

Universidad Autónoma de Nuevo León  
Facultad de Ciencias Biológicas  
Instituto de Biotecnología



**ANÁLISIS DEL SISTEMA DE TRANSPORTE DE CARBOHIDRATOS EN  
DOS LEVADURAS *Saccharomyces pastorianus***

Por

**MC. César Ignacio Hernández Vásquez**

**Como requisito parcial para obtener el grado de**

**Doctor en Ciencias con Orientación en Biotecnología**

# ANÁLISIS DEL SISTEMA DE TRANSPORTE DE CARBOHIDRATOS EN DOS LEVADURAS *Saccharomyces pastorianus*



## Comité de Tesis

A handwritten signature in blue ink, appearing to be 'BP'.

---

**Dr. Benito Peñeyra Alférez**  
Director

A handwritten signature in blue ink, appearing to be 'JHG'.

---

**Dr. Jorge Hugo García García**  
Secretario

A handwritten signature in blue ink, appearing to be 'LJG'.

---

**Dr. Luis Jesús Galán Wong**  
Vocal

A handwritten signature in blue ink, appearing to be 'MES'.

---

**Dra. Myriam Elías Santos**  
Vocal

A handwritten signature in blue ink, appearing to be 'CEL'.

---

**Dr. Carlos Eduardo Hernández Luna**  
Vocal

A handwritten signature in blue ink, appearing to be 'KAN'.

---

**Dra. Karushka Arévalo Niño**  
Subdirectora de posgrado



DIRECCIÓN  
DE POSGRADO

Universidad Autónoma de Nuevo León  
Facultad de Ciencias Biológicas  
Instituto de Biotecnología



**ANÁLISIS DEL SISTEMA DE TRANSPORTE DE CARBOHIDRATOS EN  
DOS LEVADURAS *Saccharomyces pastorianus***

**Dirección de tesis**

---

**Dr. Benito Pereyra Alferez**  
Director

---

**Dr. Luis Cástulo Damas Buenrostro**  
Asesor Externo

Universidad Autónoma de Nuevo León  
Facultad de Ciencias Biológicas  
Instituto de Biotecnología



**ANÁLISIS DEL SISTEMA DE TRANSPORTE DE CARBOHIDRATOS EN  
DOS LEVADURAS *Saccharomyces pastorianus***

El presente trabajo se llevó a cabo en el Laboratorio 4 (L4) del Instituto de Biotecnología, Facultad de Ciencias Biológicas de la Universidad Autónoma de Nuevo León y en la Cervecería Cuauhtémoc Moctezuma bajo la dirección del Dr. Benito Pereyra Alférez y del Dr. Luis C. Damas Buenrostro, respectivamente

## AGRADECIMIENTOS

A la **Vida** por darme la oportunidad de existir, lo que me lleva al pensar y contribuir en la ciencia.

Al **CONACYT**, por brindar la beca para sustentar mis estudios de doctorado.

Al **Dr. Benito Pereyra Alférez**, por todo su apoyo. Muchas gracias por esas cultivadoras charlas que además de instructivas fueron divertidas y muy trascendentales.

Al **Dr. Jorge Hugo García García** por la excelente forma de compartir y distribuir el conocimiento. Muchas gracias por esos días de profundas pláticas, nunca lo olvidaré.

Al **Dr. Luis J. Galán Wong** por el apoyo en el proceso de la investigación.

Al **Dr. Myriam Elías Santos** por su aporte y disponibilidad valiosa en esta investigación.

A la **Dr. Carlos Eduardo Hernández Luna** por su valiosa tutoría y aportes críticos en el diseño, desarrollo y conclusión de este trabajo de investigación.

A la **Dr. Luis C. Damas Buenrostro** por sus valiosos consejos tutoriales que abonaron sustancialmente en el desarrollo y conclusión de este trabajo de investigación.

A la **Dra. Jessica Giselle Herrera Gamboa y a todo el HUB-Heineken México** por su valiosa, amable y grandiosa disponibilidad. A sus consejos para el desarrollo de experimentos que ayudaron a concluir exitosamente esta investigación.

A la **Universidad Autónoma de Nuevo León** por brindarme la maravillosa oportunidad de formar parte de su excelente ambiente académico.

A **Cervecería Cuauhtémoc Moctezuma**, por todo el sustento necesario en el diseño, desarrollo y conclusión de este proyecto.

A mi esposa **Lorena** e hija **Luna**. Por brindarme una estabilidad emocional necesaria para enfrentar el camino de la investigación. Gracias por todo el apoyo incondicional. Son lo mejor que me ha pasado.

A mis **padres, Ignacio Hernández y Escolástica Vásquez** por todo el apoyo incondicional, amor y regaños a lo largo de mi vida. Sin ustedes prácticamente sería un ente viviente sin alma ni sueños. Gracias por ser uno de los motores que me impulsa a seguir.

A mis **hermanos, Josué, Larissa y Eneida** por formar parte de mi vida y compartir tantas experiencias juntos. Los amo.

A mis amigos, **Óscar, Mario y Cecy**, del L4 (Hugo, Adrián, Eduardo, Carlos, Astrid, Dani, Óscar, y los que faltaron por nombrar). Gracias por formar parte en este bonito camino universitario y académico.

A mis **hermanos de otra madre**, Eduardo, Yair, Martín. Gracias, hermanos por todas las aventuras y experiencias.

## DEDICATORIA

Para Lorena y Luna,  
Fragua del equilibrio.

Para mis padres, Ignacio y Escolástica,  
Hipérboles del infinito.

*“Todo nuestro conocimiento arranca en el sentido, pasa al entendimiento y termina en la razón”*

Immanuel Kant.

# ÍNDICE

ÍNDICE DE TABLAS .....	vii
ÍNDICE DE FIGURAS.....	viii
1. RESUMEN.....	1
2. ABSTRACT.....	2
3. INTRODUCCIÓN.....	3
4. ANTECEDENTES .....	6
4.1 El proceso de la fermentación cervecera; composición de un mosto estándar y el primer acercamiento al problema de los azúcares residuales.....	6
4.2 Los principales exponentes microbianos en la elaboración de cerveza (levaduras tipo ale y lager) .....	8
4.3 Asimilación de carbohidratos y sistemas de transporte en levaduras cerveceras 15	
4.4 El centro de operaciones del transporte de alfa glucósidos; los <i>loci</i> MAL ..	17
4.5 Principales aproximaciones al problema de azúcares residuales .....	20
4.6 La regulación entre reinos mediada por RNA y sus principales mecanismos .....	24
4.7 Principales investigaciones en el tema de la regulación entre reinos.....	27
4.8 La relación bidireccional entre la interacción planta-hongo fitopatógeno ...	28
5. JUSTIFICACIÓN .....	31
6. HIPÓTESIS .....	32
6.1 Hipótesis relacionada con la toma diferencial de maltosa y maltotriosa en las levaduras SpIB1 y SpIB2 32	
6.2 Hipótesis relacionadas con el efecto en la expresión genética de las levaduras de estudio debido a la presencia de RNA no codificante presente en el mosto cervecero.....	32
7. OBJETIVOS DEL TRABAJO .....	33
7.1 Objetivo general relacionado con la hipótesis de la sección 6.1 .....	33
7.2 Objetivos específicos relacionados con la hipótesis de la sección 6.1.....	33
7.3 Objetivo general relacionado con la hipótesis de la sección 6.2 .....	33
7.4 Objetivos específicos relacionados con la hipótesis de la sección 6.2.....	34
8. CAPÍTULO 1 Patrones de expresión de los genes Mal y su asociación con la tasa de transporte de maltosa y maltotriosa de dos levaduras <i>Saccharomyces pastorianus</i> .....	35
I. Resumen.....	36
II. Introducción .....	37

<b>III.</b>	<b>Materiales y métodos .....</b>	<b>39</b>
a)	<b>Cepas .....</b>	<b>39</b>
b)	<b>Condiciones y medios de cultivo .....</b>	<b>39</b>
c)	<b>Inspección y comparaciones por pares de permeasas presentes en los datos genómicos y transcriptómicos de las levaduras estudiadas .....</b>	<b>40</b>
c)	<b>Análisis de la estructura de la permeasa .....</b>	<b>40</b>
d)	<b>Análisis de la región reguladora de los genes Mal.....</b>	<b>41</b>
e)	<b>Valores de FPKM de los genes mal encontrados en cada cepa .....</b>	<b>41</b>
f)	<b>Análisis de expresión global.....</b>	<b>42</b>
g)	<b>Prueba de tasa de transporte celular .....</b>	<b>43</b>
h)	<b>Extracción de ADN y ARN .....</b>	<b>43</b>
i)	<b>Detección molecular de permeasas en cada cepa. ....</b>	<b>44</b>
j)	<b>Análisis del consumo de mosto y alfa glucósidos.....</b>	<b>45</b>
<b>IV</b>	<b>Resultados.....</b>	<b>45</b>
	<b>Abundancia de las permeasas en las levaduras estudiadas .....</b>	<b>45</b>
	<b>Análisis estructural, niveles de expresión y distribución filogenética de las permeasas.....</b>	<b>46</b>
	<b>Análisis de expresión global de genes .....</b>	<b>51</b>
	<b>Análisis bioquímico de la tasa de transporte y detección molecular de los transportadores de maltosa y maltotriosa .....</b>	<b>5</b>
	<b>4</b>	
<b>V</b>	<b>Discusión.....</b>	<b>56</b>
	<b>La variabilidad estructural explica la variación fenotípica en el transporte de alfa glucósidos. ....</b>	<b>5</b>
	<b>6</b>	
	<b>Los patrones de expresión y la identidad de los transportadores explican eficientemente la variación fenotípica del transporte de maltosa y maltotriosa en las cepas analizadas. ....</b>	<b>5</b>
	<b>8</b>	
	<b>En las mismas condiciones de fermentación, las levaduras exhiben diferencias en la tasa de transporte de maltosa y maltotriosa. ....</b>	<b>60</b>
	<b>Evidencia física sugiere una regulación genética diferencial entre las levaduras analizadas.....</b>	<b>61</b>
<b>VI</b>	<b>Conclusión .....</b>	<b>62</b>
	<b>Las diferencias en la regulación genética juegan un papel significativo en la variabilidad observada en el transporte de maltosa y maltotriosa en ambas cepas de <i>S. pastorianus</i></b>	

.....	6
2	
<b>9. CAPÍTULO 2. Datos de RNA-seq sugieren un efecto en la expresión genética de dos levaduras lager causado por la presencia de RNA en el mosto cervecero</b>	<b>80</b>
<b>I Resumen</b> .....	<b>81</b>
<b>II Introducción</b> .....	<b>82</b>
<b>III Materiales y métodos</b> .....	<b>85</b>
a) Cepas .....	85
b) Diseño experimental.....	85
c) Manipulación del material biológico y extracción de ARN total .....	86
d) Secuenciación de ARN y procesamiento de datos.....	87
e) Estimación de abundancia de transcritos, expresión génica diferencial, ontología génica y análisis de vías metabólicas KEGG .....	88
<b>IV Resultados</b> .....	<b>89</b>
El ARN presente en el mosto de cerveza induce cambios en la expresión global en dos levaduras lager.....	89
<b>V Discusión</b> .....	<b>98</b>
<b>VI Conclusión</b> .....	<b>101</b>
<b>10. CAPÍTULO 3. Otras contribuciones producto de la colaboración científica</b>	<b>102</b>
Artículo de investigación 1 .....	103
Comunicación corta 1 .....	111
Comunicación corta 2.....	113
Artículo de investigación 2.....	115
Capítulo de libro .....	132
<b>11. CONCLUSIONES GENERALES</b> .....	<b>175</b>
<b>12. PERSPECTIVAS</b> .....	<b>176</b>
<b>13. BIBLIOGRAFÍA</b> .....	<b>177</b>

## ÍNDICE DE TABLAS

<b>Tabla</b>	<b>Descripción</b>	<b>Página</b>
<b>1</b>	Los diez principales países productores de cerveza. Nótese como en el año 2018 México reportó le mayor crecimiento en la producción cervecera	<b>6</b>
<b>2</b>	Condensado de las afinidad y velocidad máxima de distintos transportadores. Para una revisión de las referencias ir a Vidgren, (2010).	<b>20</b>
<b>3</b>	Localización y tamaño de las permeasas para maltosa y maltotriosa presentes en las levaduras de estudio.	<b>47</b>

## ÍNDICE DE FIGURAS

<b>Figura</b>	<b>Descripción</b>	<b>Página</b>
<b>1</b>	Principales pasos en el proceso de malteado y elaboración de cerveza. Tomado y modificado de Bamforth (2016)	<b>7</b>
<b>2</b>	Proceso de la toma de carbohidratos complejos por la levadura. Primero, transportadores como Mal11p (AGT1), Mal61p y MAL31p transportan maltotriosa y maltosa. Dentro de la célula una enzima maltasa hidroliza los enlaces alfa 1→4 de los azúcares complejos y los convierte en unidades de glucosa. A su vez la maltosa y maltotriosa activan el factor transcripcional que activa los genes necesarios para el transporte y metabolismo de los carbohidratos complejos. Tomado y modificado de Gallone y cols (2018).	<b>10</b>
<b>3</b>	Modelo hipotético sugerido para el surgimiento de las levaduras lager tipo I y II. Tomado de Monerawela y Bond, (2018)	<b>12</b>
<b>4</b>	A) Principales mecanismos de transporte en la célula de levadura. B) Orden de consumo de los principales carbohidratos del mosto cervecero; glucosa→Fructosa →Maltosa→Maltotriosa.	<b>15</b>
<b>5</b>	Estructura y organización de un locus MAL canónico. Se observan los puntos de regulación por el factor transcripcional MALx3 que se une a un promotor bidireccional para activa a MALx1 y MALx2. Además, se observan los puntos de represión por catabolito impartidos en la presencia de Glucosa. Tomado de Vidgren (2010).	<b>18</b>
<b>6</b>	Principales interacciones mediadas por RNA pequeños que soportan una comunicación entre reinos con evidencia experimental. Los signos de interrogación (?) representan poca o nula evidencia que respalde la comunicación mediada por ARN's pequeños. Tomado y modificado de Rabuma y cols (2022).	<b>24</b>
<b>7</b>	Principales mecanismos de la comunicación entre reinos que tiene lugar en la relación bidireccional de algunas plantas contra los hongos fitopatógenos. Tomado y modificado de Mahanty y cols (2023).	<b>28</b>
<b>8</b>	Secciones del alineamiento múltiple de proteínas transportadoras de maltosa y maltotriosa que muestran los aminoácidos cargados involucrados en el mecanismo de translocación. Solo se muestran los dominios transmembranales DTM-1, DTM-2 y DTM-11.	<b>49</b>
<b>9</b>	Niveles de expresión de los transportadores de maltosa y maltotriosa mostrados en FPKM. Los datos son representados como gráfico de caja conteniendo los	<b>50</b>

	valores FPKM, de tres experimentos independientes, para cada gen MAL encontrado en cada cepa. Nótese que en los genes A) AGT1, B) MAL31, D) MPH2 y F) MPH3 la cepa SpIB2 presenta mayor nivel de expresión que la cepa SpIB1. La cepa SpIB1 solo tiene mayor expresión en el gene F) MAL61 al quinto día de fermentación.	
10	Filogenia estimada con el algoritmo "Maximum likelihood" mediante un Bootstrap de 1,000 para los transportadores de maltosa y maltotriosa reportados en <i>S. cerevisiae</i> , <i>S. eubayanus</i> , <i>S. pastorianus</i> , SpIB1 y SpIB2. La escala de las barras indica el número de sustituciones aminoacídicas por sitio. Nótese como las proteínas MAL31 de la cepa SpIB2, marcadas en rojo, se agrupan con las permeases MTT1/MTY1 reportadas para <i>S. pastorianus</i> .	52
11	Gráfico tipo Volcano que expone los genes diferencialmente expresados en A) el segundo y B) quinto día de la fermentación para las cepas SpIB1 y SpIB2. Los genes al alta (up) y a la baja (down) están representados en azul y rojo, respectivamente (valor absoluto del log2 fold change >1 y false-discovery rate [FDR] < 0.05). Genes con significancia estadística, pero con bajo nivel de expresión (s_but_low) (valor absoluto del log2 fold change <1 and FDR < 0.05) son expuestos en gris y los genes sin significancia estadística (ns) en negro (FDR > 0.05).	53
12	Tasa de transporte celular determinada por la toma de pNP-glucosa y pNP-maltosa en un periodo de 7 minutos de reacción en las cepas A) SpIB1 y B) SpIB2. C) Concentración en %p/p de maltosa y maltotriosa en el segundo y quinto día de la fermentación para la cepa SpIB1. Los datos para la cepa SpIB2 no fueron detectables.	56
13	Detección molecular de los genes Mal presentes en las cepas SpIB1 y SpIB2. A) Detección genómica mediante PCR de los genes Mal ScMALx1, SeMALx1, ScAGT1, SeAGT1, MPHx, and MTT1/MTY1 indicando que cada levadura contiene todos los transportadores analizados en este estudio. B) Detección de transcritos por RT-PCR en el primer y segundo día de la fermentación de las levaduras estudiadas. Nótese que, en el primer día de la fermentación, la cepa SpIB2 mostró todos los transcritos Mal, mientras que la levadura SpIB1 sólo exhibió ScMalx1, ScAGT1 y MPHx.	58
14	Análisis de componentes principales (PCA) de las librerías de ARNasas de a) SpIB1 y b) SpIB2 en los diferentes días de recolección de muestras y tratamiento. En la leyenda de la figura, la primera letra indica c = muestras de control o t = muestras de tratamiento con ARNasas, seguida de f2 o f5, que significan el segundo y quinto día de fermentación, y por último las repeticiones del experimento.	91
15	Gráficos de volcán que representan los genes DE en el a) segundo y b) quinto día de fermentación para la cepa SpIB1 y, de manera similar, en la cepa SpIB2 se muestran los genes DE obtenidos en el c) segundo y d) quinto día de fermentación. Los datos se representan como los genes regulados al alza en azul (cambio de pliegue $\geq 2$ y tasa de falsos descubrimientos [FDR] < 0,05), los genes regulados a la baja en rojo (cambio de pliegue $\leq -2$ y FDR < 0,05), los genes con un cambio significativo, pero bajo (cambio de pliegue absoluto $\geq 2$ y FDR < 0,05) se representan en gris y los genes no significativos en negro (FDR > 0,05).	93
16	Análisis ontológico de los genes DE en el segundo día de fermentación para la cepa SpIB1. Los genes DE se representan como a) genes regulados a la baja y b) genes regulados al alza que respaldan los términos ontológicos con significación estadística (tasa de descubrimiento falso [FDR] < 0,05). Los	94

	términos ontológicos en los genes regulados a la baja significan que estos genes están regulados positivamente por el ARN presente en el mosto de cerveza y los términos ontológicos de los genes regulados al alza sugieren una posible regulación a la baja por el ARN presente en el mosto de cerveza. Todas las comparaciones se realizan entre fermentaciones tratadas con ARNasa frente a fermentaciones de control (tratamiento frente a control)	
<b>17</b>	Sobrerrepresentación de los términos KEGG observados en los datos transcriptómicos para el quinto día de fermentación de la cepa SpIB1. Los términos KEGG están enriquecidos con a) los genes regulados a la baja y b) los genes regulados al alza observados. Esto significa que los términos KEGG subregulados se ven afectados positivamente por la presencia de ARN en el mosto de cerveza y viceversa, los términos KEGG regulados al alza se ven regulados negativamente por la presencia de ARN en el mosto de cerveza (comparaciones del tratamiento frente al control).	<b>96</b>
<b>18</b>	Análisis ontológico de los genes DE en el segundo día de fermentación para la cepa SpIB2. Los genes DE se representan como a) genes regulados a la baja en el tratamiento y b) genes regulados al alza que respaldan los términos ontológicos con significación estadística (tasa de descubrimiento falso [FDR] < 0,05). El término de función molecular en a) genes regulados a la baja significa que estos genes están regulados positivamente por el ARN presente en el mosto de cerveza y los términos de proceso biológico de b) genes regulados al alza sugieren una posible regulación a la baja por el ARN presente en el mosto de cerveza (comparaciones tratamiento vs control).	<b>98</b>
<b>19</b>	Análisis de enriquecimiento de los datos transcriptómicos en el quinto día de fermentación en la cepa SpIB2. Los términos GO enriquecidos en el proceso biológico con los genes a) regulados a la baja y b) regulados al alza. Nótese que las proteínas reguladoras involucradas en la autofosforilación y fosforilación de proteínas están enriquecidas por los genes regulados a la baja, lo que sugiere que estos procesos biológicos son inducidos por la presencia de ARN en el mosto. Por el contrario, el ciclo celular mitótico se ve afectado negativamente por la presencia de este ARN en el mosto (Comparaciones del tratamiento frente al control).	<b>99</b>

## 1. RESUMEN

La elaboración de cerveza es un proceso biotecnológico más antiguo que aún enfrenta grandes desafíos, como el consumo total de azúcares presentes en los medios de fermentación. Además, el estudio científico de otros factores como la naturaleza biológica del medio de cultivo quedan limitados a aproximaciones reducidas. La cerveza estilo lager, una de las más importantes a nivel mundial, se elabora con levadura *Saccharomyces pastorianus* (Sp), que debe fermentar mostos con alto contenido de maltotriosa, pero su consumo representa un problema notable, especialmente entre las cepas Sp que pertenecen al grupo I. Factores como las condiciones de fermentación, la presencia de transportadores de maltotriosa, la variación del número de copias del transportador y las variaciones en la regulación genética contribuyen a este problema. Aunado a esto, el estudio de la regulación entre reinos mediada por RNA es un tema de investigación relativamente nuevo que sugiere que los ARN no codificantes (ARNnc) de un organismo, como las plantas, pueden regular la expresión génica en otros organismos, incluidos hongos, bacterias y parásitos. En este estudio, también probamos la hipótesis de que los ARN, potencialmente presentes en el mosto de cerveza originado a partir de cebada y/o lúpulo, influyen en la expresión génica en dos cepas de levadura lager industrial. Por un lado, se determinaron los factores que más afectan la fermentación en dos cepas de levadura Sp: SpIB1, con absorción limitada de maltotriosa, y SpIB2, conocida por su eficiente transporte de maltotriosa. Aquí, SpIB2 transportó significativamente más maltosa (28%) y maltotriosa (32%) en comparación con SpIB1. Además, SpIB2 expresó todos los transportadores MAL (*ScMALx1*, *SeMALx1*, *ScAGT1*, *SeAGT1*, *MTT1* y *MPHx*) el primer día de fermentación, mientras que SpIB1 solo exhibió los genes *ScMalx1*, *ScAGT1* y *MPH2/3*. Algunos transportadores SpIB2 tenían dominios transmembrana polimórficos (TMD) que se asemejaban a *MTT1*, acompañados de una mayor expresión de estos transportadores y sus genes reguladores positivos, como MAL63. Finalmente, después del tratamiento con ARNasa en experimentos independientes, observamos efectos importantes en la expresión génica en el segundo día de fermentación, un total de 671 genes DE (expresados diferencialmente) en SpIB1 y 536 en SpIB2. En el quinto día, observamos un efecto reducido con un total de 64 genes DE en SpIB1 y 29 en SpIB2. En ambas levaduras, observamos en el segundo día de fermentación, una inducción positiva del ciclo celular y la floculación por la presencia de ARN en el mosto de cerveza, junto con la posible represión del regulador transcripcional ARO80 para el crecimiento de pseudohifas. Al quinto día, el metabolismo de los aminoácidos se vio afectado negativamente por la presencia de ARN en el medio, lo que indica una posible reducción de la aptitud en la última etapa de la fermentación. Estos hallazgos sugieren que, además de los factores mencionados anteriormente, los reguladores positivos de los transportadores Mal contribuyen significativamente a la diversidad fenotípica en el consumo de maltosa y maltotriosa entre las cepas de levadura lager estudiadas. Y a su vez indican un posible efecto del ARN presente en el mosto sobre la expresión genética de dos levaduras industriales respaldando la hipótesis de la comunicación entre reinos mediada por ARNnc.

## 2. ABSTRACT

Brewing is one of the oldest biotechnological processes, yet it continues to face significant challenges, such as ensuring the complete consumption of sugars during fermentation. Additionally, scientific research into other factors, like the biological composition of the beer wort, remains limited in scope. Lager beer, one of the most popular beer types globally, is brewed using the yeast *Saccharomyces pastorianus* (Sp). This yeast must efficiently ferment wort with a high maltose and maltotriose content, but this process poses challenges, particularly for Sp strains in group I. Factors such as fermentation conditions, the presence of maltotriose transporters, variations in transporter copy numbers, and differences in genetic regulation all contribute to these difficulties. In this study, we explore the metabolic differences in maltose and maltotriose consumption between two lager yeast strains. We also investigate a relatively new area of research: RNA-mediated cross-kingdom regulation, which suggests that non-coding RNAs (ncRNAs) from one organism, such as plants, can regulate gene expression in other organisms, including fungi, bacteria, and parasites. Specifically, we test the hypothesis that RNAs present in beer wort, originating from barley and/or hops, influence gene expression in two industrial lager yeast strains. Our findings identify key factors affecting fermentation in the two Sp strains: SpIB1, which exhibits limited maltotriose uptake, and SpIB2, known for its efficient maltotriose transport. Notably, SpIB2 transported significantly more maltose (28%) and maltotriose (32%) compared to SpIB1. Furthermore, SpIB2 expressed all relevant maltose transporters (ScMALx1, SeMALx1, ScAGT1, SeAGT1, MTT1, and MPHx) on the first day of fermentation, whereas SpIB1 only expressed ScMALx1, ScAGT1, and MPH2/3. Additionally, some transporters in SpIB2, like MTT1, exhibited polymorphic transmembrane domains (TMDs) and were accompanied by higher expression levels of these transporters and their positive regulatory genes, such as MAL63. Moreover, in independent experiments involving RNase treatment, we observed significant effects on gene expression by the second day of fermentation, with 671 differentially expressed (DE) genes in SpIB1 and 536 in SpIB2. By the fifth day, these effects diminished, with only 64 DE genes in SpIB1 and 29 in SpIB2. On the second day, both yeasts exhibited positive induction of the cell cycle and flocculation due to the presence of RNA in the wort, while a potential repression of the transcriptional regulator ARO80, which is involved in pseudohyphae growth, was noted. By the fifth day, amino acid metabolism was negatively impacted by the presence of RNA, suggesting a possible reduction in fitness during the final fermentation stages. These findings suggest that, alongside other factors, positive regulators of maltose transporters play a critical role in the phenotypic diversity of maltose and maltotriose consumption in the studied lager yeast strains. They also indicate a potential influence of RNA present in wort on gene expression in these industrial yeast strains, supporting the hypothesis of ncRNA-mediated cross-kingdom communication.

### 3. INTRODUCCIÓN

La elaboración de cerveza es uno de los procesos biotecnológicos más antiguos, manufacturado hace aproximadamente 6000-8000 años (Hornsey. 2003) y hoy en día es un importante sector económico en la industria. En el 2018 la producción de cerveza resultó en 191.10 millones de kilolitros, un aumento del 0.6% respecto al año anterior.

En este contexto, de los 1.09 millones de kilolitros del aumento respecto al año previo, México representó el 90% de esta producción. En ese mismo año en México se reportó una producción de cerveza de 11.98 millones de kilolitros, un incremento del 8.8% respecto al año anterior. Con esto México se posicionó en el cuarto país productor de cerveza a nivel mundial, tan solo por debajo de China, Estados Unidos y Brasil (Kirin Beer University Report. 2019).

Dentro de este contexto económico, el tipo de cerveza más popular, a nivel mundial es, sin duda, la cerveza tipo lager, representando el 90% de participación en el mercado. Una de sus principales características es el uso de levaduras tipo *Saccharomyces pastorianus* (*Sp*). Estas levaduras son el resultado de un proceso de domesticación y la hibridación interespecífica entre *S. cerevisiae* (*Sc*) y *S. eubayanus* (*Se*) (Gallone *et al.*, 2018; Libkind *et al.*, 2011). Existen dos linajes genéticamente distintos dentro del grupo taxonómico de *Sp*, el grupo I (Saaz) y el grupo II (Frohberg).

La distinción se basa parcialmente en la utilización diferencial de los carbohidratos del mosto cervecero, resaltando la incapacidad de las cepas del grupo I en utilizar maltotriosa. También, otros aspectos como la baja producción de compuestos de importancia aromática y sensorial en cepas del grupo I (Dunn y Sherlock. 2008; Gibson *et al.*, 2013).

El mosto cervecero contiene, típicamente, ca. 50-60% y 15-20% de maltosa y la maltotriosa respectivamente (Stambuk *et al.*, 2006). Es por esto que uno de los pasos limitantes en una fermentación eficiente es la tasa de transporte

y metabolismo de estos carbohidratos dentro de la célula de levadura (Alves et al., 2007; Meneses y Jiranek, 2002).

El reporte de Magalhães y cols (2016), expone una serie de resultados interesantes al comparar los procesos de transporte entre cepas del grupo I y grupo II. Los resultados demostraron que dos cepas del grupo I presentan un fenotipo de transporte muy similar al observado en cepas del tipo II, evidenciando así la variabilidad de este fenotipo dentro de las cepas de *Sp* del grupo I. Adicionalmente, comprobaron la presencia de un transportador específico con la fermentación eficiente observada en las cepas del grupo I.

Aunado a los factores más evidentes que pueden afectar la calidad del producto cervecero final, existe un novedoso campo de investigación relacionado con la regulación entre reinos a través de RNA's no codificantes pequeños. Este fenómeno se caracteriza por el proceso en el que RNA, normalmente no codificante, proveniente de un organismo influye en la regulación genética de otro organismo que pertenece a otro reino (Mahanty *et al.*, 2023).

Este tipo de regulación ha sido observada entre plantas y sus respectivos patógenos, animales e incluso entre distintos microorganismos. El ARN que participa en este mecanismo puede ser ARN mensajero (ARNm), ARN pequeño interferente (siRNA), microARN (miRNA) o ARN no codificante de cadena larga (lncRNA), y puede intervenir en la expresión genética y otros procesos biológicos en los organismos receptores (Leitão & Enguita. 2016).

La exploración de este fenómeno en el contexto industrial cervecero tiene mucho sentido ya que la producción de cerveza se caracteriza por la interacción de tres organismos vivos, principalmente: la cebada, el lúpulo y la levadura. La interacción de estos organismos mediante dicho fenómeno, tendría más importancia en el contexto de las relaciones fitopatógenas de algunos hongos, ya que estos pueden tener afecciones en su expresión genética a través de RNA no codificante proveniente de la planta (Mahanty *et al.*, 2023).

Es claro que uno de los pasos trascendentales implicados en una fermentación eficiente es la toma y metabolismo de la maltotriosa, el cual es

altamente influenciado por la gran variabilidad fenotípica observada en cepas del grupo I. Por esta razón, decidimos investigar este fenómeno en cepas de *S. pastorianus* pertenecientes al grupo I y II.

Además, la exploración del campo relacionado con la regulación entre reinos a través de RNA brindará información en el comportamiento general relacionado con la fermentación eficiente. A diferencia de las investigaciones anteriores, en esta investigación se empleará información proveniente de la secuenciación del genoma y transcriptoma de las cepas de estudio para abordar este fenómeno de forma holística (De León-Medina *et al.*, 2016), es decir integrando la información genómica, fisiológica y bioquímica de las cepas de estudio.

#### 4. ANTECEDENTES

##### 4.1 El proceso de la fermentación cervecera; composición de un mosto estándar y el primer acercamiento al problema de los azúcares residuales

El proceso de elaboración de cerveza es una práctica antigua que se remota a 6000-8000 años de antigüedad (Hornsey. 2003). Desde las últimas décadas del siglo diecinueve el proceso de elaboración de cerveza avanzó hasta convertirse, hoy en día, en un proceso tecnológico controlado que participa activamente como un importante sector económico en México y el mundo (Kirin Beer University Report. 2019) (Tabla 1).

**Tabla 1.** Principales países productores de cerveza. Nótese como en el año 2018 México reportó le mayor crecimiento en la producción cervecera.

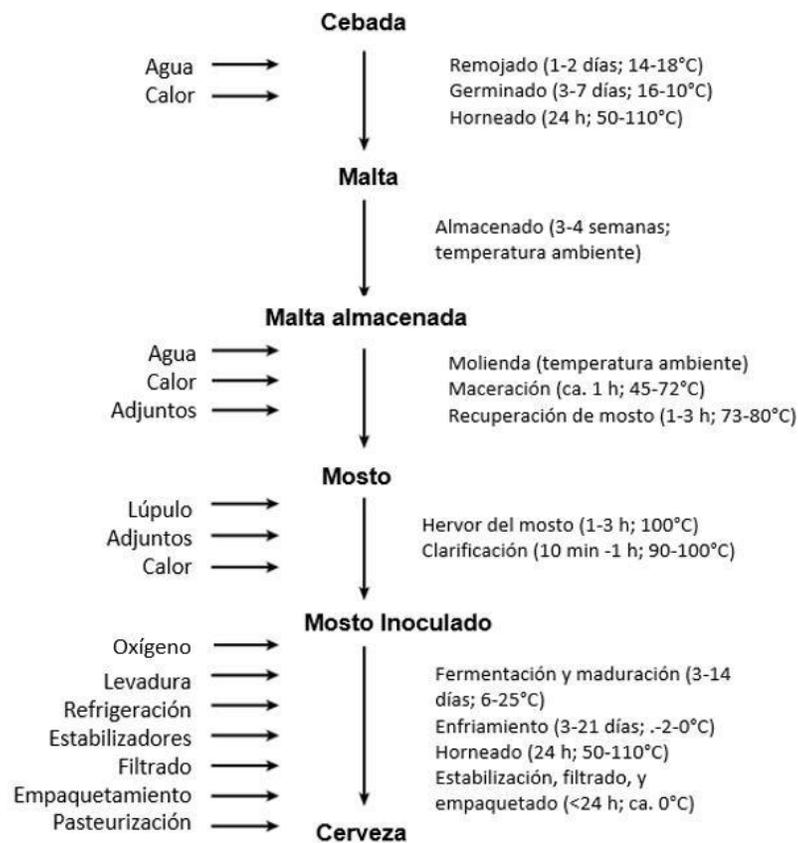
Posición 2018	País	Producción (kl)	Crecimiento del año previo (%)
1	China	38,927,200	-2.2
2	Estados Unidos	21,460,700	-1.7
3	Brasil	14,137,900	1.0
4	México	11,980,000	8.8
5	Alemania	9,365,200	0.7
6	Rusia	7,747,000	4.1
7	Japón	5,108,300	-2.7
8	Vietnam	4,300,000	-1.7
9	Reino Unido	4,228,200	4.5
10	Polonia	4,093,000	1.4

Es importante resaltar que, a pesar del asombroso avance tecnológico y científico que ha sufrido el proceso cervecero, la forma fundamental del proceso no ha sufrido cambios significativos (Bamforth, 2000).

La mayor parte del sector cervecero sigue las mismas operaciones unitarias fundamentales para alcanzar un producto final deseable. Una proporción importante de cervezas son producidas a partir de granos de cebada (*Hordeum vulgare*) que han sido malteados. En dicho proceso el grano de cebada es germinado en condiciones de humedad específica con el fin de promover la

activación de enzimas hidrolíticas que de otra forma no están presentes en el grano sin germinar. Estas enzimas son de suma importancia pues son las encargadas, en procesos siguientes, de degradar los carbohidratos complejos y proteínas en componentes más asimilables para la levadura (Kok *et al.*, 2019).

Posteriormente la cebada malteada es molida para la extracción acuosa a temperaturas altas. En esta parte, las enzimas hidrolíticas convierten el almidón a azúcares asimilables para la levadura. El líquido resultante (mosto), se recupera, separándolo del grano gastado, y es hervido en presencia de lúpulo para brindar aroma y amargor. Después del enfriamiento, el mosto es inoculado con levadura para que el proceso de fermentación sea llevado a cabo durante varios días. En esta parte, los carbohidratos son convertidos a etanol, dióxido de carbono y algunos compuestos sensoriales (Figura 1. [Bamforth, 2017]).



**Figura 1.** Principales pasos en el proceso de malteado y elaboración de cerveza. Tomado y modificado de Bamforth (2016).

La composición de un mosto estándar, abarca un 90% de azúcares; sacarosa, fructosa, glucosa, maltosa y maltotriosa (He *et al.*, 2014). No obstante, los principales azúcares como glucosa, maltosa y maltotriosa representan un 10-15%, 50-60% y 15-20% respectivamente (Stambuk *et al.*, 2006).

A pesar de la preponderante abundancia de maltosa y maltotriosa en los mostos cerveceros, su consumo es evitado hasta que la glucosa ha sido metabolizada en su totalidad, proceso que se denomina represión por catabolito de carbono (Gancedo. 1998; Kayikci y Nielsen. 2015).

En este mismo contexto jerárquico de consumo, la maltosa es fácilmente captada por la mayoría de las levaduras cerveceras y la maltotriosa no solo es el último carbohidrato consumido, sino que algunas levaduras no lo utilizan eficientemente (Meneses y Jiranek. 2002; Yoon *et al.*, 2003; Zastrow *et al.*, 2000). Esta dificultad que presentan algunas cepas de levaduras cerveceras exhibe áreas de oportunidad, tal como el contenido de azúcares fermentables en la cerveza, bajos rendimientos de etanol y perfiles sensoriales atípicos (Gibson *et al.*, 2013).

Este problema ha sido asociado principalmente a la capacidad de transporte de las levaduras cerveceras más que a su capacidad hidrolítica (Meneses y Jiranek. 2002) y es por esto que algunas investigaciones se han enfocado en entender este problema mediante distintos enfoques que involucran transportadores transmembranales en la levadura (Brickwedde *et al.*, 2017, 2018; Brouwers *et al.*, 2019; Z. Liu *et al.*, 2008; Stambuk *et al.*, 2006).

#### **4.2 Los principales exponentes microbianos en la elaboración de cerveza (levaduras tipo ale y lager)**

*Saccharomyces cerevisiae* (Sc) es, sin duda, el principal microorganismo en la producción de cerveza tipo Ale cuya temperatura de fermentación ronda entre los 16-26 °C. Este microorganismo ha sufrido algunos eventos genéticos típicos de un fenómeno de domesticación; variaciones estructurales del genoma,

aneuploidía, variación en el número de copias de genes y decaimiento genético en la reproducción sexual.

También se han observado variaciones pequeñas como la duplicación, delección, recombinación y conversión genética en regiones subteloméricas, muy importantes en la evolución (Gallone *et al.*, 2016; Selmecki *et al.*, 2015). Junto con la historia evolutiva de *Sc*, que le permite convertir etanol a partir de glucosa aún en presencia de oxígeno (efecto Crabtree), y los eventos de domesticación mencionados anteriormente, han permitido que este microorganismo haya colonizado un nicho industrial y desarrollado fenotipos deseables para el interés humano (Pis *et al.*, 2006).

De estos fenotipos, el más destacable es la habilidad de utilizar maltotriosa, azúcar abundante en el mosto cervecero y presente en mínimas concentraciones o ausente en ambientes naturales. Junto a estos eventos de domesticación, surgieron eventos de formación de pseudogenes o pérdida de genes relacionados con la reproducción sexual resultando en la incapacidad de sobrevivir en ambientes naturales (Gallone *et al.*, 2016; 2018).

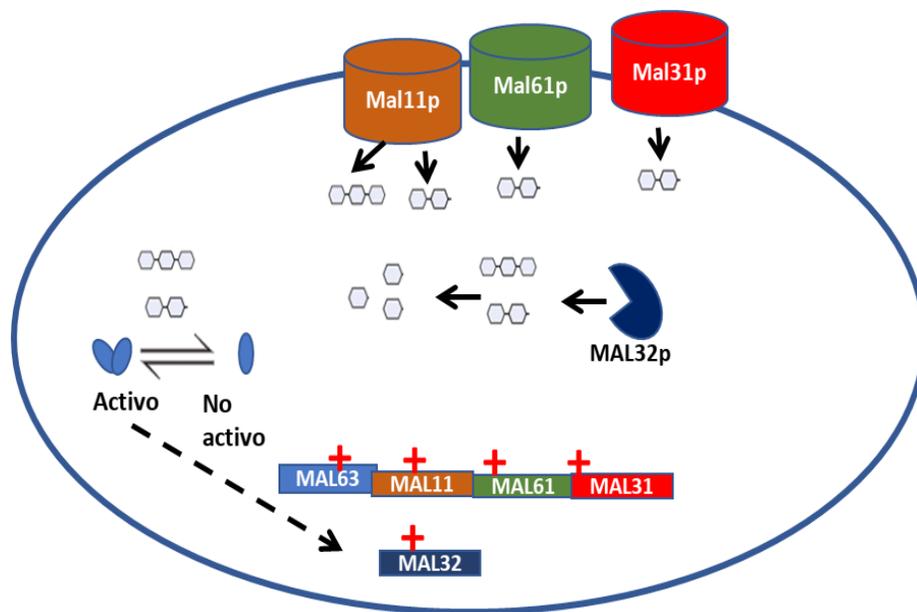
Además de colonizar nichos industriales relacionados con la producción de cerveza, *Sc* tienen una trascendencia económica más amplia, al estar presente en ambientes como la panificación, producción de otras bebidas fermentadas como el vino, sake e incluso se ha observado que es un microorganismo importante en la industria del chocolate (Parapouli *et al.*, 2020; Randez-Gil *et al.*, 2013).

Para el beneficio de todo el sector industrial y económico mencionado anteriormente *Sc* resulta ser un microorganismo modelo cuyo genoma, transcriptoma y proteoma han sido caracterizados en distintas condiciones fisiológicas, resultando en información importante para campos como la medicina, biología sintética y por supuesto para el sector económico que depende total o parcialmente de este fantástico microorganismo (Hanson. 2018).

A pesar de “el derecho por antigüedad” que tiene *Sc*, *Saccharomyces pastorianus* (*Sp*) domina el mercado cervecero con un 90% de participación. *Sp*

es utilizada en la elaboración de cerveza tipo lager cuya temperatura de fermentación se encuentra entre los 10-15°C. A partir de evidencia proveniente de la secuenciación y análisis del genoma completo, se sabe que *Sp* es un híbrido interespecífico entre *Sc* y una levadura tolerante al frío *Saccharomyces eubayanus* (*Se*) (Baker *et al.*, 2015; Libkind *et al.*, 2011; Peris *et al.*, 2016).

Entre los fenotipos importantes, se sugiere que la parte genómica de *Se* brinda la característica de tolerancia al frío, mientras que la parte *Sc* proporciona los fenotipo antes mencionado; la utilización eficiente de maltotriosa y por tanto una fermentación eficiente (Herbst *et al.*, 2017. Figura 2).



**Figura 2.** Proceso de la toma de carbohidratos complejos por la levadura. Primero, transportadores como Mal11p (*AGT1*), Mal61p y MAL31p transportan maltotriosa y maltosa. Dentro de la célula una enzima maltasa hidroliza los enlaces alfa 1→4 de los azúcares complejos y los convierte en unidades de glucosa. A su vez la maltosa y maltotriosa activan el factor transcripcional que activa los genes necesarios para el transporte y metabolismo de los carbohidratos complejos. Tomado y modificado de Gallone y cols (2018).

En este sentido, existe algo de discusión en esta tesis pues se ha observado que transportadores de maltotriosa provenientes de *Se* promueven el

metabolismo de maltotriosa en levaduras lager (Baker y Hittinger. 2019; Brouwers *et al.*, 2019).

Igualmente, cepas de *Sc* tienen rasgos que les permiten tolerar bajas temperaturas lo que implica que el aporte de tolerancia al frío, no necesariamente provienen del subgenoma de *Se* (García-Ríos *et al.*, 2017). Así mismo, recientemente se reportó que el rasgo involucrado en la adaptación a las temperaturas bajas está fuertemente asociado con el contenido genético proveniente en el ADN mitocondrial (ADNmt) de la levadura, brindando apoyo a la idea de que el aporte del ADNmt proveniente de *Se*, está involucrado en la adaptación a bajas temperatura y que el ADNmt de *Sc* orquesta el crecimiento óptimo en altas temperaturas (Baker *et al.*, 2019).

Debido a su naturaleza híbrida, *Sp* ha tenido un éxito sin precedentes en la colonización de nichos industriales en un periodo relativamente corto. Los eventos antes mencionados, sin duda, se han implicado en este éxito rotundo. Por lo tanto, la complejidad genómica de estas levaduras, lejos de representar eventos de inestabilidad, sustentan fenómenos de plasticidad genética que brindan una ventaja selectiva en ambientes estresantes como lo es la fermentación cervecera (Selmecki *et al.*, 2015).

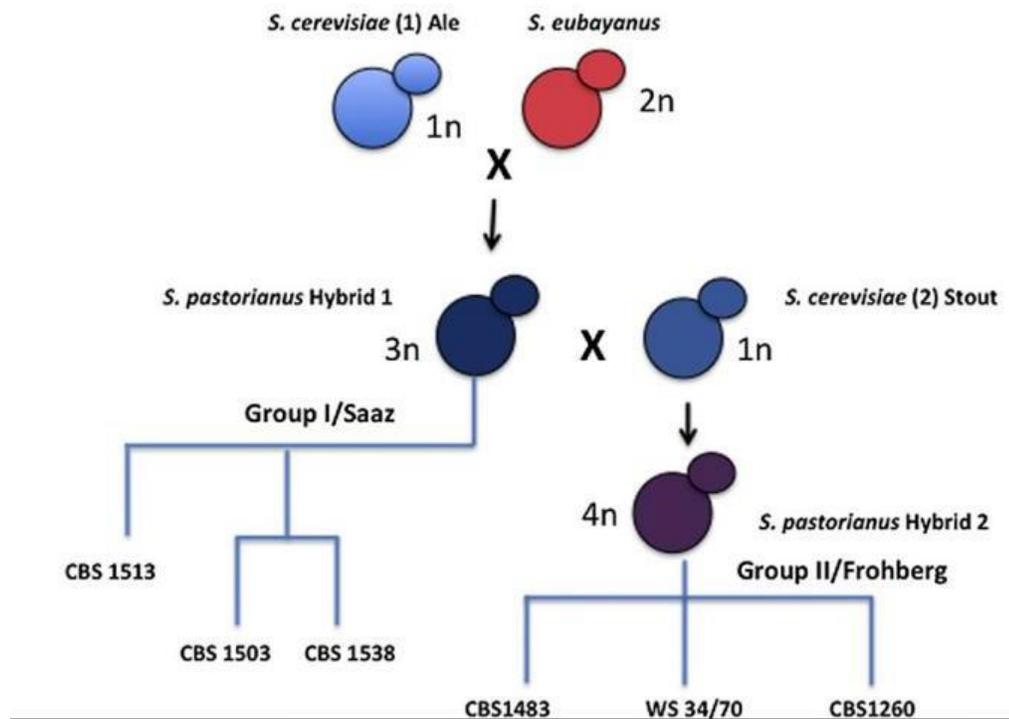
Muchos esfuerzos se han realizado para elucidar la naturaleza híbrida de *Sp* (ver Monerawela y Bond (2017)). Una de las pruebas contundentes fue la publicación de la primera secuencia del genoma completo de una levadura lager, cepa Weihenstephan 34/70. Dicho reporte reveló un genoma poliploide con un número estimado de 36 cromosomas y mediante el análisis de la secuencia se estableció que la parte subgenómica que no pertenecía a *Sc* provenía de *S. bayanus* (Nakao *et al.*, 2009).

Así las primeras consideraciones que aceptaban a *S. bayanus* como cepa parental de las levaduras lager, Libkind y cols (2011) demostraron que una nueva levadura, *S. eubayanus*, era responsable del subgenoma no-*cerevisiae* de las levaduras lager. También propusieron el modelo para la formación del híbrido de

*Sp*, que inició con la contaminación de fermentaciones cerveceras con levaduras *Se*, permitiendo la interacción de *Sc* y *Se*.

Esta investigación sugiere una hipótesis que indica el posible origen Patagónico de las levaduras lager. Aunque, se ha reportado que aislados de *Se* de la meseta tibetana comparten mayor identidad en secuencia que *Se* patagónica. En apoyo a estos hallazgos, se ha visto que el aporte subgenómico de *Se* originaria del Tíbet explica eficientemente los fenotipos en consumo de maltosa y maltotriosa de las cepas Lager (Bing *et al.*, 2014; Brouwers *et al.*, 2019).

Sea cual sea el origen de subgenoma no-*cerevisiae*, en la literatura existen distintas hipótesis que describen el origen de los tipos I y II de levaduras lager. Una de las ideas más aceptadas sugiere dos eventos de hibridación completamente independientes, cada uno con la participación de diferente levadura domesticada Ale (*Sc*) y diferente levadura salvaje *Se* (Figura 3).



**Figura 3.** Modelo hipotético sugerido para el surgimiento de las levaduras lager tipo I y II. Tomado de Monerawela y Bond, (2018).

Esta hipótesis es apoyada por análisis filogenéticos que indican que los subgenomas de las levaduras lager difieren significativamente entre los grupos I y II (Baker *et al.*, 2015). La distribución de la pérdida o retención de genes específicos en regiones subteloméricas provenientes del subgenoma de *Sc* también difieren significativamente entre estos dos grupos, sugiriendo distintas cepas Ale responsables del subgenoma *S.c* y por tanto abonando en la hipótesis de dos eventos de hibridación independientes (Monerawela *et al.*, 2015; Monerawela & Bond, 2017).

Pese a esto, investigaciones recientes discuten fuertemente el origen de los grupos I y II, sugiriendo que ambos grupos de levaduras lager comparten un único evento de hibridación, y la evolución resultante de esa hibridación dio forma a las divergencias genómicas entre ambos grupos (Salazar *et al.*, 2019).

Debido a la variación en el contenido genético y probablemente a los eventos de hibridación independiente mencionados anteriormente, se conocen dos grupos principales dentro del taxón de *Sp*. Las levaduras lager del tipo I y tipo II, conocidas comúnmente como Saaz y Froberg, respectivamente (Dunn y Sherlock, 2008).

Similarmente, el análisis de secuencias genómicas estableció que las cepas tipo I contienen un rango de entre 31 a 47 cromosomas, mientras que las del tipo II entre 36-79 cromosomas. Típicamente, las cepas del grupo I y II difieren en su contenido de ADN de *Sc*, siendo las cepas del grupo I con el menor contenido genético de *Sc* (Monerawela y Bond. 2017).

Las diferencias genómicas mencionadas anteriormente se traducen en rendimientos y actividad fisiológica diferencial entre ambos tipos de levaduras. Dentro de las diferencias más importantes, está la utilización poco eficiente de la maltotriosa por parte de las cepas del grupo I, a diferencia de las cepas tipo II que la utilizan eficientemente. Otras características importantes es que las cervezas producidas a partir de levaduras del grupo I, presentan un perfil sensorial seis veces menor comparado con las fermentaciones de levaduras del grupo II (Gibson *et al.*, 2013).

Los datos experimentales sugieren que las cepas tipo I presentan menor eficiencia en las fermentaciones debido a su bajo contenido genético relacionado con *Sc* y que, por su contraparte, las del grupo II presenta altos valores de eficiencia que posiblemente se relacionen con su mayor contenido genético proveniente de *Sc*. A pesar de esto, una investigación interesante ha expuesto que dentro del grupo I existe una gran variación respecto al fenotipo de utilización de maltotriosa. De hecho, se observó que algunas cepas del grupo I consumen la maltotriosa de forma similar a las cepas del grupo II. Lo que sugiere la presencia o ausencia de determinados transportadores específicos para el consumo de maltotriosa (Magalhães *et al.*, 2016).

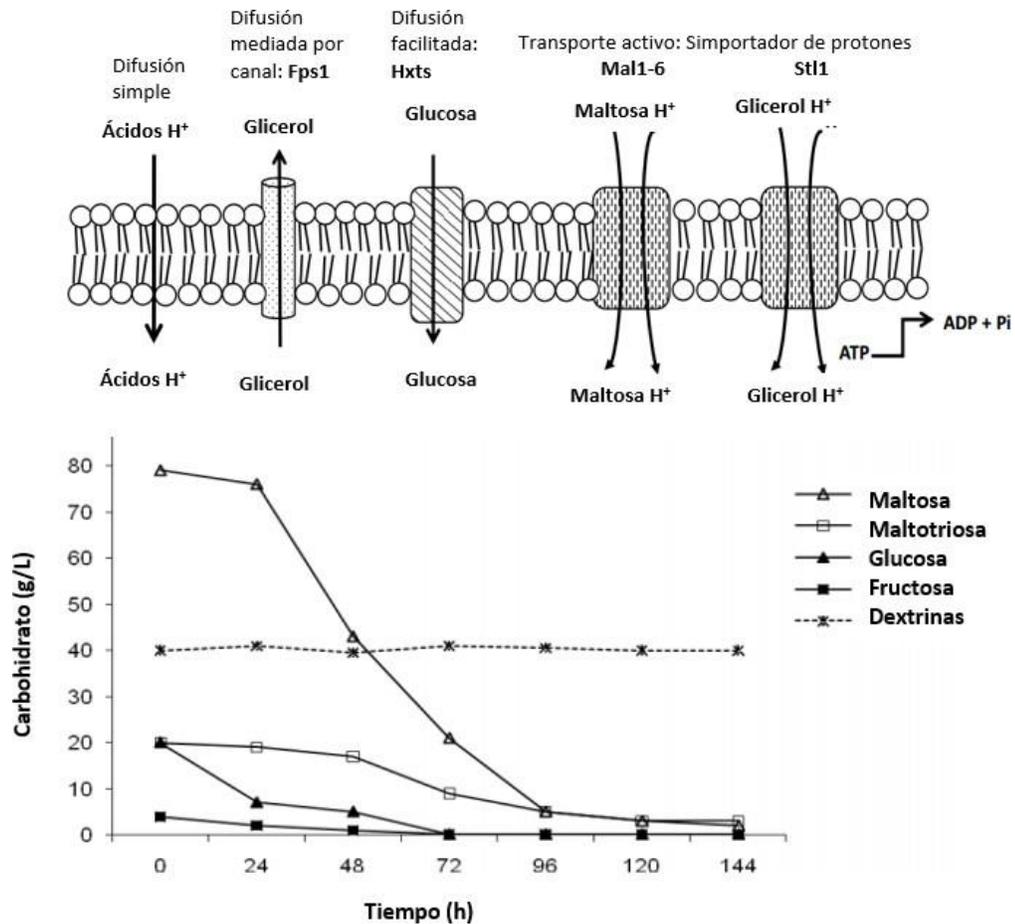
### **4.3 Asimilación de carbohidratos y sistemas de transporte en levaduras cerveceras**

La interfase entre el medio ambiente y el citoplasma de la célula de levadura está constituida principalmente por la membrana plasmática, aunque podemos considerar la pared celular y el espacio periplásmico. En este aspecto la pared celular es porosa y los azúcares son capaces de difundir a través de ella. Por lo tanto, la principal barrera que tiene la levadura frente al medio ambiente es la membrana plasmática. Los nutrientes como azúcares, presentes en el medio ambiente, no pasan deliberadamente la membrana plasmática.

La levadura emplea proteínas embebidas en la membrana plasmática denominados transportadores, para el consumo de distintas fuentes de carbono, entre ellos los azúcares como la glucosa. Existen tres principales mecanismos de transporte que permiten transportar nutrientes a través de la membrana plasmática; difusión simple, difusión facilitada y transporte activo, en donde las últimas dos categorías emplean proteínas transportadoras (Figura 4A) (Ramos *et al.*, 2016).

La levadura cervecera puede asimilar los carbohidratos presentes en el mosto, pero cuando la glucosa está presente, la captación jerárquica entra en operación. La glucosa y la fructosa son internalizados primero por una familia de transportadores de hexosas (HXT) (Ramos *et al.*, 2016).

Estos transportadores realizan la translocación de glucosa mediante difusión facilitada independiente de energía. Debido a que la represión por catabolito es un mecanismo genético altamente conservado en los hongos (Ronne. 1995), no es sorprendente que en una fermentación cervecera la sacarosa y la glucosa son consumidos por la levadura en las primeras etapas de la fermentación (Figura 4B) (Stewart. 2009).



**Figura 4.** A) Principales mecanismos de transporte en levadura. B) Orden de consumo de los principales carbohidratos; Glucosa→Fructosa →Maltosa→Maltotriosa. Figura 3A tomado de Ramos y cols (2016) y Figura 3B tomado de Stewart (2009).

La glucosa es el sustrato preferido de la levadura, resultando en un retardo considerable en el transporte de la maltosa y la maltotriosa (Figura 4B). El mecanismo por el cual la maltosa, la maltotriosa y cualquier otra fuente alternativa

de carbono se ve reprimida es la represión o inhibición por catabolito (Kayikci y Nielsen, 2015).

Por otro lado, la maltosa y la maltotriosa son internalizadas en el citoplasma de la levadura por un transporte acoplado a energía a través de un mecanismo por simportador, en el cual un protón es cotransportado con una molécula de maltosa o maltotriosa (Henderson y Poolman, 2017).

La energía necesaria para este transporte es proporcionada por un gradiente electroquímico de protones presente en la membrana, generado por una ATPasa que bombea un protón fuera de la célula por ATP hidrolizado (Figura 4A). A pesar de que existen levaduras que consumen ambos, maltosa y maltotriosa, es un hecho que, dentro del repertorio de transportadores reportados hasta el momento, la maltotriosa no tiene su propio transportador, lo que significa que algunos, pero no todos, transportadores de maltosa también transportan maltotriosa (Cousseau *et al.*, 2013; Dietvorst *et al.*, 2005, 2010; Stambuk & Araujo. 2001; Vidgren *et al.*, 2005; Yoon *et al.*, 2003).

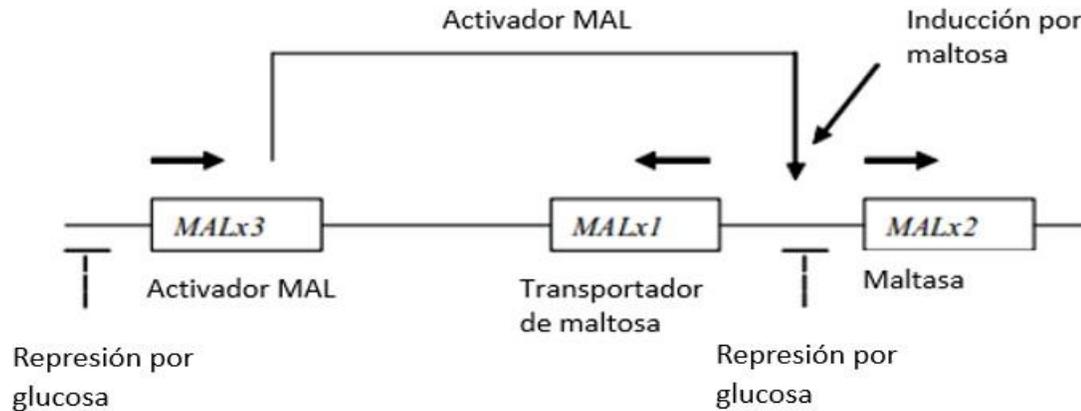
La mayoría de los transportadores reportados transportan más eficientemente la maltosa que la maltotriosa determinando así que la captación de la maltotriosa suceda en las etapas finales de la fermentación.

#### **4.4 El centro de operaciones del transporte de alfa glucósidos; los *loci* MAL**

La toma eficiente de maltosa, requiere la participación de, al menos, uno de los cinco *loci* MAL (*MAL1-4*, *MAL6*) los cuales han sido caracterizados en regiones subteloméricas (Brown *et al.*, 2010; Monerawela *et al.*, 2015).

El *locus* MAL canónico comprende tres componentes fundamentales; al menos una copia de los genes *MALx1*, *MALx2* y *MALx3* (x representa el *locus*) que codifican a un transportador Malx1p, una maltasa intracelular (alfa-glicosidasa) Malx2p que se encarga de hidrolizar los azúcares complejos en monómeros de glucosa y, finalmente un factor transcripcional Malx3p requerido

para la inducción de la expresión genética de Malx1p y Malx2p en presencia de maltosa (Figura 5) (Horák, 2013).



**Figura 5.** Estructura y organización de un locus MAL canónico. Se observan los puntos de regulación por el factor transcripcional MALx3 que se une a un promotor bidireccional para activa a MALx1 y MALx2. Además, se observan los puntos de represión por catabolito impartidos en la presencia de Glucosa. Tomado y modificado de Vidgren (2010).

En levaduras *Sp* los *loci* ScMAL1-4 y ScMAL6 de Sc así como SeMAL1-SeMAL4 provenientes de Se también han sido mapeados en regiones subteloméricas (Baker *et al.*, 2015b; Brickwedde *et al.*, 2018; Salazar *et al.*, 2017).

Trabajos previos evidenciaron la existencia de cuatro transportadores con homología codificados por los genes MALx1, AGT1, MPH2 y MPH3. En general las proteínas Malx1p exponen alta afinidad ( $K_m$  2-4 mM) hacia la maltosa y además aceptan otros sustratos como turanosa, pero no otros como la maltotriosa (Alves *et al.*, 2008; Stambuk y Araujo, 2001).

Usualmente las cepas ale y lager poseen varias copias de MALx1 en sus genomas (Jespersen *et al.*, 1999). Cabe resaltar que, además de los transportadores provenientes del subgenoma Sc, las levaduras lager también poseen transportadores de maltosa Malx1p provenientes del subgenoma de *Seubayanus* (Baker *et al.*, 2015; Libkind *et al.*, 2011; Nakao *et al.*, 2009).

Los transportadores YDL247w y YJR160c fueron descritos originalmente en levaduras *Sc*, estos marcos de lectura abiertos (ORF, por sus siglas en inglés) se asociaron con transportadores de maltosa y se denominaron *MPH2* (YDL247w) y *MPH3* (YJR160c). Lo característico de estos genes es que comparten un alto grado de identidad, pero se localizan en distintas cromosomas; *MPH2* en el cromosoma IV y *MPH3* en el cromosoma X. Estos transportadores tienen alta afinidades hacia maltosa y maltotriosa con una  $K_m$  4.4 y 7.2 respectivamente (Day *et al.*, 2002). No obstante, su afinidad hacia la maltotriosa ha sido muy discutida y se menciona que, de hecho, es ambigua o totalmente nula (Brown *et al.*, 2010; Magalhães *et al.*, 2016).

Algunos trabajos mostraron resultados discrepantes. Por un lado, ensayos de transporte de alfa glucósidos que *MPH2* y *MPH3* no evidenciaron un transporte eficiente de maltotriosa (Alves *et al.*, 2008). Incluso se propuso que estas proteínas no tienen un rol principal y por tanto exponen mínima influencia en el consumo de maltotriosa en levaduras lager (Duval *et al.*, 2010).

Por otro lado, el gen *AGT1* que codifica para un transportador de alfa glucósidos, alelo natural de *MAL11*, con un 57% de identidad entre ellos, es un transportador de alfa glucósidos de amplio espectro con una afinidad importante hacia trealosa y sacarosa ( $K_m$  8 mM); menor afinidad para maltosa ( $K_m$  5-18 mM), maltotriosa ( $K_m$  18 mM); y la menor afinidad hacia isomaltosa, melecitosa y palatinosa (Han *et al.*, 1995). Estos trabajos en cepas de *Sc*, por tanto, estos datos no son aplicables a las cepas de *Sp*. Estudios en levaduras lager indicaron que el gen *AGT1* proveniente del subgenoma de *Sc* contenía un codón de paro en el nucleótido 1183, resultando en un producto proteínico trunco de 394 aminoácidos mientras que las levaduras *Ale* analizadas exhibieron una proteína de 610 aminoácidos totalmente funcional (Vidgren *et al.*, 2005; 2009).

Los análisis genómicos de una levadura lager revelaron otro gen presente en la región subtelomérica del cromosoma XV que compartía un 79% de identidad con *Sc-AGT1* (Nakao *et al.*, 2009). Posteriormente, este transportador fue caracterizado, demostrando ser un transportador de relativa especificidad

para maltosa y maltotriosa (Km 17 y 22 mM, respectivamente), a pesar de su divergencia en secuencia con solo 85% de identidad a nivel de aminoácidos con Sc-*AGT1* (Vidgren y Londesborough, 2012).

