

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE BIOCÊNCIAS  
BACHARELADO EM BIOTECNOLOGIA - ÊNFASE EM BIOINFORMÁTICA

Fernando Junior Biedermann

**Construção do pangenoma de leveduras do gênero *Hanseniaspora sp.* e  
prospecção de genes associados com processo fermentativos industriais**

**Porto Alegre  
2024**

Fernando Junior Biedermann

**Construção do pangenoma de leveduras do gênero *Hanseniaspora* sp. e  
prospecção de genes associados com processo fermentativos industriais**

Trabalho de Conclusão de Curso  
apresentado como pré-requisito obrigatório  
à obtenção do título de bacharel em  
Biotecnologia - Ênfase em Bioinformática  
pela Universidade Federal do Rio Grande  
do Sul.

Área de Habilitação: Bioinformática  
Orientador: Prof. Dr. Diego Bonatto

Porto Alegre

2024

### CIP - Catalogação na Publicação

Biedermann, Fernando Junior

Construção do pangenoma de leveduras do gênero *Hanseniaspora* sp. e prospecção de genes associados com processo fermentativos industriais / Fernando Junior Biedermann. -- 2024.

46 f.

Orientador: Diego Bonatto.

Trabalho de conclusão de curso (Graduação) --  
Universidade Federal do Rio Grande do Sul, Instituto  
de Biociências, Curso de Biotecnologia:  
Bioinformática, Porto Alegre, BR-RS, 2024.

1. *Hanseniaspora*. 2. Fermentação. 3. Prospecção de genes. 4. Processos industriais. 5. Leveduras não convencionais. I. Bonatto, Diego, orient. II. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os dados fornecidos pelo(a) autor(a).

Fernando Junior Biedermann

**Construção do pangenoma de leveduras do gênero *Hanseniaspora* sp. e  
prospecção de genes associados com processo fermentativos industriais**

Trabalho de Conclusão de Curso apresentado como requisito parcial para obtenção do título de bacharel em Biotecnologia - Ênfase em Bioinformática pela Universidade Federal do Rio Grande do Sul.

Aprovado em: 21 / 01 / 2024

BANCA EXAMINADORA

---

Prof. Dr. Diego Bonatto

---

Profa. Dra. Fabiana Quoos Mayer

---

Dra. Lovaine Silva Duarte

## **AGRADECIMENTOS**

Acredito ser muito justo agradecer ao meu orientador Prof. Dr. Diego Bonatto pela enorme paciência comigo. E também ao meu amigo e colega Rafael de Matos que tem me acompanhado e me dado apoio moral e psicológico quando necessário.

## RESUMO

O vinho é uma bebida alcoólica produzida a partir da fermentação alcoólica da uva. Muitas espécies de leveduras estão envolvidas no processo de fermentação espontânea da uva, sendo *S. cerevisiae* a mais relevante para a produção do vinho. Outras espécies, consideradas como leveduras não convencionais, tinham um papel secundário na fabricação do vinho. Recentemente, essa visão tem mudado. Leveduras não convencionais vêm sendo estudadas a fim de verificar o seu potencial para conferir propriedades distintas ao vinho, como aroma e sabor. Dentre as leveduras mais promissoras, encontra-se *Hanseniaspora sp.*, um gênero de leveduras apiculadas abundante na microbiota natural da uva. Esse gênero, em particular, apresenta dois grupos distintos em sua árvore evolutiva: um grupo de lenta evolução e baixa diversificação, *Slow Evolving Lineages* (SEL), e um grupo de rápida evolução, *Fast Evolving Lineages* (FEL), caracterizado por extensa perda de genes ao longo da história evolutiva, porém grande diversificação de espécies. A erosão genômica experimentada pelo gênero *Hanseniaspora sp.* evidencia uma tendência para especialização ao ambiente fermentativo e de microbiota da uva. Conseqüentemente, o gênero *Hanseniaspora sp.* pode abrigar grande diversidade de genes com potencial biotecnológico. A fim de testar essa hipótese, genomas de *Hanseniaspora sp.* foram obtidos e anotados usando diferentes ferramentas computacionais. As sequências ortólogas foram identificadas entre os genomas de *Hanseniaspora sp.* e *S. cerevisiae*. Uma análise filogenômica foi empregada para classificar os genomas dentro dos grupos FEL e SEL. Os ortólogos foram anotados funcionalmente, atribuindo-lhes termos de ontologia gênica. De forma individual, para cada grupo, os transcritos de *Hanseniaspora sp.* sem ortólogos em *S. cerevisiae* foram submetidos à análise de enriquecimento de ontologia identificando as funções biológicas mais significativas para cada grupo. Obteve-se como resultado que ambos os grupos apresentam um perfil de enriquecimento muito semelhante entre si. Dentre os termos enriquecidos, alguns podem estar relacionados ao processo fermentativo de *Hanseniaspora sp.*, possuindo potencial para aplicação biotecnológica.

**Palavras-chave:** *Hanseniaspora*; fermentação; leveduras não convencionais; anotação de genomas; enriquecimento funcional; filogenia; ontologia gênica.

## ABSTRACT

Wine is an alcoholic beverage produced from the alcoholic fermentation of grapes. Many yeast species are involved in the spontaneous grape fermentation process, with *S. cerevisiae* being the most relevant for wine production. Other species, considered unconventional yeasts, were considered to have a secondary role in the manufacture of wine. Recently, this view has changed. Unconventional yeasts have been studied in order to verify their potential to impart distinct properties to wine, such as aroma and flavor. Among the most promising yeasts is *Hanseniaspora sp.*, a genus of apiculate yeasts abundant in the natural microbiota of grapes. This genus, in particular, presents two distinct groups in its evolutionary tree: a group of slow evolution and low diversification, Slow Evolving Lineages (SEL), and a group of rapid evolution, Fast Evolving Lineages (FEL), characterized by extensive loss of genes throughout evolutionary history, but great diversification of species. The genomic erosion experienced by the genus *Hanseniaspora sp.* highlights a tendency towards specialization to the fermentative and microbiota environment of the grape. Consequently, the genus *Hanseniaspora sp.* can harbor a great diversity of genes with biotechnological potential. In order to test this hypothesis, genomes of *Hanseniaspora sp.* were obtained and annotated using different computational tools. Orthologous sequences were identified among the genomes of *Hanseniaspora sp.* and *S. cerevisiae*. A phylogenomic analysis was employed to classify the genomes into the FEL and SEL groups. The orthologs were functionally annotated, assigning them gene ontology terms. Individually, for each group, the *Hanseniaspora sp.* without orthologs in *S. cerevisiae* were subjected to ontology enrichment analysis identifying the most significant biological functions for each group. The result was that both groups present a very similar enrichment profile to each other. Among the enriched terms, some may be related to the fermentative process of *Hanseniaspora sp.*, having potential for biotechnological application.

**Keywords:** *Hanseniaspora*; fermentation; unconventional yeasts; genome annotation; functional enrichment; phylogeny; gene ontology.

## LISTA DE FIGURAS

**Figura 1:** Aplicações biotecnológicas que fazem uso de leveduras

**Figura 2:** Morfologia de leveduras apiculadas do gênero *Hanseniaspora*

**Figura 3:** Árvore filogenética do gênero *Hanseniaspora*

**Figura 4:** Gráfico log(N50) x Completude.

**Figura 5:** Distribuição dos ortólogos identificados dentro dos grupos FEL e SEL de *Hanseniaspora*.

**Figura 6:** Porcentagem médias de ortólogos compartilhados com *S. cerevisiae* e ortólogos exclusivos de *Hanseniaspora* com e sem função biológica predita.

**Figura 7:** Gráfico de barras de enriquecimento do grupo FEL.

**Figura 8:** Gráfico de barras de enriquecimento do grupo SEL.

**Figura 9:** *Treemap* representando os principais termos de ontologia associados ao gênero *Hanseniaspora*.



## LISTA DE ABREVIATURAS

a.C. : Antes de Cristo

BH : *Benjamini-Hochberg Procedure*

DNA : *ácido desoxirribonucleico*

FASTA : Formato de texto para representação de sequências biológicas

FEL : *Fast Evolving Lineages*

GO : *Gene Ontology*

MUSCLE : *Multiple Sequence Comparison by Log-Expectation*

NGS : *Next Generation Sequencing*

NCBI : *National Center for Biotechnology Information*

Phylip : *PhyLogeny Inference Package*

PhyML : Phylogenetic tree inference using Maximum Likelihood

SEL : *Slow Evolving Lineages*

TSV : *Tabulation Separated Values*

## LISTA DE TABELAS

**Tabela S1:** Genomas utilizados neste trabalho.

**Tabela S2:** Parâmetros utilizados pelos anotadores.

## SUMÁRIO

<b>1. INTRODUÇÃO.....</b>	<b>11</b>
1.1. Leveduras.....	11
1.2. Vinicultura.....	12
1.3. O gênero <i>Hanseniaspora</i> .....	14
1.4. Pangenomas e prospecção de genes.....	16
<b>2. JUSTIFICATIVA.....</b>	<b>18</b>
<b>3. OBJETIVOS.....</b>	<b>18</b>
3.1. Objetivo Geral.....	18
3.2. Objetivos Específicos.....	18
<b>4. MATERIAIS E MÉTODOS.....</b>	<b>19</b>
4.1. Seleção de genomas.....	19
4.2. Anotação dos genomas.....	19
4.3. Identificação de ortólogos.....	19
4.4. Filogenômica do grupo <i>Hanseniaspora</i> .....	20
4.5. Anotação funcional dos ortólogos.....	21
4.6. Análise de enriquecimento de ontologia.....	21
<b>5. RESULTADOS E DISCUSSÃO.....</b>	<b>23</b>
5.1. Classificação dos genomas nos grupos FEL e SEL.....	23
5.2. Anotação funcional dos ortólogos.....	27
5.3. Enriquecimento de ontologia em <i>Hanseniaspora</i> .....	29
<b>6. CONCLUSÕES E PERSPECTIVAS.....</b>	<b>35</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>36</b>
<b>ANEXOS.....</b>	<b>44</b>

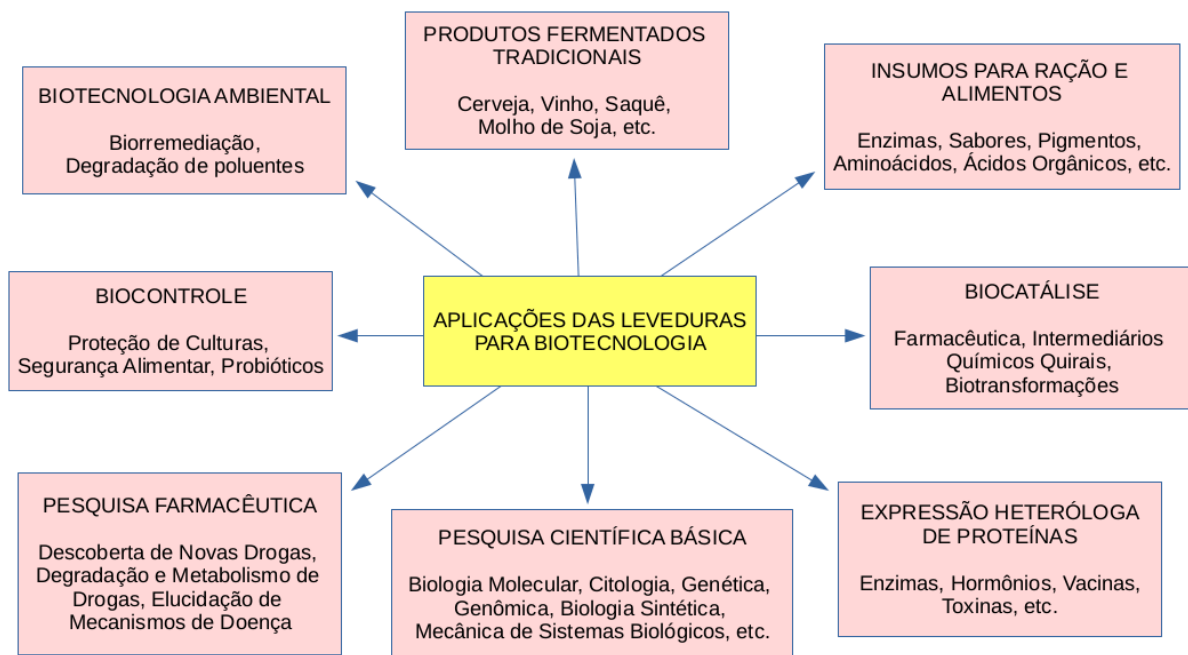
## 1. INTRODUÇÃO

### 1.1 Leveduras

As leveduras são organismos unicelulares pertencentes ao reino Fungi, com alguns membros classificados dentro do filo Ascomycota e outros pertencendo a Basidiomycota. Suas células variam entre formatos ovais e alongados, capazes de reproduzir-se assexuadamente através de brotamento ou fissão, porém muitas espécies realizam reprodução sexual (BOEKHOUT, 2022; KNOP, 2011). De acordo com “The Yeasts: a Taxonomic Study” (KURTZMAN, 2011), em sua edição mais recente, havia cerca de 1301 espécies conhecidas, distribuídas entre 186 gêneros.

