



UNIVERSIDAD NACIONAL DEL COMAHUE
Centro Regional Universitario Bariloche

Levaduras de ambientes extremos

caracterizaciones genómicas implicadas en sus adaptaciones

Trabajo de Tesis para optar por el Título de Doctora en Biología

Lic. Paula Nizovoy

Director Dr. Martín Moliné
Codirectores Dr. Nicolás Bellora
Dr. Diego Libkind

• 2023 •



Ἰησοῦς ὁ υἱὸς τοῦ ἀποστόλου
Παύλου ἐπέστειλε τὴν ἐπιστολήν
αὐτὴν πρὸς τὸν ἀδελφὸν
Καρπούριον τὸν υἱὸν τοῦ

Agradecimientos

Imposible concebir este trabajo sin la presencia de muchas personas que, en distintas etapas y de diferentes maneras, me acompañaron a lo largo de todo este proceso.

A todas ellas, toda mi gratitud.

Silvi, George, Andrés, Esteban, Nati, Juli, Frani, Ame, Blavi, Juani, Agus, Lola, Mica, Tina, Ro, Juanca, Eiza, Saibe, Edy, Maxi, Juli, Moli, Nico, Vito, Andriu, Eli, Andre, Cris y al mágico mixto de los miércoles y anexos.

Resumen

Los ambientes extremos albergan una amplia biodiversidad de microorganismos con adaptaciones únicas que les permiten sobrevivir y prosperar en estas condiciones: los extremófilos/extremotolerantes. El estudio de estos organismos permite entender y caracterizar mecanismos de adaptación, información de gran riqueza científica y con un gran potencial biotecnológico. La obtención de secuencias genómicas de organismos provenientes de estos ambientes, y su comparación con genomas de especies cercanas pero no vinculadas a estas condiciones, constituyen un material de partida rico para dicho propósito.

En esta Tesis Doctoral se presentan los genomas de 5 especies de levaduras, provenientes de 2 ambientes extremos distintos: *Naganishia vishniacii*, de valles secos en Antártida, y 4 especies del género *Goffeauzyma*, aisladas de drenajes ácidos de rocas con alta concentración de metales en solución. Para ambos grupos y 9 especies de referencia se ensamblaron y anotaron los genomas y se asignaron grupos de ortología al repertorio génico codificado. El análisis comparativo revela una reducción en el material genético para todas las especies extremotolerantes, tanto en cantidad de secuencias codificantes, como en la proporción de grupos funcionales multicopia, coherente con su asociación ambiente-específico y con una optimización en el uso de recursos en contextos de escasez de nutrientes. Estos organismos poseen también una gran cantidad de genes de función desconocida, cuyo estudio podrá proveer nuevas pautas para entender sus adaptaciones. Presentan una alta redundancia en transportadores de familia *Major Facilitator Superfamily* (MFS); una mayor representatividad/particularidades en genes vinculados con la biogénesis de la pared/membrana celular, constituyendo la defensa primaria para la tolerancia de estas especies; genes relacionados con la fotoprotección (PIKK) y la xerofilia (Sho1) en *N. vishniacii* y con la respuesta a daño oxidativo inducido por metales, el mantenimiento de la pared (proteínas RhoGAP) y la precipitación de metales (fosfatasas ácidas) en ambientes ácidos, en las especies del género *Goffeauzyma*. En estas últimas especies, también resaltan la importancia de las ATPasas de membrana vinculadas con el bombeo de iones y de nitrato reductasas, posiblemente implicadas en la síntesis de nanopartículas metálicas.

El análisis de uso de aminoácidos para *N. vishniacii* señala modificaciones globales del proteoma con un enriquecimiento en glutamina y una disminución de prolina, consistente con una mayor flexibilidad en las proteínas como adaptación a las bajas temperaturas. Finalmente, no se encontraron evidencias de estrategias de tolerancia (como las proteínas anticongelantes en relación a ambientes fríos y las metalotioneína en la detoxificación de metales), caracterizadas para otros microorganismos, lo que indica que los mecanismos de adaptación descritos para diversas especies no son universales.

Todos los rasgos genómicos caracterizados en esta Tesis mediante una estrategia bioinformática robusta permiten explicar la asociación de estas especies a sus ambientes restrictivos.

Abstract

Extreme environments host a wide biodiversity of microorganisms with unique adaptations that enable them to survive and thrive in such conditions – these are known as extremophiles. Studying these organisms provides insights into and characterizes mechanisms of adaptation, offering valuable scientific information of significant biotechnological potential. Obtaining genomic sequences from organisms inhabiting these environments and comparing them with genomes of phylogenetically related species not associated with those conditions provides a rich starting material for this purpose.

This Doctoral Thesis presents the genomes of 5 yeast species from 2 distinct extreme environments: Naganishia vishniacii from dry valleys in Antarctica, and 4 species of the Goffeauzyma genus isolated from acidic rock drainages with high metal concentrations. Genomes were assembled, annotated, and gene orthology groups were assigned for both taxa as well as 9 reference species. Comparative analysis reveals a reduction in genetic material for all extremophilic species, both in the quantity of coding sequences and the proportion of multicopy functional groups. This reduction is consistent with their environment-specific associations and resource optimization in nutrient-scarce contexts. These organisms also possess a significant number of genes with unknown functions, the study of which may provide new insights into their adaptations. They exhibit a high redundancy in Major Facilitator Superfamily (MFS) transporter genes, a greater representation of genes linked to cell wall/membrane biogenesis – forming the primary defense for species tolerance – and genes related to photoprotection (PIKK) and xerophily (Sho1) in N. vishniacii, as well as genes associated with metal-induced oxidative damage response, maintenance of the cell wall (RhoGAP proteins), and metal precipitation (acid phosphatases) in acidic environments in Goffeauzyma species. Also in that species, the importance of membrane ATPases linked to ion pumping and nitrate reductases, potentially involved in metal nanoparticle synthesis is highlighted.

Amino acid usage analysis for N. vishniacii indicates global proteome modifications characterized by an enrichment of glutamine and a decrease in proline, consistent with greater protein flexibility as an adaptation to low temperatures. Ultimately, the absence of evidence for classical tolerance strategies (such as antifreeze proteins for cold environments and metallothioneins for metal detoxification) suggests that the adaptation mechanisms described for various species are not universal.

All the genomic traits characterized in this Thesis through a robust bioinformatic strategy help explain the association of these species with their restrictive environments.

Índice general

Agradecimientos	ii
Resumen	iii
<i>Abstract</i>	v
Índice general	vi
Índice de tablas y figuras	ix
Capítulo 1 - Ambientes extremos: introducción general para acceder a mundos distintos	1
1.1 INTRODUCCIÓN	2
1.1.1 Las levaduras	2
1.1.2 Levaduras de ambientes extremos	3
1.1.3 Abordajes genómicos para el estudio de la biodiversidad microbiana	4
1.1.4 Importancia biotecnológica	7
1.1.5 Hipótesis	10
1.1.6 Objetivo general	10
Capítulo 2 - La evolución no se congela: lecciones de <i>Naganishia vishniacii</i>, una levadura antártica	11
2.1 INTRODUCCIÓN	12
2.1.1 Los ambientes fríos	12
2.1.2 Microorganismos en ambientes fríos	13
2.1.3 El género <i>Naganishia</i>	14
2.1.4 Mecanismos de adaptación	17
2.1.4.1 Reestructuración de la membrana	17
2.1.4.2 Producción de proteínas anticongelantes (AFPs)	18
2.1.4.3 Compuestos crioprotectores	19
2.1.4.4 Compuestos y pigmentos fotoprotectores	19
2.1.4.5 Adaptaciones enzimáticas	21
2.1.5 Objetivos específicos	23
2.2 MATERIALES Y MÉTODOS	24
2.2.1 Secuenciación y ensamblado del genoma de <i>N. vishniacii</i>	24
2.2.1.1 Predicciones de genes y anotaciones	25
2.2.2 Reconstrucción filogenómica	26
2.2.3 Grupos de ortología y asignaciones funcionales	27
2.2.4 Análisis del cociente dN/dS	28
2.2.5 Búsqueda de genes involucrados en la respuesta a frío	28
2.2.5.1 Confirmación exp. de la producción de carotenoides y micosporinas en <i>N. vishniacii</i>	30
2.2.6 Análisis del uso de aminoácidos	30

2.3 RESULTADOS	31
2.3.1 Secuenciación del genoma, ensamblado, predicciones de genes y anotaciones	31
2.3.2 Reconstrucción filogenómica	31
2.3.3 Grupos de ortología y asignaciones funcionales	32
2.3.4 Análisis del cociente dN/dS	35
2.3.5 Búsqueda de genes involucrados en la respuesta a frío	38
2.3.5.1 Confirmación exp. de la producción de carotenoides y micosporinas en <i>N. vishniacii</i>	39
2.3.6 Análisis del uso de aminoácidos	39
2.4 DISCUSIÓN	41
2.4.1 Secuenciación del genoma, ensamblado, predicciones de genes y anotaciones	41
2.4.3 Grupos de ortología y asignaciones funcionales	42
2.4.4 Análisis del cociente dN/dS	44
2.4.5 Búsqueda de genes involucrados en la respuesta a frío	47
2.4.6 Análisis del uso de aminoácidos	50
2.5 CONCLUSIONES	53
Capítulo 3 - Sobrevivir con menos, resistiendo más: testimonio genómico de un clado ácido y metal-tolerante	54
3.1 INTRODUCCIÓN	55
3.1.1 Ambientes acuáticos ácidos con presencia de metales	55
3.1.1.1 La acidificación de los ambientes	56
3.1.1.2 Sitios representativos	57
3.1.2 Microorganismos en ambientes acuáticos ácidos	59
3.1.2.1 Levaduras en ambientes acuáticos ácidos y con presencia de metales	61
3.1.3.1 El género <i>Goffeauzyma</i>	61
3.1.4 Mecanismos de adaptación	64
3.1.4.1 Adaptaciones a bajo pH	65
3.1.4.2 Adaptación a metales	65
3.1.5 Objetivos específicos	70
3.2 MATERIALES Y MÉTODOS	71
3.2.1 Secuenciación y ensamblado de genomas de especies de género <i>Goffeauzyma</i>	71
3.2.1.1 Predicciones de genes y anotaciones	72
3.2.2 Reconstrucción filogenómica	73
3.2.3 Grupos de ortología y asignaciones funcionales	73
3.2.3.1 Análisis de duplicaciones en los grupos funcionales multicopia	74
3.2.4 Análisis de cociente <i>dN/dS</i>	75
3.2.5 Búsqueda de genes de interés	76
3.3 RESULTADOS	79
3.3.1 Secuenciación y ensamblado de genomas de especies de <i>Goffeauzyma</i>	79

3.3.2 Reconstrucción filogenómica	80
3.3.3 Grupos de ortología y asignaciones funcionales	81
3.3.3.1 Análisis de duplicaciones en los grupos funcionales multicopia	85
3.3.4 Análisis de cociente dN/dS	92
3.3.5 Búsqueda de genes de interés	96
3.4 DISCUSIÓN	102
3.4.1 Secuenciación del genoma, ensamblado, predicciones de genes y anotaciones	102
3.4.3 Grupos de ortología y asignaciones funcionales	105
3.4.3.1 Análisis de duplicaciones en los grupos funcionales multicopia	111
3.4.4 Análisis de cociente dN/dS	114
3.4.5 Búsqueda de genes involucrados en la respuesta a metales y acidez	115
3.5 CONCLUSIONES	117
Conclusiones finales – Genómica al extremo	119
REFERENCIAS	121
TRABAJO DE RESPALDO	144

Índice de tablas y figuras

Tablas

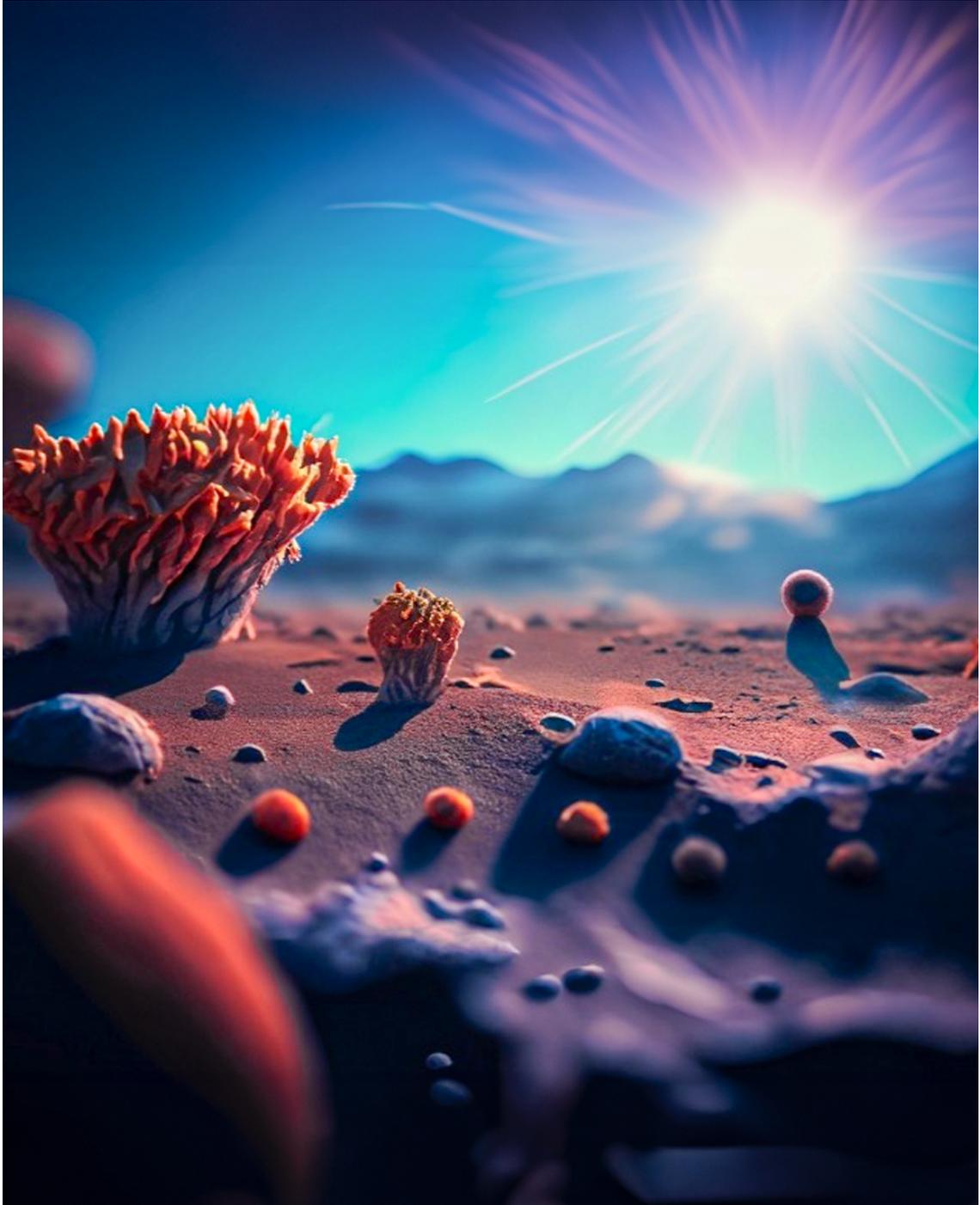
2.1.1	Lugares de aislamiento y rangos de temperatura de crecimiento para las especies del género <i>Naganishia</i>	15
2.2.1	Secuencias <i>query</i> utilizadas para las búsquedas de genes de interés involucrados en la respuesta a frío	29
3.2.1	Cepas del género <i>Goffeauzyma</i> secuenciadas en este trabajo	71
3.2.2	Secuencias <i>query</i> utilizadas para las búsquedas de genes de interés de vías no enzimáticas de respuesta a estrés oxidativo	77
3.2.3	Secuencias <i>query</i> utilizadas para las búsquedas de genes de interés de vías enzimáticas de respuesta a estrés oxidativo	78
3.3.1.	Resumen de estadísticos del ensamblado y anotación de las cuatro especies del género <i>Goffeauzyma</i> y las 3 especies hermanas estudiadas	80
3.3.2	Grupos de ortología que se encontraron únicamente en las especies del clado ARD	83
3.3.3	Grupos de ortología que se encontraron únicamente en las especies del clado MSD	84
3.3.4	Grupos de ortología que se encontraron únicamente en las especies del clado Tolerans	85
3.3.5	Grupos de ortología para los que se verificaron expansiones en nodos de la filogenia del clado ARD	91
3.3.6	Resultados de las búsquedas por homologías para vías de respuesta a estrés oxidativo	98
3.3.7	Grupos de ortología relevados en las 7 especies en función de su posible vinculación con la resistencia a metales/acidez extrema	101

Figuras

2.1.1	Valles Secos de McMurdo	13
2.1.2	<i>N. vishniacii</i> creciendo en laboratorio	16
2.3.1	Árbol filogenómico para <i>N. vishniacii</i> y grupo externo	32
2.3.2	Análisis de ortología para las 7 cepas del orden Filobasidiales	33
2.3.3	Grupos de ortología en los que el n° de copias de <i>N. vishniacii</i> fue mayor que el de <i>N. albida</i>	34
2.3.4	Distribución de valores promedio de dN/dS para <i>N. vishniacii</i> y <i>N. albida</i>	36
2.3.5	Cambios en la frecuencia de uso de aminoácidos para <i>N. vishniacii</i> , <i>N. albida</i> Y-1402-NT2 y <i>N. adeliensis</i> con respecto a <i>C. neoformans</i>	40
3.1.1	Árbol filogenético del género <i>Goffeauzyma</i> y las especies más cercanas del orden Filobasidiales	63
3.2.1	Esquema de trabajo para la comparación de las especies de interés	76
3.3.1	Árbol filogenómico para el clado ARD y grupo externo	81
3.3.2	Análisis de ortología para las especies de <i>Goffeauzyma</i> y las 3 especies del grupo externo	82
3.3.3	Distribución de grupos funcionales multicopia especies de género <i>Goffeauzyma</i>	86
3.3.4	Mapa de calor de número de copias para los grupos de ortología en los que una o más especies del género <i>Goffeauzyma</i> supera a las especies externas	89
3.3.5	Distribución de valores promedio de dN/dS para <i>G. aciditolerans</i>	93
3.3.6	Distribución de valores promedio de dN/dS para <i>G. metallitolerans</i>	94
3.3.7	Distribución de valores promedio de dN/dS para <i>G. iberica</i>	95
3.3.8	Distribución de valores promedio de dN/dS para <i>G. agrionensis</i>	96

Las ilustraciones de las portadas fueron generadas a demanda con Adobe Firefly

Capítulo 1



Ambientes extremos

introducción general para acceder a mundos distintos

1.1 INTRODUCCIÓN

«I guess I'm trying to subjectify the universe, because look where objectifying it has gotten us. To subjectify is not necessarily to co-opt, colonize, exploit. Rather it may involve a great reach outward of the mind and imagination.»

Ursula K. Le Guin

1.1.1 Las levaduras

Las levaduras son un grupo de hongos ampliamente distribuidos, caracterizados por un crecimiento predominantemente unicelular (Kurtzman et al., 2011). Son principalmente organismos descomponedores de vida libre que cumplen un rol muy importante en los ecosistemas donde habitan, tanto regenerando la materia orgánica, como mineralizando nutrientes, lo cual los vuelve disponibles para otros organismos y posibilita así, su crecimiento y desarrollo (Péter et al., 2017). Si bien son inmóviles, la capacidad de liberar esporas hace que puedan dispersarse largas distancias, incluso entre continentes (Vincent, 2000), gracias a la acción de agentes bióticos y abióticos (por ejemplo, plantas o animales, viento, cursos de agua, etc.). Además, tienen la capacidad de generar interacciones ecológicas con otros organismos actuando como simbioses, parásitos o patógenos (Rossit et al., 2007; Polvi et al., 2015; Dias et al., 2019). Por otro lado, son un grupo de organismos con el cual los humanos han interactuado durante milenios, principalmente por su capacidad de fermentar sustratos azucarados, capacidad extensamente aprovechada para la preparación de alimentos y bebidas fermentadas (Burini, 2023). Sin embargo, aunque son generalmente aisladas de medios ricos en nutrientes, su gran versatilidad fisiológica y metabólica les posibilita adaptarse a una amplia variedad de condiciones, pudiendo ser aisladas de prácticamente todos los ambientes estudiados: tanto de los naturales, como de los de origen antrópico, e incluso de aquellos considerados inhóspitos para otras formas de vida.

1.1.2 Levaduras de ambientes extremos

Todos los organismos deben responder a las condiciones de su entorno. Si bien pequeñas fluctuaciones ambientales ocurren en muchos hábitats, una desviación significativa de un óptimo ecológico resulta estresante para la mayoría de las formas de vida. La definición de ambientes extremos no es absoluta y aquello que para un organismo son condiciones subóptimas, puede no serlo para otro que las encuentre ideales para crecer. Las adaptaciones de algunos microorganismos a lo que los humanos percibimos como hostil, son notables y es lo que les permite colonizar todos los ambientes terrestres. En este sentido los organismos poseen un rango característico de condiciones en el cual ocurre su crecimiento para cada parámetro ambiental. Este rango se encuentra definido por valores mínimos y máximos, más allá de los cuales no pueden sobrevivir, estableciendo límites a su capacidad de colonización y dispersión: lo que comúnmente se define como la Ley de tolerancia de Shelford (Shelford, 1913). Dentro de este rango se encuentra un valor óptimo que garantiza una tasa de crecimiento y viabilidad máximas. Los factores que condicionan la vida pueden ser tanto físicos (como por ejemplo la temperatura, la radiación, o la luz), químicos (por ejemplo, la acidez, la salinidad y el estrés osmótico), o incluso una combinación de ambos. La disponibilidad de nutrientes también es considerada un factor condicionante y es determinante en los ambientes oligotróficos (aquellos donde los recursos nutricionales son escasos).

Los organismos que crecen en condiciones adversas para la mayoría de las formas de vida son conocidos como extremófilos y normalmente se suele distinguir entre verdaderos extremófilos (aquellos que requieren de esas condiciones para vivir) y extremotolerantes (los que tienen la capacidad de tolerarlas hasta cierto punto). Algunos autores sugieren que la línea entre estos dos términos es a veces muy delgada y que a menudo no se puede discernir entre uno y otro sin tener un conocimiento profundo de la biología y ecología de los organismos (Raspor & Zupan, 2006). Aunque la clasificación de las levaduras como extremófilas o extremotolerantes no es tan frecuente como la de los procariotas (arqueas y bacterias), algunas especies crecen bajo condiciones que son comúnmente consideradas extremas. En particular, una especie de levadura es extremófila si: (a) ha sido aislada repetidamente de un hábitat extremo, (b) muestra capacidades fisiológicas para superar el estrés ambiental particular en cultivo y (c) tiene su óptimo de crecimiento en el rango

correspondiente a las condiciones encontradas en el hábitat particular. Es importante destacar, que el hecho de que una levadura sea aislada o detectada en un ambiente extremo no es condición suficiente para que sea considerada extremófila, debido a que puede no cumplir con el resto de los requisitos. Por otra parte, las levaduras extremotolerantes son aquellas que tienen la capacidad de crecer en condiciones físicas o químicas extremas aunque éstas no se encuentren en los valores óptimos de su crecimiento. Algunas levaduras también pueden ser clasificadas como poliextremófilas o politolerantes, en función de su capacidad de prosperar o tolerar una combinación de condiciones consideradas extremas (Buzzini et al., 2018).

Los ambientes extremos se caracterizan por poseer condiciones que requieren adecuaciones particulares para que la biodiversidad que habita en ellos pueda crecer y sobrevivir (Aguilera et al., 2020). Las adaptaciones de los organismos a sus ambientes surgen mediante la selección natural actuando sobre la variabilidad preexistente.

En el caso particular de los microorganismos, la tasa de mutación suele ser mayor que la observada para plantas y animales, lo que posibilita generar variantes alélicas o el surgimiento de nuevos locus con mayor velocidad, que en conjunción con las condiciones de selección impuestas por el ambiente, les permite adaptarse (Rudakiya & Patel, 2021). En este sentido los miembros autóctonos de estos ambientes poseen rasgos particulares, que usualmente se encuentran en consonancia con las condiciones impuestas por el ambiente, que no se encuentran presentes en sus congéneres y para los cuales la huella de la selección natural puede ser identificada. Es por ello, que estos ambientes son medios favorables para el estudio de mecanismos de adaptación únicos y de una biodiversidad específicamente asociada a ellos (Aguilera et al. 2019).

1.1.3 Abordajes genómicos para el estudio de la biodiversidad microbiana

La taxonomía y la descripción de nuevas especies de microorganismos tradicionalmente se basó en características morfológicas y fenotípicas, como la forma y el aspecto del crecimiento en medios de cultivo, las condiciones de crecimiento y la asimilación de sustratos (Kurtzman et al., 2011). Estos enfoques fueron esenciales para establecer la identidad de las levaduras pero presentaban limitaciones y eran fuente de subjetividades, impidiendo discernir entre especies muy similares y dificultando la comprensión de la

diversidad en estos microorganismos. El surgimiento de técnicas moleculares (principalmente la secuenciación de marcadores y el uso de herramientas comparativas como BLAST - Basic Local Alignment Tool -) posibilitó un abordaje más refinado a la taxonomía de estos organismos, que complementó las primeras caracterizaciones (White et al., 1990; Altschul et al., 1997). Más recientemente, los avances en el campo de la genómica, con las nuevas tecnologías de secuenciación, posibilitaron la obtención de genomas completos o casi completos y sumado a la metagenómica permitieron modificar nuestra capacidad de describir, diferenciar y aprehender una vasta biodiversidad de microorganismos, incluyendo levaduras (Hittinger et al., 2015).

Se estima que existen cerca de 50000 especies de levaduras (Boekhout et al., 2022), de las cuales más de 1500 han sido identificadas y descritas en distinto grado. De todas éstas, tan solo un quinto cuentan con su genoma secuenciado (Libkind et al., 2020a). Sin embargo, este número crece de forma sostenida (Boekhout et al., 2022). La genómica de levaduras ha permitido evidenciar una extensa divergencia de secuencias observada entre diferentes linajes. Divergencia que llega hasta el nivel de especie y refleja cambios genómicos intensos que contrastan con la conservación de propiedades biológicas de las levaduras durante períodos evolutivos muy largos. En lugar de ofrecer un rango continuo de adaptaciones evolutivas graduales, como se propone en la teoría darwiniana clásica, los genomas de distintos clados de levaduras, o incluso de especies dentro del mismo género, pueden diferir entre sí de manera abrupta (Dujon, 2010). El volumen creciente de estudios genómicos -y de otras disciplinas ómicas-, permite la obtención de filogenias más robustas, fundamentales para comprender la evolución, filogeografía, ecología e incluso el potencial biotecnológico de las levaduras. La caracterización de genes o vías metabólicas completas permite establecer relaciones entre el acervo génico y los rasgos fenotípicos (Libkind et al., 2020b).

Originalmente los estudios en levaduras se han centrado en el género *Saccharomyces*, particularmente en *Saccharomyces cerevisiae* -el organismo modelo por excelencia y primer eucariota del que se obtuvo la secuencia de genoma completo (Goffeau et al., 1996). La genómica de *S. cerevisiae* es bien conocida, con un genoma de aproximadamente 12 Mpb organizado en 16 cromosomas y un genoma mitocondrial en torno a los 85 kpb. En la actualidad existen más de 1600 ensamblados disponibles correspondientes a centenares de cepas de distintos orígenes (Boekhout et al., 2022; datos de NCBI). Esta sobrerrepresentación de genomas para una especie es producto de la importancia

biotecnológica de *S. cerevisiae*, la cual es usada para la producción de alimentos y bebidas fermentadas (Radecka et al., 2015), así como para la producción de bioetanol (Argueso et al., 2009). Sin embargo, la biotecnología y la necesidad de producir nuevos compuestos y metabolitos para diversas industrias, como la producción de bioetanol, bioplásticos o metabolitos activos, así como la necesidad de realizar tratamiento de efluentes o procesos de biorremediación, encuentran solución en el estudio de otras especies de microorganismos.

Las levaduras no convencionales se refieren a todas aquellas especies que no son organismos modelos y cuentan con rasgos promisorios para su utilización en distintos procesos biotecnológicos (Godinho et al., 2021). La diversidad de éstas incluye a miles de especies pertenecientes a los *phyla* basidiomycota y ascomycota que contienen todos los hongos dicariontes (Choi & Kim, 2017). La disponibilidad de mayor cantidad de genomas de estas levaduras permite el desarrollo de herramientas y estrategias de ingeniería genética y metabólica para explotar su potencial como biofábricas en relación con entornos estresantes de distintos bioprocesos (Godinho et al., 2021). En este sentido, los abordajes genómicos y las herramientas moleculares ofrecen la posibilidad de revelar las bases moleculares de los rasgos particulares de estas levaduras y explorar sus potencialidades (Radecka et al., 2015).

La comparación de genomas de especies estrechamente relacionadas y la determinación de las relaciones filogenéticas entre secuencias génicas proveen de un marco integral para entender la evolución y la biodiversidad, posibilitando la extrapolación de conocimiento biológico entre organismos (Emms & Kelly, 2019) y generando las bases para revelar las diferencias genéticas que subyacen a diferentes estrategias de adaptación y para desentrañar su historia evolutiva (Tshisekedi et al., 2021). Para ampliar el conocimiento sobre los genes en una especie modelo a otras especies relacionadas, es importante distinguir entre genes directamente relacionados a través de la especiación (ortólogos) de aquellos genes que se han duplicado de manera independiente a la especiación (parálogos). Estos últimos pueden ser el resultado de diversos tipos de duplicación génica (recombinación desigual, duplicaciones mediadas por transposones o poliploidía), y pueden haber ocurrido recientemente o mucho antes de algún evento de especiación de interés (Cannon & Young, 2003). Las duplicaciones se consideran uno de los principales motores de la evolución al proporcionar material genético adicional, que posteriormente puede remodelarse en productos génicos “nuevos” (Ohno, 1970). En el caso particular de las

levaduras es destacable el evento de duplicación genómica completa ocurrido dentro de los Sacaromycetales y que ha sido fundamental para explicar el origen y la evolución del linaje de *S. cerevisiae* (Wolfe, 2015). Sin embargo, otros mecanismos de duplicación de menor magnitud también son fundamentales para explicar el origen de nuevos genes y rasgos evolutivos en todas las levaduras (Dujon, 2010). Del mismo modo, mecanismos evolutivos importantes para la evolución de distintos linajes de levaduras son producidas por frecuentes eventos de transferencias horizontales (Shen et al., 2018), inserciones, deleciones y evolución de secuencias nucleotídicas mediante presiones de selección positivas (Yue et al., 2017). Todos estos mecanismos han moldeado los genomas de los hongos y levaduras durante 1800 millones de años produciendo una plétora de características y una diversidad comparable, o aún mayor, que la observada para animales y vegetales (Choi & Kim, 2017; Li et al., 2021).

A pesar de los grandes avances en nuestra comprensión de los hongos logrados mediante la genómica comparativa, las aproximaciones de secuenciación por sí solas no son suficientes para resolver la mayoría de los problemas biológicos. La eficacia de un enfoque comparativo depende del conocimiento funcional previo, el cual luego puede extrapolarse a otros organismos. La mayor parte del conocimiento funcional en hongos proviene de unas pocas especies modelo y es muy limitado fuera de Ascomycota. Por lo que es necesario obtener información empírica directa de una mayor diversidad de especies. Esto no solo aumentará nuestra capacidad para inferir la función basada en homología, sino que también nos permitirá explorar el ámbito de los genes específicos de linajes (Naranjo-Ortiz & Gabaldón, 2020).

1.1.4 Importancia biotecnológica

Originalmente considerados curiosidades científicas, los organismos extremófilos y extremotolerantes son hoy en día ampliamente reconocidos por su potencial biotecnológico. Las extremozimas (enzimas adaptadas para funcionar en condiciones extremas) y los compuestos bioactivos que estos organismos producen generan un gran aporte a diversos campos de aplicación (Sharma et al., 2016). Entre las levaduras, las no-sacaromicéticas representan un nicho poco explorado con un gran potencial en la producción de compuestos de alto valor agregado, como los biocombustibles, los carotenoides, proteínas

recombinantes, diversas enzimas, toxinas biocontroladoras, vacunas, entre otros (Segal-Kischinevsky et al., 2022). Cada ambiente extremo representa, por tanto, más que una limitación, una ventana abierta a nuevas potencialidades.

En esta Tesis, se abordan en particular 2 tipos de ambientes extremos vinculados con distintos factores estresantes: los Valles Secos de Mckurdo, en Antártida, desafiantes principalmente por su baja disponibilidad de agua en estado líquido y sus bajas temperaturas; y 2 sistemas análogos en sus condiciones, pero de distinto origen, que representan ambientes acuáticos ácidos con alta concentración de metales en solución: el Sistema Río Agrio - Lago Caviahue, en Neuquén, de origen volcánico, y los drenajes ácidos de minas de la Faja Pirítica Ibérica, en la Península Ibérica, asociados con la actividad minera. Ambos tipos de ambientes poseen características que los vuelven blancos de interés para el estudio de su biodiversidad y de los mecanismos de adaptación que dichos organismos albergan para sustentar su subsistencia, tanto desde una perspectiva puramente científica, como aplicada, debido a que son fuente de un gran potencial biotecnológico.

Los ambientes fríos, en particular, proveen la posibilidad de explotar la producción de enzimas activas a bajas temperaturas. En levaduras psicrófilas, los estudios se han centrado principalmente en la búsqueda de detergentes orgánicos que posibiliten un lavado más amigable con el ambiente, la producción de celulasas para el procesamiento de textiles y biocombustibles, y de enzimas para la industria alimentaria (Buzzini et al., 2012). La degradación de compuestos aromáticos e hidrocarburos a bajas temperaturas, así como la producción de biodiesel en levaduras oleaginosas, también han sido potenciales explorados en estas extremófilas (Amaretti et al., 2010; Fernández et al., 2017). El principal interés en estas extremozimas radica en que son requeridas en menores concentraciones; no requieren ayuda térmica adicional, lo cual supone un ahorro energético; evitan transformaciones químicas indeseadas; previenen la pérdida de compuestos volátiles y permiten obtener altos rendimientos en reacciones con componentes termosensibles; y pueden ser selectivamente inactivadas con una leve aplicación de calor (Margesin, 2002; Kuddus, 2018).

Los ambientes ácidos, por otro lado, son fuente de extremozimas tolerantes a pHs bajos que resultan beneficiosas para satisfacer los requerimientos de los procesos industriales actuales, muchos de los cuales ocurren en esas condiciones (Hassan et al., 2019). Además, luego de millones años de evolución en ambientes con alta concentración de metales, las levaduras que allí prosperan son resistentes a este tipo de contaminantes y

tienen la capacidad de remediar la contaminación por metales. Los abordajes tradicionales (precipitación química, ósmosis inversa, precipitación por óxido-reducción, coagulación con cal, electrodiálisis, evaporación e incineración) suponen un alto costo, gran consumo de solventes químicos y de electricidad, ineficiencia en la remoción y generación de barros tóxicos residuales (Rudakiya & Patel, 2021). Frente a este escenario, las tecnologías basadas en biorremediación han atraído gran interés en base a su bajo costo, eficiencia de remoción incluso a bajas concentraciones de metales y a ser abordajes amigables con el ambiente. Por otra parte, una de las formas en las que sucede la captación de metales del ambiente es mediante la formación de nanopartículas de metal en la superficie de las células, o bien en su interior. Esta vía, además de ser vehículo para la remoción de los contaminantes, es aprovechada para el desarrollo de nanotecnología verde, que, mediante la combinación de principios biológicos con procesos físicos y químicos permite generar nanopartículas de manera eco-amigable y con características promisorias para múltiples campos de aplicación, como la medicina, la bioingeniería, la catálisis y las ciencias ambientales (Zhao et al., 2018).

En vistas de su gran potencial biotecnológico, el estudio y caracterización de los mecanismos de adaptación de los organismos de estos ambientes extremos resulta promisorio y constituye la base de este trabajo.

1.1.5 Hipótesis

Las características que permiten la supervivencia y crecimiento de levaduras aisladas de ambientes con condiciones consideradas extremas -y restrictivas- para muchas otras formas de vida pueden ser evidenciadas a nivel genómico. Esto supone que la comparación de los genomas de especies de ambientes extremos con los de otras que no se hallen asociadas a estas condiciones, pero que sean filogenéticamente cercanas, permitirá encontrar características particulares que sustenten la adaptación de las primeras. Dicha premisa puede reflejarse en: números de copia diferenciales en genes que posibiliten una mejor sobrevivencia; presencia de genes vinculados con mecanismos de tolerancia ya reportados en otros organismos adaptados a condiciones análogas (y ausentes en las especies que no se hallan en esos ambientes); regiones génicas que hayan estado sujetas a evolución positiva o atestigüen expansiones por duplicaciones en genes vinculados con la adaptación/tolerancia.

1.1.6 Objetivo general

A partir de abordajes de genómica comparativa, identificar las bases genéticas y genómicas de la extremofilia de dos grupos de especies:

- la levadura psicrófila *Naganishia vishniacii* (o. Filobasidiales, cl. Tremellomycetes, div. Basidiomycota), hasta el momento aislada únicamente de ambientes fríos, en contraste con especies filogenéticamente emparentadas pero de distinta distribución y

- 4 especies del género *Goffeauzyma* (o. Filobasidiales, cl. Tremellomycetes, div. Basidiomycota) que se han visto asociadas a ambientes acuáticos ácidos y con alta concentración de metales, en contraposición con los genomas disponibles de especies cercanas sin reportes en dichas condiciones.

Identificar características diferenciales en las especies de interés respecto de las especies cercanas, y características generales que sean comunes a ambos modelos, realizando inferencias evolutivas respecto de la adquisición de dichas particularidades adaptativas.



Capítulo 2

La evolución no se congela

lecciones de *Naganishia vishniacii*, una levadura antártica

2.1 INTRODUCCIÓN

2.1.1 Los ambientes fríos

Una gran proporción de la superficie terrestre está cubierta por ambientes con temperaturas permanentemente bajas: desde el Ártico a la Antártida y desde la profundidad de los océanos hasta las altas montañas. Buena parte de los océanos tienen temperaturas entre 2,5 y 5 °C, ocupando cerca del 70 % del planeta; las regiones polares constituyen el 20 %, con tan solo 0,5 % estacionalmente libre de nieve y/o hielo; el resto de los ambientes fríos se distribuyen principalmente entre las regiones montañosas de Europa (los Alpes), Asia (los Himalayas) y América (las Montañas Rocosas y los Andes) (Cowan et al., 2014).

En muchos de los ecosistemas fríos, las bajas temperaturas están asociadas a otros factores limitantes, como baja disponibilidad de agua líquida, ciclos periódicos de congelamiento y deshielo, condiciones de estrés osmótico y oxidativo y alta radiación solar, *inter alia*, lo que hace que sean ambientes extremadamente inhóspitos para muchas formas de vida. Sin embargo, han sido exitosamente colonizadas por organismos psicrófilos y psicrotolerantes que constituyen una biota nativa adaptada capaz de prosperar en esas condiciones e incluso de mantener actividad metabólica a temperaturas bajo cero (Cowan et al., 2007; Margesin & Miteva, 2011; Convey & Peck, 2019). En general, se considera que los organismos psicrófilos no pueden crecer por encima de los 20 °C y su temperatura óptima es de 15 °C, mientras que los psicrotolerantes tienen un óptimo de crecimiento más alto (20 - 30 °C) pero pueden crecer a 0 °C (Cowan et al., 2014). Las bajas temperaturas influyen sobre los microorganismos tanto de forma indirecta (debido al cambio del estado físico del agua), como directa (en función de una menor tasa metabólica debido a una actividad enzimática reducida). Es por ello que, la habilidad de estos organismos para sobrevivir y proliferar a bajas temperaturas indica que tienen la capacidad de sortear una serie de desafíos inherentes a los ambientes permanentemente fríos.

Los Valles Secos de McMurdo (Figura 2.1.1), localizados en la región desglaciada de Antártida oriental, constituyen el ejemplo más extremo de suelos polares y son considerados los desiertos más fríos y secos de la Tierra, con contenidos de agua en el suelo no mayor al 2 % en peso (Wynn-Williams et al., 1990). En invierno, un período de oscuridad continua, las temperaturas pueden alcanzar los -60 °C y los vientos, velocidades de hasta 100 km/h. En el verano, con temperaturas de -30 °C a -15 °C, las temperaturas promedio de

la superficie del suelo pueden alcanzar los 15 °C en cuestión de minutos debido a la radiación solar directa, intercalada con cobertura nubosa (Wynn-Williams et al., 1990). Las nevadas, única forma de precipitación posible, son extremadamente escasas y se concentran mayormente en el verano. Sin embargo, las rocas humedecidas por el deshielo de la nieve retienen agua líquida durante varios días (Musilova et al., 2015). Sus suelos, además, contienen los niveles más bajos de carbono orgánico en la Tierra, junto con las elevaciones más altas de las regiones desérticas de Atacama y Patagonia en Chile y Argentina (Burkins et al., 2000). La combinación de todos estos factores contribuye a generar un ambiente de vida extremadamente hostil que se ha propuesto como un posible análogo de las condiciones ambientales en Marte (de la Torre et al., 2006; Gilichinsky et al., 2007; Musilova et al., 2015).

2.1.2 Microorganismos en ambientes fríos

A pesar de las limitantes que imponen estas condiciones sobre buena parte de las formas de vida, representantes de los 2 dominios, tanto arqueas, como bacterias y eucariotas (hongos filamentosos, levaduras, líquenes, protistas y pequeños invertebrados), han sido aislados de ambientes fríos (Margesin & Miteva, 2011; Connell et al., 2014; Zalar & Gunde-Cimerman, 2014; Margesin & Collins, 2019).

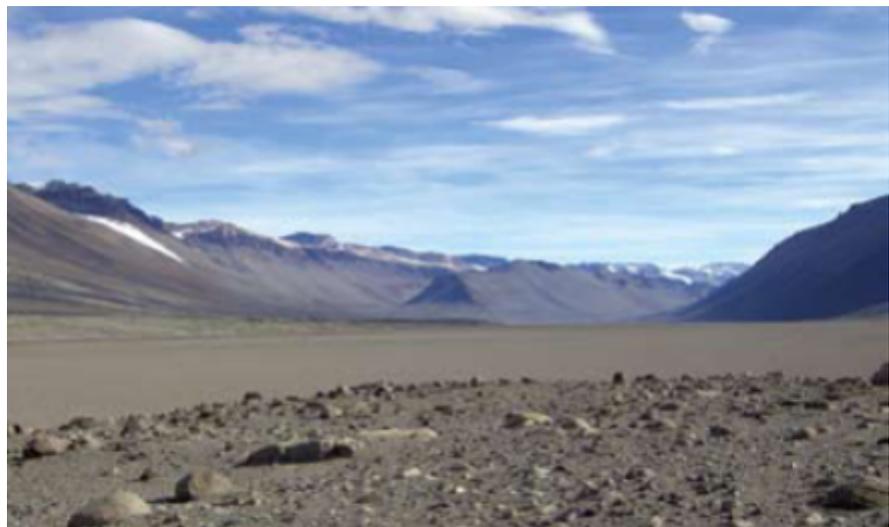


Figura 2.1.1. Valles Secos de McMurdo. Tomada de Connell et al. 2014

Entre las especies de levaduras aisladas de hábitats

fríos, las levaduras basidiomicéticas de los géneros *Glaciozyma*, *Leucosporidium*, *Mrakia*, *Naganishia* y *Goffeauzyma* suelen ser las más frecuentes. Sin embargo, a pesar de que numerosos estudios destacaron su capacidad para colonizar este tipo de ambientes, estas especies también han sido reportadas en hábitats no extremos en climas boreales y templados, como turberas, suelos forestales, filoplano, lagos y arroyos (Buzzini et al., 2018).

El conocimiento de levaduras adaptadas al frío habitando estos ambientes es escaso (Alcaíno et al. 2015). Sin embargo, el advenimiento de las tecnologías ómicas, con la secuenciación de genoma completo de organismos que habitan lugares fríos, representa un nuevo enfoque para su estudio, posibilitando la comprensión e identificación de las particularidades de estas formas de vida, así como el descubrimiento y caracterización de genes responsables de su adaptación (Yadav et al. 2018). En este sentido, la creciente disponibilidad de genomas secuenciados de levaduras de ambientes extremos fríos y sus congéneres mesófilos permiten incrementar y profundizar nuestro conocimiento sobre las adaptaciones particulares, identificando nuevos mecanismos de adaptación específicos y al mismo tiempo caracterizando tendencias que sean generales para organismos no emparentados que habitan dichos hábitats.

2.1.3 El género *Naganishia*

El género *Naganishia* (Goto, 1963) fue reformulado incorporando especies previamente designadas como *Cryptococcus* y enmendado para acomodar al clado Albidus de orden Filobasidiales, clase Tremellomycetes, división Basidiomycota (Fungi) (Fell et al., 2000; Scorzetti et al., 2002; Liu et al., 2015). Actualmente comprende 19 especies legítimas de acuerdo a MycoBank y se han descrito 3 más: *N. floricola* (Zhou et al., 2020), *N. tulchinskyi* (Bijlani et al., 2022), y *N. kalamii* (Leo et al., 2023) aún pendientes de validación. Todas las especies de este género se reproducen de manera asexual, pero difieren en su morfología, fisiología y hábitat. En efecto, y como se observa en la Tabla 2.1.1, han sido encontradas en una variedad de ambientes. Muchas especies comprenden cepas aisladas de la Antártida (subrayadas en la tabla) y prácticamente todas cuentan con registros de aislamientos provenientes de sitios con bajas temperaturas, lo cual desde un comienzo llevó a postular una adaptación de las cepas o especies a los ambientes fríos (Butinar et al. 2007, Zhou et al., 2020).

Tabla 2.1.1 Lugares de aislamiento y rangos de temperatura de crecimiento para las especies del género *Naganishia*. Se muestran subrayadas las especies con registros de aislamientos en el continente Antártico

Especie	Rango de crecimiento (°C)	Lugares de aislamiento*
<i>N. adeliensis</i>	4-25 30 variable	Deshielo y glaciares: Ártico, ríos: Patagonia Argentina, muestras clínicas, guano, materia fecal, superficies exteriores Estación Espacial Internacional (Bijlani et al., 2020)
<i>N. albida</i>	4-30 35 variable	Piel de perro y humana, agua marina: Portugal, aire: Países Bajos y Japón, sake, agua de lavado de botella de cerveza, leche, musgos, hielo: Groenlandia
<u><i>N. albidosimilis</i></u>	4-15: ND 19-30	Suelo: Antártida, plumaje de aves, drenaje ácido de minas, hielo de glaciar, cebada
<u><i>N. antarctica</i></u>	4-20 25-35 variable	Suelo y deshielo: Valles Secos de Antártida-Tierra del Fuego (Cavello et al., 2019)
<i>N. bhutanensis</i>	4-21: ND 25	Suelo: Asia y EE.UU., rumen de buey (ID de la cepa incierta)
<i>N. cerealis</i>	4-25 (con crecimiento significativo a 4 y 10)	cereales
<i>N. diffluens</i>	4-25	Muestras clínicas (piel), aire: Japón, agua contaminada: Uruguay, bosque aluvial: Austria, tripas de escarabajo
<u><i>N. friedmannii</i></u>	4-15 variable 19-21	Suelo: Antártida-Siberia-Islandia-Irán (Vishniac, 2006)
<u><i>N. globosa</i></u>	4-15: ND 19-25 , 30 variable	Espuito de tuberculosis, queso azul, aire: Países Bajos, filoplano: Portugal, suelo: Antártida
<i>N. indica</i>	No informado	Polvo arrastrado por el viento: Himalaya (Crous et al., 2019)
<i>N. liquefaciens</i>	4-15 variable 19-30	Sake, piel humana, océano profundo: Japón, hielo de glaciares: Groenlandia
<i>N. nivalis</i>	No informado	Alpes: Francia (Kachalkin, et al. 2019)
<i>N. onofrii</i>	4-30 35-37 variable	Deshielo, nieve, suelo: Alpes-EE.UU, jugo de manzana (Andreolli et al., 2021)
<i>N. qatarensis</i>	4-35	Agua marina: Catar
<i>N. randhawae</i>	4-15: ND 25-35 , 37 variable	Guano, instalaciones de ensamblaje de naves espaciales de la NASA (Leo et al., 2023)
<i>N. uzbekistanensis</i>	4-15 ND 19-30 , 35 variable	Takir: Uzbekistán, sangre de paciente VIH positivo
<i>N. vaughanmartinae</i>	4-25 30 variable	Núcleos de hielo, nieve, deshielo, suelo: Alpes-Islandia-Rusia-Antártida, hollejo de uva
<u><i>N. vishniacii</i></u>	4-19 21 variable	Suelos: Valles Secos de McMurdo Antártida - Everest (Dragone et al., 2023)

*ND: no se especifica en el rango. La información se obtuvo de The Yeasts (Boekhout et al., 2022) y se complementó con las referencias en los casos en los que no se detallaran allí dichos lugares de aislamiento.