Por último, el gen *MTT1* (*MTY1*), descrito por dos investigaciones independientes, de ahí la variación en el nombre. Salema-Oom y cols (2005) denominaron a este gen *MTY1* (de las siglas en inglés de “maltose transporter in yeast”) mientras que Dietvorst y cols (2005) los denominaron *MTT1* (por las siglas en inglés de “mty-like transporter”). Este gene comparte un 90% y 40% de identidad con genes *MALx1* y *AGT1*, respectivamente. De forma interesante, este transportador expone mayor afinidad hacia maltotriosa (Km 16-27 mM) que maltosa (Km 61-88 mM), representando una característica de suma importancia para el sector industrial. Al igual que la mayoría de los transportadores de maltosa y *AGT1*, *MTT1* transporta su sustrato mediante el mecanismo de simportador donde los protones son cotransportados con maltosa o maltotriosa en una relación estequiométrica 1:1 (Salema-Oom *et al.*, 2005).

Otro aspecto interesante mostró que *MTT1* posee variantes, como la denominada *MTT1 alt* (Dietvorst *et al.*, 2005). Esta versión contiene una delección de 66 pb del extremo 3' del gen. Esta versión resultó más eficiente en el transporte de maltotriosa que la versión original de *MTT1*. Esta observación es de particular importancia ya que, en investigaciones anteriores en nuestro laboratorio, mostraron que la secuencia de *MTT1* de una cepa industrial muestra un gen del tamaño de 594 aminoácidos y en otra cepa el tamaño es de 610 (Hernández-Vásquez, 2020). Dicho evento puede abonar en el camino deductivo hacia la respuesta a un problema generalizado como el azúcar residual en la cerveza.

#### **4.5 Principales aproximaciones al problema de azúcares residuales**

Debido a que el paso limitante en una fermentación eficiente es el transporte de alfa-glucósidos dentro de la levadura (Figura 2), muchas investigaciones se han dirigido hacia la comprensión de los factores relacionados con estos fenotipos.

Day y cols (2002) determinaron que en levaduras *Ale* (*Sc*), el transporte de maltosa y maltotriosa es un evento competitivo y excluyente, siendo la maltosa el sustrato preferido para el transporte. Esto sugiere que la presencia residual maltotriosa en la cerveza no es debido a la incapacidad genética y fisiológica de la levadura, sino más bien a que el sistema de transporte presenta baja afinidad hacia maltotriosa en conjunto con las condiciones estresantes en las últimas fases de la fermentación resulta en su captación ineficiente.

Otras investigaciones han realizado análisis de la cinética de transporte de cepas de *Sc* y a diferencia de lo observado en otras cepas *Sc*, la maltotriosa es eficientemente utilizada debido a la presencia de genes específicos como *AGT1* y *MTT1*, que presentan distinta afinidad frente a la maltosa y la maltotriosa (Tabla 2) (Alves *et al.*, 2008; Salema-Oom *et al.*, 2005).

**Tabla 2.** Condensado de las afinidad y velocidad máxima de distintos transportadores. Para una revisión de las referencias ir a Vidgren, (2010). Lo resaltado en color naranja indica la afinidad hacia la maltotriosa de los transportadores *AGT1* y *MTT1*.

Transportador	Maltosa		Maltotriosa	
	Km (mM)	V <sub>max</sub>	Km (mM)	V <sub>max</sub>
<i>Mal21</i>	4.7 ±0.2	115 ±3	n.d.	n.d.
<i>Mal31</i>	4.2 ± 1.1	42	5.7 ± 1.0	41
	4.3 ± 0.6	n.d.	335 ± 165	n.d.
<i>Mal61</i>	2–4	n.d.	n.d.	n.d.
	2.7 ± 0.6	36	7.2 ± 0.9	40
<i>Agt1</i>	5–10	n.d.	n.d.	n.d.
	17.8 ±1.3	13.2 ± 0.5	18.1 ± 1.8	7.8 ± 0.3
	5.1 ± 0.6	41	4 ± 0.7	39
	30	n.d.	36 ±2	n.d.

<i>Mphx</i>	4.4 ± 0.5	39 ± 4.2	7.2 ± 1.0	49 ± 5.4
<i>Mtt1</i>	61–88	n.d.	16-27	n.d.
	41 ± 0	n.d.	23 ± 4	n.d.

Por otra parte, las investigaciones en levaduras lager han comprobado rotundamente que la presencia o ausencia de genes específicos promueven fenotipos eficientes en la captación de maltotriosa. De estos resultados, interesantes conclusiones han surgido respecto a la gran variación existente de este fenotipo dentro de las levaduras del grupo I.

Se ha reportado que levaduras lager del grupo I presentan un rendimiento en fermentación muy similar a las levaduras del grupo II. Mediante la caracterización de la cinética de transporte con el uso de maltosa y maltotriosa radio marcada y con el apoyo de información proveniente del mapeo cromosómico de cada transportador y determinación del número de copias, se concluyó que el gen *MTT1* es el principal responsable de la fermentación eficiente en levaduras lager del grupo I (Magalhães *et al.*, 2016).

Esto pone en evidencia que dentro del grupo I la distribución diferencial de genes específicos está relacionada con el consumo eficiente de la maltotriosa. Por otra parte, resultados transcriptómicos provenientes de nuestro laboratorio (Hernández, 2020), indican que nuestra levadura de estudio perteneciente al grupo I, contienen un el gen anotado *MAL61* el cual comparte 100% de identidad con el gen *MTT1* reportado por (Dietvorst *et al.*, 2005).

Y de forma interesante, se observó que este gen presenta mayor nivel de expresión relativa en la cepa del grupo I. Contrario a esto, los rendimientos en la utilización de maltotriosa varían entre las cepas siendo menor en la cepa del grupo I.

Experimentos de evolución asistida en levaduras *Se*, generaron genes quiméricos capaces de transportar maltotriosa. Brouwers y cols (2019) determinaron que en la cepa *Se* CBS 12357 expresó un transportador nuevo, *SeMALT413*, el cual se generó por la recombinación entre tres diferentes

transportadores de maltosa de esta misma cepa. Esto indica que las condiciones de fermentación inducen procesos de recombinación que resulta en la neofuncionalización de genes que les permite adaptarse a las condiciones.

De la misma forma, Baker y Hittinger (2019) demostraron la presencia de un transportador quimérico que permitía el transporte de maltotriosa en *Se* FM1318. Se presume que el transportador surgió por recombinación de genes parálogos, permitiendo rápidamente la adaptación de *Se* al consumo de maltotriosa.

De hecho, el gen específico de levaduras lager *MTY1/MTT1*, ha sido caracterizado y expone un 90% de identidad en secuencia con genes *MALx1* (Dietvorst *et al.*, 2005; Salema-Oom *et al.*, 2005) y de forma interesante también muestra identidad en secuencia en fragmentos específicos con transportadores de maltosa de *Se* (Cousseau *et al.*, 2013). Estos eventos muestran que *MTT1* tiene una estructura mosaico similar a los genes quiméricos reportados antes, y posiblemente el surgimiento involucró eventos de recombinación durante el proceso de domesticación de las levaduras lager (E. P. Baker & Hittinger, 2019; Brouwers, Vries, *et al.*, 2019).

Debido a los factores involucrados en los fenotipos diferenciales expuestos anteriormente y al hueco en el conocimiento sobre el transporte de maltotriosa de las levaduras lager del grupo I, nos propusimos investigar si los eventos de expresión diferencial, número de copias o variación estructural están involucrados en los fenotipos diferenciales de consumo de maltosa y maltotriosa en nuestras cepas de estudio.

Esto mediante la caracterización del transporte de maltosa y maltotriosa de nuestras cepas de estudio, así como el análisis de la secuencia nucleotídica del gen estructural y regulador de los transportadores mencionados. Con el fin de aportar información valiosa en uno de los pasos más importantes para la determinación de una fermentación eficiente.

#### 4.6 La regulación entre reinos mediada por RNA y sus principales mecanismos

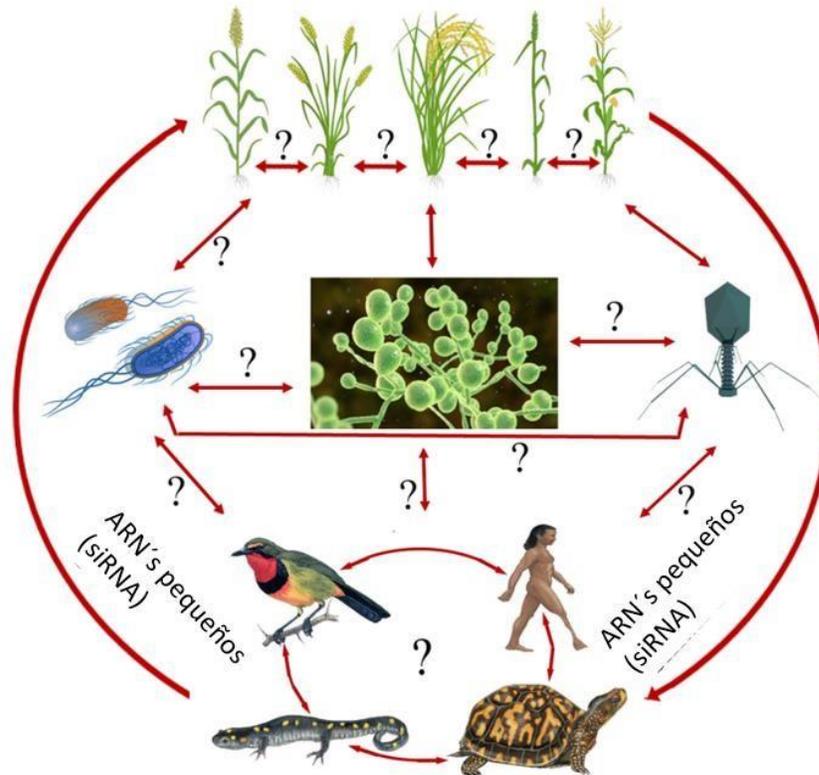
La regulación entre reinos mediante ARN es una hipótesis que trata de explicar cómo las moléculas de ARN de un organismo pueden influir en la biología de otro organismo perteneciente a un reino diferente. Este fenómeno presume de ser muy importante en varias interacciones ecológicas, agrícolas y médicas, y se está investigando cada vez más debido a sus potenciales aplicaciones en biotecnología y salud (Samad *et al.*, 2021). El mecanismo más estudiado es la interferencia de ARN (RNAi). En este proceso, moléculas de ARN pequeñas, como ARN pequeños interferentes (siRNA) y microARN (miRNA), juegan un papel crucial en la regulación de la expresión génica.

Estas moléculas pueden moverse entre organismos de diferentes reinos y modificar la expresión de genes específicos, afectando procesos fisiológicos y patológicos (Figura 6). Este mecanismo se basa en que el organismo receptor contenga el sistema efector de la interferencia como podría ser DICR y el complejo RISC para efectuar la regulación (Wang *et al.*, 2018).

De los ejemplos más destacados en este tipo de regulación, es la defensa de las plantas contra patógenos. Las plantas producen siRNAs que se transfieren a patógenos como hongos y virus, donde ejercen su efecto regulatorio en genes críticos para la virulencia del patógeno. Este proceso es conocido como "silenciamiento génico inducido por huésped" (Host-Induced Gene Silencing, HIGS) y es una estrategia que las plantas utilizan para protegerse de dichos fitopatógenos (Bilir *et al.*, 2022; Hua *et al.*, 2018).

En esta regulación también sucede lo contrario, es decir algunos fitopatógenos secretan RNA regulatorios que tienen efecto en la expresión genética de las plantas. Se ha demostrado que hongos patógenos secretan pequeños ARN que interfieren con las vías de RNAi en plantas hospedadoras, debilitando sus defensas y facilitando la infección (Weiberg *et al.*, 2013).

Por no dejar de mencionar el gran impacto biológico de este fenómeno, otro ámbito de estudio es la interacción entre plantas e insectos herbívoros. Se ha observado que los miRNAs de las plantas pueden ser absorbidos por los insectos a través de la alimentación, y estos mismos pueden regular la expresión génica en los insectos, afectando su desarrollo, reproducción y comportamiento (Chi *et al.*, 2023).



**Figura 6.** Principales interacciones mediadas por RNA pequeños que soportan una comunicación entre reinos con evidencia experimental. Los signos de interrogación (?) representan poca o nula evidencia que respalde la comunicación mediada por ARN's pequeños. Tomado y modificado de Rabuma y cols (2022).

Un ejemplo es la capacidad de los miRNAs de plantas para interferir con los genes que regulan el metabolismo en insectos, lo que podría reducir su capacidad para digerir ciertos compuestos vegetales o para reproducirse eficazmente.

En otro contexto, los miRNAs también desempeñan un papel en la regulación de los microbiomas. Los miRNAs de las plantas pueden afectar la

expresión génica en microorganismos del suelo, influenciando la composición y función del microbioma de la rizosfera. Este tipo de regulación puede tener implicaciones significativas en la salud del suelo y en la productividad de las plantas, ya que los microbiomas son cruciales para la disponibilidad de nutrientes y la protección contra patógenos del suelo (Middleton *et al.*, 2021, 2023).

Este tipo de regulación entre reinos podría ser útil en el control de plagas en la agricultura, promoviendo enfoques más sustentables que la utilización de plaguicidas químicos. Se ha demostrado que los miRNAs de plantas pueden regular genes en insectos y en algunos patógenos microbianos, lo que sugiere que esta tecnología podría ser utilizada para diseñar plantas con resistencia mejorada (Hernández-Soto *et al.*, 2021).

En términos de salud humana, existe un creciente interés en la posibilidad de que los miRNAs de plantas consumidos en la dieta puedan influir en la expresión génica en humanos. Aunque este tema es aún controvertido y requiere más investigación, algunos estudios preliminares han sugerido que los miRNAs vegetales pueden tener efectos biológicos en mamíferos, incluyendo posibles roles en la regulación de genes implicados en el metabolismo y la inmunidad (Cavalieri *et al.*, 2016; Zhang *et al.*, 2012).

Finalmente, la regulación entre reinos también podría tener aplicaciones en la biotecnología y la medicina, especialmente en el desarrollo de terapias basadas en ARN. El diseño de siRNAs específicos para genes patógenos podría ofrecer nuevas formas de tratar enfermedades infecciosas, mientras que la modulación de miRNAs podría ser una estrategia para influir en el microbioma humano y tratar trastornos metabólicos o inflamatorios. La regulación entre reinos mediada por ARN es un campo fascinante y en rápido desarrollo que promete revolucionar nuestra comprensión de la biología y abrir nuevas vías para la intervención biotecnológica. A medida que la investigación avanza, se espera que surjan nuevas aplicaciones que aprovechen esta forma de comunicación molecular para mejorar la agricultura, la salud y la biotecnología.

#### 4.7 Principales investigaciones en el tema de la regulación entre reinos

En los últimos años, el campo de la regulación entre reinos mediada por ARN ha avanzado significativamente, con nuevas investigaciones que han ampliado nuestro entendimiento de este fenómeno y sus aplicaciones potenciales en agricultura, medicina y ecología. Estudios recientes han expuesto cómo las plantas utilizan siRNAs para defenderse contra patógenos. De hecho, se ha descubierto que las plantas de tabaco pueden transferir siRNAs a patógenos fúngicos para suprimir genes de virulencia, lo que fortalece la hipótesis de que el silenciador génico inducido por el huésped (HIGS) es una defensa generalizada en plantas contra una variedad de patógenos.

Este estudio destacó la versatilidad de las plantas en la producción de siRNAs específicos para diferentes patógenos, lo cual es crucial para el desarrollo de cultivos resistentes a enfermedades (Wang *et al.*, 2016). Por otro lado, también se ha avanzado en la comprensión de cómo los miRNAs vegetales pueden influir en los insectos. Zha y cols. (2018) demostraron que los miRNAs de arroz pueden ser absorbidos por el insecto plaga *Nilaparvata lugens* (saltahoja marrón), donde regulan la expresión de genes que controlan el desarrollo y la fecundidad del insecto.

Este estudio sugirió que los miRNAs vegetales podrían ser una estrategia biotecnológica viable para el control de plagas, reduciendo la dependencia de pesticidas químicos y contribuyendo a prácticas agrícolas más sostenibles (Jing *et al.*, 2023). La interacción entre plantas y microbiomas también ha sido un foco de investigación reciente. Se ha discutido que los miRNAs exudados por plantas de *Arabidopsis* pueden afectar la expresión génica en bacterias del suelo, influyendo en la estructura y función del microbioma. Estas aproximaciones proporcionan evidencia de que las plantas pueden modular activamente su entorno microbiano a través de la liberación de moléculas de ARN, lo que podría tener implicaciones para la agricultura sostenible y la gestión del suelo (Huang *et al.*, 2019).

La investigación en este fenómeno de la regulación entre reinos mediada por ARN continúa expandiéndose, con un enfoque en aplicaciones prácticas y la comprensión de los mecanismos subyacentes. Desde 2017, la investigación en regulación entre reinos mediada por ARN ha avanzado notablemente, revelando la complejidad y el potencial de esta forma de comunicación molecular.

A medida que la tecnología y los métodos de análisis avanzan, es probable que descubramos más ejemplos de regulación entre reinos y que estas interacciones se aprovechen para aplicaciones prácticas en agricultura y medicina. Hasta cierto punto, queda mucho por entender especialmente en lo relacionado con la funcionalidad y la relevancia biológica de estos ARN en organismos no vegetales, lo que continuará siendo un área importante de investigación futura.

#### **4.8 La relación bidireccional entre la interacción planta-hongo fitopatógeno**

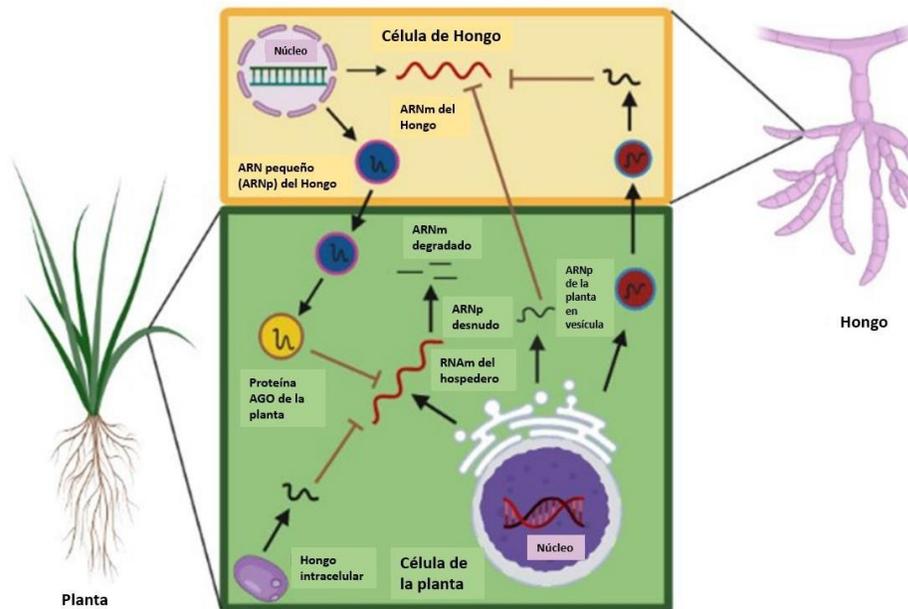
En el contexto específico de la interacción planta-hongo fitopatógeno, la comunicación entre reinos juega un papel crucial en la defensa de las plantas contra infecciones y en la capacidad de los hongos para colonizar y causar enfermedades. Las plantas han desarrollado mecanismos sofisticados para defenderse contra los patógenos, incluidos los hongos fitopatógenos (Figura 7).

Uno de estos mecanismos es el silenciamiento génico inducido por huésped (HIGS), en el cual las plantas producen siRNAs específicos que se transfieren a los hongos patógenos y silencian genes esenciales para su virulencia. Un brillante estudio sugiere que los hongos como *Botrytis cinerea* pueden liberar ARN pequeños que interfieren con las vías de RNAi en sus plantas hospedadoras, debilitando sus defensas y facilitando la infección.

Este hallazgo resaltó la capacidad de los hongos para utilizar ARN como una herramienta de patogenicidad, así como el potencial de las plantas para

contrarrestar estas estrategias patógenas mediante la ingeniería genética para producir siRNAs dirigidos a genes específicos de virulencia en hongos (Weiberg *et al.*, 2013).

Una revisión propone los posibles mecanismos de cómo el hongo *Fusarium graminearum*, un patógeno significativo en cultivos de cereales, interactúa con su hospedador. Se describe que las plantas pueden utilizar siRNAs para regular la expresión de genes en el hongo, reduciendo así su capacidad para causar enfermedades.



**Figura 7.** Principales mecanismos de la comunicación entre reinos que tiene lugar en la relación bidireccional de algunas plantas contra hongos fitopatógenos. Tomado y modificado de Mahanty y cols (2023).

Este tipo de interferencia de ARN bidireccional subraya la complejidad de las interacciones planta-patógeno y sugiere que los sistemas de RNAi son un campo fértil para la investigación y el desarrollo de cultivos resistentes a enfermedades (Machado *et al.*, 2018).

La regulación entre reinos mediada por ARN en la relación planta-hongo fitopatógeno tiene importantes implicaciones para la agricultura. Los enfoques basados en HIGS y la manipulación de miRNAs ofrecen nuevas estrategias para desarrollar cultivos resistentes a enfermedades. Al generar plantas que produzcan siRNAs y miRNAs específicos para genes clave en patógenos, es posible reducir significativamente el impacto de enfermedades fúngicas en la producción agrícola, mejorando la seguridad alimentaria y reduciendo la dependencia de fungicidas químicos (Qiao *et al.*, 2021).

La regulación entre reinos mediada por ARN en la interacción planta-hongo fitopatógeno es un campo de investigación prometedor que está revelando nuevas dimensiones de la defensa de las plantas y la patogenicidad de los hongos. A medida que avanzamos en nuestra comprensión de estos mecanismos, es probable que surjan nuevas aplicaciones en biotecnología agrícola, proporcionando herramientas más eficaces y sostenibles para el manejo de enfermedades de los cultivos.

La investigación continua en esta área no solo contribuirá al desarrollo de nuevas estrategias de protección de cultivos, sino que también ampliará nuestro conocimiento fundamental de las complejas interacciones ecológicas y evolutivas entre plantas y patógenos. Que, en última instancia, pueden tener repercusiones importantes incluso en el ámbito de la biotecnología aplicada como la elaboración de cerveza.

## 5. JUSTIFICACIÓN

Debido a que la producción de cerveza es un sector industrial importante para México y el mundo que impacta social y económicamente, es de interés aportar información, que soporte a la teoría existente sobre el problema de azúcares residuales, y se involucre en el desarrollo de procesos más eficientes.

El transporte de alfa glucósidos como la maltosa y la maltotriosa es uno de los pasos limitantes más importantes en el desarrollo de una fermentación eficiente, por lo tanto, el desarrollo de investigación que aporte información relevante es de importancia capital en contextos de desarrollo de nuevos productos o en el control y desarrollo de procesos más eficientes.

Aunado a lo anterior, la elaboración de cerveza es el proceso biotecnológico más antiguo de la humanidad y es de importancia capital estudiar la posible relación molecular existente entre los tres organismos presentes en dicho proceso: la levadura, la cebada y el lúpulo.

Con el fin de evidenciar la posible regulación genética entre reinos y abonar en el vasto conocimiento existente en la biotecnología de la cerveza, el estudio de la regulación mediante RNA no codificante proveniente de las plantas, malta y lúpulo, hacia la expresión genética de la levadura representa una idea innovadora que posiblemente brinde información importante sobre los mecanismos de variabilidad en la fermentación cervecera.

## 6. HIPÓTESIS

### **6.1 Hipótesis relacionada con la toma diferencial de maltosa y maltotriosa en las levaduras SplB1 y SplB2.**

1. La variación nivel de expresión de los genes *MALx1*, *MPHx*, *AGT1* y *MTT1*, involucrados en el transporte de maltosa y maltotriosa determinan la toma diferencial de alfa glucósidos en dos levaduras de proceso.

2.- La variación en la estructura de las proteínas transportadoras afecta aminoácidos importantes en el mecanismo de translocación de maltosa o maltotriosa.

### **6.2 Hipótesis relacionadas con el efecto en la expresión genética de las levaduras de estudio debido a la presencia de RNA no codificante presente en el mosto cervecero.**

1- Existirá un efecto significativo en la expresión genética global de las levaduras de estudio debido a la presencia de RNA en el mosto cervecero.

2.- Una fracción de los genes diferencialmente expresados y regulados a la baja en cada levadura pertenecerá a alguna categoría con importancia industrial.

## **7. OBJETIVOS DEL TRABAJO**

### **7.1 Objetivo general relacionado con la hipótesis de la sección 6.1**

Demostrar que las diferencias en la expresión de los genes *MALx1*, *MPHx*, *AGT1* y *MTT1*, están involucrados en el fenotipo diferencial de consumo de maltosa y maltotriosa en dos cepas de *Spastorianus* y que las variaciones en la estructura de proteínas transportadoras afectan aminoácidos importantes en el transporte de maltosa y maltotriosa.

### **7.2 Objetivos específicos relacionados con la hipótesis de la sección 6.1**

- 1.- Caracterizar la región reguladora, codificante de los genes *MALx1*, *MPHx*, *AGT1* y *MTT1*.
- 2.- Caracterizar la eficiencia de fermentación de las levaduras de estudio en las mismas condiciones de cultivo.
- 3.- Caracterizar la tasa de transporte de maltosa y maltotriosa en las dos cepas de estudio.
- 4.- Determinar la expresión de los transportadores *MALx1*, *MPHx*, *AGT1* y *MTT1* en condiciones de fermentación lager.

### **7.3 Objetivo general relacionado con la hipótesis de la sección 6.2**

Determinar que la expresión genética global de las levaduras de estudio está afectada por RNA presente en el mosto cervecero.

#### **7.4 Objetivos específicos relacionados con la hipótesis de la sección 6.2**

- 1.- Realizar un experimento de fermentación para cada cepa que permita evaluar el posible efecto de RNA presente en el mosto cervecero.
- 2.- Obtener el RNA total de las levaduras para su posterior envío a una plataforma de secuenciación de nueva generación.
- 3.- Realizar un preprocesamiento de las lecturas de secuenciación.
- 4.- Obtener los genes diferencialmente expresados en las condiciones de fermentación diseñadas para cada cepa.
- 5.- Analizar los genes diferencialmente expresados y asociarlos a un proceso biológico de interés.

**8. CAPÍTULO 1 Patrones de expresión de los genes Mal y su asociación con la tasa de transporte de maltosa y maltotriosa de dos levaduras *Saccharomyces pastorianus***

Este trabajo fue aceptado para su publicación en la revista **Applied and Environmental Microbiology** de la editorial **American Society of Microbiology** (ASM) bajo el título:

**“Expression patterns of Mal genes and association with differential maltose and maltotriose transport rate of two *Saccharomyces pastorianus* yeasts”**

Nota: todo el material suplementario mencionado en este capítulo está disponible en la publicación en línea del artículo. Por favor dirigirse a <https://journals.asm.org/doi/10.1128/aem.00397-24> o consultar el DOI siguiente.

**DOI:** 10.1128/aem.00397-24

## I. Resumen

La elaboración de cerveza es un proceso bien conocido que aún enfrenta grandes desafíos, como el consumo total de azúcares presentes en los medios de fermentación. La cerveza estilo lager, una de las más importantes a nivel mundial, se elabora con levadura *Saccharomyces pastorianus* (Sp), que debe fermentar mostos con alto contenido de maltotriosa, pero su consumo representa un problema notable, especialmente entre las cepas Sp que pertenecen al grupo I.

Factores como las condiciones de fermentación, la presencia de transportadores de maltotriosa, la variación del número de copias del transportador y las variaciones en la regulación genética contribuyen a este problema.

Aquí evaluamos los factores que afectan la fermentación en dos cepas de levadura Sp: SpIB1, con absorción limitada de maltotriosa, y SpIB2, conocida por su eficiente transporte de maltotriosa. SpIB2 transportó significativamente más maltosa (28%) y maltotriosa (32%) en comparación con SpIB1.

Además, SpIB2 expresó todos los transportadores MAL (*ScMALx1*, *SeMALx1*, *ScAGT1*, *SeAGT1*, *MTT1* y *MPHx*) el primer día de fermentación, mientras que SpIB1 solo exhibió los genes *ScMax1*, *ScAGT1* y *MPH2/3*. Algunos transportadores SpIB2 mostraban dominios transmembrana polimórficos (TMD) que se asemejaban a *MTT1*, acompañados de una mayor expresión junto con sus genes reguladores positivos, como *MAL63*. Esto último siendo un factor de gran importancia para el estudio de este problema.

Estos hallazgos sugieren que, además de los factores mencionados anteriormente, los reguladores positivos de los transportadores Mal contribuyen significativamente a la diversidad fenotípica en el consumo de maltosa y maltotriosa entre las cepas de levadura lager estudiadas.

## II. Introducción

La cerveza es una de las bebidas fermentadas más popular a nivel mundial (Baiano, 2021), con una producción que alcanzó los 189 millones kilolitros en 2022 ([https://www.statista.com/statistics/270275/worldwide-beer-production/#:~:text=In%202022%2C%20the%20global%20beer,the%20United%20States%20and%20Brazil](https://www.statista.com/statistics/270275/worldwide-beer-production/#:~:text=In%202022%2C%20the%20global%20beer,the%20United%20States%20and%20Brazil).)).

El papel de los microorganismos en la biotecnología de alimentos fermentados, depende de la conversión enzimática de sustratos crudos en productos nutritivos y comestibles para el ser humano, especialmente en la elaboración de cerveza (Dimidi *et al.*, 2019; Tamang *et al.*, 2016).

En este sentido, las levaduras son cruciales en la elaboración de cerveza; *Saccharomyces cerevisiae* (Sc) se utiliza para la cerveza estilo ale, con temperaturas de fermentación que oscilan entre 16 °C y 26 °C. Mientras que *Saccharomyces pastorianus* (Sp) es utilizada en la cerveza estilo lager y fermentan entre 5 °C y 15°C (Bonatto, 2021; Petruzzi *et al.*, 2016).

Sp, es un híbrido natural entre especies de Sc y una levadura criotolerante, *Saccharomyces eubayanus* (Se) (Baker *et al.*, 2015; Libkind *et al.*, 2011; Peris *et al.*, 2016), que se divide en dos grupos metabólica y genéticamente distintos; el tipo Saaz (grupo I) y tipo Frohberg (grupo II) (Gibson *et al.*, 2013).

Dos hipótesis explican el origen de la levadura lager: **(i)** la existencia de dos eventos de hibridación independientes que involucran diferentes levaduras, Sc y Se (Baker *et al.*, 2015; Gallone *et al.*, 2018; Monerawela *et al.*, 2015); **(ii)** ambos grupos de levaduras lager comparten un único evento de hibridación, y la evolución resultante de esa hibridación dio forma a las divergencias genómicas entre ambos grupos (Salazar *et al.*, 2019).

La prevalencia de Sp en la industria cervecera implica que este híbrido posee una ventaja selectiva sobre sus cepas originales. Las características clave que contribuyen a esta ventaja incluyen una mejor aptitud biológica en condiciones de estrés y un consumo eficiente de maltotriosa (Gallone *et al.*, 2016, 2018; Herbst *et al.*, 2017).

Pero, una de las principales diferencias fisiológicas entre las levaduras lager es la absorción limitada de maltotriosa en algunos miembros del grupo I (Gibson *et al.*, 2013). El transporte de maltosa y maltotriosa es un “cuello de botella” para lograr una buena fermentación (Meneses & Jiranek, 2002), ya que representan entre el 60% y el 50% de los azúcares fermentables en un mosto cervecero estándar, respectivamente (He *et al.*, 2014; Stambuk *et al.*, 2006).

A pesar de su abundancia, el transporte de maltosa y maltotriosa está regulado por un mecanismo de represión por catabolitos de carbono, donde se prioriza la glucosa sobre la maltosa y la maltotriosa (Gancedo, 1998; Horák, 2013; Kayikci y Nielsen. 2015). En consecuencia, la mayoría de las levaduras cerveceras consumen más fácilmente la maltosa, lo que hace que la maltotriosa prevalezca más en las etapas posteriores de la fermentación, donde algunas cepas de levadura presentan una deficiencia en su absorción (Meneses & Jiranek, 2002; Monerawela *et al.*, 2015; Yoon *et al.*, 2003; Zastrow *et al.*, 2000).

En el fenotipo de consumo deficiente de maltotriosa intervienen diferentes factores los cuales incluyen: **(i)** las condiciones estresantes en los últimos días de la fermentación (Day *et al.*, 2002); **(ii)** la presencia o ausencia de transportadores específicos de maltotriosa, como *AGT1* y *MTT1/MTY1* (Dietvorst *et al.*, 2005; Magalhães *et al.*, 2016; Salema-Oom *et al.*, 2005); y **(iii)** diferencias regulatorias en genes transportadores de maltosa y maltotriosa (Dietvorst *et al.*, 2010; Vidgren *et al.*, 2011).

Además, la aparición de genes quiméricos, resultantes de la recombinación genómica, también participa en el mecanismo de generación de variación en levaduras estrechamente relacionadas, lo que lleva al desarrollo de transportadores de maltotriosa altamente eficientes (Baker y Hittinger, 2019; Brouwers *et al.*, 2019).

Como se indicó anteriormente, el consumo de alfa glucósido varía entre la levadura lager y su estudio carece de una imagen global como en el análisis a nivel transcriptómico. Aquí, investigamos el sistema de transporte de azúcar en dos levaduras lager, específicamente SpIB1, caracterizada por una absorción limitada de maltotriosa, y SpIB2, que se distingue por su eficaz transporte de maltotriosa.

Investigamos si las condiciones de fermentación y/o las diferencias en la arquitectura genética de los genes MAL están asociadas con esta diversidad fenotípica del consumo de maltosa y maltotriosa. Mediante el análisis integrativo del transcriptoma, las detecciones cualitativas de genes y transcritos MAL, la determinación bioquímica de la tasa de transporte de maltosa y maltotriosa aproximamos eventos moleculares, como regiones polimórficas en los transportadores de maltosa y maltotriosa, y los patrones diferenciales de expresión del gen MAL que están asociados con los fenómenos subyacentes involucrados en la diversidad fenotípica de la utilización del azúcar en dos levaduras lager.

Hasta donde sabemos, este trabajo muestra el primer enfoque de datos de RNA-seq para obtener información sobre el consumo de azúcar en dos levaduras lager, lo que nos permitió detectar los genes reguladores MAL como una característica clave en esta diferencia.

### **III. Materiales y métodos**

#### **a) Cepas.**

Las cepas SpIB1 y SpIB2, que representan los grupos I y II, se obtuvieron de Cervecería Cuauhtémoc Moctezuma (Monterrey, NL, México). Las cepas de levadura se almacenaron en medio YP-M (1% extracto de levadura, 2% peptona y 2% maltosa) con 40% glicerol a -20°C después de 48 h o al alcanzar la fase estacionaria para uso futuro.

#### **b) Condiciones y medios de cultivo.**

Ambas cepas fueron inoculadas en medio YPM y mosto cervecero con la siguiente composición: 138 ppm de nitrógeno amino libre, 0,11% (p/p) de fructosa, 4,85% (p/p) de glucosa, 11% (p/p) maltosa y 3% (p/p) de maltotriosa. La fermentación se desarrolló durante 5 días a 16°C y el experimento se realizó por triplicado.

Los muestreos fueron realizados en el primer, segundo y quinto día, elegidos por su significativa expresión fenotípica en la producción de dióxido de carbono y el fenotipo de floculación.

### **c) Inspección y comparaciones por pares de permeasas presentes en los datos genómicos y transcriptómicos de las levaduras estudiadas.**

Utilizamos los datos genómicos obtenidos previamente (León-Medina *et al.*, 2016) para recopilar información sobre los transportadores de maltosa y maltotriosa dentro de las cepas SpIB1 y SpIB2. Al examinar los archivos de anotaciones del genoma, identificamos genes transportadores en cada cepa.

Empleando la función `getfasta` de `bedtools` versión 2.30.0 (Quinlan & Hall, 2010), la utilidad `grep` UNIX y el paquete `ggplot2` R, generamos gráficos de genes e información de los loci.

### **d) Análisis de la estructura de la permeasa.**

Fue elaborado un análisis estructural mediante alineamiento múltiple de las proteínas transportadoras para obtener los aminoácidos importantes en el mecanismo de translocación de la maltosa (Trichez *et al.*, 2019). El análisis de conservación de secuencia de aminoácidos se realizó con el paquete `msaver` versión 1.32.0 (Bodenhofer *et al.*, 2015), junto con el algoritmo CLUSTAL-W y matriz de sustitución BLOSUM80.

Los criterios de inclusión para la secuencia de aminoácidos de las permeasas implicaron secuencias completas y/o predicciones completas del dominio transmembrana. La distribución filogenética de cada permeasa se obtuvo mediante el método de máxima verosimilitud (del inglés Maximum

Likelihood) con los paquetes ape versión 5.7-1 (Paradis *et al.*, 2004) y phangorn versión 2.11.1, con la implementación de un espacio muestral Bootstrap de 1000 réplicas (Schliep. 2011).

#### **e) Análisis de la región reguladora de los genes Mal.**

El software YEASTRACT fue empleado para identificar supuestos sitios de unión de los factores de transcripción *MAL63* y *MIG1* para cada región reguladora río arriba de los genes Mal (Monteiro *et al.*, 2020).

Basado en un modelo científico previo (Dietvorst *et al.*, 2010; Vidgren *et al.*, 2011), en ausencia de sitios de unión al activador *MAL63* de los genes MAL, los transportadores muestran niveles de transcripción reducidos para los genes indicadores o un bajo crecimiento en medios de cultivo con maltotriosa como única fuente de carbono y viceversa.

#### **f) Valores de FPKM de los genes mal encontrados en cada cepa.**

A partir de los datos transcriptómicos generados previamente (Hernández, 2020), fueron obtenidos los valores de FPKM para los transportadores *MALx1*, *AGT1*, *MTT1* y *MPH2/3*. Los datos transcripcionales fueron tomados del segundo y quinto día de fermentación.

Debido a la naturaleza híbrida de la levadura, los posibles eventos de recombinación específicos y la ausencia de un genoma de referencia confiable, se desarrolló un ensamblaje *de novo*. Las lecturas transcriptómicas pareadas para ambas cepas de levadura se procesaron con Trimmomatic versión 0.38 (parámetros LEADING:5 TRAILING:5 MINLEN:50) (Bolger *et al.*, 2014) y Sortmerna para la eliminación de ARN contaminante (Kopylova *et al.*, 2012).

Posteriormente, las lecturas forward y reverse para cada cepa fueron concatenada de forma independiente y ensambladas utilizando la herramienta Trinity versión 2.14.0 con parámetros predeterminados (Haas *et al.*, 2013; Raychowdhury *et al.*, 2011). Los transcriptomas ensamblados fueron anotados utilizando la herramienta Trinotate versión 3.1.0 (Bryant *et al.*, 2017).

Para mapear lecturas de alta calidad contra el transcriptoma ensamblado *de novo* de cada cepa, fue empleada la herramienta Bowtie2 versión 2.4.4 con parámetros predeterminados (Langmead & Salzberg, 2012). Para determinar la abundancia de cada transcrito el software RSEM versión 1.3.3 fue empleado con valores predeterminados (Li y Dewey, 2011).

En particular, en el contexto del resultado del ensamblaje Trinity, se utiliza el término "isoforma", dejando el significado biológico relacionado con el empalme alternativo. Además, en este estudio se enfatiza el uso inequívoco de la nomenclatura genética *AGT1/MAL11* y *MTT1/MTY1*.

#### **g) Análisis de expresión global.**

Un análisis diferencial de expresión génica para fue elaborado para ambas cepas de levadura utilizando muestras recolectadas en el segundo y quinto día de fermentación.

Para evaluar las diferencias estadísticas en la expresión de los genes *Mal*, empleamos el alineamiento de las lecturas con el genoma de referencia de *Sp CBS1483* (genoma y anotaciones disponibles en NCBI bajo el bioproject prjna522669). El alineamiento de lecturas fue realizado con el programa Star versión 2.7.3a con parámetros predeterminados (Dobin *et al.*, 2013). Htseq-count versión 0.11.1 (Anders *et al.*, 2015) se utilizó para la obtención de archivos de conteo a partir de los archivos de alineamiento *bam*. Para el análisis de expresión global el software edgeR versión 3.34.0 (Robinson *et al.*, 2010) fue utilizado.

Se filtraron genes con niveles de expresión bajos ( $\text{cpm} > 5$ ), contrastando los grupos en el segundo día de fermentación (f2) para ambas cepas (grupo de contraste f2\_820-f2\_790), y de manera similar en el quinto día de fermentación (f5) (grupo de contraste f5\_820-f5\_790).

Consideramos que un gen se expresaba significativamente si mostraba al menos una abundancia doble en cada cepa ( $\log_2(2) = 1$ ) con una tasa de descubrimiento de falsos positivos (FDR) de  $< 0,05$ . Para efectos prácticos de la

investigación, solamente nos centramos en los genes Mal de la lista de genes expresados diferencialmente.

#### **h) Prueba de tasa de transporte celular.**

La tasa de transporte de pNP- $\alpha$ -glucosa (estructuralmente relacionada con la maltosa) y pNP- $\alpha$ -maltosa (estructuralmente relacionada con la maltotriosa) fue medida como se describió anteriormente (Hollatz y Stambuk, 2001).

En breve: las cepas SpIB2 y SpIB1 fueron suspendidas a 16 gL<sup>-1</sup> en succinato-Tris 50 mM, pH 5,0 a 30 °C durante 5 min; añadimos pNP-glucosa o pNP-maltosa (40 mM en agua) a concentración final de 5 mM e incubados. Muestras de 100  $\mu$ L fueron retiradas cada min y transferidas a baño de agua hirviendo durante 3 min.

Después de enfriar a temperatura ambiente, agregamos un volumen de NaHCO<sub>3</sub> 2 M, las células fueron eliminadas por centrifugación y se midió el p-nitrofenol presente en el sobrenadante a 400 nm.

Los experimentos fueron hechos por triplicado y controles de células previamente hervidas fueron implementados. La prueba t de Student para las comparaciones de las pendientes de transporte fue elaborada en R versión 4.2.3

#### **i) Extracción de ADN y ARN.**

La purificación del ADN fue llevada a cabo como se mencionó anteriormente (Amer *et al.*, 2011). Brevemente: las cepas SpIB1 y SpIB2 fueron inoculadas, por separado, en medio YPM e incubadas 48 h.

Las células fueron lavadas y 2 ml de células fueron recuperadas a 3000 rpm; agregamos 400  $\mu$ L de solución de extracción estéril e incubamos 10 min a 65°C. Posteriormente, añadimos 130  $\mu$ l de acetato de sodio 3 M, pH 5,2, a cada muestra, se agitó durante 30 seg e incubó a -20 °C.

Finalmente, el lisado fue centrifugado a 13000 rpm a 4°C durante 15 min; el sobrenadante fue sometido a precipitación con isopropanol. Para la extracción de ARN se realizó el protocolo de fenol caliente (Green y Sambrook, 2021), con

la única diferencia del uso de solución AES (acetato de sodio 50 mM, EDTA 10 mM y SDS al 0,5%) en lugar de la solución TES propuesta.

Brevemente, muestras de 10 ml del primer y segundo día de fermentación fueron centrifugadas durante 5 minutos a 3000 rpm a 4°C; los sedimentos celulares fueron lavadas con 5 ml de agua fría estéril y se centrifugaron a 3000 rpm a 4°C durante 5 min; el sobrenadante fue descartado y el sedimento celular congelado en nitrógeno líquido y almacenado a -70 °C hasta su uso.

La muestra fue colocada en hielo (para descongelar) y resuspendida en 400 µl de solución AE; 400 µL de fenol ácido fue agregado seguido de 30 s de agitación; las muestras de extracción fueron incubadas durante 60 minutos a 65°C y una centrifugación de 5 minutos a 13000 rpm a 4°C fue empleada. El sobrenadante fue colocado en un nuevo tubo de microcentrífuga estéril y un segundo procedimiento de extracción con fenol ácido caliente fue realizado.

El sobrenadante resultante fue tratado con un volumen de cloroformo y centrifugado durante 5 minutos a 13000 rpm a 4°C; el sobrenadante resultante fue sujeto a precipitación con dos volúmenes de isopropanol enfriado con hielo y 40 µl de acetato de sodio 3 M a pH 5,3, seguido de una incubación en refrigeración durante 60 min.

Las muestras resultantes fueron centrifugadas durante 10 minutos a 13000 rpm a 4 °C y el sedimento de ARN total se lavó con etanol al 70 % enfriado con hielo en agua estéril tratada con pirocarbonato de dietilo (DPCE). Finalmente, el sedimento lavado resultante fue resuspendido en 50 µl de agua tratada con DPCE. La calidad de la extracción de ARN fue evaluada mediante electroforesis en gel de agarosa, y las concentraciones determinadas en un equipo tipo nanodrop.

**j) Detección molecular de permeasas en cada cepa.**

Para la detección genómica de los genes que codifican las permeasas *MALx1*, *AGT1*, *MTT1* y *MPH2/3* en ambas levaduras, los cebadores y experimentos de

PCR como se describió anteriormente fueron empleados (Magalhães *et al.*, 2016).

En breve: el siguiente programa de PCR fue implementado: preincubación (95°C durante 5 min), ciclo de amplificación repetido 35 veces (95°C durante 10 s,  $T_m$  °C para cada cebador específico durante 10 s, 72°C durante 10 s). Para la detección de los transcritos mensajeros, un experimento de PCR con transcripción inversa (RT PCR) para detectar las transcripciones de *Malx1*, *AGT1*, *MTT1* y *MPH2/3* fue ejecutado.

Para esto, fue utilizada una transcriptasa inversa comercial (M-MLV Reverse Transcriptase, Promega), siguiendo las instrucciones del fabricante y utilizando las mismas condiciones de PCR mencionadas anteriormente (Magalhães *et al.*, 2016). Los cebadores para *ACT1*, *TAF10* y *TFC1* fueron empleados como genes de control y/o referencia (Teste *et al.*, 2009).

#### **k) Análisis del consumo de mosto y alfa glucósidos.**

Las muestras de mosto fueron analizadas mediante HPLC Serie 1200 de Agilent Technologies utilizando un detector de infrarrojos (IR). La fase móvil fue H<sub>2</sub>SO<sub>4</sub> (5 mM). El caudal se estableció en 1 ml/min utilizando una columna de carbohidratos Aminex HPX-87C Hi-Plex Na, 300 × 7,7 mm, y la columna fue equilibrada con H<sub>2</sub>SO<sub>4</sub> 5 mM en agua a 55 °C (Magalhães *et al.*, 2016).

Los experimentos fueron realizados por triplicado y la prueba t de Student fue empleada para determinar la significancia estadística en la diferencia entre las medias de transporte de maltosa y maltotriosa.

## **IV Resultados**

### **Abundancia de las permeasas en las levaduras estudiadas.**

Las posibles diferencias en el número de copias entre las permeasas, fueron determinadas con datos genómicos para identificar los loci asociados con los genes transportadores *MAL31*, *AGT1*, *MPH2/3* y *MTT1*, en ambas levaduras.

Aunque una variación del número de copias es congruente entre ambas cepas para los genes *AGT1*, *MPH2/3*. Los genes *MTT1*, *MAL31* mostraron diferencias exhibiendo cuatro loci para la cepa SpIB2 y seis loci para la cepa SpIB1 (Tabla 3; Fig. S1). A pesar de que ambas cepas tenían el transportador *MTT1* en su genoma, la secuencia de nucleótidos estaba incompleta en comparación con la longitud de referencia (Dietvorst et al., 2005; Salema-Oom et al., 2005).

La longitud de los nucleótidos y aminoácidos en las permeasas varió significativamente (Tabla 3), y solo las isoformas específicas de *AGT1* y *MAL31* estaban completas en cada genoma. Además, la distribución estructural de los loci MAL difirió del patrón canónico, que se observó solo en el scaffold15 en SpIB2 (Fig. S1) (Monerawela et al., 2015; Monerawela y Bond, 2018).

Los datos transcriptómicos confirmaron la integridad de las permeasas, mostrando que ambas cepas contienen el transportador *MAL31* completo en cada transcriptoma, mientras que la presencia del transportador *MTT1* parecía estar ausente. Para resolver estos resultados contradictorios, se realizaron análisis sobre los alineamientos múltiples para determinar la distribución filogenética de estos transportadores utilizando un algoritmo de máxima verosimilitud (del inglés Maximum likelihood).

### **Análisis estructural, niveles de expresión y distribución filogenética de las permeasas.**

Para detectar variaciones específicas en la estructura, el nivel de expresión y las relaciones filogenéticas de los transportadores de maltosa y maltotriosa que podrían aproximar la eficiencia diferencial del consumo de azúcar, realizamos un alineamiento múltiple de secuencias aminoacídicas para detectar regiones polimórficas relacionadas con los mecanismos de translocación de maltosa y maltotriosa.

**Tabla 3.** Localización y tamaño de las permeasas para maltosa y maltotriosa presentes en las levaduras de estudio.

<b>SpIB2 permeasas</b>						
<b>Nombre del gen</b>	<b>Nombre sistemático</b>	<b>Locus (número de scaffold)</b>	<b>Coordenadas</b>	<b>Cadena</b>	<b>Tamaño en genoma (nt)</b>	<b>Tamaño en transcriptoma (aa)</b>
<i>AGT1</i>	YGR289C	Scaffold3	4536 - 6368	+	1851/610	558
<i>AGT1</i>	YGR289C	Scaffold50	71200 - 73051	-	1185/394	
<i>MAL31</i>	YBR298C	Scaffold102	3 - 581	-	581/193	615
<i>MAL31</i>	YBR298C	Scaffold6	612004 - 613788	-	1784/594	
<i>MAL31</i>	YBR298C	Scaffold11	492091 - 492738	-	2165/721	
<i>MAL31</i>	YBR298C	Scaffold15	20813 - 21508	+	1793/597	
<i>MPH3</i>	YJR160C	Scaffold29	946600 - 948408	-	1808/602	374
<i>MTY1</i>	MTY1	Scaffold13	658031 - 659338	-	1307/435	No detectada
<b>Sp820 permeases</b>						
<i>AGT1</i>	YGR289C	Scaffold11	3868 - 5700	+	1833/611	610
<i>AGT1</i>	YGR289C	Scaffold23	4299 - 6150	+	1185/394	
<i>MAL31</i>	YBR298C	Scaffold8	14887 - 15582	+	1825/608	614
<i>MAL31</i>	YBR298C	Scaffold10	282142 - 283926	-	1784/594	
<i>MAL31</i>	YBR298C	Scaffold26	9677 - 11624	+	1947/614	
<i>MAL31</i>	YBR298C	Scaffold30	645561 - 645869	-	309/103	
<i>MAL31</i>	YBR298C	scaffold82	1 - 792	+	792/264	
<i>MAL31</i>	YBR298C	scaffold85	308 - 757	+	448/149	
<i>MPH3</i>	YJR160C	Scaffold24	556421 - 558229	-	1809/602	462
<i>MTY1</i>	MTY1	Scaffold64	204 - 1553	-	1349/449	No detectada

Además, realizamos análisis transcriptómicos del segundo y quinto día de fermentación para cada cepa de levadura con el fin de detectar diferencias que podrían explicar la eficiencia del transporte de maltosa y maltotriosa. Finalmente, llevamos a cabo una filogenia de máxima verosimilitud para estimar las

relaciones estructurales/funcionales entre los transportadores de maltosa y maltotriosa de ambas cepas.

Detectamos que los transportadores *AGT1* y *MAL31* de ambas cepas mostraron 12 dominios transmembrana (TMD1-12), lo cual es característico de los transportadores MSF (Drew *et al.*, 2021; Madej *et al.*, 2014; Yan, 2013). A pesar de su secuencia parcial (558aa), el transportador SpIB2 *AGT1* mostró el conjunto completo de dominios transmembrana típicos de los miembros de la familia MSF.

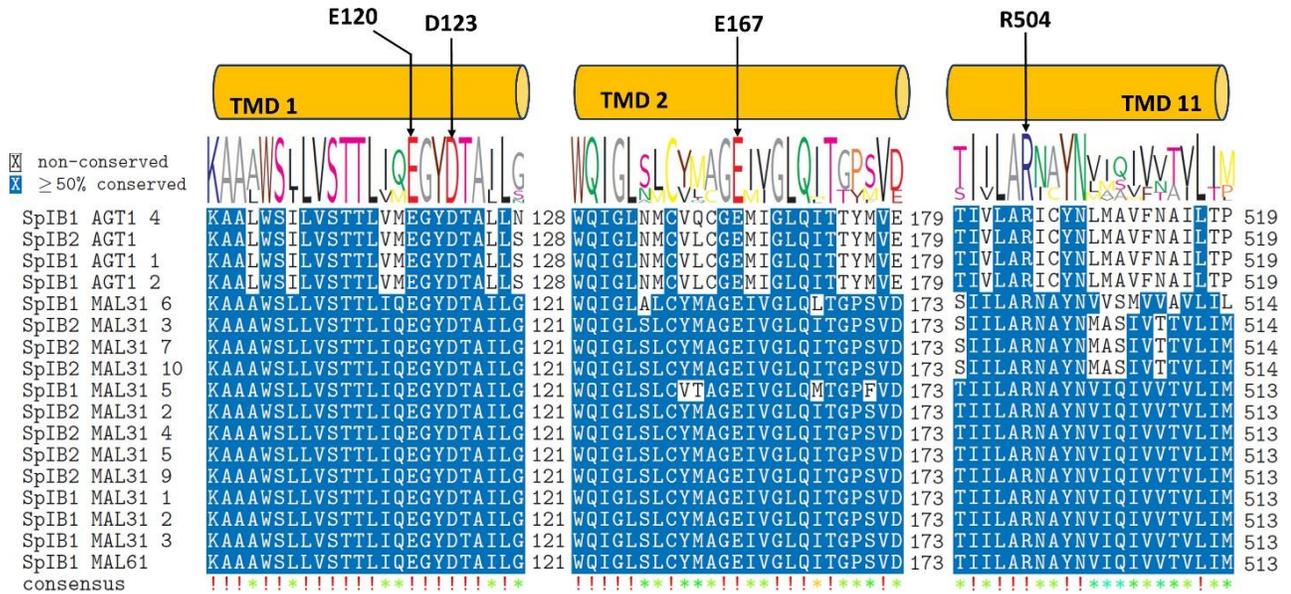
A pesar de ello, algunas transcritos asociados con los genes *MPH2/3* y *MAL31* en ambas cepas carecían de predicciones del dominio transmembrana, mientras que otras exhibieron deleciones que afectaron la estructura de TMD (Fig. S2 y S3). Para las permeasas restantes, la predicción de la estructura 3D coincide con los transportadores de alfa-glucósido, lo que muestra un alto valor de confianza entre la estructura TMD y las distribuciones de error alineadas predichas son similares entre ellas (Fig. S2 y S3).

Teniendo en cuenta estos resultados, aquellas proteínas que mostraron deleciones o falta de dominios transmembrana fueron excluidos de los alineamientos múltiples y de los análisis filogenéticos (Philippe *et al.*, 2011).

Con respecto al patrón de conservación, los dominios TMD1 (110–129 aa), TMD2 (157–178 aa), TMD7 (371–399), TMD11 (501–524 aa) y TMD12 (532–550 aa) mostraron >50% de conservación. En cierta medida, también observamos algunas variaciones en las regiones específicas de *MAL31* y *AGT1* evidenciando una región polimórfica que depende del transportador Mal (Fig. 8; Fig. S4).

Los patrones de conservación de aminoácidos de las permeasas revelaron una presencia constante ( $\geq 50\%$ ) de E120, D123, E167 y R505 en sus respectivos TMD (Fig. 8). Como se muestra en la Fig. S4 (revisar en el artículo publicado en línea), la asparagina (N) se conserva  $\geq 50\%$  en todas las secuencias de TMD11, excepto las isoformas *MAL11/AGT1* específicas presentes en SpIB2 y SpIB1, donde hay isoleucina (I). A su vez, una alanina (A) se conserva  $\geq 50\%$  en las secuencias TMD12, pero no para SpIB2 y SpIB1 *MAL11/AGT1*, donde se observó

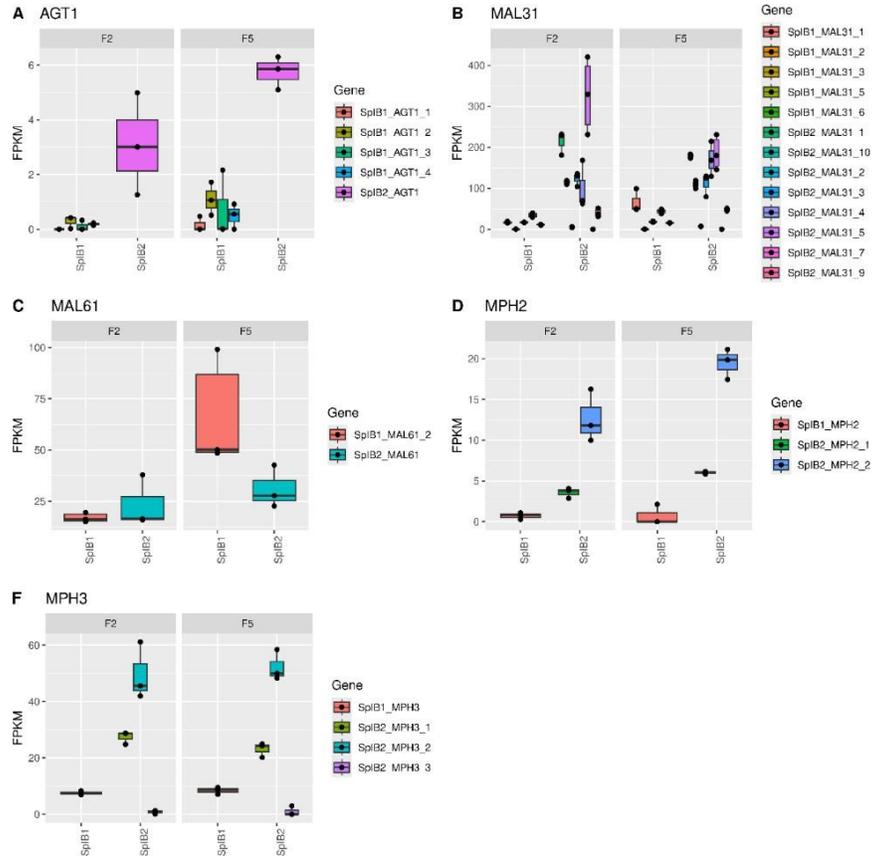
treonina (T557) en la isoforma SpIB2, mientras que las isoformas de SpIB1 tenían A379 e Y384 (Fig. S5).



**Figura 8.** Secciones del alineamiento múltiple de proteínas transportadoras de maltosa y maltotriosa que muestran los aminoácidos cargados involucrados en el mecanismo de translocación. Solo se muestran los dominios transmembranales DTM-1, DTM-2 y DTM-11.

En el análisis del nivel de expresión, SpIB2 exhibe mayor actividad transcripcional para la mayoría de los transportadores, excepto *MAL61*, que muestra mayor expresión en la cepa SpIB1. Específicamente, la isoforma *AGT1* de SpIB2 demuestra una mediana más alta de fragmentos por kilo base de transcrito por millón de lecturas mapeadas (FPKM), 2,6 y 6,08, en el segundo y quinto día de fermentación, en comparación con la isoforma de SpIB1 con valores de FPKM de 0,22 y 1,39 en los mismos días (Fig. 9A).

Para los genes *MAL31*, la isoforma SpIB2 muestra los valores de expresión más altos en el segundo y quinto día (329,26 y 180,08 FPKM), en contraste con la isoforma SpIB1 que muestra 34,52 y 50,12 FPKM en los días respectivos (Fig. 9B). Estos resultados son consistentes con el análisis de la región reguladora donde, en general, la levadura SpIB2 tiene más sitios de unión funcionales para las proteínas reguladoras *MAL63* (Tablas S2 y S3).



**Figura 9.** Niveles de expresión de los transportadores de maltosa y maltotriosa mostrados en FPKM. Los datos son representados como gráfico de caja conteniendo los valores FPKM, de tres experimentos independientes, para cada gen MAL encontrado en cada cepa. Nótese que en los genes **A)** AGT1, **B)** MAL31, **D)** MPH2 y **F)** MPH3 la cepa SpIB2 presenta mayor nivel de expresión que la cepa SpIB1. La cepa SpIB1 solo tiene mayor expresión en el gene **F)** MAL61 al quinto día de fermentación.

Con respecto a los transportadores *MAL61*, la isoforma SpIB1 exhibe la expresión más alta (67,27 FPKM) en el quinto día de fermentación (Fig. 9C). La actividad transcripcional de *MPH2* y *MPH3* es mayor en SpIB2 que en SpIB1 (Fig. 8D y F). Empero, ambas cepas muestran isoformas con deleciones que afectan a numerosos dominios transmembrana (Fig. S2).

Para obtener la relación entre los transportadores MAL de SpIB2, SpIB1 y los reportados para Sc, Se y Sp, se estimó un árbol filogenético de máxima verosimilitud (del inglés Maximum Likelihood). La Figura 10 mostró la presencia de tres clados principales observados previamente (Baker y Hittinger. 2019;

Brown *et al.*, 2010), donde las proteínas *AGT1* muy divergentes formaron su propio grupo.

Aquí, observamos que las proteínas *MAL11* de *SpIB2* y *SpIB1* se agruparon en este clado, lo que sugiere una función similar. Los grupos *MPH3* y *MPH2* están constituidos por dos proteínas de *Sc* y una proteína de *Sp*. No obstante, no incluimos las proteínas de *SpIB2* y *SpIB1* ya que observamos eliminaciones en los dominios transmembrana importantes para la translocación de maltosa (Fig. S2).

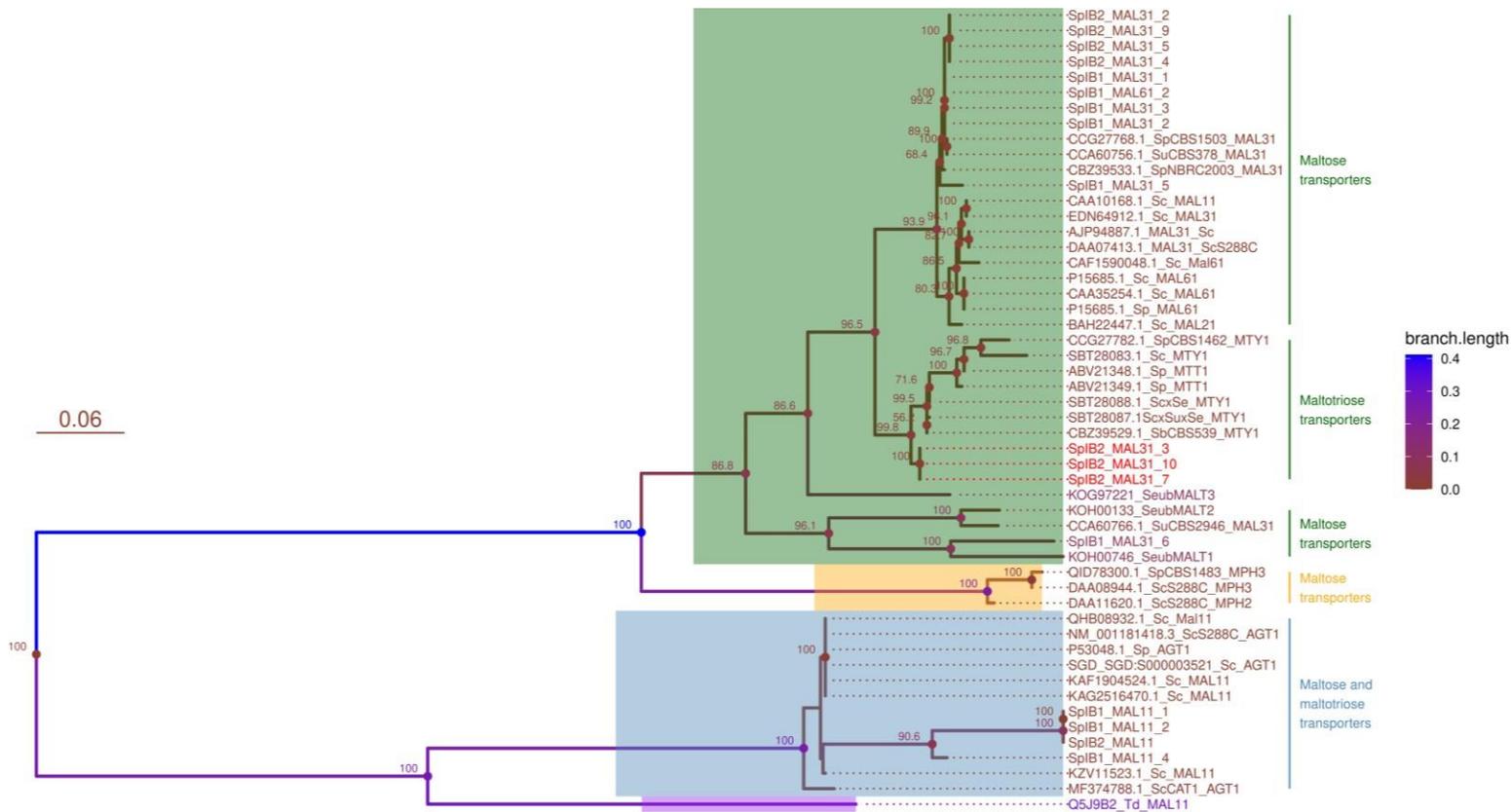
El clado más grande está constituido por transportadores *MALT* de *Se* y *MAL31\_6* de *SpIB1*, *MALx1* de *Sc* y *Sp* y transportadores *MTT1/MTY1* específicos de híbridos lager. Es importante señalar que las proteínas *SpIB2* *MAL31* se agruparon, respaldadas con un valor de bootstrap de 98,6, en el grupo *MTT1*, lo que sugiere una función de transporte de maltotriosa muy relacionada.

