.

Table 3- Reagent's composition of mineral solution 10000x (V=25mL), vitamin solution 20x (V=100mL), ergosterol solution 10x (V=25mL), components of synthetic grape juice medium (GJM).

Mineral solution 10000x (V_{stock}=25mL)		**Vitamin solution 20x (V_{stock}=100mL)**		**Ergosterol solution 10x (V_{stock}=25mL)**	
Reagents	Mass (mg)	**Reagents**	Mass (mg)	**Reagents**	Mass (mg)
MnCl2.4H2O	49.55	**Myo-inositol**	200	**Ergosterol**	125
ZnCl2	33.875	**Pyridoxine.HCl**	4	**Tween 80**	*
FeCl2.4H2O	12.525	**Nicotinic acid**	4		
CuCl2	3.4	**Calcium pantothenate**	2		
H3BO3	1.425	**Thiamine.HCl**	1		
Co(NO3)2.6H2O	7.275	**p-aminobenzoic acid**	0.4		
Na2MoO4.2H2O	6.05	**Riboflavin**	0.4		
KIO3	2.7	**Biotin**	0.25		
		Folic acid	0.4		

*Tween 80 is liquid: V= 6.5 mL.

Table 4- Reagent's composition of amino acid solution 25x (V=25mL), a component of synthetic grape juice medium (GJM).

Amino acid solution 25x (V_{stock}=25mL)			
Reagents	Mass (mg)	**Reagents**	Mass (mg)
Alanine	250	**Lysine.HCl**	780
Arginine	2275	**Methionine**	375
Aspargine	425	**Phenilalanine**	375
Aspartic acid	875	**Proline**	1250
Glutamic acid	1555.5	**Serine**	1000
Glutamine	500	**Tryptophan**	250
Glicine.HCl	185	**Tyrosine**	50
Histidine	510	**Threonine**	875
Isoleucine	500	**Valine**	500
Leucine	750		

WL medium [86] [87] was also prepared for use of WL agar plates in viable cell plate counting. Reagent composition is presented in **Table 5**. After the mixture of all reagents, pH is adjusted to 4.5.

Table 5- Reagent's composition of WL medium, per litre.

Reagents	Mass (mg)
Yeast extract	4000
Tryptone	5000
Glucose	50000
Potassium dihydrogen phosphate	550
Potassium chloride	425
Calcium Chloride	125
Magnesium sulphate	125
Ferric chloride anhydrous	2.5
Manganese sulphate	2.5
Bromocrysol green	22
Agar	15000

2.2 Assessment of *S. diversa* MJT240 to use different sugars as the sole source of carbon and energy.

To test the capability of *S. diversa* MJT240 cells to grow in different growth media with different sugars as sole carbon sources, spot assays were used. For this, cells were cultivated in YPD liquid medium until mid-exponential phase (OD600nm~10 for *S. diversa* and OD600nm~3.5 for *S. ludwigii*) and then used to prepare, in water, a cell suspension having an OD600nm of 0.05. Two subsequent dilutions of this initial cell suspension, 1:5 and 1:10, were prepared. The initial suspension and the diluted ones were applied as spots of 2uL onto the surface of solid selective minimal medium plates with glucose replaced for fructose or mannose, galactose, mannitol, sorbitol, saccharose, maltose, raffinose, rhamnose, xylose and acetate as the sole carbon sources. In all cases, the final concentration of sugar used was 2%. Sodium acetate was used as a source of acetate. The inoculated plates were incubated at 30ºC for 2-4 days, depending on the detectable growth.

2.3 Assessment of *S. diversa* MJT240 and *S. ludwigii* UTAD17 growth parameters along growth in synthetic grape juice medium.

S. diversa MJT240 and *S. ludwigiii* UTAD17 were profiled for their growth rates in synthetic grape juice medium. For both assays, a pre-inoculum of the strains was performed overnight in GJM medium in 50mL shake flasks containing 25mL of medium. Cells were incubated at 30ºC with orbital agitation of 250rpm. On the next day, an appropriate volume of the pre-culture was used to inoculate fresh GJM medium, aiming to obtain a cell suspension with an OD600nm of 0.1. These cultures were then incubated at 30ºC and 250rpm for 48-72h until the stationary phase was reached. Growth was followed based on the increase of OD600nm and of viable cell plate counting in appropriate time intervals. Using the values of OD600nm from the exponential phase, the specific growth rates (μ) and doubling time (Dt) were calculated for *S. diversa* MJT240 and *S. ludwigiii* UTAD17.

2.4 Growth of *S. diversa* MJT240, alone or in combination with *S. ludwigii* UTAD17, in liquid GJM medium.

Growth trials in synthetic GJM involving co-cultivation of *S. diversa* MJT240 and *S. ludwigii* UTAD17 were conducted using three experimental setups: (1) a single culture of *S. ludwigii* UTAD17 (U17), (2) a single culture of *S. diversa* MJT240 (Sd240) and 3) a mixed culture with *S. ludwigii* UTAD17 and *S. diversa* inoculated simultaneously. The procedures used to inoculate the growth medium, in the single and the co-cultures, were those described above for growth parameters determination. In the co-culture assays, the growth of the yeasts was followed by measuring the increase of CFUs in appropriate time intervals until the stationary phase was reached. To monitor the growth of *S. diversa* MJT240 cells an aliquot of the culture was plated on solid WL medium being the *S. diversa* colonies distinguishable from those of *S. ludwigii* UTAD17 due to differences in colour and shape. To accompany the growth of *S. ludwigii* UTAD17, an aliquot of the culture was plated on YPD medium + 6mM SO_2, a condition in which *S. diversa* MJT240 exhibited no visible growth.

2.5 Genomic DNA karyotyping based on PFGE (Pulse Field Gel Electrophoresis).

Separation of *S. diversa* MJT240 chromosomal DNA was carried out as described by Sipiczki et al. [88] and as modified by El Hage & Houseley[89]. Briefly, yeast chromosomes were separated in 1% agarose gels in 0.5 xTBE buffer cooled at 12 °C in a BioRad CHEF-DRIII electrophoresis apparatus (Bio-Rad, Hercules, CA, USA). Electrophoresis was conducted in different conditions. For separation of low size chromosomal bands, system conditions were, after adjusting run time from 22h to 26h and buffer temperature from 14ºC to 12ºC, voltage gradient of 6 V/cm, switch interval 60-120s and angle 120º. For separation of higher size bands, system conditions were, in a first attempt, voltage gradient of 2 V/cm for 72 h, switch interval of 1800s and angle 106º. After optimization, system conditions were a single run with three consecutive run blocks and voltage gradient of 2 V/cm: block 1, with a run time of 24h with angle 96º and switch interval 1200s; block 2 with a run time of 24h with angle of 100º and switch interval 1500s and block 3, with a run time of 24h with angle 106º and switch interval 1800s. The size markers used to calculate the molecular sizes of *S. diversa* MJT240 chromosomal bands were based on the chromosomes of *S. pombe* (for chromosomal bands ranging 3.5 to 5.7 Mbp) and *S. cerevisiae* (for chromosomal bands ranging 0.225- 2.2 Mbp). Gels were photographed after PFGE using an image acquisition system, model ALLIANCE 4.7, from UVITEC Cambridge.

2.6 Genome sequencing, assembly and annotation of *S. diversa* MJT240.

Genomic DNA extraction of *S. diversa* MJT240, as well as subsequent Illumina sequencing, were performed as a paid service at CD Genomics (Shirley, New York, United States). Genomic DNA of *S. diversa* MJT240 was extracted using Quiagen Magattract HMW kit according to the manufacturer's instructions. After Illumina sequencing, 43,659,976 reads were obtained (with approximately 150 bp on average) and were de novo assembled in SPAdes. The annotation of the resulting contigs was performed in the Geneious software framework (version 2019.2.3) using as a gene finder Augustus trained in *Saccharomyces cerevisiae, Schizosaccharomyces pombe* and *Candida tropicalis*. Manual curation of the three predictions performed was based on BLASTP using the predicted protein sequences as input for searches across the entire dataset of proteins described at NCBI. Only those models for which an orthologue in other species could be identified (having an alignment with an associated e-value of $<10^{-20}$) were considered valid. Using the predicted ORFeome of *S. diversa* MJT240, a putative metabolic reconstruction was performed making use of the KEGG BlastKoala annotation tool and 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 set at the default parameters was also used. For the comparative proteomic analysis between *S. diversa* MJT240 and *S. cerevisiae* EC1118*, S. ludwigii* UTAD17*, H. guilliermondii* UTAD222*, H. uvarum* AWRI3580*, T. delbrueckii* CBS1146*, B. Bruxellensis Z. baili* clib213 and *L. fermentati* CBS6772 pairwise BLASTP analyses were performed using the sets of proteins publicly available at UNIPROT and NCBI 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 other cases, the protein pairs were considered dissimilar.

Chapter 3 Results and discussion

3.1 *S. diversa* MJT240 karyotyping.

To establish the structure of the *S. diversa* MJT240 genome, karyotyping was performed based on pulse-field gel electrophoresis (PFGE). This technology was selected based on the successful prior experience of the team in analysing the genome of other Non-*Saccharomyces* wine yeasts[9, 10]. Considering this information was not previously available for other *S. diversa* strains, the first step was the optimization of the electrophoretic steps to optimize the separation of chromosomal bands. Therefore, different experimental setups were attempted (as detailed in materials and methods), with change of several system conditions, including run and switch time, voltage gradient, buffer temperature, angle and number of blocks. After several attempts, an optimized protocol, previously described to successfully separate *S. cerevisiae* chromosomal bands[90] was defined, which included: a) a single block with a run time of 26h, voltage gradient 6V/cm, switch interval 60-120s and angle 120º, for separation of lower molecular weight bands; b) three consecutive blocks, each with a run time of 24h and voltage gradient of 2 V/cm, with increasing switch interval and angle, for separation of high molecular weight bands. The use of this protocol resulted in the separation of chromosomal bands of *S. diversa* MJT240 that is shown in **Figure 5**.

The results show a good separation of the chromosomes located in the range of 1.0-2.2 Kbp (**Fig.5 panel A**), however, the separation of chromosomal bands with higher molecular weight didn't occur properly. Particularly, it is observed the presence of a high molecular weight band with much higher intensity than the remaining ones, which suggests that in this location there could be more than one chromosomal band, due to low resolution (note the arrow in **Fig.5 panel A**). To confirm this hypothesis, a second PFGE was performed to optimize the separation of the higher molecular weight band, this time increasing the parameters run and switch interval time and decreasing the parameters angle and voltage gradient. The first attempt resulted in a disproportionate migration and diffusion of the high weight bands (**Appendix B**). Therefore, new system conditions were set up, consisting of three-run blocks with increasing angle, resulting in the gel shown in **Figure 5 panel B**. It is evident that the modifications made allowed a better resolution in the separation of the higher molecular weight band, which now appears as two independent chromosomal bands (**Fig.5 panel B**). However, although this change improved the separation of the higher molecular weight bands, it lacked the necessary resolution to separate the lower molecular weight bands that appeared overlapped in the bottom of the gel. Therefore, the final karyotyping of the genomic DNA was determined as a combination of the results obtained in gels shown in **Fig.5** panels A and B. Thus, we can conclude that the *S. diversa* MJT240 genome includes 5 chromosomal bands, totaling a genome size of 9.68 Mb. Sizes of each chromosomal band are presented in **Table 6**.

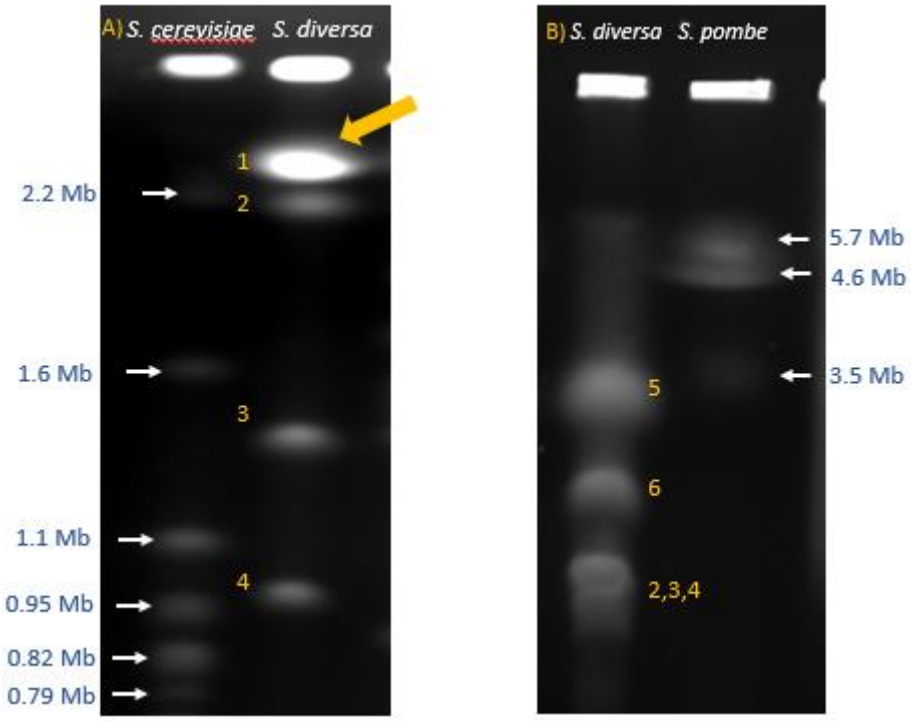


Figure 5- Gel image results after PFGE were performed. Left image (A) refers to the result of PFGE in adjusted conditions of S. cerevisiae PFGE and right image (B) refers to the result in the following conditions. Separated chromosomal bands are named from 1 to 6. *S. cerevisiae* and *S. pombe* markers' chromosome sizes are displayed in the gels.

Table 6- Chromosomal band sizes of *S. diversa* MJT240 strain, separated using PFGE technology, calculated using as a pattern for calibration curves two different markers: *S. cerevisiae* marker[91] for bands 2-4; *S. pombe* marker[92] for bands 5, 6.

Chromosomal band	Size (kb)
2	2164
3	1386
4	1027
5	2933
6	2173

The five chromosomal bands observed for the *S. diversa* strain and the estimated genome size, are considerably lower compared to other yeast species classified in the *Candida* genus, ranging from 8-14 chromosomes and 14.1-30.7 Mb in total genome size (*C. albicans, C. tropicalis, C. glabrata, C. krusei*, etc).[93]. Among wine yeasts, *T. delbrueckii*[94] has 8 chromosomes and 9.2-11.5 Mb of total genome size, *Candida zemplinina*[95] has 3 chromosomes and a total genome size of 9.8Mb, which is highly similar to the obtained genome size of *S. diversa.*

Interestingly, *H. vinae* has been described to have 5 chromosomes and a total genome size ranging 8-9Mb[96], albeit the other members of the *Hanseniaspora* genus have 7 or 8 chromosomes and genome size ranging 9-10.5Mb [96]. These comparisons with other yeasts make it evident that this strain of *S. diversa* has a distinct karyotype, appearing to have fewer chromosomes compared to other known yeasts and a total DNA size in the range of other known yeasts. A noticeable aspect was that the assembled DNA from the whole-genome sequencing resulting from SPAdes de novo assembly was approximately 9.86 Mb, which is similar to the total amount of DNA estimated using the PFGE technique. Nevertheless, it should be noted that a total DNA quantification experiment should be performed to confirm if the sum of the sizes of the described chromosomal bands correspond to the total DNA size of the strain, an information that will be relevant to accurately determine the ploidy of the organism.

3.2 Whole-genome sequencing of *S. diversa* MJT240.

To obtain the whole-genome sequencing of *S. diversa* MJT240, two rounds of Illumina sequencing were performed. Around 43659976 reads were obtained, with an average size of 150 bp. De novo assembly of these reads was performed resorting to two different algorithms, the one embedded in CLC Genomics Workbench software and SPAdes assembler[97], as a plugin in Geneious software. The results of both assemblies are presented in **Table 7** and respective frequency distribution of the contigs assembled with both algorithms is presented in a Histogram, as seen in **Figure 6**.

Table 7- Parameters of the result of de novo assemble of *S. diversa* MJT240 reads performed by two different algorithms, CLC and SPAdes.

Parameters	CLC	SPAdes
Number of contigs	579	1150
% GC	33.8	34.0
Contig minimum size (bp)	200	78
Contig maximum size (bp)	1,021,160	970,026
N50 (excluding scaffolded regions) (bp)	272,524	374,723
Sum of assembled contigs (bp)	9,773,418	9,861,525

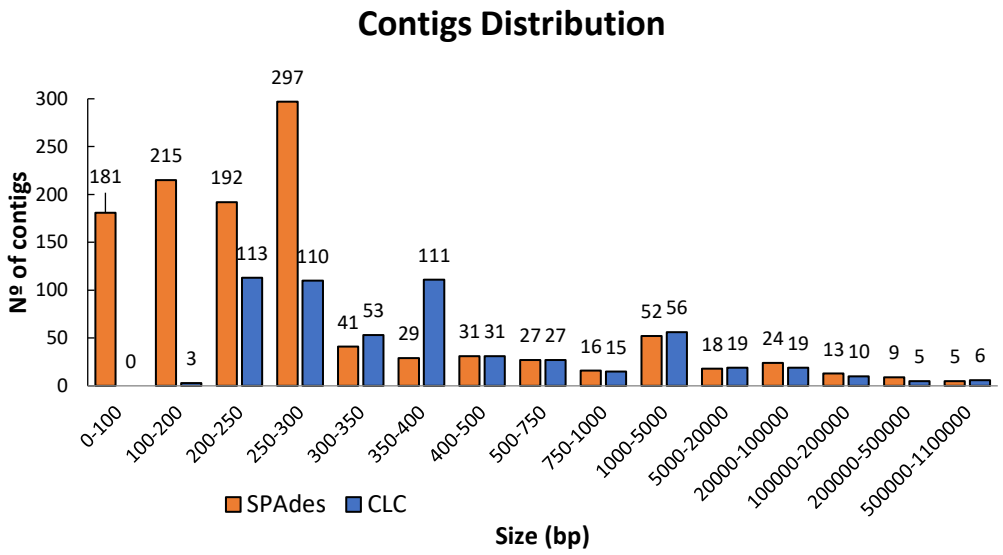


Figure 6- Frequency distribution of contigs assembled using SPAdes de novo assemble algorithm and CLC Genomics incorporated de novo assemble algorithm. SPAdes and CLC de novo assemble resulted in a total of, respectively, 1150 and 579 contigs.

The number of contigs was higher in the assembly provided by SPAdes, however, as it is possible to observe from the histogram, the difference is essentially seen in the smaller contigs range, especially those that are smaller than 300 bp. It is possible that this difference results from parameters embedded in the algorithms, an information that couldn't be retrieved from the software and that may exclude smaller contigs in the case of CLC. The N50 parameter, which takes into account the assembled size of the contigs, was considerably larger in the results provided by this algorithm, SPAdes compared to CLC, as it can be seen by the results shown in **Table 7**. Comparative analysis of the results obtained with both algorithms, involving BLASTN analysis, revealed no significant differences in the information included in contigs assembled by each algorithm. Therefore, taking all parameters into account, the contigs provided by SPAdes were selected for the annotation procedure, since these had a higher N50 and, therefore, are larger, a trait that facilitates gene finding. Using as the criterion of annotation only contigs bigger than 400 bp, 195 contigs were selected for annotation, resulting in total genome size of 9.66 Mb, approximately 98% of the total genome assembled initially by SPAdes.