A maioria das leveduras é anaeróbia facultativa, utilizando várias rotas metabólicas para obtenção de energia na ausência de oxigênio. O processo mais comum é a fermentação alcoólica que converte carboidratos em etanol e dióxido de carbono. Esses organismos possuem distribuição ubíqua, sendo elementos importantes no processo de decomposição da matéria orgânica. Em seus nichos ecológicos podem sobreviver utilizando a fermentação, enquanto houver disponibilidade suficiente de nutrientes em seu substrato (MAICAS, 2020; SPENCER, 1997).

O uso de leveduras para alimentação e produção de bebidas alcoólicas desempenhou um papel fundamental na história da humanidade. A evidência arqueológica indica que a cerca de 4.000 a.C, os egípcios já dominavam o uso do fermento para a produção de vinhos e pães (ROSS, 2002). O constituinte do fermento são leveduras, especialmente *Saccharomyces cerevisiae*, que é a levedura modelo mais estudada pela ciência e amplamente usada na indústria de panificação e produção de bebidas alcoólicas (PARAPOULI, 2020). As leveduras possuem inúmeras outras aplicações biotecnológicas que vão além da alimentação humana, tais como a produção de biocombustíveis e outros compostos químicos, a expressão de proteínas recombinantes, biologia sintética, etc. (NANDY, 2018; TULLIO, 2022). Um resumo dessas aplicações pode ser observado na **Figura 1**.



**Figura 1:** Aplicações biotecnológicas que fazem uso de leveduras (Modificado de KURTZMAN, 2011).

## 1.2 Vinicultura

A vinicultura é o processo de produção do vinho. O vinho é uma das bebidas alcoólicas mais antigas do mundo. As evidências arqueológicas apontam para o surgimento do vinho em aproximadamente 6.000 a.C., período neolítico, na região do Oriente Médio. Há cerca de 4.000 a.C., o cultivo da uva por sumérios, egípcios e babilônios para produção de vinho ou consumo *in natura* era generalizado. O vinho foi incorporado a diversos aspectos da cultura, mitologia e religiosidade desses povos, sendo consumido durante celebrações, cerimônias religiosas e explorado para fins medicinais (MCGOVERN, 2003).

Os gregos desenvolveram uma profunda compreensão sobre o cultivo da uva e da vinicultura, a partir das técnicas que foram aprendidas dos egípcios que mantinham extensa cultura de videiras ao longo do Delta do Rio Nilo. Os romanos herdaram esse conhecimento dos gregos e, durante a expansão do Império Romano, introduziram o cultivo da uva em todo o continente Europeu (LI, 2018).

Apesar de ser quase tão antigo quanto a própria história da humanidade, só desenvolvemos uma verdadeira compreensão sobre o processo que dá origem ao vinho no século XIX graças aos trabalhos do químico e microbiologista francês Louis Pasteur e alguns de seus contemporâneos, que identificaram as leveduras, especialmente *Saccharomyces cerevisiae* e outras espécies aparentadas a esta, como os agentes responsáveis pela fermentação do vinho, do pão e de outras bebidas alcoólicas (BARNETT, 2000). Pasteur foi um dos primeiros a caracterizar o processo de fermentação por microrganismos, o primeiro a isolar e estabelecer monoculturas de microrganismos para inoculação de diferentes produtos fermentados e o criador da pasteurização (MORTIMER, 2000; EL-MANSI, 2019).

Os primeiros trabalhos no estudo da microbiologia foram essenciais para o desenvolvimento da indústria das bebidas alcoólicas, como vinho e cerveja. Até então, o processo de produção do vinho era muito artesanal e dependente da fermentação espontânea, que devido a presença de culturas heterogêneas e condições muito variáveis, é muito propenso à contaminação por outros microrganismos indesejados e pouco escalável, limitando sua aplicação industrial. Ao final do século XIX, entretanto, várias inovações foram desenvolvidas que permitiram o progresso industrial, tais como o estabelecimento de monoculturas puras de leveduras, o surgimento dos primeiros biorreatores e métodos de esterilização, como a pasteurização, que permitem um rigoroso controle de qualidade da produção (BERGER, 1972; MORTIMER, 2000; BUCHHOLZ, 2013; MAICAS, 2023).

Atualmente, o mercado do vinho é global, sendo os Estados Unidos e países europeus, especialmente a França, Itália e Espanha, os maiores produtores e consumidores (OHANA-LEVI, 2023). O vinho destaca-se de outras bebidas alcoólicas devido ao apelo de suas características sensoriais. Acredita-se que variações no clima, terra e estresse experimentadas pelas videiras possam alterar as propriedades sensoriais do vinho, um conceito que os franceses denominam de "*terroir*". Assim, ao contrário de outras *commodities*, o mercado do vinho possui um componente que apela para sua região de origem, movimentando o turismo e investimentos econômicos em algumas regiões (BISSON, 2002).

Nesse contexto, origina-se a enologia, o estudo dos vinhos e de cada etapa

da sua produção. A fim de inovar neste campo de pesquisa, e aprimorar as características sensoriais do vinho, atenção tem sido dada a outros elementos no processo de fermentação que foram desprezados previamente (CAPOZZI, 2015). Em particular, o efeito que leveduras não-convencionais apresentam sobre as propriedades organolépticas (aroma, coloração, sabor, etc.) do vinho tem despertado particular interesse científico (GAMERO, 2020).

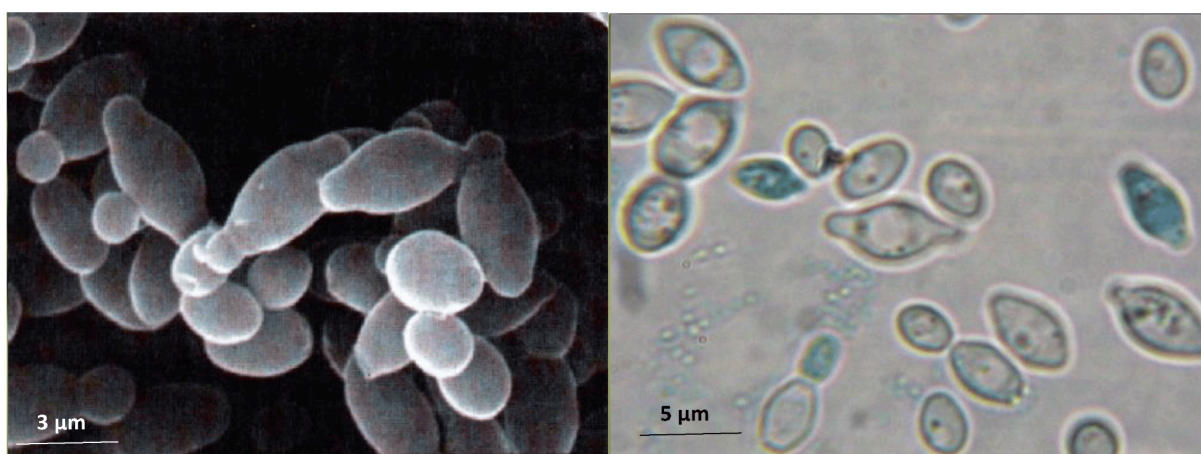
Naturalmente, os frutos frescos de uva contêm uma microbiota muito diversa, o que inclui diversos gêneros distintos de leveduras tais como *Brettanomyces*, *Candida*, *Hanseniaspora*, *Pichia*, entre muitos outros (RASPOR, 2006). A fermentação alcoólica do vinho é realizada principalmente por leveduras do gênero *Saccharomyces*, pois esse grupo apresenta alta taxa de fermentação e tolerância a concentrações elevadas de etanol. Entretanto, leveduras *Saccharomyces* não são naturalmente encontradas na microbiota da uva, pois estas são incapazes de degradar a parede celular dos frutos frescos, dependendo de outros microrganismos para isso, surgindo apenas tardiamente no mosto durante o processo de fermentação espontânea do vinho (VADKERTIOVÁ, 2012; WATANABE, 2023).

Outras leveduras presentes no mosto do vinho, porém menos relacionadas com o gênero *Saccharomyces*, foram denominadas pelo termo não-*Saccharomyces* (em inglês, “*non-Saccharomyces yeasts*”) e apresentavam um papel secundário na fermentação da uva, recebendo pouco interesse científico. Algumas dessas leveduras não-*Saccharomyces* compõem grande parte da microbiota natural da uva (BINATI, 2020). Recentemente, essa visão mudou na comunidade científica. Alguns trabalhos têm observado efeitos positivos no uso de leveduras não-*Saccharomyces* para alteração da composição química, melhoria do conteúdo nutricional, controle de acidez e na introdução de variações no aroma, coloração e gosto do vinho (MORATA, 2021; WINDHOLTZ, 2021; AKAN, 2023).

### 1.3 O gênero *Hanseniaspora*

*Hanseniaspora* é o gênero de leveduras com células apiculadas mais comum, sendo naturalmente encontradas em alguns frutos, especialmente uvas onde algumas espécies compõem a maior parte da microbiota do fruto. *Hanseniaspora* é um dos gêneros não-*Saccharomyces* mais importantes na fermentação do vinho e

algumas outras bebidas alcólicas (ROMANO, 1992; CARRAU, 2015; STERNES, 2016; VARELA, 2017). Esse gênero de leveduras ocupa um nicho ecológico muito específico e estreito, conhecido pela capacidade de metabolizar poucas fontes de carbono, principalmente açúcares simples, além de possuírem alta demanda de nutrientes mais complexos como vitaminas, o que sugere que estão sempre associados às plantas onde são usualmente encontradas (KURTZMAN, 2011). O gênero é muito diverso, mas facilmente confundido com outros grupos de leveduras relacionados ou morfológicamente semelhantes, como *Kloeckera*, que também são naturalmente encontrados na microbiota dos frutos fermentáveis (CADEZ, 2002). Na **Figura 2**, é exibida a morfologia celular de duas espécies pertencentes ao gênero *Hanseniaspora*.



**Figura 2:** Morfologia de leveduras apiculadas do gênero *Hanseniaspora*. À esquerda, uma microscopia eletrônica de *H. uvarum*, à direita, células da espécie *H. vineae* coloridas com corante azul de metileno (Modificado de MARTIN, 2018).

Inicialmente, o gênero foi pouco estudado, com alguns estudos focados na identificação molecular de espécies de *Hanseniaspora* (ESTEVE-ZARZOSO, 2001; CADEZ, 2002). Recentemente, graças à tecnologia de sequenciamento e análise genômica, um trabalho muito significativo para a elucidação da história evolutiva desse gênero foi conduzido por STEENWYK *et al.* (2019), que sequenciaram e compararam vários genomas de diferentes espécies de *Hanseniaspora*, descobrindo dentro desse gênero alguns dos menores genomas entre leveduras estudadas, o que evidencia uma evolução marcada por extensiva perda de genes. Notadamente,



muitos genes responsáveis pela regulação do ciclo celular e mecanismos de reparação do genoma estão ausentes em todos os genomas estudados, o que sugere que a diversificação do gênero *Hanseniaspora* se deu a partir de uma linhagem ancestral possuindo fenótipo de hipermutação. Além disso, o autor identificou dois grupos claramente distintos dentro do gênero: um grupo de lenta evolução e menor diversificação, nomeado “*Slow Evolving Lineages*” (SEL), e um grupo de rápida evolução e alta diversificação, cujos genomas são excepcionalmente erodidos, nomeado “*Fast Evolving Lineages*” (FEL) (STEENWYK, 2019).