Finalmente, incluimos un transportador *MAL11* muy distante de *Torulaspora delbrueckii* para obtener una interpretación más sólida del árbol (Philippe *et al.*, 2011).

### **Análisis de expresión global de genes.**

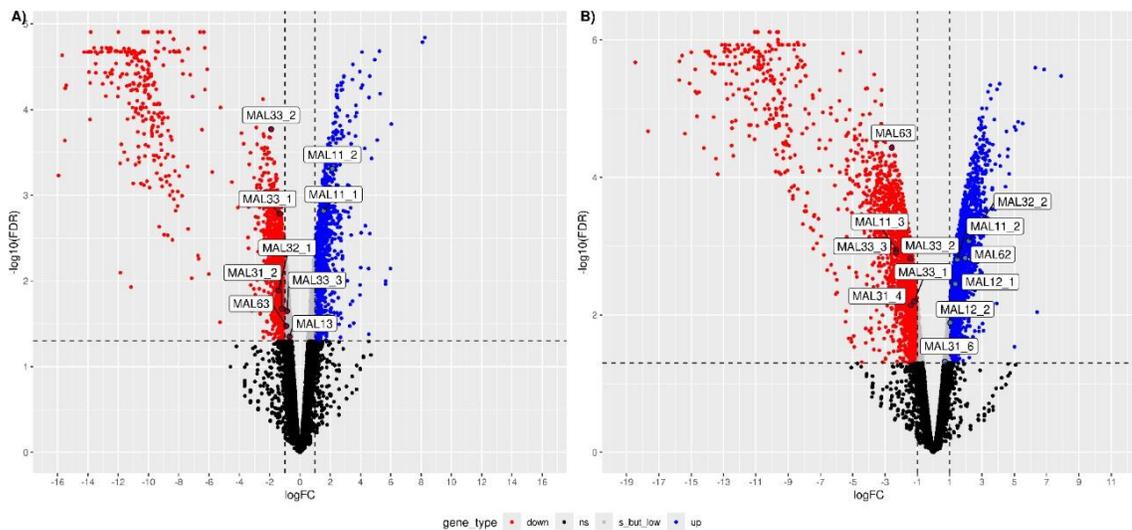
Para determinar la expresión génica afectada solo por el tipo de cepa de levadura y no por las condiciones de fermentación y también para obtener información estadística de la sección de análisis del nivel de expresión (FPKM), decidimos realizar un análisis diferencial de la expresión génica mediante el alineamiento de las lecturas transcriptómicas de *SpIB1* y *SpIB2* contra el genoma de referencia de la levadura *Sp*.

Los loci *MAL* analizados en este trabajo mostraron una identidad del 99% al 100% y una cobertura del 87% al 100%, con regiones cromosómicas del *Sp* CBS1483 (Tabla S1). En términos generales, observamos un efecto considerable según el tipo de cepa. En el segundo día de fermentación, determinamos 1613 genes regulados negativamente y 1215 genes regulados positivamente en *SpIB1*.



**Figura 10.** Filogenia estimada con el algoritmo “Maximum likelihood” mediante un Bootstrap de 1,000 para los transportadores de maltosa y maltotriosa reportados en *S. cerevisiae*, *S. eubayanus*, *S. pastorianus*, SpIB1 y SpIB2. La escala de las barras indica el número de sustituciones aminoacídicas por sitio. Nótese como las proteínas MAL31 de la cepa SpIB2, marcadas en rojo, se agrupan con las permeases *MTT1/MTY1* reportadas para *S. pastorianus*.

En el quinto día de fermentación, observamos 2494 genes regulados negativamente y 2183 genes regulados positivamente en SpIB1. Aquí podemos interpretar a la inversa estos genes expresados diferencialmente (Anders *et al.*, 2013; Robinson *et al.*, 2010), donde los genes regulados negativamente en SpIB1 están regulados positivamente en SpIB2, y los genes regulados positivamente en SpIB1 están regulados negativamente en SpIB2.



**Figura 11.** Gráfico tipo Volcano que expone los genes diferencialmente expresados en A) el segundo y B) quinto día de la fermentación para las cepas SpIB1 y SpIB2. Los genes al alta (*up*) y a la baja (*down*) están representados en azul y rojo, respectivamente (valor absoluto del log<sub>2</sub> fold change >1 y false-discovery rate [FDR] < 0.05). Genes con significancia estadística, pero con bajo nivel de expresión (*s\_but\_low*) (valor absoluto del log<sub>2</sub> fold change <1 and FDR < 0.05) son expuestos en gris y los genes sin significancia estadística (*ns*) en negro (FDR > 0.05).

A partir de estos genes expresados diferencialmente, nos centramos en los genes de loci MAL, lo que dio como resultado la comparación de fermentación del segundo día (F2\_820 frente a F2\_790), con dos genes expresados diferencialmente en SpIB1; *MAL11\_2* y *MAL11\_1* (Fig. 11A).

Por el contrario, la cepa SpIB2 exhibe expresión diferencial en los genes *MAL33\_1*, *MAL33\_2*, *MAL32\_1*, *MAL31\_2*, *MAL33\_3*, *MAL63* y *MAL13* (Fig. 11A; Tabla S4). Estos genes realizan funciones reguladoras positivas y su expresión global se confirma mediante mediciones de FPKM, lo que indica una mayor expresión en SpIB2 (Fig. 9; Fig. S5).

A pesar de ello, es importante tener en cuenta que los genes *MAL63*, *MAL13* y *MAL33\_3* son estadísticamente significativos, pero muestran una expresión baja en SpIB2 (log<sub>2</sub> fold change <1, puntos grises en la Fig. 11A), lo que indica que su abundancia es menos del doble en la cepa SpIB2.

El mismo patrón de expresión se observa en el quinto día de fermentación (F5), donde la cepa SpIB1 muestra expresión de los transportadores de maltosa

y maltotriosa, *MAL11 (AGT1)* y *MAL31*, así como algunas maltasas, como *MAL32* y *MAL62* (Fig. 11B; Tabla S5).

De manera similar, la cepa SpIB2 tiene más genes reguladores positivos, como *MAL63* y *MAL33\_1–3*, y algunos transportadores, como *MAL11\_3* y *MAL31\_4*. Los primeros cuatro genes, que son reguladores positivos, mostraron más del doble de abundancia en SpIB2 en comparación con SpIB1, lo que destaca la utilización significativa de estos genes por parte de SpIB2.

### **Análisis bioquímico de la tasa de transporte y detección molecular de los transportadores de maltosa y maltotriosa.**

Para demostrar si las condiciones de fermentación tienen un efecto significativo sobre el transporte de maltosa y maltotriosa, determinamos las curvas de captación celular para los sustratos *p*-nitrofenil- $\alpha$ -D-glucopiranososa (pNP-glucosa) y *p*-nitrofenil- $\alpha$ -D-glucopiranosil-(1,4)-D-glucopiranosido (pNP-maltosa) para las cepas SpIB2 y SpIB1 en las mismas condiciones de reacción.

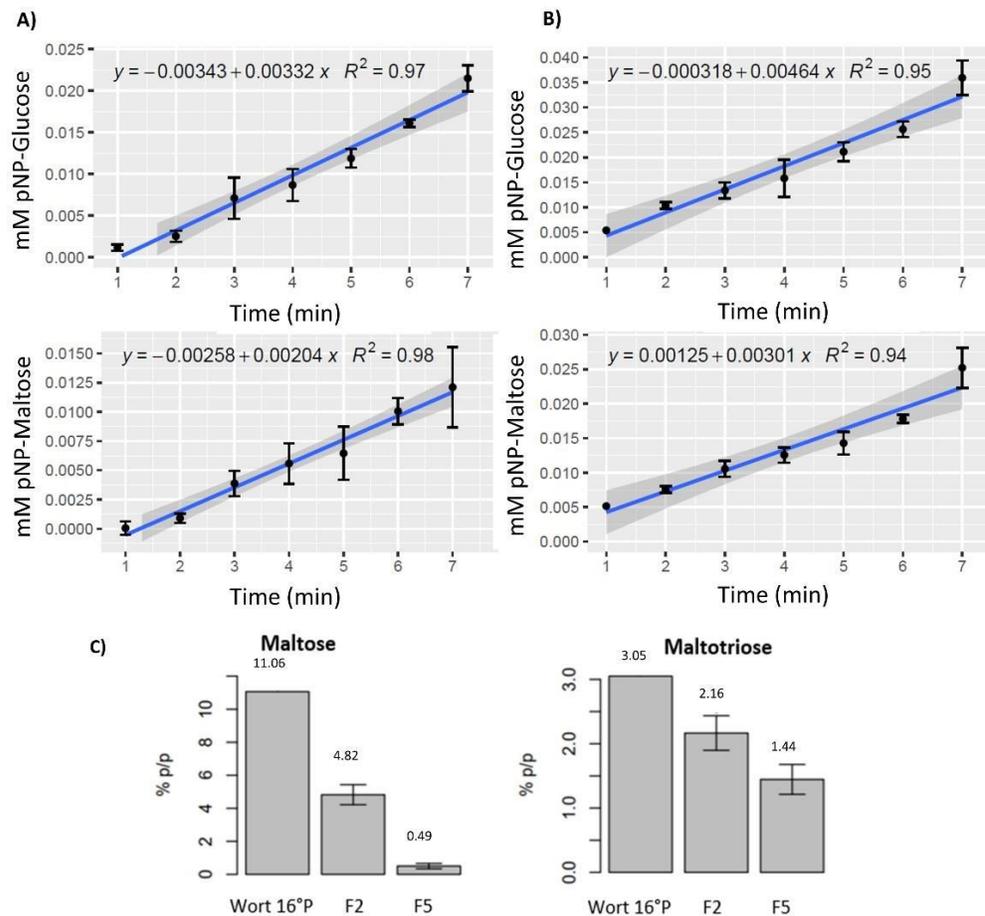
Los fundamentos de este ensayo son que los sustratos de pNP marcados químicamente se transportan al interior de la célula mediante transportadores de alfa-glucósidos (*MALx1*, *AGT1*, *MPHx* y *MTT1*), una vez en el citoplasma de la célula las alfa-amilasas hidrolizan el enlace alfa de la maltosa y la maltotriosa, liberando el pNP. Después de la permeabilización y la extracción de pNP, se realizaron mediciones espectrofotométricas a 400 nm para una determinación cuantitativa (Hollatz & Stambuk, 2001).

Debido a que la relación estequiométrica de pNP al alfa-glucósido de interés es 1:1 (Henderson & Poolman, 2017), se puede extrapolar la concentración de pNP transportado a la concentración de alfa-glucósido marcado. Aunque ambas cepas toman maltosa a un ritmo mayor que la maltotriosa, SpIB2 demuestra tasas de transporte más altas de ambos alfa glucósidos que SpIB1 (Fig. 12A y 12B).

Específicamente, para la maltosa, SpIB2 tiene una tasa de transporte de  $282 \text{ nmol min}^{-1} \text{ mg}^{-1}$  de peso seco, mientras que SpIB1 tiene  $202 \text{ nmol min}^{-1} \text{ mg}^{-1}$  de peso seco (Fig. S6). De manera similar, en el transporte de maltotriosa,

SpIB2 exhibe una tasa más alta de  $183 \text{ nmol min}^{-1} \text{ mg}^{-1}$  de peso seco en comparación con SpIB1 de  $124 \text{ nmol min}^{-1} \text{ mg}^{-1}$  de peso seco.

Estos resultados bioquímicos también reflejan el comportamiento fenotípico en la misma condición de fermentación donde observamos  $0,49 \pm 0,04\%$  p/p y  $1,44 \pm 0,27\%$  P/P de maltosa residual y maltotriosa en el quinto día, respectivamente, en SpIB1, mientras que los azúcares fueron no detectables en SpIB2 (Fig. 12C).



**Figura 12.** Tasa de transporte celular determinada por la toma de pNP-glucosa y pNP-maltosa en un periodo de 7 minutos de reacción en las cepas A) SpIB1 y B) SpIB2. C) Concentración en %p/p de maltosa y maltotriosa en el segundo y quinto día de la fermentación para la cepa SpIB1. Los datos para la cepa SpIB2 no fueron detectables.

Los datos de amplificación genética sugieren que ambas cepas portan los genes *ScMALx1*, *SeMALx1*, *ScAGT1*, *SeAGT1*, *MTT1* y *MPH2/3*, donde se

confirmó la presencia de un gene tipo *MTT1* en SpIB1 (Fig. 13A). La detección de transcritos indica que ambas cepas expresan todas las permeasas analizadas, aunque con un desfase temporal observado en la cepa SpIB1. En particular, SpIB2 exhibe todas las permeasas en el primer día, mientras que SpIB1 solo muestra *ScMax1*, *ScAGT1* y *MPHx*.

En el segundo día de fermentación, observamos *SeMax1*, *ScAGT1*, *MPHx* y *MTT1* en SpIB2 y todas las transcritos para los transportadores analizadas en SpIB1 (Fig. 13B). De estos resultados, destacamos que SpIB2 expresa todos los transportadores MAL analizados en estas condiciones en el primer día de la fermentación, mientras que con un desfase temporal la cepa SpIB1 expresa todos los transportadores MAL, especialmente el *MTT1*. Esto sugiere un probable retraso en la regulación de la transcripción causado por diferencias en el nivel de expresión de las proteínas reguladoras MAL.

En general, estos hallazgos respaldan nuestras inferencias bioinformáticas, destacando la regulación genética como el factor predominante en la variación fenotípica del transporte de maltosa y maltotriosa en las cepas estudiadas.

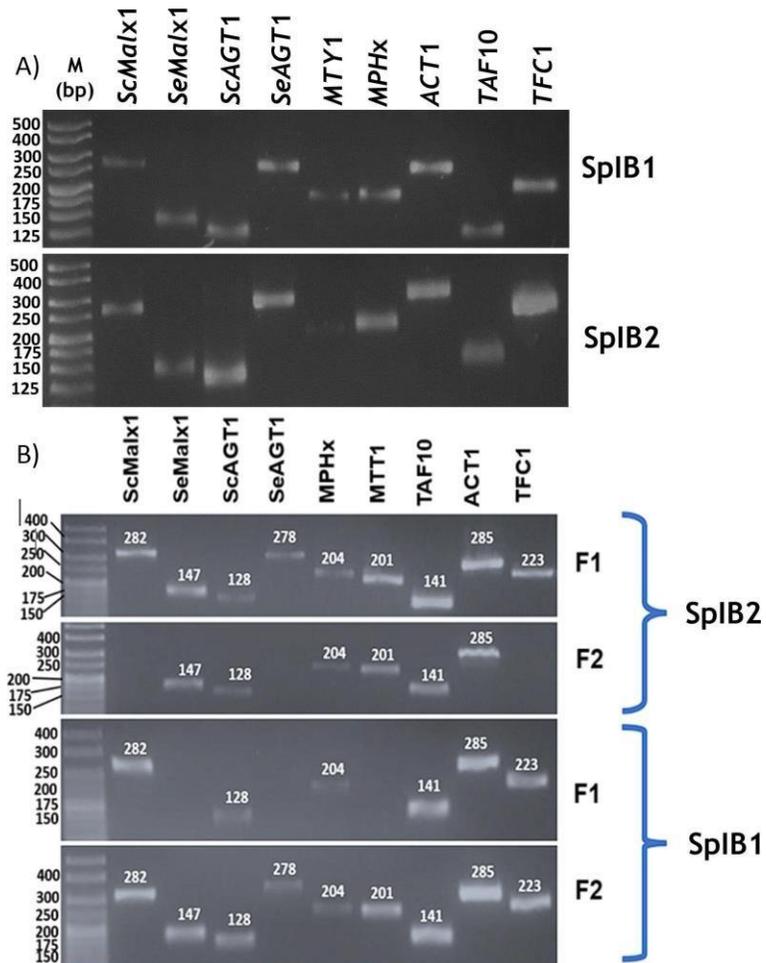
## V Discusión

### **La variabilidad estructural explica la variación fenotípica en el transporte de alfa glucósidos.**

La mutación espontánea es una fuerza impulsora importante detrás de la diversidad de proteínas bajo la selección natural, lo que potencialmente permite una divergencia de ~50% a 80% sin afectar la estructura de las proteínas (Gong y Blundell, 2010; Schaefer y Rost, 2012).

Pese lo cual, los cambios en residuos específicos, particularmente aquellos relacionados con los sitios de unión al sustrato, podrían afectar la función de las proteínas (Gong & Blundell, 2010). Las permeasas Mal analizadas en este estudio muestran un patrón de aminoácidos altamente conservado ( $\geq 50\%$ ) en TMD1, TMD2 y TMD11, que se informaron como relevantes para el

transporte de maltotriosa (Henderson y Poolman, 2017; Trichez *et al.*, 2019). Además, una alanina (A) se conserva  $\geq 50\%$  en las secuencias de TMD12, pero no en SpIB2 y SpIB1 *MAL11/AGT1*, donde se observó treonina (T557) (Smit *et al.*, 2008).



**Figura 13.** Detección molecular de los genes Mal presentes en las cepas SpIB1 y SpIB2. **A)** Detección genómica mediante PCR de los genes Mal *ScMALx1*, *SeMALx1*, *ScAGT1*, *SeAGT1*, *MPHx*, and *MTT1/MTY1* indicando que cada levadura contiene todos los transportadores analizados en este estudio. **B)** Detección de transcritos por RT-PCR en el primer y segundo día de la fermentación de las levaduras estudiadas. Nótese que, en el primer día de la fermentación, la cepa SpIB2 mostró todos los transcritos Mal, mientras que la levadura SpIB1 sólo exhibió *ScMALx1*, *ScAGT1* y *MPHx*.

Con todo esto, es importante mencionar que estos aminoácidos se modelan en el contexto de los transportadores *AGT1* y pueden no representar el panorama

general de la especificidad de la maltotriosa de los transportadores similares a *MTT1/MTY1*. Estudios recientes han informado sobre algunos sitios polimórficos cruciales en TMD7 y TMD11 en *MAL61*, *MTT1* y transportadores similares a *MTT1*, donde la presencia de residuos T379 y N384 en TMD7 están implicados en la especificidad de la maltotriosa (Hatanaka *et al.*, 2022).

Aquí, SpIB2 *MAL31* mostró estas sustituciones de aminoácidos, pero no estuvo presente en ninguna proteína *MAL31* de SpIB1. Este resultado sugiere fuertemente que SpIB2 tiene un eficiente sistema de transporte de maltotriosa presente en su transcriptoma, mientras que SpIB1 tiene un paisaje similar a *AGT1* en su arquitectura de transporte de alfa glucósido.

En general, los cambios de aminoácidos y sintenia encontrados en los loci Mal (Fig. S1) indican que las condiciones de fermentación de la cerveza industrial actúan como una fuerza selectiva, y la naturaleza híbrida concomitante de *Sp* puede causar la plasticidad del transporte de maltosa y maltotriosa observada en estas levaduras (Brown *et al.*, 2010; Hatanaka *et al.*, 2022; Monerawela *et al.*, 2015; Monerawela y Bond, 2018; Salazar *et al.*, 2019).

### **Los patrones de expresión y la identidad de los transportadores explican eficientemente la variación fenotípica del transporte de maltosa y maltotriosa en las cepas analizadas.**

Actualmente, diversas cepas de *Sp* utilizan proteínas *MTT1* y/o *AGT1* para el transporte de maltotriosa (Day *et al.*, 2002; Dietvorst *et al.*, 2005; Magalhães *et al.*, 2016; Salema-Oom *et al.*, 2005; Yoon *et al.*, 2003; Zastrow *et al.*, 2000). El agotamiento del azúcar en el segundo (F2) y quinto (F5) día para las cepas SpIB2 y SpIB1 refleja las diferencias en la regulación genética de los genes MAL. Esta noción está respaldada por los datos transcriptómicos de los genes putativos *MTT1* y *AGT1* en SpIB2, pero solo *AGT1* en SpIB1.

Los resultados transcriptómicos confirmaron que *MAL31* y *MTT1* tienen más valores de expresión (FPKM) en SpIB2 que en SpIB1. Pese a esto, los valores de expresión (FPKM) para *AGT1* fueron bajos en ambas cepas, aún con

su gran importancia reportada previamente (Fig. 9) (Vidgren *et al.*, 2005, 2009, 2011).

Este resultado tiene sentido porque estas levaduras lager tienen *AGT1* en su base genética, lo que sugiere un camino evolutivo relacionado con una menor expresión de *AGT1* y la presencia concomitante de transportadores similares a *MTT1* (Baker & Hittinger, 2019). Los transportadores de maltosa *MAL61* se expresaron más en SpIB1 en el quinto día de fermentación (Fig. 9C), y esto es consistente con su consumo de maltosa (Fig. 12C), que se asemeja al intervalo de maltosa y maltotriosa residual observado previamente (Magalhães *et al.*, 2016).

Finalmente, el nivel de expresión de los genes *MPHx* fue mayor en SpIB2 que en SpIB1. Aun así, los genes *MPHx* son sólo transportadores específicos de maltosa y su función en el transporte de maltotriosa es ambigua (Brown *et al.*, 2010; Magalhães *et al.*, 2016).

Con respecto a las relaciones filogenéticas de los transportadores de maltosa y maltotriosa, nuestro árbol es consistente con informes anteriores (Baker & Hittinger, 2019; Brown *et al.*, 2010). En este sentido, la cepa SpIB2 mostró tres proteínas *MAL31* agrupadas con transportadores *MTT1/MTY1*, lo que evidencia que esta cepa probablemente tiene un sistema de transporte de maltotriosa eficiente, mientras que los genes transportadores de la cepa SpIB1 no exhibieron agrupación con los transportadores *MTT1/MTY1*.

Es importante señalar que SpIB1 *MAL31\_6* se agrupa con transportadores de maltosa de *Se*, asemejándose una función similar a la permeasa de maltosa de esta levadura. Como se observa aquí y en otros estudios, los transportadores de maltosa y maltotriosa *AGT1* son un grupo muy distante del clado *MAL31* más grande, que tiene un subclado de transportadores específicos de maltotriosa *MTY1/MTT1*, lo que sugiere que la evolución de la característica de consumo de maltotriosa tiene redundancia y ocurrencia frecuente en el panorama evolutivo de los transportadores *MAL*.

Contrariamente a la rápida evolución de los loci *MAL*, las proteínas transportadoras específicas de maltotriosa son un evento de baja frecuencia en

la vía de evolución de este rasgo industrial crucial (Baker y Hittinger, 2019; Brouwers *et al.*, 2019; Brown *et al.*, 2010), lo que explica la rica diversidad fenotípica observada en el transporte de alfa glucósido en SpIB2, SpIB1 y otras levaduras industriales importantes (Gibson *et al.*, 2013; Magalhães *et al.*, 2016).

**En las mismas condiciones de fermentación, las levaduras exhiben diferencias en la tasa de transporte de maltosa y maltotriosa.**

Para apoyar las predicciones bioinformáticas relacionadas con la expresión diferencial de transportadores y reguladores MAL y probar el escenario hipotético donde las condiciones de fermentación ejercen un efecto en el consumo de maltosa y maltotriosa de SpIB2 y SpIB1, la tasa de transporte celular de pNP-glucosa (análogo estructural de maltosa) y pNP-maltosa (análogo estructural de la maltotriosa) se calculó en las mismas condiciones de reacción (Hollatz & Stambuk, 2001; Smit *et al.*, 2007; Trichez *et al.*, 2019; Zastrow *et al.*, 2000).

A pesar de que la hipótesis nula es la no diferencia en la velocidad de transporte celular en las mismas condiciones de reacción, los resultados del transporte de carbohidratos mostraron diferencias importantes entre las cepas analizadas. La cepa SpIB2 fue un 28 % y un 32 % más eficiente para ingresar a la maltosa y la maltotriosa, respectivamente (Fig. 12; Fig. S6).

Contrariamente a resultados anteriores, las condiciones de fermentación contribuyen poco o nada al fenotipo de transporte de alfa glucósido (Day *et al.*, 2002). En el contexto de la regulación genética impartida por la represión de catabolitos y la afinidad previa del transportador, podemos reevaluar la importancia del transporte de maltosa en lugar del transporte de maltotriosa.

Aunque, *AGT1* es responsable del transporte de ambos azúcares, muestra dos sistemas de afinidad, un sistema superior para la maltosa ( $K_m = 14$  mM) y un sistema inferior para la maltotriosa ( $K_m = 27$  mM) (EK *et al.*, 1995; Smit *et al.*, 2007; Stambuk y Araujo, 2001; Vidgren *et al.*, 2005). Esta información bioquímica y predicciones bioinformáticas podrían ayudarnos a explicar que el transporte

reducido de maltotriosa observado en SpIB1 es causado por las diferencias en los patrones de expresión de los transportadores y reguladores MAL (Fig. 12C).

### **Evidencia física sugiere una regulación genética diferencial entre las levaduras analizadas.**

La detección genómica indica que ambas cepas contienen la misma maquinaria molecular para el transporte de maltosa y maltotriosa (Fig. 13A). La detección de transcritos mediante ensayos de PCR inversa demostró que todos los genes analizados se inducen en SpIB2 el primer día, incluso con presencia de glucosa, incluido *MTT1*. No así para la cepa SpIB1, donde el gen de importancia *MTT1* fue ausente.

La tasa de transporte reducida observada en SpIB1 puede estar asociada con una menor expresión de proteínas similares a *MTT1* y otras supuestas proteínas auxiliares como el alelo *ScAGT1* (Han *et al.*, 1995; Vidgren *et al.*, 2005) y *MPH2/3* (Magalhães *et al.*, 2016).

Aunado a esto, se ha informado que el alelo *ScAGT1* está truncado en la levadura Sp (Vidgren *et al.*, 2009), y los genes *MPHx* se han asociado con la maltosa, pero no con el transporte de maltotriosa (Alves *et al.*, 2008; Brown *et al.*, 2010; Magalhães *et al.*, 2016).

Este comportamiento molecular coincide con las concentraciones de maltosa en el segundo día de fermentación (Fig. 12C), donde solo los transcritos de *ScMalx1* y *MPHx*, expresadas el primer día, son funcionalmente competentes para este consumo (Fig. 13B).

El segundo día, SpIB1 expresó todos los genes analizados (Fig. 13B) que reflejan el agotamiento casi completo de maltosa, pero no de maltotriosa (Fig. 12C). En estas circunstancias, es razonable especular que ambas cepas tienen patrones de regulación diferentes relacionados con la jerarquía del consumo de azúcar, permitiendo así la variación en el consumo de maltosa y maltotriosa.

## VI Conclusión

**Las diferencias en la regulación genética juegan un papel significativo en la variabilidad observada en el transporte de maltosa y maltotriosa en ambas cepas de *S. pastorianus*.**

Se han realizado muchos esfuerzos para mitigar las limitaciones de las nuevas tecnologías de secuenciación para resolver regiones repetitivas y conflictivas en los genomas de algunas levaduras (Magalhães *et al.*, 2016; Salazar *et al.*, 2019).

En cierto modo, complementar los enfoques bioinformáticos con el transcriptoma de interés y la evidencia experimental permite un poder de resolución aún mayor de la información ómica en cuestión (Nobre *et al.*, 2016; Wei *et al.*, 2020). El presente estudio investiga los factores que influyen en el consumo de maltosa y maltotriosa en dos levaduras lager industriales.

Pudimos identificar no sólo variaciones genéticas en dos regiones importantes involucradas en la especificidad de la maltotriosa, sino también diferencias en los niveles de expresión del transportador MAL y los genes reguladores MAL positivos. Estos resultados indican que la diversidad fenotípica relacionada con el consumo de maltotriosa es más compleja que el fenómeno de un solo factor, destacando la utilización significativa de genes reguladores MAL.

La mayor expresión de genes reguladores positivos para genes MAL constituye un nuevo contribuyente a la complejidad de los fenotipos de consumo de maltosa y maltotriosa en la levadura estudiada. Este factor adicional debe incluirse en el análisis holístico para mejorar el proceso de elaboración de cerveza.

# Expression patterns of *Mal* genes and association with differential maltose and maltotriose transport rate of two *Saccharomyces pastorianus* yeasts

César I. Hernández-Vásquez,<sup>1</sup> Jorge H. García-García,<sup>1</sup> Esmeralda R. Pérez-Ortega,<sup>2</sup> Adriana G. Martínez-Segundo,<sup>2</sup> Luis C. Damas-Buenrostro,<sup>2</sup> Benito Pereyra-Alfárez<sup>1</sup>

**AUTHOR AFFILIATIONS** See affiliation list on p. 15.

**ABSTRACT** Beer brewing is a well-known process that still faces great challenges, such as the total consumption of sugars present in the fermentation media. Lager-style beer, a major worldwide beer type, is elaborated by *Saccharomyces pastorianus* (Sp) yeast, which must ferment high maltotriose content worts, but its consumption represents a notable problem, especially among Sp strains belonging to group I. Factors, such as fermentation conditions, presence of maltotriose transporters, transporter copy number variation, and genetic regulation variations contribute to this issue. We assess the factors affecting fermentation in two Sp yeast strains: SpIB1, with limited maltotriose uptake, and SpIB2, known for efficient maltotriose transport. Here, SpIB2 transported significantly more maltose (28%) and maltotriose (32%) compared with SpIB1. Furthermore, SpIB2 expressed all MAL transporters (*ScMALx1*, *SeMALx1*, *ScAGT1*, *SeAGT1*, *MTT1*, and *MPHx*) on the first day of fermentation, whereas SpIB1 only exhibited *ScMalx1*, *ScAGT1*, and *MPH2/3* genes. Some SpIB2 transporters had polymorphic transmembrane domains (TMD) resembling *MTT1*, accompanied by higher expression of these transporters and its positive regulator genes, such as *MAL63*. These findings suggest that, in addition to the factors mentioned above, positive regulators of Mal transporters contribute significantly to phenotypic diversity in maltose and maltotriose consumption among the studied lager yeast strains.

**IMPORTANCE** Beer, the third most popular beverage globally with a 90% market share in the alcoholic beverage industry, relies on *Saccharomyces pastorianus* (Sp) strains for lager beer production. These strains exhibit phenotypic diversity in maltotriose consumption, a crucial process for the acceptable organoleptic profile in lager beer. This diversity ranges from Sp group II strains with a notable maltotriose-consuming ability to Sp group I strains with limited capacity. Our study highlights that differential gene expression of maltose and maltotriose transporters and its upstream trans-elements, such as MAL gene-positive regulators, adds complexity to this variation. This insight can contribute to a more comprehensive analysis needed to the development of controlled and efficient biotechnological processes in the beer brewing industry.

**KEYWORDS** lager yeast, maltose and maltotriose uptake, differential gene expression

Beer represents a highly popular fermented beverage globally (1), with production reaching 189 million kiloliters in 2022 (<https://www.statista.com>). The important role of microorganisms in fermented food biotechnology depends on the enzymatic conversion of raw substrates to human edible and nutritive products, especially in beer brewing (2, 3). Yeasts are crucial in beer brewing, with *Saccharomyces cerevisiae* (Sc) used for ale-style beer, with fermentation temperature ranging from 15 °C to 26 °C, and

**Editor** Edward G. Dudley, The Pennsylvania State University, University Park, Pennsylvania, USA

Address correspondence to Benito Pereyra-Alfárez, [benito.pereyraal@uanl.edu.mx](mailto:benito.pereyraal@uanl.edu.mx).

The authors declare no conflict of interest.

See the funding table on p. 15.

**Received** 6 March 2024

**Accepted** 22 May 2024

**Published** 8 July 2024

Copyright © 2024 American Society for Microbiology. All Rights Reserved.

*Saccharomyces pastorianus* for lager-style beer, with fermentation temperatures between 5 °C and 16 °C (4, 5). *S. pastorianus* (Sp), an interspecies hybrid of *S. cerevisiae* (Sc), and a cryotolerant yeast, *Saccharomyces eubayanus* (Se) (6-8), is divided into Saaz type (group I) and Froberg type (group II) based on chromosome number content and sugar consumption (9). There are two hypotheses that explain the lager yeast's origin: (i) the existence of two independent hybridization events involving different yeasts, Sc, and Se (6, 10, 11); and (ii) both groups of lager yeast share a single hybridization event, and the evolution resulting from that hybridization shaped genomic divergences between both groups (12). The prevalence of Sp in the brewing industry implies that this hybrid possesses a selective advantage over its parent strains. Key characteristics contributing to this advantage include enhanced biological fitness under stress conditions and efficient maltotriose consumption (10, 13, 14). However, physiological differences between lager yeast highlight the defective maltotriose uptake seen in some members of group I (9). The transport of maltose and maltotriose stands as a bottleneck for achieving effective fermentation (15), given their substantial presence, accounting for 50%-60% and 15%-20% of the fermentable sugars in a standard brewing wort, respectively (16, 17). Despite their abundance, the transport of maltose and maltotriose is regulated by a carbon catabolite repression mechanism, where glucose is prioritized over maltose and maltotriose (18-20). Consequently, maltose is more easily consumed by most brewing yeasts, leaving maltotriose more prevalent in later stages of fermentation, where some yeast strains exhibit a deficiency in its uptake (11, 15, 21, 22).

Different factors are involved in deficient maltotriose consumption phenotype. These include (i) the stressful conditions on the last days of fermentation (23), (ii) the presence or absence of specific transporters for maltotriose, such as *AGT1* and *MTT1/MTY1* (24-26), and (iii) regulatory differences in maltose and maltotriose transporter genes (27, 28). Furthermore, the emergence of chimeric genes, resulting from genomic recombination, are also involved in the mechanism of generating variation in closely related yeast, leading to the development of highly efficient maltotriose transporters (29, 30).

As set before, alpha glycoside consumption varies among lager yeast, and its study lack global picture as in RNA-seq transcriptomic level analysis. Here, we investigate the sugar transport system in two lager yeasts, specifically SpIB1, characterized by a limited maltotriose uptake, and SpIB2, distinguished by its effective maltotriose transport. We investigate if fermentations conditions and/or differences in the genetic architecture of *MAL* genes are associated with this phenotypic diversity of maltose and maltotriose consumption. By integrative analysis of transcriptome, qualitative detections of *MAL* genes and transcripts, and biochemical determination of maltose and maltotriose transport rate, we discuss molecular events, such as polymorphic regions in maltose and maltotriose transporters, and differential *MAL* gene expression patterns are associated with the underlying phenomena involved in the phenotypic diversity of sugar utilization in two lager yeasts. To our knowledge, this work shows the first RNA-seq data approach to get insight into sugar consumption in two lager yeasts, enabling us to detect the regulatory *MAL* genes as a key feature in this difference.

## RESULTS

### Abundance of permeases in the strains studied

To investigate the possible gene copy number disparities among permeases, potentially influencing distinct sugar utilization patterns, we leveraged genome sequencing data to identify gene loci associated with *MAL31*, *AGT1*, *MPH2/3*, and *MTT1* in both SpIB2 and SpIB1 strains. Although congruence in copy number was observed for *AGT1*, *MPH2/3*, and *MTT1*, *MAL31* exhibited variation in copy number between SpIB2 (four loci) and SpIB1 (six loci) (Table 1; Fig. S1). Despite both strains having the *MTT1* transporter in their genome, the nucleotide sequence was incomplete compared with the reference length (25, 26). The length of nucleotides and amino acids in the permeases significantly varied (Table 1), with only *AGT1*- and *MAL31*-specific isoforms being complete in each genome. Additionally, the structural distribution of *MAL* loci differed from the canonical

**TABLE 1** The location and size of permeases for maltose and maltotriose in the studied strains

Gene name	Systematic name	Locus (scaffold number)	Coordinates	Strand	Length in genome nt/aa	Length in transcriptome (aa)
SpIB2 permeases						
<i>AGT1</i>	YGR289C	Scaffold3	4536-6368	+	1,851/610	558
<i>AGT1</i>	YGR289C	Scaffold50	71200-73051	-	1,185/394	
<i>MAL31</i>	YBR298C	Scaffold102	3-581	-	581/193	615
<i>MAL31</i>	YBR298C	Scaffold6	612004-613788	-	1,784/594	
<i>MAL31</i>	YBR298C	Scaffold11	492091-492738	-	2,165/721	
<i>MAL31</i>	YBR298C	Scaffold15	20813-21508	+	1,793/597	
<i>MPH3</i>	YJR160C	Scaffold29	946600-948408	-	1,808/602	374
<i>MTY1</i>	MTY1	Scaffold13	658031-659338	-	1,307/435	Not detected
SpIB1 permeases						
<i>AGT1</i>	YGR289C	Scaffold11	3868-5700	+	1,833/611	610
<i>AGT1</i>	YGR289C	Scaffold23	4299-6150	+	1,185/394	
<i>MAL31</i>	YBR298C	Scaffold8	14887-15582	+	1,825/608	614
<i>MAL31</i>	YBR298C	Scaffold10	282142-283926	-	1,784/594	
<i>MAL31</i>	YBR298C	Scaffold26	9677-11624	+	1,947/614	
<i>MAL31</i>	YBR298C	Scaffold30	645561-645869	-	309/103	
<i>MAL31</i>	YBR298C	Scaffold82	1-792	+	792/264	
<i>MAL31</i>	YBR298C	Scaffold85	308-757	+	448/149	
<i>MPH3</i>	YJR160C	Scaffold24	556421-558229	-	1,809/602	462
<i>MTY1</i>	MTY1	Scaffold64	204-1553	-	1,349/449	Not detected

pattern, which was observed only at scaffold15 in SpIB2 (11, 31) (Fig. S1). Transcriptomic data confirmed the completeness of permeases, showing that both strains had the MAL31p transporter complete in each transcriptome, whereas the presence of the MTT1p transporter seemed to be absent. To resolve these conflicting results, multiple alignment analyses were conducted to determine the phylogenetic distribution of these transporters using a maximum likelihood algorithm.

### Structural analysis, expression levels, and phylogenetic distribution of the permeases

To detect specific variations in the structure, expression level, and phylogenetic relationships of maltose and maltotriose transporters that could approximate the differential sugar consumption efficiency, we conduct multiple sequence alignment to detect polymorphic regions related to maltose and maltotriose translocation mechanism. Also, we conduct transcriptomic analysis of the second and fifth days of fermentation for each yeast strain to detect differences that could explain the maltose and maltotriose transport efficiency. Finally, we conduct a maximum likelihood phylogeny to estimate structural/function relationships between maltose and maltotriose transporters of both strains.

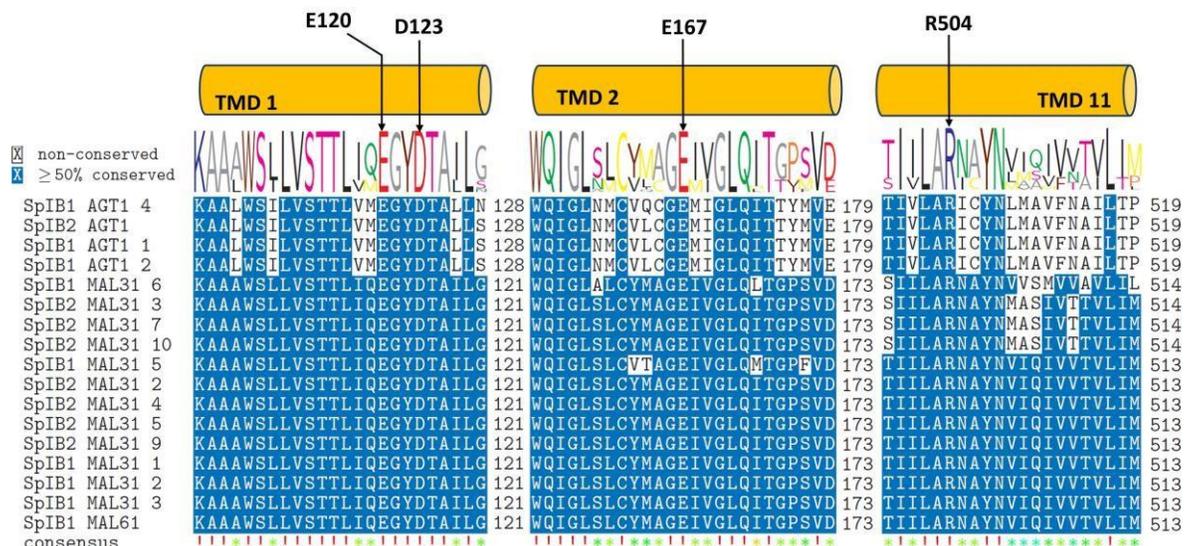
We detect that the *AGT1* and *MAL31* transporters of both strains showed 12 transmembrane domains (TMD1-12), which is characteristic of MSF transporters (32-34). Despite its partial sequence (558 aa), the SpIB2 *AGT1* transporter displayed the complete set of canonical transmembrane domains typical of MSF family members. However, some transcripts associated with *MPH2/3* and *MAL31* genes in both strains lacked transmembrane domain predictions, whereas others exhibited deletions affecting TMD structure (Fig. S2 and S3). For the remaining permeases, the 3D structure prediction matches with alpha-glycoside transporters, showing a high confidence value between TMD structure and predicted aligned error distributions similar between them (Fig. S2 and S3). Taking into account these results, those proteins that showed deletions or not transmembrane domains predictions were excluded from multiple sequence analysis and phylogenetic tree estimations (35).

Regarding the conservation pattern, TMD1 (110-129 aa), TMD2 (157-178 aa), TMD7 (371-399), TMD11 (501-524 aa), and TMD12 (532-550 aa) showed >50% of conservation. However, we also observed some variations in MAL31- and AGT1-specific regions evidencing a polymorphic region that depends on Mal transporter (Fig. 1; Fig. S4). Permease amino acid conservation patterns revealed a consistent presence (≥50%) of E120, D123, E167, and R504 in their respective TMD (Fig. 1). As drawn at Fig. S4, asparagine (N) is conserved ≥50% in all TMD11 sequences, except for specific MAL11/AGT1 isoforms present in SpIB2 and SpIB1, where there is an isoleucine (I). Additionally, an alanine (A) is conserved ≥50% in TMD12 sequences, but not for SpIB2 and SpIB1 MAL11/AGT1, where threonine (T557) was observed in SpIB2 isoform, whereas SpIB1 isoforms had A379 and Y384 (Fig. S5).

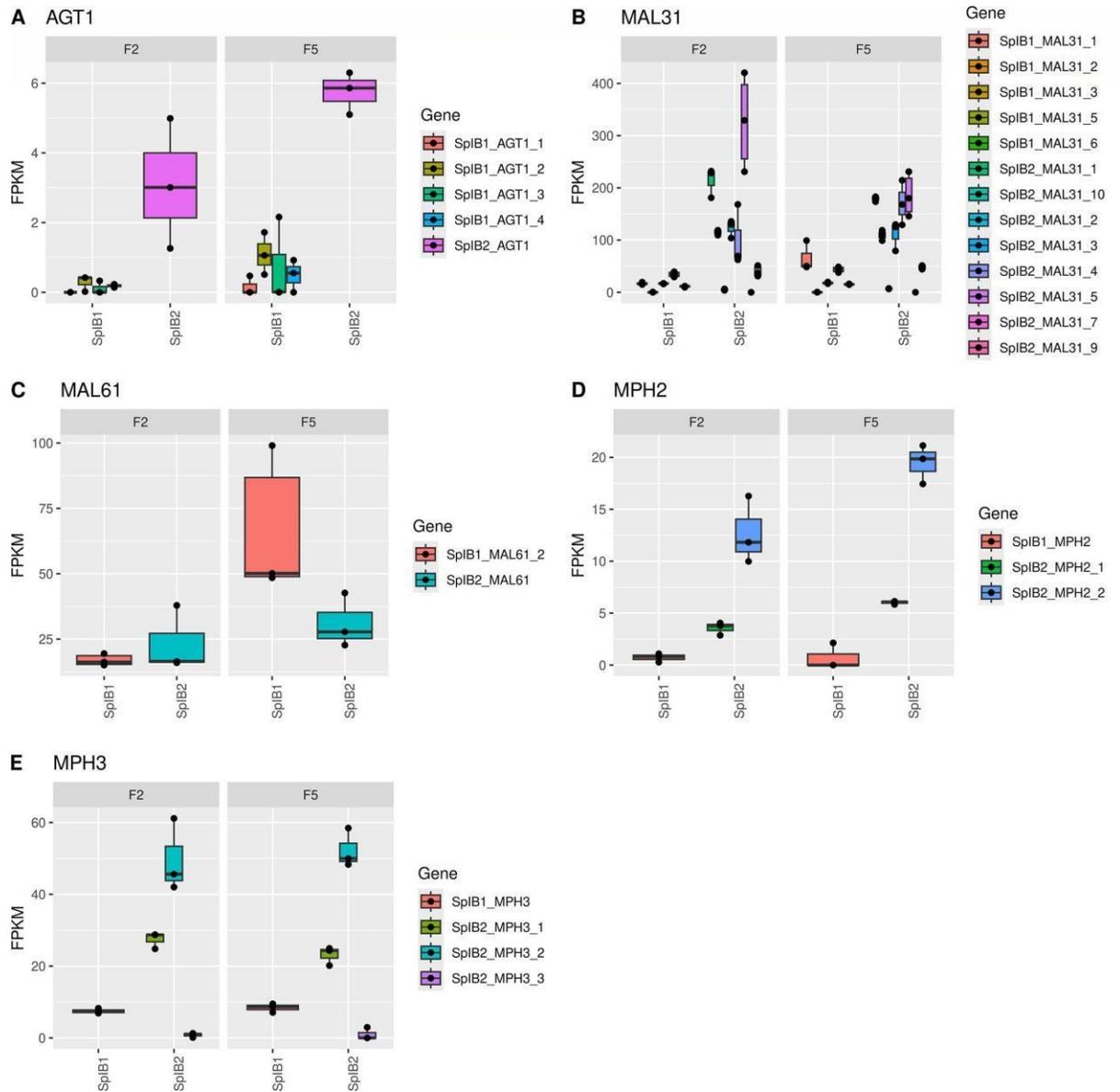
In the expression level analysis, SpIB2 exhibits higher transcriptional activity for most transporters, except for MAL61, which shows more expression in the SpIB1 strain. Specifically, the AGT1 isoform from SpIB2 demonstrates a higher median of fragment per kilobase of transcript per million of mapped reads values (FPKM), 2.6 and 6.08, on the second and fifth days of fermentation, compared with the isoform from SpIB1 with FPKM values of 0.22 and 1.39 on the same days (Fig. 2A). For MAL31 genes, the SpIB2 isoform displays the highest expression values on the second and fifth days (329.26 and 180.08 FPKM), contrasting with the highest expressed SpIB1 isoform showing 34.52 and 50.12 FPKM on the respective days (Fig. 2B). These results are consistent with the regulatory region analysis where, in general, the yeast SpIB2 has more functional binding sites for MAL63 regulatory proteins (Tables S2 and S3).

Concerning MAL61 transporters, the SpIB1 isoform exhibits the highest expression (67.27 FPKM) at the fifth day of fermentation (Fig. 2C). Transcriptional activity of MPH2 and MPH3 is higher in SpIB2 than SpIB1 (Fig. 2D and E). However, both strains display transcript isoforms with deletions affecting numerous transmembrane domains (Fig. S2).

To obtain the relationship between MAL transporters of SpIB2, SpIB1, and those reported for *S. cerevisiae*, *S. eubayanus*, and *S. pastorianus*, a maximum likelihood phylogenetic tree was estimated. Figure 3 reveals the presence of three principal clades observed previously (30, 36), where very divergent AGT1 proteins formed their own group. Here, we observed that MAL11 proteins from SpIB2 and SpIB1 grouped in this cluster, suggesting a similar function. The MPH3 and MPH2 groups are constituted of

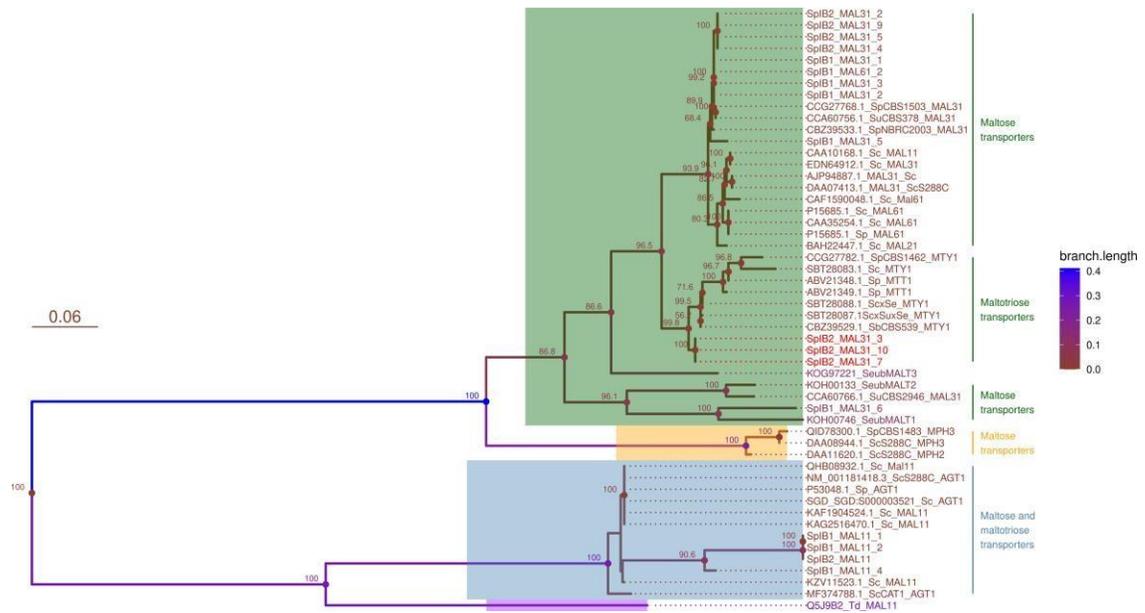


**FIG 1** Sections of multiple amino acid sequence alignments of maltose and maltotriose transporters showing the charged residues involved in the maltose translocation mechanism. Only the predicted transmembrane domains TMD-1, TMD-2, and TMD-11 are represented.



**FIG 2** Expression of maltose and maltotriose transporters depicted in FPKM values. Date is represented as boxplot MAL gene expression FPKM values extracted from triplicate transcriptomic data of each strain. Note that in (A) AGT1, (B) MAL31, (D) MPH2, and (E) MPH3 genes the strain SplB2 have more expression than the SplB1 strain. The strain SplB1 has major expression of (C) MAL61 transporter only on the fifth day of fermentation.

two proteins from *S. cerevisiae* and one protein from *S. pastorianus*; however, we do not include the proteins from SplB2 and SplB1 as we observed deletions in the important transmembrane domains for maltose translocation (Fig. S2). The largest clade constituted of MALT transporters form *S. eubayanus* and MAL31\_6 from SplB1, MALx1 from both Sc and Sp and lager hybrid-specific MTT1/MTY1 transporters. It is important to note that SplB2 MAL31 proteins grouped, supported with 98.6 bootstrap value, in the MTT1 cluster, suggesting a very related maltotriose transport function. Finally, we include a very distant MAL11 transporter from *Torulaspora delbrueckii* to obtain a more robust interpretation of the tree (35).



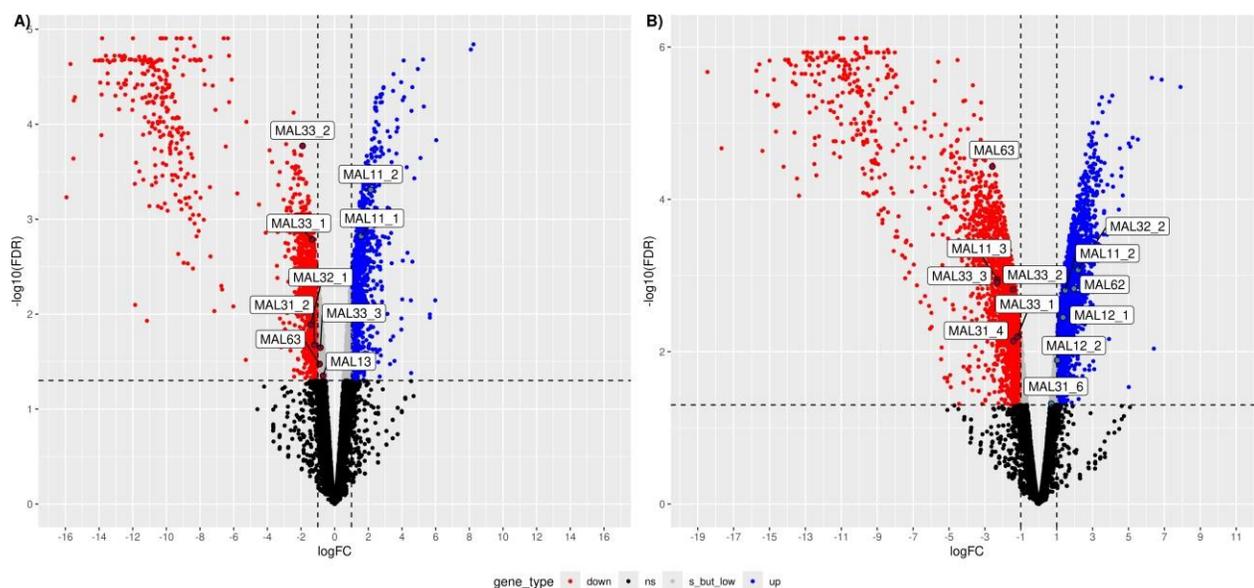
**FIG 3** Maximum likelihood phylogeny with 1,000 times bootstrapping for maltose and maltotriose transporter proteins reported for *S. cerevisiae*, *S. eubayanus*, *S. pastorianus*, SpiB2, and SpiB1. The scale bar indicates the number of amino acid substitutions per site. Note how the SpiB2 MAL31 proteins, marked in red, grouped with the MTT1/MTY1 permeases reported in Sp.

### Global expression analysis

To determine the gene expression affected only by the type of yeast strain and not by the fermentation condition and also to obtain statistical insights from the expression level analysis (FPKM) section, we decided to perform differential gene expression analysis by reading alignment of transcriptomic reads from SpiB1 and SpiB2 against the reference genome of the yeast *S. pastorianus* CBS 1483. The MAL loci analyzed in this work showed 99%-100% identity and a coverage of 87%-100%, with chromosomal regions of the Sp CBS1483 (Table S1). In general terms, we observed a considerable effect depending on the strain type. In the second day of fermentation, we determined 1613 down-regulated genes and 1215 up-regulated genes in SpiB1. On the fifth day of fermentation, we observed 2494 down-regulated genes and 2183 up-regulated genes in SpiB1. Here we can conversely interpret these differentially expressed genes (37-39), where the down-regulated genes in SpiB1 are up-regulated in SpiB2, and up-regulated genes in SpiB1 are down-regulated genes in SpiB2.

From these differentially expressed genes, we focused on MAL loci genes, resulting in the second-day fermentation comparison (F2\_820 vs F2\_790), with two differentially expressed genes in SpiB1; MAL11\_2 and MAL11\_1 (Fig. 4A). In contrast, strain SpiB2 exhibits differential expression in genes MAL33\_1, MAL33\_2, MAL32\_1, MAL31\_2, MAL33\_3, MAL63, and MAL13 (Fig. 4A; Table S4). These genes perform positive regulatory functions, and its global expression is confirmed by FPKM measurements, indicating major expression in SpiB2 (Fig. 2; Fig. S5). However, it is important to note that the MAL63, MAL13, and MAL33\_3 genes are statistically significant but show a low expression in SpiB2 ( $\log_2$  of fold change < 1, gray dots in Fig. 4A), which indicate that their abundance is less than double in the strain SpiB2.

The same pattern of expression is seen on the fifth day of fermentation (F5), where strain SpiB1 shows expression of maltose and maltotriose transporters, MAL11 (AGT1) and MAL31, as well as some maltases, such as MAL32 and MAL62 (Fig. 4B; Table S5).



**FIG 4** Volcano plot representing differentially expressed genes in the (A) second and (B) fifth days of fermentation. Up- and down-regulated genes are represented in blue and red, respectively (absolute  $\log_2$  fold change  $>1$  and false-discovery rate [FDR]  $< 0.05$ ). Genes with significant but low expression change (absolute  $\log_2$  fold change  $<1$  and FDR  $< 0.05$ ) are depicted in gray, and non-significant genes are in black (FDR  $> 0.05$ ).

Similarly, strain SpIB2 has more positive regulatory genes, such as *MAL63* and *MAL33\_1-3*, and some transporters, such as *MAL11\_3* and *MAL31\_4*. The first four genes, which are positive regulators, showed more than double the abundance in SpIB2 compared with SpIB1, highlighting the significant utilization of these genes by SpIB2.

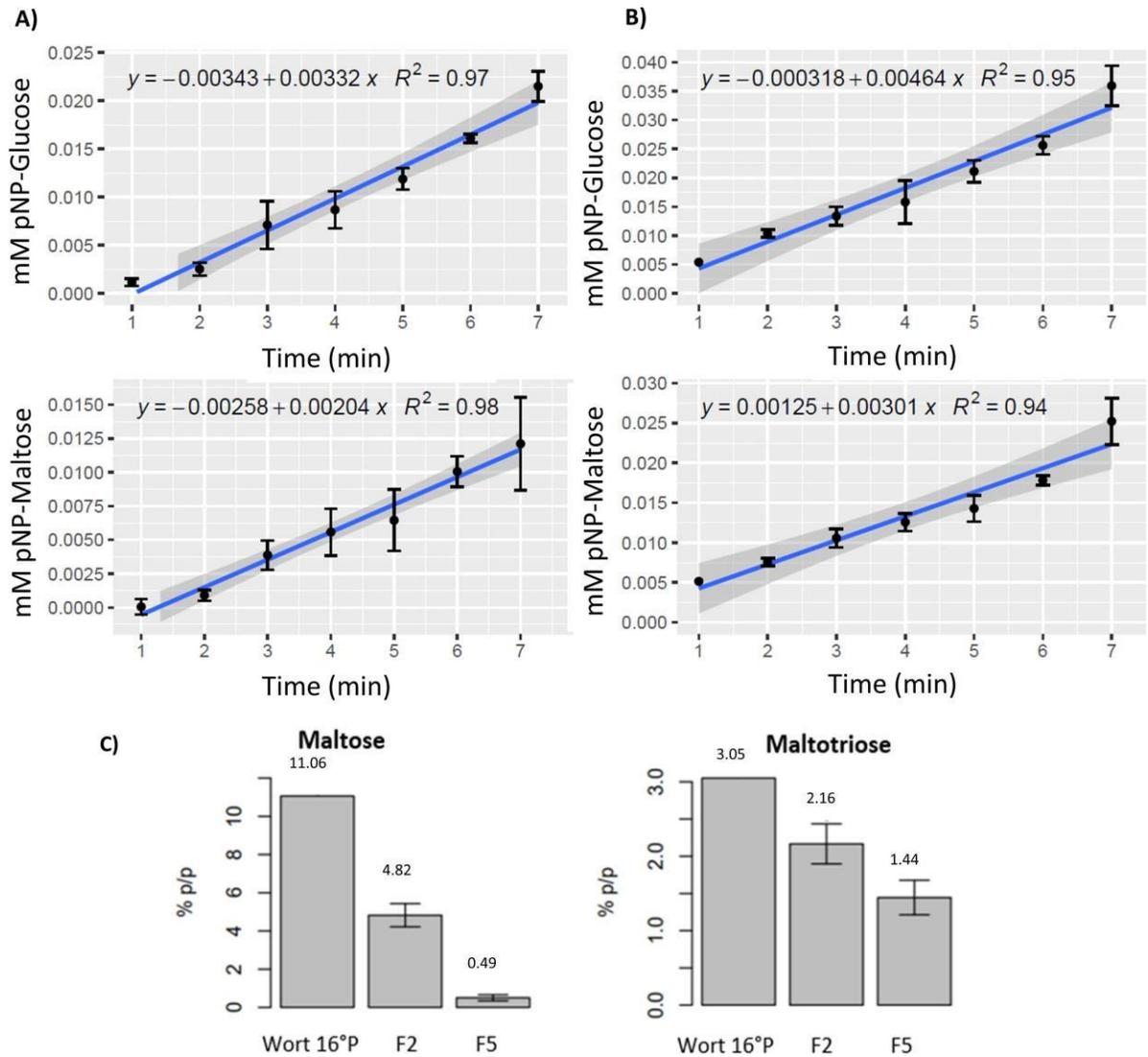
### Biochemical analysis of the transport rate and molecular detection of maltose and maltotriose transporters

To demonstrate if fermentation conditions have a significant effect in the maltose and maltotriose transport, we determined the *p*-nitrophenyl- $\alpha$ -D-glucopyranoside (pNP-glucose) and *p*-nitrophenyl- $\alpha$ -D-glucopyranosyl-(1,4)-D-glucopyranoside (pNP-maltose) transport rate curves for the strains SpIB2 and SpIB1 in the same reaction conditions. The fundamentals of this assay are that chemically labeled pNP substrates are transported into the cell by alpha-glycoside transporters (*Malx1*, *AGT1*, *MPHx*, and *MTT1*), where alpha-amylases hydrolyze the alpha bond of maltose and maltotriose, releasing the pNP. After permeabilization and pNP extraction, spectrophotometric measurements were performed at 500 nm for quantitative determination (40). Because the stoichiometric ratio of pNP to the alpha-glucoside of interest is 1:1 (41), it can extrapolate the concentration of transported pNP to the concentration of labeled alpha glycoside.

Although both strains take maltose at a higher rate than maltotriose, SpIB2 demonstrates higher transport rates of both alpha-glycosides than SpIB1 (Fig. 5A and B). Specifically, for maltose, SpIB2 has a transport rate of  $282 \text{ nmol min}^{-1} \text{ mg}^{-1}$  dry weight, whereas SpIB1 has  $202 \text{ nmol min}^{-1} \text{ mg}^{-1}$  dry weight (Fig. S6).

Similarly, in maltotriose transport, SpIB2 exhibits a higher rate of  $183 \text{ nmol min}^{-1} \text{ mg}^{-1}$  dry weight compared with SpIB1  $124 \text{ nmol min}^{-1} \text{ mg}^{-1}$  dry weight. These biochemical results also reflect the phenotypic behavior in the same fermentation condition where we observed  $0.49 \pm 0.04\%$  p/p and  $1.44 \pm 0.27\%$  P/P of residual maltose and maltotriose at the fifth day, respectively, in SpIB1, whereas the sugars were not detectable in SpIB2 (Fig. 5C).