3.3 Annotation of *S. diversa* MJT240 contigs.

The annotation of the 195 contigs provided by SPADES was performed at the Geneious computational framework which allows ab initio gene prediction using various gene finders. In this work, we resorted to the utilization of the Augustus[98] gene finder plugin and used three different reference datasets: *Saccharomyces cerevisiae*; *Schizosaccharomyces pombe* and *C. tropicalis.* This type of analysis leads to the prediction, for every genomic region, of three gene models (see **Figure 7**) out of which only one was selected and confirmed as a possible *S. diversa* gene. For that selection, manual validation of the gene models was performed resorting to

BLASTP analysis with proteins described at the NCBI protein database. Only those gene models having a described orthologue at NCBI were considered valid, regardless of the reference dataset they were identified from.

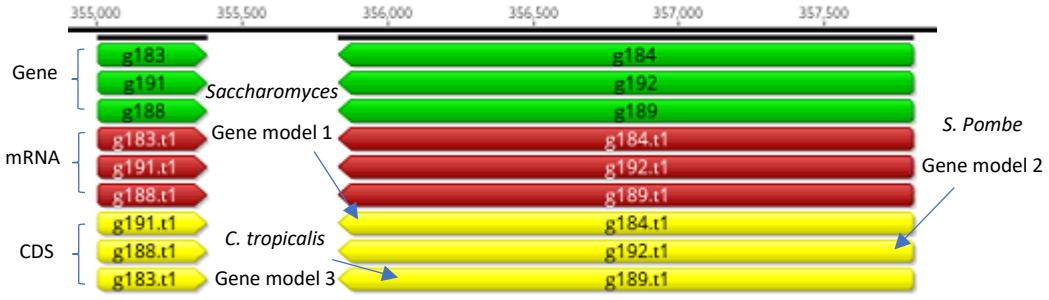
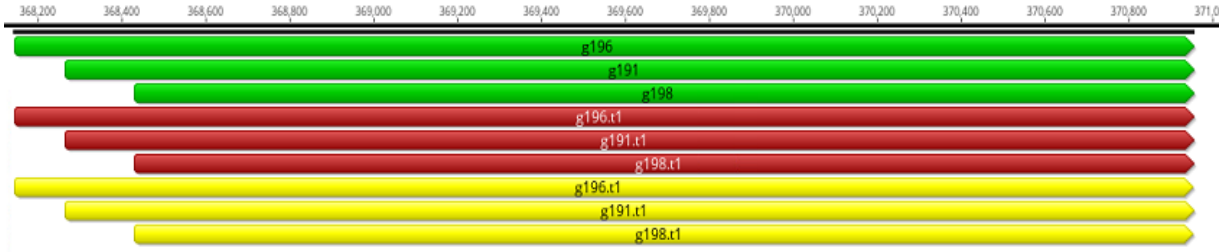


Figure 7- Visualization of the annotation set up in Geneious software, using Find genes with Augustus tool and as reference datasets *S. Pombe, S. cerevisiae and C. tropicalis*. Green lines refer to the gene, red lines refer to mRNA and yellow lines refer to the coding sequence (CDS). Each gene model predicts one gene and corresponding mRNA and CDS.

Below are detailed two examples of situations where the manual validation of gene models was necessary to select what was the better-suited model, among the different predictions made by the algorithms.

Example 1:



Blastp result of Gene 196:

hypothetical protein C6P42_004663 [[Candida] californica]

Sequence ID: KAG0680801.1 Length: 1072 Number of Matches: 1

See 1 more title(s) See all Identical Proteins(IPG)

Range 1: 1 to 798 GenPept Graphics Next Match Pre

Score	Expect	Method	Identities	Positives	Gaps
415 bits(1067)	6e-125	Compositional matrix adjust.	267/807(33%)	443/807(54%)	80/807(9%)

```
Query  41   MTRSSSFKGVKKSRNRIPVSCEECRKKKMKCDKGKPCSSCIKNDCTEHCVY-LRDSMQPA  99
            M ++S    VKK+RNRIP SC+ CRK+K+KCD+ KPCS+C+K+   + C Y +
Sbjct  1    MDKTSPSDRVKKTRNRIPTSCDHCRKRKLKCDRQKPCSNCVKSQTEDMCKYAVSTKTNTN  60

Query  100  VKNSPKKANLKTELLKLKMRLNQMEKVMKAHSIDMSNLDDANLFDSMTKSDSDE------  153
             NS  K N+  E++KLK+++N++EK+++ ++ID+S+   +N+   +    +D+
Sbjct  61   SSNSIPKINMNNEIIKLKLKINKLEKILEMNNIDISSY--SNMLPGVLDDSNDDDDDDS  118

Query  154  ---DEEVSDLLNKFDTMIIKENNTWRSGIVGCWTLTSKDLHIKSLIHPIYSHSESQFSAY  210
               D  VS L NKFD MIIKEN    SG     T   +D  + ++              + +Y
Sbjct  119  NIVDPMVS-LSNKFDAMIIKENKILHSGTTSYVTFIVRDKQLANIFEAYAKRHIMIYESY  177
```

Blastp result of Gene 191:

hypothetical protein C6P42_004663 [[Candida] californica]

Sequence ID: KAG0680801.1 Length: 1072 Number of Matches: 1

See 1 more title(s) See all Identical Proteins(IPG)

Range 1: 1 to 798 GenPept Graphics Next Match Prev

Score	Expect	Method	Identities	Positives	Gaps
416 bits(1070)	9e-126	Compositional matrix adjust.	267/807(33%)	443/807(54%)	80/807(9%)

```
Query  1    MTRSSSFKGVKKSRNRIPVSCEECRKKKMKCDKGKPCSSCIKNDCTEHCVY-LRDSMQPA  59
            M ++S    VKK+RNRIP SC+ CRK+K+KCD+ KPCS+C+K+   + C Y +
Sbjct  1    MDKTSPSDRVKKTRNRIPTSCDHCRKRKLKCDRQKPCSNCVKSQTEDMCKYAVSTKTNTN  60

Query  60   VKNSPKKANLKTELLKLKMRLNQMEKVMKAHSIDMSNLDDANLFDSMTKSDSDE------  113
             NS  K N+  E++KLK+++N++EK+++ ++ID+S+   +N+   +    +D+
Sbjct  61   SSNSIPKINMNNEIIKLKLKINKLEKILEMNNIDISSY--SNMLPGVLDDSNDDDDDDS  118

Query  114  ---DEEVSDLLNKFDTMIIKENNTWRSGIVGCWTLTSKDLHIKSLIHPIYSHSESQFSAY  170
               D  VS L NKFD MIIKEN    SG     T   +D  + ++              + +Y
Sbjct  119  NIVDPMVS-LSNKFDAMIIKENKILHSGTTSYVTFIVRDKQLANIFEAYAKRHIMIYESY  177
```

Blastp result of Gene 198:

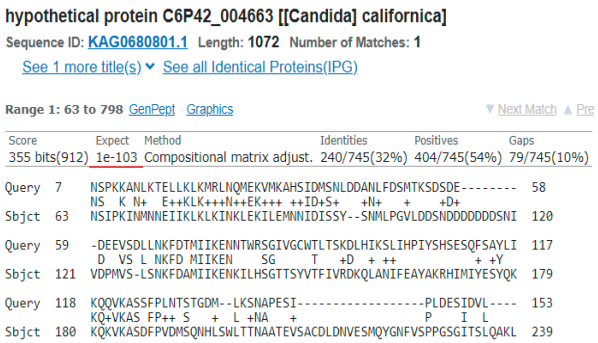


Figure 8- Example 1 describes a situation in manual curation of annotation, where the three gene models have a sequence with the same START codon and different STOP codon. The first image refers to the visualization in the Geneious framework of the three gene model predictions. The other images refer to the protein alignment result in NCBI for each amino acid sequence of gene models. In alignment images, the e-value is identified with a red line below.

In this example, the three predicted gene models have the same STOP codon but have a different START codon. Therefore, the three automatic amino acid sequences predicted by each reference dataset were used as inputs to search the NCBI protein database for putative homologues and the one having associated the better alignment were selected as the valid model. From the visualization of the results shown in **Figure 8,** it is clear the three predicted gene models have associated good *e-value*s, however, gene model 198 has a considerably higher e-value, suggesting this is not the correct gene prediction. To differentiate between the other two models, it was necessary to observe the alignment between the automatic amino acid translation (Query) and the amino acid sequence that had a better alignment (subject). That analysis revealed the alignment obtained with gene model 191 was better compared to the one obtained with gene model 196, as it was observed a coincidence of the START codons in both subject and query. On the contrary, Query alignment with subject sequence using gene model 196 only starts at amino acid number 41. In view of all this, from the three possible predictions, the one validated was gene model 191, as it is the most complete. The validated CDS is represented below, with respective gene and mRNA:

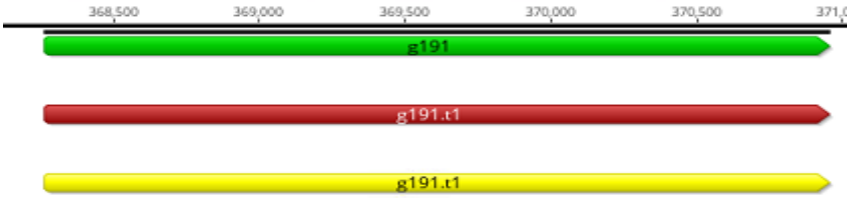
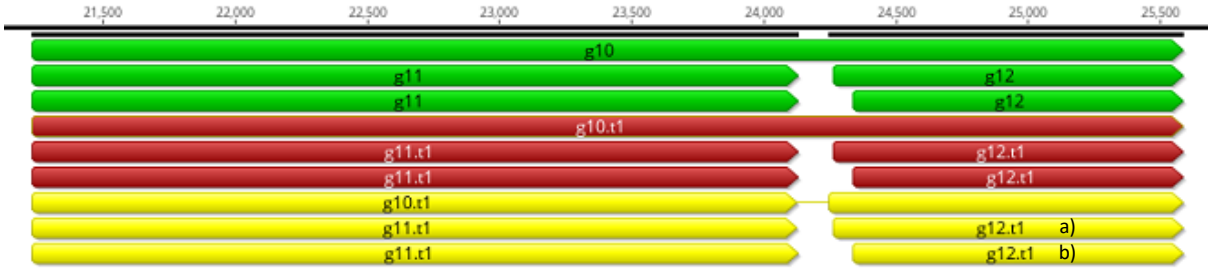
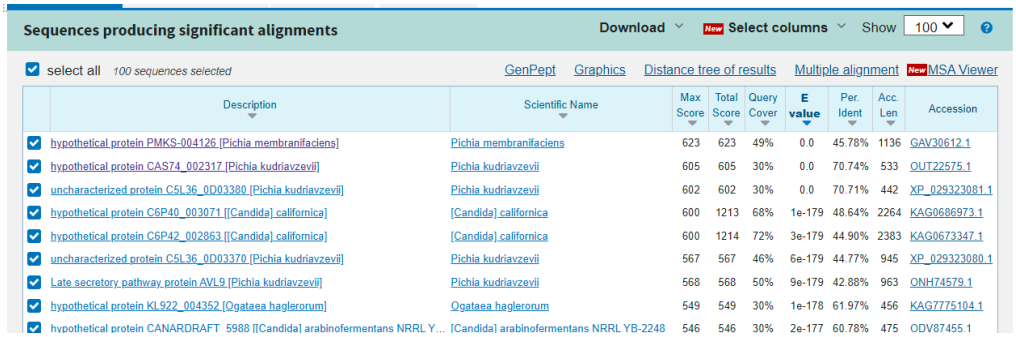


Figure 9- Visualization in Geneious framework of the validated gene model 191.

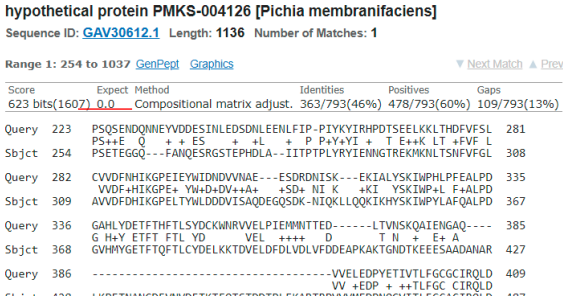
Example 2:



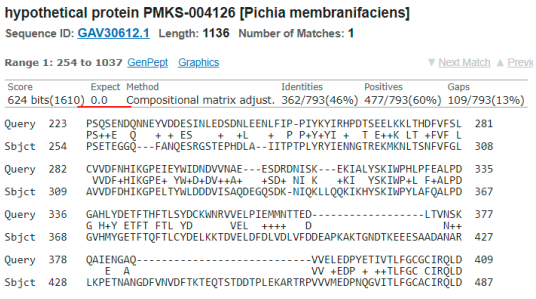
Blastp result of Gene 10:



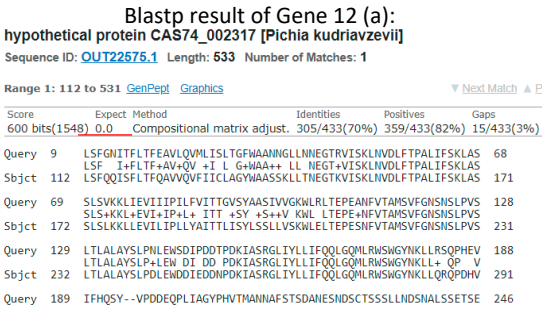
Blastp result of Gene 10:



Blastp result of Gene 11:



Blastp result of Gene 12 (a):



Blastp result of Gene 12 (b):

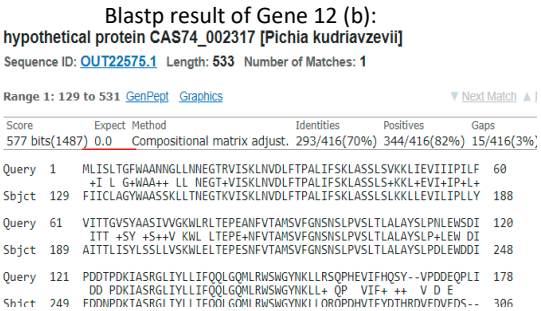


Figure 10- Example 2 describes a situation in manual curation of annotation, where one gene model is a possible gene with an intron (g10) and the other gene models are two separated possible genes. The first image refers to the visualization in the Geneious framework of the three gene model predictions. The other images refer to the protein alignment result in NCBI for each amino acid sequence of gene models. In alignment images, the e-value is identified with a red line below.

In this second example, one gene model predicted a larger gene with an intron, and the other two gene models predicted two separated genes, albeit with a minor change in the START codon. BLASTP of the automatic amino acid translations obtained from the gene models was performed to determine which gene structure should be validated, the one having an intron or the one considering two independent genes. Since all gene model results in BLASTP have a great e-value, it is not possible to take any conclusion using this tool. By visualizing the sequences producing significant alignments with the amino acid sequence of gene 10, those do not include an intron and have similar sizes to gene model 11. Furthermore, alignment in gene models 11 and 12 is better, as also identity is much higher. These observations suggest the gene with an intron is not the correct gene model and may be discarded. Therefore, gene model 11 is validated as the possible gene. It is still necessary to determine which of gene models 12 a) and 12 b) is more accurate, as they have a different START codon. Although in both gene models, the subject sequence doesn't start at START codon, query sequence of gene 12 a) has more amino acids aligned to subject sequence, compared to gene model 12 b), thus having a better alignment. Therefore, gene model 12 a) is more complete and, therefore, is validated as the possible gene, whereas gene 12 b) is discarded. The final validated CDSs are represented below:

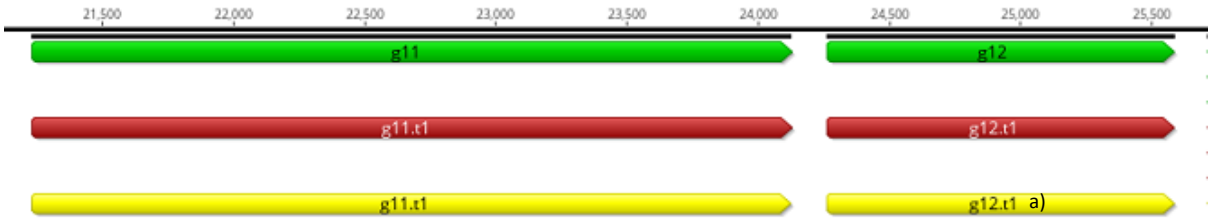


Figure 11- Visualization in Geneious framework of the validated gene models g11 and g12 a).

At the end of annotation, after manual curation of all predicted gene models, the predicted ORFeome is estimated in 5070 protein-encoding genes. The incidence of introns is ~ 3.16%, which is in line with the intron's normal percentage in other yeasts such as *S. kluyvery*, *Z. rouxii* and *H. guilliermondi*[9, 99]. More specifically, there were identified 150 genes with two exons, 5 genes with three exons and other five with more than three exons. The mean length of all gene sequences is 484 amino acids, also similar to other yeasts[99]. These and other features are observable in **Table 8**.

Table 8- General features of *S. diversa* MJT240 genome annotation performed in Geneious software framework, after manual curation.

Annotation features	Value
Nº of genes	5070
max length (amino acid)	4980
min length (amino acid)	34
average length (amino acid)	484
GC %	34.5
nº of introns	160
% introns	3.16
genes with 2 exons	150
genes with 3 exons	5
genes with +3 exons	5

3.4 Metabolic reconstruction of *S. diversa* MJT240.

The predicted ORFeome of *S. diversa* MJT240 was uploaded (in a txt file) to the BlastKoala tool[100]. This tool, provided by Kyoto Encyclopedia of Genes and Genomes (KEGG)[101], performs an in silico prediction of a biological function for the proteins, distributing them in functional classes and, particularly, among the pathways that are described in the database. With this tool, it is possible to perform expedite conclusions about the completeness of the different metabolic pathways. Using this methodology it was possible to conclude, among other things, that the *Hanseniaspora guilliermondii* species has incomplete neoglucogenesis and glyoxylate cycle[9] and that *Saccharomycodes ludwigii* harbours genes for catabolism of N-acetylglucosamine[10]. A total of 67 pathway modules with all blocks completed were obtained including the central metabolic pathways (glycolysis, gluconeogenesis, pyruvate oxidation, citrate cycle, glyoxylate cycle, pentose phosphate pathway [PPP], oxidative phosphorylation and fatty acid catabolism). An overview of the metabolic reconstruction of *S. diversa* MJT240 is presented in **Figure 12**. Genes required for catabolism of starch, sucrose, xylose and mannose were found, this being consistent with the demonstrated ability of *S. diversa* cells *to* grow in media using these sugars (or others) as sole carbon sources tested in our laboratory, as seen in **Figure 13**. In the case of xylose, a divergence was found because in the assay performed it was not visible growth, which could result from the absence of genes encoding for a non-specific aldose reductase, part of the cofactor associated with xylitol formation that is necessary for xylose assimilation (see schematic representation in **Figure 14**).