Leveduras do gênero *Hanseniaspora* são frequentemente encontradas durante o processo de fermentação natural do vinho, aparecendo nos primeiros estágios da fermentação e sendo, posteriormente, substituídas por *Saccharomyces cerevisiae* (SPENCER, 1997). Em particular, *Hanseniaspora uvarum* parece ser uma das espécies mais abundantes em frutos maduros de uva e no mosto nos primeiros estágios da fermentação, e por isso uma das leveduras mais bem estudadas desse gênero e com mais cepas conhecidas (ALBERTIN, 2016). Ultimamente, o interesse na fermentação por leveduras não-*Saccharomyces* tem aumentado devido ao potencial destas para conferir características únicas de aroma e sabor aos vinhos e outras bebidas alcólicas, bem como a possibilidade co-fermentação utilizando múltiplas leveduras, com algumas espécies pertencentes a *Hanseniaspora* sendo muito estudadas para esse fim (GRANCHI, 2002; PIETRAFESA, 2020; VALERA, 2021).

#### 1.4 Pangenomas e prospecção de genes

A Genômica é o campo da genética que se concentra no estudo e análise de genomas completos, a fim de compreender as informações hereditárias, as funções de suas partes, o desenvolvimento de um organismo e sua evolução (THANGADURAI, 2015). Inicialmente a genômica foi limitada devido à dificuldade técnica de se sequenciar genomas inteiros, porém o surgimento das técnicas de “*Next Generation Sequencing*” (NGS) revolucionou o estudo da genômica, tornando possível o sequenciamento de genomas inteiros de maneira rápida e a um custo cada vez mais acessível (PERVEZ, 2022). Atualmente, o principal desafio da

genômica consiste no desenvolvimento de melhores ferramentas de bioinformática para estudar a enorme quantidade de dados de sequenciamento disponíveis (GOLDMAN, 2016; BIANCONI, 2023).

Uma vez que mais espécies têm seu genoma sequenciado, o paradigma da análise genômica tem mudado do estudo da genética do indivíduo para a genética de populações. Nesse contexto, um conceito emergente é o pangenoma, caracterizado como o conjunto completo de genes e variações na sequência genética, incluindo eventos estruturais, dentro de todas as espécies ou cepas pertencentes a um clado (BROCKHURST, 2019). Pangenoma é portanto a representação de toda a diversidade genética presente em um clado (HICKEY, 2023). Essa abordagem é útil, entre outras coisas, para estudar a história e as tendências evolutivas, a plasticidade genômica e adaptabilidade das espécies ou cepas pertencentes ao clado representado como um todo (BARH, 2020).

Identificar novas sequências genéticas é uma etapa fundamental no processo de prospecção de genes, ou compostos produzidos por eles, com potencial para exploração biotecnológica em diversos setores, como a agricultura, alimentação, farmacêutica, indústria, etc. (SCHLOSS, 2003; ALBARANO, 2020). Atualmente, somos capazes de explorar a diversidade biológica através de ferramentas de bioinformática usando unicamente informações de sequência. Essa abordagem tem sido útil para prospecção de genes a partir de genomas individuais e mesmo de metagenomas, isto é, genomas de comunidades microbianas heterogêneas (ZERIKLY, 2009; ROUMPEKA, 2017). A mesma estratégia também se aplica a outras comunidades de organismos. Nesse sentido, a construção de pangenomas também é útil para prospecção de genes com potencial biotecnológico a partir de um grupo de organismos de interesse (GONG, 2023).

## 2. JUSTIFICATIVA

O gênero *Hanseniaspora* ocupa um papel fundamental na vinicultura, pois através da fermentação, essas leveduras contribuem para a formação de compostos aromáticos e metabólitos que conferem aroma, sabor e complexidade ao vinho, o que as tornam um interessante alvo de vários estudos. Dada a sua história evolutiva única, recentemente descoberta, e a grande diversidade de espécies, o gênero *Hanseniaspora* apresenta potencial para prospecção de genes. A compreensão mais profunda da diversidade genética exclusiva desse grupo, pode revelar novos genes com possíveis aplicações biotecnológicas em processos industriais diversos, especialmente na produção de vinhos.

## 3. OBJETIVOS

### 3.1 Objetivo Geral

Caracterizar a diversidade de ortólogos presentes no gênero *Hanseniaspora*, identificar tendências evolutivas dentro do gênero e conjecturar sobre a possível influência destas no processo fermentativo conduzido por essas leveduras.

### 3.2 Objetivos Específicos

- Selecionar genomas em fase de *scaffold* representativos do gênero *Hanseniaspora* e conduzir a anotação gênica de cada um deles.
- Identificar ortólogos comuns entre *Saccharomyces cerevisiae* e *Hanseniaspora* e proteínas presentes unicamente em *Hanseniaspora*.
- Realizar a anotação funcional dos ortólogos exclusivos de *Hanseniaspora* e identificar funções biológicas enriquecidas no gênero.
- Obter *insights* sobre a diversidade genética exclusiva de *Hanseniaspora* e o possível efeito destes no processo fermentativo realizado por essas leveduras.

## 4. MATERIAIS E MÉTODOS

### 4.1 Seleção de genomas

Um total de 60 genomas de *Hanseniaspora* foram selecionados a partir do banco de dados GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>, acesso a 21/08/2023) utilizando o termo de busca "Hanseniaspora[Organism]", representando o total de genomas associados ao gênero presentes no banco de dados. Os genomas obtidos e seus dados estão representados na **Tabela S1**. Estatísticas de qualidade foram geradas, para cada montagem, usando o software QUAST (GUREVICH, 2013). O software BUSCO (SIMÃO, 2015) foi usado para acessar os valores de completude (uma estatística que corresponde à porcentagem de ortólogos identificados entre a montagem e um genoma de referência, sendo útil para avaliar a qualidade de uma montagem) em modo de sequência genômica em relação ao conjunto de dados saccharomycetes\_odb10 de *S. cerevisiae*.

### 4.2 Anotação dos genomas

Os *softwares* de predição gênicas Augustus (STANKE, 2004), Genemark-ES (BORODOVSKY, 2011) e Geneid (PARRA, 2000) foram utilizados para anotação dos genomas de *Hanseniaspora*. Dado que modelos de espécie são necessários para o uso das ferramentas Augustus e Geneid, foram utilizados os modelos de leveduras "saccharomyces" e "sjaponicus.param\_Oct\_12\_2006", para Augustus e Geneid, respectivamente. A **Tabela S2** possui os parâmetros utilizados para cada anotador. Os arquivos em formato .gff3 contendo as anotações geradas pelas ferramentas para cada um dos genomas foram concatenados e genes redundantes foram removidos usando o *software* gffcompare (PERTEA, 2020). Peptídeos com extensão superior a 100 aminoácidos foram identificados através do utilitário TransDecoder.LongOrfs (<https://github.com/TransDecoder/TransDecoder>, acesso em 30/09/2023). Para cada espécie, os proteomas resultantes contendo as sequências de aminoácidos para todas as proteínas preditas foram salvos em formato FASTA.

### 4.3 Identificação de ortólogos

O *software* OrthoFinder (EMMS, 2019) foi utilizado para identificar famílias de ortólogos entre *Hanseniaspora* e *Saccharomyces cerevisiae* e famílias de não-ortólogos exclusivas para o gênero *Hanseniaspora*. Para a execução da ferramenta, o proteoma completo de *Saccharomyces cerevisiae* foi obtido a partir da plataforma Ensembl ([https://ftp.ensembl.org/pub/release-110/fasta/saccharomyces\\_cerevisiae/pep/Saccharomyces\\_cerevisiae.R64-1-1.pep.all.fa.gz](https://ftp.ensembl.org/pub/release-110/fasta/saccharomyces_cerevisiae/pep/Saccharomyces_cerevisiae.R64-1-1.pep.all.fa.gz), acesso em 01/10/2023) e usado como referência na identificação de ortólogos e não-ortólogos com os proteomas anotados de *Hanseniaspora*. Utilizando um *script* em linguagem Python simples, os ortogrupos (a denominação dada pela ferramenta aos agrupamentos de genes relacionados) identificados por OrthoFinder foram divididos em três categorias: ortólogos compartilhados com *Saccharomyces*, ortólogos exclusivos do gênero *Hanseniaspora* e genes que não puderam ser associados a nenhum ortogrupo pela ferramenta.

#### 4.4 Filogenômica do grupo *Hanseniaspora*

A fim de construir a relação filogenética entre os genomas estudados e classificá-los dentro dos grupos FEL e SEL identificados por STEENWYK *et al.* (2019), isolou-se um grupo de 23 ortogrupos de gene único compartilhados entre *S. cerevisiae* e todos os genomas analisados. Utilizando o método de concatenação (LUO, 2021), as sequências polipeptídicas desses ortólogos foram reunidas e concatenadas em um arquivo multi-*fasta* e submetidas ao alinhamento múltiplo de sequências usando o *software* MUSCLE (EDGAR, 2004) v.1.3.8 usando parâmetros *default*. O alinhamento em formato Phylip resultante foi utilizado para performar a inferência filogenética usando o método de máxima verossimilhança (GUINDON, 2010) através da interface Web de PhyML 3.0 (<http://www.atgc-montpellier.fr/phyml/>, acesso a 03/12/2023) utilizando seleção automática de modelo e suporte estatístico dos ramos por aLRT SH-Like (ANISIMOVA, 2011). A árvore resultante foi visualizada usando a ferramenta FigTree (<https://github.com/rambaut/figtree/>, acesso em 07/12/2023), destacando-se os genomas pertencentes aos grupos FEL e SEL. Um *script* em linguagem R foi usado para plotar as médias dos números totais de transcritos preditos, ortólogos compartilhados e não-compartilhados com *S.*

*cerevisiae* (Figura 5).

#### 4.5 Anotação funcional de sequências ortólogas

Os ortólogos pertencentes a ortogrupos comuns entre *S. cerevisiae* e *Hanseniaspora* tiveram sua anotação funcional atribuída por meio da associação com os termos GO (*Gene Ontology*) com os genes de *S. cerevisiae* através de um script Python usando como referência as anotações de termos GO para todos os genes de *S. cerevisiae*, obtidas através da interface *Web BioMart* dentro da plataforma Ensembl (<https://www.ensembl.org/biomart/martview>, acesso em 24/11/2023). Para os ortogrupos exclusivos de *Hanseniaspora* e os genes não atribuídos por OrthoFinder, a anotação funcional foi executada através da ferramenta InterProScan5 (JONES, 2014), com arquivo de saída em formato \*.TSV com identificação dos termos GO. Domínios PFAM de genes mapeados por InterProScan5, que não puderam ter termos de ontologia anotados pela ferramenta, foram resgatados e os termos de ontologia associados foram atribuídos aos transcritos de forma manual usando um script python, usando como referência um arquivo de associação entre domínios PFAM e seus termos GO proveniente de GODomainMiner (<https://godm.loria.fr/download/>, acesso a 29/12/2023).

#### 4.6 Análise de enriquecimento de ontologia

Para cada genoma de *Hanseniaspora*, os termos GO associados a cada um dos transcritos preditos foram listados em um arquivo de formato TSV, utilizando um *script* python. A fim de produzir perfis de enriquecimento para os grupos FEL e SEL, termos de ontologia e suas associações foram obtidos a partir da biblioteca GO.db (<https://bioconductor.org/packages/release/data/annotation/html/GO.db.html>, acesso em 29/12/2023) para a linguagem R. Preliminarmente, os ortólogos compartilhados com *S. cerevisiae* foram removidos do conjunto de genes analisados, a fim de identificar termos enriquecidos apenas no conjunto de ortólogos exclusivos de *Hanseniaspora*. Anotações para transcritos identificadas por InterProScan5 ou mapeados a partir de domínios PFAM contendo termos GO obsoletos também foram removidos do conjunto de transcritos analisados.

A análise de enriquecimento para os genomas pertencentes aos grupos FEL e SEL foi realizada utilizando a biblioteca clusterProfiler (WU, 2021) para a linguagem R, utilizando o método *Benjamini-Hochberg Procedure* (BH) para ajuste de valores p (HAYNES, 2013). Termos de ontologia com p ajustado  $< 0.01$  foram considerados enriquecidos dentro dos conjuntos de dados analisados. A biblioteca enrichplot (<https://bioconductor.org/packages/release/bioc/html/enrichplot.html>, acesso em 29/12/2023) foi utilizada para produzir as **Figuras 7** e **Figura 8** do enriquecimento funcional para os grupos FEL e SEL, e a biblioteca rrvgo (<https://bioconductor.org/packages/release/bioc/html/rrvgo.html>, acesso em 29/12/2023) foi utilizada para compor a **Figura 9**, que exibe os principais processos associados aos ortólogos exclusivos dentro do gênero *Hanseniaspora* como um todo, através da redução semântica dos GO Terms enriquecidos em ambos os grupos FEL e SEL.