Naganishia vishniacii (Figura 2.1.2) fue aislada por primera vez de muestras de arenisca en el supergrupo de Beacon, Antártida, y su descubrimiento en el desierto de Ross proveyó el primer indicio de microorganismos nativos habitando los suelos desglaciados de estos altos valles (Vishniac & Hempfling, 1979a). Hasta comienzos de 2023, la especie sólo había sido aislada en los Valles Secos de McMurdo, llevando a postularla como una buena candidata de

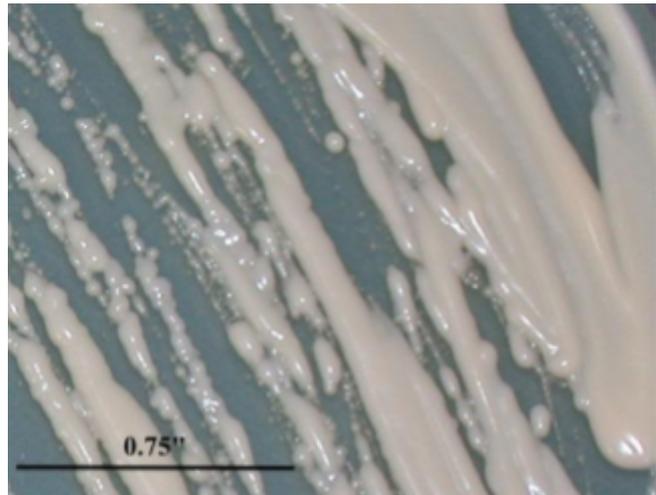


Figura 2.1.2. *N. vishniacii* creciendo en laboratorio
Foto: Katherine Earle

microbiota autóctona (Vishniac & Hempfling, 1979a) en función de la imposibilidad de encontrarla en suelos similares; su capacidad de crecer a 4 °C e incluso a -3 °C, pero no a 21 °C o más, y al hecho de que muchas especies halladas en estas regiones han sido también aisladas de otros ambientes (*ibídem*). Recientemente, sin embargo, 3 aislamientos de esta especie fueron obtenidos de sedimentos del collado Sur del Sagarmatha, en el Monte Everest (~7900 m.s.n.m.) (Dragone et al., 2023), constituyendo el primer registro de esta especie fuera del continente antártico.

Existen 17 aislamientos registrados para esta especie en MycoCosm, con una similaridad promedio de 99,6 % (de acuerdo a un análisis polifásico incluyendo regiones ITS y LSU). La cepa sobre la que se desarrolló el trabajo presentado en esta Tesis (CBS ANT03-052/CBS 10616), al igual que primera descrita (Vishniac & Hempfling, 1979a), se aisló de la zona alta del valle central de McMurdo (Valle de Wright) en la campaña de verano austral 2003-2004 (Conell et. al., 2008) a partir de muestras de suelo tomadas en esterilidad y luego cultivadas en medio YPD a 15 °C, durante 8 semanas. El aislamiento fue identificado en base a una identidad de 100% con la cepa tipo (CBS 7110) (Westerdijk Fungal Biodiversity Institute, n.d.).

2.1.4 Mecanismos de adaptación

La capacidad de los psicrófilos para sobrevivir y proliferar a bajas temperaturas implica que han superado barreras clave inherentes a los entornos permanentemente fríos generadas por una reducida actividad enzimática y una consiguiente disminución en las tasas de transcripción, traducción y división celular; desnaturalización por frío y plegamiento inapropiado de las proteínas, que alteran su funcionamiento; y formación de hielo intracelular, que genera lesiones en las diferentes organelas (D'Amico et al., 2006). Más aún, en general, estos organismos se ven expuestos a una combinación de otros factores estresantes asociados con esos ambientes, como son la baja disponibilidad de agua (que por debajo de los 0 °C se congela y se vuelve al menos parcialmente inaccesible), estrés osmótico y oxidativo y una alta radiación solar (Buzzini & Margesin, 2014). Los organismos psicrófilos, cuentan pues con una compleja gama de adaptaciones estructurales y funcionales que les permiten sobrellevar esas condiciones, muchas de las cuales han sido descritas como mecanismos comunes a una amplia variedad de formas de vida tolerantes al frío, desde vertebrados hasta microorganismos (Clark & Worland, 2008; Singh et al., 2017; Collins & Margesin 2019).

2.1.4.1 Reestructuración de la membrana

La membrana plasmática puede ser considerada el sensor inicial de las condiciones de estrés externo (Buzzini & Margesin, 2014). Más aún, se ha propuesto que la rigidez de los lípidos de la membrana a bajas temperaturas y bajo estrés hiperosmótico generan cambios conformacionales en un sensor que desencadena el resto de las respuestas específicas de adaptación (Los & Murata, 1997; Panadero et al., 2006; Russell, 2008). Esto se asocia a que un funcionamiento adecuado de la membrana está directamente ligado a su fluidez y ésta depende de su composición lipídica (Gunde-Cimerman et al., 2014; Gostinčar et al., 2019). De hecho, la temperatura mínima de crecimiento de un organismo se alcanza cuando el grado de fluidez de su membrana se vuelve demasiado bajo para permitir el transporte de nutrientes (Buzzini & Margesin, 2014). Lógicamente, este factor impone límites para el desarrollo y distribución de los organismos en el medio ambiente en función de su mayor o menor capacidad de remodelar la composición de la membrana frente a cambios en el entorno. Los principales componentes de las membranas eucariotas son los esteroides y los

fosfolípidos y son éstos los blancos de los mecanismos de reestructuración más usuales (Al-Fageeh & Smales, 2006). Los esteroides son responsables de varias características físicas de las membranas, como la regulación de su permeabilidad, su fluidez y su estabilidad. Un decremento en la proporción de los esteroides a favor de los fosfolípidos posibilita un incremento en la fluidez de membrana. Por otro lado, los distintos tipos de fosfolípidos difieren en la naturaleza de sus grupos polares y en la longitud y grado de insaturación de sus cadenas. En este sentido, la introducción de uno o más dobles enlaces, así como cambios en la posición de las insaturaciones y modificaciones en el largo de las cadenas de estos componentes alteran su configuración facilitando el mantenimiento de la estructura y permeabilidad de las membranas. De manera resumida, en las membranas de las levaduras los fosfolípidos confieren fluidez y los esteroides, rigidez (Walker, 1998). De esta manera, a bajas temperaturas se induce una respuesta coordinada de desaturasas y deshidratasas que aumentan la proporción de ácidos grasos poliinsaturados frente a los saturados y/o (de manera menos frecuente) reducen el largo de las cadenas (Al-Fageeh & Smales, 2006) permitiendo superar los efectos de rigidez generados por el frío. Los mecanismos precisos mediante los cuales las levaduras ajustan la composición de fosfolípidos de membrana en función de la temperatura no han sido del todo caracterizados, pero la regulación de desaturasas y sintetasas de ácidos grasos parecieran ser la base de dichas adaptaciones (Buzzini et al., 2014).

2.1.4.2 Producción de proteínas anticongelantes (AFPs)

Estas proteínas tienen dos actividades: por un lado, disminuyen de manera no coligativa el punto de congelamiento de las soluciones: una propiedad denominada histéresis térmica que impide el crecimiento de los cristales; y por otro, inhiben la formación de cristales grandes a expensas de los más pequeños: una actividad denominada inhibición de la recristalización del hielo. Ambos efectos se logran gracias a su capacidad de unirse irreversiblemente a la superficie de los cristales de hielo, protegiendo así a las células de lesiones por congelamiento al impedir la formación de conglomerados de cristales que suelen generarse rápidamente ante cambios bruscos de temperatura (Lauersen et al., 2011; Park et al., 2012). Las AFPs han sido reportadas en numerosos organismos, incluyendo bacterias, diatomeas, plantas, hongos, insectos, crustáceos y peces (Baksaran et al., 2021). En hongos, la actividad anticongelante fue demostrada por primera vez por Olsen y Duman

en 1993 y desde entonces se han descrito centenares de cepas albergando dicha capacidad, aunque una mínima proporción están representadas por levaduras (géneros *Glaciozyma*, *Leucosporidium* y *Rhodotorula*) (Kim et al., 2014). Se ha propuesto que los genes de AFPs fueron adquiridos mediante transferencia horizontal de genes bacterianos o de otros eucariotas que coexisten con las levaduras (Richards et al., 2011). Sin embargo, el conocimiento sobre los genes, la regulación, la función y la estructura de las AFPs en levaduras es escaso.

2.1.4.3 Compuestos crioprotectores

Los compuestos crioprotectores previenen los daños en las células generados por bajas en la temperatura promoviendo la exclusión del agua y reduciendo la formación de cristales de hielo (Tulha et al., 2010). La acumulación de trehalosa y glicerol juega un papel fundamental en la protección de las levaduras frente al congelamiento.

La acumulación de altos niveles de trehalosa en respuesta a diferentes condiciones de estrés ha demostrado ejercer un rol protector en la estabilización de la membrana (Aguilera et al., 2007) y se ha visto asociada, particularmente con la resistencia a muy bajas temperaturas (Kandror et al., 2004). De hecho, en temperaturas cercanas al congelamiento se induce la síntesis de las enzimas involucradas en su producción (trehalosa-6-fosfato sintasa -Tps1- y trehalosa-6-fosfato fosfatasa -Tps2-) en *Saccharomyces cerevisiae* (*ibidem*). El mecanismo exacto mediante el cual la trehalosa ejerce su rol protector no ha sido del todo caracterizado.

La acumulación de glicerol, de igual manera, confiere protección contra los daños causados por congelamiento y la enzima responsable de su síntesis en *S. cerevisiae* es activada ante un descenso de la temperatura (Panadero et al., 2006). El glicerol es el único soluto osmoprotector que se acumula en las levaduras frente al estrés hiperosmótico y su rol protector contra el congelamiento pareciera estar vinculado con su capacidad de evitar la contracción osmótica que resulta de los procesos de congelamiento/descongelamiento (Tulha et al., 2010).

2.1.4.4 Compuestos y pigmentos fotoprotectores

La radiación solar es indispensable para la vida en la Tierra, pero la radiación ultravioleta es extremadamente dañina. Su acción directa sobre biomoléculas (ácidos

nucleicos y proteínas) altera su función biológica. A este efecto se le suma un daño indirecto debido a que induce la generación de especies reactivas de oxígeno (ROS) que oxidan componentes vitales de la célula (Buzzini & Margesin, 2014). La radiación UV es uno de los factores más condicionantes asociado a ambientes fríos. Es por ello que, entre los mecanismos de adaptación, los organismos suelen también albergar defensas fotoprotectoras. Entre las levaduras psicrófilas, la producción de carotenoides es un rasgo común e incluso algunas producen compuestos únicos que no se encuentran en especies mesófilas, como la plectanixantina o la astaxantina (Moliné, 2010). La producción de micosporinas también es un rasgo estudiado en levaduras de ambientes fríos que les otorga protección frente a esta amenaza (Libkind et al., 2006; Brandão et al., 2011).

Los carotenoides son pigmentos naturales que se encuentran de manera transversal a los organismos fotosintéticos y esporádicamente en bacterias y hongos, incluyendo en levaduras. Se han descrito más de 700 tipos de compuestos carotenoides distintos en la naturaleza, asociados a una amplia gama de funciones. En los microorganismos, el rol principal de estos compuestos es la captura de ROS, capacidad explorada en bacterias, algas, hongos y plantas (Britton, 2008). En levaduras, tres grupos distintos de Basidiomycota y algunas especies de Ascomycota son capaces de sintetizarlos, incluyendo (pero no limitado a) los géneros *Sporobolomyces*, *Rhodospiridium*, *Rhodotorula*, *Cystofilobasidium*, *Dioszegia*, y *Phaffia* (*Xanthophyllomyces*), que comprenden especies adaptadas al frío (Buzzini & Margesin, 2014). Los carotenoides de levadura son químicamente diversos y cada grupo filogenético produce compuestos pigmentados específicos (Libkind et al., 2006; Moliné et al., 2009). La relación entre el contenido de carotenoides y la sobrevivencia de las levaduras en condiciones de alta exposición solar ha sido demostrada (Méndez-Álvarez et al., 2000).

Las micosporinas son compuestos hidrosolubles que contienen una unidad cíclica (ciclohexenona o ciclohexenimina) con, al menos, dos doble enlaces conjugados que ejercen la absorción de rayos UV, evitando el daño directo de la radiación. La síntesis de micosporinas fue descubierta en levaduras adaptadas al frío (Libkind et al., 2004) y su producción es un rasgo específico de ciertos grupos taxonómicos, aunque su distribución filogenética es más amplia que la de los carotenoides. Su acumulación bajo condiciones de radiación UV ha sido demostrada, observándose una correlación negativa entre ésta y la aparición de daños en el ADN luego de la irradiación (Moliné, 2010; Moliné et al., 2011).

2.1.4.5 Adaptaciones enzimáticas

A bajas temperaturas, la tasa de reacción de los procesos bioquímicos catalizados por enzimas se reduce de manera exponencial. Además, en estas condiciones, las interacciones hidrofóbicas se debilitan, desestabilizando el plegado de las proteínas e impulsando un proceso de desnaturalización en frío (Irwin, 2020). La flexibilidad de las proteínas es clave en la interacción enzima-sustrato y resulta fundamental para la catálisis. Las fluctuaciones térmicas inducen cambios en la organización estructural de las proteínas y sus complejos que impactan directamente sobre su funcionalidad. Resulta evidente, por lo tanto, que las enzimas psicrófilas han evolucionado mecanismos para lidiar con las tasas reducidas de reacción bioquímica inherentes a bajas temperaturas. Como características generales de las enzimas psicrófilas, se ha visto que son óptimamente activas a temperaturas por debajo de 20 °C (más bajo que las mesófilas y termófilas); son débilmente estables y generalmente se desnaturalizan a 37 °C; y mantienen su alta actividad catalítica a bajas temperaturas (Salwam & Sharma, 2020). Para mantener la actividad catalítica a bajas temperaturas, las proteínas psicrófilas generalmente son más flexibles, adaptándose fácilmente a las modificaciones estructurales requeridas para funcionar de manera eficiente. Sin embargo, se ha visto que distintas enzimas psicrófilas, provenientes de diferentes fuentes o condiciones, adoptan diferentes mecanismos, por lo que no pareciera existir una estrategia general de adaptación (Siddiqui et al., 2006). No obstante, se ha propuesto que ciertas modificaciones en las enzimas adaptadas al frío las vuelven más flexibles o menos plegadas y más eficientes en su unión a sustratos, en comparación con sus homólogas mesófilas. Entre ellas se incluyen generalmente: aminoácidos más pequeños y menos hidrofóbicos; menor cantidad de puentes salinos; mayor agrupamiento de residuos de glicina, que posibilitan una mayor movilidad local; menor número de residuos de prolina en bucles, mejorando la flexibilidad de la cadena entre estructuras secundarias; y residuos no polares en la superficie de la proteína y en contacto con el medio, aumentando las cargas negativas generales y favoreciendo la interacción con el solvente (Irwin, 2020). Todas estas modificaciones respecto de sus contrapartes activas a media o alta temperatura contribuyen a una mayor flexibilidad de las proteínas a temperaturas cercanas al congelamiento, optimizando su actividad catalítica. También se ha visto que estas enzimas tienen una entalpía más baja y una entropía de activación más negativa que sus homólogas mesófilas o termófilas, por lo

que la velocidad de reacción es menos dependiente de la temperatura (Isaksen et al., 2014; Irwin, 2020).

Al margen de las estrategias hasta aquí descritas, muchas de estas respuestas y mecanismos de adaptación son desconocidos o poco estudiados en levaduras tolerantes al frío. El conocimiento y comprensión de los mismos constituye un desafío y un tema de interés, no solo debido a su potencial biotecnológico, sino también por su importancia en los ciclos biogeoquímicos de nutrientes, su rol ecológico y su injerencia en el entendimiento de los orígenes de la vida en nuestro planeta (Hébraud & Potier, 1999; Albanese et al., 2020).

2.1.5 Objetivos específicos

- Caracterizar el genoma de *N. vishniacii* con respecto a 6 especies del orden Filobasidiales que se encuentren más cercanas pero que no estén estrechamente asociadas a las condiciones ambientales de la psicrotolerante. Identificando diferencias generales de tamaño, número de genes, y otras características genómicas mediante la obtención de un ensamblado genómico de calidad que incluya predicciones genéticas, anotaciones funcionales e inferencias de ortología.
- Establecer las relaciones de parentesco entre *N. vishniacii* y las especies de referencia que permitan su comparación y la identificación de los eventos evolutivos mediante un abordaje filogenómico.
- Identificar genes que hayan estado sometidos a selección positiva reciente entre *N. vishniacii* y sus especies hermanas o que hayan sido reportados para otros organismos adaptados a frío, y relacionar su función con con adaptaciones nicho-específica.
- Identificar modificaciones de los proteomas predichos que faciliten la actividad enzimática a bajas temperaturas mediante la identificación de sesgos en el uso de aminoácidos.

2.2 MATERIALES Y MÉTODOS

2.2.1 Secuenciación y ensamblado del genoma de *N. vishniacii*

El genoma de *N. vishniacii* ANT03–052/CBS 10616 se secuenció utilizando una biblioteca Illumina Regular Fragments (RF, tamaño de inserto: 300 pb) y una biblioteca Illumina Long Mate Pairs (LMP, 4 kpb). Para la construcción de la biblioteca RF, 100 ng de ADN se sonicaron para obtener fragmentos de aproximadamente 300 pb, utilizando un ultrasonicador Covaris LE220 (Covaris, Woburn, MA, EE. UU.). Los fragmentos del tamaño deseado se seleccionaron con perlas SPRI (*Solid Phase Reversible Immobilization magnetic beads*) (Beckman Coulter, Indianapolis, IN, EE. UU.) y se utilizó el kit de creación de bibliotecas *KAPA-Illumina* (KAPA Biosystems, Wilmington, MA, EE. UU.) de acuerdo a lo indicado por el fabricante, que posibilita la reparación de los extremos solapantes y el añadido de adenosinas para posterior ligación de adaptadores Illumina-compatibles (IDT, Inc, Coralville, IA, EE. UU.) Para la biblioteca LMP, se utilizó un fragmentador Covaris g-TUBE (Covaris) sobre 5 µg de ADN y se seleccionaron los fragmentos de tamaño adecuado mediante separación por gel de agarosa. El ADN cortado se trató para la reparación de sus extremos y la ligación de adaptadores biotinilados conteniendo sitios LoxP. Los fragmentos ligados con adaptadores fueron circularizados mediante reacción de escisión con recombinasa *Cre* (NEB Inc., Massachusetts, EE.UU.). Los templados de ADN circularizados se fragmentaron aleatoriamente utilizando el fragmentador Covaris LE220 (Covaris), se trataron para la reparación de sus extremos y prolongación con adenosina utilizando el kit de creación de bibliotecas *KAPA-Illumina* (KAPA Biosystems) y se inmovilizaron en perlas de estreptavidina (Invitrogen, Carlsbad, CA, EE. UU.). Por último, se ligaron adaptadores Illumina-compatibles (IDT, Inc) a los fragmentos y se efectuaron 8 ciclos de PCR para enriquecer la biblioteca final (KAPA Biosystems).

Todas las bibliotecas se cuantificaron utilizando el kit de qPCR *KAPA Biosystem next generation sequencing library* (KAPA Biosystems) en un ciclador Roche LightCycler 480 para PCR en tiempo real, multiplexadas con otras bibliotecas. La preparación final de las bibliotecas para su secuenciación en la plataforma Illumina HiSeq, se realizó con el kit *TruSeq paired-end cluster kit v3* (Illumina) y un sistema de generación de clústeres Illumina *cBot*. La secuenciación de las celdas de flujo se realizó mediante un secuenciador Illumina HiSeq2000

utilizando el kit de secuenciación TruSeq SBS v3, siguiendo un esquema de corridas indexadas de 2 × 150 pb (2 × 100 pb para LMP).

Se obtuvieron 2 ensamblados utilizando 2 estrategias. El primer ensamblado fue realizado en el JGI con AllPathsLG vR41043 (Gnerre et al., 2011) y depositado en el portal de hongos de JGI MycoCosm (mycocosm.jgi.doe.gov/Nagvi1) (Grigoriev et al., 2014). Posteriormente (este trabajo) se obtuvo un segundo ensamblado. Para ello, a calidad de los *reads* se analizó mediante el *software* FASTQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>) y se limpiaron las librerías de posibles contaminantes utilizando la base de datos UniVec de NCBI (disponible en <ftp://ftp.ncbi.nlm.nih.gov/pub/UniVec/>), a fin de descartar segmentos insertados en las secuencias nucleicas que pudiesen tener su origen en vectores de clonado. Los *reads* genómicos se ensamblaron utilizando el *software* SPAdes v3.6.0 (Bankevich al., 2012) incluyendo la remoción de adaptadores, filtrado de secuencias de baja calidad y corrección de errores. Ambas versiones de ensamblados fueron fusionadas mediante el *pipeline* NGAP4 (Nizovoy et al., 2018) que localiza regiones con nucleótidos indefinidos del ensamblaje primario (identificados con la letra N repetidas veces) y recupera las regiones flanqueantes a esas zonas. Este par de fragmentos es luego utilizado como señuelo sobre el ensamblado de SPAdes escaneado con BLAT (Kent, 2002) para verificar si en dicho caso la región delimitada por los mismos, se encuentra definida (sin Ns). En tal caso, la región es recuperada e integrada al ensamblado original, quedando de esta manera resuelta la isla de nucleótidos indefinidos.

El genoma obtenido como resultado de este abordaje (*Whole Genome Shotgun*) fue ingresado en los repositorios DDBJ/ENA/GenBank bajo la identificación de acceso: JABEVT000000000 (versión descrita en este trabajo), único ejemplar de la especie secuenciado

2.2.1.1 Predicciones de genes y anotaciones

Se realizaron predicciones de genes *ab initio* con GeneMark-ES v2.3e (Ter-Hovhannisyan et al., 2008) auto-entrenado con los *scaffolds* genómicos (parámetros --min_contig --max_nnn 1000). El proceso de anotación automática se realizó mediante el *pipeline* Funannotate v1.5.0 (Love et al., 2018), a partir de los *scaffolds* y las predicciones obtenidas. La anotación de dominios proteicos fue realizada con Pfam v32.0 en los casos que

fue posible (Bateman et al., 2004). La clasificación en grupos de ortología de copia única se realizó utilizando la base de datos BUSCO (*Benchmarking Universal Single-Copy Orthologs*) v1.0 para Basidiomycota (Waterhouse et al., 2018), la agrupación de las predicciones dentro de clústeres de genes ortólogos de acuerdo a la base de datos eggNOG v4.5.1 (Huerta-Cepas et al., 2017) y la asignación de categorías funcionales COG (Tatusov et al., 2000). Para las anotaciones finales, se utilizó la identidad Nv_00#### para cada una de las 6183 predicciones génicas (con #### entre 0001 y 6183).

El proceso de predicción y anotación se realizó de igual forma para las especies filogenéticamente más cercanas disponibles pertenecientes al orden Filobasidiales: *Naganishia albida* Y-1402 y NT2002 (acc. DDBJ/EMBL/GenBank: PRJNA291988 y PRJNA298369; Vajpeyi & Chandran, 2016; Yong et al., 2016), *Naganishia adeliensis* IF1SW-F1 (PRJNA623412; Bijlani et al., 2020), *Filobasidium wieringae* UCDFST 05–544 (PRJDB3683), *Solicoccozyma phenolicus* JCM11743 (PRJDB3680) y *Solicoccozyma terricola* JCM24523 (PRJNA327103). La elección de las mismas se basó en que al momento del análisis representaban los genomas disponibles filogenéticamente más cercanos a *N. vishniacii*, priorizando aquellos que ya estuvieran publicados (para las especies sin publicaciones disponibles, los ensamblajes fueron incluidos previa autorización de sus autores). Dentro del género *Naganishia*, en particular, se seleccionaron 2 cepas de *N. albida* (de las 4 disponibles) en base a que representaban 2 sustratos de aislamientos muy distintos (pulmón humano y suelo) y 1 de *N. adeliensis*, la única para la cual se pudo verificar una correcta afiliación (reconstrucciones filogenéticas realizadas *ad hoc* en el momento de la selección de los genomas más representativos -datos no mostrados- indicaron que la otra cepa disponible tenía mayor cercanía a *N. albida* que a *N. adeliensis*). Asimismo, se realizaron predicciones y anotaciones sobre el genoma de *Cryptococcus neoformans* JEC21 (acc.: AE017341.1; Stajich et al., 2012) a fin de comparar los resultados con esta especie de referencia para la división Basidiomycota.

2.2.2 Reconstrucción filogenómica

La reconstrucción filogenómica se realizó en base a 2130 genes de copia única compartidos entre todas especies del orden Filobasidiales y *C. neoformans* JEC21. La asociación de los genes a los correspondientes grupos de ortología se determinó mediante

eggNOG-Mapper v 4.5.1 (Huerta-Cepas et al., 2017) (ver “Grupos de ortología y asignaciones funcionales”). Las predicciones proteicas seleccionadas se alinearon mediante MUSCLE con parámetros por defecto (Edgar, 2004) y los alineamientos obtenidos se concatenaron resultando en un largo de 1.093.927 columnas. El árbol filogenómico se obtuvo empleando RAxML-HPC2 (Stamatakis et al., 2014), utilizando el modelo de sustituciones aminoacídicas PROTGAMMALG4M. El modelo evolutivo de sustitución se seleccionó empleando ModelTest-NG con parámetros por defecto (Darriba et al., 2020). La fiabilidad de las ramas se determinó realizando 1000 réplicas *bootstrap*. Las corridas de RAxML y ModelTest-NG se realizaron en la plataforma científica CIPRES (Miller et al., 2010).

Por otro lado, se estimó el nivel de divergencia genómica entre pares de especies estrechamente relacionadas utilizando la métrica Kr a partir de la función *genomediff* implementada en GenomeTools (Gremme et al., 2013). Kr es un estimador independiente de alineamiento del número de sustituciones de nucleótidos por sitio, basado en la longitud promedio de la subsecuencia individual más corta, lo que resulta en divergencias corregidas de Jukes-Cantor entre secuencias genómicas (Parra et al., 2023).

2.2.3 Grupos de ortología y asignaciones funcionales

Para la búsqueda de genes específicos que puedan estar implicados en la tolerancia al frío, los genes de *N. vishniacii* fueron clasificados como compartidos o únicos. Para cada predicción se realizó la asignación de ortología a partir de clústeres pre calculados mediante eggNOG v4.5.1 (Huerta-Cepas et al., 2017) empleando la base de datos de hongos. Las secuencias de *N. vishniacii* sin correlato en la base de datos de eggNOG se clasificaron en genes únicos o compartidos mediante su comparación contra los 7 genomas del orden Filobasidiales estudiados empleando tblastn (eval < 1e-05, identidad ≥ 50 %). Las predicciones que no se agruparon con ninguna de las del resto de las especies se consideraron exclusivas de *N. vishniacii* y sus posibles funciones se dedujeron por homología empleando bases de datos no redundantes del NCBI, en los casos en los que fuera posible.

2.2.4 Análisis del cociente dN/dS

El cociente de sustituciones no sinónimas (dN) y sinónimas (dS) se computó mediante KaKs Calculator V2.0 (Wang et al., 2010) para todos los genes de copia única comunes entre *N. vishniacii* y ambas cepas de *N. albida*. Para ello, las predicciones proteicas comunes (en base a lo determinado por eggNOG-Mapper v4.5.1 - “Grupos de ortología y asignaciones funcionales”) fueron alineadas de a pares con MUSCLE empleando parámetros por defecto (Edgar, 2004) y convertidas a alineamientos de codones libres de *gaps* mediante PAL2NAL (Suyama et al., 2006). Las secuencias alineadas resultantes fueron particionadas en fragmentos de 34 codones de largo y analizadas mediante una estrategia de tipo *sliding windows* (con saltos de 8 codones) empleando el modelo YN (Yang & Nielsen, 2000). Este análisis se realizó también sobre los alineamientos no fragmentados (secuencias completas). Todas las predicciones con regiones positivas (cociente dN/dS > 1, eval < 1e-02) fueron subsecuentemente comparadas entre *N. vishniacii* y *F. wieringae*. Los genes con regiones positivas para ambos pares de especies fueron evaluados gráficamente con JCoDa (Steinway et al., 2010).

Se consideraron portadoras de señales de selección positiva únicamente a las secuencias que compartieran regiones con cociente mayor a la unidad para los alineamientos de *N. vishniacii* vs. una o ambas cepas de *N. albida* y para los de *N. vishniacii* vs. *F. wieringae*, siempre y cuando dicha región no presentara un cociente mayor a 1 entre las cepas de *N. albida* vs. *F. wieringae*

2.2.5 Búsqueda de genes involucrados en la respuesta a frío

Los genes implicados en vías normalmente relacionadas con la tolerancia al frío se buscaron mediante blastp sobre los proteomas de todas las especies de orden Filobasidiales, empleando secuencias descritas por otros autores. En la Tabla 2.2.1 se detallan las secuencias utilizadas para cada una de las vías/genes analizados.

Además de la búsqueda de desaturasas, se observaron diferencias en el número de copias de enzimas implicadas en el metabolismo lipídico para los grupos de ortología asociados (de acuerdo a lo determinada por eggNOG).

Tabla 2.2.1. Secuencias *query* utilizadas para las búsquedas de genes de interés involucrados en la respuesta a frío

Vía/Gen de interés		Secuencia <i>query</i>	Referencia / base de datos	Método	
PROTEÍNAS ANTICONGELANTES AFPs		166 secuencias de origen fúngico 3832 secuencias de origen bacteriano	obtenidas de NCBI empleando "antifreeze" como palabra clave (Material Suplementario: *_antifreeze.faa)	blastp	
FOTOPROTECCIÓN	síntesis de carotenoides	idi isopentenil difosfato isomerasa	BAA33979.1	NCBI	blastp
		FPS farnesil pirofosfatasa sintetasa	AHW57996.1	NCBI	
		crtE farnesil transferasa	AAY33922.1	NCBI	
		crtI fitoeno deshidrogenasa	G01847_P	Bellora et al., 2016	
		crtYB fitoeno beta-caroteno sintasa	AAO47570.1 G05006_P	NCBI Bellora et al., 2016	
		crtR NADP-citocromo P450 reductasa	ACI43098.1	NCBI	
		beta-caroteno 15,15-monooxigenasa	G04735_P	Bellora et al., 2016	
	síntesis de micosporinas	EEVS-like 2-epi-5-epi-valiolo na sintasa	G05217_P	Bellora et al., 2016	
		O-met catecol o-metiltransferasa	G05218_P		
		ATP-grasp carbamoilposfato sintasa	G05219_P		
TREHALOSAS	síntesis	TPS1 trehalosa-6-fosfat o-sintetasa	XP_025548824.1 PF00982	NCBI pfam	blastp y corroboración de dominios
		TPS2 trehalosa 6-fosfato fosfatasa	NP_594975.1 PF02358		
	degradación	NTH1 trehalasa neutra	NP_010284.1 PF01204		
		NTH2 trehalasa neutra	NP_009555.1		
		ATH1 trehalasa ácida	NP_015351.1		
DESATURASAS de ÁCIDOS GRASOS		Δ4	XP_018259658.1; RXK41666.1; XP_019000774.1; XP_566973.1	NCBI pfam	blastp y corroboración de dominios
Δ8	XP_014181797.1; EKD01784.1; XP_019014123.1; OCF60046.1; ORY23563.1; XP_569193.1				
Δ9	OCF45831.1; XP_019035553.1; OXB34927.1; NP_011460.3; XP_567496.1				
Δ12	XP_019004193.1; XP_018262343.1; XP_0245126				

2.2.5.1 Confirmación experimental de la producción de carotenoides y micosporinas en *N. vishniacii*

La levadura se cultivó en placas de petri con medio MMS modificado (Moliné et al., 2011) bajo tres tratamientos: (a) **radiación**, (b) **estrés oxidativo** y (c) **baja temperatura**. Las placas fueron inoculadas con 100 µL de una suspensión de 10^6 células, esparcidas con una espátula de Drigalski hasta secado y cultivadas: (a) a 19 °C bajo radiación fotosintéticamente activa en una cámara para prueba ambiental con luz provista por 10 tubos fluorescentes de luz blanca (Sanyo, 40 W); (b) a 19 °C y agregado de 1 mL de H₂O₂ 0,01 mM y (c) a 0,5 °C. La extracción de carotenoides se realizó de acuerdo a lo descrito por Sedmak (Sedmak et al., 1990), con modificaciones propuestas por Moliné y colaboradores (2009). La absorbancia de las muestras suspendidas en n-hexano se midió en un espectrofotómetro UV-visible (Shimadzu UV1800) a 473 nm. La extracción de micosporinas se realizó de acuerdo a lo descrito previamente (Moliné et al., 2011) y su absorbancia se midió a 310 nm.

2.2.6 Análisis del uso de aminoácidos

Para evaluar el uso de aminoácidos, se realizó un análisis de composición aminoacídica comparativa. Para ello se escaneó la presencia del conjunto de 248 genes eucarióticos centrales (*Core of Eukaryotic Genes*: CEGs; Parra et al., 2009) en todas las predicciones de especies de Filobasidiales y en *C. neoformans*. Este abordaje se realizó empleando HMMER v 3.1b1 (Eddy, n.d.) aplicando los modelos específicos con un umbral definido de acuerdo a Parra y colaboradores (2009). Los conjuntos de CEGs comunes entre cada una de las especies de Filobasidiales y *C. neoformans* (entre 239 y 243) se alinearon de a pares utilizando MUSCLE (Edgar, 2004) con parámetros por defecto. Las secuencias resultantes para cada especie fueron concatenadas y se determinó su composición aminoacídica mediante Composition Profiler con parámetros predeterminados (Vacic et al., 2007), refiriendo estos valores a los de *C. neoformans*. La misma metodología se aplicó sobre los concatenados de alineamientos aminoacídicos resultantes de los 2130 genes comunes a todas las especies en función de lo determinado por eggNOG.

2.3 RESULTADOS

2.3.1 Secuenciación del genoma, ensamblado, predicciones de genes y anotaciones

El genoma de 19,7 Mbp de *N. vishniacii* (ANT03–052; CBS 10616) se ensambló en 38 *scaffolds*, con longitudes desde 2,1 kb a 1,8 Mb, con un *coverage* de 146,7x y un número de nucleótidos indefinidos de 25.706 Ns (0,13 %). Su grado de completitud, en base a la representatividad de los genes BUSCO de Basidiomycota, fue de 90,9 % y de 99,2 % en relación al set de genes CEG. Las predicciones génicas resultaron en 6183 genes codificantes putativos, que representan el 53,4 % de la secuencia del genoma. A partir de estos modelos, 4956 genes putativos (80,3 %) fueron asignados a 23 categorías funcionales diferentes de eggNOG y clasificados en 4609 grupos de ortólogos. De estas predicciones, 2613 (42,3 %) pudieron ser anotadas con una función putativa.

Los tamaños genómicos, números de predicciones génicas y el contenido de G + C de todas las especies del orden Filobasidiales ensambladas para este trabajo se presentan en la Figura 2.3.1.

2.3.2 Reconstrucción filogenómica

La reconstrucción del árbol filogenómico confirmó que *N. vishniacii* es la especie hermana de *N. albida*, y ambas especies se relacionan con *N. adeliensis*, confirmando la monofilia del género para las especies utilizadas. Los tamaños de genomas para todas las especies de este género fueron similares (19,4 - 24,5 Mpb) aunque el número de predicciones de genes para *N. vishniacii* fue entre un 4 y un 17 % menor que el de las especies hermanas. En la Figura 2.3.1 se muestra el árbol filogenómico para las 8 cepas estudiadas, construido en base a grupos de ortología de copia única comunes a todas.

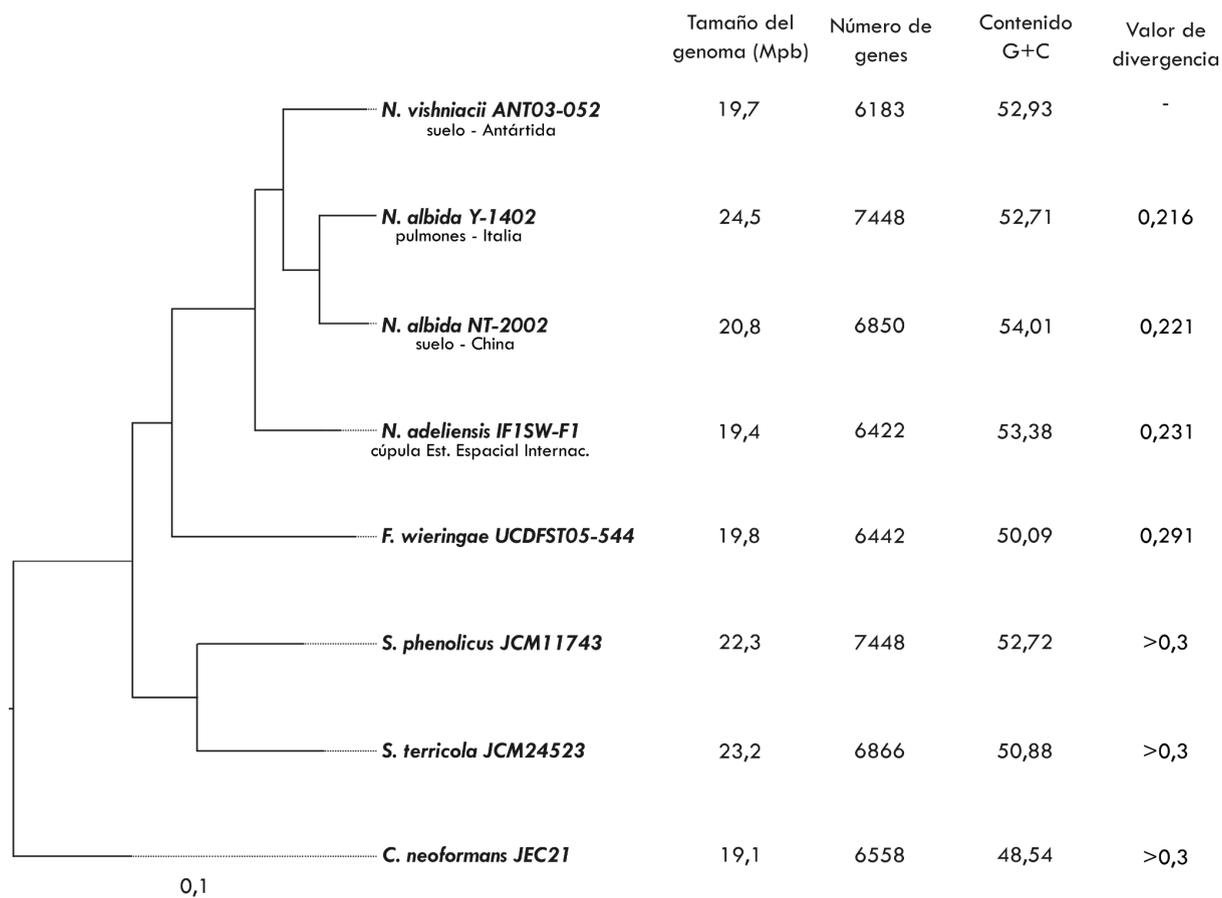


Figura 2.3.1. Árbol filogenómico de los organismos analizados computado a partir de las secuencias proteicas de 2130 grupos de ortología de copia única compartidos entre todos. Construido empleando RAxML v. 8.2.12 con el modelo de sustitución aminoacídica PROTGAMMALG4M. Nodos sustentados por un *bootstrap* de 100 %.

2.3.3 Grupos de ortología y asignaciones funcionales

El análisis de inferencia de ortología de los 4609 grupos diferentes en *N. vishniacii* reveló que 4593 (~99,7 %) eran compartidos entre esta especie y una o más de las especies relacionadas y 2926 (~63,5 %) resultaron comunes a las 7 cepas de Filobasidiales. Mientras que 16 grupos se encontraron únicamente en el genoma de *N. vishniacii* y cerca de un 18 % de los genes predichos en esta especie ($N = 1073$) no fue asignado a ningún grupo de ortología de la base de datos de eggNOG.

El análisis de las proteínas derivadas de este conjunto de genes realizado mediante tblastn contra el resto de los genomas de las especies estudiadas, mostró que en 753 predicciones (70 %) existían coincidencias con alguna o varias especies. Para el resto de las

320 predicciones sin coincidencias significativas y sin asignación de ortología, la búsqueda en la base de datos no redundante del NCBI resultó en 297 (casi el 93 %) sin correlato significativo. Las 23 restantes (~7 %) pudieron vincularse a proteínas hipotéticas o coincidieron parcialmente con una secuencia depositada y seis de ellas (~2 %) pudieron ser vinculadas a funciones putativas. La lista completa se presenta como Material Suplementario (Tabla S2.1) y la información se resume en la Figura 2.3.2.

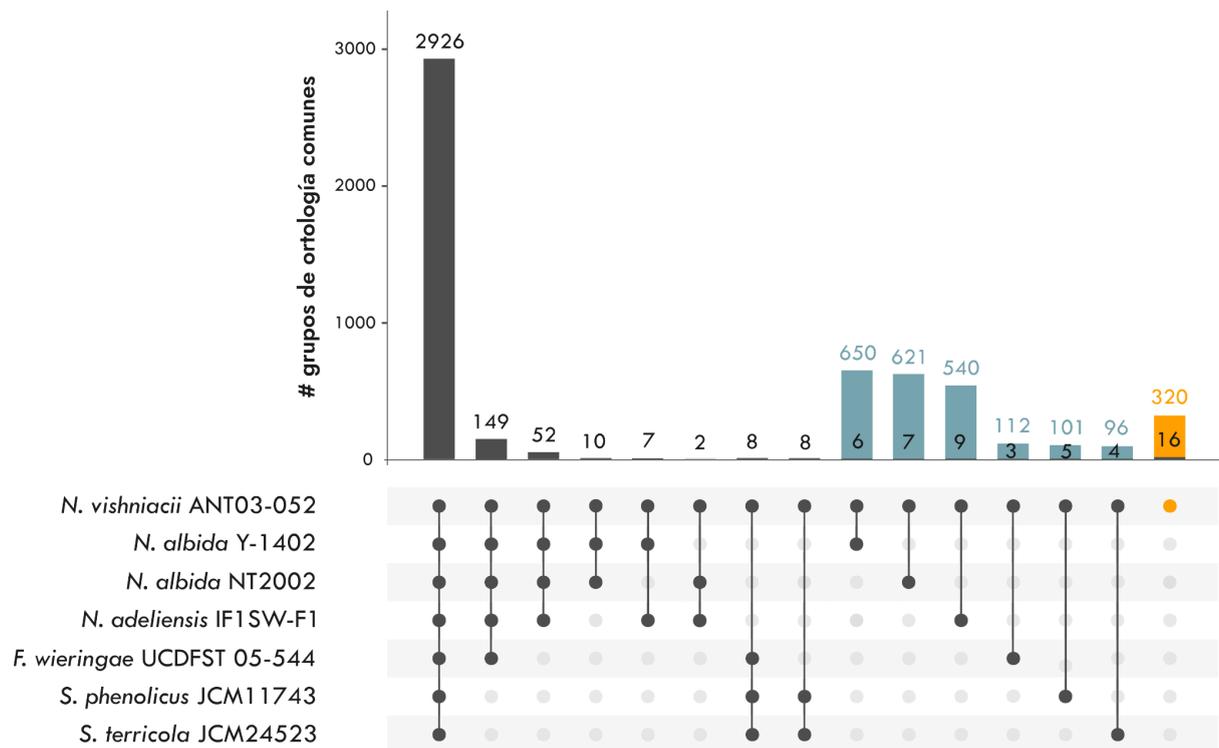


Figura 2.3.2. Análisis de ortología para las 7 cepas del orden Filobasidiales. Las barras negras representan el número de genes presentes en grupos de ortología compartidos. Los puntos oscuros rellenos denotan presencia y los vacíos (gris claro) indican ausencia de grupos de ortología comunes. Las barras celestes sobre cada intersección pareada representan predicciones de *N. vishniacii* que no fueron asignadas a grupos de ortología pero que sí arrojaron coincidencias al escanearse en los otros genomas mediante tblastn. Las barras amarillas representan predicciones de *N. vishniacii* que no fueron asignados a grupos de eggNOG y tampoco tuvieron correlato en el resto de los Filobasidiales. Se muestran únicamente las intersecciones más relevantes relativas a *N. vishniacii*. Gráfico visualizado con UpSetR (Conway et al., 2017).

Entre los 16 grupos de ortología que se encontraron únicamente en el genoma de *N. vishniacii*, las categorías funcionales de eggNOG más representadas fueron: Función desconocida (43 %), Mecanismos de transducción de señales (13 %), Biosíntesis, transporte y catabolismo de metabolitos secundarios (13 %) y Transcripción (13 %). También se encontraron grupos asociados a Transporte y metabolismo de carbohidratos y a

Procesamiento y modificación del ARN. La descripción funcional resultó, entre otras, en: transferasas putativas (Nv_005274), transportadores de solutos (Nv_002329) y chaperonas (Nv_005012). La lista completa se encuentra disponible como Material Suplementario (Tabla S2.2).

Grupo de ortología	ID N. <i>Vishniacii</i>	Categoría funcional COG	Desc. funcional	#Nv	#NaY	#NaN	#Nad	#rF
ENOG410PGGB	Nv_000525	Procesamiento & modificaciones del ARN	Endonucleasa <i>Dicer-like</i>	3	2	2	1	2
	Nv_001541							
	Nv_003488							
ENOG410PGKM	Nv_002463	Chaperonas, modificaciones post-traduccionales, recambio de proteínas	Aminopeptidasa	4	2	1	1	1-2
	Nv_003321							
	Nv_003922							
ENOG410PMCP	Nv_003280	Producción & conversión de energía	Oxidorreductasa	5	4	3	3	2-4
	Nv_003328							
	Nv_004462							
	Nv_005845							
ENOG410PII4	Nv_000220	Transporte & metabolismo de carbohidratos	α -1-6 manosiltransferasa	4	2	1	2	1-3
	Nv_003422							
	Nv_005291							
	Nv_005418							
ENOG410PNF7	Nv_003563	Función desconocida	Glicosil hidrolasa flia. 71	3	1	1	3	1-2
	Nv_004049							
	Nv_004279							
ENOG410PNRT	Nv_002066	Función desconocida	Flia. SUR7/ Pall	4	3	3	3	1-2
	Nv_002068							
	Nv_002408							
	Nv_004073							
ENOG410PWG1	Nv_001317	Función desconocida	No disponible	3	2	2	5	2
	Nv_001333							
	Nv_003778							
ENOG410PJ40	Nv_001387	Chaperonas, modificaciones post-traduccionales, recambio de proteínas	Amidohidrolasa	3	2	1	2	1-4
	Nv_001586							
	Nv_003277							
ENOG410PFJG	Nv_000658	Transporte & metabolismo de carbohidratos	Transportador MFS	3	2	1	2	1-3
	Nv_003680							
	Nv_003681							
ENOG410PH2P	Nv_000862	Transporte & metabolismo de carbohidratos	Quinato permeasa	3	2	2	2	2-3
	Nv_001298							
	Nv_005799							
ENOG410PFKH	Nv_000275	Transporte & metabolismo de carbohidratos	Quinato permeasa	14	11	10	11	9-16
	Nv_000284							
	Nv_002412							
ENOG410PF97	Nv_005830	Transporte & metabolismo de carbohidratos	Maltosa permeasa, proteína relacionada con el transporte de trehalosa	5	3	1	4	5-6
	Nv_001371							
	Nv_002166							
	Nv_002181							
	Nv_002182							
Nv_005515								

Figura 2.3.3. Grupos de ortología en los que el número de copias de *N. vishniacii* resultó mayor que el de *N. albida*. Se muestran: números de copias en *N. vishniacii* (#Nv), *N. albida* Y-1402 (#NaY), *N. albida* NT2002 (#NaN), *N. adeliensis* (#Nad) y el rango observado en el resto de los Filobasidiales (#rF). Resaltado en verde: grupos en los que el resto de los Filobasidiales mostró una redundancia mayor o igual a la de *N. vishniacii*.

Para los grupos de ortología compartidos, se compararon los números de copias en cada una de las especies. Este análisis reveló que 2366 grupos (~81 %) se componían de 1 copia por especie, 19 (~1 %), de 2 copias y 1 de 3 copias, mientras que 540 grupos (18 %) mostraron diferencias en el número de copias entre especies.

Al comparar los números de copias entre *N. vishniacii* y *N. albida*, se observó que para 12 grupos de ortología la especie antártica superó a su congénere (Figura 2.3.3). En particular, genes implicados en el Procesamiento y modificación del ARN, Transporte y metabolismo de aminoácidos y solutos y Producción de energía, fueron más abundantes en *N. vishniacii* que en ambas cepas de *N. albida*. Para los grupos ENOG410PJ40, ENOG410PFJG, ENOG410PH2P, ENOG410PFKH y ENOG410PF97 (resaltados en verde), si bien *N. vishniacii* mostró un número de copias mayor que *N. albida*, su redundancia fue igual o incluso menor que la del resto de las especies de Filobasidiales.

2.3.4 Análisis del cociente dN/dS

La distribución de los valores medios del cociente dN/dS para los 3942 / 3991 genes de copia única comunes entre *N. vishniacii* y ambas cepas de *N. albida* (Y-1402 / NT2002 respectivamente) se presentan en la Figura 2.3.4(A). El análisis del cociente dN/dS osciló entre 0 y 0,526 / 0,590, aunque el 98 % de los genes arrojaron un valor de dN/dS < 0,2 (81 / 79 % con dN/dS < 0,1 respectivamente). Ningún gen mostró un cociente medio mayor a 1 (análisis de secuencia completa), sin embargo, el análisis mediante *sliding windows* permitió detectar 65 / 77 anotaciones (1,9 / 1,6 % respectivamente) con dN/dS > 1 en alguna región y 49 / 23 de ellas que mostraron además un valor > 1 en las comparaciones pareadas de *N. vishniacii* vs. *F. wieringae*. Al comparar una a una de estas anotaciones entre las 3 especies utilizando JCoDA, pudieron evidenciarse 31 / 13 grupos de ortología (0,79 / 0,30 %) con posibles señales de selección positiva: dN/dS > 1 en regiones compartidas entre *N. vishniacii* vs. *N. albida* y *N. vishniacii* vs. *F. wieringae*, pero no entre *F. wieringae* vs. *N. albida*. Los números se muestran entre paréntesis en la Figura 2.3.4(A). En 9 de ellos se evidenció selección positiva para ambas cepas de *N. albida*.