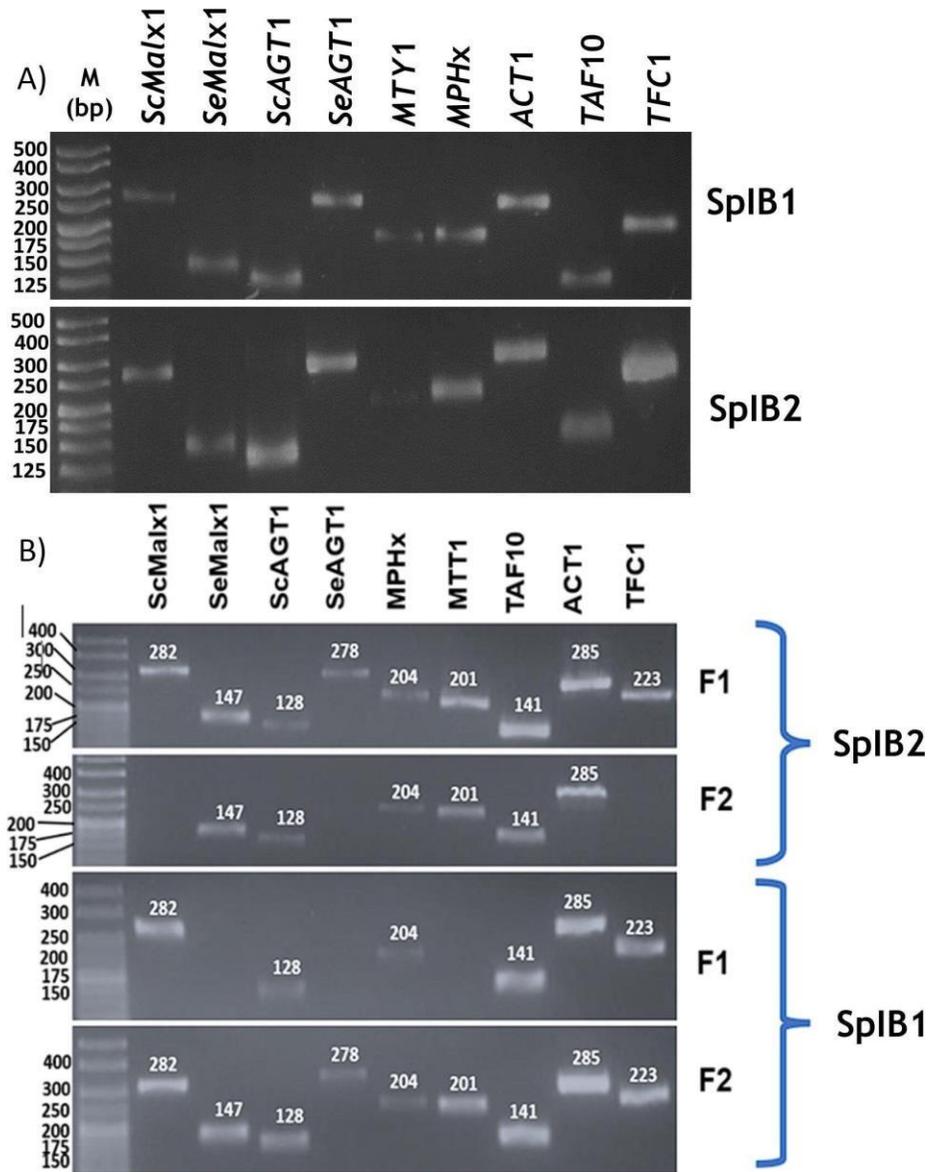
Genetic amplification data suggest that both strains carry *ScMALx1*, *SeMALx1*, *ScAGT1*, *SeAGT1*, *MTT1*, and *MPH2/3* genes, where the presence of a *MTT1*-like in SpIB1



**FIG 5** Cellular transport rate determined by uptake of pNP-glucose and pNP-maltose for 7 min in the strain (A) SplB1 and (B) SplB2. (C) Residual maltose and maltotriose on the second and fifth days of fermentation for strain SplB1. Data for strain SplB2 were undetectable.

was confirmed (Fig. 6A). Transcript gene detection indicates that both strains express all analyzed permeases. Notably, SplB2 exhibits all permeases on the first day, whereas SplB1 only shows *ScMalx1*, *ScAGT1*, and *MPHx*. On the second day of fermentation, we observed *SeMalx1*, *ScAGT1*, *MPHx*, and *MTT1* in SplB2 and all analyzed transcripts in SplB1 (Fig. 6B). From these outputs, we highlight that SplB2 expresses all MAL transporters analyzed in these conditions, and SplB1 expresses all MAL transporters, especially the *MTT1* on the second day of fermentation, suggesting a probable transcription regulation delay caused by differences in the expression level of the MAL regulatory proteins.

Overall, these findings support our bioinformatic inferences, standing out genetic regulation as the predominant factor in the phenotypic variation of maltose and maltotriose transport in the studied strains.



**FIG 6** Molecular detection of maltose and maltotriose permeases in SpIB2 and SpIB1. (A) Genomic detection with PCR of permeases ScMALx1, SeMALx1, ScAGT1, SeAGT1, MPHx, and MTT1/MTY1 indicates that each strain has all the MAL transporters included in this analysis. (B) Transcript detection by RT-PCR on the first and second days of the fermentation of the studied strains. Notice that on the first day, the strain SpIB2 showed all Mal transcripts, whereas SpIB1 only showed ScMALx1, ScAGT1, and MPHx.

## DISCUSSION

### Structural variability accounts for the phenotypic variation in alpha-glycoside transport

Spontaneous mutation is a significant driving force behind protein diversity under natural selection, potentially allowing for ~50%-80% divergence without effecting protein structure (42, 43). However, changes in specific residues, particularly those

related to substrate-binding sites, could impact protein function (42). Mal permease displays a highly conserved amino acid pattern ( $\geq 50\%$ ) in TMD1, TMD2, and TMD11, which were reported as relevant for maltotriose transportation (41, 44). Additionally, an alanine (A) is conserved  $\geq 50\%$  in TMD12 sequences but not for SpIB2 and SpIB1 MAL11/AGT1, where threonine (T557) was observed (45). However, it is important to mention that these amino acids are modeled in the context of AGT1 transporters and may not represent the big picture of maltotriose specificity of MTT1/MTY1-like transporters.

Recent studies have reported some crucial polymorphic sites in TMD7 and TMD11 in MAL61, MTT1, and MTT1-like transporters, where the presence of T379 and N384 residues in TMD7 are involved in maltotriose specificity (46). Here, the SpIB2 MAL31 showed these amino acid substitutions but not present in any SpIB1 MAL31 proteins. This result strongly suggests that SpIB2 has an efficient maltotriose transport system present in its transcriptome, whereas SpIB1 has an AGT1-like landscape in its alpha glycoside transport architecture.

Overall, amino acid and synteny changes found in the *Mal* loci (Fig. S1) indicate that industrial brewing fermentation conditions act as a selective force, and the concomitant hybrid nature of Sp may cause the maltose and maltotriose transport plasticity observed in this yeast (11, 12, 31, 36, 46).

### **Expression patterns and identity of the transporters efficiently explain the phenotypic variation of maltose and maltotriose transport in the analyzed strains**

Currently, diverse *S. pastorianus* strains use MTT1 and/or AGT1 proteins for maltotriose transport (21-26). Sugar depletion on the second (F2) and fifth (F5) days for strain SpIB2 and SpIB1 reflects an illustration of differences in genetic regulation of MAL-type genes. This notion is supported by the transcriptomic data of putative MTT1 and AGT1 genes in SpIB2 but just AGT1 in SpIB1. Transcriptomic results confirmed that MAL31 and MTT1 have more expression values (FPKM) in SpIB2 than in SpIB1. However, the expression values (FPKM) for AGT1 were low in both strains, even though its importance has been previously reported (Fig. 2) (28, 47, 48). This outcome makes sense because these lager yeasts have AGT1 in its genetic background, suggesting an evolutionary path related to lower expression of AGT1 and the concomitant presence of MTT1-like transporters (30).

The maltose transporters MAL61 were more expressed in SpIB1 on the fifth day of fermentation (Fig. 2C), and this is consistent with its maltose consumption (Fig. 5C), which resemble the interval of maltose and maltotriose residuals observed previously (24). Finally, the expression level of MPHx genes was more in SpIB2 than in SpIB1; however, the MPHx genes were only maltose-specific transporters and its function in maltotriose transport is ambiguous (24, 36).

Regarding the phylogenetic relationships of maltose and maltotriose transporters, our tree is consistent with previous reports (30, 36). In this sense, the SpIB2 strain showed three MAL31 proteins clustered with MTT1/MTY1 transporters, evidencing that this strain probably has an efficient maltotriose transport system, whereas no MTT1/MTY1 clustering was observed in SpIB1 strain. It is important to note that SpIB1 MAL31\_6 grouped with *S. eubayanus* maltose transporters, resembling a function similar to maltose permease of this yeast.

As observed here and in other studies, the maltose and maltotriose transporters AGT1 are a very distant group from the largest MAL31 clade, which has a MTY1/MTT1 maltotriose-specific transporter subclade, suggesting that the evolution of the maltotriose consumption feature has redundancy and frequent occurrence in MAL transporter landscape. However, contrary to the rapid evolution of MAL loci, the maltotriose-specific transporter proteins are a low-frequency event in the evolution pathway of this crucial industrial trait (29, 30, 36), which explains the rich phenotypic diversity observed in the alpha glycoside transport in SpIB2, SpIB1, and other important industrial yeasts (9, 24).

### **Under the same conditions, the yeasts show differences in their maltose and maltotriose transport rates**

To support the bioinformatic predictions related with the differential expression of MAL transporters and regulators and prove the hypothetic scenario where the fermentation conditions exert an effect in the maltose and maltotriose consumption of SpIB2 and SpIB1, the cellular transport rate of pNP-glucose (structural analog of maltose) and pNP-maltose (structural analog of maltotriose) was calculated in the same reaction conditions (22, 40, 44, 49).

Despite the null hypothesis being not different in the cellular transport rate in the same reaction conditions, carbohydrate transport results showed important differences between the analyzed strains. SpIB2 strain was 28% and 32% more efficient to ingress maltose and maltotriose, respectively (Fig. 5; Fig. S6). Contrary to previous results, fermentation conditions make little or no contribution to the alpha-glucoside transport phenotype (23).

In the context of genetic regulation imparted by catabolite repression and the previous transporter affinity, we can reevaluate the importance of maltose rather than maltotriose transport. Even so, *AGT1* is responsible for transport of both sugars, and it shows two affinity systems, a higher system for maltose ( $K_m = 14$  mM) and a lower system for maltotriose ( $K_m = 27$  mM) (47, 49-51). These biochemical information and bioinformatic predictions could help us to explain that the reduced maltotriose transport observed in SpIB1 is caused by the differences in the expression patterns of MAL transporters and regulators (Fig. 5C).

### **Physical evidence of permeases suggests differential regulation among yeasts studied**

Genomic detection indicates that both strains contain the same molecular machinery for maltose and maltotriose transport (Fig. 6A). Transcript detection through reverse PCR assays demonstrated that all tested genes are induced in SpIB2 on the first day, even with glucose presence, including *MTT1*. However, in SpIB1, the last gene was absent. Reduced transport rate seen in SpIB1 may be associated with a lower expression of *MTT1*-like and other putative auxiliary proteins as the *ScAGT1* allele (47, 50) and *MPH2/3* (24). However, it has been reported that the *ScAGT1* allele is truncated in *S. pastorianus* yeast (48), and the *MPHx* genes have been associated with maltose but not with maltotriose transport (24, 36, 52). This molecular behavior matches the maltose concentrations at the second day of fermentation (Fig. 5C), where only the *ScMalx1* and *MPHx* transcripts, expressed on the first day, are functionally competent for this consumption (Fig. 6B).

On the second day, SpIB1 expressed all analyzed genes (Fig. 6B) that reflect the nearly complete depletion of maltose but not maltotriose (Fig. 5C). Under these circumstances, it is reasonable to speculate that both strains have different regulations patterns related to the hierarchy of sugar consumption.

### **Conclusion: Difference in genetic regulation plays a significant role in the variability of maltose and maltotriose transport in both strains of *S. pastorianus***

Many efforts have been addressed to mitigate the limitations of the new sequencing technologies in resolving repetitive and conflicting regions in the genomes of some yeasts (12, 24). Nonetheless, complementing bioinformatics approaches with the transcriptome of interest and experimental evidence allows for even higher resolving power of the omics information in question (53, 54). The present study investigates factors influencing maltose and maltotriose consumption in two industrial lager yeasts. We can identify not only genetic variations in two important regions involved in maltotriose specificity but also differences in expression levels of MAL transporter and positive MAL regulatory genes. These results indicate that the phenotypic diversity related to maltotriose consumption is more complex than a single-factor phenomenon,

highlighting the significant utilization of *MAL* regulatory genes. The major expression of *MAL*-positive regulatory genes stands as novel contributor to the complexity of maltose and maltotriose consumption phenotypes in the studied yeast. This additional factor must be included in the holistic analysis to improve the beer brewing process.

## **MATERIALS AND METHODS**

### **Strains**

Strains SpIB1 and SpIB2, representing groups I and II, were obtained from Cervecería Cuauhtémoc Moctezuma (Monterrey, NL, México). Inoculated into YP-M medium (1% yeast extract, 2% peptone, and 2% maltose), the yeast strains were stored in YP-M medium with 40% glycerol at  $-20^{\circ}\text{C}$  after 48 h or upon reaching stationary phase for future use.

### **Cultivation conditions and media**

Both strains were inoculated in YP-M medium and brewing wort with the following composition: 138 ppm free amino nitrogen, 0.11% (w/w) fructose, 4.85% (w/w) glucose, 11% (w/w) maltose, and 3% (w/w) maltotriose. Fermentation proceeded for 5 days at  $16^{\circ}\text{C}$ , and the experiment was performed in triplicates. Sampling occurred on the first, second, and fifth days, chosen for their significant phenotypic expression in carbon dioxide production and flocculation phenotype.

### **Inspection and pair-based comparisons of permeases present in the genomic and transcriptomic data of the studied yeasts**

We utilized previously obtained genomics (55) to gather insights into maltose and maltotriose transporters within SpIB1 and SpIB2 strains. By examining genome annotation files (31), we identified transporter genes in each strain. Employing the `getfasta` function of `bedtools` version 2.30.0 (56), `grep` UNIX utility, and `ggplot2` R package, we generated gene plots and loci information.

### **Permease structure analysis**

We conducted structural analysis by performing multiple sequence alignment for obtaining amino acids in the maltose translocation mechanism (44). Sequence conservation analysis was performed using the `msa` package version 1.32.0 (57), with `CLUSTAL-W` and `BLOSUM80` substitution matrix, to identify these amino acids in the *MAL* transporter proteins of both yeasts. The inclusion criteria for permease amino acid sequence involved complete sequences and/or with complete transmembrane domain predictions.

The phylogenetic distribution of each permease was obtained through the maximum likelihood method in `ape` version 5.7-1 (58) and `phangorn` version 2.11.1 packages with a bootstrap sample space of 1,000 replicates (59). The topology of transmembrane alpha-helix domains (TMD) in the *MAL* permeases of the studied strains was determined using `CCTOP` software (60). Finally, 3D models of the *MAL* permeases were constructed using the `AlphaFold DB` methodology (61, 62).

### ***Mal* gene regulatory region analysis**

We used `YEASTRACT` to identify putative binding sites of the transcription factors *MAL63p* and *Mig1p* for each *Mal* upstream regulatory region (63). Based on a previous scientific model (27, 28), in the absence of binding sites to the activator of *Mal* genes, *MAL63* shows reduced transcription levels for reporter genes or low growth in culture media with maltotriose as the sole carbon source and vice versa.

### FPKM values of the *mal* genes found in each strain

From the previously generated transcriptomic data (64), we derived FPKM values for the transporters *MALx1*, *AGT1*, *MTT1*, and *MPH2/3* to obtain *MAL* gene expression. Transcriptomic data were obtained from the second and fifth days of fermentation. Due to the hybrid nature of the yeast, potential specific recombination events, and the absence of a reliable reference genome, we performed *de novo* assembly.

Transcriptomic paired-end reads for both yeast strains were quality-trimmed using Trimmomatic version 0.38 (parameters LEADING:5 TRAILING:5 MINLEN:50) (65) and Sortmerna for contaminant RNA removal (66). Subsequently, forward and reverse reads for each strain were concatenated independently and assembled using Trinity tool version 2.14.0 with default parameters (67, 68). The assembled transcripts were annotated using Trinotate tool version 3.1.0 (69). For mapping trimmed reads to the *de novo* assembled transcript of each strain, Bowtie2 tool version 2.4.4 with default parameters (70) was employed. RSEM software version 1.3.3 with default values determined expression values (FPKM) (71). Notably, in the context of Trinity assembly output, the term “isoform” is used, leaving the biological significance regarding alternative splicing. Additionally, the unambiguous use of gene nomenclature *AGT1/MAL11* and *MTT1/MTY1* is emphasized in this study.

### Global expression analysis

We conducted a differential gene expression analysis for both yeast strains using samples collected on the second and fifth days of fermentation. To assess statistical differences in *Mal* gene expression, we employed read alignment with the reference genome of *S. pastorianus* cbs 1483 (genome and annotations available at NCBI under bioproject prjna522669). Star version 2.7.3a with default parameters was used for alignment (72). Htseq-count version 0.11.1 (73) converted bam alignment files to read count matrices, and edgeR version 3.34.0 (39) performed differential expression analysis. Genes with low expression levels (cpm > 5) were filtered, contrasting the groups on the second day of fermentation (f2) for both strains (contrast group f2\_820-f2\_790), and similarly on the fifth day of fermentation (f5) (contrast group f5\_820-f5\_790). A gene was considered significantly expressed if it showed at least a twofold abundance in each strain [ $\log_2(2) = 1$ ] with a false discovery rate (FDR) of <0.05. We focus on *Mal* genes from the list of differentially expressed genes.

### Cellular transport rate test

To determine the transport rate of pNP- $\alpha$ -glucose (structurally related to maltose) and pNP- $\alpha$ -maltose (structurally related to maltotriose) in our studied strains, we performed a cellular transport experiment in standard conditions as previously described (40). Briefly, SpIB2 and SpIB1 were suspended at 16 gL<sup>-1</sup> in 50 mM succinate-Tris pH 5.0 at 30°C for 5 min; pNP-glucose or pNP-maltose (40 mM in water) was independently added to each strain at a final concentration of 5 mM. After 1-min interval, 100- $\mu$ L aliquots were sampled and immediately placed in boiling water bath for 3 min. After cooling at room temperature, one volume of 2 M of NaHCO<sub>3</sub> was added to each sample, and the cells were removed by centrifugation, and the *p*-nitrophenol present in supernatant was measured at 400 nm. The experiments were performed in triplicate, and controls of previously boiled cells were used. Student's *t*-test was performed in R version 4.2.3.

### DNA and RNA extraction

DNA purification was carried out as previously mentioned (74). Briefly, the SpIB1 and SpIB2 yeasts were inoculated in YP-M medium, and after 48 h, 2-mL aliquots of cells were spun down at 3,000 rpm; 400  $\mu$ L of sterile extraction buffer was added and incubated at 65°C for 10 min; 130  $\mu$ L of 3 M sodium acetate pH5.2 was added in each sample followed vortexing for 30 s and incubated at -20°C. Finally, the lysate was centrifugated

at 13,000 rpm at 4 °C for 15 min; the supernatant was precipitated with isopropanol. For RNA extraction, the hot phenol protocol was performed (75), with the only difference being the use of AES solution (50 mM sodium acetate, 10 mM EDTA, and 0.5% SDS) instead of the proposed TES solution. Briefly, 10-mL samples from first and second fermentation days were centrifugated for 5 min at 3,000 rpm at 4 °C; the cell pellets were washed with 5 mL of sterile cold water and centrifugated at 3,000 rpm at 4 °C for 5 min; the supernatant was removed, and the cell pellet was frozen in liquid nitrogen and stored at -70 °C until use. The frozen cell pellet was thawed on ice and resuspended in 400 µL of AE solution; 400 µL of acid phenol was added followed by 30 s vortexing; the extraction samples were incubated for 60 min at 65 °C and then centrifugated for 5 min at 13,000 rpm at 4 °C. The supernatant was placed in new sterile microcentrifuge tubes, and a second hot acid phenol extraction procedure was performed. The resultant supernatant was treated with one volume of chloroform and centrifugated for 5 min at 13,000 rpm at 4 °C; the resulting supernatant was precipitated with two volumes of ice-cold isopropanol and 40 µL of 3 M sodium acetate at pH 5.3, then ice-cold incubated for 60 min. Then, the samples were centrifugated for 10 min at 13,000 rpm at 4 °C, and the total RNA pellet was washed with 70% ice-cold ethanol in diethyl pyrocarbonate (DPCE)-treated sterile water. Finally, the resultant washed pellet was resuspended in 50 µL DPCE-treated water, seen in agarose gel electrophoresis, and concentrations were determined in nanodrop equipment.

### **Molecular detection of permeases in each strain**

For the genomic detection of gene encoding *Malx1*, *AGT1*, *MTT1*, and *MPH2/3* permeases in both genomes, primers, and PCR experiments were performed as described previously (24). Briefly, the following PCR program was used: pre-incubation (95 °C for 5 min), amplification cycle repeated 35 times (95 °C for 10 s, Tm °C for each specific primer for 10 s, 72 °C for 10 s).

A reverse transcription PCR (RT-PCR) experiment was used to detect the *Malx1*, *AGT1*, *MTT1*, and *MPH2/3* transcripts. Briefly, a commercial reverse transcriptase (M-MLV Reverse Transcriptase, Promega) was used, following the manufacturer instructions, and using the same PCR conditions as mentioned above (24). The primers for *ACT1*, *TAF10*, and *TFC1* were used as reference control housekeeping genes (76).

### **Wort and alpha glycoside consumption analytics**

Wort samples were analyzed by HPLC Series 1200 Agilent Technologies using infrared detector (IR). The mobile phase was H<sub>2</sub>SO<sub>4</sub> (5 mM). The flow rate was established to 1 mL/min using carbohydrate column Aminex HPX-87C Hi-Plex Na, 300 × 7.7 mm, and the column was equilibrated with 5 mM H<sub>2</sub>SO<sub>4</sub> in water at 55 °C (24). The experiments were done in triplicates and the Student's *t*-test was done for statistical significance in the difference of maltose and maltotriose uptake means.

### **ACKNOWLEDGMENTS**

We thank Viktor Boer (HEINEKEN Research & Development) and Toni Gabaldón (Institute for Research in Biomedicine, Barcelona) for critical inputs and reading of the manuscript. In addition, we thank “Cervecería Cuauhtémoc Moctezuma” for kindly providing us the yeast strains and wort used in this work.

E.R.P.-O., L.C.D.-B., B.P.-A., and C.I.H.-V. conceived general idea and project administration. J.H.G.-G. and C.I.H.-V. conceived and designed experiments. C.I.H.-V., J.H.G.-G., and A.G.M.-S. performed the experiments. B.P.-A., J.H.G.-G., and C.I.H.-V. analyzed the data. B.P.-A., J.H.G.-G., and C.I.H.-V. wrote the paper.

The study outlined in this article received financial support from “Programa de Apoyo a la Ciencia, Tecnología e Innovación ProACTI 2023” under the project number 3-BQ-2023.

## AUTHOR AFFILIATIONS

<sup>1</sup>Facultad de Ciencias Biológicas, Universidad Autónoma de Nuevo León, Instituto de Biotecnología, Nuevo León, Mexico

<sup>2</sup>Department of Research and Development, Cervecería Cuauhtémoc Moctezuma, Nuevo León, Mexico

## AUTHOR ORCIDS

César I. Hernández-Vásquez  <http://orcid.org/0000-0002-2067-1122>

Benito Pereyra-Alfárez  <http://orcid.org/0000-0002-3726-0389>

## FUNDING

Funder	Grant(s)	Author(s)
Universidad Autonoma de Nuevo Leon	3-BQ-2023	César I. Hernández-Vásquez Jorge H. García-García Adriana G. Martínez-Segundo Luis C. Damas-Buenrostro Benito Pereyra-Alfárez
Universidad Autonoma de Nuveo Leon	3-BQ-2023	Esmeralda R. Pérez-Ortega

## AUTHOR CONTRIBUTIONS

César I. Hernández-Vásquez, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Writing - original draft, Writing - review and editing | Jorge H. García-García, Data curation, Formal analysis, Methodology, Validation, Writing - review and editing | Esmeralda R. Pérez-Ortega, Conceptualization, Funding acquisition, Project administration, Resources, Validation | Adriana G. Martínez-Segundo, Data curation, Methodology | Luis C. Damas-Buenrostro, Conceptualization, Funding acquisition, Project administration, Resources, Validation | Benito Pereyra-Alfárez, Conceptualization, Formal analysis, Funding acquisition, Project administration, Resources, Validation, Writing - review and editing

## DATA AVAILABILITY

The data are available on the GEO NCBI platform, and the accession number for the data is [GSE269873](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE269873).

## ADDITIONAL FILES

The following material is available [online](#).

### Supplemental Material

**Supplemental material (AEM00397-24-s0001.docx).** Figures S1 to S6; Tables S1 to S5.

## REFERENCES

- Baiano A. 2021. Craft beer: an overview. *Compr Rev Food Sci Food Saf* 20:1829-1856. <https://doi.org/10.1111/1541-4337.12693>
- Tamang JP, Watanabe K, Holzapfel WH. 2016. Review: diversity of microorganisms in global fermented foods and beverages. *Front Microbiol* 7:377. <https://doi.org/10.3389/fmicb.2016.00377>
- Dimidi E, Cox SR, Rossi M, Whelan K. 2019. Fermented foods: definitions and characteristics, impact on the gut microbiota and effects on gastrointestinal health and disease. *Nutrients* 11:1806. <https://doi.org/10.3390/nu11081806>
- Petruzzi L, Rosaria Corbo M, Sinigaglia M, Bevilacqua A. 2016. Brewer's yeast in controlled and uncontrolled fermentations, with a focus on novel, nonconventional, and superior strains. *Food Rev Int* 32:341-363. <https://doi.org/10.1080/87559129.2015.1075211>
- Bonatto D. 2021. The diversity of commercially available ale and lager yeast strains and the impact of brewer's preferential yeast choice on the fermentative beer profiles. *Food Res Int* 141:110125. <https://doi.org/10.1016/j.foodres.2021.110125>
- Baker EC, Wang B, Bellora N, Peris D, Hulfachor AB, Koshalek JA, Adams M, Libkind D, Hittinger CT. 2015. The genome sequence of *Saccharomyces eubayanus* and the domestication of lager-brewing yeasts. *Mol Biol Evol* 32:2818-2831. <https://doi.org/10.1093/molbev/msv168>

1. Libkind D, Hittinger CT, Valério E, Gonçalves C, Dover J, Johnston M, Gonçalves P, Sampaio JP. 2011. Microbe domestication and the identification of the wild genetic stock of lager-brewing yeast. *Proc Natl Acad Sci U S A* 108:14539-14544. <https://doi.org/10.1073/pnas.1105430108>
2. Peris D, Langdon QK, Moriarty RV, Sylvester K, Bontrager M, Charron G, Leducq JB, Landry CR, Libkind D, Hittinger CT. 2016. Complex ancestries of lager-brewing hybrids were shaped by standing variation in the wild yeast *Saccharomyces eubayanus*. *PLoS Genet* 12:e1006155. <https://doi.org/10.1371/journal.pgen.1006155>
3. Gibson BR, Storgårds E, Krogerus K, Vidgren V. 2013. Comparative physiology and fermentation performance of Saaz and Froberg lager yeast strains and the parental species *Saccharomyces eubayanus*. *Yeast* 30:255-266. <https://doi.org/10.1002/yea.2960>
4. Gallone B, Mertens S, Gordon JL, Maere S, Verstrepen KJ, Steensels J. 2018. Origins, evolution, domestication and diversity of *Saccharomyces* beer yeasts. *Curr Opin Biotechnol* 49:148-155. <https://doi.org/10.1016/j.copbio.2017.08.005>
5. Monerawela C, James TC, Wolfe KH, Bond U. 2015. Loss of lager specific genes and subtelomeric regions define two different *Saccharomyces cerevisiae* lineages for *Saccharomyces pastorianus* group I and II strains. *FEMS Yeast Res* 15:fou008. <https://doi.org/10.1093/femsyr/fou008>
6. Salazar AN, Gorter de Vries AR, van den Broek M, Brouwers N, de la Torre Cortés P, Kuijpers NGA, Daran J-M, Abeel T. 2019. Chromosome level assembly and comparative genome analysis confirm lager-brewing yeasts originated from a single hybridization. *BMC Genomics* 20:916. <https://doi.org/10.1186/s12864-019-6263-3>
7. Gallone B, Steensels J, Prah T, Soriaga L, Saels V, Herrera-Malaver B, Merlevede A, Roncoroni M, Voordeckers K, Miraglia L, Teiling C, Steffy B, Taylor M, Schwartz A, Richardson T, White C, Baele G, Maere S, Verstrepen KJ. 2016. Domestication and divergence of *Saccharomyces cerevisiae* beer yeasts. *Cell* 166:1397-1410. <https://doi.org/10.1016/j.cell.2016.08.020>
8. Herbst RH, Bar-Zvi D, Reikhav S, Soifer I, Breker M, Jona G, Shimoni E, Schuldiner M, Levy AA, Barkai N. 2017. Heterosis as a consequence of regulatory incompatibility. *BMC Biol* 15:38. <https://doi.org/10.1186/s12915-017-0373-7>
9. Meneses FJ, Jiranek V. 2002. Expression patterns of genes and enzymes involved in sugar catabolism in industrial *Saccharomyces cerevisiae* strains displaying novel fermentation characteristics. *J Inst Brew* 108:322-335. <https://doi.org/10.1002/j.2050-0416.2002.tb00557.x>
10. He Y, Dong J, Yin H, Zhao Y, Chen R, Wan X, Chen P, Hou X, Liu J, Chen L. 2014. Wort composition and its impact on the flavour-active higher alcohol and ester formation of beer - a review. *J Inst Brew* 120:157-163. <https://doi.org/10.1002/jib.145>
11. Stambuk BU, Alves SL, Hollatz C, Zastrow CR. 2006. Improvement of maltotriose fermentation by *Saccharomyces cerevisiae*. *Lett Appl Microbiol* 43:370-376. <https://doi.org/10.1111/j.1472-765X.2006.01982.x>
12. Gancedo JM. 1998. Yeast carbon catabolite repression. *Microbiol Mol Biol Rev* 62:334-361. <https://doi.org/10.1128/MMBR.62.2.334-361.1998>
13. Kayikci Ö, Nielsen J. 2015. Glucose repression in *Saccharomyces cerevisiae*. *FEMS Yeast Res* 15:1-8. <https://doi.org/10.1093/femsyr/fov068>
14. Horák J. 2013. Regulations of sugar transporters: insights from yeast. *Curr Genet* 59:1-31. <https://doi.org/10.1007/s00294-013-0388-8>
15. Yoon SH, Mukerjee R, Robyt JF. 2003. Specificity of yeast (*Saccharomyces cerevisiae*) in removing carbohydrates by fermentation. *Carbohydr Res* 338:1127-1132. [https://doi.org/10.1016/s0008-6215\(03\)00097-1](https://doi.org/10.1016/s0008-6215(03)00097-1)
16. Zastrow CR, Mattos MA, Hollatz C, Stambuk BU. 2000. Maltotriose metabolism by *Saccharomyces cerevisiae*. *Biotechnol. Lett* 22:455-459. <https://doi.org/10.1023/A:1005691031880>
17. Day RE, Rogers PJ, Dawes IW, Higgins VJ. 2002. Molecular analysis of maltotriose transport and utilization by *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 68:5326-5335. <https://doi.org/10.1128/AEM.68.11.5326-5335.2002>
18. Magalhães F, Vidgren V, Ruohonen L, Gibson B. 2016. Maltose and maltotriose utilisation by group I strains of the hybrid lager yeast *Saccharomyces pastorianus*. *FEMS Yeast Res* 16:fow053. <https://doi.org/10.1093/femsyr/fow053>
19. Dietvorst J, Londesborough J, Steensma HY. 2005. Maltotriose utilization in lager yeast strains: MTT1 encodes a maltotriose transporter. *Yeast* 22:775-788. <https://doi.org/10.1002/yea.1279>
20. Salema-Oom M, Valadão Pinto V, Gonçalves P, Spencer-Martins I. 2005. Maltotriose utilization by industrial *Saccharomyces* strains: characterization of a new member of the  $\alpha$ -glucoside transporter family. *Appl Environ Microbiol* 71:5044-5049. <https://doi.org/10.1128/AEM.71.9.5044-5049.2005>
21. Dietvorst J, Walsh MC, van Heusden GPH, Steensma HY. 2010. Comparison of the MTT1- and MAL31-like maltose transporter genes in lager yeast strains. *FEMS Microbiol Lett* 310:152-157. <https://doi.org/10.1111/j.1574-6968.2010.02056.x>
22. Vidgren V, Kankainen M, Londesborough J, Ruohonen L. 2011. Identification of regulatory elements in the AGT1 promoter of ale and lager strains of brewer's yeast. *Yeast* 28:579-594. <https://doi.org/10.1002/yea.1888>
23. Brouwers N, Gorter de Vries AR, van den Broek M, Weening SM, Elink Schuurman TD, Kuijpers NGA, Pronk JT, Daran J-M. 2019. *In vivo* recombination of *Saccharomyces eubayanus* maltose-transporter genes yields a chimeric transporter that enables maltotriose fermentation. *PLoS Genet* 15:e1007853. <https://doi.org/10.1371/journal.pgen.1007853>
24. Baker EP, Hittinger CT. 2019. Evolution of a novel chimeric maltotriose transporter in *Saccharomyces eubayanus* from parent proteins unable to perform this function. *PLoS Genet* 15:e1007786. <https://doi.org/10.1371/journal.pgen.1007786>
25. Monerawela C, Bond U. 2018. The hybrid genomes of *Saccharomyces pastorianus*: a current perspective. *Yeast* 35:39-50. <https://doi.org/10.1002/yea.3250>
26. Madej MG, Sun L, Yan N, Kaback HR. 2014. Functional architecture of MFS D-glucose transporters. *Proc Natl Acad Sci U S A* 111:E719-E727. <https://doi.org/10.1073/pnas.1400336111>
27. Drew D, North RA, Nagarathinam K, Tanabe M. 2021. Structures and general transport mechanisms by the major facilitator superfamily (MFS). *Chem Rev* 121:5289-5335. <https://doi.org/10.1021/acs.chemrev.0c00983>
28. Yan N. 2013. Structural advances for the major facilitator superfamily (MFS) transporters. *Trends Biochem Sci* 38:151-159. <https://doi.org/10.1016/j.tibs.2013.01.003>
29. Philippe H, Brinkmann H, Lavrov DV, Littlewood DTJ, Manuel M, Wörheide G, Baurain D. 2011. Resolving difficult phylogenetic questions: why more sequences are not enough. *PLoS Biol* 9:e1000602. <https://doi.org/10.1371/journal.pbio.1000602>
30. Brown CA, Murray AW, Verstrepen KJ. 2010. Rapid expansion and functional divergence of subtelomeric gene families in yeasts. *Curr Biol* 20:895-903. <https://doi.org/10.1016/j.cub.2010.04.027>
31. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15:550. <https://doi.org/10.1186/s13059-014-0550-8>
32. Anders S, McCarthy DJ, Chen Y, Okoniewski M, Smyth GK, Huber W, Robinson MD. 2013. Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. *Nat Protoc* 8:1765-1786. <https://doi.org/10.1038/nprot.2013.099>
33. Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26:139-140. <https://doi.org/10.1093/bioinformatics/btp616>
34. Hollatz C, Stambuk BU. 2001. Colorimetric determination of active  $\alpha$ -glucoside transport in *Saccharomyces cerevisiae*. *J Microbiol Methods* 46:253-259. [https://doi.org/10.1016/s0167-7012\(01\)00281-0](https://doi.org/10.1016/s0167-7012(01)00281-0)
35. Henderson R, Poolman B. 2017. Proton-solute coupling mechanism of the maltose transporter from *Saccharomyces cerevisiae*. *Sci Rep* 7:14375. <https://doi.org/10.1038/s41598-017-14438-1>
36. Schaefer C, Rost B. 2012. Predict impact of single amino acid change upon protein structure. *BMC Genomics* 13 Suppl 4:S4. <https://doi.org/10.1186/1471-2164-13-S4-S4>
37. Gong S, Blundell TL. 2010. Structural and functional restraints on the occurrence of single amino acid variations in human proteins. *PLoS One* 5:e9186. <https://doi.org/10.1371/journal.pone.0009186>
38. Triche D, Knychala MM, Figueiredo CM, Alves SL, da Silva MA, Miletti LC, de Araujo PS, Stambuk BU. 2019. Key amino acid residues of the AGT1 permease required for maltotriose consumption and fermentation by *Saccharomyces cerevisiae*. *J Appl Microbiol* 126:580-594. <https://doi.org/10.1111/jam.14161>

1. Smit A, Moses SG, Pretorius IS, Cordero Otero RR. 2008. The Thr505 and Ser557 residues of the AGT1-encoded A-glycoside transporter are critical for maltotriose transport in *Saccharomyces cerevisiae*. *J Appl Microbiol* 104:1103-1111. <https://doi.org/10.1111/j.1365-2672.2007.03671.x>
2. Hatanaka H, Toyonaga H, Ishida Y, Mizohata E, Ono E. 2022. Functional diversity and plasticity in the sugar preferences of *Saccharomyces MALT* transporters in domesticated yeasts. *FEMS Yeast Research* 22. <https://doi.org/10.1093/femsyr/foac055>
3. Vidgren V, Ruohonen L, Londesborough J. 2005. Characterization and functional analysis of the MAL and MPH Loci for maltose utilization in some ale and lager yeast strains. *Appl Environ Microbiol* 71:7846-7857. <https://doi.org/10.1128/AEM.71.12.7846-7857.2005>
4. Vidgren V, Huuskonen A, Virtanen H, Ruohonen L, Londesborough J. 2009. Improved fermentation performance of a lager yeast after repair of its AGT1 maltose and maltotriose transporter genes. *Appl Environ Microbiol* 75:2333-2345. <https://doi.org/10.1128/AEM.01558-08>
5. Smit A, Cordero Otero RR, Pretorius IS. 2007. Differences among AGT1-encoded A-glycoside transporters and their ability to transport maltotriose in *Saccharomyces* yeasts. *Ann. Microbiol* 57:77-84. <https://doi.org/10.1007/BF03175054>
6. Han EK, Cotty F, Sottas C, Jiang H, Michels CA. 1995. Characterization of AGT1 encoding a general alpha-glycoside transporter from *Saccharomyces*. *Mol Microbiol* 17:1093-1107. [https://doi.org/10.1111/j.1365-2958.1995.mmi\\_17061093.x](https://doi.org/10.1111/j.1365-2958.1995.mmi_17061093.x)
7. Stambuk BU, de Araujo PS. 2001. Kinetics of active A-glycoside transport in *Saccharomyces cerevisiae*. *FEMS Yeast Res* 1:73-78. <https://doi.org/10.1111/j.1567-1364.2001.tb00015.x>
8. Alves SL, Herberts RA, Hollatz C, Trichez D, Miletti LC, de Araujo PS, Stambuk BU. 2008. Molecular analysis of maltotriose active transport and fermentation by *Saccharomyces cerevisiae* reveals a determinant role for the AGT1 permease. *Appl Environ Microbiol* 74:1494-1501. <https://doi.org/10.1128/AEM.02570-07>
9. Wei X, Zhang C, Freddolino PL, Zhang Y. 2020. Detecting gene ontology misannotations using taxon-specific rate ratio comparisons. *Bioinformatics* 36:4383-4388. <https://doi.org/10.1093/bioinformatics/btaa548>
10. Nobre T, Campos MD, Lucic-Mercy E, Arnholdt-Schmitt B. 2016. Misannotation awareness: a tale of two gene-groups. *Front Plant Sci* 7:868. <https://doi.org/10.3389/fpls.2016.00868>
11. De León-Medina PM, Elizondo-González R, Damas-Buenrostro LC, Geertman J-M, Van den Broek M, Galán-Wong LJ, Ortiz-López R, Pereyra-Alferez B. 2016. Genome annotation of a *Saccharomyces* sp. lager brewer's yeast. *Genom Data* 9:25-29. <https://doi.org/10.1016/j.gdata.2016.05.009>
12. Quinlan AR, Hall IM. 2010. Bedtools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26:841-842. <https://doi.org/10.1093/bioinformatics/btq033>
13. Bodenhofer U, Bonatesta E, Horejš-Kainrath C, Hochreiter S. 2015. Msa: an R package for multiple sequence alignment. *Bioinformatics* 31:3997-3999. <https://doi.org/10.1093/bioinformatics/btv494>
14. Paradis E, Claude J, Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20:289-290. <https://doi.org/10.1093/bioinformatics/btg412>
15. Schliep KP. 2011. phangorn: phylogenetic analysis in R. *Bioinformatics* 27:592-593. <https://doi.org/10.1093/bioinformatics/btq706>
16. Dobson L, Reményi I, Tusnády GE. 2015. CCTOP: a consensus constrained TOPology prediction web server. *Nucleic Acids Res* 43:W408-W412. <https://doi.org/10.1093/nar/gkv451>
17. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, et al. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature* 596:583-589. <https://doi.org/10.1038/s41586-021-03819-2>
18. Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, Yuan D, Stroe O, Wood G, Laydon A, et al. 2022. AlphaFold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res* 50:D439-D444. <https://doi.org/10.1093/nar/gkab1061>
19. Monteiro PT, Oliveira J, Pais P, Antunes M, Palma M, Cavalheiro M, Galocha M, Godinho CP, Martins LC, Bourbon N, Mota MN, Ribeiro RA, Viana R, Sá-Correia I, Teixeira MC. 2020. YEASTRACT+: a portal for cross-species comparative genomics of transcription regulation in yeasts. *Nucleic Acids Res* 48:D642-D649. <https://doi.org/10.1093/nar/gkz859>
20. Hernández-Vásquez CI. 2020. Análisis del transcriptoma, basado en RNA-seq, de una levadura cervecera tipo lager en fermentación a nivel industrial y laboratorio. San Nicolás de los Graza, N.L. México Universidad Autónoma de Nuevo León
21. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114-2120. <https://doi.org/10.1093/bioinformatics/btu170>
22. Kopylova E, Noé L, Touzet H. 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* 28:3211-3217. <https://doi.org/10.1093/bioinformatics/bts611>
23. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, MacManes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, Williams T, Dewey CN, Henschel R, LeDuc RD, Friedman N, Regev A. 2013. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc* 8:1494-1512. <https://doi.org/10.1038/nprot.2013.084>
24. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. 2011. Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nat Biotechnol* 29:644-652. <https://doi.org/10.1038/nbt.1883>
25. Bryant DM, Johnson K, DiTommaso T, Tickle T, Couger MB, Payzin-Dogru D, Lee TJ, Leigh ND, Kuo TH, Davis FG, Bateman J, Bryant S, Guzikowski AR, Tsai SL, Coyne S, Ye WW, Freeman RM, Peshkin L, Tabin CJ, Regev A, Haas BJ, Whited JL. 2017. A tissue-mapped axolotl *de novo* transcriptome enables identification of limb regeneration factors. *Cell Rep* 18:762-776. <https://doi.org/10.1016/j.celrep.2016.12.063>
26. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357-359. <https://doi.org/10.1038/nmeth.1923>
27. Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. *BMC Bioinformatics* 12:323. <https://doi.org/10.1186/1471-2105-12-323>
28. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29:15-21. <https://doi.org/10.1093/bioinformatics/bts635>
29. Anders S, Pyl PT, Huber W. 2015. HTSeq-a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31:166-169. <https://doi.org/10.1093/bioinformatics/btv638>
30. Osama EA, Mohamed AM, AbdEl Rahim M, Shaban RMS. 2011. Non liquid nitrogen-based-method for isolation of DNA from filamentous fungi. *Afr J Biotechnol* 10:14337-14341. <https://doi.org/10.5897/AJB11.1401>
31. Green MR, Sambrook J. 2021. Total RNA extraction from *Saccharomyces cerevisiae* using hot acid phenol. *Cold Spring Harb Protoc* 2021:db. <https://doi.org/10.1101/pdb.prot101691>
32. Teste MA, Duquenne M, François JM, Parrou JL. 2009. Validation of reference genes for quantitative expression analysis by real-time RT-PCR in *Saccharomyces cerevisiae*. *BMC Mol Biol* 10:99. <https://doi.org/10.1186/1471-2199-10-99>

**9. CAPÍTULO 2. Datos de RNA-seq sugieren un efecto en la expresión genética de dos levaduras lager causado por la presencia de RNA en el mosto cervecero**

Este capítulo ha sido preparado para su publicación en la revista **Scientific Reports** de la editorial **Nature**. Actualmente se encuentra en el estatus de enviado y se está a la espera de las primeras observaciones. El trabajo fue sometido con el título en inglés siguiente:

**“RNA-seq transcriptome data support effect on genetic expression in two industrial yeasts caused by RNA present in brewing wort”**

## I Resumen

La comunicación entre reinos mediada por ARN sugiere que los ARN no codificantes (ARNnc) de un organismo, como las plantas, pueden regular la expresión génica en otros organismos, incluidos hongos, bacterias y parásitos.

En este estudio, exploramos la hipótesis de que el ARN, potencialmente presente en el mosto de cerveza originado a partir de cebada y/o lúpulo, influyen en la expresión génica en dos cepas de levadura lager industrial. Después del tratamiento con ARNasa, observamos efectos importantes en la expresión génica en el segundo día de fermentación, un total de 671 genes DE (expresados diferencialmente) en SpIB1 y 536 en SpIB2. En el quinto día, observamos un efecto reducido con un total de 64 genes DE en SpIB1 y 29 en SpIB2.

En ambas levaduras, observamos en el segundo día de fermentación una inducción del ciclo celular y la floculación por la presencia de ARN en el mosto de cerveza, mientras que se observó una posible represión del regulador transcripcional *ARO80* para el crecimiento de pseudohifas.

Al quinto día, el metabolismo de los aminoácidos se vio afectado negativamente por la presencia de ARN en el medio, lo que indica una posible reducción de la aptitud en la última etapa de la fermentación. Estos resultados indican un posible efecto del ARN presente en el mosto sobre la expresión genética de dos levaduras industriales y respaldan la hipótesis de la comunicación entre reinos mediada por ARNnc.

## II Introducción

Existen funciones básicas bien caracterizadas que involucran al ARN no codificante (ARNnc). Estos procesos biológicos van desde la regulación transcripcional y el procesamiento del ARN hasta la interferencia mediada por ARN (miARN). Los principales tipos de ARNnc incluyen el ARN de transferencia (ARNt), las ribozimas, el microARN (miARN) y el ARN no codificante largo (ARNlnc).

Los ARN no codificantes largos (ARNlnc) son particularmente importantes para los eventos de silenciamiento génico, especialmente en organismos eucariotas (Cech & Steitz, 2014). Estos eventos biológicos son cruciales para el silenciamiento de transposones, la defensa viral, la compactación de la heterocromatina y la regulación postranscripcional de los genes codificantes de proteínas (Craig y Darryl, 2004; Dang *et al.*, 2011).

La interferencia de ARN (ARNi) es un proceso eucariota conservado que requiere pequeños ARN no codificantes acoplados a proteínas como Argonaute/Piwi, un componente nuclear del complejo de silenciamiento inducido por ARN (de las siglas en inglés RISC). Este complejo interviene a través de la complementariedad de bases entre el ARN y los genes que codifican a proteínas (Cerutti & Casas-Mollano, 2006; Meister & Tuschl, 2004).

Existe evidencia acumulada que sugiere que los ncRNAs pueden transferirse dentro de células y tejidos e incluso a través de especies, dominios y reinos (Leitão y Enguita, 2016; Zhou *et al.*, 2017). Se ha encontrado una relación bidireccional en la transferencia de ncRNA entre plantas y hongos en contextos fitopatológicos.

Esta evidencia muestra que los ncRNAs fúngicos pueden regular negativamente los sistemas inmunes de las plantas, mientras que los ncRNAs de las plantas pueden reducir los fenotipos patógenos de ciertos hongos (Weiberg y Jin, 2015; Zeng *et al.*, 2019).

Aunque la relación planta-hongo está menos documentada, investigaciones previas indican que las plantas de algodón producen ncRNAs específicos, que exportan a las hifas de los hongos, induciendo el silenciamiento de genes que reduce el fenotipo patógeno del hongo (Zhang *et al.*, 2016). Este tipo de evidencia nos llevó a plantear la hipótesis de que las levaduras de importancia industrial pueden capturar ARN exógeno y regular su expresión genética en consecuencia.

Es probable que estos eventos de regulación genética mediados por ncRNA ocurran en entornos antropogénicos, como la manipulación microbiológica de la levadura en el mosto de cerveza. Si bien el uso sistemático de levaduras lager en procesos industriales es relativamente reciente, el uso de levaduras ale se remonta a entre 6000 y 13000 años atrás (Cabras y Higgins, 2016; Hornsey, 2003; Liu *et al.*, 2018).

En particular, el uso de levaduras híbridas interespecíficas, como *Saccharomyces pastorianus* (*Sp*), es común en la industria cervecera. El origen exacto de estos híbridos no se conoce con certeza, pero se especula que pueden haber resultado de la contaminación de las fermentaciones de *Saccharomyces cerevisiae* (*Sc*) con levaduras de la especie *Saccharomyces eubayanus* (*Se*) (Gibson & Liti, 2015; Krogerus *et al.*, 2015; Libkind *et al.*, 2011).

Esta levadura es un ejemplo de domesticación exitosa de híbridos interespecíficos ya que no puede sobrevivir en el entorno natural, lo que la hace completamente dependiente de la propagación humana. En este contexto, la levadura no conoce otro entorno que el del mosto de cerveza, que ha dado forma a su trasfondo evolutivo. Los únicos desencadenantes del cambio son el propio mosto de cerveza y las condiciones de propagación y fermentación utilizadas en el proceso de elaboración de la cerveza (Dunn y Sherlock, 2008; Libkind *et al.*, 2011).

Como se ha mencionado anteriormente, una de las interacciones importantes que pueden haber desarrollado las levaduras cerveceras es la posible regulación de la expresión génica mediada por ARN exógenos presentes en el mosto de cerveza. Esta hipótesis se basa también en la evidencia disponible

que indica que el arquetipo de interferencia mediada por ARN de doble cadena (ARNi) está evolutivamente ausente en la levadura en ciernes *S. c*

A pesar de esto, este sistema de interferencia está presente en otros organismos como *Saccharomyces castelli* y *Candida albicans*, donde el proceso se lleva a cabo por vías alternativas a Dicer (Dang *et al.*, 2011; Drinnenberg *et al.*, 2009; Nakayashiki *et al.*, 2006).

La pérdida evolutiva de este sistema de interferencia en las levaduras gemantes representa un aspecto importante en nuestra investigación, ya que se ha demostrado experimentalmente que nuestras cepas son híbridos interespecíficos funcionales entre *Sc* y *Se* (León-Medina *et al.*, 2016).

No obstante, esto no implica que el sistema de silenciamiento génico esté completamente ausente en las levaduras en ciernes. En respuesta a la falta de RNAi, levaduras como *Sc* han experimentado una expansión específica del sistema de ARN antisentido largo no codificante (ASlncRNA) y otros ARN no codificantes importantes que realiza funciones sinónimas del RNAi (Alcid & Tsukiyama, 2016; Balarezo-Cisneros *et al.*, 2021).

Aunque existe cierto debate sobre la funcionalidad del sistema de interferencia del ARN antisentido, con indicios de que no todos los ARN antisentido desempeñan un papel en el silenciamiento y/o reducción de la abundancia de proteínas en levaduras (Huber *et al.*, 2016), la evidencia apoya la idea de que el ARN antisentido es importante en la regulación de la expresión génica.

Por ejemplo, el ARN antisentido está involucrado en el control de la entrada en el ciclo celular meiótico y la dinámica molecular de la pared celular en *Sc* (Hongay *et al.*, 2006; Novačić *et al.*, 2020). Es posible que debido a la ausencia de un sistema de interferencia mediado por ARN pequeño y la presencia de un sistema regulador genético mediado por ARN largo no codificante (lncRNA), nuestras levaduras hayan desarrollado una forma de comunicación funcional con los componentes presentes en los medios de cultivo a base de malta de cebada.

Esto podría amortiguar potencialmente la ausencia del sistema regulador mediado por RNAi. Aunque el concepto de comunicación entre reinos mediada por ncRNA es controvertido, la investigación pionera sobre este tema nos ha llevado a explorar la hipótesis de que una comunicación molecular similar entre reinos ocurre en entornos antropogénicos, como el manejo microbiológico de levaduras cerveceras.

Como resultado de los datos presentados anteriormente, decidimos probar la hipótesis de que los cambios en la expresión genética en levaduras lager son inducidos por la transferencia horizontal de ARN no codificante. Para ello, empleamos tecnología de secuenciación de ARN y enfoques computacionales.

### **III Materiales y métodos**

#### **a) Cepas**

Las cepas utilizadas en este estudio fueron cordialmente brindadas por CCM. Estas levaduras son utilizadas en procesos industriales y forman parte de la especie *Sp* y la nomenclatura en este artículo corresponde a SpIB1 y cepa SpIB2, esto con base en su capacidad de utilizar alfa-glucósidos expuestos en la clasificación fenotípica del grupo I, Saaz, y grupo II, Froghber, respectivamente (Gibson *et al.*, 2013).

Estas cepas fueron mantenidas en un medio YP-M (1% extracto de levadura, 2% peptona y 2% maltosa) con 40% de glicerol a -20 °C después de 48 h en YP-M líquido o al alcanzar fase estacionaria hasta su uso.

#### **b) Diseño experimental**

Para probar la hipótesis de posibles cambios en la expresión génica en las dos cepas de levaduras lager inducidas por la transferencia horizontal de ARN

presente en el mosto de cerveza, fue implementado un experimento con tratamiento enzimático (ARNasa).

El experimento consistió en 4 matraces cilíndricos con 450 mL de mosto estándar de 16 °P (140 ppm de nitrógeno amínico libre, 0,11% [p/p] de fructosa, 4,80% [p/p] de glucosa, 11% [p/p] de maltosa y 3% [p/p] de maltotriosa). Dos de ellos tratados con 5 mL de ARNasa (ARNasa A Bio Basic, 50 kU/mg) y los dos restantes como control. Luego, los 4 matraces, 2 tratamientos y 2 controles, fueron incubados con agitación a 120 rpm durante 4 h a temperatura ambiente.

Después del tratamiento, 200 mL de cada matraz fueron distribuidos en cuatro matraces cilíndricos de 1 L e inoculados a una tasa de  $1,5 \times 10^6$  células/ml/°P de cada levadura (SplB1 y SplB2) (Thesseling *et al.*, 2019). Resultando en 8 fermentaciones, 2 repeticiones del tratamiento con ARNasa y 2 repeticiones del control para cada cepa utilizada.

Luego, las condiciones de fermentación fueron establecidas a 16°C por 5 días. Las muestras del material biológico fueron recuperadas el segundo y quinto día de fermentación. Las comparaciones estadísticas fueron dentro de cada cepa; tratamiento el día 2 contra control el día 2 y tratamiento el día 5 contra control el día 5.

### **c) Manipulación del material biológico y extracción de ARN total**

Los muestreos de los días 2 y 5 de la fermentación consistieron en 15 mL de material biológico el cual fue centrifugado a 3000 rpm por 5 min. Luego cada pellet celular fue resuspendido en un buffer 5 mL de PBS e inmediatamente congelado en nitrógeno líquido y almacenado hasta su uso.

Para la extracción de ARN, un protocolo previamente reportado fue implementado (Rio *et al.*, 2010), con algunas modificaciones. En resumen; el material biológico fue descongelado en baño metabólico a 37°C y centrifugado a

3000 rpm por 5 min. Luego 3 volúmenes de TriZol fueron adicionados para solubilizar la muestra durante 10 min a temperatura ambiente (TA).

Para la formación de la fase 600 µl de cloroformo fueron agregados, fue aplicado un vórtice durante 15 segundos y una posterior incubación durante 5 min a TA. Luego fue centrifugado a 10000 g durante 15 min a TA y la fase acuosa fue recuperada. Fueron añadidos 1,5 mL de isopropanol, seguido de agitación suave y una incubación durante 10 min a TA.

El resultado fue centrifugado a 12000 g durante 10 min y descartado el sobrenadante. Luego, 6 mL de etanol al 75% en agua estéril tratada con dietil pirocarbonato (DEPC) fue adicionada, mezclada y centrifugada a 7500 g durante 5 min. El etanol fue eliminado y el procedimiento fue repetido una vez más. Las muestras de ARN total en etanol al 75% en agua-DEPC fueron enviadas a la empresa GENEWIZ para la secuenciación del ARN.

#### **d) Secuenciación de ARN y procesamiento de datos**

Las bibliotecas de ARN fueron preparadas con selección de poli A y secuenciadas en una máquina Illumina HiSeq 2x150 pb. Lo que obtuvimos fueron 16 archivos fasta con lecturas pareadas (F y R) con una calidad de 80% por encima del valor Phred 30.

Para obtener datos de alta calidad, las lecturas fueron evaluadas con el software FASTQC (disponible en: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>), luego la eliminación de lecturas de baja calidad, recorte de adaptadores y limpieza de ARN ribosomal con el software fue realizado con los programas Trimmomatic (Bolger *et al.*, 2014) y SortMeRNA (Kopylova *et al.*, 2012), respectivamente.

Para revisar los datos procesados, FASTQC fue empleado para obtener una visión generalizada de los posibles cambios en los datos de alta calidad generados (Delhomme *et al.*, 2014).

Una vez que los datos pasaron los controles de calidad descritos anteriormente, un ensamblaje *de novo* con el uso de la herramienta Trinity versión 2.8.4 fue implementado, utilizando los parámetros por defecto, modificando únicamente el número de hilos de cómputo y la cantidad de memoria RAM a utilizar (Haas *et al.*, 2013; Raychowdhury *et al.*, 2011). Los transcritos ensamblados fueron posteriormente anotados con el software Trinotate (Bryant *et al.*, 2017).

#### **e) Estimación de abundancia de transcritos, expresión génica diferencial, ontología génica y análisis de vías metabólicas KEGG**

En primera instancia el software Bowtie2 fue empleado para el alineamiento de las lecturas procesadas contra el transcriptoma ensamblado, esto con los parámetros por defecto en el programa (Langmead & Salzberg, 2012). Para estimar la abundancia de transcritos el software RSEM fue utilizado con los parámetros por defecto (Li & Dewey, 2011).

Para obtener los genes expresados diferencialmente en las comparaciones establecidas entre el tratamiento y el control el software edgeR fue implementado (Robinson *et al.*, 2010). Los genes expresados diferencialmente (GED's) obtenidos para cada comparación y levaduras siguieron el criterio de al menos 2 veces su abundancia,  $\log_2(2) = 1$ , con un valor  $p$  ajustado menor a 0,05 (FDR <0,05).

Para el análisis ontológico y de vías metabólicas de los GED's el programa ClusterProfiler fue utilizado (Yu *et al.*, 2012) junto con los datos de anotaciones de *Scerevisiae* 288C, parte importante del paquete org.Sc.sgd.db (disponible en: <https://bioconductor.org/packages/release/data/annotation/html/org.Sc.sgd.db.html>).

El método para ajustar los p-valores, necesarios en los múltiples tests realizados en la expresión global y el análisis ontológico, fue el procedimiento de Benjamini-Hochberg.

## **IV Resultados**

### **El ARN presente en el mosto de cerveza induce cambios en la expresión global en dos levaduras lager**

Para obtener los conjuntos de genes afectados por el ARN presente en el mosto de cerveza, se analizaron los perfiles de expresión de SpIB1 y SpIB2 en fermentaciones lager con tratamientos de ARNasa y controles sin tratamiento.

Para determinar los patrones de agrupamiento de la expresión global entre las fermentaciones de cada levadura y sus condiciones, se realizó un análisis de componentes principales (PCA) de las bibliotecas de RNAseq.

Cuando comparamos las réplicas, observamos, como se esperaba, un agrupamiento global entre las muestras de control para ambas levaduras. Hasta cierto punto, se observa cierta varianza en las muestras de tratamiento, especialmente en el segundo día de fermentación de ambas levaduras.

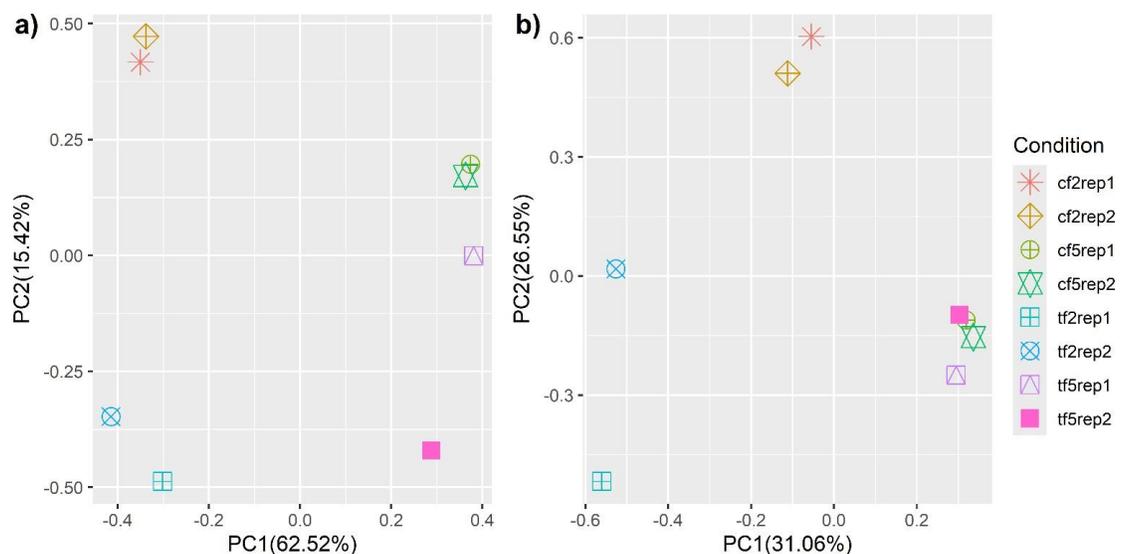
En este sentido, observamos en la Figura 14 un efecto provocado por el tratamiento en el segundo día de fermentación para ambas levaduras. Aunque, en la cepa SpIB1 observamos que, en el PC1 que explica el 62% de la varianza observada en los datos, las muestras del segundo día no exhiben una dispersión considerable, pero en la PC2, que explica el 15% de la varianza, observamos una distancia considerable entre las muestras control y tratamiento.

Para las muestras obtenidas en el quinto día de fermentación de SpIB1, observamos que en la proyección de la PC1 se detectó una dispersión reducida, y en la proyección de la PC2 observamos algunas de las varianzas entre las muestras control y tratamiento. Para la levadura SpIB2 observamos cierta dispersión en la proyección del PC1, que explica el 31% de la varianza, entre las muestras control y tratamiento del segundo día de fermentación.

En estas mismas muestras, también observamos cierto grado de variación en la proyección del PC2, que explica el 26% de la varianza. Por último, se

detectó un patrón de agrupamiento considerable, en la proyección de ambos componentes principales, entre las muestras control y tratamiento a partir del quinto día de fermentación, lo que significa que el patrón de expresión global muestra una variabilidad reducida en estas muestras.

En conjunto, estos resultados sugieren que el tratamiento enzimático con ARNasa del mosto de cerveza tiene un efecto en el transcriptoma de la levadura lager analizada, siendo el efecto evidente en el segundo día de fermentación, pero menor en el quinto día.



**Figura 14.** Análisis de componentes principales (PCA) de las librerías de ARNasas de **a)** SpIB1 y **b)** SpIB2 en los diferentes días de recolección de muestras y tratamiento. En la leyenda de la figura, la primera letra indica c = muestras de control o t = muestras de tratamiento con ARNasas, seguida de f2 o f5, que significan el segundo y quinto día de fermentación, y por último las repeticiones del experimento.

