





"Explotación del genoma de Issatchenkia terricola para la identificación de glicosidasas con potencial aplicación en enología"

Tesis de Maestría – PEDECIBA Biología Biología Celular y Molecular

Lic. Juliette Dourron

Tutoras: Dra. Ana Ramón Dra. Paula González-Pombo

Tribunal: Dra. Luisa Berná Dra. Gabriela Irazoqui Dr. Andrés Iriarte

> Sección Bioquímica – Facultad de Ciencias Sección Bioquímica – Facultad de Química Universidad de la República

Noviembre 2020

AGRADECIMIENTOS

A mis tutoras, Ana y Paula, por darme la oportunidad de formar parte de este lindo proyecto y equipo de trabajo, del que he aprendido un montón de cosas en todos los ámbitos posibles.

A las chicas del proyecto: Ali, Ana, Andrea, Caro, Paula, Stefi. Gracias a cada una en particular porque de todas he aprendido algo.

A Ana por la confianza en mí, porque eso se convirtió en confianza en mí misma.

A Paula y a Stefi por introducirme en el mundo de las levaduras vínicas, por todo lo enseñado, la ayuda y paciencia.

A Ali, gracias por ser mi mentora informática, por tenerme toda la paciencia del mundo y hacer que me entusiasme y aprenda un montón de algo a lo que le tenía terror.

A Andi, por ayudarme con la parte de expresión de proteínas y por el apoyo moral.

A los integrantes del 313, por hacer que sea el mejor laboratorio y porque siempre que se necesitan están ahí, dispuestos a dar una mano y con la mejor disposición. Manu prometo algún día dejarte en paz.

A la sección Bioquímica de Ciencias, por generar un lindo lugar de trabajo y por los intercambios que siempre aportan, además de la ayuda que he recibido en muchas ocasiones.

A la sección Bioquímica de Química, siempre que me tocó hacer algo por allá el ambiente es precioso y se pasa bárbaro (incluye actividades extraacadémicas), me hubiera gustado pasar por ahí un rato más.

A Mati, por las tardes de estudio, por bancarme la cabeza, por siempre tirar para adelante y por ser tremendo amigo.

A ti, Chuco, gracias por todo el amor siempre.

A Mamá, Papá y Alain, porque sé que sin ellos no estaría donde estoy. Gracias por el apoyo siempre, las palabras quedan escasas para lo agradecida que estoy, los amo.

A mis compañeras y amigas, Paula y Carla, por ser como son y por quererme como soy. Porque vivir con ustedes es de lo más lindo que me ha pasado, y como cada persona que menciono acá son parte de este resultado, gracias.

A mis amigos, por todo, porque la vida sin amigos no sé qué sería.

A los miembros del tribunal, por aceptar corregir este trabajo.

A la ANII y la CAP por financiar este posgrado, ya que sin el apoyo económico no me hubiera sido posible realizarlo.

RESUMEN

El uso de β -glucosidasas en la industria del vino reviste especial interés, ya que estas enzimas pueden promover la liberación de compuestos aromáticos a partir de los precursores glicosídicos no aromáticos presentes en vinos jóvenes, incrementando el aroma de los mostos, vinos y jugos de frutas (Winterhalter & Skouroumounis 1997). Dado que la mayoría de las cepas de *Saccharomyces* no poseen actividad β -glucosidasa, se ha destacado la importancia de las β -glucosidasas de las cepas no-*Saccharomyces*. Además, se ha visto que estas cepas de levaduras nativas, son esenciales para la autenticidad del vino, impartiendo distintas características regionales y otras características deseables (Jolly & Pretorius 2006, Palmeri & Spagna 2007).

En trabajos previos del grupo de la Dra. Paula González-Pombo se aisló una β -glucosidasa (4180) de la cepa nativa *Issatchenkia terricola*, que se destaca por impartir características aromáticas propias a los vinos locales y que presenta propiedades muy promisorias en condiciones enológicas (González-Pombo *et al.* 2011). Con el fin de identificar otros genes codificantes para β -glucosidasas así como para otras actividades enzimáticas con potencial interés biotecnológico, se propuso llevar a cabo la secuenciación masiva del genoma de esta levadura.

Se realizó la secuenciación y anotación del genoma de *l. terricola*. Se logró un ensamblado de tamaño esperado, aproximadamente 13 Mpb, formado por 15 contigs (de los cuales 5 tienen un largo mayor a 1.000.000 pb) y un N50 de 3.383.105 pb. Se predijeron 5.230 genes codificantes para proteínas entre las cuales se identificaron aproximadamente 40 glicosidasas, de las cuales 10 son β -glucosidasas. Uno de éstos genes es el correspondiente al gen de la β -glucosidasa 4180 (González-Pombo *et al.* 2011). No se encontraron otras enzimas de interés enológico, a excepción de una posible arabinosidasa.

Tras un análisis filogenético de las β -glucosidasas identificadas se seleccionaron dos para su estudio (989 y 3296). Para la enzima 989 se han probado diferentes condiciones para su expresión en *Escherichia coli*. Se logró una expresión modesta de la enzima en condiciones insolubles. Queda pendiente realizar la optimización de las condiciones de expresión para logar purificarla en niveles adecuados para llevar a cabo su caracterización bioquímica. En paralelo se realizaron construcciones para el clonado en *Saccharomyces cerevisiae* de ambas enzimas. Los resultados obtenidos hasta el momento son promisorios y constituyen un avance hacia la identificación de β glucosidasas de una cepa de levadura nativa y su potencial aplicación enológica.

ÍNDICE

1. INTR	ODUCCIÓN	7
1.1.	Breve historia del vino	7
1.2.	Flavour, gusto y aroma del vino	7
1.3.	Clasificación de los aromas	8
1.4.	Composición del aroma	9
1.5.	Compuestos glicosilados	10
1.5	5.1. Liberación de la aglicona odorante	12
1.6. E	Enzimas glicosidasas	13
1.6	5.1. β-glucosidasas	13
1.7. l	evaduras vínicas	16
1.7	7.1. Diversidad de levaduras nativas	17
1.7	7.2. Aplicaciones industriales de enzimas glicosidasas	
1.8. <i>1</i>	ssatchenkia terricola	19
1.9 N	luevas tecnologías de secuenciación	20
1.10.	Importancia del estudio genético de levaduras	21
2. OBJE	TIVOS	22
2.1. (Dbjetivo general	22
2.2. (Dbjetivos específicos	22
3. MAT	ERIALES Y MÉTODOS	23
3.1.	Obtención de ácidos nucleicos de <i>I. terricola</i>	23
3.1	I.1. Cepa utilizada	23
3.1	I.2. Medios de cultivo para <i>I. terricola</i>	23
3.1 20	L.3. Ensayo para determinar cuantitativamente actividad β-glucosidasa (de O 17)	valle <i>et al.</i> 24
3.1	I.4. Cultivo para la extracción de ácidos nucleicos	24
3.2. E	Ensamblado y anotación del genoma de <i>I. terrícola</i>	25
3.2	2.1 Secuenciación	25
3.2	2.2. Tratamiento de los datos	25
3.2	2.3. Ensamblado	25
3.2	2.4. Anotación de genes	27
3.2	2.5. Búsqueda de enzimas	28
3.2	2.6. Selección de enzimas para expresar	29
Análi	sis filogenético	29
3.3.1	Fécnicas de rutina de biología molecular	29

	3.3.1. Amplificación de secuencias por PCR	29
	3.3.2. Purificación de fragmentos de ADN	30
	3.3.3. Minipreparación de ADN plasmídico	30
	3.3.4. Visualización de ácidos nucleicos	31
	3.3.5. Análisis de secuenciación de plásmidos	31
	3.3.6. Preparación de células competentes	32
	3.3.7. Medios de cultivo	33
	3.3.8. Transformación de <i>E. coli</i>	33
	3.4. Expresión de proteínas en <i>E. coli</i>	33
	3.4.1. Clonado en vectores plasmídicos mediante Restriction Free Cloning (RF-cloning)	33
	3.4.2. Cultivos para la expresión recombinante de la enzima 989	35
	3.4.3. Evaluación de la expresión y purificación de las proteínas recombinantes	36
	3.4.4. Evaluación de la presencia de la β-glucosidasa recombinante en las distintas fracciones	37
	3.5. Expresión de las enzimas 989 y 3296 en Saccharomyces cerevisiae	38
	3.5.1. Clonado de las enzimas seleccionadas en los cassettes de expresión	40
4.	RESULTADOS Y DISCUSIÓN	43
	4.1. Extracción de ácidos nucleicos	43
	4.2. Ensamblado y anotación del genoma de <i>I. terricola</i>	43
	4.2.1. Tratamiento de los datos del servicio de secuenciación	43
	4.2.2. Ensamblado del genoma	45
	4.2.3. Anotación de genes	48
	4.2.4. Selección de β-glucosidasas para su clonado	56
	4.3. Clonado de las enzimas 989 y 3296 para su expresión recombinante en Escherichia c	oli
		58
	4.3.1. Clonado en vector pGM-T	58
	4.3.2. Clonado en vectores de expresión pET-32a mediante RF- <i>cloning</i>	60
	4.4. Expresión de la enzima 989 en <i>E. coli</i>	62
	4.4.1. Expresión de la proteína 989 fusionada a His-tag y DsbC	64
	4.4.2. Expresión de la proteína 989 sin el péptido de tránsito	68
	4.5. Confección de construcciones génicas para la expresión de las enzimas 989 y 3296 e cerevisiae	n <i>S.</i> 71
	Extracción de ADNg	71
	Construcción de los cassettes de expresión por Fusion-PCR	71
	Clonado de los cassettes en el vector pWJ1042	73
	Amplificación de regiones para dirigir la recombinación homóloga	73
5.	CONCLUSIONES Y PERSPECTIVAS	74

6.	REFERENCIAS	. 78
7.	ANEXO	. 91
	7.1. Materiales y métodos	. 91
	7.1.1. Tablas de cebadores	. 91
	7.1.2. Vectores	. 93
	pET-32a con promotor T7	. 93
	pET-32a con promotor T7	. 94
	7.1.2.2.Clonado en S. cerevisiae	. 94
	7.2. Resultados	. 95

1. INTRODUCCIÓN

1.1. Breve historia del vino

El vino es una bebida alcohólica resultado de la fermentación de la uva (Vitis vinífera). Es la bebida alcohólica más extensamente consumida por la humanidad desde hace miles de años (Fehér et al. 2007). Se han encontrado fósiles de vid que datan de hace aproximadamente 65 millones de años (This et al. 2006). La vid presenta una gran facilidad de adaptación (Levadoux 1956), factor que seguramente ayudó a su expansión. La viticultura data de aproximadamente 7000 a.C. La primera representación del proceso de vinificación encontrada fue realizada por los egipcios alrededor del año 3000 a.C., reproduciendo escenas de prensado y cosecha (Fleming 1996). Los egipcios fueron quienes enseñaron a los griegos a cultivar la viña que se volvió esencial en la agricultura de éstos, quienes además lo establecieron en toda la cuenca del Mediterráneo.

Luego, a partir de aproximadamente el año 100 a.C. los romanos extendieron el cultivo en toda Europa (Levadoux 1956; Fleming 1996; This *et al.* 2006). Con la caída del imperio Romano el cultivo de vid y la producción de vino pasaron a la Iglesia Católica.

A partir de la Edad Media la calidad de los vinos mejora: antes eran cortados con agua y hierbas aromáticas (Fehér *et al.* 2007), pero a partir de esta época se comienza a producir el vino como lo conocemos hoy en día (Fleming 1996). A finales del siglo X comenzó a desarrollarse Burdeos como única región vinícola que no estaba bajo la influencia de la Iglesia.

Con la colonización, la vid se esparció por el resto del mundo y fue a partir del siglo XX que se empieza a desarrollar la enología como ciencia del vino (Fleming 1996; Fehér *et al.* 2007).

1.2. Flavour, gusto y aroma del vino

El *flavour* del vino es una de las características más importantes que determinan la calidad del mismo. El *flavour* es el resultado de la sensación total obtenida por la percepción de los compuestos del aroma (sensados por la nariz y boca) y las sensaciones del gusto o sabor (acidez, dulzura, alcohol, astringencia, etc.) (Marais 1983; Rapp & Mandery 1986; Lambrechts & Pretorius 2000; Styger 2011).

Las sustancias que componen el *flavour* incluyen dentro de los responsables del sabor: ácidos orgánicos, azúcares, compuestos fenólicos y minerales. Dentro de los compuestos que componen el aroma se encuentran: alcoholes, alcoholes superiores, ésteres, norisoprenoides, aldehídos, cetonas, compuestos sulfurados e hidrocarburos, entre los cuales se destacan los terpenos y sus derivados (Schreier & Walter 1979; Rapp & Mandery 1986; Styger 2011).

Los compuestos del gusto en concentraciones por debajo de su umbral de percepción (10⁻³ a 10 g/L) pueden ser percibidos cuando se encuentran formando parte de una mezcla con otros compuestos. Y de esta misma manera, cuando están presentes en

mezclas, pero en concentraciones por encima de su umbral de percepción, pueden reducir o enmascarar la detección de otros compuestos (Francis & Newton 2005).

Los compuestos volátiles pueden ser detectados a concentraciones muy bajas (normalmente entre 10⁻⁴ y 10⁻¹² g/L) y son responsables del aroma percibido orto y retronasalmente (Schreier & Walter 1979; Rapp & Mandery 1986; Francis & Newton 2005). El aroma es asociado puramente a los compuestos odorosos es decir, compuestos volátiles libres. Miles de estos compuestos se encuentran presentes en el vino (Schreier & Walter 1979; Marais 1983; Ortega-Heras 2001; Jeromel *et al.* 2019) pero solo unos pocos contribuyen al aroma. Los compuestos del aroma al igual que los del sabor interactúan entre sí, atenuando o aumentando su percepción como resultado de estas interacciones y de las interacciones con la matriz donde se encuentran (Marais 1983; Francis & Newton 2005; Robinson *et al.* 2009; Villamor & Ross 2013).

1.3. Clasificación de los aromas

Los aromas del vino se clasifican según la etapa del proceso de vinificación en que se generen. Los aromas se dividen en varietales, pre-fermentativos, fermentativos y post-fermentativos (Marais 1983; Rapp & Mandery 1986; Lambrechts & Pretorius 2000).

Los **aromas varietales** son los que provienen de la propia uva y son característicos de cada variedad, aportando tipicidad aromática. Las prácticas vitivinícolas y el clima pueden afectar la cantidad de estos compuestos (Schreier & Walter 1979; Anagnostopoulos *et al.* 2019). Principalmente son compuestos del tipo terpénicos y norisoprenoides y pueden encontrarse en su forma libre, volátil, o en forma glicosilada, como precursores de aroma.

Los **aromas pre-fermentativos** son los que se originan desde el momento de la cosecha hasta el inicio de la fermentación. Durante esta etapa de prensado y estrujado, las enzimas endógenas de la uva catalizan reacciones enzimáticas sobre algunos precursores aromáticos generando compuestos de seis átomos de carbono, a menudo denominados como compuestos C6. La generación de compuestos también incluye las reacciones provocadas por la microbiota nativa de la uva.

Los **aromas fermentativos** se originan durante el propio proceso de fermentación. Este proceso consiste en la metabolización de los azúcares por parte de los microorganismos fermentativos para producir etanol y CO₂. Estos aromas son producto del metabolismo de levaduras durante la fermentación alcohólica, y de bacterias lácticas durante la fermentación maloláctica. En esta etapa ocurre la síntesis de algunos compuestos así como la transformación de estos por la microbiota presente. Dentro de los compuestos químicos generados en esta etapa se encuentran principalmente alcoholes, alcoholes superiores, ácidos grasos, ésteres de acetato y de etilo, lactonas y compuestos azufrados (Boido *et al.* 2003; Carrau *et al.* 2016).

Los **aromas post-fermentativos**, o de añejamiento, se generan durante el proceso de conservación y crianza, debido a reacciones fisicoquímicas que ocurren según el tipo de ambiente donde se realiza, pudiendo ser reductor (botella) u oxidativo (barrica) lo cual llevará a distintos resultados. En este proceso ocurre la disminución de los compuestos volátiles frutales, característicos de vinos jóvenes, para evolucionar hacia aromas más complejos característicos de los vinos añejados (Styger *et al.* 2011).

1.4. Composición del aroma

Los primeros estudios sobre aromas del vino datan de 1942, donde se identificaron unos pocos compuestos usando métodos químicos clásicos. A finales de los 50 se comienza a aplicar la cromatografía de gases para el estudio del aroma, lo cual permite la identificación de una gran cantidad de compuestos (Rapp & Mandery, 1986). Actualmente, el análisis de los compuestos volátiles del aroma se realiza por cromatografía de gases con detección por espectrometría de masas y evaluación por olfatometría y análisis sensorial. Estas técnicas han permitido grandes avances en el estudio de los compuestos del aroma del vino (Ortega-Heras *et al.* 2001; Culleré *et al.* 2004; Gurbuz *et al.* 2006).

Algunos aromas se encuentran presentes en todos los vinos, independientemente de su origen y tipo, y son los que definen el aroma base de los mismos. Dentro de estos compuestos se incluyen alcoholes, alcoholes superiores y sus ésteres y acetatos, ácidos y ésteres de éstos, los mismos se detallan en la siguiente tabla (Figura 1):

Miscellaneous	Fusel alcohols	Organic acids	Isoacids	Organic acid ethyl esters	Fusel alcohol acetates	Ethyl esters of isoacids
Ethanol Diacetyl Acetaldehyde	Isobutanol Isoamyl alcohol Hexanol β-Phenylethanol Methionol	Acetic acid Butyric acid Hexanoic acid Octanoic acid Decanoic acid	Isobutyric acid 2-Methyl butyric acid Isovaleric acid	Ethyl acetate Ethyl butyrate Ethyl hexanoate Ethyl octanoate Ethyl decanoate	Isobutyl acetate Isoamyl acetate β-Phenylethyl acetate	Ethyl isobutyrate Ethyl 2-methyl butyrate Ethyl isovalerate

Figura 1. Compuestos químicos que conforman la base del aroma a vino. Extraído de Ferreira 2010.

Dentro del grupo de compuestos que constituyen el aroma de uvas y vinos, las pirazinas y los terpenos son los que se relacionan con las características varietales del aroma, destacándose principalmente el aporte de estos últimos al perfil aromático de vinos jóvenes (Strauss *et al.* 1986; Rapp & Mandery 1986; Mateo & Jimenez 2000).

Los terpenos son compuestos aromáticos sintetizados como productos del metabolismo de las plantas a partir de acetil-coenzima A (Mateo & Jimenez 2000). Se encuentran presentes en las diferentes variedades de uva, pero lo que hace característica a cada variedad son las concentraciones y combinaciones de los mismos. Se han identificado muchos compuestos terpénicos en uva y vino (Strauss *et al.* 1986; Rapp & Mandery 1986; Mateo & Jimenez 2000). Entre ellos se destacan hidrocarburos monoterpénicos, alcoholes monoterpénicos, numerosos óxidos, aldehídos y dioles terpénicos. En la Figura 2 se pueden apreciar las estructuras de los principales alcoholes monoterpénicos presentes en jugos y vinos.



Figura 2. Principales alcoholes monoterpénicos presentes en jugos y vinos. Extraído de Mateo & Jimenez 2000.

El linalol, el geraniol, el nerol y el citronerol son los más abundantes naturalmente, y junto con sus derivados constituyen la base de muchos aceites esenciales. El linalol y el geraniol comprenden entre un 75-80% del contenido total de terpenos de jugos y vinos (Marais 1983).

Una importante cantidad de estos terpenos se encuentran en forma de glicósidos, constituyendo una reserva potencial de compuestos aromáticos no volátiles (Gunata *et al.* 1985; Hjelmeland & Ebeler 2014).

1.5. Compuestos glicosilados

La estructura de estos compuestos consta de dos partes (Figura 3): una aglicona, correspondiente al compuesto aromático y una parte glucídica. La parte aglicona consiste en monoterpenos, C₁₃-norisoprenoides, derivados de benceno o alcoholes aromáticos. La parte glucídica está compuesta por glucosa o disacáridos (ramnosa-glucosa, arabinosa-glucosa, apiosa-glucosa, glucosa-glucosa) (Gunata *et al.* 1985; Winterhalter & Skouroumounis 1997).



Figura 3. Esquema general de los precursores glicosilados. Extraído y modificado de Winterhalter & Skouroumounis 1997

Como se mencionó anteriormente, muchos de los compuestos (la gran mayoría) que aportan al aroma varietal se encuentran en la uva y el mosto como compuestos glicosilados (Gunata *et al.* 1985; Gunata *et al.* 1990; Mateo & Jimenez 2000; Baumes 2009). Por lo tanto no aportan directamente al aroma, pero constituyen una reserva potencial del mismo.

En la siguiente tabla (Figura 4) se evidencia cómo la concentración de compuestos glicosilados es mayor que la de su forma libre para distintas agliconas en distintas variedades de uvas:

TABLE II
LEVELS OF BOUND AND FREE TERPENOLS AND AROMATIC ALCOHOLS IN VARIOUS GRAPE VARIETIES
Results are given as $\mu g/l$ of juice.

Grape variety	Total terpenols		Geraniol		Linalool		Nerol		a-Terpineol		Citronellol		2-Phenylethanol		Benzyl alcohol	
	Bound	Free	Bound	Free	Bound	Free	Bound	Free	Bound	Free	Bound	Free	Bound	Free	Bound	Free
Muscat of Alexandria	4040	1513	1507	342	1839	1084	618	59	61	21	21	7.5	157	58	109	41
Muscat of Frontignan	1398	1 640	396	107	207	1409	658	74	75	26	62	24	96	25	93	38
Muscat of Hamburg	1047	594	426	241	172	281	318	52	86		45	19.7	48	18	158	42
Muscat Ottonel	2873	1679	1291	172	722	1449	635	35	186	12	39	10.7	265	24	326	66
Gewürztraminer	4325	282	3356	218	22.5	5.6	617	43	183	3.2	146	12	159	37	185	35
Riesling	276	58	65	26	87	19.4	10.3	5.4	114	7.4	-	-	249	49	312	64
Cabernet-Sauvignon	26	5.3	12	3.6	4.2	1.7			9.9	-			88	16	144	38
Carignane	81	7.4	40	4.8	26	2.6	14.8	_	-	-	-	_	102	24	124	38
Cinsaut	314	13	69	13	5.4		6.9	-	233	_	_	_	177	9	135	8
Clairette	105	2.6	34	2.6	9.5	-	4	-	57	-	_	_	129	6	121	12
Grenache	71	11.8	40	5.2	5.4	6.6	8.2		17.4	_	_	-	81	18	160	52
Picpoul	105	5.8	25	5.8	2.2	_	3.0	_	75	_	_		163	38	181	46
Syrah	36	1.7	36	1.7	-	_	_	_	-	_	-	-	93	6	183	8
Terret	96	_	40	_	1.9	_	3.2	_	51	-	_	_	123	5	90	5
Ugni blanc	83	2.8	54	2.6	8.5	0.2	6.9		13.3	-	-	-	94	4	200	10

Figura 4. Comparación entre la concentración de compuestos volátiles libres y su forma conjugada para distintas variedades de uva. Extraído de Gunata et al. 1985.

1.5.1. Liberación de la aglicona odorante

Es de interés, entonces, que en el proceso de elaboración del vino, pueda obtenerse el aroma volátil a partir de los precursores aromáticos. Los precursores aromáticos glicosilados pueden ser hidrolizados durante distintas etapas del proceso de elaboración del vino mediante dos mecanismos por los cuales se da la liberación de la aglicona odorante de los glicósidos: la hidrólisis ácida y la enzimática (Gunata *et al.* 1985; Gunata *et al.* 1990; Mateo & Jimenez 2000; Baumes 2009). Los mismos se describen a continuación:

1.5.1.1. Hidrólisis ácida

La hidrólisis ácida se da naturalmente durante el proceso de añejamiento del vino debido a su pH ácido (2,8 - 3,5). Como este proceso es lento, no se puede apreciar su efecto en vinos jóvenes, ya que éstos salen al mercado previo al proceso de crianza, por lo cual su reserva de aromas permanece como tal (Gunata *et al.* 1985). La hidrólisis ácida puede dar lugar, además, a modificaciones de la aglicona, generando productos no deseados (Figura 5).

Odor descriptio	n Compound(s) responsible for the odor	Literature	Odor description	Compound(s) responsible for the odor	Literature
Green pepper, grassy Green Strawberry	3-isobutyl-2-methoxypyrazine 3-isopropyl-2-methoxypyrazine 3-ethyl-2-methoxypyrazine hexanal, hexenals 2,5-dimethyl-4-methoxy-2,3- dihydro-3-furanone 2,5-dimethyl-4-hydroxy-2,3- dihydro-3-furanone	Bayonove et al. ¹¹ Augustyn et al. ⁴ Drawert ²¹ Rapp et al. ⁵⁸	Rotten eggs Onion-like/ garlic-like Vinegary Sauerkraut Corky	hydrogen sulfide mercaptans ethyl acetate, acetic acid lactic acid, diacetyl + other components methyl tetrahydronaphthalene	Dubois and Rigaud ²⁴
Geranium-like Mousy	2-ethoxy-hexa-3,5-diene 2-ethyl-3,4,5,6-tetrahydro- pyridine 2-acetyl-3,4,5,6-tetrahydro- pyridine 2-acetyl-1,4,5,6-tetrahydro- pyridine	Crowell and Gujmon ²⁰ Tucknott ⁹⁹ Strauss and Heresztyn ⁹³ , Rapp et al., unpubl.	Woody	2,4,6-trichloroanisole sesquiterpenes produced by <i>P.roquefortii</i> 2,3,4,6-tetrachloroanisole and pentachloroanisole 3-methyl-y-octalactone phenols of oak wood	Tanner and Zannier ^{94,95} Heimann, Rapp et al. ³⁹ Tanner et al. ⁹⁵ Kepner et al. ³² Singleton and Noble ⁸

Figura 5. Compuestos no deseados en vinos. Extraído y modificado de Rapp & Mandery 1986.

1.5.1.2. Hidrólisis enzimática secuencial

La hidrólisis enzimática de los glicósidos disacarídicos se da de forma secuencial y es llevada a cabo primero por la acción de disacaridasas (α -L-arabinosidasas, α -Lramnosidasas, β -D-apiosidasas) que remueven la unidad terminal no reductora hidrolizando el enlace ($1\rightarrow 6$) inter azúcar y liberando el azúcar correspondiente (α -Larabinosa, α -L-ramnosa, β -D-apiosa) y los β -D-glucósidos. En una segunda etapa la aglicona es liberada por acción de una enzima β -D-glucosidasa que actúa sobre el β -Dglucósido (Gunata *et al.* 1985; Gunata *et al.* 1988; Sarry & Gunata 2004; Belda *et al.* 2015).

Esta hidrólisis es llevada a cabo en diversas etapas del proceso de vinificación, previo, durante y después de la fermentación. Las enzimas responsables de la misma provienen de la propia uva y de la microbiota asociada a ésta, así como de la presente en el proceso de vinificación (Gunata *et al.* 1990).

1.5.1.3. Hidrólisis ácida vs enzimática

En conclusión, la hidrólisis enzimática presenta la ventaja de ser más eficiente en cuanto a la especificidad de sustrato y la velocidad de la reacción. A diferencia de lo que ocurre en la hidrólisis ácida, la aglicona es liberada sin sufrir modificaciones en su estructura que pueden dar lugar a productos secundarios no deseados, lo cual además le confiere un aroma más natural al vino (Mateo y Jimenez, 2000). Es por esto que la industria vitivinícola se ha interesado en el uso de enzimas en sus preparaciones, buscando explotar el potencial aromático presente en las uvas mediante la liberación de compuestos volátiles a partir de sus precursores glicosilados (Lambrechts y Pretorius, 2000).

1.6. Enzimas glicosidasas

Las glicosidasas son enzimas clave en el metabolismo de carbohidratos. Catalizan la hidrólisis de enlaces O-glicosídicos en carbohidratos y polímeros que los contengan. Son clasificadas en la familia EC 3.2.1, como O-glicosil hidrolasas (Sarry & Gunata 2004; Divakar 2013; Sahay 2019).

Se subdividen en 82 familias basadas en las similitudes de sus secuencias aminoacídicas. En todas éstas se identifican una serie de residuos conservados de histidina, ácido glutámico y ácido aspártico, que se ha propuesto que estarían implicados en la hidrólisis enzimática (Bhatia *et al.* 2002; Sarry & Gunata 2004).

Como se mencionó anteriormente, muchos de los precursores de aroma se encuentran unidos a azúcares mediante enlaces O-glicosídicos, por lo cual se destacan las glicosidasas (ver Hidrólisis enzimática secuencial) para la liberación de la aglicona odorante.

1.6.1. β-glucosidasas

Dentro de las glicosidasas utilizadas en la elaboración del vino se destacan las β -glucosidasas que pertenecen a la familia EC 3.2.1.21. Esta enzima es la responsable del

último paso en la liberación de agliconas, hidrolizando la unión enlace (β -1) entre la molécula de glucosa y el sustituyente sobre el carbono anomérico (Villena *et al.* 2007; González-Pombo *et al.* 2011; Sahay 2019). La actividad de esta enzima es fundamental ya que por más que se encuentren presentes las otras glicosidasas que participan de la hidrólisis secuencial, si no hay actividad β -glucosidasa, no se liberará la aglicona odorante.

En la Figura 6 se puede apreciar el mecanismo de hidrólisis propuesto para las β glucosidasas. Una vez que ocurre la unión inicial del sustrato a la enzima, mediante catálisis ácida se da el ataque nucleofílico al carbono anomérico, formándose así el intermediario glicósido-enzima. La hidrólisis del intermediario se da por catálisis básica, donde una molécula de agua ataca al carbono anomérico. El resultado final son la glucosa y la aglicona como productos y la enzima restaurada (Kempton & Withers 1992; Sanz-Aparicio *et al.* 1998).



Figura 6. Mecanismo propuesto para la hidrólisis mediada por β-glucosidasas. Extraído de Kempton & Withers 1992.

1.6.1.1. Estructura

A la fecha muy pocas estructuras se encuentran disponibles para esta familia de enzimas, pero estudios sugieren que pertenecen a una superfamilia de proteínas con estructuras de barriles $(\alpha/\beta)_8$ con similitudes en los residuos de aminoácidos de sus sitios activos. En la Figura 7 se muestra la estructura cristalográfica de BglA, una β -glucosidasa de *Bacillus polymyxa*. En la misma se puede observar una subunidad de la enzima octamérica dispuesta en forma de barril $(\alpha/\beta)_8$ (Sanz-Aparicio *et al.* 1998).



Figura 7. Estructura de BglA de B. polymyxa con un ligando en su sitio de unión. Se representan con color rosado las hélices alfa, en azul las hojas β y los loops en cyan. En el centro se observa un gluconato con sus átomos de carbono en blanco y átomos de oxígeno en rojo. Extraído de Sanz-Aparicio et al. 1998.

1.6.1.2. Presencia de 6-glucosidasas en la naturaleza

La hidrólisis de glucósidos es parte esencial de muchos procesos metabólicos en distintos tipos de organismos, por lo cual podemos hallar este tipo de enzimas tanto en bacterias como en hongos, plantas, insectos y mamíferos (Sanz-Aparicio *et al.* 1998). Estas enzimas presentan diferencias en la especificidad de sustrato y en su origen evolutivo que reflejan su participación en diversos y distintos procesos biológicos. Las secuencias de estas enzimas, a pesar de sus diversos orígenes comparten entre sí un alto grado de homología.