La clasificación en categorías funcionales COG para los grupos de ortología con señales de selección positiva se presenta en la Figura 2.3.4(B). La mayoría de los genes predichos (42,9 %) identificados mediante este abordaje fueron de Función desconocida, seguidos por genes implicados en la Modificación postraducciona, recambio de proteínas y chaperonas

(20 %), el Procesamiento y modificación del ARN (8,6 %) y la Replicación, recombinación y reparación (8,6 %).

Un total de 10 de estos 35 genes fueron además asignados a genes de la base de datos BUSCO de Basidiomycota (nomenclatura EOG~ en la Figura 2.3.4(B)).

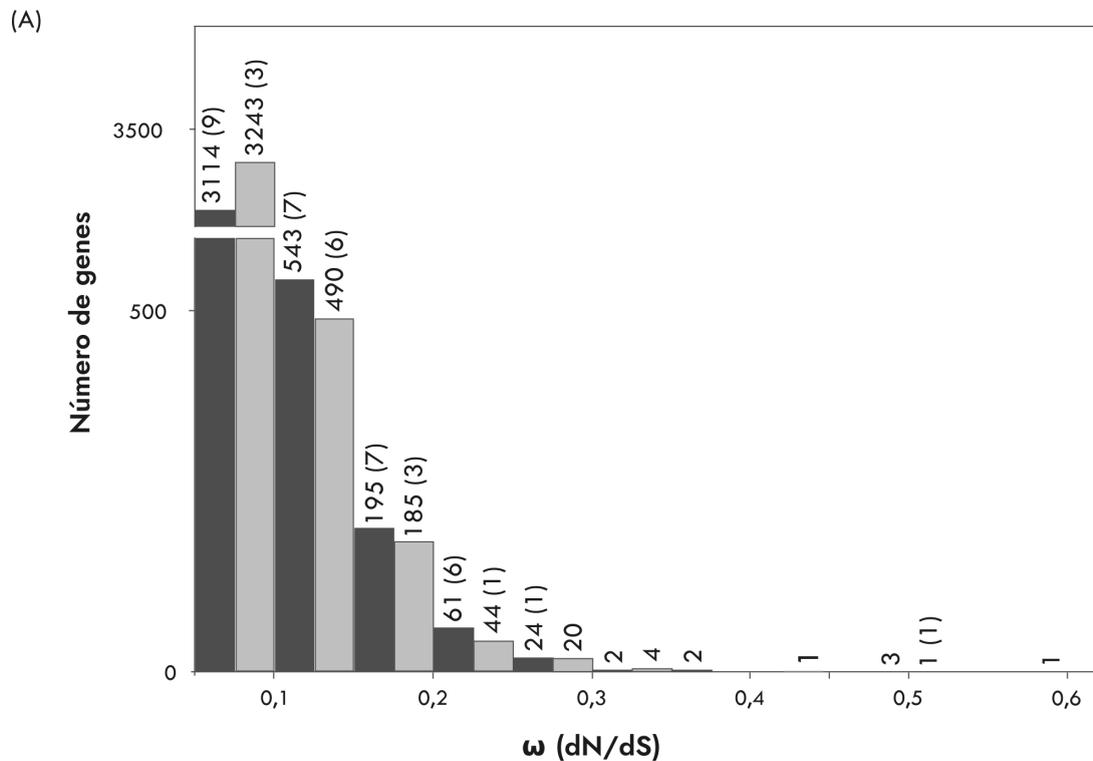


Figura 2.3.4. Análisis del cociente dN/dS para los grupos de ortología de copia única comunes entre *N. vishniacii* y *N. albida*. (A) Distribución de valores medios de dN/dS para los 3942 / 3991 grupos compartidos entre *N. vishniacii* y las cepas Y-1402 (barras oscuras) / NT2002 (barras claras) de *N. albida*. Entre paréntesis se muestra el número de genes con dN/dS > 1 en alguna región, considerados bajo selección positiva en base a la estrategia de *sliding windows* (N = 35).

(B)

2,9	ENOG410PN9W	EOG92R0GHG	Estructura y dinámica de la cromatina
8,6	ENOG410PM4Y	EOG92R02F8	Procesamiento & modificaciones del ARN
	ENOG410PMXS		
	ENOG410PS3A		
2,9	ENOG410PHFH		Transcripción
8,6	ENOG410PJ61		Replicación, recombinación & reparación
	ENOG410PGIW	EOG92R003Y	
	ENOG410PGTD	EOG092R01C6	
2,9	ENOG410PN8T		Transducción de señales
2,9	ENOG410PFNE		Tráfico intracelular, secreción & trans. vesicular
20,0	ENOG410PHMD	EOG092R01O5	Chaperonas, modificaciones post-traduccionales, recambio de proteínas
	ENOG410PHS9		
	ENOG410PR2H		
	ENOG410PWCK		
	ENOG410PNUA		
	ENOG410PGYH		
	ENOG410PNMI		
2,9	ENOG410PGBT	EOG092R00O2	Producción & conversión de energía
2,9	ENOG410PPG8	EOG092R0LJS	Transporte & metabolismo de nucleótidos
2,9	ENOG410PH9R	EOG092R03OR	Transporte & metabolismo de iones inorgánicos
42,9	ENOG410PXCK		Función desconocida
	ENOG410PJ2P		
	ENOG410PM0K	EOG092R063F	
	ENOG410PQZN		
	ENOG410PSCK		
	ENOG410PSFZ		
	ENOG410PT9V		
	ENOG410PQEC	EOG092R04D6	
	ENOG410PS2W		
	ENOG410PS3A		
	ENOG410Q0Y1		
	ENOG410PGVK		
	ENOG410PRE7		
	ENOG410QC40		
ENOG410Q13C			

N = 35

Figura 2.3.4 (cont). Análisis del cociente dN/dS para los grupos de ortología de copia única comunes entre *N. vishniacii* y *N. albida*. (B) Categorías funcionales COG para los genes bajo selección positiva. Se muestran los grupos que arrojaron señales positivas entre *N. vishniacii* y ambas cepas de *N. albida* (celeste), solo cepa Y-1402 (oscuro) y solo cepa NT2002 (claro). Los genes que pudieron asignarse a la base de datos BUSCO se muestran con su nomenclatura correspondiente (EOG~).

2.3.5 Búsqueda de genes involucrados en la respuesta a frío

El estudio de genes específicos relacionados con la adaptación al frío demostró que *N. vishniacii* posee genes implicados en la fotoprotección, respuestas antioxidantes, síntesis de trehalosa y genes de interés vinculados con el metabolismo y transporte de lípidos, pero carece de proteínas anticongelantes.

La búsqueda por homología de **proteínas anticongelantes** dio resultados positivos únicamente en el genoma de *N. adeliensis*. Esta especie posee un gen con una alta homología ($< 1e-11$) a una glicoproteína anticongelante caracterizada en *Rhodotorula toruloides*. Para el resto de las especies de Filobasidiales, incluyendo *N. vishniacii*, no se encontraron evidencias de homología con ninguna de las 3998 secuencias utilizadas como *query*. En cuanto a los genes vinculados a **fotoprotección** y **respuestas antioxidantes**, se encontró un grupo de tres genes previamente implicados en la síntesis de micosporina-glutaminol-glucosido (MGG) (Bellora et al., 2016), consistente en un homólogo de DHQS (2-epi-5-epi-valiolona sintasa, *EEVS-like*; Nv_000059), una O-metiltransferasa (O-MT; Nv_000060) y una ATP-grasp (Nv_000058). Estos genes pudieron localizarse en todos los genomas del género *Naganishia*, así como en *F. wieringae* y *S. terricola*, pero no en *S. phenolicus*. La biosíntesis de MGG se confirmó experimentalmente en todas las especies positivas: *F. wieringae* (Muñoz, 2010; Brandão et al., 2011), *S. terricola* (Libkind et al., 2009) y *Naganishia* spp. (Guzmán, 2016 y en las 3 condiciones ensayadas en este trabajo, descritas en la metodología “Confirmación experimental de la producción de carotenoides y micosporinas en *N. vishniacii*”). Además se constató la presencia del conjunto completo de genes implicados en la síntesis de carotenoides (Nv_003146: isopentenil difosfato isomerasa; Nv_004897: farnesiltransferasa; Nv_000648: farnesil pirofosfatasa sintetasa; Nv_002341: fitoeno deshidrogenasa; Nv_005201: fitoeno beta-caroteno sintasa; Nv_004862: NADP-citocromo P450 reductasa; Nv_000692: beta-caroteno 15,15-monooxigenasa), así como en las dos cepas de *N. albida* y en las de *N. adeliensis* y *F. wieringae*. Los genes correspondientes a las otras especies estudiadas no fueron encontrados (Material Suplementario, Tabla S2.3). También se identificaron 4 genes implicados en el metabolismo de **trehalosa** en las 7 cepas analizadas. Las enzimas para la síntesis (TPS1: Nv_002644; Nv_002645 y TPS2: Nv_002256) y degradación (NTH1:

Nv_002469) de este disacárido de glucosas fueron identificadas en todos los genomas. (Material Suplementario, Tabla S2.4).

Con respecto a los genes vinculados con el metabolismo de lípidos, la búsqueda de dominios proteicos específicos posibilitó localizar genes de **desaturasas** $\Delta 4$ (Nv_004586), $\Delta 8$ (Nv_004525) y $\Delta 12$ (Nv_000815) en todos los genomas de Filobasidiales. La presencia de desaturasas $\Delta 9$ (Nv_003331) fue constatada en todos los genomas analizados excepto en *S. terricola*. El análisis de inferencia de ortología para las enzimas implicadas en el **metabolismo lipídico** arrojó entre 119 y 136 grupos compuestos por proteínas hipotéticas no identificadas en las 7 cepas de Filobasidiales, con 78 grupos comunes a todas ellas. Por otro lado, 49 grupos pudieron ser anotados con funciones asociadas, con 37 grupos compartidos entre todas las especies. Si bien la redundancia génica para las enzimas dentro de cada grupo resultó similar entre especies, en ninguna de las especies de *Naganishia* pudo localizarse un gen correspondiente a una enoyl-CoA-hidratasa, a la vez que una esterol 3-beta-glucosiltransferasa no pudo ser encontrada en ninguno de sus genomas, ni en el de *F. wieringae* (Material Suplementario, Tabla S2.5).

2.3.5.1 Confirmación experimental de la producción de carotenoides y micosporinas en *N. vishniacii*

Las mediciones espectrofotométricas a 473 nm no mostraron señales de lectura para ninguno de los extractos obtenidos a partir de las 3 condiciones ensayadas (radiación, estrés oxidativo y bajas temperaturas), por lo que la producción de compuestos carotenoides no fue evidenciada. Por el contrario, las mediciones efectuadas a 310 nm posibilitaron la resolución de un pico característico asociado a las micosporinas.

2.3.6 Análisis del uso de aminoácidos

Las diferencias en el uso de aminoácidos de las 7 cepas estudiadas en relación a *C. neoformans*, se calcularon para los grupos de ortólogos de copia única comunes a todas. Además, se comparó la composición aminoacídica entre *N. vishniacii* y el resto de las especies del género *Naganishia*. En todas las especies de este género se observó un enriquecimiento en el uso de 6 aminoácidos (Ala, Arg, Asp, Gln, Met y Thr) y un menor uso de otros 5 (Glu, Lys, Pro, Ser y Trp), respecto de lo correspondiente en *C. neoformans* ($p < 0,001250$; Figura 2.3.5).

La comparación entre especies del género *Naganishia* mostró una menor frecuencia en el uso de Ala para *N. vishniacii* respecto de las demás. Además, se observó un menor uso de Pro y un enriquecimiento en Ser al comparar a la especie antártica con las dos cepas de *N. albida*.

Por el contrario, cuando la comparación se realizó sobre el conjunto de genes pertenecientes a la base de datos CEG, sólo se encontraron diferencias significativas en el uso de Gln (valor $p = 0,001250$), con un enriquecimiento para las especies *N. vishniacii* y *N. adeliensis* respecto de *C. neoformans*, pero no se observaron diferencias significativas en la composición de aminoácidos entre ambas especies de género *Naganishia*. La magnitud del cambio en el uso de este aminoácido resultó similar para ambos conjuntos de genes evaluados. En la Tabla S2.6 (Material Suplementario) se muestra una lista completa con los cambios composicionales para ambos conjuntos de genes.

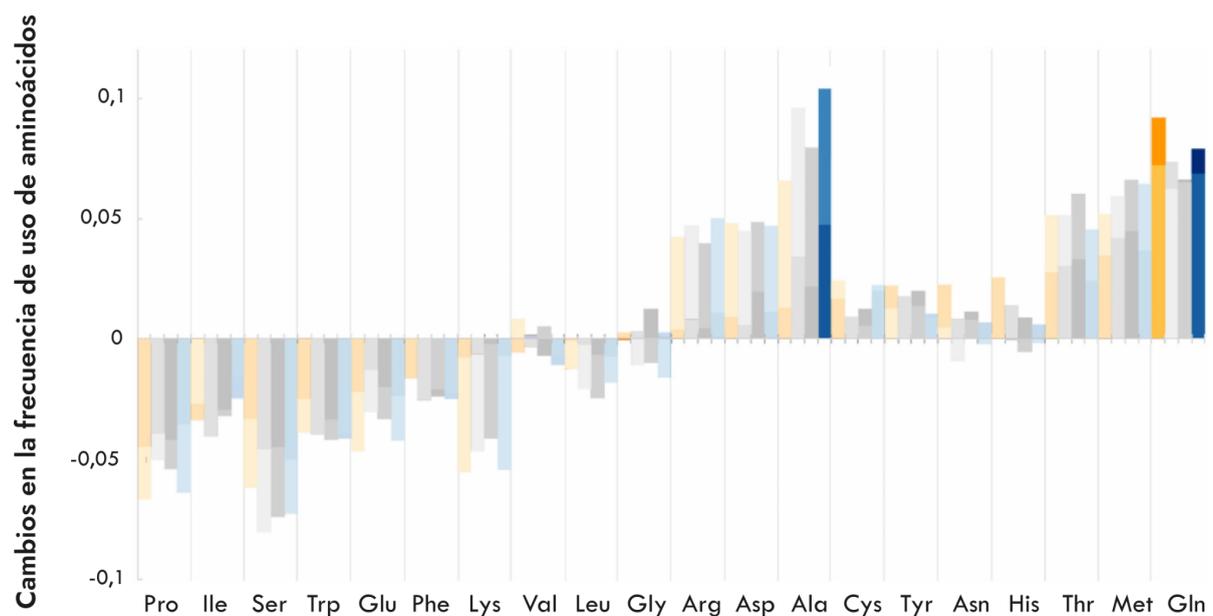


Figura 2.3.5. Cambios en la frecuencia de uso de aminoácidos para *N. vishniacii* (naranja), *N. albida* Y-1402 (gris oscuro) - NT2002 (gris claro) y *N. adeliensis* (azul), con respecto a *C. neoformans*. Se muestran los valores del análisis sobre genes conservados de la base de datos CEG (colores oscuros) y sobre los grupos de ortología de copia única comunes a todas las especies (colores claros). Las barras resaltadas corresponden a los aminoácidos que mostraron diferencias significativas para ambos conjuntos de genes.

2.4 DISCUSIÓN

2.4.1 Secuenciación del genoma, ensamblado, predicciones de genes y anotaciones

En este trabajo se describió por primera vez el genoma de la levadura extremotolerante adaptada al frío *N. vishniacii* y se identificaron las posibles adaptaciones de la misma al ambiente del cual fue aislada. El abordaje combinatorio diseñado para obtener el ensamblado genómico permitió reconstruir un genoma de alta calidad para *N. vishniacii* (ANT03-052; CBS 10616), con un bajo número de *scaffolds* (38), una muy baja proporción de regiones indefinidas (0,13 %) y una alta representatividad de genes completos de las bases de datos BUSCO (90,9 %) y CEG (99,2 %).

El tamaño del genoma de *N. vishniacii* resultó menor que el de las cepas de *N. albida*, pero comparable al de *F. wieringae* y *N. adeliensis*. Sin embargo, el número de genes predichos para esta especie fue inferior (entre 239 y 1265 genes menos) al observado para todas las demás utilizadas en este estudio. La pérdida de genes es un fenómeno pervasivo en todos los organismos (Albalat & Cañestro, 2016) y puede ser beneficiosa para los microorganismos de vida libre que habitan en ambientes extremos u oligotróficos (García-Fernández et al., 2004; Ullrich et al., 2016; Zhang et al., 2017), o bien que se encuentran en complejas comunidades microbianas sobre todo en ambientes homogéneos (Morris et al., 2012; Mas et al., 2016). Esta pérdida de genes, que ocurre por selección y no por deriva, fue propuesta como la “Hipótesis de la Reina Negra”, en relación al juego de cartas de nombre homónimo, en el que el objetivo es obtener el menor puntaje por lo que se buscará desprenderse de la reina de espadas: la carta de mayor valor. En esta teoría coevolutiva, se propone que ciertos genes, o de manera más amplia, funciones biológicas, son análogos a la reina de espadas: costosas e indeseables. Por lo que resulta ventajoso para un organismo dejar de ejercerlas, siempre y cuando, otros miembros de la comunidad lo hagan. Al mismo tiempo, esta función debe proveer de bienes comunes indispensables, de manera que su retención sea obligada en, al menos, una serie de individuos en la comunidad. Lógicamente, si bien se buscará desprenderse de la reina de espadas, el juego es inviable sin la carta. Esta hipótesis explica que ciertos organismos sigan siendo funcionales y aún más, adquieran ventajas adaptativas al emplear menos recursos propios, a pesar de la reducción de sus genomas y de la pérdida de genes (Morris et al., 2012).

Si bien el impacto que genera la pérdida de genes sobre el *fitness* de las levaduras sigue siendo objeto de debate y algunos autores consideran que las ventajas son sólo marginales o están sujetas a condiciones experimentales puntuales, o bien que las ganancias en *fitness* no se deben a las deleciones selectivas, sino a beneficios espontáneos derivados de las mutaciones que las generan (Thatcher et al., 1998; Sliwa & Korona, 2005; Lang et al., 2009), el número considerablemente bajo de genes en *N. vishniacii* con respecto a las otras especies relacionadas sugiere que ésta podría ser una característica adaptativa a un ambiente con baja disponibilidad de recursos y muchos factores de estrés, pero sujeto a condiciones relativamente constantes (bajas temperaturas, alta radiación solar, baja actividad de agua) (Selbmann et al., 2014). Como se discutirá en el capítulo siguiente y en las conclusiones finales, este rasgo es común a las extremotolerantes evaluadas en esta Tesis, y sea probablemente un rasgo adaptativo de los extremotolerantes en general.

2.4.3 Grupos de ortología y asignaciones funcionales

La comparación de los grupos de ortología presentes en los genomas de Filobasidiales disponibles hasta la fecha de análisis, permitió identificar que la mayoría de las predicciones de *N. vishniacii* se encontraban representadas en las especies relacionadas. La inferencia de ortología permitió caracterizar 44 genes candidatos, correspondientes a 12 grupos de ortología (Figura 2.3.3) con un mayor número de copias (redundancia génica) en el genoma de *N. vishniacii* respecto de ambas cepas de *N. albida* y, en la mayoría de los casos, también respecto del resto de especies del orden. El interés particular volcado en estos grupos de ortología se centra en que la redundancia de genes suele estar relacionada con un aumento en la dosis de los productos de los genes codificados (Schuster-Böckler et al., 2010) y constituye un mecanismo de adaptación a condiciones ambientales novedosas o estresantes (Kondrashov, 2012), por lo que podía ser un carácter esperable en el genoma de *N. vishniacii*. Entre los grupos funcionales con mayor redundancia génica para esta levadura, la superfamilia de proteínas de transporte de membranas (MFS) reportada en este trabajo ha sido propuesta como parte de un importante mecanismo de adaptación al frío en la levadura psicrófila obligada *Mrakia psychrophila* (Su et al., 2016), probablemente contribuyendo a la acumulación de nutrientes del medio para dar lugar a todas las demás modificaciones metabólicas y reorganizaciones necesarias para contrarrestar las condiciones impuestas por el frío.

Para el resto de los grupos funcionales encontrados, la relación con el frío no está clara. Diferentes organismos contrarrestan los cambios de temperatura sobreexpresando genes involucrados en la modulación del procesamiento del ARN por medio de distintas vías (Phadtare et al., 2004; Zhu et al., 2007), como por ejemplo mediante la producción y acumulación de carbohidratos de reserva que actúan como metabolitos protectores (Schade et al., 2004), y mediante la modificación de la composición de la membrana lipídica, a fin de mantener una apropiada funcionalidad y estructura (véase más adelante). La redundancia génica de endonucleasas tipo *Dicer*, oxidorreductasas (enoil reductasas) y proteínas transportadoras de trehalosa podría estar relacionada con esas respuestas generales.

El abordaje comparativo aquí descrito, permitió además, desentrañar 16 grupos de ortología cuya presencia se constató únicamente en el genoma de *N. vishniacii* (con respecto a los otros analizados en este trabajo). La mayoría de ellos pudieron asociarse funcionalmente con Mecanismos de transducción de señales, Biosíntesis/transporte y catabolismo de metabolitos secundarios y Transcripción. También se encontraron grupos relacionados con el Transporte y metabolismo de carbohidratos, Procesamiento y modificación del ARN y Biogénesis de pared/membrana celular. Entre ellos, es destacable la presencia de un gen con dominio MFS (Nv_002329), que aporta indicios adicionales respecto del posible rol de este grupo en la adaptación al frío; una transferasa (Nv_005274): grupo funcional cuyo importancia en la tolerancia al frío de bacterias antárticas había sido previamente sugerido, posiblemente por estar implicado en la biogénesis de membranas y en modificaciones de carbohidratos, proteínas o lípidos (Grzymiski et al., 2006) y un gen relacionado con la actividad de chaperona (Nv_005274): enzimas necesarias para el correcto plegamiento de proteínas, para el mantenimiento de su conformación en frío (Ferrer et al., 2003) y que, de acuerdo a estudios previos, participaría de un efecto sinérgico conjuntamente con la trehalosa durante la adaptación de levaduras a temperaturas cercanas al punto de congelación (Kandror et al., 2004).

Finalmente, respecto de los 320 genes de *N. vishniacii* que no pudieron ser asignados a grupos de ortología y para los cuales no se encontró correlato en las especies analizadas, una gran mayoría (N = 297) no mostraron coincidencias significativas con la base de datos no redundante de NCBI. Un número similar de genes sin coincidencias en bases de datos fue reportado para especies de basidiomicetes y ascomicetes (Bellora et al., 2016; Levasseur et al., 2014; Blevins et al., 2019) y son usualmente categorizados como putativos genes

huérfanos o genes linaje-específicos. En algunos casos, puede tratarse de genes nuevos para los cuales no logran encontrarse homologías detectables con otras secuencias (Weisman et al., 2020) y suelen ser vinculados a adaptaciones ecológicas (Wilson et al., 2005). Para determinar si alguna/s de estas predicciones están involucradas en la adaptación de *N. vishniacii* a las extremas particularidades de su hábitat, son necesarios análisis específicos.

2.4.4 Análisis del cociente dN/dS

La distribución de valores medios de dN/dS para los genes de copia única comunes entre *N. vishniacii* y *N. albida*, reveló que una gran mayoría se encuentra bajo selección negativa (dN/dS < 1). La presión selectiva positiva sobre genes suele ocurrir durante ciertos lapsos acotados de tiempo a lo largo de la evolución, afectando únicamente a algunos aminoácidos en regiones específicas de dominios, mientras que el resto de las regiones suelen estar bajo selección purificadora. Por estos motivos el cociente para la mayoría de los genes suele ser menor a 1 para especies divergentes. En este sentido, la comparación de valores de cocientes promedio (computados para secuencia completa) suele ser engañosa y rara vez supera a la unidad (Gillespie, 1991). Más aún, diversos autores han reportado valores promedio menores o mayores a 1, tanto bajo selección negativa como bajo selección positiva (Kryazhimskiy & Plotkin, 2008; Spielman & Wilke, 2015) cuestionando la validez de estos análisis. Incluso se ha señalado que se ven afectados por artefactos estadísticos, *hitchhiking* provocado por otros genes, y particularmente por el tiempo de divergencia, limitando estos análisis a especies de divergencia reciente (Rocha et al., 2006). Sin embargo, el cociente dN/dS es un estadístico general muy extendido en los análisis de genómica comparativa que permite, no solo tener un panorama general de cómo se encuentra actuando la selección sobre todos los genes, sino también, ocasionalmente identificar genes que hayan sufrido una fuerte selección positiva reciente. En este trabajo se observó una distribución de genes respecto a su cociente dN/dS coincidente con lo reportado habitualmente para otros microorganismos. Como se destacó en los resultados, el 98 % de los genes ortólogos evaluados presentan valores de cociente menores a 0,2, lo que indica que, para los genes comunes compartidos con *N. albida*, la selección ha sido principalmente depuradora desde que compartieron un ancestro común. Ésto es esperable dado que la mayoría de estos genes pertenecen a funciones celulares comunes que en la mayoría de los casos no requieren adecuaciones particulares para ambientes extremos. Sin embargo, con el

objetivo de identificar regiones de genes particulares que hayan requerido una adecuación es que se procedió a analizar qué ocurría dentro de las regiones de cada uno de ellos.

En base a lo anterior se aplicó un análisis mediante *sliding windows*, donde el gen se analiza de forma fraccionada en una ventana de análisis predeterminada (de 100 codones) que se desplaza río abajo a lo largo del mismo con saltos (de 5 codones) con el objetivo de identificar regiones donde el cociente supere la unidad. Mediante esta estrategia pudieron identificarse 35 genes con regiones con un cociente dN/dS >1, evidenciando selección positiva. Debido a que se empleó una estrategia altamente conservativa, considerando únicamente aquellas regiones comunes que resultaban en cocientes > 1 al comparar a *N. vishniacii* con las especies hermanas más cercanas y con la especie externa *F. wieringae*, siempre y cuando, el valor en dicha región fuese < 1 para la estimación de la especie hermana frente a *F. wieringae*, únicamente se tomaron en consideración los genes con señales de selección presentes en *N. vishniacii*, sugiriendo que el cambio adaptativo habría ocurrido en esta especie, y no antes. Entre ellos, cerca de un 40 % de los genes se relacionan funcionalmente a: Modificaciones transcripcionales; recambio de proteínas y chaperonas; Procesamiento y modificación del ARN; y Recombinación y reparación. Todos estos genes están involucrados de alguna manera en la expresión y plegado de proteínas, roles que, como se mencionó previamente, han sido previamente reportados en relación a la tolerancia al frío de organismos psicotolerantes (Phadtare et al., 2004). Más aún, entre las adaptaciones de los organismos psicrófilos se ha descrito el desafío de mantener los niveles transcripcionales y de traducción, así como el plegado de proteínas en temperaturas bajas (Kumar et al., 2018). El hallazgo de que gran parte de los genes identificados por tener presión de selección positiva se encuentran relacionados con funciones de expresión y reestructuración general del transcriptoma y el proteoma resulta de gran relevancia para comprender cómo los microorganismos pueden adaptarse a ambientes extremos.

El análisis particular de cada uno de los genes demuestra que dos de ellos poseen especial importancia considerando el particular ambiente del cual fue aislada *N. vishniacii*. El gen Nv_005203 (ENOG410PSCK) mostró el mayor valor medio. La búsqueda por HMMER de este candidato reveló que se trataría de un gen relacionado con A0A1E3HK01 de *C. depauperatus*, y Q5KHT7 de *C. neoformans* (accesos correspondientes a la base de datos Uniprot). Ambos genes están relacionados con la función de sensado de la osmolaridad extracelular, y que disparan la expansión de la membrana celular. En este sentido las

adaptaciones extremófilas de *N. vishniacii* no se limitan a su carácter psicrófilo, debido a que también ha sido descrita como una especie xero-tolerante (Vishniac, 2006).

Teniendo en cuenta que los Valles Secos de McMurdo constituyen los desiertos fríos más extremos conocidos (Vishniac & Hempfling, 1979b) y la ausencia de agua en los suelos xéricos es uno de los factores principales que limitan el crecimiento microbiano en estos ambientes (Klingler & Vishniac, 1988), la presión selectiva sobre el gen Nv_005203 está probablemente más vinculada con la necesidad de *N. vishniacii* de sobrevivir en ambientes con baja actividad de agua que con las condiciones de temperatura. Una de las respuestas de supervivencia más estudiadas en los hongos xerófilos para crecer frente a baja disponibilidad de agua es la acumulación intracelular de solutos compatibles como el glicerol, a fin de mantener el balance de agua entre el interior y el exterior de la célula (Pettersson & Leong, 2011). El osmosensor de membrana Nv_005203 que se encontró en este análisis es altamente similar al gen Sho1 descrito en *S. cerevisiae* y *Aspergillus* spp., involucrado en la síntesis y acumulación de glicerol (Salazar et al., 2009). En *S. cerevisiae*, esta proteína forma parte de la cascada de transducción de señales que bajo diferentes condiciones de estrés (particularmente estrés hiperosmótico, oxidativo y frío) desencadena una serie de respuestas transcripcionales (Rodríguez Peña et al., 2010). Por todo esto el gen Nv_005203 resulta el principal candidato entre los genes con señales de selección positiva para explicar una de las adaptaciones de *N. vishniacii* a la xerofilia. Por otro lado, se destaca la señal de selección positiva sobre el gen Nv_003572 (ENOG410PGIW). La búsqueda de la secuencia por HMMER reveló similitud con una serina/treonina proteína kinasa relacionada con A0A1E3IMU4 (Uniprot) de *C. depauperatus*. Esta proteína es una kinasa de respuesta a daños en el ADN que pertenece a la familia de fosfoinositol 3-quinasas (PIKK *family*). Miembros de esta familia de enzimas son activados bajo daños en el ADN, amplificando aún más la señal mediante la activación de otras kinasas (Durocher & Jackson, 2001). En células con daños en el ADN debidos a estrés genotóxico, como la radiación ionizante, la luz UV o bloqueos en la replicación del ADN (Kozhina et al., 2011), la propagación de la señal resulta en el arresto del ciclo celular y la estimulación de mecanismos de reparación del ADN (Bøe et al., 2018). La selección en este gen, por lo tanto, podría estar relacionada con la alta incidencia de radiación solar característica del ambiente propio de *N. vishniacii* (Cary et al., 2010), como será discutido más adelante.

Las funciones de los demás genes con fuertes señales de selección positiva no pudieron ser inferidas, por lo que fueron anotados como proteínas hipotéticas. La relevancia funcional de estas predicciones en la respuesta de *N. vishniacii* a las condiciones extremas de su hábitat son tema de futuros estudios.

2.4.5 Búsqueda de genes involucrados en la respuesta a frío

La búsqueda de genes previamente relacionados con la tolerancia a frío en levaduras y otros organismos se realizó bajo la hipótesis de que las estrategias más frecuentemente utilizadas por otros microorganismos debían estar presentes en nuestra levadura en estudio. Sin embargo, si bien el genoma de *N. vishniacii* revela que las secuencias que codifican para estos rasgos se encuentran presentes en su mayoría, en base a los resultados aquí presentados, esta especie pareciera carecer de una estrategia de gran importancia como son las proteínas anticongelantes (AFPs). Las AFPs han sido identificadas en varios organismos psicrófilos (Ewart et al. 1999), incluídas levaduras tales como especies basidiomicetes de géneros *Glaciozyma*, *Leucosporidium* y *Rhodotorula* (Kim et al., 2014). Estas proteínas actúan previniendo el daño celular debido a la cristalización de hielo extracelular que mecánicamente produce daños sobre las células (Lee et al., 2010; Hashim et al., 2014). Las AFPs son capaces de unirse a estos cristales evitando el crecimiento de los mismos y por lo tanto evitando el daño. La búsqueda intensiva utilizando todas las secuencias *query* disponibles en el NCBI, tanto de levaduras como de bacterias no permitió hallar secuencias similares en el genoma de *N. vishniacii*. Debido a que esta especie vive en los suelos antárticos, es esperable que cuente con estrategias para evitar la formación de cristales de hielo. Aunque los resultados obtenidos en este trabajo indican que *N. vishniacii* no posee AFPs similares a ninguna de las conocidas, es importante destacar que estas proteínas han evolucionado muchas veces en diferentes organismos a partir de diferentes péptidos (Cheng, 1998) y que se caracterizan por poseer una región extensa relativamente plana e hidrofóbica en su superficie generalmente conformada por motivos repetitivos (Wang et al., 2017). Por estas razones, no es posible asegurar que *N. vishniacii* no posea otras AFP aún desconocidas diferentes de las que se encuentran descritas en otros organismos.

Más allá de la ausencia de este tipo de proteínas anticongelantes, la búsqueda (*genome mining*) de otros genes habitualmente relacionados con la supervivencia de levaduras extremófilas permitió identificar secuencias homólogas para la síntesis de

antioxidantes, fotoprotectores, metabolitos crioprotectores, y de la síntesis de lípidos. Respecto a los genes relacionados con respuestas antioxidantes y moléculas de protección UV, se identificó la ruta completa para la síntesis de carotenoides y micosporinas. La producción de carotenoides y compuestos de protección UV es un rasgo muy extendido, presente en muchas especies de levaduras (Libkind et al., 2011; Moliné et al., 2014) y ambos compuestos juegan un rol importante en la fotoprotección y poseen propiedades antioxidantes (Moliné et al., 2009; 2011). La Antártida comprende regiones con alta incidencia de radiación solar como resultado de los largos períodos de exposición durante el verano y de la reducción de la capa de ozono (Madronich et al., 1998; McKenzie et al., 2011). Esta radiación es un factor de estrés que limita las posibilidades de vida (Onofri et al., 2007). En los ecosistemas rocosos fríos, como son los Valles Secos de McMurdo, suelen aislarse especies de levaduras de los órdenes Filobasidiales y Tremellales (Selbmann et al., 2014). La presencia de los genes responsables de la síntesis tanto de carotenoides como de micosporinas (particularmente MGG), fueron inequívocamente localizados en *N. vishniacii*, así como en el resto de las especies del género *Naganishia* y en *F. wieringae*, pero no en el resto de las especies bajo estudio. La presencia de compuestos fotoprotectores en levaduras de estos ambientes es, por lo expuesto, un rango esperable. La confirmación de los genes responsables de la producción de MGG, junto con la demostración experimental de su producción, permite hipotetizar que esta especie de levadura usa estos metabolitos para sobrevivir a la alta exposición solar de dichos ambientes.

Por otro lado, especies del género *Naganishia* han sido reportadas entre los organismos más resistentes a la radiación UV (Schmidt et al., 2017). Por el contrario, si bien se hallaron los 6 genes involucrados en la síntesis de carotenoides, en este trabajo no pudo confirmarse experimentalmente la síntesis de estos pigmentos en *N. vishniacii* para ninguna de las condiciones de foto-estimulación, estrés oxidativo y temperaturas evaluadas. Las colonias de las tres especies del género *Naganishia*, así como las de *F. wieringae*, adquieren en cultivo un color blanco o crema (Savoldelli & Tortorano, 2009; Kurtzman et al., 2011) y en algunas condiciones ligeramente rosáceo. Si bien en este trabajo esta última coloración fue observada, la misma no se correlaciona con pigmentos carotenoides. La ausencia de compuestos carotenoides en los cultivos de *N. vishniacii* solo puede deberse a la falta de condiciones necesarias para estimular su síntesis. La presencia de pigmentos carotenoides en levaduras adaptadas al frío ha sido exclusivamente asociada a cambios de temperatura,

protección antioxidante o respuestas fotoprotectoras (Moliné et al., 2014). La ausencia de la acumulación de estos pigmentos en *N. vishniacii* ante las condiciones ensayadas sugiere que en estos microorganismos extremotolerantes podrían cumplir otros roles aún desconocidos.

Por otra parte, la presencia de las enzimas involucradas en la síntesis y degradación de trehalosa también fue confirmada para todos los genomas de los Filobasidiales estudiados. La trehalosa es un metabolito protector que ayuda a mantener la integridad y funcionalidad de la membrana (Rudolph et al., 1986) frente a condiciones ambientales extremas, como son las temperaturas de congelamiento (Wiemken, 1990). En ciertas levaduras, un aumento en la acumulación intracelular de trehalosa es suficiente para conferir alta tolerancia a la desecación (Tapia et al., 2015). La presencia de los genes de trehalosa, conjuntamente con una expansión en proteínas transportadoras de este disacárido (mencionada más arriba) sugiere que la capacidad de sintetizar este compuesto podría ser de utilidad para sobrellevar las frías condiciones y baja disponibilidad de agua características de los Valles Secos de McMurdo.

En el mismo sentido, la presencia del set completo de genes de desaturasas en el genoma de *N. vishniacii* podría guardar relación con su capacidad de responder al frío. La fluidez de la membrana puede ser regulada mediante cambios en su composición lipídica (Ernst et al., 2016). El aumento de los dobles enlaces en las cadenas de ácidos grasos de los fosfolípidos, permite mantener un estado físico apropiado, aún a bajas temperaturas (Chintalapati et al., 2004). Este hecho se evidencia mediante la expresión de desaturasas $\Delta 12$ recombinantes, tras lo cual la supervivencia de *S. cerevisiae* a bajas temperatura se incrementa (Shi et al., 2012). En este trabajo fue posible localizar el set completo de desaturasas para todas las especies de Filobasidiales, incluyendo la presencia de desaturasas $\Delta 12$, lo cual indica que estas especies podrían tener la capacidad de regular la fluidez de membrana bajo condiciones extremas. Asimismo, cuando se comparó individualmente la redundancia génica de una variedad de enzimas involucradas en el metabolismo de lípidos, resultó imposible hallar la enzima esteroil 3-beta-glucosiltransferasa tanto en los genomas de las especies del género *Naganishia* analizadas en este trabajo, como en el de *F. wieringae*. Esta enzima, que cataliza la síntesis de esteroil glucósidos y acil esteroil glucósidos, es ubicua en hongos (Grille et al., 2010). La adición de una molécula de azúcar al grupo 3'-hidroxilo de un esteroil incrementa el tamaño de la cabeza hidrofílica del lípido, alterando sus propiedades biofísicas y se ha observado que una disminución en la proporción de

esterol-glucósidos y esterol-glucósidos acilados se correlaciona con una mayor resistencia a estrés por congelamiento (Thieringer et al., 1998). En este contexto, y dado que la glicosilación de los esteroides no resultaría en un beneficio para las especies adaptadas al frío, puede esperarse que *N. vishniacii* carezca de dicho gen. Sin embargo, este resultado debería ser corroborado experimentalmente. Igualmente, en ninguna de las especies del género *Naganishia* pudo localizarse la enzima enoyl-CoA-hidratasa. En base a la bibliografía consultada, este resultado no pudo ser interpretado en un contexto adaptativo. Su estudio y validación, siendo un rasgo compartido únicamente por las especies de este género, quedan pendientes.

La búsqueda de genes específicos relacionados con la psicofilia resultó un método adecuado para confirmar o rechazar la producción de compuestos y estrategias previamente descritas en otros organismos, e incluso para confirmar potencialidades, como la posibilidad de sintetizar carotenoides en *N. vishniacii*; pero también posibilitó demostrar que estrategias presentes en multitud de especies psicrófilas, como es la síntesis de AFPs, no necesariamente son características extendidas a todos los microorganismos adaptados al frío. La evolución de las adaptaciones no posee pautas predeterminadas y si bien las estrategias de *genome mining* permiten comprender qué pueden y qué no pueden hacer estos microorganismos, la ausencia de determinadas características poco tienen que ver con la imposibilidad de éstos para ocupar estos nichos.

2.4.6 Análisis del uso de aminoácidos

Finalmente, en este trabajo se realizó un análisis del uso de aminoácidos para todos los genes del conjunto de ortólogos, así como para genes eucariotas altamente conservados (CEG), con el fin de determinar si la evolución favoreció el uso de determinados aminoácidos por sobre todos, durante la evolución de la psicofilia. Los resultados permitieron demostrar un considerable enriquecimiento en el uso de glutamina en el proteoma de *N. vishniacii*, fenómeno que fue muy evidente para el set de genes eucariotas altamente conservados (CEGs) y no solo respecto a otras especies de Filobasidiales lejanas sino también con respecto a la especie psicrotolerante hermana, *N. adeliensis* (Scorzetti et al., 2000). Algunos autores han reportado que la principal característica estructural para enzimas adaptadas al frío es su mayor flexibilidad, un rasgo que posibilita compensar la reducción de la velocidad de las reacciones químicas inherente a las bajas temperaturas

(Gerday et al., 2000; Gianese et al., 2001). Por esta razón, dichas enzimas suelen estar constituidas por un mayor uso de residuos de menor tamaño (Adekoya et al., 2005), lo que permite reducir el costo energético de los cambios conformacionales que se requieren para interactuar con los sustratos. El incremento en la eficiencia catalítica constituye una característica clave para las enzimas de bajas temperaturas, que deben compensar las bajas tasas metabólicas debidas a la baja energía cinética disponible, para superar las barreras de reacción (Georlette et al., 2004). En este sentido, se ha observado en Arqueas que existe una tendencia lineal en toda la gama de temperaturas óptimas de crecimiento, desde psicrófilas hasta hipertermófilas, con respecto al uso de aminoácidos como la glutamina (Saunders et al., 2003). Consistentemente con esta observación realizada en un grupo muy distante evolutivamente, el porcentaje de uso de glutamina en *N. vishniacii* resultó ser el mayor entre todas las especies comparadas. Por otro lado, en el proteoma de esta especie también se observó una marcada tendencia a reducir el contenido de prolina, al compararlo con el de *C. neoformans* y los de ambas cepas de *N. albida* (aunque no frente a la especie psicrófila *N. adeliensis*). Si bien este cambio no resultó estadísticamente significativo, resulta destacable debido a que una tendencia similar ha sido reportada como atributo de enzimas adaptadas al frío (Feller et al., 1997). El uso de prolina en los péptidos, al contrario que la glutamina, restringe las rotaciones del esqueleto de las moléculas y reduce su flexibilidad (Siddiqui & Cavicchioli, 2006). Por esta razón, dicho aminoácido suele ser más común en proteínas adaptadas a temperaturas altas (Aghajari et al., 1998; Spiwok et al., 2007; Tronelli et al., 2007; Sælensminde et al., 2009). La disminución en el uso de prolina en *N. vishniacii* y *N. adeliensis* está vinculada con su adaptación a ambientes fríos. Si bien el menor uso de prolina y el mayor de glutamina han sido sugeridos desde modelos teóricos e *in vitro* con confirmación en Arqueas, y el uso preferencial de algunos aminoácidos se ha observado en algunas levaduras psicrófilas, este trabajo es el primero en demostrar una tendencia tan clara dentro de la evolución de levaduras psicrotolerantes de un clado y comparando con especies hermanas tan cercanas.

En este trabajo, la exploración del genoma de la levadura extremotolerante *N. vishniacii* permitió revelar una amplia gama de características genómicas que podrían constituir las bases subyacentes a su capacidad para crecer en los ambientes extremos de los suelos secos de Antártida. En base a la bibliografía consultada, desde que esta especie fue aislada por primera vez de areniscas del supergrupo de Beacon (Vishniac & Hempfling,

1979a), no se había reportado su presencia fuera de esta zona en Antártida (Buzzini & Margesin, 2014). Si bien algunos autores han propuesto que gran parte de la microbiota de este continente podría proceder de otros lugares y colonizar este territorio mediante dispersión aérea (Vincent, 2000; Schmidt et al., 2017), la habilidad de *N. vishniacii* para soportar los cambios ambientales diarios y anuales de su peculiar hábitat y su repetido aislamiento de este lugar, sugerían que podría constituir una especie endémica (Vishniac & Hempfling, 1979b). Aunque la falta de estudios sistemáticos e independientes del cultivo en esta región hacen que su biogeografía, su relevancia ecológica y los patrones de distribución espacial de los hongos sean cuestiones aún no del todo reveladas (Dreesen et al., 2014). En este sentido, un estudio reciente (Dragone et al., 2023) indicaría que esta especie está presente también en los glaciares del Monte Everest, lo cual es un nuevo indicativo de su picrotolerancia y adaptación exclusiva a ambientes fríos pero ampliaría la distribución del hábitat de la especie.

La genómica comparativa en base a la anotación normalizada de 6 genomas de especies cercanas, permitió identificar un conjunto de genes candidatos probablemente implicados en la adaptación al frío y también diferentes tendencias durante la evolución del genoma de *N. vishniacii* en relación con su capacidad para hacer frente a condiciones de frío extremo, incluyendo genes que son únicos, con mayor número de copias, que albergan señales de selectividad positiva, implicados en vías metabólicas específicas, y con cambios únicos en la composición de aminoácidos.

En base a la bibliografía consultada, tal combinación de características genómicas de tolerancia al frío no había sido reportada previamente en hongos. Si bien el análisis del genoma constituye un primer paso para comprender las adaptaciones de un dado organismo para sobrevivir bajo condiciones extremas, la integración de una mayor variedad de abordajes ómicos *downstream* es necesaria para inferir posibles mecanismos subyacentes a estas potencialidades.

2.5 CONCLUSIONES

- El genoma obtenido para *N. vishniacii* mediante la fusión de dos estrategias de ensamblado fue de muy alta calidad, con un bajo número de *scaffolds* y un muy bajo contenido de regiones indefinidas. El empleo de la misma metodología de predicción génica y anotación, tanto para *N. vishniacii* como para las especies de referencia, permitió realizar comparaciones consistentes.
- El análisis filogenómico permitió establecer las relaciones con las especies hermanas psicrotolerantes *N. albida*, y *N. adeliensis*.
- El genoma de *N. vishniacii* posee entre 239 y 1265 genes menos que las especies de referencia. Esta disminución de información podría relacionarse con el ambiente extremo en el que vive.
- Los genes relacionados con la traducción, transcripción, y modificación de productos de expresión, así como los relacionados con la biogénesis de la pared y la membrana celular y transportadores de membrana de familia MFS poseen alta redundancia, sugiriendo de relevancia su estudio funcional en relación a la adaptación al frío de *N. vishniacii*.
- Existe una gran cantidad de genes potencialmente huérfanos, cuyo estudio podrá proveer nuevas pautas para entender la tolerancia de *N. vishniacii* a las condiciones del ambiente o a su biología.
- Las predicciones genéticas demuestran que el 98 % de los genes están bajo selección depuradora y que sólo 35 estuvieron bajo selección positiva reciente. La tolerancia a la xerofilia y la resistencia a la radiación son dos rasgos seleccionados en esta especie.
- La levadura cuenta con estrategias de supervivencia a estos ambientes extremos ampliamente descritos para otros microorganismos, como son la producción de compuestos antioxidantes y fotoprotectores, del crioprotector trehalosa, y un metabolismo de lípidos orientado al mantenimiento de una composición de membrana adaptada al frío.
- *N. vishniacii* pareciera carecer de genes conocidos necesarios para la síntesis esterol-glucósidos y esterol-glucósidos acilados, lo que podría permitir producir

membranas celulares con mayor fluidez como adaptación a las bajas temperaturas. Este resultado y su relevancia adaptativa requieren de una validación.

- Existe un uso de codones preferencial que aumenta la proporción relativa de la glutamina y disminuye la de prolina, de modo que las proteínas presenten una mayor flexibilidad como adaptación a las bajas temperaturas.



Capítulo 3

Sobrevivir con menos, resistiendo más
testimonio genómico de un clado ácido y metal-tolerante

3.1 INTRODUCCIÓN

3.1.1 Ambientes acuáticos ácidos con presencia de metales

Los ambientes acuáticos ácidos representan un grupo de ecosistemas extremadamente singulares caracterizados por un pH muy bajo (generalmente inferior a 3) y usualmente con altas concentraciones de especies metálicas, ya que esta condición permite la solubilización de las mismas. Estos entornos se encuentran en muchas partes del mundo y pueden tener dos orígenes diferentes: natural o antropogénico (creado por el ser humano) (Johnson, 1998). Los primeros existen como resultado de procesos naturales, como es el caso de los cuerpos acuáticos que rodean géiseres y volcanes, donde altas concentraciones de compuestos inorgánicos reducidos del azufre y ácido sulfhídrico proporcionan condiciones para un pH muy bajo. Tal es el caso del Río Agrio en Argentina (Cabrera et al., 2020) y el lago Kawah Idjen en Indonesia (Varekamp et al., 2000), ambos generados por actividad volcánica, y los sistemas Seltun y Hveradalir en Islandia (Aguilera et al., 2010) y el Parque Nacional Yellowstone (Spear et al., 2005), de origen geotérmico. También existen hábitats ácidos naturales producto del drenaje ácido de rocas (Bird et al., 2006; Gammons et al., 2021; Zarroca et al., 2021). Sin embargo, estos últimos son raros en comparación con los de origen antropogénico. Los hábitats ácidos antropogénicos son el resultado de actividades humanas, principalmente la actividad minera, los vertederos de carbón y el vertido de efluentes industriales y cloacales (López-Archilla et al., 2004), y son muy comunes en diferentes países alrededor del mundo, generando fuentes de contaminación que representan graves problemas ambientales (Akcil & Koldas, 2006). En estos casos, la exposición de minerales sulfurosos en su forma reducida, como la pirita (FeS_2), combinada con agua y oxígeno, dan lugar a reacciones que conducen a la generación de ácidos con acumulación de sulfatos e iones metálicos, como el hierro ferroso (Dragisic et al., 1999, Johnson, 2001; Kuyucak, 2002; Johnson & Bangor, 2009).