Para obtener el número de genes expresados diferencialmente (DG) en cada condición y levadura, analizamos estos genes bajo el criterio de que tienen el doble del cambio (valor absoluto del fold change  $\geq 2$  y FDR  $< 0,05$ ). En el SpIB1 en el segundo día de fermentación, se observó un total de 671 genes DE los cuales 250 estaban regulados negativamente (puntos rojos en la Fig. 15a), 133 genes DE estaban regulados positivamente (puntos azules en la Fig. 15a) y 288

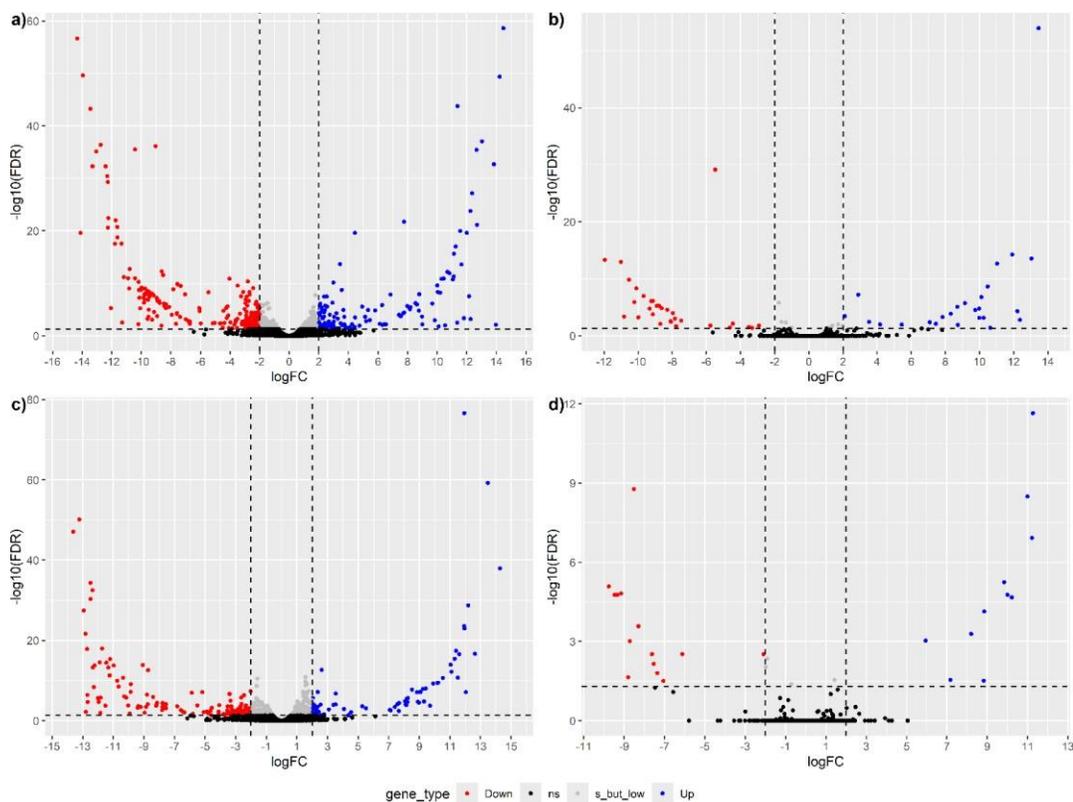
mostraron un cambio significativo, pero de baja expresión (puntos grises en la Fig. 15a).

En el quinto día de fermentación para la misma levadura, observamos un total de 64 genes DE donde 30 mostraron una regulación negativa (puntos rojos en la Fig. 15b) y 25 genes DE estaban regulados positivamente (puntos azules en la Fig. 15b) y, 9 genes DE se observaron como significativos, pero con un cambio de baja expresión (puntos grises en la Fig. 15b).

Por otra parte, la cepa SpIB2 mostró un total de 536 genes DE en el segundo día de fermentación, donde 135 genes DE fueron regulados a la baja (puntos rojos en la Fig. 15c), 75 se regularon al alza (puntos azules en la Fig. 15c) y 326 genes DE fueron significativos, pero con un cambio bajo en su expresión (puntos grises en la Fig. 15c).

Finalmente, en el quinto día de fermentación, la cepa SpIB2 mostró un total de 29 genes DE, donde 14 genes DE mostraron una regulación a la baja (puntos rojos en la Fig. 15d), 11 genes DE regulados al alza (puntos azules en la Fig. 15d) y 4 genes DE con un cambio significativo, pero bajo en su expresión (puntos grises en la Fig. 15d).

En conjunto, estos hallazgos indican que los transcriptomas de las dos cepas de levadura lager analizadas en este estudio exhiben una variación significativa atribuible al tratamiento con ARNasa. Esta variación sugiere fuertemente que la presencia de moléculas de ARN influye en los perfiles de expresión génica global de las cepas de levadura bajo investigación.



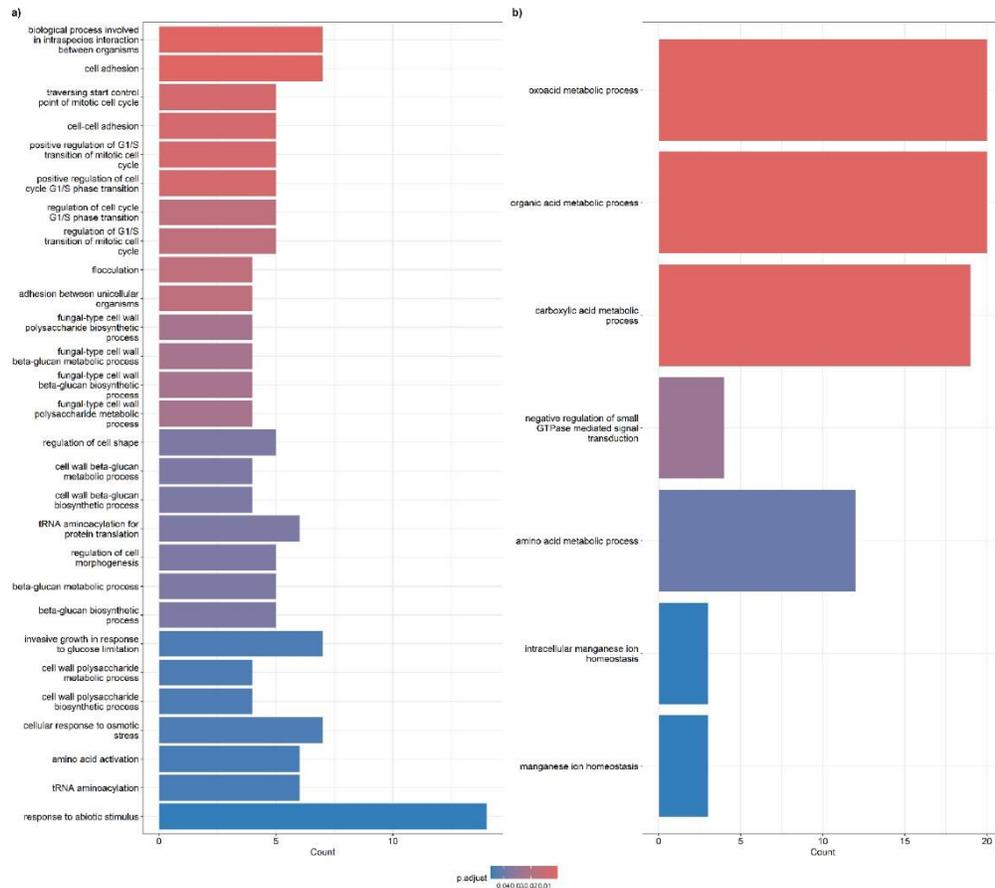
**Figura 15.** Gráficos de volcán que representan los genes DE en el a) segundo y b) quinto día de fermentación para la cepa SpIB1 y, de manera similar, en la cepa SpIB2 se muestran los genes DE obtenidos en el c) segundo y d) quinto día de fermentación. Los datos se representan como los genes regulados al alza en azul (fold change  $\geq 2$  y tasa de falsos descubrimientos [FDR]  $< 0,05$ ), los genes regulados a la baja en rojo (fold change  $\leq -2$  y FDR  $< 0,05$ ), los genes con un cambio significativo, pero bajo (fold change absoluto  $\geq 2$  y FDR  $< 0,05$ ) se representan en gris y los genes no significativos en negro (FDR  $> 0,05$ ).

## El ARN presente en el mosto de cerveza es necesario para la capacidad de fermentación óptima de las cepas de levadura lager estudiadas

Para obtener una visión más general y completa de los genes DE observados en cada cepa, se realizó un análisis por ontología génica y vías metabólicas para obtener las sobrerrepresentaciones significativas de los términos GO y KEGG.

En las comparaciones del segundo día para la cepa SpIB1, observamos que los términos GO como proceso biológico involucrado en interacciones intraespecies (GO:0051703), adhesión celular (GO:0007155), floculación (GO:0000128), adhesión entre organismos unicelulares (GO:0098610) y crecimiento invasivo en respuesta a la limitación de glucosa (GO:0022604)

estuvieron regulados positivamente por la presencia de ARN en el mosto de cerveza (Fig. 16a).



**Figura 16.** Análisis ontológico de los genes DE en el segundo día de fermentación para la cepa SpIB1. Los genes DE se representan como a) genes regulados a la baja y b) genes regulados al alza que respaldan los términos ontológicos con significación estadística (tasa de descubrimiento falso [FDR] < 0,05). Los términos ontológicos en los genes regulados a la baja significan que estos genes están regulados positivamente por el ARN presente en el mosto de cerveza y los términos ontológicos de los genes regulados al alza sugieren una posible regulación a la baja por el ARN presente en el mosto de cerveza. Todas las comparaciones se realizan entre fermentaciones tratadas con ARNasa frente a fermentaciones de control (tratamiento frente a control).

En los términos GO mencionados anteriormente se observaron genes como *FLO9*, *FLO5*, *FLO1*, *FLO10*, *SNF2*, *CYC8*. Junto con el sistema de floculación, también se observaron los términos GO relacionados con la vía de

crecimiento filamentosos, donde el proceso metabólico y biosintético del betaglucano estaban sobrerrepresentados (GO:0034406 y GO:0034410 respectivamente).

Otros términos GO importantes de mencionar están relacionados con el ciclo celular mitótico, donde la regulación positiva de la transición G1/S del ciclo celular mitótico (GO:1900087), la regulación positiva de la transición de fase G1/S del ciclo celular (GO:1902808), la regulación de la transición de fase G1/S del ciclo celular (GO:1902806) y la regulación de la transición G1/S del ciclo celular mitótico (GO:2000045), es importante mencionar que estos términos GO del ciclo celular mitótico están enriquecidos por el mismo conjunto de genes DE.

En conjunto, estos resultados sugieren que el posible efecto del ARN presente en el mosto de cerveza regula positivamente el ciclo celular mitótico de la cepa SpIB1, lo que indica una relación positiva con el ARN no codificante presente en el mosto de cerveza.

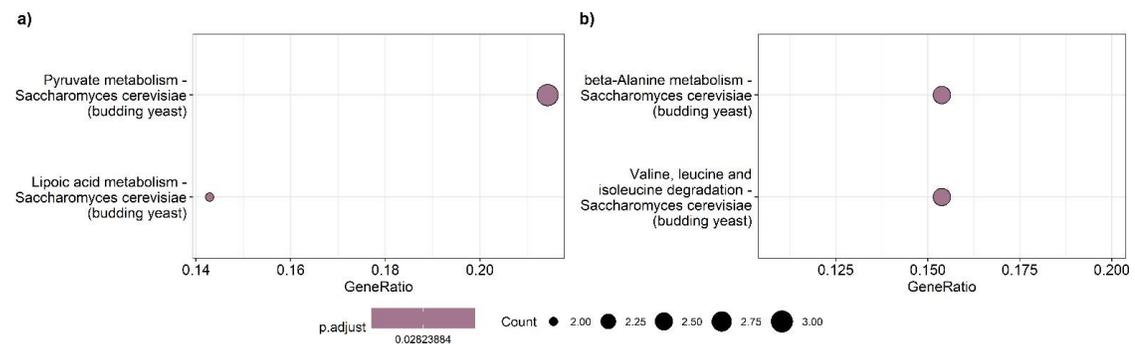
Por otro lado, los genes DE que fueron afectados negativamente por el ARN presente en el mosto, alimentaron términos GO relacionados con los compuestos de nitrógeno, fósforo y azufre (Fig. 16b), términos GO como proceso metabólico de oxoácidos (GO:0043436), proceso metabólico de ácidos orgánicos (GO:0006082) y proceso metabólico de ácidos carboxílicos (GO:0019752) estuvieron sobrerrepresentados en este análisis.

Junto con estos términos GO, el proceso metabólico de aminoácidos (GO:0006520) también se observó como afectado negativamente. Estos resultados sugieren que la capacidad biosintética de la cepa SpIB1 se redujo por la presencia del ARN en el mosto de cerveza. En general, la promoción del crecimiento filamentosos y la supresión de la capacidad biosintética de SpIB1 fueron el comportamiento principal del cambio de datos transcriptómicos observado en esta levadura en el segundo día de fermentación.

El comportamiento transcriptómico del quinto día de fermentación para SpIB1 no mostró sobrerrepresentación de términos GO, lo que sugiere un estado metabólico bajo que coincide con los últimos días de fermentación, donde la

levadura mostró una baja aptitud biológica en la expresión del transcriptoma (Olesen *et al.*, 2002).

A pesar de que los datos transcriptómicos del quinto día no muestran sobrerrepresentación de términos GO, observamos dos términos metabólicos KEGG que mostraron regulaciones positivas por la presencia de ARN en el mosto de cerveza. Como se esperaba, el comportamiento metabólico en el quinto día de fermentación coincide con los términos GO observados en el segundo día de fermentación, donde el metabolismo del ácido lipoico y del piruvato (sce00785 y sce00620, respectivamente) se observaron como vías metabólicas reguladas positivamente (Fig. 17a).



**Figura 17.** Sobrerrepresentación de los términos KEGG observados en los datos transcriptómicos para el quinto día de fermentación de la cepa SpIB1. Los términos KEGG están enriquecidos con a) los genes regulados a la baja y b) los genes regulados al alza observados. Esto significa que los términos KEGG subregulados se ven afectados positivamente por la presencia de ARN en el mosto de cerveza y viceversa, los términos KEGG regulados al alza se ven regulados negativamente por la presencia de ARN en el mosto de cerveza (comparaciones del tratamiento frente al control).

Estos términos KEGG sugieren un comportamiento fermentativo ordinal en presencia de ARN en el mosto de cerveza. Por otro lado, debido al agotamiento del ARN durante el tratamiento con ARNasa, el comportamiento metabólico en el quinto día de fermentación sugiere una reducción de la capacidad catabólica de aminoácidos de la levadura, lo que coincide con los términos GO anteriores que indican una capacidad reducida para el metabolismo del nitrógeno.

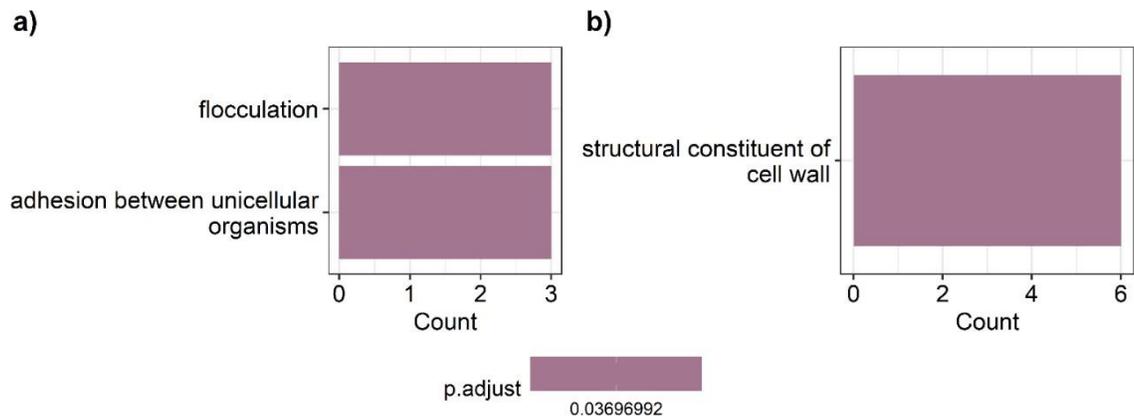
En general, estos resultados sugieren que la presencia de ARN en el mosto de cerveza es importante para la correcta capacidad de fermentación de la levadura SpIB1, al promover el ciclo celular mitótico, la floculación y el crecimiento filamentosos en el segundo día de fermentación. Mientras que, en el quinto día de fermentación, se asemeja al comportamiento de fermentación ordinal relacionado con el metabolismo del piruvato.

Esto, también respaldado por los términos GO y KEGG observados en los genes regulados positivamente, causados por el agotamiento del ARN presente en el mosto de cerveza, que indican una reducción del proceso metabólico del nitrógeno y los aminoácidos.

La respuesta transcriptómica de SpIB2 parece ser más caótica, ya que solo observamos términos GO relacionados con la función molecular en los genes sobreexpresados del segundo día de fermentación (Fig. 18a) y no se enriquecieron términos GO relacionados con el componente celular y el proceso biológico.

Esta información podría ayudarnos a aproximarnos al estado metabólico de la cepa SpIB2, como lo indica el constituyente estructural de la pared celular (GO:0005199) observado en los genes sobreexpresados del segundo día de fermentación. En este término GO observamos genes como *PAU1*, *PAU2*, *PAU5*, *PAU22*, *PAU19* y *PAU17* que indican las condiciones estresantes del desempeño fermentativo vigoroso de esta levadura, donde se debe mantener la integridad de la pared celular (Luo y van Vuuren, 2008).

Por otra parte, los genes regulados a la baja afectados por la eliminación del ARN presente en el mosto de cerveza, enriquecieron procesos biológicos como la floculación (GO:0000128) y la adhesión entre organismos unicelulares (GO:0098610). Esto es consistente con el resultado transcripcional para SpIB2 en comparación con SpIB1, donde la floculación y la adhesión entre organismos unicelulares se vieron afectadas positivamente por la presencia de ARN en el mosto.



**Figura 18.** Análisis ontológico de los genes DE en el segundo día de fermentación para la cepa SpIB2. Los genes DE se representan como a) genes regulados a la baja en el tratamiento y b) genes regulados al alza que respaldan los términos ontológicos con significación estadística (tasa de descubrimiento falso [FDR] < 0,05). El término de función molecular en a) genes regulados a la baja significa que estos genes están regulados positivamente por el ARN presente en el mosto de cerveza y los términos de proceso biológico de b) genes regulados al alza sugieren una posible regulación a la baja por el ARN presente en el mosto de cerveza (comparaciones tratamiento vs control).

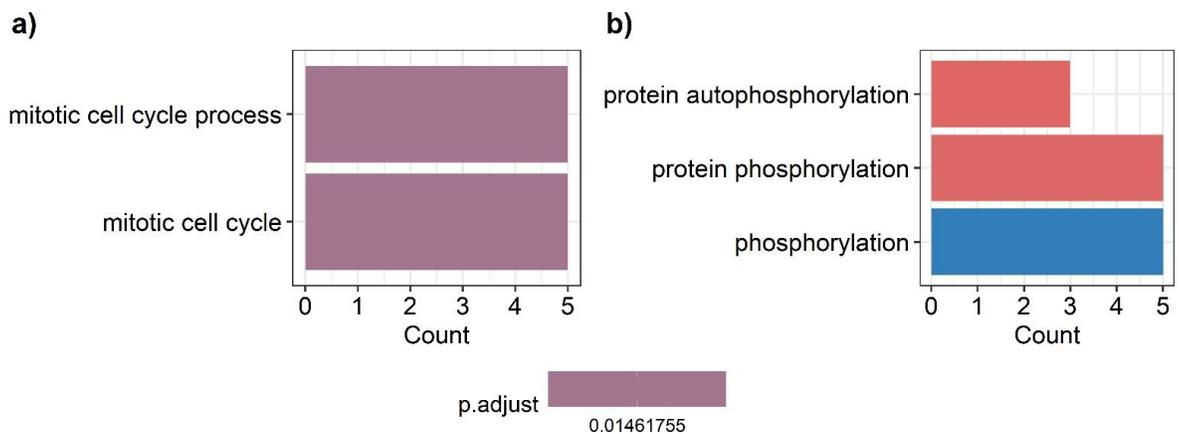
Finalmente, el resultado transcriptómico en el quinto día de fermentación indica que las proteínas reguladoras involucradas en la autofosforilación de proteínas (GO:0046777) y fosforilación (GO:0006468) están significativamente enriquecidas con los genes sobreexpresados (Fig. 19).

Estos genes, involucrados principalmente en el estrés oxidativo y nutricional, se ven afectados negativamente por la presencia de ARN en el mosto, lo que sugiere una funcionalidad adecuada para enfrentar estas condiciones estresantes en la etapa final de la fermentación.

Y, como se esperaba, el proceso biológico como el ciclo celular mitótico (GO:1903047) se enriqueció con los genes sobreexpresados, lo que significa que, por la presencia de ARN en el mosto de cerveza, la expresión de estos genes se activó. El cuadro general del resultado transcriptómico de SpIB2 sugiere que la alta capacidad fermentativa de esta levadura se ve afectada positivamente por la presencia de ARN en el mosto de cerveza, al inducir genes involucrados

en la floculación y el ciclo celular mitótico y al reprimir genes involucrados en la respuesta a condiciones estresantes.

Estos resultados consistentes entre estas levaduras lager indican que la presencia de ARN en el mosto es necesaria para el proceso adecuado de reproducción asexual y floculación, mientras que procesos como el metabolismo de aminoácidos y los genes inducidos por estrés fueron reprimidos.



**Figura 19.** Análisis de enriquecimiento de los datos transcriptómicos en el quinto día de fermentación en la cepa SpIB2. Los términos GO enriquecidos en el proceso biológico con los genes a) regulados a la baja y b) regulados al alza. Nótese que las proteínas reguladoras involucradas en la autofosforilación y fosforilación de proteínas están enriquecidas por los genes regulados a la baja, lo que sugiere que estos procesos biológicos son inducidos por la presencia de ARN en el mosto. Por el contrario, el ciclo celular mitótico se ve afectado negativamente por la presencia de este ARN en el mosto (Comparaciones del tratamiento frente al control).

## V Discusión

En este trabajo, exploramos el supuesto efecto que el ARN exógeno, presente en el mosto cervecero, tiene sobre la expresión genética de dos levaduras cerveceras (cepas SpIB1 y SpIB2). El comportamiento de los datos transcriptómicos mostró un mayor efecto en el segundo día de fermentación y un efecto reducido en el quinto día en ambas cepas utilizadas en el estudio.

La explicación general de este comportamiento está relacionada con el estado metabólico de las levaduras es más activo en el segundo día, mientras que en el quinto día el estrés por etanol, oxidativo, osmótico y nutricional ejercen efectos que promueven la adhesión célula-célula (Stewart, 2018). Los principales resultados representativos están en el proceso de regulación positiva en cada levadura donde se eliminó el ARN total en el mosto, ya que estos procesos metabólicos se ven afectados negativamente por la presencia de ARN en el mosto de cerveza.

El panorama general de la respuesta transcriptómica al tratamiento con RNasa en el segundo día de fermentación mostró una regulación negativa en el proceso metabólico de ácidos orgánicos y carboxílicos y el proceso de aminoácidos en SpIB1. El gen *ARO80*, que enriquece los procesos metabólicos de oxo y ácidos orgánicos en SpIB1 y SpIB2, es regulado a la baja por la presencia del ARN en el mosto, es particularmente interesante ya que está involucrado en la activación transcripcional de *ARO10* y *9*.

Los genes *ARO10* y *9* se han relacionado con el sistema de *quorum sensing* en *S. cerevisiae*, siendo estas las enzimas productoras de moléculas señal como el triptofol y el 2-feniletanol. En asociación con esta idea, estas moléculas se han relacionado con el proceso de formación de pseudohifas en *Sc* (Avbelj *et al.*, 2015), una estructura posiblemente análoga a los mecanismos de penetración de hongos fitopatógenos en los tejidos vegetales, como lo hace un ascomiceto patógeno del arroz *Magnaporthe oryzae* (Sakulkoo *et al.*, 2018).

Por lo tanto, la inhibición por RNA del factor de transcripción del gen *ARO80*, tiene un profundo significado que ilustra los mecanismos de defensa conservados por las plantas contra el crecimiento micelial de hongos fitopatógenos. Este mismo escenario se observa en SpIB2, a pesar de que este importante gen *ARO80* no enriquece ningún proceso biológico o metabólico, observamos que este gen se regulaba a la baja en el tratamiento con RNasa de SpIB2.

Por último, el resultado transcriptómico en el quinto día de fermentación muestra que la cepa SpIB1 tiene una regulación positiva de los metabolismos de piruvato y ácido lipóico en presencia del ARN en el mosto, lo que sugiere un efecto positivo en el curso normal de la fermentación. Esto es consistente con la regulación negativa del proceso metabólico relacionado con el metabolismo de la beta-alanina y la degradación de valina, leucina e isoleucina. Estos procesos metabólicos son necesarios cuando la fuente de carbono es limitada y son muy importantes en el resultado organoléptico de la producción de cerveza (Hazelwood *et al.*, 2008).

Esto puede estar asociado con la regulación negativa del gen *ARO80* en el segundo día de fermentación, donde es importante para la inducción de *ARO8* y *ARO9* involucrados en la vía de Ehrlich (Hazelwood *et al.*, 2008).

Para el SpIB2, el efecto en el tratamiento del quinto día de fermentación se reduce, debido al estado metabólico característico en la fase de muerte o floculación. Los procesos biológicos en el tratamiento con ARNasa (Fig. 19b) se enriquecieron con el gen *GCN2*, una quinasa sensora de estrés nutricional requerida para la adaptación en ambientes con bajo contenido de nitrógeno (Cambiaghi *et al.*, 2014).

En el contexto de la relación patológica planta-hongo, especulamos que el evento de inhibición de una quinasa sensora traería importantes ventajas a la planta. La regulación negativa de *GCN2* en presencia de posible ARNnc tiene consecuencias negativas para la levadura, ya que presenta dificultades para adaptarse a ambientes con bajos niveles de nitrógeno.

En el mismo contexto, otro gen interesante es *RIM15*, una quinasa que regula positivamente la entrada apropiada a la fase estacionaria bajo estrés nutricional (Cameroni *et al.*, 2004). La intervención transcripcional de este gen puede traer ventajas a la planta al interrumpir el proceso apropiado de entrada a dicha fase de algunos hongos fitopatógenos.

Bajo el supuesto de que existen ARN en el mosto cervecero con efecto sobre la expresión genética de las levaduras cerveceras, estos datos

posiblemente reflejen los eventos defensivos que desarrolla la planta, *Hordeum vulgare* y/o *Humulus lupulus*, frente a hongos fitopatógenos presentes en sus ambientes de crecimiento, los cuales probablemente comparten genes ortólogos con otros ascomicetos como las levaduras gemantes (Kiss *et al.*, 2019).

Si bien es cierto que el tema relacionado con la comunicación a través de especies y reinos mediada por ncRAN es controversial debido a la falta de evidencia experimental y teórica sobre los mecanismos de acción de este sistema, aquí exponemos mediante experimentos de secuenciación de ARN y métodos computacionales un posible efecto sobre la expresión genética de dos levaduras lager debido a la acción del RNA presente en el mosto cervecero utilizado en la producción industrial de cerveza lager.

## **VI Conclusión**

En este trabajo, determinamos el posible efecto del RNA presente en el mosto de cerveza sobre la expresión genética de dos levaduras lager. El efecto principal se observó el día 2, cuando se presentó el mayor número de DEG. Mientras que el día 5 se observó un efecto reducido.

En el contexto de la relación planta-patógeno, algunos genes posiblemente inhibidos por la acción del RNA tienen sentido, ya que los genes diana en hongos fitopatógenos pueden compartir homología con los otros ascomicetos.

Este trabajo aporta información sobre el controvertido tema relacionado con la comunicación entre reinos mediada por RNA y más específicamente nuestros descubrimientos pueden ayudar en la toma de decisiones en el proceso de manipulación microbiológica de cepas de levaduras con importancia industrial.

Somos conscientes de que son necesarios más experimentos para demostrar el efecto del RNA. Aun así, bajo la secuenciación de ARN y las herramientas computacionales utilizadas en este trabajo, determinamos un efecto del RNA sobre la expresión genética de dos levaduras con importancia industrial.

### **10. CAPÍTULO 3. Otras contribuciones producto de la colaboración científica**

En este apartado se muestran los productos de una sólida colaboración científica que surgieron durante el desarrollo doctoral de esta investigación.



Original Research Article

## Identification of potential target genes in *Homo sapiens*, by miRNA of *Triticum aestivum*: A cross kingdom computational approach

Daniel Sánchez-Romo, César I. Hernández-Vásquez, Benito Pereyra-Alfárez, Jorge H. García-García\*

Facultad de Ciencias Biológicas, Universidad Autónoma de Nuevo León, Av Universidad S/N, 66455, San Nicolás de los Garza, Nuevo León, Mexico



### ARTICLE INFO

**Keywords:**  
microRNA  
mRNA  
*Triticum aestivum*  
Human  
Functional enrichment

### ABSTRACT

Plant-derived miRNAs can be found in the human body after dietary intake, and they can affect post-transcriptional gene regulation in human. It is important to identify targets to determine the possible effects in human genes by using computational approach. In this study, 787 possible mRNA human targets were predicted by 84 miRNAs of wheat. A total of 14 miRNAs were identified with individual binding to 33 mRNAs associated with schizophrenia, epilepsy, neurodevelopmental disorders, and various cancers, located in the 3'UTR of the mRNA. A functional enrichment was carried out, where the results showed associations to pathways such as dopaminergic synapse (hsa04728), and signaling pathways, significantly associated with the target genes. The prediction of target mRNAs in humans by wheat miRNAs, offer candidates that could facilitate the search and verification, which could be of relevance for future projects and therefore contribute in the therapeutic treatment of various human diseases.

### 1. Introduction

The microRNAs (miRNAs) of plants and animals are an abundant class of small non-coding RNAs sequences around ~22 nucleotides long [1]. They are derived from hairpin precursors, which have remained evolutionarily similar in various species, and have the ability to negatively regulate gene expression at the post-transcriptional level by complementary base binding in various regions of the mRNA (3'UTR, 5'UTR and ORF) and via the RNA-induced silencing complex (RISC) whose binding causes mRNA cleavage or translation inhibition [2–4].

It is expected that 30–60% of the genes in humans are miRNA regulatory targets [5], involved in various developmental processes, molecular, cellular pathways and human disease [6]. When crossing the intestinal barrier after ingestion miRNAs can be packaged in microvesicles, circulate through the human body and be identified in fluids such as human plasma [7] and breast milk [8].

Identifying mammalian gene targets in a timely manner by plant miRNA offers a reliable, effective and economical way to determine the possible effects of interactions given by exogenous molecules, regardless of their origin, on human mRNA [9–14].

Recently, plant miRNAs and their interactions have been actively studied, contributing to the knowledge of coevolution mechanisms

between plant miRNAs and human mRNAs. For example, the *in silico* prediction on the possible regulation of plant miRNA on human mRNA was tested, coincided with an increase in cholesterol-LDL levels *in vivo* and *in vitro*, caused by the decrease of LDLRAP1 protein expression, providing the first evidence that food miRNA could bind to a target mRNA and influence the regulation of interspecies expression [15]. Similarly, through bioinformatics analysis, found a number of potential target genes for MIR2911 from the *Lonicera japonica* plant (honeysuckle) were found in different viruses such as H1N1, H3N2, H5N1 and H7N9. Later, in cultures of cellular lines infected with H1N1 and transfected with synthetic MIR2911, they managed to significantly decrease the viral count because the target genes: PB2 and NS1, are essential for replication. Subsequent functional studies showed that miR2911 inhibited several IAVs (influenza A viruses) including H1N1, H5N1 and H7N9, decreasing mortality in an animal model [16]. Suggesting the first reports that exogenous miRNAs regulate gene expression in endogenous mRNAs similar to mammalian miRNAs and in turn the use of the bioinformatics approach, playing an important role in target identification.

In the present study we have selected the monocotyledonous plant *Triticum aestivum* (wheat) because it is a common source in human nutrition, we performed an *in silico* prediction analysis of potential

\* Corresponding author.

E-mail address: [jorge.garciagr@uanl.edu.mx](mailto:jorge.garciagr@uanl.edu.mx) (J.H. García-García).

<https://doi.org/10.1016/j.ncrna.2022.03.002>

Received 10 December 2021; Received in revised form 31 January 2022;

Accepted 14 March 2022

Available online 18 March 2022

2468-0540/© 2022 The Authors.

Published by KeAi

Communications Co., Ltd.

This is an open access article

under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

mRNAs target in *Homo sapiens* (human) by mature miRNAs of wheat, and its relationship to diseases through an enrichment of the possible target genes. In addition, we identified target genes with diverse functions in cellular processes, providing information on regulatory mechanisms given by wheat miRNAs. Also, a functional enrichment has been carried out to explore the participation of the target human mRNAs involved in biological processes and associated pathways, including apelin signaling pathway, fanconi anemia pathway and dopaminergic synapse.

## 1. Material and methods

### 1.1. Data collection

The complete set of known mature miRNA sequences, which includes a total of 122 miRNAs from wheat, were downloaded from the registered database (version 22.1) (<http://www.mirbase.org>) in FASTA format. To reduce errors in prediction and future detection, we applied rigorous filtering for mature miRNAs, i) The mature miRNA should have more than 1400 reads during sequencing reported in its bioinformatic analysis; ii) The miRNA size should be in the range of 19–25 nucleotides and iii) All mature miRNAs should be in the same direction from 5' to 3'. Following this, 84 miRNAs from wheat were selected as reference data set for the study.

### 1.2. Prediction of potential targets of miRNAs

The prediction of possible targets on human was carried out by the sequence similarity search tool BLASTn (2.9.0) on its NCBI website (<http://www.ncbi.nlm.nih.gov>) using the default settings, an E-value threshold of 10; and the search coordinates were set in the nucleotide sequence 2 to 13 of each miRNA. From the RefSeq RNA database the predicted miRNA target sites were identified by aligning the sequences in a complementary way to the miRNA of wheat with human mRNA. RNAhybrid [17] was used to scan the hybridization patterns of the individual miRNAs/mRNA following Zhou Z. et al. methodology with some modifications [16]. i) The seed region had to perfectly match the nucleotide 2 to 7 of the mature miRNA. ii) The minimum free energy (MFE) for the hybridization was below  $-25$  kcal/mol iii) No G-U bonds in the seed region. We also used the MIR168a as a target prediction control [15].

### 1.3. miRNA homology

We searched for homology using MirCompare between 2042 *H. sapiens* and 84 *T. aestivum* matured miRNAs from miRBase was used [18]. The parameters used for this analysis were taken from Minutolo et al., 2018 [19], with modification in the threshold of the seed region according to the guidelines provided by the default settings of the tool. The possible target genes of the resulting human miRNAs with homology to wheat miRNAs as reported by MirCompare were found by an analysis with TargetScan 8.0 algorithms (<http://www.targetscan.org>).

### 1.4. Functional enrichment analysis

The functional analysis of the targets, including biological processes was done using clusterProfiler [20], package inside programming language R version 4.0.3 [21]; with only those target sequences with miRNA binding sites in the 5'UTR, 3'UTR and CDS regions. Designating the potential processes with an adjusted p-value lower than 0.05 (FDR <0.05) that involve the predicted mRNA sequences and suggesting the probable influence in the regulation of plant miRNAs in the human organism.

## 2. Results

### 2.1. Characteristics of the binding of wheat mature miRNAs with human mRNAs

According to the methodology previously mentioned, the results obtained were similar to the ones described by Zhang et al. 2012 [15]. We selected 84 mature miRNAs from wheat and predict putative target genes in humans, based on the criteria specified in methods section, which resulted in a total 889 miRNA/target interactions with 787 genes with diverse functions, the characteristics and specific target sequences of the 84 miRNAs analyzed in this study are provided in the (Supplementary Table 1). 70 of these genes had interactions with more than one miRNA and 6 come from miRNAs families, as previously reported by Han R. et al. [22] (Table 1). These miRNAs were divided in two groups; in group one, 14 miRNAs identified with one to four target genes, giving a total of 30 target genes. In group two 70 miRNAs with five or more target genes, giving a total of 757 target genes. The tae-miR9652–3p and tae-miR9652–5p of group one, originated from the same precursor (pre-miRNA) in miRNA biogenesis, have binding sites in the mRNAs of different genes. In the second group the tae-miR9657b–3p and tae-miR9657b5–p, were found as a pair originating from the same pre-miRNA, obtaining 18 targets for miR-3p, and 10 for miR-5p strand, all being different. The miRNAs with the highest number of targets were tae-miR9657c (19 genes), tae-miR9654a (21 genes), tae-miR160 (22 genes), tae-miR396 (23 genes), tae-miR9662a (23 genes), tae-miR9666a (24 genes), tae-miR9664 (28 genes), tae-miR9655 (31 genes), tae-miR9666b–3p (34 genes), tae-miR164 (35 genes).

In order to show the targets and MFE for the binding between some wheat miRNAs with human mRNAs, data was filtered to show the targets genes with lower value of MFE and were the sites union are located in 3'UTR binding (Table 2). 10 of the miRNAs show 1 target gene, 2 miRNAs show two targets, and 2 miRNAs show three and six targets, with a MFE between  $-34.1$  and  $-43$  kcal/mol. The tae-miR160 shows 10 target genes with a free energy value between  $-36.9$  and  $-45.9$  kcal/mol.

Because genes are nucleotide sequences coding for diffusible elements that participate in all cellular processes, it is important to show information about some human target genes by wheat miRNAs, involved in various biological processes (Supplementary Table 2): RALY - lung cancer, CNTNAP4 - epilepsy, CCDC88C - neurodevelopment, PCARE - pigmentary retinosis, FRMD8 - inflammation, CRIBB33 - cataracts, PLEKHA6 - schizophrenia, BTNL9 - ocular melanoma, KAT5 - anaplastic thyroid cancer, PARBV - oral squamous cell carcinoma, PRKAG2 - nonalcoholic steatohepatitis, CHST10 - tumor suppressor, RTN4RL1 - nasopharyngeal carcinoma, MTCL1 - cerebellar ataxia, GNAI2 - ovarian cancer, SMARCC2 - embryogenesis and corticogenesis, CSRP1 - lung fibrosis, B3GNT7 - colon cancer, KCNA2 - epileptic encephalopathy, NOX5 - diabetic nephropathy, CBX7 - hepatocellular carcinoma, BCL7B - Wnt signaling, MBD3 - pancreatic cancer, CORO1B - TGF $\beta$ 1 signaling, ATPAF2 - dementia risk, DIAPH3 - lung adenocarcinoma, and TPM2 - aortic dissection.

Interactions between wheat miRNA and human mRNA can show how efficiently these molecules act. Some tae-miRNA160/mRNAs binding patterns in the 3'UTR of the mRNA, resulting from the bioinformatic tool RNAhybrid are shown in (Fig. 1). Taking into consideration the nucleotide sequence after the seed region of the mature miRNA, in the *in silico* prediction through the RNAhybrid tool, it is possible to observe the type of interaction between the tae-miRNAs with their targets and provide an approach to the type of silencing like repression and/or degradation through the complementary base pairing. For this reason, when examining the type of binding or complementarity that was performed in each miRNA/mRNA result, in a complete or incomplete complementary binding, it was found that 100% of the 889 predicted tae-miRNA/target interactions, were shown as imperfect hybrids in at least one nucleotide of the complementary mRNA after the seed region.

**Table 1**  
Potential target genes of 84 wheat miRNAs identified in human.

miRNA ID	Number of targets	miRNA family	Target Gene
Tae-miR1117	9		TULP4, CHSY3, SMARCC1, C5orf52, DSCAML1, BZW1, UBR5, LOC102724740, ACVR1B
Tae-miR1118	7		ARHGAP32, PRUNE2, XKRX, HOOK3, LOC105372263, ANTXR2, MAP1B
Tae-miR1120b	1		RPAP2
Tae-miR1122a	6		SERINC5, ZBTB39, GDNF, SPAG9, ZC3H12C, CDC42BPB,
Tae-miR1122b	2		LOC105375144, SPTLC3
Tae-miR1122c	1		PPP2R3A
Tae-miR1123	9		EPO, LOC105375220, ASGR2, HPCAL1, CNFN, COL16A1, IGF2BP2, BEST1, LOC100507516
Tae-miR1125	5		CHRM3, SLC38A8, RALY, TIMM10, RYR2
Tae-miR1127a	4		MSRB3, IPO13, RRAD, MFAP3
Tae-miR1127b	9		XKR6, STXBPSL, NOL8, ZNF691, KRIT1, RP1L1, ZNF510, SNX13, RGS6
Tae-miR1128	12		CNTNAP4, LOC105371840, ESR1, RAPGEF6, LOC107986008, TMEM161B-AS1, BRINP1, INTS4, THAP9, ADCY1, RUSC2, KLHL36
Tae-miR1130b	2		PRORY, ASPH
Tae-miR1131	7		SMARCC1, DSCAML1, UBR5, CHSY3, BZW1, TULP4, STXBPSL
Tae-miR1133	7		WDR75, MALT1, RIMKLB, SKAP1, ACAD11, GPR176, ST8SIA4
Tae-miR1135	5		LOC112268243, LOC105371988, SPACA1, THSD7A, DDX5
Tae-miR1136	11		OBSCN, DHX37, MRC1, STAB1, LOC105379283, SNORA14B, ZNF174, CALN1, TMEFF2, CALN1, UQCC2
Tae-miR1137a	3		ATP6V1B2, ZFPM2, ATXN1L
Tae-miR1137b	8		RBX1, STK35, CREB5, RNF121, KIF27, JAG1, PKHD1, ITGA11
Tae-miR1138	8		KDM3B, USP35, PRKACB, MAK, ZBED2, TESK2, BMP3, ZSCAN25,
Tae-miR1139	5		ZFH3, ZC3H8, CDON, MTD1, DZIP1L
Tae-miR156	14		CCDC88C, F1R, FAM13A, KIF5A, LOC105369388, MAPK11, ZBTB40, LINCO0312, CCDC15, TMEM201, ADGRD1, CAPN13, TGOLN2, CCNH
Tae-miR159a	8		CNOT9, GPR4, TEX13C, LOC105377862, MDH1B, LINC02789, MAN1A2, GPR135,
Tae-miR159b	8	miR159	CNOT9, GPR4, TEX13C, LOC105377862, MDH1B, LINC02789, MAN1A2, GPR135,
Tae-miR160	22	miR160	PCARE, FAM118B, BCAS1, BCAP29, TMEM241, CAMTA1, FMRD8, PSMD7, ASPDH, CRYBB3, PLEKHA6, BTNL9, ANKRD35, KAT5, DAAM2, PARVB, PRKAG2, RTN4RL1, CHST10, RHPN1, GHRL, LOC100996662
Tae-miR164	35	miR164	FGFRL1, ZNF701, SLC9A9, MTCL1, PPRC1, GREM1, SYNE1, C8orf33, GRM2, PCYT2, GLI3, SNED1, GNAI2, SMARCC2, CSRPI, B3GNT7, KCNA2, LOC107986317, RFPL4B, FSTL4, SLC25A1, RASSF6, KCNH6, APOBEC3B, APOB3CA, ABCB9, BIRC6, NCAPH, SETD4, SLC6A7, LRRC55, MNT, ACBD4, TRIM41, PBXIP1
Tae-miR167a	13	miR167	PRICKLE2, WASHC2A, TM2D3, CACNA2D4, LOC105375492, NOX5, CBX7, TNFRSF10D, BCL7B, MBD3, NCBP3, SLC9B2, THBS3
Tae-miR167b	9		EME1, ZNF20, PLCE1, LOC107986558, CACNA1C-IT1, LOC440292, CAMK2D, GPD1, SNRNP200
Tae-miR167c	12		TM2D3, WASHC2A, CACNA2D4, NCBP3, BCL7B, LOC105375492, MBD3, NOX5, CBX7, TNFRSF10D, SLC9B2, THBS3
Tae-miR171a	1	miR171	IGF1, C15orf56
Tae-miR171b	2		TMEM8A, RNF112, TMEM151A, NAT16, RPAP1, LIMK2, GCC1, LOC107986282, MPRIP, RAPH1, LOC112268024
Tae-miR1847	15		CRISP2, CYP27A1, ALG3, LAMA5, N4BP1, ART3, ZNF619, UBQLN4, TMEM178B, LMOD1, LG3, LOC105377347, SH3PXD2A, CRISPLD1, MRM1
Tae-miR319	14		SCART1, HOXB3, SLC1A4, TIGD7, FAT4, FAM205C, KLC2, KMT5A, ACLY, MUC16, CCDC134, LOC105375153, RAB5B, SUOX
Tae-miR395a	7		CLCN5, LOC107984275, VAPA, LOC105373908, CERS4, PPARGC1B, ITGB8
Tae-miR395b	11	miR395	CLCN5, USP6NL, LINC02739, FN1, CASP5, WDFY3, CEP63, METRN, UBAP1L, LOC112267986, LYPLAL1
Tae-miR396	23		SLC25A25, CPOG1, CYFIP1, NOTCH2NLB, ZNF814, KDM1A, PELI3, GCDH, ZNF32, ADAMTS18, ZNF776, CL11A1, PTCH2, UNC119B, ZBTB14, LOC107984192, REST, FOXP2, SLITRK9, CDC5L, EFHB, KIF2A, GRIN1
Tae-miR397	5		PALM2AKAP2, SURF6, BCAR1, C1orf67, RNF175
Tae-miR398	14		PPM1H, CORO1B, ANKK1, LOC105374534, FBXO10, DNMT1, SLC37A1, PCNX3, LOC105370363, LOC105378740, ZNF691, KIAA0408, SOGA3, PPIPSK2,
Tae-miR399	16		DRICH1, ANKRD36, SET, LINC02142, ARIH2, FAM166C, ZFP2, SCCPDH, ADPRM, LRR1Q1, AOC3, NUP210, CALM1, MANBA, MYLK4, ANKFNI
Tae-miR408	16		USP43, ATPAF2, PSMD11, ADGRF5, DIAPH3, AGO3, ABCG8, LOC105376475, ZNF518A, EPS15L1, ANXA11, ROS1, RYR3, KCNB2, LOC112268372, KSR1
Tae-miR444a	14		LOC112268280, RUND1, LOC105371800, ZC3H4, LOC107986334, CDC25A, ZC3H4, CCDC171, SUPT20HL1, LOC101929566, SYNJ2, TEKT3, FRMD3, YPEL5
Tae-miR444b	13		LOC112268280, RUND1, LOC107986334, LOC105371800, ZC3H4, CDC25A, CCDC171, SUPT20HL1, LOC101928566, SYNJ2, TEKT3, FRMD3, YPEL5
Tae-miR5048	5		CHRM3, TBC1D4, ADAM22, GRM4, LOC105378532
Tae-miR5049	2		LOC147004, PNPT1
Tae-miR5050	9		PEG3, IPPK, ITPA, TRIO, TEX30, SCFD2, FOSL2, LINC01226, BCLAF1
Tae-miR5062	9		ARMC9, YEATS2, LOC107985153, CD2BP2, AB12, RNF183, CACNA2D4, AHNAK, HS6ST2
Tae-miR5175	16		SETBP1, RBBBP6, GPRC5D, CHST2, C4orf54, CENPS-CORT, CORT, RNLS, TRIM37, NSL1, LINC01588, SNRNP27, LOC101929270, SUCC, STK31, DNJAC6
Tae-miR5200	5		LINC02842, LOC101927610, LOC107986016, NRG4, LOC101928421
Tae-miR531	18		REPIN1, LOC107985320, SLC5A10, GBX2, GPC6, NINL, BACH2, B3GLCT, PLXNB3, DAB2IP, GCK, DCP2, ZNF282, NECTIN4, RFTN1, GPR22, ZNF618, GPR153
Tae-miR6197	3		EXOC6, ABCD2, ZNF519
Tae-miR7757	2		RPR6KA2, HDAC8
Tae-miR9652-3p	1		PSMD13
Tae-miR9652-5p	1		POU2F1
Tae-miR9653b	15		ADCK5, SHE, TPST2, CPNE2, FGGY, P4HA2, TPM2, P4HA2, LOC105372821, VPS26C, HES7, ABCG8, RAB37, MTOR, RAB37
Tae-miR9654a	21		DTX2, SGO2, LOC10537819, TBRG1, LOC107984429, NT5DC2, LOC107986811, ELP6, LYRM9, LOC105375206, CFAP61, CHI3L2, GEMIN5, AP3M2, ZSCAN26, FBXO40, BZW2, TNRC6B, SESTD1, EBF2, TBL2, ZNF740, CNF740, CNST, HOXD4, ARHGAP1, SLITRK3, KALRN, UFL1-AS1, SLC04A1, PCDH8, MLYCD, PDE1A
Tae-miR9654b	12		

(continued on next page)

**Table 1** (continued)

miRNA ID	Number of targets	miRNA family	Target Gene
Tae-miR9655	31		KPNA6, ELMO1, GEMINS, TOMM20, TRIM71, CCBE1, CMBL, TAX1BP3, FGFR10P, CCR9, ELOVL2, ADH6, ARHGAP33, RNF8, TTC39B, CHD1, HMBOX1, MYADM, PLSR3, HIPK1, STX7, LOC105372520, LOC105377622, WBP2, NIPAL3, TRPC3, VAMP2, PRRT2, USP12, HOOK3, AP1S3
Tae-miR9656	7		NPSR1-AS1, DAAMI, LDLRAP1, TRIM72, FAM49B, TRAK1, FAM49B
Tae-miR9657a	18		MECR, C17orf80, PCNT, STOX1, MAGI2, MELTF, PITRM1, LASP1, EPHA10, FYCO1, RGS6, SNX29, LZTS3, ZNF395, LINC00473, TTF1, MYO10, CCDC40
Tae-miR9657b-3p	17		C17orf80, MELTF, PCNT, STOX1, MAGI2, LZTS3, PITRM1, EPHA10, RGS6, SMX29, LASP1, TTF1, ZNF395, FYCO1, LINC00473, MYO10, GAD1
Tae-miR9657b-5p	10		MAFB, SPON1, MAF, TBC1D8, MAFA, FSTL4, CEBPB, DMXL1, SCAMP4, ELFN2
Tae-miR9657c	19		MECR, C17orf80, MELTF, PCNT, STOX1, MAGI2, LZTS3, PITRM1, EPHA10, RGS6, SNX29, LASP1, TTF1, ZNF395, FYCO1, LINC00473, MYO10, CCDC40, GAD1
Tae-miR9662a	23		LOC101929698, LRCOL1, LOC105370502, KCNA7, RBCK1, CLDN20, HRH2, TAF5L, MIR155HG, LINC02568, RIBC2, PFKFB2, TOP3A, LOC107986445, MOB3C, CLTRN, JAKMIP2, HIRA, LOC105374262, PCNT, SS18L2, CASP9, ARNTL
Tae-miR9662b	12		LOC101929698, PIGV, FAM83H-AS1, LOC105370502, AOX3P-AOX2P, HRH2, LOC105371449, LINC02568, HIRA, LOC107986445, LOC102723784, GNB5
Tae-miR9663	5		TAF1, ZC3H7B, PSMB4, DHP5, SGF29
Tae-miR9664	28		IQGAP1, TNXB, RIN3, USP40, LOC112268055, TNC, SOX30, ABCC1, FAHD2B, ABCC1, FAHD2CP, TNR, MLST8, CLEC3B, CCN1, CASZ1, KRT31, COGPI, NTNG2, CORIN, ZBTB48, CFAP46, TENM3, TELO2, LOC112268408, ALPK1, DENND3, KIAA0586
Tae-miR9666a	24		RAPGED4-AS1, BOC, OVOL1, USH2A, NACA, C8orf34, ANXA11, DOK1, CDH11, MST1R, CSRN1P, MBD1, PRR23D1, STON2, DCHS1, MARC1, WDR46, MBD1, DEFB119, ZFAND5, CDH6, SFMBT1, SEC14L2, DLX5, RAD51AP2
Tae-miR9666b-3p	34		SEC14L2, DLX5, RAD51AP2, POLR2A, YEATS4, CC2D2A, PCDHA6, PCDHA2, PCDHA10, PCDHA4, PCDHA13, IFT22, RAET1G, BSN, PDLIM5, ADAMTS10, KCNIP3, GABBR1, PCDHA8, PDZD2, ATP13A1, RBM6, ARID5B, ARHGAP26, PHF2, PCDHA12, PCDHA5, SGPP2, PCDHA7, ARHGAP26, CLOCK, WDR1, SLC24A1, PEX5L
Tae-miR9668	6		LIN02362, TMEM185B, POLE2, LOC101928985, LOC105376956, SLX4
Tae-miR9669	8		TOP3A, RXRG, SSH2, ENGASE, LRP6, GAB1, SMUG1, BAHCC1
Tae-miR9670	7		LOC105372542, NCAPD3, TPM2, GDAP1, SMG7, SCTR, NRG1
Tae-miR9672a	12		FRG2C, DCTN5, BMP8A, PKHD1, PKHD1L1, LOC107987108, NELL2, CEP152, SEC22B, WDFY3, PLXNA4, TACR3
Tae-miR9672b	6		TOX4, LOC102724245, MYO18B, GNL2, APMAP, KIAA1217
Tae-miR9674a	18		ACSS1, FANK1, TJP2, LOC105370291, EIF3A, DYRK1A, ATP9B, ZNF629, PDZRN3, LIPG, RWDD1, PRMT3, TMEM246, ELF4, VPS33A, ERICH1, FARI, LINGO2
Tae-miR9674b	15		TBX19, DSTYK, TXNDC9, FANCE, NSDHL, GALNT3, TCF4, FAAH2, TOP2B, LOC112268009, TCF4, TSTD2, GCLC, SBF2, LINC03094
Tae-miR9675	5		BTRC, TMEM135, MRRF, ATRIP, ATRIP-TREX1
Tae-miR9676	3		TMEM259, GRID1, FOXD1
Tae-miR9677a	6		MAN2A2, RRAS2, OLFML2B, LOC100132249, ABCC8, SLC8B1
Tae-miR9679	11		DPH7, ABLIM1, ZMIZ1, BCL7B, MBOAT2, CBX5, DRP2, INPP5B, ZNF704, TBC1D12, RELN
Tae-miR9772	5		AAK1, RAP1B, NUCB2, SRGAP2, LOC105374129
Tae-miR9774	6		CTB-3M24.3, H6PD, OTUD6B, EDNRB, RGS3, EXPH5
Tae-miR9775	3		MGAT4D, MED27, USP11
Tae-miR9776	5		OLIG2, HECA, ZBTB11, RNF213, LOC105370091
Tae-miR9782	16		LOC112267915, RCSD1, RNLS, ZNF184, SAMD12, CYP24A1, GASK1A, LOC105371867, LOC101928819, TRDMT1, LOC105375471, NIPAL2, SRGAP1, ALG10B, CD38, LOC107987129
Tae-miR9783	7		EIF4EBP2, FGF2BP2, DHX8, GINS2, ARPC5, TRPC3, FAM209B

### 1.1. Analysis of human homologs of wheat miRNAs

In the analysis we found homology candidates in 9 human miRNA sequences, with 7 wheat miRNA sequences. Supporting this, we identified target gene predictions of human miRNAs using the TargetScan tool and determined those targets in common predicted in this study as a whole (Table 3).

### 1.2. Functional enrichment of target genes by wheat miRNAs

Functional annotation of target genes is an important step in understanding the respective functions. A gene ontology (GO) enrichment analysis was performed in the Kyoto Encyclopedia of Genes and Genomes (KEGG) to identify the functions that miRNAs could have on their predicted targets in humans, where all mRNAs that had binding sites in the prediction were used. From the complete target search data, 670 different genes were used, selected according to probable binding site, including only those in the 5' UTR, 3'UTR, and CDS region and excluding those in non-coding sequences for functional enrichment analysis. The main GO terms classified in biological process are shown, with

significance adjusted p-values (<0.05) indicating that a total of 17 mRNAs enriched the biological process of homophilic cell adhesion (GO:0007156), through adhesion molecules to the plasma membrane, and 22 mRNAs enriched the biological process of cell-cell adhesion through adhesion molecules to the plasma membrane (GO:0098742), by miRNAs known from wheat in the prediction (Fig. 2).

Enrichment analysis of the KEGG pathway showed that the 670 predicted genes of the 84 selected wheat miRNAs were significantly enriched in 8 pathways, with an adjusted p value (<0.05) (Supplementary Table 3). 18 genes were enriched in calcium signaling pathway (hsa04020), 10 genes were enriched in the gastric acid secretion pathway (hsa04971), 14 genes were enriched in the apelin signaling pathway (hsa04371), 9 genes were enriched in the long-term potentiation pathway (hsa04720), 11 genes were enriched in the circadian entrainment pathway (hsa04713), 8 genes were enriched in the fanconi anemia pathway (hsa03460), 13 genes were enriched in the dopaminergic synapse pathway (hsa04728), and 13 genes enriched the adrenergic signaling pathway in cardiomyocytes (hsa04261) (Fig. 3).

**Table 2**  
Individual interaction between wheat miRNAs and human mRNAs at the 3' UTR site.

Gene	miRNA ID	kcal/mol
EPO	Tae-MIR1123	-39.1
RALY	Tae-MIR1125	-35.3
CNTNAP4	Tae-MIR1128	-38.7
CCDC88C	Tae-MIR156	-34.9
PCARE	Tae-MIR160	-45.9
FRMD8	Tae-MIR160	-40.3
CRYBB3	Tae-MIR160	-38.8
PLEKHA6	Tae-MIR160	-38.7
BTNL9	Tae-MIR160	-38.6
KAT5	Tae-MIR160	-37.7
PARVB	Tae-MIR160	-37.4
PRKAG2	Tae-MIR160	-37.2
CHST10	Tae-MIR160	-36.9
RTN4RL1	Tae-MIR160	-36.9
MTCL1	Tae-MIR164	-37.5
GNAI2	Tae-MIR164	-35
SMARCC2	Tae-MIR164	-34.9
CSRP1	Tae-MIR164	-34.3
B3GNT7	Tae-MIR164	-34.2
KCN A2	Tae-MIR164	-34
NOX5	Tae-MIR167a	-34.5
CBX7	Tae-MIR167a	-34.2
BCL7B	Tae-MIR167c	-35
MBD3	Tae-MIR167c	-34.6
SURF6	Tae-MIR397	-34.1
CORO1B	Tae-MIR398	-36.1
DRICH1	Tae-MIR399	-34.8
ATPAF2	Tae-MIR408	-36.8
DIAPH3	Tae-MIR408	-35.6
ZNF518A	Tae-MIR408	-34.2
LOC107985320	Tae-MIR531	-43
TPM2	Tae-MIR9653b	-37.4
LRCOL1	Tae-MIR9662a	-35.4

## 1. Discussion

Based on the results of this work and the extensive study on the mechanisms of plant miRNAs when entering the circulatory system through the diet, entering the cells to achieve a target and regulate the expression of a gene [15]. Some authors [23,24], suggest the use of available and evaluated tools for predictive analyses, since they recognize the fundamental importance of identifying the target genes in the human genome by exogenous miRNAs, in order to infer their functions and regulatory mechanisms, however they agree that it is convenient to explore and combine predictive tools because individually, they may not reflect the abundance and possible regulatory capacity of miRNAs, improving the speed and accuracy of the targets. In our research we used two search engines, that way we were able to simulate a possible interaction between species. By using BLASTn as the first instance as a search engine for complementary sequences with the updated version of the RefSeq human transcriptome database, we were able to find several targets and their variants, which is important because factors related to expression variation in target genes can provide a solid foundation and more systematic view of the genetic mechanisms that are related to human genetic variation [25].

The results obtained by the BLASTn analysis were filtered in a second computational method using RNAhybrid tool as used successfully used in several studies of cross kingdom miRNAs [16,26,27], which allows us to determine the MFE, characteristics and complementary regions of the miRNA/mRNA interaction, which allowed filtering of those interactions that we consider energetically favorable (MFE-25 kcal/mol). This

provided us with a combination of algorithms in the study that allowed us to obtain a total of 787 genes and their variants with different functions as shown in (Supplementary Table 1). In addition, Table S1 shows the transcripts of non-coding RNAs (ncRNAs) that were targeted by the 84 miRNAs, since miRNA targets are not limited to coding transcripts, but can also be found in ncRNAs, indicating that miRNA targeting mechanisms can be complex by binding to a wide variety of targets [28].

A total of target genes for 84 miRNA of wheat, divided into two groups according to the number of targets, and families that each miRNA obtained, since miRNAs with a larger number of target genes should be monitored in human fluids, due to their possible ability to significantly regulate human receptor cells [9]. These targets were located in various regions along the mRNA (5'UTR, CDS, 3'UTR), increasing the area of study on the interactions given by exogenous molecules, it could help understand the role of binding sites in the mRNA with the miRNAs [29]. We also evaluated the type of binding, obtaining 100% of the possible binding as imperfect base-pair by at least 1 base, however the analysis was made with the criterion of perfect binding of the seed region which would suggest that the possible mechanism of regulation of expression according to the reported silencing by mature miRNAs [2] could be translation inhibition and not degrading by cleavage the mRNA, which is sufficient for the miRNAs to suppress their targets [23,30]. In plants, miRNA bind to their targets with high complementarity, causing messenger degradation, however in animals, miRNA recognition in just their seed region is extremely important in target recognition, this complementarity can trigger expression regulation, [31,32]. Even in humans, the seed region of some miRNA can define toxicity or tumor-suppressive for some cancer cell lines [33]. This is due to seed region target recognition mechanisms [34,35], show that an initial search for target sites transitions from being transient to kinetically stable, when target complementarity binds to the full length of the seed region of miRNA [36]. Furthermore, several authors have identified and demonstrated regulation of target gene expression, even when miRNAs have not bound in a fully complementary manner, showing evidence of their function [16,37,38].

In this study we included a homology analysis, identifying 7 wheat-derived miRNAs show homology with 9 human miRNAs, revealing a series of target genes in common, through our prediction. Although plant sequences do not show high homology with human sequences, results obtained by author [39], show that they can perform similar functional activities to endogenous sequences, in a cross-kingdom interaction [19]. Therefore, we consider this dataset provides a solid basis for future projects.

In our mRNA and miRNA interaction analysis we aimed to identify the genes that are the target of multiple miRNAs and their possible health effect through KEGG pathways. We found that the target genes are associated with the calcium signaling pathway, showing participation of the CD38 gene, associated with diseases such as asthma, neoplasia and neuroimmune diseases [40], as well as the ADCY1 gene, involved in insulin secretion, related to diabetes mellitus and its correlation with pancreatic cancer [41].

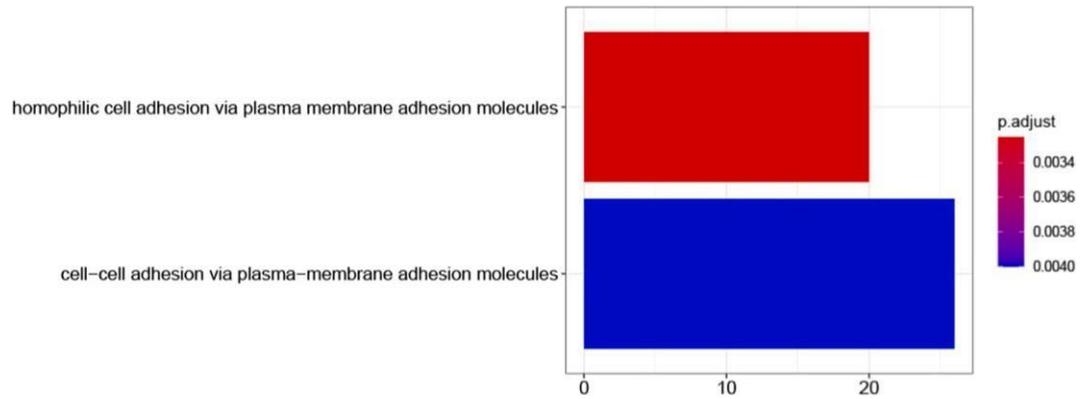
Among the pathways, three stand out, the pathway of fanconi anemia (hsa03460), circadian entrainment (hsa04713), and dopaminergic synapse (hsa04728), which we have found a relation with a couple of publications [42,43](Gluten-Free Diet for the Treatment of ADHD; Pilot Study) where in a case study they describe the difficulty of finding a diagnosis to a series of neuropsychiatric manifestations, described as "non-celiac gluten sensitivity" (NCGS), where a series of symptoms related to the consumption of gluten-containing foods (wheat) are characterized in people not affected by celiac disease (CD) or wheat allergy (AW). Attention deficit hyperactivity disorder (ADHD) related



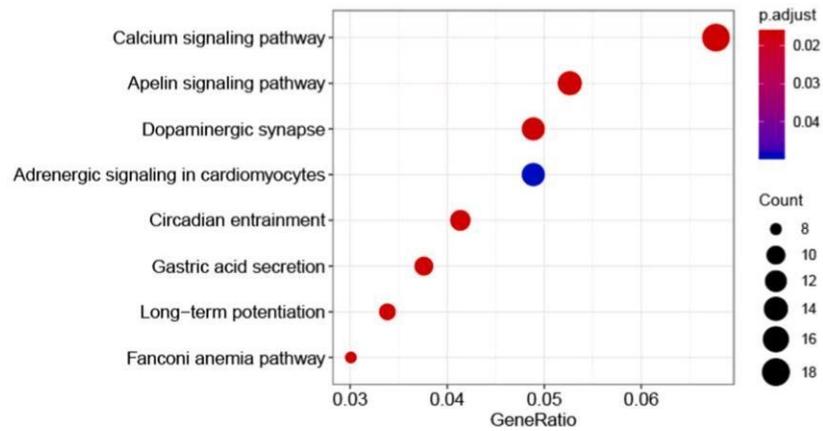
health, such as cancer, schizophrenia, and metabolic diseases.