## 5. RESULTADOS E DISCUSSÃO

### 5.1 Classificação dos genomas nos grupos FEL e SEL

No estudo conduzido por STEENWYK *et al.* (2019), o gênero *Hanseniaspora* revelou possuir alguns dos menores genomas de leveduras conhecidos, indicando grande redução genômica ao longo da história evolutiva do gênero. Durante a construção da filogenia de *Hanseniaspora* foi revelada a clara distinção de dois grupos distintos dentro do gênero, um grupo de lenta evolução e menos diversificado, SEL, e um grupo de rápida evolução, genomas muito erodidos, porém com grande diversificação de espécies, FEL. No estudo em questão, entretanto, apenas 25 genomas pertencentes ao gênero foram analisados.

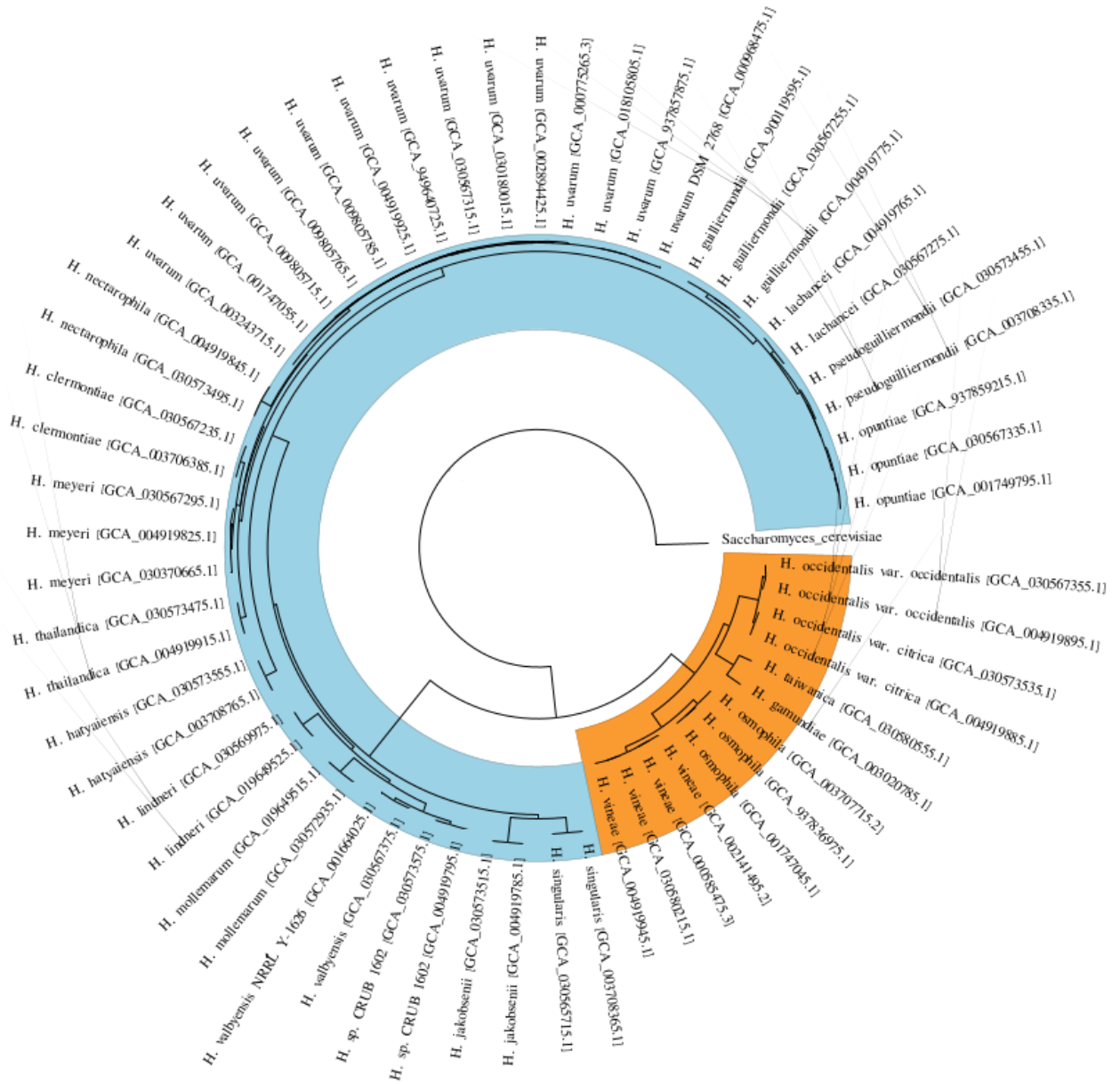
A fim de classificar os 60 genomas obtidos para esse trabalho dentro dos grupos de rápida e lenta evolução, adotaram-se os procedimentos Seleção de genomas-Filogenômica do grupo *Hanseniaspora* narrados na seção de métodos. A árvore filogenética gerada pode ser visualizada na **Figura 3**. Nela, os grupos FEL e SEL podem ser facilmente distinguidos, a posição de cada genoma dentro da árvore filogenética foi usada para agrupá-los dentro dos grupos FEL e SEL.

A **Figura 4** exibe um gráfico do log(N50) pela completude de cada genoma, com cada ponto colorido de acordo com o grupo ao qual pertence. Como a estatística de completude de BUSCO é baseada na identificação de ortólogos com uma espécie aparentada, valores baixos dessa estatística podem indicar perda de genes no genoma ou uma baixa qualidade da montagem. A sobreposição dessa estatística com o log(N50) permite diferenciar os genomas de boa qualidade, porém baixos valores de completude. É notável a divisão clara entre os grupos, mesmo genomas do grupo SEL com menor qualidade são significativamente mais completos que genomas do grupo FEL devido à maior perda de genes dentro de FEL.

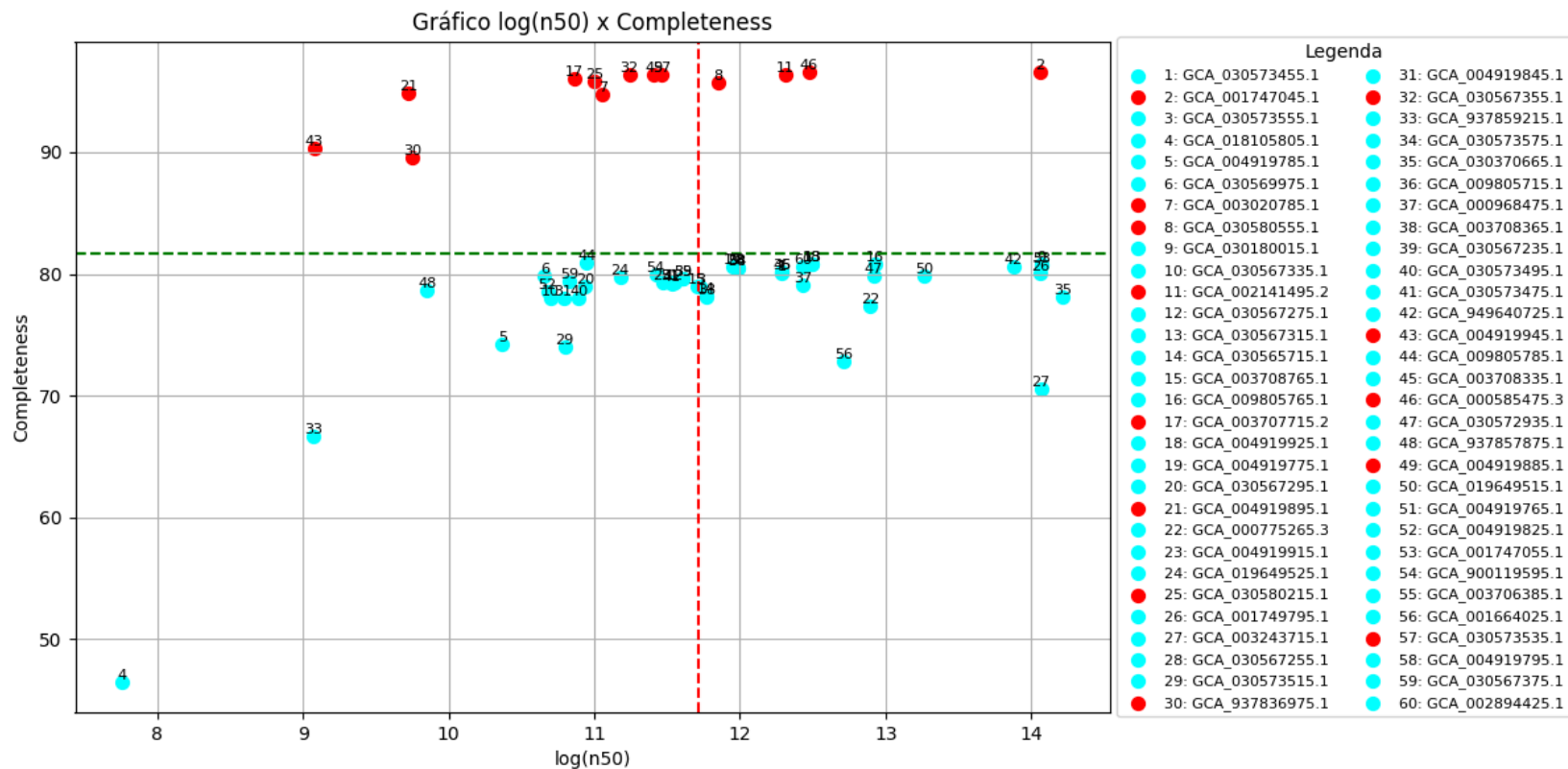
A **Figura 5** exibe as diferenças nas médias do número de genes anotados, ortólogos compartilhados com *S. cerevisiae* e ortólogos exclusivos de *Hanseniaspora* entre os grupos FEL e SEL. Em todas as métricas, as médias do grupo SEL são significativamente maiores que as do grupo FEL. Levando em consideração ambas as figuras em conjunto, em concordância com STEENWYK *et al.* (2019), esses dados reafirmam a maior redução genômica experimentada pelas



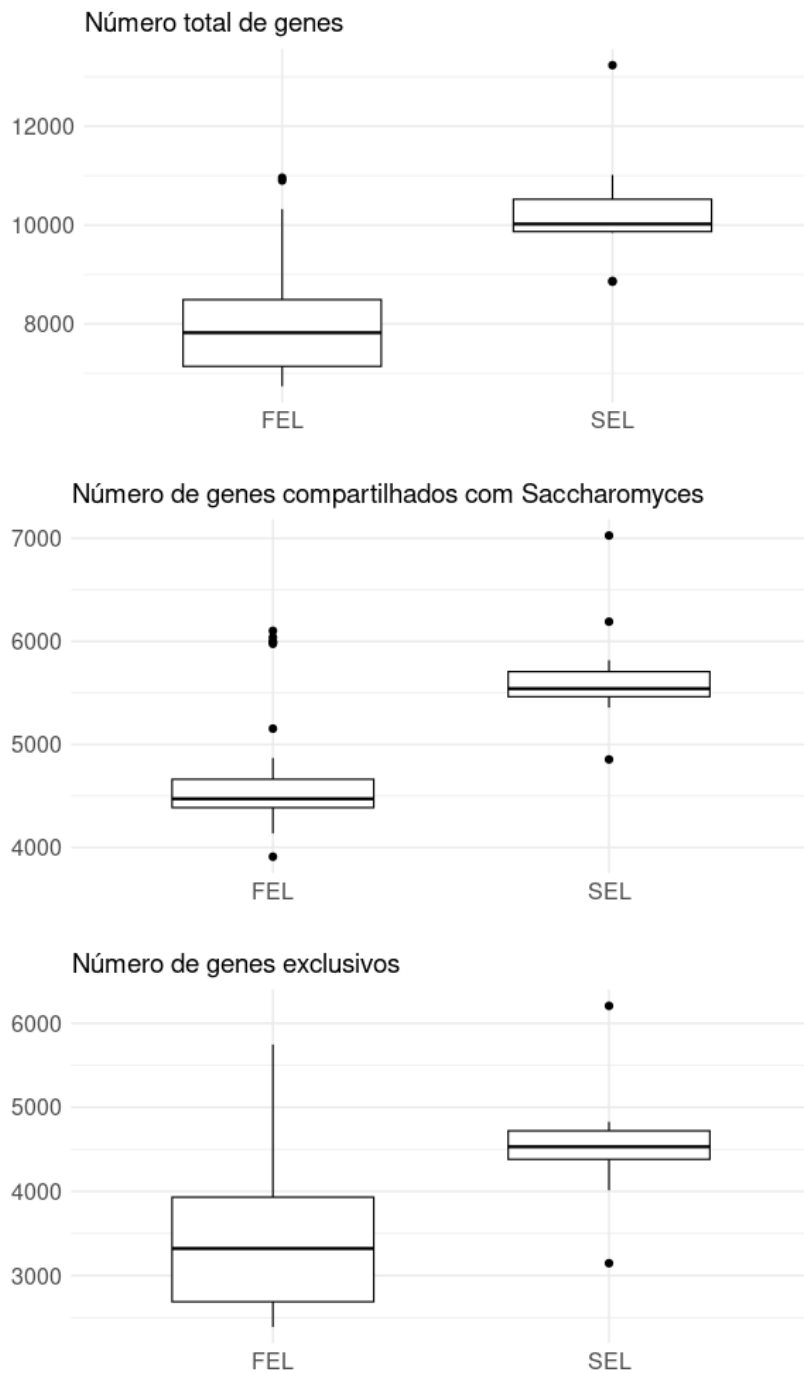
espécies pertencentes ao grupo FEL.