En estos ambientes, donde la acidez facilita la solubilización de elevadas concentraciones de especies metálicas, suelen encontrarse elementos como hierro, aluminio, cobre, zinc, cadmio, entre otros, y su presencia en concentraciones tóxicas plantea desafíos significativos para la vida acuática. Los metales y metaloides representan la mayoría de los elementos químicos (Gadd, 1992) y pueden clasificarse como esenciales y no

esenciales para un dado organismo, dependiendo si cumplen o no funciones biológicas (Castro et al., 2019). En tales casos, los metales esenciales actúan como cofactores de enzimas, establecen gradientes de carga y concentración a través de las membranas, intervienen en la estabilidad de estructuras celulares, y forman parte constitutiva de proteínas y ácidos nucleicos (Gadd 1992). Sin embargo, aún éstos, en altas concentraciones pueden resultar tóxicos (Castro et al., 2019).

En los microorganismos, los metales ejercen su toxicidad desplazando a otros metales esenciales de sus sitios de unión o interactuando con grupos funcionales de proteínas y ácidos nucleicos, generando de esta forma cambios conformacionales y provocando la disrupción de su funcionalidad. También alteran la permeabilidad de la membrana y generan especies reactivas de oxígeno, ocasionando más daños sobre las biomoléculas. Todas estas interacciones pueden limitar el crecimiento de los microorganismos, generar cambios morfológicos anormales e inhibir procesos bioquímicos indispensables (Roane et al., 2015).

Debido, entonces, a la alta concentración de protones y de metales disueltos, los ambientes acuáticos ácidos se consideran extremos por presentar condiciones poco propicias para la mayoría de las formas de vida conocidas. Sin embargo, a pesar de su aparente inhospitalidad, estos hábitats albergan una biodiversidad adaptada y exclusiva que ha evolucionado para prosperar en condiciones extremas y de la cual los microorganismos son protagonistas.

3.1.1.1 La acidificación de los ambientes

Muchos procesos biológicos propios del metabolismo generan acidez en el medio. Sin embargo, determinados procesos de oxidorreducción en el metabolismo de ciertas especies de bacterias poseen un marcado efecto en la acidez. La oxidación de azufre (elemental o mineral) puede resultar en la producción de ácido sulfúrico. Este proceso, que puede ser abiótico, es fuertemente acelerado por la acción biológica de ciertas bacterias y es responsable de la generación de la mayoría de los ambientes extremadamente ácidos del planeta (Johnson, 2014). Conocido como Drenaje Ácido de Rocas (ARD por su sigla en inglés), el mismo se desencadena tras la exposición al agua y al aire de minerales sulfurosos (como la pirita o la calcopirita) que posteriormente se oxidan provocando la acidificación del ambiente acuático (Nordstrom & Southam, 1997). Por otra parte, en las áreas volcánicas el azufre elemental, que se produce por condensación del sulfuro de hidrógeno y el dióxido de

sulfuro de los gases emanados, es también oxidado por arqueas y bacterias sulfo-oxidadoras (Johnson, 2014). Además, un proceso análogo, el Drenaje Ácido de Minas (AMD), sucede en torno a minas de metales y yacimientos de carbón, donde la actividad extractiva expone minerales sulfurosos que en contacto con el oxígeno o el hierro, se oxidan generando ácido sulfúrico, cuya producción es fuertemente acelerada por organismos litotróficos que regeneran el hierro propiciando una mayor disolución oxidativa (Johnson & Quatrini, 2020).

Tanto los AMD, los ARD¹, y los sitios de áreas volcánicas, además de ser extremadamente ácidos contienen una alta concentración de metales que son movilizados tras la disolución de sulfuros y otros minerales asociados y se vuelven más solubles cuanto mayor es la concentración de iones hidronio en el medio (Johnson & Halberg, 2003). Esta combinación de factores los vuelve ambientes sumamente hostiles para el desarrollo de la vida. Sin embargo, ambos son fuentes de una interesante biodiversidad, principalmente de microorganismos (Gadanhó et al., 2006; Gadanhó & Sampaio, 2009; Russo et al., 2010; Libkind et al., 2014; Rasuk et al., 2016; González-Toril & Aguilera, 2019; Johnson & Quatrini, 2020).

3.1.1.2 Sitios representativos

La existencia de ambientes ácidos de origen natural y antrópico es relevante para estudios comparativos acerca de las comunidades microbianas que se establecen en estos sitios. Dos de ellos resultan de particular importancia en este sentido. El ARD de la Faja Pirítica Ibérica (IPB por su sigla en inglés) que da origen al Río Tinto en España y el Volcán Copahue del cual se origina el Río Agrio en Argentina. Ambos sitios poseen un pH similar y una composición de metales disueltos parecida, aunque su origen, sus lugares geográficos y el tiempo desde el cual existen son muy diferentes.

La IPB se localiza en el sudoeste de la Península Ibérica y comprende parte de Portugal y de las provincias de Huelva y Sevilla, en España. Este sistema es uno de los distritos de sulfuros masivos volcanogénicos más importantes del mundo y ha sido minado desde hace más de 5000 años (Tornos et al., 2008). Si bien actualmente no hay minas activas operando, aquellas ya abandonadas e incorrectamente cerradas contienen diversas fuentes de aguas ácidas entre las que destacan las escombreras (Grande et al., 2018), con la existencia de al

¹ En lo próximo, se emplea ARD, indistintamente, sin importar su génesis

menos 57 vertederos que representan una de las mayores acumulaciones de residuos piríticos del mundo. Los drenajes que afloran de estos desechos sobrecargan la red fluvial de la IPB con metales, arsénico y sulfatos (Grande et al., 2018) y los cursos de agua que reciben estos lixiviados se caracterizan por su color rojizo, extrema acidez (valores de pH comprendidos entre 2 y 5), y contenidos anómalos de sulfatos, hierro, manganeso y elementos traza potencialmente tóxicos, como As, Bi, Cd, Cu, Hg, Pb, Sb, Sn, Ti, Zn, entre otros (Fernández-Caliani, 2008).

De estos ambientes, el Río Tinto resulta emblemático. Algunas crónicas del siglo XVI ya se referían a éste como uno de los más fuertemente impactados de la red fluvial que irriga la zona (Grande et al., 2018). La alusión a su color característico, dado por la alta concentración de pirita, lo hacía verse como un ambiente que «no cría peces ni otra cosa viva» (Nocete et al., 2005). Sin embargo, este río representa un ecosistema extremófilo único en el mundo, con una gran biodiversidad microbiana (Tornos et al., 2008; Amils & Fernández-Remolar, 2020).

Un ambiente análogo, pero de distinta génesis, es el sistema acuático Río Agrío - Lago Caviahue (RAC), situado en el Parque Provincial Caviahue - Copahue, Neuquén. Éste, a diferencia de la IPB, constituye un ambiente acuático ácido de origen geotermal, que comprende ambos cuerpos de agua (Baffico et al., 2010) y cuya acidez proviene de los gases volcánicos de dióxido de azufre, cloruro y fluoruro de hidrógeno (Pedrozo et al., 2010). El río, en su vertiente superior, nace en las laderas del volcán Copahue en dos afluentes que luego se unifican a un único curso que desemboca en el Lago Caviahue. El pH en los afluentes oscila entre 0,5 y 1,5 en sus nacientes y 2,6-2,9, en su desembocadura. El lago (con una superficie aproximada de 1000 hectáreas) mantiene su pH relativamente constante entre 2,7 y 3,0 y da origen al Río Agrío Inferior, que fluye hasta converger con el Río Neuquén. Las aguas ácidas del RAC contienen grandes cantidades de metales disueltos, entre los que son destacables el hierro (69,5 g/L) y el arsénico (0,2 g/L) (Pedrozo et al., 2001).

Ambos ambientes, antes considerados inhóspitos, albergan una gran diversidad microbiana adaptada a las condiciones extremas a las que se encuentran expuestas. A pesar de sus similitudes, varían mucho en sus características fisicoquímicas y, como consecuencia, en su ecología microbiana. Esto se debe a que los ambientes ácidos asociados a operaciones mineras son muy recientes en términos de escalas geológicas y evolutivas, a pesar de que,

como es el caso de la IPB, algunas actividades mineras tienen un largo historial (Aguilera & González-Toril, 2019).

3.1.2 Microorganismos en ambientes acuáticos ácidos

En los ambientes acuáticos ácidos, a diferencia de lo que sucede en otro tipo de ambientes extremos, las características fisicoquímicas son parcialmente consecuencia del metabolismo microbiano y no una condición inherente del sistema (como sí lo son las temperaturas extremas, alta radiación, baja presión, etc.) (González-Toril et al., 2003). Debido a su amplia gama de actividades metabólicas, los microorganismos contribuyen fuertemente con los ciclos biogeoquímicos naturales de nutrientes y materia orgánica, así como de metales y minerales (Whitman et al., 1998; Madsen, 2011). En particular, juegan un papel central en la formación y disolución de minerales, transformaciones redox de metales y metaloides, formación de complejos organometálicos, etc. que impactan en la movilidad y toxicidad de los elementos metálicos (Gadd & Raven, 2010). Esto se debe a que, en su forma soluble, los metales se vuelven biodisponibles, término que implica que pueden ser incorporados por animales, plantas o microorganismos (Roane et al. 2015). En el mismo sentido, los compuestos reducidos de hierro y azufre originados por actividad geotermal o meteorización de minerales ligada a actividades mineras, proveen fuentes de energía para ciertas bacterias y arqueas sulfo y ferro-oxidadoras que exacerban las condiciones ácidas de estos ambientes, modificando a su vez la movilidad de las especies metálicas presentes (Nordstrom & Southam, 1997; Quatrini et al., 2005).

Los principales microorganismos responsables de esta actividad son bacterias y arqueas. La bacteria *Acidithiobacillus ferrooxidans* es capaz de acelerar la oxidación de la pirita 10^6 veces respecto de lo que sucedería de manera espontánea (Mirete et al., 2017). Bacterias de este género, junto con *Leptospirillum* spp., *Sulfobacillus* spp., *Acidimicrobium* spp. y arqueas del género *Ferroplasma* spp. suelen ser de las más reportadas en ambientes ácidos tanto de origen natural, como antrópico (Mirete et al., 2017).

Entre las arqueas, el orden Thermoplasmatales es el más usual, comprendiendo las ferro-oxidadoras de género *Ferroplasma* (Mohapatra et al., 2007). También existen acidófilos heterótrofos asociados con estos ambientes, incluidos en el género *Acidobacterium*, *Acidiphilium*, y *Acidisphaera*, entre otros, involucrados en la disolución reductiva de

minerales (reducción de hierro férrico a ferroso) que pueden formar asociaciones comensales con las ferro-sulfo-oxidadoras, removiendo compuestos orgánicos tóxicos para las bacterias quimiolitotrofas (Mirete et al., 2017; Gupta et al., 2019).

Si bien la mayoría de los abordajes a los microorganismos de estos ambientes están enfocados en bacterias por ser las principales responsables de estos procesos biogeoquímicos (López-Archilla et al., 2001) y debido a que las altas concentraciones de metales limitan el crecimiento y diversidad eucariota (González-Toril & Aguilera, 2019), en las últimas décadas se ha demostrado que los ecosistemas acuáticos ácidos contienen una alta biodiversidad de organismos eucariotas, siendo éstos los principales contribuyentes a la biomasa en dichos ambientes (López-Archilla et al., 2001; Aguilera et al., 2006; Amaral Zettler et al., 2002).

Estudios basados en ecología molecular demostraron que los organismos eucariotas son tan frecuentes como los procariotas (González-Toril & Aguilera, 2019). Las microalgas (productores primarios), los protozoos (ciliados, flagelados, rotíferos y amebas, que contribuyen a la producción primaria o secundaria) y los hongos y levaduras (que actúan como descomponedores e intervienen en el reciclado de carbono (López-Archilla et al., 2001; Méndez-García et al., 2015) resultan comunes. Con respecto a los hongos en general, las especies aisladas de ambientes ácidos son consideradas ácido-tolerantes más que acidófilos estrictos, debido a que también son capaces de crecer a pH neutro o incluso alcalino (Gross & Robbins, 2000). Sin embargo, también se han aislado hongos filamentosos como *Acontium cylatium*, *Trichosporon cerebriae* y *Cephalosporium* sp. que son considerados acidófilos (Schleper et al., 1995a). Independientemente de esto, los hongos juegan un importante rol en estos ambientes debido a que, junto con otros microorganismos, forman biofilms en la superficie de las rocas donde ocurre la precipitación de minerales y metales, conformando además un soporte físico para otras poblaciones microbianas (Fournier et al., 1998). Por ejemplo, se ha reportado que la oxidación de hierro por parte de *At. ferrooxidans* es estimulada por la levadura basidiomicética *Rhodotorula mucilaginosa* (Aguilera & González, 2019). Además, algunos hongos tienen la capacidad de secuestrar metales específicos, lo cual propicia la existencia de especies menos tolerantes (Cánovas et al., 2003).

Con respecto a las levaduras de ambientes ácidos, los reportes son escasos y usualmente están enfocados en la descripción de su biodiversidad más que en los

mecanismos que les permiten adaptarse al estrés ácido. Sin embargo, muchas especies acidófilas y ácido-tolerantes han sido aisladas de la IPB (Gadanho et al., 2006; Gadanho & Sampaio, 2009; Buzzini et al., 2018), así como del sistema del Río Agrio (Russo et al., 2010).

3.1.2.1 Levaduras en ambientes acuáticos ácidos y con presencia de metales

Muchas especies de levaduras han sido aisladas de diversos sustratos ácidos, como piel de animales, fluidos gástricos, jugos de frutas, suelos y aguas volcánicas, drenajes ácidos de minas, aguas termales, lagos y ríos ácidos. Al igual que para otros hongos, la mayoría son especies ácido-tolerantes y no acidófilas estrictas (Gross & Robbins, 2000).

Las comunidades de levaduras de ambientes acuáticos ácidos suelen estar dominadas por basidiomicetes (Gadanho et al., 2006; Russo et al., 2008; Libkind et al., 2014; Gruzdev et al., 2020). El género *Rhodotorula* suele ser de los más reportados (Gadanho et al., 2006; Gadanho & Sampaio, 2009; González-Toril & Aguilera, 2019), sin embargo, el mismo ha sufrido una reducción importante en el número de especies por su naturaleza polifilética siendo muchas de sus especies re-clasificadas en géneros como *Glaciozyma*, *Leucosporidiella*, *Farysizyza*, *Microbotryozyma* y *Meredithblackwellia* (Yurkov et al., 2015). Entre los ascomicetes, se han descrito especies pertenecientes a Hemiascomycetes y Eusascomycetes (*Glomerella* sp. y *Lecythophora* sp.) (Gadanho et al., 2006). Especies que también han sido descritas en áreas geotermales (Russo et al., 2008).

Muchas de las levaduras de ambientes acuáticos ácidos no han sido registradas en otros lugares. Incluso se han observado similitudes en la diversidad de especies provenientes de diferentes sitios, lo cual sugiere la existencia de un grupo de organismos con adaptaciones específicas a estas condiciones como por ejemplo, especies del género *Goffeauzyma* (Russo et al., 2008).

3.1.3.1 El género *Goffeauzyma*

El género *Goffeauzyma* (Phylum Basidiomycota, Subph. Agaricomycotina, Clase Tremellomycetes, Orden Filobasidiales) fue propuesto en base a caracterizaciones moleculares y fisiológicas, reclasificando a las 6 especies que hasta entonces conformaban el clado Gastricus y pertenecían al género *Cryptococcus* (Liu et al., 2015). El género *Goffeauzyma* pertenece a la familia Filobasidiaceae, junto a los géneros *Filobasidium*, *Naganishia* y *Heterocephalacria* (Liu et al., 2015), Figura 3.1.1.

La mayoría de las especies de este género han sido aisladas de una gama de hábitats asociados a condiciones ácidas. La especie tipo, *G. gastricus* fue hallada de lavados gástricos de un paciente con tuberculosis y de suelo y hojarasca provenientes de hayedos ácidos en Australia, y *G. gilvescens* se encontró en suelo de tundra del Cabo Barrow, en Alaska (Fonseca et al., 2011). La única excepción a esta regla es la especie recientemente descrita *G. siamensis*, aislada de hojas de *Ananas comosus* (ananá) en Tailandia (Nutaratat et al., 2022). Las 4 especies restantes, que son las que competen a este capítulo: *G. aciditolerans*, *G. metallitolerans*, *G. iberica* y *G. agrionensis*, han sido aisladas únicamente en ambientes acuáticos ácidos (Gadannho & Sampaio, 2009; Russo et al., 2010). La asociación de *G. aciditolerans*, *G. metallitolerans* y *G. iberica*, con los drenajes ácidos de minas en la IPB y de *G. agrionensis* con el sistema acuático ácido RAC, sugiere que estas especies emparentadas poseen adaptaciones fisiológicas ligadas a las condiciones fisicoquímicas de los ambientes de los cuales fueron aisladas. Esta agrupación de especies en particular, recibió el nombre de ecoclado ARD y la incorporación de *G. agrionensis*, aislada de un ambiente ácido geográficamente distante al de las restantes, sumado a sus distintas génesis, llevó a postular la naturaleza autóctona del clado en asociación a ambientes acuáticos ácidos (Libkind et al., 2014). Esta idea, además, encontró sustento en la imposibilidad de aislarla de otros ambientes y en su repetida ocurrencia en aguas de ARD durante distintos esfuerzos de muestreo (Gadannho & Sampaio, 2009; Russo et al., 2010). Sin embargo, la reciente caracterización de *G. siamensis*, filogenéticamente muy cercana a *G. iberica*, sugiere una revisión de esta clasificación, debido a su aislamiento de un ambiente ácido (pH 3-4) (Nutaratat et al., 2022), pero no acuático, como en el caso de las 4 restantes.

Las 3 especies aisladas de IPB (*G. aciditolerans*, *G. metallitolerans* y *G. iberica*) mostraron ser resistentes a altas concentraciones de hierro, arsenito, arsenato, cadmio, cobre y zinc, y fueron halladas en muestras con un pH medio de 2,2 (Gadannho & Sampaio, 2009). *G. aciditolerans*, recuperada de las zonas con condiciones más extremas de las Minas de Santo Domingo (MSD) del lado portugués de la IPB (pH 1,8 y concentraciones de Fe y Cu mayores a 18 y 1,5 g/L respectivamente) resultó ser la especie más tolerante de las 3. Esta especie, además, pudo aislarse junto con *G. iberica* en muestras del Río Tinto, el cual discurre del lado español en torno a yacimientos hidrotermales de pirita. Por otro lado *G. metallitolerans*, a diferencia de las anteriores, sólo pudo encontrarse en MSD (Gadannho et al., 2006; Gadannho & Sampaio, 2009), mientras que la congénere sudamericana *G.*

agrionensis, fue aislada únicamente de aguas del Río Agrio con un rango de pH entre 1,8 y 6,7, con mayor abundancia en muestras provenientes de los sitios más ácidos de RAC (pH 1,8 a 2,7). Esta especie, así como las otras, demostró una alta tolerancia a metales, en particular a Cd^{2+} , Co^{2+} , Cu^{2+} , Li^+ , Ni^{2+} y Zn^{2+} . *G. agrionensis* fue recuperada de todas las secciones del RAC muestreadas, durante distintas épocas del año y a través de distintos métodos de aislamiento (Russo et al., 2010; Russo, 2011). Ensayos de laboratorio indicaron que los aislamientos de *G. agrionensis* son capaces de sobrevivir a las condiciones extremas del agua del Río Agrio por más de 260 días (Russo et al., 2008). Esto, así como los perfiles de resistencia polimetálica similares para las cuatro especies refuerzan su estrecha cercanía y su naturaleza extremófila en asociación con este tipo de ambientes.

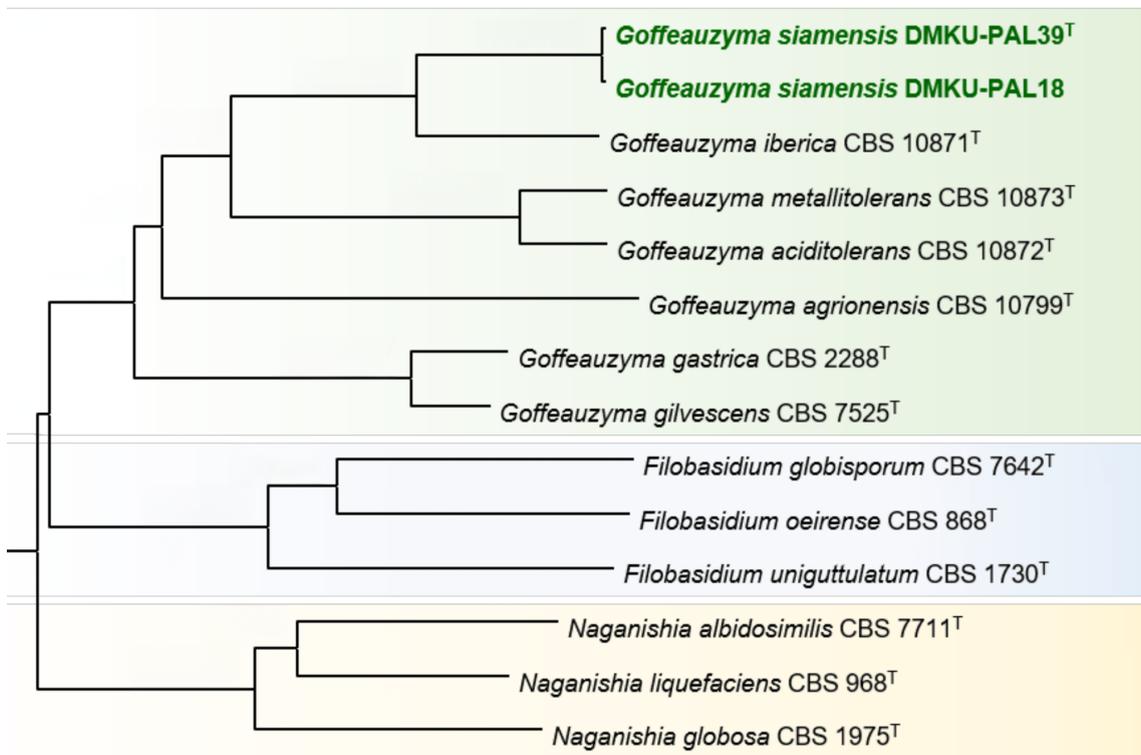


Figura 3.1.1 Árbol filogenético mostrando al género *Goffeauzyma* y las especies más cercanas del orden Filobasidiales. Basado en secuencias de ARNr SSU, dominios D1/D2 de ARNr LSU, regiones ITS, RPB2 y TEF1. Tomado con modificaciones de Nutaratat et al., 2022.

Los trabajos existentes referidos a estas especies abarcan principalmente tipificaciones fisiológicas y caracterizaciones en relación a su capacidad de tolerar las condiciones particulares de sus hábitats, incluyendo perfiles de resistencia a distintos metales. Sin

embargo, no ahondan en los posibles mecanismos subyacentes ni incorporan abordajes ómicos, objetivos que se integran en este capítulo.

3.1.4 Mecanismos de adaptación

La mayoría de los procesos fisiológicos celulares de los seres vivos se dan en valores de pH cercanos a la neutralidad (5-7), rango característico de buena parte de los ambientes naturales (Buzzini et al., 2018), sin embargo, los microorganismos son capaces de crecer a pHs extracelulares más amplios adecuando su pH intracelular a valores fisiológicos. La mayoría de las levaduras son capaces de crecer en un amplio rango de valores de pH entre 3 y 8, aunque las condiciones óptimas suelen estar entre 4,5 y 5,5 (Péter et al., 2017). Por lo que, para crecer en condiciones de pH ácidas, los microorganismos requieren adaptaciones particulares.

La acidez extrema, al igual que la alta concentración de metales, poseen varios efectos fisiológicos negativos. Estas condiciones pueden afectar la estructura y función de las proteínas al provocar cambios en la carga eléctrica de los aminoácidos alterando su plegamiento, además pueden interferir con la integridad de las membranas celulares, generando su desestabilización y alterando su permeabilidad. Muchas especies metálicas pueden interactuar con enzimas celulares inhibiendo su actividad; el equilibrio de iones en la célula también se ve afectado, y por último estas condiciones favorecen una mayor producción de especies reactivas de oxígeno y el consecuente daño oxidativo en las células (Bánfalvi, 2011).

En ambientes ácido-metálicos, las levaduras son capaces de sobrevivir merced a la existencia de mecanismos que les permiten mantener su pH interno en valores compatibles con la vida y preservar su homeostasis celular aún en presencia de metales y metaloides. En este sentido es importante remarcar que la tolerancia a condiciones de vida estresantes es un rasgo complejo que no suele estar controlado únicamente por uno o unos pocos genes (Tan et al., 2022).

Existen distintos tipos de adaptaciones que permiten sobrevivir y desarrollarse a la alta concentración de protones del medio o a la presencia de metales en concentraciones inhibitorias para otras formas de vida. Los mecanismos que subyacen a estas adaptaciones no son del todo conocidos. La mayor parte de los trabajos fueron publicados en bacterias.

Sin embargo, la tolerancia a algunos factores ambientales, como la acidez, puede llegar a ser mayor en las levaduras que en la mayoría de las bacterias que compiten en las mismas condiciones (Péter et al., 2017).

3.1.4.1 Adaptaciones a bajo pH

Para mantener la homeostasis de protones, las células emplean mecanismos de alto consumo energético que permiten extruir los iones al espacio extracelular o a vacuolas (Fletcher et al., 2017). En levaduras y hongos la tolerancia a pH ácido se logra mediante diversas estrategias, que incluyen: mayor presencia de bombas de protones en membrana plasmática que permiten mantener un pH intracelular relativamente neutro (Nicolay et al., 1987). En estos procesos, las bombas de protones (P-ATPasas) y las ATPasas vacuolares (V-ATPasas), junto con otros transportadores, juegan un papel clave (Rodrigues et al., 2007), la disminución en la permeabilidad de membrana a los protones (Salvadori et al., 2014), donde un incremento en la concentración de fosfolípidos de membrana altera su densidad contribuyendo a evitar una mayor penetración de protones (Buzzini et al., 2018) y mediante reducción de la difusión de éstos al interior por gradiente electroquímico desfavorable (protonación de distintas biomoléculas que aumentan la carga positiva intracelular) (Cavicchioli, 2000). A pesar de la existencia de estos mecanismos conocidos, existen muy pocos casos reportados de levaduras y hongos estrictamente acidófilos (capaces de crecer en pH menores a 3) (Schleper et al., 1995b) y la mayoría de los trabajos se centran en hongos acidotolerantes (Gross & Robbins, 2000). Si bien los mecanismos de homeostasis de pH no se encuentran completamente caracterizados en microorganismos acidófilos, la creciente disponibilidad de secuencias de genomas de acidófilos y de estudios funcionales metagenómicos, pueden proveer nuevos conocimientos que nos permiten entender mejor los mismos.

3.1.4.2 Adaptación a metales

Para evitar los efectos tóxicos de los metales se requiere de mecanismos de regulación que permitan asegurar que no se superen ciertos umbrales de concentración. A fin de garantizar el funcionamiento celular y mantener dicha homeostasis de metales, los microorganismos cuentan con, por lo menos, 4 mecanismos, la Biosorción; un proceso

pasivo en el que grupos funcionales en el área externa de la célula (principalmente carboxilos, aminos, hidroxilos, fosfatos y sulfhidrilos) interactúan con los iones metálicos evitando su internalización y mediando su complejación, adsorción física, intercambio iónico y precipitación (Pal & Paul, 2008). La Bioacumulación; un mecanismo metabólicamente activo que involucra la captura de compuestos tóxicos, seguida de su transporte al interior de las células donde son secuestrados por proteínas ricas en grupos sulfhidrilos o bien compartimentalizados dentro de vacuolas (Brady & Duncan, 1994). La Biotransformación; que mediante reacciones de óxido-reducción en el citoplasma convierte los metales tóxicos a formas menos dañinas. Y finalmente los Sistemas de eflujo; un mecanismo que posibilita reducir la acumulación de metales en el citoplasma, a partir de sistemas de bombeo de los mismos al exterior (Dopson & Holmes, 2014).

Las levaduras son capaces de resistir a varias especies metálicas mediante alguno o varios de estos mecanismos (Chojnacka, 2010). Esta capacidad depende tanto de adaptaciones genéticas y fisiológicas, como de parámetros ambientales (Bahafid et al., 2017). En las levaduras, la biosorción se produce mediante mecanismos de intercambio iónico o interacciones electrostáticas de los cationes metálicos con la pared celular negativamente cargada (Rahbar Saadat et al., 2021). Por ejemplo, Chen & Wang (2007) reportaron la biosorción de Zn (II) en células de *S. cerevisiae* (valores desde 74,8 a 654,8 $\mu\text{mol/g}$) vía intercambio iónico de K (I), Mg (II), Na (I) y Ca (II). *S. cerevisiae* también fue reportada como biosorbente de Pb (II) > Zn (II) > Cr > Co > Cd (II) > Cu (II) disueltos en solución (Farhan & Khadom, 2015). Kaszycki y colaboradores (2004a) estudiaron la diferente tolerancia a Cr (III) y Cr (VI) de levaduras de géneros *Pichia*, *Candida*, *Debaryomyces* y *Kluyveromyces*, donde un lento paso de captación ocurre luego de una rápida biosorción a la superficie de la célula. En esos casos se vio que la tolerancia depende fuertemente del estado fisiológico de crecimiento, la densidad de biomasa y el tiempo de contacto con el metal.

El mecanismo de biosorción puede involucrar también productos secretados, con la capacidad de unir e inmovilizar cationes en solución (Guibaud et al., 2012). En este sentido los exopolisacáridos (EPS, sustancias producidas y liberadas al entorno extracelular, formando una matriz alrededor de las células) juegan un rol fundamental al inmovilizar y retener iones metálicos (Teitzel & Parsek, 2003; Harrison et al., 2006). La biosorción y bioacumulación de metales debida a EPS es uno de los mecanismos más relevantes de

formación de biofilms y agregados celulares que contribuyen a la protección de las células del ambiente externo (Ferreira et al., 2017). Los EPS consisten en una mezcla compleja de polímeros entre los cuales se encuentran polisacáridos, proteínas, compuestos extracelulares ácidos y lípidos. En las levaduras, las estructuras de EPS más efectivas como agentes de unión a metales son los mananos, pululanos, glucooligosacáridos, galactooligosacáridos y otros heteropolisacáridos en los cuales la manosa y la glucosa son los más activos en la interacción con estos cationes (Gientka et al., 2015; Rahbar Saadat et al., 2021). Bankar y colaboradores (2018) describieron que *Yarrowia lipolytica* produce grandes cantidades de EPS, lo cual puede explicar su alta tolerancia a Zn y a Cr, mientras que un estudio con *Rhodotorula mucilaginosa* reveló gran eficiencia de remoción de Hg(II), Cu(II) y Pb(II), (Grujić et al., 2017). Por otro lado, Li y colaboradores (2019) mostraron que el glutatión (GSH, un tripéptido de acción antioxidante) puede ser un componente crítico de los EPS para la unión de Pb en *Rhodotorula mucilaginosa*. Las proteínas y los polisacáridos presentes en los EPS, resultan entonces fundamentales en la reducción de la toxicidad de metales (Chen & Wang, 2007).

La precipitación de metales es otro proceso involucrado en la biosorción, que implica la formación de metales inorgánicos insolubles (hidróxidos, sulfuros o fosfatos). La formación de biominerales de plomo luego del crecimiento de levaduras con fuentes de fósforo orgánico permitió demostrar que cepas de ciertas especies (*Pichia acacia*, *Kluyveromyces lactis*, *Candida sake*, *Candida argentea*, *Cryptococcus filicatus* y *Cryptococcus podzolicus*) son capaces de mediar la precipitación de fosfato de plomo ($Pb_3(PO_4)_2$), piromorfita ($Pb_5(PO_4)_3Cl$), anglesita ($PbSO_4$), y óxidos de plomo (PbO) (Liang et al., 2016a).

La bioacumulación, por el contrario, involucra sistemas activos de transporte de membranas que internalizan al metal presente en el ambiente celular. Este mecanismo puede iniciarse por medios pasivos o activos (Raspor & Zupan, 2006). Los metales esenciales generalmente ingresan a la célula mediante transportadores específicos, mientras que aquellos que no son requeridos, y por lo tanto carecen de sus propios transportadores, son capaces de utilizar los de los metales esenciales. El cadmio, por ejemplo, ingresa a la célula mediante proteínas que están involucradas en la captación de Zn, Ca, Mn, y Cr (VI) y es internalizado mediante transportadores de sulfato (Wysocki & Tamás, 2010). Dentro de la célula, los cationes son transformados en compuestos tiolados, o son precipitados. En general, los metales no esenciales tienen una alta reactividad por los grupos sulfhidrilos y

reaccionan con compuestos que contienen grupos sulfuros (como las metalotioneínas -MT-, el glutatión -GSH- y las fitoquelatinas -PC-), normalmente implicados en los mecanismos de detoxificación (Assis Fidalgo, 2011). En *S. cerevisiae* y en levaduras no convencionales, el GSH es un tripéptido involucrado en la modulación del estado redox y cumple un papel protector frente al estrés oxidativo. El GSH puede reducir el estado de valencia de los metales y formar complejos, siendo uno de los medios más importantes para reducir la toxicidad de los metales (Li et al., 2019). Los residuos Cys presentes en el GSH reaccionan con Cd(II) y con Cu(II), formando un complejo metal-GS. Estos complejos metal-tiolatos son expulsados fuera de la célula, o acumulados en compartimientos internos como las vacuolas (Penninckx, 2002). De manera similar al GSH, en las PC, unos pequeños péptidos tiolados unen metales vía residuos Cys, reduciendo la reactividad del ion metálico y de esa forma, su toxicidad (Clemens et al., 1999). En *S. cerevisiae PC2* se relaciona con la detoxificación de Cd, Cu y Zn. Por otra parte, las metalotioneínas (MTs) son proteínas igualmente ricas en residuos Cys. La detoxificación de Zn en *S. cerevisiae* ocurre vía MTs codificadas por *Crs5*, de manera análoga a *Zym1* en *Sch. pombe* (Assis Fidalgo, 2011). En *S. cerevisiae*, *Cup1* es la MT mayormente descrita involucrada en la detoxificación de Cu, aunque también se ha asociado con la formación de complejos y detoxificación de Cd (Tamás & Wysocki, 2001). La levadura *Cryptococcus humicola* acumula Cd (50 % del contenido en el medio) y algunos autores sugieren que compuestos fosforados -ortofosfatos y polifosfatos- participan en la detoxificación de los iones metálicos vía formación de complejos metal-fósforo en el citoplasma (Kulakovskaya et al., 2018). En *Zygosaccharomyes rouxii*, se observó la acumulación de Cd(II) y de Zn(II) tanto dentro de la célula como en su superficie, vía la expresión de transportadores, reductasas, oxidasas y permeasas (Li et al., 2014). La incorporación de cromo por levaduras también ha demostrado ser una consecuencia de ambos mecanismos: biosorción y acumulación activa (Krauter et al., 1996; Muter et al., 2001; Kaszycki et al., 2004a; Kaszycki et al., 2004b).

Finalmente, muchos autores señalan la detoxificación vía vacuolas como un mecanismo importante en la captura de Zn, Cd, Cu y Ni (Bode et al., 1995; Nishimura et al., 1998; Paidhungat & Garret, 1998; El-Nour et al., 2010). Las vesículas contienen una variedad de proteínas y compuestos orgánicos capaces de unir iones metálicos y vinculados al transporte de metales del medio intracelular al exterior de la célula (Wang & Chen, 2006).

La biotransformación también ocurre en levaduras, generando un cambio químico en el metal, tal como su estado de oxidación o metilación. La reducción de iones metálicos vía actividad enzimática puede resultar en la formación de especies menos tóxicas (Igiri et al., 2018). La reducción intracelular de Cr(VI) a Cr(III), una forma catiónica mucho menos dañina, puede ocurrir por sistemas enzimáticos (cromato reductasas) o no enzimáticos. Muchas especies de levaduras, tales como *Hansenula polymorpha*, *Saccharomyces cerevisiae*, *Yarrowia lipolytica*, *Rhodotorula pilimanae*, *Pichia guilliermondii*, *Candida fukuyamaensis* y *Rhodotorula mucilaginosa*, han sido usadas para convertir Cr(VI) a Cr(III) (Villegas et al., 2011; Chatterjee et al., 2012; Igiri et al., 2018). Por otro lado, la tolerancia de *P. guilliermondii* al cromato depende de su capacidad de reducir extracelularmente al Cr(VI) (Ksheminska et al., 2008). Se ha visto, además, que GSH y Cys se relacionan con la reducción no enzimática de cromato para su biotransformación (Bahafid et al., 2017). Adicionalmente, ciertos grupos funcionales de los EPS pueden ser activos en la reducción de altos grados de oxidación de especies de Cr(VI) y U(VI) (Jian et al., 2020).

3.1.5 Objetivos específicos

- Caracterizar los genomas de 4 especies del género *Goffeauzyma* (clado ARD) y 3 especies del orden Filobasidiales de distribución más amplia, identificando similitudes dentro del clado y diferencias respecto a las especies de referencia, de tamaño, número de genes, y otras características genómicas, mediante la obtención de ensamblados genómicos de calidad que incluyan predicciones genéticas, anotaciones funcionales e inferencias de ortología.
- Establecer las relaciones de parentesco entre las 4 especies del género *Goffeauzyma* y las especies de referencia que permitan su comparación y la identificación de los eventos evolutivos, mediante un abordaje filogenómico.
- Identificar genes que hayan estado sometidos a selección positiva entre el género *Goffeauzyma* y sus especies hermanas, ya sea mediante selección positiva, duplicaciones específicas ocurridas dentro del clado o que hayan sido reportados para otros organismos adaptados a ambientes ácidos extremos y relacionar su función con con adaptaciones nicho-específicas.

3.2 MATERIALES Y MÉTODOS

3.2.1 Secuenciación y ensamblado de genomas de especies de género *Goffeauzyma*

Se utilizaron 4 especies de género *Goffeauzyma* aisladas de ambientes acuáticos ácidos, según se listan en la Tabla 3.2.1, 3 de estas fueron aisladas de las Minas de Santo Domingo, mientras que *G. agrionensis* fue aislada en el Río Agrio.

Tabla 3.2.1. Cepas del género *Goffeauzyma* secuenciadas en este trabajo

Espece	#ID cepa	Lugar de aislamiento
<i>Goffeauzyma iberica</i> ¹	CBS10871 [†] (SDY 190)	Drenaje ácido de rocas, Minas de piritita de Santo Domingo, Portugal
<i>Goffeauzyma aciditolerans</i> ¹	CBS10872 [†] (SDY 081)	MSD
<i>Goffeauzyma metallitolerans</i> ¹	CBS10873 [†] (SDY 022)	MSD
<i>Goffeauzyma agrionensis</i> ²	CBS10799 [†] (CRUB 1317)	Sistema Río Agrio, Neuquén, Argentina

¹Gadanho et al., 2009

²Russo et al., 2010

Los genomas de cada aislamiento se secuenciaron en el laboratorio del Dr. Hittinger en la Universidad de Wisconsin - Madison, en el marco de un proyecto de cooperación bilateral CONICET - National Science Foundation. Las cepas se cultivaron en medio YPD (peptona 3g/L, dextrosa 20 g/L, agar bacteriológico 20 g/L, extracto de levadura 10 g/L), hasta alcanzar fase estacionaria y se colectaron mediante centrifugación con 500 µl de perlas de vidrio de 0,5 mm lavadas con ácido (Sigma #G8772) para luego ser resuspendidas en *buffer* de lisis (Tris 10 mM, EDTA 1 mM, NaCl 100 mM, SDS 1 % y Triton X-100 2 % en agua). Se realizó una doble extracción en una mezcla de fenol:cloroformo:isoamílico 25:24:1 (Sigma #P2069) seguida de precipitación en etanol 100 % durante toda la noche a - 80 °C. Los extractos se colectaron por centrifugación, se lavaron dos veces con etanol 70 %, se disolvieron en Tris-Cl 10 mM (pH 8) y se trataron con ARNasa A (VWR #97064-064) durante 30 minutos a 37 °C. Un volumen de 5 µg del ADNg fue sonificado empleando un ultrasonicador Covaris E220 Focused-ultrasonicator (Covaris, Massachusetts, EE. UU.). A continuación, se realizó una reparación terminal, ligación de adaptadores y selección por tamaño (500 pb) empleando el kit *DNA Library Prep* (NEB #E7370L; NEB Inc., Massachusetts,

EE.UU.) de acuerdo al protocolo indicado por el fabricante, con la excepción de que se utilizaron perlas Axygen AxyPrep Mag PCR (Fisher Sci #14-223-152) en reemplazo de las perlas Beckman Coulter AMPure XP. Las librerías obtenidas se sometieron a 2 lecturas de 250 pb en un equipo de secuenciación Illumina HiSeq 2500 NGS (Illumina). La calidad de los *reads* se analizó mediante el *software* FASTQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>) y se limpiaron las librerías de posibles contaminantes utilizando la base de datos UniVec de NCBI (disponible en <ftp://ftp.ncbi.nlm.nih.gov/pub/UniVec/>), a fin de descartar segmentos insertados en las secuencias nucleicas que pudiesen tener su origen en vectores de clonado. Los *reads* genómicos se ensamblaron utilizando el *software* SPAdes v3.6.1 (Bankevich et al., 2012) incluyendo la remoción de adaptadores, filtrado de secuencias de baja calidad y corrección de errores.

3.2.1.1 Predicciones de genes y anotaciones

Se realizaron predicciones de genes *ab initio* con GeneMark-ES v2.3e (Ter-Hovhannisyan et al., 2008) auto-entrenado con los *scaffolds* genómicos (parámetros --min_contig --max_nnn 1000). El proceso de anotación automática se realizó mediante el *pipeline* Funannotate v1.5.0 (Love et al., 2018), a partir de los *scaffolds* y las predicciones obtenidas. Se incluyó, en los casos en los que fuese posible, la anotación de dominios proteicos de Pfam v32.0 (Bateman et al., 2004), la clasificación en grupos de ortología de copia única incluidos en la base de datos BUSCO v1.0 para Basidiomycota (Waterhouse et al., 2018), la agrupación de las predicciones dentro de clústeres de genes ortólogos de acuerdo a la base de datos eggNOG v5.0 (Huerta-Cepas et al., 2019) y la asignación de categorías funcionales COG (Tatusov et al., 2000). A fin de realizar comparaciones genómicas que pudieran dar cuenta de los rasgos característicos de estas especies en asociación a sus altas tolerancias a metales y acidez, el proceso de predicción y anotación se realizó de igual forma para las especies *Filobasidium wieringae* UCDFST 05-544 (acc.: GCA_001600055.1), *Naganishia albida* NT2002 (acc.: GCA_001468955.1; Yong et al., 2016) y *Papiliotrema flavescens* NRRL Y-50378 (acc.: GCA_000442785.1; Rong & McSpadden, 2013). La elección de las mismas se basó en que al momento del análisis representaban los genomas disponibles filogenéticamente más cercanos de especies sin reportes en estas condiciones extremas. La inclusión del genoma de *F. wieringae* se realizó previa obtención de permiso

por parte de sus autores. Para las anotaciones finales, se utilizaron las identidades **_#### para cada una de las predicciones génicas de cada organismo, con **: Ga: *G. aciditolerans*, Gm: *G. metallitolerans*, Gi: *G. iberica*, Gg: *G. agrionensis*, Fw: *F. wieringae*, Na: *N. albida* y Pf: *P. flavescens*.

3.2.2 Reconstrucción filogenómica

Se realizó una reconstrucción filogenómica en base a 2031 genes de copia única compartidos entre las 4 especies de interés, *F. wieringae*, *N. albida* y *P. flavescens* (todas especies de la clase Tremellomycetes), incluyendo, además, a la especie de referencia *C. neoformans* JEC21 (acc. AE017341.1; Stajich et al., 2012) para obtener la raíz del árbol. La asociación de los genes a los correspondientes grupos de ortología se determinó mediante eggNOG-Mapper v5.0 (Huerta-Cepas et al., 2019) (ver “Grupos de ortología y asignaciones funcionales”). Las predicciones proteicas seleccionadas se alinearon mediante MUSCLE con parámetros por defecto (Edgar, 2004) y los alineamientos obtenidos se concatenaron resultando en un largo de 2.273.6016 columnas. El árbol filogenómico enraizado se obtuvo empleando RAxML-HPC2 (Stamatakis et al., 2014), utilizando el modelo de sustituciones aminoacídicas PROTGAMMALG. La fiabilidad de las ramas se determinó realizando 1000 réplicas bootstrap. El modelo evolutivo se seleccionó empleando ModelTest-NG con parámetros por defecto (Darriba et al., 2020). Las corridas de RAxML y ModelTest-NG se realizaron en la plataforma científica CIPRES (Miller et al., 2010).

3.2.3 Grupos de ortología y asignaciones funcionales

A fin de encontrar genes que pudiesen estar asociados con la tolerancia de estas especies a la alta concentración de metales y a la acidez propias de sus ambientes, las predicciones proteicas fueron categorizadas por medio de un método rápido de asignación de ortología, empleando clústeres precomputados incluidos en la base de datos de hongos de eggNOG v5.0 (fuNOG) (Huerta-Cepas et al., 2019). Esto permitió categorizar los grupos de ortología en comunes/exclusivos de las especies adaptadas a ambientes extremos respecto de las especies del grupo externo, así como estudiar posibles señales de evolución y realizar comparaciones de la cantidad de copias génicas en cada grupo (“Análisis de duplicaciones en

los grupos funcionales multicopia”). Las corridas, acotadas al grupo taxonómico Fungi, se realizaron con un e-val de $1e-3$ y una cobertura de 20 %.

3.2.3.1 Análisis de duplicaciones en los grupos funcionales multicopia

Para todos los clústeres identificados en base a eggNOG que presentaran más de una copia en alguna de las especies del género *Goffeauzyma* (N = 480), se realizaron reconstrucciones de la filogenia del gen con el fin de analizar la existencia de duplicaciones de genes que pudieran haber ocurrido dentro del clado. Para ello, se analizó secuencialmente el número de copias en el nodo/especies de interés, respecto del respectivo en las demás especies. En este abordaje, se buscó que la cantidad de copias en el nodo de interés fuese conservada (mismo valor para todas las especies comprendidas) pero distinta al resto de las especies (= número en nodo de interés, \neq respecto de los demás nodos). Para cada uno de los grupos de ortólogos seleccionados (N = 319) se construyeron árboles filogenéticos a partir de secuencias proteicas predichas utilizando el *software* MEGA11 (Tamura et al., 2021). Las secuencias de todas las copias pertenecientes a cada una de las especies fueron alineadas con MUSCLE utilizando los parámetros por defecto (Edgar, 2004). Todos los alineamientos proteicos fueron curados manualmente revisando los correspondientes a cada grupo de ortología e identificando secuencias “partidas”, o sea aquellas secuencias que para un genoma hubiesen sido anotadas como dos genes, no solapantes, cuando en realidad formaban parte de una única secuencia. Los árboles de genes se construyeron utilizando el método de máxima verosimilitud y el modelo de sustitución aminoacídica WAG y tasa de sustitución uniforme entre sitios, el resto de los parámetros fueron los sugeridos por defecto (Zmasek & Eddy, 2001; Tamura et al., 2021). La inferencia de duplicaciones se realizó en base al algoritmo de Zmasek (Zmasek & Eddy, 2001), como se encuentra implementado en el *software* MEGA11. Para ello, se introdujo el árbol de genes obtenido, sin raíz, y el mismo se concilió con el árbol filogenómico enraizado obtenido para las especies del análisis. El resultado consiste en una o más hipótesis que minimicen el número de eventos evolutivos explicando la redundancia mediante eventos de especiación y de duplicación. En aquellos casos donde se obtuviera más de una opción, se seleccionó la que se encontrara sustentada por el mayor número de hipótesis. Todas las duplicaciones ocurridas dentro del clado de las especies de *Goffeauzyma*, aún cuando sólo ocurrieran en una especie, fueron listadas.

3.2.4 Análisis de cociente dN/dS

Se estimó el cociente de sustituciones no sinónimas (dN) y sinónimas (dS) para todos los grupos de ortología de copia única comunes a cada una de las 4 especies del género *Goffeauzyma* respecto de *F. wieringae*, *N. albida* y *P. flavescens* (N = 2337). Para ello, se realizaron alineamientos múltiples de las secuencias proteicas predichas empleando MUSCLE con parámetros por defecto (Edgar, 2004). Los alineamientos proteicos obtenidos fueron convertidos a codones por medio de PAL2NAL (Suyama et al., 2006) eliminando los *gaps*. Las estimaciones se realizaron con el *software* KaKs Calculator V2.0 (Wang et al., 2010) sobre archivos derivados de los alineamientos agrupados de a pares de especies. Las secuencias fueron particionadas en fragmentos de 100 codones de longitud y se siguió un abordaje de tipo *sliding windows* con saltos de 5 codones, utilizando el modelo YN (Yang & Nielsen, 2000). El cociente se calculó también para los pares de secuencias sin particionar (secuencias completas). Se consideraron posibles señales de evolución positiva a las regiones con cocientes dN/dS > 1 ($p_{val} < 5e-02$) en ventanas comunes a la especie de interés y cada una de las especies del grupo externo, lo cual llevaría a pensar que las sustituciones habrían ocurrido en algún momento en la evolución de las especies del género *Goffeauzyma* a partir de regiones conservadas en las especies del grupo externo. Los genes seleccionados en base a estos criterios fueron transcritos a proteínas y realineados (MUSCLE, *-gapopen -2*) con las secuencias de referencia de *C. neoformans* de sus correspondientes grupos de ortología, para ser evaluados gráficamente empleando JCoDA con la opción *prealignment* y descartando los primeros 25 codones (Steinway et al., 2010).

En la Figura 3.2.1 se esquematizan los procedimientos descritos en esta sección.