**Table 3**  
Potential target genes in common by wheat and human miRNAs.

**Fig. 3.** Pathway enrichment of predicted human target genes derived from



**Fig. 2.** Gene ontology analysis of predicted human target genes derived from miRNAs of wheat.



miRNAs of wheat in 8 significant KEGG categories.

**1. Conclusion**

Considering that wheat and humans are different species, it is difficult to determine the possible alterations given by exogenous molecules caused by cross-kingdom interactions, however, it is necessary to highlight the importance of the use of bioinformatics tools, for the prediction of possible targets, because this study demonstrates the possible effects of wheat miRNAs in regulating complex disease networks like a schizophrenia, epilepsy, tumor suppressor, inflammation and diverse type of cancer. Due to the fact that the information collected could facilitate the search and verification, that could be of importance for future projects. In addition, these findings can generate new lines of research in various areas, such as nutrition, medicine, biotechnology, agriculture, etc. It can also show possible miRNAs candidates that can further contribute in therapeutic treatment of various human diseases.

**CRedit authorship contribution statement**

**Daniel Sánchez-Romo:** Writing – original draft, Methodology, Investigation, Case Resources. **I. Hernández-Vásquez:** Formal analysis **Benito Pereyra-Alferez:** Conceptualization, Resources. **Jorge H. García-García:** Project administration, Writing – review & editing

Tae-miRNA	Hsa-miRNA	Seed Score	Genes
Tae-miR5772	Hsa-miR-143-3p	0.8125	AAK1
Tae-miR164	Hsa-miR-139-3p	0.8125	ZNF701, ESTL4, RASSF6, LRRC55, MNT
Tae-miR164	Hsa-miR-5703	0.8125	ZNF701, SLC9A9, GREM1, GLI3, SNED1, SMARCC2, CSRP1, B3GNT7, RASSF6, SLC25A1, NCAPH, LRRC55, MNT
Tae-miR666h	Hsa-miR-4259	0.8125	PCDH41, PBXIP1
Tae-miR5048-5p	Hsa-miR-1205	0.8125	PCDH6, PCDHA10, PCDHA13, PCDHA2, PCDHA4, PCDHAC2, PCDHA3, RBM33, ARID5B, PCDHA12, PCDHA5, PCDHA7, CLOCK
Tae-miR395a	Hsa-miR-203a-3p	0.8125	CHRM3, ADAM22, GRM4
Tae-miR395a	Hsa-miR-203a-3p	0.8125	VAPA, PPARGC1B

## Declaration of competing interest

None.

## Acknowledgments

Non applicable.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ncrna.2022.03.002>.

## References

- [1] V. Ambros, The functions of animal microRNAs, *Nature* 431 (2004) 350–355, <https://doi.org/10.1038/nature02871>.
- [2] R.C. Friedman, K.K.-H. Farh, C.B. Burge, D.P. Bartel, Most mammalian mRNAs are conserved targets of microRNAs, *Genome Res.* 19 (2009) 92–105, <https://doi.org/10.1101/gr.082701.108>.
- [3] P. Graves, Y. Zeng, Biogenesis of mammalian MicroRNAs: a global view, *Genomics, Proteomics Bioinforma* 10 (2012) 239–245, <https://doi.org/10.1016/j.gpb.2012.06.004>.
- [4] I. Lee, S.S. Ajay, I.Y. Jong, S.K. Hyun, H.H. Su, H.K. Nam, S.M. Dhanasekaran, A. M. Chinnaiyan, B.D. Athey, New class of microRNA targets containing simultaneous 5'-UTR and 3'-UTR interaction sites, *Genome Res.* 19 (2009) 1175–1183, <https://doi.org/10.1101/gr.089367.108>.
- [5] P. Perge, Z. Nagy, A. Decmann, I. Igaz, P. Igaz, Potential relevance of microRNAs in inter-species epigenetic communication, and implications for disease pathogenesis, *RNA Biol.* 14 (2017) 391–401, <https://doi.org/10.1080/15476286.2016.1251001>.
- [6] Y. Zhao, L. Cong, W.J. Lukiw, Plant and animal microRNAs (miRNAs) and their potential for inter-kingdom communication, *Cell. Mol. Neurobiol.* 38 (2018) 133–140, <https://doi.org/10.1007/s10571-017-0547-4>.
- [7] H. Liang, S. Zhang, Z. Fu, Y. Wang, N. Wang, Y. Liu, C. Zhao, J. Wu, Y. Hu, J. Zhang, X. Chen, K. Zen, C.Y. Zhang, Effective detection and quantification of dietetically absorbed plant microRNAs in human plasma, *J. Nutr. Biochem.* 26 (2015) 505–512, <https://doi.org/10.1016/j.jnutbio.2014.12.002>.
- [8] B.J. Stephen, N. Pareek, M. Saeed, M.A. Kausar, S. Rahman, M. Datta, XenomiRNA in maternal-infant immune crosstalk: an aid to disease alleviation, *Front. Immunol.* 11 (2020) 1–8, <https://doi.org/10.3389/fimmu.2020.00404>.
- [9] A. Rakhmetullina, A. Pyrkova, D. Aisina, A. Ivashchenko, In silico prediction of human genes as potential targets for rice miRNAs, *Comput. Biol. Chem.* 87 (2020), 107305, <https://doi.org/10.1016/j.compbiolchem.2020.107305>.
- [10] D. Pirim, B. Dogan, In silico identification of putative roles of food-derived xenomirs on diet-associated cancer, *Nutr. Cancer* 72 (2020) 481–488, <https://doi.org/10.1080/01635581.2019.1670854>.
- [11] M. Patel, N. Mangukia, N. Jha, H. Gadhavi, K. Shah, S. Patel, A. Mankad, H. Pandya, R. Rawal, Computational identification of miRNA and their cross kingdom targets from expressed sequence tags of *Ocimum basilicum*, *Mol. Biol. Rep.* 46 (2019) 2979–2995, <https://doi.org/10.1007/s11033-019-04759-x>.
- [12] J. He, T. Chen, Q. Xi, J. Sun, J. Luo, M. Li, H. Zhang, B. Zeng, J. Wu, Y. Zhang, Identification of microRNA in *Houttuynia cordata* Thunb and prediction of cross kingdom functions, *ExRNA* 1 (2019), <https://doi.org/10.1186/s41544-019-0028-7>.
- [13] D. Kumar, S. Kumar, G. Ayachit, S.B. Bhairappanavar, A. Ansari, P. Sharma, S. Soni, J. Das, Cross-kingdom regulation of putative miRNAs derived from happy tree in cancer pathway: a systems biology approach, *Int. J. Mol. Sci.* 18 (2017) 1–21, <https://doi.org/10.3390/ijms18061191>.
- [14] H. Gadhavi, M. Patel, N. Mangukia, K. Shah, K. Bhadrasha, S.K. Patel, R.M. Rawal, H.A. Pandya, Transcriptome-wide miRNA identification of *Bacopa monnieri*: a cross-kingdom approach, *Plant Signal. Behav.* 15 (2020), <https://doi.org/10.1080/15592324.2019.1699265>.
- [15] L. Zhang, D. Hou, X. Chen, D. Li, L. Zhu, Y. Zhang, J. Li, Z. Bian, X. Liang, X. Cai, Y. Yin, C. Wang, T. Zhang, D. Zhu, D. Zhang, J. Xu, Exogenous Plant MIR168a Specifically Targets Mammalian LDLRAP1: Evidence of Cross-Kingdom Regulation by microRNA, 2012, pp. 107–126, <https://doi.org/10.1038/cr.2011.158>.
- [16] Z. Zhou, X. Li, J. Liu, L. Dong, Q. Chen, J. Liu, H. Kong, Q. Zhang, X. Qi, D. Hou, L. Zhang, G. Zhang, Y. Liu, Y. Zhang, J. Li, J. Wang, X. Chen, H. Wang, J. Zhang, H. Chen, K. Zen, C.Y. Zhang, Honeysuckle-encoded atypical microRNA2911 directly targets influenza A viruses, *Cell Res.* 25 (2015) 39–49, <https://doi.org/10.1038/cr.2014.130>.
- [17] R. Giegerich, M. Rehmsmeier, P. Steffen, M. Ho, Fast and effective prediction of microRNA/target duplexes, *RNA* 10 (2004) 1507–1517, <https://doi.org/10.1261/rna.5248604.and>.
- [18] S. Pirro, A. Minutolo, A. Galgani, M. Potestà, V. Colizzi, C. Montesano, Bioinformatics prediction and experimental validation of MicroRNAs involved in cross-kingdom interaction, *J. Comput. Biol.* 23 (2016) 976–989, <https://doi.org/10.1089/CMB.2016.0059>.
- [19] A. Minutolo, M. Potestà, A. Gismondi, S. Pirro, M. Cirilli, F. Gattabria, A. Galgani, L. Sessa, M. Mattei, A. Canini, R. Muleo, V. Colizzi, C. Montesano, Olea europaea small RNA with functional homology to human miR34a in cross-kingdom interaction of anti-tumoral response, *Sci. Rep.* 8 (2018) 1–14, <https://doi.org/10.1038/s41598-018-30718-w>.
- [20] G. Yu, L.G. Wang, Y. Han, Q.Y. He, ClusterProfiler: an R package for comparing biological themes among gene clusters, *OMICS A J. Integr. Biol.* 16 (2012) 284–287, <https://doi.org/10.1089/omi.2011.0118>.
- [21] R. R core Team, A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, 2018. <https://www.r-project.org/>.
- [22] R. Han, C. Jian, J. Lv, Y. Yan, Q. Chi, Z. Li, Q. Wang, J. Zhang, X. Liu, H. Zhao, Identification and characterization of microRNAs in the flag leaf and developing seed of wheat (*Triticum aestivum* L.), *BMC Genome.* 15 (2014) 289, <https://doi.org/10.1186/1471-2164-15-289>.
- [23] B. Liu, J. Li, M.J. Cairns, Identifying miRNAs, targets and functions, *Briefings Bioinf.* 15 (2014) 1–19, <https://doi.org/10.1093/bib/bbs075>.
- [24] H. Min, S. Yoon, Got target?: computational methods for microRNA target prediction and their extension, *Exp. Mol. Med.* 42 (2010) 233–244, <https://doi.org/10.3858/emm.2010.42.4.032>.
- [25] C. Liu, W.A. Rennie, C.S. Carmack, S. Kanoria, J. Cheng, J. Lu, Y. Ding, Effects of genetic variations on microRNA: target interactions, *Nucleic Acids Res.* 42 (2014) 9543–9552, <https://doi.org/10.1093/nar/gku675>.
- [26] H. Zhang, Y. Li, Y. Liu, H. Liu, H. Wang, W. Jin, Y. Zhang, C. Zhang, D. Xu, Role of plant MicroRNA in cross-species regulatory networks of humans, *BMC Syst. Biol.* 10 (2016) 1–10, <https://doi.org/10.1186/s12918-016-0292-1>.
- [27] Y. Luo, P. Wang, X. Wang, Y. Wang, Z. Mu, Q. Li, Y. Fu, J. Xiao, G. Li, Y. Ma, Y. Gu, L. Jin, J. Ma, Q. Tang, A. Jiang, X. Li, M. Li, Detection of dietetically absorbed maize-derived microRNAs in pigs, *Sci. Rep.* 7 (2017) 1–10, <https://doi.org/10.1038/s41598-017-00488-y>.
- [28] A. Helwak, G. Kudla, T. Dudnakova, D. Tollervy, Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding, *Cell* 153 (2013) 654–665, <https://doi.org/10.1016/j.cell.2013.03.043>.
- [29] W. Gu, Y. Xu, X. Xie, T. Wang, J.H. Ko, T. Zhou, The role of RNA structure at 5' untranslated region in microRNA-mediated gene regulation, *RNA* 20 (2014) 1369–1375, <https://doi.org/10.1261/rna.044792.114>.
- [30] J.K.W. Lam, M.Y.T. Chow, Y. Zhang, S.W.S. Leung, siRNA versus miRNA as therapeutics for gene silencing, *Mol. Ther. Nucleic Acids* 4 (2015) e252, <https://doi.org/10.1038/mtna.2015.23>.
- [31] A.F.A. Samad, M.F. Kamaruddin, M. Sajad, Cross-kingdom regulation by plant microRNAs provides novel insight into gene regulation, *Adv. Nutr* (2020) 1–15, <https://doi.org/10.1093/advances/nmaa095>.
- [32] E. Huntzinger, E. Izaurralde, Gene silencing by microRNAs: contributions of translational repression and mRNA decay, *Nat. Rev. Genet.* 12 (2011) 99–110, <https://doi.org/10.1038/nrg2936>.
- [33] Q.Q. Gao, W.E. Putzbach, A.E. Murmann, S. Chen, A.A. Sarshad, J.M. Peter, E. T. Bartom, M. Hafner, M.E. Peter, 6mer seed toxicity in tumor suppressive microRNAs, *Nat. Commun.* 9 (2018), <https://doi.org/10.1038/s41467-018-06526-1>.
- [34] S.M. Klum, S.D. Chandradoss, N.T. Schirle, C. Joo, I.J. Macrae, Helix-7 in Argonaute2 shapes the microRNA seed region for rapid target recognition, *EMBO J.* 37 (2018) 75–88, <https://doi.org/10.15252/embj.201796474>.
- [35] M. Klein, S.D. Chandradoss, M. Depken, C. Joo, Why Argonaute is needed to make microRNA target search fast and reliable, *Semin. Cell Dev. Biol.* 65 (2017) 20–28, <https://doi.org/10.1016/j.semcdb.2016.05.017>.
- [36] S.D. Chandradoss, N.T. Schirle, M. Szczepaniak, I.J. Macrae, C. Joo, A dynamic search process underlies MicroRNA targeting, *Cell* 162 (2015) 96–107, <https://doi.org/10.1016/j.cell.2015.06.032>.
- [37] M. Li, T. Chen, J.J. He, J.H. Wu, J.Y. Luo, R.S. Ye, M.Y. Xie, H.J. Zhang, B. Zeng, J. Liu, Q.Y. Xi, Q.Y. Jiang, J.J. Sun, Y.L. Zhang, Plant MIR167e-5p inhibits enterocyte proliferation by targeting  $\beta$ -catenin, *Cells* 8 (2019) 1–14, <https://doi.org/10.3390/cells8111385>.
- [38] A.R. Chin, M.Y. Fong, G. Somlo, J. Wu, P. Swiderski, X. Wu, S.E. Wang, Cross-kingdom inhibition of breast cancer growth by plant miR159, *Cell Res.* 26 (2016) 217–228, <https://doi.org/10.1038/cr.2016.13>.
- [39] B. Avsar, Y. Zhao, W. Li, W.J. Lukiw, *Atropa belladonna* expresses a microRNA (aba-miRNA-9497) highly homologous to Homo sapiens miRNA-378 (hsa-miRNA-378); both miRNAs target the 3'-untranslated region (3'-UTR) of the mRNA encoding the neurologically relevant, zinc-finger transcription factor, *Cell. Mol. Neurobiol.* 40 (2020) 179–188, <https://doi.org/10.1007/s10571-019-00729-w>.
- [40] D.A. Deshpande, A.G.P. Guedes, R. Graeff, S. Dogan, S. Subramanian, T.F. Walseth, M.S. Kannan, CD38/cADPR Signaling Pathway in Airway Disease: Regulatory Mechanisms, *Mediators Inflamm.* 2018, <https://doi.org/10.1155/2018/8942042>, 2018.
- [41] W. Pang, W. Yao, X. Dai, A. Zhang, L. Hou, L. Wang, Y. Wang, X. Huang, X. Meng, L. Li, Pancreatic cancer-derived exosomal microrna-19a induces  $\beta$ -cell dysfunction by targeting *adcyl1* and *epac2*, *Int. J. Biol. Sci.* 17 (2021) 3622–3633, <https://doi.org/10.7150/ijbs.56271>.
- [42] C. Catassi, J.C. Bai, B. Bonaz, G. Bouma, A. Calabro, A. Carroccio, G. Castillejo, C. Ciacci, F. Cristofori, J. Dolinsek, R. Francavilla, L. Elli, P. Green, W. Holtmeier, P. Koehler, S. Koletzko, C. Meinhold, D. Sanders, M. Schumann, D. Schuppman, R. Ullrich, A. V'ecsei, U. Volta, V. Zevallos, A. Sapone, A. Fasano, Non-celiac gluten



## Draft Genome Sequence of *Bacillus toyonensis* Strain GM18, Isolated from Agricultural Soil

Carlos J. Moreno-Peña,<sup>a</sup>  Jorge H. García-García,<sup>a</sup> Cesar I. Hernández-Vásquez,<sup>a</sup> Luis J. Galán-Wong,<sup>a</sup>  Benito Pereyra-Alfárez<sup>a</sup>

<sup>a</sup>Instituto de Biotecnología, Facultad de Ciencias Biológicas, Universidad Autónoma de Nuevo León, San Nicolas de los Garza, Nuevo León, Mexico

**ABSTRACT** *Bacillus toyonensis* is a recently described species related to *Bacillus cereus* and *Bacillus thuringiensis*. The GM18 strain previously identified as *B. thuringiensis* is now classified as *B. toyonensis* based on the RNA 16S sequence and whole-genome average nucleotide identity. The genome analysis revealed the presence of insecticide, nematocidal, and antitumoral proteins.

The GM18 bacterium was isolated from agricultural soil collected in the state of Guanajuato, Mexico, using the methodology described by Travers et al. (1), and based on its H antigen characteristics, this strain was recognized as a new serotype (H-24) and named *Bacillus thuringiensis* subsp. *neolonensis*. In an early description, Rodriguez-Padilla et al. (2) showed that GM18 possesses characteristics of the *B. thuringiensis* group; it is Gram positive and sporulated and synthesizes a parasporal crystal. However, it had no insecticidal activity. Later, when the inclusion body was tested against cancer cell lines, solubilized proteins displayed biological activity, and cytotoxicity ranged from low against 1c1c7 and A431 to high against A375 and BpRc1 cell lines, but no cytotoxicity for sheep erythrocytes was detected (3).

For the genome sequencing of *B. toyonensis* GM18, cells obtained from our laboratory cryostock were grown in 25 ml of LB broth for 6 h at 30°C, and then the pellet was recovered after centrifugation and washed with sterile water. The pellet was shipped chilled to Genewiz (New Jersey, USA), where total DNA extraction was performed, and the DNA library was prepared with the Illumina DNA prep, (M) tagmentation, and IDT for Illumina DNA/RNA indexes set A tagmentation kits (Illumina, Inc., California, USA). Whole-genome sequencing of *B. toyonensis* GM18 was performed with paired-end Illumina MiSeq sequencing, which resulted in 11,175,537 paired-end reads of 137 to 150 bp with a coverage of 250×; for genome assembly and annotation, we used Galaxy (4). In brief, the sequence quality was determined with FastQC (Galaxy version 0.721galaxy1) (5). Then, low-quality sequences and adapters were removed with Trimmomatic (Galaxy version 0.038.0) (6). The processed sequence reads were assembled with the Shovill pipeline, including the SPAdes assembler (Galaxy version 1.1.01galaxy0) (7); the resulting contigs were annotated with Prokka (Galaxy version 1.14.51galaxy0) (8) and with the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) using the best-placed reference protein set method (GeneMarkS-21) (9, 10); 91.59% of all sequences were assembled in 156 contigs with an  $N_{50}$  value of 171,900 bp. Default parameters were used for all software.

We found that GM18 has a 6.13-Mb genome containing 6,129 open reading frames (ORF) with a GC content of 34.93%. The genome has 96 operons for tRNAs and 16 for rRNAs. The average nucleotide identity (ANI) of the whole genome determined during the GenBank submission process (11) showed 99.018% identity with the type genome of *Bacillus toyonensis*; this and the sequence identity of the rRNA 16S (GenBank accession number [JADNYS010000095.1](https://doi.org/10.1128/MRA.00008-21)), *gyrB* ([MBF7148837.1](https://doi.org/10.1128/MRA.00008-21)), *groEL* ([MBF7150603.1](https://doi.org/10.1128/MRA.00008-21)), and

Citation Moreno-Peña CJ, García-García JH, Hernández-Vásquez CI, Galán-Wong LJ, Pereyra-Alfárez B. 2021. Draft genome sequence of *Bacillus toyonensis* strain GM18, isolated from agricultural soil. *Microbiol Resour Announc* 10:e00008-21. <https://doi.org/10.1128/MRA.00008-21>.

Editor Frank J. Stewart, Georgia Institute of Technology

Copyright © 2021 Moreno-Peña et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Benito Pereyra-Alfárez, [benito.pereyraal@uanl.edu.mx](mailto:benito.pereyraal@uanl.edu.mx).

Received 11 January 2021

Accepted 5 February 2021

Published 4 March 2021

*XRE* (MBF7150840.1) genes suggests that this isolate belongs to the *B. toyonensis* species (12).

In brief, the sequence of the GM18 genome could be key to understanding the *Bacillus* group and its members' evolutionary relations. This genome shows greater identity with *Bacillus toyonensis*, although it has protein-coding genes with identity to *cry* (GenBank accession number WP\_195755641.1, WP\_195755715.1), *vip* (WP\_195755724.1, WP\_195755747.1), and parasporin (WP\_195755716.1) that are characteristic of *B. thuringiensis* species; it also has insertion sequences with identity to those of *B. thuringiensis*, but this could be because *B. toyonensis* is a relatively newly described species of the *Bacillus* group (13).

Data availability. The *Bacillus toyonensis* GM18 whole-genome shotgun sequence has been deposited in GenBank under the accession number JADNYS000000000 and BioProject number PRJNA676032. The raw sequence files can be found under SRA number SRP301246.

#### ACKNOWLEDGMENT

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

#### REFERENCES

- Travers RS, Martin PA, Reichelderfer CF. 1987. Selective process for efficient isolation of soil *Bacillus* spp. *Appl Environ Microbiol* 53:1263–1266. <https://doi.org/10.1128/AEM.53.6.1263-1266.1987>.
- Rodríguez-Padilla C, Galan-Wong L, de Barjac H, Roman-Calderon E, Tamez-Guerra R, Dulmage H. 1990. *Bacillus thuringiensis* subspecies neoleonensis serotype H-24, a new subspecies which produces a triangular crystal. *J Invertebr Pathol* 56:280–282. [https://doi.org/10.1016/0022-2011\(90\)90111-1](https://doi.org/10.1016/0022-2011(90)90111-1).
- Espino-Vázquez AN. 2014. Caracterización biológica de parasporinas en cepas nativas de *Bacillus thuringiensis*. PhD thesis. Universidad Autónoma de Nuevo León, Nuevo León, Mexico. <http://eprints.uanl.mx/id/eprint/4042>.
- Blankenberg D, Von Kuster G, Coraor N, Ananda G, Lazarus R, Mangan M, Nekrutenko A, Taylor J. 2010. Galaxy: a Web-based genome analysis tool for experimentalists. *Curr Protoc Mol Biol* 89:19.10.1–19.10.21. <https://doi.org/10.1002/0471142727.mb1910s89>.
- Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
- Seemann T. 2017. Shovill: faster SPAdes assembly of Illumina reads. <https://github.com/tseemann/shovill>.
- Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
- Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.
- Haft DH, DiCuccio M, Badretdin A, Brover V, Chetvernin V, O'Neill K, Li W, Chitsaz F, Derbyshire MK, Gonzales NR, Gwadz M, Lu F, Marchler GH, Song JS, Thanki N, Yamashita RA, Zheng C, Thibaud-Nissen F, Geer LY, Marchler-Bauer A, Pruitt KD. 2018. RefSeq: an update on prokaryotic genome annotation and curation. *Nucleic Acids Res* 46:D851–D860. <https://doi.org/10.1093/nar/gkx1068>.
- Ciufo S, Kannan S, Sharma S, Badretdin A, Clark K, Turner S, Brover S, Schoch CL, Kimchi A, DiCuccio M. 2018. Using average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the NCBI. *Int J Syst Evol Microbiol* 68:2386–2392. <https://doi.org/10.1099/ijsem.0.002809>.
- Chelliah R, Wei S, Park BJ, Rubab M, Banan-Mwine Dalirri E, Barathkannan K, Jin YG, Oh DH. 2019. Whole genome sequence of *Bacillus thuringiensis* ATCC 10792 and improved discrimination of *Bacillus thuringiensis* from *Bacillus cereus* group based on novel biomarkers. *Microb Pathog* 129:284–297. <https://doi.org/10.1016/j.micpath.2019.02.014>.
- Jiménez G, Urdiain M, Cifuentes A, López-López A, Blanch AR, Tamames J, Kämpfer P, Kolstø AB, Ramón D, Martínez JF, Codoñer FM, Rosselló-Móra R. 2013. Description of *Bacillus toyonensis* sp. nov., a novel species of the *Bacillus cereus* group, and pairwise genome comparisons of the species of the group by means of ANI calculations. *Syst Appl Microbiol* 36:383–391. <https://doi.org/10.1016/j.syapm.2013.04.008>.



## Draft Genome Sequence of *Bacillus albus* Strain IB84, Isolated from Mexican Soil

Jorge H. García-García,<sup>a</sup> Carlos J. Moreno-Peña,<sup>a</sup> Cesar I. Hernández-Vásquez,<sup>a</sup> Luis J. Galán-Wong,<sup>a</sup> Benito Pereyra-Alfárez<sup>a</sup>

<sup>a</sup>Instituto de Biotecnología, Facultad de Ciencias Biológicas, Universidad Autónoma de Nuevo León, San Nicolas de los Garza, Nuevo León, Mexico

**ABSTRACT** *Bacillus albus* is a new species, but it lies on the borderline with *Bacillus thuringiensis*. In this work, we report a strain previously identified as *Bacillus thuringiensis* IB84, which now, based on average nucleotide identity and rRNA 16S, *gyrB*, *groEL*, and *xre* gene sequences, must be identified as *Bacillus albus*.

The IB84 strain was originally recognized as *Bacillus thuringiensis* because it is Gram positive and sporulated and produces a small parasporal crystal. This bacterium was originally isolated from agricultural soil collected in the state of San Luis Potosí, Mexico, using the methodology described by Travers et al. (1), and it belongs to the *B. thuringiensis* collection of the Instituto de Biotecnología of the Facultad de Ciencias Biológicas-UANL. Preliminary bioassays revealed that crystal inclusion proteins were not toxic against several insects (2).

For the whole-genome sequencing of *Bacillus albus* IB84, cells obtained from our laboratory cryostock were grown in 25 ml of LB broth for 6 h at 30°C, and then the pellet was recovered after centrifugation and washed with sterile water. The pellet was shipped under cold conditions to Genewiz (New Jersey, USA), where total DNA extraction was performed using the EasyPure genomic DNA kit (Genewiz), and the DNA library was prepared with the DNA prep (M) tagmentation, and IDT for Illumina DNA/RNA indexes set A tagmentation kits (Illumina, Inc., California, USA). The whole genome was sequenced using paired-end Illumina MiSeq sequencing, which resulted in 10,065,062 paired-end reads of 118 to 150 bp with a coverage of 250×. For genome assembly and annotation, we used Galaxy at <http://usegalaxy.org> (3). In brief, the sequence quality was determined with FastQC (Galaxy version 0.721galaxy1) (4). Then, low-quality sequences and adapters were removed with Trimmomatic (Galaxy version 0.038.0) (5). Processed sequence reads were assembled with the Shovill pipeline, which includes the SPAdes assembler (Galaxy version 1.1.01galaxy0) (6); the resulting contigs were annotated with the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) using the best-placed reference protein set method (GeneMarkS-21) (7, 8); 92.03% of all sequence reads were assembled in 156 contigs. Default parameters were used for all software.

We identified a 5.96-Mb genome with an  $N_{50}$  value of 155,983 bp, 6,015 open reading frames (ORF), 35.06% GC content, 74 tRNA operons, 6 rRNAs, and 5 noncoding RNAs (ncRNAs). The average nucleotide identity (ANI) of the whole genome determined during the GenBank submission process (9) showed 97.907% identity with the type genome of *Bacillus albus*; this and the sequence identities of rRNA 16S, *gyrB*, *groEL*, and *xre* genes suggest that this isolate belongs to the *Bacillus albus* species (10). This species has been described recently as having an important salt tolerance and the ability to degrade poultry waste processing industrial (PWPI) effluent (10, 11). The IB84 whole-genome analysis revealed sequences associated with assimilation and detoxification of nitrogen and arsenate compounds, as well as salt tolerance, antibiotic resistance, and oxidative stress genes.

The whole-genome sequence, assembly, and annotation of IB84 help to further understand *B. albus* and its evolutionary relationships with the *Bacillus* group. Furthermore, it

Citation García-García JH, Moreno-Peña CJ, Hernández-Vásquez CI, Galán-Wong LJ, Pereyra-Alfárez B. 2021. Draft genome sequence of *Bacillus albus* strain IB84, isolated from Mexican soil. Microbiol Resour Announc 10:e00274-21. <https://doi.org/10.1128/MRA.00274-21>.

Editor Steven R. Gill, University of Rochester School of Medicine and Dentistry

Copyright © 2021 García-García et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Benito Pereyra-Alfárez, [benito.pereyraal@uanl.edu.mx](mailto:benito.pereyraal@uanl.edu.mx).

Received 22 March 2021

Accepted 23 March 2021

Published 29 April 2021

allows us to study the possible advantages of its genetics, its biochemical machinery, and its biotechnological potential.

Data availability. The *Bacillus albus* IB84 whole-genome shotgun sequence has been deposited in GenBank under the accession number [JADNYT000000000](https://doi.org/10.1128/AEM.53.6.1263-1266.1987) and BioProject number [PRJNA676033](https://doi.org/10.1128/AEM.53.6.1263-1266.1987). The raw sequence files can be found under SRA number [SRP301583](https://doi.org/10.1128/AEM.53.6.1263-1266.1987).

#### ACKNOWLEDGMENT

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

#### REFERENCES

1. Travers RS, Martin PA, Reichelderfer CF. 1987. Selective process for efficient isolation of soil *Bacillus* spp. *Appl Environ Microbiol* 53:1263–1266. <https://doi.org/10.1128/AEM.53.6.1263-1266.1987>.
2. Espino-Vázquez AN, Gómez-Treviño A, Galán-Wong LJ, Pereyra-Alfárez B. 2012. Isolation of *Bacillus thuringiensis* strains with cytotoxic activity against MOLT-4, a leukemia cell line, p 147–151. *In* Mendez-Vilas A (ed), *Microbes in applied research: current advances and challenges*. World Scientific Publishing Co., Singapore. <https://doi.org/10.1142/8462>.
3. Blankenberg D, Von Kuster G, Coraor N, Ananda G, Lazarus R, Mangan M, Nekrutenko A, Taylor J. 2010. Galaxy: a Web-based genome analysis tool for experimentalists. *Curr Prot Mol Biol* 89:19.10.1–19.10.21. <https://doi.org/10.1002/0471142727.mb1910s89>.
4. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
5. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
6. Seemann T. 2017. Shovill: faster SPAdes assembly of Illumina reads. <https://github.com/tseemann/shovill>.
7. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.
8. Haft DH, DiCuccio M, Badretdin A, Brover V, Chetvernin V, O'Neill K, Li W, Chitsaz F, Derbyshire MK, Gonzales NR, Gwadz M, Lu F, Marchler GH, Song JS, Thanki N, Yamashita RA, Zheng C, Thibaud-Nissen F, Geer LY, Marchler-Bauer A, Pruitt KD. 2018. RefSeq: an update on prokaryotic genome annotation and curation. *Nucleic Acids Res* 46:D851–D860. <https://doi.org/10.1093/nar/gkx1068>.
9. Ciufu S, Kannan S, Sharma S, Badretdin A, Clark K, Turner S, Brover S, Schoch CL, Kimchi A, DiCuccio M. 2018. Using average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the NCBI. *Int J Syst Evol Microbiol* 68:2386–2392. <https://doi.org/10.1099/ijsem.0.002809>.
10. Chelliah R, Wei S, Park B-J, Jin Y-G, Rubab M, Daliri E, Kannan B, Oh D-H. 2019. Whole genome sequence of *Bacillus thuringiensis* ATCC 10792 and improved discrimination of *Bacillus thuringiensis* from *Bacillus cereus* group based on novel biomarkers. *Microb Pathog* 129:284–297. <https://doi.org/10.1016/j.micpath.2019.02.014>.
11. Ben Mefteh F, Bouket AC, Daoud A, Luptakova L, Alenezi FN, Gharsallah N, Belbahri L. 2019. Metagenomic insights and genomic analysis of phosphogypsum and its associated plant endophytic microbiomes reveals valuable actors for waste bioremediation. *Microorganisms* 7:382. <https://doi.org/10.3390/microorganisms7100382>.

## CAPÍTULO 9

### EL TIPO DE MALTA IMPACTA EN EL PERFIL Y ACTIVIDAD DE PROTEASAS<sup>1</sup>

Data de submissão: 30/01/2023

Data de aceite: 17/01/2023

**Claudia Berenice López-Alvarado**  
Instituto de Biotecnología  
Facultad de Ciencias Biológicas  
Universidad Autónoma de Nuevo León  
San Nicolás de los Garza  
Nuevo León, México  
ORCID: 0000-0002-9797-0120

**Jessica Giselle Herrera-Gamboa**  
Instituto de Biotecnología  
Facultad de Ciencias Biológicas  
Universidad Autónoma de Nuevo León  
San Nicolás de los Garza  
Nuevo León, México  
ORCID: 0000-0002-1524-7386

**Jorge Hugo García-García**  
Instituto de Biotecnología  
Facultad de Ciencias Biológicas  
Universidad Autónoma de Nuevo León  
San Nicolás de los Garza  
Nuevo León, México  
ORCID: 0000-0001-8758-9919

**César Ignacio Hernández-Vásquez**  
Instituto de Biotecnología  
Facultad de Ciencias Biológicas  
Universidad Autónoma de Nuevo León  
San Nicolás de los Garza  
Nuevo León, México  
ORCID: 0000-0002-2067-1122

**Esmeralda Pérez-Ortega**  
Laboratorio de Investigación y  
Desarrollo, Cervecería Cuauhtémoc  
Moctezuma S. A de C. V.  
Monterrey, Nuevo León, México

**Luis Cástulo Damas-Buenrostro**  
Laboratorio de Investigación y  
Desarrollo, Cervecería Cuauhtémoc  
Moctezuma S. A de C. V.  
Monterrey, Nuevo León, México

**Benito Pereyra-Alfárez<sup>2</sup>**  
Instituto de Biotecnología  
Facultad de Ciencias Biológicas  
Universidad Autónoma de Nuevo León  
San Nicolás de los Garza  
Nuevo León, México  
ORCID: 0000-0002-3726-0389

<sup>1</sup> FUENTE DE FINANCIAMIENTO: Agradecemos al Consejo Nacional de Ciencia y Tecnología (CONACYT) por su apoyo financiero a través del proyecto apoyado por el Programa de Estímulos a la Investigación, Desarrollo Tecnológico e

**RESUMEN:** La cebada malteada representa el principal componente del mosto cervecero y las enzimas proteasas juegan un papel fundamental en la calidad. Fue comparada la actividad proteolítica en cebadas malteadas,





(CCM), México. Las muestras fueron almacenadas a -20 °C hasta su uso. Adicionalmente, fue utilizada la malta comercial 2-ROW como control, obtenida de la empresa Malteurop (Milwaukee, WI, USA).

## 1.1 EXTRACCIÓN DE PROTEÍNAS

Las muestras de malta (10 g) fueron molidas hasta un polvo fino (<50 µm) usando un molino de marca Foss Cyclotec 1093. Las proteínas fueron extraídas con buffer de extracción (Tris-HCl 25 mM; MgCl<sub>2</sub> 5 mM pH 7.5), según lo reportado por López et al. (2016). La concentración de proteína soluble total se determinó mediante el método de Bradford.

## 1.2 EFECTO DEL PH EN LAS HIDROLASAS PRESENTES EN LA CEBADA MALTEADA

La evaluación del efecto del pH sobre la actividad enzimática tipo amilasa fue realizada en microplaca de 96 pozos de acuerdo con la metodología descrita por Oliveira et al. (2019), con algunas modificaciones. Las pruebas fueron realizadas bajo diferentes condiciones de pH utilizando buffer de acetato de sodio 0.05 M (pH 3.0, 4.0 y 5.0) y buffer de fosfato de sodio 0.1 M (pH 6.0, 7.0 y 9.0). En cada pozo fueron colocados: 20 µl de cada buffer a sus respectivos niveles de pH, 10 µl de solución de almidón de papa al 0.2 % y 10 µl de cada muestra. La microplaca fue incubada a 40 °C por 30 min. Terminado el tiempo de incubación, fueron añadidos 30 µl de la solución de Lugol a una dilución 1:10 y 170 µl de agua bidestilada. Finalmente, la microplaca fue leída a 595 nm. El efecto del pH en la actividad de proteasa fue realizada de acuerdo con la metodología descrita por Secades y Guijarro (1999) con algunas modificaciones. Para este ensayo fue empleada azocaseína 1.0 % como sustrato disuelta en buffer Tris-HCl 25 mM con MgCl<sub>2</sub> 5mM en un rango de pH de 3.0 a 11.0. A 60 µl de cada muestra de malta fueron agregados 240 µl de azocaseína 1.0 %, mezclados por vórtex e incubadas a 40°C por 30 min. La reacción fue detenida con 300 µl de TCA 10 %, incubado en hielo por 30 min y centrifugación a 15,000 g por 10min a 4 °C. Pasado el tiempo fueron tomados 400 µl del sobrenadante y fueron neutralizados con 100 µl de NaOH 1.8 N. Finalmente, se leyó la absorbancia a 420 nm. Una unidad de actividad enzimática se definió como la cantidad que produjo un aumento en A420 de 0.01 en 30 min a 30 °C. Se realizaron triplicados biológicos para todas las pruebas.

## 2.4 ACTIVIDAD PROTEOLÍTICA

La actividad enzimática fue evaluada mediante zimogramas de una dimensión como lo describen Herrera et al. (2018) y López et al. (2016) con algunas modificaciones.



#### 241 Nitrógeno amino libre (FAN)

Fue utilizado el método colorimétrico de ninhidrina para determinar la cantidad de FAN en el mosto. (Método de Wort-12). Los mostos (1 ml) se diluyeron a 100 ml con agua destilada y 2 ml de la dilución fueron transferidos a tubos de ensayo por triplicado. Después, se añadió 1 ml de reactivo de ninhidrina ( $\text{Na}_2\text{HPO}_4$  2.8 mM, ninhidrina 0.3 mM,  $\text{KH}_2\text{PO}_4$  4.4 mM, pH 6.6 – 6.8) a cada muestra y se colocó en un baño de agua hirviendo durante 16 minutos. La reacción se dejó enfriar durante 20 min en un baño de agua a 20 °C. Posteriormente, fueron añadidos 5 ml de la solución de dilución ( $\text{KIO}_3$  9.4 mM, etanol al 38.4 %) a cada muestra. Después de mezclar bien la reacción, la absorbancia es medida a 570 nm contra agua bidestilada. Fue utilizada solución de glicina 0.14 mM como solución estándar. Para los cálculos se utilizó la siguiente fórmula:  $\text{FAN (mg/ml)} = (\text{absorbancia neta de la solución de prueba} / \text{absorbancia neta de la solución estándar}) \times 2 \times \text{dilución}$ .

#### 242 Determinación de azúcares en el mosto

Las muestras de mosto fueron analizadas por HPLC Agilent 1100 con detección ELSD. El solvente de HPLC fue 75 % de acetonitrilo en agua, a una tasa de flujo de 1 ml/min utilizando una columna de hidratos de carbono de Alltech de 4.6 mm × 250 mm. El ELSD se ajustó a un caudal de nitrógeno de 2 ml/min (a 87 ° DO). Las mediciones fueron realizadas por duplicado.

#### 243 Color

El color fue determinado siguiendo el método Beer-10. 100 ml de volumen final de mosto es filtrado y mezclado con 5 g de tierra de diatomeas grado analítico. Se deja reposar por 5 min y se filtra nuevamente para clarificar completamente el mosto. La absorbancia se mide a 430 nm en una cubeta de espectrofotómetro (1/2 in.). El color es calculado por la fórmula:  $\text{Color} = 10 (\text{Absorbancia en } \frac{1}{2} \text{ cubeta a } 430 \text{ nm})$ .

#### 244 pH

La concentración de iones hidrógeno se mide empleando un potenciómetro a temperatura ambiente.

#### 245 Viscosidad

La viscosidad de los mostos se midió de acuerdo con el método Wort-13. Este método calcula la viscosidad absoluta o dinámica (en unidades centipoise) utilizando un viscosímetro Ostwald, con un rango de agua de 50 - 150 s y de acuerdo con el principio

de Hoesppler al medir el tiempo de rodadura de una bola a través de un capilar de vidrio inclinado lleno de 5 ml de agua o mosto a 20 °C. Cálculos: Viscosidad = Tiempo de flujo de mosto x Densidad de mosto x 1.002.

#### 24.1 Análisis de aminoácidos libres

Fue determinado el perfil y concentración de aminoácidos presentes en cada mosto mediante el método de AccQ Tag Waters, utilizando el reactivo AccQ Flour (6-aminoquinolyl-N-hydroxysuccinimidyl carbamato, AQC); seguido de cromatografía de líquidos de alto desempeño reversa (High Performance Liquid Chromatography, HPLC) para separar los derivados; finalmente, la cuantificación se realizó por medio de un detector de fluorescencia a 395 nm.

#### 24.2 Determinación de DMS (libre, total y precursor)

Para la determinación de dimetil sulfuro (DMS) en mosto se emplea la cuantificación por cromatografía de gases acoplada a Headspace (HS/GC). Para la medición de DMS libre se utilizó directamente el mosto, mientras que para DMS-p (DMS precursor) se realizó un pretratamiento térmico en rotavapor con la adición de hidróxido de sodio, esto convierte los DMS-p en DMS, luego la solución se mide con HS/GC utilizando una columna GS-Q. El DMS total es la suma del DMS libre y DMS-p.

#### 24.3 Monitoreo de las hidrolasas presentes en los mostos 100 % cebada malteada

Para el monitoreo de las hidrolasas presentes en los mostos 100 % cebada malteada fue empleada la metodología descrita en la sección 2.4.

## 1 RESULTADOS Y DISCUSIÓN

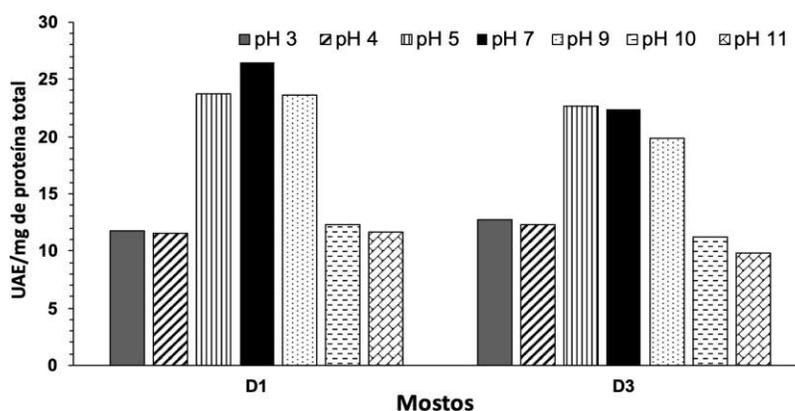
### 1.1 PARÁMETROS DE CALIDAD DE LA MALTA CERVECERA

Las características bioquímicas de la cebada malteada influirán directamente sobre la elaboración y calidad de la cerveza. En este trabajo, caracterizamos dos maltas cerveceras y relacionamos sus características con la calidad del mosto. Las cebadas malteadas provienen de diferentes ciclos de cultivo (invierno y verano), ambas son comúnmente empleadas en la industria cervecera en México. Las diferencias en sus perfiles bioquímicos influirán en los siguientes parámetros: enzimas, concentración de proteína total soluble, niveles FAN, viscosidad, color, DMS, perfil de aminoácidos y de azúcares.

## 11 EFECTO DEL PH

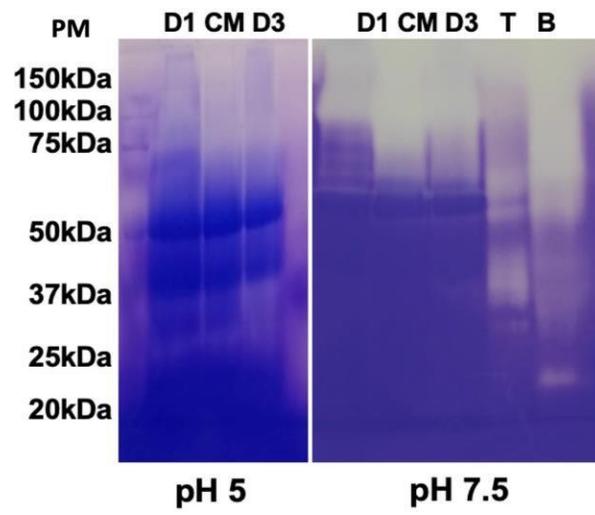
El efecto del pH fue medido con el fin de caracterizar las proteasas presentes en dos maltas, nombradas D1 y D3 , en un rango de 3.0 a 11.0. Ambas presentaron mayor actividad enzimática entre pH 5.0 a 9.0 (Fig. 1). Esto resultados demuestran que las maltas contienen proteasas ácidas (tipo cisteinil y aspártico proteasas) y alcalinas (tipo serino proteasas, carboxipeptidasas y aminopeptidasas) y coincide con lo reportado por Jones y Marinac (2002). La malta D3 presentó un mayor contenido de enzimas ácidas, mientras que la malta D1 contiene mayor contenido de enzimas alcalinas.

Figura 1. Efecto del pH en la actividad de proteasa en muestras de malta.



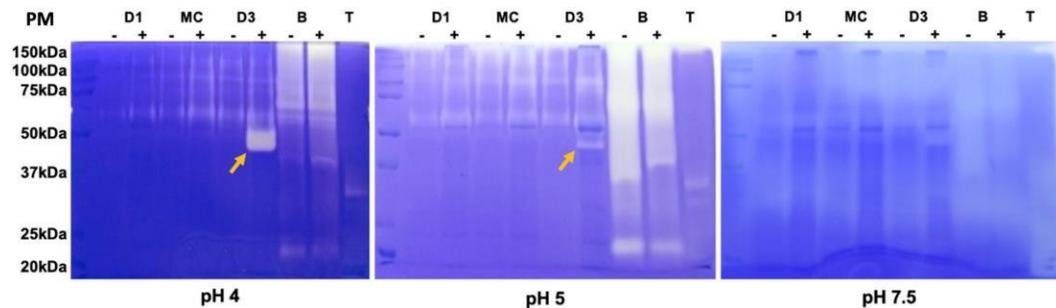
## 12 ACTIVIDAD PROTEOLÍTICA

La cebada contiene reservas de nutrientes, almidón y proteínas, que son necesarios para su desarrollo. Durante el malteado son activadas las enzimas que convertirán estas reservas en nutrientes asimilables por la levadura, logrando un buen desempeño durante la fermentación (Schmitt y Marinac, 2008). Cada malta varía de acuerdo con su variedad y condición de cultivo, por eso es necesario entender los sistemas hidrolíticos presentes en cada una. Con este estudio, ayudaremos a predecir el desempeño de la fermentación y tomar acciones preventivas, ya sea combinando lotes de malta o mediante la suplementación de enzimas durante el proceso. Se evaluó la actividad tipo proteasa utilizando zimogramas con gelatina como sustrato, las dos maltas mostraron un perfil diferente al ser incubadas a diferentes valores de pH. La malta D1 a pH 7.5 muestra 6 bandas con actividad proteolítica entre 50 y 150 kDa, mientras que la malta D3 mostró 4 bandas con actividad proteolítica con diferentes pesos moleculares 100 kDa, 60 kDa, 45 kDa y 30 kDa. Grudkowska et al. (2013) reportan una metaloproteinasa (150 kDa) y una proteínasa aspártica (110 kDa) similares. Las metaloproteinasas son relacionadas con la



una cerveza con características deseables. Se puede observar que a pH 7.5 la actividad es inhibida en todas las maltas y controles (Fig. 3). Esto indica que las maltas poseen actividad de tipo cisteinil proteasas las cuales son inhibidas a este valor de pH incluso con el uso de inductores.

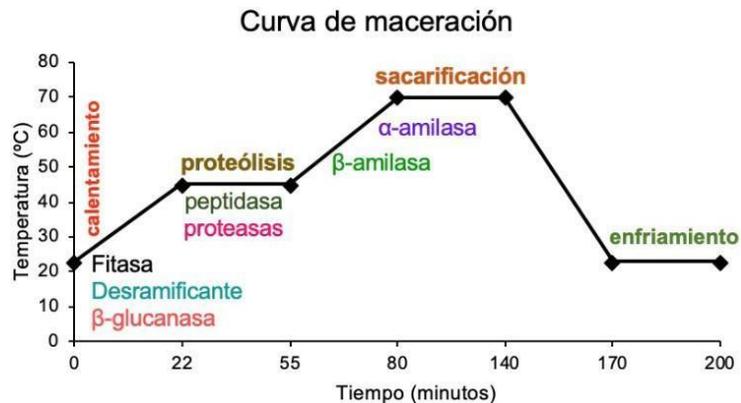
Figura 3. Inducción de la actividad proteolítica en muestras de malta **D1**, Malta D1. **MC**, Malta comercial. **D3**, Malta D3. **B**, Bromelina de tallo de piña 30µg/µl. **T**, Tripsina de páncreas bovina 20 µg /µl. - sin inductor. + con inductor.



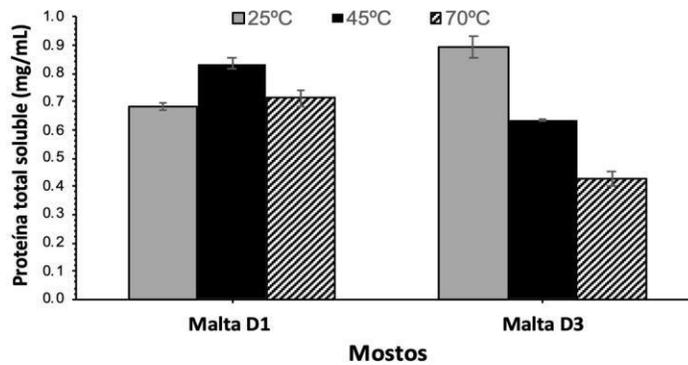
## 1.1 PREPARACIÓN DEL MOSTO

Diferentes grupos de enzimas hidrolíticas participan durante el proceso de maceración y son activadas mediante los cambios de temperatura que ocurren en las diferentes etapas del proceso (Fig. 4).

Figura 4. Curva de maceración de los mostos con 100 % malta.



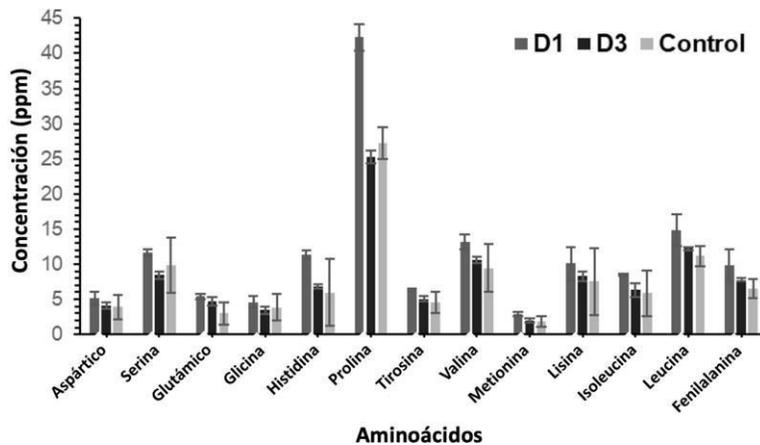
Para relacionar las hidrolasas provenientes de la malta con la calidad del mosto, se prepararon mostos 100 % de cebada malteada (Iimure et al., 2012). La concentración de proteína total soluble fue medida en cada etapa de la maceración (Fig. 5) con el fin de estudiar cambios en el contenido de proteínas presentes durante el proceso. Los resultados muestran que en el mosto D1 no hubo disminución en el contenido de



Parámetros de análisis	Mostos		
	D1	D3	Malta comercial
Proteína total soluble (mg/ L)	700	400	400
Alfa amino nitrógeno (mg/L)	350	273	269
Viscosidad (cps)	1,72	1,67	1,51
pH	5,9	6,0	5,9

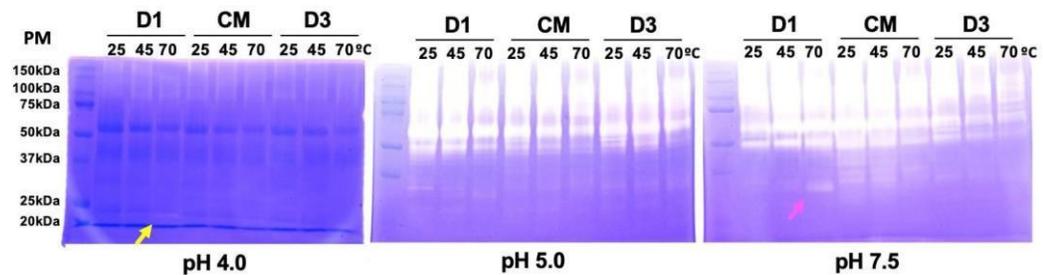
Estos resultados también concuerdan con los valores obtenidos respecto al color: mosto D1 ( $20.80 \pm 0.57$ ), mosto D3 ( $20.90 \pm 0.14$ ) y mosto control (8.4). Guzmán-Ortiz et al. (2018), reportaron resultados similares con mostos preparados con la variedad de cebada Doña Josefa y presenta un color tipo melano ( $16.26 \pm 0.19$ ) y tipo chocolate ( $27.73 \pm 0.79$ ), clasificándolas como cervezas tipo Weissbier y Double IPA, respectivamente. En cuanto al perfil de aminoácidos libres, el mosto D1 mostró mayor concentración con respecto al control y al mosto D3 (Fig. 6). Lingzhen et al. (2016), encontraron que la prolina y el ácido glutámico son importantes constituyentes en la formación de turbidez en la cerveza.

Figura 6. Perfil de aminoácidos de mostos 100 % malta.



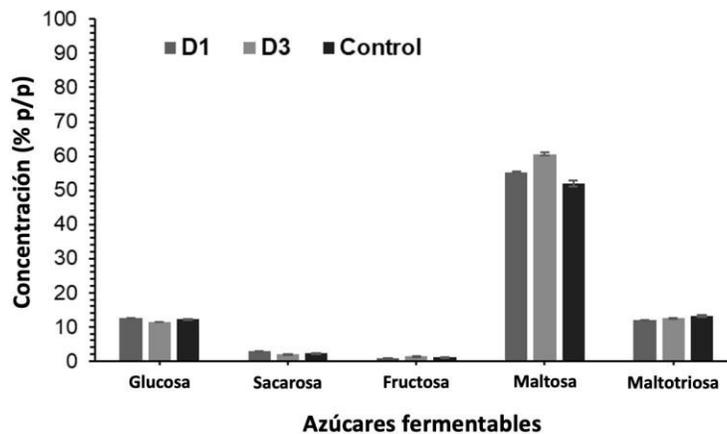
Mediante zimogramas, observamos que las principales diferencias entre la malta D1 respecto a las maltas D3 y la comercial, son bandas de actividad presentes entre 25 y 37 kDa observadas principalmente a pH 5.0 y 7.5 (Fig. 7). Una pequeña banda de actividad de ca. 22 kDa fue observada en la etapa de maceración a 70 °C (pH 4.0), la cual está presente solo en esta malta (flecha amarilla). Otra de ca. 33 kDa (pH 7.0) que es más intensa en la malta D1 (flecha rosa). Estas dos bandas podrían estar relacionadas a los resultados descritos anteriormente.

Figura 7. Actividad proteolítica presente en el proceso de maceración **D1**. Mosto preparado con malta D1. **CM**. Mosto preparado con malta comercial. **D3**. Mosto preparado con malta D3. **25, 45 y 70** representa la temperatura de cada etapa a diferentes valores de pH.



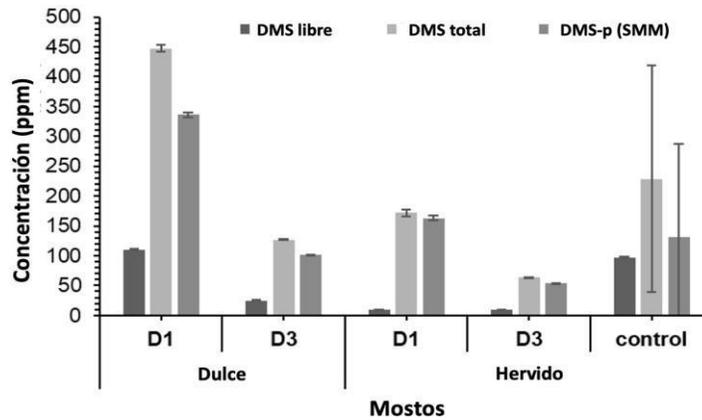
El perfil de azúcares de los dos mostos evaluados está dentro de los parámetros de calidad establecidos por la European Brewery Convention, 1998; donde casi el 60 % es maltosa, seguido de maltotriosa 15-17 %, glucosa 12-15 % y 1.3-1.9 % de fructosa (Fig. 8). Estos resultados indican que las condiciones de cultivo no afectan de manera significativa el perfil de azúcares fermentables, ya que los mostos preparados con esta variedad, producen el perfil óptimo para una buena fermentación y poseen niveles similares al mosto preparado con malta comercial. Sin embargo, el mosto D3 presentó niveles ligeramente más altos de fructosa, maltosa y maltotriosa.

Figura 8. Determinación de azúcares fermentables en los mostos.



Finalmente, evaluamos el DMS libre, total y el s-metilmetionina (SMM) precursor del DMS. El DMS es un importante contribuyente al sabor de la cerveza tipo lager, es beneficioso para el sabor y aroma en concentraciones moderadas entre 30 y 100 ppb pero a concentraciones superiores a su umbral aporta notas desagradables al aroma de la cerveza tal como aroma a pan de elote (Ferreira y Guido, 2018). Encontramos que el mosto D3 dulce cumple con los parámetros de calidad para este compuesto, detectándose niveles bajos (127 ppb mosto sin ebullición y 63 ppb con ebullición) y cae dentro del rango que no son percibidos. Por el contrario, el mosto D1 presentó niveles moderados (172 ppb) de DMS solo en los mostos después de la etapa de ebullición. Los mostos hervidos preparados con malta D1 y D3 presentaron niveles menores de DMS en comparación con el mosto de la malta comercial (229ppb). Por lo tanto, la malta D3 puede ser utilizada para preparar mostos dulces reduciendo el tiempo de maceración, mientras que, utilizando la malta D1 requerirá de la etapa de ebullición después de la maceración para disminuir los niveles de DMS (Fig. 10).

Figura 9. Determinación de Dimetil sulfuro en mostos 100 % malta.



No se había demostrado como impacta la aplicación de azufre y nitrógeno en los procesos de fertilización de la cebada, hasta que el grupo de Zhao et al. (2006) comprobaron que al aplicar concentraciones altas de azufre y nitrógeno puede afectar la concentración de DMS precursor en la cebada malteada, la calidad del malteo, el tamaño del grano y a su vez el sabor de la cerveza. Ellos recomiendan una adecuada concentración para que no se vea afectados los parámetros de calidad de la cebada. Estos resultados se relacionan con lo observado con el mosto D1, donde nuevamente las condiciones de cultivo impactan en las características de la cebada malteada, lo que significa que la cebada D1 en su condición de cultivo está más expuesta a suelos ricos en azufre y nitrógeno, de ahí sus niveles altos de DMS precursor y FAN.

## 1 CONCLUSIONES

La producción de enzimas hidrolíticas depende principalmente de la variedad de cebada, condiciones de cultivo y condiciones ambientales. La actividad enzimática es activada durante el proceso de malteado de la cebada y maceración del mosto, y son pieza clave para la producción de cerveza. De estas actividades dependerán la cantidad de aminoácidos y azúcares disponibles para la levadura durante la fermentación. La malta D1 mostro una proteasa de 150 kDa similar a una metaloproteinasa reportada relacionada con el KI y FAN. Presenta niveles superiores de FAN y una mayor cantidad del aminoácido prolina relacionado con la presencia de turbidez en la cerveza. Este resultado se relaciona también con la presencia de mayor actividad proteolítica alcalina. La malta D3 muestra mayor actividad proteolítica ácida que se relaciona con un óptimo nivel de FAN y perfil de aminoácidos. Presenta una banda tipo tiolproteasa EP-A muy prominente ca. 37-40 kDa inducida con ácido giberélico, principal enzima de digestión de las hordeínas. Utilizando







## CAPITULO 5

---

# ANÁLISIS DE LA EXPRESIÓN GENÉTICA: EL TRANSCRIPTOMA

César I. Hernández-Vásquez, Ramiro Elizondo-González, Jessica G. Herrera-Gamboa, Jorge Hugo García-García, Claudia B. López-Alvarado, Adrián García-Martínez y Benito Pereyra-Alfárez

### ~~1. Descripción y diseño experimental~~

---

Como cualquier enfoque científico, el diseño experimental juega importantes roles en la obtención de resultados reproducibles, confiables y con la menor cantidad posible de variación técnica; error sistemático. En un experimento de alto rendimiento que implique tecnologías de secuenciación de nueva generación, como es el caso de secuenciación de ARN (RNA-seq, por sus siglas en Inglés), la salida de las observaciones representa millones de mediciones en forma de fragmentos de secuencias (Reuter *et al.*, 2015) por lo que un diseño experimental apropiado permite tomar en cuenta la variación biológica del material experimental así como la variación explicada por el error sistemático. El diseño experimental apropiado se traduce en inferencias biológicas cuantitativas más cercanas a la realidad.

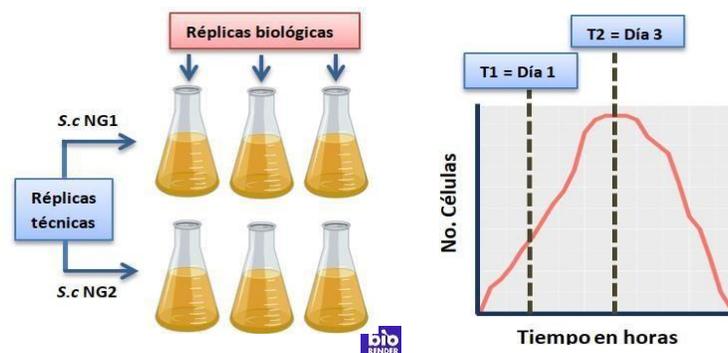
En esencia, un diseño experimental podría variar dependiendo de las necesidades de las preguntas del investigador y por supuesto del tipo de experimento a realizar. Sin embargo, como requisito crucial para un estudio transcriptómico basado en RNA-seq, es necesario la manipulación rápida y correcta de las muestras biológicas, un número apropiado de réplicas para el objeto biológico de estudio, así como estandarizar un protocolo de extracción de ARN que permita obtener muestras de ácidos nucleicos lo más homogéneas posible. Además, tener cuidado a la hora de seleccionar el tipo de librería y profundidad de secuenciación, es un aspecto importante en el diseño experimental (Conesa *et al.*, 2016).

Dentro del primer requisito, la manipulación de las muestras, es de importancia capital que la recopilación del material biológico se realice de la forma más rápida y homogénea posible. Ante cualquier cambio en las condiciones ambientales y fisiológicas, los organismos, a nivel celular, orquestan cambios en la abundancia relativa de ARN total para amortiguar las condiciones cambiantes. Esta regulación dinámica que en parte depende de procesos estocásticos y condiciones ambientales, activan mecanismos de degradación o sobre producción de ARN específicos. Aunado a esto los distintos tipos de ARN, ribosomales, mensajeros y no codificantes presentan vidas medias relativamente heterogéneas (Duffy *et al.*, 2018; Raj y van Oudenaarden, 2008; Schofield *et al.*, 2018; Schwalb *et al.*, 2016; Yang *et al.*, 2003), evidenciando así la importancia crítica del manejo adecuado de la muestra biológica.