Enzymes for the anabolism/catabolism of glucuronate, related to interconversion of less common sugars, are also missing, thus this pathway is also interrupted, as well as galactose (it was not possible to detect a homologue for an aldose 1-epimerase and galactokinase) and acetate catabolism (it was not identified a homologue for an acetate kinase). The lack of these genes is consistent with the inability of *S. diversa* cells to grow in media with galactose as the carbon source or sodium acetate as the acetate source (**Figure 13**). However, it should be noted that *S. diversa* MJT240 should have capacity to assimilate acetate, as a complete glyoxylate cycle was

determined. It is possible no growth was registered due to the down-regulation of some genes in this pathway, or due to yeast's sensitivity to sodium.

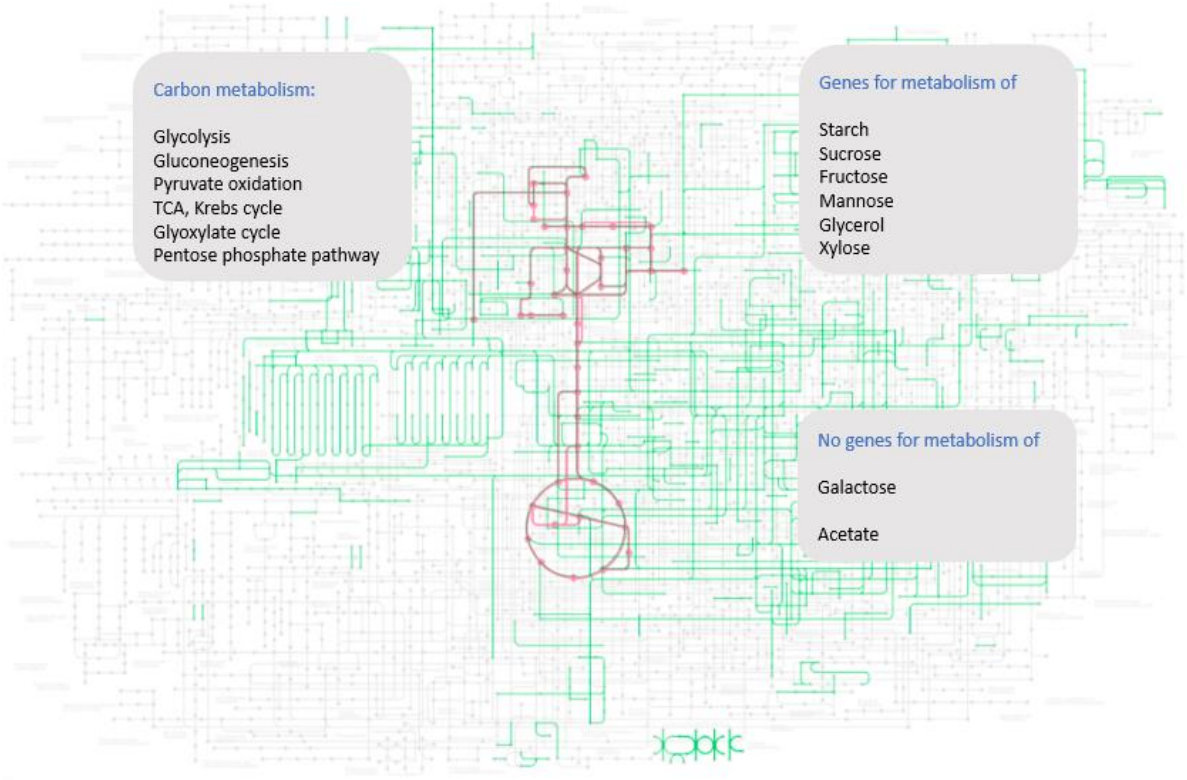


Figure 12- Metabolic reconstruction of *S. diversa* MJT240 focused on carbohydrate metabolism, as uncovered by the KEGG reconstruction mapper tool. Yeast has all central metabolism pathways - glycolysis, gluconeogenesis, pyruvate oxidation, citrate cycle, glyoxylate cycle, pentose phosphate pathway, oxidative phosphorylation – and genes were found for fructose, mannose, starch, glycerol and xylose metabolism, although not for galactose nor acetate.

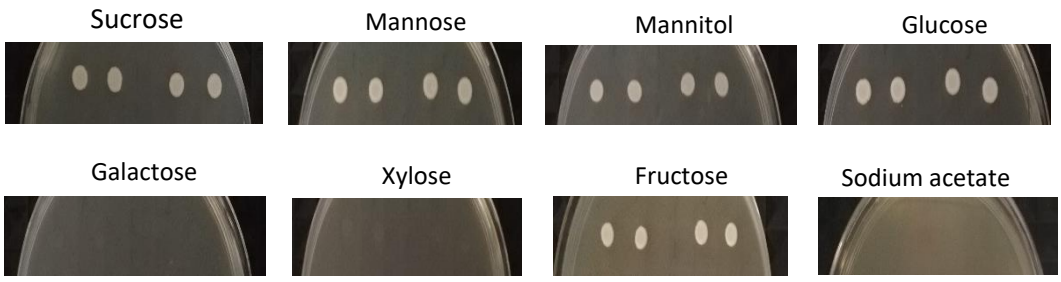


Figure 13- Growth of *S. diversa* MJT240 in YPD agar plates with different sugars as sole carbon sources, using as technique spot assays containing *S. diversa* (DO=0.02). Replicates were used for each spot. Used carbon sources were fructose mannose, galactose, mannitol, sorbitol, sucrose, maltose, raffinose, rhamnose and xylose. Only some are visualized here. For full visualization of experiment, visit **Appendix E.**

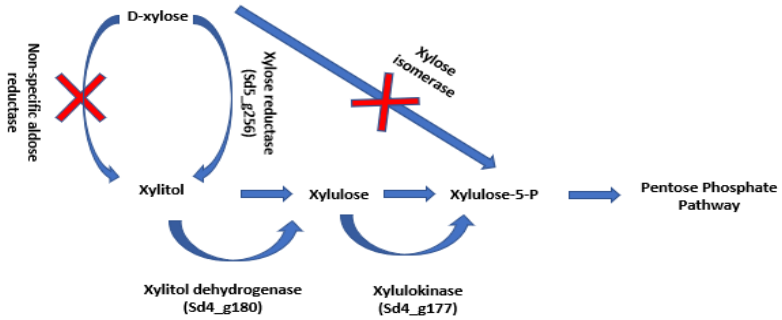


Figure 14– Scheme reaction of D- xylose assimilation pathway. Genes encoding for a nonspecific aldose reductase or xylose isomerase were not found in *S. diversa* MJT240 ORFeome.

The pathway for catabolism/anabolism of Inositol phosphate metabolism is interrupted, as it lacks one gene encoding for the enzyme myo-inositol-1-monophosphatase. A reaction scheme for both pathways is presented in **Figure 15**.

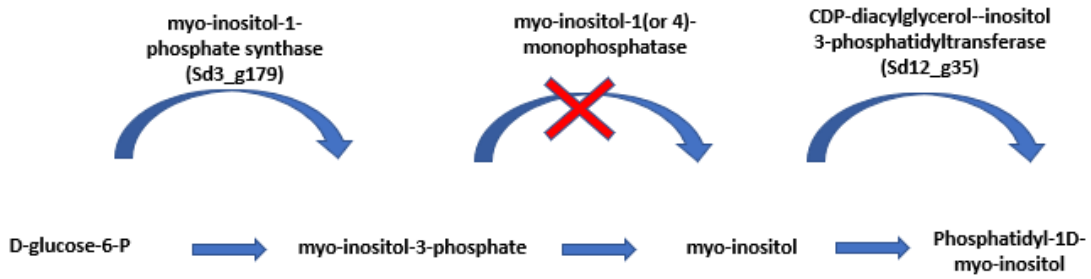


Figure 15– Schematic reaction of Myo-inositol biosynthesis pathway. All genes encoding for the necessary enzymes were found except for the enzyme myo-inositol-1-monophosphatase, as visible in the scheme.

Concerning glycerolipid metabolism, *S. diversa* MJT240 has several genes encoding for enzymes related to glycerol biosynthesis, mainly through the acylglycerol degradation pathway. All genes encoding for enzymes necessary in the triacylglycerol biosynthesis pathway were found, as seen in **Figure 16**. It is a relatively interesting finding that *S. diversa* is suggested to produce glycerol since the compound is reported to have a favourable impact in terms of sweetness and fullness in alcoholic beverages like wine. [102] Expectedly, all the enzymes required for the biosynthesis of ergosterol were found among the *S. diversa* MJT240 ORFeome.

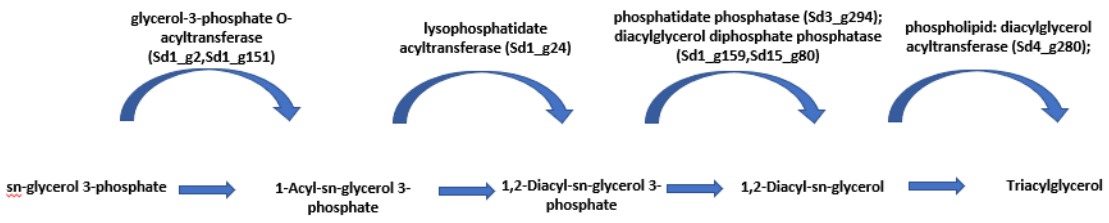


Figure 16– Scheme reaction of triacylglycerol biosynthesis pathway. All genes encoding for the necessary enzymes were found, as visible in the scheme.

With respect to nitrogen metabolism[103], it was not possible to find in the genome of *S. diversa* genes encoding for enzymes nitrate and nitrite reductase, which suggests that it may be unable to assimilate nitrate or nitrite. Furthermore, no orthologue genes encoding for urea carboxylase were detected, suggesting that it might also be unable to use urea as the sole nitrogen source. Genes encoding for enzymes required for the biosynthesis of all the proteogenic amino acids were found as well as four ribose-phosphate pyrophosphokinase encoding genes, involved in the biosynthesis of phosphoribosyl diphosphate (PRPP). This metabolite is an important intermediate in cellular metabolism, particularly in nitrogen metabolism, as it is required for the biosynthesis of the amino acids histidine and tryptophan[104].

Concerning sulphur metabolism, *S. diversa* MJT240 was found to encode all the genes necessary for assimilatory sulphate reduction pathway, as also for the methionine salvage pathway, something that appears to be interesting since this pathway was not found in *Hanseniaspora* yeasts such as *H. guilliermondi* and an *S. ludwigii* strain[9, 10, 105]. One of the functions of the methionine salvage pathway is to supply precursors for the synthesis of polyamines, such as putrescine, spermidine and spermine. Genes encoding for enzymes involved in the synthesis of these polyamines, with exception of spermine, were all found. The yeast's ability to synthesize spermidine is very interesting as the latter has been demonstrated to be an essential nutrient for *S. cerevisiae* growth[106], which could be a determining factor for its high growth rate (see results below).

Similarly, all the genes required for the synthesis of other important vitamins/cofactors, such as pyridoxal-5-phosphate (PLP), pantothenate, CoA, lipoic acid and ubiquinone were found. PLP is known to be an essential cofactor in *S. cerevisiae* for enzymes involved in glucose, amino acid and lipid metabolism as well as for thiamine regulation[107]. Pantothenate is also considered an important growth factor for yeasts and a metabolic precursor for CoA, also an essential cofactor for pathways in lipid metabolism and oxidative respiration. These vitamins are all normally synthesized by *S. cerevisiae*.[108] With respect to the biosynthesis of thiamine, only genes encoding for a thiamine pyrophosphokinase and alkaline and acid phosphatase were detected among the ORFeome of *S. diversa* MJT240, whereas to biotin biosynthesis, only a gene encoding for biotin synthase was found, thus suggesting this yeast species may be auxotrophic for these nutrients. A schematic representation of both biosynthetic pathways is represented in **Figure 17**, with the respective enzymes and compounds. Thiamine has a vital role in driving fermentation due to its central role as a cofactor of several key enzymes in central carbon metabolism, namely pyruvate decarboxylase and pyruvate dehydrogenase, among others[109]. Biotin also plays a key role as a coenzyme of several important coenzymes involved in fatty acids synthesis, amino acids and sugar metabolism[107]. Other wine yeasts, like those belonging to the *Hanseniaspora* genus, were also found to be auxotrophic for thiamine and biotin biosynthesis, this being suggested to underlie their lower fermentation rates[9]. In the case of *S. diversa*, although no fermentations were carried on, growth experiments were performed in synthetic GJM, in both single and co-culture with *S. ludwigii* (see results below). The results showed *S. diversa* MJT240 growth rate is higher than other wine yeasts, including an *S. cerevisiae* strain[110], thus its growth rate appears to not be affected by the absent biosynthesis of these vitamins. It should be noted this relatively high growth rate of *S. diversa* was observed in the GJM medium, which is enriched with all vitamins, including thiamine and biotin, thus masking what can be a difficulty of this yeast in synthesizing these nutrients.

It will be essential to determine if the yeast's possible auxotrophy for thiamine and biotin will have an impact on the growth rate in real must, a highly competitive environment.

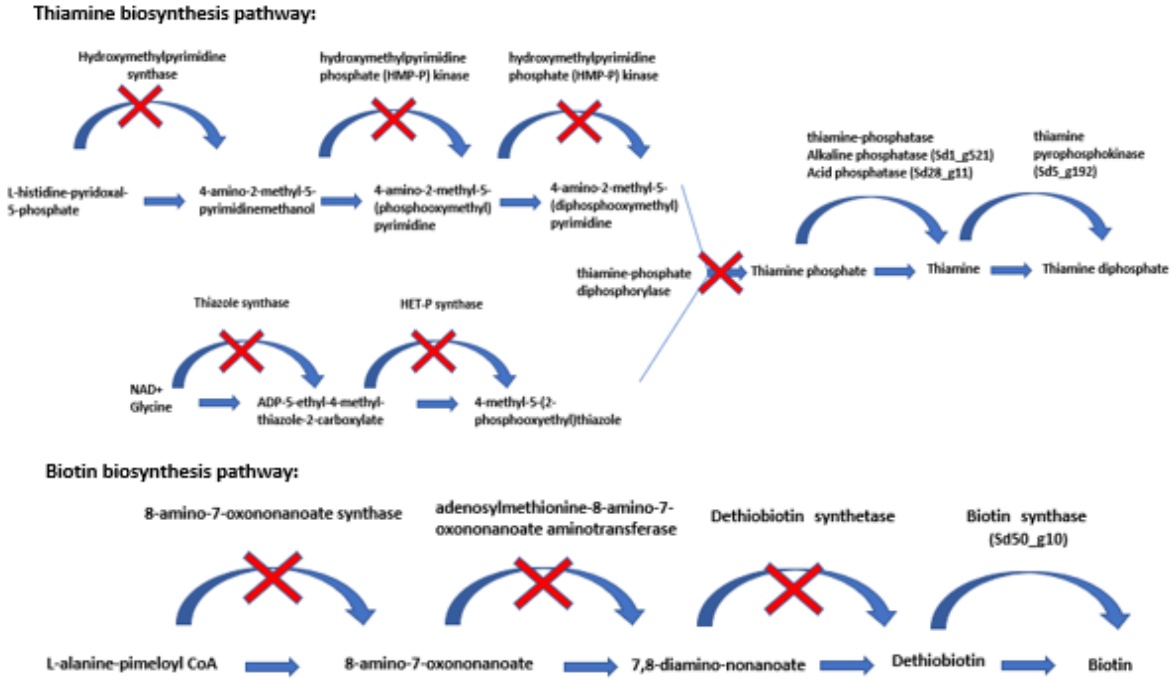


Figure 17- Scheme reaction of vitamins Thiamine and Biotin in *S. diversa MJT240*. In the thiamine pathway, only genes encoding for alkaline and acid phosphatase and thiamine pyrophosphokinase were found, whereas in the Biotin pathway only one gene encoding for biotin synthase was found.

An interesting observation was the detection of complete riboflavin (RF) biosynthesis pathway in the genome of *S. diversa MJT240*, as seen in **Figure 18**. Therefore, it was important to analyse this finding to determine if the yeast has flavinogenic traits, or if it only produces riboflavin like almost all other yeasts. The only flavinogenic yeast that produced industrial concentrations of RF is *C. Famata*, however, due to the low stability of the yeast, the industrial process was stopped[111]. *S. diversa* contains all genes encoding for necessary proteins for RF synthesis. Industrial producers of RF, *B. subtillis* and *A. gossypii* share the biosynthesis of two important precursors for the production of this vitamin: guanine triphosphate (GTP) and ribulose-5-P[112]. All the genes for the synthesis of these precursors were identified in *S. diversa* MJT240. The biosynthetic pathways that lead to the production of RF are, as seen in **Figure 19**, β-oxidation in the peroxisome, TCA cycle, gluconeogenesis, Entner Doudoroff pathway, PPP, purine pathway and finally flavin synthesis[111]. *S. diversa* has all these complete pathways, except for the Entner Doudoroff pathway, which is incomplete, as no gene encoding for enzyme necessary the synthesis of 2-dehydro-3-deoxy-phosphogluconate, an important precursor of riboflavin biosynthesis, was detected in the genome of *S. diversa* MJT240. Also, in flavinogenic molds, the transcriptional repression of RF synthesis is not caused by flavins, but by iron ions, which is regulated by the putative transcription factor encoded by the SEF1 gene[113], that was not found in the *S. diversa* MJT240 genome. In other yeasts used for industrial production of RF, overexpression of genes RIB1 and RIB7, which are also present in *S.*

diversa, resulted in an overproduction of RF[111]. Increment of RF synthesis is obtained in industrial RF microbial cells by metabolic engineering of the microorganisms: overexpression of purine biosynthesis, optimization of central metabolism, enhanced synthesis of glycine, etc[111]. It is not possible to confirm, with full certainty, that this yeast is not flavinogenic, as there are several factors, described above, that impact synthesis of RF and can only be determined experimentally,like the level of expression of flavinogenic genes. Nevertheless, the incomplete Entner Doudoroff pathway and the absence of the SEF1 gene suggest the yeast shouldn't be flavinogenic.

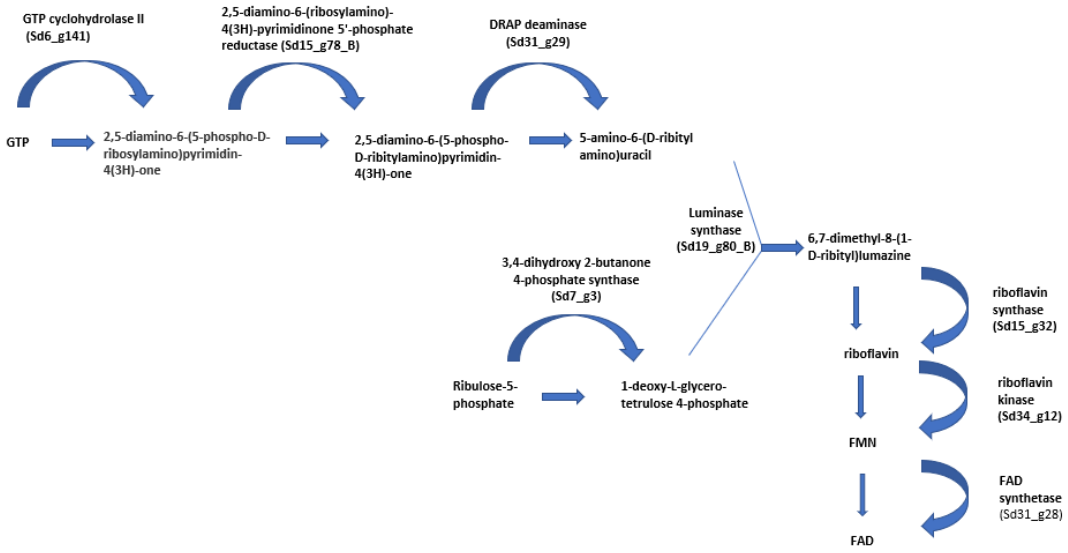


Figure 18– Scheme reaction of riboflavin biosynthesis, and cofactors FMN and FAD, in *S. diversa* MJT240. The pathway is complete, as genes encoding for necessary enzymes required were all found.