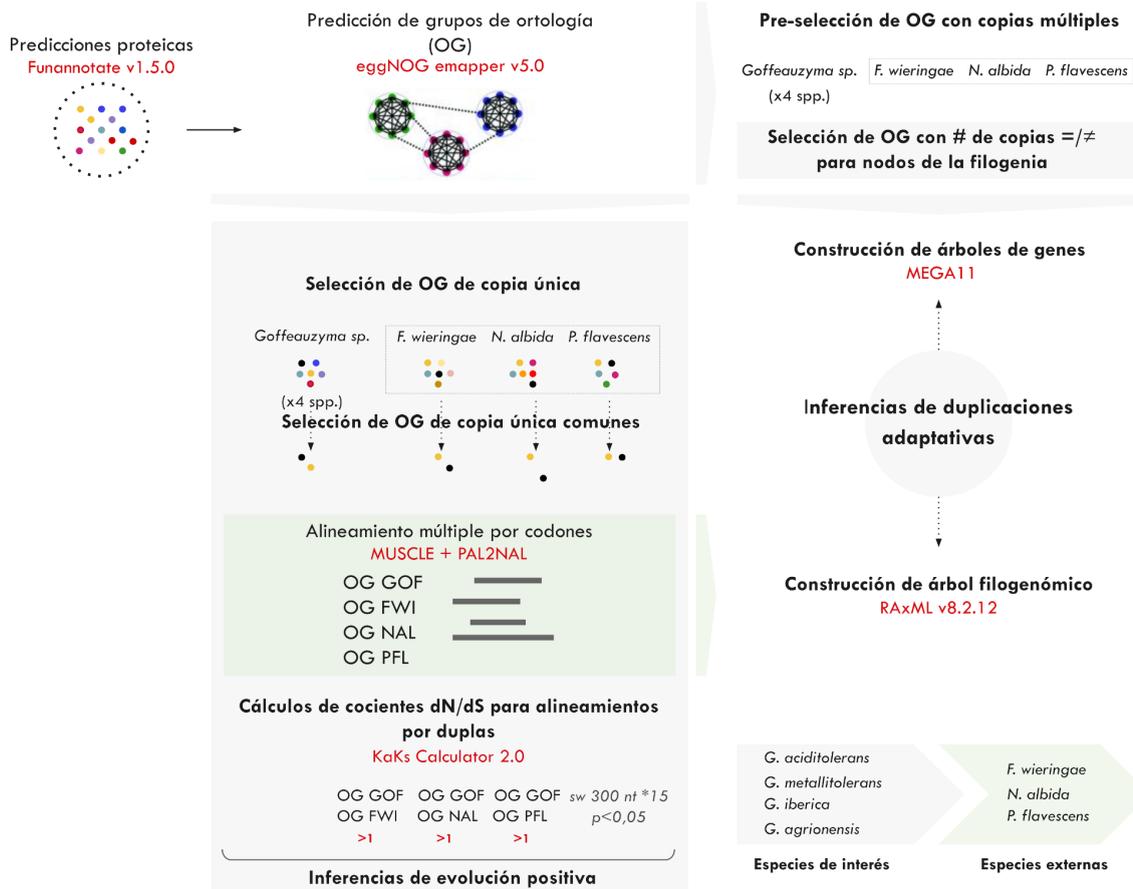


Figura 3.2.1. Esquema de trabajo para la comparación de las especies de interés respecto de aquellas tomadas como grupo externo. Abreviaturas: OG: grupo de ortología, GOF: especie de género *Goffeaulyma*, FWI: *F. wieringae*, NAL: *N. albida*, PFL: *P. flavescens*

3.2.5 Búsqueda de genes de interés

Se buscaron genes previamente caracterizados en levaduras, hongos y bacterias que tuviesen relación con la homeostasis, transporte y detoxificación de metales y con el mantenimiento del pH intracelular y eflujo de protones, a partir de secuencias descritas por otros autores. Las búsquedas se realizaron utilizando las *suites* tblastn y blastp sobre los genomas y proteomas derivados de todas las especies de orden Filobasidiales. Esta búsqueda con secuencias *query* específicas se sumó a un abordaje más generalizado consistente en la minería de funciones de interés dentro de los grupos de ortología designados. Las secuencias *query* utilizadas para las distintas vías relevadas se detallan en las Tablas 3.2.2 y 3.2.3.

Tabla 3.2.2 Secuencias *query* utilizadas para las búsquedas de genes de interés de vías no enzimáticas de respuesta a estrés oxidativo

V I A	Gen de interés	Secuencia <i>query</i>	Referencia / base de datos	Método	
R O S · N O E N Z I M Á T I C A	síntesis de glutatión	GSH1 γ-glutamilcisteína sintetasa	NP_012434.1	NCBI	blastp
		GSH2 glutatión-sintetasa	NP_014593.1	NCBI	
	síntesis de carotenoides	idi isopentenil difosfato isomerasa	AAY33922.1	NCBI	blastp /tblastn
		FPS farnesil pirofosfatasa sintetasa	G01847_P	Bellora et al., 2016	
		crtE farnesil transferasa	AAO47570.1 G05006_P	NCBI Bellora et al., 2016	
		crtI fitoeno deshidrogenasa	ACI43098.1	NCBI	
		crtYB fitoeno beta-caroteno sintasa	BAA33979.1	NCBI	
	crtR NADP-citocromo P450 reductasa	AHW57996.1	NCBI		
	PC fitoquelatina sintasa	6 secuencias de levaduras	obtenidas de NCBI (Material Suplementario: *_PCS.faa)	blastp y presencia de dominio PF005023 (fitoquelatina sintasa)	
	MT metalotioneína	13 secuencias de levaduras y bacterias	obtenidas de NCBI (Material Suplementario: *_MT.faa)	blastp y presencia de dominios PFAM PF11403 (MT de levaduras); PF12809 (flia. MT eucariotas); PF00649 (unión a ADN en FT que activan la transcripción de MT)	

Para el relevamiento de metalotioneínas también se calculó la proporción de residuos cisteína por proteína a fin de detectar un uso por encima de lo usual, dado que, tanto en hongos filamentosos como en levaduras el porcentaje de residuos cisteína en las metalotioneínas es ~20 %, siendo el aminoácido más abundante en esas proteínas (Gutiérrez et al., 2019).

Tabla 3.2.3. Secuencias *query* utilizadas para las búsquedas de genes de interés de vías enzimáticas de respuesta a estrés oxidativo

V I A	Gen de interés	Secuencias <i>query</i> / OG	Referencia / base de datos	Método
R O S . E N Z I M Á T I C A	superóxido dismutasas SOD	Cu G04828 Mn G03640	Bellora et al., 2016	blastp
	catalasas	G01478+G04411, G05570, G05702	Bellora et al., 2016	blastp /tblastn
	GRX glutarredoxinas	ENOG503P6QI, ENOG503NXX9, ENOG50P2U5, ENOG503P72N; PF00462	Base de datos de eggNOG v5.0 (db: taxa level 4751, 15-Sep-2018) http://eggnog5.embl.de/#/app/downloads	búsqueda de grupos de ortología asociados a las funciones en la base de datos de eggNOGv5.0
	GPX glutatión peroxidasa	ENOG503P24F		
GST glutatión transferasa	ENOG503P3EX, ENOG503NTZJ, ENOG503NZ93, ENOG503P2DJ, ENOG503NZ0D, ENOG503NYFG; PPF02798, PF00043			

Además de las búsquedas por homologías, se relevó la presencia y número de copias de grupos de ortología que, en base a la bibliografía, podrían estar vinculados con mecanismos de adaptación a las condiciones extremas estudiadas. Esto se realizó previa selección de grupos de ortología pertinentes de la base de datos de eggNOG v5.0 para hongos (db: *taxa level 4751*, 15-Sep-2018). De esta manera, se relevaron: 23 proteínas de resistencia/transportadores de cobre, 6 proteínas de resistencia/transportadores de zinc, 5 proteínas de resistencia/eflujo de arsénico, 3 proteínas de resistencia/transportadores de cadmio, 28 proteínas relacionadas con el transporte/gradiente de protones, 26 canales iónicos y 9 transportadores de cationes (eflujo/difusión).

3.3 RESULTADOS

3.3.1 Secuenciación y ensamblado de genomas de especies de *Goffeauzyma*

Los genomas de las 4 especies de género *Goffeauzyma* mostraron tamaños entre ~16 y ~25 Mpb, con coberturas genómicas entre 44 y 127x. Las longitudes de las 3 congéneres de MSD fueron superiores a las del grupo externo, mientras que *G. agrionensis* fue la especie con el genoma más pequeño de todos (27-33 % menor). Comparaciones con genomas del orden Filobasidiales derivados del Capítulo 2 y depositados en la base de datos de NCBI (N = 16, consulta realizada en julio de 2023, datos incluidos como Material Suplementario, Tabla S3.3.1) revelaron que la especie de Río Agrio es además la más pequeña de todas las disponibles hasta la fecha. El grado de completitud de los ensamblados, en base a la representatividad de los genes BUSCO de Basidiomycota, varió entre ~84 y ~95 % y se obtuvieron números comparables de secuencias codificantes para las cuatro especies (~5700 a ~5900), pero menores (~20 %) a los predichos para el grupo externo (~6400 a ~8100). La comparación de estos estadísticos con los disponibles para secuencias del orden Filobasidiales (N = 12, consulta realizada en julio de 2023, datos incluidos como Material Suplementario, Tabla S3.3.1) permitió ver que la cantidad de secuencias codificantes para el clado ARD es, en promedio (5778), un 13 % más pequeña que el promedio del orden (6645). Cerca del 80 % de los modelos predichos para cada especie fue clasificado en grupos de ortología de acuerdo a la base de datos de hongos de eggNOG. En estos grupos, 5 % albergaron múltiples copias (≥ 2), mientras que dicha proporción en el grupo externo varió entre 7 y 9 %. Los principales resultados derivados del proceso de ensamblado y anotación se listan en la Tabla 3.3.1, donde se incluyen también los correspondientes a las especies del grupo externo.

Tabla 3.3.1. Resumen de estadísticos del ensamblado y anotación de las cuatro especies del género *Goffeauzyma* y las tres especies hermanas estudiadas

Especie	<i>Goffeauzyma</i>				Grupo externo		
	<i>aciditolerans</i>	<i>metallitolerans</i>	<i>iberica</i>	<i>agrionensis</i>	<i>F. wieringae</i>	<i>N. albida</i>	<i>P. flavescens</i>
Coverage promedio	60x	44x	45x	127x	277x	95x	45x
Longitud (Mpb)	25,09	25,07	22,97	16,72	19,78	20,77	22,64
# contigs >2kpb	1196	879	826	2204	35	161	550
N50 (kpb)	315	455	527	9,33	952	332	90
% completitud*	86,4	87,6	87,5	88,2	84,2	87,9	95,2
# CDS	5707	5715	5727	5961	6448	6455	8153
% CDSs con asignación de OGs**	83,9	83,5	84,9	82,5	80,4	81,9	81,09
% OGs** multicopia	5,5	5,1	5,6	5,3	7,3	7,0	9,1

OG: grupo de ortología

* según clasificación en grupos de ortología de copia única incluidos en la base de datos BUSCO para Basidiomycota

** de acuerdo a la base de datos de hongos de eggNOG 5.0

3.3.2 Reconstrucción filogenómica

La reconstrucción filogenómica basada en grupos de ortología de copia única compartidos por las 7 levaduras de clase Tremellomycetes mostró una agrupación de las especies de género *Goffeauzyma* en un clado, denominado en este trabajo ARD. Dentro de éste, *G. agrionensis* fue la especie de divergencia más temprana, mientras que las 3 cepas provenientes de drenajes ácidos de minas conformaron un subclado más reciente, denominado MSD. *G. aciditolerans* y *G. metallitolerans* fueron las últimas especies en divergir, conformando el clado Tolerans. Las relaciones entre estas especies y el grupo externo se presentan en la Figura 3.3.1

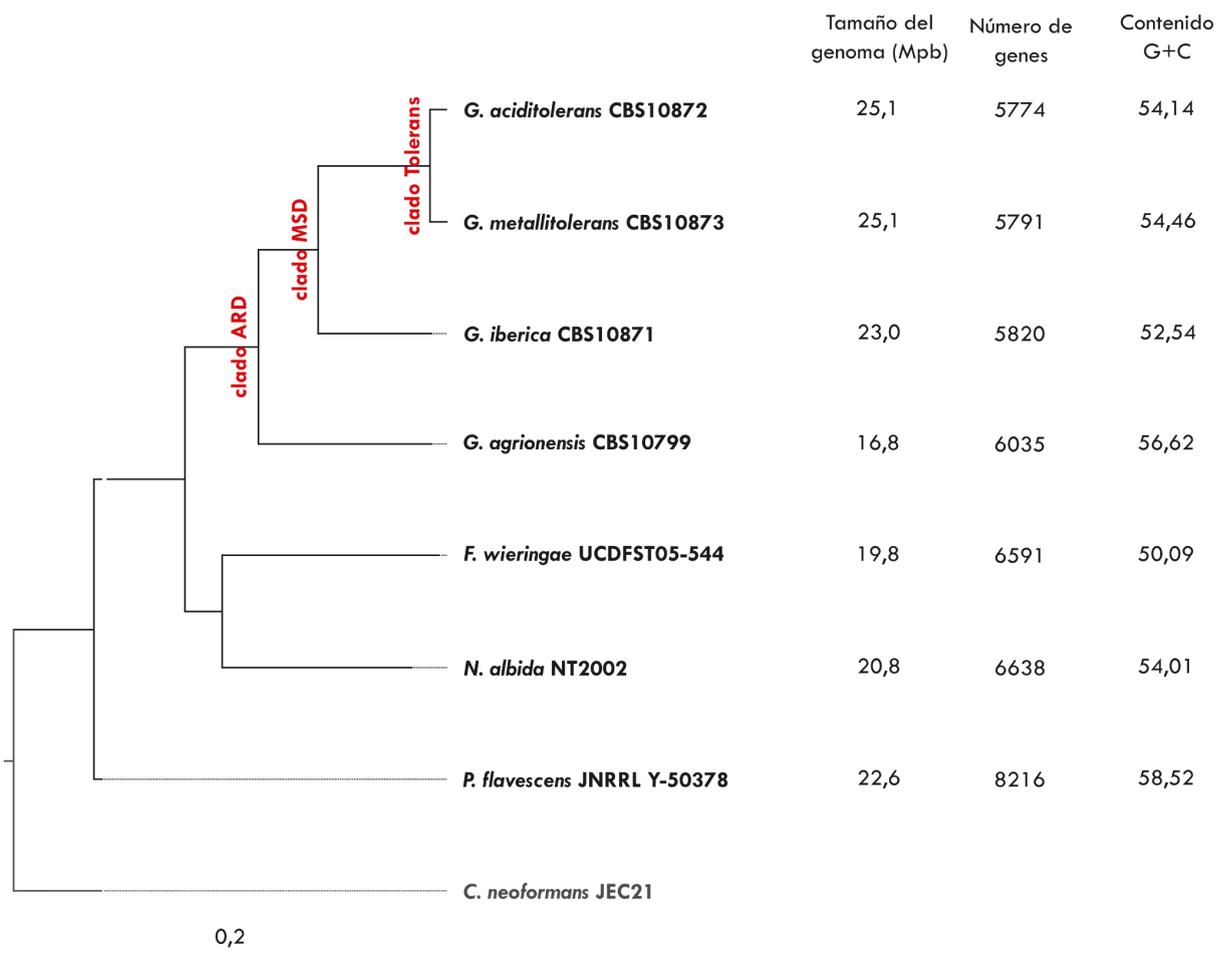


Figura 3.3.1. Árbol filogenómico de los organismos analizados computado a partir de las secuencias proteicas de 2031 grupos de ortología de copia única compartidos entre todos. Construido empleando RAxML v. 8.2.12 con el modelo de sustitución aminoacídica PROTGAMMALG4M. Nodos sustentados por un *bootstrap* de 100 %. La especie *C. neoformans* se utilizó para establecer la raíz del árbol

3.3.3 Grupos de ortología y asignaciones funcionales

El análisis comparativo de ortología para las siete especies de clase Tremellomycetes permitió evidenciar que 2966 anotaciones (aproximadamente el 70 %) de los grupos de ortología son comunes a todas las especies. Un total de 351 grupos (~7 %) se encontraron únicamente en el género *Goffeauzyma*, ya fuese de manera exclusiva a alguno de los genomas, o bien de forma conjunta a varios o todos ellos. En el clado ARD, 37 (~0,8 %) fueron comunes a todas las especies, mientras que en los subclados MSD y Tolerans, se observaron 24 y 33 grupos de ortología compartidos, respectivamente. Entre 37 y 83 grupos, por otro lado, se clasificaron como únicos de cada especie de ARD, no hallándose en sus congéneres ni en las especies hermanas.

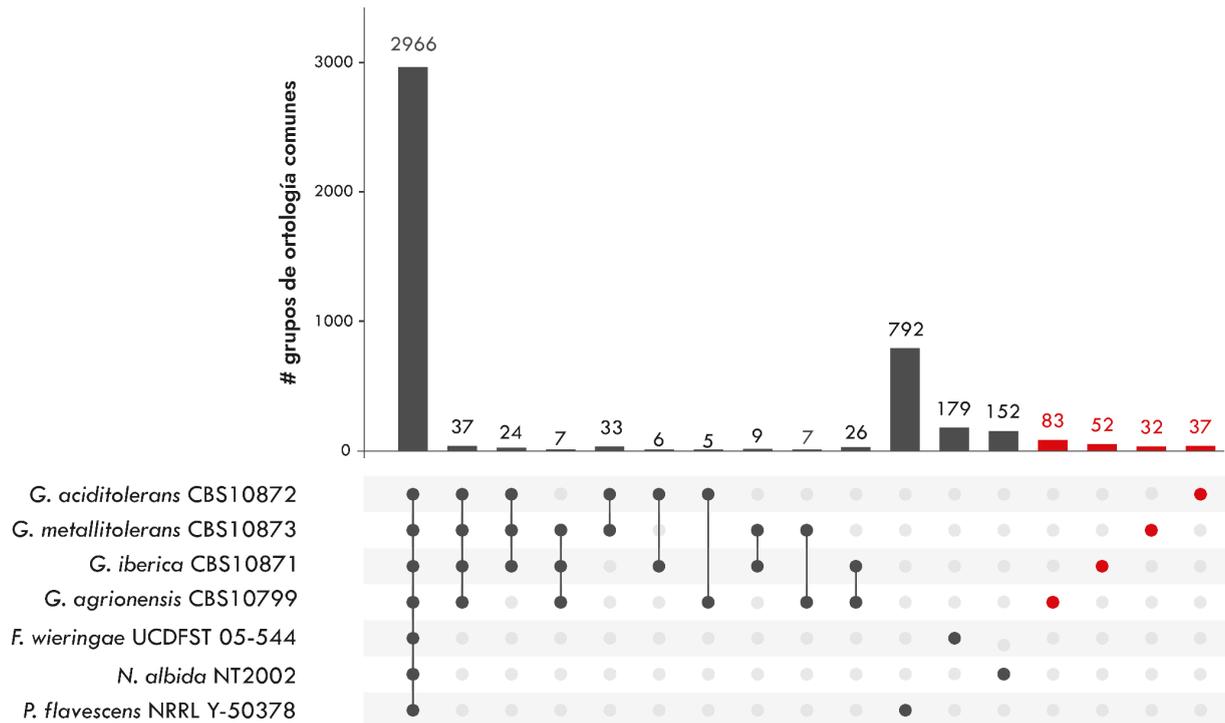


Figura 3.3.2. Análisis de ortología para las 4 especies de interés y las 3 especies del grupo externo. El gráfico de barras verticales muestra el número de grupos de ortología comunes/exclusivos a la/es especie/s detalladas en el gráfico de interacciones. Los puntos oscuros rellenos denotan presencia y los vacíos (gris claro) indican ausencia de grupos de ortología en cada una de las especies. Se muestran únicamente las intersecciones más relevantes a los fines de este trabajo. Gráfico visualizado con UpSetR (Conway et al., 2017).

La información se resume en la Figura 3.3.2 y en las Tablas 3.3.2, 3.3.3 y 3.3.4 se detallan los grupos de ortología encontrados exclusivamente en cada uno de los clados, con sus respectivas categorías funcionales (COG). En negrita se destacan aquellos que, en base a la bibliografía, podrían estar implicados en mecanismos de tolerancia. Los mismos se discutirán más adelante.

La categoría funcional COG referida a anotaciones involucradas en el Transporte y metabolismo de carbohidratos fue la que presentó un número de anotaciones mayormente representada en todos los clados. Mientras que entre 43 y 54 % de los grupos de ortología no tienen asignación funcional (Función desconocida).

Tabla 3.3.2 Grupos de ortología que se encontraron únicamente en las especies del clado ARD (N = 37). Las IDs separadas por “:” corresponden a predicciones fragmentadas. En **negrita** se destacan los grupos con funcionalidad relevante, discutidos más abajo

Grupo de ortología	ID gen por especie	Categoría funcional COG	Descripción funcional	
ENOG503P642	Ga_2834,Gm_5312,Gi_1592,Gg_5418	2,7 % Transducción de señales	RhoGAP	
ENOG503Q66Y	Ga_3589,Gm_1552,Gi_2128,Gg_5203	2,7 % Biogénesis de membrana / pared celular	Dom. RmlD	
ENOG503Q48C ENOG503NY2M	Ga_1139,Gm_5467,Gi_2573,Gg_1970 Ga_4989,Gm_5367,Gi_5590,Gg_2998	5,4 % Tráfico intracelular, secreción & transporte vesicular	ERMES/MDM Transportador	
ENOG503Q3UA	Ga_1906,Gm_1852,Gi_1526,3186, Gg_3758	2,7 % Chaperonas, modificaciones post-traduccionales, recambio de proteínas	Proteína <i>subtilisin-like</i>	
ENOG503NXPJ	Ga_4128:4156,Gm_4122, Gi_4170,Gg_1106,2757	24,3 % Transporte & metabolismo de carbohidratos	Fosfatasa ácida	
ENOG503NYXT ENOG503NTVN	Ga_3433,Gm_1866,Gi_2583,Gg_1715 Ga_4033,Gm_3739,Gi_2069:3254, Gg_5537		Amido-hidrolasa Carboxilesterasa tipo B	
ENOG503PCVC ENOG503Q5BK ENOG503PCS5 ENOG503NV11	Ga_0643,Gm_3217,Gi_5015,Gg_5306 Ga_0407,Gm_2883,Gi_2694,Gg_5886 Ga_2136,Gm_0477,Gi_4948,Gg_1168 Ga_3580,Gm_5679,Gi_4100,Gg_2213		Carboxilesterasa tipo B Glicosil-hidrolasa Fucosil-transferasa Glicosil-hidrolasa	
ENOG503NUCP ENOG503P42S	Ga_5190,Gm_1588,Gi_0790,Gg_3408 Ga_2029,Gm_5119,Gi_5797,Gg_4119		MFS MFS	
ENOG503NZIY	Ga_1836,Gm_3775,Gi_2494,Gg_0723		2,7 % Transporte & metabolismo de aminoácidos	Acetil-transferasa
ENOG503P29S ENOG503NUT6	Ga_4928,Gm_3839,Gi_5114,Gg_1379 Ga_2276,Gm_1654,Gi_1228,1229, Gg_2961		5,4 % Transporte & metabolismo de nucleótidos	Citosina.deaminasa Purín nucleósido permeasa
ENOG503NXTP	Ga_5566,Gm_3987,Gi_3467,Gg_1427		2,7 % Transporte & metabolismo de coenzimas	Farnesil-transferasa
ENOG503NY1X	Ga_4169,Gm_1430,Gi_1221,Gg_4824		2,7 % Transporte & metabolismo de lípidos	α/β hidrolasa
ENOG503NVJX	Ga_4806,Gm_4288,Gi_1842,Gg_004698		2,7 % Transporte & metabolismo de iones inorgánicos	Purina-citosina permeasa
ENOG503NYC2	Ga_4806,Gm_4288,Gi_1842,Gg_004698		2,7 % Transporte & catabolismo, biosíntesis de metabolitos secundarios	Flavin-monooxigenasa
ENOG503NYVJ ENOG503P2G6 ENOG503P6Z7 ENOG503P8Z9 ENOG503PBW9 ENOG503PBT4 ENOG503PIAV ENOG503Q50P ENOG503NUBH ENOG503NVGS ENOG503NZ10 ENOG503P5UH ENOG503NXOD ENOG503P155 ENOG503PAQ0 ENOG503NYGF	Ga_2905,Gm_3550,Gi_0225,Gg_2807 Ga_4599,Gm_1281,Gi_2908,Gg_5945 Ga_2971,Gm_2141,Gi_4938,Gg_5691 Ga_4954,Gm_3165,Gi_4365,Gg_4432 Ga_3748,Gm_4031,Gi_2562,Gg_0204 Ga_5079,Gm_3423,Gi_0852,Gg_5414 Ga_2181,Gm_1981,Gi_2691,Gg_1707 Ga_0321,Gm_3951,Gi_2297,Gg_3423 Ga_0785,Gm_3659,Gi_2205,Gg_3881 Ga_0801,Gm_4206,Gi_3294,Gg_5970 Ga_3879,Gm_4064,Gi_1668,Gg_5005 Ga_0041,Gm_3458,Gi_1897,Gg_4083 Ga_5158,Gm_3904,Gi_2327,Gg_0429 Ga_5329,Gm_2868,Gi_4919,Gg_2537 Ga_5611,2909,Gm_0769,4340, Gi_0793,Gg_2907 Ga_1254,Gm_4233,Gi_1079,Gg_2985	43,3 % Función desconocida	'Aminoácido' Transp. de aminoácidos Proteína CCH Zn-fg DUF1960 Flavín-nucleótido Glicosil-transferasa Serín-peptidasa flia.S33 Prot. con dom. WD	

Tabla 3.3.3. Grupos de ortología que se encontraron únicamente en las especies del clado MSD (N = 24). En **negrita** se destacan los grupos con funcionalidad relevante, discutidos más abajo.

Grupo de ortología	ID gen por especie	Categoría funcional COG	Descripción funcional
ENOG503P1AP	Ga_2152,Gm_1590,Gi_0788	8,3 % Transcripción	Factor de transcripción
ENOG503P0PK	Ga_5180,Gm_0762,Gi_2727		Factor de transcripción
ENOG503Q57P	Ga_4889,Gm_2163,Gi_0345	4,2 % Transducción de señales	RhoGAP
ENOG503PCK9	Ga_2579,Gm_4118,Gi_1131	4,2 % Mecanismos de defensa	β -lactamasa
ENOG503PC9A	Ga_0231,Gm_4925,Gi_4527	12,5 % Transporte & metabolismo de carbohidratos	MFS
ENOG503NVFS	Ga_1075,Gm_1834,Gi_2467		Glucosamina-6-fosfato
ENOG503PONU	Ga_5244,Gm_2772,Gi_5370		Gluconolactonasa
ENOG503NWZ7	Ga_0236,Gm_2964,Gi_4217	4,2 % Transporte & metabolismo de aminoácidos	Linoleato diol sintasa
ENOG503P7KN	Ga_0232,Gm_4367,Gi_4528	8,3 % Transporte & metabolismo de lípidos	Esterol desaturasa
ENOG503PBHH	Ga_2446,Gm_2911,Gi_4756		α/β hidrolasa
ENOG503P16V	Ga_0234,Gm_4368,Gi_5244	4,2 % Transporte & catabolismo, biosíntesis de metabolitos secundarios	Familia citocromo P450
ENOG503Q672	Ga_0341,Gm_4917,Gi_1295	54,1 % Función desconocida	DUF1691
ENOG503PAK0	Ga_0684,Gm_3328,Gi_3188		familia α/β hidrolasa
ENOG503NX28	Ga_2087,Gm_0693,Gi_0202		MFS
ENOG503P59A	Ga_2273,Gm_1651,Gi_5165		familia citocromo b5
ENOG503NVJR	Ga_2415,5713,Gm_4814, Gi_0523		fosfatasa ácida
ENOG503P4UB	Ga_2491,Gm_5442,Gi_5363		fosfohidrolasa dom. 'HD'
ENOG503P6RE	Ga_3430,Gm_1863,Gi_0924		
ENOG503P2J7	Ga_3609,Gm_3059,Gi_5443		
ENOG503NYSX	Ga_1113,Gm_1570,Gi_4571		
ENOG503PEJJ	Ga_1352,4433,4793,Gm_3990,		
ENOG503NVSX	Gm_4245,Gi_2944:5450,5161, 5215		
ENOG503P07X	Ga_2841,Gm_3809,Gi_4486		
ENOG503P07X	Ga_4034,4987,Gm_1406,3738,		
ENOG503P07X	Gi_4662		
ENOG503NWB2	Ga_4266,Gm_0396,Gi_0761		

Tabla 3.3.4. Grupos de ortología que se encontraron únicamente en las especies del clado Tolerans (N = 33). En **negrita** se destacan los grupos con funcionalidad relevante, discutidos más abajo.

Grupo de ortología	ID gen por especie	Categoría funcional COG	Descripción funcional
ENOG503Q5YJ	Ga_1758,Gm_1542	3 % Transcripción	Motivo rico en Cys luego de subconjunto de dominios SET
ENOG503P2F6	Ga_4884,Gm_1207	3 % Estructura y dinámica e la cromatina	Histona h3
ENOG503Q4HW	Ga_4065,Gm_2271	3 % Ctrl del ciclo cel., división, partición. de cromosomas. / Tr. de señales / Citoesqueleto	Familia septin GTPasa
ENOG503Q2YS	Ga_0249,Gm_3395	3 % Mecanismos de defensa	Epimerasa dpte. de NAD
ENOG503PBR2	Ga_0173,Gm_2509	3% Transducción de señales	Transportador de oligop.
ENOG503NYQE	Ga_0880,Gm_4028	3% Biogénesis de membrana / pared celular	Udp-glucosa 4-epimerasa
ENOG503P3Q3	Ga_2148,Gm_1594	3 % Chaperonas, modificaciones post-traduccionales, recambio de proteínas	Tiorredoxina
ENOG503PAJ6 ENOG503NX80	Ga_5217,Gm_2194 Ga_3290,Gm_3534	6 % Producción & conversión de energía	Oxidorreductasa Complejo ubiq. cito. C reduct.
ENOG503P7BZ ENOG503Q3G3 ENOG503Q4CA	Ga_1704,5085,5086, Gm_0478,1277 Ga_1536,Gm_5740 Ga_3553,Gm_1756	9,1 % Transporte & metabolismo de carbohidratos	O-fucosil-transferasa N-acetilglucosamina quinasa Epimerasa/deshidratasa
ENOG503P1AJ	Ga_4966,Gm_4821	3 % Transporte & metabolismo de nucleótidos / Transporte & metabolismo de carbohidratos	Bis(5'-adenosil)-trifosfatasa
ENOG503Q55U	Ga_5238,Gm_5490	3 % Transporte & metabolismo de lípidos	Delta-isomerasa
ENOG503NYCB	Ga_4212,Gm_4384	3 % Transporte & metabolismo de iones inorgánicos	Sulfatasa
ENOG503NZ3D	Ga_4235,Gm_4328	3 % Transporte & metabolismo de iones inorgánicos / Bios. transp & catab. de metabolitos sec.	Oxidasa dpte. de NADPH
ENOG503Q4FB	Ga_5129,Gm_4346	3 % Transporte & catabolismo, biosíntesis de metabolitos secundarios	Dominio KR
ENOG503P4EI ENOG503P8BP ENOG503PCQ8 ENOG503PJTN ENOG503Q52R ENOG503PEAR ENOG503P3R8 ENOG503PNI4 ENOG503P9C1 ENOG503NZFZ ENOG503NX36 ENOG503P5KM ENOG503P6IH ENOG503NZUX ENOG503P1GY ENOG503PEOG	Ga_2050,Gm_3864 Ga_1341,Gm_2476 Ga_4190,Gm_5475 Ga_4186,Gm_1832 Ga_1017,Gm_1093 Ga_2949,Gm_5744 Ga_1950,Gm_4466 Ga_4751,Gm_0455 Ga_2461,Gm_5079 Ga_2590,Gm_2221 Ga_2364,Gm_4053 Ga_3533,Gm_5037 Ga_0806,Gm_4201 Ga_3915,Gm_5689 Ga_4873,Gm_2040 Ga_3882,Gm_5483	43,3 % Función desconocida	acetil-transferasa (GNAT) DUF4396 transport. de Fe de ↓ afin regul. de membrana in. mito

3.3.3.1 Análisis de duplicaciones en los grupos funcionales multicopia

El curado manual de las secuencias correspondientes a los 480 grupos funcionales para los cuales había mayor número de copias en alguna especie, permitió identificar indudablemente 319 grupos con más de una copia, mientras que 161 eran producto de

errores de anotación (genes partidos anotados como 2, cuando en realidad es sólo 1 y algún intrón es mal interpretado como sitio de finalización). Del total de los grupos analizados, un cuarto (N = 87) presentaron múltiples copias en las 4 especies de ARD, 25 sólo en el clado MSD, 34 en el clado Tolerans y entre 7 y 53 en alguno de los 4 genomas. *G. agrionensis* resultó la especie con mayor número de genes en copias múltiples, aún teniendo el genoma más pequeño del clado. A partir de los análisis de inferencia de eventos de duplicación génica, se pudieron identificar 27 expansiones (duplicaciones) ocurridas en alguno de los nodos del clado ARD, siendo 8 las duplicaciones ocurridas en el nodo común a todas las especies de ARD, 2 en el clado MSD y 7 en el clado Tolerans. *G. aciditolerans* fue la única especie que no presentó ninguna duplicación propia, mientras que para el resto se verificaron entre 2 y 5. En la Figura 3.3.3 se muestra la distribución de los grupos funcionales con múltiples copias estudiados para las 4 especies. Entre puntos se indican las expansiones por duplicaciones detectadas en cada uno de los nodos de la filogenia del grupo.

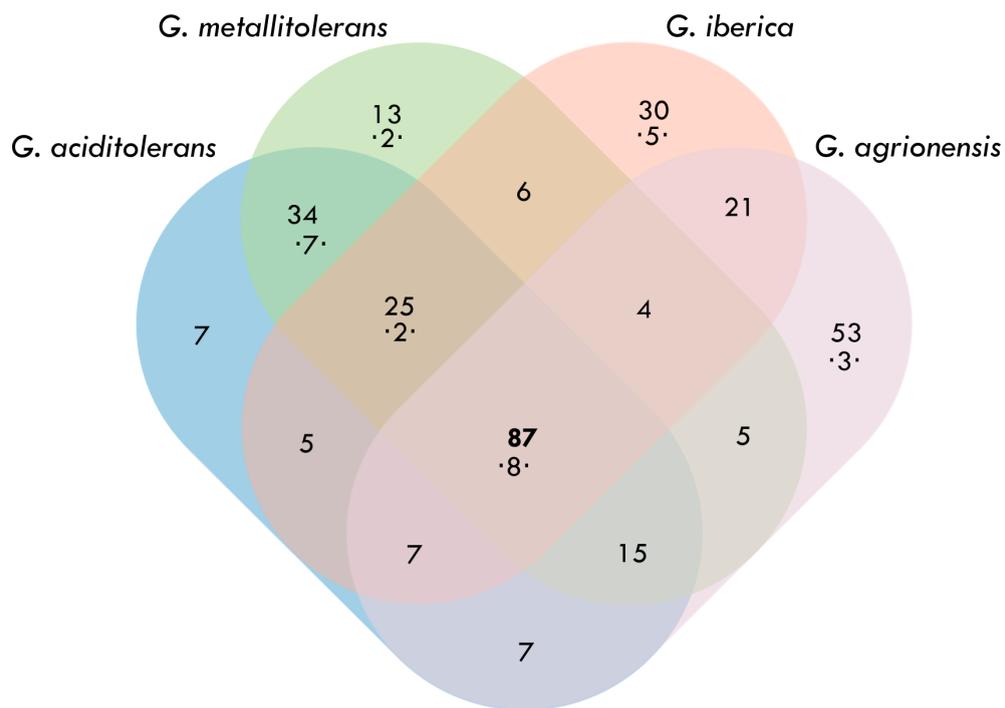


Figura 3.3.3. Distribución de grupos funcionales multicopia entre las 4 especies de género *Goffeauzyma*. Valores entre puntos: expansiones génicas evidenciadas mediante la inferencia de eventos de duplicaciones.

En 89 grupos funcionales, el número de copias en todas o alguna/s de las especies del género *Goffeauzyma* supera/n a los respectivos en todas las especies del grupo externo. Para estos grupos se analizaron sus descripciones funcionales a fin de encontrar aquellos que pudiesen estar vinculados con adaptaciones a las condiciones ambientales. Los resultados se presentan en el mapa de calor ilustrado en la Figura 3.3.4, donde se muestra el número de copias por grupo de ortología para cada una de las especies. Los bloques con números de copia homogéneos para el clado ARD o subclados dentro de él (enumerados en la figura) y los grupos de ortología de interés (cuyas identidades se resaltan allí también), se desarrollan a continuación. En la figura, además se señalan (destacadas en rojo) las expansiones génicas por duplicaciones, que se abordarán más adelante (ver “Expansiones génicas por duplicación dentro del clado”).

En el bloque n° 1, todo el clado ARD tiene el doble o más copias que las especies de referencia más cercanas (*F. wieringae* y *N. albida*) y supera en una o dos a la más distante (*P. flavescens*). Este bloque está comprendido por el grupo de ortología ENOG503NXMN, perteneciente a la superfamilia de transportadores con dominio MFS (*Major Facilitator Superfamily*), responsables de transportar un amplio espectro de sustratos, tanto a favor, como en contra del gradiente de concentración. La inferencia de duplicaciones permitió identificar en este bloque tres eventos de duplicación ocurridos en distintos nodos de la filogenia del grupo (ver “Expansiones génicas por duplicación dentro del clado”). De esta forma la redundancia se debe principalmente a eventos ocurridos en el clado. El bloque n° 2, para el cual todo el clado ARD duplica a las especies externas, se compone de 7 grupos de ortología pertenecientes a categorías funcionales variadas. Entre ellos, se destaca la presencia de ENOG503NVR2, con dominio *Vhs* y relacionado con el proceso biológico de transporte intracelular de proteínas inducido por cobre (GO:0006886). Por último, el bloque n° 3 está comprendido por grupos para los cuales el clado Tolerans tienen un número de copias mayor a los de las especies externas. En este bloque, se destacan los grupos ENOG503NUJY, una glutamato decarboxilasa involucrada en la respuesta a estrés oxidativo y ENOG503NXKX, una ATPasa tipo P (proteína que acopla la hidrólisis de ATP al bombeo de iones través de la membrana en contra de su concentración), cuya inferencia de duplicaciones indica que se expandió dentro del clado. Dentro de este bloque, se identificaron otras 2 expansiones por duplicación, que se mencionarán más adelante (ver “Expansiones génicas por duplicación dentro del clado”).

Por fuera de estos bloques principales se destacan los grupos: ENOG503NXPJ y ENOG503NVJR, relacionados con fosfatasas ácidas, enzimas que catalizan reacciones de hidrólisis a bajo pH y que se encuentran ausentes en todas las especies externas, y presentan dos copias en *G. aciditolerans* y una en el resto de las *Goffeauzyma* (salvo en *G. agrionensis*, donde se encontró 1 copia para el primer grupo pero no pudo encontrarse el segundo); ENOG503PV7J y ENOG503PKJU, ambos con dominio de función desconocida DUF563, el primero de ellos encontrado únicamente en la especie externa más distante (*P. flavescens*) y con una duplicación verificada en *G. iberica* (ver “Expansiones génicas por duplicación dentro del clado”); ENOG503Q4JQ, perteneciente a la familia citocromo P450 (CYP450), responsables de un amplio rango de reacciones, entre ellas la oxidación de compuestos para su posterior detoxificación; ENOG503NVHV, un transportador con dominio MFS y ENOG503P237, un transportador de sideróforos (compuestos quelantes de hierro) presente en dos copias para *G. agrionensis* y una para *G. iberica* y la especie externa *F. wieringae*.

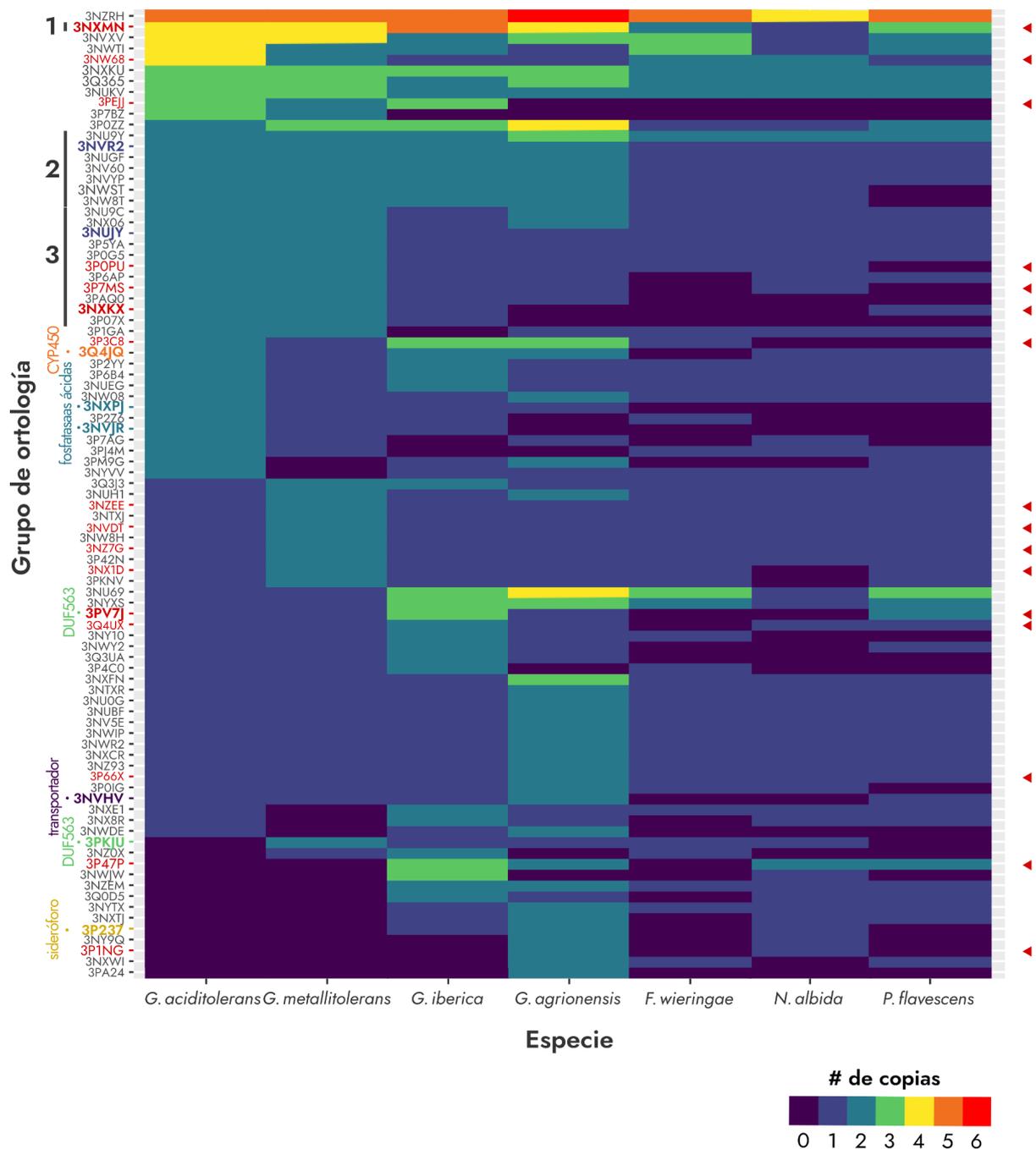


Figura 3.3.4. Mapa de calor expresando el número de copias para los grupos de ortología en los que una o más especies del género *Goffeauzyma* supera a las especies externas. Las identidades de los grupos de interés resaltados en el cuerpo del texto se muestran en **negrita** y los bloques discutidos, numerados. Las flechas rojas señalan a los grupos para los que se corroboraron expansiones génicas por duplicaciones en el clado y sus identidades se indican con el mismo color. Se utiliza la nomenclatura ENOG50~(5 caracteres) para facilitar la visualización.

Expansiones génicas por duplicación dentro del clado

El análisis de los 319 árboles de genes permitió inferir 27 grupos de ortología con al menos una expansión ocurrida por un evento de duplicación dentro del clado (~8 % de todos los casos analizados), con ~30 % involucrando a todo el clado ARD (N = 8), ~7 % al subclado MSD (N = 2), ~26 % al subclado Tolerans (N = 7) y cerca de la mitad (~37 %) a las especies en particular: *G. metallitolerans* (N = 2, ~7 %), *G. iberica* (N = 5, ~19 %) y *G. agrionensis*, (N = 3, ~11 %). Entre ellos, en base a su descripción funcional se destacan: ENOG503NUYQ y ENOG503NXMN con presencia de dominios MFS, ambos con expansiones en todo ARD; y ENOG503PIXM con evidencia de expansiones en el clado MSD, con dominio de función desconocida DUF563; ENOG503NXXKX, ENOG503NW68 y ENOG503NXJD con expansiones en el clado Tolerans, el primero correspondiente a una ATPasa tipo P y los otros dos ,homólogos a transportadores (el último, con presencia de dominio MFS); y ENOG503PV7J, con evidencia de expansión en *G. iberica*, correspondiente un grupo con dominio de función desconocida DUF563. Estos resultados se muestran en la Tabla 3.3.5, donde se detallan los grupos de ortología con al menos una expansión, los nodos más basales donde se localizaron las mismas y las categorías y descripciones funcionales de cada uno. Los 27 árboles de genes correspondientes se incluyen como Material Suplementario (Tabla S3.3.5).

Tabla 3.3.5. Grupos de ortología para los que se verificaron expansiones en nodos de la filogenia del clado ARD. Se indica en qué nodo se localizaron y sus respectivas categorías y descripciones funcionales. En **negrita** se destacan los grupos de mayor interés.

	Grupo de ortología	Categoría funcional COG	Descripción funcional
A R D	ENOG503NW3E	Transcripción	Regulador de nitrógeno
	ENOG503P5JJ	Tráfico intracelular, secreción & transporte vesicular	Deshidrogenasa de celobiosa
	ENOG503NX3C	Producción & conversión de energía	Oxidorreductasa (Zn bind.)
	ENOG503NUUF ENOG503NX4R	Transporte & metabolismo de carbohidratos	α-1,6-manosiltransferasa Alantoato permeasa
	ENOG503NUYQ	Transporte & metabolismo de aminoácidos	MFS
	ENOG503NXMN	Transporte & metabolismo de iones inorgánicos	MFS
	ENOG503P3C8	Función desconocida	
M S D	ENOG503Q4UX	Transporte & metabolismo de iones inorgánicos	Intercambiador Na ⁺ /H ⁺
	ENOG503PIXM	Función desconocida	DUF563
T o l e r a n s	ENOG503NZEE	Estructura nuclear	Proteína de matriz nuclear
	ENOG503NX1D	Chaperonas, modificaciones post-traduccionales, recambio de proteínas	Transamidasa GPI (GAA1)
	ENOG503POPU	Producción & conversión de energía / Transporte & metabolismo de coenzimas	Nitrato reductasa
	ENOG503NXJD	Transporte & metabolismo de carbohidratos	Transportador (dom MFS)
	ENOG503NW68 ENOG503NXKX	Transporte & metabolismo de iones inorgánicos	Transportador ATPasa tipo P
	ENOG503P7MS	Función desconocida	Ciclina
G. metallitolerans	ENOG503NZ7G	Citoesqueleto	Proteína asociada a microtúbulos
	ENOG503NVDT	Transporte & metabolismo de coenzimas	Familia PdxS SNZ
G. iberica	ENOG503PE71	Chaperonas, modificaciones post-traduccionales, recambio de proteínas	Proteína con dominio GXXXQXW
	ENOG503NYJ2	Transporte & metabolismo de carbohidratos	Glicosil-hidrolasa (flia. 79)
	ENOG503PV7J ENOG503PEJJ ENOG503P47P	Función desconocida	DUF563
	ENOG503P3C8	Transducción de señales	<i>PsbP-like protein</i>
G. agrionensis	ENOG503P1NG ENOG503P1ZZ	Función desconocida	

3.3.4 Análisis de cociente dN/dS

Los cocientes dN/dS computados para los 2337 grupos de ortología de copia única comunes a todas las especies del género *Goffeauzyma* y a las tres especies del grupo externo revelaron que, para las estimaciones sobre secuencia completa, únicamente tres tuvieron un cociente mayor a 1 en relación a una sola de las especies de referencia. El resto de los cocientes presentó una distribución similar a la observada en el capítulo 2 para *N. vishniacii*. El abordaje mediante *sliding windows* permitió identificar entre 12 y 16 grupos de ortología por especie (29 grupos distintos) con regiones comunes para las que el valor del cociente fue mayor a 1 frente a las 3 especies externas. Entre ellos, 4 grupos fueron comunes a todo el clado ARD; 1 a las especies de la Península Ibérica (MSD); 1 a las del clado Tolerans; 3 fueron compartidos por *G. aciditolerans*, *G. metallitolerans* y *G. agrionensis*; 2 por *G. aciditolerans* y *G. agrionensis*; 1 por *G. aciditolerans* y *G. iberica* y 1 por *G. metallitolerans* y *G. iberica*. Debido a que el enfoque de este trabajo consiste en la búsqueda de estrategias comunes de adaptación, todos los grupos compartidos por 2 o más especies (N = 14) fueron evaluados manualmente con JCoDA. Las regiones positivas para todas las especies verificadas por homología (HMMER y blast) no tuvieron correspondencia con ninguna secuencia depositada, por lo que se asume que son fragmentos mal anotados. Los resultados se muestran en las Figuras 3.3.5 (A) - 3.3.8 (A), donde se grafican las distribuciones del cociente promedio para las duplas de cada una de las especies del clado ARD respecto de las 3 especies del grupo externo y se indican entre paréntesis la cantidad de genes con regiones bajo supuesta presión selectiva, listados debajo de las distribuciones (Figuras 3.3.5 (B) - 3.3.8 (B)).