**Figura 3:** Árvore filogenética do gênero *Hanseniaspora*. A raiz da árvore é *S. cerevisiae*. Dois grupos claramente distintos irradiam a história evolutiva do gênero, FEL e SEL. Em laranja, temos o grupo de lenta evolução SEL e, em azul, o grupo de rápida evolução FEL.



**Figura 4:** Gráfico log(N50) x Completude. Os genomas estão coloridos de acordo com o grupo ao qual pertencem. Pontos em vermelho representam genomas pertencentes ao grupo SEL e pontos em azul representam genomas pertencentes ao grupo FEL. As linhas pontilhadas vermelha e verde, representam as médias das métricas log(N50) e Completude, respectivamente.

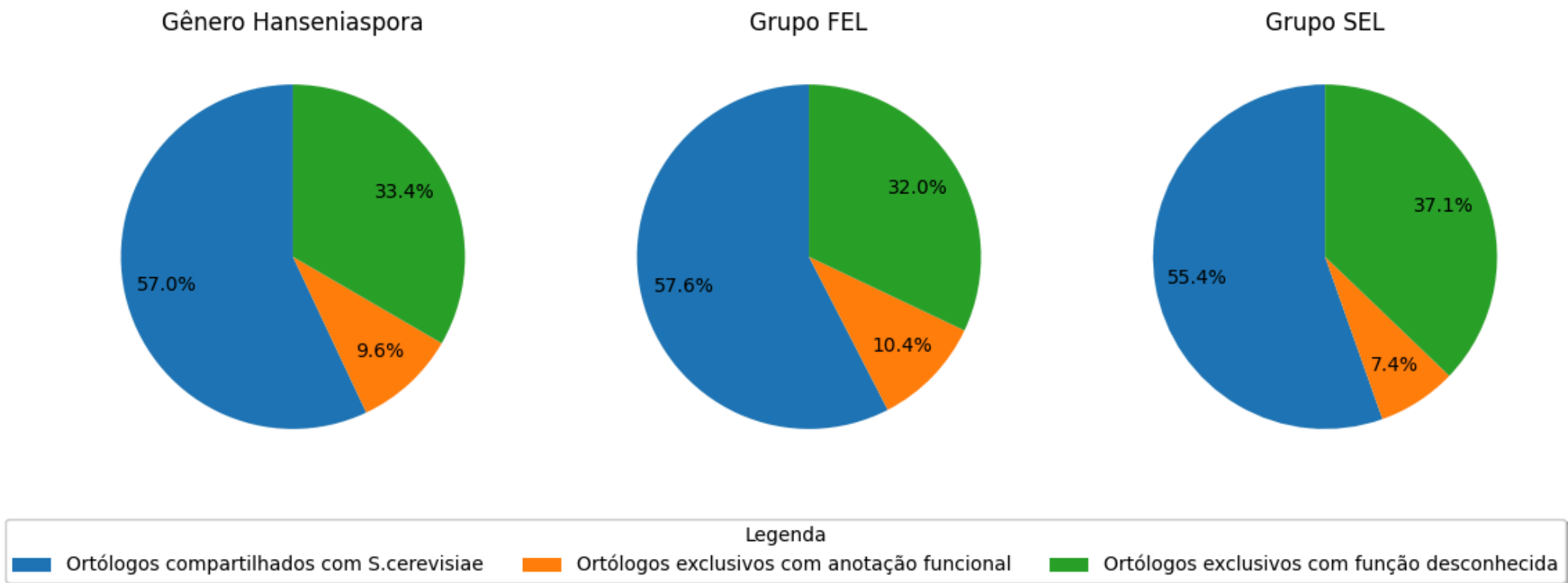


**Figura 5:** Distribuição dos ortólogos identificados dentro dos grupos FEL e SEL de *Hanseniaspora*. De cima para baixo, os gráficos boxplot representam a média em: número de genes anotados, número ortólogos compartilhados com *S. cerevisiae* e número de ortólogos exclusivos para cada um dos genomas analisados.

## 5.2 Anotação funcional dos ortólogos

Após a identificação dos ortogrupos compartilhados com *S. cerevisiae*, os transcritos exclusivos de *Hanseniaspora* foram anotados por InterProScan5. A ferramenta conseguiu atribuir caracterização funcional apenas para uma pequena porcentagem dos transcritos de *Hanseniaspora*, como pode ser observado na **Figura 6**. Essa identificação irregular de anotações funcionais por InterProScan5 pode ser explicada, em parte, por uma limitação da própria ferramenta, que é sujeita a vários cenários, onde mudanças em seus bancos de dados podem impactar significativamente na atribuição de termos de ontologia, como foi explorado por SANGRADOR-VEGAS *et al.* (2016).

Para fins de comparação, em CADEZ *et al.* (2019), onde o genoma da espécie *Hanseniaspora gamundiae*, o qual também é representado aqui, foi caracterizado, o número de genes preditos foi de 4.624, um número bem próximo de outras espécies de *Hanseniaspora* analisadas no trabalho em questão, em comparação com os 11.016 transcritos preditos para *Hanseniaspora gamundiae*, neste trabalho (**Tabela S1**). Pode-se supor que isso é principalmente devido à metodologia mais flexível adotada aqui, pois com o objetivo de identificar o maior número possível de transcritos para cada genoma, optou-se por realizar múltiplas anotações com 3 ferramentas distintas: Augustus, Genemark-ES e Geneid (ver **Tabela S2**) e reunir transcritos não redundantes, enquanto CADEZ, *et al.* (2019) optou por realizar a anotação de seus genomas usando apenas a ferramenta Genemark-ES com parâmetros mais exigentes. Como resultado, dos 4.624 genes preditos, 4.128 (89.3%) foram anotados com sucesso através da ferramenta Blast2GO (CONESA, 2005) utilizada no estudo citado.

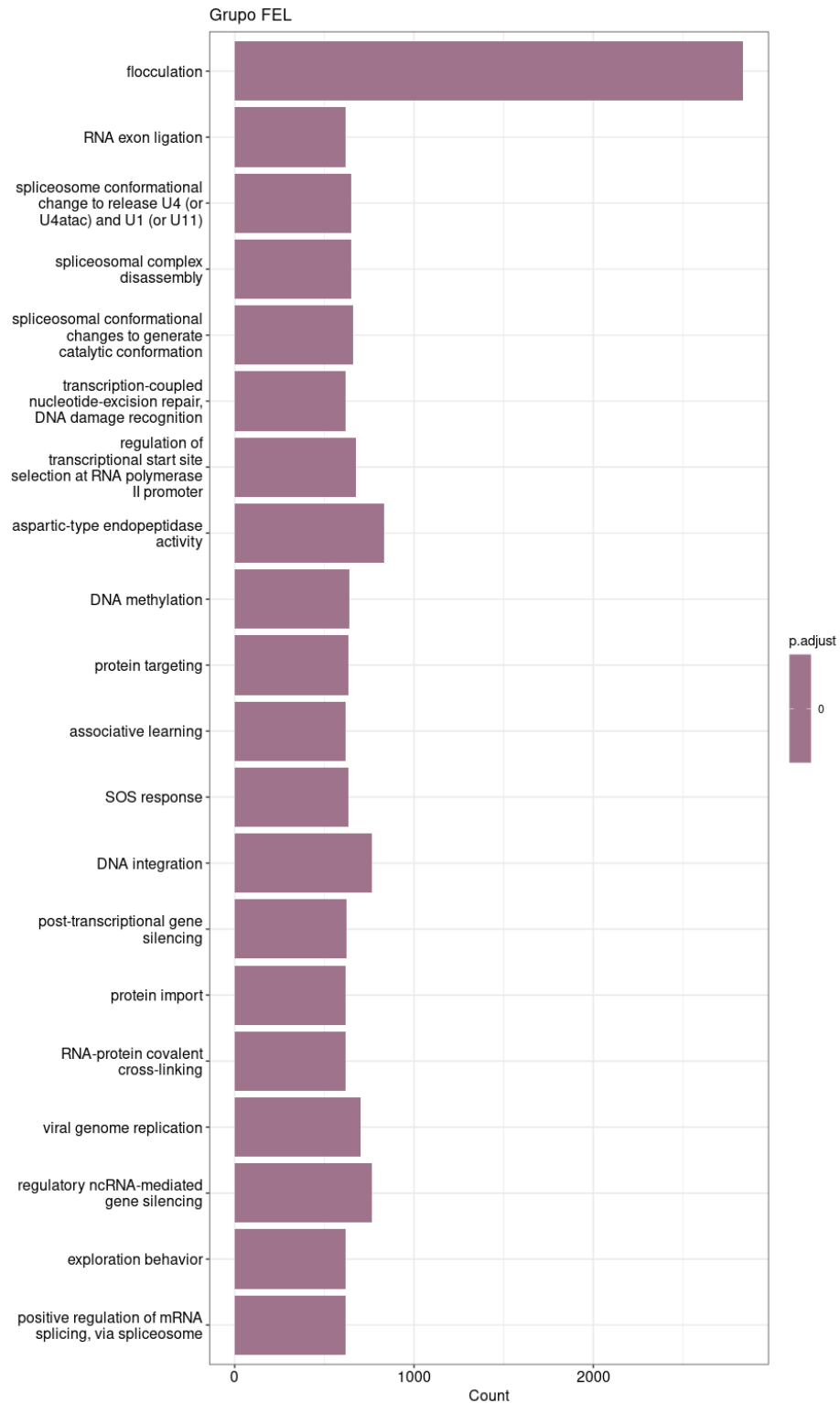


**Figura 6:** Porcentagem médias de ortólogos compartilhados com *S. cerevisiae* e ortólogos exclusivos de *Hanseniaspora* com e sem função biológica predita. Para ambos os grupos FEL e SEL, apenas uma pequena parte dos transcritos pôde ser anotada, a maior parte dos transcritos ainda possui função desconhecida.