En el segundo punto, si el costo o dificultad de la recopilación de muestras es reducida, el protocolo puede demandar material en exceso que pueda apoyar, en caso de fallos, etapas posteriores. Si habitáramos un mundo donde la vida fuera perfecta, viviríamos en un mundo homogéneo y sin variación. Por tanto, si la variación biológica y técnica no existiera, los científicos serían tremendamente felices sabiendo que sus investigaciones bajarían

en costo al solo requerir una sola unidad experimental en cada combinación de factores. Sin embargo, la variabilidad de nuestro sistema natural nos exhorta la necesidad de la replicación la cual nos brinda un panorama poblacional más preciso y nos permite vislumbrar la variabilidad que existe en nuestro experimento al brindarnos una aproximación precisa sobre la variabilidad y una disminución considerable del error estándar de la media muestral (Mathe y Davis, 2016).

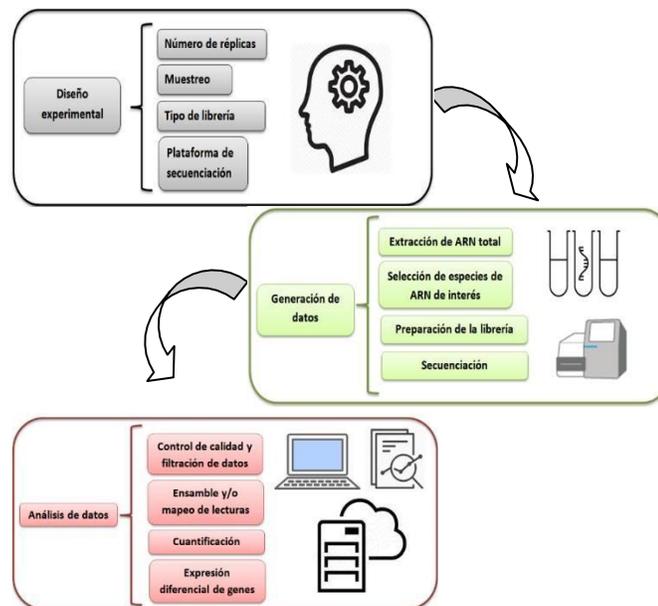
Para ejemplificar la replicación en un experimento RNA-seq expondremos los pasos y protocolos empleados en nuestra investigación sobre el transcriptoma de *Saccharomyces cerevisiae* (S.c) en distintos tiempos de su curva de crecimiento como nuestras condiciones fisiológicas. La hipótesis de nuestro experimento implica que en distintos tiempos del ciclo de crecimiento de S.c el perfil transcripcional presentará discrepancias importantes entre los tiempos de crecimiento. Para cumplir con el cometido, nuestro diseño experimental implicó la realización de crecimiento de S.c en tubos con medio de cultivo. En nuestro ejemplo, las réplicas técnicas representan un mismo inóculo, refiriéndonos a la misma generación celular, de levadura en distintos tubos con medio de cultivo exponiendo así múltiples curvas de crecimiento para un solo inóculo de levadura.



**Figura 1.** Diseño experimental. S.c NG1 y NG2 indica la cepa con número de generación 1 y 2 respectivamente. En cada experimento se recolectaron muestras en los tiempos 1 y 2 (Día 1 y Día 3 de la fermentación) de la curva de crecimiento. En este arreglo, las columnas corresponden a las réplicas

La réplica biológica se realizó con la misma cepa de levadura, pero con distinto número de generación. En resumen, nuestro diseño expone que las réplicas técnicas son el conjunto de matraces con cepas del mismo número de generación y las réplicas biológicas son el conjunto matraces con distinto número de generación.

Posteriormente, el protocolo de tratamiento de las muestras constó de centrifugación de las muestras de levadura y reconstitución en amortiguador PBS, de forma rápida y cuidadosa para cumplir con un apropiado manejo de las muestras. Luego se efectuó una congelación rápida en nitrógeno líquido para mantener el estado fisiológico de las muestras lo más intacto posible hasta su uso en los ensayos de extracción de ARN. Con este arreglo experimental podemos obtener información biológica importante sobre nuestro sistema biológico además de disminuir nuestro error sistemático a tal grado que podamos observar homogeneidad a la hora de analizar los datos transcriptómicos.



**Figura 2.** Panorama general del flujo de trabajo típico en un experimento de RNA-seq

Una vez diseñado el experimento, procede su ejecución y análisis posteriores de los datos de secuenciación. La Figura 2 muestra el panorama general de la generación de muestras para su secuenciación masiva y sus posteriores análisis. En las siguientes secciones se explica con detalle los pasos fundamentales en un experimento de RNA-seq.

## 12 Extracción y calidad de ARN

Antes de que el ARN pueda ser secuenciado, debe ser extraído y purificado del ambiente intracelular que consiste principalmente en proteínas, ADN y carbohidratos. El transcriptoma de una célula eucariota alberga una amplia complejidad que incluye ARN codificante y no codificante por lo que este aspecto debe de tomarse en cuenta a la hora de los requerimientos de la investigación. En la literatura, existe una vasta variedad de kits utilizados para la extracción de ARN. Sin embargo, métodos que emplean el reactivo TriZol y fenol-cloroformo son ampliamente utilizados en algunas investigaciones que emplean secuenciación de alto rendimiento (Shekhawat *et al.*, 2019; Sweet *et al.*, 2018). Sin importar el protocolo empleado, los fundamentos bioquímicos de una extracción y purificación de ARN permanecen constantes en cada enfoque metodológico existente y esto se debe a que dichos fundamentos están basados en la estructura química del ARN (Fig 3A).

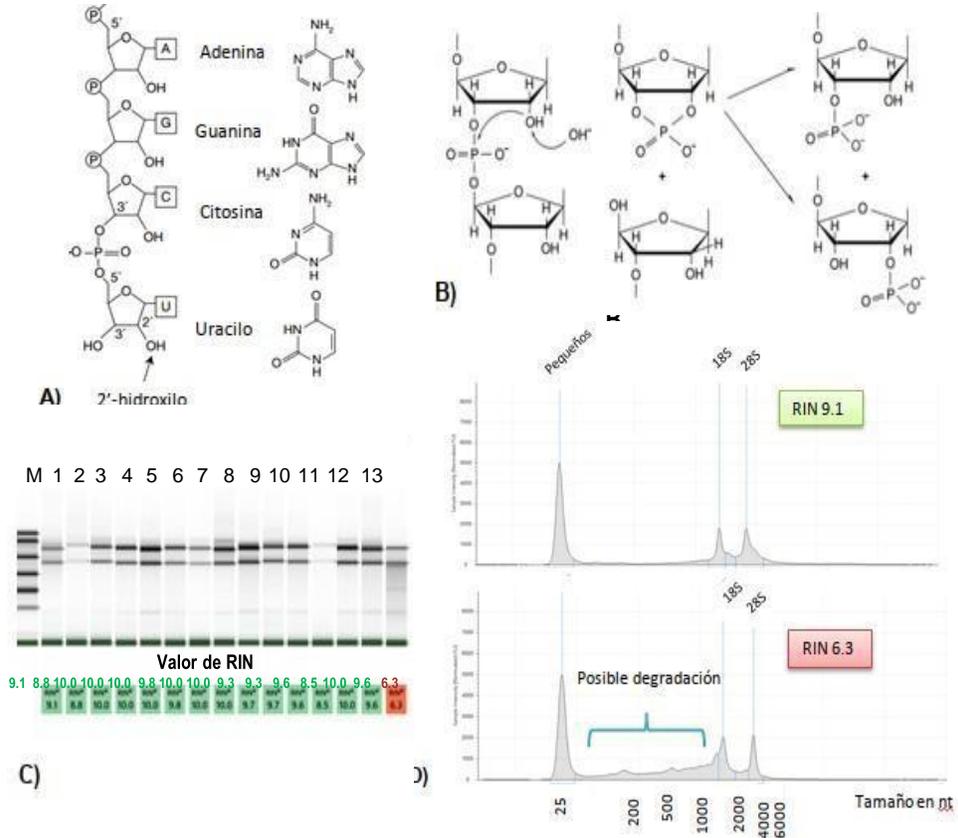
Es importante mencionar que la estructura química de ARN lo hace más susceptible a degradación que el ADN. El grupo 2'-hidroxilo unido a la ribosa puede actuar como un nucleófilo intramolecular atacando el enlace fosfodiéster adyacente, creando un 2',3' fosfato cíclico escindiendo el esqueleto del RNA (Li y Breaker, 1999) (Fig 3B). Aunado a la susceptibilidad intrínseca del ARN, existen tres causas principales de degradación de ARN; pH alcalino, iones metálicos divalente y ribonucleasas, que son totalmente prevenibles. Como se mencionó anteriormente, existen muchos

enfoques diferentes para purificar RNA que funciona de forma correlativa si se ejecutan apropiadamente. En cualquier protocolo de purificación, el primer paso es la solubilización del material del cual se extraerá el ARN; células, normalmente.

En dicho proceso de solubilización es importante inactivar las nucleasas existentes mediante el uso de compuestos caotrópicos como isotiocianato de guanidina o detergentes desnaturizantes como SDS. Posterior a la solubilización, es necesario emplear compuestos orgánicos apolares como el cloroformo para inducir la separación de la solución en dos fases, fase orgánica y acuosa. Normalmente, la fase orgánica contiene la mayor parte de las proteínas y otros componentes solubles en cloroformo como lípidos. La fase acuosa contiene moléculas de ARN hidrosolubles. Dicha fase acuosa se recupera cuidadosamente en un recipiente nuevo y libre de nucleasas. Después de la purificación, el ARN debe ser recuperado y la forma estándar de hacerlo es empleando la precipitación por etanol al 95% o isopropanol. Una vez centrifugado, la pastilla de ARN debe ser identificada para la remoción de exceso de sales con etanol al 75%. Lavada la pastilla, debe de recuperarse nuevamente por centrifugación y re suspendida en agua libre de nucleasas (RNAsa) (El-Ashram *et al.*, 2016; Nilsen, 2013; Rio *et al.*, 2010).

Como se mencionó anteriormente, el ARN es más susceptible a degradación por lo que debe cumplir con criterios de calidad y pureza que permita buenos resultados en la preparación de las librerías y propiamente en la secuenciación. Tradicionalmente, la calidad de ARN se comprueba vía electroforesis en gel por la inspección visual de las bandas principales de ARN ribosomal (Sambrook y Russell, 2001). Dado que la interpretación humana es subjetiva y sesgada, se han desarrollado algoritmos informáticos para el cálculo de un valor numérico que indique la integridad del ARN (RIN). Esta tecnología desarrollada por la empresa Agilent Technologies, permite dichos cálculos mediante la utilización de un

dispositivo, Agilent Bioanalyzer, que brinda resultados cuantitativos sobre la calidad de nuestro ARN (RIN) (Masotti y Preckel, 2006). El bioanalyzer es un dispositivo bioanalítico que usa tecnología de micro fluidos para la separación electroforética de moléculas biológicas, ARN y ADN, de forma automática y altamente reproducible (Mueller *et al.*, 2000).



**Figura 3. A)** Estructura química del ARN. **B)** Hidrólisis alcalina responsable de la degradación de ARN inducida por pH. El análisis de calidad puede ser observado mediante electroferograma en gel simulado **C)** o electroferograma capilar **D).**

Cantidades exiguas de ARN son separadas en canales, dentro chips microfabricados, de acuerdo a su peso molecular y subsecuentemente detectados por medio de fluorescencia. El

resultado es visualizado en un electroferograma (Figs. 3C y 3D) donde la cantidad de fluorescencia es proporcional a la cantidad de ARN de una talla específica. Debido a que la degradación de ARN es un proceso continuo que implica nulas categorías de integridad, algoritmos basados en aprendizaje automático (Machine Learning) pueden ser aplicados a los datos de observaciones experimentales para deducir un modelo implícito de integridad de ARN, el cual provee un método de obtención de datos de integridad de ARN independientes de las observaciones subjetivas del usuario (Schroeder *et al.*, 2006).

Los valores del RIN van de 10 a 1 siendo los valores de 10 los que representa las más alta calidad y pobre o nula degradación de la muestra. Valores RIN por debajo de 6 representa un panorama de baja calidad que puede afectar experimentos posteriores como la preparación de las librerías y sesgos en la secuenciación, significando esto conclusiones biológicas erróneas (Kukurba y Montgomery, 2015). Para brindar un escenario pragmático de los conceptos abordados, exponemos los resultados obtenidos de extracciones de ARN realizado en nuestro proyecto de investigación. El ARN purificado fue analizado en un Bioanalyzer 2100 (Agilent Technologies, CA). En la Figura 3C se ejemplifica un electroferograma en gel simulado con los datos digitales obtenidos en el bioanalyzer y en la Figura 3D se observan dos electroferogramas capilares pertenecientes a distintas muestras. Es importante notar que una de las muestras presenta indicios de degradación por lo que el algoritmo le asigna un RIN de 6.3, mientras que otra muestra presenta un buen criterio de calidad exponiendo un valor RIN de 9.1. En este caso, la muestra correspondiente a RIN = 6.3 se decidió secuenciar a pesar de su valor de calidad reducido. Algunas investigaciones apoyan la idea de que; muestras de ARN relativamente degradadas pueden ser utilizadas en estudios de RNA-seq, sin perder información biológica importante (Reiman *et al.*, 2019).

## **1.2 Tipo de librería, profundidad de secuenciación y fundamentos de secuenciación por síntesis**

Una vez que el ARN ha sido aislado y verificado de cumplir los criterios de calidad ( $RIN \geq 8$ ), el siguiente paso en la secuenciación del transcriptoma es la creación de una librería de RNA-seq. En términos que circulan dentro de proyectos que impliquen secuenciación de nueva generación, el concepto librería hace referencia a una colección, preferiblemente aleatoria, de fragmentos de DNA que están listos para ser secuenciados mediante un protocolo o plataforma específica. La construcción de librerías de secuenciación involucra: i) aislamiento de moléculas de ARN deseada, ii) fragmentación de moléculas a un tamaño deseado, iii) transcripción reversa de fragmentos de ARN a ADNc y iv) la ligación de adaptadores de secuenciación (Head *et al.*, 2014).

En el primer paso, el ARN de interés de ser separado o enriquecido respecto a otras moléculas de ARN no deseadas. Como ejemplo, algunos ARN abundantes como el ribosomal que comprende cerca del 80%, debe ser removido ya que puede introducir sesgos en la preparación de la librería. Si los ARN ribosomales no son removidos antes de la construcción de la librería, entonces consumirá la mayor parte del conjunto de lecturas de secuenciación, reduciendo la cobertura general y limitando la detección de algunos ARN mensajero (ARNm) de bajo nivel de expresión (Kukurba y Montgomery, 2015). Para la mayoría de los organismos eucariotas el enriquecimiento de la porción mensajera del ARN puede ser realizada mediante el uso de selección por poli(A). La selección por poli(A) requiere una proporción relativamente alta de ARNm sin degradación. En algunos casos, como las biopsias de tejidos, no se puede obtener muestras de ARN con altos valores de calidad y por tanto se requiere de la eliminación directa del ARNr. En el caso de los organismos procariotas, en los cuales los ARNm no están poliadenilados, la

única alternativa es la eliminación directa de ARNr (Conesa *et al.*, 2016). En la tabla 1 se muestran algunos diseños en el tipo de librerías, los usos y descripciones de los objetivos experimentales e investigación que abordan.

**Tabla 1.** Protocolos y diseños de librerías para secuenciación de ARN.  
Tomado y modificado de Kukurba y Montgomery (2015)

<b>Diseño de librería</b>	<b>Uso</b>	<b>Descripción</b>
Selección por poly(A)	Secuenciación de ARNm	Enriquecimiento específico de ARNm
Eliminación directa de ARNr	Secuenciación de ARNm, pre-ARNm y ARNnc	Elimina ARNr y enriquece ARNm, pre-ARNm y ARNnc
Selección por tamaño	Secuenciación de miARN	Selecciona especies de ARN usando fragmentación de tamaño por electroforesis en gel
Nucleasa duplex-específica	Reduce transcritos altamente abundantes	Elimina transcritos altamente abundantes, incluyendo ARNr y otros genes altamente expresados
Cadena específica	Ensamble <i>de novo</i> del transcriptoma	Preserva información de la cadena del transcrito
Multiplex	Secuenciación de múltiples muestras juntas	Marcaje genético que permite la secuenciación de múltiples muestras juntas

En cuanto al segundo paso general, la fragmentación de ácidos nucleicos no siempre se emplea en proyectos de secuenciación, esto depende del enfoque y las preguntas que se quieran responder durante la investigación. Sin embargo, la preparación de librerías mediante fragmentación, shot gun, son ampliamente utilizadas y la plataforma más utilizada es Illumina. De hecho, de acuerdo a la base de datos de SRA (Sequence Read Archive), más del 80% de las secuencias almacenadas en dicha base de datos provienen de la plataforma Illumina (Kodama *et al.*, 2012; Leinonen *et al.*, 2011). Por tanto, tomaremos como tema principal las librerías de secuencias cortas generadas típicamente por “shot gun” vía

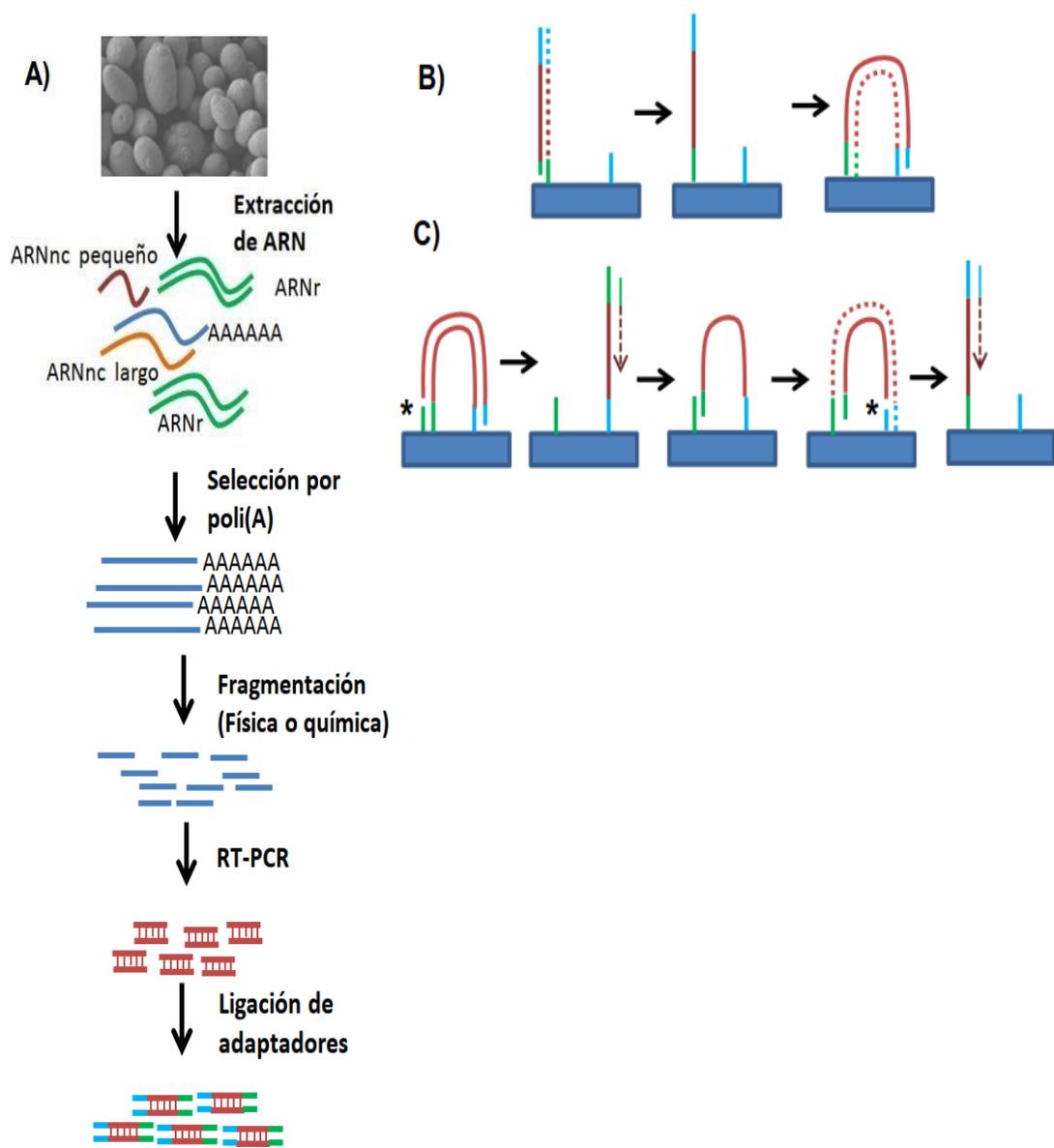
plataforma Illumina. La talla de fragmentos específicos en la librería final es un parámetro de suma importancia, sin importar las plataformas de secuenciación que empleen fragmentación, en proyectos de secuenciación de nueva generación (NGS, por sus siglas en inglés). Existen tres enfoques principales para cumplir con el paso de la fragmentación del ADN o ARN en el caso de RNA-seq: físico, enzimático y químico. La fragmentación de ácidos nucleicos se hace típicamente mediante métodos físicos, ej. Sonicación, o por métodos enzimáticos, eje. Cócteles de endonucleasas no específicas y reacciones de “tagmentación” con transposones (Head *et al.*, 2014).

Seguido de la fragmentación, en el caso de ARN-seq, las muestras deben ser convertidas a ADNc para poder entrar en los protocolos de secuenciación por síntesis y esto se hace mediante retrotranscripción (RT-PCR). Otra parte importante son los adaptadores de secuenciación que funcionan como indicadores de muestras en procesos paralizados, además de funcionar como medio de anclaje a la fase sólida en la celda de secuenciación, por apareamiento de bases (Illumina, 2015.). En la Figura 4A se muestran los pasos importantes en la preparación de las librerías para RNA-seq.

Una vez listas las librerías, el procedimiento de secuenciación puede llevarse a cabo. Dado que hemos estado hablando de secuenciación por fragmentación de moléculas de ácidos nucleicos (shotgun) y la plataforma más utilizada al momento es Illumina, abordaremos los fundamentos básicos de dicha plataforma. Sin dejar de mencionar que existen otros tipos de RNA-seq donde el objetivo es producir lecturas más largas para cumplir con objetivos de optimización de ensamblaje y descubrimientos de isoformas (Stark *et al.*, 2019).

En primera instancia, las librerías deben de amplificarse de forma clonal para generar conjuntos enriquecidos de cada molécula, esta amplificación clonal es importante ya que los detectores de fluorescencia no están diseñados para detectar señales de una





## **1. Preprocesamiento: control de calidad y limpieza de los datos transcriptómicos**

En los ejemplos posteriores se asume que el lector presenta conocimiento básico de sistemas operativos basados en GNU/LINUX como Debian y Ubuntu, por mencionar algunos. En esta parte expondremos algunos programas de personal preferencia. Sin embargo, existen múltiples opciones que satisfacen distintos objetivos específicos de investigación por lo que exhortamos al lector explorar las alternativas y verificar la literatura de tal manera que se tome en cuenta herramientas actualizadas o completamente nuevas. En cuanto a la infraestructura computacional se asume que el lector tiene acceso a una computadora con un mínimo de 10 núcleos y una memoria RAM de 32 GB. Si no es el caso, recomendamos la utilización de servicios computacionales en la nube brindados por Amazon Web Service (<https://aws.amazon.com/es/>) y Google Cloud (<https://cloud.google.com/>) que solucionan problemas relacionados con la falta de hardware computacional.

### **2.1 Control de calidad: revisión de estadísticos de calidad de los datos crudo mediante FastQC**

Una vez que los datos RNA-seq en formato FASTAQ son recibidos de la plataforma de secuenciación, es de suma importancia verificar la calidad de dichos datos. De los aspectos más importantes, se debe revisar la calidad general de las secuencias, el porcentaje de guanina y citocna (%GC) y el porcentaje de secuencias sobrerrepresentadas. El programa FastQC se ha convertido en la herramienta estándar para realizar este trabajo. El siguiente comando, realizado en la terminal de UBUNTU 18.04 LTSrealizar el reporte de calidad de todos los archivos fastq.

```
fastqc -o qc/raw -t 10 *.fq.gz
```

En el comando anterior asumimos que los directorios *qc* y *raw* ya han sido creados además de estar posicionados en el directorio de trabajo, lugar en el sistema de archivos donde se encuentran los archivos FASTQ correspondientes a las lecturas de secuenciación. En nuestro caso de estudio tenemos 24 archivos FASTQ con un promedio de 22M de lecturas. La parte *\*.fq.gz* procesa todos los archivos que terminen con *fq.gz*, esto es, todos nuestros archivos FASTQ que se encuentran en el directorio de trabajo. La salida del comando anterior es un archivo Zip que contiene un reporte en formato HTML el cual se subdivide en diferentes secciones que describen distintos aspectos de calidad que hemos analizado en las lecturas crudas. Las secciones son las siguientes:

#### **a) Estadísticas básicas**

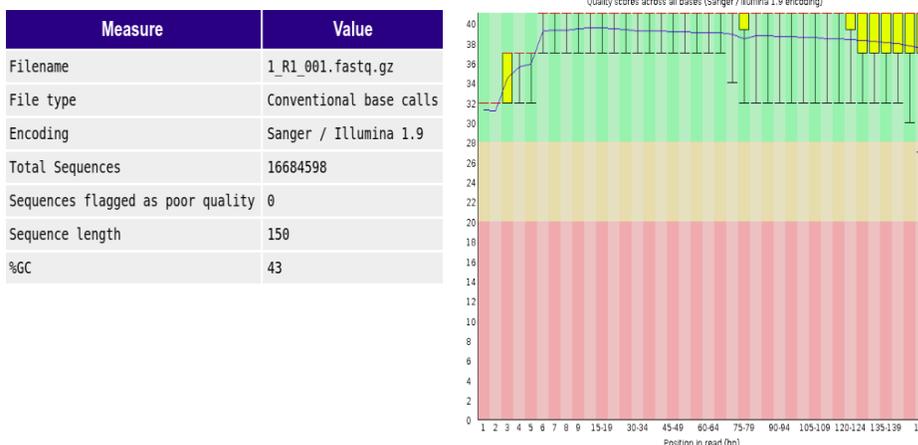
Este módulo genera un conjunto de mediciones simples en los archivos analizados. La descripción incluye:

- El nombre del archivo: El nombre original del archivo que está siendo analizados
- Tipo de archivo: indica si el archivo contiene el llamado de bases o espacios de color que se convierten en llamadas de bases
- Codificación: indica el código ASCII de la calidad de las bases en esos archivos
- Secuencias totales: Número total de secuencias para cada archivo analizado
- Secuencias de baja calidad: Numero de secuencias que contienen bases de baja calidad
- Longitud de secuencias: El tamaño de las secuencias analizadas

- Porcentaje GC: Proporción de guanina y citosina calculada para todas las secuencias del archivo analizados

### b) Calidad de las secuencias por base

En este módulo se representa gráficamente la calidad de cada base en todas las lecturas de secuenciación. El eje y representa los valores Phred de calidad. Cuanto más alto el valor mejor la calidad de la base. La escala de calidad de Phred representa la probabilidad  $p$  de que una base es incorrecta. Por ejemplo, un valor Phred de 10 indica un error por cada 10 bases o 90% de exactitud. En el eje x se representa la posición de las bases en la lectura. El fondo de la gráfica se divide en calidad muy buena de bases (verde), bases de calidad razonable (naranja) y bases de baja calidad (rojo) (Fig. 5).



**Figura 5.** Estadísticas básicas de nuestros datos RNA-seq crudos a la izquierda. Gráfica de calidad de cada base dentro de todas las secuencias de la librería.

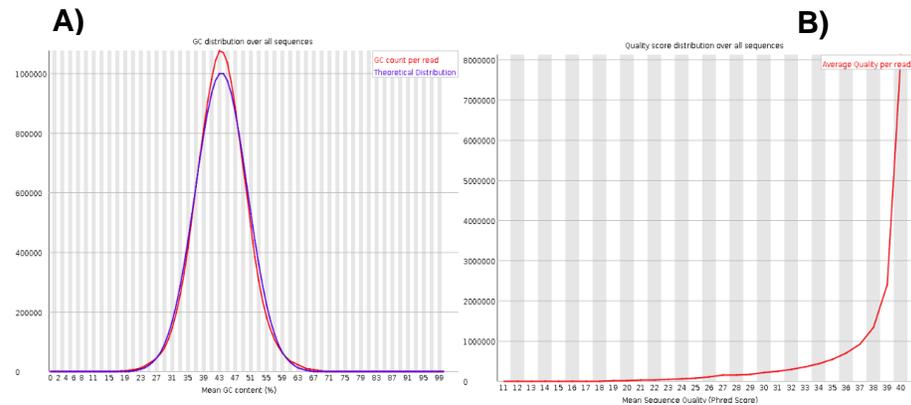
### c) Contenido GC por secuencia

Este apartado mide el contenido GC a través de la longitud completa de cada secuencia en el archivo analizado y la compara con una distribución normal teórica. En una librería que consta de segmentos aleatorios de secuencias se espera que la distribución de contenido GC corresponda al organismo correspondiente. Sin

embargo, al no tener el organismo en cuestión el contenido GC modal se calcula a partir de los datos observados y se usa para construir una distribución teórica como referencia. Se arrojan avisos cuando la distribución observada contiene sesgos en comparación con la distribución teórica.

### a) Valores de calidad por secuencia

En este módulo se grafica los valores de calidad en el eje x y el número de secuencias que contienen dicho valor de calidad en el eje y. Esto permite observar si un conjunto de lecturas contiene, universalmente, bajos valores de calidad. Si los datos contienen valores de calidad buenos, el histograma se sesga hacia la izquierda (Fig. 6).



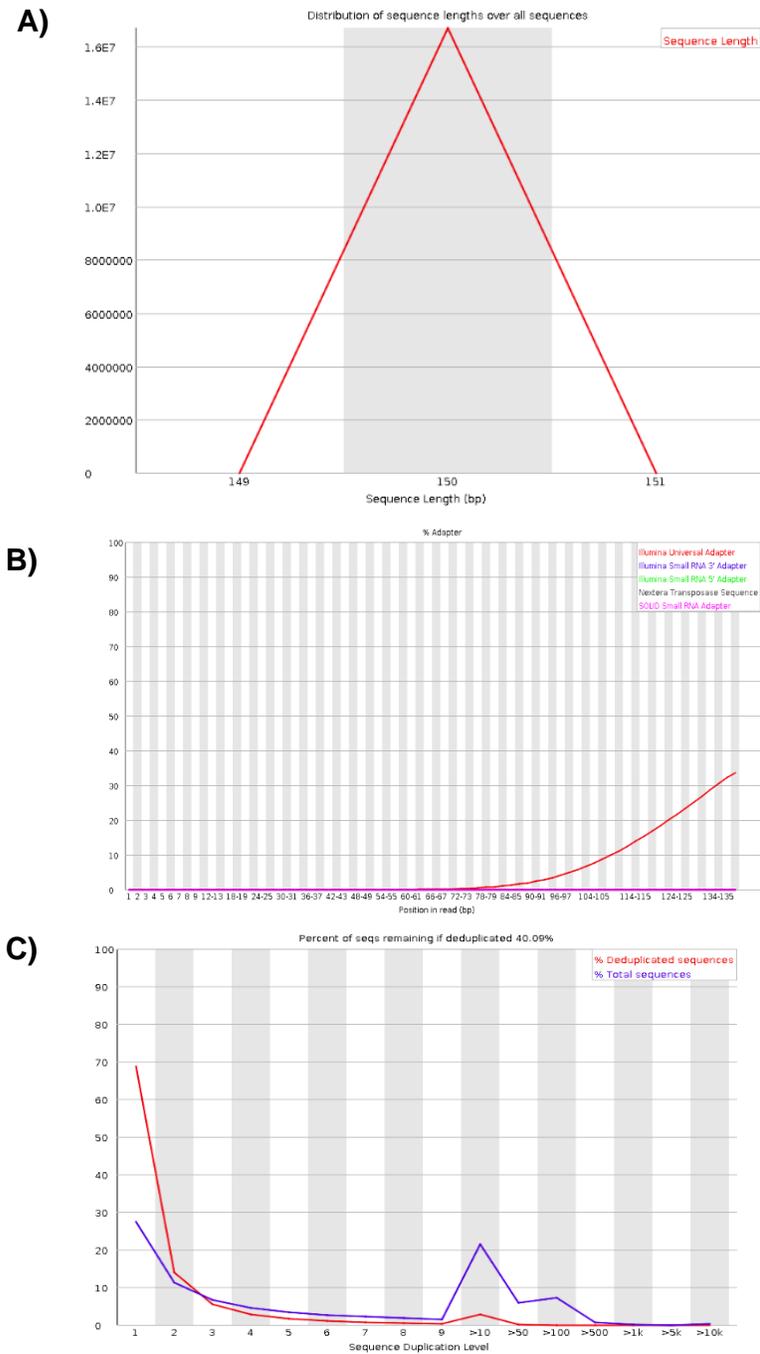
**Figura 6. A)** Distribución teórica (Azul) y observada (roja) del contenido GC del archivo FASTQ analizado. **B)** Valores de calidad por secuencia. Se puede observar que los valores de calidad no contienen subgrupos en bajos niveles ya que la mayor proporción de las lecturas contienen valores entre 25-40 valor Phred, la gráfica muestra asimetría negativa.

### b) Distribución de la longitud de secuencias

Algunas plataformas de secuenciación generan fragmentos de secuencias con una longitud uniforme, pero otros pueden generar datos con longitud de secuencia variable. Esto se puede observar mediante la comparación de la longitud de las secuencias contra el conjunto de lecturas que tienen dicha longitud. Una distribución uniforme debe de contener un rango muy estrecho de longitud donde la mayor proporción de lecturas de situarse.

---

---



**Figura 7. A)** Distribución del tamaño de la librería. Nótese el rango estrecho de tamaño de 149-150 pb. **B)** Grafico de contenido de adaptadores en las librerías analizadas. Se observa una proporción importante en el extremo de las lecturas. Este adaptador corresponde a oligonucleótidos universales empleados por la plataforma Illumina. **C)** Grafico que expone los niveles de duplicación de secuencias. En datos RNA-seq es común observar niveles de duplicación >10 que corresponden a la alta cobertura de genes con altos niveles de expresión.



```

#Creamos un directorio donde pondremos la base de datos
DBDIR=-/db/sortmerna
mkdir -p ${DBDIR}
#Copiar los archivos FASTA en el directorio de la base de
datos
cp rRNA_databases/*.fasta ${DBDIR}
#Indexamos todas las bases de datos
for f in ${DBDIR}/*.fasta; do    indexdb_rna --ref
${f},${f}/.fasta/ & done
#Creamos una variable de bash para todas las bases de datos
SORTMERNADB=$(for f in ${DBDIR}/*.fasta; do echo -n
"${f},${f}/.fasta/"; done)
#Eliminamos el colon final incorrecto de la variable
SORTMERNADB=${SORTMERNADB/%/:}
#Agregamos la variable al archivo ~/.profile para que la
base de datos este #siempre disponible

```

Con los comandos anteriores construimos la base de datos necesaria para el filtrado de ARNr de nuestras librerías. El paso siguiente es el filtrado de nuestras librerías y recordando tenemos un conjunto de 24 archivos fasta correspondiente a tres replicas técnicas y tres replicas biológicas de nuestras fermentaciones. Son 24 archivos porque tenemos datos pareados (paired-end reads). Sin embargo, el programa SortMeRna necesita un archivo por muestra por lo que tendremos que unir los dos archivos de cada muestra y realizar el filtrado de ARNr de la forma siguiente.

```

# unimos los dos archivos de cada muestra
merge-paired-reads.sh    x_R1_001.fastq    x_R2_001.fastq
x_001_merged.fastq
# Filtramos los archivos unidos para cada muestra
sortmerna --ref $SORTMERNADB --reads x_001_merged.fastq -
paired_in \    -a 18 -log --fastx --aligned 11_001_rRNA --
other x_001_sortmerna -v

```

En los comandos anteriores reemplazamos la *x* con el número del 1 al 12 que representan la nomenclatura de nuestras librerías. Seguido del filtrado, debemos separar los archivos originalmente unidos para generar un reporte FastQC para cada uno de los dos archivos por muestra. Esto se cumple con los siguientes comandos:

```
#Separamos los archivos filtrados
unmerge-paired-reads.sh x_001_merged_sortmerna_fastq \
x_R1_sortmerna.fastq x_R2_sortmerna.fastq

#Eliminamos nuestros archivos merged y movemos los
archivos separados #a un directorio nuevo (/sortmerna/)
rm *merged*.fastq
mkdir sortmerna
mv *sortmerna* /sortmerna
mv *rRNA*.log /sortmerna
mv *rRNA*.fastq /sortmerna

#Creamos un directorio qc/sortmerna para guardar los
archivos del #reporte
mkdir qc/sortemrna

#Generamos un reporte de calidad con FastQC para todos los
archivos que #contengan en su nombre la palabra *sortmerna*
fastqc -o qc/sortmerna/ -t 10 sortmerna/*sortmerna*.fastq
```

Hasta el momento hemos eliminado las posibles contaminaciones con RNAr presentes en la base de datos que creamos. En el reporte de FastQC no vemos ningún cambio aparente en la distribución del contenido GC indicando que la contaminación que se observó fue relativamente reducida. Sin embargo, si tabulamos la proporción de secuencias de nuestras librerías que coinciden con la base de datos de ARNr obtenemos la siguiente Tabla 2. Para los análisis posteriores los archivos filtrados se utilizan como entrada para los algoritmos. Por ejemplo, en el paso siguiente,

remover las lecturas de baja calidad y los adaptadores de secuenciación presentes en nuestras librerías, los archivos de salida de SortMeRNA serán utilizados como entrada para el algoritmo de limpieza.

**Tabla 2.** Porcentaje de coincidencia de las librerías de RNA-seq con la base de datos de ARN ribosomal construida con las utilidades de software SortMeRNA.

Archivo	Base de datos de ARN ribosomal			
	(% de coincidencia)			
1_001_merge.fastq (F2 Lab 1)	0.01	0	0.69	1.65
2_001_merge.fastq (F2 Lab 2)	0.01	0	0.9	2.00
3_001_merge.fastq (F2 Lab 3)	0.01	0	1.28	2.68
4_001_merge.fastq (F5 Lab 1)	0.01	0	0.93	1.95
5_001_merge.fastq (F5 Lab 2)	0.01	0.01	2.11	4.24
6_001_merge.fastq (F5 Lab 3)	0.01	0	0.89	1.78
7_001_merge.fastq (F2 Lab 4)	0.01	0	0.87	2.02
8_001_merge.fastq (F2 Lab 5)	0.01	0.01	1.25	2.70
9_001_merge.fastq (F2 Lab 6)	0.01	0	1.28	2.74
10_001_merge.fastq (F5 Lab 4)	0.14	0.24	2.03	4.21
11_001_merge.fastq (F5 Lab 5)	0.10	0.10	1.67	3.34
12_001_merge.fastq (F5 Lab 6)	0.01	0	0.77	1.68

## 2.1 Recorte de lecturas de baja calidad y eliminación de adaptadores de secuenciación

Es común que en los secuenciadores Illumina, la calidad de las bases está relacionada con su posición en la lectura. Un ejemplo clásico es que las bases de los últimos ciclos de la secuenciación, normalmente presentan baja calidad con respecto a las posiciones anteriores, esto es evidente en el reporte de calidad que se generó anteriormente (Fig. 4). Uno de los principales enfoques frente a este problema, con el fin de aumentar la tasa de alineamiento de

las lecturas, es la eliminación de las bases de baja calidad de los extremos 3' hasta que la calidad especificada por el usuario sea alcanzada. Otro problema adicional, es la presencia de adaptadores de secuenciación en los extremos de nuestras secuencias que intervienen de forma importante en los análisis posteriores si no se eliminan. Existen una gran variedad de herramientas disponibles para la realización de estas tareas. En este libro presentamos ejemplos con la herramienta Trimmomatic, una herramienta que cumple satisfactoriamente con ambas tareas (Bolger *et al.*, 2014). Los comandos utilizados para nuestros datos paired-end son los siguientes:

```
#Eliminación de bases de baja calidad y adaptadores  
java -jar /home/cesarihv7/trimmomatic/Trimmomatic-  
0.39/trimmomatic-0.39.jar PE -threads 16 -phred33  
sortmerna/x_R1_sortmerna.fastq  
sortmerna/x_R2_sortmerna.fastq  
x_R1_sortmerna_trimmomatic.fastq  
x_R1_sortmerna_unpaired.fastq  
x_R2_sortmerna_trimmomatic.fastq  
x_R2_sortmerna_unpaired.fastq  
ILLUMINACLIP:/home/cesarihv7/trimmomatic/Trimmomatic-  
0.39/adapters/TruSeq3-PE1.fa:3:30:10 SLIDINGWINDOW:5:20  
MINLEN:50
```

Es importante notar que la ruta de ejecución de los archivos jar deben de ser adaptados el entorno general de nuestra terminal. Los archivos TrueSeq3-PE1.fa son parte de la instalación de Trimmomatic. Al igual que en los comandos anteriores solo reemplazamos la letra x con el número de nuestras librerías. En general nuestro comando elimina los adaptadores de secuenciación presentes en el archivo TrueSeq3-PE1.fa. También corta las secuencias, mediante el algoritmo SLIDINGWINDOW:

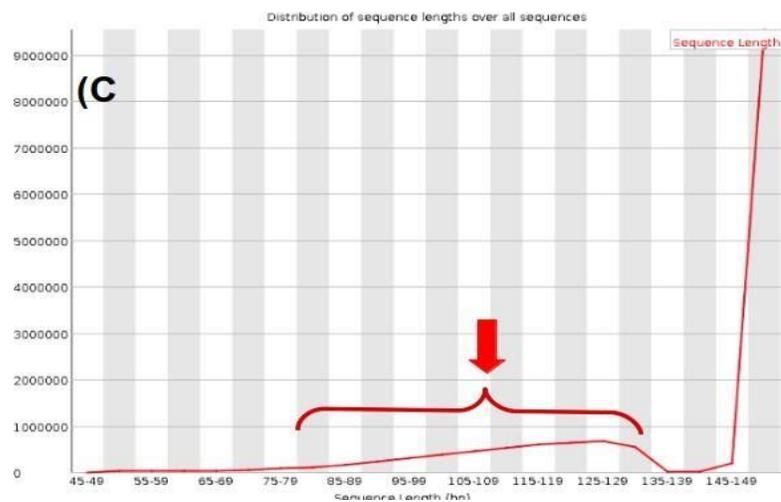
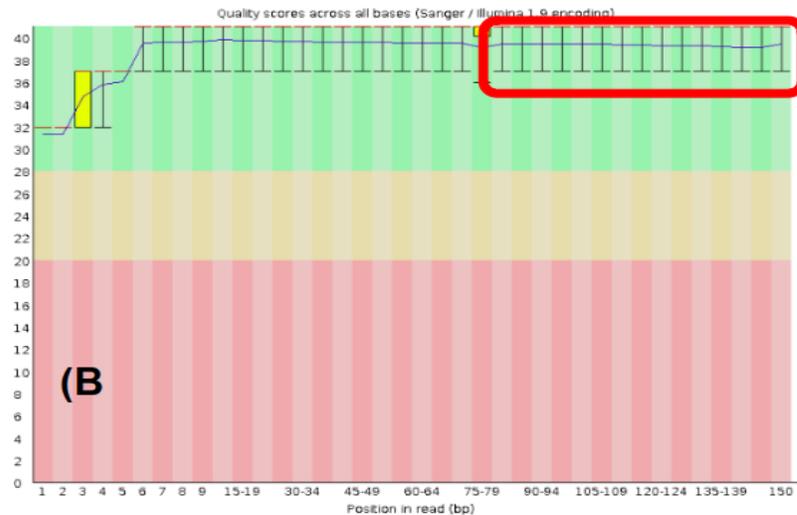
esto es que busca 5 bases y promedia su calidad, valor Phred, y si este valor no alcanza 20 entonces corta la secuencia, aceptando solo secuencias de 50 pb (MINLEN:50). Para la explicación detallada de estos parámetros y otros se exhorta al lector revisar el manual de usuario de Trimmomatic disponible en [http://www.usadellab.org/cms/uploads/supplementary/Trimmomatic/TrimmomaticManual\\_V0.32.pdf](http://www.usadellab.org/cms/uploads/supplementary/Trimmomatic/TrimmomaticManual_V0.32.pdf).

Luego de realizada esta tarea se ejecuta un último análisis de calidad, FastQC, para verificar que el pre-procesamiento de nuestros datos se realizó de forma adecuada. Esto es conservando lecturas de alta calidad y sin introducir sesgo técnico aparente. Algunos cambios deberán ser evidentes con respecto al reporte de calidad de nuestros datos crudos. Por ejemplo, si observamos el reporte FastQC para la calidad de las bases podemos notar que el número total de secuencias se redujo por el efecto de eliminar aquellas que no cumplen con los criterios de calidad, además de ver un aumento de calidad en las bases al final de cada lectura, la sobrerrepresentación de secuencias prácticamente desaparece, la distribución de la longitud cambia debido al corte de lecturas por SLIDINGWINDOW y un pequeño cambio en el contenido GC, que representarían los cambios más importantes en esta primera etapa de control de calidad (Fig. 8).

Con este procesamiento de calidad, contamos con datos listos para los análisis posteriores. Es importante enfatizar que dependiendo de la plataforma de secuenciación o protocolo utilizado estos pasos pueden ser prescindibles y sin ningún problema en los análisis posteriores. Sin embargo, es de suma importancia verificar la calidad de las librerías para tomar decisiones con respecto a la calidad de los datos.

**(A)**

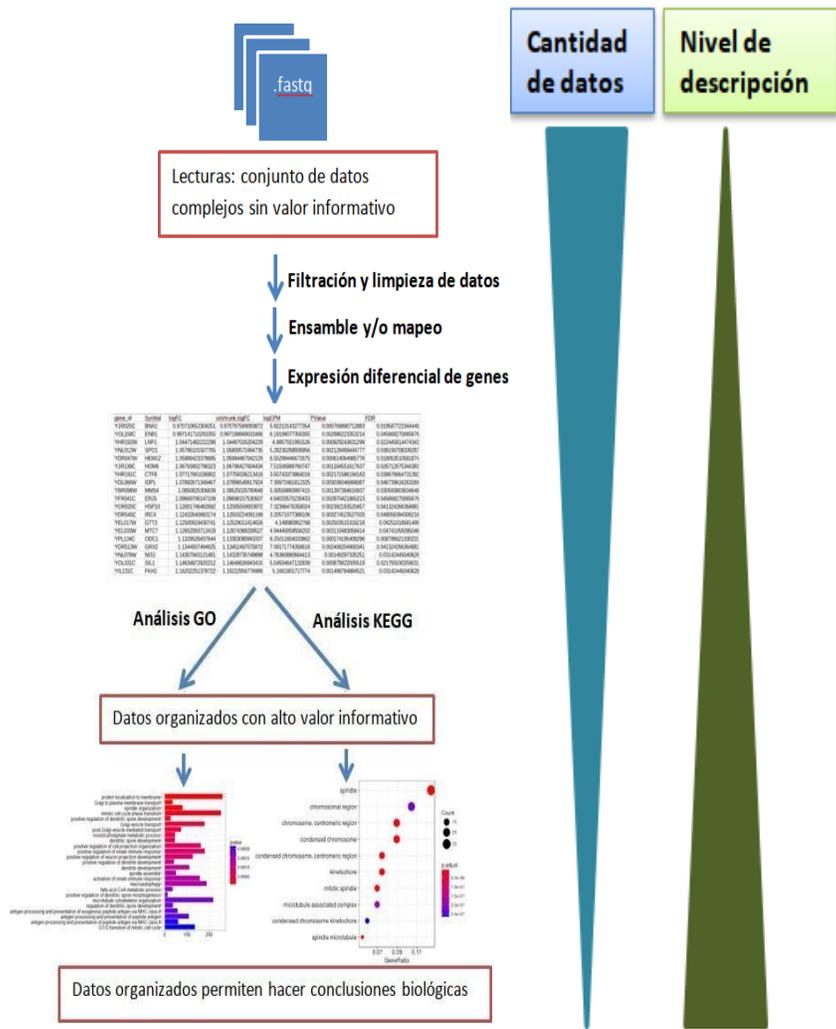
Measure	Value
Filename	1_R1_sortmerna_trimmomatic.fastq
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	14935051
Sequences flagged as poor quality	0
Sequence length	50-150
%GC	42



**Figura 8.** Cambios más importantes en el análisis de calidad, eliminación de bases de baja calidad y adaptadores de secuenciación. **A)** Estadísticas básicas que muestran, en este ejemplo representativo, la disminución de lecturas en la librería de 16684598 a 14935051 y un cambio en el contenido GC de 43 a 42. **B)** Aumento de calidad en las bases del extremo 3' de las lecturas. **C)** Cambio en la distribución de la longitud de secuencias debido al corte de aquellas que no cumplen con el valor de calidad especificado.

## **Análisis integral de los GDE's por ontología de genes y rutas metabólicas**

Como se mencionó en el capítulo anterior, el incremento de la utilización de las ciencias OMICAS ha permitido abordar estudios a escalas genómicas, proteómicas, transcriptómicas, metabolómicas, entre otras ómicas (Joyce y Palsson, 2006). Estos términos, recordando un poco, hacen referencia a todo el contenido molecular de un sistema biológico. Tomando como ejemplo nuestro estudio de caso, el transcriptoma de la levadura *S.c* es todo el conjunto de transcritos presentes en cada tiempo de muestreo. Enfatizando en la cantidad masiva de información que se obtiene de un experimento de RNA-seq, un problema fundamental es el manejo integral de dicha información. En los subtemas anteriores hemos filtrado por análisis de calidad, ensamblado o mapeado las lecturas de secuenciación y hemos realizado un análisis de expresión diferencial de genes comparando nuestros tiempos de fermentación. Estos procesos han permitido, de cierta forma, acortar y moldear la información de tal manera que de una compleja librería de secuencias se obtienen genes, ensamblados o mapeados, que se expresan en una condición. No obstante, esta información ya analizada no presenta ningún indicio de valor biológico, es decir solo tenemos una lista de genes diferencialmente expresados (GDE's). El enfoque más utilizado para organizar aún más la información y poder hacer conclusiones con alto valor biológico, es agrupar todos los GDE's en categorías ontológicas (GO) y rutas metabólicas (KEGG) (Carbon *et al.*, 2019; Kanehisa *et al.*, 2017). Después de todo, la ciencia se trata de hacer conclusiones sencillas a partir de un conjunto de datos complejos (Fig 9).



**Figura 9.** El procesamiento general de los datos se traduce en niveles de organización y descripción donde el más alto nivel de organización se representa en los análisis GO y KEGG. Esta figura representa una de las características más importantes de la

En el análisis ontológico (GO) se brinda un léxico integrativo para la comunidad en la anotación de genes. En datos de secuenciación como péptidos, ARN, ADN, este tipo de análisis da un término GO definido para cada gen en nuestro conjunto de datos. En esencia el análisis GO nos permite describir, clasificar y analizar datos. La tecnología que subyace a GO tiene su origen en inteligencia artificial basada en lógica que permite una inferencia sofisticada y detección de errores. El nivel más alto de la clasificación los términos GO

pueden representar eventos como componente celular, función molecular y proceso biológico. Donde el **proceso biológico** hace referencia al objetivo biológico al cual un gen determinado contribuye. Una subcategoría de proceso biológico describe una serie de eventos o un conjunto de eventos vistos en función molecular. **Función molecular** hace referencia a la actividad bioquímica del producto génico. Por ejemplo, hidrólisis, unión a ligandos específicos, entre otros. Por último, **componente celular**, describe el lugar en la célula donde el producto génico es activo (Hastings, 2017; Liu, 2017).

Por otro lado, el análisis mediante rutas metabólicas con la utilización de la base de datos de la Enciclopedia de Genes y Genomas de Kioto (KEGG por sus siglas en inglés), nos brinda un campo semántico altamente descriptivo de la funcionalidad de nuestros GDE's. La base de datos KEGG es una compilación de mapas verificados manualmente de las interacciones biológicas entre moléculas de un organismo. Fundamentalmente, es una fuente de conocimiento para el análisis sistemático de la función de genes.

Esto se realiza mediante la relación de información genómica con información funcional de alto nivel. La información genómica es una colección de genes para todos los genomas completamente secuenciado, anotados y se almacena en la base de datos GENE. La información de funcional de alto nivel contiene representación gráfica de procesos celulares y esta información se guarda en las bases de datos PATHWAY. Por último, la base de datos LIGAND proporciona información acerca de compuestos químicos, enzimas y reacciones enzimáticas (M Kanehisa y Goto, 2000; Manyam *et al.*, 2015).

Hablando de forma general, estas herramientas de clasificación, nos permite organizar la información de acuerdo a la funcionalidad de nuestro conjunto de genes. Esto nos permite obtener una imagen más detallada de lo que representa nuestra lista de GDE's



2012) dentro del entorno del proyecto Bioconductor (Gentleman *et al.*, 2004; Huber *et al.*, 2015).

Nuestro material de trabajo será una lista de genes diferencialmente expresados en las condiciones del tiempo fermentación, esta tabla contiene, entre otras cosas, los identificadores de los genes necesarios para realizar los análisis de enriquecimiento y agrupación. Específicamente, tomaremos como ejemplo, el día uno de la fermentación (Fig. 10).

entrezID	ORF_ID	Symbol	logFC	unshrunk.logFC	logCPM	PValue	FDR
853482	YJR025C	BNA1	0.970710652304251	0.970787569093872	5.92213143277354	0.000769890712883	0.019587722344449
854007	YOL158C	ENB1	0.997141710291055	0.997198669031666	6.19199077356355	0.002886223353214	0.045668270065676
856599	YHR192W	LNP1	1.04471482222298	1.04487026204229	4.8857551991526	0.000929243631299	0.022445814474343
855720	YNL012W	SPO1	1.05786101507765	1.05800572494735	5.28230268936956	0.002128469444777	0.038156708339287
851617	YDR047W	HEM12	1.05989423378685	1.05994467042129	6.55299446672675	0.000614064985778	0.018053510581874
853604	YJR139C	HOM6	1.06793802798323	1.06796427604434	7.51506989760747	0.001164551617637	0.026712675344383
856598	YHR191C	CTF8	1.07717661036902	1.07759336213416	3.50743373864016	0.002171586194163	0.038676664731392
851493	YDL066W	IDP1	1.07892671349467	1.07896548917924	7.39972461812325	0.003036046999087	0.046739616263269
852395	YBR098W	MMS4	1.0850825306839	1.08525026780648	5.00559893997415	0.001397384616937	0.030593803634648
850602	YFR041C	ERJ5	1.09669706147109	1.09698157530507	4.64033570230433	0.002870421865223	0.045668270065676
854185	YOR020C	HSP10	1.12001746482692	1.12005559003972	7.32396476358024	0.002392193520457	0.041324266364881
852153	YDR540C	IRC4	1.12432640983174	1.12503224091199	3.20571077388106	0.003274523527033	0.048056394306218
856698	YEL017W	GTT3	1.12500919430741	1.12529011414626	4.149080952768	0.002503515316218	0.04251018581499
856678	YEL033W	MTC7	1.12652093713418	1.12674388208527	4.94440059556202	0.003110483058414	0.04741055095046
855969	YPL134C	ODC1	1.1329526437644	1.13303080993337	6.25011604033862	0.000174195408296	0.008786621330231
852124	YDR513W	GRX2	1.1344937494925	1.13452497075872	7.09171774358616	0.002406204900341	0.041324266364881
855646	YNL078W	NIS1	1.14307943121481	1.14328735749898	4.76360880664413	0.00149297335251	0.03142445040628
854126	YOL031C	SIL1	1.14634872920212	1.14648626943415	5.54504647132839	0.000879922005519	0.021755030259031
854675	YIL131C	FKH1	1.16202251378722	1.16222956776986	5.1661801717774	0.001498784884521	0.03142445040628

**Figura 10.** Lista de genes positivamente regulados en el día uno de la fermentación. Es posible observar que se presentan los identificadores en las tres primeras columnas

Lo primero que tenemos que hacer es exportar la tabla de Excel al entorno de R, esto se consigue con los siguientes comandos:

```
#Exportar la tabla de GDE's de excel a R
> matriz_GDE_F1 <- read.csv("F1_GDE.csv", sep =
",", header = TRUE)
```

Con el comando anterior tenemos disponible los datos dentro del ambiente de R. Lo que se tiene que hacer luego es acceder solo a los identificadores genéticos para poder introducirlos en el programa ClusterProfile. En el primer análisis, GO, utilizaremos los

identificadores ENTREZ y los obtendremos de la matriz GDE's de la siguiente manera:

```
#Extraer los identificadores ENTREZ
> ENTREZ_ID_F1 <- matriz_GDE_F1[,2]

#Convertimos los identificadores a vector
necesario para el análisis GO
> ENTREZ_ID_F1 <- as.vector(ENTREZ_ID_F1)
```

Con esto obtenemos los datos necesarios para realizar el análisis de enriquecimiento por ontología de genes (GO). Es importante mencionar que necesitamos tener instalado el paquete de Bioconductor y otros paquetes para poder realizar los análisis de enriquecimiento. Esto se consigue ejecutando los siguientes comandos en R:

```
#Instalar todos los paquetes necesarios para el
análisis de enriquecimiento
#Primero instalamos Bioconductor
> if (!requireNamespace("BiocManager", quietly =
TRUE))
  install.packages("BiocManager")

#Luego instalamos los paquetes necesarios
> BiocManager::install("clusterProfiler")
> library(clusterProfiler)
> BiocManager::install("org.Sc.sgd.db")
> library(org.Sc.sgd.db)
> BiocManager::install("topGO")
> BiocManager::install("Rgraphviz")
> library(topGO)
> library(Rgraphviz)
> BiocManager::install("enrichplot")
> library(enrichplot)
```

A partir de esto, el entorno de R está listo para realizar los análisis de enriquecimiento por GO y KEGG. Para realizar el análisis necesitamos los identificadores ENTREZ, obtenidos anteriormente, para poder alimentar el software ClusterProfile. El análisis de enriquecimiento GO se consigue de la siguiente manera:

```
#Análisis de enriquecimiento por Ontología de Genes
#Análisis para Función Molecular
> GO_analisis_F1 <- enrichGO(ENTREZ_ID_F1, OrgDb
= "org.Sc.sgd.db",
KeyType = "ENTREZID", ont = "MF", pvalueCutoff =
0.05, pAdjustMethod = "BH")

#Extraemos una tabla que resume nuestros
resultados y la imprimimos en #nuestro directorio
de trabajo
> GO_MF_F1_resumen <- summary(GO_analisis_F1)
> write.table(GO_MF_F1_resumen, sep = '\t', file
= "GO_MF_F1_resumen.tab")

# Ahora graficamos nuestros resultados
> barplot(GO_analisis_F1, showCategory = 50)
```

Esto se hace para las tres categorías ontológicas de más alto nivel: función molecular, proceso biológico y componente celular. Esto se hace solo cambiando el parámetro *ont* en el comando *enrichGO*. En el ejemplo realizamos función molecular (MF), solo restaría cambiar a CC, componente celular y BP, para proceso biológico (Fig. 11).

Con estos resultados completamos el análisis de enriquecimiento por ontología de genes. Esta información nos da un indicio de posibles eventos moleculares en nuestras condiciones. Sin embargo, la información contundente es la brindada por una evaluación de rutas metabólicas mediante análisis de enriquecimiento de datos KEGG. Esto se consigue mediante la ejecución de los siguientes comandos:

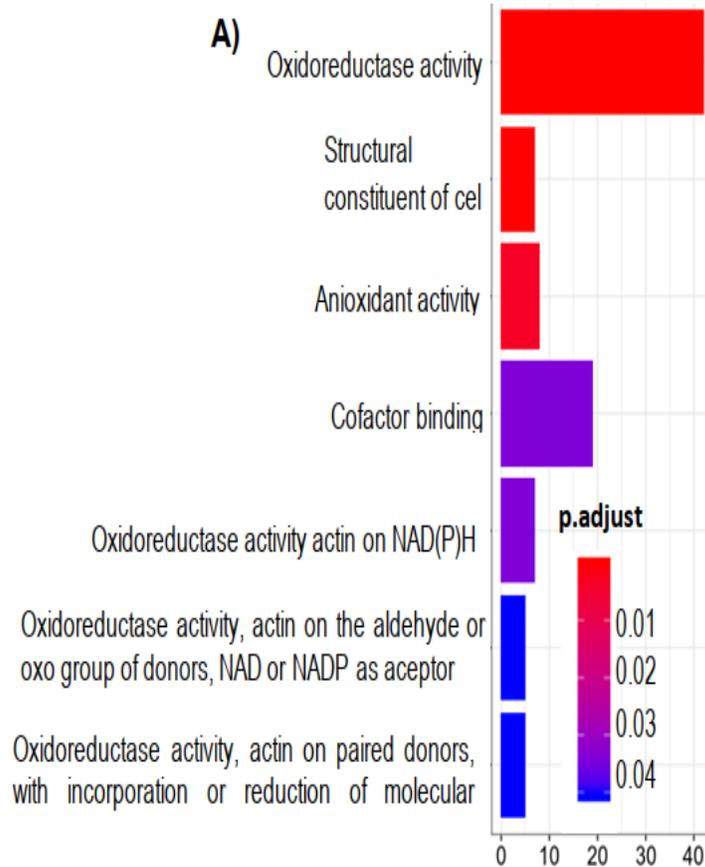
```

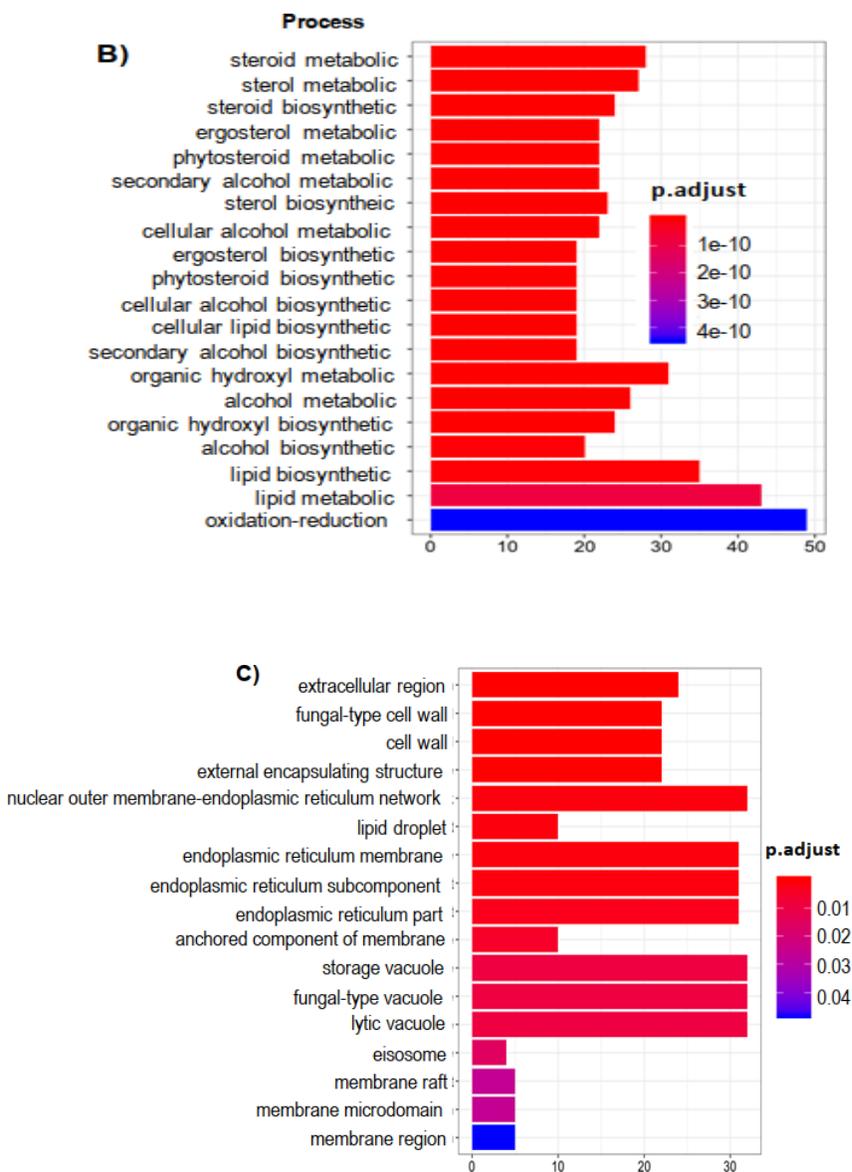
#Análisis de enriquecimiento de rutas metabólicas
con el uso de KEGG
#Primero recuperamos los identificadores para
KEGG de nuestra matriz de expresión
> KEGG_ID_F2 <- IND_F2_entrez_id_ref_genome[,3]

#Ejecutamos el análisis de rutas metabólicas
> f2_kegg_enrich <- enrichKEGG(KEGG_ID_F2,
organism = "sce", keyType = "kegg",
pvalueCutoff = 0.05, pAdjustMethod =
"BH")

#Realizamos una tabla resumen de nuestras rutas
metabólicas significativas
#y la imprimimos en nuestro directorio de trabajo
> F2_kegg_summary <- summary(f2_kegg_enrich)
> write.table(F2_kegg_summary, sep = "\t", file =
"F2_kegg_summary.tab")
# Graficamos nuestros resultados
> dotplot(f2_kegg_enrich, showCategory=30)

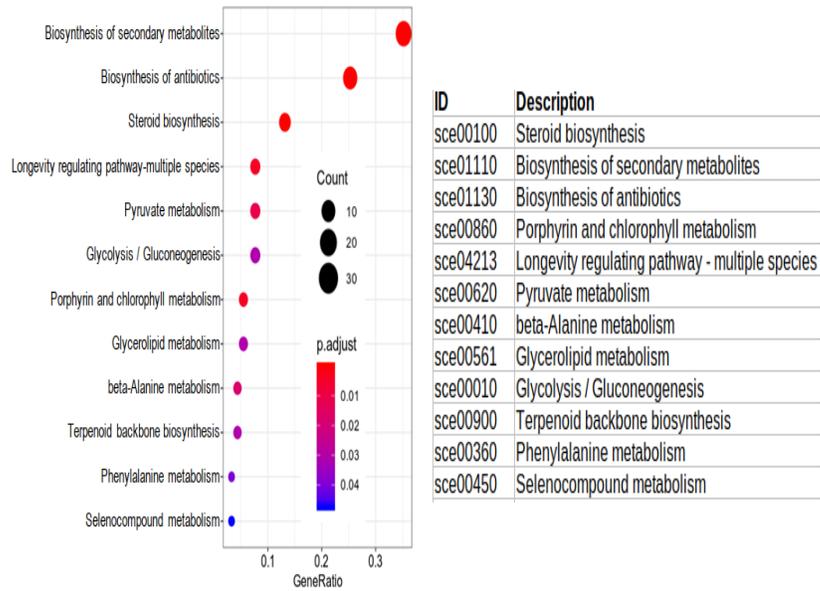
```





**Figura 11.** Ejemplos de resultados de los análisis de enriquecimientos por GO. **A)** Términos significativamente enriquecidos en función molecular, **B)** proceso biológico y **C)** componente celular en el día 1 de la fermentación.