Relacionado con el vino se han reportado β -glucosidasas en la planta de la uva, en bacterias del género *Oenococcus*, en diversos géneros de levaduras (*Saccharomyces* y no-*Saccharomyces*) y en hongos filamentosos. Se ha observado que las enzimas provenientes de la propia uva no toleran el pH ácido, ni las concentraciones de etanol y glucosa presentes en el mosto y en el vino (Winterhalter & Skouroumounis 1997; Mateo & Jimenez 2000). Entre las bacterias productoras de β -glucosidasas se destaca *Oenococcus oeni* que participa de la fermentación malolática.

Muchas de las preparaciones comerciales de β -glucosidasas son obtenidas de *Aspergillus niger*. Este hongo filamentoso también presenta actividad arabinosidasa y ramnosidasa, pero además puede ser fuente de actividades adicionales que pueden ocasionar transformaciones no deseadas de los compuestos presentes en el vino. Una forma de evitar este problema es la purificación de enzimas de estos preparados. Se ha

visto que β -glucosidasas aisladas de preparaciones comerciales tienen buenos resultados en la liberación de aromas (Barbagallo *et al.* 2004; Sarry & Gunata 2004; Andrades *et al.* 2019), pero esta práctica tiene la desventaja de aumentar los costos del proceso de producción.

En cuanto a las β -glucosidasas de levaduras se han realizado numerosos estudios de búsqueda de estas enzimas en diferentes cepas. Éstas están presentes en los géneros *Hanseniaspora, Pichia, Candida, Saccharomycodes, Metschinkowia, Zygosaccharomyces y Brettanomyces,* entre otros, por lo cual se ha centrado la atención en estos grupos de microoganismos para el estudio de las β -glucosidasas (Villena *et al.* 2007; Wanapu *et al.* 2012).

1.7. Levaduras vínicas

En sus inicios, el vino era producido por fermentación espontánea, es decir, a partir de la fermentación realizada por la microbiota proveniente de las uvas y el equipamiento de la bodega, lo cual se conoce como microbiota nativa o indígena (Querol & Ramón 1996; Pretorius 2000; Vontrobova *et al.* 2019). Fue en 1863 que Luis Pasteur descubrió que las levaduras eran las responsables del proceso de vinificación (Pretorius 2000). Las concentraciones de levaduras en un mosto fresco son típicamente de $10^3 - 10^5$ ufc/mL (Lambrechts & Pretorius 2000; Vontrobova *et al.* 2019). Solo algunos de los microorganismos presentes en el jugo son capaces de crecer en él (son varios los factores que afectan su crecimiento: inhibición por etanol, temperatura de la fermentación, pH del vino, falta de O₂, alta concentración de azúcares del mosto, etc.) y muy pocos son capaces de competir con las levaduras fermentativas (levaduras *Saccharomyces*, donde predomina *Saccharomyces cerevisiae*), que dominan la fermentación alcohólica (Henick-Kling 1988; Querol & Ramón 1996).

En las fermentaciones espontáneas ocurren dos fermentaciones sucesivas (o en ocasiones simultáneas): la fermentación primaria o fermentación alcohólica donde las levaduras convierten el azúcar del mosto en etanol y CO₂. Le sigue una fermentación secundaria, la fermentación maloláctica, donde las bacterias acidolácticas convierten el ácido málico en ácido láctico y CO₂ (Henick-Kling 1988; Gambaro *et al.* 2001; Boido *et al.* 2002) (Figura 8). Esta última es más frecuente en vinos tintos y no siempre tiene lugar.

Durante los primeros 3-4 días de fermentación predominan las levaduras indígenas. Luego, a raíz del aumento en la concentración de etanol producida por la fermentación, comienza a predominar *S. cerevisiae* (Querol & Ramón 1996) (Figura 8).



Figura 8. Composición de microorganismos durante la fermentación del mosto. A- Levaduras no-Saccharomyces; B- Levaduras Saccharomyces; C- Bacterias acidolácticas; D- Bacterias o levaduras que deterioran la calidad el vino. Extraído de Vontrobova et al. 2019.

1.7.1. Diversidad de levaduras nativas

Las poblaciones nativas imparten características particulares a los vinos (Vontrobova *et al.* 2019) con la desventaja que su crecimiento no controlado, así como el de otros microorganismos, puede afectar negativamente la calidad del vino (Henick-Kling 1988). Son muchos los factores que influyen en la composición de la población de levaduras. Estos factores están determinados por las condiciones climatológicas de cultivo (suelo, horas de sol, humedad, etc.), la variedad de uva, la maduración del grano al momento de cosecha, las prácticas agrícolas (fungicidas, daños en las uvas, etc.) y la higiene de los equipos y materiales utilizados, entre otros, y terminan afectando la calidad el producto final (Querol & Ramón 1996). El uso combinado de especies por lo general resulta en un comportamiento impredecible. Es por esto que en la mayoría de las regiones productoras de vino, desde mediados del siglo 20 se inocula el mosto con cepas de levaduras seleccionadas para asegurar que la fermentación no se enlentezca ni se detenga. Esta práctica permite además lograr reproducibilidad entre las distintas producciones y asegurar por lo tanto un producto uniforme y de calidad.

S. cerevisiae es la cepa fermentadora por excelencia. Sin embargo, en general ésta carece de las actividades enzimáticas que promuevan la hidrólisis de los precursores de aromas, que, en cambio, sí están presentes en diferentes especies de levaduras nativas no-*Saccharomyces*.

La población de levaduras nativas se divide en dos grandes grupos: levaduras *Saccharomyces* y no-*Saccharomyces*. Entre las levaduras no-*Saccharomyces* se encuentran los siguientes géneros: *Kloeckera, Hanseniaspora* (siendo estas dos las mayoritarias, representan entre un 50-75% de la población (Querol & Ramón 1996; Giorello et al. 2018)), *Metschnikowia, Pichia, Schizozaccharomyces, Candida, Cryptococcus, Rhodotorula Kluyveromyces, Hansenula, Debaryomyces, Torulaspora*,

Brettanomyces, Zygosaccharomyces y Saccharomycodes, Issatchenkia, entre otros (Henick-Kling 1988; Querol & Ramón 1996; Sánchez-Torres *et al.* 1998; Pretorius 2000; Strauss *et al.* 2001; Clemente-Jimenez *et al.* 2003; Romano, Ciani & Fleet 2019; Vontrobova *et al.* 2019; Vilela 2020). Como se mencionó antes, estas levaduras son de interés porque a diferencia de las levaduras *Saccharomyces*, producen y secretan varias enzimas (estearasas, glicosidasas, lipasas, β -glucosidasas, etc.) al espacio periplásmico y al medio, donde pueden interactuar con precursores de aroma impartiendo características particulares, en general beneficiosas, al aroma varietal (Lambrechts & Pretorius 2000; Strauss *et al.* 2001; Maicas et Mateo 2005; Carrau *et al.* 2016; Belda *et al.* 2017; Sun *et al.* 2018; Jeromel *et al.* 2019; Morata *et al.* 2019; Romano, Ciani & Fleet 2019; Roudil *et al.* 2019; Vilela 2020).

Últimamente, los consumidores demandan vinos de mayor complejidad y características sensoriales distintivas (Swiegers *et al.* 2005). De hecho, en 2016 la producción global de vino fue de 271 billones de litro y el consumo fue de 241 billones, lo cual indica un exceso en la oferta que hace que la competencia en el mercado sea aún mayor (Organización Internacional de la Vid y el Vino, OIV: <u>http://www.oiv.int/es/statistiques/recherche</u>, consultado en octubre de 2020). En respuesta a una demanda cada vez más exigente y buscando tipicidad aromática, se están explorando otras herramientas como son las co-fermentaciones de *Saccharomyces* junto con levaduras no-*Saccharomyces* (Swiegers & Pretorius 2005; Jolly *et al.* 2006; González-Pombo *et al.* 2011; Padilla *et al.* 2016; de Ovalle *et al.* 2017; Morata *et al.* 2019; Vontrobova *et al.* 2019), así como la aplicación durante la vinificación de enzimas producidas por estos microorganismos (Ottone *et al.* 2019). Sin embargo, aún queda mucho por estudiar en relación a las potenciales aplicaciones industriales de las levaduras no-*Saccharomyces* (Gamero 2016; Padilla *et al.* 2019).

1.7.2. Aplicaciones industriales de enzimas glicosidasas

1.7.2.1. Co-fermentaciones

Se ha estudiado el uso de cultivos puros de levaduras no-*Saccharomyces* para la fermentación, pero por lo general éstas tienen baja capacidad fermentativa, por lo cual no son capaces de completar la fermentación. Además, se ha observado que producen altos niveles de compuestos no deseados, como ácido acético, acetaldehídos, acetona y etil acetato. Le estrategia de elección es, en consecuencia, realizar co-fermentaciones de cepas *Saccharomyces* y no-*Saccharomyces* (Ciani *et al.* 2010). Son varios los trabajos que demuestran que cuando se realizan vinificaciones con cultivos mixtos de *Saccharomyces* y no-*Saccharomyces*, la composición de volátiles del vino resultante es mejor frente a cultivos exclusivos de *S. cerevisiae* comercial (Gil *et al.* 1996; Esteve-Zarzoso *et al.* 1998; Andorra 2010; Andorra *et al.* 2012; Ciani & Comitini 2015; Romano, Ciani & Fleet 2019; Li *et al.* 2020; Vilela 2020). Son claves en el proceso de inoculación mixta la selección de las levaduras más adecuadas y el control de las condiciones de

fermentación para lograr buenos vinos con características distintivas (Andorra 2010; Wanapu *et al.* 2012; Roudil *et al.* 2019) y reproducibilidad entre las producciones.

1.7.2.2. Uso de preparaciones comerciales

Como una estrategia alternativa a la co-fermentación también se suelen utilizar preparados enzimáticos comerciales durante el proceso de vinificación. Como mencioné anteriormente (ver Presencia de β -glucosidasas en la naturaleza), en el caso particular de las glicosidasas, se ha extendido el uso de preparados de pectinasas generalmente provenientes de *Aspergillus niger* que contienen actividad glicosidasa (Mateo & Jimenez 2000). La utilización de estas enzimas tiene como resultado un aumento de los niveles de compuestos volátiles, monoterpenos, norisoprenoides y derivados bencénicos (Cabaroglu *et al.* 2003). Sin embargo, muchos de estos preparados no fueron específicamente diseñados para su aplicación en vino y pueden tener actividades secundarias que resulten en detrimento del vino, además de que pueden presentar problemas de inhibición en las condiciones de elaboración del mismo (Ottone *et al.* 2019).

Otra desventaja del uso de preparados comerciales y la principal razón para la búsqueda de alternativas, es que la oferta de estos preparados es limitada en el mercado. Por lo tanto, las opciones reducidas provocan que el uso de una misma preparación por parte de distintos productores otorgue cierta homogeneidad entre los productos finales, perdiendo tipicidad, que es lo que se busca a la hora de elaborar un vino de calidad y original.

1.8. Issatchenkia terricola

Como se ha mencionado, son de interés en la industria enológica las cepas de levaduras no-*Saccharomyces* y las enzimas capaces de actuar sobre sustratos glicosídicos producidas por éstas, de forma de permitir la liberación de compuestos volátiles durante la elaboración de vinos. Se destacan especialmente las β -glucosidasas ya que son las enzimas responsables del último paso en la liberación de agliconas glicosiladas (Villena *et al.* 2007) y que la mayoría de los compuestos glicosilados se encuentran como β -D-glucósidos.

Como antecedente, nuestro grupo de trabajo, y en la línea de investigación que dirige la Dra. González-Pombo desde hace varios años, se enfoca hacia la búsqueda y estudio de biocatalizadores de glicosidasas de cepas nativas de aplicación biotecnológica. A partir de la colección de cepas de Enología de Facultad de Química, se explora la diversidad de nuestro patrimonio para seleccionar enzimas con propiedades adecuadas para su aplicación biotecnológica. Las β -glucosidasas tienen numerosas aplicaciones biotecnológicas, desde conversión de biomasa para la producción de biocombustible (Srivastava *et al.* 2019), degradación de materiales lignocelulósicos (Sørensen *et al.* 2013) hasta su uso en la industria alimenticia para el mejorarmiento de aromas en jugos y bebidas alcohólicas (Pogorzelski & Wilkowska 2007). En este sentido, se han desarrollado nuevos biocatalizadores de β -glucosidasas a partir de cepas de esta colección (*Metschnikowia pulcherrima, Issatchenkia orientalis e Issatchenkia terricola*)

(González-Pombo *et al.* 2008; Gonzalez-Pombo 2010; González-Pombo *et al.* 2011). Como resultado de estos estudios se destacó la cepa *I. terricola*, por presentar una β -glucosidasa con características prometedoras en cuanto a su estabilidad en condiciones enológicas así como la actividad y especificidad aromática observada en estudios de aromatización de vinos jóvenes (González-Pombo *et al.* 2011; de Ovalle *et al.* 2017). Dado que los bajos rendimientos de producción de esta enzima por parte *I. terricola* limitan la posibilidad de su aplicación biotecnológica, se intentó mejorar el rendimiento mediante el escalado del volumen de cultivo y el ensayo de distintos medios de cultivo. Lamentablemente, los resultados obtenidos no fueron satisfactorios (de Ovalle *et al.* 2017). Sin embargo, el potencial biotecnológico de esta enzima nos llevó a plantear un abordaje molecular para la búsqueda de alternativas que hagan posible la aplicación enológica de esta β -glucosidasa.

1.9 Nuevas tecnologías de secuenciación

La secuenciación de ácidos nucleicos permite la determinación del orden exacto de nucleótidos presentes en una secuencia de una molécula de ADN o ARN. En los últimos años el uso de la secuenciación ha crecido exponencialmente. Esto se debe al desarrollo de nuevos métodos de secuencación conocidos como métodos de próxima generación (NGS, Next-Generation Sequencing).

Desde que surgió en 1977 el método de *Sanger* era el utilizado para secuenciar. En 2005 se lanza la primera plataforma de secuenciación masiva comenzando una nueva era de análisis genómico. Las plataformas de los NGS realizan secuenciación masiva paralela, en la cual millones de fragmentos de una muestra son secuenciados al mismo tiempo. Estos métodos permiten que un genoma entero sea secuenciado en menos de un día y tienen un costo menor en términos de tiempo y dinero, haciéndolo accesible para más laboratorios y mejorarndo por tanto la cantidad de investigación y diagnóstico a partir de la utilización de estas tecnologías (Voelkerding *et al.* 2009; Grada & Weinbrecth 2013).

Las primeras tecnologías en surgir luego de *Sanger* (que se consideran de segunda generación) tienen la limitación de que sus lecturas cortas dificultan la resolución de determinados problemas biológicos, como es el ensamblado *de novo* de genomas.

Sin embargo, en los últimos años se han desarrollado otras tecnologías como *Oxford Nanopore* y *Pacific Bioscience (PacBio)*, que se conocen como de tercera generación, ya que se distinguen de las tecnologías de segunda generación. *PacBio* se distingue por realizar la secuenciación a tiempo real a nivel de moléculas únicas, dando lecturas largas, lo que permite por ejemplo, el estudio de regiones extendidas de repetidos, entre otras aplicaciones. Éste método tiene menor rendimiento y una mayor tasa de errores que las tecnologías de segunda generación. Dado que las ventajas de ambos tipos de NGS se complementan se propone una estrategia híbrida que utiliza ambas para sobrepasar las desventajas de cada una. En el caso del ensamblado *de novo* las lecturas largas de *PacBio* proporcionan *scaffolds* largos mientras que su tasa de error es corregida con lecturas cortas de segunda generación que poseen menor tasa de errores (Rhoads & Fai 2005; Voelkerding *et al.* 2009; Grada & Weinbrecth 2013).

1.10. Importancia del estudio genético de levaduras

En el mercado del vino, cada día más exigente, es ventajoso recurrir a estrategias de mejora genética de cepas previamente seleccionadas y con antecedentes de aplicación eficiente (Sánchez-Torres *et al.* 1998). Se han intentado realizar mejoras génicas mediante métodos clásicos, pero actualmente se cuenta con herramientas que permiten aislar genes concretos, modificarlos y re introducirlos en el organismo de origen o en otro hospedero (OMG, organismos modificados genéticamente). Si bien el uso comercial de microorganismos genéticamente modificados genera controversias cada vez es mayor su aceptación. Existen en la actualidad levaduras modificadas genéticamente, reconocidas como seguras a nivel alimentario (del inglés GRAS, *Generally Recognized as Safe*) por la FDA (*Food and Drug Administration*) disponibles en el mercado (Husnik *et al.* 2007; Tobergte & Curtis 2013).

Varios estudios proponen como alternativa la modificación genética de S. cerevisiae (Bisson & Karpel 2010), fermentadora por excelencia, para otorgarle características deseables de otras levaduras, como puede ser la incorporación de un gen de una glicosidasa. Esto permitiría promover la liberación de aromas a partir de precursores, a la vez que se asegura una fermentación completa de forma exitosa. Varios genes de glicosidasas de bacterias y hongos (incluido S. cerevisiae) han sido clonados (Perez-González-Pombo et al. 1993; Margolles-Clark et al. 1996; Sánchez-Torres et al. 1998; Manzanares et al. 2003; Kitagawa et al. 2010; Chang et al. 2011; Zietsman et al. 2011; Aftab et al. 2012; Delgado et al. 2020; Gaya et al. 2020) con la intención de producir enzimas a gran escala para diversas aplicaciones biotecnológicas. Por otra parte, la ingeniería genética podría constituir una herramienta para superar las limitaciones de la aplicación práctica de estas enzimas debido a que su actividad, por lo general, en condiciones enológicas es baja. Las enzimas de interés deberían ser capaces de funcionar de manera eficiente en las condiciones de elaboración del vino, esto quiere decir, en concentraciones altas de azúcar en el mosto, en un pH de trabajo de 3,0 – 4,0 y en concentraciones por encima del 10% de etanol, además de presentar tolerancia a la inhibición por SO₂ y taninos.

Las levaduras no-*Saccharomyces* se destacan como fuente de enzimas capaces de impartir características deseables en la producción de aromas en el vino. Sin embargo, la mayoría de los genomas de estas especies no se encuentran secuenciados, lo que constituye un impedimento para la búsqueda e identificación de estas enzimas, ya sea para producirlas de forma recombinante o para insertar esos genes en *S. cerevisiae* (Carrau *et al.* 2016; Giorello *et al.* 2018; Romano, Ciani & Fleet 2019). Hasta el momento a nivel internacional son pocas las cepas vínicas estudiadas desde un punto de vista genético (Woolfit *et al.* 2007; Borneman *et al.* 2008; Sternes *et al.* 2016; Seixas *et al.* 2018; Guaragnella *et al.* 2020).

Dentro de las levaduras estudiadas a nivel genómico, en el trabajo de Douglass y colaboradores (Douglass et al. 2018) logran obtener un ensamblado de buena calidad,

resuelto a nivel cromosómico, para la levadura *Pichia Kudriavzevii*, conocida también como *Issatchenkia orientalis*. Esta es la levadura más emparentada con *I. terricola* para la cual se cuenta con genomas disponibles, para así tener un organismo con el cual realizar comparaciones a falta de un genoma de referencia.

Actualmente en nuestro país es escaso el conocimiento disponible en cuanto a genomas de cepas nativas, hasta la fecha se ha caracterizado únicamente el genoma de *Hanseniaspora vinae* (Giorello *et al.* 2014; Giorello *et al.* 2018). La carencia de información genética sobre estas cepas resulta en que muchas enzimas permanezcan inexploradas hasta el momento. A la fecha de inicio de este trabajo no se contaba con datos de secuenciación para *I. terricola*. Actualmente se encuentra disponible un solo genoma reportado (Shen *et al.* 2018), el cual se hizo en el marco de un proyecto de secuenciación de múltiples levaduras utlizando tecnología *Illumina*, generando un genoma en calidad de borrador.

En este trabajo de tesis, se llevó a cabo la secuenciación del genoma de *I. terricola* como punto de partida para la identificación de genes de β -glucosidasas y de otras enzimas de valor enológico, potencialmente aplicables en la industria vitivinícola para aportar a la calidad e identidad de los vinos nacionales. Esto permitirá llevar a cabo abordajes moleculares para la producción de estas enzimas en sistemas de expresión heterólogos, así como también su caracterización bioquímica. A estos aspectos aplicados, se suma asimismo el aporte al conocimiento desde un punto de vista básico de la biología de una cepa de levadura nativa.

2. OBJETIVOS

2.1. Objetivo general

Obtención del genoma de la levadura nativa *Issatchenkia terricola* y su explotación para la identificación de enzimas glicosidasas de interés enológico.

2.2. Objetivos específicos

- Ensamblado *de novo* del genoma y anotación de genes de la levadura nativa *I. terricola*
- \bullet Identificación en este genoma de enzimas glicosidasas, en particular de $\beta\text{-}$ glucosidasas
- Clonado de β-glucosidasas seleccionadas en vectores de expresión para *E. coli*
- Expresión recombinante de la β-glucosidasa 989 en *E. coli*
- Generación de construcciones génicas para la expresión de las enzimas seleccionadas en *S. cerevisiae*

3. MATERIALES Y MÉTODOS

3. 1. Obtención de ácidos nucleicos de I. terricola

3.1.1. Cepa utilizada

La cepa de *I. terricola* utilizada en este trabajo corresponde al número de referencia 0621 de la colección de la Sección Enología de Facultad de Química, Montevideo, Uruguay. Fue aislada de uvas Tannat de viñedos uruguayos en 2006 (González-Pombo *et al.* 2011).

3.1.2. Medios de cultivo para I. terricola

YPD (Yeast extract Peptone Dextrose)

Peptona 1 g/L, extracto de levadura 25 g/L, glicerol 8 mL/L. Se lleva a volumen final en buffer fosfato citrato 50 mM pH = 5,0 y se esteriliza por calor húmedo a 121°C durante 20 minutos.

<u>Simil mosto</u>

Glucosa 100 g/L, fructosa 100 g/L, NH₄Cl 460 mg/L. Se lleva a volumen final con agua, se ajusta el pH a 3,5 y se esteriliza por calor húmedo a 121°C durante 20 minutos. Luego de esterilizado se agrega la solución de aminoácidos.

Solución de aminoácidos 100X (100 mL):

Prolina 205 mg, L-glutamina 169 mg, L-arginina 12,5 mg, L-triptófano 60 mg, L-alanina 49 mg,

L-Ácido glutámico 40 mg, L-serina 26 mg, L-treonina 26 mg, L-leucina 16 mg, L- Ácido aspártico 15 mg, L-valina 15 mg, L-fenilalanina 13 mg, L-isoleucina 11 mg, L-histidina 11 mg, L-metionina 11 mg, L-tirosina 6 mg, L-glicina 6 mg, L-lisina 6 mg, L-cisteína 4 mg. Esterilización por filtración con filtro de 0,2 μm.

Medio sintético

Se prepara una solución 10 g/L de glucosa en buffer citrato pH 5, 50 mM. Se esteriliza por calor húmedo a 121°C durante 20 minutos. Luego de esterilizado se agrega la fuente de nitrógeno (que en este caso va a ser urea), y las soluciones de vitaminas y elementos traza.

Soluciones:

Solución de urea 10X: Se prepara una solución de urea 40g/L y se esteriliza por filtración.

<u>Solución de vitaminas 100X</u>: ácido fólico 600 mg/L, mioinositol 600 mg/L, d-biotina 600 mg/L, pantotenato de calcio 80 g/L, ácido p-aminobenzoico 80 g/L, riboflavina 80 g/L, piridoxina 160 g/L. Se esteriliza por filtración con filtro de 0,2 μm.

Solución de elementos traza:

i) Solución A (aniones) 100X

Na₂MoO₄.2H₂O 65 g/L, KI 10 g/L, H₃BO₃ 10 g/L. Se justa el pH a 1,5 y se esteriliza por filtración con filtro de 0,2 μ m.

ii) Solución C (cationes) 50X

Ácido cítrico 30 g/L, CoCl₂ 7,5 g/L, MnSO₄.H₂O 150 g/L, ZnSO₄.7H₂O 250 g/L, FeSO₄.7H₂O 750 g/L, CuSO₄.5H₂O 75 g/L. Se justa el pH a 1,5 y se esteriliza por filtración con filtro de 0,2 μ m.

3.1.3. Ensayo para determinar cuantitativamente actividad β -glucosidasa (de Ovalle *et al.* 2017)

El ensayo se realiza utilizando el sustrato artificial 4-nitrofenil β -D-glucopiranósido (pNPG), midiendo la aparición de 4-nitrofenol (pNP) a 405 nm. Se tomó un 1 mL del cultivo de la levadura en condiciones asépticas y se transfirió a un tubo *Eppendorf*. Se centrifugó el mismo durante 1 minuto a 5000 g y se determinó actividad β -glucosidasa al sobrenadante. La muestra de 500 µl de sobrenadante se incubó con 1,25 mM del sustrato preparado en buffer de actividad (acetato de sodio 0.1 M, pH 4.5) y se tomaron alícuotas a diferentes tiempos, durante 30 minutos y temperatura ambiente (23°C). La reacción se detuvo mediante la adición de buffer carbonato-bicarbonato 0.2 M, pH 10. A continuación se midió la absorbancia a 405 nm. Se realizó el blanco de igual forma que la mezcla reactiva pero en vez de utilizar el sobrenadante del cultivo se utilizó medio estéril.

Una unidad de enzima se define como la cantidad de ésta necesaria para liberar 1 μ mol de pNP por minuto a temperatura ambiente (23°C) y pH 4,5.

Se calculó la velocidad de la reacción utilizando la pendiente del gráfico de Absorbancia405nm vs tiempo (minutos) con la siguiente fórmula:

velocidad (uM.min⁻¹) =
$$\frac{\text{pendiente del gráfico (min^{-1})}}{\epsilon_{pNP} (\mu M^{-1} \text{cm}^{-1}). b (\text{cm}^{-1})}$$

Se calcularon las UE/mL teniendo en cuenta los volúmenes de reacción y los factores de dilución mediante la siguiente fórmula:

$$\frac{\text{UE}}{\text{mL}} = \frac{\text{velocidad .volumen de reacción (mL). factor de dilución}}{\text{volumen de enzima (mL)}}$$

3.1.4. Cultivo para la extracción de ácidos nucleicos

A partir de colonias aisladas de levaduras de placas de YPD agar crecidas durante 48hs a 28°C, se realiza un pre-cultivo de 20 mL de YPD, durante 72 hs a 28°C y 100 rpm. 200 mL de YPD fueron inoculados con el pre-cultivo para obtener una DO600 de ~0,2. Se realizó el monitoreo diario de DO₆₀₀ y de actividad β -glucosidasa (ver ensayo de actividad β -glucosidasa).

3.1.4.1. Extracción de ADN

El ADN fue extraído utilizando el kit *Quick-DNA™ Fungal/Bacterial Miniprep Kit Catalog No. D6005 (Zymo Research)*. Las células fueron lisadas durante 2 minutos con el equipo *Mini-Beadbeater-16 (Biospec products)*. Se evaluó la concentración y calidad del ADN por Nanodrop (*NanoDrop Lite Spectrophotometer, Thermo Scientific*) y gel de agarosa 1%.

3.1.4.2. Extracción de ARN

Se creció la levadura en dos medios distintos, para obtener ARN mensajero en distintas condiciones: en medio Simil Mosto (por ser el más parecido al mosto a partir del cual se realiza la vinificación) y en medio Sintético (por ser un medio en que *l. terricola* presenta un buen crecimiento y una buena actividad β -glucosidasa) (de Ovalle *et al.* 2017). La extracción de ARN fue realizada el día de mayor actividad β -glucosidasa (ver ensayo de actividad β -glucosidasa).

El ARN total fue extraído utilizando el kit *Quick-RNA™ Miniprep (Zymo Research*). Se sumó un paso adicional de disrupción celular (mismas condiciones que para la extracción de ADN) después del agregado del buffer de lisis para mejorar la calidad y la cantidad de ARN extraído.

Se evaluó la calidad del ARN por Nanodrop y migración en gel de agarosa 1%.

3.2. Ensamblado y anotación del genoma de I. terrícola

3.2.1 Secuenciación

Las construcciones de librerías y secuenciación fueron realizadas en *Macrogen* (Corea), siguiendo protocolos estándar. El ADN genómico fue secuenciado siguiendo una estrategia híbrida (lecturas de *Pacific Bioscience (PacBio*) e *Illumina*). Las librerías de *Illumina* fueron construidas usando *TruSeq DNA Sample Prep Kit* y secuenciadas en el equipo Hiseq2500 produciendo lecturas de 100 pb y extremos apareados. En la secuenciación con *PacBio* se construyeron librerías de 20 kb y se secuenciaron en una celda de *PacBio RSII SMRT*.

Para asistir a la anotación génica, el ARN obtenido a partir del cultivo en los medios Sintético y Simil Mosto (ver obtención de ácidos nucleicos de *I. terricola*) fue secuenciado usando tecnología *Illumina*. Las librerías fueron construidas usando *Truseq mRNA library* después de la purificación de ARN poliA y fueron secuenciadas en un equipo *Hiseq2500* produciendo lecturas de 100 pb y extremos apareados.

3.2.2. Tratamiento de los datos

La calidad de las lecturas cortas de *Illumina* fue analizada con el programa FastQC v0.11.5 y se descartaron las diez primeras bases con el programa Trimomatic v0.36 (Bolger *et al.* 2014) con el parámetro HEADCROP: 10.

Las lecturas largas de *PacBio* fueron corregidas utilizando las lecturas cortas de *Illumina* con el programa LoRDEC v0.9 (Salmela & Rivals 2014), utilizando los parámetros –k 19 y –s 3.

3.2.3. Ensamblado

El ensamblado del genoma se hizo con las lecturas largas corregidas con las lecturas cortas utilizando el programa Flye (v2.3.6b-g37b4b0) (Kolmogorov *et al.* 2019) con parámetros por defecto y un tamaño de genoma de 12 Mb. Se realizó también el ensamblado con el programa Miniasm (v0.3-r179) (Li 2016), para lo cual antes del

ensamblado se mapearon las lecturas cortas a las lecturas largas con el programa minimap2 (v2. 2.12-r829-dirty) utilizando los parámetros por defecto.

3.2.3.1. Evaluación de los genomas

La calidad de los ensamblados fue evaluada utilizando el programa Quast v5.0.0 (Gurevich *et al.* 2013). Se utilizaron los parámetros --*gene-finding* para la predicción de genes, --*conserved-genes-finding* para evaluar la completitud del genoma a través del programa BUSCO (v3.0.2.) (Seppey *et al.* 2019) y --*fungus* para indicar que se trata del genoma de un hongo.

Se buscó en NCBI (National Center for Biotechnology Information, www.ncbi.nlm.nih.gov) los genomas de *Issatchenkia orientalis* (*Pichia Kudriavzevii*, Douglass *et al.* 2018), levadura reportada más emparentada con *I. terricola* (Tabla 1, descargados en abril de 2018).

Сера	Acceso	Nivel	Tamaño (Mb)	GC%	Scaffolds	CDS	Genes	Fecha	Ensamblador	Cobertura	Lecturas
CBS573	GCA_003054445.1	Complete	10,81	38.2	5	5154	5366	04/18	HGAP v. 3	93,8X	Ρ
CBS5147	GCA_003054405.1	Complete	10,78	38.3	5	0	0	04/18	HGAP v. 3	94X	Ρ
SJP	GCA_003033855.1	Complete	10,92	38.3	11	0	0	04/18	Falcon;HGAP3	-	Ρ
Ckrusei653	GCA_002166775.1	Contig	10,91	38.4	11	4949	5130	06/17	HGAP v. 3	24X	Ρ
129	GCA_001983325.1	Scaffold	11,68	38.5	260	5470	5470	02/17	DBG2OLC v.	115X	I+P
NG7	GCA_002798095.1	Scaffold	10,64	38.3	345	0	0	11/17	-	58,7X	I
M12	GCA_000286515.1	Contig	10,45	38.3	621	0	0	08/12	CLC	175X	I
NBRC 1279	GCA_000787615.1	Contig	10,18	38.4	669	0	0	11/14	GS De Novo	45X	454
SD108	GCA_000764455.1	Scaffold	12,94	38.5	2482	7107	7107	10/14	Newbler v. 2.6	115X	454
KMBL5774	GCA_001879685.1	Contig	7,25	-	5788	0	0	11/16	GS De Novo	100X	Р

 Tabla 1. Genomas ensamblados para Pichia Kudriavzevii descargados del NCBI.