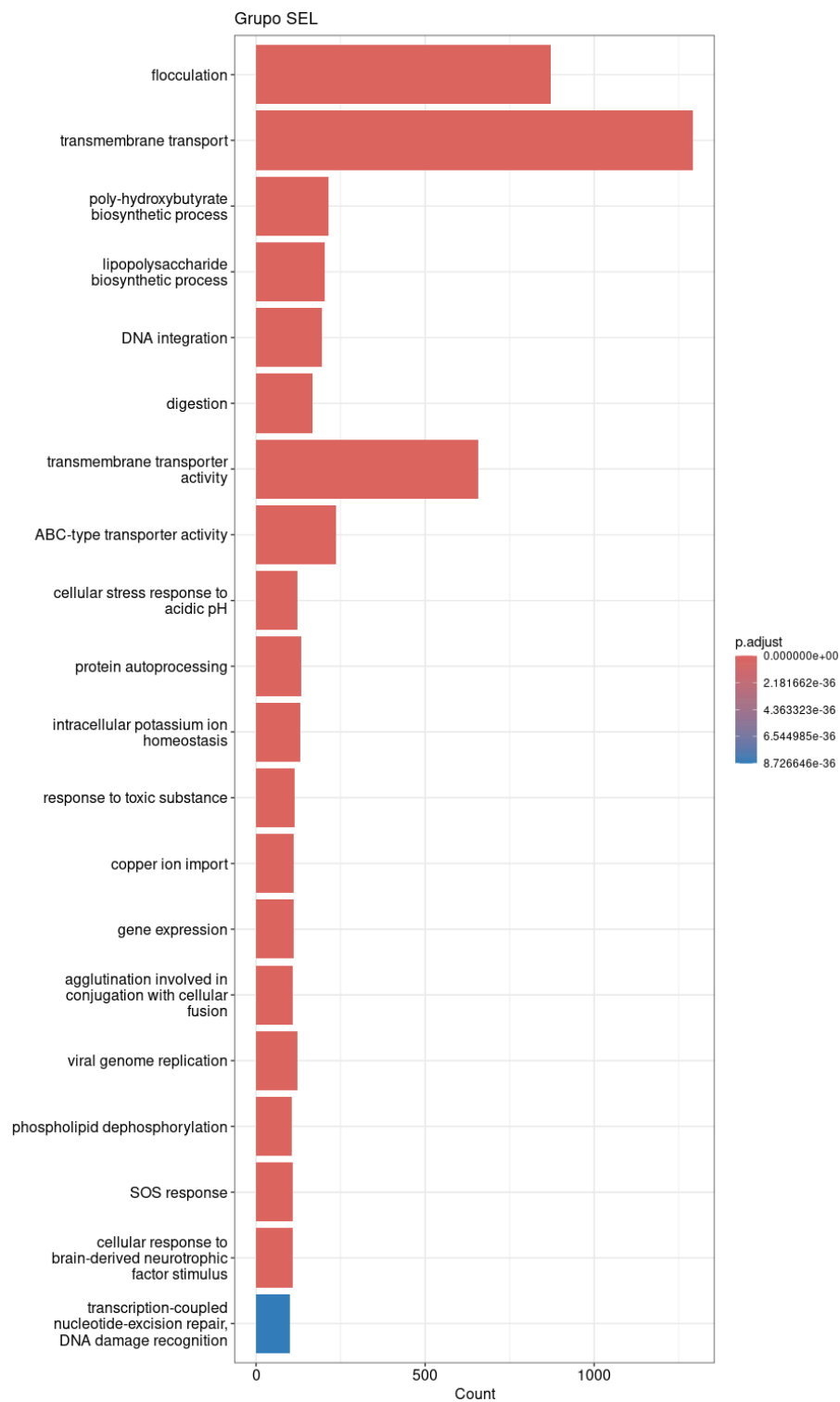
### 5.3 Enriquecimento de ontologia em *Hanseniaspora*

Com o objetivo de identificar o enriquecimento funcional no contexto dos grupos FEL e SEL operou-se uma análise de enriquecimento sobre os termos de ontologia presentes nos grupos seguindo os procedimentos da seção “Análise de enriquecimento de ontologia”. Os termos associados com  $p$  ajustado  $< 0.01$  foram considerados como sendo significativamente enriquecidos dentro dos grupos. A análise retornou um conjunto de 1.233 e 1.053 termos de ontologia significativamente enriquecidos nos grupos FEL e SEL, respectivamente. Os 20 termos associados a funções biológicas mais significativos nos grupos FEL e SEL são apresentados em **Figura 7** e **Figura 8**, respectivamente.

As assinaturas nos dois grupos, em nível global, são muito semelhantes entre si. Em ambos, o termo mais enriquecido é floculação, um processo observado em algumas leveduras que, como explicado por GUARAGNELLA *et al.* (2020), consiste na agregação das células formando aglomerados denominados flocos. Esse processo é de grande interesse para a indústria vinícola, pois aumenta a resistência às condições de estresse do meio de fermentação e facilita a filtragem e remoção das células do vinho produzido (VARELA, 2020). No estudo citado, várias cepas de *H. uvarum* isoladas de fermentações espontâneas de uvas, tiveram seus genomas sequenciados e anotados, e análises de enriquecimento de domínios mostraram que muitas proteínas preditas para as linhagens apresentaram assinaturas relacionadas ao processo de floculação, embora apenas duas das linhagens estudadas exibisse a formação de flocos em testes experimentais (GUARAGNELLA, 2020). Proteínas com função associadas ao processo de floculação, devem ser portanto características do gênero *Hanseniaspora*, ainda que nem todos os seus membros sejam capazes de realizar esse processo.



**Figura 7:** *Barplot* de enriquecimento do grupo FEL. As funções biológicas associadas aos 20 termos de ontologia mais representados dentro do grupo FEL. Todos os 20 primeiros termos associados ao grupo FEL possuíam  $p$  ajustado  $< 0.0001$ .



**Figura 8:** *Barplot* de enriquecimento do grupo SEL. As funções biológicas associadas aos 20 termos de ontologia mais representados dentro do grupo SEL.



Outros processos enriquecidos em *Hanseniaspora* estão associados à regulação e biossíntese de componentes da membrana e parede celular, em especial, fosfatidilinositol e manoproteínas. Como evidenciado nos trabalhos de HENDERSON *et al.* (2013) e LÓPEZ-MALO *et al.* (2015), há crescente evidência de que a composição lipídica da membrana se altera durante o prosseguimento da fermentação e que o conteúdo de manoproteínas na parede celular desempenha um papel importante na tolerância à fermentação em baixas temperaturas, que é tipicamente empregada para manter os componentes voláteis que conferem o perfil aromático de alguns vinhos. Entretanto, muitas linhagens de leveduras atualmente são pouco adaptadas para proceder a fermentação de forma satisfatória em condições de baixa temperatura, o que pode levar a problemas como lento crescimento da levedura e fermentação reduzida (MASSERA, 2021). De fato, *H. vineae* e outras leveduras não-*Saccharomyces* se mostraram promissoras em operar a fermentação a temperaturas mais baixas durante as fases iniciais de co-cultura com *S. cerevisiae* (MATURANO, 2015; ZHANG, 2018).

Na revisão de RIBEIRO *et al.* (2022) é apontado o papel que a parede celular fúngica exerce na resposta e tolerância a múltiplos tipos de agentes estressores, portanto podemos conjecturar que o enriquecimento de termos relacionados à organização da parede celular pode indicar possíveis genes de *Hanseniaspora* relacionados com a resposta ao estresse. Outros termos enriquecidos em ambos os grupos estão principalmente relacionados com funções regulatórias, controle da expressão gênica, transporte de íons e substâncias, metabolismo de açúcares e aminoácidos, replicação, reparo por recombinação homóloga, processos de biossíntese, resposta a nutrientes, tolerância a estresses, etc. Esses termos são especialmente relevantes porque podem estar associados ao ciclo de vida e ao processo de fermentação em leveduras do gênero *Hanseniaspora*.

A colonização de *Hanseniaspora*, bem como outras espécies de microrganismos que compõem a microflora da uva, ocorre inicialmente na casca do fruto, trazidos muito provavelmente por insetos atraídos pelos compostos voláteis produzidos por essas leveduras ou partículas do solo, permitindo a propagação das leveduras para novos hospedeiros (SIPICZKI, 2016). Frutos como a uva, possuem uma espessa casca composta principalmente por componentes da parede celular

vegetal como celulose, lignina, hemicelulose, pectina, etc., que age como uma barreira de proteção contra a dessecação, dano físico e invasão de microrganismos. Algumas espécies de fungos e bactérias conseguem se manter de maneira estável nesse ambiente, devido a capacidade de obter nutrição a partir da casca por meio da secreção de enzimas, como celulasas, pectinases e cutinases. A degradação da casca expõe a polpa do fruto, permitindo que leveduras como *S. cerevisiae*, que não sobrevivem na casca do fruto, tenham acesso a essa reserva nutricional, dando início à fermentação. O mosto do vinho é composto principalmente pelo conteúdo da polpa da uva, que ao contrário da casca, é mais adequada para a reprodução das leveduras por ser rica em açúcares e outros nutrientes importantes (WATANABE, 2023).

A secreção de enzimas importantes como beta-glucosidase, glucanases, pectinase e proteases já havia sido identificada em algumas espécies pertencentes ao gênero *Hanseniaspora*, o que está de acordo com a capacidade dessas leveduras de subsistir na superfície das uvas (MANE, 2017). No entanto, várias evidências sugerem uma tendência à adaptação de *Hanseniaspora* às condições de estresse do ambiente fermentativo de forma análoga a algumas adaptações vistas em *S. cerevisiae*, como resposta à privação de nitrogênio, produção de compostos inibitórios contra outros organismos e duplicação de genes relacionados ao processo fermentativo (LLEIXÀ, 2019; CARRAU, 2023; VAN WYK, 2024). De fato, a maioria das leveduras do gênero *Hanseniaspora*, com a notável exceção de *Hanseniaspora smithiae* sp. nov., uma espécie que perdeu várias rotas metabólicas relacionadas à fermentação e não se encontra tipicamente associada com plantas, parece ter se especializado como fermentadores naturais em diversas variedades de frutos, conservando muitos dos genes envolvidos nesse processo (CADEZ, 2021; CARRAU, 2023).

Também foi sugerido que a perda de genes dentro de *Hanseniaspora* associados ao controle do ciclo celular pode estar associada a uma estratégia competitiva de crescimento acelerado (STEENWYK, 2019). Além disso, como abordado por TONDINI *et al.* (2020), a variabilidade das condições ao longo do processo de fermentação do mosto demanda respostas rápidas da levedura a essas mudanças, o que pode envolver muitos dos processos de regulação da expressão gênica e resposta ao estresse enriquecidos em *Hanseniaspora*. Como exemplo

temos que análises transcriptômicas de *H. uvarum*, uma das espécies mais promissoras para vinicultura, em meio fermentativo exibiram grandes variações no número de genes diferencialmente expressos ao longo de diferentes etapas da fermentação (GIORELLO, 2018).

Muitos dos termos identificados neste trabalho estão relacionados aos mesmos processos associados à estratégia de fermentação conduzida por *Hanseniaspora*, como reprodução acelerada, tolerância aos agentes estressores do mosto e competição interespecífica. Em menor número, termos associados com replicação viral também podem ser encontrados dentro do conjunto de transcritos exclusivos de *Hanseniaspora*, o que não é incomum, pois a presença de genoma viral parece ser comum em outras leveduras (BECKER, 2017). A **Figura 9** apresenta um resumo das principais funções biológicas enriquecidas no gênero *Hanseniaspora*.

Funções biológicas enriquecidas em *Hanseniaspora*



**Figura 9:** Treemap representando os principais termos de ontologia associados ao gênero *Hanseniaspora*. Essa figura foi produzida por meio da redução semântica dos GO Terms enriquecidos para ambos os grupos, FEL e SEL.

## 6. CONCLUSÃO E PERSPECTIVAS

No presente trabalho, diferentes genomas em fase de *scaffold* de *Hanseniaspora* foram selecionados e anotados. Uma filogenia foi construída para classificá-los dentro dos grupos de evolução rápida ou evolução lenta, FEL e SEL, respectivamente. O objetivo principal era realizar a anotação funcional dos genes preditos para os genomas e observar funções biológicas enriquecidas nos grupos.

A partir dos dados gerados, observa-se que o gênero *Hanseniaspora* apresenta uma grande diversidade, com muitos genes ainda não caracterizados que podem exercer atividades biológicas de interesse para o processo de fermentação desses organismos. Apesar de filogeneticamente distintos, os resultados obtidos neste trabalho sugerem uma uniformidade dentro do gênero, com muitas funções biológicas comuns entre os grupos FEL e SEL. Termos de interesse para a indústria, como a floculação e tolerância a estresses, são comuns a muitos dos transcritos exclusivos de *Hanseniaspora*.

Uma vez determinadas as funções biológicas enriquecidas dentro do gênero *Hanseniaspora*, a perspectiva seria proceder com a caracterização molecular dos transcritos preditos associados a essas funções, através de alinhamentos com sequências conhecidas em bancos de dados com NCBI, a determinação de sua estrutura e identificação de domínios proteicos por meio de ferramentas de bioinformática. É nessa etapa que podemos prospectar individualmente os genes candidatos mais promissores para uso industrial. Por exemplo, as enzimas, que alteram perfis de sabor e aroma no processo fermentativo ou fatores de transcrição envolvidos na resposta ao estresse, possibilitando seu isolamento, modificação, condução de testes experimentais e aplicação transgênica em outros organismos.

No cenário atual, o interesse por leveduras não convencionais vem crescendo significativamente, devido ao potencial destas em conferir alterações nas propriedades físico-químicas do vinho. Em particular, o gênero *Hanseniaspora* são algumas das leveduras mais promissoras nesse sentido. Um estudo mais aprofundado desse gênero oferece uma oportunidade única para aprimorarmos nossas técnicas de vinicultura e fermentação.