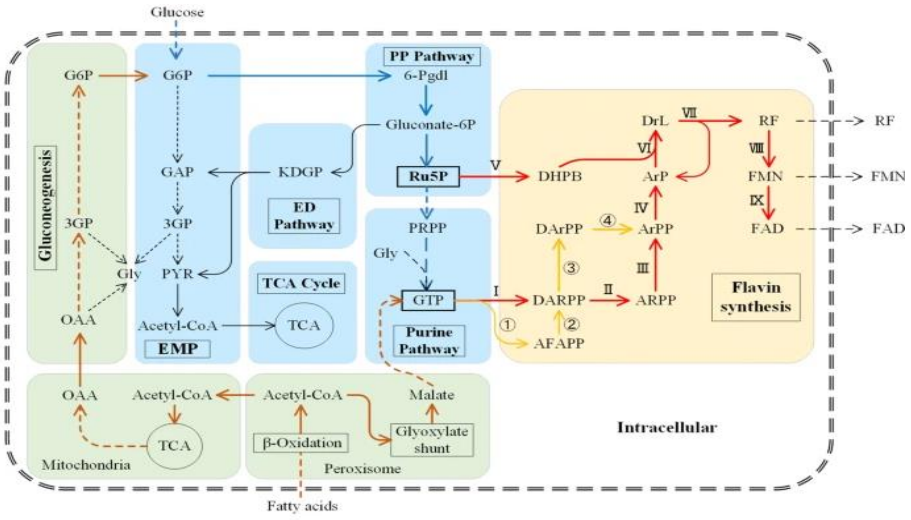


Figure 19– Schematic overview of riboflavin biosynthesis process in microorganisms, including the pathways leading from the substrate to flavins: β-oxidation in peroxisome, TCA cycle, gluconeogenesis, Entner Doudoroff pathway, PPP and purine biosynthesis pathway.

3.5 Comparative analysis of the predicted ORFeome of *Saturnispora diversa* MJT240 with other wine yeast species

To determine the extent to which the ORFeome of this *S. diversa's* strain is similar to the proteomes described for other wine yeast species, a comparative proteomic analysis was performed. For that, the sequences of proteins predicted to be encoded by *S. diversa* MJT240 were compared, using BLASTP, with those described to be encoded by *H. uvarum, H. guilliermondi* and *S. ludwigii* from the *Saccharomycodeacea* family, the *S. cerevisiae* strain EC1118, *T. delbrueckii, L. fermentati*, all from the family *Saccharomycetacea*. The spoilage yeasts *B. bruxellensis* and *Z. bailli* were also selected for this analysis. Some yeasts were of particular interest to be compared with *S. diversa* MJT240 ORFeome, but the annotated Proteome wasn't available. That was the case of *Metschnikowia pulcherrima* (anamorph *Candida pulcherrima*) and *Candida zemplinina/ Candida stellata*, as they belong to the previous genus of *S. diversa* and are considered wine yeasts. The pairwise protein comparison resulted in the classification of proteins in three datasets: highly similar (whenever the e-value of the associated alignment was below e-50), similar (whenever the e-value was between e^{-20} and e^{-50}) and dissimilar (whenever the e-value of the associated alignment was higher than e-20). The results obtained are summarized in the chart shown in **Figure 20.**

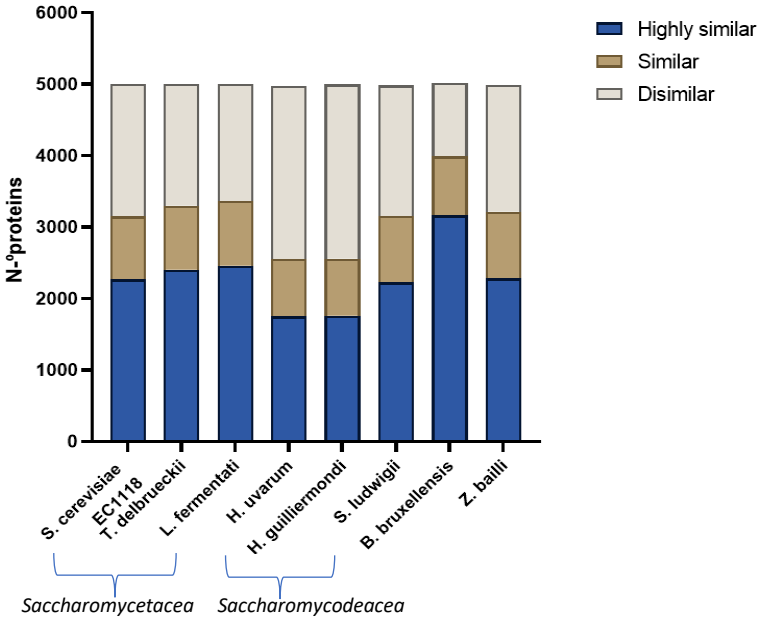


Figure 20– Comparative analysis of *S. diversa* MJT240 Proteome with other wine yeasts using pairwise Blastp alignments. The graphic shows proteins with high similarity (E-value < 10^{-50}), similarity (E-value $10^{-50} > x < 10^{-20}$) and dissimilarity (E-value > 10^{-20}). Wine yeast to whom *S. diversa 240* was compared is *S. cerevisiae EC1118, T. delbrueckii* CBS1146*, L. fermentati* CBS6772*,* from *Saccharomycetacea* family, *H. guilliermondi* UTAD222*, H. uvarum* AWRI3580 *and S. ludwigii* UTAD17, from *Saccharomycodeacea* family, and spoilage wine yeasts *B. bruxellensis* and *Z, bailii clib213*.

The results obtained showed the lowest degree of similarity was obtained with the *H. guilliermondii* and *H. uvarum* proteomes that together, comprise a group, that seems more dissimilar compared to all the other yeast species examined. Notably, a similar conclusion was drawn upon the comparison of the proteome of *S. ludwigii* with several yeast species, including *Hanseniaspora* species. The highest degree of similar proteins was obtained with *B. bruxellensis,* while for the remaining species the number of orthologues pairs was somewhat identical. Further studies are now required to better elucidate exactly the proteins that are identical in *S. diversa* and *B. bruxellensis* and that diverge more in the other yeasts, as well as the subsequent physiological outcomes of that. To pinpoint proteins that could be more specific of *S. diversa*, the dataset of all dissimilar proteins identified from the pairwise comparison analysis undertaken with *T. delbrueckii, L. fermentati, H. guilliermondi* and *S. ludwigii* were analysed. The results are presented in the Venn diagram shown in **Figure 21**.

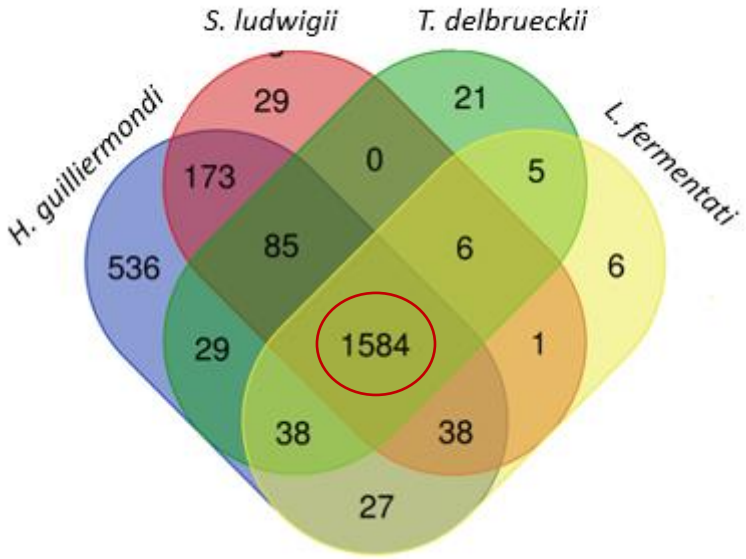


Figure 21– *S. diversa* MJT240 proteins considered to be dissimilar compared to others found in yeast species *T. delbrueckii, H. uvarum, S. ludwigii, C. glabrata* and *L. fermentati* were compared to determine the set of proteins specific to *S. diversa* MJT240. The results are shown in Venn's plot. A set of 1584 proteins were determined to be unique for *S. diversa* MJT240.

Thereby, it was possible to identify the set of *S. diversa* proteins dissimilar from those found in the other yeast species. From the original 5070 annotated ORFs, 1584 were only identified in *S. diversa* MJT240. A substantial part of these are hypothetical proteins or have poorly characterized functions. The set of proteins with characterized function is visible in **Appendix C**. It will now be interesting to compare this set of proteins with those found in *B. bruxellensis*, to identify which could be the set of *S. diversa* more "specific" proteins. It was interesting to find in the set of ORFs identified only in *S. diversa,* eight enzymes involved in protein glycosylation: two oligosaccharyltransferases (Sd48_g3 and Sd5_g14), involved in the first step of protein glycosylation - transfer of a glycan to an asparagine residue-; a beta-1,4-N-acetylglucosaminyltransferase (Sd3_g98), linked to the generation of N-acetyl-D-glucosamine-beta-1,4-O-D-mannosylprotein; a phosphatidylinositol N-acetylglucosaminyltransferase subunit P (Sd2_g29), involved in the first step of GPI biosynthesis; a GPI-anchor transamidase subunit S (Sd3_g259) and a GPI ethanolamine phosphate transferase subunit F (Sd27_g6). Also, it

was uncovered the presence of a dolichol-phosphate mannosyltransferase (Sd5_g238), whose function is to form dolichol phosphate mannose, the mannosyl donor in pathways that lead to N-glycosylation and a dolichol phosphate-mannose regulatory protein (Sd1_g45_B), that regulates the function of dolichol-phosphate mannosyltransferase and is also involved in the first step of GPI biosynthesis. Interestingly, an ethanolamine kinase (Sd36_g17) was also identified in the set of proteins specific to *S. diversa*, this being an enzyme that catalyzes the synthesis of phosphatidylethanolamine, a pivotal compound for GPI biosynthesis, involved in several steps of this pathway. Further studies will now be required to investigate whether these proteins are also absent from the proteomes of other wine yeast species, and, more importantly, what could be their physiological function in the context of cell physiology. Seven proteins related to oxidative phosphorylation metabolism, namely in Complex I - NADH dehydrogenase, were also identified in the set of *S. diversa* proteins but not in *T. delbrueckii, L. fermentati, H. guilliermondi* and *S. ludwigii*. Notably, these enzymes are also not present in *S. cerevisiae*[114]. However, it is not clear whether the presence of these proteins has any effect on oxidative phosphorylation, as Complex I is composed of 42 subunits and several genes encoding for necessary proteins were not identified in *S. diversa* MJT240 ORFeome, mainly those related to NADH dehydrogenase alpha and beta subcomplex.

3.6 Prediction of genes of *S. diversa* MJT240 with impact on wine aroma

Several Non-*Saccharomyces* wine strains have shown potential to improve wine aroma by producing aromatic molecules that are not produced, or that are produced in very low concentrations by *S. cerevisiae*. At least one study has reported the usefulness of *S. diversa* in this context[56]. Therefore, although it is not known whether such a beneficial effect could also be observed for the strain used in this work, the genome sequence was searched for genes encoding for enzymes involved in the production of volatile aroma compounds. Reconstruction of *S. diversa* MJT240 metabolic network showed that this species is equipped with genes encoding for enzymes required for the synthesis of higher alcohols, ethyl esters, acetaldehydes (see **Appendix D**), volatile compounds that have a positive impact on wine aroma. However, the predicted proteome didn't include beta-glucosidases or alcohol acetyltransferases, involved in the production of acetate esters. One β-mannosyltransferase (Sd22_g55), previously described to mediate the incorporation of β-1,2-linked oligomannosides in the cell wall[115], and 5 dolichol-phosphate-mannose-protein mannosyltransferases were uncovered (Sd1_g56, Sd6_g20, Sd7_g136, Sd9_g56, Sd10_g9), known to be involved in cell wall incorporation of several proteins[116], which could be linked with the ability to produce mannoproteins, a relevant trait since these improve mouthfeel sensation. Notably, these findings are consistent with the aroma profile derived from the fermentation of Touriga nacional musts with *S. diversa* [56], in which an intense aroma was detected in comparison with other wines from other yeasts. Specifically, the associated volatile compounds to the aroma descriptors that had more impact in those wines (nutty and dried fruits), benzaldehyde, esters and isopentyl alcohol, are synthesized by the enzymes above described to be encoded by *S. diversa* MJT240. In the future, it

will be relevant to determine experimentally if these volatile aromatic compounds are in fact produced by this strain of *S. diversa* and, if so, it will be important to determine if an impact in wine resulting aroma is observed.

3.7 Effect of the presence of *S. diversa* MJT240 in the growth of the contaminating yeast *Saccharomycodes ludwigii* during fermentation.

Considering the previously described potential of *S. diversa* MJT240 as a bio-control agent of several phytopathogenic fungal species[80-83], including those that target grape vines, it was herein decided to examine whether the strain in use in this thesis also shows an interesting biocontrol potential against the contaminating yeast *Saccharomycodes ludwigii* UTAD17. The assays were undertaken in synthetic grape juice medium (GJM) since it was not possible to use grape must. To better characterize the physiology of *S. diversa* and *S. ludwigii*, preliminary single culture assays were performed to determine the growth kinetics of these species in synthetic grape juice medium. The results obtained are shown in **Figures 22 and 23**. As in the co-culture setting it was impossible to distinguish between *S. diversa* and *S. ludwigii* cells based on OD, a method was devised based on the counting of CFUs. For that, aliquots of the co-culture were plated in YPD+ 6mM SO_2 to count *S. ludwigii* cells (*S. diversa* cells do not grow under these conditions) while to count *S. diversa* cells, an aliquot of the mixed-culture was plated in WL medium, in which the yeast species can be distinguished based on their different morphology and colour [87]. This difference is visible in **Figure 24**, where *S. diversa* colonies exhibit an intense green colour with a dark centre and a flat round morphology, whereas *S. ludwigii* colonies have a creamy/white colour and exhibit a convex, pointy morphology. Furthermore, *S. diversa* growth in agar plates was shown to be much faster than *S. ludwigii* growth which facilitated counting of the colonies.

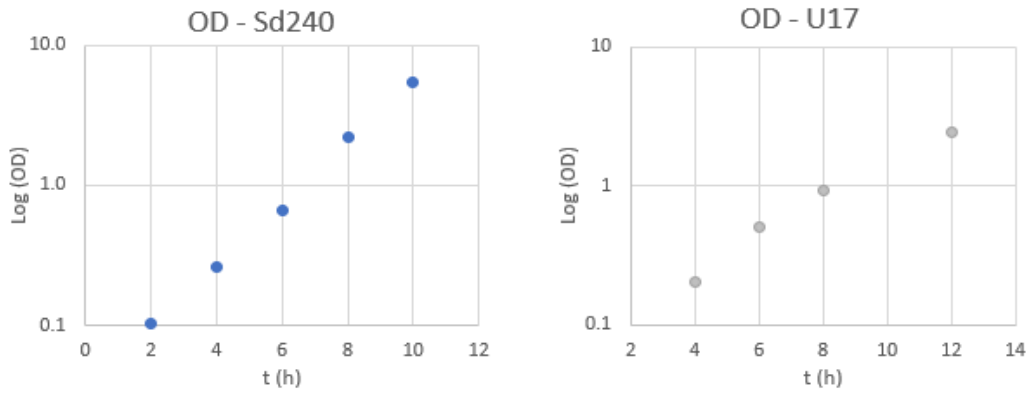


Figure 22- Graphic representation of OD vs time for *S. diversa* MJT240 (right) and *S. ludwigii* UTAD17 (left). Data is presented on a logarithmic scale. Only data of the exponential phase is shown. Three replicates were made.

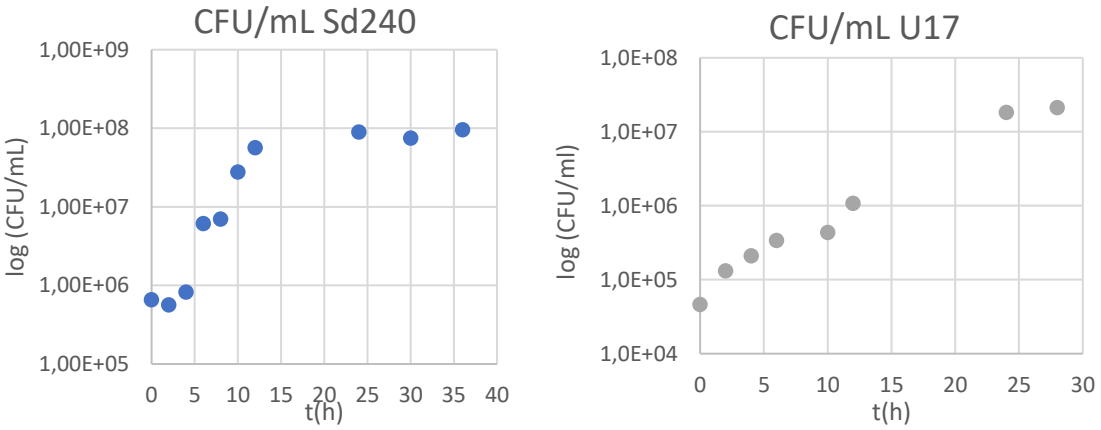


Figure 23- Graphic representation of CFU/mL vs time for *S.diversa* MJT240 (right) and UTAD17 (left). Data is presented on a logarithmic scale. Data is shown until the stationary phase is reached. Three replicates were made.

Trendlines of *S. diversa240* and *UTAD17* are, respectively:

$$y = 0.6267x - 2.0318 \qquad y = 0.282x - 1.1023$$

Specific growth rate (μ) is retrieved directly from the trendline slope, and the doubling time (Td) is calculated with **Equation 1**. Parameters are presented in **Table 9**.

$$Td = \frac{\mathrm{Ln}\,2}{\mu} \qquad (\mathbf{1})$$

Table 9- Growth kinetics of *S. diversa* MJT240 and *S. ludwigii* UTAD17: specific growth rate, μ, presented in h^{-1} and min^{-1}, and doubling time, Td (min).