I- Illumina, P- PacBio

https://www.ncbi.nlm.nih.gov/genome/browse/#!/eukaryotes/14557/ Consultado en abril de 2018

Se compararon los genomas ensamblados entre sí utilizando MUMmer (Nucmer v4.0.0beta2, mummerplot v3.5) (Kurtz *et al.* 2004), que permite alinear genomas enteros. Los archivos generados se visualizaron utilizando la herramienta online <u>https://dnanexus.github.io/dot/</u>.

3.2.3.2. Selección de un ensamblado para trabajar sobre el mismo

A partir de los métodos de evaluación se eligió aquel genoma que presentó los mejores parámetros de ensamblado, teniendo en cuenta principalmente la integridad del mismo, determinada por el número de *contigs*, el largo de éstos y el N50 del ensamblado.

Se le llama *contig* a una secuencia continua de ADN. Refiere a la superposición de lecturas de secuenciación que forman un mapa físico de un fragmento de ADN. Un

contig permitirá establecer la ubicación espacial de los genes que se encuentren el mismo. Un genoma estará mejor ensamblado cuando el mismo se resuelva en el menor número posible de *contigs*. En un ensamblado ideal el número de *contigs* para el ADN cromosómico coincide con el número de cromosomas del ser vivo.

El N50 es una medida de longitud media de un conjunto de secuencias, corresponde a la longitud del *contig* tal que, usando éste y los de mayor tamaño se logra cubrir la mitad de bases del genoma. Se calcula ordenando todos los *contigs* de mayor a menor y determinando el conjunto mínimo de éstos que representen el 50% del genoma.

Como respaldo se tuvieron en cuenta los resultados de la búsqueda de genes con el programa BUSCO.

3.2.3.3. Pulido del ensamblado seleccionado

Para el pulido del genoma se utilizó el programa Pilon (v1.22) (Walker *et al.* 2014) utilizando el mapeo de las lecturas cortas sobre el genoma en formato bam. Este mapeo se realizó utilizando Bowtie2 (v2.3.4.2) (Langmead & Salzberg 2012) con los parámetros por defecto.

3.2.4. Anotación de genes

Para realizar la anotación de genes se utilizó el programa Braker 1.9 (Hoff *et al.* 2019) con la opción --*fungus* y los demás parámetros por defecto. Con HISAT2 (versión 2.0.0-beta, Kim *et al.* 2015) se realizó el mapeo de las lecturas de RNAseq contra el genoma ensamblado, el bam del archivo generado fue utilizado como entrada para entrenar a los predictores Genemark y Augustus.

3.2.4.1. Evaluación de la anotación

Se evaluó la calidad de la anotación utilizando el programa BUSCO con las proteínas anotadas como entrada (parámetro –m prot) contra los linajes *ascomycete, saccharomycetes* y *saccharomycetales* descargados el 12 de noviembre de 2018 de la base de datos OrthoDB (v9.1, Waterhouse *et al.* 2012) utilizando los parámetros por defecto. Este programa busca los genes anotados en un grupo de genes ortólogos dentro de un determinado grupo taxonómico. Una correcta anotación tiene un alto porcentaje de genes identificados por BUSCO.

3.2.4.2. Anotación funcional del genoma

Se utilizó el programa InterProScan (v 5.31-70.0, Jones *et al.* 2014) para la asignación de dominios proteicos para las proteínas anotadas con el modo *protein*, la opción *–goterm* para reportar los términos GO (*GeneOntology*) y los demás parámetros por defecto. InterPro es una plataforma que tiene acceso a distintas bases de datos (PANTHER, Pfam, PROSITE, entre otras) con información sobre familias, dominios y sitios funcionales de proteínas. Las características identificables en proteínas reportadas son utilizadas para la identificación de proteínas desconocidas.

Se utilizó el programa Blast (blastp 2.7.1+, Camacho *et al.* 2009) contra la base de datos *Swissprot* (descargada de <u>https://www.uniprot.org/downloads</u> en febrero de 2019) para las proteínas anotadas con los parámetros -*e-value 1e-5*, *—maxtarget seq* 1 y demás parámetros por defecto.

Se realizó el análisis de las vías KEGG (*Kyoto Encyclopedia of Genes and Genomes*). Para ello se realizó la asignación de términos KEGG con la herramienta online KAAS (<u>https://www.genome.ip/kaas-bin/kaas main</u>) utilizando el programa GHOSTX y el archivo con la secuencia de las proteínas anotadas. Se seleccionaron los organismos *Saccharomycetes* para la búsqueda y como método de asignación BBH (*bi-directional best hit*). Con el programa *online KEGG Mapper – Reconstruct Pathway* (<u>https://www.genome.ip/kegg/tool/map pathway.html</u>) se determinaron que genes se encuentran involucrados en las distintas vías metabólicas y cuantas de éstas vías se encuentran completas en la designación de términos KEGG. Se realizó en análisis sobre el proteoma de *I. terricola* y de las cepas reportadas: *S. cerevisiae* (S288C) y *P. kudriavzevii* (129).

3.2.5. Búsqueda de enzimas

3.2.5.1. Identificación glicosidasas

La identificación de glicosidasas se llevó a cabo mediante 2 métodos distintos:

• A partir de la anotación funcional realizada con InterProScan se identificaron como glicosidasas aquellas enzimas a las cuales les fue asignada el término "glycosyl hydrolase".

• Se utilizó la herramienta dbCAN (<u>http://bcb.unl.edu/dbCAN2/blast.php</u>, Zhang *et al.* 2018) para la identificación de glicosidasas. dbCAN (*DataBase for automated Carbohydrate-active enzyme ANnotation*) es un servidor web y base de datos para la anotación automática de enzimas que actúen sobre carbohidratos. Los datos generados se basan en la clasificación de familias de la base de datos CAZy (http://www.cazy.org/) basada en similitudes de secuencia. Esta base describe las familias relacionadas estructuralmente en su sitio catalítico y en sus sustratos o dominios funcionales de enzimas que degradan, modifican o crean enlaces glicosídicos. Esta herramienta proporciona un dominio característico definido para cada familia CAZy. Para la búsqueda se utilizaron las tres herramientas disponibles: DIAMOND, HMMER y Hotpep. Se filtraron las glicosil hidrolasas (GH).

3.2.5.2. Identificación de 6-glucosidasas

La identificación de glicosidasas se llevó a cabo mediante 3 métodos distintos:

• A partir de la anotación funcional realizada con InterProScan, se identificaron aquellas enzimas a las cuales les fue asignada el término "β-glucosidase".

• Se descargaron del NCBI proteínas anotadas como β-glucosidasas de levaduras (búsqueda realizada el 21 de agosto de 2018, término de búsqueda: *beta glucosidase*

Yeast, Filtros: *Species= "Fungi*", "*Ascoymycetes*", "*Budding yeasts*" y *Source databases*=RefSeq, Base de datos "Proteins"). Se utilizó el programa Blast (blastp 2.6.0+) para buscar homología de estas secuencias con las proteínas anotadas, utilizando los parámetros -*e-value 1e-5*, *—maxtarget seq* 1 y demás parámetros por defecto.

• Se utilizó el programa Blast (blastp 2.6.0+) contra la base de datos de Swissprot para buscar proteínas anotadas que presentaran homología con β -glucosidasas de esta base de datos, se utilizaron los parámetros -*e*-value 1*e*-5, –maxtarget seq 1 y demás parámetros por defecto.

3.2.5.3. Identificación de otras enzimas de interés enológico

Se descargaron del NCBI proteínas anotadas como apiosidasas, ramnosidasas y arabinosidasas buscadas con los parámetros: *protein, refseq, fungi, ascomycetes, budding yeasts* (realizada el 21 de agosto de 2018). Se realizó un Blast contra el proteoma de *I. terricola* con los parámetros -*e-value 1e-5, –maxtarget seq* 1 y demás parámetros por defecto.

Se utilizó el programa InterProScan para determinar los dominios InterPro (IPR) asignados a cada grupo de enzimas. Se filtraron los códigos de InterPro que correspondieran a términos característicos de ese grupo y se buscó que enzimas de las anotadas tenían asignados alguno de estos términos.

3.2.6. Selección de enzimas para expresar

Análisis filogenético

Se realizó un análisis filogenético de las enzimas identificadas como β -glucosidasas por alguno de los métodos de búsqueda.

Se alinearon las secuencias de las glicosidasas con el programa muscle v3.8.31 (utilizando parámetros por defecto). El alineamiento se subió al servidor <u>https://www.genome.jp/tools-bin/ete</u>, las reconstrucciones filogenéticas fueron realizadas con el programa ETE3 v3.1.1 (Huerta-Cepas *et al.* 2016) por máxima verosimilitud con dos métodos distintos: FastTree (v2.1.8, Price *et al.* 2009) utilizando los parámetros: *Aligner - none, Alignment cleaner-none, Model tester- none y Tree builder – fasttree_default*. Para el otro método, PhyML boostrap (v20160115, Guindon *et al.* 2010) se utilizaron los parámetros: *Aligner - none, Aligner - none, Aligner - none, Alignment cleaner-none, Model tester- none, Model tester- prottest_default y Tree builder – phyml_default_bootstrap*. Los árboles generados se analizaron con el programa FigTree v1.4.4.

3.3. Técnicas de rutina de biología molecular

3.3.1. Amplificación de secuencias por PCR

3.3.1.1. Kapa-HiFi (BioSystems)

Se realizó la reacción de PCR en un volumen final de 25 μ L conteniendo el buffer de uso provisto con la enzima, y concentraciones finales de 2 mM de Mg²⁺, 0,03 mM de dNTPs, 0,03 mM de cada cebador y 1 unidad de la enzima. Se utiliza como molde 100 ng de ADNg.

Se utilizó el programa de ciclado de la Tabla 2 modificando temperatura de hibridación, tiempo de extensión y número de ciclos según el producto a amplificar.

	Temperatura (°C)	Tiempo	
Desnaturalización inicial	95	3 minutos	
Desnaturalización	98	20 segundos	
Hibridación de los cebadores	T hibridación	15 segundos	nº ciclos
Extensión	72	60 segundos/kb	
Extensión final	72	5 minutos	

Tabla 2. Programa general de ciclado para la enzima Kapa-Hifi

3.3.1.2. Phusion High-Fidelity DNA Polymerase (ThermoFisher)

Las reacciones de PCR se realizaron en un volumen final de 50 μ L conteniendo el buffer provisto con la enzima y concentraciones finales de 2 mM de Mg²⁺, 0,03 mM de dNTPs, 0,03 mM de cada cebador y 2 unidades de la enzima para el RF-*cloning*. Se tuvo en cuenta para la selección del tiempo de extensión una velocidad de 60 segundos/kb a amplificar.

3.3.1.3. HighTaq DNA Polymerase (Bioron)

Se utilizó esta enzima para las reacciones de PCR colonia para verificar presencia de inserto en vectores.

Se realizó la reacción de PCR en un volumen final de 25 μ L conteniendo el buffer provisto con la enzima y concentraciones finales de 2 mM de Mg²⁺, 0,01 mM de dNTPs, 0,03 mM de cada cebador y 1 unidad de la enzima. Se utilizó como molde 1 μ L de una suspensión de células de la colonia a analizar en 10 μ L de agua milliQ.

Se tuvo en cuenta para la selección del tiempo de extensión una velocidad de 60 segundos/kb a amplificar.

3.3.2. Purificación de fragmentos de ADN

Se utilizó el kit *Monarch DNA Gel Extraction Kit* (*BioLabs*) para purificar los productos de reacción de PCR y otros fragmentos de ADN (por ejemplo, productos de digestiones con enzimas de restricción), previa migración en geles de agarosa. El kit *GeneJET PCR Purification Kit* (*Thermo*) fue utilizado para purificar directamente productos únicos de PCR.

3.3.3. Minipreparación de ADN plasmídico

Se incubó durante toda la noche una colonia aislada de *E. coli* en 3 mL de medio LB (ver medios de cultivo) suplementado con el antibiótico apropiado según la cepa. Para conservar el clon se replicó la colonia en una placa de LB en presencia del antibiótico apropiado. Se transfirió el precultivo a un tubo *Eppendorf* que se centrifugó durante 1 minuto a 12000 g, descartándose el sobrenadante y repitiendo luego este paso hasta completar el volumen total de precultivo. El *pellet* resultante se resuspendió con pipeta en 200 µL de Solución I. Se agregó 5 µL de ARNasa (10 mg/mL) y se incubó a 37°C durante

5 minutos. Se agregó 200 μ L de Solución II preparada en el momento, mezclando luego suavemente por inversión. Se agregó 200 μ L de Solución III fría y nuevamente se mezcló por inversión, sin agitar. Se dejó 5 minutos en hielo. Luego de agregar 200 μ L de una mezcla de cloroformo:alcohol isoamílico (24:1), se mezcla por *vórtex*, centrifugando luego 15 minutos a 12000 g. El sobrenadante fue recuperado y el ADN precipitado por agregado de un volumen igual de isopropanol, mezclado por inversión e incubación por 5 minutos a temperatura ambiente. Se centrifugó 15 minutos a 12000 g y se descartó el sobrenadante. El *pellet* y las paredes del tubo fueron lavadas con 500 μ L de etanol (EtOH) 70% sin resuspensión, centrifugando luego durante 5 minutos a 12000 g, tras lo cual se retiró el sobrenadante y se dejó secar el *pellet* al aire. Luego éste se resuspendió en TE 10:1 pH 8 o en agua MilliQ.

Soluciones:

Solución I: Tris-HCl 50 mM pH = 7,5; EDTA 20 mM pH =8 Solución II: NaOH 0,2 N; SDS 1% Solución III: AcK 3 M; ácido acético glacial 11,5% TE 10:1: Tris-HCl 10 mM pH = 8; EDTA 1 mM

3.3.4. Visualización de ácidos nucleicos

Los ácidos nucleicos se analizaron en geles de agarosa 1% con buffer TAE 1X como buffer de corrida en una cuba de electroforesis a 90 V. Para cargar las muestras se utilizó buffer de carga concentración final 1X. Se utilizaron los marcadores de peso molecular *GeneRulerTM 1 kb Plus DNA Ladder (Thermo), GeneRulerTM DNA Ladder Mix (Thermo)* y *1 kb DNA ladder (New England Biolabs*). Se visualizaron los ácidos nucleicos mediante exposición del gel a luz ultravioleta.

En el caso de los geles para ARN se tuvieron los cuidados de utilizar reactivos libres de RNAsas y soluciones preparadas a partir de agua tratada con DEPC (pirocarbonato de dietilo).

Reactivos y soluciones:

Gel de agarosa

Se disolvió la cantidad de agarosa necesaria para obtener el porcentaje masa/volumen deseado en buffer TAE 1X. Se utilizaron 2,5 µL/50 mL de gel de agente intercalante SYBR Safe DNA Gel Stain (Invitrogen).

TAE 50X: Tris Base 2M; Ácido acético glacial 1M; EDTA 0,05 M pH 8,0 en agua.

Buffer de carga 6X: 0,25% azul de bromofenol; 0,25% xileno cianol; 30% glicerol en agua.

3.3.5. Análisis de secuenciación de plásmidos

Se analizaron los resultados de secuenciación utilizando el programa *Blast* (blastn 2.7.1+, Camacho *et al.* 2009) con los parámetros por defecto contra las secuencias de las construcciones.

Se analizaron los cromatogramas de los datos de secuenciación con el programa *Chromas 2.5 (Technelysium Pty. Ltd.*).

3.3.6. Preparación de células competentes

3.3.6.1. Preparación de células quimiocompetentes

(Tom Knight and the Registry of Standard Biological Parts (http://partsregistry.org))

Se inoculó 250 mL de medio LB con 1 mL de células de la línea XL-1 provenientes de un precultivo y se dejó crecer a temperatura ambiente hasta una DO de 0,3 (aproximadamente 16 horas). Se centrifugó a 3000 rpm durante 10 minutos a 4°C y se descartó el sobrenadante. Se resuspendió el *pellet* en 80 mL de buffer CCMB80 previamente enfriado en hielo y se incubaron las células en hielo durante 20 minutos, tras lo cual se centrifugó nuevamente a 3000 rpm durante 10 minutos a 4°C, descartándose el sobrenadante. El *pellet* se resuspendió en 10 mL de buffer CCMB80 frío y se midió la DO de una dilución de 50 µL de células en 200 µL de SOC (ver medios de cultivo), para ajustar luego la DO a 1-1,5 con buffer CCMB80, tras lo cual las células se congelaron inmediatamente en N₂ líquido en alícuotas de 50 µL y se almacenaron a -80°C. Se calculó la eficiencia de transformación con un plásmido de concentración conocida.

Soluciones:

Buffer CCMB80: KAc 10 mM pH 7,0; CaCl₂.2H₂O 80 mM; MnCl₂.4H₂O 20 mM; MgCl₂.6H₂O 10 mM; glicerol 10%. Se ajustó el pH a 7,0, se esterilizó por filtración y se almacenó a 4°C.

Preparación del stock:

Se estriaron células en placas de medio SOB (ver medios de cultivo) (o LB) y se incubó a temperatura ambiente durante aproximadamente 48 horas. Se picaron varias colonias en 2 mL de LB y se dejó en agitación overnight a 23°C, alicuotando luego en criotubos con glicerol al 15% de concentración final.

3.3.6.2. Preparación de células electrocompetentes

Las células XL-1 blue tienen resistencia a tetraciclina, por lo cual fue agregada al precultivo y al cultivo a una concentración final de 12,5 µg/mL. En el caso de las células DH5 α no se utilizó antibiótico. Se preparó un precultivo en 5 mL de LB a partir de una colonia fresca de la bacteria. Se dejó crecer durante la noche a 37°C y 200 rpm. Con 4,5 mL del precultivo se inoculó 1 L de LB, que se creció a 37°C y 200 rpm hasta una DO de 0,65. Se enfrió el cultivo 15-30 minutos en hielo y se centrifugaron las células a 4000 rpm durante 15 minutos a 4°C. El *pellet* de células se resuspendió en un volumen final de 1 L de agua estéril fría. Se centrifugaron las células a 4000 rpm durante 15 minutos a 4°C, resuspendiéndose el *pellet* en 20 mL de glicerol estéril y frio al 10%. Se centrifugó a 4000 rpm durante 15 minutos a 4°C y se descartó el sobrenadante. El *pellet* se resuspendió en 2,5 mL de glicerol al 10% estéril frío. Se alicuotaron las células de a 50 µL en *Eppendorfs* previamente enfriados y se almacenaron a -80°C. Se calculó la eficiencia de transformación con un plásmido de concentración conocida.

3.3.7. Medios de cultivo

3.3.7.1. LB (Lysogeny Broth)

10 g de bacto-triptona, 5 g de extracto de levadura, 10g de NaCl. Se lleva a un litro con agua destilada, se ajusta el pH a 7 y se esteriliza por calor húmedo a 121° C durante 20 minutos.

3.3.7.2. SOB (Super Optimal Broth)

0.5% extracto de levadura, 2% triptona, NaCl 10 mM, KCl 2,5 mM, $MgCl_2$ 10 mM. Se esteriliza por calor húmedo a 121° C durante 20 minutos.

3.3.7.3. SOC (Super Optimal broth with Catabolites repression)

Medio SOB con el agregado de glucosa, concentración final 20 mM. La glucosa se prepara y esteriliza separadamente y se añade luego de autoclavado el medio SOB.

3.3.8. Transformación de E. coli

3.3.8.1. Transformación química de E. coli

A una alícuota de 50 μ L de células competentes de la cepa XL-1 de *E. coli* descongelada en hielo, se añadió no más de 50 ng del plásmido de interés, se incubó 20-30 minutos en hielo. Se realizó el shock térmico colocando las células 90 segundos a 42°C. Inmediatamente se colocaron 1-2 minutos en hielo, luego se añadió 500 μ L de LB y se incubó durante 1 hora a 37°C y 200 rpm. Se sembraron 100 μ L de la mezcla de transformación en placas LB-Ampi. El resto de la mezcla se centrifugó durante 1 minuto a 5000 g para concentrar las células, descartándose la mayor parte del sobrenadante excepto un remanente de 100-150 μ L para resuspender las células que finalmente se sembraron en placas de LB-Ampi. Las placas se incubaron en una estufa a 37°C hasta el día siguiente, cuando se evaluó la presencia de transformantes.

3.3.8.2. Transformación mediante electroporación

Se utilizaron células electrocompetentes de *E. coli* de la cepa XL-1. A una alícuota de 40 μ L de células descongelada en hielo se le añadió 0,7 μ L del ADN de interés. Se incubó por 1-2 minutos en hielo y se transfirió la mezcla a una celda de electroporación de 2 mm de espesor enfriada previamente. La electroporación se realizó utilizando el electroporador *Gene Pulser X Cell Electroporation System* (*BioRad*) siguiendo el protocolo indicado para bacterias (pulso de 2,5 kV por 5 ms). Inmediatamente se agregó 1 mL de medio LB a la celda y se mezcló por inversión. La suspensión de células se transfirió a un tubo *Eppendorf* y se incubó durante 1 h a 37°C a 200 rpm. El plaqueo y la incubación se realizaron como se describió para el caso de la transformación de células quimiocomptentes.

3.4. Expresión de proteínas en E. coli

3.4.1. Clonado en vectores plasmídicos mediante Restriction Free *Cloning* (RF-*cloning*) Esta estrategia consiste en el clonado de una secuencia de interés en un vector de expresión sin el uso de enzimas de restricción. En una primera etapa el fragmento a ser insertado es amplificado con extensiones en sus extremos que hibridan con el vector en el lugar donde se desea realizar la inserción; a este fragmento se le llamará "megaprimer". En una segunda etapa mediante PCR, las extensiones del megaprimer hibridarán con la región del vector donde será insertado el fragmento mediante la ADN polimerasa (Figura 9, van den Ent & Lowe 2006).

1: Vector prep and PCR



2: Linear amplification



Figura 9. Esquema de RF-cloning. 1. Se muestra en el vector en rojo y azul los sitios donde se clonará el fragmento de interés. En el fragmento linear se indica la amplificación del fragmento de interés con las extensiones complementarias a las regiones roja y azul añadidas en los cebadores (megaprimer). 2. El megaprimer actúa como cebador en la amplificación linear del vector. Unión a los sitios de inserción por hibirdación de las regiones roja y azul del megaprimer con las regiones correspondientes en el vector. 3. Digestión de la hebra parental metilada con DpnI. Extraído y modificado de van den Ent & Lowe 2006.

3.4.1.1. PCR para amplificar megaprimers para el RF-cloning

Se utilizó la enzima *KapaHiFi* en las condiciones de reacción descritas en 3.3.1.1., con temperaturas de hibridación de 62ºC o 65ºC y 25 ciclos de amplificación.

Se utilizaron los siguientes juegos de cebadores según el fragmento a amplificar:

	Fragmento a amplificar							
	989	989 sin péptido señal	3296	3296 sin péptido señal				
Cebador Fw	989_vector _ecoli_F	989_vector_ecoli_F_sin PS	3296_vecto r_ecoli_F	3296_vector_ecoli_F_sin PS				
Cebador Rv	989_vector _ecoli_R	989_vector_ecoli_R	3296_vecto r_ecoli_R	3296_vector_ecoli_R				
Tamaño esperado	993 pb	927 pb	1158 pb	1107 pb				

 Tabla 3. Cebadores y tamaño esperado para cada fragmento a amplificar

Se visualizaron los productos de las PCR en gel de agarosa 1% y se purificaron como se describió en la sección 3.3.4.

3.4.1.2. Clonado mediante RF-cloning

Mediante una segunda reacción de PCR, se introdujo el *megaprimer* correspondiente en el vector de elección, utilizando para ello la enzima *Phusion*. Las cantidades utilizadas fueron las sugeridas por el programa <u>https://www.rf-cloning.org/</u> (Bond & Naus 2012).

Se utilizó el programa genérico de PCR descripto en el punto 3.3.1.2. para la amplificación con la enzima *Phusion*. Se utilizó una temperatura de hibridación de 60ºC y por tratarse de RF-*cloning*, se utilizó un tiempo de hibridación de un minuto, para optimizar el pegado del *megaprimer*. Se realizaron 30 ciclos de amplificación.

Paso seguido, se lleva a cabo una digestión con DpnI para eliminar la hebra parental del vector, metilada. A 20 μ L del producto de PCR se agregó 2 μ L (20 unidades) de DpnI (10 U/ μ L) (*Thermo*) y se realizó la digestión durante toda la noche a 37°C. Se inactivó la enzima por incubación durante 20 minutos a 80°C. 5 μ L de esta digestión se utilizaron para transformar células quimiocompetentes y 0,7 μ L para transformar células electrocompetentes.

3.4.2. Cultivos para la expresión recombinante de la enzima 989

Se realizaron cultivos de expresión en medio LB con la cepa de *E. coli* BL-21 transformada con los plásmidos obtenidos por el RF-*cloning*. Se agregó en todos los pasos ampicilina ya que los vectores utilizados otorgan resistencia a este antibiótico.

Se ensayaron volúmenes de cultivo de 25 y 100 mL, siempre utilizando un matraz adecuado para que el volumen de medio de cultivo fuera 1/10 del volumen del matraz. Inicialmente se realizó un precultivo con el que luego se inoculó el matraz donde se llevó a cabo la expresión, con un volumen de precultivo igual a 1/100 del volumen de medio de cultivo. Se crecieron las bacterias hasta una DO de aproximadamente 0,6 y se añadió IPTG para la inducción de la expresión. Fueron probadas también distintas temperaturas de crecimiento (20°C, 30°C y 37°C) y tiempos de inducción (3 hs, 5 hs, *overnight* (16 horas)) tras los cuales se recolectaron las células por centrifugación a 5000 rpm a 4°C, durante 30 minutos. El *pellet* se lavó con PBS 1X, se centrifugó en las mismas condiciones durante 15 minutos y se suspendió en 5 mL de buffer de lisis por cada 25 mL de cultivo. Se agregaron los siguientes inhibidores de proteasas a las siguientes concentraciones finales: Benzamidina 1 mM, Inhibidores de tripsina 1 µg/mL y PMSF 1 mM.

El *pellet* suspendido y se sonicó en baño de hielo en tres intervalos de 10 minutos para evitar el calentamiento de la muestra, utilizando el equipo *SonicRuptor250 (Omni International)* con los parámetros *Power* = 40 y *Pulser* = 20. Entre intervalos se realizó la homogenización de la suspensión. Al final del tratamiento se tomó una alícuota para analizar la fracción total (ET) y la suspensión se centrifugó a 12000 rpm a 4°C, durante 60 para separar a continuación la fracción soluble (FS) e insoluble (FI).

La FI se lava con PBS 1X para eliminar cualquier resto que pueda haber quedado de proteínas de la FS, tras lo que se centrifugó nuevamente a 12000 rpm a 4°C, durante 15 minutos. Se descartó el sobrenadante y se suspendió el *pellet* en 5 mL de urea 2 M por cada 25 mL de cultivo, conteniendo inhibidores de proteasas.

En todas las condiciones se realizó en paralelo un control sin inducir con IPTG, cultivo que se trató siguiendo los mismos protocolos. Todas las muestras de ET, FS y FI se analizaron por SDS-PAGE (*Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis*) y eventualmente alguna de estas fracciones se utilizó para ensayos de actividad glicosidasa por zimograma e inmunodetección por *western-blot*.

3.4.3. Evaluación de la expresión y purificación de las proteínas recombinantes

3.4.3.1. Cálculo de los pesos moleculares para las construcciones

Para la predicción del peso molecular de las proteínas expresadas recombinantemente se utilizó la herramienta <u>https://www.bioinformatics.org/sms/prot_mw.html</u> con la secuencia de aminoácidos de la construcción.

3.4.3.2. Análisis de las muestras por SDS-PAGE

Para el análisis de proteínas en condiciones desnaturalizantes se realizó el método descrito por Laemmli (Laemmli 1970), con un gel concentrador al 6% de poliacrilamida y un gel separador de poliacrilamida al 12%. Previo al cargado de las muestras se agregó a las mismas buffer de carga a concentración final 1X y se las hirvió durante 5 minutos, enfriándolas en hielo. Se cargaron 4 μ L del marcador de peso molecular *AccuRuler RGB Prestained Protein Ladder (Meastrogen*). Se realizó la electroforesis en buffer de corrida 1X a 15 mA por gel durante la corrida en el gel concentrador y a 30 mA por gel durante la corrida en el gel concentrador y a 30 mA por gel durante la corrida en el separador.

Se realizó la visualización de proteínas mediante tinción con *Coomassie Blue R-250* (*Sigma-Aldrich*) incubando el gel en la solución de tinción de un día para el otro. Luego se decoloró con solución decolorante hasta la visualización de las bandas.

Soluciones y reactivos

Buffer de carga 4X: Tris-HCl 0,25 M pH 6,8; Glicerol 40%; SDS 4%; 2-Mercaptoetanol 4%; Azul de bromofenol 0,02%

Buffer de corrida 1X: Tris 25 mM; Glicina 192 mM; SDS 0,1% pH 8,3

Solución de Coomassie Blue: Coomassie Blue R-250 0,2%; Etanol 30%; Ácido acético 7%

Solución decolorante: Etanol 30%; Ácido acético 7%

3.4.3.3. Cromatografía de afinidad por iones metálicos

a) Preparación de la matriz

Las proteínas fueron clonadas en vectores de expresión que permiten el agregado a la proteína recombinante de interés de un His-tag (en este caso agregado en el extremo N-terminal). Por lo cual, para la purificación se utilizó la matriz comercial *Chelating Sepharose Fast Flow (BioProcess)* previamente cargada con cobre. Los grupos metálicos de dicha matriz forman complejos no covalentes con las histidinas presentes en las proteínas, en especial con el His-tag adicionado a la proteína de interés. Se utilizó 1 mL de la matriz comercial la cual se lavó con 2 volúmenes de agua MilliQ, luego se aplicó 0,2 volúmenes de la solución del metal (CuSO₄ 0,2 M). Se lavó con 5 volúmenes de agua MilliQ y luego con 5 volúmenes (o hasta que el pH del efluente sea de 4,0) de acetato de
sodio 0,02 M pH = 4,0 suplementado con NaCl 0,5 M. A continuación se lavó con 5 volúmenes de agua MilliQ y se equilibró la matriz con 5 volúmenes de buffer de lisis.

b) Preparación de la muestra

El sobrenadante obtenido de lisar las bacterias provenientes de los cultivos de expresión se filtró por un filtro de 0,45 μ m previo a ser colocado en la columna conteniendo la matriz de afinidad.

c) Cromatografía

Se colocó la muestra en la matriz de afinidad y se recuperó el percolado. Se realizaron dos series consecutivas de 5 lavados con buffer de lavado 1 y 3 lavados con buffer de lavado 2, recogiendo los percolados de cada uno por separado. Finalmente se llevó a cabo la elución utilizando un buffer conteniendo una alta concentración de imidazol y recogiéndose seis fracciones de 1 mL en tubos *Eppendorf*.

d) Regeneración de la columna

Para regenerar la columna se lavó con un volumen de solución de EDTA 0,2 M y NaCl 0,5 M para despegar el metal de la matriz. Se lavó con 4 volúmenes de NaCl 0,5 M para retirar el EDTA residual. Se tapó la columna, se llenó con NaCl 2 M, se homogeneizó y se incuba durante 10 minutos. Luego se destapó, se eluyó la fase líquida y se repitió la operación con NaOH 1 M. Finalmente, se almacenó en EtOH 20%.

e) Soluciones

Buffer de lisis: Tris-HCl 20 mM; NaCl 0,5 M; Glicerol 10%; pH = 8

Buffer de lavado 1: Tris-HCl 20 mM; NaCl 150 mM; pH = 7,4

Buffer de lavado 2: Tris-HCl 20 mM; NaCl 150 mM; Imidazol 20 mM; pH = 7,4

Buffer de elución 1: Tris-HCl 20 mM; NaCl 50 mM; Glicerol 10%; Imidazol 300 mM; pH = 7,4

Buffer de elución 2: Tris-HCl 20 mM; NaCl 50 mM; Glicerol 10%; Imidazol 500 mM; pH = 7,4

3.4.4. Evaluación de la presencia de la β -glucosidasa recombinante en las distintas fracciones

3.4.4.1. Zimograma (ensayo cualitativo de actividad 6-glucosidasa)

Se realizan los geles en las mismas condiciones que para la SDS-PAGE, a diferencia de ésta el buffer de carga de las proteínas no contiene agentes reductores y las muestras no se calientan para no desnaturalizarlas.