## REFERÊNCIAS BIBLIOGRÁFICAS

- AKAN, Madina *et al.* Exploring the Potential of Non-Conventional Yeasts in Wine Fermentation with a Focus on *Saccharomyces fermentans*. **Fermentation**, [s. l.], v. 9, n. 9, p. 786, 2023.
- ALBARANO, Luisa *et al.* Genome Mining as New Challenge in Natural Products Discovery. **Mar Drugs**, [s. l.], v. 18, n. 4, 2020.
- ALBERTIN, Warren *et al.* *Hanseniaspora uvarum* from Winemaking Environments Show Spatial and Temporal Genetic Clustering. **Frontiers in Microbiology**, [s. l.], v. 6, n. 1569, 2016.
- ANISIMOVA, Maria *et al.* Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. **Systematic Biology**, [s. l.], v. 60, n. 5, p. 685-699, 2011.
- BARH, Debmalya (ed.) *et al.* Pan-genomics: Applications, Challenges, and Future Prospects, 1th ed. **Academic Press**, [s. l.], 2020.
- BARNETT, James A. A history of research on yeasts 2: Louis Pasteur and his contemporaries, 1850–1880. **Yeast**, [s. l.], v. 16, n. 8, p. 755-771, 2000.
- BECKER, Björn *et al.* Yeast Killer Toxin K28: Biology and Unique Strategy of Host Cell Intoxication and Killing. **Toxins**, [s. l.], v. 9, n. 10, p. 333, 2017.
- BERGER, Donald G. QUALITY CONTROL IN THE BREWING INDUSTRY. **Journal of Milk and Food Technology**, [s. l.], v. 35, n. 12, p. 719-724, 1972.
- BIANCONI, Irene *et al.* Current Uses and Future Perspectives of Genomic Technologies in Clinical Microbiology. **Antibiotics**, [Basel, Switzerland], v. 12, n. 11, p. 1580, 2023.
- BINATI, Renato L. *et al.* Contribution of non-*Saccharomyces* yeasts to wine volatile and sensory diversity: A study on *Lachancea thermotolerans*, *Metschnikowia* spp. and *Starmerella bacillaris* strains isolated in Italy. **International Journal of Food Microbiology**, [s. l.], v. 318, p. 108470, 2020.
- BISSON, Linda F. *et al.* The present and future of the international wine industry. **Nature**, [s. l.], v. 418, n. 6898, p. 696-699, 2002.
- BOEKHOUT, Teun *et al.* Trends in yeast diversity discovery. **Fungal Diversity**, [s. l.],

v. 114, n. 1, p. 491-537, 2022.

BORODOVSKY, Mark *et al.* Eukaryotic gene prediction using GeneMark.hmm-E and GeneMark-ES. **Current Protocols In Bioinformatics**, [s. l.], v. 4, n. 6, p. 1-4, 2011.

BROCKHURST, Michael A. *et al.* The Ecology and Evolution of Pangenomes. **Current Biology**, [s. l.], v. 29, n. 20, p. R1094-R1103, 2019.

BUCHHOLZ, Klaus *et al.* The roots—a short history of industrial microbiology and biotechnology. **Applied Microbiology and Biotechnology**, [s. l.], v. 97, n. 9, p. 3747-3762, 2013.

CADEZ, Neza *et al.* *Hanseniaspora smithiae* sp. nov., a Novel Apiculate Yeast Species From Patagonian Forests That Lacks the Typical Genomic Domestication Signatures for Fermentative Environments. *Frontiers in Microbiology*, [s. l.], v. 12, 2021.

CADEZ, Neza *et al.* Genomic content of a novel yeast species *Hanseniaspora gamundiae* sp. nov. from fungal stromata (Cytaria) associated with a unique fermented beverage in Andean Patagonia, Argentina. **PloS One**, [s. l.], v. 14, n. 1, p. e0210792, 2019.

CADEZ, Neza *et al.* Molecular identification and genetic diversity within species of the genera *Hanseniaspora* and *Kloeckera*. **FEMS Yeast Research**, [s. l.], v. 14, n. 4, p. 279-289, 2002.

CAPOZZI, Vittorio *et al.* Microbial terroir and food innovation: The case of yeast biodiversity in wine. **Microbiological Research**, [s. l.], v. 181, p. 75-83, 2015.

CARRAU, Francisco *et al.* Biology and physiology of *Hanseniaspora vineae*: metabolic diversity and increase flavour complexity for food fermentation. **FEMS Yeast Research**, [s. l.], v. 23, p. foad010, 2023.

CARRAU, Francisco *et al.* Yeast diversity and native vigor for flavor phenotypes. **Trends in biotechnology**, [s. l.], v. 33, n. 3, p. 148-154, 2015.

CONESA, Ana *et al.* Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. **Bioinformatics**, [Oxford, England], v. 21, n. 18, p. 3674-3676, 2005.

EDGAR, Robert C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. **Nucleic Acids Research**, [s. l.], v. 32, n. 5, p. 1792-1797, 2004.

EL-MANSI, Mansi (ed.) *et al.* Fermentation Microbiology and Biotechnology, 4th ed. **CRC Press**, Boca Raton, 2019.

EMMS, David M. OrthoFinder: phylogenetic orthology inference for comparative genomics. **Genome Biology**, [s. l.], v. 20, n. 1, p. 238, 2019.

ESTEVE-ZARZOSO, Braulio *et al.* Molecular characterisation of *Hanseniaspora* species. **Antonie van Leeuwenhoek**, [s. l.], v. 80, n. 1, p. 85-92, 2001.

GAMERO, Amparo *et al.* Aromatic Potential of Diverse Non-Conventional Yeast Species for Winemaking and Brewing. **Fermentation**, [s. l.], v. 6, n. 2, p. 50, 2020.

GIORELLO, Facundo *et al.* Genomic and Transcriptomic Basis of *Hanseniaspora vineae*'s Impact on Flavor Diversity and Wine Quality. **Applied and Environmental Microbiology**, [s. l.], v. 85, n. 1, p. e01959-18, 2018.

GOLDMAN, Aaron David *et al.* What Is a Genome? **PLoS Genetics**, [s. l.], v. 12, n. 7, p. e1006181, 2016.

GONG, Ying *et al.* A review of the pangenome: how it affects our understanding of genomic variation, selection and breeding in domestic animals?. **Journal of Animal Science and Biotechnology**, [s. l.], v. 14, n. 1, p. 73, 2023.

GRANCHI, Lisa *et al.* Oenological properties of *Hanseniaspora osmophila* and *Kloeckera corticis* from wines produced by spontaneous fermentations of normal and dried grapes. **FEMS Yeast Research**, [s. l.], v. 2, n. 3, p. 403-407, 2002.

GUARAGNELLA, Nicoletta *et al.* Genome Sequencing and Comparative Analysis of Three *Hanseniaspora uvarum* Indigenous Wine Strains Reveal Remarkable Biotechnological Potential. **Frontiers in Microbiology**, [s. l.], v. 10, p. 3133, 2020.

GUINDON, Stéphane *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. **Systematic Biology**, [s. l.], v. 59, n. 3, p. 307-321, 2010.

GUREVICH, Alexey *et al.* QUASt: quality assessment tool for genome assemblies. **Bioinformatics**, [Oxford, England], v. 29, n. 8, p. 1072-1075, 2013.

HAYNES, Winston. Benjamini–Hochberg Method. In: Dubitzky, W., Wolkenhauer, O., Cho, KH., Yokota, H. (eds.). Encyclopedia of Systems Biology, **Springer**, New York, 2013.

HENDERSON, Clark M. *et al.* Fermentation temperature modulates

phosphatidylethanolamine and phosphatidylinositol levels in the cell membrane of *Saccharomyces cerevisiae*. **Applied and Environmental Microbiology**, [s. l.], v. 79, n. 17, p. 5345-5356, 2013.

HICKEY, Glenn *et al.* Pangenome graph construction from genome alignments with Minigraph-Cactus. **Nature Biotechnology**, 2023.

JONES, Philip *et al.* InterProScan 5: genome-scale protein function classification. **Bioinformatics**, [s. l.], v. 30, n. 9, p. 1236-1240, 2014.

KNOP, Michael. Yeast cell morphology and sexual reproduction – A short overview and some considerations. **Comptes Rendus Biologies**, [s. l.], v. 334, n. 8, p. 599-606, 2011.

KURTZMAN, Cletus P. (ed.) *et al.* The yeasts, a taxonomic study, v. 1, 5th ed. **Elsevier**, New York, 2011.

LI, Hua *et al.* The worlds of wine: Old, new and ancient. **Wine Economics and Policy**, [s. l.], v. 7, n. 2, p. 178-182, 2018.

LLEIXÀ, Jessica *et al.* Analysis of the NCR Mechanisms in *Hanseniaspora vineae* and *Saccharomyces cerevisiae* During Winemaking. **Frontiers in Genetic**, [s. l.], v. 9, p. 747, 2019.

LÓPEZ-MALO, María *et al.* Evolutionary engineering of a wine yeast strain revealed a key role of inositol and mannoprotein metabolism during low-temperature fermentation. **BMC Genomics**, [s. l.], v. 16, p. 537, 2015.

LUO, Jie *et al.* Reassessment of *Annamocarya sinensis* (*Carya sinensis*) Taxonomy through Concatenation and Coalescence Phylogenetic Analysis. **Plants**, [Basel, Switzerland], v. 11, n. 1, p. 52, 2021.

MAICAS, Sergi *et al.* The Life of *Saccharomyces* and Non-*Saccharomyces* Yeasts in Drinking Wine. **Microorganisms**, [s. l.], v. 11, n. 5, p. 1178, 2023.

MAICAS, Sergi. The Role of Yeasts in Fermentation Processes. **Microorganisms**, [s. l.], v. 8, 2020.

MANE, Sarika S. *et al.* Diversity of Natural Yeast Flora of Grapes and Its Significance in Wine Making. In: *Yeast Diversity in Human Welfare*. **Springer**, Singapore, 2017.

MARTIN, Valentina *et al.* Oenological Impact of the *Hanseniaspora/Kloeckera* Yeast Genus on Wines—A Review. **Fermentation**, [s. l.], v. 4, n. 3, p. 76, 2018.



MASSERA, Ariel *et al.* Effect of low temperature fermentation on the yeast-derived volatile aroma composition and sensory profile in Merlot wines. **LWT**, [s. l.], v.142, p. 111069, 2021.

MATURANO, Y. Paola *et al.* Yeast population dynamics during prefermentative cold soak of Cabernet Sauvignon and Malbec wines. **International Journal of Food Microbiology**, [s. l.], v. 199, p. 23-32, 2015.

MCGOVERN, Patrick E. *et al.* Ancient Wine: The Search for the Origins of Viniculture. **Princeton University Press**, 2003.

MORATA, Antonio *et al.* Non-Saccharomyces as Biotools to Control the Production of Off-Flavors in Wines. **Molecules**, [Basel, Switzerland], v. 26, n. 15, p. 4571, 2021.

MORTIMER, Robert K. Evolution and variation of the yeast (Saccharomyces) genome. **Genome Research**, [s. l.], v. 10, n. 4, p. 403-409, 2000.

NANDY, Subir Kumar *et al.* A review on sustainable yeast biotechnological processes and applications. **Microbiological Research**, [s. l.], v. 207, p. 83-90, 2018.

OHANA-LEVI, Noa *et al.* Long-Term Trends of Global Wine Market. **Agriculture**, [s. l.], v. 13, n. 1, p. 224, 2023.

PARAPOULI, Maria *et al.* Saccharomyces cerevisiae and its industrial applications. **AIMS Microbiology**, [s. l.], v. 6, n. 1, p. 1-31, 2020.

PARRA, Genis *et al.* GeneID in Drosophila. **Genome Research**, [s. l.], v. 10, n. 4, p. 511-515, 2000.

PERTEA, Geo *et al.* GFF Utilities: GffRead and GffCompare. **F1000Research**, [s. l.], v. 9, 2020.

PERVEZ, Muhammad Tariq *et al.* A Comprehensive Review of Performance of Next-Generation Sequencing Platforms. **BioMed Research International**, [s. l.], v. 2022, p. 3457806, 2022.

PIETRAFESA, Angela *et al.* Saccharomyces cerevisiae and Hanseniaspora uvarum mixed starter cultures: Influence of microbial/physical interactions on wine characteristics. **Yeast**, [Chichester, England], v. 37, n. 11, p. 609-621, 2020.

RASPOR, Peter *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**, [s. l.], v. 109, n. 1, p. 97-102, 2006.

RIBEIRO, Ricardo A. *et al.* The cell wall and the response and tolerance to stresses of biotechnological relevance in yeasts. **Frontiers in Microbiology**, [s. l.], v. 13, 2022.

ROMANO, Patrizia *et al.* Higher alcohol and acetic acid production by apiculate wine yeasts. **Journal of Applied Bacteriology**, [s. l.], v. 73, n. 2, p. 126-130, 1992.