Sd240		**U17**	
μ (h-1)	0.6267	**μ (h-1)**	0.2820
μ (min-1)	0.0104	**μ (min-1)**	0.00470
Td (min)	66.4	**Td (min)**	147

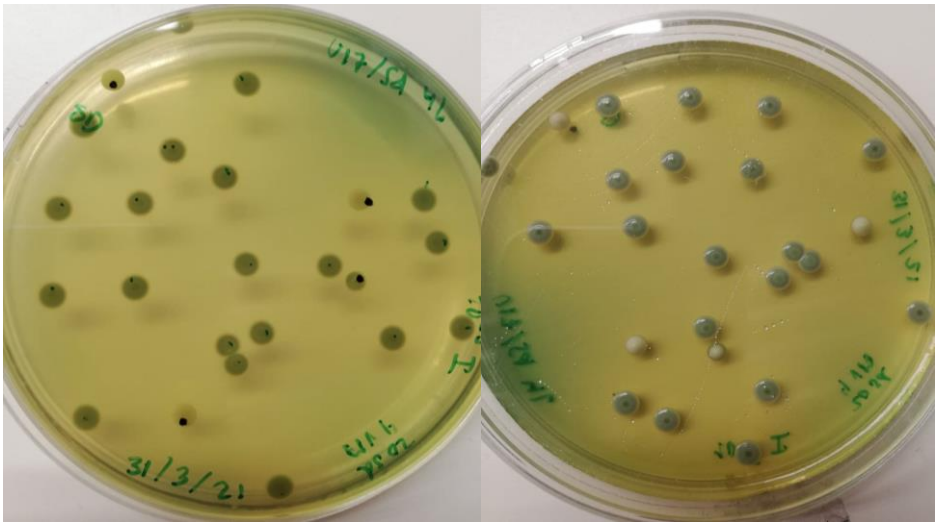


Figure 24- Growth of *S. diversa* MJT240 and *S. ludwigii* UTAD17colonies in WL agar medium. *S. diversa* MJT240colonies present an intense green colour with a dark centre, and *S. ludwigii* UTAD17 colonies present a creamy/white colour.

The results obtained concerning growth dynamics of the co-cultures established between *S. diversa* MJT240 and *S. ludwigii* UTAD17 are shown in **Figures 25 and 26**.

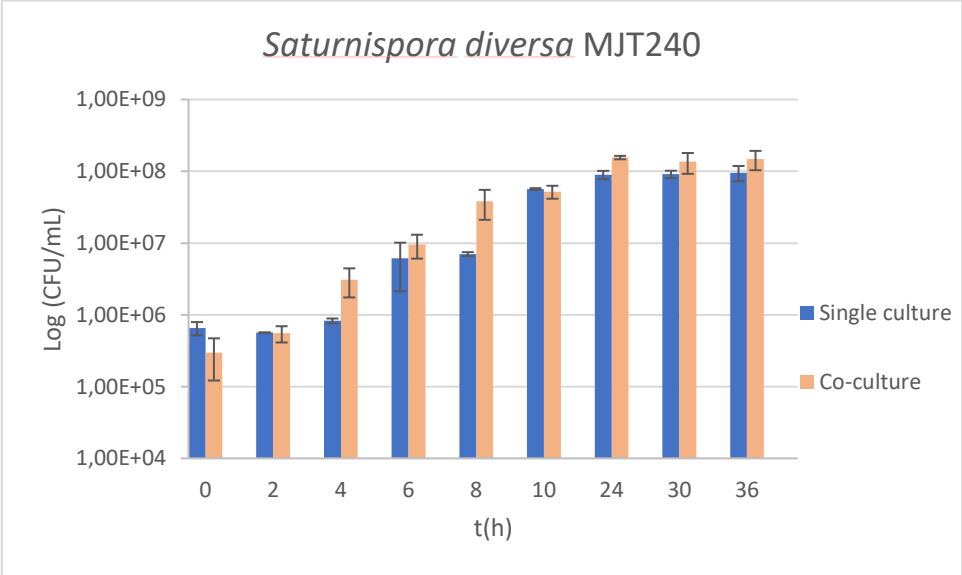


Figure 25- Graphic representation of *S. diversa's* growth (Sd240) in co-culture assays along time (h) using CFU counting as a growth measuring tool. Data is presented on a logarithmic scale. *S. diversa* MJT240 cells in co-culture were plated in WL medium. Single culture growth assay is presented for comparison. Three replicates were made.

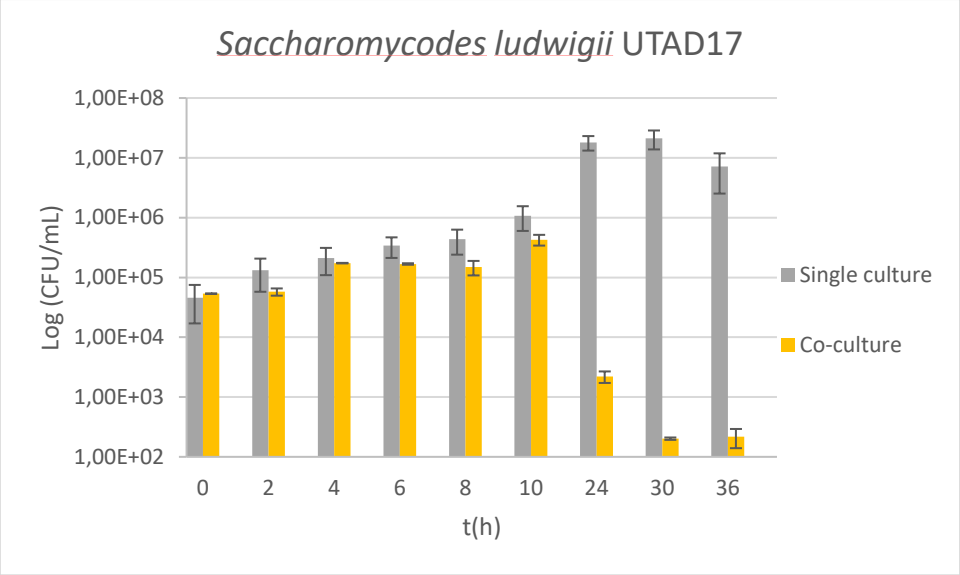


Figure 26- Graphic representation of *S. ludwigii's* growth (U17) in co-culture assays along time (h) using CFU counting as a growth measuring tool. Data is presented on a logarithmic scale. U17 cells in co-culture were plated in YPD+ 6mM SO2 medium. Single culture growth assay is presented for comparison. Three replicates were made.

The results obtained show no significant differences in growth of *S. diversa* MJT240 while in single or in co-culture (**Fig.24**), thus, there was little effect in this species by the presence of *S. ludwigii* UTAD17. Differently, in the case of *S. ludwigii* a marked reduction in cellular viability was registered after 10h of co-cultivation with the viability decreasing significantly, compared to the values observed in single culture (**Fig.25**)- from $1.82x10^7$ CFU/mL to $2.02x10^2$ CFU/mL, a percentage loss of 99.998%. This loss of viability was very interesting and appears to point in the direction towards a biocontrol activity by the *S. diversa* strain, however, further assays should be performed to determine that. It is possible that the reduction in *S. ludwigii* viability results from a competition of nutrients since *S. diversa* has a doubling time much lower than UTAD17 and, therefore, consumes sugar at a much faster rate. However, this alone shouldn't account for a reduction in cellular viability that appears more to result from a toxicity mechanism that may involve, for example, the production by *S. diversa* of an antimicrobial peptide or a compound whose accumulation is deleterious for *S. ludwigii*.

Final Considerations

In this work, the first genomic sequence, dully annotated, of a strain belonging to the species *Saturnispora diversa*, MJT240, was determined. A brief exploration of the ORFeome uncovered the finding of several traits of *S. diversa* MJT240, including the presence of all central metabolic pathways and its suggested auxotrophy for the vitamins thiamine and biotin. With the genomic sequence determined, it will now be possible to perform a more extensive metabolic analysis of the yeast, for instance, facilitating the execution of Proteome and Transcriptome assays. As for future perspectives, it will be important to evaluate whether the volatile aromatic compounds described before are in fact produced by *Saturnispora diversa* MJT240 and if the impact on the aroma profile is significant. Concerning its phenotypic behaviour, it will also be essential to determine whether the phenotype of biocontrol by *S. diversa* MJT will also be observed in real musts and not only in GJM. It will also be interesting to determine if *S. diversa* MJT240 presence has the potential to inhibit the growth of other spoilage yeasts, namely *B. bruxellensis* and *Z. baili.* And, necessarily, it will be essential to further detail the underlying inhibitory mechanism, something that the availability of the genomic sequence now turns possible to inspect, for example, at the transcriptomic level.

Chapter 4 References

1. Drumonde-Neves, J., et al., *Association between Grape Yeast Communities and the Vineyard Ecosystems.* PLOS ONE, 2017. **12**(1): p. e0169883.
2. Čadež, N., J. Zupan, and P. Raspor, *The effect of fungicides on yeast communities associated with grape berries.* FEMS Yeast Research, 2010. **10**(5): p. 619-630.
3. Raspor, P., et al., *Yeasts isolated from three varieties of grapes cultivated in different locations of the Dolenjska vine-growing region, Slovenia.* International Journal of Food Microbiology, 2006. **109**(1): p. 97-102.
4. Mezzasalma, V., et al., *Geographical and Cultivar Features Differentiate Grape Microbiota in Northern Italy and Spain Vineyards.* Frontiers in Microbiology, 2018. **9**(946).
5. Zott, K., et al., *Dynamics and diversity of non-Saccharomyces yeasts during the early stages in winemaking.* International Journal of Food Microbiology, 2008. **125**(2): p. 197-203.
6. Guildsomm. *Winemaking*. 2020 15/10/2021].
7. Jolly, N.P., C. Varela, and I.S. Pretorius, *Not your ordinary yeast: non-Saccharomyces yeasts in wine production uncovered.* FEMS Yeast Research, 2014. **14**(2): p. 215-237.
8. Liu, Y., et al., *Wine microbiome: A dynamic world of microbial interactions.* Critical Reviews in Food Science and Nutrition, 2017. **57**(4): p. 856-873.
9. 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, 2018. **26**(1): p. 67-83.
10. Tavares, M.J., et al., *Genome sequencing, annotation and exploration of the SO2-tolerant non-conventional yeast Saccharomycodes ludwigii.* BMC Genomics, 2021. **22**(1): p. 131.
11. Comitini, F., et al., *New insights on the use of wine yeasts.* Current Opinion in Food Science, 2017. **13**: p. 44-49.
12. Mateo, J.J. and S. Maicas, *Application of Non-Saccharomyces Yeasts to Wine-Making Process.* Fermentation, 2016. **2**(3): p. 14.
13. Wang, C., A. Mas, and B. Esteve-Zarzoso, *The Interaction between Saccharomyces cerevisiae and Non-Saccharomyces Yeast during Alcoholic Fermentation Is Species and Strain Specific.* Frontiers in Microbiology, 2016. **7**(502).
14. Morata, A., *Enological Repercussions of Non-Saccharomyces Species in Wine Biotechnology.* Fermentation, 2019. **5**(3): p. 72.
15. Comitini, F., et al., *Selected non-Saccharomyces wine yeasts in controlled multistarter fermentations with Saccharomyces cerevisiae.* Food Microbiology, 2011. **28**(5): p. 873-882.
16. Ciani, M. and L. Ferraro, *Enhanced Glycerol Content in Wines Made with Immobilized Candida stellata Cells.* Applied and Environmental Microbiology, 1996. **62**(1): p. 128-132.
17. Bely, M., et al., *Impact of mixed Torulaspora delbrueckii–Saccharomyces cerevisiae culture on high-sugar fermentation.* International Journal of Food Microbiology, 2008. **122**(3): p. 312-320.
18. Ciani, M., et al., *Controlled mixed culture fermentation: a new perspective on the use of non-Saccharomyces yeasts in winemaking.* FEMS Yeast Research, 2010. **10**(2): p. 123-133.
19. Branco, P., et al., *Antimicrobial peptides (AMPs) produced by Saccharomyces cerevisiae induce alterations in the intracellular pH, membrane permeability and culturability of Hanseniaspora guilliermondii cells.* International Journal of Food Microbiology, 2015. **205**: p. 112-118.
20. Sipiczki, M., *Overwintering of Vineyard Yeasts: Survival of Interacting Yeast Communities in Grapes Mummified on Vines.* Frontiers in Microbiology, 2016. **7**(212).
21. Garofalo, C., et al., *From grape berries to wine: population dynamics of cultivable yeasts associated to "Nero di Troia" autochthonous grape cultivar.* World J Microbiol Biotechnol, 2016. **32**(4): p. 59.

22. Di Maro, E., D. Ercolini, and S. Coppola, *Yeast dynamics during spontaneous wine fermentation of the Catalanesca grape.* International Journal of Food Microbiology, 2007. **117**(2): p. 201-210.
23. Pinto, C., et al., *Wine fermentation microbiome: a landscape from different Portuguese wine appellations.* Frontiers in Microbiology, 2015. **6**(905).
24. Castrillo Cachón, D., et al., *Yeast diversity on grapes from Galicia, NW Spain: biogeographical patterns and the influence of the farming system: Macrowine 2018.* OENO One, 2019. **53**(3).
25. Borren, E. and B. Tian, *The Important Contribution of Non-Saccharomyces Yeasts to the Aroma Complexity of Wine: A Review.* Foods, 2021. **10**(1): p. 13.
26. Pina, C., et al., *Ethanol tolerance of five non-Saccharomyces wine yeasts in comparison with a strain of Saccharomyces cerevisiae—influence of different culture conditions.* Food Microbiology, 2004. **21**(4): p. 439-447.
27. Capozzi, V., et al., *Microbial terroir and food innovation: The case of yeast biodiversity in wine.* Microbiological Research, 2015. **181**: p. 75-83.
28. Rojas, V., et al., *Studies on acetate ester production by non-saccharomyces wine yeasts.* Int J Food Microbiol, 2001. **70**(3): p. 283-9.
29. Ciani, M. and F. Fatichenti, *Selective sugar consumption by apiculate yeasts.* Letters in Applied Microbiology, 1999. **28**(3): p. 203-206.
30. Englezos, V., et al., *Exploitation of the non-Saccharomyces yeast Starmerella bacillaris (synonym Candida zemplinina) in wine fermentation: physiological and molecular characterizations.* Int J Food Microbiol, 2015. **199**: p. 33-40.
31. Magyar, I. and T. Tóth, *Comparative evaluation of some oenological properties in wine strains of Candida stellata, Candida zemplinina, Saccharomyces uvarum and Saccharomyces cerevisiae.* Food Microbiology, 2011. **28**(1): p. 94-100.
32. Rodríguez, M.E., et al., *Selection and preliminary characterization of β-glycosidases producer Patagonian wild yeasts.* Enzyme and Microbial Technology, 2007. **41**(6): p. 812-820.
33. Guth, H., *Identification of Character Impact Odorants of Different White Wine Varieties.* Journal of Agricultural and Food Chemistry, 1997. **45**(8): p. 3022-3026.
34. Morata, A., et al., *Contribution of Non-Saccharomyces Yeasts to Wine Freshness. A Review.* Biomolecules, 2020. **10**(1): p. 34.
35. Oro, L., M. Ciani, and F. Comitini, *Antimicrobial activity of Metschnikowia pulcherrima on wine yeasts.* Journal of Applied Microbiology, 2014. **116**(5): p. 1209-1217.
36. van Breda, V., N. Jolly, and J. van Wyk, *Characterisation of commercial and natural Torulaspora delbrueckii wine yeast strains.* International Journal of Food Microbiology, 2013. **163**(2): p. 80-88.
37. Ramírez, M. and R. Velázquez, *The Yeast Torulaspora delbrueckii: An Interesting But Difficult-To-Use Tool for Winemaking.* Fermentation, 2018. **4**(4): p. 94.
38. Renault, P., et al., *Enhanced 3-Sulfanylhexan-1-ol Production in Sequential Mixed Fermentation with Torulaspora delbrueckii/Saccharomyces cerevisiae Reveals a Situation of Synergistic Interaction between Two Industrial Strains.* Frontiers in Microbiology, 2016. **7**(293).
39. Malfeito-Ferreira, M., *Yeasts and wine off-flavours: a technological perspective.* Annals of Microbiology, 2011. **61**(1): p. 95-102.
40. Blomqvist, J. and V. Passoth, *Dekkera bruxellensis—spoilage yeast with biotechnological potential, and a model for yeast evolution, physiology and competitiveness.* FEMS Yeast Research, 2015. **15**(4).
41. Malfeito-Ferreira, M. and A.C. Silva, *Spoilage Yeasts in Wine Production*, in *Yeasts in the Production of Wine*, P. Romano, M. Ciani, and G.H. Fleet, Editors. 2019, Springer New York: New York, NY. p. 375-394.
42. Agnolucci, M., et al., *Brettanomyces bruxellensis yeasts: impact on wine and winemaking.* World Journal of Microbiology and Biotechnology, 2017. **33**(10): p. 180.

43. Capozzi, V., et al., *Viable But Not Culturable (VBNC) state of Brettanomyces bruxellensis in wine: New insights on molecular basis of VBNC behaviour using a transcriptomic approach.* Food Microbiology, 2016. **59**: p. 196-204.
44. Vejarano, R., *Saccharomycodes ludwigii, Control and Potential Uses in Winemaking Processes.* Fermentation, 2018. **4**(3): p. 71.
45. Malfeito-Ferreira, M., *Chapter 15 - Spoilage Yeasts in Red Wines*, in *Red Wine Technology*, A. Morata, Editor. 2019, Academic Press. p. 219-235.
46. Aranda, A., *Enological Repercussions of Non-Saccharomyces Species.* Fermentation, 2019. **5**(3): p. 68.
47. Zuehlke, J.M., B. Petrova, and C.G. Edwards, *Advances in the Control of Wine Spoilage by Zygosaccharomyces and Dekkera/Brettanomyces.* Annual Review of Food Science and Technology, 2013. **4**(1): p. 57-78.
48. de Barros Pita, W., et al., *The biotechnological potential of the yeast Dekkera bruxellensis.* World Journal of Microbiology and Biotechnology, 2019. **35**(7): p. 103.
49. 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.* International Journal of Food Microbiology, 2011. **147**(3): p. 170-180.
50. Lemos Junior, W.J.F., et al., *Whole genome comparison of two Starmerella bacillaris strains with other wine yeasts uncovers genes involved in modulating important winemaking traits.* FEMS Yeast Research, 2018. **18**(7).
51. Lachance, M.-A., et al., *Chapter 90 - Candida Berkhout (1923)*, in *The Yeasts (Fifth Edition)*, C.P. Kurtzman, J.W. Fell, and T. Boekhout, Editors. 2011, Elsevier: London. p. 987-1278.
52. Kurtzman, C.P., *Description of Martiniozyma gen. nov. and transfer of seven Candida species to Saturnispora as new combinations.* Antonie van Leeuwenhoek, 2015. **108**(4): p. 803-809.
53. Holloway, P., R.E. Subden, and M.A. Lachance, *The Yeasts in a Riesling Must From the Niagara Grape-Growing Region of Ontario.* Canadian Institute of Food Science and Technology Journal, 1990. **23**(4): p. 212-216.
54. Zabukovec, P., N. Čadež, and F. Čuš, *Isolation and Identification of Indigenous Wine Yeasts and Their Use in Alcoholic Fermentation.* Food technology and biotechnology, 2020. **58**(3): p. 337-347.
55. Kumla, J., et al., *Evaluation of Multifarious Plant Growth Promoting Trials of Yeast Isolated from the Soil of Assam Tea (Camellia sinensis var. assamica) Plantations in Northern Thailand.* Microorganisms, 2020. **8**(8): p. 1168.
56. Teixeira, A., I. Caldeira, and F.L. Duarte, *Molecular and oenological characterization of Touriga Nacional non-Saccharomyces yeasts.* Journal of Applied Microbiology, 2015. **118**(3): p. 658-671.
57. OHARA, Y., NONOMURA, I-L and YUNOME, I-[*Candida vhtaria Ohara, Nonomura et Yunome, Candida fimetaria Soneda, Torulopsis bacillaris (Kroemer et Krumbholz) Lodder*

var. obesa (en japonais). Agr. Chem. Soc. Japan 34, 1960.
58. VAN UDEN, N.a.B., H, *Candida Berkhout, p.893-1087.* The Yeasts, a taxonomic study, 2nd ed., 1970.
59. Kurtzman, C.P. and C.J. Robnett, *Identification and phylogeny of ascomycetous yeasts from analysis of nuclear large subunit (26S) ribosomal DNA partial sequences.* Antonie van Leeuwenhoek, 1998. **73**(4): p. 331-371.
60. Kurtzman, C.P., *New species and new combinations in the yeast genera Kregervanrija gen. nov., Saturnispora and Candida.* FEMS Yeast Research, 2006. **6**(2): p. 288-297.
61. Liu, Z. and C.P. Kurtzman, *Phylogenetic relationships among species ofWilliopsis andSaturnospora gen. nov. as determined from partial rRNA sequences.* Antonie van Leeuwenhoek, 1991. **60**(1): p. 21-30.
62. Kurtzman, C.P., *Saturnospora ahearnii, a new salt marsh yeast from Louisiana.* Antonie Van Leeuwenhoek, 1991. **60**(1): p. 31-4.