Una vez realizada la corrida electroforética se incubó el gel durante una hora a temperatura ambiente con agitación en presencia de TRITON x100 (1,5%) para eliminar el SDS y lograr el *refolding* de las proteínas. Luego se enjuagó con agua destilada y se le colocó el sustrato MUG para la reacción de la β -glucosidasa. Se incubó durante 15 minutos a 37° y se reveló con luz ultravioleta. Como control positivo se sembraron 10 µL de una dilución 1/100 de un extracto comercial con una β -glucosidasa (*Cytolase PCL5 DSM Food Specialities, USA*).

Soluciones

Solución de MUG (4-metilumbeliferil-\beta-D-glucósido): MUG (Sigma) 5 mM en buffer fosfato de sodio 0,1 M pH = 7,0

3.4.4.2. Inmunodetección del His-tag por Western-blot

Se migraron las muestras a analizar en un gel de SDS-PAGE, incluyendo un marcador de peso molecular y un control positivo (PTPB fosfatasa de *Mycobacterium* con His-tag). El gel se equilibró 15 minutos en buffer TBST antes de la transferencia.

Se realizó la transferencia a una membrana de PVDF (#88518 *Thermo*) durante 1 hora a 100 V en buffer de transferencia en una cuba de *BioRad Mini Trans-Blot® Electrophoretic Transfer Cell*.

Se dejó bloqueando la membrana con Solución de Bloqueo (000105 *Life Membrane Blocking Solution, Thermo*) durante toda la noche a 4°C. Se realizaron 3 lavados de 5 minutos con TBST y se incubó toda la noche a 4°C con el anticuerpo primario anti-HisTag (His-Tag 27E8 *Mouse mAb* (*CellSignaling*), dilución 1:1000 en solución de bloqueo). Luego de repetir los lavados, se incubó 1 hora a temperatura ambiente con el anticuerpo secundario (HRP *anti-mouse*; dilución 1:10000 en TBST). Finalmente, luego de 3 lavados de 10 minutos con se procedió al revelado. Con este fin se incubó con el reactivo *SuperSignal™ West Pico PLUS Chemiluminescent Substrate (Thermo*) según las instrucciones del fabricante y se reveló la quimioluminiscencia desarrollada con el equipo *G-Box (Syngene*), registrándose la señal luego de 6 exposiciones de 10 segundos cada una.

Soluciones

TBST: Tris-HCl 0,01 M pH 7,5; NaCl 0,075 mM; Tween 20 0,001%

Buffer de transferencia: Tris 25 mM, glicina 200 mM, EtOH 20%, SDS 0,04%

3.4.4.3. Espectrometría de masas

Se realizó la identificación de proteínas por mapeo peptídico con tripsina y medida por MALDI-TOF incluyendo los MS/MS. El análisis de las muestras fue realizado por la Unidad de Bioquímica y Proteómica Analítica del Instituto Pasteur de Montevideo. Las muestras para el análisis fueron obtenidas a partir de una corrida de SDS-PAGE con los cuidados necesarios para evitar la contaminación con queratina.

3.5. Expresión de las enzimas 989 y 3296 en Saccharomyces cerevisiae

Para el clonado en *S. cerevisiae* su utilizó la estrategia reportada en el trabajo de Bai Flagfeldt *et al.* (2009). La misma consiste en la integración cromosómica dirigida de un *cassette* de expresión mediante recombinación homóloga. El *cassette* consiste de tres regiones: promotor, gen a insertar y terminador. Estas tres regiones son amplificadas independientemente y luego son unidas para formar el *cassette* mediante *Fusion* PCR (Reid *et al.* 2002).



Figura 10. Esquema de PCR-fusion. 1. Los cebadores se ilustran con las flechas. Los cebadores con las extensiones (adaptámeros) se ilustran como flechas con triángulos en el extremo 5', los triángulos representan secuencias que se complementan entre sí. 2. La PCR con estos adaptámeros incorporan la extensión en el extremo 5' del fragmento amplificado (rombo). 3. En una segunda reacción de PCR con los productos de las PCR anteriores las secuencias complementarias incorporadas hibridan entre si y los fragmentos son amplificados por la polimerasa (líneas punteadas). 4. En los sucesivos ciclos de PCR los cebadores de los extremos amplifican el producto fusionado. Extraído de Reid et al. 2002.

Este *cassette* es clonado con enzimas de restricción en un plásmido que contiene el gen URA3 de *Kluveromyces lactis*. Luego mediante digestión con otro par de enzimas se libera la construcción que va a ser introducida en *S. cerevisiae*, la misma está conformada por el *cassette* construido por *Fusion* PCR seguido del gen de URA3 para la selección de transformantes. Se utilizará este gen de selección porque la cepa a transformar es la T73-4 que porta una doble deleción de ambas copias del gen URA3 (Puig *et al.* 1998), permitiendo la selección de transformantes por su capacidad de crecer en un medio sin uracilo. Se amplifican además las regiones *upstream* y *downstream* no codificantes del *locus* de inserción para dirigir la recombinación homóloga al sitio del cromosoma elegido. Estas regiones cuentan con extensiones que presentan homología con los extremos del *cassette* final, lo que permite finalmente dirigir la inserción por homología dentro de *S. cerevisiae*.



Figura 11. Esquema la estrategia de clonado en S. cerevisiae. Integración del cassette con el gen de la enzima seleccionada. Up/down – Regiones upstream y downstream el sitio de integración (int), P – Promotor (ADH, ACT o ENDO), T- Terminador CYC, R- Repetidos directos. Extraído y modificado de Bai Flagfeldt et al. (2009).

3.5.1. Clonado de las enzimas seleccionadas en los cassettes de expresión

3.5.1.1. PCR para amplificar los fragmentos para la construcción del cassette

Se utilizó el programa genérico de PCR descripto en 3.3.1.1. para la *KapaHiFi*. Se utilizó una temperatura de hibridación de 60ºC y por tratarse de cebadores largos se realizó la hibridación del durante un minuto. Se realizaron 25 ciclos de amplificación y se utilizaron los siguientes juegos de cebadores según el fragmento a amplificar:

		Fragmento a amplificar			
	Promotor ADH	Promotor ACT	Terminador CYC		
Cebador Fw	ADH1_Fw ^a	ACT1_Fw ^a	CYC1_Fw		
Cebador Rv	ADH1_Rv	ACT1_Rv	CYC1_Rv ^b		
Tamaño esperado	1437 pb	424 pb	250 pb		
		Fragmento a amplificar			
	989extADH*	989extACT*	989promendo*		
Cebador Fw	ADH1_989_Fw	ACT1_989_Fw	Sacll_prom_989_Fw ^a		
Cebador Rv	CYC1_989_Rv	CYC1_989_Rv	CYC1_989_Rv		
Tamaño esperado	991 pb	988 pb	1473 pb		
		Fragmento a amplificar			
	2206-0044004*		2206 m m an d a *		
	3296extADH*	SZ9BEXTACT	5296promendo*		
Cebador Fw	ADH1_3296_Fw	ACT1_3296_Fw	SacII_prom_3296_Fw ^a		
Cebador Rv	CYC1_3296_Rv	CYC1_3296_Rv	CYC1_3296_Rv		
Tamaño esperado	1156 pb	1153 pb	1638 pb		
*Todos los fragmentos generados tienen una extensión del terminador CYC y del promotor correspondiente en cada caso para realizar la PCR de fusión de promotor, enzima y terminador. a- Contiene sitio de restricción para SacII; b- Contiene sitio de restricción para Xmal					

Tabla 4. Cebadores y tamaño esperado para cada fragmento a amplificar

Se visualizaron los productos de las PCR en gel de agarosa 1% y se purificaron como fue descripto anteriormente.

3.5.1.2. PCR fusión para generar los cassettes

En la reacción de fusión, se utilizaron 25 ng de cada fragmento y la enzima *KapaHiFi* con el programa descripto 3.3.1.1., excepto que se agregó el doble de dNTPs. Se utilizó una temperatura de hibridación de 65ºC. Se realizaron 25 ciclos, añadiéndose a partir del ciclo 15, 20 segundos al tiempo de extensión. Y se utilizaron los siguientes juegos de cebadores según la fusión a amplificar:

	Cassette 989			
Promotor del cassette	ADH	ACT	ENDO	
Cebador Fw	ADH1_Fw	ACT1_Fw	Sacll_prom_989_Fw	
Cebador Rv	CYC1_Rv	CYC1_Rv	CYC1_Rv	
Tamaño esperado	2620 pb	1607 pb	1683 pb	
		Cassette 3296		
Promotor del cassette	ADH	ACT	ENDO	
cebador Fw	ADH1_Fw	ACT1_Fw	Sacll_prom_3296_Fw	
cebador Rv	CYC1_Rv	CYC1_Rv	CYC1_Rv	
Tamaño esperado	2785 pb	1772 pb	1848 pb	

Tabla 5. Cebadores	∕ tamañc	esperado	para cado	a cassette a	amplificar
--------------------	----------	----------	-----------	--------------	------------

Se visualizaron los productos de las PCR en gel de agarosa 1% y se purificaron como fue descripto anteriormente.

3.5.1.3. PCR para amplificar las regiones upstream y downstream del gen YPRCΔ15 para dirigir la recombinación homóloga

En el artículo Bai Flagfeldt *et al.* 2009 reportan que para una misma construcción varían los niveles de expresión del gen reportero según el sitio del genoma donde haya sido insertado el mismo. Basado es sus resultados se seleccionó el locus del gen YPRCΔ15 para la inserción del *cassette* por presentar los niveles más altos de expresión dentro de los 20 sitos que son evaluados en el artículo.

Se utilizó la enzima *KapaHiFi* en las condiciones de reacción que se describen en 3.3.1.1. Se utilizó una temperatura de hibridación de 62ºC y se realizaron 25 ciclos de amplificación.

Y se utilizaron los siguientes juegos de cebadores según el fragmento a amplificar:

	Fragmento a amplificar		
	Región upstream Región downstrean		
Cebador Fw	up_fw	down_fw	
Cebador Rv	up_rv	down_rv	
Tamaño esperado	680 pb	650 pb	

Tabla 6. Cebadores y tamaño esperado para cada fragmento a amplificar

Se visualizaron los productos de las PCR en gel de agarosa 1% y se purificaron como fue descripto anteriormente.

3.5.1.4. Clonado del cassette en el vector pWJ1042

Se realizó la digestión del plásmido pWJ1042 y de los *cassettes* con *Sac*II y *Xma*I (*BioLabs*) bajo las recomendaciones del fabricante, en un volumen final de 50 μL. Se digirió 1 μg de cada muestra durante 4 horas a 37°.

Se visualizaron los productos de la digestión en gel de agarosa 1% y se purificaron como fue descripto anteriormente, eluyéndose con agua y no con buffer de elución para no afectar la ligación.

Se realizó la ligación de cada *cassette* en el plásmido pWJ1042 en un volumen final de 20 μ L con la enzima T4 (BioLabs) siguiendo las instrucciones del fabricante. Para el cálculo de las cantidades de cada fragmento se utilizó la siguiente fórmula:

relación inserto: vector. tamaño inserto (pb). cantidad de vector tamaño vector (pb)

Teniendo en cuenta una relación inserto:vector de 10 y 25 ng de vector a utilizar.

Se dejó ligando durante toda la noche a 4°C y se inactivó la ligasa por calentamiento durante 10 minutos a 65°C.

Se transformaron 50 μ L de células quimiocompetentes como es descripto en la sección de técnicas de rutina de biología molecular con 5 μ L del producto de la ligación.

Se realizó la extracción de ADN plasmídico de los transformantes para analizar la presencia del inserto. Se migraron los plásmidos en gel de agarosa 1% y se compararon los tamaños con el plásmido sin inserto. Además se realizó una digestión con *Sac*II y *Xma*I como fue descripto anteriormente para verificar efectivamente la presencia del inserto.

3.5.1.5. Digestión para obtener el cassette final para la transformación

Se realiza la digestión del plásmido pWJ1042 ligado al *cassette* que corresponda en cada caso con las enzimas *Ahd*I y *PspOM*I para liberar el fragmento que contiene el *cassette* junto con el gen URA3 de *Kluveromyces lactis* para la co transformación de *S. cerevisiae* junto con los fragmentos *up* y *down*.

4. RESULTADOS Y DISCUSIÓN

En esta primera sección se presentarán los resultados relativos a la obtención del genoma de I. terricola. Esto incluye la obtención de ácidos nucleicos de calidad adecuada para su secuenciacón masiva, el análisis de los datos de secuenciación para el ensamblado, la evaluación de los ensamblados obtenidos y anotación del genoma. Finalmente se seleccionan dos glucosidasas con las cuales se continuó el trabajo.

4.1. Extracción de ácidos nucleicos

Los servicios de secuenciación masiva requieren una cantidad determinada de material y una calidad adecuada del mismo. Por esta razón, los protocolos de extracción de ácidos nucleicos fueron optimizados de modo de obtener al menos 1,0 µg de ADN (concentración \ge 10 ng/µL) para la secuenciación con Illumina, 12 µg de ADN (concentración \ge 50 ng/µL) para la secuenciación con PacBio y 1 µg de ARN (concentración \ge 20 ng/µL) de cada condición para la secuenciación de ARN. Se verificó la integridad de las muestras mediante electroforesis en gel de agarosa (Figura 1), y se evaluó la relación de absorbancias 260/280 con Nanodrop obteniendo para cada caso un valor de calidad mayor al mínimo aceptable (1,8 para ADN y 2,0 para ARN).

Para la extracción de ARN se creció la levadura en dos medios distintos para obtener ARN mensajero en distintas condiciones: en medio Simil Mosto (por ser el más parecido al mosto a partir del cual se realiza la vinificación) y en medio Sintético (por ser un medio en que *l. terricola* presenta un buen crecimiento y una buena actividad β -glucosidasa) (de Ovalle *et al.* 2017). Los ácidos nucleicos extraídos fueron enviados al servicio de secuenciación de Macrogen.



Figura 12. Extracción de ácidos nucleicos. a) Carril 1- MPM GeneRuler[™] 1kb Plus DNA Ladder, Carril 2-5 ADNg de I. terricola; b) Carril 1- MPM GeneRuler[™] 1kb Plus DNA Ladder, Carril 2- ARN de I. terricola en medio Sintético, Carril 3- ARN de I. terricola en medio Simil Mosto

4.2. Ensamblado y anotación del genoma de I. terricola

4.2.1. Tratamiento de los datos del servicio de secuenciación

La secuenciación con PacBio dio como resultado 138.868 lecturas largas.

Con Illumina se secuenció un total de 2.273.483.639 bases en 22.509.739 lecturas cortas con extremos apareados con un largo medio de 91 pares de bases, logrando una cobertura de 170X. Se evaluó la calidad de las lecturas cortas mediante FastQC, y debido a que tenían calidad inicial muy buena (Figura 13) solo se descartaron las primeras diez primeras bases de cada lectura por presentar sesgos en su composición y menor calidad.



Figura 13. Resultado del análisis de FastQC para las lecturas cortas a) Calidad inicial de las lecturas crudas b) Calidad de las lecturas procesadas.



Figura 14. Contenido GC para las lecturas de Illumina. Gráfico de la izquierda: Primer juego de lecturas; gráfico de la derecha: Segundo juego de lecturas.

En la Figura 14 se muestran los gráficos de contenido GC para los dos juegos de lecturas de Illumina. En ambos casos se observa un contenido de aproximadamente 36% GC y una distribución normal, a excepción de una menor cantidad de lecturas (representadas por un pequeño hombro en la campana) que representa en ambos casos lecturas con un contenido GC de aproximadamente 16%. Esto podría corresponder a las lecturas del ADN mitocondrial, ya que en *S. cerevisiae* el porcentaje GC del ADN mitocondiral es de 18% (Fangman & Dujon 1984) y se esperan valores similares a éste porcentaje para otras levaduras.

4.2.2. Ensamblado del genoma

Se realizó el ensamblado mediante dos programas distintos (Flye y Miniasm) y se analizó la calidad de los ensamblados utilizando el programa Quast. En la Tabla 7 se detallan los parámetros obtenidos para cada uno de los ensamblados.

Programa	Largo (bases)	%GC	Contigs	Contigs	<i>Contig</i> más N50 (bases) largo (bases)	Genes	BUS	со	
Ū			5	> 1Mb		. ,		Completos	Parciales
Flye	12.522.592	36,42	15	5	3.439.421	3.383.105	5.646	97,93%	0,34%
Miniasm	14.239.185	35,10	63	5	3.452.613	2.377.647	6.660	91,38%	4,83%

Tabla 7. Métricas de los genomas ensamblados.

El genoma ensamblado con Miniasm se encuentra fraccionado en 63 *contigs*. De éstos, 5 tienen un tamaño mayor a 1Mb y 2 son mayores a 500.000 b. El N50 de es 2,38 Mb para un genoma de 14,2 Mb y se predijeron para el mismo 6.660 genes. El análisis de completitud muestra que el 91,38% de los genes del linaje *fungi* fueron hallados completos y un 4,83% fue hallado fragmentado.

El genoma obtenido con Flye se encuentra menos fragmentado, en 15 *contigs*, de los cuales los 5 más largos son mayores a 1,5 Mb y podrían corresponder a los cromosomas de la levadura, ya que para la levadura emparentada *P. kudriavzevii* se ha reportado la presencia de 5 cromosomas de tamaño similar (Douglass *et al.* 2018). En la Tabla 8 se puede apreciar que, además, la cantidad de genes anotados para cada cromosoma en ambos organismos es similar, lo cual reafirma la calidad del genoma obtenido para *I. terricola*.

	I. terricola		P. kudriavze	evii CBS573
Contig	Largo (Mb)	Nº genes	Largo (Mb)	Nº genes
1	3.442	1423	2.852	1346
2	3.385	1426	2.746	1322
3	2.364	995	2.542	1204
4	1.698	712	1.384	670
5	1.617	674	1.289	598

Tabla 8. Largo y número de genes para los contigs más largos del genoma ensamblado de I. terricola y P. kudriavzevii CBS573 (Douglass et al. 2018).

Se ensambló el genoma mitocondrial con las lecturas de Illumina utilizando el programa Norgal v0.1 (Al-Nakeeb *et al.* 2017). Se realizó una búsqueda por homología (blastn) entre los fragmentos generados en el ensamblado del genoma obtenido con Flye y el genoma mitocondrial. Los fragmentos más cortos del genoma ensamblado con Flye presentan alta homología con el genoma mitocondrial, por lo que se infiere que esos fragmentos no corresponden al genoma nuclear (Alicia Costábile, comunicación personal).

El N50 es de 3,38 Mb para un genoma de 12,5 Mb y se predijeron 5.646 genes. El análisis de completitud realizado con BUSCO, muestra que el 97,93% de los genes del linaje *fungi*

fueron hallados completos y un 0,34% fue hallado fragmentado. Este resultado es mejor que el de Miniasm (91,38% de genes completos y 4,83% fragmentado).

El tamaño total del genoma obtenido con Flye de 12,52 Mb, es más parecido al tamaño de genomas reportados de levaduras emparentadas que el obtenido con Miniasm (14,2 Mb) (Tabla 1) (Douglass *et al.* 2018).

Al comparar la integridad y la contigüidad de los genomas, el obtenido con Flye es sensiblemente mejor en ambas características respecto al Miniasm. Sin embargo, para este último programa el ensamblado también es de buena calidad.

Se compararon los ensamblados con el programa MUMer. En la Figura 15 se observa como los dos ensamblados son bastante similares entre sí, confirmando que ambos son de buena calidad. Puede observarse que los 5 *contigs* largos de Flye se corresponden con los 7 *contig* más largos de Miniasm (Figura 15, Tabla 9). Los *contigs* cortos de un ensamblado se corresponden con los *contigs* cortos del otro, que corresponde a la línea vertical que se observa en el gráfico.

Tabla 9.	Correspondencia	entre los	contia	de los	aenomas	ensamblados.
	00110000011010101	0	contrag .		genenne	0110011101010101010101

<i>Contig</i> Flye	Largo (pb)	<i>Contig</i> Miniasm	Largo (pb)
contig_1	1.616.677	utg000004	1.615.247
contig_4	3.441.690	utg000001	3.452.613
	1.698.012	utg000002	1.021.108
contig_5		utg000006	678.372
contig_6	2.364.215	utg000007	2.377.647
scaffold_2	2 205 022	utg000003	2.570.719
	3.385.022	utg000005	868.091



Figura 15. Comparación entre los genomas ensamblados.

La mayor diferencia entre ambos ensamblados se presenta a nivel de los fragmentos más cortos, correspondientes al ADN mitocondrial. Puede que esté allí la diferencia entre la predicción inicial de genes para ambos ensamblados realizada por el programa *Quast* (6660 para Miniasm y 5646 para Flye). Este programa no genera un archivo de coordenadas génicas para los genes predichos, por lo que para poder verificar que la diferencia en la cantidad de genes radique en esto se tendría que realizar la anotación para el genoma ensamblado con Miniasm y de esta manera ver cuantos genes predice en cada fragmento largo (posiblemente cromosomas) y en los fragmentos más cortos (correspondientes a ADN mitocondrial) y comparar con el ensamblado con Flye. Ambos programas difieren en las exigencias al momento de enmasacarar repetidos (Li 2016, Kolmogorov *et al.* 2019). Se debería investigar el número y tamaño de las regiones colapsadas y repetitivas en ambos ensamblados.

Como resultado de la comparación de ambos ensamblados, se optó por utilizar el obtenido con Flye para el cual se realizó un pulido con el programa Pilon. Pilon dió como resultado la corrección de 137 poliformismos de nucleótido único (SNPs), 1793 insersiones pequeñas conformando un total de 6911 bases y 119 deleciones pequeñas conformando un total de 195 bases. No se encontraron bases ambiguas. El resultado del pulido no dio diferencias mayores en las métricas de calidad respecto al genoma ensamblado original (Tabla 10).

Tabla 10. Diferencias ent	tre los parámetros de	l ensamblado con F	Flye luego y antes del	pulido con Pilon.
---------------------------	-----------------------	--------------------	------------------------	-------------------

			Fragmento más	GC%		BUSCO (% genes)	
Ensamblado	Fragmentos	Largo (pb)	largo (pb)	(%)	N50 (pb)	Completos	Parciales
Sin pulir	15	12.522.592	3.439.421	36,42	3.383.105	97,93	0,34
Pulido	15	12.529.621	3.441.690	36,40	3.385.022	97,93	0,34

4.2.3. Anotación de genes

Con el programa Braker y los datos de RNAseq se anotaron 5.230 genes que codifican para 5.378 proteínas en las condiciones ensayadas. De esos genes 5.082 presentan una única isoforma, 132 presentan 2 isoformas y 16 presentan 3 o más isoformas.

Completitud del genoma

Se evaluó la completitud de la anotación del genoma comparando los resultados de BUSCO para los genomas reportados de *P. kudriazevii* (Tabla 1) que presentaban modelos génicos disponibles públicamente. Para la búsqueda se seleccionaron los linajes específicos de *I. terricola: Ascomycetes, Saccharomycetes* y *Saccharomycetales.* En la Tabla 11 se puede observar que el genoma de *I. terricola* tiene la misma cantidad de *contigs* que el genoma CBS573, donde confirman la presencia de 5 cromosomas (Douglass *et al.* 2018) en *P. kudriavzevii*. El resto de los genomas reportados presentan ensamblados más fraccionados. Además, para todos los linajes utilizados, el genoma de *I. terricola* presenta un porcentaje mayor de genes completos respecto a los genomas reportados. El genoma de CBS573 es comparable al de *I. terricola* en los parámetros evaluados: número de *contigs*, genes, secuencias codificantes (CDS) y porcentaje de genes BUSCO para cada linaje.

		lssatchenkia terricola	GCA_00305444 5 (CBS573)	GCA_00216677 5 (Ckrusei653)	GCA_00198 3325 (129)	GCA_0007644 55 (SD108)
Contig	IS	5	5	11	260	2482
Gene	S	5230	5366	5130	5470	7107
CDS		5378	5154	4949	5470	7107
Linaje			Resultado	os BUSCO		
Ascomycete	Completos	92,3%	91,7%	86,3%	79,6%	82,4%
	Fragmentados	3,1%	4,3%	7,1%	11,3%	9,8%
	Ausentes	4,6%	4,0%	6,6%	9,1%	7,8%
	Total genes buscados			1315		
Saccharomycetes	Completos	89,2%	90,4%	85,1%	77,3%	80,3%
	Fragmentados	4,3%	4,0%	6,4%	10,9%	9,4%
	Ausentes	6,5%	5,6%	8,5%	11,8%	10,3%
	Total genes buscados			1759		
Saccharomycetales	Completos	80,9%	80,0%	74,7%	66,6%	71,0%
	Fragmentados	8,5 %	9,4%	11,9%	14,2%	12,8%
	Ausentes	10,6%	10,6%	13,4%	19,2%	16,2%
	Total genes buscados			1711		

Tabla 11. Comparación entre genomas reportados de P.	kudriavzevii con anotación de genes y los genes
anotados para I. terricola	

4.2.3.1. Anotación funcional

InterProScan

De las 5.378 proteínas anotadas, a un 97% (5.216) de ellas se les predijo algún dominio mediante InterProScan: al 87% (4.676) se les predijo un dominio IPR, 81% (4.379) fueron identificadas con la base de datos Pfam y 82% (4.417) con la base de datos PANTHER. Se identificó péptido señal en un 6,5% (349) de las proteínas y 18,6% (1000) proteínas presentan regiones transmembrana. Del total de proteínas al 67,8% (3.644) le fue asignado un término GO (Figura 16).



Figura 16. Diagrama de Venn para los resultados de InterProScan realizado con <u>http://www.interactivenn.net/</u> (Heberle et al. 2015).

Búsqueda de homólogos en bases de datos

Para determinar la cantidad de genes que presentan genes homólogos en otros organismos se realizó un blastp contra la base de datos Swissprot de proteínas curadas. Se encontró que del total de proteínas 83,8% (4.508) se alinearon con alguna proteína de la base de datos, de ellas 2.605 pertenecen a levaduras.

Análisis de las vías KEGG

Se realizó el análisis de las vías KEGG y se determinó qué genes se encuentran involucrados en las distintas vías metabólicas y cuántas de estas vías se encuentran completas. Las principales vías metabólicas (metabolismo de carbohidratos, lípidos, aminoácidos, etc.) se encuentran completas, lo cual confirma la calidad del ensamblado y de la anotación de genes. Además, de interés en la producción de aromas, se identificaron 16 genes involucrados en la biosíntesis del esqueleto terpenoide y 2 genes involucrados en la biosíntesis de sesquiterpenoides y triterpenoides. Las vías de biosíntesis de isoprenoides C5 y C10-C20 (Figura 17) se encuentran completas.

TERPENOID BACKBONE BIOSYNTHESIS



Figura 17. Vías de síntesis de terpenos. Extraído de KEGG.

De las 5.378 proteínas predichas en *I. terricola* 2.834 fueron asignadas a términos de ortología KEGG y clasificadas según la jerarquía Brite. Entre éstas, 1.164 fueron identificadas como enzimas y 43 como glicosiltransferasas (Figura 18). A modo comparativo, se realizó el análisis para los organismos reportados *S. cerevisiae* y *P. kudriavzevii*. Se puede observar en la Figura 18 que la cantidad de genes de cada clase en 3 organismos son comparables, reforzando la calidad de la anotación realizada.



Figura 18. Comparación términos KEGG entre I. terricola, S. cerevisiae y P. kudriavzevii. Se muestra los genes clasificados para cada organismo en las categorías Brite de interés.

4.2.3.2. Búsqueda de glicosidasas

Para una búsqueda primaria de la presencia de glicosidasas en el genoma de *I. terricola* se utilizó el programa InterProScan. Este programa logró identificar 40 proteínas como glicosidasas dentro de las 5.378 proteínas anotadas.

Se utilizó el programa dbCAN para la identificación de proteínas que actúen sobre carbohidratos. Este programa tiene disponible 3 algoritmos distintos para la búsqueda (Diamond, HMMER y Hotpep). Se asumió que aquellas proteínas reconocidas por más de una herramienta tienen una mayor probabilidad de poseer la función asignada. Para lo cual, del total de proteínas anotadas, se identificaron 118 proteínas que actúan sobre carbohidratos (Figura 19), de las cuales 37 fueron identificadas como glicosilhidrolasas. Cada una de estas enzimas fue asignada a una familia de glicosilhidrolasas.



Figura 19. Identificación de enzimas mediante las 3 herramientas de dbCAN. Se consideraron reconocidas aquellas proteínas identificadas por al menos dos de las tres herramientas utilizadas (Diamond, HHMER, Hotpep).

4.2.3.2.1. Identificación de β-glucosidasas <u>i) Búsqueda de enzimas β-glucosidasas</u>

Al realizar un Blast de las proteínas anotadas contra β -glucosidasas descargadas de NCBI (Materiales y métodos 3.2.4.2), diez secuencias codificantes de *I. terricola* presentaron homología con al menos una de las β -glucosidasas reportadas (Tabla 12). El *query* son las proteínas anotadas, cinco de ellas tienen una cobertura de la secuencia aminoacídica igual o mayor a 88% (945, 989, 3296, 3698, 4598), y para cuatro de ellas la proteína con mejor homología es de *P. kudriavzevii*. Los porcentajes de identidad con la secuencia de la β -glucosidasa "homóloga" para estas cinco proteínas van desde el ~40% al ~76%. Tres enzimas presentan una cobertura de la secuencia entre 51 y 68% (1025, 1375 y 4410) y la enzima 2300 presenta una cobertura menor (36%), estas cuatro enzimas presentan porcentajes de identidad con el gen homólogo que van desde el ~24% al ~32%. Para las proteínas 1375 y 3832 los dominos Pfam e InterPro predichos corresponden a una región de repetidos de leucina. Para la mayoría de los resultados de blast se obtuvieron buenos valores de e-value (entre 0 y e-8).

Con el resultado de blast contra la base de Swissprot se encontraron 5 enzimas identificadas como β -glucosidasas (989, 1395, 3296, 3698, 4180). 3 de ellas (989, 3296 y 3698) se identificaron mediante InterProsScan con dominio Pfam β -glucosidasa.

De las 12 enzimas identificadas por algún método como β -glucosidasas, 5 no tienen péptido señal. Las restantes 7 sí tienen péptido señal, que presenta entre 17 y 22 aminoácidos de extensión.