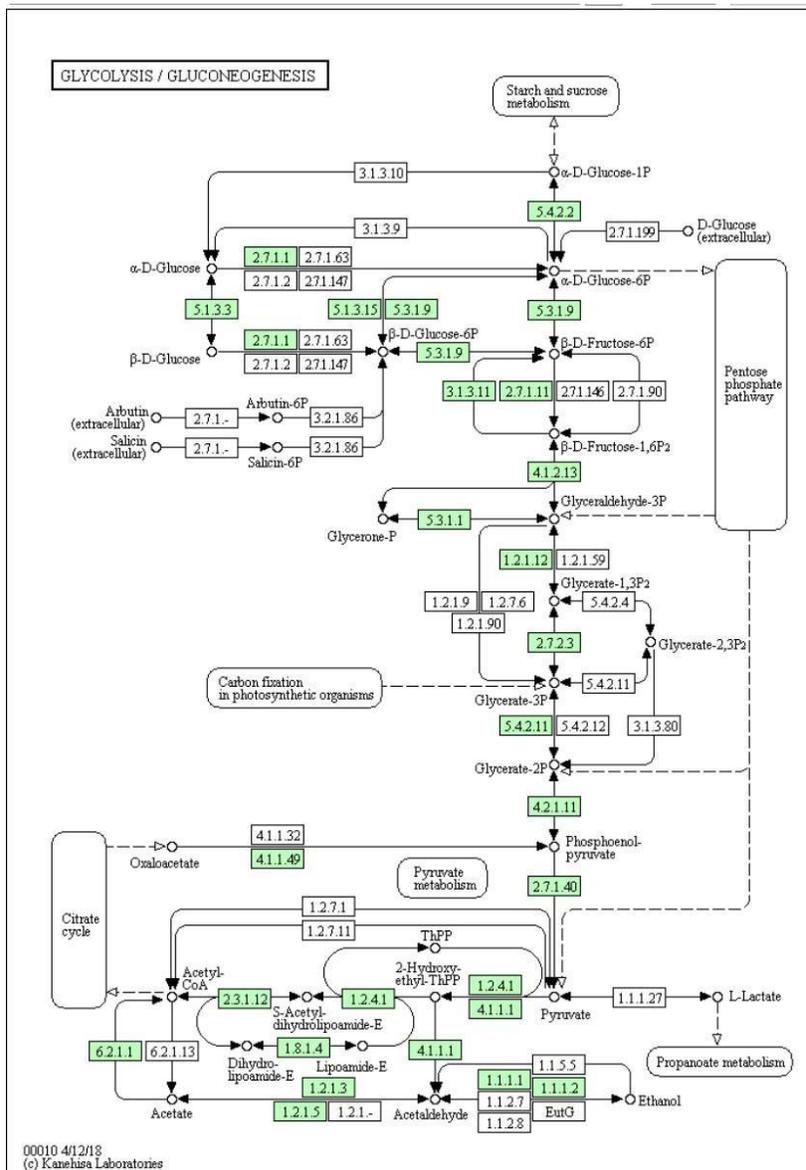
Esto se puede visualizar de forma gráfica (Fig. 12) o en forma de diagrama (Fig. 13). La forma gráfica es la que obtenemos al ejecutar el comando *dotplot* en la sección de arriba y los diagramas de rutas metabólicas se obtienen con los identificadores presentes en nuestra tabla resumida.



**Figura 12.** Se muestra gráficamente las rutas significativamente enriquecidas y una tabla con la descripción de la ruta y el identificador KEGG para la búsqueda de esquemas de rutas metabólicas en KEGG.

Si tomamos un identificador y hacemos la búsqueda en la base de datos KEGG (la consulta se realiza en la liga siguiente: <https://www.genome.jp/kegg/pathway.html>) para una representación esquemática de la ruta metabólica obtenemos el diagrama expuesto en la figura x4. Que corresponde a la ruta metabólica glucolisis/gluconeogénesis significativamente enriquecida en el día 1 de la fermentación. El identificador KEGG que utilizamos fue sce00010 a partir de nuestra tabla resumida en la Figura 13. Con esta información detalladamente organizada podemos hacer conclusiones sobre el estado metabólico de nuestro sistema biológico de estudio. Hasta este nivel nuestra información es un campo fértil para las propuestas de hipótesis y





## Literatura Citada

Carbon, S., Douglass, E., Dunn, N., Good, B., Harris, N. L., Lewis, S. E., M. Westerfield. 2019. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Research*, **47(D1)**, D330-D338. <https://doi.org/10.1093/nar/gky1055>

Gaudet, P., y C. Dessimoz. 2017. Gene Ontology: Pitfalls, biases, and remedies (pp. 189-205). Humana Press, New York, NY. [https://doi.org/10.1007/978-1-4939-3743-1\\_14](https://doi.org/10.1007/978-1-4939-3743-1_14)

Gentleman, R. C., Carey, V. J., Bates, D. M., Bolstad, B., Dettling, M., Dudoit, S., J. Zhang. 2004. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology*, **5(10)**.

Hastings, J. 2017. Primer on ontologies (pp. 3-13). Humana Press, New York, NY. [https://doi.org/10.1007/978-1-4939-3743-1\\_1](https://doi.org/10.1007/978-1-4939-3743-1_1)

Huber, W., Carey, V. J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B. S., and M. Morgan. 2015. Orchestrating high-throughput genomic analysis with Bioconductor. *Nature Methods*, **12(2)**, 115-121. <https://doi.org/10.1038/nmeth.3252>

Joyce, A. R., & Palsson, B. (2006). The model organism as a system: Integrating “omics” data sets. *Nature Reviews Molecular Cell Biology*, **7(3)**, 198-210. <https://doi.org/10.1038/nrm1857>

Kanehisa, M., & Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, **28(1)**, 27-30. <https://doi.org/10.1093/nar/28.1.27>

Kanehisa, Minoru, Furumichi, M., Tanabe, M., Sato, Y., & Morishima, K. (2017). KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Research*, **45(D1)**, D353-D361. <https://doi.org/10.1093/nar/gkw1092>

Liu, Z. J. (Ed.). (2017). *Bioinformatics in Aquaculture*. Chichester, UK: John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118782392>

Manyam, G., Birerdinc, A., & Baranova, A. (2015). KPP: KEGG Pathway Painter. *BMC Systems Biology*, **9(2)**, 2-5. <https://doi.org/10.1186/1752-0509-9-S2-S3>

## Bibliografía

Bentley, D. R., Balasubramanian, S., Swerdlow, H. P., Smith, G. P., Milton, J., Brown, C. G., ... Smith, A. J. (2008). Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*, **456(7218)**, 53-59. <https://doi.org/10.1038/nature07517>

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, **30(15)**, 2114-2120. <https://doi.org/10.1093/bioinformatics/btu170>

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., ... Mortazavi, A. (2016a). A survey of best practices for RNA-seq data analysis. *Genome Biology*, 17(1), 1-19. <https://doi.org/10.1186/s13059-016-0881-8>

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., ... Mortazavi, A. (2016b). A survey best practices for RNA-seq data analysis. *Genome Biology*, 17 1-19. <https://doi.org/10.1186/s13059-016-0881-8>

Delhomme, N., Niklas Mähler, N., Sundell, D., Chan Hvidsten, T. R., & Street, N. R. (2014). Guidelines for RN analysis. *Epigenesys*, (September 201 <https://doi.org/10.1109/HiPC.2015.19>

Duffy, E. E., Schofield, J. A., & Simon, M. D. (20 into transcriptome-wide RNA population dy chemistry of 4-thiouridine. *Wiley Interdisci 10(1)*, 1-18. <https://doi.org/10.1002/wrna>.

El-Ashram, S., Al Nasr, I., & Suo, X. (2 Extraction and optimization. *Biote <https://doi.org/10.1016/j.btre.2016>*

Fedurco, M., Romieu, A., Willi (2006). BTA, a novel reage efficient generation of soli *Acids Resea <https://doi.org/10.109>*

Head, S. R., Kiyom Nieuwerburgh, Library const and ch <https://do>

Illumin Tec ge

- Bioinformatics*, 28(24), 3211-3217.  
<https://doi.org/10.1093/bioinformatics/bts611>
- Kukurba, K. R., & Montgomery, S. B. (2015). Topic Introduction RNA Sequencing and Analysis, 951-970.  
<https://doi.org/10.1101/pdb.top084970>
- Leinonen, R., Sugawara, H., & Shumway, M. (2011). The sequence read archive. *Nucleic Acids Research*, 39(SUPPL. 1), 2010-2012.  
<https://doi.org/10.1093/nar/gkq1019>
- Li, Y., & Breaker, R. R. (1999). Kinetics of RNA degradation by specific base catalysis of transesterification involving the 2'-hydroxyl group. *Journal of the American Chemical Society*, 121(23), 5364-5372. <https://doi.org/10.1021/ja990592p>
- Masotti, A., & Preckel, T. (2006). Analysis of small RNAs with the Agilent 2100 Bioanalyzer. *Nature Methods*, 3(8), 2-3.  
<https://doi.org/10.1038/nmeth908>
- Mathe, E. A., & Davis, S. (2016). *Statistical genomics : methods and protocols*.
- Metzker, M. L. (2010). Sequencing technologies the next generation. *Nature Reviews Genetics*, 11(1), 31-46.  
<https://doi.org/10.1038/nrg2626>
- Mueller, O., Hahnenberger, K., Dittmann, M., Yee, H., Dubrow, R., Nagle, R., & Ilesley, D. (2000). A microfluidic system for high-speed reproducible DNA sizing and quantitation. *Electrophoresis*, 21(1), 128-134.  
[https://doi.org/10.1002/\(SICI\)1522-2683\(20000101\)21:1<128::AID-ELPS128>3.0.CO;2-M](https://doi.org/10.1002/(SICI)1522-2683(20000101)21:1<128::AID-ELPS128>3.0.CO;2-M)
- Nilsen, T. W. (2013). The Fundamentals of RNA Purification. *Cold Spring Harbor Protocols*, 8(7), 618-624.  
<https://doi.org/10.1101/pdb.top075838>
- Raj, A., & van Oudenaarden, A. (2008). Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences. *Cell*, 135(2), 216-226. <https://doi.org/10.1016/j.cell.2008.09.050>
- Reiman, M., Laan, M., Rull, K., & Söber, S. (2019). Effects of RNA integrity on transcript quantification by total RNA sequencing of clinically collected human placental samples. *The FASEB Journal*, 3298-3308. <https://doi.org/10.1096/fj.201601031RR>

Reuter, J. A., Spacek, D. V., & Snyder, M. P. (2015). High-Throughput Sequencing Technologies. *Molecular Cell*, 58(4), 586-597. <https://doi.org/10.1016/j.molcel.2015.05.004>

Rio, D. C., Ares, M., Hannon, G. J., & Nilsen, T. W. (2010). Purification of RNA using TRIzol (TRI reagent). *Cold Spring Harbor Protocols*, 2010(6), pdb.prot5439. <https://doi.org/10.1101/pdb.prot5439>

Sambrook, J., & Russell, D. W. (David W. (2001). *Molecular cloning : a laboratory manual*. Cold Spring Harbor Laboratory Press. Retrieved from [http://93.174.95.29/\\_ads/08ED769A253765B81383D10063642E47](http://93.174.95.29/_ads/08ED769A253765B81383D10063642E47)

Schofield, J. A., Duffy, E. E., Kiefer, L., Sullivan, M. C., & Simon, M. D. (2018). TimeLapse-seq: adding a temporal dimension to RNA sequencing through nucleoside recoding. *Nature Methods*, 15(3), 221-225. <https://doi.org/10.1038/nmeth.4582>

Schroeder, A., Mueller, O., Stocker, S., Salowsky, R., Leiber, M., Gassmann, M., ... Ragg, T. (2006). The RIN: An RNA integrity number for assigning integrity values to RNA measurements. *BMC Molecular Biology*, 7, 1-14. <https://doi.org/10.1186/1471-2199-7-3>

Schwalb, B., Michel, M., Zacher, B., Frühauf, K., Demel, C., Tresch, A., ... Cramer, P. (2016). TT-seq maps the human transient transcriptome. *Science (New York, N. Y.)*, 352(6290), 1225-1228. <https://doi.org/10.1126/science.aad9841>

Shekhawat, K., Patterton, H., Bauer, F. F., & Setati, M. E. (2019). RNA-seq based transcriptional analysis of *Saccharomyces cerevisiae* and *Lachancea thermotolerans* in mixed-culture fermentations under anaerobic conditions. *BMC Genomics*, 20(1), 145. <https://doi.org/10.1186/s12864-019-5511-x>

Stark, R., Grzelak, M., & Hadfield, J. (2019). RNA sequencing: the teenage years. *Nature Reviews Genetics*. <https://doi.org/10.1038/s41576-019-0150-2>

Sweet, M. E., Cocciolo, A., Slavov, D., Jones, K. L., Sweet, J. R., Graw, S. L., ... Taylor, M. R. G. (2018). Transcriptome analysis of human heart failure reveals dysregulated cell adhesion in dilated cardiomyopathy and activated immune pathways in ischemic heart failure. *BMC Genomics*, 19(1), 812. <https://doi.org/10.1186/s12864-018-5213-9>



## 11. CONCLUSIONES GENERALES

El presente trabajo arroja luz sobre los factores subyacentes a los diferentes fenotipos de consumo de maltosa y maltotriosa que existen en algunas levaduras cerveceras. Nuestras observaciones sugieren que el evento principal asociado con las diferencias en el consumo de alfa-glucósidos en las levaduras estudiadas es la expresión génica diferencial de los reguladores positivos de los transportadores.

Ambas levaduras tienen la misma maquinaria para una fermentación eficiente, pero desde una perspectiva evolutiva, la levadura SpIB2 ha producido un fenotipo de consumo de alfa-glucósidos eficiente. Esto sugiere que la levadura SpIB1 podría lograr una fermentación eficiente y posiblemente un cambio en su panorama general de fermentación a través de un experimento evolutivo asistido.

Por último, las aportaciones que se hacen en el campo del tema relacionado con la regulación entre reinos es que los mecanismos de interferencia también pueden suceder en ambientes de importancia biotecnológica, como la elaboración de cerveza. Cuyo efecto puede incidir en el proceso estándar de una fermentación lager.

Si bien, no se presenta el mecanismo de acción de dicho fenómeno, se presenta evidencia empírica de su posible contribución en la fermentación eficiente de cada cepa analizada. Esto abre la discusión para futuras investigaciones que se planteen el estudio de la comunicación entre reinos a través de RNA en contextos puramente industriales.

## 12. PERSPECTIVAS

A lo largo de esta investigación hemos analizado profundamente los aspectos, principalmente moleculares, relacionados con la diferencia en la toma de alfa glucósidos, así como el posible efecto de RNA presente en el mosto cervecero en la expresión genética global en dos levaduras *Saccharomyces pastorianus* empleadas en el proceso cervecero.

Los resultados expuestos en este trabajo impactan directamente el diseño de medios de cultivo y/o mostos cerveceros que implique distinta cantidad de metabolitos represores como la glucosa y el correcto monitoreo de la presencia de RNA en el mosto. También, estos resultados sugieren que, a pesar de que las dos levaduras contienen el mismo potencial genético de consumo de maltosa y maltotriosa, la represión por catabolito está impartiendo diferentes interacciones en cada levadura.

Es por esto que dentro de las principales perspectivas se encuentra caracterizar la respuesta molecular de los loci MAL en presencia de distintas concentraciones de glucosa en un medio químicamente definido. Además de la caracterización de la expresión genética de los transportadores, se debe de implementar un ensayo bioquímico de transporte frente a distintas concentraciones de los sustratos y el represor.

De esta forma, se contribuiría sólidamente a las principales conclusiones de esta investigación. Por otro lado, debido a que las aproximaciones computacionales sobre la determinación del número copias está limitada por la estructura y tipo de datos empleados, el diseño y ejecución de un experimento de PCR cuantitativa (qPCR) es de vital importancia para la confirmación de los resultados sobre la variación en el número de copias.

Además, como lo sugiere este trabajo, los reguladores positivos son de vital importancia y su estudio traería beneficios importantes a la industria cervecera. Por último, lo relacionado con el efecto de RNA en el medio debe ser

probado por experimentos de qPCR para validar las observaciones realizadas en este trabajo.

Tomando como base este trabajo, se propone que se realicen experimentos relacionados con la inducción de los loci MAL, esto es, experimentos de crecimiento de las levaduras en cuestión en medios químicamente definidos con maltosa o maltotriosa y analizar la expresión genética de los loci MAL, además, es necesario el desarrollo de experimentos en donde se evalúe la expresión genética de los genes afectados por la presencia de RNA en el mosto cervecero como control de calidad para el curso de una fermentación estándar.

### 13. BIBLIOGRAFÍA

- Alcid, E. A., & Tsukiyama, T. (2016). Expansion of antisense lncRNA transcriptomes in budding yeast species since the loss of RNAi. *Nature Structural and Molecular Biology*, 23(5), Article 5. <https://doi.org/10.1038/nsmb.3192>
- Alves, S. L., Herberts, R. A., Hollatz, C., Miletti, L. C., & Stambuk, B. U. (2007). Maltose and maltotriose active transport and fermentation by *Saccharomyces cerevisiae*. *Journal of the American Society of Brewing Chemists*, 65(2), Article 2. <https://doi.org/10.1094/ASBCJ-2007-0411-01>
- Alves, S. L., Herberts, R. A., Hollatz, C., Trichez, D., Miletti, L. C., Araujo, P. S. D., & Stambuk, B. U. (2008). Molecular analysis of maltotriose active transport and fermentation by *Saccharomyces cerevisiae* reveals a determinant role for the AGT1 permease. *Applied and Environmental Microbiology*, 74(5), Article 5. <https://doi.org/10.1128/AEM.02570-07>
- Amer, O. E., Mahmoud, M. A., El-Samawaty, A. E. R. M. A., & Sayed, S. R. M. (2011). Non liquid nitrogen-based-method for isolation of DNA from filamentous fungi. *African Journal of Biotechnology*, 10(65), Article 65. <https://doi.org/10.5897/ajb11.1401>
- Anders, S., McCarthy, D. J., Chen, Y., Okoniewski, M., Smyth, G. K., Huber, W., & Robinson, M. D. (2013). Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. *Nature Protocols*, 8(9), 1765-1786. <https://doi.org/10.1038/nprot.2013.099>
- Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics (Oxford, England)*, 31(2), Article 2. <https://doi.org/10.1093/bioinformatics/btu638>
- Avbelj, M., Zupan, J., Kranjc, L., & Raspor, P. (2015, septiembre 22). *Quorum-Sensing Kinetics in Saccharomyces cerevisiae: A Symphony of ARO Genes and Aromatic Alcohols* (world) [Research-article]. ACS Publications; American Chemical Society. <https://doi.org/10.1021/acs.jafc.5b03400>

- Baiano, A. (2021). Craft beer: An overview. *Comprehensive Reviews in Food Science and Food Safety*, 20(2), 1829-1856. <https://doi.org/10.1111/1541-4337.12693>
- Baker, E. C. P., Peris, D., Moriarty, R. V., Li, X. C., Fay, J. C., & Hittinger, C. T. (2019). Mitochondrial DNA and temperature tolerance in lager yeasts. *Science Advances*, 5(1), Article 1. <https://doi.org/10.1126/sciadv.aav1869>
- Baker, E. C., Wang, B., Bellora, N., Peris, D., Hulfachor, A. B., Koshalek, J. A., Adams, M., Libkind, D., & Hittinger, C. T. (2015). The genome sequence of *Saccharomyces eubayanus* and the domestication of lager-brewing yeasts. *Molecular Biology and Evolution*, 32(11), Article 11. <https://doi.org/10.1093/molbev/msv168>
- Baker, E. P., & Hittinger, C. T. (2019). Evolution of a novel chimeric maltotriose transporter in *Saccharomyces eubayanus* from parent proteins unable to perform this function. *PLoS Genetics*, 15(4), e1007786. <https://doi.org/10.1371/journal.pgen.1007786>
- Balarezo-Cisneros, L. N., Parker, S., Fraczek, M. G., Timouma, S., Wang, P., O'Keefe, R. T., Millar, C. B., & Delneri, D. (2021). Functional and transcriptional profiling of non-coding RNAs in yeast reveal context-dependent phenotypes and in trans effects on the protein regulatory network. *PLoS Genetics*, 17(1), e1008761. <https://doi.org/10.1371/journal.pgen.1008761>
- Bamforth, C. W. (2000). Brewing and brewing research: Past, present and future. *Journal of the Science of Food and Agriculture*, 80(9), Article 9. [https://doi.org/10.1002/1097-0010\(200007\)80:9<1371::AID-JSFA654>3.0.CO;2-K](https://doi.org/10.1002/1097-0010(200007)80:9<1371::AID-JSFA654>3.0.CO;2-K)
- Bamforth, C. W. (2017). Progress in Brewing Science and Beer Production. *Annual Review of Chemical and Biomolecular Engineering*, 8(1), Article 1. <https://doi.org/10.1146/annurev-chembioeng-060816-101450>
- Bilir, Ö., Göl, D., Hong, Y., McDowell, J. M., & Tör, M. (2022). Small RNA-based plant protection against diseases. *Frontiers in Plant Science*, 13. <https://doi.org/10.3389/fpls.2022.951097>
- Bing, J., Han, P. J., Liu, W. Q., Wang, Q. M., & Bai, F. Y. (2014). Evidence for a far east asian origin of lager beer yeast. *Current Biology*, 24(10), Article 10. <https://doi.org/10.1016/j.cub.2014.04.031>
- Bodenhofer, U., Bonatesta, E., Horejš-Kainrath, C., & Hochreiter, S. (2015). Msa: An R package for multiple sequence alignment. *Bioinformatics*, 31(24), Article 24. <https://doi.org/10.1093/bioinformatics/btv494>
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), Article 15. <https://doi.org/10.1093/bioinformatics/btu170>
- Bonatto, D. (2021). The diversity of commercially available ale and lager yeast strains and the impact of brewer's preferential yeast choice on the fermentative beer profiles. *Food Research International*, 141(July 2020), Article July 2020. <https://doi.org/10.1016/j.foodres.2021.110125>
- Brickwedde, A., Broek, M. van den, Geertman, J. M. A., Magalhães, F., Kuijpers, N. G. A., Gibson, B., Pronk, J. T., & Daran, J. M. G. (2017). Evolutionary engineering in chemostat cultures for improved maltotriose fermentation kinetics in *Saccharomyces pastorianus* lager brewing yeast. *Frontiers in Microbiology*, 8(SEP), Article SEP. <https://doi.org/10.3389/fmicb.2017.01690>
- Brickwedde, A., Brouwers, N., Broek, M. van den, Murillo, J. S. G., Fraiture, J. L., Pronk, J. T., & Daran, J. M. G. (2018). Structural, physiological and regulatory analysis of maltose transporter genes in *Saccharomyces eubayanus* CBS 12357T. *Frontiers in Microbiology*, 9(AUG), Article AUG. <https://doi.org/10.3389/fmicb.2018.01786>
- Brouwers, N., Brickwedde, A., Vries, A. R. G. de, Broek, M. van den, Weening, S. M., Eijnden, L. van den, Diderich, J. A., Bai, F. Y., Pronk, J. T., & Daran, J. M. G. (2019). Himalayan *Saccharomyces eubayanus* genome sequences reveal genetic markers explaining

- heterotic maltotriose consumption by *Saccharomyces pastorianus* hybrids. *Applied and Environmental Microbiology*, 85(22), Article 22. <https://doi.org/10.1128/AEM.01516-19>
- Brouwers, N., Vries, A. R. G. de, Broek, M. van den, Weening, S. M., Schuurman, T. D. E., Kuijpers, N. G. A., Pronk, J. T., & Daran, J. M. G. (2019). In vivo recombination of *Saccharomyces eubayanus* maltose-transporter genes yields a chimeric transporter that enables maltotriose fermentation. *PLoS Genetics*, 15(4), Article 4. <https://doi.org/10.1371/journal.pgen.1007853>
- Brown, C. A., Murray, A. W., & Verstrepen, K. J. (2010). Rapid Expansion and Functional Divergence of Subtelomeric Gene Families in Yeasts. *Curr Biol.*, 20(10), Article 10. <https://doi.org/10.1016/j.cub.2010.04.027>. Rapid
- Bryant, D. M., Johnson, K., DiTommaso, T., Tickle, T., Couger, M. B., Payzin-Dogru, D., Lee, T. J., Leigh, N. D., Kuo, T. H., Davis, F. G., Bateman, J., Bryant, S., Guzikowski, A. R., Tsai, S. L., Coyne, S., Ye, W. W., Freeman, R. M., Peshkin, L., Tabin, C. J., ... Whited, J. L. (2017). A Tissue-Mapped Axolotl De Novo Transcriptome Enables Identification of Limb Regeneration Factors. *Cell Reports*, 18(3), Article 3. <https://doi.org/10.1016/j.celrep.2016.12.063>
- Cabras, I., & Higgins, D. M. (2016). Beer, brewing, and business history. *Business History*, 58(5), Article 5. <https://doi.org/10.1080/00076791.2015.1122713>
- Cambiaghi, T. D., Pereira, C. M., Shanmugam, R., Bolech, M., Wek, R. C., Sattlegger, E., & Castilho, B. A. (2014). Evolutionarily conserved IMPACT impairs various stress responses that require GCN1 for activating the eIF2 kinase GCN2. *Biochemical and Biophysical Research Communications*, 443(2), Article 2. <https://doi.org/10.1016/j.bbrc.2013.12.021>
- Cameroni, E., Hulo, N., Roosen, J., Winderickx, J., & Virgilio, C. D. (2004). The novel yeast PAS kinase Rim15 orchestrates G0-associated antioxidant defense mechanisms. *Cell Cycle*, 3(4), Article 4. <https://doi.org/10.4161/cc.3.4.791>
- Cavalieri, D., Rizzetto, L., Tocci, N., Rivero, D., Asquini, E., Si-Ammour, A., Bonechi, E., Ballerini, C., & Viola, R. (2016). Plant microRNAs as novel immunomodulatory agents. *Scientific Reports*, 6(1), 25761. <https://doi.org/10.1038/srep25761>
- Cech, T. R., & Steitz, J. A. (2014). The noncoding RNA revolution—Trashing old rules to forge new ones. *Cell*, 157(1), Article 1. <https://doi.org/10.1016/j.cell.2014.03.008>
- Cerutti, H., & Casas-Mollano, J. A. (2006). On the origin and functions of RNA-mediated silencing: From protists to man. *Current Genetics*, 50(2), Article 2. <https://doi.org/10.1007/s00294-006-0078-x>
- Chi, X., Wang, Z., Wang, Y., Liu, Z., Wang, H., & Xu, B. (2023). Cross-Kingdom Regulation of Plant-Derived miRNAs in Modulating Insect Development. *International Journal of Molecular Sciences*, 24(9), 7978. <https://doi.org/10.3390/ijms24097978>
- Cousseau, F. E. M., Alves, S. L., Trichez, D., & Stambuk, B. U. (2013). Characterization of maltotriose transporters from the *Saccharomyces eubayanus* subgenome of the hybrid *Saccharomyces pastorianus* lager brewing yeast strain Weihenstephan 34/70. *Letters in Applied Microbiology*, 56(1), Article 1. <https://doi.org/10.1111/lam.12011>
- Craig, C. M., & Darryl, C. J. (2004). Revealing the world of RNA interference. *Nature*, 431(September), Article September.
- Dang, Y., Yang, Q., Xue, Z., & Liu, Y. (2011). RNA interference in fungi: Pathways, functions, and applications. *Eukaryotic Cell*, 10(9), Article 9. <https://doi.org/10.1128/EC.05109-11>
- Day, R. E., Rogers, P. J., Dawes, I. W., & Higgins, V. J. (2002). Molecular analysis of maltotriose transport and utilization by *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology*, 68(11), Article 11. <https://doi.org/10.1128/AEM.68.11.5326-5335.2002>

- Delhomme, N., Mähler, N. N., Sundell, D., Chanaka, M., Hvidsten, T. R., & Street, N. R. (2014). Guidelines for RNA-Seq data analysis. *Epigenesys*, September 2015, Article September 2015. <https://doi.org/10.1109/HiPC.2015.19>
- Dietvorst, J., Londesborough, J., & Steensma, H. Y. (2005). Maltotriose utilization in lager yeast strains: MTTI encodes a maltotriose transporter. *Yeast*, 22(10), Article 10. <https://doi.org/10.1002/yea.1279>
- Dietvorst, J., Walsh, M. C., Heusden, G. P. H. V., & Steensma, H. Y. (2010). Comparison of the MTT1- and MAL31-like maltose transporter genes in lager yeast strains. *FEMS Microbiology Letters*, 310(2), Article 2. <https://doi.org/10.1111/j.1574-6968.2010.02056.x>
- Dimidi, E., Cox, S. R., Rossi, M., & Whelan, K. (2019). Fermented Foods: Definitions and Characteristics, Impact on the Gut Microbiota and Effects on Gastrointestinal Health and Disease. *Nutrients*, 11(8), 1806. <https://doi.org/10.3390/nu11081806>
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1), Article 1. <https://doi.org/10.1093/bioinformatics/bts635>
- Drew, D., North, R. A., Nagarathinam, K., & Tanabe, M. (2021). Structures and General Transport Mechanisms by the Major Facilitator Superfamily (MFS). *Chemical Reviews*, 121(9), Article 9. <https://doi.org/10.1021/acs.chemrev.0c00983>
- Drinnenberg, I. A., Weinberg, D. E., Xie, K. T., Mower, J. P., Wolfe, K. H., Fink, G. R., & Bartel, D. P. (2009). RNAi in budding yeast. *Science*, 326(5952), Article 5952. <https://doi.org/10.1126/science.1176945>
- Dunn, B., & Sherlock, G. (2008). Reconstruction of the genome origins and evolution of the hybrid lager yeast *Saccharomyces pastorianus*. *Genome Research*, 18(10), Article 10. <https://doi.org/10.1101/gr.076075.108>
- Duval, E. H., Alves, S. L., Dunn, B., Sherlock, G., & Stambuk, B. U. (2010). Microarray karyotyping of maltose-fermenting *Saccharomyces* yeasts with differing maltotriose utilization profiles reveals copy number variation in genes involved in maltose and maltotriose utilization. *Journal of Applied Microbiology*, 109(1), Article 1. <https://doi.org/10.1111/j.1365-2672.2009.04656.x>
- EK, H., F, C., C, S., H, J., & CA, M. (1995). Characterization of AGT1 encoding a general alpha-glucoside transporter from *Saccharomyces*. *Molecular microbiology*, 17(6), Article 6. [https://doi.org/10.1111/J.1365-2958.1995.MMI\\_17061093.X](https://doi.org/10.1111/J.1365-2958.1995.MMI_17061093.X)
- Gallone, B., Mertens, S., Gordon, J. L., Maere, S., Verstrepen, K. J., & Steensels, J. (2018). Origins, evolution, domestication and diversity of *Saccharomyces* beer yeasts. *Current Opinion in Biotechnology*, 49, 148-155. <https://doi.org/10.1016/j.copbio.2017.08.005>
- Gallone, B., Steensels, J., Prahl, T., Soriaga, L., Saels, V., Herrera-Malaver, B., Merlevede, A., Roncoroni, M., Voordeckers, K., Miraglia, L., Teiling, C., Steffy, B., Taylor, M., Schwartz, A., Richardson, T., White, C., Baele, G., Maere, S., & Verstrepen, K. J. (2016). Domestication and Divergence of *Saccharomyces cerevisiae* Beer Yeasts. *Cell*, 166(6), Article 6. <https://doi.org/10.1016/j.cell.2016.08.020>
- Gancedo, J. M. (1998). Yeast Carbon Catabolite Repression. *Microbiology and Molecular Biology Reviews*, 62(2), Article 2. <https://doi.org/10.1128/mnbr.62.2.334-361.1998>
- García-Ríos, E., Morard, M., Parts, L., Liti, G., & Guillamón, J. M. (2017). The genetic architecture of low-temperature adaptation in the wine yeast *Saccharomyces cerevisiae*. *BMC Genomics*, 18(1), Article 1. <https://doi.org/10.1186/s12864-017-3572-2>
- Gibson, B., & Liti, G. (2015). *Saccharomyces pastorianus*: Genomic insights inspiring innovation for industry. *Yeast*, 32(1), Article 1. <https://doi.org/10.1002/yea.3033>
- Gibson, B. R., Storgårds, E., Krogerus, K., & Vidgren, V. (2013). Comparative physiology and fermentation performance of Saaz and Froberg lager yeast strains and the parental

- species *Saccharomyces eubayanus*. *Yeast*, 30(7), Article 7.  
<https://doi.org/10.1002/yea.2960>
- Gong, S., & Blundell, T. L. (2010). Structural and functional restraints on the occurrence of single amino acid variations in human proteins. *PLoS ONE*, 5(2), Article 2.  
<https://doi.org/10.1371/journal.pone.0009186>
- Green, M. R., & Sambrook, J. (2021). Total RNA Extraction from *Saccharomyces cerevisiae* Using Hot Acid Phenol. *Cold Spring Harbor Protocols*, 2021(12), Article 12.  
<https://doi.org/10.1101/PDB.PROT101691>
- Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., Eccles, D., Li, B., Lieber, M., MacManes, M. D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C. N., ... Regev, A. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, 8(8), Article 8.  
<https://doi.org/10.1038/nprot.2013.084>
- Han, E.-K., Cotty, F., Sottas, C., Jiang, H., & Michels, C. A. (1995). Characterization of AGT1 encoding a general  $\alpha$ -glucoside transporter from *Saccharomyces*. *Molecular Microbiology*, 17(6), Article 6. [https://doi.org/10.1111/j.1365-2958.1995.mmi\\_17061093.x](https://doi.org/10.1111/j.1365-2958.1995.mmi_17061093.x)
- Hanson, P. K. (2018). *Saccharomyces cerevisiae*: A Unicellular Model Genetic Organism of Enduring Importance. *Current Protocols Essential Laboratory Techniques*, 16(1), Article 1. <https://doi.org/10.1002/cpet.21>
- Hatanaka, H., Toyonaga, H., Ishida, Y., Mizohata, E., & Ono, E. (2022). Functional diversity and plasticity in the sugar preferences of *Saccharomyces* MALT transporters in domesticated yeasts. *FEMS Yeast Research*, 22(1), foac055. <https://doi.org/10.1093/femsyr/foac055>
- Hazelwood, L. A., Daran, J.-M., Maris, A. J. A. van, Pronk, J. T., & Dickinson, J. R. (2008). The Ehrlich pathway for fusel alcohol production: A century of research on *Saccharomyces cerevisiae* metabolism. *Applied and environmental microbiology*, 74(8), Article 8.  
<https://doi.org/10.1128/AEM.02625-07>
- He, Y., Dong, J., Yin, H., Zhao, Y., Chen, R., Wan, X., Chen, P., Hou, X., Liu, J., & Chen, L. (2014). Wort composition and its impact on the flavour-active higher alcohol and ester formation of beer—A review. *Journal of the Institute of Brewing*, 120(3), Article 3.  
<https://doi.org/10.1002/jib.145>
- Henderson, R., & Poolman, B. (2017). Proton-solute coupling mechanism of the maltose transporter from *Saccharomyces cerevisiae*. *Scientific Reports*, 7(1), Article 1.  
<https://doi.org/10.1038/s41598-017-14438-1>
- Herbst, R. H., Bar-Zvi, D., Reikhav, S., Soifer, I., Breker, M., Jona, G., Shimoni, E., Schuldiner, M., Levy, A. A., & Barkai, N. (2017). Heterosis as a consequence of regulatory incompatibility. *BMC Biology*, 15(1), Article 1. <https://doi.org/10.1186/s12915-017-0373-7>
- Hernández-Soto, A., Echeverría-Beirute, F., & Guzmán-Hernández, T. (2021). The RNAi as a tool to control tropical pathogens. *Agronomía Mesoamericana*, 326-337.  
<https://doi.org/10.15517/am.v32i1.40896>
- Hernández-Vásquez César Ignacio. (2020). *Análisis del transcriptoma, basado en RNA-seq, de una levadura cervecera tipo lager en fermentación a nivel industrial y laboratorio* [Universidad Autónoma de Nuevo León]. <http://eprints.uanl.mx/id/eprint/21059>
- Hollatz, C., & Stambuk, B. U. (2001). Colorimetric determination of active  $\alpha$ -glucoside transport in *Saccharomyces cerevisiae*. *Journal of Microbiological Methods*, 46(3), Article 3.  
[https://doi.org/10.1016/S0167-7012\(01\)00281-0](https://doi.org/10.1016/S0167-7012(01)00281-0)
- Hongay, C. F., Grisafi, P. L., Galitski, T., & Fink, G. R. (2006). Antisense Transcription Controls Cell Fate in *Saccharomyces cerevisiae*. *Cell*, 127(4), Article 4.  
<https://doi.org/10.1016/j.cell.2006.09.038>

- Horák, J. (2013). Regulations of sugar transporters: Insights from yeast. *Current genetics*, 59(1-2), Article 1-2. <https://doi.org/10.1007/s00294-013-0388-8>
- Hornsey, I. S. (2003). *History of Beer and Brewing*. Royal Society of Chemistry. <https://doi.org/10.1039/9781847550026>
- Hua, C., Zhao, J.-H., & Guo, H.-S. (2018). Trans-Kingdom RNA Silencing in Plant–Fungal Pathogen Interactions. *Molecular Plant*, 11(2), 235-244. <https://doi.org/10.1016/j.molp.2017.12.001>
- Huang, C.-Y., Wang, H., Hu, P., Hamby, R., & Jin, H. (2019). Small RNAs – Big Players in Plant-Microbe Interactions. *Cell Host & Microbe*, 26(2), 173-182. <https://doi.org/10.1016/j.chom.2019.07.021>
- Huber, F., Bunina, D., Gupta, I., Khmelinskii, A., Meurer, M., Theer, P., Steinmetz, L. M., & Knop, M. (2016). Protein Abundance Control by Non-coding Antisense Transcription. *Cell Reports*, 15(12), Article 12. <https://doi.org/10.1016/j.celrep.2016.05.043>
- Jespersen, L., Cesar, L. B., Meaden, P. G., & Jakobsen, M. (1999). Multiple  $\alpha$ -glucoside transporter genes in brewer's yeast. *Applied and Environmental Microbiology*, 65(2), Article 2. <https://doi.org/10.1128/aem.65.2.450-456.1999>
- Jing, S., Xu, J., Tang, H., Li, P., Yu, B., & Liu, Q. (2023). The roles of small RNAs in rice-brown planthopper interactions. *Frontiers in Plant Science*, 14, 1326726. <https://doi.org/10.3389/fpls.2023.1326726>
- Kayikci, Ö., & Nielsen, J. (2015). Glucose repression in *Saccharomyces cerevisiae*. *FEMS Yeast Research*, 15(6), Article 6. <https://doi.org/10.1093/femsyr/fov068>
- Kiss, E., Hegedüs, B., Virágh, M., Varga, T., Merényi, Z., Kószó, T., Bálint, B., Prasanna, A. N., Krizsán, K., Kocsubé, S., Riquelme, M., Takeshita, N., & Nagy, L. G. (2019). Comparative genomics reveals the origin of fungal hyphae and multicellularity. *Nature Communications*, 10(1), 4080. <https://doi.org/10.1038/s41467-019-12085-w>
- Kok, Y. J., Ye, L., Muller, J., Ow, D. S. W., & Bi, X. (2019). Brewing with malted barley or raw barley: What makes the difference in the processes? *Applied Microbiology and Biotechnology*, 103(3), Article 3. <https://doi.org/10.1007/s00253-018-9537-9>
- Kopylova, E., Noé, L., & Touzet, H. (2012). SortMeRNA: Fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics*, 28(24), Article 24. <https://doi.org/10.1093/bioinformatics/bts611>
- Krogerus, K., Magalhães, F., Vidgren, V., & Gibson, B. (2015). New lager yeast strains generated by interspecific hybridization. *Journal of Industrial Microbiology and Biotechnology*, 42(5), Article 5. <https://doi.org/10.1007/s10295-015-1597-6>
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9(4), Article 4. <https://doi.org/10.1038/nmeth.1923>
- Leitão, A. L., & Enguita, F. J. (Eds.). (2016). *Non-coding RNAs and Inter-kingdom Communication*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-39496-1>
- León-Medina, P. M. D., Elizondo-González, R., Damas-Buenrostro, L. C., Geertman, J. M., Broek, M. V. den, Galán-Wong, L. J., Ortiz-López, R., & Pereyra-Alfárez, B. (2016). Genome annotation of a *Saccharomyces* sp. Lager brewer's yeast. *Genomics Data*, 9, 25-29. <https://doi.org/10.1016/j.gdata.2016.05.009>
- Li, B., & Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*, 12(1), Article 1. <https://doi.org/10.1186/1471-2105-12-323>
- Libkind, D., Hittinger, C. T., Valefió, E., Gonçá Ives, C., Dover, J., Johnston, M., Gonçá Ives, F., & Sampaio, J. P. (2011). Microbe domestication and the identification of the wild genetic stock of lager-brewing yeast. *Proceedings of the National Academy of Sciences of the United States of America*, 108(35), Article 35. <https://doi.org/10.1073/pnas.1105430108>

- Liu, L., Wang, J., Rosenberg, D., Zhao, H., Lengyel, G., & Nadel, D. (2018). Fermented beverage and food storage in 13,000 y-old stone mortars at Raqefet Cave, Israel: Investigating Natufian ritual feasting. *Journal of Archaeological Science: Reports*, 21(May), Article May. <https://doi.org/10.1016/j.jasrep.2018.08.008>
- Liu, Z., Zhang, G., & Sun, Y. (2008). Mutagenizing brewing yeast strain for improving fermentation property of beer. *Journal of Bioscience and Bioengineering*, 106(1), Article 1. <https://doi.org/10.1263/jbb.106.33>
- Luo, Z., & van Vuuren, H. J. J. (2008). Stress-induced production, processing and stability of a seripauperin protein, Pau5p, in *Saccharomyces cerevisiae*. *FEMS Yeast Research*, 8(3), 374-385. <https://doi.org/10.1111/j.1567-1364.2008.00355.x>
- Machado, A. K., Brown, N. A., Urban, M., Kanyuka, K., & Hammond-Kosack, K. E. (2018). RNAi as an emerging approach to control *Fusarium* head blight disease and mycotoxin contamination in cereals. *Pest Management Science*, 74(4), 790-799. <https://doi.org/10.1002/ps.4748>
- Madej, M. G., Sun, L., Yan, N., & Kaback, H. R. (2014). Functional architecture of MFS D-glucose transporters. *Proceedings of the National Academy of Sciences of the United States of America*, 111(7), Article 7. <https://doi.org/10.1073/pnas.1400336111>
- Magalhães, F., Vidgren, V., Ruohonen, L., & Gibson, B. (2016). Maltose and maltotriose utilisation by group I strains of the hybrid lager yeast *Saccharomyces pastorianus*. *FEMS Yeast Research*, 16(5), Article 5. <https://doi.org/10.1093/femsyr/fow053>
- Mahanty, B., Mishra, R., & Joshi, R. K. (2023). Cross-kingdom small RNA communication between plants and fungal phytopathogens-recent updates and prospects for future agriculture. *RNA Biology*, 20(1), 109-119. <https://doi.org/10.1080/15476286.2023.2195731>
- Meister, G., & Tuschl, T. (2004). Mechanisms of gene silencing by double-stranded RNA. *Nature*, 431(7006), Article 7006. <https://doi.org/10.1038/nature02873>
- Meneses, F. J., & Jiraneek, V. (2002). Expression patterns of genes and enzymes involved in sugar catabolism in industrial *saccharomyces cerevisiae* strains displaying novel fermentation characteristics. *Journal of the Institute of Brewing*, 108(3), Article 3. <https://doi.org/10.1002/j.2050-0416.2002.tb00557.x>
- Middleton, H., Monard, C., Daburon, V., Clostres, E., Dozois, J., Tremblay, J., Yergeau, É., & Amrani, A. E. (2023). *Plant miRNAs in the rhizosphere target microbial genes*. <https://doi.org/10.21203/rs.3.rs-3363027/v1>
- Middleton, H., Yergeau, É., Monard, C., Combier, J.-P., & El Amrani, A. (2021). Rhizospheric Plant–Microbe Interactions: miRNAs as a Key Mediator. *Trends in Plant Science*, 26(2), 132-141. <https://doi.org/10.1016/j.tplants.2020.09.005>
- Monerawela, C., & Bond, U. (2017). Brewing up a storm: The genomes of lager yeasts and how they evolved. *Biotechnology Advances*, 35(4), Article 4. <https://doi.org/10.1016/j.biotechadv.2017.03.003>
- Monerawela, C., & Bond, U. (2018). The hybrid genomes of *Saccharomyces pastorianus*: A current perspective. *Yeast*, 35(1), Article 1. <https://doi.org/10.1002/yea.3250>
- Monerawela, C., James, T. C., Wolfe, K. H., & Bond, U. (2015). Loss of lager specific genes and subtelomeric regions define two different *Saccharomyces cerevisiae* lineages for *Saccharomyces pastorianus* Group I and II strains. *FEMS yeast research*, 15(2), Article 2. <https://doi.org/10.1093/femsyr/fou008>
- Monteiro, P. T., Oliveira, J., Pais, P., Antunes, M., Palma, M., Cavalheiro, M., Galocha, M., Godinho, C. P., Martins, L. C., Bourbon, N., Mota, M. N., Ribeiro, R. A., Viana, R., Sá-Correia, I., & Teixeira, M. C. (2020). YEASTRACT+: A portal for cross-species comparative genomics of transcription regulation in yeasts. *Nucleic Acids Research*, 48(D1), Article D1. <https://doi.org/10.1093/nar/gkz859>

- Nakao, Y., Kanamori, T., Itoh, T., Kodama, Y., Rainieri, S., Nakamura, N., Shimonaga, T., Hattori, M., & Ashikari, T. (2009). Genome sequence of the lager brewing yeast, an interspecies hybrid. *DNA Research*, *16*(2), Article 2. <https://doi.org/10.1093/dnares/dsp003>
- Nakayashiki, H., Kadotani, N., & Mayama, S. (2006). Evolution and diversification of RNA silencing proteins in fungi. *Journal of Molecular Evolution*, *63*(1), Article 1. <https://doi.org/10.1007/s00239-005-0257-2>
- Nobre, T., Campos, M. D., Lucic-Mercy, E., & Arnholdt-Schmitt, B. (2016). Misannotation Awareness: A Tale of Two Gene-Groups. *Frontiers in Plant Science*, *7*. <https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2016.00868>
- Novačić, A., Vučenović, I., Primig, M., & Stuparević, I. (2020). Non-coding RNAs as cell wall regulators in *Saccharomyces cerevisiae*. *Critical Reviews in Microbiology*, *0*(0), Article 0. <https://doi.org/10.1080/1040841X.2020.1715340>
- Olesen, K., Felding, T., Gjermansen, C., & Hansen, J. (2002). The dynamics of the *Saccharomyces carlsbergensis* brewing yeast transcriptome during a production-scale lager beer fermentation. *FEMS Yeast Research*, *2*(4), Article 4. [https://doi.org/10.1016/S1567-1356\(02\)00155-1](https://doi.org/10.1016/S1567-1356(02)00155-1)
- Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of Phylogenetics and Evolution in R language. *BIOINFORMATICS APPLICATIONS NOTE*, *20*(2), Article 2. <https://doi.org/10.1093/bioinformatics/btg412>
- Parapouli, M., Vasileiadis, A., Afendra, A. S., & Hatziloukas, E. (2020). *Saccharomyces cerevisiae* and its industrial applications. En *AIMS Microbiology* (Vol. 6, Número 1). <https://doi.org/10.3934/microbiol.2020001>
- Peris, D., Langdon, Q. K., Moriarty, R. V., Sylvester, K., Bontrager, M., Charron, G., Leducq, J. B., Landry, C. R., Libkind, D., & Hittinger, C. T. (2016). Complex Ancestries of Lager-Brewing Hybrids Were Shaped by Standing Variation in the Wild Yeast *Saccharomyces eubayanus*. *PLoS Genetics*, *12*(7), Article 7. <https://doi.org/10.1371/journal.pgen.1006155>
- Petruzzi, L., Corbo, M. R., Sinigaglia, M., & Bevilacqua, A. (2016). Brewer's yeast in controlled and uncontrolled fermentations, with a focus on novel, nonconventional, and superior strains. *Food Reviews International*, *32*(4), Article 4. <https://doi.org/10.1080/87559129.2015.1075211>
- Philippe, H., Brinkmann, H., Lavrov, D. V., Littlewood, D. T. J., Manuel, M., Wörheide, G., & Baurain, D. (2011). Resolving Difficult Phylogenetic Questions: Why More Sequences Are Not Enough. *PLOS Biology*, *9*(3), e1000602. <https://doi.org/10.1371/journal.pbio.1000602>
- Pis, J., Polakova, S., Merico, A., & Compagno, C. (2006). *How did Saccharomyces evolve to become a good brewer?* (4). *22*(4), Article 4.
- Qiao, Y., Xia, R., Zhai, J., Hou, Y., Feng, L., Zhai, Y., & Ma, W. (2021). Small RNAs in Plant Immunity and Virulence of Filamentous Pathogens. *Annual Review of Phytopathology*, *59*(1), 265-288. <https://doi.org/10.1146/annurev-phyto-121520-023514>
- Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*(6), Article 6. <https://doi.org/10.1093/bioinformatics/btq033>
- Rabuma, T., Gupta, O. P., & Chhokar, V. (2022). Recent advances and potential applications of cross-kingdom movement of miRNAs in modulating plant's disease response. *RNA Biology*, *19*(1), 519-532. <https://doi.org/10.1080/15476286.2022.2062172>
- Ramos, J., Sychrová, H., & Kschischo, M. (2016). Yeast Membrane Transport. En *Yeast Membrane Transporter* (Vol. 892). <http://link.springer.com/10.1007/978-3-319-25304-6>
- Randez-Gil, F., Corcoles-Saez, I., & Prieto, J. A. (2013). Genetic and phenotypic characteristics of baker's yeast: Relevance to baking. *Annual Review of Food Science and Technology*, *4*(1), Article 1. <https://doi.org/10.1146/annurev-food-030212-182609>

- Raychowdhury, R., Gnirke, A., Fan, L., Yassour, M., Regev, A., Palma, F. di, Thompson, D. A., Chen, Z., Hacohen, N., Nusbaum, C., Lindblad-Toh, K., Levin, J. Z., Haas, B. J., Rhind, N., Grabherr, M. G., Birren, B. W., Amit, I., Mauceli, E., Friedman, N., ... Adiconis, X. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*(7), Article 7. <https://doi.org/10.1038/nbt.1883>
- Rio, D. C., Ares, M., Hannon, G. J., & Nilsen, T. W. (2010). Purification of RNA using TRIzol (TRI reagent). *Cold Spring Harbor protocols*, *2010*(6), Article 6. <https://doi.org/10.1101/pdb.prot5439>
- Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, *26*(1), Article 1. <https://doi.org/10.1093/BIOINFORMATICS/BTP616>
- Ronne, H. (1995). Glucose repression in fungi. *Trends in Genetics*, *11*(1), 12-17. [https://doi.org/10.1016/S0168-9525\(00\)88980-5](https://doi.org/10.1016/S0168-9525(00)88980-5)
- Sakulkoo, W., Osés-Ruiz, M., Garcia, E. O., Soanes, D. M., Littlejohn, G. R., Hacker, C., Correia, A., Valent, B., & Talbot, N. J. (2018). A single fungal MAP kinase controls plant cell-to-cell invasion by the rice blast fungus. *Science*, *359*(6382), Article 6382. <https://doi.org/10.1126/science.aaq0892>
- Salazar, A. N., Vries, A. R. G. D., Broek, M. V. D., Brouwers, N., Cortès, P. D. L. T., Kuijpers, N. G. A., Daran, J. M. G., & Abeel, T. (2019). Chromosome level assembly and comparative genome analysis confirm lager-brewing yeasts originated from a single hybridization. *BMC Genomics*, *20*(1), Article 1. <https://doi.org/10.1186/s12864-019-6263-3>
- Salema-Oom, M., Pinto, V. V., Gonçalves, P., & Spencer-Martins, I. (2005). Maltotriose utilization by industrial *Saccharomyces* strains: Characterization of a new member of the  $\alpha$ -glucoside transporter family. *Applied and Environmental Microbiology*, *71*(9), Article 9. <https://doi.org/10.1128/AEM.71.9.5044-5049.2005>
- Samad, A. F. A., Kamaroddin, M. F., & Sajad, M. (2021). Cross-Kingdom Regulation by Plant microRNAs Provides Novel Insight into Gene Regulation. *Advances in Nutrition*, *12*(1), 197-211. <https://doi.org/10.1093/advances/nmaa095>
- Schaefer, C., & Rost, B. (2012). Predict impact of single amino acid change upon protein structure. *BMC genomics*, *13 Suppl 4*(Suppl 4), Article Suppl 4. <https://doi.org/10.1186/1471-2164-13-S4-S4>
- Schliep, K. P. (2011). phangorn: Phylogenetic analysis in R. *Bioinformatics*, *27*(4), 592-593. <https://doi.org/10.1093/bioinformatics/btq706>
- Selmecki, A. M., Maruvka, Y. E., Richmond, P. A., Guillet, M., Shores, N., Sorenson, A. L., De, S., Kishony, R., Michor, F., Dowell, R., & Pellman, D. (2015). Polyploidy can drive rapid adaptation in yeast. *Nature*, *519*(7543), Article 7543. <https://doi.org/10.1038/nature14187>
- Smit, A., Moses, S. G., Pretorius, I. S., & Otero, R. R. C. (2008). The Thr505 and Ser557 residues of the AGT1-encoded  $\alpha$ -glucoside transporter are critical for maltotriose transport in *Saccharomyces cerevisiae*. *Journal of Applied Microbiology*, *104*(4), Article 4. <https://doi.org/10.1111/j.1365-2672.2007.03671.x>
- Smit, A., Otero, R. R. C., & Pretorius, I. S. (2007). Differences among AGT1-encoded  $\alpha$ -glucoside transporters and their ability to transport maltotriose in *Saccharomyces* yeasts. En *Annals of Microbiology* (Vol. 57, Número 1, pp. 77-84).
- Stambuk, B. U., Alves, S. L., Hollatz, C., & Zastrow, C. R. (2006). Improvement of maltotriose fermentation by *Saccharomyces cerevisiae*. *Letters in Applied Microbiology*, *43*(4), Article 4. <https://doi.org/10.1111/j.1472-765X.2006.01982.x>
- Stambuk, B. U., & Araujo, P. S. D. (2001). Kinetics of active  $\alpha$ -glucoside transport in *Saccharomyces cerevisiae*. *FEMS Yeast Research*, *1*(1), Article 1. [https://doi.org/10.1016/S1567-1356\(01\)00007-1](https://doi.org/10.1016/S1567-1356(01)00007-1)

- Stewart, G. G. (2009). The Horace Brown Medal lecture: Forty years of brewing research. *Journal of the Institute of Brewing*, 115(1), Article 1. <https://doi.org/10.1002/j.2050-0416.2009.tb00340.x>
- Stewart, G. G. (2018). Yeast flocculation—Sedimentation and flotation. *Fermentation*, 4(2), Article 2. <https://doi.org/10.3390/fermentation4020028>
- Tamang, J. P., Watanabe, K., & Holzapfel, W. H. (2016). Review: Diversity of Microorganisms in Global Fermented Foods and Beverages. *Frontiers in Microbiology*, 7, 377. <https://doi.org/10.3389/fmicb.2016.00377>
- Teste, M. A., Duquenne, M., François, J. M., & Parrou, J. L. (2009). Validation of reference genes for quantitative expression analysis by real-time RT-PCR in *Saccharomyces cerevisiae*. *BMC Molecular Biology*, 10, 99. <https://doi.org/10.1186/1471-2199-10-99>
- Thesseling, F. A., Bircham, P. W., Mertens, S., Voordeckers, K., & Verstrepen, K. J. (2019). A Hands-On Guide to Brewing and Analyzing Beer in the Laboratory. *Current Protocols in Microbiology*, 54(1), e91. <https://doi.org/10.1002/cpmc.91>
- Trichez, D., Knychala, M. M., Figueiredo, C. M., Alves, S. L., Silva, M. A. da, Miletti, L. C., Araujo, P. S. de, & Stambuk, B. U. (2019). Key amino acid residues of the AGT1 permease required for maltotriose consumption and fermentation by *Saccharomyces cerevisiae*. *Journal of Applied Microbiology*, 126(2), Article 2. <https://doi.org/10.1111/jam.14161>
- Vidgren, V., Huuskonen, A., Virtanen, H., Ruohonen, L., & Londesborough, J. (2009). Improved fermentation performance of a lager yeast after repair of its AGT1 maltose and maltotriose transporter genes. *Applied and environmental microbiology*, 75(8), Article 8. <https://doi.org/10.1128/AEM.01558-08>
- Vidgren, V., Kankainen, M., Londesborough, J., & Ruohonen, L. (2011). Identification of regulatory elements in the AGT1 promoter of ale and lager strains of brewer's yeast. *Yeast*, 28(8), Article 8. <https://doi.org/10.1002/yea.1888>
- Vidgren, V., & Londesborough, J. (2012). Characterization of the *saccharomycesbayanus*-type *Agt1* transporter of lager yeast. *Journal of the Institute of Brewing*, 118(2), Article 2. <https://doi.org/10.1002/jib.22>
- Vidgren, V., Ruohonen, L., & Londesborough, J. (2005). Characterization and functional analysis of the MAL and MPH Loci for maltose utilization in some ale and lager yeast strains. *Applied and environmental microbiology*, 71(12), Article 12. <https://doi.org/10.1128/AEM.71.12.7846-7857.2005>
- Wang, M., Weiberg, A., Lin, F.-M., Thomma, B. P. H. J., Huang, H.-D., & Jin, H. (2016). Bidirectional cross-kingdom RNAi and fungal uptake of external RNAs confer plant protection. *Nature Plants*, 2(10), 16151. <https://doi.org/10.1038/nplants.2016.151>
- Wang, W., Liu, D., Zhang, X., Chen, D., Cheng, Y., & Shen, F. (2018). Plant MicroRNAs in Cross-Kingdom Regulation of Gene Expression. *International Journal of Molecular Sciences*, 19(7), Article 7. <https://doi.org/10.3390/ijms19072007>
- Wei, X., Zhang, C., Freddolino, P. L., & Zhang, Y. (2020). Detecting Gene Ontology misannotations using taxon-specific rate ratio comparisons. *Bioinformatics*, 36(16), 4383-4388. <https://doi.org/10.1093/bioinformatics/btaa548>
- Weiberg, A., & Jin, H. (2015). Small RNAs—the secret agents in the plant-pathogen interactions. *Current Opinion in Plant Biology*, 26, 87-94. <https://doi.org/10.1016/j.pbi.2015.05.033>
- Weiberg, A., Wang, M., Lin, F.-M., Zhao, H., Zhang, Z., Kaloshian, I., Huang, H.-D., & Jin, H. (2013). Fungal Small RNAs Suppress Plant Immunity by Hijacking Host RNA Interference Pathways. *Science*, 342(6154), 118-123. <https://doi.org/10.1126/science.1239705>
- Yan, N. (2013). Structural advances for the major facilitator superfamily (MFS) transporters. *Trends in Biochemical Sciences*, 38(3), Article 3. <https://doi.org/10.1016/j.tibs.2013.01.003>

- Yoon, S. H., Mukerjea, R., & Robyt, J. F. (2003). Specificity of yeast (*Saccharomyces cerevisiae*) in removing carbohydrates by fermentation. *Carbohydrate Research*, 338(10), Article 10. [https://doi.org/10.1016/S0008-6215\(03\)00097-1](https://doi.org/10.1016/S0008-6215(03)00097-1)
- Yu, G., Wang, L. G., Han, Y., & He, Q. Y. (2012). ClusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS A Journal of Integrative Biology*, 16(5), Article 5. <https://doi.org/10.1089/omi.2011.0118>
- Zastrow, C. R., Mattos, M. A., Hollatz, C., & Stambuk, B. U. (2000). Maltotriose metabolism by *Saccharomyces cerevisiae*. *Biotechnology Letters*, 22(6), Article 6. <https://doi.org/10.1023/A:1005691031880>
- Zeng, Gupta, Jiang, Yang, Gong, & Zhu. (2019). Cross-Kingdom Small RNAs Among Animals, Plants and Microbes. *Cells*, 8(4), Article 4. <https://doi.org/10.3390/cells8040371>
- Zhang, L., Hou, D., Chen, X., Li, D., Zhu, L., Zhang, Y., Li, J., Bian, Z., Liang, X., Cai, X., Yin, Y., Wang, C., Zhang, T., Zhu, D., Zhang, D., Xu, J., Chen, Q., Ba, Y., Liu, J., ... Zhang, C. Y. (2012). Exogenous plant MIR168a specifically targets mammalian LDLRAP1: Evidence of cross-kingdom regulation by microRNA. *Cell Research*, 22(1), Article 1. <https://doi.org/10.1038/cr.2011.158>
- Zhang, T., Zhao, Y. L., Zhao, J. H., Wang, S., Jin, Y., Chen, Z. Q., Fang, Y. Y., Hua, C. L., Ding, S. W., & Guo, H. S. (2016). Cotton plants export microRNAs to inhibit virulence gene expression in a fungal pathogen. *Nature Plants*, 2(10), Article 10. <https://doi.org/10.1038/nplants.2016.153>
- Zhou, G., Zhou, Y., & Chen, X. (2017). New insight into inter-kingdom communication: Horizontal transfer of mobile small RNAs. *Frontiers in Microbiology*, 8(MAY), Article MAY. <https://doi.org/10.3389/fmicb.2017.00768>