63. James, S.A., et al., *Saturnispora quitensis sp. nov., a yeast species isolated from the Maquipucuna cloud forest reserve in Ecuador.* International Journal of Systematic and Evolutionary Microbiology, 2011. **61**(12): p. 3072-3076.
64. Morais, C.G., et al., *Saturnispora bothae sp. nov., isolated from rotting wood.* International Journal of Systematic and Evolutionary Microbiology, 2016. **66**(10): p. 3810-3813.
65. Kurtzman, C.P., *Chapter 64 - Saturnispora Liu & Kurtzman (1991)*, in *The Yeasts (Fifth Edition)*, C.P. Kurtzman, J.W. Fell, and T. Boekhout, Editors. 2011, Elsevier: London. p. 765-772.
66. El-Samawaty, A.E.-R.M.A., et al., *Saturnispora mangrovi f.a., sp. nov. from Syhat mangrove, Saudi Arabia.* International Journal of Systematic and Evolutionary Microbiology, 2020. **70**(2): p. 977-981.
67. Drumonde-Neves, J., et al., *Learning from 80 years of studies: a comprehensive catalogue of non-Saccharomyces yeasts associated with viticulture and winemaking.* FEMS Yeast Research, 2021. **21**(3).
68. Khunnamwong, P. and S. Limtong, *Saturnispora kantuleensis f.a., sp. nov., a novel yeast species isolated from peat in a tropical peat swamp forest in Thailand.* International Journal of Systematic and Evolutionary Microbiology, 2018. **68**(4): p. 1160-1164.
69. Morais, P.B., M.-A. Lachance, and C.A. Rosa, *Saturnispora hagleri sp. nov., a yeast species isolated from Drosophila flies in Atlantic rainforest in Brazil.* International Journal of Systematic and Evolutionary Microbiology, 2005. **55**(4): p. 1725-1727.
70. Nielsen, D.S., M. Jakobsen, and L. Jespersen, *Candida halmiae sp. nov., Geotrichum ghanense sp. nov. and Candida awuaii sp. nov., isolated from Ghanaian cocoa fermentations.* International Journal of Systematic and Evolutionary Microbiology, 2010. **60**(6): p. 1460-1465.
71. Boonmak, C., et al., *Candida siamensis sp. nov., an anamorphic yeast species in the Saturnispora clade isolated in Thailand.* FEMS Yeast Research, 2009. **9**(4): p. 668-672.
72. Morais, P.B., F.M.P. de Sousa, and C.A. Rosa, *Yeast in plant phytotelmata: Is there a "core" community in different localities of rupestrian savannas of Brazil?* Brazilian Journal of Microbiology, 2020. **51**(3): p. 1209-1218.
73. Lentz, M., et al., *Genetic and physiological characterization of yeast isolated from ripe fruit and analysis of fermentation and brewing potential.* Journal of the Institute of Brewing, 2014. **120**(4): p. 559-564.
74. Jaiboon, K., et al., *Yeasts from peat in a tropical peat swamp forest in Thailand and their ability to produce ethanol, indole-3-acetic acid and extracellular enzymes.* Mycological Progress, 2016. **15**(7): p. 755-770.
75. Berta Gonçalves, I.O., Eunice Bacelar, Maria Cristina Morais, Alfredo Aires, Fernanda Cosme, Jorge Ventura-Cardoso, Rosário Anjos and Teresa Pinto, *Aromas and Flavours of Fruits, Generation of Aromas and Flavours*. 2018, IntechOpen.
76. Russo, F., et al., *Characterisation of volatile compounds in Cilento (Italy) figs (Ficus carica L.) cv. Dottato as affected by the drying process.* International Journal of Food Properties, 2017. **20**(sup2): p. 1366-1376.
77. Buzzini, P., et al., *A study on volatile organic compounds (VOCs) produced by tropical ascomycetous yeasts.* Antonie van Leeuwenhoek, 2003. **84**(4): p. 301-311.
78. Petronilho, S., et al., *Revealing the Usefulness of Aroma Networks to Explain Wine Aroma Properties: A Case Study of Portuguese Wines.* Molecules, 2020. **25**(2): p. 272.
79. Ferreira, V., *1 - Volatile aroma compounds and wine sensory attributes*, in *Managing Wine Quality*, A.G. Reynolds, Editor. 2010, Woodhead Publishing. p. 3-28.
80. Li, G., et al., *Stress tolerance and biocontrol performance of the yeast antagonist, Candida diversa, change with morphology transition.* Environ Sci Pollut Res Int, 2016. **23**(3): p. 2962-7.
81. Liu, J., et al., *Transcriptome profiling reveals differential gene expression associated with changes in the morphology and stress tolerance of the biocontrol yeast, Pichia cecembensis.* Biological Control, 2018. **120**: p. 36-42.

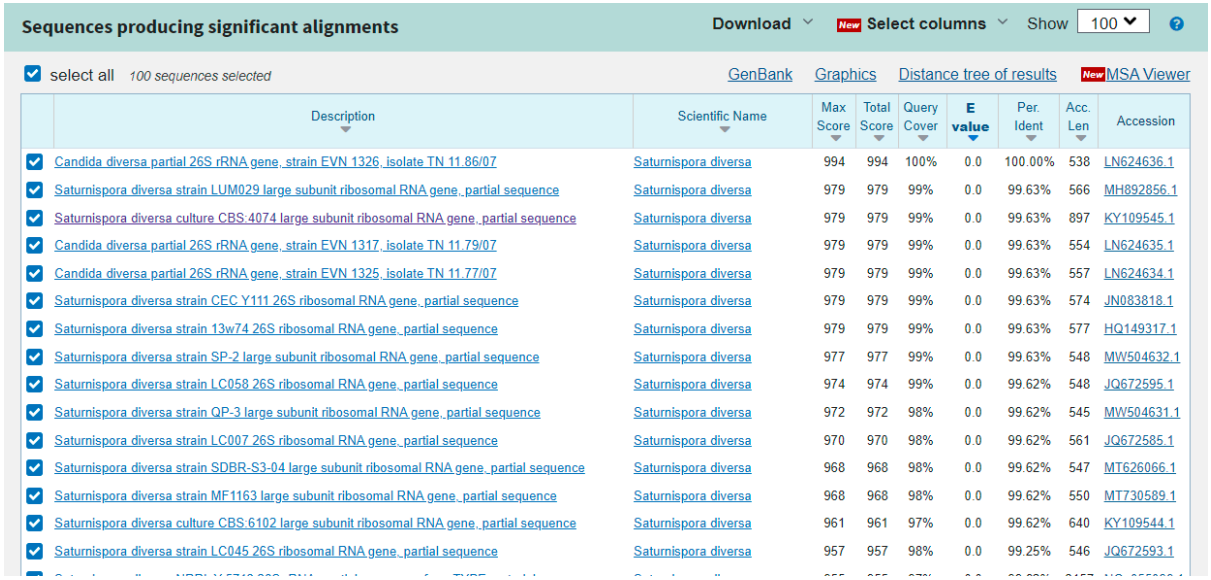
82. Tang, J., et al., *Combining an antagonistic yeast with harpin treatment to control postharvest decay of kiwifruit.* Biological Control, 2015. **89**: p. 61-67.
83. Peter Raspor*'^, D.M.-M., Martina Avbelj and Neza Cadez, *Biocontrol of Grey Mould Disease on Grape Caused by*

Botrytis cinérea with Autochthonous Wine Yeasts. 2010.
84. Tavares, M.J., et al., *Genome Sequence of the Wine Yeast Saccharomycodes ludwigii UTAD17.* Microbiology Resource Announcements, 2018. **7**(18): p. e01195-18.
85. Henschke, P.A. and V. Jiranek, *Yeasts-metabolism of nitrogen compounds in Wine Microbiology and Biotechnology.* Harwood Academic Publishers, 1993: p. 77-164.
86. Scientific, T. *Dehydrated Culture Media WL Nutrient Agar*. 10/03/2021].
87. Pallmann, C., et al., *Use of WL Medium to Profile Native Flora Fermentations.* American Journal of Enology and Viticulture, 2001. **52**.
88. Sipiczki, M., et al., *Genetic and Chromosomal Stability of Wine Yeasts*, in *Food Microbiology Protocols*, J.F.T. Spencer and A.L. de Ragout Spencer, Editors. 2001, Humana Press: Totowa, NJ. p. 273-281.
89. Hage, A.E. and J. Houseley, *Resolution of budding yeast chromosomes using pulsed-field gel electrophoresis.* Methods Mol Biol, 2013. **1054**: p. 195-207.
90. Maringele, L. and D. Lydall, *Pulsed-Field Gel Electrophoresis of Budding Yeast Chromosomes*, in *Yeast Protocol*, W. Xiao, Editor. 2006, Humana Press: Totowa, NJ. p. 65-73.
91. BIO-RAD. *S. cerevisiae PFGE marker*. 25/09/21].
92. BIO-RAD. *S. pombe PFGE marker*. 25/09/2021].
93. Doi, M., et al., *Estimation of chromosome number and size by pulsed-field gel electrophoresis (PFGE) in medically important Candida species.* Microbiology, 1992. **138**(10): p. 2243-2251.
94. Masneuf-Pomarede, I., et al., *The Genetics of Non-conventional Wine Yeasts: Current Knowledge and Future Challenges.* Frontiers in Microbiology, 2016. **6**(1563).
95. Sipiczki, M., *Species identification and comparative molecular and physiological analysis of Candida zemplinina and Candida stellata.* Journal of Basic Microbiology, 2004. **44**(6): p. 471-479.
96. Esteve-Zarzoso, B., et al., *Molecular characterization of Hanseniaspora species.* Antonie van Leeuwenhoek, 2001. **80**: p. 85-92.
97. Bankevich, A., et al., *SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing.* Journal of computational biology : a journal of computational molecular cell biology, 2012. **19**(5): p. 455-477.
98. Mario Stanke, O.K. *Augustus plugin for Geneious software*. 2021.
99. Consortium, T.G., et al., *Comparative genomics of protoploid Saccharomycetaceae.* Genome Research, 2009. **19**(10): p. 1696-1709.
100. (KEGG), K.E.o.G.a.G., *Blastkoala*. 2019.
101. Laboratories, K., *Kyoto Encyclopedia of Genes and Genomes (KEGG)*. 1995.
102. Zhao, X., S. Procopio, and T. Becker, *Flavor impacts of glycerol in the processing of yeast fermented beverages: a review.* Journal of Food Science and Technology, 2015. **52**(12): p. 7588-7598.
103. Linder, T., *Nitrogen Assimilation Pathways in Budding Yeasts*, in *Non-conventional Yeasts: from Basic Research to Application*, A. Sibirny, Editor. 2019, Springer International Publishing: Cham. p. 197-236.
104. Hove-Jensen, B., et al., *Phosphoribosyl Diphosphate (PRPP): Biosynthesis, Enzymology, Utilization, and Metabolic Significance.* Microbiology and molecular biology reviews : MMBR, 2016. **81**(1): p. e00040-16.
105. Steenwyk, J.L., et al., *Extensive loss of cell-cycle and DNA repair genes in an ancient lineage of bipolar budding yeasts.* PLOS Biology, 2019. **17**(5): p. e3000255.
106. Chattopadhyay, M.K., C.W. Tabor, and H. Tabor, *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. Proceedings of the National Academy of Sciences, 2003. **100**(24): p. 13869-13874.
107. Perli, T., et al., *Vitamin requirements and biosynthesis in Saccharomyces cerevisiae.* Yeast, 2020. **37**(4): p. 283-304.
108. Perli, T., *Engineering of vitamin and cofactor synthesis in yeasts*. 2021.
109. Labuschagne, P.W.J. and B. Divol, *Thiamine: a key nutrient for yeasts during wine alcoholic fermentation.* Applied Microbiology and Biotechnology, 2021. **105**(3): p. 953-973.
110. Ciani, M. and G. Picciotti, *The growth kinetics and fermentation behaviour of some non-Saccharomyces yeasts associated with wine-making.* Biotechnology Letters, 1995. **17**(11): p. 1247-1250.
111. Liu, S., et al., *Production of riboflavin and related cofactors by biotechnological processes.* Microbial Cell Factories, 2020. **19**(1): p. 31.
112. Schwechheimer, S.K., et al., *Biotechnology of riboflavin.* Appl Microbiol Biotechnol, 2016. **100**(5): p. 2107-19.
113. Averianova, L.A., et al., *Production of Vitamin B2 (Riboflavin) by Microorganisms: An Overview.* Frontiers in Bioengineering and Biotechnology, 2020. **8**(1172).
114. Fang, J. and D.S. Beattie, *External alternative NADH dehydrogenase of Saccharomyces cerevisiae: a potential source of superoxide.* Free Radical Biology and Medicine, 2003. **34**(4): p. 478-488.
115. 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, 2008. **283**(15): p. 9724-9736.
116. Bourdineaud, J.-P., et al., *Pmt1 mannosyl transferase is involved in cell wall incorporation of several proteins in Saccharomyces cerevisiae.* Molecular Microbiology, 1998. **27**(1): p. 85-98.

Appendix A

Figure 1- Blastn result of isolate 11.86.

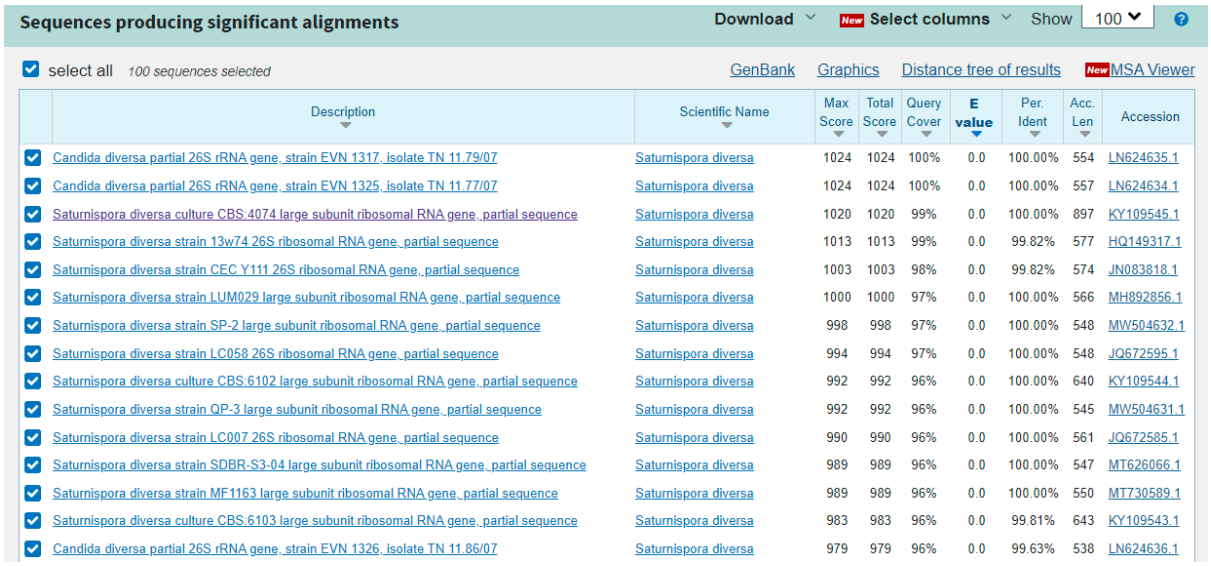


Sequences producing significant alignments Download Select columns Show 100

select all *100 sequences selected* GenBank Graphics Distance tree of results MSA Viewer

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
Candida diversa partial 26S rRNA gene, strain EVN 1326, isolate TN 11.86/07	Saturnispora diversa	994	994	100%	0.0	100.00%	538	LN624636.1
Saturnispora diversa strain LUM029 large subunit ribosomal RNA gene, partial sequence	Saturnispora diversa	979	979	99%	0.0	99.63%	566	MH892856.1
Saturnispora diversa culture CBS:4074 large subunit ribosomal RNA gene, partial sequence	Saturnispora diversa	979	979	99%	0.0	99.63%	897	KY109545.1
Candida diversa partial 26S rRNA gene, strain EVN 1317, isolate TN 11.79/07	Saturnispora diversa	979	979	99%	0.0	99.63%	554	LN624635.1
Candida diversa partial 26S rRNA gene, strain EVN 1325, isolate TN 11.77/07	Saturnispora diversa	979	979	99%	0.0	99.63%	557	LN624634.1
Saturnispora diversa strain CEC Y111 26S ribosomal RNA gene, partial sequence	Saturnispora diversa	979	979	99%	0.0	99.63%	574	JN083818.1
Saturnispora diversa strain 13w74 26S ribosomal RNA gene, partial sequence	Saturnispora diversa	979	979	99%	0.0	99.63%	577	HQ149317.1
Saturnispora diversa strain SP-2 large subunit ribosomal RNA gene, partial sequence	Saturnispora diversa	977	977	99%	0.0	99.63%	548	MW504632.1
Saturnispora diversa strain LC058 26S ribosomal RNA gene, partial sequence	Saturnispora diversa	974	974	99%	0.0	99.62%	548	JQ672595.1
Saturnispora diversa strain QP-3 large subunit ribosomal RNA gene, partial sequence	Saturnispora diversa	972	972	98%	0.0	99.62%	545	MW504631.1
Saturnispora diversa strain LC007 26S ribosomal RNA gene, partial sequence	Saturnispora diversa	970	970	98%	0.0	99.62%	561	JQ672585.1
Saturnispora diversa strain SDBR-S3-04 large subunit ribosomal RNA gene, partial sequence	Saturnispora diversa	968	968	98%	0.0	99.62%	547	MT626066.1
Saturnispora diversa strain MF1163 large subunit ribosomal RNA gene, partial sequence	Saturnispora diversa	968	968	98%	0.0	99.62%	550	MT730589.1
Saturnispora diversa culture CBS:6102 large subunit ribosomal RNA gene, partial sequence	Saturnispora diversa	961	961	97%	0.0	99.62%	640	KY109544.1
Saturnispora diversa strain LC045 26S ribosomal RNA gene, partial sequence	Saturnispora diversa	957	957	98%	0.0	99.25%	546	JQ672593.1

Figure 2- Blastn result of isolate 11.79.