Se asignó mediante dbCAN a que familia de glicosilhidrolasas pertenece cada enzima y se encontró que se distribuyen en 4 familias: GH3, GH5, GH17 y GH132. Las dos proteínas con repetidos de leucina no fueron identificadas por dbCAN por lo cual se descartaron como glicosidasas a pesar de los resultados obtenidos por blast con proteínas descargadas de NCBI.

 Tabla 12. Resultados de la búsqueda de 6-glucosidasas mediante Blast e InterProScan. Query – proteínas anotadas para I. terrícola.

	gen	g945.t1	g989.t1	g1025.t1	g1375.t1	g2300.t1	g3296.t1	g3698.t1	g3832.t1	g4410.t1	g4598.t1	g4180.t1		g1395.t1
Pfam	Código	PF03856	PF00332	PF00332	PF12799	-	PF03856	PF03856	PF13855	-	PF00933	PF00150		PF00150
	Descripción	Beta- glucosidase (SUN family)	Glycosyl hydrolases family 17	Glycosyl hydrolases family 17	Leucine Rich repeats	-	Beta- glucosidase (SUN family)	Beta- glucosidase (SUN family)	Leucine rich repeat	-	Glycosyl hydrolase family 3 N terminal domain	Cellulase (glycos family 5)	syl hydrolase	Cellulase (glycosyl hydrolase family 5)
Inter Pro	Código	IPR005556	IPR000490	IPR000490	IPR001611, IPR003591, IPR025875, IPR032675	IPR017853	IPR005556	IPR005556	IPR001611, IPR003591, IPR025875, IPR032675	IPR017853	IPR017853	IPR018087, I IPR001547	IPR017853	IPR001547
	Descripción	SUN family	Glycoside hydrolase family 17	Glycoside hydrolase family 17	Leucine- rich repeat	Glycoside hydrolase superfamily	SUN family	SUN family	Leucine- rich repeat	Glycoside hydrolase superfamily	Glycoside hydrolase superfamily	Glycoside G hydrolase, H family 5, s conserved site	Glycoside hydrolase superfamily	Glycoside hydrolase, family 5
		No	1-22	1-16	No	1-18	1-17	1-21	No	1-22	No	1-18		No
Blast NCBI	Código	XP_029323 920.1	XP_500465. 1	XP_500465. 1	XP_024716 243.1	XP_500465. 1	XP_029319 821.1	XP_029322 839.1	XP_024716 243.1	XP_500465. 1	XP_029321 798.1			
	Descripción	uncharacter ized protein C5L36_0C0 2630	YALIOB0356 4p	YALIOB0356 4p	hypothetica l protein C7M61_00 1227	-	uncharacter ized protein C5L36_0A0 9360	uncharacter ized protein C5L36_0D0 1000	hypothetica l protein C7M61_00 1227	YALIOBO356 4p	uncharacte rized protein C5L36_0C0 2630			
	Organismo	Pichia kudriavzevii	Yarrowia lipolytica	Yarrowia lipolytica	Candida pseudohae mulonis	Yarrowia lipolytica	Pichia kudriavzevii	Pichia kudriavzevii	Candida pseudohae mulonis	Yarrowia lipolytica	Pichia kudriavzevii			
	% cov query	100.0	100.0	60.0	51.0	36.0	95.0	100.0	34.0	68.0	88.0			
	evalue	3.42e-152	1.08e-113	1.06e-12	4.23e-09	5.76e-08	2.14e-150	0.0	1.84e-40	1.82e-19	8.10e-146			
	% ID	75.7	51.7	24.5	32.6	24.1	65.1	72.8	40.0	28.2	40.6			
Blast Swiss	Código	YM79_YEAS T	BGL2_YEAS T	SCW4_YEAS T	SDS22_YEA ST	SCW11_YE AST	UTH1_YEAS 7	SUN4_YEAS T	YFT6_SCHP O	SCW4_YEA ST	NAGZ_PHO PR	EXG_PICAN		EXG_YARLI
prot	Descripción	Uncharacte rized protein	Glucan 1,3- beta- glucosidase	Probable family 17 glucosidase	Protein phosphatas e 1 regulatory subunit	Probable family 17 glucosidase	Probable secreted beta- glucosidase	Probable secreted beta- glucosidase	Uncharacte rized leucine-rich repeat- containing protein	Probable family 17 glucosidase	Beta- hexosamini dase	Glucan 1,3-beta	ı-glucosidase	Glucan 1,3- beta- glucosidase
	Organismo	Saccharomy ces cerevisiae	Saccharom yces cerevisiae	Saccharom yces cerevisiae	Saccharom yces cerevisiae	Saccharom yces cerevisiae	Saccharom yces cerevisiae	Saccharom yces cerevisiae	Schizosacch aromyces pombe	Saccharom yces cerevisiae	Photobacte rium profundum	Pichia angusta		Yarrowia lipolytica
	% cov query	95.0	94	60	85	36	72	100	31	70	26	96		74
	evalue	1.82e-96	2.42e-132	2.07e-69	6.73e-98	9.26e-111	2.72e-134	3.34e-157	7.15e-44	1.37e-109	4.72e-19	0.0		3.37e-128
	% ID	58	61.6	44.8	54.5	58.3	69.5	61.5	42.0	57.0	26.7	63.5		49.9
Cazy		GH132	GH17	GH17	-	GH17	GH132	GH132	-	GH17	GH3	GH5		GH5

ii) Identificación de la β-glucosidasa purificada por nuestro grupo

La β -glucosidasa a partir de la cual surgió el interés por *l. terricola* fue purificada por el grupo de la Dra. González-Pombo (González-Pombo *et al.* 2011; de Ovalle *et al.* 2017) y se contaba con datos de espectrometría de masa de la misma (Stefani de Ovalle, comunicación personal). De estos resultados se encontró que 4 péptidos coinciden con la proteína 4180, identificándose por lo tanto el gen codificante para dicha β -glucosidasa. Esta proteína fue identificada como β -glucosidasa por los resultados del Blast contra *Swissprot*, pero no con el Blast contra secuencias de β -glucosidasa de NCBI ni a partir de los dominios identificados con el programa InterProScan. Sin embargo, fue asignada a la familia GH5 por dbCAN.

4.2.3.2.2. Búsqueda de otras enzimas de interés enológico

Otras enzimas de interés en la liberación de aromas a partir de precursores glicosilados son las arabinosidasas, ramnosidasas y apiosidasas (ver introducción).

Para lo cual se procedió a hacer una búsqueda por Blast de secuencias descargadas del NCBI (Materiales y métodos 3.2.4.3). Se encontró 1 arabinosidasa y no se encontraron ramnosidasas. Con el método de búsqueda utilizado no se hallaron en el NCBI apiosidasas anotadas (Tabla 13).

Enzima	Genes anotados (NCBI)	Gen	Cobertura del gen	Código	Descripción	Organismo	Cobertura del gen	evalue
β-glucosidasas	234	Ver tabla	beta-glucosida	asas (Tabla 4)				
Arabinosidasas	116	g952.t1	87%	XP_031009954.1	Xylosidase/arabinosidase	Lachnellula hyalina	60%	8.72e-39
Ramnosidasas	17	-	-	-	-	-	-	-
Apiosidasas	0	-	-	-	-	-	-	-

Tabla 13. Búsqueda de enzimas de interés enológico mediante Blast.

Se realizó una búsqueda con InterProScan para cada grupo de enzimas de interés para determinar qué dominios IPR les eran asignados. Se identificaron los dominios característicos de cada grupo y se buscó si esos dominios se encontraban presentes en alguna proteína anotada en *I. terricola*. Se encontraron pocos dominios característicos, los dominios asignados correspondían a glicosidasas (muy general) o a otras regiones que no son de interés para la actividad enzimática (por ejemplo: Transcription factor domain, Immunoglobulin-like fold, Leucine rich repeat, etc). Con los dominios característicos hallados no se encontraron resultados distintos a los del blast.

4.2.4. Selección de β -glucosidasas para su clonado

Las secuencias de las diez proteínas identificadas mediante Blast como β -glucosidasas fueron sometidas a alineamiento (Figura 20) y a un análisis filogenético (Figura 21), para analizar las relaciones evolutivas entre los distintos genes.

	10	20	30	40	50	60	70				
al180 ±1	.			• • • • • • • •		1	• • • • • • • •				
g4100.01											
g1595.01	MFTRVSENSGVNPTHDIGKLLFAGFOGTDFSNDCHAAKLIRDHGVHCFLL										
$q_{45} = 0.01$	HE TRADEWOOANE THD TOKINE VOL ÖATDE ONDOUWUVPTKDUOAUOL TH										
a3296 ±1	MKE SO	LSVI.TFAAL									
a3698 ±1	MKT.ST	TSLLATTSL	STTVK								
α989.t1	MKLSVASTTALATSLAAGPV										
g2300.t1	MNFKSLFSI	TACLSVAKA	LPIPIPOPD	IVTVEDISHTT	TYVPVAEVI	ISNGITTIE	LTTIAPTTSPI				
g4410.t1	MLFKNYIAN	TLLLACAAT	VLASP								
q1025.t1	MKLTV	WLPIFIGTA	LAAPLF								
2											
	260	270	280	290	300	310	320				
g4180.t1				V	KLQKKNSRW	DYESEKIYGV	N				
g1395.t1	VQEYFNKNLQSQLKAR	SMEVENHIN	NIKLGESNA	YNEEMDDSSIT	RDSINLATE	DYKATKVNGI	N				
g4598.tl	AVTFPGLGNSFISNIT	EAPVVIDDM	AQFVNVTSK	PFKAMIDHDII	DSVLVTAAF	LEAIQDDDS P	AFLNPKIVDGL				
g945.t1					MV/	VYNTEDSPS-					
g3296.tl	ETRTVVVDAADVTATA	AAIDVAAAT	IDSSESTTA	PTTLATSTTSA	ASAETSTGA	TEKKSTSSGS	G				
g3698.tl	<u>Q</u> TSTSTLNPATQVSTE	SASASSTST	STSESSSSA	STSSSSSSSES	ESSSSTSSS	SSSSSSSSSSS	S				
g989.t1					- <u>SA</u> IASLGF	DLGVQDNGGA					
g2300.t1	SSAVVETTSSASSSSE	SSATSTESS	ATSTESPAT	STESSTTSTES	PTTATTSTS	SFSSSSSSSS	SSSFSDFVTTW				
g4410.t1	TLGGENTELASAPAAT	QTTISSQAD	TYTAYAYS	VATASSSDSSA	AVAATSTES	TAAASSTASS	D				
g1025.t1	SAVAPTTASATAASPT	LSLMLDSGS	SVIDHNVAS	EVTLVTSTTTS	SSEESPSSS	DYEEYSSTSS	A				
	510	520	530	540	550	560	570				
1100 11				• • • • • • • •			1 • • • • 1 • • • • 1				
g4180.tl	EPYIVPSLFEAFGG	EDSTRPVDE	THE		-SATIGKER	AEEVL					
g1395.tl	EPFITPSLYEAASVGGTEASTPVDEYHYCKSLGFTEAHFRLCKSLGFTEAHFRL										
g4598.tl	SNFNCSVNYVKYSPYGLTVSIKEAISNCSMLIFFCGNTSILCYQAELLKQISLLLQRDGKYLI										
g945.tl											
g3296.tl	STFEDGTIPCSEFPSSIDGMVSIDWLGFGGWASVMAMDGSTSE										
g3696.LI	QEFQDGT1KCSDFPSG-QGVVSLDYLGFGGWSGLYHPGTTATG										
g909.L1											
$g_{2300.01}$	DNIDATTSSTSSTSSATSAAATSSSSGILTKVPGTIVISPYENDGGCKSYATVLSDLTLIK										
$g_{4410.01}$	TECCOTCONNECCCO				-CKSLADIA		FLIS				
g1025.01	110001000000000	r indiridiride di			ORDHDELY						
	7.00		700	200							
			/80	/90		1	820				
g4180.t1	SASEYNDVVIGIELVN	EPLGPILND'	FELKEYYNN	GYEI		VRKEGD	VPVIIHDAFLQ				
g1395.t1	GGPEYNDVVIGYQNLN	EPLVGAYPM	EKILRFDEI	TYDMF		RKRNSN	NWFVIHDAFID				
g4598.tl	SKNNDVNNDDCNCNGN	NSNVGSKRNS	SSTAKRKSW	KEK PWVVEAFD	PHRDVSAMS	MIRTSTISDI	YYPLNNDLLER				
g945.t1	PGDESMLIPTNVLGGD	TSPL				SVPGTD	YWASTA				
g3296.tl	PGSEDMVVPTELQAGS	TEPL				SVVNED	NYFDWEGKKTS				
g3698.tl	PGTENMVIPTIVEAGT	ENYL				TTVDED	NYYTWEGLKTS				
g989.tl	ADIPVGTVDSWNILVN	GEAY				DVIEQS	TVMYAN				
g2300.t1	TG-DVITSEPPVTFEN	NPEL				CTESEI	DFVGIN				
g4410.t1	TG-PVVSVDTHVAILN	NPGL				CDIS	DYMAFN				
g1025.t1	TG-NVVTVDTVPVFED	NPSL				CDHV	DYVTVN				
	1010	1020	1030	1040	1050	1060	1070				
al180 +1			· · · · · · · ·	• • • • • • • •		1	1 • • • • [• • • •]				
g4100.L1	FETERT FRACENER			W							
$a_{1598} + 1$	IDALOKICI SELIOD	ARDAFFICO	ר ד.ד. ג ייג⊖יייי	DEEVI-FSCMV	AKUCEVIKA	ETELCANDAND	BRSTTTVAVT				
a945 + 1	VNENDNEGIBLECDDE	SKCNGGTCF	F		CY						
a3296 +1	PNYNVKTVATDC	STINCD-CS	· · · · · · · · · · · · · · · · · · ·	TNCFF							
a3698 +1		SSVVNCECV	· · · · · · · · · · · · · · · · · · ·								
a989.+1	FWKEGTCATEC	WGVNTEVER	- A	FDEDW							
g2300.t1	ALO-SIFDV	VGTDVTILT	- [FDDYW							
g4410.t1	AIS-SIDET	CGSDTLVFT	A	FDDI.W							
g1025.t1	AIK-AIVENSD	VAENTLMFT	- [YNDYW							

Figura 20. Alineamiento mediante el programa muscle de las secuencias aminoacídicas de las diez 6-glucosidasas halladas. Las secuencias subrayadas con la línea negra corresponden a los péptidos de señalización.



Figura 21. Árbol filogenético para las diez *B*-glucosidasas halladas realizado con PhyML. En amarillo se encuentran agrupadas las *B*-glucosidasas pertenecientes a la familia GH132, en verde las pertenecientes a la familia GH17 y en anaranjado las pertenecientes a la familia GH5.

En la Figura 21 se observa que las diez β -glucosidasas se separan en cuatro grupos de acuerdo a la familia asignada por el dbCAN (marcadas en color en la Figura 21). En el grupo amarillo se encuentran agrupadas las de la familia GH132 y en el grupo verde se encuentran las β -glucosidasas pertenecientes a la familia GH17. La β -glucosidasa purificada y la 1395 pertenecen a la familia GH5 y se encuentran agrupadas en el grupo anaranjado. La enzima 4598 pertenece a la familia GH3 y no se agrupa con ninguna otra β -glucosidasa. Además se realizó otro árbol filogenético con FastTree y el perfil del mismo coincide con el que se muestra, apoyando los resultados obtenidos en ambos casos.

Como se puede observar en la matriz de identidad para las enzimas (Figura 22), la proteína 4598 es la que menor similitud presenta con el resto de las otras proteínas y es la que se encuentra en el árbol separada del resto. Además los valores de correspondencia entre proteínas de una misma familia son los más altos en cada caso.

Seq->	g4180 GH5	g1395 GH5	g4598 GH3	g945 GH132	g3296 GH132	g3698 GH132	g989 GH17	g2300 GH17	g4410 GH17	g1025 GH17
g4180 GH5	ID	0,306	0,023	0,06	0,055	0,062	0,081	0,041	0,066	0,078
g1395 GH5	0,306	ID	0,043	0,054	0,057	0,063	0,061	0,044	0,079	0,077
g4598 GH3	0,023	0,043	ID	0,027	0,042	0,044	0,033	0,056	0,043	0,038
g945 GH132	0,06	0,054	0,027	ID	0,192	0,159	0,072	0,036	0,059	0,055
g3296 GH13	0,055	0,057	0,042	0,192	ID	0,469	0,119	0,068	0,158	0,14
g3698 GH13	0,062	0,063	0,044	0,159	0,469	ID	0,081	0,104	0,141	0,149
g989 GH17	0,081	0,061	0,033	0,072	0,119	0,081	ID	0,084	0,159	0,189
g2300 GH17	0,041	0,044	0,056	0,036	0,068	0,104	0,084	ID	0,143	0,16
g4410 GH17	0,066	0,079	0,043	0,059	0,158	0,141	0,159	0,143	ID	0,286
g1025 GH17	0,078	0,077	0,038	0,055	0,14	0,149	0,189	0,16	0,286	ID

Figura 22. Matriz de identidad para las distintas 6-glucosidasas. Se muestra en escala de colores la identidad entre las secuencias aminoacídicas. Rojo: menor identidad, Verde: mayor identidad.

Para seleccionar las dos enzimas para ser expresadas se eligió una β -glucosidasa de la familia GH17 (989) y otra de la familia GH132 (3296). Ambas pertenecen a familias distintas a la familia GH5 (familia a la que pertenece la β -glucosidasa purificada anteriormente por nuestro grupo (4180)), es de interés que pertenezcan a distintos grupos para estudiar proteínas que pudieran diferir en el mecanismo de acción o afinidad de sustrato. Además, los dos genes selecccionados fueron identificados como genes codificantes de β -glucosidasas por los tres métodos de búsqueda utilizados.

En la siguiente sección se presentarán los resultados relativos al clonado de las secuencias codificantes de las enzimas 989 y 3296 en vectores de expresión de E. coli. Además se detallan los resultados de los ensayos de expresión de la proteína recombinante 989 con péptido señal fusionada a His-tag y a His-tag/DsbC y de la proteína 989 sin péptido señal fusionada a His-tag.

4.3. Clonado de las enzimas 989 y 3296 para su expresión recombinante en *Escherichia coli*

Para un abordaje molecular es necesario conocer si los genes de estas enzimas presentan intrones. Esto determina si se puede amplificar la secuencia codificante a partir de ADN genómico (ADNg) o si es necesario partir de ARNm y amplificar la secuencia mediante RT-PCR. En este caso se pudo verificar mediante la anotación del genoma que los genes que codifican para las enzimas seleccionadas no presentan intrones. Por lo tanto se realizó el diseño de cebadores para la amplificación de las distintas construcciones a partir de las secuencias anotadas para las proteínas 989 y 3296 (ver secuencias y cebadores en Anexo).

4.3.1. Clonado en vector pGM-T

Inicialmente se plantó el clonado del gen de las enzimas en un vector pGM-T (*Tiangen*) para generar un respaldo de estos en un vector de propagación. Para ello, las secuencias correspondientes a las secuencias codificantes de las enzimas 989 e 3296 fueron amplificadas por PCR (ver materiales y métodos) utilizando como molde ADNg. Se obtuvieron fragmentos del tamaño esperado (933 y 1098 pb respectivamente, Figura 23). En un primer paso, se clonaron los fragmentos amplificados y purificados en un vector pGM-T.



Figura 23. Amplificación de las secuencias codificantes para las enzimas a) 989 b) 3296, a partir de ADNg. MPM: GeneRuler[™] 1kb Plus DNA Ladder

Posteriormente se seleccionaron dos de los clones resultantes para cada construcción, y una vez extraídos los plásmidos éstos fueron chequeados mediante digestión con enzimas de restricción. Se realizó un ensayo de digestión simple con la enzima de restricción EcoRI, que escinde el inserto y también una digestión doble con EcoRI-EcoRV, ya que esta última enzima corta dentro del inserto y por tanto contribuye a la identificación de los clones buscados.



Figura 24. Verificación de la inserción de la secuencia codificante para la enzima 989 en el vector pGM-T. a) Digestión con EcoRI. Tamaño esperado: 933 pb b) Digestión con EcoRI y EcoRV. Tamaños esperados: 657 y 336 pb. 1 − Control negativo, 2 y 3- 989 en pGM-T, A- sin digerir, B- digerido, MPM: GeneRulerTM DNA Ladder Mix

Efectivamente, al digerir los plásmidos extraídos de los clones correspondientes a la enzima 989 (pGMT-989.1 y pGMT-989.2) con EcoRI se escinde un inserto del tamaño adecuado, 933 pb (Figura 24a). La obtención de dos bandas de 657 y 336 pb al digerir con ERI-ERV (Figura 24b), reafirma la identidad del fragmento clonado. La presencia del inserto se verificó finalmente por secuenciación.

Se señala que el plásmido utilizado como control negativo (extraído de una colonia azul) posee inserto, lo que no es esperado. Evidentemente se cometió un error en la apreciación del color de la colonia, o se arrastró material de alguna colonia blanca vecina.

En el caso de los plásmidos extraídos de los clones correspondientes a la enzima 3296 (pGMT-3296.1 y pGMT-3296.2), la digestión doble con ERI y ERV (Figura 25) arroja el perfil esperado: dos bandas de 351 y 676 pb (se genera una tercera banda de 74 pb que no se visualiza en el gel), por lo que ambos clones fueron enviados a secuenciar para confirmar su identidad.



Figura 25. Verificación de inserción de la secuencia codificante para la enzima 3296 en el vector pGM-T. Digestión con EcoRI y EcoRV. Tamaños esperados: 657 y 336 pb. 1 – Control negativo, 2 y 3- 3296 en pGM-T, A- sin digerir, B- digerido. MPM: GeneRulerTM DNA Ladder Mix

En todos los plásmidos la secuenciación puso en evidencia la presencia de mutaciones (2-3 en cada plásmido). Esto puede deberse a que la polimerasa *Ranger* (*Bioline*) utilizada en la amplificación no es de alta fidelidad.

En consecuencia, se procedió a la amplificación de las secuencias de interés con una enzima de alta fidelidad (*KapaHiFi, BioSystems*). Se procedió además a intentar el clonado directo en los vectores de expresión pET-32a y pET32a-DsbC por el método de RF-*cloning* (Reid *et al.* 2002).

4.3.2. Clonado en vectores de expresión pET-32a mediante RF-cloning

Para realizar el clonado con el método de RF-*cloning* se utilizó el procedimiento descrito en Correa *et al.* 2009. Los plásmidos donde se clonaron las enzimas fueron generados en ese trabajo y los mismos poseen en el sitio de inserción la secuencia correspondiente a la GFP (ver Vectores en Materiales y métodos). Se diseñó el método de clonado de modo tal que la secuencia que codifica la GFP es sustituida por el inserto de interés.

En una primera reacción de PCR se amplificó la secuencia codificante de cada enzima con extensiones en sus extremos que hibridan con la región del vector donde se realizará la inserción. Los *"megaprimers"* generados en esta reacción presentaron los tamaños esperados (tamaño de la secuencia de la enzima + 60 pb de las extensiones) de 993 pb (enzima 989) y 1158 pb (enzima 3296) (Figura 26).



Figura 26. Productos de la amplificación por PCR de los megaprimers correspondientes a las enzimas 989 y 3296 para el clonado mediante RF-cloning. MPM: NEB 1Kb DNA Ladder

En una segunda PCR, los megaprimers fueron introducidos en dos vectores diferentes: el vector de expresión pET32-a (permite el agregado del His-tag) y además una variante de este mismo vector que permite la síntesis de la proteína fusionada al His-tag y a la proteína DsbC (pET32a-DsbC). DsbC (Disulfide bond C) es una disulfuro isomerasa que puede favorecer el plegamiento correcto de proteínas que presentan puentes disulfuro necesarios para el correcto plegamiento de las mismas (Nozach et al. 2013) y por tanto para su producción como proteínas solubles (Arnau *et al.* 2005, Correa & Oppezzo 2011). Se seleccionó como proteína de fusión la DsbC particularmente ya que en la secuencia de la proteína 989 hay 4 cisteínas para las cuales se predijo la formación de 2 puentes disulfuro entre los residuos 39 y 261, y entre los residuos 67 y 308 (ver Figura 45 en Anexo). En la secuencia de la proteína 3296 hay 10 cisteínas para las cuales se predijo la formación de 5 puentes disulfuro entre los residuos 115 y 350, 150 y 160, 156 y 278, 185 y 201 y entre los residuos 221 y 336 (ver Figura 46 en Anexo). Para determinar si las proteínas a expresar presentan formación de enlaces disulfuro se utilizó el programa DiANNA (http://clavius.bc.edu/~clotelab/DiANNA/) con la secuencia aminoacídica de las proteínas.

Por razones de tiempo al comenzar la pandemia, se decidió continuar únicamente con la producción recombinante de la enzima 989. Algunos de los clones resultantes del RFcloning en los dos vectores mencionados (pET-32a-989.1 a .6 y pET-32a-DsbC-989.1 a .6) se analizaron por digestión con BamHI para verificar la presencia del inserto (993 pb) (Figura 27). Se realizó para cada caso la digestión del plásmido original como control.



Figura 27. Verificación de inserción de la secuencia codificante para la enzima 989 en a) vector pET32-a b) vector pET32a-DsbC mediante digestión con BamHI. MPM: MPM: NEB 1Kb DNA Ladder

El vector pET32-a de partida cuenta con un único sitio de corte para BamHI por lo que la digestión linearizó el plásmido. El vector pET32a-DsbC posee dos sitios BamHI (a los lados del sitio de inserción) por lo que la digestión liberó el inserto que tenía este plásmido (secuencia que codifica la GFP, de 711 pb). Para todos los clones analizados se observa la banda de aproximadamente 1000 pb correspondiente a la inserción de la secuencia que codifica la 989.

Se verificó además la presencia del inserto mediante amplificación por PCR (Figura 28). El control en cada caso es el plásmido sobre el cual se realizó el RF-*cloning*. En la Figura 7 se puede observar cómo para todos los clones analizados, la PCR amplificó un fragmento correspondiente a la secuencia codificante para la enzima 989, mientras que para los controles esta banda está ausente.



Figura 28. Verificación de inserción de la secuencia codificante para la enzima 989 mediante PCR en a) vector pET32-a b) vector pET32a-DsbC mediante PCR. MPM: NEB 1Kb DNA Ladder

Se enviaron 2 clones (plásmidos pET-32a-989 y pET32a-DsbC-989) de cada construcción para verificar su identidad por secuenciación. El resultado de la misma confirmó que la secuencia de la enzima 989 se encuentra inserta en el plásmido con ausencia de mutaciones respecto de la secuencia anotada, lo cual además confirma la calidad de la anotación. La secuenciación confirmó también la identidad de los vectores utilizados.

Se guardaron en glicerol células *E. coli* XL-1 transformadas con el plásmido como respaldo. Finalmente, se realizó con éxito la transformación de células *E. coli* BL-21 DE3, para la expresión de la enzima.

4.4. Expresión de la enzima 989 en E. coli

En un primer ensayo, se realizaron cultivos de células de *E. coli* BL-21 transformadas con el plásmido pET-32a-989 y en paralelo, como control de la expresión, se utilizó una cepa transformada con el plásmido pET-32a que porta la secuencia que codifica para la

proteína verde fluorescente (GFP). Los cultivos se llevaron a cabo en un volumen de 25 mL de LB, en matraces de 250 mL, a dos temperaturas diferentes, 30º y 37ºC.

A partir de cada uno de los cultivos, se preparó un extracto total (ET), a partir del cual se separó por centrifugación una fracción soluble (FS) y una fracción insoluble (FI). En la Figura 29 se muestra un análisis por SDS-PAGE de las proteínas presentes en cada uno de éstos.

Se puede observar que para ambas temperaturas, en el caso del control que expresa la GFP, aparece una banda en el ET y en la FS de aproximadamente 27 kDa, que coincide con el peso molecular esperado para la GFP, lo cual demuestra que tanto el sistema como las condiciones de expresión son adecuados. Esta banda está ausente en las muestras obtenidas del ensayo de expresión de las células portando el plásmido pET-32a-989, lo que confirma que el gen de la GFP fue sustituido por el gen de interés 989.

Sin embargo, al analizar las distintas fracciones para la expresión de la enzima 989, no se observan bandas que difieran respecto al control, ni que sugieran una sobreexpresión de alguna proteína. El peso calculado para la proteína 989 con el sitio TEV y el His-tag es de aproximadamente 35 kDa.



Figura 29. SDS-PAGE del cultivo de expresión de His-tag-989 en un volumen de 25 mL, inducción con IPTG 1 mM durante 5 horas. a) 30°C b) 37°C. MPM: AccuRuler RGB Prestained Protein Ladder

A partir de estos resultados, se realizó una serie de modificaciones al protocolo de expresión como se detalla en la Tabla 14. Estas modificaciones incluyeron variaciones del tiempo y concentración de inductor, escalado del volumen de cultivo, realización de cromatografías de afinidad por iones metálicos (IMAC) con la finalidad de purificar, concentrar y detectar la proteína si ésta se expresara a bajos niveles.

Modificación del protocolo de expresión/purificación	Justificación de la estrategia	Resultados (Anexo)
-Aumento del tiempo de inducción de 5 hs a <i>O.N.*</i> -Aumento de la concentración de IPTG usada de 1 a 5 mM -Temperatura 30ºC -Control: cultivo en paralelo sin el agregado de IPTG	El tiempo de inducción y la concentración de IPTG utilizadas pueden no ser suficientes para inducir la expresión de la proteína a niveles detectables. Se utilizó la temperatura de 30ºC ya que se observó mayor expresión de la GFP a esta temperatura que a 37ºC.	Figura 45
-Toma de alícuotas a distintos tiempos	La proteína, de estar expresándose podría estar siendo rápidamente degradada y por lo tanto no ser detectada en los geles realizados. Este ensayo también sirve para determinar el tiempo óptimo de expresión.	Figura 46
-IMAC de la FS del primer cultivo (25 mL, 37ºC, 5 hs inducción)	La proteína se expresa en niveles muy bajos para ser detectada por SDS-PAGE en la FS. Se realiza la cromatografía para concentrarla.	Figura 47
-Escalado a 100 mL de cultivo	La proteína se expresa en niveles muy bajos para ser purificada por IMAC a partir de un volumen de cultivo de 25 mL.	Figura 48
-IMAC de la FS del cuarto cultivo (100 mL a 37ºC, 5 hs de inducción)	La proteína se expresa en niveles muy bajos para ser detectada por IMAC en un cultivo de 25 mL. Se espera de esta forma concentrar una cantidad suficiente de proteína para su visualización.	Figura 49
*ON	1C h	

Tabla 14. Variaciones en el protocolo de expresión de la enzima His-tag-989 en células de E. coli BL-21.

*O.N. – overnight (inducción durante 16 horas)

Para el protocolo de monitoreo de la expresión (toma de alícuotas a distintos tiempos) se analizaron los ET para cada tiempo, se normalizó la alícuota tomada con la DO (densidad óptica) del cultivo para tratar de sembrar en todos los casos la misma cantidad de proteína.

Para ninguno de los protocolos descritos en la Tabla 14 se observaron diferencias significativas entre el cultivo inducido y el control sin inducir. Se realizó además la IMAC del cultivo control (sin inducir) para conocer el patrón de las proteínas de *E. coli* que podrían co-purificar con la proteína de interés (ver Figura 51 en el anexo) y facilitar así la diferenciación de éstas de la eventual expresión de la proteína de interés. Sin embargo, en las muestras obtenidas en los ensayos de inducción del gen de interés y posterior purificación por IMAC se observa únicamente el patrón de proteínas de *E. coli* (ver Figura 51 en el anexo) no detectándose nuevas bandas que correspondan a la proteína His-tag-989 (35 kDa).