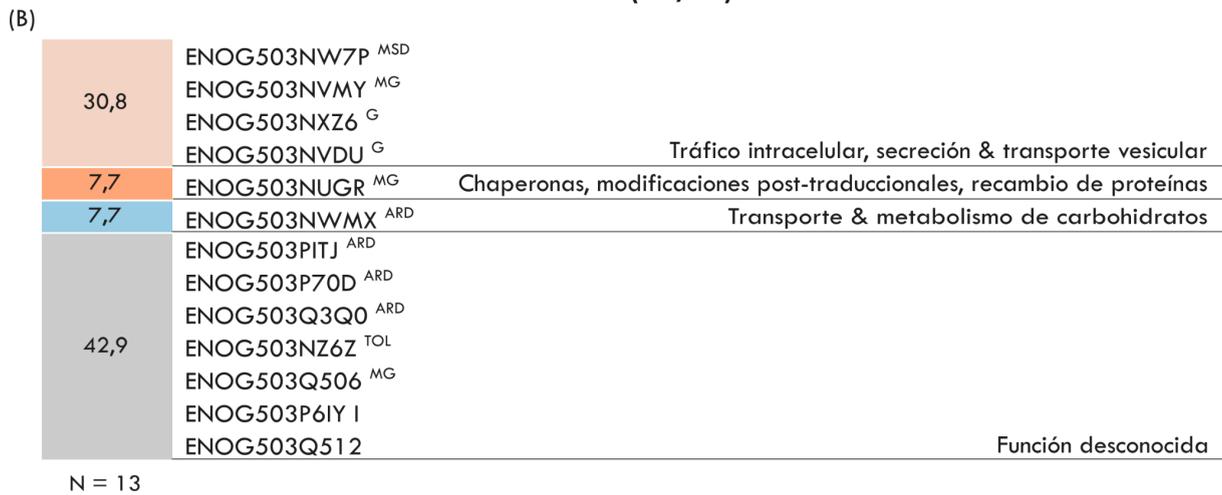
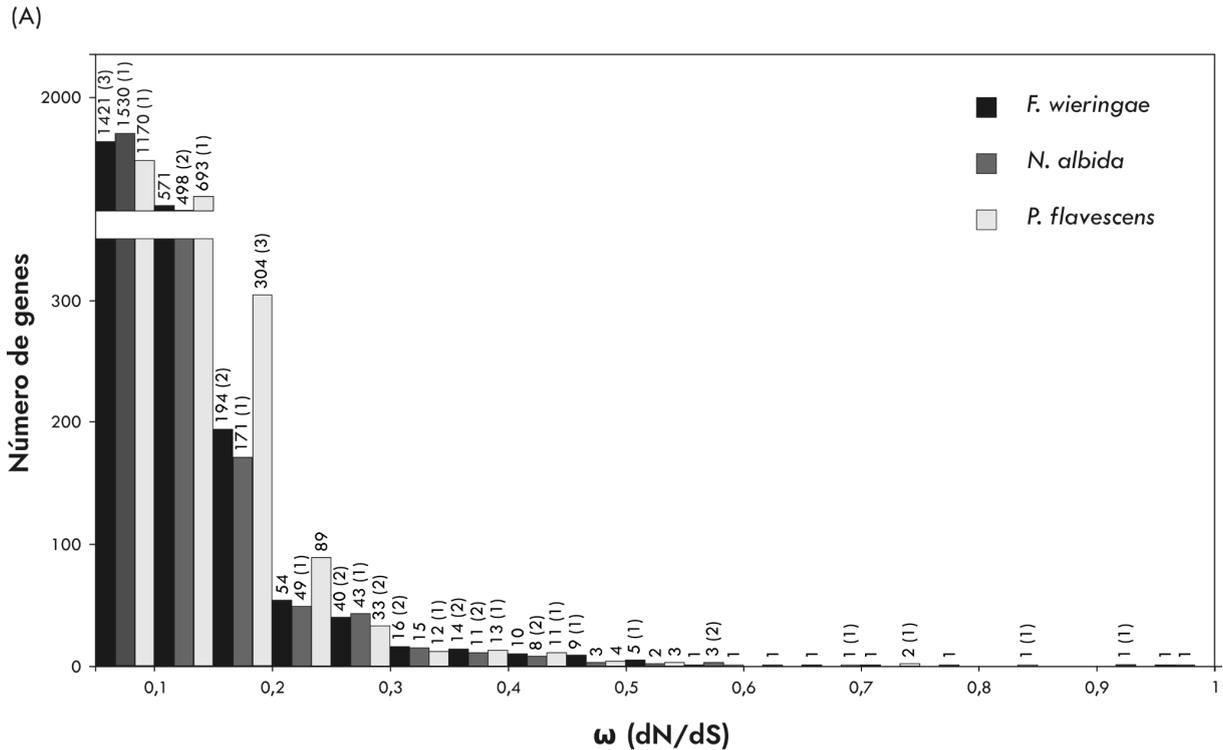
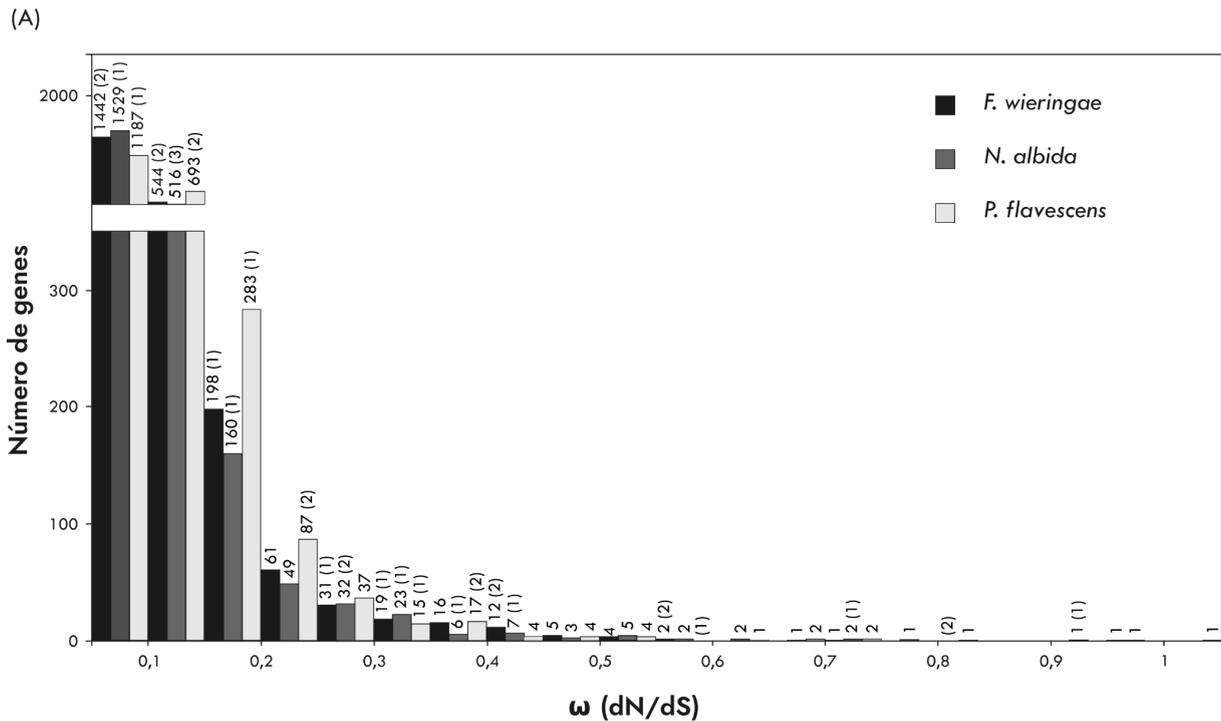


Figura 3.3.5. Análisis de dN/dS para grupos de ortología comunes a las siete especies (N = 2337). (A) Distribución de valores promedio de dN/dS para *G. aciditolerans* y las tres especies del grupo externo. Entre paréntesis: número de genes con dN/dS > 1 (pval <0,05) en alguna región, considerados bajo presión selectiva (N = 13). (B) Categorías funcionales (COG) de los genes bajo presión selectiva. Se indican los grupos que también presentaron señal de evolución positiva en: ^{ARD}ARD, ^{MSD}MSD, ^{TOL}Tolerans, ^M*G. metallitolerans*, ^I*G. iberica*, ^G*G. agrionensis*

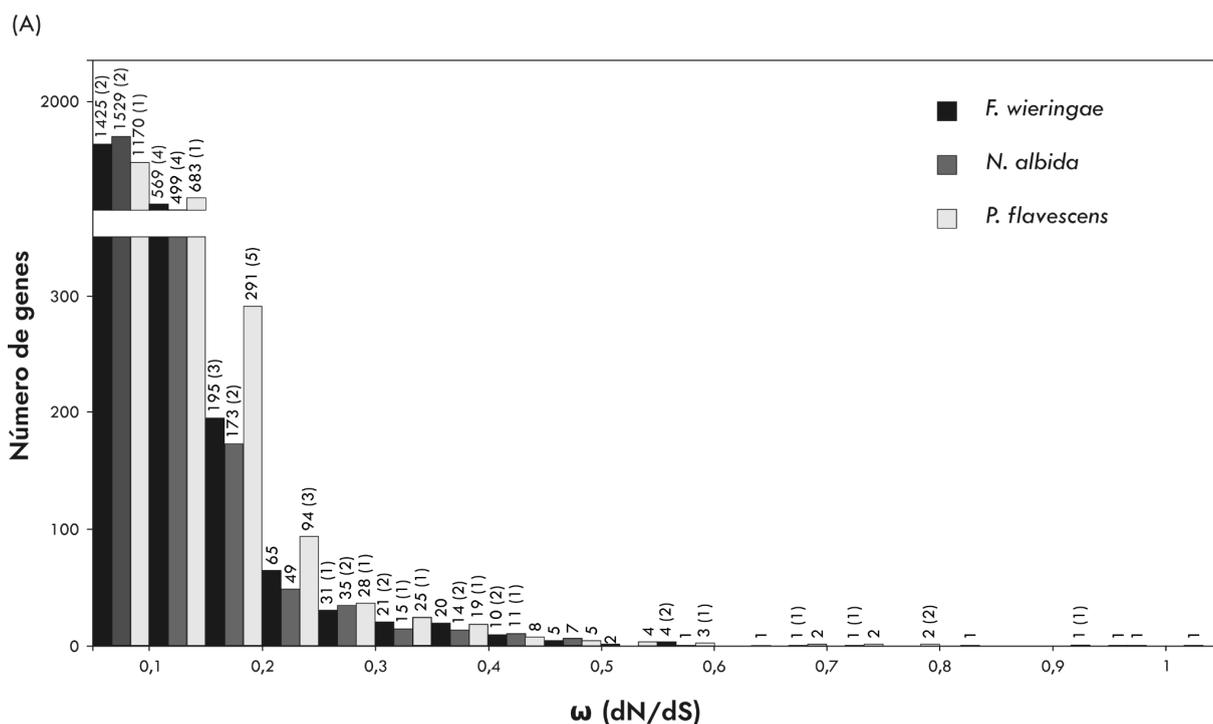


(B)

8,3	ENOG5033NTXG ^G	Transcripción
8,3	ENOG503NUGR ^{AG}	Tráfico intracelular, secreción & transporte vesicular
16,7	ENOG503NW7P ^{MSD} ENOG503NVMY ^{AG}	Chaperonas, modificaciones post-traduccionales, recambio de proteínas
8,3	ENOG503NWMX ^{ARD}	Transporte & metabolismo de carbohidratos
58,4	ENOG503PITJ ^{ARD}	Función desconocida
	ENOG503P70D ^{ARD}	
	ENOG503Q3Q0 ^{ARD}	
	ENOG503NZ6Z ^{TOL}	
	ENOG503Q506 ^{AG}	
	ENOG503NTYE ^I	
ENOG503Q320		

N = 12

Figura 3.3.6. Análisis de dN/dS para grupos de ortología comunes a las siete especies (N = 2337). (A) Distribución de valores promedio de dN/dS para *G. metallitolerans* y las tres especies del grupo externo. Entre paréntesis: número de genes con dN/dS > 1 (pval <0,05) en alguna región, considerados bajo presión selectiva (N = 12). (B) Categorías funcionales (COG) de los genes bajo presión selectiva. Se indican los grupos que también presentaron señal de evolución positiva en: ^{ARD}ARD, ^{MSD}MSD, ^{TOL}Tolerans, ^A*G. aciditolerans*, ^I*G. iberica*, ^G*G. agrionensis*

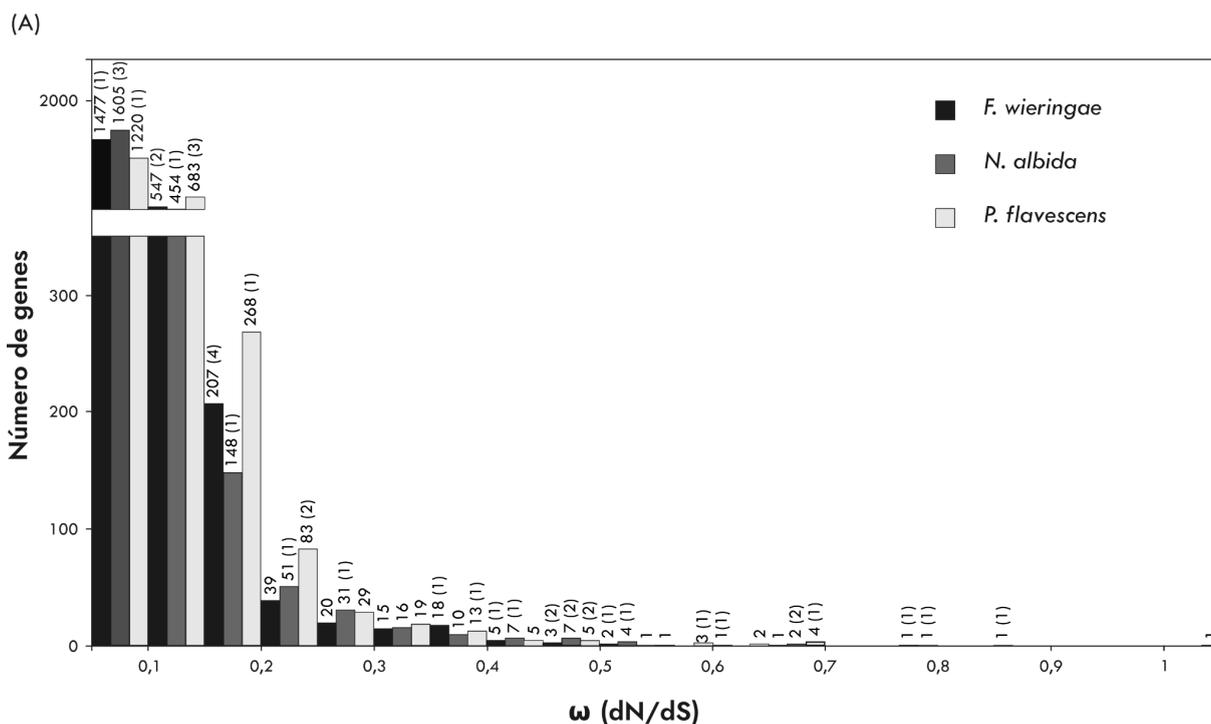


(B)

12,4	ENOG503NU44 ENOG503NYMC	Transcripción
6,3	ENOG503NU5F	Procesamiento & modificaciones del ARN
6,3	ENOG503NW30	Replicación, recombinación & reparación
6,3	ENOG503NVQQ	Estructura & dinámica de la cromatina
6,3	ENOG503NZ9T	Transducción de señales
6,3	ENOG503NX16	Citoesqueleto
12,4	ENOG503NW7P ^{MSD} ENOG503NUKF	Tráfico intracelular, secreción & transporte vesicular
2,9	ENOG503PBR4	Chaperonas, modificaciones post-traduccionales, recambio de proteínas
2,9	ENOG503NWMX ^{ARD}	Transporte & metabolismo de carbohidratos
42,9	ENOG503PITJ ^{ARD}	Función desconocida
	ENOG503P70D ^{ARD}	
	ENOG503Q3Q0 ^{ARD}	
	ENOG503P6IY ^A	
	ENOG503NTYE ^M	

N = 16

Figura 3.3.7. Análisis de dN/dS para grupos de ortología comunes a las siete especies (N = 2337). (A) Distribución de valores promedio de dN/dS para *G. iberica* y las tres especies del grupo externo. Entre paréntesis: número de genes con dN/dS > 1 (pval < 0,05) en alguna región, considerados bajo presión selectiva (N = 16). (B) Categorías funcionales (COG) de los genes bajo presión selectiva. Se indican los grupos que también presentaron señal de evolución positiva en: ^{ARD}ARD, ^{MSD}MSD, ^A*G. aciditolerans*, ^M*G. metallitolerans*



(B)

7,1	ENOG5033NTXG ^A	Transcripción
28,7	ENOG503NVMY ^{AM}	Tráfico intracelular, secreción & trans. vesicular
	ENOG503NXZ6 ^A	
	ENOG503NVDU ^A	
	ENOG503NWSX	
7,1	ENOG503NZK8	Tráfico intracel., secreción & trans. vesicular / Estructura nuclear
7,1	ENOG503NUGR ^{AM}	Chaperonas, modificaciones post-traduccionales, recambio de proteínas
7,1	ENOG503NWMX ^{ARD}	Transporte & metabolismo de carbohidratos
42,9	ENOG503PITJ ^{ARD}	Función desconocida
	ENOG503P70D ^{ARD}	
	ENOG503Q3Q0 ^{ARD}	
	ENOG503Q506 ^{AM}	
	ENOG503P4N7	
	ENOG503P7P4	

N = 14

Figura 3.3.8. Análisis de dN/dS para grupos de ortología comunes a las siete especies (N = 2337). (A) Distribución de valores promedio de dN/dS para *G. agrionensis* y las tres especies del grupo externo. Entre paréntesis: número de genes con dN/dS > 1 (pval < 0,05) en alguna región, considerados bajo presión selectiva (N = 14). (B) Categorías funcionales (COG) de los genes bajo presión selectiva. Se indican los grupos que también presentaron señal de evolución positiva en: ^{ARD} *G. aciditolerans*, ^A *G. aciditolerans*, ^M *G. metallitolerans*

3.3.5 Búsqueda de genes de interés

La búsqueda de genes específicos relacionados con las adaptaciones al bajo pH y alta concentración de metales demostró que las especies del género *Goffea* poseen múltiples sistemas relacionados con la resistencia al estrés oxidativo (aunque bajo número

de genes catalasa), transportadores específicos de metales, chaperonas y proteínas de shock, pero carecen de metalotioneínas y fitoquelatinas.

Con respecto a las respuestas de defensa ante estrés oxidativo, los 4 genomas del clado ARD, presentaron secuencias correspondientes a genes involucrados en 7 de las 9 vías exploradas. El análisis de genes relacionados con vías no enzimáticas permitió identificar en copia única, los genes responsables de la **síntesis de glutatión** y de **carotenoides** para todas las especies, incluyendo las del grupo externo. No se encontraron secuencias homólogas a ninguna **fitoquelatina** (PC) y sólo pudo encontrarse una secuencia homóloga a **metalotioneínas** (MTs) en el genoma de la especie externa *P. flavescens*. El análisis de la única predicción proteica reveló una cantidad de residuos de cisteínas por encima de lo usual (superior al 20 %). La comparación de la secuencia contra la base de datos de Uniprot, resultó homóloga a una metalotioneína de *Kockovaella imperatae* (acc: A0A1Y1UF93, cobertura: 40,4 %, eval: 9,6 e-21). Para el resto de los modelos proteicos de todas las especies, la proporción de residuos cisteína mostró una distribución pareja, con un 95-96 % de todas las predicciones con un uso menor al 2,5 % y el resto, con un uso debajo del 5 % (disponible como Material Suplementario, Tabla S3.3.6). Independientemente de esto, todos los genomas presentan una predicción perteneciente al grupo de ortología ENOG503P6XA (*copper fist DNA binding*), con la presencia del dominio PF00649 presente en factores de transcripción fúngicos que activan la transcripción de MTs. Estas predicciones son ortólogas al gen YPR008W de *S. cerevisiae*: un activador transcripcional involucrado en la adaptación al estrés ácido. Con respecto a las vías enzimáticas, se halló gran variabilidad de grupos de ortología representados y números de copia diferenciales entre las especies. Para las enzimas de familia **glutatión transferasa**, se encontraron 11 grupos distintos, de los cuales sólo 5 tuvieron representantes en una o más especies de género *Goffeauzyma* (con 3 de ellos comunes a todas las especies estudiadas). *G. agrionensis* fue la especie de ARD con mayor cantidad de grupos asignados a esta función, incluyendo una copia extra en uno de los grupos comunes, y uno único de este genoma (correspondiente a la GST exclusiva de levaduras URE2). También se observaron diferencias en las enzimas de oxidación/reducción de glutatión, hallándose en el genoma de *G. agrionensis* únicamente 2 representantes de los 4 grupos de ortología correspondientes a **glutaredoxinas** presentes en el resto de los genomas. De manera similar, en los genomas de clado MSD, así como en el de *N. albida*, pudieron detectarse 2 grupos de ortología distintos correspondientes a **glutatión**

peroxidasas, de los cuales 1 solo pudo ser relevado en las otras especies y se halló 1 grupo de función glutatión peroxidasa, compartido entre todo el clado MSD y *N. albida*, pero ausente en el resto de las especies. La presencia de **superóxido dismutasas** fue igual para todos los genomas. Mientras que las **catalasas** se englobaron en un único grupo de ortología para todas las especies de *Goffeauzyma*, hallándose en todas ellas 1 secuencia homóloga a la secuencia *query* G05570. Tanto en *G. agrionensis* como en todo el grupo externo se encontró además 1 versión homóloga a G05702, con 2 copias en *P. flavescens*, mientras que la presencia de 1 versión homóloga a la secuencia *query* G01478-G04411 se constató en el grupo externo y en *G. metallitolerans*.

En la Tabla 3.3.6 se resumen los resultados de las búsquedas de genes vinculados con estas respuestas.

Tabla 3.3.6. Resultados de las búsquedas por homologías para vías de respuesta a estrés oxidativo (vías no enzimáticas). Se muestran los grupos de ortología (GO) representados y su identidad en cada una de las especies. *CAM: curado a mano. Para los GO, Se utiliza la nomenclatura ENOG50~(5 caracteres) para facilitar la visualización.

V I A	Gen de interés	GO	<i>Goffeauzyma</i>				Grupo externo			
			<i>aciditolerans</i>	<i>metallitolerans</i>	<i>iberica</i>	<i>agrionensis</i>	<i>F. wieringae</i>	<i>N. albida</i>	<i>P. flavescens</i>	
R O S · N O E N Z I M Á T I C A	GSH1 γ-glutamilsteína sintetasa	3NVBR	Ga_1327	Gm_0502	Gi_3416	Gg_0970	Fw_3473	Na_2528	Pf_7756	
	GSH2 glutatión-sintetasa	3NUOF	Ga_5729	Gm_0751	Gi_2741	Gg_5831	Fw_1998	Na_2285	Pf_2693	
	síntesis de caroten oides	idi	3NU8E	Ga_5276	Gm_3398	Gi_0590	Gg_4082	Fw_2856	Na_3260	Pf_3856
		FPS	3NXTX	Ga_4039	Gm_5550	Gi_3999	Gg_0908	Fw_2452	Na_5708	Pf_7284
		crtE	3NVHG	Ga_3436	Gm_3816 *CAM	Gi_097 *CAM	Gg_5367	<i>scaffold 449</i> (88789-88441)	Na_6403	Pf_0852
		crtI	3NY26	Ga_4753	Gm_4978	Gi_2294	Gg_4446	<i>scaffold 6</i> (757164-755392)	Na_3243	Pf_4952
		crtYB	3Q1TM	Ga_5013	Gm_4489	Gi_5477	Gg_0974	Fw_2475	Na_2754	Pf_7906
crtR	3NUZ8	Ga_3143	Gm_0741	Gi_2369	Gg_2351	Fw_6270	Na_5258	Pf_0986		
PC fitoquelatina	-	-	-	-	-	-	-	-		
MT metalotioneina	-	-	-	-	-	-	-	Pf_3895 (% Cys >20)		

Tabla 3.3.6 (cont.). Resultados de las búsquedas por homologías para vías de respuesta a estrés oxidativo (vías enzimáticas). Se muestran los grupos de ortología (GO) representados y su identidad en cada una de las especies. *CAM: curado a mano. Para los GO, se utiliza la nomenclatura ENOG50~(5 caracteres) para facilitar la visualización.

V I A	Gen de interés	GO	<i>Goffeauzyma</i>				Grupo externo		
			<i>aciditolerans</i>	<i>metallitolerans</i>	<i>iberica</i>	<i>agrionensis</i>	<i>F. wieringae</i>	<i>N. albida</i>	<i>P. flavescens</i>
R O S · E N Z I M Á T I C A	SOD superóxido dismutasa	Cu 3P1VF	Ga_4977	Gm_5639	Gi_5048	Gg_5866	Fw_6435	Na_4561	Pf_2360
		Mn 3NZG9	Ga_5632	Gm_5006	Gi_4991	Gg_6012	Fw_0661	Na_0284	Pf_3345
	catalasas	3NUEX	Ga_4455	Gm_3588	Gi_4903	Gg_3197	Fw_0215	Na_6424	Pf_3686
		-	-	-	-	Gg_1563	Fw_6588	Na_2570	Pf_6174 Pf_1717
		-	-	Gm_1759	-	-	Fw_2517	Na_2555	Pf_2051
	GRX glutarredoxinas	3P6QI	Ga_3001	Gm_5484	Gi_0985	-	Fw_4416	Na_1007	Pf_0977
		3NXX9	Ga_1668	Gm_5608	Gi_1185	Gg_2178	Fw_3301	Na_5929	Pf_3446
		P2U5	Ga_2664	Gm_3015	Gi_3935	-	Fw_1209	Na_5225	Pf_3425
		3P72N	Ga_3030	Gm_2543	Gi_2543	Gg_2067	Fw_3796	Na_3735	Pf_5463
	GPX peroxidasa	3P24F	Ga_2395	Gm_1683	Gi_1319	Gg_3890	Fw_2041	Na_3400	Pf_6941
			Ga_1582	Gm_1810	Gi_1958	-	-	Na_6011	-
	GST glutación transferasa	3P3EX	Ga_4080	Gm_1401	Gi_4890	Gg_5544	Fw_5115	Na_0304	Pf_4069
		3NTZJ	Ga_0424	Gm_2180	Gi_4085	Gg_5584	Fw_0153	Na_4038	Pf_6169
		3NZ93	Ga_5144	Gm_2987	Gi_2162	Gg_0826	Fw_2674	Na_5500	Pf_4994
			-	-	-	Gg_2793	-	-	-
		3P2DJ	Ga_5288	Gm_3204	-	-	-	Na_1080	-
		3NZOD	-	-	-	Gg_2658	Fw_1254	Na_6447	-
		3NYFG	-	-	-	Gg_5277	-	-	-
		Otras	-	-	-	-	1	-	4

El relevamiento de otros grupos de ortología relacionados con tolerancia permitió hallar homólogos dentro del género *Goffeauzyma* para todas las búsquedas realizadas, con excepción de los transportadores de cadmio, sin embargo, no existen diferencias significativas respecto a las especies externas, aunque los grupos de ortología pueden ser diferentes (solo se presentan números generales sin distinción entre ellos). La información detallada se encuentra como Material Suplementario (Tabla S3.3.7) y la información resumida en la Tabla 3.3.7. El relevamiento de proteínas de tolerancia/homeostasis de los metales analizados fue virtualmente idéntico para todas las especies con entre 2 y 4 copias

para cada grupo, dependiendo la especie. Entre los resultados a destacar se pudo observar que ninguna especie de *Goffeauzyma* presentó homólogos de **transportadores de cadmio**, los cuales sólo se identificaron en la especie de referencia *P. flavescens*. La búsqueda de secuencias de 28 grupos de ortología relacionados con el **transporte, eflujo y difusión de protones**, permitió identificar entre 7 y 8 grupos de ortología de copia única para todas las especies. Para los **canales iónicos**, de 26 grupos relevados, se encontraron entre 7 y 9 por especie, con secuencias parálogas en *G. aciditolerans* y *G. metallitolerans* en 2 de los grupos (correspondientes a canales de potasio). Respecto de los **transportadores y proteínas de eflujo y difusión de cationes**, se encontraron 3 grupos comunes a todas las especies, con secuencias parálogas en *G. aciditolerans* y en *N. albida*. Por último, el relevamiento de **chaperonas, chaperoninas y proteínas de shock** mostró en líneas generales, números más elevados (ya fuese en número de genes y/o en grupos entre los que se distribuyeron) para las especies externas que para las especies de interés. El grupo ENOG503NYVV fue el único con múltiples representantes en genomas de *Goffeauzyma* (2 copias en *G. aciditolerans*). Este grupo de ortología está validado por YGL073W de *S. cerevisiae*: un factor de transcripción de shock térmico que activa múltiples genes en respuesta a diversas fuentes de estrés y a la acidificación intracelular transitoria.

Tabla 3.3.7. Grupos de ortología (GOs) relevados en las 7 especies en función de su posible vinculación con la resistencia a metales/acidez extrema. Se indica el número de genes que se encontró para cada categoría en cada especie y entre paréntesis el número de grupos de ortología a las que pertenecen.

Tipo de proteína	# GOs relevados	<i>Goffeauzyma</i>				Grupo externo		
		<i>aciditolerans</i>	<i>metallitolerans</i>	<i>iberica</i>	<i>agrionensis</i>	<i>F. wieringae</i>	<i>N. albida</i>	<i>P. flavescens</i>
Cu transp./resist. familia <i>Ctr</i> y otros	23	2 (2)	3 (3)	4 (4)	2 (2)	3 (2)	2 (2)	4 (4)
Zn transp./resist.	6	2 (2)	2 (2)	2 (2)	2 (2)	2 (2)	3 (2)	2 (2)
Cd transp./resist.	3	-	-	-	-	-	-	3 (1)
As transp./resist.	5	3 (3)	3 (3)	3 (3)	3 (3)	4 (3)	3 (3)	3 (3)
H ⁺ transp./eflujo/dif.	27	8 (8)	8 (8)	8 (8)	8 (8)	7 (7)	8 (8)	8 (8)
Canales iónicos	26	10 (8)	10 (8)	7 (7)	8 (7)	9 (8)	8 (8)	9 (9)
Cationes transp./eflujo/dif.	10	3 (3)	3 (3)	3 (3)	3 (3)	4 (3)	3 (3)	3 (3)
Chaperon(in)as	153	26 (26)	26 (26)	27 (27)	25 (25)	24 (24)	29 (28)	29 (28)
Proteínas de shock	61	15 (14)	13 (13)	15 (15)	13 (13)	14 (14)	20 (17)	17 (17)

3.4 DISCUSIÓN

3.4.1 Secuenciación del genoma, ensamblado, predicciones de genes y anotaciones

Este trabajo es la primera descripción de genomas, ensamblados y anotados, de las cuatro especies del género *Goffeauzyma* aisladas de ARD. La gran similitud de adaptaciones presentes entre las 4 especies y la similitud de los ambientes extremos que ocupan posibilitó realizar un abordaje novedoso con el objetivo, no sólo de describir, sino de identificar adaptaciones generales dentro del grupo y así comprender cómo los genomas de las levaduras evolucionan y permiten que éstas habiten nichos de características restrictivas. Si bien los estadísticos genómicos obtenidos en este capítulo son ligeramente menores que los de *N. vishniacii*, el ensamblado de las 4 especies fue en todos los casos de alta calidad, con un *coverage* relativamente alto (mayor a 40X para todas), bajo número de *scaffolds* y un grado de completitud que permite garantizar que las anotaciones obtenidas sustentan los análisis realizados. En este sentido, es importante destacar que el valor de completitud obtenido para un genoma está sujeto a la base de datos BUSCO empleada. Estas bases de datos se encuentran construidas utilizando genomas de especies de referencia que no siempre logran abarcar de forma acabada las características del clado que representan. Los recursos genómicos disponibles no presentan distribución equitativa para todos los taxones que contienen y esto resulta frecuentemente en que los valores de completitud de los genomas, sobre todo para especies no convencionales, se encuentran subestimados (Waterhouse et al., 2018). Por ejemplo, al evaluar la completitud de cepas de *S. cerevisiae* y *Schizosaccharomyces pombe* Zhang y colaboradores (2022) concluyeron que las diferencias del estadístico no tenían que ver con la calidad de los ensamblados sino con las bases de datos empleadas. La base de datos más cercana para *S. cerevisiae* corresponde a la clase *saccharomycetes*, mientras que para *Sch. pombe*, ésta se amplía al *phylum* *ascomycota*. Al ser de un nivel taxonómico mayor, dicha base de datos contiene más genes específicos de otras especies por lo que su representatividad global en el genoma de *Sch. pombe* es menor (Zhang et al., 2022). En este trabajo se observó que, la completitud de los genomas para las especies de género *Goffeauzyma* presenta valores entre 86 y 88 %, empleando la base de datos de *Basidiomycota*. Estos valores son comparables a los obtenidos para 90 especies de *Fungi* donde Li y colaboradores (2018) indicaron valores de completitud entre 69,3 % y

98,6%. La completitud estimada mediante BUSCO es consistente con lo esperado para levaduras no convencionales e indica que los genomas obtenidos son de alta calidad.

La cantidad de secuencias codificantes predichas para cada una de las levaduras del género *Goffeauzyma* fue menor que las obtenidas para las del grupo externo. La búsqueda *ad hoc* de nuevos genomas, publicados posteriormente a este análisis, permitió verificar que es incluso menor al promedio de todos los genomas disponibles a la fecha del orden Filobasidiales (N = 12), permitiendo constatar que son las levaduras de este género las que redujeron la información. La pérdida de genes es un motor de cambios evolutivos y un elemento pervasivo en todos los reinos de la vida (Albalat & Cañestro, 2016). En levaduras, particularmente, muchas de las mutaciones que surgen en condiciones de crecimiento restrictivas como la escasez de nutrientes, ocasionan pérdidas de función (*ibídem*). En este sentido, Lu y colaboradores mostraron que al exponer a distintas cepas de *Saccharomyces cerevisiae* a condiciones de acidez extrema (pH 2), alta concentración de etanol y elevadas temperaturas, la pérdida de genes se correlacionaba con un mayor *fitness* de las mismas (Lu et al. 2015). En el mismo sentido, en bacterias y arqueas acidófilas, la racionalización del genoma (*streamlining*) vía reducción de su tamaño y pérdida de genes no esenciales, también se ha propuesto como un mecanismo que posibilita reducir los costos de energía y sobrevivir en ambientes de pH extremadamente bajo (Hou et al., 2008; Vergara et al., 2020; Cortez et al., 2022). Además de desplegar múltiples estrategias de resistencia energéticamente demandantes para mantener un citoplasma circumneutral, los acidófilos deben sobrevivir en condiciones de escasos nutrientes y, generalmente, abundancia de metales (Cortez et al., 2022). Relacionado con estas condiciones particulares, Olson propuso la hipótesis “menos es más”, sugiriendo que la pérdida de genes representa una respuesta evolutiva común en poblaciones que experimentan un cambio en el patrón de presiones selectivas impuestas por el ambiente (Olson, 1999). En este sentido esta Tesis ya ha destacado la disminución de genes en la levadura psicrofilia *N. vishniacii* (Capítulo 2) y su relación con la “Hipótesis de la reina negra”. La marcada reducción de secuencias codificantes en el clado ARD, no sólo es consistente con los trabajos previamente mencionados, sino que refuerza la evidencia de la importancia adaptativa de reducir el número de genes en adaptación de los extremófilos. Si bien los ambientes que ocupan estos microorganismos poseen características particulares y complejas, los mismos son estables en sus condiciones. Tanto en *Naganishia vishniacii* como en las especies de género *Goffeauzyma*

discutidas en este capítulo, existe una clara señal de que son especies fuertemente especializadas a estos ambientes, por lo cual es esperable que sólo conserven genes que les permitan sobrevivir y competir eficientemente en las condiciones impuestas, y pierdan todos aquellos relacionados con respuestas fisiológicas y/o metabólicas que no requieran. El estudio de la pérdida de números de copias funcionales global (los casos puntuales se discutirán más adelante), permitieron verificar que las especies de ARD tienen un promedio de 33 % menos de grupos funcionales con múltiples copias que las levaduras del grupo externo. La ventaja adaptativa en relación a la pérdida de genes podría radicar entonces en la pérdida de variantes redundantes en aquellas familias de genes cuyas funciones bioquímicas se encuentren cubiertas en las condiciones del ambiente. Futuros abordajes permitirán profundizar sobre cómo la dispensabilidad de genes se relaciona con los ambientes extremos, y qué tipo de información es más susceptible de ser perdida.

Por otro lado, los tamaños de los genomas de las especies de *Goffeauzyma* fueron similares a los de las especies de referencia, con la excepción del genoma de *Goffeauzyma agrionensis*, que no sólo fue el más pequeño de todos, sino que además resultó ser el más pequeño de todos los disponibles hasta la fecha para el orden Filobasidiales (N = 12 en la base de datos de NCBI, según una revisión *ad hoc* realizada al momento de la presentación de este manuscrito). El tamaño de los genomas dentro de los Hongos muestra una muy amplia variación que va desde las 9Mb hasta las 180Mb, con un promedio que ronda las 40 Mb, tanto para basidiomicetes como ascomycetes (Mohanta et al., 2015). Las grandes fluctuaciones observadas a nivel Macroevolutivo de todo este gran grupo son generalmente explicadas mediante procesos de deriva y de selección (Kelkar et al., 2012). Aunque los cambios en el tamaño del genoma se producen de forma independiente de la información contenida en genes (pueden haber variaciones importantes con misma cantidad de secuencias codificantes), y ciertos autores aún sostienen que el ADN “basura” no posee efectos significativos en el *fitness*, existe una mirada creciente que indica que los cambios en el tamaño del genoma siempre tendrán impacto sobre el *fitness* por lo cual las variaciones deben ser interpretadas como adaptativas o maladaptativas (Mohanta et al., 2015). En este sentido, la reducción del tamaño de un genoma puede entenderse como parte del mecanismo de racionalización debido a la carga energética y de recursos que implica su replicación (Koskiniemi et al., 2012). A su vez, el tamaño del genoma se correlaciona con el tamaño del núcleo, por lo que el primero puede ser resultado de un *trade-off* entre disminuir

el gasto energético y asegurar un tamaño mínimo. Por otro lado, la reducción de los genomas es considerada una especialización supeditada a la previsibilidad del entorno y que ha sido vinculada con la vida en condiciones extremas (Wolf & Koonin, 2013; Salwan & Sharma, 2022). En este sentido, resulta interesante remarcar que *G. agrionensis* es la única especie del clado ARD aislada de un ambiente de origen natural y geográficamente distante al de las demás, al que la acción de los glaciares moldeó hasta el Pleistoceno tardío (hace cerca de 12.000 años), si bien la actividad volcánica que dio origen a la caldera Copahue-Caviahue se inició en el Plioceno (hace unos 2 millones de años) (Baffico et al., 2010). El resto de las especies, por el contrario, están asociadas a un ambiente de transformación y acidificación más reciente, con actividad minera desde la dominación romana pero con mayor impacto a partir de su desarrollo a cielo abierto (finales de siglo XIX) (Fernández-Caliani, 2008). En este sentido, el tamaño genómico considerablemente menor de esta especie sugiere que puede estar asociado con tiempos de evolución muy prolongados o con características muy particulares de este hábitat que hayan conducido a la reducción observada.

3.4.3 Grupos de ortología y asignaciones funcionales

La asignación y comparación de los grupos de ortología presentes en el clado ARD - y/o subclados dentro del mismo - y ausentes en las especies del grupo externo, permitieron identificar elementos clave en relación a la adaptación de estas especies a medios acuáticos ácidos y con alta concentración de metales. Los grupos relacionados con el transporte y metabolismo de carbohidratos y una importante proporción de familias sin categoría funcional asociada (Función desconocida) resultan los más destacados y dan indicio de su importancia en la evolución de las especies de *Goffeauzyma*. El primer rasgo puede entenderse en relación a las necesidades energéticas que requieren las estrategias y mecanismos de tolerancia y adaptación. Kan y colaboradores, expusieron una cepa de *Rhodotorula mucilaginosa* a altas concentraciones de cobre, y observaron que su adecuación se vinculaba con una aceleración en la descomposición de carbohidratos para la obtención de energía, con una mayor producción de polisacáridos involucrados en la unión a metales y con un incremento de polisacáridos en las membranas celulares (Kan et al., 2019). De manera similar, la comparación de cepas de *Penicillium* resistentes con cepas sensibles a metales resultó en una expansión en la cantidad de anotaciones concernientes al metabolismo de carbohidratos (Chi et al., 2021). El predominio de grupos encontrados

dentro de esta categoría funcional para todas las especies de *Goffeauzyma* resalta la necesidad de éstas de optimizar el uso de recursos, obtener intermediarios necesarios para contrarrestar las condiciones adversas reinantes y satisfacer la alta demanda energética requerida para mantener, además, un pH interno circumneutral en un contexto de alta concentración de protones. Por otro lado, la alta representatividad de predicciones con función desconocida es sugerente sobre cómo las levaduras se adaptan a sus ambientes. Las proteínas de función desconocida y los genes huérfanos representan un gran problema para la bioinformática por la imposibilidad de asignarles una función o comprender su importancia para los seres vivos (Hanson et al., 2010). La proporción de estas secuencias en los genomas eucariotas y en particular dentro de hongos es muy significativa y sin embargo sus secuencias se encuentran poco conservadas y su distribución es limitada a los organismos dentro de un clado sin correlato con secuencias homólogas en otros clados (Fakhar et al., 2023). La bioinformática es una herramienta que permite asignar funciones putativas a las secuencias presentes dentro del genoma mediante el reconocimiento de homología con secuencias similares cuya funcionalidad haya sido caracterizada en otro organismo (Pearson, 2013). En este sentido, los genes huérfanos no poseen correlato con otras especies estudiadas y por tanto la predicción de su función se encuentra limitada. Sin embargo, se ha propuesto que estos genes poseen una importancia fundamental para la comprensión de las características particulares de cada especie en relación a los nichos que ocupan (Gollery et al., 2010) y que se encuentran implicados en procesos biológicos como el metabolismo y las respuestas a estrés (Fakhar et al., 2013). El hecho de que más de un 40 % de los grupos funcionales del clado ARD ausentes en las especies de referencia pertenezcan a esta categoría, y que incluso proporciones mayores se observan en subclados más recientes como el MSD, sugieren que su importancia adaptativa para el género es significativa. Las herramientas bioinformáticas utilizadas en este trabajo no permiten sugerir cuál sería su rol biológico y si bien en los últimos años han surgido algunas estrategias de predicción de funcionalidad, el uso de éstas y su validación experimental es aún limitado (Fakhar et al., 2023). Sin embargo, el análisis experimental y futuros modelados podrían aportar valiosas pistas sobre mecanismos de tolerancia al frío actualmente ignorados.

Para los otros grupos de ortología, el análisis de las descripciones funcionales condujo la atención sobre 3 particulares, debido a que se encontraron representadas en las 4 especies, pero también por grupos hallados al interior del subclado MSD (correspondiente a

especies provenientes de la Península Ibérica). Estos grupos, relacionados con (1) proteínas activadoras RhoGAP, (2) predicciones con dominios MFS y (3) fosfatasa ácida, se vinculan a distintos niveles de acción con la adaptación de estas especies y su tolerancia a las condiciones de sus ambientes de proveniencia. Su relevancia se discute a continuación:

(1) El primero guarda relación con la modulación de distintas vías de transducción de señales que posibilitan una delicada comunicación entre los entornos interno y externo. En este nexo, las proteínas activadoras de Rho detectan estímulos extracelulares y modulan la acción de proteínas río abajo que desencadenan acciones hacia el interior de la célula (Bos et al., 2007). Esta familia de efectores actúa en una variedad de respuestas, entre ellas, la vía de integridad de la pared celular: una cascada de transducción responsable del mantenimiento de la integridad de la célula frente a condiciones estresantes y que además mantiene un diálogo activo con otras cascadas de señalización, posibilitando expandir su respuesta frente a distintos estímulos adversos (Fuchs & Mylonakis, 2009; García et al., 2019). Gatti y colaboradores demostraron que mutaciones que inactivan la RhoGAP involucrada en la vía de integridad de pared celular de *Saccharomyces cerevisiae* generan cepas sensibles a condiciones de bajo pH (Gatti et al., 2005). Si bien la familia es amplia y modula la acción de proteínas involucradas en diversas respuestas, la presencia de secuencias inducibles por estrés en la región promotora de esta familia de efectores, sugiere su activación frente a condiciones desfavorables (*ibídem*). El hallazgo de proteínas activadoras RhoGAP de manera exclusiva a todo el clado ARD y la existencia de grupos ortólogos también al interior del subclado MSD, indica que estos moduladores están implicados en la adaptación de estas especies a un entorno que requiere de un continuo monitoreo de las condiciones externas y de un delicado mantenimiento de la homeostasis interna.

(2) El segundo está vinculado con la presencia de dominios MFS (*Major Facilitator Superfamily*) y guarda relación con distintos mecanismos de transporte que posibilitan la tolerancia de las especies a la alta concentración de metales y acidez extrema. Esta familia de proteínas integrales de membrana constituye uno de los 2 grupos más importantes dentro del reino Fungi encargados de mediar el transporte de un amplio espectro de sustratos a través de las membranas biológicas (Del Sorbo et al., 2000), entre ellos, del simporte de glucosa con protones, utilizando para tal fin la fuerza protón motriz (Chen et al., 2017). Si bien estos transportadores actúan tanto en la entrada como en la salida de solutos, suelen ser clasificados principalmente como transportadores de eflujo, asistiendo en la

detoxificación de compuestos por medio de su compartimentalización en organelas o su extrusión fuera de la célula (Khatiwada et al., 2020). Transportadores de esta gran familia, en particular, han sido descritos en levaduras como mediadores de respuesta a metales, ya sea por su conducción aislada o conjugados con glutatión (Dhaoui et al., 2011). A su vez se ha visto que cepas deficientes de transportadores de este tipo pueden ser defectuosas en la homeostasis de pH y de glutatión y su acción es requerida para la resistencia celular a especies reactivas de oxígeno. Existe, además, evidencia que indica que transportadores MFS actúan indirectamente en el control del potencial de membrana, alterando la homeostasis de lípidos y regulando el pH interno de la célula y la maquinaria de respuesta a estrés en fungi (Chen et al., 2017). Debido a que son responsables del transporte de sustratos ampliamente diversos (Kaback et al., 2001), una caracterización funcional y molecular de los grupos encontrados es requerida para determinar las vías específicas en la que están involucrados. Sin embargo, su alta incidencia, tanto en el clado ARD como en el subclado MSD, sumado a las evidencias que existen en relación a sus roles diversos en medios ácidos y con alta concentración de metales, aportan sustento a su relevancia en la adaptación de estas especies.

(3) Por último, el tercero está vinculado con un mecanismo más específico en la adaptación del clado a la presencia de metales en un contexto de alta concentración de protones. Las fosfatasas ácidas catalizan la hidrólisis de varios ésteres de fosfatos en ambientes ácidos, liberando aniones que reaccionan con los iones metálicos, lo cual provoca su precipitación y previene su entrada a la célula (Crusberg et al., 2004). Esta acción protectora, y una mayor producción de dichas enzimas en ambientes con alta concentración de metales fueron reportadas en bacterias y posteriormente en levaduras y hongos filamentosos (Liang et al., 2016b). La presencia de fosfatasas ácidas en los genomas de ARD y MSD constituye un nuevo indicio de la implicancia de estas enzimas en la tolerancia a metales en ambientes ácidos y de su particular importancia en la resistencia de las especies de este clado.

Más allá de estas 3 funcionalidades destacadas, otros grupos poseen funciones que sugieren potenciales roles adaptativos importantes, por ejemplo, la presencia de secuencias referidas a flavin-monooxigenasas (FMO) en todo el clado ARD en virtud de la alta producción de especies reactivas de oxígeno que ocurre en presencia de metales (Pocsi et al., 2004). Las FMO actúan en el sistema búfer que garantiza un adecuado mantenimiento del

potencial óxido-reductor citoplasmático (Suh et al., 1999; Catucci et al., 2019) y su accionar cobra especial relevancia en especies expuestas a una sobreproducción de especies reactivas de oxígeno inducida por su exposición a metales. Dentro del clado MSD, además de lo ya mencionado, se destaca la presencia de enzimas de la familia citocromo P450 (CYP450) y de grupos con función esterolesaturasa, implicados en la respuesta a la presencia de radicales de oxígeno y en el mantenimiento de la integridad estructural de las levaduras frente a las condiciones de estrés en las que se encuentran, respectivamente. Las enzimas de familia CYP450, al igual que las FMO descritas, son monooxigenasas de membrana cuya acción como antioxidantes y confiriendo tolerancia a contaminantes ha sido bien documentada en varios organismos (van Gorcom et al., 1998; Črešnar & Petrič, 2011) y en particular en varias especies del reino Fungi expuestas a distintos metales (Bazzicalupo, 2020; Yu et al., 2020). Debido a que la superfamilia CYP450 es relativamente amplia, la implicancia de los subgrupos encontrados en las acciones de tolerancia debe ser confirmada mediante un estudio funcional de las moléculas. Por otro lado, la estructura y funcionalidad de la membrana está vinculada con su composición y la regulación de la misma sirve como respuesta adaptativa a muchos cambios ambientales (Zhang & Rao, 2010). Alteraciones en la composición de esteroides de membrana vuelven a las levaduras hipersensibles o resistentes a diferentes tipos de estrés (Kodedova & Sychrova, 2015). Fletcher y colaboradores reportaron que la tolerancia a acidez a largo plazo en una cepa de *Saccharomyces cerevisiae* expuesta a evolución experimental, se adquiría mediante la modificación de la composición de esteroides y la modulación de los niveles de hierro, en función de su necesidad para la biosíntesis de esteroles (Fletcher et al., 2017). La presencia de estos grupos en el clado MSD refuerza nuevamente la importancia de estas dos vías en la tolerancia de las especies de ARD: la primera línea de defensa celular, en la que la membrana plasmática junto con la pared celular protegen a las levaduras de las condiciones del ambiente y la existencia de mecanismos que las protejan de la exacerbada generación de especies reactivas de oxígeno en condiciones de alta concentración de metales.