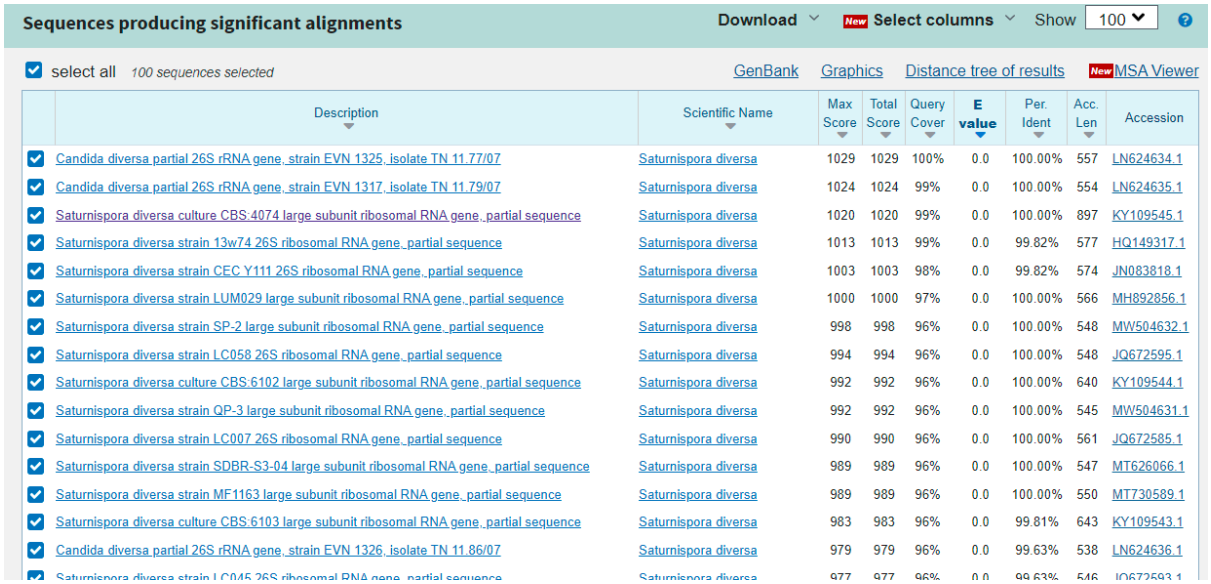


Sequences producing significant alignments Download Select columns Show 100

select all *100 sequences selected* GenBank Graphics Distance tree of results MSA Viewer

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
Candida diversa partial 26S rRNA gene, strain EVN 1317, isolate TN 11.79/07	Saturnispora diversa	1024	1024	100%	0.0	100.00%	554	LN624635.1
Candida diversa partial 26S rRNA gene, strain EVN 1325, isolate TN 11.77/07	Saturnispora diversa	1024	1024	100%	0.0	100.00%	557	LN624634.1
Saturnispora diversa culture CBS:4074 large subunit ribosomal RNA gene, partial sequence	Saturnispora diversa	1020	1020	99%	0.0	100.00%	897	KY109545.1
Saturnispora diversa strain 13w74 26S ribosomal RNA gene, partial sequence	Saturnispora diversa	1013	1013	99%	0.0	99.82%	577	HQ149317.1
Saturnispora diversa strain CEC Y111 26S ribosomal RNA gene, partial sequence	Saturnispora diversa	1003	1003	98%	0.0	99.82%	574	JN083818.1
Saturnispora diversa strain LUM029 large subunit ribosomal RNA gene, partial sequence	Saturnispora diversa	1000	1000	97%	0.0	100.00%	566	MH892856.1
Saturnispora diversa strain SP-2 large subunit ribosomal RNA gene, partial sequence	Saturnispora diversa	998	998	97%	0.0	100.00%	548	MW504632.1
Saturnispora diversa strain LC058 26S ribosomal RNA gene, partial sequence	Saturnispora diversa	994	994	97%	0.0	100.00%	548	JQ672595.1
Saturnispora diversa culture CBS:6102 large subunit ribosomal RNA gene, partial sequence	Saturnispora diversa	992	992	96%	0.0	100.00%	640	KY109544.1
Saturnispora diversa strain QP-3 large subunit ribosomal RNA gene, partial sequence	Saturnispora diversa	992	992	96%	0.0	100.00%	545	MW504631.1
Saturnispora diversa strain LC007 26S ribosomal RNA gene, partial sequence	Saturnispora diversa	990	990	96%	0.0	100.00%	561	JQ672585.1
Saturnispora diversa strain SDBR-S3-04 large subunit ribosomal RNA gene, partial sequence	Saturnispora diversa	989	989	96%	0.0	100.00%	547	MT626066.1
Saturnispora diversa strain MF1163 large subunit ribosomal RNA gene, partial sequence	Saturnispora diversa	989	989	96%	0.0	100.00%	550	MT730589.1
Saturnispora diversa culture CBS:6103 large subunit ribosomal RNA gene, partial sequence	Saturnispora diversa	983	983	96%	0.0	99.81%	643	KY109543.1
Candida diversa partial 26S rRNA gene, strain EVN 1326, isolate TN 11.86/07	Saturnispora diversa	979	979	96%	0.0	99.63%	538	LN624636.1

Figure 3- Blastn result of isolate 11.77.



Sequences producing significant alignments Download Select columns Show 100

select all *100 sequences selected* GenBank Graphics Distance tree of results MSA Viewer

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
Candida diversa partial 26S rRNA gene, strain EVN 1325, isolate TN 11.77/07	Saturnispora diversa	1029	1029	100%	0.0	100.00%	557	LN624634.1
Candida diversa partial 26S rRNA gene, strain EVN 1317, isolate TN 11.79/07	Saturnispora diversa	1024	1024	99%	0.0	100.00%	554	LN624635.1
Saturnispora diversa culture CBS:4074 large subunit ribosomal RNA gene, partial sequence	Saturnispora diversa	1020	1020	99%	0.0	100.00%	897	KY109545.1
Saturnispora diversa strain 13w74 26S ribosomal RNA gene, partial sequence	Saturnispora diversa	1013	1013	99%	0.0	99.82%	577	HQ149317.1
Saturnispora diversa strain CEC Y111 26S ribosomal RNA gene, partial sequence	Saturnispora diversa	1003	1003	98%	0.0	99.82%	574	JN083818.1
Saturnispora diversa strain LUM029 large subunit ribosomal RNA gene, partial sequence	Saturnispora diversa	1000	1000	97%	0.0	100.00%	566	MH892856.1
Saturnispora diversa strain SP-2 large subunit ribosomal RNA gene, partial sequence	Saturnispora diversa	998	998	96%	0.0	100.00%	548	MW504632.1
Saturnispora diversa strain LC058 26S ribosomal RNA gene, partial sequence	Saturnispora diversa	994	994	96%	0.0	100.00%	548	JQ672595.1
Saturnispora diversa culture CBS:6102 large subunit ribosomal RNA gene, partial sequence	Saturnispora diversa	992	992	96%	0.0	100.00%	640	KY109544.1
Saturnispora diversa strain QP-3 large subunit ribosomal RNA gene, partial sequence	Saturnispora diversa	992	992	96%	0.0	100.00%	545	MW504631.1
Saturnispora diversa strain LC007 26S ribosomal RNA gene, partial sequence	Saturnispora diversa	990	990	96%	0.0	100.00%	561	JQ672585.1
Saturnispora diversa strain SDBR-S3-04 large subunit ribosomal RNA gene, partial sequence	Saturnispora diversa	989	989	96%	0.0	100.00%	547	MT626066.1
Saturnispora diversa strain MF1163 large subunit ribosomal RNA gene, partial sequence	Saturnispora diversa	989	989	96%	0.0	100.00%	550	MT730589.1
Saturnispora diversa culture CBS:6103 large subunit ribosomal RNA gene, partial sequence	Saturnispora diversa	983	983	96%	0.0	99.81%	643	KY109543.1
Candida diversa partial 26S rRNA gene, strain EVN 1326, isolate TN 11.86/07	Saturnispora diversa	979	979	96%	0.0	99.63%	538	LN624636.1
Saturnispora diversa strain LC045 26S ribosomal RNA gene, partial sequence	Saturnispora diversa	977	977	96%	0.0	99.63%	546	JQ672593.1

Appendix B

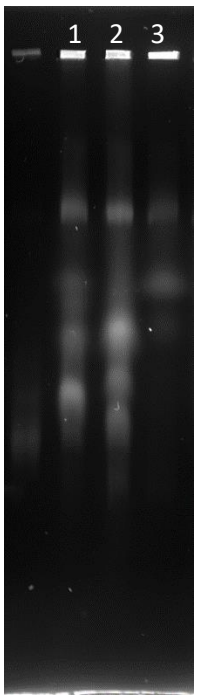


Figure 4- Image result of second PFGE performed, in system conditions of voltage gradient of 2 V/cm for 72 h, switch interval of 1800s and angle 106º, which led to disproportionate diffusion and migration of chromosomal bands. Column 1 refers to *S. cerevisiae* chromosomal bands, 2 to *S. diversa* MJT240 chromosomal bands and 3 to *S. pombe* chromosomal bands.

Appendix C

After pairwise comparative analysis with Proteome of *T. delbrueckii, L. fermentati, H. guilliermondi* and *S. ludwigii,* a set of proteins specific to *S. diversa* MJT240 was determined. From this set, a total of 320 were assigned to an enzyme/ function in Kegg Reconstruction Mapper tool, presented in **Table 1.**

Table 1- Set of 320 proteins specific to *S. diversa* MJT240*,* with the corresponding gene and with characterized function.

Gene	Function
Sd23_g41	splicing factor U2AF 65 kDa subunit
Sd2_g79	large subunit ribosomal protein L54
Sd4_g246	sporulation-specific protein 4
Sd14_g20	AP-1-like transcription factor
Sd25_g28	-pre-mRNA-processing factor SLU7
Sd24_g34	trafficking protein particle complex II-specific subunit 65
Sd19_g52	palmitoyl-Protein thioesterase
Sd3_g263	pre-mRNA 3'-end-processing factor FIP1
Sd4_g228	- vacuolar protein sorting-associated protein 54
Sd12_g22	DNA damage checkpoint Protein
Sd7_g91	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex subunit 12
Sd19_g59	methenyltetrahydrofolate synthase domain-containing protein
Sd47_g23	transcription initiation factor TFIIA large subunit
Sd7_g14	alcohol dehydrogenase
Sd12_g57	pre-mRNA-splicing factor 38A
Sd18_g13	4-amino-4-deoxychorismate lyase
Sd31_g37	elongation factor 1-gamma
Sd14_g82	plasminogen activator inhibitor 1 RNA-binding protein
Sd2_g65	gem associated protein 2
Sd35_g13	mediator of RNA polymerase II transcription subunit 13, fungi type
Sd3_g174	NADH dehydrogenase (ubiquinone) flavoprotein 1
Sd5_g147	charged multivesicular body protein 6
Sd20_g58	mitochondrial import receptor subunit TOM5
Sd2_g10	RAB6A-GEF complex partner protein 2
Sd2_g286	polyadenylate-binding protein 2
Sd46_g3	biogenesis of lysosome-related organelles complex 1 subunit KXD1
Sd7_g186	origin recognition complex subunit 3
Sd14_g72	nucleoporin NDC1
Sd40_g17	NADH dehydrogenase (ubiquinone) 1 beta subcomplex subunit 7
Sd1_g425_B	dolichol phosphate-mannose biosynthesis regulatory protein
Sd7_g201	protein SFI1

Sd4_g109	DASH complex subunit DAD2
Sd3_g45	ribonuclease P/MRP protein subunit POP5
Sd1_g366	GINS complex subunit 3: DNA replication
Sd1_g241	protein transport protein DSL1/ZW10
Sd33_g5	protein PET117: mitochondrial cytochrome c oxidase assembly
Sd92_g2	beta-lactamase class A TEM: Antibiotic resistance
Sd2_g309	tubulin-specific chaperone E:
Sd7_g196	exosome complex component RRP46
Sd29_g18	Myb-like DNA-binding protein RAP1
Sd9_g59	kinetochore protein Spc25, fungi type
Sd1_g221	transcriptional repressor OPI1
Sd3_g42	large subunit ribosomal protein L49
Sd4_g15	TBC1 domain family member 14: GTPase activation
Sd6_g192	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex subunit 13
Sd3_g193	1-aminocyclopropane-1-carboxylate deaminase: pyridoxal phosphate binding
Sd21_g32	U3 small nucleolar RNA-associated protein 3
Sd5_g14	oligosaccharyltransferase complex subunit gamma: protein O-linked mannosylation
Sd42_g1	NADH-ubiquinone oxidoreductase chain 1
Sd1_g57	cohesin complex subunit SCC1
Sd2_g96	small subunit ribosomal protein YMR-31
Sd3_g177 -	EKC/KEOPS complex subunit PCC1/LAGE3
Sd4_g105	inhibitor of growth protein 4
Sd17_g39	condensin complex subunit 2
Sd2_g138	peroxin-2: protein import into peroxisome matrix
Sd4_g128	protein CWC15: mRNA processing, mRNA splicing
Sd4_g239	sorting nexin-3/12
Sd5_g238	dolichol-phosphate mannosyltransferase subunit 3
Sd6_g101	ribosome biogenesis protein ALB1: Ribosome biogenesis, Transport
Sd20_g8	large subunit ribosomal protein L17
Sd6_g108	autophagy-related protein 33
Sd7_g68	guanine nucleotide exchange factor
Sd50_g2	kinetochore protein NNF1
Sd1_g126	NADH dehydrogenase (ubiquinone) 1 beta subcomplex subunit 8
Sd29_g7	coiled-coil-helix-coiled-coil-helix domain-containing protein 2
Sd23_g26	G protein-coupled receptor GPR1
Sd13_g38	golgi SNAP receptor complex member 2
Sd1_g197	HSP20 family protein
Sd7_g156	phosphatidylinositol-3,4,5-trisphosphate 3-phosphatase PTEN
Sd4_g165	NADH dehydrogenase (ubiquinone) Fe-S protein 3
Sd33_g19_B	vacuolar ATPase assembly integral membrane protein VMA21
Sd1_g346	importin-7: Host-virus interaction, Protein transport, Transport
Sd9_g36	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex assembly factor 3
Sd9_g122	M-phase phosphoprotein 6, fungi type
Sd18_g30	tubulin-specific chaperone B

Sd5_g158	NADH dehydrogenase (ubiquinone) 1 beta subcomplex subunit 9
Sd4_g190	BET1blocked early in transport 1
Sd51_g2	prefoldin subunit 4
Sd6_g7	anaphase-promoting complex subunit 2
Sd3_g90	SWI5-dependent HO expression protein 3
Sd20_g26	26 proteasome complex subunit DSS1
Sd26_g35	mitochondrial import inner membrane translocase subunit TIM13
Sd34_g38	protein disulfide-isomerase A6: protein folding
Sd15_g83	F-type H+-transporting ATPase subunit k
Sd27_g27	cleavage stimulation factor subunit 2: mRNA processing
Sd25_g19_B	signal peptidase complex subunit 2
Sd12_g123	protein transport protein SEC20: ER-Golgi transport protein,
Sd37_g18	DNA damage checkpoint protein LCD1
Sd30_g32	ribonuclease P protein subunit POP4
Sd4_g77	ribosome biogenesis protein UTP30
Sd2_g267	mediator of RNA polymerase II transcription subunit 18, fungi type
Sd9_g151	NADH dehydrogenase (ubiquinone) Fe-S protein 6
Sd1_g123	arrestin-related trafficking adapter 2/8
Sd37_g27	mediator of replication checkpoint protein 1
Sd23_g13	protein phosphatase inhibitor 2
Sd4_g80	trafficking protein particle complex subunit 10
Sd3_g259	GPI-anchor transamidase subunit S
Sd5_g43	CTD kinase subunit gamma
Sd10_g58	mediator of RNA polymerase II transcription subunit 21
Sd19_g43	translation initiation factor 3 subunit M
Sd8_g158	unconventional prefoldin RPB5 interactor 1
Sd10_g92	protein-serine/threonine kinase
Sd3_g123	DNA excision repair protein ERCC-1
Sd5_g253	Ase1/PRC1/MAP65 family protein
Sd11_g140	SNW domain-containing protein 1
Sd8_g115	E3 ubiquitin-protein ligase complex SLX5-SLX8 subunit SLX5
Sd7_g202	translation initiation factor 3 subunit D
Sd11_g151	pre-rRNA-processing protein IPI1
Sd9_g71	large subunit ribosomal protein L34
Sd21_g6	Derlin-2/3
Sd17_g33	UV radiation resistance-associated gene protein (UVRAG)
Sd1_g326	NADH dehydrogenase (ubiquinone) Fe-S protein 8
Sd30_g35	histone deacetylase complex subunit SAP30:
Sd27_g15	acetylornithine deacetylase
Sd7_g148	origin recognition complex subunit 6
Sd17_g19 -	zinc finger HIT domain-containing protein 1
Sd19_g47	prefoldin subunit 1
Sd6_g97_B	ribonuclease P protein subunit RPR2
Sd7_g107	mediator of RNA polymerase II transcription subunit 8, fungi type

Sd31_g1	glutathione S-transferase
Sd1_g14_B	4'-phosphopantetheinyl transferase
Sd16_g24	TBC1 domain family member 5
Sd1_g278	NADH dehydrogenase (ubiquinone) Fe-S protein 7
Sd5_g86	U6 snRNA-associated Sm-like protein LSm7
Sd42_g13	NADH-ubiquinone oxidoreductase chain 4
Sd8_g150	N-alpha-acetyltransferase 50
Sd9_g85	CTD kinase subunit beta
Sd9_g93	pre-mRNA-splicing factor ISY1
Sd2_g103	origin recognition complex subunit 5
Sd15_g49	proteasome chaperone 3
Sd37_g7	splicing factor 3A subunit 3
Sd4_g176_B	signal peptidase complex subunit 1
Sd9_g73	U3 small nucleolar RNA-associated protein 16
Sd11_g77	iron-sulfur cluster assembly
Sd25_g8	kinetochore protein Spc24, fungi type
Sd3_g268	TBC1 domain family member 20
Sd24_g13	magnesium transporter (NIPA, SLC57A2S)
Sd19_g27_B	solute carrier family 45, member 1/2/4
Sd1_g105	RAD50-interacting protein 1
Sd4_g214_A	mediator of RNA polymerase II transcription subunit 5
Sd13_g83	small nuclear ribonucleoprotein E
Sd4_g37	something about silencing protein 4 (SAS4)
Sd36_g12	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex subunit 9
Sd25_g26	protein SPT2
Sd3_g98	beta-1,4-N-acetylglucosaminyltransferase
Sd5_g204_B	small nuclear ribonucleoprotein B and B'
Sd1_g461_B	F-type H+-transporting ATPase subunit f
Sd10_g76	NADH dehydrogenase (ubiquinone) Fe-S protein 5
Sd31_g42	central kinetochore subunit Mal2/MCM21
Sd2_g312	splicing factor 3B subunit 5
Sd36_g17	ethanolamine kinase
Sd8_g111	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex assembly factor 6
Sd1_g299	transcription initiation factor TFIID subunit 13
Sd12_g121	small subunit ribosomal protein MRP21
Sd7_g165	F-type H+-transporting ATPase subunit k
Sd5_g116	cell growth-regulating nucleolar protein (LYER)
Sd1_g72	DNA-directed RNA polymerase I subunit RPA34
Sd16_g86	endoplasmic reticulum junction formation protein lunapark
Sd1_g384	THO complex subunit 4:
Sd9_g131	protein EAP1
Sd8_g155_B	autophagy-related protein 5
Sd8_g58	cystinosin
Sd10_g112	optic atrophy 3 protein