Siendo todos estos intentos infructuosos, se procedió a ensayar la expresión de la enzima 989 como una fusión con His-tag y DsbC, que presenta las ventajas expuestas anteriormente (ver Clonado en vectores de expresión pET-32a mediante RF-*cloning*).

4.4.1. Expresión de la proteína 989 fusionada a His-tag y DsbC

La expresión de la proteína 989 fusionada a His-tag-DsbC se ensayó en diferentes condiciones, similares a las descritas para la enzima His-tag-989: cultivo inicial de 25 mL con 5 horas de inducción a 30º y 37ºC, cultivo *O.N.* a 30ºC y monitoreo a distintos

tiempos a 30°C. En ninguna de estas condiciones fue posible detectar una expresión diferencial significativa de ninguna proteína en ninguna de las fracciones provenientes del cultivo inducido (Figura 52, Figura 53 y Figura 54 respectivamente en Anexo).

Previendo la posibilidad de que la proteína se estuviera expresando a bajos niveles, se pasó a una escala de cultivo de 100 mL en matraz de 1 L, donde se probó la expresión durante 3 horas de inducción a 20º, 30º y 37ºC (Figura 30).





Para los cultivos a 20º y 30ºC no se observan diferencias respecto al cultivo sin inducir. Sin embargo, en la fracción insoluble del cultivo a 37ºC se observa, entre las pocas bandas presentes, una de mayor intensidad de un peso molecular de aproximadamente 53 kDa. Esta banda parece estar ausente en el control y su peso molecular se aproxima al esperado para la fusión His-tag/DsbC-989 (58 kDa).

Para evaluar si esta proteína pudiese estar presente en niveles no detectables en el conjunto de proteínas de la fracción soluble se sometió la misma a una IMAC. No se observa en los eluídos ninguna banda que corresponda al mismo peso molecular de la banda que se destaca en la fracción insoluble (Figura 30). Tampoco se observan otras bandas que pudieran ser de interés más allá del patrón normal de la IMAC para un cultivo de *E. coli* (Figura 51 en Anexo).

A continuación se procedió a realizar un *western-blot* anti His-tag para las distintas fracciones del cultivo a 37°C para determinar si la banda observada en la Figura 30 se corresponde efectivamente a la proteína His-tag-989 recombinante que está siendo expresada. Como se mencionó anteriormente, la construcción realizada permite la expresión de la proteína con una etiqueta de histidinas, la cual se usó para la detección

por *western-blot* utilizando anticuerpos anti-His. Se utilizó como control positivo una fosfatasa de *Mycobacterium* con His-tag de aproximadamente 60 kDa.



Figura 31. Western-blot del cultivo de expresión de His-tag/DsbC-989 en un volumen de 100 mL, inducción con IPTG 5 mM durante 3 horas a 37ºC. MPM: AccuRuler RGB Prestained Protein Ladder

El revelado de la inmunodetección muestra que únicamente en las fracciones provenientes del cultivo inducido se detectan proteínas, y que por tanto portarían un His-tag (Figura 31). En el extracto total se observan dos bandas, una a la altura de aproximadamente 53 KDa y otra de menor peso molecular cercano a los 23 KDa. Esta misma banda se observa en la fracción soluble, pero no en la insoluble. En la fracción insoluble se observa una banda correspondiente a 53 KDa, que coincide con la banda observada en la SDS-PAGE (Figura 30). Este ensayo permitiría concluir que, en las condiciones ensayadas, fue posible expresar una proteína cuyo peso molecular coincide con el esperado para la fusión de 989 a His-tag-DsbC, estando la misma en forma insoluble. La presencia de la banda de aproximadamente 23 kDa en la FS y en el ET podría ser un indicio de degradación de la proteína de interés por proteasas bacterianas. La ausencia de esta banda de menor masa molecular en la FI podrías deberse a que en los agregados insolubles la región digerida no se encuentra expuesta.

Para confirmar la identidad de la proteína se llevó a cabo una SDS-PAGE de la fracción insoluble y se recortó la banda de 53 kDa que fue enviada al servicio de espectrometría de masas del Institut Pasteur de Montevideo. En la Figura 32 se observa el espectro de masas de los fragmentos obtenidos por MALDI-TOF.



Figura 32. Fragmentos obtenidos del análisis por MALDI-TOF de la proteína de 54 KDa obtenida en la FI

4 de estos péptidos fueron secuenciados e identificados en la secuencia de Histag/DsbC-989 (Figura 33) pudiéndose así confirmar la identidad de la misma.

Matche	d peptides	shown in Bo	ld Red					
1	HHHHHHGSSG	MKKGFMLFTL	LAAFSGFAQA	DDAAIQQTLA	KMGIKSSDI	Q		
51	PAPVAGMKTV	LINSGVLYIT	DDGKHIIQGP	MYDVSGTAPV	NVTNKMLLK	Q		
101	LNALEKEMIV	YKAPQEKHVI	TVFTDITCGY	CHKLHEOMAD	YNALGITVR	ΩY.		
151	LAFPRQGLDS	DAEKEMKAIW	CAKDKNKAFD	DVMAGKSVAP	ASCDVDIAD	H		
201	YALGVQLGVS	GTPAVVLSNG	TLVPGYQPPK	EMKEFLDEHQ	KMISGKGSG	S		
251	ENLYFQGSMK	LSVASITALA	ISLAAGPVSA	IASLGFDLGV	QDNGGACKV	A		
301	SDYEADLAIL	APYSKMVKTY	AVSTCNTLEV	LGPVVESSGF	QLTLGVWPT	P		
351	QDKFTAEKEG	LQSYLPTISK	SAISSVLVGS	EALYRGDLTG	PELASYISE	I		
401	KSLLATINDK	DGNSYADIPV	GTVDSWNILV	NGEAYDVIEQ	STVMYANAF	P		
451	YWQGQTQQNS	SFSFFDDIMQ	ALQTVQTIKG	STDFDFYIGE	TGWPTGGSN	ΥY		
501	GASVPSVDNA	ETFWKEGICA	IRGWGVNIFV	FEAFDEDWKP	DISGSSVET	Н		
551	WGVFDSSRNL	KYSLDCDF						
Start	- End O	bserved	Mr(expt) 1	Mr(calc)	Delta	Miss	Sequence	
134	- 149 18	30.8706 18	29.8633 18	29.9094	-0.0461	0	K. LHEQMADYNALGITVR.Y	(Ions score 44)
150	- 155 7	66.3832 7	65.3759 7	65.4173	-0.0414	0	R.YLAFPR.Q (Ions sco	re 29)
371	- 385 15	51.7900 15	50.7827 15	50.8304	-0.0477	0	K.SAISSVLVGSEALYR.G	(Ions score 51)
386	- 401 16	92.8369 16	91.8296 16	91.8618	-0.0321	0	R.GDLTGPELASYISEIK.S	(Ions score 22)

Figura 33. Secuencia mostrando los péptidos de His-tag/DsbC-989 secuenciados por MS/MS que permitieron comprobar la identidad de la proteína de interés.

Con estos resultados se confirma que, efectivamente, se logró expresar la proteína 989 fusionada a His-tag/DsbC en un cultivo de 100 mL a 37°C con una inducción de 3 horas y que la misma se halla mayoritariamente en la fracción insoluble. A su vez, a través del *western-blot* se observan signos de degradación en la fracción soluble, lo cual explicaría

por qué no es posible detectar la proteína en esta fracción. No se observaron signos de expresión para las otras condiciones ensayadas.

4.4.2. Expresión de la proteína 989 sin el péptido de tránsito

Mediante la herramienta SignalP se identificó en la secuencia de la enzima 989 una secuencia de 22 aminoácidos predicho como péptido de tránsito (secuencias en Anexo). Si bien es un péptido de eucariota, estos no difieren mucho de los péptidos de señalización de los procariotas y podría estar siendo reconocido por *E. coli*. Es importante señalar que en el artículo Hao *et al.* (2020) los autores reportan la expresión extracelular de una β -glucosidasa eucariota y de origen fúngico expresada en *E. coli* BL-21. En ese trabajo, la remoción del péptido señal lleva a la expresión de la proteína con mayor actividad que la silvestre, pero que se encuentra insoluble, acumulada en los cuerpos de inclusión.

Dado que en los ensayos descritos en el punto anterior en ningún caso se analizó el sobrenadante en los cultivos de expresión realizados, no se puede descartar que la proteína haya sido secretada y que esta sea la razón por la cual no se detectó la proteína en algunas de las condiciones ensayadas. Para descartar esta posibilidad, se repitió la estrategia de clonado pero amplificando la enzima de modo de eliminar el péptido de tránsito. Esta variante se clonó en los vectores pET32-a-His-tag y pET32a-His-tag-DsbC.

La expresión de las bacterias transformadas con el vector pET32-a-His-tag-989 se ensayó en cultivos de 100 mL sometidos a 3 horas de inducción, (condiciones en que se observó expresión de His-tag/DsbC-989), probándose distintas temperaturas (20^o, 30^o y 37^oC) (Figura 34). Se realizó para cada temperatura un cultivo control sin inducir, pero por cuestiones prácticas para la visualización se muestra un solo control para 20^o y 30^oC ya que el perfil es el mismo para ambas temperaturas (ver Figura 56 en el anexo).



Figura 34. SDS-PAGE del cultivo de expresión de 989 sin péptido señal en un volumen de 100 mL, inducción con IPTG 5 mM durante 3 horas a 20º, 30º y 37ºC. MPM: AccuRuler RGB Prestained Protein Ladder

En dichos ensayos no se observan diferencias significativas en los cultivos a 20º y 30ºC respecto a los cultivos control. Sin embargo a 37ºC se observa en la FI una banda de peso molecular ligeramente inferior a 32 kDa (marcada con la flecha), que parece tener mayor intensidad respecto al cultivo control y a los cultivos a 20º y 30ºC. Con la

esperanza de recuperar algo de la proteína recombinante en la fracción soluble se decide someter la FS a una purificación por IMAC.

En la IMAC realizada (Figura 35) se observa una banda intensa entre 23 y 32 kDa. Si bien el tamaño aparenta ser menor que el de la banda señalada de la Figura 34, como llamó la atención su intensidad se decidió enviar esta banda para su análisis en el espectrómetro de masas.



Figura 35. Expresión de His-tag-989 sin péptido señal. IMAC de la FS del cultivo a 37°C, 3 horas de inducción con 5 mM de IPTG. P – Percolado, L1 – Primeros lavados, sin imidazol, L2 – Segundos lavados con 20 mM de imidazol, E – Eluídos. MPM: AccuRuler RGB Prestained Protein Ladder

El resultado del MALDI-TOF permitió la identificación de la proteína como SlyD, una peptidil-prolil cis-trans isomerasa de *E. coli*, reportada como una de las proteinas de *E. coli* comúnmente co-purificadas por IMAC. Esta es una proteína regulada por iones metálicos (Hottenrott *et al.* 1997) que contiene dos regiones ricas en histidinas, lo cual explica su interacción con la matriz (Bolanos-Garcia & Davies 2006).

En paralelo se realizó un *western-blot* de los ET, FS y FI obtenidos de los cultivos realizados en todas las condiciones de expresión ensayadas para eventualmente identificar la proteína expresada en alguna de las fracciones, que no sea apreciable en los geles de SDS-PAGE. Los resultados se muestran en la Figura 36.



Figura 36. Western-blot del cultivo de expresión de His-tag-989 sin péptido señal en un volumen de 100 mL, inducción con IPTG 5 mM durante 3 horas a 20º, 30º y 37ºC. MPM: AccuRuler RGB Prestained Protein Ladder

Como control positivo se utilizó la FI del cultivo de expresión de 989-His-tag-DsbC donde se confirmó la expresión de la proteína. En esta muestra se puede observar la banda correspondiente a la proteína recombinante a la altura de 53 kDa (tamaño esperado) y se observan otras bandas, indicadas también con asteriscos anaranjados que estarían evidenciando un proceso de degradación de la proteína.

En la Figura 36 se pueden apreciar, señalados con asteriscos verdes y rojos dos bandas a la altura de 32 kDa que aparecen tanto en el control como en los cultivos inducidos. La intensidad de las bandas es mayor en los cultivos inducidos a medida que se aumenta la temperatura. Se destaca la banda marcada con el asterisco rojo en la fracción insoluble a 37°C, que coincide con el tamaño esperado de la enzima 989 His-tag sin péptido de tránsito (Figura 34). Sin embargo, no se puede afirmar que exista una verdadera inducción de la expresión de la proteína dado que el ensayo no incluyó un control de carga. Hubiera sido deseable enviar al espectrómetro de masas la banda mencionada para confirmar su identidad. En resumen, si bien existen indicios de que la enzima 989 His-tag sin péptido de tránsito se exprese en forma insoluble a 37°C, como ocurre en el caso de la fusión 989-His-tag-DsbC, se necesitarían experimentos adicionales para poder afirmarlo.

Finalmente, para verificar si existiese actividad β-glucosidasa en alguna de las fracciones donde se confirmó o se sospecha haber expresado la enzima 989, se ensayó un zimograma con las fracciones de interés como muestras y un control positivo comercial.

Para todas las muestras el resultado fue negativo para actividad β -glucosidasa (no se muestra).

En esta última sección se describe la obtención de las construcciones destinadas a la generación de cepas de S. cerevisiae genéticamente modificadas para expresar las enzimas de interés.

4.5. Confección de construcciones génicas para la expresión de las enzimas 989 y 3296 en *S. cerevisiae*

La estrategia para expresar la enzima 989 en *S. cerevisiae* se explica en detalle en Materiales y métodos. Resumidamente, se realiza por Fusion-PCR la construcción del *cassette* de expresión conteniendo el promotor de interés, la enzima a expresar y un terminador. Luego ese *cassette* es clonado en el vector pWJ1042 que porta el gen de selección URA3. El plásmido resultante es digerido con las enzimas Ahdl y PspOMI, lo que da por resultado la escisión de un fragmento que incluye el *cassette* y el marcador URA3. Éste es transformado en la cepa de *S. cerevisiae* T73-4 (URA3 -), junto con las regiones *upstream* y *downstream* correspondientes del *locus* del gen YPRCΔ15, que son además parcialmente solapantes. El ensamblado dentro de la célula de los tres fragmentos dirige la recombinación homóloga a este *locus*, seleccionado como sitio blanco porque se ha reportado que genes clonados en este sitio presentan altos niveles de expresión (Bai Flagfeldt *et al.* 2009).

Extracción de ADNg

Se realizó la extracción de ADNg de *S. cerevisiae* (concentración 65 ng/ μ L) de la cepa T73-2 (Puig *et al.* 1998) (Figura 37) para poder amplificar los promotores y terminador, así como las regiones *upstream* y *downstream* del locus del gen YPRC15.



Figura 37. Análisis por electroforesis en gel de agarosa del ADNg de S. cerevisiae

Construcción de los cassettes de expresión por Fusion-PCR

Las secuencias correspondientes a los promotores ADH (1437 pb) y ACT (424 pb) y al terminador CYC (250 pb) fueron amplificadas por PCR utilizando como molde ADNg de *S. cerevisiae* con los cebadores y las condiciones indicadas en Materiales y métodos. En

el gel de la Figura 38 se constata la obtención de fragmentos del tamaño esperado para cada caso.



Figura 38. Amplificación de ADH, ACT y CYC de S. cerevisiae, MPM: GeneRuler[™] 1kb Plus DNA Ladder

Las secuencias codificantes correspondientes a las enzimas 989 y 3296 fueron amplificadas a partir de ADNg de *I. terricola* con cebadores que introducen extensiones solapantes con el promotor y el terminador. Estas extensiones solapantes permitirán en la PCR de fusión la unión del promotor con enzima y enzima con terminador. En el caso de la enzima con el promotor endógeno solo es necesario fusionar un fragmento, que incluye 500 pb *upstream* del gen además de la secuencia codificante, al terminador CYC. En la Figura 39 se puede observar los fragmentos amplificados, en todos los casos del tamaño esperado (ver Materiales y métodos).



Figura 39. Amplificaciones de las enzimas 989 (MPM: NEB 1Kb DNA Ladder) y 3296 (MPM: GeneRulerTM DNA Ladder Mix) con las extensiones necesarias para realizar la PCR fusión del cassette. a) enzima con extensión ADH y CYC, b) enzima con extensión ACT y CYC, c) enzima con promotor endógeno y extensión CYC

Una vez purificados todos los fragmentos, se realizó la PCR de fusión para cada *cassette* (Figura 40). Una vez purificadas las bandas de tamaño esperado, se procedió a su digestión y posteriormente a la incorporación en el vector pWJ1042.


Figura 40. Cassettes de las enzimas 989 (MPM: NEB 1Kb DNA Ladder) y 3296 (MPM: GeneRulerTM DNA Ladder Mix) tras realizar la PCR fusión del cassette. a) cassette con promotor ADH, b) cassette con promotor ACT c) cassette con promotor endógeno

Clonado de los cassettes en el vector pWJ1042

Se realizaron las digestiones de los *cassettes* y del vector pWJ1042 con las enzimas SacII y XmaI. Los fragmentos de interés se purificaron a partir de geles de agarosa y se realizó la ligación de cada *cassette* en el vector. Se transformó *E. coli* para cada producto de ligación, se seleccionaron varios transformantes de cada construcción y se realizaron las extracciones de ADN plasmídico.

Amplificación de regiones para dirigir la recombinación homóloga

Se amplificaron las secuencias correspondientes a las regiones *upstream* (*up*) y *downstream* (*down*) del gen YPRC Δ 15 (Figura 41) con cebadores diseñados de forma tal que añaden una extensión que presenta homología con el *cassette* que se desea dirigir al *locus* (Bai Flagfeldt *et al.* 2009). En el gel mostrado en la Figura 41 se muestra la obtención de los fragmentos con los tamaños esperados: *up* 680 pb, *down* 650 pb.



Figura 41. Amplificaciones de las regiones upstream y downstream del gen YPRC∆15 MPM: GeneRuler[™] DNA Ladder Mix

En la amplificación del fragmento *down* se obtuvo una segunda banda inespecífica de mayor peso molecular, por lo cual el fragmento es purificado a partir de gel y no se llevó a cabo la purificación directa del producto de PCR.

5. CONCLUSIONES Y PERSPECTIVAS

En nuestro país el estado actual de avance en cuanto al conocimiento de genomas de cepas nativas es escaso, se ha caracterizado el genoma de *Hanseniaspora vinae* (Giorello *et al.* 2014; Giorello *et al.* 2018). Las cepas nativas representan un potencial interesante en el mejoramiento de la complejidad aromática de vinos, y desde hace algunos años se les viene prestando especial atención. La carencia de información genética sobre estas cepas hace que muchas potenciales enzimas permanezcan inexploradas hasta el momento. En este trabajo se obtuvo un ensamblado y anotación de buena calidad del genoma de la levadura nativa *I. terricola* para la cual no se contaba con datos de secuenciación. Cabe destacar que hasta el momento, a nivel internacional, pocas cepas vínicas fueron estudiadas desde un punto de vista genético (Woolfit *et al.* 2007; Borneman *et al.* 2008; Sternes *et al.* 2016; Seixas *et al.* 2018; Guaragnella *et al.* 2020).

Se identificó dentro de las proteínas anotadas en el genoma la secuencia de la β glucosidasa purificada previamente por nuestro grupo (González-Pombo *et al.* 2011; de Ovalle *et al.* 2017), permitiendo un abordaje molecular para la producción de la enzima. Además, se identificaron 10 posibles β -glucosidasas y una posible arabinosidasa, que podrán ser estudiadas en profundidad por su posible rol enológico. La disponibilidad de las secuencias codificantes de estas proteínas abre la posibilidad de clonar, modificar, expresar heterólogamente, purificar y caracterizar estas proteínas.

Sería de interés explotar el genoma de *I. terricola* desde un punto de vista de las vías metabólicas, con miras a entender la fisiología de este organismo, analizando las vías de utilización de fuentes carbonadas y nitrogenadas, vías de síntesis de compuestos que generan aromas, genes que codifican transportadores, etc. Estos estudios podrían incluir un análisis comparativo del genoma de esta levadura frente a *S. cerevisiae* y a otras levaduras vínicas para evaluar diferencias en el contenido génico que expliquen las características que imparten las distintas cepas (modulación de aromas y sabores) en el proceso de vinificación.

En este trabajo se incluyeron datos de secuenciación de ARNm para asistir a la anotación del genoma. En un futuro, para evaluar los niveles de transcripción de los distintos genes se deberían realizar réplicas de las distintas condiciones de cultivo para la extracción y secuenciación de ARNm. De esta manera, se podrá evaluar estadísticamente la expresión diferencial de genes en distintas condiciones de interés enológico.

Del análisis del genoma de *I. terricola* se seleccionaron dos β -glucosidasas para su expresión como proteínas recombinantes. Durante esta tesis las secuencias codificantes de ambas enzimas fueron clonadas y una de ellas expresada en forma recombinante en *E. coli*. Tras ensayar una multiplicidad de condiciones de cultivo e inducción, se logró

expresar exitosamente a 37ºC la proteína 989 fusionada a His-tag/DsbC, pero en forma insoluble. DsbC por lo tanto no estaría cumpliendo su rol de favorecer la solubilización actuando como chaperona en la formación de puentes disulfuro. Dado que en condiciones similares no fue posible expresar la proteína sin fusionar a DsbC, esta última favorecería de algún modo la síntesis y estabilización de la proteína 989.

La agregación de la proteína His-tag/DsbC-989 podría explicarse por la presencia del péptido señal de la proteína 989. En general los péptidos señal están constituidos por trechos de aminoácidos hidrofóbicos (von Heijne 1990). Si estas regiones no cumplen su función de direccionamiento hacia la membrana y son sintetizadas en el entorno citoplasmático, esto podría provocar la agregación de la proteína sintetizada. Sería interesante entonces, probar la expresión de His-tag/DsbC-989 en ausencia del péptido señal.

En este trabajo también se obtuvo la construcción de la proteína His-tag-989 con péptido señal. En este contexto no se pudo detectar proteína recombinante en los extractos celulares, en ninguna de las condiciones ensayadas. Como ha sido reportado en la literatura (Hao *et al.* 2020), no se puede descartar que el péptido señal de la proteína 989 pueda ser reconocido como tal en *E. coli.* La estructura de los péptidos de señalización eucariotas es similar a la de procariotas (von Heijne 1990). De hecho, el predictor de péptido señal (SignalP) reconoce la región de la proteína que corresponde al péptido señal y es clasificada tanto cómo péptido eucariota como procariota. La posibilidad de que el péptido señal de la enzima 989 sea funcional en *E. coli* sugiere que sería importante analizar los sobrenadantes de cultivo. Esto permitiría evaluar si la proteína pudiese estar siendo secretada, y que por eso no fuera detectada en los extractos analizados.

Teniendo en cuenta estas evidencias se realizó una construcción para ensayar la expresión de la proteína 989 sin el péptido señal. Nuestros resultados muestran que no se puede descartar que la proteína se encuentre en la FI de un cultivo de expresión a 37ºC (Figura 34). Sin embargo, hacen falta más experimentos para confirmarlo, como por ejemplo, verificar por espectrometría de masas su identidad.

Dadas las ventajas de expresar proteínas recombinantes en forma soluble, sería aconsejable probar otras fusiones con otras proteínas o *tags* que puedan influir en la solubilidad y/o estabilidad de la proteína producida. Se ha reportado que para la expresión de una misma proteína los resultados varían dependiendo de la proteína de fusión de elección (Correa & Oppezzo 2011, Correa *et al.* 2014). Nuestro grupo está evaluando la fusión con la MBP (*maltose binding protein*), la cual se ha reportado como eficaz en la solubilización de proteínas insolubles (Burguess & Deutscher 2009). Existen reportes donde la expresión de β -glucosidasas solubles ha sido dificultosa (Raza *et al.* 2020). En algunos casos ha sido necesario la co-expresión con chaperonas para lograr al menos un 20% de proteína recombinante soluble (Machida *et al.* 1997). Si ninguna de las alternativas llevara a la expresión de la proteína en forma soluble, sería necesario proceder a purificarla a partir de la fracción insoluble y posteriormente renaturalizarla.

Es posible también que *E. coli* no sea el sistema de expresión adecuado para esta proteína eucariota (Raza *et al.* 2020).

Luego de expresadas en *E. coli*, se procedería a la caracterización bioquímica y estructural de las enzimas mediante diversas metodologías: cromatografía de exclusión molecular y electroforesis (nativa y SDS-PAGE uni y bidimensional). Se estudiarían sus propiedades fisicoquímicas tales como punto isoeléctrico, condiciones óptimas de actividades y estabilidad (pH y temperatura) y su cinética enzimática (determinación de Km y Vmáx para el sustrato pNPG. En pos de su aplicación biotecnológica, se haría especial énfasis en la caracterización en cuanto a su actividad y estabilidad en condiciones enológicas (bajos pHs, presencia de etanol hasta 12% y concentración de glucosa 100 g/L), como se realizó para la β-glucosidasa 4180 (de Ovalle *et al.* 2017). Se podría estudiar también la especificidad de las enzimas obtenidas frente a sustratos sintéticos (con configuración alfa y beta) y naturales (disacáridos y otros). Para determinar el posible rol de estas enzimas como liberadoras de aromas se podrían realizar tratamientos de vinos jóvenes con las enzimas puras en forma soluble. El impacto en el perfil aromático del vino se determinaría mediante cromatografía de gases acoplada a un espectrómetro de masas (GC-MS).

Una alternativa al uso de las proteínas recombinantes es el uso de cepas de levadura genéticamente modificadas para expresar las actividades de interés. En este sentido, uno de los objetivos de esta tesis fue la obtención de cepas de *S. cerevisiae* expresando las glucosidasas seleccionadas. La estrategia elegida permite la obtención de cepas recombinantes libres de secuencias de origen bacteriano lo cual facilita su aplicación biotecnológica en el área alimentaria (Bai Flagfeldt *et al.* 2008). Se obtuvieron para las dos enzimas de interés las construcciones necesarias para el desarrollo de esta estrategia.

En paralelo al desarrollo de esta tesis se intentó aplicar esta estrategia a la β-glucosidasa 4180, previamente purificada por el grupo de la Dra. González-Pombo. Dado que fue imposible obtener transformantes por este método, se optó por clonar las construcciones para la expresión bajo diferentes promotores (ADH, ACT y endógeno) en el vector pESC-URA (Agilent Technologies). Con esta nueva estrategia sí fue posible obtener transformantes de S. cerevisiae y los primeros resultados de expresión de la βglucosidasa 4180 son promisorios. Hay indicios de actividad enzimática en los transformantes obtenidos de las construcciones para expresar la enzima bajo el control de los diferentes promotores. Los primeros resultados muestran, además, que en los transformantes que portan el promotor ADH la enzima sería expresada extracelularmente en algunas condiciones (Stefani de Ovalle, comunicación personal). La desventaja del uso de esta estrategia es que implica la introducción de secuencias de origen bacteriano (entre éstas el gen de resistencia a ampicilina) lo que haría a la cepa de S. cerevisiae generada no apta para el uso en la industria alimenticia. Sin embargo, la cepa obtenida podrá ser utilizada para evaluar la capacidad de las cepas generadas en la producción de compuestos volátiles durante el proceso de fermentación y su comportamiento en microvinificación. De ser los resultados satisfactorios, se podría

proceder a la construcción de una cepa libre de secuencias heterólogas, más aceptable para ser aplicada en la producción real de vinos.

En base a estos resultados, y suponiendo un comportamiento similar para las enzimas 989 y 3296, se está evaluando la estrategia a seguir, ya que probablemente la expresión en *S. cerevisiae* sea la más sencilla y directa para obtener las enzimas puras y en cantidad suficiente para proceder a la caracterización bioquímica de las mismas.

6. REFERENCIAS

S. Aftab. (2012). *Cloning* and expression of endo-1,4–β-glucanase gene from Bacillus licheniformis ATCC 14580 into Escherichia coli BL21 (DE 3). African Journal of Biotechnology, 11(12), 2846–2854. https://doi.org/10.5897/ajb11.2902

Al-Nakeeb, K., Petersen, T. N., & Sicheritz-Pontén, T. (2017). Norgal: Extraction and de novo assembly of mitochondrial DNA from whole-genome sequencing data. BMC Bioinformatics, 18(1), 1–7. https://doi.org/10.1186/s12859-017-1927-y

A. Anagnostopoulos, D., Kamilari, E., & Tsaltas, D. (2019). Contribution of the Microbiome as a Tool for Estimating Wine's Fermentation Output and Authentication. In Advances in Grape and Wine Biotechnology. https://doi.org/10.5772/intechopen.85692

Andorrà, I., Berradre, M., Rozès, N., Mas, A., Guillamón, J. M., & Esteve-Zarzoso, B. (2010). Effect of pure and mixed cultures of the main wine yeast species on grape must fermentations. In European Food Research and Technology (Vol. 231, Issue 2). https://doi.org/10.1007/s00217-010-1272-0

Andorrà, I., Berradre, M., Mas, A., Esteve-Zarzoso, B., & Guillamón, J. M. (2012). Effect of mixed culture fermentations on yeast populations and aroma profile. LWT - Food Science and Technology, 49(1), 8–13. https://doi.org/10.1016/j.lwt.2012.04.008

Andrades, D. de, Graebin, N. G., Ayub, M. A. Z., Fernandez-Lafuente, R., & Rodrigues, R. C. (2019). Physico-chemical properties, kinetic parameters, and glucose inhibition of several β -glucosidases for industrial applications. In Process Biochemistry (Vol. 78, Issue January). Elsevier. https://doi.org/10.1016/j.procbio.2019.01.008

Arnau, J., Lauritzen, C., Petersen, G. E., & Pedersen, J. (2006). Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. Protein Expression and Purification, 48(1), 1–13. https://doi.org/10.1016/j.pep.2005.12.002

Bai Flagfeldt, D., Siewers, V., Huang, L., & Nielsen, J. (2008). Characterization of chromosomal integration sites for heterologous gene expression in Saccharomyces cerevisiae. Yeast, August, 191–198. https://doi.org/10.1002/yea

Barbagallo, R. N., Spagna, G., Palmeri, R., Restuccia, C., & Giudici, P. (2004). Selection, characterization and comparison of β -glucosidase from mould and yeasts employable for enological applications. In Enzyme and Microbial Technology (Vol. 35, Issue 1). https://doi.org/10.1016/j.enzmictec.2004.03.005

Barbosa, C., Mendes-Faia, A., Lage, P., Mira, N. P., & Mendes-Ferreira, A. (2015). Genomic expression program of *Saccharomyces* cerevisiae along a mixed-culture wine fermentation with Hanseniaspora guilliermondii. In Microbial Cell Factories (Vol. 14, Issue 1). BioMed Central. https://doi.org/10.1186/s12934-015-0318-1

Baumes (2009). Chapter 8^a. Wine aroma precursors. Wine chemistry and biochemistry. In Wine Chemistry and Biochemistry. https://doi.org/10.1007/978-0-387-74118-5

Belda, I., Ruiz, J., Alonso, A., Marquina, D., & Navascués, E. (2015). Actividades enzimáticas de levaduras no *Saccharomyces* para su aplicación enológica

Belda, I., Ruiz, J., Esteban-Fernández, A., Navascués, E., Marquina, D., Santos, A., & Moreno-Arribas, M. V. (2017). Microbial contribution to Wine aroma and its intended use for Wine quality improvement. In Molecules (Vol. 22, Issue 2). https://doi.org/10.3390/molecules22020189

Bhatia, Y., Mishra, S., & Bisaria, V. S. (2002). Microbial β -Glucosidases : *Cloning*, Properties, and Applications (Vol. 22, Issue 4).