ROSS, Reynolds Paul *et al.* Preservation and fermentation: past, present and future. **International Journal of Food Microbiology**, [s. l.], v. 79, n. 2, p. 3-16, 2002.

ROUMPEKA, Despoina D. *et al.* A Review of Bioinformatics Tools for Bio-Propecting from Metagenomic Sequence Data. **Frontiers in Genetics**, [s. l.], v. 8, n. 23, 2017.

SANGRADOR-VEGAS, Amaia *et al.* GO annotation in InterPro: why stability does not indicate accuracy in a sea of changing annotations. **Database**, [s. l.], v. 2016, p. baw027, 2016.

SCHLOSS, Patrick D. *et al.* Biotechnological prospects from metagenomics. **Current Opinion In Biotechnology**, [s. l.], v. 14, n. 3, p. 303-310, 2003.

SIMÃO, Felipe A. *et al.* BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. **Bioinformatics**, [Oxford, England], v. 31, n. 19, p. 3210-3212, 2015.

SIPICZKI, Matthias. Overwintering of Vineyard Yeasts: Survival of Interacting Yeast Communities in Grapes Mummified on Vines. **Frontiers in Microbiology**, [s. l.], v. 7, p. 212, 2016.

SPENCER, John F.T.; SPENCER, Dorothy M. Yeasts in Natural and Artificial Habitats. **Springer-Verlag**, Berlin, 1997.

STANKE, Mario *et al.* AUGUSTUS: a web server for gene finding in eukaryotes. **Nucleic Acids Research**, [s. l.], v. 32, p. W309-12, 2004.

STEENWYK, Jacob L. *et al.* Extensive loss of cell-cycle and DNA repair genes in an ancient lineage of bipolar budding yeasts. **PLoS Biology**, [s. l.], v. 17, n. 5, p. e3000255, 2019.

STERNES, Peter R. *et al.* Genome Sequences of Three Species of *Hanseniaspora* Isolated from Spontaneous Wine Fermentations. **Genome Announcements**, [s. l.], v. 4, n. 6, p. e01287-16, 2016.

THANGADURAI, Devarajan (ed.) *et al.* Genomics and Proteomics: Principles, Technologies, and Applications, 1th ed. Apple Academic Press, New York, 2015.

TONDINI, Federico *et al.* Early adaptation strategies of *Saccharomyces cerevisiae* and *Torulaspora delbrueckii* to co-inoculation in high sugar grape must-like media. **Food Microbiology**, [s. l.], v. 90, 2020.

TULLIO, Vivian. Yeast Genomics and Its Applications in Biotechnological Processes: What Is Our Present and Near Future?. **Journal of Fungi**, [Basel, Switzerland], v. 8, n. 7, p. 752, 2022.

VADKERTIOVÁ, Renáta *et al.* Yeasts and yeast-like organisms associated with fruits and blossoms of different fruit trees. **Canadian Journal of Microbiology**, [s. l.], v. 58, n. 12, p. 1344-1352, 2012.

VALERA, María José *et al.* Wine Aroma Characterization of the Two Main Fermentation Yeast Species of the Apiculate Genus *Hanseniaspora*. **Fermentation**, [s. l.], v. 7, n. 3, p. 162, 2021.

VAN WYK, Niël *et al.* Exploring future applications of the apiculate yeast *Hanseniaspora*. **Critical Reviews in Biotechnology**, [s. l.], v. 44, n. 1, p. 100-119, 2024.

VARELA, Cristian *et al.* Identification of flocculant wine yeast strains with improved filtration-related phenotypes through application of high-throughput sedimentation rate assays. **Scientific Reports**, [s. l.], v. 10, n. 1, p. 2738, 2020.

VARELA, Cristian *et al.* Yeasts found in vineyards and wineries. **Yeast**, [Chichester, England], v. 34, n. 3, p. 111-128, 2017.

WATANABE, Daisuke *et al.* Adaptation of yeast *Saccharomyces cerevisiae* to grape-skin environment. **Scientific Reports**, [s. l.], v. 13, n. 1, p. 9279, 2023.

WINDHOLTZ, Sara *et al.* Non-*Saccharomyces* yeasts as bioprotection in the composition of red wine and in the reduction of sulfur dioxide. **LWT**, [s. l.], v. 149, p. 111781, 2021.

WU, Tianzhi *et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. **The Innovation**, [s. l.], v. 2, n. 3, p. 100141, 2021.

ZERIKLY, Malek *et al.* Strategies for the discovery of new natural products by genome mining. **ChemBiochem : a European Journal of Chemical Biology**, [s. l.],

v. 10, n. 4, p. 625-633, 2009.

ZHANG, Bo-Qin *et al.* Use of Indigenous *Hanseniaspora vineae* and *Metschnikowia pulcherrima* Co-fermentation With *Saccharomyces cerevisiae* to Improve the Aroma Diversity of Vidal Blanc Icewine. **Frontiers in Microbiology**, [s. l.], v. 9, 2018.

## ANEXOS

**Tabela S1:** Genomas utilizados neste trabalho. As primeiras duas colunas, Entrada NCBI e Organismo, apresentam a entrada do genoma em NCBI e o organismo ao qual pertence, respectivamente. A coluna Grupo de evolução indica a qual dos grupos, o genoma foi classificado neste trabalho, com base no relacionamento das espécies dentro da árvore filogenética do gênero (**Figura 3**). As colunas Genes anotados e Tamanho do genoma (bp), apresentam o número de transcritos distintos preditos para o genoma, neste trabalho, e o tamanho do genoma em pares de bases.

Entrada NCBI	Organismo	Grupo de evolução	Genes anotados	Tamanho do genoma (bp)
GCA_019649515.1	<i>Hanseniaspora mollemarum</i>	FEL	6963	8962907
GCA_004919785.1	<i>Hanseniaspora jakobsenii</i>	FEL	10323	14051380
GCA_003708335.1	<i>Hanseniaspora pseudoguilliermondii</i>	FEL	7973	9077954
GCA_009805765.1	<i>Hanseniaspora uvarum</i>	FEL	7042	8745848
GCA_030370665.1	<i>Hanseniaspora meyeri</i>	FEL	9486	8750633
GCA_030567315.1	<i>Hanseniaspora uvarum</i>	FEL	7139	8761644
GCA_004919765.1	<i>Hanseniaspora lachancei</i>	FEL	8504	9948763
GCA_030573515.1	<i>Hanseniaspora jakobsenii</i>	FEL	7757	11018984
GCA_003708365.1	<i>Hanseniaspora singularis</i> (nom. inval.)	FEL	7272	9236024
GCA_004919925.1	<i>Hanseniaspora uvarum</i>	FEL	7119	8801344
GCA_001747055.1	<i>Hanseniaspora uvarum</i>	FEL	9263	8807786
GCA_030569975.1	<i>Hanseniaspora lindneri</i>	FEL	7821	10913158
GCA_030572935.1	<i>Hanseniaspora mollemarum</i>	FEL	6953	8960409
GCA_030573495.1	<i>Hanseniaspora nectarophila</i>	FEL	8254	8660224
GCA_003708765.1	<i>Hanseniaspora hatyaiensis</i> (nom. inval.)	FEL	8477	10009638
GCA_030573455.1	<i>Hanseniaspora pseudoguilliermondii</i>	FEL	7885	8795275
GCA_002894425.1	<i>Hanseniaspora uvarum</i>	FEL	9450	8827137
GCA_030567255.1	<i>Hanseniaspora guilliermondii</i>	FEL	7026	9061334
GCA_019649525.1	<i>Hanseniaspora lindneri</i>	FEL	7702	10792737

Entrada NCBI	Organismo	Grupo de evolução	Genes anotados	Tamanho do genoma (bp)
GCA_030573555.1	<i>Hanseniaspora hatyaiensis</i> (nom. inval.)	FEL	8287	9643601
GCA_004919825.1	<i>Hanseniaspora meyeri</i>	FEL	10901	10244664
GCA_030567295.1	<i>Hanseniaspora meyeri</i>	FEL	9330	8839385
GCA_030567335.1	<i>Hanseniaspora opuntiae</i>	FEL	8391	8896164
GCA_004919795.1	<i>Hanseniaspora</i> sp. CRUB 1602	FEL	7152	9548207
GCA_030573475.1	<i>Hanseniaspora thailandica</i> (nom. inval.)	FEL	7805	8992097
GCA_030567275.1	<i>Hanseniaspora lachancei</i>	FEL	8124	8900485
GCA_004919915.1	<i>Hanseniaspora thailandica</i> (nom. inval.)	FEL	8217	9976088
GCA_003243715.1	<i>Hanseniaspora uvarum</i>	FEL	9974	8966449
GCA_003706385.1	<i>Hanseniaspora clermontiae</i>	FEL	9206	9003383
GCA_001749795.1	<i>Hanseniaspora opuntiae</i>	FEL	10958	8831957
GCA_004919845.1	<i>Hanseniaspora nectarophila</i>	FEL	9198	9955993
GCA_009805785.1	<i>Hanseniaspora uvarum</i>	FEL	7081	8751237
GCA_030180015.1	<i>Hanseniaspora uvarum</i>	FEL	7126	8887533
GCA_004919775.1	<i>Hanseniaspora guilliermondii</i>	FEL	7029	9135640
GCA_937859215.1	<i>Hanseniaspora opuntiae</i>	FEL	7537	7685595
GCA_001664025.1	<i>Hanseniaspora valbyensis</i> NRRL Y-1626	FEL	7838	11464036
GCA_030567235.1	<i>Hanseniaspora clermontiae</i>	FEL	8994	8711982
GCA_937857875.1	<i>Hanseniaspora uvarum</i>	FEL	6735	8221548
GCA_949640725.1	<i>Hanseniaspora uvarum</i>	FEL	7240	8904028
GCA_030573575.1	<i>Hanseniaspora</i> sp. CRUB 1602	FEL	7142	9447694
GCA_900119595.1	<i>Hanseniaspora guilliermondii</i>	FEL	6894	9037850
GCA_030567375.1	<i>Hanseniaspora valbyensis</i>	FEL	7312	9528379
GCA_000775265.3	<i>Hanseniaspora uvarum</i>	FEL	7078	8085558
GCA_009805715.1	<i>Hanseniaspora uvarum</i>	FEL	7209	9073990
GCA_030565715.1	<i>Hanseniaspora singularis</i> (nom. inval.)	FEL	7152	8883820
GCA_000968475.1	<i>Hanseniaspora uvarum</i> DSM 2768	FEL	7881	9503661
GCA_018105805.1	<i>Hanseniaspora uvarum</i>	FEL	8433	8650206
GCA_002141495.2	<i>Hanseniaspora vineae</i>	SEL	9850	11373809
GCA_000585475.3	<i>Hanseniaspora vineae</i>	SEL	9868	11338948
GCA_030580215.1	<i>Hanseniaspora vineae</i>	SEL	10538	11887022

Entrada NCBI	Organismo	Grupo de evolução	Genes anotados	Tamanho do genoma (bp)
GCA_004919945.1	<i>Hanseniaspora vineae</i>	SEL	10015	11103790
GCA_001747045.1	<i>Hanseniaspora osmophila</i>	SEL	13235	11458415
GCA_937836975.1	<i>Hanseniaspora osmophila</i>	SEL	8869	9686042
GCA_003707715.2	<i>Hanseniaspora osmophila</i>	SEL	10021	11218515
GCA_004919885.1	<i>Hanseniaspora occidentalis</i> var. <i>citrica</i>	SEL	10522	12166029
GCA_030567355.1	<i>Hanseniaspora occidentalis</i> var. <i>occidentalis</i>	SEL	10113	11515705
GCA_004919895.1	<i>Hanseniaspora occidentalis</i> var. <i>occidentalis</i>	SEL	9928	11273150
GCA_030573535.1	<i>Hanseniaspora occidentalis</i> var. <i>citrica</i>	SEL	10073	11340894
GCA_030580555.1	<i>Hanseniaspora taiwanica</i> (nom. <i>inval.</i> )	SEL	8853	10019776
GCA_003020785.1	<i>Hanseniaspora gamundiae</i>	SEL	11016	10183153

**Tabela S2:** Parâmetros utilizados pelos anotadores. Para todos os anotadores as anotações foram geradas em arquivos formato \*.gff3.

Programa	Comando	Parâmetros
Augustus	augustus	--strand=both --genemodel=intronless --species=saccharomyces --gff3=on
Genemark-ES	gmes_petap.pl	--ES --fungus --gff3=on
Geneid	geneid	-P sjaponicus.param_Oct_12_2006 -s -3