Sd10_g151_B	telomere length regulation protein
Sd10_g157	RNA polymerase-associated protein LEO1
Sd15_g98	pre-mRNA-splicing factor SYF2
Sd11_g138	enhancer of yellow 2 transcription factor:
Sd1_g33	replication factor A3
Sd41_g22	RNA polymerase I-specific transcription initiation factor RRN6
Sd27_g6	GPI ethanolamine phosphate transferase 2/3 subunit F:
Sd1_g119	sentrin-specific protease 8
Sd38_g24	syntaxin 18
Sd48_g4	SAP domain-containing ribonucleoprotein
Sd3_g152	DNA replication regulator SLD3
Sd1_g124	ribonuclease HI
Sd42_g9	NADH-ubiquinone oxidoreductase chain 3
Sd7_g198	U4/U6 small nuclear ribonucleoprotein PRP31
Sd1_g271	E3 ubiquitin-protein ligase BAH
Sd12_g137	[acyl-carrier-protein] S-malonyltransferase
Sd11_g5	ubiquinol-cytochrome c reductase subunit 9
Sd16_g29	peroxin-13
Sd6_g48	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex subunit 5
Sd6_g91	U6 snRNA-associated Sm-like protein LSm4
Sd15_g3	kinesin family member 5
Sd2_g107	protein IFH1
Sd4_g73	exosome complex component MTR3
Sd3_g63	exonuclease V
Sd12_g95	exocyst complex component 3
Sd4_g185	solute carrier family 24 (sodium/potassium/calcium exchanger), member 6
Sd13_g67	cell cycle checkpoint protein
Sd42_g10	NADH-ubiquinone oxidoreductase chain 2
Sd9_g182	SHO1 osmosensor
Sd1_g142	protein YIPF6
Sd32_g18	ribonuclease P/MRP protein subunit POP3
Sd30_g13	ESCRT-I complex subunit TSG101
Sd40_g5	nuclear mRNA export protein PCID2/THP1
Sd3_g256B	U2 small nuclear ribonucleoprotein B''
Sd14_g88	protein Cut8
Sd7_g59	protein transport protein SEC39
Sd18_g27	programmed cell death protein 5
Sd4_g99	chromatin assembly factor 1 subunit A
Sd15_g101 –	calmodulin
Sd5_g134	signal recognition particle subunit SRP14
Sd34_g4	smad nuclear-interacting protein 1
Sd24_g10	26S proteasome regulatory subunit N4
Sd12_g100_B	pre-mRNA-processing factor 39
Sd6_g26	mediator of RNA polymerase II transcription subunit 19, fungi type

Sd5_g72	exocyst complex component 7
Sd10_g89	RAB6A-GEF complex partner protein 1
Sd4_g261	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex assembly factor 2
Sd27_g2	nucleoporin ASM4
Sd23_g6	anaphase-promoting complex subunit 4
Sd26_g28	protein NUD1
Sd19_g62	protein unc-45
Sd19_g77	ATPase inhibitor, mitochondrial
Sd17_g37	transcription initiation factor TFIID subunit 11
Sd11_g141	catabolite repression protein CreC
Sd12_g41	prenyl protein peptidase
Sd4_g251	mitochondrial transcription factor 1
Sd33_g1	nuclear pore complex protein Nup133
Sd31_g11	coatomer subunit épsilon
Sd7_g150	mitochondrial import receptor subunit TOM22
Sd2_g310	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex subunit 2
Sd9_g72	solute carrier family 39 (zinc transporter), member 9
Sd6_g231	prefoldin subunit 2
Sd31_g8	- B-cell receptor-associated protein 31
Sd17_g64	RAT1-interacting protein
Sd13_g107	protein CMS1
Sd27_g10	inositol-pentakisphosphate 2-kinase
Sd1_g529	BolA-like protein 1
Sd19_g48	autophagy-related protein 27
Sd30_g27	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex assembly factor 7
Sd16_g96	dynein light chain LC8-type
Sd5_g37	peroxin-19
Sd8_g160	tubulin-specific chaperone A
Sd3_g58	lectin, mannose-binding 2
Sd4_g30	WD repeat-containing protein 26
Sd13_g36	translocation protein SEC72: Protein transport, Transport
Sd1_g461	protein SOV1, mitochondrial
Sd48_g3	oligosaccharyltransferase complex subunit delta (ribophorin II)
Sd4_g194	charged multivesicular body protein 7
Sd43_g7	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex assembly factor 1
Sd26_g27	cytochrome c oxidase assembly factor 14
Sd9_g30	regulator of Ty1 transposition protein 103
Sd7_g117	TATA element modulatory factor
Sd2_g352	cullin-associated NEDD8-dissociated protein 1
Sd6_g243	1-acylglycerol-3-phosphate O-acyltransferase
Sd35_g24	cytochrome c oxidase assembly factor 4
Sd3_g209	membrane-associated progesterone receptor componente
Sd20_g9	translation initiation factor 3 subunit E
Sd14_g98	U6 snRNA-associated Sm-like protein LSm6

Sd2_g106	translation initiation factor 3 subunit F
Sd1_g121	GINS complex subunit 2
Sd8_g189	G2-specific checkpoint protein
Sd29_g19	peptide chain release factor
Sd1_g413	enhancer of mRNA-decapping protein 3
Sd14_g36_B	mitochondrial distribution and morphology protein 34
Sd2_g223	kinetochore protein Spc7/SPC105
Sd2_g231	NADH dehydrogenase (ubiquinone) Fe-S protein 1
Sd8_g9	rRNA-processing protein CGR1
Sd25_g7	COX assembly mitochondrial protein 1
Sd7_g97	HAT1-interacting factor 1
Sd4_g59	tRNA(Ile)-lysidine synthase
Sd62_g3	tail-anchored protein insertion receptor
Sd2_g104	protein BUR2
Sd46_g16	U3 small nucleolar RNA-associated protein 8
Sd10_g88	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex subunit 6
Sd2_g124	peptidyl-tRNA hydrolase ICT1
Sd2_g153	CCR4-NOT transcription complex subunit 7/8
Sd2_g29	phosphatidylinositol N-acetylglucosaminyltransferase subunit P
Sd27_g36	elongin-C
Sd26_g6	transmembrane protein 33
Sd24_g15	macrophage erythroblast attacher
Sd2_g353	outer membrane protein insertion porin family
Sd2_g228	inhibitor of Bruton tyrosine kinase
Sd8_g70	DNA polymerase delta subunit 3
Sd7_g157	MICOS complex subunit MIC26
Sd16_g50	MICOS complex subunit MIC26
Sd16_g54_B	Ran-binding protein 9/10
Sd10_g119	pyruvate dehydrogenase kinase 2/3/4
Sd1_g125	ribonuclease HI
Sd1_g238	signal transducing adaptor molecule
Sd2_g348	cruciform cutting endonuclease 1
Sd5_g220	vacuolar protein sorting-associated protein VTA1
Sd1_g129	zinc finger HIT domain-containing protein 3
Sd6_g8	DCN1-like protein 1,2
Sd11_g18	Vam6/Vps39-like protein vacuolar protein sorting-associated protein 39
Sd13_g114	mediator of RNA polymerase II transcription subunit 4
Sd8_g97	DNA polymerase epsilon subunit 3
Sd2_g256	elongation factor 1-gamma
Sd5_g176	DnaJ homolog subfamily C member 3
Sd38_g20	N-terminal acetyltransferase B complex non-catalytic subunit
Sd6_g187	staphylococcal nuclease domain-containing protein 1
Sd13_g34	F-type H+-transporting ATPase subunit h
Sd3_g159	cytochrome c oxidase subunit 7

Sd3_g21	meiotic recombination protein SPO11
Sd1_g88	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex assembly factor 5
Sd34_g30	ribonuclease MRP protein subunit RMP1
Sd1_g205	large subunit ribosomal protein MRP49
Sd27_g29	cleavage stimulation factor subunit 2
Sd29_g28	anaphase-promoting complex subunit 5
Sd14_g25	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex subunit 8
Sd1_g209	peroxisomal membrane protein 4
Sd1_g54	protein transport protein SEC61 subunit beta
Sd14_g60	tubulin-specific chaperone D
Sd19_g53	ribosome biogenesis protein UTP30
Sd19_g82	ASTRA-associated protein 1
Sd5_g140	large subunit ribosomal protein L30
Sd11_g142	STAM-binding Protein
Sd23_g11	DNA replication regulator SLD2
Sd56_g4	NADH dehydrogenase (ubiquinone) Fe-S protein 4
Sd42_g2	F-type H+-transporting ATPase subunit c
Sd5_g209	ubiquinol-cytochrome c reductase subunit 10
Sd12_g20	RecQ-mediated genome instability protein 1
Sd1_g170	COP9 signalosome complex subunit 2
Sd2_g23	non-structural maintenance of chromosomes element 1
Sd29_g14	regulator of vacuolar morphogenesis
Sd15_g57	translation initiation factor 3 subunit H
Sd38_g9	P-type Ca2+ transporter type 2C
Sd31_g3	glutathione S-transferase

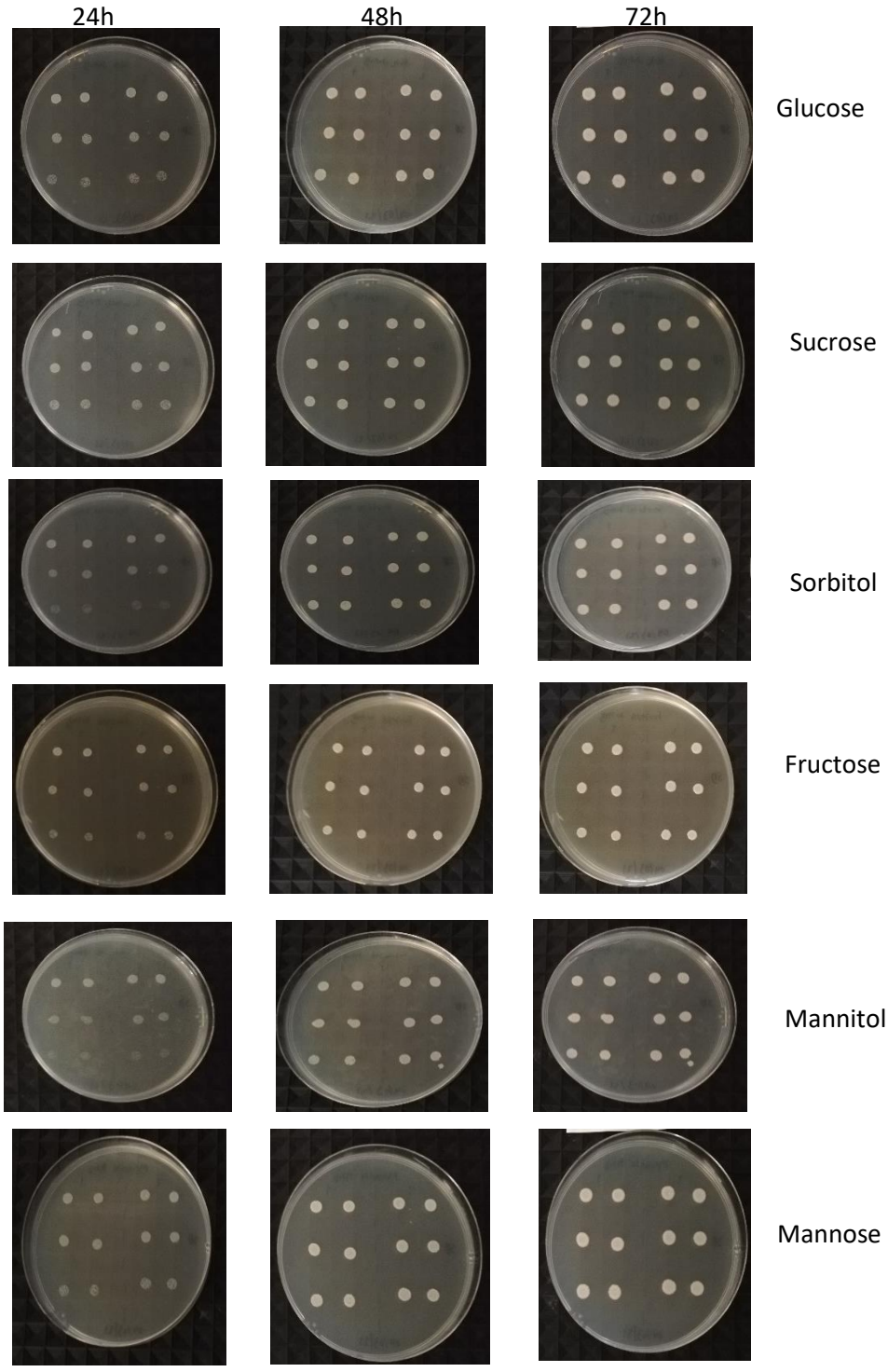
Appendix D

Table 2- Genes identified in *S. diversa* MJT240 genome, whose function is to produce volatile aromatic compounds.

Volatile compound	Gene	Function	S. cerevisiae S288c orthologue
Acetaldehyde	Sd6_g11	Pyruvate decarboxylase	*PDC1,5,6*
Acetate esters	-	probable alcohol acetyl-transferase	-
	-	probable alcohol acetyl-transferase	-
	Sd11_g32	isoamyl acetate esterase	-
	Sd11_g31	isoamyl acetate esterase	-
Ethyl esters	Sd7_g159	related to Medium-chain fatty acid ethyl ester synthase/esterase 2	*EHT1*
	Sd46_g6	related to Medium-chain fatty acid ethyl ester synthase/esterase 2	*EHT1*
Fatty Acids	Sd2_g135	related to acetyl-CoA hydrolase	*ACH1*
Fatty Acids/Ethyl esters	Sd1_g317	probable 2-enoyl thioester reductase	*ETR1*
	Sd12_g137	probable malonyl-CoA:ACP transferase	*MCT1*
Higher alcohols	Sd7_g14	probable Alcohol dehydrogenase	*ADH1,2,5,6*
	Sd8_g52	probable Alcohol dehydrogenase	*ADH1,2,5,6*
	Sd11_g31	related to Isoamyl acetate-hydrolyzing esterase	*IAH1*
	Sd11_g32	related to Isoamyl acetate-hydrolyzing esterase	*IAH1*
	Sd8_g52	probable S-(hydroxymethyl)glutathione dehydrogenase	*SFA1*
Terpenes	-	Beta-glucosidase	-
	-	Beta-glucosidase	-
	-	Beta-glucosidase	-
Thiols	Sd30_g42	related to Gamma-glutamyltransferase	*ECM38*
	-	related to putative cystathionine beta-lyase	*CYS3*
	-	probable cystathionine gamma-lyase	*CYS3*
	Sd19_g61	proton-dependent oligopeptide transporter	*OPT1*
	Sd30_g3	cysteine-S-conjugate beta-lyase	*IRC7*
	Sd1_g440	probable Cys-Gly metallodipeptidase DUG1	*DUG1*
	Sd8_g26	related to Lactoylglutathione lyase	*GLO1*
	Sd1_g69	Probable kynurenine--oxoglutarate transaminase BNA3	*BNA3*

Appendix E

To test capability of S. diversa MJT240 cells to grow in different growth media having different sugars as carbon sources spot assays were used. Herein are presented the image results of *S. diversa* MJT240 growth in all different carbon sources, after 72h.



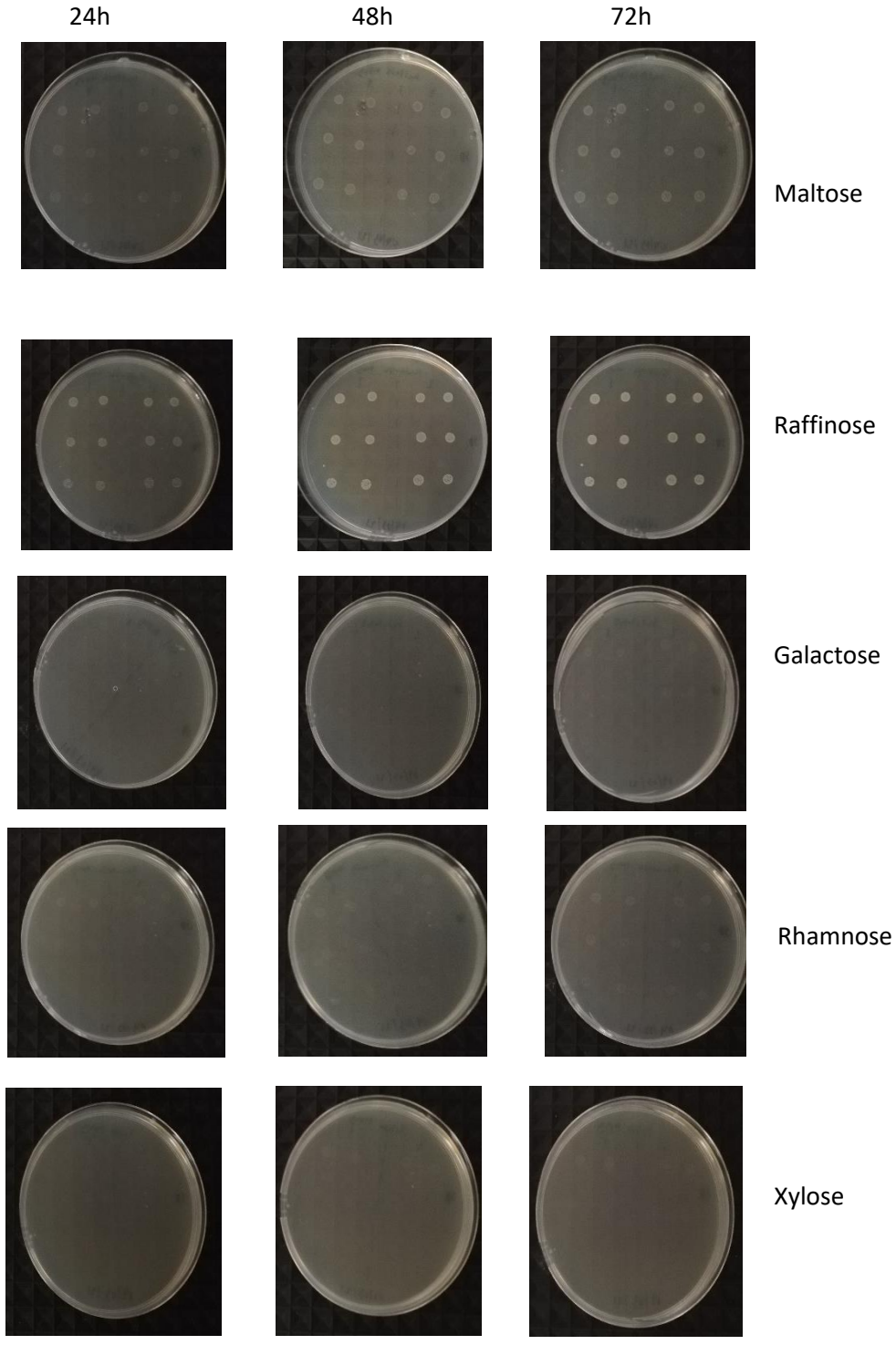


Figure 5- *S. diversa* MJT240 growth in different sugars as sole carbon source.