Esto último también permite explicar los grupos encontrados al interior del clado Tolerans, destacándose la presencia de tiorredoxinas, implicadas en la homeostasis redox en presencia de metales (Vido et al., 2001) y reconocidas como parte del repertorio de respuesta desplegado por arqueas y bacterias acidófilas implicadas en la biolixiviación de metales, frente al estrés oxidativo inducido (Cárdenas et al., 2012; Norambuena et al., 2012);

un grupo con motivo rico en cisteínas antecedido de un dominio SET, presente en distintos reguladores en los que posibilita moderar la expresión génica frente a numerosos contextos y cuya presencia pudo vincularse con la protección de *Saccharomyces cerevisiae* al estrés oxidativo (Tran et al., 2019) y de hongos ectomicorrícicos expuestos a distintos metales (Bellion et al., 2006); y una oxidasa dependiente de NADPH homóloga a la de *Alternaria alternata*, un hongo patógeno oportunista en el cual Chung y colaboradores la asociaron a la generación de niveles basales de especies reactivas de oxígeno que posibilitan la activación de reguladores involucrados con la detoxificación necesaria para hacer frente a la respuesta oxidativa de su hospedador (Chung et al., 2014). Resulta interesante mencionar que en el patógeno el mecanismo también involucra la biosíntesis de sideróforos que posibilitan la captación de hierro, esencial para la acción de la catalasa y la superóxido dismutasa (Chen et al., 2014). Es posible que la presencia de un transportador de hierro de baja afinidad en este subclado esté también vinculada con dicha respuesta. Otros autores señalaron a la homeostasis de hierro como una condición necesaria para la tolerancia de *Saccharomyces cerevisiae* a níquel y cadmio, por mecanismos aún no del todo determinados (Ruotolo et al., 2018) al corroborar alteraciones en transportadores de hierro examinando una colección de cepas mutantes de delección. Asimismo, su implicancia en modificaciones de la composición de la membrana, en función de su necesidad para la biosíntesis de esterol (Fletcher et al., 2017), fue esbozada más arriba. Una vez más, así como en los taxones inferiores, en el subclado de divergencia más temprana los grupos destacados se vinculan con el mantenimiento de la estructura celular y con la respuesta al estrés oxidativo.

Por último, en este trabajo se destaca la presencia del grupo con dominio de función desconocida DUF4396 hallado en el clado Tolerans. Si bien este dominio no pertenece a una clasificación funcional particular, Straw y colaboradores observaron una expresión 6 veces mayor en una secuencia conteniendo dicho dominio al exponer a *Streptomyces lividans* a altas concentraciones de cobre. Estos autores tampoco encontraron homología cercana al realizar una búsqueda en bases de datos proteicas (Straw et al., 2018). Sin embargo, la aparición de este dominio en secuencias de especies de dos géneros no emparentados expuestas a condiciones estresantes similares constituye un fuerte indicio de la implicancia de este dominio en algún tipo de mecanismo de adaptación y refuerza la importancia de ahondar sobre la funcionalidad y caracterización de predicciones con descripción funcional vacante como punto de partida interesante en la comprensión de posibles nuevos

mecanismos de tolerancia. Todos los mecanismos de resistencia evidenciados mediante el abordaje de asignación de grupos de ortología apuntan a que las principales estrategias en estas especies están vinculadas con el mantenimiento de la membrana/pared celular en condiciones que dificulten la alteración del medio interno en un contexto extracelular desfavorable e impidan el ingreso de especies tóxicas. Por otro lado, frente a la amenaza de que ocurra, este abordaje demuestra la importancia de la existencia de medios búfer que posibiliten contrarrestar sus consecuencias, principalmente la producción de especies reactivas de oxígeno, así como el control del potencial de membrana y la extrusión de los agentes tóxicos.

3.4.3.1 Análisis de duplicaciones en los grupos funcionales multicopia

A partir de la estrategia de análisis de grupos con múltiples copias se evidenciaron dos tendencias opuestas, ambas relacionadas con la evolución a condiciones extremófilas específicas. Por un lado, la menor proporción de grupos de grupos funcionales multicopia para las especies del clado ácido ($5,4 \pm 0,2$ %) respecto a las especies del grupo externo (8 ± 1). Como contrapartida, las familias de genes multicopia se relacionan con estrategias adaptativas particulares; hecho que encuentra sustento, además, en que la adaptación a metales, es a menudo poligénica y ocurre, tanto por divergencia nucleotídica, como por variaciones en el número de copias (Bazzicalupo et al., 2020). De hecho, Gerstein y colaboradores (Gerstein et al., 2015) demostraron, en consonancia con otros autores, que la tolerancia a cobre en cepas de *Saccharomyces cerevisiae* expuestas al metal en exceso se lograba en parte debido a un aumento en el número de copias en genes vinculados con su detoxificación y un consecuente incremento en la expresión los productos codificados. Respuestas similares se observaron también en plantas expuestas a zinc y cadmio (Talke et al., 2006; Craciun et al., 2012). El análisis de grupos funcionales multicopia para el clado ARD refuerza lo mencionado más arriba, incorporando, además, a partir de la inferencia de duplicaciones, una base evolutiva en la adquisición de las mismas. La alta representatividad de genes con dominios MFS (ENOG503NXMN, ENOG503NUYQ, ENOG503NXJD), verificándose al menos una expansión para 4 de ellos en nodos de la filogenia, vuelve a señalar su importancia (discutida más arriba), probablemente en relación con la respuesta a estrés oxidativo, al transporte y eflujo de especies metálicas y al mantenimiento del potencial de membrana. De igual manera, se reiteran alto número de copias de otros genes

involucrados en la respuesta a estrés oxidativo (familia citocromo P450: ENOG503Q4JQ, glutamato descarboxilasas: ENOG503NUJY) confirmando la importancia de estos mecanismos de defensa frente al estrés oxidativo inducido por la presencia de metales. La presencia de genes relacionados con el transporte de sideróforos (ENOG503P237) probablemente también tenga relación con estas vías dada la importancia de incrementar su captación en un ambiente pobre de nutrientes. Este abordaje, permitió además evidenciar la presencia de genes de respuesta específica a cobre en todo el clado, incluída una ATPasa Tipo P (ENOG503NXXK), homóloga a la bomba de translocación de cobre de *T. asahii* (EKC98179.1), lo cual fue verificado mediante blast recíprocos realizados *ad hoc*. Este grupo de proteínas se encuentran representadas en todos los reinos de la vida y son las principales determinantes de la resistencia al cobre en bacterias (Antsotegi-Uskola et al., 2017) mediando su eflujo en ambientes ricos en este metal. Distintos abordajes bioinformáticos han confirmado la relevancia de estas enzimas en la resistencia de distintas especies de *Acidithiobacillus* presentes en medios ácidos con alto contenido de cobre (Quatrini et al., 2007; González et al., 2014); para este grupo se verificó, además, al menos una expansión por duplicación dentro del clado Tolerans. El indicio de que este tipo de eventos sucedieron en nodos de la filogenia del clado resalta aún más el carácter adaptativo de estos grupos de genes.

Expansiones génicas por duplicación ocurridas en el clado

La duplicación de genes es un tipo de mutación espontánea que suele ocurrir con efectos neutros sobre el organismo, sin embargo, el destino de estas duplicaciones depende de procesos de mutaciones, deriva y selección (Kondrashov, 2012). Por lo general los fragmentos duplicados, al ser redundantes, quedan sujetos a que las mutaciones que ocurren sobre el mismo terminen llevando a la pérdida de su función (pseudogenización), sin embargo en determinadas ocasiones la duplicación en sí misma puede tener efectos de dosaje y ser mantenida por selección natural dependiendo de su efecto sobre el *fitness*, mientras que en otras ocasiones pueden ocurrir mutaciones que actúen sobre su función y que eso produzca efectos positivos sobre el *fitness* (Innan & Kondrashov, 2010). Por estos motivos la duplicación de genes es un proceso fundamental y de gran importancia evolutiva para el desarrollo de nuevas funciones génicas y nuevas adaptaciones (Ohno, 1970; Chauve et al., 2008; Demuth & Hahn, 2009). Por todo esto, la identificación de 27 eventos de

duplicación ocurridos dentro del clado ARD resulta de particular relevancia para comprender la evolución de estas especies.

Varias de las funciones representadas por estos grupos ya fueron discutidas como parte de los abordajes anteriores, reincidencia que refuerza aún más la direccionalidad de la evolución de estas especies en asociación con sus ambientes. En este sentido, se verificaron al menos 4 expansiones por duplicaciones para genes con dominio MFS, 3 de ellas implicando a todo el clado ARD (ENOG503NX4R, ENOG503NUYQ y ENOG503NXMN) y 1 al subclado Tolerans (ENOG503NXJD). La localización de 3 de estos eventos en el nodo basal de este clado constituye una evidencia más de su importancia en relación a la adaptación de estas especies como rasgo general. Para el clado MSD, se destaca además el grupo con dominio de función desconocido DUF563 (ENOG503PIXM), también presente en un grupo expandido en el genoma de *G. iberica* (ENOG503PV7J), insistiendo en el estudio de los genes de función desconocida. Por último, para las especies del clado Tolerans 2 grupos cobraron especial relevancia: el grupo homólogo a una ATPasa tipo P (ENOG503NXKX) responsable del bombeo de iones, cuya importancia fue ya discutida más arriba y un grupo homólogo a una nitrato reductasa (ENOG503P0PU). Esta enzima es considerada una de las más importantes para la biogénesis de nanopartículas metálicas (Zomorodian et al., 2016; Baymiller et al., 2017), principalmente de plata, metal minado en la zona (aunque con intermitencia) desde las primeras civilizaciones de Occidente, hasta el cierre de las minas, en el siglo XXI (Tornos et al., 2008). De hecho, Mourato y colaboradores (2011) verificaron la biosíntesis de partículas de plata y oro por una levadura extremófila aislada de este sistema acuático. El mecanismo de síntesis por vía de esta enzima ha sido bien caracterizado en el hongo filamentoso *Fusarium oxysporum* (Durán et al., 2005; Guilguer-Casagrande et al., 2019). Al realizar búsquedas por homología *ad hoc* de las secuencias de nitrato reductasas del clado Tolerans contra el genoma de este hongo depositado en NCBI, se pudo corroborar su correspondencia con la del hongo. La biosíntesis de nanopartículas es uno de los mecanismos de resistencia a estrés por metales en microorganismos (Roychoudhury, 2020), razón por la cual su evolución adaptativa en las especies de género *Goffeauzyma* es esperable y las vuelven de particular importancia biotecnológica para procesos que requieran la formación de nanopartículas de este metal. Para el resto de los genes con duplicaciones al interior del clado, su asociación con mecanismos de tolerancia no es clara, ni evidente, y no existen reportes de que dichos genes se relacionen con mecanismos de

resistencia, sin embargo, son genes relevantes para que futuros estudios establezcan función en relación a las condiciones evolutivas del grupo. Por otro lado, las expansiones ocurridas en ramas individuales (al menos 2 en *G. metallitolerans*, 5 en *G. iberica* y 3 en *G. agrionensis*) se listaron, pero quedan por fuera del análisis de esta Tesis. Su estudio puede dar cuentas de las adaptaciones especie específicas ocurridas en la evolución de estos organismos.

En un contexto de racionalización genómica, las evidencias de duplicaciones en nodos de la filogenia del clado y el mantenimiento de múltiples copias contrario a la tendencia global que presentan estas especies por reducir la redundancia génica, refuerza que las principales respuestas a metales en estas especies tienen relación con su protección frente al estrés oxidativo y una mayor disponibilidad de transportadores y bombas, ya sea para garantizar una adecuada provisión de nutrientes en un contexto de escasez, o bien para posibilitar la extrusión de especies tóxicas y garantizar la homeostasis de pH. A su vez, este enfoque sugiere la existencia de mecanismos específicos y promisorios en términos de su campo de aplicación, como es la síntesis de nanopartículas.

3.4.4 Análisis de cociente dN/dS

El cálculo de cocientes dN/dS proporciona evidencias de adaptación o restricción funcional en los genes codificantes. Sin embargo, la interpretación y validación de los resultados no siempre es directa y requiere de la consideración de distintas variables, debido a que factores como la calidad de las secuencias, sesgos por contenidos de GC variables (Ratnakumar et al., 2010), o las escalas temporales (Rocha et al., 2006) y parámetros poblacionales (Kryazhimskiy & Plotkin, 2008) pueden derivar en cocientes no representativos de la direccionalidad de la evolución. De manera consistente con lo evidenciado para la gran mayoría de genes de *N. vishniacii* y otros genomas de levaduras (Freel et al., 2014; Salathé et al., 2013; Brion et al., 2015), las comparaciones pareadas de ortólogos de las especies de *Goffeauzyma* y las especies de referencias sugieren que los cuatro genomas se encuentran principalmente bajo selección depuradora, indicando que el *core* de genes compartidos se encuentra conservado. El abordaje por medio de *sliding windows*, a diferencia de lo reportado para la especie psicrotolerante, no resultó en candidatos con evidencias de selección positiva. Esto puede deberse a que, entre secuencias distantes la tasa de evolución de sustituciones sinónimas es mucho más rápida que la de sustituciones no sinónimas por lo que el dN suele saturar entre especies muy divergentes (Rocha et al., 2006). Esta situación,

que no se evidenció para levadura antártica debido a la disponibilidad de secuencias genómicas filogenéticamente más cercanas, resultó determinante para las levaduras estudiadas en este Capítulo. El surgimiento creciente de datos genómicos, especialmente de especies no modelo, proveerán nuevos recursos para sortear estos desafíos y resolver enigmas evolutivos complejos, como la adaptación nicho-específica de organismos extremos.

3.4.5 Búsqueda de genes involucrados en la respuesta a metales y acidez

De forma análoga a lo que se observó para *N. vishniacii* (Capítulo 2), la búsqueda de genes reportados para especies tolerantes a condiciones similares a las del clado ARD, mostró que ciertas características esperables, en tanto constituyen mecanismos específicos de respuesta en otros organismos, como es la presencia de las metalotioneínas o fitoquelatinas, no necesariamente están presentes de manera extendida en todos los microorganismos adaptados a este tipo de condicionantes ambientales. La lógica de lo esperable queda sujeta entonces a otras consideraciones y la caracterización de estrategias de tolerancia requiere de la complementación de distintos abordajes que permitan evidenciar particularidades de otra manera pasadas por alto. En efecto, la relevancia de los mecanismos de respuesta a estrés oxidativo, un carácter sobresaliente del análisis de grupos de ortología únicos y de aquellos con evidencias de expansiones en los nodos de la filogenia del clado, quedaría solapada si el abordaje se acotara únicamente *genome mining* de vías generales bien caracterizadas, dado que esta estrategia no permitió evidenciar diferencia alguna entre los genomas de interés y los del grupo externo.

Dos tipos de oligopéptidos han sido extensamente caracterizados como estrategias para eliminar la acción tóxica de los metales: el primero compuesto por los péptidos enzimáticamente sintetizados glutatión (GSH), de distribución ubicua en todos los reinos de la vida, y fitoquelatinas (PCs), solo detectadas hasta ahora en plantas, levaduras, nematodos y microalgas; y el segundo, codificados por una familia de genes, las metalotioneínas (MTs), que, al igual que el GSH, son de distribución ubicua (Gutiérrez et al., 2019). El GSH y las vías enzimáticas relacionadas con este tripéptido (glutarredoxinas, glutatión peroxidasas y transferasas), se encontraron representadas en todas las especies, sin distinción clara entre el clado y el grupo externo. Sin embargo, la imposibilidad de encontrar evidencias PCs y MTs en las especies expuestas a altas concentraciones de metales sugieren que estas estrategias no son universales para las levaduras. En el sentido opuesto, especies más generalistas, como

Saccharomyces cerevisiae y *Schizosaccharomyces pombe*, cuentan con ambos efectores moleculares, aún cuando su resistencia a metales no es un rasgo sobresaliente (Culotta et al., 1994; Borrelly et al., 2002; Adamo et al., 2012; Yong et al., 2014). Este hecho no se limita únicamente a levaduras y el cuestionamiento de por qué los genes de las sintasas de PCs, o las MTs se encuentran ampliamente distribuídas, a pesar de que los organismos que las llevan no se encuentran, ni se encontraron evolutivamente, expuestos a estrés por metales ha sido planteado numerosas veces (Hamer, 1986; Palmiter, 1998; Clemens et al., 2006). Tentativas de respuestas han demostrado que las MTs no son esenciales y el *knockout* de uno o más genes de MTs no genera mortalidad (Gutiérrez et al., 2019). El uso de distintas estrategias de búsqueda (*genome mining* mediante secuencias homólogas y proporción de residuos cisteínas), sumado al hecho de que en microorganismos las MTs suelen encontrarse en varias isoformas codificadas por genes parálogos (Gutiérrez et al., 2009), por lo que resulta poco probable que todas las versiones queden encubiertas a este tipo de estrategias, da sustento a los resultados y refuerza la ausencia de esta estrategia en las especies de género *Goffeauzyma*.

Por otro lado, para otros genes vinculados con transporte y mecanismos de resistencia específicos para distintas especies metales relevados, así como eflujo de protones, canales iónicos, chaperonas y proteínas de shock térmico, las funciones se encontraron representadas en el clado ARD, pero no se evidenciaron diferencias sustantivas con respecto a las especies externas en cuanto a número de ortólogos o expansiones de parálogos. Ambos resultados permiten sugerir que las estrategias comunes que posibilitan la tolerancia a altas concentraciones de metales y acidez no son universales y la evolución de respuestas en asociación a estas condiciones restrictivas adquiere rasgos generales en los genomas de estas especies, algunos de ellos evidenciados por los abordajes hasta aquí discutidos, más que mecanismos puntuales o específicos de resistencia.

3.5 CONCLUSIONES

- Se obtuvieron genomas de calidad para las 4 especies de género *Goffeauzyma* aisladas de ARD. Su anotación, conjuntamente con la de especies cercanas no asociadas a estos ambientes, permitió un análisis comparativo comprehensivo que develó una serie de adaptaciones de interés para comprender su carácter extremófilo.
- Los genomas de las 4 especies poseen un menor número de genes y una menor proporción de grupos funcionales multicopia que las especies de referencia. Esta reducción puede entenderse en relación a una racionalización genómica vinculada a la íntima asociación de estas especies con los ambientes acuáticos ácidos y a una optimización de recursos en dicho contexto.
- El tamaño del genoma de *G. agrionensis* es el menor de todos los analizados hasta la fecha para el orden Filobasidiales. Esta característica no respondería a una tendencia general del clado y estaría vinculada con un tiempo más prolongado de evolución de la misma en relación a ambientes restrictivos, o bien a particularidades de su ambiente, de génesis y ubicación geográfica distintos al del resto de las especies de *Goffeauzyma*.
- Los genes relacionados con el transporte y metabolismo de carbohidratos se encuentran altamente representados en las 4 especies, posiblemente asociado a la oligotrofia de sus ambientes.
- El predominio de genes con función desconocida constituye un rico sustrato para futuros abordajes que permitan caracterizar adaptaciones aún desconocidas de estas especies.
- El análisis de grupos de ortología ausentes en las especies de referencia sugiere que las principales estrategias en este clado están vinculadas con el mantenimiento de la membrana/pared celular, la protección frente al daño oxidativo y la extrusión de especies tóxicas.
- Resulta imperioso profundizar sobre las proteínas activadoras RhoGAP, los transportadores de la superfamilia MFS y las fosfatasas ácidas. El monitoreo de las condiciones externas y del mantenimiento de la homeostasis interna, un eficiente uso de sustratos energéticos y transporte de metabolitos e iones, así como la protección

frente a especies reactivas de oxígeno, y la reducción de la toxicidad de los metales del medio en un contexto de acidez extrema, permiten explicar la relevancia de estos grupos exclusivos, pero ameritan una confirmación molecular y funcional.

- El estudio de grupos funcionales con múltiples copias permitió proponer estrategias adaptativas coherentes con los mecanismos descritos por el resto de los abordajes (MFS, bombeo de iones, protección frente a estrés oxidativo).
- Se verificaron 27 eventos de expansiones génicas por duplicación en nodos de la filogenia del grupo, destacando el carácter adaptativo de los transportadores MFS, las ATPasas de membrana y una enzima de acción nitrato reductasa, posiblemente vinculada con la biosíntesis de nanopartículas.
- El estudio de dominios de función desconocida, evidenciados por varios de los abordajes aquí presentados puede resultar interesante para comprender nuevas estrategias de adaptación.
- La inexistencia de evidencias de estrategias clásicas de tolerancia a metales (como las metalotioneínas) indica que estos mecanismos de adaptación no son universales.

Conclusiones finales



Genómica comparativa al extremo

Comprender las adaptaciones al ambiente ha sido uno de los problemas principales de la Biología, desde Darwin hasta el presente. Científicos y científicas buscan entender cómo la evolución da forma a estas características y cuáles son los cambios concretos que hacen que los organismos se adapten mejor a su entorno. La genómica comparativa viene arrojando luz sobre la comprensión, no sólo de las relaciones filogenéticas entre especies, sino también sobre los mecanismos moleculares que subyacen a la evolución de las secuencias, la especiación, los reordenamientos genómicos, el origen de nuevos genes y la evolución de aquellos implicados en distintas adaptaciones (Koonin et al., 2000; Miller et al., 2004; Wolfe, 2006). La disponibilidad de una gama cada vez mayor de genomas de levaduras es una poderosa fuente de información a la hora de determinar las bases genéticas de las variaciones fenotípicas, los mecanismos de rearrreglos genómicos y aquellos involucrados en el aislamiento y especiación. También posibilitan dilucidar los procesos que ocurren en escalas de tiempo pequeñas, como la adaptación reciente de especies a sus ambientes (Wolfe, 2006; Riley et al., 2016; Libkind et al., 2020a).

Los resultados resumidos en esta Tesis representan un aporte significativo para nuestra comprensión de las bases genómicas que sustentan la supervivencia y el crecimiento de las levaduras en entornos inhóspitos para buena parte de las formas de vida que habitan nuestro planeta. Tal como fue planteado en la hipótesis, la genómica comparativa posibilitó identificar muchos de estos rasgos en base a la distinción de características comunes a las especies adaptadas, frente a las de especies de distribución más amplia. La mayor riqueza de estos hallazgos radica sin dudas en su carácter transversal a dos poliextremotolerancias distintas, debido a que los factores estresantes configuran escenarios totalmente distintos en ambos ambientes extremos. Esto sugiere que la evolución de respuestas adaptativas se produce en algunas ocasiones sobre rasgos generales cuyas huellas pueden rastrearse en los genomas testigo y que podrían no ser específicas del tipo de condicionamiento ambiental particular. Entre ellos, las pérdidas de genes observadas para los linajes utilizados en este trabajo constituyen un rasgo cuya relevancia es aún mayor a la luz de teorías evolutivas recientes, como la racionalización genómica y la Hipótesis de la Reina Negra. Este efecto resultó marcado tanto a nivel global, en la cantidad de predicciones génicas y secuencias proteicas derivadas, como en la proporción de grupos funcionales multicopia. El análisis *ad hoc* incentivado por estas observaciones, permitió además ver que los resultados no se limitan a las especies empleadas como grupo externo en cada uno de los abordajes, sino que

se extienden a otros representantes del orden Filobasidiales (Material Suplementario, Tabla S3.3.1). En un contexto de racionalización genómica, dicho fenómeno puede interpretarse como una optimización de recursos en ambientes oligotróficos dada la carga genética y de recursos que implica la replicación de material genético dispensable para la adaptación y sobrevivencia de estas especies a las condiciones específicas de sus ambientes extremos, pero estables. De interés, a futuro, resulta promisorio el análisis de dichos genes y caracterización de las funcionalidades dispensadas en cada (o en ambas) condición(es). De manera aparentemente contradictoria, ciertas duplicaciones que expanden el número de *loci* dentro del genoma, también se destacan en este trabajo y juegan un rol fundamental en la evolución de estas especies, resaltando la importancia del abordaje comparativo para identificar cómo estos eventos ocurrieron en las filogenias.

Entre los mecanismos generales cabe destacar que la membrana celular y genes vinculados con su integridad y detección de las condiciones externas tienen un rol protagónico en la extremotolerancia de todos los organismos aquí relevados, propiciando una reestructuración adaptativa y como desencadenantes de respuestas *downstream*. Asimismo los transportadores de membrana de familia MFS poseen alta redundancia y un rol relevante, tanto en la adaptación a ambientes fríos, como a los ambientes acuáticos ácidos y con alta concentración de metales. Finalmente, en todos los genomas extremotolerantes existe una gran cantidad de genes con actividad desconocida cuyo estudio se propone como una línea interesante a abordar para caracterizar nuevos mecanismos de adaptación y tolerancia.

La mayoría (sino todos) los genes compartidos entre especies extremotolerantes y especies de referencia de otros ambientes, se encuentran bajo selección depuradora. El análisis de casos puntuales por fuera de esta tendencia, brinda elementos clave para entender las adaptaciones al ambiente (como quedó evidenciado para *N. vishniacii*), gracias a la valiosa impronta que la evolución imprime en los genomas y que puede ser evidenciada mediante la herramienta de dN/dS. Por el contrario, la inexistencia de evidencias de estrategias clásicas de tolerancia al frío en *N. vishniacii* (proteínas anticongelantes) y a metales, en el clado ARD (metalotioneínas o fitoquelatinas), indica que estos mecanismos de adaptación no son universales y que cada clado es capaz de solucionar los problemas de adecuación al ambiente de variadas formas.

A estos hallazgos, que contribuyen a nuestra comprensión de las formas que adopta la evolución de rasgos adaptativos, se suma la puesta en disponibilidad de los genomas obtenidos como fruto de este trabajo, que incrementa nuestros recursos y con ello aporta a nuestro conocimiento de las levaduras no convencionales extremotolerantes.

A lo largo de todos los años que dieron lugar a los resultados aquí presentados, los ensamblados genómicos que inicialmente se me presentaron como una desafiante aproximación al inhóspito mundo de la Bioinformática fueron desplegando una multiplicidad de caminos a explorar. En mi trayectoria, condensada en esta Tesis, me fui adentrando en el conocimiento de los organismos extremotolerantes y de las huellas que, como fue propuesto en la hipótesis inicial, sus adaptaciones al ambiente imprimen en sus genomas. Las evidencias desarrolladas en ese sentido, posibilitaron la reconstrucción de una historia evolutiva, entre otras posibles. La hipótesis planteada no solo fue confirmada en sus aspectos generales, sino que fue expandida a partir de las observaciones y resultados obtenidos durante todo este tiempo.

REFERENCIAS



A hombros de gigantes

- Adamo, G. M., Lotti, M., Tamás, M. J., & Brocca, S. (2012). Amplification of the CUP1 gene is associated with evolution of copper tolerance in *Saccharomyces cerevisiae*. *Microbiology*, 158(9), 2325–2335
- Adekoya, O. A., Helland, R., Willassen, N. P., & Sylte, I. (2005). Comparative sequence and structure analysis reveal features of cold adaptation of an enzyme in the thermolysin family. *Proteins: Structure, Function, and Bioinformatics*, 62(2), 435–449.
- Aghajari, N., Feller, G., Gerday, C., & Haser, R. (1998). Structures of the psychrophilic *Alteromonas haloplanctis* α -amylase give insights into cold adaptation at a molecular level. *Structure*, 6(12), 1503–1516.
- Aguilera, A., Manrubia, S. C., Gómez, F., Rodríguez, N., & Amils, R. (2006). Eukaryotic community distribution and its relationship to water physicochemical parameters in an extreme acidic environment, Río Tinto (Southwestern Spain). *Applied and Environmental Microbiology*, 72(8), 5325–5330.
- Aguilera, J., Rande-Gil, F., & Prieto, J. A. (2007). Cold response in *Saccharomyces cerevisiae*: new functions for old mechanisms. *FEMS microbiology reviews*, 31(3), 327–341.
- Aguilera, A., Souza-Egipsy, V., González-Toril, E., Rendueles, O., & Amils, R. (2010). Eukaryotic microbial diversity of phototrophic microbial mats in two Icelandic geothermal hot springs. *Int Microbiol*, 13(1), 21–32.
- Aguilera, A., & González-Toril, E. (2019). Eukaryotic life in extreme environments: acidophilic fungi. *Fungi in extreme environments: ecological role and biotechnological significance* (pp. 21–38). Springer.
- Aguilera, Á., Souza-Egipsy, V., & González-Toril, E. (2020). La vida en Río Tinto. *Ministerio de Defensa - Gobierno de España*.
- Akcil, A., & Koldas, S. (2006). Acid Mine Drainage (AMD): causes, treatment and case studies. *Journal of cleaner production*, 14(12–13), 1139–1145.
- Albalat, R., & Cañestro, C. (2016). Evolution by gene loss. *Nature Reviews Genetics*, 17(7), 379–391.
- Albanese, D., Coleine, C., Rota-Stabelli, O., Onofri, S., Tringe, S. G., Stajich, J.E., Selbmann, L., & Donati, C. (2020). Antarctic cryptoendolithic bacterial lineages of pre-Cambrian origin as proxy for Mars colonization. *bioRxiv*, 2020-02.
- Alcaíno, J., Cifuentes, V., & Baeza, M. (2015). Physiological adaptations of yeasts living in cold environments and their potential applications. *World Journal of Microbiology and Biotechnology*, 31, 1467–1473.
- Al-Fageeh, M. B., & Smales, C. M. (2006). Control and regulation of the cellular responses to cold shock: the responses in yeast and mammalian systems. *Biochemical Journal*, 397(2), 247–259.
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research*, 25(17), 3389–3402.
- Amaral Zettler, L. A., Gómez, F., Zettler, E., Keenan, B. G., Amils, R., & Sogin, M. L. (2002). Eukaryotic diversity in Spain's River of Fire. *Nature*, 417(6885), 137–137.
- Amaretti, A., Raimondi, S., Sala, M., Roncaglia, L., De Lucia, M., Leonardi, A., & Rossi, M. (2010). Single cell oils of the cold-adapted oleaginous yeast *Rhodotorula glacialis* DBVPG 4785. *Microbial cell factories*, 9(1), 1–6.
- Amils, R., & Fernández-Remolar, D. (2020). Río Tinto: An extreme acidic environmental model of astrobiological interest. *Extremophiles as Astrobiological Models*, 21–44.

- Andreolli, M., Lampis, S., Lorenzini, M., & Zapparoli, G. (2021). Features of basidiomycetous yeasts from grapes and apples associated with crop environment and fermenting juice. *Journal of Applied Microbiology*, 131(4), 1932-1941.
- Antsotegi-Uskola, M., Markina-Iñarrairaegui, A., & Ugalde, U. (2017). Copper resistance in *Aspergillus nidulans* relies on the PI-type ATPase CrpA, regulated by the transcription factor AceA. *Frontiers in Microbiology*, 8, 912.
- Argueso, J. L., Carazzolle, M. F., Mieczkowski, P. A., Duarte, F. M., Netto, O. V., Missawa, S. K., Galzerani, F., Costa, G.G., Vidal, R. O., Noronha, M. F. & Pereira, G. A. (2009). Genome structure of a *Saccharomyces cerevisiae* strain widely used in bioethanol production. *Genome research*, 19(12), 2258-2270.
- Assis Fidalgo, C. I. (2011). *Heavy metal resistance in extremophilic yeasts: A molecular and physiological approach*. [Doctoral dissertation, Universidade de Lisboa].
- Baffico, G. D., Díaz, M. M., Temporetti, P. F., Beamud, S. G., & Pedrozo, F. L. (2010). El lago Caviahue: un ambiente único. *Desde la patagonia. Difundiendo saberes*, 7(10), 18-23.
- Bahafid, W., Joutey, N. T., Asri, M., Sayel, H., Tirry, N., El Ghachtouli, N., & Sayel, N. T. H. (2017). Yeast biomass: an alternative for bioremediation of heavy metals. *Yeast-Industrial Applications*, 559.
- Baskaran, A., Kaari, M., Venugopal, G., Manikkam, R., Joseph, J., & Bhaskar, P. V. (2021). Antifreeze proteins (Afp): Properties, sources and applications—A review. *International Journal of Biological Macromolecules*, 189, 292-305.
- Bánfalvi, G. (2011). Heavy metals, trace elements and their cellular effects. In *Springer eBooks* (pp. 3–28).
- Bankar, A., Zinjarde, S., Shinde, M., Gopalghare, G., & Ravikumar, A. (2018). Heavy metal tolerance in marine strains of *Yarrowia lipolytica*. *Extremophiles*, 22, 617-628.
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Pribelski, A. D., Pyshkin, A. V., Sirotkin, A., Vyahhi, N., Tesler, G., Alekseyev, M. A., & Pevzner, P. A. (2012). SPADES: A new genome assembly algorithm and its applications to Single-Cell sequencing. *Journal of Computational Biology*, 19(5), 455–477.
- Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S., Khanna, A. K., Marshall, M., Moxon, S., Sonnhammer, E. L. L., Studholme, D. J., Yeats, C., & Eddy, S. R. (2004). The Pfam protein families database. *Nucleic Acids Research*, 32(90001), 138D – 141.
- Baymiller, M., Huang, F., & Rogelj, S. (2017). Rapid one-step synthesis of gold nanoparticles using the ubiquitous coenzyme NADH. *Matters*.
- Bazzicalupo, A. L., Ruytinx, J., Ke, Y. H., Coninx, L., Colpaert, J. V., Nguyen, N. H., Vilgalys, R. & Branco, S. (2020). Fungal heavy metal adaptation through single nucleotide polymorphisms and copy-number variation. *Molecular Ecology*, 29(21), 4157-4169.
- Bellion, M., Courbot, M., Jacob, C., Blaudez, D., & Chalot, M. (2006). Extracellular and cellular mechanisms sustaining metal tolerance in ectomycorrhizal fungi. *FEMS microbiology letters*, 254(2), 173-181
- Bellora, N., Moliné, M., David-Palma, M., Coelho, M. A., Hittinger, C. T., Sampaio, J. P., Gonçalves, P. & Libkind, D. (2016). Comparative genomics provides new insights into the diversity, physiology, and sexuality of the only industrially exploited tremellomycete: *Phaffia rhodozyma*. *BMC genomics*, 17(1), 1-16.
- Bijlani, S., Singh, N. K., Mason, C. E., Wang, C. C., & Venkateswaran, K. (2020). Draft genome sequences of Tremellomycetes strains isolated from the International Space Station. *Microbiology Resource Announcements*, 9(26), 10-1128.

- Bird, D. A., Sares, M. A., Policky, G. A., Schmidt, T. S., & Church, S. E. (2006). Impacts on water quality and biota from natural acid rock drainage in Colorado's Lake Creek watershed. In *Seventh International Conference on Acid Rock Drainage* (pp. 158-126).
- Blevins, W. R., Ruiz-Orera, J., Messeguer, X., Blasco-Moreno, B., Villanueva-Cañas, J.L., Espinar, L., Díez, J., Carey, L. B. & Mar Albà, M. (2019). Frequent birth of de novo genes in the compact yeast genome. *bioRxiv*, 575837.
- Bode, H. P., Dumschat, M., Garotti, S., & Fuhrmann, G. F. (1995). Iron sequestration by the yeast vacuole: a study with vacuolar mutants of *Saccharomyces cerevisiae*. *European journal of biochemistry*, 228(2), 337-342.
- Bøe, C. A., Håland, T. W., Boye, E., Syljuåsen, R. G., & Grallert, B. (2018). A novel role for ATR/Rad3 in G1 phase. *Scientific reports*, 8(1), 6880.
- Boekhout, T., Bai, F., Daniel, H. M., Groenewald, M., Robert, R., Tan, S., & Yurkov, A. (2022, October). *The Yeasts: A Taxonomic Study*. Retrieved November 8, 2023, from <https://theyeasts.org/species-search>
- Borrelly, G. P., Harrison, M. D., Robinson, A. K., Cox, S. G., Robinson, N. J., & Whitehall, S. K. (2002). Surplus zinc is handled by Zym1 metallothionein and Zhf endoplasmic reticulum transporter in *Schizosaccharomyces pombe*. *Journal of Biological Chemistry*, 277(33), 30394-30400.
- Bos, J. L., Rehmann, H., & Wittinghofer, A. (2007). GEFs and GAPs: critical elements in the control of small G proteins. *Cell*, 129(5), 865-877.
- Brady, D., & Duncan, J. R. (1994). Bioaccumulation of metal cations by *Saccharomyces cerevisiae*. *Applied Microbiology and Biotechnology*, 41, 149-154.
- Brandao, L. R., Libkind, D., Vaz, A. B., Espírito Santo, L. C., Moliné, M., de García, V., van Broock, M. & Rosa, C. A. (2011). Yeasts from an oligotrophic lake in Patagonia (Argentina): diversity, distribution and synthesis of photoprotective compounds and extracellular enzymes. *FEMS microbiology ecology*, 76(1), 1-13.
- Brion, C., Pflieger, D., Friedrich, A., & Schacherer, J. (2015). Evolution of intraspecific transcriptomic landscapes in yeasts. *Nucleic acids research*, 43(9), 4558-4568.
- Britton, G., Liaaen-Jensen, S., & Pfander, H. (2008). Special molecules, special properties. In *Carotenoids*, 4 (pp. 1-6).
- Burini, J. A. (2023). *Desarrollo de cultivos iniciadores cerviceros basados en levaduras nativas para la producción de bebidas fermentadas con alto valor agregado*. [Doctoral dissertation, Universidad Nacional del Comahue].
- Burkins, M. B., Virginia, R. A., Chamberlain, C. P., & Wall, D. H. (2000). Origin and distribution of soil organic matter in Taylor Valley, Antarctica. *Ecology*, 81(9), 2377-2391.
- Butinar, L., Spencer-Martins, I., & Gunde-Cimerman, N. (2007). Yeasts in high Arctic glaciers: the discovery of a new habitat for eukaryotic microorganisms. *Antonie Van Leeuwenhoek*, 91, 277-289.
- Buzzini, P., Branda, E., Goretti, M., & Turchetti, B. (2012). Psychrophilic yeasts from worldwide glacial habitats: diversity, adaptation strategies and biotechnological potential. *FEMS microbiology ecology*, 82(2), 217-241.
- Buzzini, P., & Margesin, R. (2014). Cold-adapted yeasts: a lesson from the cold and a challenge for the XXI century. *Cold-adapted yeasts: biodiversity, adaptation strategies and biotechnological significance*, 3-22.
- Buzzini, P., Turchetti, B., & Yurkov, A. (2018). Extremophilic yeasts: the toughest yeasts around?. *Yeast*, 35(8), 487-497.
- Cabrera, J. M., Temporetti, P., & Pedrozo, F. (2020). Trace metal partitioning and potential mobility in the naturally acidic sediment of Lake Caviahue, Neuquén, Argentina. *Andean Geology*, 47(1), 46.

- Cannon, S. B., & Young, N. D. (2003). OrthoParaMap: distinguishing orthologs from paralogs by integrating comparative genome data and gene phylogenies. *BMC bioinformatics*, 4(1), 1-15.
- Cánovas, D., Durán, C., Rodríguez, N., Amils, R., & De Lorenzo, V. (2003). Testing the limits of biological tolerance to arsenic in a fungus isolated from the River Tinto. *Environmental Microbiology*, 5(2), 133-138.
- Cary, S. C., McDonald, I. R., Barrett, J. E., & Cowan, D. A. (2010). On the rocks: the microbiology of Antarctic Dry Valley soils. *Nature Reviews Microbiology*, 8(2), 129-138.
- Castro, C., Urbietta, M. S., Cazón, J. P., & Donati, E. R. (2019). Metal biorecovery and bioremediation: Whether or not thermophilic are better than mesophilic microorganisms. *Bioresource technology*, 279, 317-326.
- Catucci, G., Querio, G., Sadeghi, S. J., Gilardi, G., & Levi, R. (2019). Enzymatically produced trimethylamine N-oxide: Conserving it or eliminating it. *Catalysts*, 9(12), 1028.
- Cavello, I. A., Bezus, B., Martinez, A., Garmendia, G., Vero, S., & Cavalitto, S. (2019). Yeasts from Tierra Del Fuego Province (Argentina): biodiversity, characterization and bioprospection of hydrolytic enzymes. *Geomicrobiology Journal*, 36(9), 847-857.
- Chatterjee, S., Chatterjee, N. C., & Dutta, S. (2012). Bioreduction of chromium (VI) to chromium (III) by a novel yeast strain *Rhodotorula mucilaginosa* (MTCC 9315). *African Journal of Biotechnology*, 11(83), 14920-14929.
- Chauve, C., Doyon, J. P., & El-Mabrouk, N. (2008). Gene family evolution by duplication, speciation, and loss. *Journal of Computational Biology*, 15(8), 1043-1062.
- Chen, C., & Wang, J. L. (2007). Characteristics of Zn²⁺ biosorption by *Saccharomyces cerevisiae*. *Biomedical and Environmental Sciences: BES*, 20(6), 478-482.
- Chen, L. H., Yang, S. L., & Chung, K. R. (2014). Resistance to oxidative stress via regulating siderophore-mediated iron acquisition by the citrus fungal pathogen *Alternaria alternata*. *Microbiology*, 160(5), 970-979.
- Chen, L. H., Tsai, H. C., Yu, P. L., & Chung, K. R. (2017). A major facilitator superfamily transporter-mediated resistance to oxidative stress and fungicides requires *Yap1*, *Skn7*, and MAP kinases in the citrus fungal pathogen *Alternaria alternata*. *PLoS One*, 12(1), e0169103.
- Cheng, C. H. C. (1998). Evolution of the diverse antifreeze proteins. *Current opinion in genetics & development*, 8(6), 715-720.
- Chi, B. B., Lu, Y. N., Yin, P. C., Liu, H. Y., Chen, H. Y., & Shan, Y. (2021). Sequencing and comparative genomic analysis of a highly metal-tolerant *Penicillium janthinellum* P1 provide insights into its metal tolerance. *Frontiers in microbiology*, 12, 663217.
- Chintalapati, S., Kiran, M. D., & Shivaji, S. (2004). Role of membrane lipid fatty acids in cold adaptation. *Cellular and molecular biology* (Noisy-le-Grand, France), 50(5), 631-642.
- Choi, J., & Kim, S. H. (2017). A genome Tree of Life for the Fungi kingdom. *Proceedings of the National Academy of Sciences*, 114(35), 9391-9396.
- Chojnacka, K. (2010). Biosorption and bioaccumulation—the prospects for practical applications. *Environment international*, 36(3), 299-307.
- Chung, K. R. (2014). Reactive oxygen species in the citrus fungal pathogen *Alternaria alternata*: the roles of NADPH-dependent oxidase. *Physiological and molecular plant pathology*, 88, 10-17
- Clark, M. S., & Worland, M. R. (2008). How insects survive the cold: molecular mechanisms—a review. *Journal of Comparative Physiology B*, 178, 917-933.

- Clemens, S., Kim, E. J., Neumann, D., & Schroeder, J. I. (1999). Tolerance to toxic metals by a gene family of phytochelatin synthases from plants and yeast. *The EMBO journal*, 18(12), 3325-3333.
- Clemens, S. (2006). Evolution and function of phytochelatin synthases. *Journal of plant physiology*, 163(3), 319-332.
- Collins, T., & Margesin, R. (2019). Psychrophilic lifestyles: mechanisms of adaptation and biotechnological tools. *Applied microbiology and biotechnology*, 103, 2857-2871.
- Connell, L., Redman, R., Craig, S., Scorzetti, G., Iszard, M., & Rodriguez, R. (2008). Diversity of soil yeasts isolated from South Victoria Land, Antarctica. *Microbial Ecology*, 56, 448-459.
- Connell, L. B., Rodriguez, R. R., Redman, R. S., & Dalluge, J. J. (2014). Cold-adapted yeasts in Antarctic deserts. *Cold-adapted yeasts: biodiversity, adaptation strategies and biotechnological significance*, 75-98.
- Convey, P., & Peck, L. S. (2019). Antarctic environmental change and biological responses. *Science Advances*, 5(11), eaaz0888.
- Conway, J. R., Lex, A., & Gehlenborg, N. (2017). UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics*, 33(18), 2938-2940.
- Cortez, D., Neira, G., González, C., Vergara, E., & Holmes, D. S. (2022). A large-scale genome-based survey of acidophilic bacteria suggests that genome streamlining is an adaptation for life at low pH. *Frontiers in Microbiology*, 13, 803241.
- Cowan, D. A., Casanueva, A., & Stafford, W. (2014). Ecology and Biodiversity of Cold-Adapted microorganisms. In *ASM Press eBooks* (pp. 117–132).
- Craciun, A. R., Meyer, C. L., Chen, J., Roosens, N., De Groot, R., Hilson, P., & Verbruggen, N. (2012). Variation in HMA4 gene copy number and expression among *Noccaea caerulea* populations presenting different levels of Cd tolerance and accumulation. *Journal of Experimental Botany*, 63(11), 4179-4189.
- Črešnar, B., & Petrič, Š. (2011). Cytochrome P450 enzymes in the fungal kingdom. *Biochimica Et Biophysica Acta (BBA) - Proteins and Proteomics*, 1814(1), 29–35.
- Crous, P. W., Wingfield, M. J., Lombard, L., Roets, F., Swart, W. J., Alvarado, P., Carnegie, A., Moreno, G. & Groenewald, J. Z. (2019). Fungal Planet description sheets: 951–1041. Persoonia: *Molecular Phylogeny and Evolution of Fungi*, 43, 223.
- Crusberg, T. C., Mark, S. S., & Dilorio, A. (2004). Biomineralization of heavy metals. *Fungal biotechnology in agricultural, food, and environmental applications*, 409-417.
- Culotta, V. C., Howard, W. R., & Liu, X. F. (1994). CRS5 encodes a metallothionein-like protein in *Saccharomyces cerevisiae*. *Journal of Biological Chemistry*, 269(41), 25295-25302.
- D'Amico, S., Collins, T., Marx, J. C., Feller, G., Gerday, C., & Gerday, C. (2006). Psychrophilic microorganisms: challenges for life. *EMBO reports*, 7(4), 385-389.
- Darriba, D., Posada, D., Kozlov, A. M., Stamatakis, A., Morel, B., & Flouri, T. (2020). ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. *Molecular biology and evolution*, 37(1), 291-294.
- de la Torre, J. R., Goebel, B. M., Friedmann, E. I., & Pace, N. R. (2003). Microbial diversity of cryptoendolithic communities from the McMurdo Dry Valleys, Antarctica. *Applied and environmental microbiology*, 69(7), 3858-3867.

- Del Sorbo, G., Schoonbeek, H. J., & De Waard, M. A. (2000). Fungal transporters involved in efflux of natural toxic compounds and fungicides. *Fungal genetics and biology*, 30(1), 1-15.
- Demuth, J. P., & Hahn, M. W. (2009). The life and death of gene families. *Bioessays*, 31(1), 29-39.
- Dhaoui, M., Auchère, F., Blaiseau, P. L., Lesuisse, E., Landoulsi, A., Camadro, J. M., Haguenaer-Tsapis, R. & Belgareh-Touzé, N. (2011). Gex1 is a yeast glutathione exchanger that interferes with pH and redox homeostasis. *Molecular biology of the cell*, 22(12), 2054-2067.
- Dias, C., Santos, J., Reis, A., & da Silva, T. L. (2019). Yeast and microalgal symbiotic cultures using low-cost substrates for lipid production. *Bioresource Technology Reports*, 7, 100261.
- Dopson, M., & Holmes, D. S. (2014). Metal resistance in acidophilic microorganisms and its significance for biotechnologies. *Applied microbiology and biotechnology*, 98, 8133-8144.
- Dragisic, V., Vrvic, M., Milenic, D., & Miladinovic, B. (1999). Mine waters and waters of copper deposit tailing disposals of Serbia and environmental protection. Hydrogeochemical Aspect. In *Mine, Water & Environment*, IMWA Congress. Sevilla, Spain (pp. 527-533).
- Dragone, N. B., Perry, L. B., Solon, A. J., Seimon, A., Seimon, T. A., & Schmidt, S. K. (2023). Genetic analysis of the frozen microbiome at 7900 m asl, on the South Col of Sagarmatha (Mount Everest). *Arctic, Antarctic, and Alpine Research*, 55(1), 2164999.
- Dreesens, L. L., Lee, C. K., & Cary, S. C. (2014). The distribution and identity of edaphic fungi in the McMurdo Dry Valleys. *Biology*, 3(3), 466-483.
- Dujon, B. (2010). Yeast evolutionary genomics. *Nature Reviews Genetics*, 11(7), 512-524.
- Durán, N., Marcato, P. D., Alves, O. L., De Souza, G. I., & Esposito, E. (2005). Mechanistic aspects of biosynthesis of silver nanoparticles by several *Fusarium oxysporum* strains. *Journal of nanobiotechnology*, 3, 1-7.
- Durocher, D., & Jackson, S. P. (2001). DNA-PK, ATM and ATR as sensors of DNA damage: variations on a theme?. *Current opinion in cell biology*, 13(2), 225-231.
- Eddy, S. R. (n.d.). HMMER: biosequence analysis using profile hidden Markov models. HMMER. <http://hmmer.org/>
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*, 32(5), 1792-1797.
- El-Nour, K. M. A., Eftaiha, A. F., Al-Warthan, A., & Ammar, R. A. (2010). Synthesis and applications of silver nanoparticles. *Arabian Journal of Chemistry*, 3(3), 135-140.
- Emms, D. M., & Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome biology*, 20, 1-14.
- Ernst, R., Ejsing, C. S., & Antonny, B. (2016). Homeoviscous adaptation and the regulation of membrane lipids. *Journal of molecular biology*, 428(24), 4776-4791.
- Ewart, K. V., Lin, Q., & Hew, C. L. (1999). Structure, function and evolution of antifreeze proteins. *Cellular and Molecular Life Sciences CMLS*, 55, 271-283.
- Fakhar, A. Z., Liu, J., Pajerowska-Mukhtar, K. M., & Mukhtar, M. S. (2023). The Lost and Found: Unraveling the Functions of Orphan Genes. *Journal of Developmental Biology*, 11(2), 27.
- Farhan, S. N., & Khadom, A. A. (2015). Biosorption of heavy metals from aqueous solutions by *Saccharomyces Cerevisiae*. *International Journal of Industrial Chemistry*, 6, 119-130.