Bisson, L. F., & Karpel, J. E. (2010). Genetics of yeast impacting wine quality. In Annual Review of Food Science and Technology (Vol. 1, Issue 1). https://doi.org/10.1146/annurev.food.080708.100734

Boido, E., Lloret, A., Medina, K., Carrau, F., & Dellacassa, E. (2002). Effect of β-glycosidase activity of Oenococcus oeni on the glycosylated flavor precursors of Tannat wine during malolactic fermentation. In Journal of Agricultural and Food Chemistry (Vol. 50, Issue 8). https://doi.org/10.1021/jf0109367

Boido, E., Lloret, A., Medina, K., Farñia, L., Carrau, F., Versini, G., & Dellacassa, E. (2003). Aroma composition of Vitis vinifera cv. Tannat: The typical red wine from Uruguay. Journal of Agricultural and Food Chemistry, 51(18), 5408–5413. https://doi.org/10.1021/jf030087i

Bolanos-Garcia, V. M., & Davies, O. R. (2006). Structural analysis and classification of native proteins from E. coli commonly co-purified by immobilised metal affinity chromatography. Biochimica et Biophysica Acta - General Subjects, 1760(9), 1304–1313. https://doi.org/10.1016/j.bbagen.2006.03.027

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

Bond, S. R., & Naus, C. C. (2012). RF-*Cloning*.org: An online tool for the design of restriction-free *cloning* projects. Nucleic Acids Research, 40(W1), 1–5. https://doi.org/10.1093/nar/gks396

Borneman, A. R., Forgan, A. H., Pretorius, I. S., & Chambers, P. J. (2008). Comparative genome analysis of a Saccharomyces cerevisiae wine strain. FEMS Yeast Research, 8(7), 1185–1195. https://doi.org/10.1111/j.1567-1364.2008.00434.x

Burguess, R. R., & Deutscher, M. P. (2009). Methods in Enzymology - Guide to protein purification. https://books.google.com/books?hl=pt-PT&lr=&id=zTiRJHpKIQoC&pgis=1

Cabaroglu, T., Selli, S., Canbas, A., Lepoutre, J. P., & Günata, Z. (2003). Wine flavor enhancement through the use of exogenous fungal glycosidases. In Enzyme and Microbial Technology (Vol. 33, Issue 5). https://doi.org/10.1016/S0141-0229(03)00179-0

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. BMC Bioinformatics, 10, 1–9. https://doi.org/10.1186/1471-2105-10-421

Carrau, F., Boido, E., & Dellacassa, E. (2016). Fungal Metabolites. In Fungal Metabolites. https://doi.org/10.1007/978-3-319-19456-1

Chang, J., Park, I. H., Lee, Y. S., Ahn, S. C., Zhou, Y., & Choi, Y. L. (2011). *Cloning*, expression, and characterization of β -glucosidase from Exiguobacterium sp. DAU5 and transglycosylation activity. Biotechnology and Bioprocess Engineering, 16(1), 97–106. https://doi.org/10.1007/s12257-010-0092-1

Ciani, M., Comitini, F., Mannazzu, I., & Domizio, P. (2010). Controlled mixed culture fermentation: A new perspective on the use of non-*Saccharomyces* yeasts in winemaking. In FEMS Yeast Research (Vol. 10, Issue 2). https://doi.org/10.1111/j.1567-1364.2009.00579.x

Ciani, M., & Comitini, F. (2015). Yeast interactions in multi-starter wine fermentation. In Current Opinion in Food Science (Vol. 1, Issue 1). Elsevier Ltd. https://doi.org/10.1016/j.cofs.2014.07.001

Clemente-Jimenez, J. M., Mingorance-Cazorla, L., Martínez-Rodríguez, S., Las Heras-Vázquez, F. J., & Rodríguez-Vico, F. (2004). Molecular characterization and oenological properties of wine yeasts isolated during spontaneous fermentation of six varieties of grape must. In Food Microbiology (Vol. 21, Issue 2). https://doi.org/10.1016/S0740-0020(03)00063-7

Correa, A., & Oppezzo, P. (2011). Tuning different expression parameters to achieve soluble recombinant proteins in E. coli: Advantages of high-throughput screening. Biotechnology Journal, 6(6), 715–730. https://doi.org/10.1002/biot.201100025

Correa, A., Ortega, C., Obal, G., Alzari, P., Vincentelli, R., & Oppezzo, P. (2014). Generation of a vector suite for protein solubility screening. 5(February), 1–9. https://doi.org/10.3389/fmicb.2014.00067

Culleré, L., Escudero, A., Cacho, J., & Ferreira, V. (2004). Gas Chromatography-Olfactometry and Chemical Quantitative Study of the Aroma of Six Premium Quality Spanish Aged Red Wines. Journal of Agricultural and Food Chemistry, 52(6), 1653–1660. https://doi.org/10.1021/jf0350820

de Ovalle, S., Cavello, I., Brena, B. M., Cavalitto, S., & González-Pombo, P. (2017). Production and characterization of a β -glucosidase from Issatchenkia terricola and its use for hydrolysis of aromatic precursors in Cabernet Sauvignon wine. In LWT - Food Science and Technology (Vol. 87). https://doi.org/10.1016/j.lwt.2017.09.026

Delgado, L., Parker, M., Fisk, I., & Paradisi, F. (2020). Performance of the extremophilic enzyme BglA in the hydrolysis of two aroma glucosides in a range of model and real wines and juices. In Food Chemistry (Vol. 323). Elsevier Ltd. https://doi.org/10.1016/j.foodchem.2020.126825

Divakar, S. (2013). Enzymatic transformation. In Enzymatic Transformation (Vol. 9788132208). https://doi.org/10.1007/978-81-322-0873-0

Douglass, A. P., Offei, B., Braun-Galleani, S., Coughlan, A. Y., Martos, A. A. R., Ortiz-Merino, R. A., Byrne, K. P., & Wolfe, K. H. (2018). Population genomics shows no distinction between pathogenic Candida krusei and environmental Pichia kudriavzevii: One species, four names. PLoS Pathogens, 14(7), 1–27 Esteve-Zarzoso, B., Manzanares, P., Ramön, D., & Quero, A. (1998). The role of non-*Saccharomyces* yeasts in industrial winemaking. In International Microbiology (Vol. 1, Issue 2). https://doi.org/10.2436/im.v1i2.59

Fangman, W. L., & Dujon, B. (1984). Yeast mitochondrial genomes consisting of only A•T base pairs replicate and exhibit suppressiveness. Proceedings of the National Academy of Sciences of the United States of America, 81(22 I), 7156–7160. https://doi.org/10.1073/pnas.81.22.7156

Fehér, J., Lengyel, G., & Lugasi, A. (2007). The cultural history of wine - Theoretical background to wine therapy. Central European Journal of Medicine, 2(4), 379–391. https://doi.org/10.2478/s11536-007-0048-9

Fia, G., Canuti, V., & Rosi, I. (2014). Evaluation of potential side activities of commercial enzyme preparations used in winemaking. In International Journal of Food Science and Technology (Vol. 49, Issue 8). https://doi.org/10.1111/ijfs.12508

Francis, I. L., & Newton, J. L. (2005). Determining wine aroma from compositional data. In Australian Journal of Grape and Wine Research (Vol. 11, Issue 2). https://doi.org/10.1111/j.1755-0238.2005.tb00283.x

Gámbaro, A., Boido, E., Zlotejablko, A., Medina, K., Lloret, A., Dellacassa, E., & Carrau, F. (2001). Effect of malolactic fermentation on the aroma properties of Tannat wine. In Australian Journal of Grape and Wine Research (Vol. 7, Issue 1). https://doi.org/10.1111/j.1755-0238.2001.tb00190.x

Gamero, A., Quintilla, R., Groenewald, M., Alkema, W., Boekhout, T., & Hazelwood, L. (2016). High-throughput screening of a large collection of non-conventional yeasts reveals their potential for aroma formation in food fermentation. In Food Microbiology (Vol. 60). Elsevier Ltd. https://doi.org/10.1016/j.fm.2016.07.006

Gaya, P., Peirotén, Á., & Landete, J. M. (2020). Expression of a β -glucosidase in bacteria with biotechnological interest confers them the ability to deglycosylate lignans and flavonoids in vegetal foods. In Applied Microbiology and Biotechnology (Vol. 104, Issue 11). Applied Microbiology and Biotechnology. https://doi.org/10.1007/s00253-020-10588-x

Gil, J. V., Mateo, J. J., Jiménez, M., Pastor, A., & Huerta, T. (1996). Aroma compounds in wine as influenced by apiculate yeasts. In Journal of Food Science (Vol. 61, Issue 6). https://doi.org/10.1111/j.1365-2621.1996.tb10971.x

Giorello, F. M., Berná, L., Greif, G., Camesasca, L., Salzman, V., Medina, K., Robello, C., Gaggero, C., Aguilar, P. S., & Carrau, F. (2014). Genome sequence of the native apiculate wine yeast Hanseniaspora vineae T02/19AF. Genome Announcements, 2(3), 237–238. https://doi.org/10.1128/genomeA.00530-14

Giorello, F., Valera, M. J., Martin, V., Parada, A., Salzman, V., Camesasca, L., Fariña, L., Boido, E., Medina, K., Dellacassa, E., Berna, L., Aguilar, P. S., Mas, A., Gaggero, C., & Carrau, F. (2019). Genomic and transcriptomic basis of Hanseniaspora vineae's impact on flavor diversity and wine quality. In Applied and Environmental Microbiology (Vol. 85, Issue 1). https://doi.org/10.1128/AEM.01959-18

Godoy, L., Acuña-Fontecilla, A., & Catrileo, D. (2020). Formation of Aromatic and Flavor Compounds in Wine; A Perspective of Positive and Negative Contributions of Non-*Saccharomyces* Yeasts. In Intech: Vol. i (Issue tourism). https://doi.org/10.1016/j.colsurfa.2011.12.014

González-Pombo, P., Pérez, G., Carrau, F., Guisán, J. M., Batista-Viera, F., & Brena, B. M. (2008). One-step purification and characterization of an intracellular β-glucosidase from Metschnikowia pulcherrima. In Biotechnology Letters (Vol. 30, Issue 8). https://doi.org/10.1007/s10529-008-9708-3

Gonzalez-Pombo (2010) Tesis de Doctorado en Química. Universidad de la República-Facultad de Química, Montevideo, Uruguay.

González-Pombo, P., Fariña, L., Carrau, F., Batista-Viera, F., & Brena, B. M. (2011). A novel extracellular β -glucosidase from Issatchenkia terricola: Isolation, immobilization and application for aroma enhancement of white Muscat wine. In Process Biochemistry (Vol. 46, Issue 1). https://doi.org/10.1016/j.procbio.2010.07.016

Grada, A., & Weinbrecht, K. (2013). Next-generation sequencing: Methodology and application. Journal of Investigative Dermatology, 133(8), e11-4. https://doi.org/10.1038/jid.2013.248

Guaragnella, N., Chiara, M., Capece, A., Romano, P., Pietrafesa, R., Siesto, G., Manzari, C., & Pesole, G. (2020). Genome Sequencing and Comparative Analysis of Three Hanseniaspora uvarum Indigenous Wine Strains Reveal Remarkable Biotechnological Potential. Frontiers in Microbiology, 10(January), 1–14. https://doi.org/10.3389/fmicb.2019.03133

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. Systematic Biology, 59(3), 307–321. https://doi.org/10.1093/sysbio/syq010

Gunata, Y. Z., Bayonove, C. L., Baumes, R. L., & Cordonnier, R. E. (1985). The aroma of grapes I. Extraction and determination of free and glycosidically bound fractions of some grape aroma components. In Journal of Chromatography A (Vol. 331, Issue C). https://doi.org/10.1016/0021-9673(85)80009-1

Günata, Z., Dugelay, I., Sapis, J.-C., Baumes, R. L., & Bayonove, C. L. (1990). Action des glycosidases exogènes au cours de la vinification : libération de l'arôme a partir de précurseurs glycosidiques. In OENO One (Vol. 24, Issue 3). https://doi.org/10.20870/oeno-one.1990.24.3.1231

Günata, Y. Z., Bayonove, C. L., Tapiero, C., & Cordonnier, R. E. (1990). Hydrolysis of Grape Monoterpenyl β-D-Glucosides by Various β-Glucosidases. In Journal of Agricultural and Food Chemistry (Vol. 38, Issue 5). https://doi.org/10.1021/jf00095a016

Z. Gunata, J. Dugelay, J.C. Sapis, Raymond Baumes, Claude Bayonove (1993). Role of the enzyme in the use of the flavour potential from grape glycosides in wine making. International symposium on flavour precursors, Sep 1992, Wurzburg, Germany. 19 p., 1993. (hal-02844337)

Gürbüz, O., Rouseff, J. M., & Rouseff, R. L. (2006). Comparison of aroma volatiles in commercial merlot and Cabernet Sauvignon wines using gas chromatography-olfactometry and gas chromatography-mass spectrometry. Journal of Agricultural and Food Chemistry, 54(11), 3990–3996. https://doi.org/10.1021/jf053278p

Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: Quality assessment tool for genome assemblies. Bioinformatics, 29(8), 1072–1075. https://doi.org/10.1093/bioinformatics/btt086

Hao, S., Liu, Y., Qin, Y., Zhao, L., Zhang, J., Wu, T., Sun, B., & Wang, C. (2020). Expression of a highly active β -glucosidase from Aspergillus niger AS3 . 4523 in Escherichia coli and its application in gardenia blue preparation.

Heberle, H., Meirelles, V. G., da Silva, F. R., Telles, G. P., & Minghim, R. (2015). InteractiVenn: A web-based tool for the analysis of sets through Venn diagrams. BMC Bioinformatics, 16(1), 1–7. https://doi.org/10.1186/s12859-015-0611-3

Henick-Kling, T. (1988). Yeast and Bacterial Control in Winemaking. https://doi.org/10.1007/978-3-642-83340-3_10

Hewitt, S. N., Choi, R., Kelley, A., Crowther, G. J., Napuli, A. J., & Van Voorhis, W. C. (2011). Expression of proteins in Escherichia coli as fusions with maltose-binding protein to rescue non-expressed targets in a high-throughput protein-expression and purification pipeline. Acta Crystallographica Section F: Structural Biology and Crystallization Communications, 67(9), 1006–1009. https://doi.org/10.1107/S1744309111022159

Hjelmeland, A. K., & Ebeler, S. E. (2015). Glycosidically bound volatile aroma compounds in grapes and wine: A review. In American Journal of Enology and Viticulture (Vol. 66, Issue 1). https://doi.org/10.5344/ajev.2014.14104

Hoff K.J., Lomsadze A., Borodovsky M., Stanke M. (2019) Whole-Genome Annotation with BRAKER. In: Kollmar M. (eds) Gene Prediction. Methods in Molecular Biology, vol 1962. Humana, New York, NY. https://doi.org/10.1007/978-1-4939-9173-0_5

Huerta-Cepas, J., Serra, F., & Bork, P. (2016). ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. Molecular Biology and Evolution, 33(6), 1635–1638. https://doi.org/10.1093/molbev/msw046

Husnik JI, Delaquis PJ, Cliff MA, Van Vuuren HJJ (2007) Functional analyses of the malolactic wine yeast ML01. Am J Enol Vitic 58: 42–52

Jeromel, A., Korenika, A. M. J., & Tomaz, I. (2019). An influence of different yeast species on wine aroma composition. In Fermented Beverages: Volume 5. The Science of Beverages. https://doi.org/10.1016/B978-0-12-815271-3.00006-3

Jolly, N. P., Augustyn, O. P. H., & Pretorius, I. S. (2006). The Role and Use of Non-*Saccharomyces* Yeasts in Wine Production. In The Role and Use of Non-*Saccharomyces* Yeasts in Wine Production (Vol. 27, Issue 1). https://doi.org/10.21548/27-1-1475

Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A. F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S. Y., Lopez, R., & Hunter, S. (2014). InterProScan 5: Genome-scale protein function classification. Bioinformatics, 30(9), 1236–1240. https://doi.org/10.1093/bioinformatics/btu031

Kempton, J. B., & Withers, S. G. (1992). Mechanism of Agrobacterium β -Glucosidase:KineticStudies.Biochemistry,31(41),9961–9969.https://doi.org/10.1021/bi00156a015

Kitagawa, T., Tokuhiro, K., Sugiyama, H., Kohda, K., Isono, N., Hisamatsu, M., Takahashi, H., & Imaeda, T. (2010). Construction of a β -glucosidase expression system using the multistress-tolerant yeast Issatchenkia orientalis. Applied Microbiology and Biotechnology, 87(5), 1841–1853. https://doi.org/10.1007/s00253-010-2629-9

Kolmogorov, M., Yuan, J., Lin, Y., & Pevzner, P. A. (2019). Assembly of long, error-prone reads using repeat graphs. Nature Biotechnology, 37(5), 540–546. https://doi.org/10.1038/s41587-019-0072-8

Kurtz, S., Shumway, M., Antonescu, C., Salzberg, S. L., Phillippy, A., Smoot, M., Delcher, A. L., & Delcher, A. L. (2004). Versatile and open software for comparing large genomes. Genome Biology, 5(2), 12. http://www.tigr.org/software/mummer.

Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680-685.

Lage, P., Barbosa, C., Mateus, B., Vasconcelos, I., Mendes-Faia, A., & Mendes-Ferreira, A. (2014). H. guilliermondii impacts growth kinetics and metabolic activity of S. cerevisiae: The role of initial nitrogen concentration. In International Journal of Food Microbiology (Vol. 172). Elsevier B.V. https://doi.org/10.1016/j.ijfoodmicro.2013.11.031

Lambrechts, M. G., & Pretorius, I. S. (2000). Yeast and its Importance to Wine Aroma. In Understanding Wine Chemistry (Vol. 21, Issue June). https://doi.org/10.1002/9781118730720

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. Nature Methods, 9(4), 357–359. https://doi.org/10.1038/nmeth.1923

Levadoux, L. (1956). Les populations sauvages et cultivées de Vitis vinifera L. Annales de l'amélioration Des Plantes, 6, 59–118.

Li, Heng (2016).Minimap and miniasm: Fast mapping and de novo assembly for noisylongsequences.Bioinformatics,32(14),https://doi.org/10.1093/bioinformatics/btw152

Li, N., Wang, Q. Q., Xu, Y. H., Li, A. H., & Tao, Y. S. (2020). Increased glycosidase activities improved the production of wine varietal odorants in mixed fermentation of P. fermentans and high antagonistic S. cerevisiae. In Food Chemistry (Vol. 332). https://doi.org/10.1016/j.foodchem.2020.127426

Machida, S., Yu, Y., Singh, S. P., Kim, J. D., Hayashi, K., & Kawata, Y. (1998). Overproduction of β -glucosidase in active form by an Escherichia coli system coexpressing the chaperonin GroEL/ES. FEMS Microbiology Letters, 159(1), 41–46. https://doi.org/10.1016/S0378-1097(97)00541-7

Maicas, S., & Mateo, J. J. (2005). Hydrolysis of terpenyl glycosides in grape juice and other fruit juices: A review. In Applied Microbiology and Biotechnology (Vol. 67, Issue 3). https://doi.org/10.1007/s00253-004-1806-0

Manzanares, P., Orejas, M., Gil, J. V., De Graaff, L. H., Visser, J., & Ramón, D. (2003).Construction of a Genetically Modified Wine Yeast Strain Expressing the Aspergillusaculeatus rhaA Gene, Encoding an α-L-Rhamnosidase of Enological Interest. In AppliedandEnvironmentalMicrobiology(Vol. 69, Issue 12).https://doi.org/10.1128/AEM.69.12.7558-7562.2003

Marais, J. (1983). Terpenes in the Aroma of Grapes and Wines: A Review. In South African Journal of Enology & Viticulture (Vol. 4, Issue 2). https://doi.org/10.21548/4-2-2370

MARGOLLES-CLARK, E., TENKANEN, M., LUONTERI, E., & PENTTILA, M. (1996). Three α -galactosidase genes of Trichoderma reesei cloned by expression in yeast. European Journal of Biochemistry, 240(1), 104–111. https://doi.org/10.1111/j.1432-1033.1996.0104h.x

Mateo, J. J., & Jiménez, M. (2000). Monoterpenes in grape juice and wines. In Journal of Chromatography A (Vol. 881, Issues 1–2). https://doi.org/10.1016/S0021-9673(99)01342-4

Medina, K., Boido, E., Dellacassa, E., & Carrau, F. (2012). Growth of non-*Saccharomyces* yeasts affects nutrient availability for *Saccharomyces* cerevisiae during wine fermentation. In International Journal of Food Microbiology (Vol. 157, Issue 2). Elsevier B.V. https://doi.org/10.1016/j.ijfoodmicro.2012.05.012

Morata, A., Escott, C., Bañuelos, M. A., Loira, I., Del Fresno, J. M., González-Pombo, C., & Suárez-lepe, J. A. (2020). Contribution of non-*Saccharomyces* yeasts to wine freshness. A review. In Biomolecules (Vol. 10, Issue 1). https://doi.org/10.3390/biom10010034

Nozach, H., Fruchart-gaillard, C., Fenaille, F., Beau, F., Henrique, O., Ramos, P., Douzi, B., Saez, N. J., Moutiez, M., Servent, D., Gondry, M., Thaï, R., Cuniasse, P., Vincentelli, R., & Dive, V. (2013). High throughput screening identifies disulfide isomerase DsbC as a very efficient partner for recombinant expression of small disulfide-rich proteins in E. coli. 1–16.

Ortega-Heras, M., González-Pombo-SanJosé, M. L., & Beltrán, S. (2001). Aroma composition of wine studied by different extraction methods. In Analytica Chimica Acta (Vol. 458, Issue 1). https://doi.org/10.1016/S0003-2670(01)01526-4

Cordero Otero, R. R., Ubeda Iranzo, J. F., Briones-Perez, A. I., Potgieter, N., Villena, M. A., Pretorius, I. S., & Van Rensburg, P. (2003). Characterization of the β -Glucosidase Activity Produced by Enological Strains of Non-*Saccharomyces* Yeasts. In Journal of Food Science (Vol. 68, Issue 8). https://doi.org/10.1111/j.1365-2621.2003.tb07062.x

Ottone, C., Romero, O., Aburto, C., Illanes, A., & Wilson, L. (2019). Biocatalysis in the winemaking industry: Challenges and opportunities for immobilized enzymes. In Comprehensive Reviews in Food Science and Food Safety (Vol. 19, Issue 2). https://doi.org/10.1111/1541-4337.12538

Padilla, B., Gil, J. V., & Manzanares, P. (2016). Past and future of non-*Saccharomyces* yeasts: From spoilage microorganisms to biotechnological tools for improving wine aroma complexity. In Frontiers in Microbiology (Vol. 7, Issue MAR). https://doi.org/10.3389/fmicb.2016.00411

Palmeri R, Spagna G (2007) "b-Glucosidase in cellular and acellular form for winemaking application"; Enzyme Microb Technol 40:382–389

PEREZ-GONZALEZ, J. A., GONZALEZ, R., QUEROL, A., SENDRA, J., & RAMON, D. (1993). Construction of a recombinant wine yeast strain expressing a fungal pectate lyase gene. FEMS Microbiology Letters, 126(3), 263–269. https://doi.org/10.1016/0378-1097(95)00021-V

Partow, S., Siewers, V., Bjorn, S., & Maury, J. (2010). Characterization of different promoters for designing a new expression vector in Saccharomyces cerevisiae. Yeast, July, 191–198. https://doi.org/10.1002/yea

Pogorzelski, E., & Wilkowska, A. (2007). Selection of Yeast strains containing ß-Glucosidase for improving wine aroma. In Journal of Agricultural Science and Technology B (Issue January).

Pretorius, I. S. (2000). Tailoring wine yeast for the new millennium: Novel approaches to the ancient art of winemaking. In Yeast (Vol. 16, Issue 8). https://doi.org/10.1002/1097-0061(20000615)16:8<675::AID-YEA585>3.0.CO;2-B

Puig, S., Ramón, D., & Pérez-Ortín, J. E. (1998). Optimized Method to Obtain Stable Food-Safe Recombinant Wine Yeast Strains. Journal of Agricultural and Food Chemistry, 46(4), 1689–1693. https://doi.org/10.1021/jf9706538

Querol, A., & Ramón, D. (1996). Application of molecular techniques in wine microbiology: A review. In Acta Alimentaria (Vol. 31, Issue 1). https://doi.org/10.1556/AAlim.31.2002.1.4

Rapp, A., & Mandery, H. (1986). Wine aroma (Vol. 42).

Raza, A., Pothula, R., Abdelgaffar, H., Bashir, S., & Jurat-Fuentes, J. L. (2020). Identification and functional characterization of a β -glucosidase from Bacillus tequelensis BD69 expressed in bacterial and yeast heterologous systems. PeerJ, 2020(3), 1–22. https://doi.org/10.7717/peerj.8792

Reid, R. J. D., Lisby, M., & Rothstein, R. (2002). *Cloning*-free genome alterations in Saccharomyces cerevisiae using adaptamer-mediated PCR. Methods in Enzymology, 350(1983), 258–277. https://doi.org/10.1016/S0076-6879(02)50968-X

Rhoads, A., & Au, K. F. (2015). PacBio Sequencing and Its Applications. Genomics, Proteomics and Bioinformatics, 13(5), 278–289. https://doi.org/10.1016/j.gpb.2015.08.002 Robinson, A. L., Ebeler, S. E., Heymann, H., Boss, P. K., Solomon, P. S., & Trengove, R. D. (2009). Interactions between wine volatile compounds and grape and wine matrix components influence aroma compound headspace partitioning. In Journal of Agricultural and Food Chemistry (Vol. 57, Issue 21). https://doi.org/10.1021/jf902586n

Romano, P., Ciani, M., & Fleet, G. H. (2019). Yeasts in the production of wine. In Yeasts in the Production of Wine. https://doi.org/10.1007/978-1-4939-9782-4

Roudil, L., Russo, P., Berbegal, C., Albertin, W., Spano, G., & Capozzi, V. (2019). Non-*Saccharomyces* Commercial Starter Cultures: Scientific Trends, Recent Patents and Innovation in the Wine Sector. In Recent Patents on Food, Nutrition & Agriculture (Vol. 11, Issue 1). https://doi.org/10.2174/2212798410666190131103713

Sahay, S. (2019). Wine enzymes: Potential and practices. In Enzymes in Food Biotechnology: Production, Applications, and Future Prospects. https://doi.org/10.1016/B978-0-12-813280-7.00006-2

Salmela, L., & Rivals, E. (2014). LoRDEC: Accurate and efficient long read error correction. Bioinformatics, 30(24), 3506–3514. https://doi.org/10.1093/bioinformatics/btu538

Sánchez-Torres, P., González-Candelas, L., & Ramón, D. (1998). Heterologous Expression of a Candida molischiana Anthocyanin-β-glucosidase in a Wine Yeast Strain. In Journal of Agricultural and Food Chemistry (Vol. 46, Issue 1). https://doi.org/10.1021/jf970570r

Sanz-Aparicio, J., Hermoso, J. A., Martínez-Ripoll, M., Lequerica, J. L., & Polaina, J. (1998). Crystal structure of β -glucosidase A from Bacillus polymyxa: Insights into the catalytic activity in family 1 glycosyl hydrolases. Journal of Molecular Biology, 275(3), 491–502. https://doi.org/10.1006/jmbi.1997.1467

Sarry, J. E., & Günata, Z. (2004). Plant and microbial glycoside hydrolases: Volatile release from glycosidic aroma precursors. In Food Chemistry (Vol. 87, Issue 4). https://doi.org/10.1016/j.foodchem.2004.01.003

Schreier, P., & Walter, G. J. (1979). Flavor composition of wines: A review. In C R C Critical Reviews in Food Science and Nutrition (Vol. 12, Issue 1). https://doi.org/10.1080/10408397909527273

Shen, X. X., Opulente, D. A., Kominek, J., Zhou, X., Steenwyk, J. L., Buh, K. V., Haase, M. A. B., Wisecaver, J. H., Wang, M., Doering, D. T., Boudouris, J. T., Schneider, R. M., Langdon, Q. K., Ohkuma, M., Endoh, R., Takashima, M., Manabe, R. ichiroh, Čadež, N., Libkind, D., ... Rokas, A. (2018). Tempo and Mode of Genome Evolution in the Budding Yeast Subphylum. Cell, 175(6), 1533-1545.e20. https://doi.org/10.1016/j.cell.2018.10.023

Seguinot, P., Bloem, A., Brial, P., Meudec, E., Ortiz-Julien, A., & Camarasa, C. (2019). Analysing the impact of the nature of the nitrogen source on the formation of volatile compounds to unravel the aroma metabolism of two non-*Saccharomyces* strains. In International Journal of Food Microbiology (Vol. 316). Elsevier B.V. https://doi.org/10.1016/j.ijfoodmicro.2019.108441 Seppey, M., Manni, M., & Zdobnov, E. M. (2019). BUSCO: Assesing Genome Assembly and Annotation Completeness. Notes on the Greek Text of Genesis, 1962, 185–201. https://doi.org/10.2307/j.ctvfxvc64.18

Seixas, I., Barbosa, C., Mendes-Faia, A., Güldener, U., Tenreiro, R., Mendes-Ferreira, A., & Mira, N. P. (2019). Genome sequence of the non-conventional wine yeast Hanseniaspora guilliermondii UTAD222 unveils relevant traits of this species and of the Hanseniaspora genus in the context of wine fermentation. DNA Research, 26(1), 67–83. https://doi.org/10.1093/dnares/dsy039

Sternes, P. R., Lee, D., Kutyna, D. R., & Borneman, A. R. (2016). Genome sequences of three species of Hanseniaspora isolated from spontaneous wine fermentations. Genome Announcements, 4(6), 1–2. https://doi.org/10.1128/genomeA.01287-16

Strauss, C. R., Wilson, B., Gooley, P. R., & Williams, P. J. (1986). Role of Monoterpenes in Grape and Wine Flavor. 222–242. https://doi.org/10.1021/bk-1986-0317.ch018

Strauss, M. L. A., Jolly, N. P., Lambrechts, M. G., & Van Rensburg, P. (2001). Screening for the production of extracellular hydrolytic enzymes by non-*Saccharomyces* wine yeasts. In Journal of Applied Microbiology (Vol. 91, Issue 1). https://doi.org/10.1046/j.1365-2672.2001.01379.x

Styger, G., Prior, B., & Bauer, F. F. (2011). Wine flavor and aroma. In Journal of Industrial Microbiology and Biotechnology (Vol. 38, Issue 9). https://doi.org/10.1007/s10295-011-1018-4

Sun, W. X., Hu, K., Zhang, J. X., Zhu, X. L., & Tao, Y. S. (2018). Aroma modulation of Cabernet Gernischt dry red wine by optimal enzyme treatment strategy in winemaking. In Food Chemistry (Vol. 245). https://doi.org/10.1016/j.foodchem.2017.11.106

Swiegers, J. H., & Pretorius, I. S. (2005). Yeast modulation of wine flavor. In Advances in Applied Microbiology (Vol. 57, Issue SUPPL. A). https://doi.org/10.1016/S0065-2164(05)57005-9

Swiegers, J. H., Bartowsky, E. J., Henschke, P. A., & Pretorius, I. S. (2005). Yeast and bacterial modulation of wine aroma and flavour. In Australian Journal of Grape and Wine Research (Vol. 11, Issue 2). https://doi.org/10.1111/j.1755-0238.2005.tb00285.x

Tabor, S. (1990). Expression Using the T7 RNA Polymerase / Promoter System. 1–11.

This, P., Lacombe, T., & Thomas, M. R. (2006). Historical origins and genetic diversity ofwinegrapes.TrendsinGenetics,22(9),511–519.https://doi.org/10.1016/j.tig.2006.07.008

Tobergte DR, Curtis S (2013) GRAS Notification for the use of a modified yeast for reduction of ethyl carbamate in fermented beverages. J Chem Inf Model 53: 1689–1699

Tolosa, J. J. M., & Prieto, S. M. (2018). Non-*Saccharomyces* yeasts: An enzymatic unexplored world to be exploited. In Enzymes in Food Biotechnology: Production, Applications, and Future Prospects. Elsevier Inc. https://doi.org/10.1016/B978-0-12-813280-7.00025-6

Van Den Ent, F., & Löwe, J. (2006). RF *cloning*: A restriction-free method for inserting target genes into plasmids. Journal of Biochemical and Biophysical Methods, 67(1), 67–74. https://doi.org/10.1016/j.jbbm.2005.12.008

Vilela, A. (2020). Modulating wine pleasantness throughout wine-yeast co-inoculation or sequential inoculation. In Fermentation (Vol. 6, Issue 1). https://doi.org/10.3390/fermentation6010022

Villamor, R. R., & Ross, C. F. (2013). Wine matrix compounds affect perception of wine aromas. In Annual Review of Food Science and Technology (Vol. 4, Issue 1). https://doi.org/10.1146/annurev-food-030212-182707

Villena, M. A., Iranzo, J. F. Ú., & Pérez, A. I. B. (2007). β-Glucosidase activity in wine yeasts: Application in enology. In Enzyme and Microbial Technology (Vol. 40, Issue 3). https://doi.org/10.1016/j.enzmictec.2006.07.013

Voelkerding, K. V., Dames, S. A., & Durtschi, J. D. (2009). Next-generation sequencing: from basic research to diagnostics. Clinical Chemistry, 55(4), 641–658. https://doi.org/10.1373/clinchem.2008.112789

von Heijne, G. (1990). The signal peptide. The Journal of Membrane Biology, 115(3), 195–201. https://doi.org/10.1007/BF01868635

Vontrobová, E., Kubizniakova, P., Fiala, J., Sochor, J., & Matoulková, D. (2019). Autochthonous yeasts as one of the tools to produce wines by original technologies. In Kvasny Prumysl (Vol. 65, Issue 1). https://doi.org/10.18832/kp2019.65.38

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C. A., Zeng, Q., Wortman, J., Young, S. K., & Earl, A. M. (2014). Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS ONE, 9(11). https://doi.org/10.1371/journal.pone.0112963

Wanapu, C., Sripunya, P., & Boonkerd, N. (2012). Selection of Yeast strains containing ß-Glucosidase for improving wine aroma. In Journal of Agricultural Science and Technology B (Issue January).

Waterhouse, R. M., Tegenfeldt, F., Li, J., Zdobnov, E. M., & Kriventseva, E. V. (2013). OrthoDB : a hierarchical catalog of animal, fungal and bacterial orthologs. 41(November 2012), 358–365. https://doi.org/10.1093/nar/gks1116

Winterhalter, P., & Skouroumounis, G. K. (1997). Glycoconjugated aroma compounds: occurrence, role and biotechnological transformation. In Advances in biochemical engineering/biotechnology (Vol. 55). https://doi.org/10.1007/bfb0102063

Woolfit, M., Rozpędowska, E., Piškur, J., & Wolfe, K. H. (2007). Genome survey sequencing of the wine spoilage yeast Dekkera (Brettanomyces) bruxellensis. Eukaryotic Cell, 6(4), 721–733. https://doi.org/10.1128/EC.00338-06

Young, C. L., Britton, Z. T., & Robinson, A. S. (2012). Recombinant protein expression and purification: A comprehensive review of affinity tags and microbial applications. Biotechnology Journal, 7(5), 620–634. https://doi.org/10.1002/biot.201100155

Zhang, H., Yohe, T., Huang, L., Entwistle, S., Wu, P., Yang, Z., Busk, P. K., Xu, Y., & Yin, Y. (2018). DbCAN2: A meta server for automated carbohydrate-active enzyme annotation. Nucleic Acids Research, 46(W1), W95–W101. https://doi.org/10.1093/nar/gky418

Zhang, W., Zhuo, X., Hu, L., & Zhang, X. (2020). Effects of crude β -glucosidases from *issatchenkia terricola, pichia kudriavzevii, metschnikowia pulcherrim*a on the flavor complexity and characteristics of wines. In Microorganisms (Vol. 8, Issue 6). https://doi.org/10.3390/microorganisms8060953

Zietsman, A. J. J., De Klerk, D., & Van Rensburg, P. (2011). Coexpression of α -larabinofuranosidase and β -glucosidase in *Saccharomyces cerevisiae*. FEMS Yeast Research, 11(1), 88–103. https://doi.org/10.1111/j.1567-1364.2010.00694.x

7. ANEXO

7.1. Materiales y métodos

7.1.1. Tablas de cebadores

Los cebadores fueron analizados con el programa OligoAnalizer de Integrated DNA Technologies <u>https://www.idtdna.com/calc/analyzer</u>

Cebador	Uso	Secuencia (5' – 3')		Largo (nt)	Tm (ʻ	°C)
989_3'_F	Amplificación del gen 989	ATGAAGCTTTCTGTTGCAT	ГСТ	21	55,5	
Ecoli989sinPS	Amplificación del gen 989 s péptido señal	sin ATCGCATCTCTCGGTTTTG	βAT	21	56,9	
prom_989_Fw	Amplificación de 989 con 50 pb <i>upstream</i> el gen (promot endógeno)	00 ACGCATATAGTATTATTTA or	AGCCATTG	26	60,0	
989_5'_R	Amplificación del gen 989	TTAGAAATCACAATCTAA	IGAGTATTTG	28	59,9	
989_int_F	Cebador interno 989 (pa	ara ACCATTTCTAAATCAGCTA	ATTICTICIG	28	62,8	
989_int_R	secuenciar)	TGACCAAGATGTTCCATG	AGTC	22	60,3	
3296_3'_F	Amplificación del gen 3296	ATGAAGTTCTCCCAATTAT	CCGTT	24	60,1	
Ecoli3296sinPS	Amplificación del gen 3296 s péptido señal	sin GCTCCAATGGCCATCAAC		18	55,9	
prom_3296_Fw	Amplificación de 3296 con 50 pb <i>upstream</i> el gen (promot endógeno)	00 CCACTCCGTACCCAATCAC or	δA	20	60,5	
3296_5'_R	Amplificación del gen 3296	TTAATTGTAGAAAACAAT	GTTGGCAGA	27	60,7	
3296_bglu_R	Cebador interno 3296 (pa	ara GGTACTATTCCATGTTCTG	βA	20	54,3	
3296_int_R	secuenciar)	CATATCTTCGGAACCAGG	GTAATC	24	63,5	
Cebador	Uso	Secuencia (5' – 3')	Largo Tm* (nt) (°C)	Refe	rencia	
Cebador 989_vector_eco F	Uso li_ Para clonar 989 (CON péptido señal) en vector T7 mediante RF- <i>cloning</i>	Secuencia (5' – 3') GGATCGGAAAACCTGTATTTT CAGGGATCCATGAAGCTTTCTG TTGCATCTATTACTGCTCT	Largo Tm* (nt) (°C) 62 59,6	Refer (Corr 2014	r encia ea <i>et</i>)	al.
Cebador 989_vector_eco F 989_vector_eco R	Uso I_ Para clonar 989 (CON péptido señal) en vector T7 mediante RF- <i>cloning</i> I_ Para clonar 989 en vector T7 mediante RF- <i>cloning</i>	Secuencia (5' – 3') GGATCGGAAAACCTGTATTTT CAGGGATCCATGAAGCTTTCTG TTGCATCTATTACTGCTCT GAACTGCGGGTGGCTCCAGCT GCCGGATCCTTAGAAATCACAA TCTAATGAGTATTTGAGGTTAC GGCTAGA	Largo Tm* (nt) (°C) 62 59,6 72 60,7	Refer (Corr 2014 (Corr 2014	r encia ea et) ea et	al.
Cebador 989_vector_eco F 989_vector_eco R 989_vector_eco F_sinPS	Uso Ii_ Para clonar 989 (CON péptido señal) en vector T7 mediante RF-cloning Ii_ Para clonar 989 en vector T7 mediante RF-cloning Ii_ Para clonar 989 (SIN péptido señal) en vector T7 mediante RF-cloning	Secuencia (5' – 3') GGATCGGAAAACCTGTATTTT CAGGGATCCATGAAGCTTTCTG TTGCATCTATTACTGCTCT GAACTGCGGGTGGCTCCAGCT GCCGGATCCTTAGAAATCACAA TCTAATGAGTATTTGAGGTTAC GGCTAGA GGATCGGAAAACCTGTATTTT CAGGGATCCATCGCATCTCTCG GTTTTGAT	Largo (nt) Tm* (°C) 62 59,6 72 60,7 51 61,0	Refer (Corr 2014 (Corr 2014 (Corr 2014	r encia ea et) ea et) ea et	al. al.
Cebador 989_vector_eco F 989_vector_eco R 989_vector_eco F_sinPS 3296_vector_eco _F	 Uso Para clonar 989 (CON péptido señal) en vector T7 mediante RF-cloning Para clonar 989 en vector T7 mediante RF-cloning Para clonar 989 (SIN péptido señal) en vector T7 mediante RF-cloning Para clonar 3296 (CON péptido señal) en vector T7 mediante RF-cloning 	Secuencia (5' – 3') GGATCGGAAAACCTGTATTTT CAGGGATCCATGAAGCTTTCTG TTGCATCTATTACTGCTCT GAACTGCGGGTGGCTCCAGCT GCCGGATCCTTAGAAATCACAA TCTAATGAGTATTTGAGGTTAC GGCTAGA GGATCGGAAAACCTGTATTTT CAGGGATCCATCGCATCTCTCG GTTTTGAT GGATCGGAAAACCTGTATTTT CAGGGATCCATGAAGTTCTCCC AATTATCCGTTTTAACTTTTGCT G	Largo (nt) Tm* (°C) 62 59,6 72 60,7 51 61,0 67 60,5	Refer (Corr 2014 (Corr 2014 (Corr 2014 (Corr 2014	rencia ea et) ea et) ea et) ea et	: al. : al. : al.
Cebador 989_vector_eco F 989_vector_eco R 989_vector_eco F_sinPS 3296_vector_eco _F 3296_vector_eco _R	UsoIi_Para clonar 989 (CON péptido señal) en vector T7 mediante RF-cloningIi_Para clonar 989 en vector T7 mediante RF-cloningIi_Para clonar 989 (SIN péptido señal) en vector T7 mediante RF-cloningIi_Para clonar 3296 (CON péptido señal) en vector T7 mediante RF-cloningIiiPara clonar 3296 (CON péptido señal) en vector T7 mediante RF-cloningIiiPara clonar 3296 (CON péptido señal) en vector T7 mediante RF-cloning	Secuencia (5' – 3') GGATCGGAAAACCTGTATTTT CAGGGATCCATGAAGCTTTCTG TTGCATCTATTACTGCTCT GAACTGCGGGTGGCTCCAGCT GCCGGATCCTTAGAAATCACAA TCTAATGAGTATTTGAGGTTAC GGCTAGA GGATCGGAAAACCTGTATTTT CAGGGATCCATCGCATCTCTCG GTTTTGAT GGATCGGAAAACCTGTATTTT CAGGGATCCATGAAGTTCTCCC AATTATCCGTTTTAACTTTTGCT G GAACTGCGGGTGGCTCCAGCT GCCGGATCCTTAATTGTAGAAA ACAATGTTGGCAGAACCAGAG GT	Largo (nt) Tm* (°C) 62 59,6 72 60,7 51 61,0 67 60,5 66 61,2	Refer (Corr 2014 (Corr 2014 (Corr 2014 (Corr 2014 (Corr 2014	rencia ea et) ea et) ea et) ea et	: al. : al. : al. : al.

Tabla 15. Cebadores utilizados en este trabajo

En negrita se indica la región del cebador que hibrida con el vector de clonado, la otra parte hibrida con la enzima de interés. *Tm correspondiente a la parte que hibrida con el gen de interés, para la PCR del RF-*cloning* se utilizó 60° como temperatura de hibridación (Correa *et al.* 2014)

Clonado en Saccharomyces cerevisae

Cebador	Uso	Secuencia (5′ – 3′)	Largo (nt)	Tm (°C)	Referencia
up_fw	Amplificación de la región <i>up</i> para dirigir	GCCAGGCGCCTTTATATCAT	20	61,3	(Bai Flagfeldt <i>et al.</i> 2009)
up_rv	recombinación homóloga al sitio YPRCΔ15, cromosoma XVI	CCGCGGTGGAGCTCCAGCTTTT GTT TTTGCGAAACCCTATGCTC T	45	59,8*	(Bai Flagfeldt <i>et al.</i> 2009)
down_fw	Amplificación de la región <i>downstream</i> para dirigir	TAAGAATTGTCGTTCATGGTGA CA CAATGGAAGGTCGGGATGA G	44	59,9*	(Bai Flagfeldt <i>et al.</i> 2009)
down_rv	recombinación homóloga al sitio YPRCΔ15, cromosoma XVI	ATAAAGCAGCCGCTACCAAA	20	59,9	(Bai Flagfeldt <i>et al.</i> 2009)
ADH1-fw	Amplificación del promotor ADH con	ATATTA <i>CCGCGG</i> AGGGGGGATC GAAGAAATGATG	33	61,2	(Partow <i>et al.</i> 2009)
ADH1-rv	sitio de corte SacII	TGTATATGAGATAGTTGATTGT ATGCTTGG	30	61,2	(Partow <i>et al.</i> 2009)
ACT1-fw	Amplificación del promotor ACT con sitio	ATATTACCGCGGCCCTCAATATT TCTCTGTCACC	34	63,6	(Bai Flagfeldt <i>et al.</i> 2009)
ACT1-rv	de corte SacII	TGTTAATTCAGTAAATTTTCGAT CTTG	27	50,8	(Bai Flagfeldt <i>et al.</i> 2009)
CYC1-fw	Amplificación del terminador CYC con	CGATATCATGTAATTAGTTATGTC ACGC	28	54,0	(Bai Flagfeldt <i>et al.</i> 2009)
CYC1-rv	sitio de corte de Xmal	ATAATCCCCGGGGCAAATTAAA GCCTTCGAGC	32	64,9	(Bai Flagfeldt <i>et al.</i> 2009)
ADH1_989_Fw	Amplificación de 989 con extensión ADH	CCAAGCATACAATCAACTATCT CATATACA ATGAAGCTTTCTGT TGCATCT	51	52,7*	-
ACT1_989_Fw	Amplificación de 989 con extensión ACT	CAAGATCGAAAATTTACTGAAT TAACA ATGAAGCTTTCTGTTGC ATCT	48	52,7*	-
CYC1_989_Rv	Amplificación de 989 con extensión CYC	GCGTGACATAACTAATTACATG ATATCG TTAGAAATCACAATCT AATGAGTATTTG	56	50,2*	-
Sacll_prom_989_Fw	Amplificación de 989 con su promotor endógeno y sitio de corte de SacII	ATATTACCGCGGACGCATATAG TATTATTTAGCCATTG	38	51,4*	-
ADH1_3296_Fw	Amplificación de 3296 con extensión ADH	CCAAGCATACAATCAACTATCT CATATACA ATGAAGTTCTCCCA ATTATCCGTT	54	54,3*	-
ACT1_3296_Fw	Amplificación de 3296 con extensión ACT	CAAGATCGAAAATTTACTGAAT TAACA ATGAAGTTCTCCCAATT ATCCGTT	51	54,3*	-
CYC1_3296_Rv	Amplificación de 3296 con extensión CYC	GCGTGACATAACTAATTACATG ATATCG TTAATTGTAGAAAACA ATGTTGGCAGA	55	53,9*	-
Sacll_prom_3296_Fw	Amplificación de 3296 con su promotor endógeno y sitio de corte de SacII	ATATTACCGCGGCCACTCCGTA CCCAATCAGA	32	56,3*	-

En negrita se indica la región del cebador que hibrida el gen/región de interés. La extensión hibrida con la región del promotor/terminador correspondiente en cada caso para construcción del *cassette*, y con las regiones terminales del *cassette* para dirigir la recombinación homóloga en el caso de la amplificación de las regiones *upstream* y *downstream*.

En cursiva se indican sitios de cortes de enzimas de restricción.

*Tm correspondiente a la parte que hibrida con el gen/región a amplificar.

7.1.2. Vectores

7.1.2.1. Clonado en E. coli

Los dos vectores utilizados fueron construidos por Correa *et al.* 2014. En este tipo de vectores la expresión del gen de interés se encuentra bajo control del promotor T7 (Tabor *et al.* 2001) que es inducible por IPTG. Los vectores que nos proporcionaron poseen la proteína verde fluorescente (GFP) como resultado de la inserción con el método descrito en el artículo, al utilizar el método de RF-*cloning* descrito en el artículo sobre estos vectores la GFP es reemplazada por la proteína de interés.



pET-32a con promotor T7 (Correa et al. 2014)

Figura 42. Vector pET32-a utilizado para RF-cloning



pET-32a con promotor T7 y proteína de fusión DsbC (Correa et al. 2014)

Figura 43. Vector pET32a-DsbC utilizado para RF-cloning

7.1.2.2.Clonado en S. cerevisiae (Bai Flagfeldt et al. 2009)



Figura 44. Vector pWJ1042

7.2. Resultados

Secuencia codificante para la enzima 989

>g989.t1 CDS=1-933 péptido señal

ATGAAGCTTTCTGTTGCATCTATTACTGCTCTTGCTATCTCATTGGCCGCTGGCCCTGT CTCAGCTATCGCATCTCCGGTTTTGATCTTGGTGTTCAAGACAACGGCGGTGCCTGTA AGGTTGCTTCTGATTACGAAGCAGATCTCGCAATTTTAGCTCCATACTCTAAGATGGTT AAGACCTATGCTGTCAGTACCTGTAACACATTAGAAGTTTTAGGTCCTGTCGTTGAAAG TTCAGGTTTCCAACTCACCCTTGGTGTCTGGCCAACCCCACAAGATAAATTCACTGCTG AAAAGGAAGGTTTACAAAGTTACTTACCAACCATTTCTAAATCAGCTATTTCTTCTGTT TTAGTTGGTTCTGAAGCTTTATACAGAGGTGATTTAACCGGTCCAGAATTAGCTTCTTA CATTTCCGAAATTAAATCCTTATTAGCAACTATTAACGATAAAGACGGTAACTCTTATG CTGATATTCCTGTCGGTACCGTCGACTCATGGAACATCTTGGTCAATGGTGAAGCTTAT GATGTTATTGAACAATCTACTGTGATGTACGCAAATGCTTTCCCATACTGGCAAGGTCA AACTCAACAAAACTCCAGTTTCTCATTCTTTGATGATATCATGCAAGCTTTACAAACTG TTCAAACTATCAAGGGTTCCACCGATTTTGATTTCTACATTGGTGAAACTGGTTGGCCA ACTGGAGGCTCCAACTATGGTGCCTCTGTTCCATCTGTTGACAACGCAGAAACTTTCTG GAAAGAGGGTATTTGTGCAATCAGAGGTTGGGGTGTTAACACTTTTGTTTTCGAAGCTT TCGATGAAGACTGGAAGCCAGATACTTCAGGCTCTTCTGTTGAAACTCATTGGGGTGTC TTTGATTCTAGCCGTAACCTCAAATACTCATTAGATTGTGATTTCTAA

Secuencia aminoacídica para la enzima 989

>g989.t1 <u>péptido señal</u>

MKLSVASITALAISLAAGPVSAIASLGFDLGVQDNGGACKVASDYEADLAILAPYSKMV KTYAVSTCNTLEVLGPVVESSGFQLTLGVWPTPQDKFTAEKEGLQSYLPTISKSAISSV LVGSEALYRGDLTGPELASYISEIKSLLATINDKDGNSYADIPVGTVDSWNILVNGEAY DVIEQSTVMYANAFPYWQGQTQQNSSFSFFDDIMQALQTVQTIKGSTDFDFYIGETGWP TGGSNYGASVPSVDNAETFWKEGICAIRGWGVNTFVFEAFDEDWKPDTSGSSVETHWGV FDSSRNLKYSLDCDF

Disulfide bond scores					
Cysteine sequence position	Distance	Bond	Score		
39 - 67	28	DNGGACKVASD-YAVSTCNTLEV	0.01044		
39 - 261	222	DNGGACKVASD-WKEGICAIRGW	0.16997		
39 - 308	269	DNGGACKVASD-KYSLDCDFXXX	0.01165		
67 - 261	194	YAVSTCNTLEV-WKEGICAIRGW	0.02443		
67 - 308	241	YAVSTCNTLEV-KYSLDCDFXXX	0.99386		
261 - 308	47	WKEGICAIRGW-KYSLDCDFXXX	0.01182		
Step 5: Weighted matching					
Predicted bonds					
39 - 261		DNGGACKVASD - WKEGICAIRGW			
67 - 308		YAVSTCNTLEV - KYSLDCDFXXX			
Predicted connectivity					
1-3, 2-4					

Predicción de puentes disulfuro para la enzima 989

Figura 45. Predicción de puentes disulfuro de la proteína 989 con el programa DiaNNA 1.1.

Secuencia codificante para la enzima 3296

>g3296.t1 CDS=1-1098

péptido señal

ATGAAGTTCTCCCAATTATCCGTTTTAACTTTTGCTGCTTTAACTGCTGCTGCTCCAAT GGCCATCAACCATCAACATCACAACCACAAGAGAGAGAGCTGAAGCCGTTGTCGTTATTG AAACTCGTACTGTCGTTGTTGACGCCGCTGATGTCACTGCCACTGCAGCTGCTATTGAC GTTGCTGCTGCAACCACCGATTCATCTGAATCAACTACTGCTCCAACTACTTTAGCAAC CTTCTTCCGGTAGTGGTTCATCCACTTTTGAAGACGGTACTATTCCATGTTCTGAATTC CCATCATCTTACGATGGTATGGTTTCTATCGACTGGTTAGGTTTTGGTGGTGGGCATC CGTTATGGCAATGGATGGTTCCACCTCAGAGTCATGTACCGATGGTTTTTACTGTTCTT AGATCAGTTGGTGGTTTATTATGTAAAGGTGGTTACCTTTACAGATCAAACACTGCTTC TGATTCTTTATGTGAATCTGACATGGACACCGGTTTTGTTGTCTCTGAATTATCAGAAT CCGTTGCTCTTTGTAGAACCGATTACCCTGGTTCCGAAGATATGGTTGTTCCAACTGAA TTACAAGCTGGTTCAACTGAACCATTGTCTGTCGTCAATGAAGATAACTATTTCGACTG GGAAGGTAAGAAGACTTCTTCTCAATACTATGTTAACAATGCTGGTGTCAGTATTGAAG ATGGTTGTATTTGGGGTACTTCTTCTGCTGGTGTCGGTAACTACGCTCCTCTTGTTATT GGTTCAGGTTACACCGATGGTTTGACTTACTTATCGTTGATTCCAAACCCAAACAACAA AAATGCTCCAAACTATAATGTTAAAATTGTTGCTACTGATGGCTCCACCATTAATGGTG ACCTCTGGTTCTGCCAACATTGTTTTCTACAATTAA

Secuencia aminoacídica para la enzima 3296

>g3296.t1 <u>péptido señal</u>

MKFSQLSVLTFAALTAAAPMAINHQHHNHKREAEAVVVIETRTVVVDAADVTATAAAID VAAATTDSSESTTAPTTLATSTTSAASAETSTGATEKKSTSSGSGSSTFEDGTIPCSEF PSSYDGMVSIDWLGFGGWASVMAMDGSTSESCTDGFYCSYACKPGMSKTQWPSNQPSDG RSVGGLLCKGGYLYRSNTASDSLCESDMDTGFVVSELSESVALCRTDYPGSEDMVVPTE LQAGSTEPLSVVNEDNYFDWEGKKTSSQYYVNNAGVSIEDGCIWGTSSAGVGNYAPLVI GSGYTDGLTYLSLIPNPNNKNAPNYNVKIVATDGSTINGDCSYINGEFSGSSDGCTVTV TSGSANIVFYN

Predicción de puentes disulfuro para la enzima 989

Disulfide bond scores				
Cysteine sequence position	Distance	Bond	Score	
115 - 150	35	DGTIPCSEFPS-STSESCTDGFY	0.01302	
115 - 156	41	DGTIPCSEFPS-TDGFYCSYACK	0.01064	
115 - 160	45	DGTIPCSEFPS-YCSYACKPGMS	0.0119	
115 - 185	70	DGTIPCSEFPS-VGGLLCKGGYL	0.69699	
115 - 201	86	DGTIPCSEFPS-ASDSLCESDMD	0.01568	
115 - 221	106	DGTIPCSEFPS-ESVALCRTDYP	0.01065	
115 - 278	163	DGTIPCSEFPS-SIEDGCIWGTS	0.1669	
115 - 336	221	DGTIPCSEFPS-TINGDCSYING	0.01038	
115 - 350	235	DGTIPCSEFPS-GSSDGCTVTVT	0.01054	
150 - 156	6	STSESCTDGFY-TDGFYCSYACK	0.01039	
150 - 160	10	STSESCTDGFY-YCSYACKPGMS	0.09731	
150 - 185	35	STSESCTDGFY-VGGLLCKGGYL	0.01042	
150 - 201	51	STSESCTDGFY-ASDSLCESDMD	0.01446	
150 - 221	71	STSESCTDGFY-ESVALCRTDYP	0.01057	
150 - 278	128	STSESCTDGFY-SIEDGCIWGTS	0.01199	
150 - 336	186	STSESCTDGFY-TINGDCSYING	0.01037	
150 - 350	200	STSESCTDGFY-GSSDGCTVTVT	0.01068	
156 - 160	4	TDGFYCSYACK-YCSYACKPGMS	0.32518	
156 - 185	29	TDGFYCSYACK-VGGLLCKGGYL	0.01153	
156 - 201	45	TDGFYCSYACK-ASDSLCESDMD	0.06415	
156 - 221	65	TDGFYCSYACK-ESVALCRTDYP	0.01083	
156 - 278	122	TDGFYCSYACK-SIEDGCIWGTS	0.79655	
156 - 336	180	TDGFYCSYACK-TINGDCSYING	0.0104	
156 - 350	194	TDGFYCSYACK-GSSDGCTVTVT	0.01108	
160 - 185	25	YCSYACKPGMS-VGGLLCKGGYL	0.01037	
160 - 201	41	YCSYACKPGMS-ASDSLCESDMD	0.01038	
160 - 221	61	YCSYACKPGMS-ESVALCRTDYP	0.01037	
160 - 278	118	YCSYACKPGMS-SIEDGCIWGTS	0.01089	
160 - 336	176	YCSYACKPGMS-TINGDCSYING	0.01037	
160 - 350	190	YCSYACKPGMS-GSSDGCTVTVT	0.01038	
185 - 201	16	VGGLLCKGGYL-ASDSLCESDMD	0.9973	
185 - 221	36	VGGLLCKGGYL-ESVALCRTDYP	0.40528	
185 - 278	93	VGGLLCKGGYL-SIEDGCIWGTS	0.01251	
185 - 336	151	VGGLLCKGGYL-TINGDCSYING	0.99687	
185 - 350	165	VGGLLCKGGYL-GSSDGCTVTVT	0.01364	
201 - 221	20	ASDSLCESDMD-ESVALCRTDYP	0.01069	
201 - 278	77	ASDSLCESDMD-SIEDGCIWGTS	0.01121	
201 - 336	135	ASDSLCESDMD-TINGDCSYING	0.01041	
201 - 350	149	ASDSLCESDMD-GSSDGCTVTVT	0.01054	
221 - 278	57	ESVALCRTDYP-SIEDGCIWGTS	0.01039	
221 - 336	115	ESVALCRTDYP-TINGDCSYING	0.9973	
221 - 350	129	ESVALCRTDYP-GSSDGCTVTVT	0.01087	
278 - 336	58	SIEDGCIWGTS-TINGDCSYING	0.01039	
278 - 350	72	SIEDGCIWGTS-GSSDGCTVTVT	0.01039	
336 - 350	14	TINGDCSYING-GSSDGCTVTVT	0.01128	
Step 5: Weighted matching				
Predicted bonds				
115 - 350		DGTIPCSEFPS - GSSDGCTV	TVT	
150 - 160		STSESCTDGFY - YCSYACKP	GMS	
156 - 278		TDGFYCSYACK - SIEDGCIW	GTS	
185 - 201		VGGLLCKGGYL - ASDSLCES	DMD	
221 - 336		ESVALCRTDYP - TINGDCSY	ING	
Predicted connectivity				
1-10, 2-4, 3-8, 5-6, 7-9				

Figura 46. Predicción de puentes disulfuro de la proteína 3296 con el programa DiaNNA 1.1.



Figura 47. SDS-PAGE del cultivo de expresión de His-tag-989 en un volumen de 25 mL, inducción con IPTG 5 mM overnight a 30°C. MPM: AccuRuler RGB Prestained Protein Ladder

Se observa una banda entre 93 y 130 kDa, marcada con la flecha, que se ve aumentada en intensidad en el inducido. Se puede observar que esta banda aparece de manera característica en los cultivos de inducidos pero no en los controles. No se observan otras diferencias.



Figura 48. SDS-PAGE del cultivo de expresión de His-tag-989 en un volumen de 25 mL, inducción con IPTG 5 mM distintos tiempos de inducción a 30°C. MPM: AccuRuler RGB Prestained Protein Ladder



Figura 49. IMAC de la FS del cultivo a 37ºC, 5 horas de inducción con 5 mM de IPTG. P − Percolado, L1 − Primeros lavados, sin imidazol, L2 − Segundos lavados con 20 mM de imidazol, E − Eluídos. MPM: AccuRuler RGB Prestained Protein Ladder



Figura 50. SDS-PAGE del cultivo de expresión de His-tag-989 en un volumen de 100 mL, inducción con IPTG 5 mM 5 horas de inducción a 37ºC. MPM: AccuRuler RGB Prestained Protein Ladder



Figura 51. IMAC de la FS del cultivo a 37ºC, 5 horas de inducción con 5 mM de IPTG. P − Percolado, L1 − Primeros lavados, sin imidazol, L2 − Segundos lavados con 20 mM de imidazol, E1 − Eluídos con imidazol 300 mM, E2 − Eluídos con imidazol 500 mM . MPM: AccuRuler RGB Prestained Protein Ladder



Figura 52. SDS-PAGE del cultivo de expresión de His-tag/DsbC-989 en un volumen de 25 mL, inducción con IPTG 1 mM durante 5 horas a 30°C y 37°C. MPM: AccuRuler RGB Prestained Protein Ladder



Figura 53. SDS-PAGE del cultivo de expresión de His-tag/DsbC-989 en un volumen de 25 mL, inducción con IPTG 5 mM overnight a 30℃. MPM: AccuRuler RGB Prestained Protein Ladder



Figura 54. SDS-PAGE del cultivo de expresión de His-tag/DsbC-989 en un volumen de 25 mL, inducción con IPTG 5 mM distintos tiempos de inducción a 30°C. MPM: AccuRuler RGB Prestained Protein Ladder



Figura 55. IMAC de la FS del cultivo de expresión de His-tag/DsbC-989 a 37℃, 3 horas de inducción con 5 mM de IPTG. P – Percolado, L1 – Primeros lavados, sin imidazol, L2 – Segundos lavados con 20 mM de imidazol, E – Eluídos. MPM: AccuRuler RGB Prestained Protein Ladder



Figura 56. SDS-PAGE del cultivo de expresión de His-tag-989 sin PS en un volumen de 100 mL, inducción con IPTG 5 mM durante 3 horas a 20º y 30ºC. MPM: AccuRuler RGB Prestained Protein Ladder