- Fell, J. W., Boekhout, T., Fonseca, A., Scorzetti, G., & Statzell-Tallman, A. (2000). Biodiversity and systematics of basidiomycetous yeasts as determined by large-subunit rDNA D1/D2 domain sequence analysis. *International journal of systematic and evolutionary microbiology*, 50(3), 1351-1371.
- Feller, G., Arpigny, J. L., Narinx, E., & Gerday, C. (1997). Molecular adaptations of enzymes from psychrophilic organisms. *Comparative Biochemistry and Physiology Part A: Physiology*, 118(3), 495-499.
- Fernández-Caliani, J. C. (2008). Una aproximación al conocimiento del impacto ambiental de la minería en la Faja Pírrica Ibérica. *Macla: revista de la Sociedad Española de Mineralogía*, (10), 24-28.
- Fernández, P. M., Martorell, M. M., Blaser, M. G., Ruberto, L. A. M., de Figueroa, L. I. C., & Mac Cormack, W. P. (2017). Phenol degradation and heavy metal tolerance of Antarctic yeasts. *Extremophiles*, 21, 445-457.
- Ferreira, M. L., Casabuono, A. C., Stacchiotti, S. T., Couto, A. S., Ramirez, S. A., & Vullo, D. L. (2017). Chemical characterization of *Pseudomonas veronii* 2E soluble exopolymer as Cd (II) ligand for the biotreatment of electroplating wastes. *International Biodeterioration & Biodegradation*, 119, 605-613.
- Ferrer, M., Chernikova, T. N., Yakimov, M. M., Golyshin, P. N., & Timmis, K. N. (2003). Chaperonins govern growth of *Escherichia coli* at low temperatures. *Nature biotechnology*, 21(11), 1266-1267.
- Fletcher, E., Feizi, A., Bisschops, M. M., Hallström, B. M., Khoomrung, S., Siewers, V., & Nielsen, J. (2017). Evolutionary engineering reveals divergent paths when yeast is adapted to different acidic environments. *Metabolic engineering*, 39, 19-28.
- Fonseca, Á.; Boekhout, T.; Fell, J.W. *Cryptococcus Vuillemin* (1901). In *The Yeasts: A Taxonomic Study* (pp. 1161–1737).
- Fournier, D., Lemieux, R., & Couillard, D. (1998). Essential interactions between *Thiobacillus ferrooxidans* and heterotrophic microorganisms during a wastewater sludge bioleaching process. *Environmental pollution*, 101(2), 303-309.
- Freel, K. C., Friedrich, A., Hou, J., & Schacherer, J. (2014). Population genomic analysis reveals highly conserved mitochondrial genomes in the yeast species *Lachancea thermotolerans*. *Genome biology and evolution*, 6(10), 2586-2594.
- Fuchs, B. B., & Mylonakis, E. (2009). Our paths might cross: the role of the fungal cell wall integrity pathway in stress response and cross talk with other stress response pathways. *Eukaryotic cell*, 8(11), 1616-1625
- Gadanhó, M., Libkind, D., & Sampaio, J. P. (2006). Yeast diversity in the extreme acidic environments of the Iberian Pyrite Belt. *Microbial ecology*, 52, 552-563.
- Gadanhó, M., & Sampaio, J. P. (2009). *Cryptococcus ibericus* sp. nov., *Cryptococcus aciditolerans* sp. nov. and *Cryptococcus metallitolerans* sp. nov., a new ecoclade of anamorphic basidiomycetous yeast species from an extreme environment associated with acid rock drainage in São Domingos pyrite mine, Portugal. *International journal of systematic and evolutionary microbiology*, 59(9), 2375-2379.
- Gadd, G. M. (1992). Metals and microorganisms: a problem of definition. *FEMS microbiology letters*, 100(1-3), 197-203.
- Gadd, G. M., & Raven, J. A. (2010). Geomicrobiology of eukaryotic microorganisms. *Geomicrobiology Journal*, 27(6-7), 491-519.
- Gammons, C. H., Doolittle, M. F., Eastman, K. A., & Poulson, S. R. (2021). Geochemistry of natural acid rock drainage in the Mt Evans area, Anaconda–Pintler Range, Montana, USA. *Geochemistry: Exploration, Environment, Analysis*, 21(4), geochem2021-068.

- García, R., Pulido, V., Orellana-Muñoz, S., Nombela, C., Vázquez de Aldana, C. R., Rodríguez-Peña, J. M., & Arroyo, J. (2019). Signalling through the yeast MAPK Cell Wall Integrity pathway controls P-body assembly upon cell wall stress. *Scientific reports*, 9(1), 3186.
- García-Fernández, J. M., de Marsac, N. T., & Diez, J. (2004). Streamlined regulation and gene loss as adaptive mechanisms in *Prochlorococcus* for optimized nitrogen utilization in oligotrophic environments. *Microbiology and Molecular Biology Reviews*, 68(4), 630-638.
- Gatti, X., de Bettignies, G., Claret, S., Doignon, F., Crouzet, M., & Thoraval, D. (2005). RGD1, encoding a RhoGAP involved in low-pH survival, is an Msn2p/Msn4p regulated gene in *Saccharomyces cerevisiae*. *Gene*, 351, 159-169.
- Georlette, D., Blaise, V., Collins, T., D'Amico, S., Gratia, E., Hoyoux, A., Marx, J. C., Sonan, G., Feller, G. & Gerday, C. (2004). Some like it cold: biocatalysis at low temperatures. *FEMS microbiology reviews*, 28(1), 25-42.
- Gerday, C., Aittaleb, M., Bentahir, M., Chessa, J. P., Claverie, P., Collins, T., D'Amico, S., Dumont, J., Garsoux, G., Georlette, D. & Feller, G. (2000). Cold-adapted enzymes: from fundamentals to biotechnology. *Trends in biotechnology*, 18(3), 103-107.
- Gerstein, A. C., Ono, J., Lo, D. S., Campbell, M. L., Kuzmin, A., & Otto, S. P. (2015). Too much of a good thing: the unique and repeated paths toward copper adaptation. *Genetics*, 199(2), 555-571.
- Gianese, G., Argos, P., & Pascarella, S. (2001). Structural adaptation of enzymes to low temperatures. *Protein engineering*, 14(3), 141-148.
- Gientka, I., Błażej, S., Stasiak-Róžańska, L., & Chlebowska-Śmigiel, A. (2015). Exopolysaccharides from yeast: insight into optimal conditions for biosynthesis, chemical composition and functional properties-review. *Acta scientiarum polonorum. Technologia alimentaria*, 14(4).
- Gilichinsky, D. A., Wilson, G. S., Friedmann, E. I., McKay, C. P., Sletten, R. S., Rivkina, E. M., Vishnivetskaya, T. A., Erokhina, L. G., Ivanushkina, N. E., Kochkina, G. A. & Tiedje, J. M. (2007). Microbial populations in Antarctic permafrost: biodiversity, state, age, and implication for astrobiology. *Astrobiology*, 7(2), 275-311.
- Gillespie, J. H. (1991). DNA evolution. In *The causes of molecular evolution* (Vol. 2) (pp. 64-104). Oxford University Press.
- Gnerre, S., MacCallum, I., Przybylski, D., Ribeiro, F. J., Burton, J. N., Walker, B. J., Sharpe, T., Hall, G., Shea, T. P., Sykes, S. & Jaffe, D. B. (2011). High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proceedings of the National Academy of Sciences*, 108(4), 1513-1518.
- Godinho, C. P., Palma, M., Oliveira, J., Mota, M. N., Antunes, M., Teixeira, M. C., Monteiro, P.T. & Sá-Correia, I. (2021). The NC Yeasttract and CommunityYeasttract databases to study gene and genomic transcription regulation in non-conventional yeasts. *FEMS yeast research*, 21(6), foab045.
- Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J. D., Jacq, C., Johnston, M. & Oliver, S. G. (1996). Life with 6000 genes. *Science*, 274(5287), 546-567.
- Gollery, M., Harper, J., Cushman, J., Mittler, T., Girke, T., Zhu, J. K., Bailey-Serres, J. & Mittler, R. (2006). What makes species unique? The contribution of proteins with obscure features. *Genome biology*, 7(7), 1-8.
- González, C., Yanquepe, M., Cardenas, J. P., Valdes, J., Quatrini, R., Holmes, D. S., & Dopson, M. (2014). Genetic variability of psychrotolerant *Acidithiobacillus ferrivorans* revealed by (meta) genomic analysis. *Research in microbiology*, 165(9), 726-734.
- González-Toril, E., Llobet-Brossa, E., Casamayor, E. O., Amann, R., & Amils, R. (2003). Microbial ecology of an extreme acidic environment, the Tinto River. *Applied and environmental microbiology*, 69(8), 4853-4865.

- González-Toril, E., & Aguilera, Á. (2019). Microbial ecology in extreme acidic environments: Use of molecular tools. In *Microbial diversity in the genomic era* (pp. 227-238). Academic Press.
- Gostinčar, C., Turk, M., & Gunde-Cimerman, N. (2019). Environmental impacts on fatty acid composition of fungal membranes. In *Fungi from different environments* (pp. 278-325). CRC Press.
- Goto, S. (1963). On a new yeast genus *Naganishia*. *J Ferment Technol*, 41, 459-462.
- Grande, J. A., De La Torre, M. L., Santisteban, M., & Fortes, J. C. (2018). Hydrochemical characterization and evaluation of the impact of AMD processes on river basin areas in the Iberian Pyrite Belt. *Water Policy*, 20(1), 146-157.
- Gremme, G., Steinbiss, S., & Kurtz, S. (2013). GenomeTools: a comprehensive software library for efficient processing of structured genome annotations. *IEEE/ACM transactions on computational biology and bioinformatics*, 10(3), 645-656.
- Grigoriev, I. V., Nikitin, R., Haridas, S., Kuo, A., Ohm, R., Otilar, R., Riley, R., Salamov, A., Zhao, X., Korzeniewski, F. & Shabalov, I. (2014). MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic acids research*, 42(D1), D699-D704.
- Grille, S., Zaslowski, A., Thiele, S., Plat, J., & Warnecke, D. (2010). The functions of steryl glycosides come to those who wait: Recent advances in plants, fungi, bacteria and animals. *Progress in lipid research*, 49(3), 262-288.
- Gross, S., & Robbins, E. I. (2000). Acidophilic and acid-tolerant fungi and yeasts. *Hydrobiologia*, 433(1-3), 91-109.
- Grujić, S., Vasić, S., Radojević, I., Čomić, L., & Ostojić, A. (2017). Comparison of the *Rhodotorula mucilaginosa* biofilm and planktonic culture on heavy metal susceptibility and removal potential. *Water, Air, & Soil Pollution*, 228, 1-8.
- Gruzdev, E. V., Beletsky, A. V., Kadnikov, V. V., Mardanov, A. V., Ivanov, M. V., Karnachuk, O. V., & Ravin, N. V. (2020). Diversity of eukaryotic microorganisms in the drainage waters of a coal open-cast mine. *Microbiology*, 89, 641-646.
- Grzymalski, J. J., Carter, B. J., DeLong, E. F., Feldman, R. A., Ghadiri, A., & Murray, A. E. (2006). Comparative genomics of DNA fragments from six Antarctic marine planktonic bacteria. *Applied and Environmental Microbiology*, 72(2), 1532-1541.
- Guibaud, G., Bhatia, D., d'Abzac, P., Bourven, I., Bordas, F., van Hullebusch, E. D., & Lens, P. N. (2012). Cd (II) and Pb (II) sorption by extracellular polymeric substances (EPS) extracted from anaerobic granular biofilms: evidence of a pH sorption-edge. *Journal of the Taiwan Institute of Chemical Engineers*, 43(3), 444-449.
- Guilger-Casagrande, M., & Lima, R. D. (2019). Synthesis of silver nanoparticles mediated by fungi: a review. *Frontiers in bioengineering and biotechnology*, 7, 287.
- Gunde-Cimerman, N., Plemenitaš, A., & Buzzini, P. (2014). Changes in lipids composition and fluidity of yeast plasma membrane as response to cold. *Cold-adapted Yeasts: Biodiversity, Adaptation Strategies and Biotechnological Significance*, 225-242.
- Gupta, A., Sarkar, J., & Sar, P. (2019). Understanding the structure and function of extreme microbiome through genomics: Scope and challenges. In *Microbial Diversity in the Genomic Era* (pp. 581-610). Academic Press.
- Gutierrez, J. C., Amaro, F., & Martín-González, A. (2009). From heavy metal-binders to biosensors: Ciliate metallothioneins discussed. *BioEssays*, 31(7), 805-816.
- Gutiérrez, J. C., de Francisco, P., Amaro, F., Díaz, S., & Martín-González, A. (2019). Structural and functional diversity of microbial metallothionein genes. *Microbial Diversity in the Genomic Era*, 387-407.

- Guzmán, L. (2016). *Distribución filogenética de la producción de micosporinas en los phyla Ascomycota y Basidiomycota*. [Undergraduate dissertation, Universidad Nacional del Comahue].
- Hamer, D. H. (1986). Metallothionein. *Annual review of biochemistry*, 55(1), 913-951.
- Hanson, A. D., Pribat, A., Waller, J. C., & Crécy-Lagard, V. D. (2010). 'Unknown' proteins and 'orphan' enzymes: the missing half of the engineering parts list—and how to find it. *Biochemical Journal*, 425(1), 1-11.
- Hashim, N. H. F., Sulaiman, S., Abu Bakar, F. D., Illias, R. M., Kawahara, H., Najimudin, N., Mahadi, N. M. & Murad, A. M. A. (2014). Molecular cloning, expression and characterisation of Afp4, an antifreeze protein from *Glaciozyma antarctica*. *Polar biology*, 37, 1495-1505.
- Harrison, J. J., Rabiei, M., Turner, R. J., Badry, E. A., Sproule, K. M., & Ceri, H. (2006). Metal resistance in *Candida* biofilms. *FEMS microbiology ecology*, 55(3), 479-491.
- Hassan, N., Rafiq, M., Rehman, M., Sajjad, W., Hasan, F., & Abdullah, S. (2019). Fungi in acidic fire: A potential source of industrially important enzymes. *Fungal Biology Reviews*, 33(1), 58-71.
- Hébraud, M., & Potier, P. (1999). Cold shock response and low temperature adaptation in psychrotrophic bacteria. *Journal of molecular microbiology and biotechnology*, 1(2), 211-219.
- Hou, S., Makarova, K. S., Saw, J. H., Senin, P., Ly, B. V., Zhou, Z., Ren, Y., Wang, J., Galperin, M. Y., Omelchenko, M. V. & Alam, M. (2008). Complete genome sequence of the extremely acidophilic methanotroph isolate V4, *Methylophilum inferorum*, a representative of the bacterial phylum Verrucomicrobia. *Biology Direct*, 3, 1-25.
- Huerta-Cepas, J., Forslund, K., Coelho, L. P., Szklarczyk, D., Jensen, L. J., Von Mering, C., & Bork, P. (2017). Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. *Molecular biology and evolution*, 34(8), 2115-2122.
- Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S. K., Cook, H., Mende, D. R., Letunic, I., Rattei, T., Jensen, L. J. & Bork, P. (2019). eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic acids research*, 47(D1), D309-D314.
- Igiri, B. E., Okoduwa, S. I., Idoko, G. O., Akabuogu, E. P., Adeyi, A. O., & Ejiogu, I. K. (2018). Toxicity and bioremediation of heavy metals contaminated ecosystem from tannery wastewater: a review. *Journal of toxicology*, 2018.
- Innan, H., & Kondrashov, F. (2010). The evolution of gene duplications: classifying and distinguishing between models. *Nature Reviews Genetics*, 11(2), 97-108.
- Irwin, J. A. (2020). Overview of extremophiles and their food and medical applications. In *Physiological and biotechnological aspects of extremophiles* (pp. 65-87). Academic Press.
- Jian, X., Li, S., Feng, Y., Chen, X., Kuang, R., Li, B., & Sun, Y. (2020). Influence of synthesis methods on the high-efficiency removal of Cr (VI) from aqueous solution by Fe-modified magnetic biochars. *ACS omega*, 5(48), 31234-31243.
- Johnson, D. B. (1998). Biodiversity and ecology of acidophilic microorganisms. *FEMS microbiology ecology*, 27(4), 307-317.
- Johnson, D. B. (2001). Importance of microbial ecology in the development of new mineral technologies. *Hydrometallurgy*, 59(2-3), 147-157.
- Johnson, D. B., & Hallberg, K. B. (2003). The microbiology of acidic mine waters. *Research in microbiology*, 154(7), 466-473.

- Johnson, D. B., & Bangor, U. K. (2009). Extremophiles: acidic environments. *The desk encyclopedia of microbiology*, 463-480.
- Johnson, D. B. (2014). Physiology and Ecology of acidophilic microorganisms. In *ASM Press eBooks* (pp. 255–270).
- Johnson, D. B., & Quatrini, R. (2020). Acidophile microbiology in space and time. *Current issues in molecular biology*, 39(1), 63-76.
- Kaback, H. R., Sahin-Tóth, M., & Weinglass, A. B. (2001). The kamikaze approach to membrane transport. *Nature reviews Molecular cell biology*, 2(8), 610-620.
- Kachalkin, A. V., Turchetti, B., Inácio, J., Carvalho, C., Mašínová, T., Pontes, A., Röhl, O., Glushakova, A.M., Akulov, A., Baldrian, P., Begerow, D., Buzzini, P., Sampaio, J.P. & Yurkov, A. M. (2019). Rare and undersampled dimorphic basidiomycetes. *Mycological Progress*, 18, 945-971.
- Kan, G., Wang, X., Jiang, J., Zhang, C., Chi, M., Ju, Y., & Shi, C. (2019). Copper stress response in yeast *Rhodotorula mucilaginosa* AN 5 isolated from sea ice, Antarctic. *Microbiologyopen*, 8(3), e00657.
- Kandrór, O., Bretschneider, N., Kreydin, E., Cavalieri, D., & Goldberg, A. L. (2004). Yeast adapt to near-freezing temperatures by STRE/Msn2, 4-dependent induction of trehalose synthesis and certain molecular chaperones. *Molecular cell*, 13(6), 771-781.
- Kaszycki, P., Fedorovych, D., Ksheminska, H., Babyak, L., Wójcik, D., & Koloczek, H. (2004a). Chromium accumulation by living yeast at various environmental conditions. *Microbiological Research*, 159(1), 11-17.
- Kaszycki, P., Fedorovych, D., Ksheminska, H., Babyak, L., Wójcik, D., & Koloczek, H. (2004b). Chromium accumulation by living yeast at various environmental conditions. *Microbiological Research*, 159(1), 11-17.
- Kelkar, Y. D., & Ochman, H. (2012). Causes and consequences of genome expansion in fungi. *Genome Biology and Evolution*, 4(1), 13-23.
- Kent, W. J. (2002). BLAT—the BLAST-like alignment tool. *Genome research*, 12(4), 656-664.
- Khatiwada, B., Hasan, M. T., & Sun, A. (2020). Proteomic response of *Euglena gracilis* to heavy metal exposure—identification of key proteins involved in heavy metal tolerance and accumulation. *Algal Res* 45: 101764.
- Kim, H. J., Lee, J. H., Do, H., & Jung, W. (2014). Production of antifreeze proteins by cold-adapted yeasts. *Cold-adapted yeasts: Biodiversity, adaptation strategies and biotechnological significance*, 259-280.
- Klingler, J. M., & Vishniac, H. S. (1988). 1.3 A Contribution of Antarctic Ecology to Yeast Systematics. *Polarforschung*, 58(2/3), 83-92.
- Kodedova, M., & Sychrova, H. (2015). Changes in the sterol composition of the plasma membrane affect membrane potential, salt tolerance and the activity of multidrug resistance pumps in *Saccharomyces cerevisiae*. *PLoS One*, 10(9), e0139306.
- Kondrashov, F. A. (2012). Gene duplication as a mechanism of genomic adaptation to a changing environment. *Proceedings of the Royal Society B: Biological Sciences*, 279(1749), 5048-5057.
- Koonin, E. V., Aravind, L., & Kondrashov, A. S. (2000). The impact of comparative genomics on our understanding of evolution. *Cell*, 101(6), 573-576.
- Koskiniemi, S., Sun, S., Berg, O. G., & Andersson, D. I. (2012). Selection-driven gene loss in bacteria. *PLoS genetics*, 8(6), e1002787.
- Kozhina, T. N., Kozhin, S. A., & Korolev, V. G. (2011). Gene RAD31 is identical to gene MEC1 of yeast *Saccharomyces cerevisiae*. *Russian Journal of Genetics*, 47, 533-537.

- Krauter, P., Martinelli, R., Williams, K., & Martins, S. (1996). Removal of Cr (VI) from ground water by *Saccharomyces cerevisiae*. *Biodegradation*, 7, 277-286.
- Kryazhimskiy, S., & Plotkin, J. B. (2008). The population genetics of dN/dS. *PLoS genetics*, 4(12), e1000304.
- Ksheminska, H., Fedorovych, D., Honchar, T., Ivash, M., & Gonchar, M. (2008). Yeast tolerance to chromium depends on extracellular chromate reduction and Cr (III) chelation. *Food Technology and Biotechnology*, 46(4), 419-426.
- Kuddus, M. (2018). Cold-active enzymes in food biotechnology: An updated mini review. *Journal of Applied Biology and Biotechnology*, 6(3), 58-63.
- Kulakovskaya, T. (2018). Inorganic polyphosphates and heavy metal resistance in microorganisms. *World Journal of Microbiology and Biotechnology*, 34, 1-8.
- Kumar, A., Alam, A., Tripathi, D., Rani, M., Khatoon, H., Pandey, S., Ehtesham, N. Z. & Hasnain, S. E. (2018, December). Protein adaptations in extremophiles: an insight into extremophilic connection of mycobacterial proteome. In *Seminars in cell & developmental biology* (Vol. 84, pp. 147-157). Academic Press.
- Kurtzman, C. P., Fell, J. W., & Boekhout, T. (2011). Definition, classification and nomenclature of the yeasts. In *The yeasts: a taxonomic study* (pp. 3-5). Elsevier.
- Kuyucak, N. (2002). Acid mine drainage prevention and control options. *CIM bulletin*, 96-102.
- Lang, G. I., Murray, A. W., & Botstein, D. (2009). The cost of gene expression underlies a fitness trade-off in yeast. *Proceedings of the National Academy of Sciences*, 106(14), 5755-5760.
- Lauersen, K. J., Brown, A., Middleton, A., Davies, P. L., & Walker, V. K. (2011). Expression and characterization of an antifreeze protein from the perennial rye grass, *Lolium perenne*. *Cryobiology*, 62(3), 194-201.
- Lee, J. K., Park, K. S., Park, S., Park, H., Song, Y. H., Kang, S. H., & Kim, H. J. (2010). An extracellular ice-binding glycoprotein from an Arctic psychrophilic yeast. *Cryobiology*, 60(2), 222-228.
- Leo, P., de Melo Teixeira, M., Chander, A. M., Singh, N. K., Simpson, A. C., Yurkov, A., Karouia, F., Stajich, J. E., Mason, C. E. & Venkateswaran, K. (2023). Genomic characterization and radiation tolerance of *Naganishia kalamii* sp. nov. and *Cystobasidium onofrii* sp. nov. from Mars 2020 mission assembly facilities. *IMA fungus*, 14(1), 15.
- Levasseur, A., Lomascolo, A., Chabrol, O., Ruiz-Dueñas, F. J., Boukhris-Uzan, E., Piumi, F., Kües, U., Ram, A. F., Murat, C., Haon, M. & Record, E. (2014). The genome of the white-rot fungus *Pycnoporus cinnabarinus*: a basidiomycete model with a versatile arsenal for lignocellulosic biomass breakdown. *BMC genomics*, 15(1), 1-24.
- Li, C., Jiang, W., Ma, N., Zhu, Y., Dong, X., Wang, D., Meng, X. & Xu, Y. (2014). Bioaccumulation of cadmium by growing *Zygosaccharomyces rouxii* and *Saccharomyces cerevisiae*. *Bioresource technology*, 155, 116-121.
- Li, H., Wu, S., Ma, X., Chen, W., Zhang, J., Duan, S., Gao, Y., Kui, L., Huang, W., Wu, P. & Dong, Y. (2018). The genome sequences of 90 mushrooms. *Scientific reports*, 8(1), 9982.
- Li, J., Jiang, Z., Chen, S., Wang, T., Jiang, L., Wang, M., Wang, S. & Li, Z. (2019). Biochemical changes of polysaccharides and proteins within EPS under Pb (II) stress in *Rhodotorula mucilaginosa*. *Ecotoxicology and Environmental Safety*, 174, 484-490.
- Li, Y., Steenwyk, J. L., Chang, Y., Wang, Y., James, T. Y., Stajich, J. E., Groenewald, M., Dunn, C. W., Hittinger, C. T. & Rokas, A. (2021). A genome-scale phylogeny of the kingdom Fungi. *Current Biology*, 31(8), 1653-1665.
- Liang, X., Csetenyi, L., & Gadd, G. M. (2016a). Lead bioprecipitation by yeasts utilizing organic phosphorus substrates. *Geomicrobiology journal*, 33(3-4), 294-307.

- Liang, X., Kierans, M., Ceci, A., Hillier, S., & Gadd, G. M. (2016b). Phosphatase-mediated bioprecipitation of lead by soil fungi. *Environmental microbiology*, 18(1), 219-231
- Libkind, D., Pérez, P., Sommarug, R., del Carmen Diéguez, M., Ferraro, M., Brizzio, S., Zagarese, H. & van Broock, M. (2004). Constitutive and UV-inducible synthesis of photoprotective compounds (carotenoids and mycosporines) by freshwater yeasts. *Photochemical & Photobiological Sciences*, 3, 281-286.
- Libkind, D., Diéguez, M. C., Moliné, M., Pérez, P., Zagarese, H. E., & van Broock, M. (2006). Occurrence of photoprotective compounds in yeasts from freshwater ecosystems of northwestern Patagonia (Argentina). *Photochemistry and photobiology*, 82(4), 972-980.
- Libkind, D., Moliné, M., Sampaio, J. P., & Van Broock, M. (2009). Yeasts from high-altitude lakes: influence of UV radiation. *FEMS microbiology ecology*, 69(3), 353-362.
- Libkind, D., Hittinger, C. T., Valério, E., Gonçalves, C., Dover, J., Johnston, M., Gonçalves, P. & Sampaio, J. P. (2011). Microbe domestication and the identification of the wild genetic stock of lager-brewing yeast. *Proceedings of the National Academy of Sciences*, 108(35), 14539-14544.
- Libkind, D., Russo, G., & van Broock, M. (2014). Yeasts from extreme aquatic environment: Hyperacidic freshwaters. In *Freshwater fungi and fungi-like organisms* (pp. 443– 463). De Gruyter.
- Libkind, D., Čadež, N., Oplente, D. A., Langdon, Q. K., Rosa, C. A., Sampaio, J. P., Gonçalves, P. & Lachance, M. A. (2020). Towards yeast taxogenomics: lessons from novel species descriptions based on complete genome sequences. *FEMS Yeast Research*, 20(6), foaa042.
- Libkind, D., Peris, D., Cubillos, F. A., Steenwyk, J. L., Oplente, D. A., Langdon, Q. K., Rokas, A. & Hittinger, C. T. (2020). Into the wild: new yeast genomes from natural environments and new tools for their analysis. *FEMS yeast research*, 20(2), foaa008.
- Liu, X. Z., Wang, Q. M., Theelen, B., Groenewald, M., Bai, F. Y., & Boekhout, T. (2015). Phylogeny of tremellomycetous yeasts and related dimorphic and filamentous basidiomycetes reconstructed from multiple gene sequence analyses. *Studies in mycology*, 81(1), 1-26.
- López-Archilla, A. I., Marín, I., & Amils, R. (2001). Microbial community composition and ecology of an acidic aquatic environment: the Tinto River, Spain. *Microbial ecology*, 41, 20-35.
- López-Archilla, A. I., González, A. E., Terrón, M. C., & Amils, R. (2004). Ecological study of the fungal populations of the acidic Tinto River in southwestern Spain. *Canadian Journal of Microbiology*, 50(11), 923-934.
- Love, J., Palmer, J., Stajich, J., Esser, T., Kastman, E., & Winter, D. (2018, August 9). *Nextgenusfs/funannotate: Funannotate v1.5.0*. Zenodo. <https://zenodo.org/records/1342272>
- Lu, X., Wu, Q., Zhang, Y., & Xu, Y. (2015). Genomic and transcriptomic analyses of the Chinese Maotai-flavored liquor yeast MT1 revealed its unique multi-carbon co-utilization. *BMC genomics*, 16(1), 1-14.
- Madronich, S., McKenzie, R. L., Björn, L. O., & Caldwell, M. M. (1998). Changes in biologically active ultraviolet radiation reaching the Earth's surface. *Journal of Photochemistry and Photobiology B: Biology*, 46(1-3), 5-19.
- Madsen, E. L. (2011). Microorganisms and their roles in fundamental biogeochemical cycles. *Current opinion in biotechnology*, 22(3), 456-464.
- Margesi R (2002) Cold-active enzymes as new tools in biotechnology. In: *Extremophiles* (pp. 164-184). Oxford (in press)
- Margesi, R., & Miteva, V. (2011). Diversity and ecology of psychrophilic microorganisms. *Research in microbiology*, 162(3), 346-361.

- Margesin, R., & Collins, T. (2019). Microbial ecology of the cryosphere (glacial and permafrost habitats): current knowledge. *Applied microbiology and biotechnology*, *103*, 2537-2549.
- Mas, A., Jamshidi, S., Lagadeuc, Y., Eveillard, D., & Vandenkoornhuys, P. (2016). Beyond the black queen hypothesis. *The ISME journal*, *10*(9), 2085-2091.
- McKenzie, R. L., Aucamp, P. J., Bais, A. F., Björn, L. O., Ilyas, M., & Madronich, S. (2011). Ozone depletion and climate change: impacts on UV radiation. *Photochemical & Photobiological Sciences*, *10*(2), 182-198.
- Méndez-Álvarez, S., Rüfenacht, K., & Eggen, R. I. (2000). The oxidative stress-sensitive yap1 null strain of *Saccharomyces cerevisiae* becomes resistant due to increased carotenoid levels upon the introduction of the *Chlamydomonas reinhardtii* cDNA, coding for the 60S ribosomal protein L10a. *Biochemical and biophysical research communications*, *267*(3), 953-959.
- Méndez-García, C., Peláez, A. I., Mesa, V., Sánchez, J., Golyshina, O. V., & Ferrer, M. (2015). Microbial diversity and metabolic networks in acid mine drainage habitats. *Frontiers in microbiology*, *6*, 475.
- Miller, W., Makova, K. D., Nekrutenko, A., & Hardison, R. C. (2004). Comparative genomics. *Annu. Rev. Genomics Hum. Genet.*, *5*, 15-56.
- Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010, November). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In *2010 gateway computing environments workshop (GCE)* (pp. 1-8). IEEE.
- Mirete, S., Morgante, V., & González-Pastor, J. E. (2017). Acidophiles: diversity and mechanisms of adaptation to acidic environments. *Adaption of Microbial Life to Environmental Extremes: Novel Research Results and Application*, 227-251.
- Mohanta, T. K., & Bae, H. (2015). The diversity of fungal genome. *Biological procedures online*, *17*, 1-9.
- Mohapatra, S., Bohidar, S., Pradhan, N., Kar, R. N., & Sukla, L. B. (2007). Microbial extraction of nickel from Sukinda chromite overburden by *Acidithiobacillus ferrooxidans* and *Aspergillus* strains. *Hydrometallurgy*, *85*(1), 1-8.
- Moliné, M., Libkind, D., del Carmen Diéguez, M., & van Broock, M. (2009). Photoprotective role of carotenoids in yeasts: response to UV-B of pigmented and naturally-occurring albino strains. *Journal of Photochemistry and Photobiology B: Biology*, *95*(3), 156-161.
- Moliné, M. (2010). *Producción de compuestos fotoprotectores (carotenoides y micosporinas) por levaduras* (Doctoral dissertation, PhD thesis, Universidad Nacional de Tucumán, Tucumán, Argentina).
- Moliné, M., Arbeloa, E. M., Flores, M. R., Libkind, D., Farías, M. E., Bertolotti, S. G., Churio, M. S. & van Broock, M. R. (2011). UVB photoprotective role of mycosporines in yeast: photostability and antioxidant activity of mycosporine-glutaminol-glucoside. *Radiation research*, *175*(1), 44-50.
- Moliné, M., Libkind, D., de Garcia, V., & Giraudo, M. R. (2014). Production of pigments and photo-protective compounds by cold-adapted yeasts. *Cold-adapted yeasts: biodiversity, adaptation strategies and biotechnological significance*, 193-224.
- Morris, J. J., Lenski, R. E., & Zinser, E. R. (2012). The Black Queen Hypothesis: evolution of dependencies through adaptive gene loss. *MBio*, *3*(2), 10-1128.
- Mourato, A., Gadanho, M., Lino, A. R., & Tenreiro, R. (2011). Biosynthesis of crystalline silver and gold nanoparticles by extremophilic yeasts. *Bioinorganic chemistry and applications*, 2011.
- Muñoz, M. (2010). Levaduras y hongos dimórficos del filoplano de *Nothofagus pumilio* y el papel de la exposición solar en su distribución y producción de metabolitos fotoprotectores. (Undergraduate thesis, Universidad Nacional del Comahue, Río Negro, Argentina).

- Murata, N., & Los, D. A. (1997). Membrane fluidity and temperature perception. *Plant physiology*, 115(3), 875.
- Musilova, M., Wright, G., Ward, J. M., & Dartnell, L. R. (2015). Isolation of radiation-resistant bacteria from Mars analog Antarctic Dry Valleys by preselection, and the correlation between radiation and desiccation resistance. *Astrobiology*, 15(12), 1076-1090.
- Muter, O., Millers, D., Grigorjeva, L., Ventina, E., & Rapoport, A. (2001). Cr (VI) sorption by intact and dehydrated *Candida utilis* cells: differences in mechanisms. *Process Biochemistry*, 37(5), 505-511.
- Naranjo-Ortiz, M. A., & Gabaldón, T. (2020). Fungal evolution: cellular, genomic and metabolic complexity. *Biological Reviews*, 95(5), 1198-1232.
- Nicolay, K., Veenhuis, M., Douma, A. C., & Harder, W. (1987). A ³¹P NMR study of the internal pH of yeast peroxisomes. *Archives of microbiology*, 147, 37-41.
- Nishimura, K. E. N., Igarashi, K., & Kakinuma, Y. (1998). Proton gradient-driven nickel uptake by vacuolar membrane vesicles of *Saccharomyces cerevisiae*. *Journal of bacteriology*, 180(7), 1962-1964.
- Nizovoy, P., Moliné, M., Libkind, D., Bellora N. (2018). *Cleaning 'Nwanted regions: a novel approach to getting rid of N content by assemblies combination*. PeerJ Preprints. <https://peerj.com/preprints/3504.pdf>
- Nocete, F., Álex, E., Nieto, J. M., Sáez, R., & Bayona, M. R. (2005). An archaeological approach to regional environmental pollution in the south-western Iberian Peninsula related to Third millennium BC mining and metallurgy. *Journal of Archaeological Science*, 32(10), 1566-1576.
- Nordstrom, D. K., & Southam, G. (1997). Geomicrobiology of sulfide mineral oxidation. In *Geomicrobiology: Interactions between Microbes and Minerals* (pp. 361-390). Reviews in Mineralogy and Geochemistry.
- Nutaratat, P., Boontham, W., & Khunnamwong, P. (2022). A Novel Yeast Genus and Two Novel Species Isolated from Pineapple Leaves in Thailand: *Savitreella phatthalungensis* gen. nov., sp. nov. and *Goffeauzyma siamensis* sp. nov. *Journal of Fungi*, 8(2), 118.
- Ohno, S. (1970). The creation of a new gene from a redundant duplicate of an old gene. In *Evolution by gene duplication* (pp. 71-82). Springer Science & Business Media.
- Olsen, T. M., & Duman, J. G. (1997). Maintenance of the supercooled state in the gut fluid of overwintering pyrochroid beetle larvae, *Dendroides canadensis*: role of ice nucleators and antifreeze proteins. *Journal of Comparative Physiology B*, 167, 114-122.
- Olson, M. V. (1999). When less is more: gene loss as an engine of evolutionary change. *The American Journal of Human Genetics*, 64(1), 18-23.
- Onofri, S., Selbmann, L., de Hoog, G. S., Grube, M., Barreca, D., Ruisi, S., & Zucconi, L. (2007). Evolution and adaptation of fungi at boundaries of life. *Advances in Space Research*, 40(11), 1657-1664.
- Paidhungat, M., & Garrett, S. (1998). Cdc1 and the vacuole coordinately regulate Mn²⁺ homeostasis in the yeast *Saccharomyces cerevisiae*. *Genetics*, 148(4), 1787-1798.
- Pal, A., & Paul, A. (2008). Microbial extracellular polymeric substances: central elements in heavy metal bioremediation. *Indian journal of microbiology*, 48, 49-64.
- Palmiter, R. D. (1998). The elusive function of metallothioneins. *Proceedings of the National Academy of Sciences*, 95(15), 8428-8430.
- Panadero, J., Pallotti, C., Rodríguez-Vargas, S., Randez-Gil, F., & Prieto, J. A. (2006). A downshift in temperature activates the high osmolarity glycerol (HOG) pathway, which determines freeze tolerance in *Saccharomyces cerevisiae*. *Journal of Biological Chemistry*, 281(8), 4638-4645.

- Parra, G., Bradnam, K., Ning, Z., Keane, T., & Korf, I. (2009). Assessing the gene space in draft genomes. *Nucleic acids research*, 37(1), 289-297.
- Parra, M., Libkind, D., Hittinger, C. T., Álvarez, L., & Bellora, N. (2023). Assembly and comparative genome analysis of a Patagonian *Aureobasidium pullulans* isolate reveals unexpected intraspecific variation. *Yeast*, 1-17.
- Park, K. S., Do, H., Lee, J. H., Park, S. I., Jung Kim, E., Kim, S. J., Kang, S. H. & Kim, H. J. (2012). Characterization of the ice-binding protein from Arctic yeast *Leucosporidium* sp. AY30. *Cryobiology*, 64(3), 286-296.
- Pearson, W. R. (2013). An introduction to sequence similarity (“homology”) searching. *Current protocols in bioinformatics*, 42(1), 3-1.
- Pedrozo, F. L., Díaz, M. M., Temporetti, P. F., Baffico, G. D., & Beamud, S. G. (2010). Características limnológicas de un sistema ácido: río Agrio-Lago Caviahue, Provincia del Neuquén, Argentina. *Ecología austral*, 20(2), 173-184.
- Penninckx, M. J. (2002). An overview on glutathione in *Saccharomyces* versus non-conventional yeasts. *FEMS yeast research*, 2(3), 295-305.
- Péter, G., Takashima, M., & Čadež, N. (2017). Yeast habitats: different but global. *Yeasts in natural ecosystems: ecology*, 39-71.
- Pettersson, O. V. & Leong, S. L. (2011). Fungal xerophiles (osmophiles). In *Encyclopedia of Life Science (Microbiology)* (pp. 1–9). John Wiley and Sons Ltd.
- Phadtare, S., Inouye, M., & Severinov, K. (2004). The mechanism of nucleic acid melting by a CspA family protein. *Journal of molecular biology*, 337(1), 147-155.
- Pocsi, I., Prade, R. A., & Penninckx, M. J. (2004). Glutathione, altruistic metabolite in fungi. *Advances in microbial physiology*, 49(1), 1-76.
- Polvi, E. J., Li, X., O’Meara, T. R., Leach, M. D., & Cowen, L. E. (2015). Opportunistic yeast pathogens: reservoirs, virulence mechanisms, and therapeutic strategies. *Cellular and Molecular Life Sciences*, 72, 2261-2287.
- Quatrini, R., Lefimil, C., Holmes, D. S., & Jedlicki, E. (2005). The ferric iron uptake regulator (Fur) from the extreme acidophile *Acidithiobacillus ferrooxidans*. *Microbiology*, 151(6), 2005-2015.
- Quatrini, R., Lefimil, C., Veloso, F. A., Pedrosa, I., Holmes, D. S., & Jedlicki, E. (2007). Bioinformatic prediction and experimental verification of Fur-regulated genes in the extreme acidophile *Acidithiobacillus ferrooxidans*. *Nucleic acids research*, 35(7), 2153-2166.
- Radecka, D., Mukherjee, V., Mateo, R. Q., Stojiljkovic, M., Foulquié-Moreno, M. R., & Thevelein, J. M. (2015). Looking beyond *Saccharomyces*: the potential of non-conventional yeast species for desirable traits in bioethanol fermentation. *FEMS yeast research*, 15(6), fov053.
- Rahbar Saadat, Y., Yari Khosroushahi, A., & Pourghassem Gargari, B. (2021). Yeast exopolysaccharides and their physiological functions. *Folia Microbiologica*, 66, 171-182.
- Raspor, P., & Zupan, J. (2006). Yeasts in extreme environments. In *Biodiversity and ecophysiology of yeasts* (pp. 371-417). Springer.
- Ratnakumar, A., Mousset, S., Glémin, S., Berglund, J., Galtier, N., Duret, L., & Webster, M. T. (2010). Detecting positive selection within genomes: the problem of biased gene conversion. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1552), 2571-2580.
- Richards, T. A., Leonard, G., Soanes, D. M., & Talbot, N. J. (2011). Gene transfer into the fungi. *Fungal Biology Reviews*, 25(2), 98-110.

- Riley, R., Haridas, S., Wolfe, K. H., Lopes, M. R., Hittinger, C. T., Göker, M., Salamov, A. A., Wisecaver, J. H., Long, T. M., Calvey, C. H. & Jeffries, T. W. (2016). Comparative genomics of biotechnologically important yeasts. *Proceedings of the National Academy of Sciences*, 113(35), 9882-9887.
- Roane, T. M., Pepper, I. L., & Gentry, T. J. (2015). Microorganisms and metal pollutants. In *Environmental microbiology* (pp. 415-439). Academic Press.
- Rocha, E. P., Smith, J. M., Hurst, L. D., Holden, M. T., Cooper, J. E., Smith, N. H., & Feil, E. J. (2006). Comparisons of dN/dS are time dependent for closely related bacterial genomes. *Journal of theoretical biology*, 239(2), 226-235.
- Rodrigues, M. L., Nimrichter, L., Oliveira, D. L., Frases, S., Miranda, K., Zaragoza, O., Alvarez, M., Nakouzi, A., Feldmesser, M. & Casadevall, A. (2007). Vesicular polysaccharide export in *Cryptococcus neoformans* is a eukaryotic solution to the problem of fungal trans-cell wall transport. *Eukaryotic cell*, 6(1), 48-59.
- Rodríguez-Peña, J. M., García, R., Nombela, C., & Arroyo, J. (2010). The high-osmolarity glycerol (HOG) and cell wall integrity (CWI) signalling pathways interplay: a yeast dialogue between MAPK routes. *Yeast*, 27(8), 495-502.
- Rong, X., & McSpadden Gardener, B. B. (2013). Draft genome sequence of *Cryptococcus flavescens* strain OH182.9_3C, a biocontrol agent against fusarium head blight of wheat. *Genome announcements*, 1(5), 10-1128.
- Rossit, A. R. B., de Almeida, M. T. G., Nogueira, C. A. M., da Costa Oliveira, J. G., Barbosa, D. M. U., Moscardini, A. C., Mascarenhas, J. D. A. P., Gabbay, Y. B., Marques, F. R., Cardoso, L. V. & Machado, R. L. D. (2007). Bacterial, yeast, parasitic, and viral enteropathogens in HIV-infected children from Sao Paulo State, Southeastern Brazil. *Diagnostic microbiology and infectious disease*, 57(1), 59-66.
- Rothschild, L. J., & Mancinelli, R. L. (2001). Life in extreme environments. *Nature*, 409(6823), 1092-1101.
- Roychoudhury, A. (2020). Yeast-mediated green synthesis of nanoparticles for biological applications. *Indian J Pharm Biol Res*, 8(03), 26-31.
- Rudakiya, D. M., & Patel, Y. (2021). Bioremediation of metals, metalloids, and nonmetals. In *Microbial Rejuvenation of Polluted Environment: Volume 2*, (pp. 33-49). Springer.
- Rudolph, A. S., Crowe, J. H., & Crowe, L. M. (1986). Effects of three stabilizing agents—proline, betaine, and trehalose—on membrane phospholipids. *Archives of biochemistry and biophysics*, 245(1), 134-143.
- Ruotolo, R., Marchini, G., & Ottonello, S. (2008). Membrane transporters and protein traffic networks differentially affecting metal tolerance: a genomic phenotyping study in yeast. *Genome biology*, 9, 1-19.
- Russell, N. J. (2008). Membrane components and cold sensing. In *Psychrophiles: from biodiversity to biotechnology* (pp. 177-190). Springer.
- Russo, G., Libkind, D., Sampaio, J. P., & Van Broock, M. R. (2008). Yeast diversity in the acidic Rio Agrio–Lake Caviahue volcanic environment (Patagonia, Argentina). *FEMS microbiology ecology*, 65(3), 415-424.
- Russo, G., Libkind, D., Ulloa, R. J., de García, V., Sampaio, J. P., & van Broock, M. R. (2010). *Cryptococcus agrionensis* sp. nov., a basidiomycetous yeast of the acidic rock drainage ecoclade, isolated from an acidic aquatic environment of volcanic origin. *International journal of systematic and evolutionary microbiology*, 60(4), 996-1000.
- Russo, G. (2011). *Levaduras de ambientes acuáticos ácidos del Parque Provincial Caviahue-Copahue. Estudios básicos y aplicados*. [Doctoral dissertation, Universidad Nacional del Comahue].
- Sælensminde, G., Halskau, Ø., & Jonassen, I. (2009). Amino acid contacts in proteins adapted to different temperatures: hydrophobic interactions and surface charges play a key role. *Extremophiles*, 13, 11-20.
- Salathé, M., Ackermann, M., & Bonhoeffer, S. (2005). The effect of multifunctionality on the rate of evolution in yeast. *Molecular biology and evolution*, 23(4), 721-722.

- Salazar, M., Vongsangnak, W., Panagiotou, G., Andersen, M. R., & Nielsen, J. (2009). Uncovering transcriptional regulation of glycerol metabolism in *Aspergilli* through genome-wide gene expression data analysis. *Molecular Genetics and Genomics*, 282, 571-586.
- Salvadori, M. R., Ando, R. A., Oller do Nascimento, C. A., & Corrêa, B. (2014). Intracellular biosynthesis and removal of copper nanoparticles by dead biomass of yeast isolated from the wastewater of a mine in the Brazilian Amazonia. *PLoS One*, 9(1), e87968.
- Salwan, R., & Sharma, V. (2022). Genomics of prokaryotic extremophiles to unfold the mystery of survival in extreme environments. *Microbiological Research*, 127156.
- Saunders, N. F., Thomas, T., Curmi, P. M., Mattick, J. S., Kuczek, E., Slade, R., Davis, J., Franzmann, P. D., Boone, D., Rusterholtz, K. & Cavicchioli, R. (2003). Mechanisms of thermal adaptation revealed from the genomes of the Antarctic Archaea *Methanogenium frigidum* and *Methanococoides burtonii*. *Genome research*, 13(7), 1580-1588.
- Savoldelli, V., & Tortorano, A. M. (2009). *Cryptococcus*. In *Clinical Mycology* (pp. 231-249). Churchill Livingstone Elsevier.
- Schade, B., Jansen, G., Whiteway, M., Entian, K. D., & Thomas, D. Y. (2004). Cold adaptation in budding yeast. *Molecular Biology of the Cell*, 15(12), 5492-5502.
- Schleper, C., Puehler, G., Holz, I., Gambacorta, A., Janekovic, D., Santarius, U. T. E., Klenk, H. P. & Zillig, W. (1995). *Picrophilus* gen. nov., fam. nov.: a novel aerobic, heterotrophic, thermoacidophilic genus and family comprising archaea capable of growth around pH 0. *Journal of bacteriology*, 177(24), 7050-7059.
- Schleper, C., Piihler, G., Kuhlmoorgen, B., & Zillig, W. (1995b). Life at extremely low pH. *Nature*, 375(6534), 741-742.
- Schmidt, S. K., Vimercati, L., Darcy, J. L., Arán, P., Gendron, E. M., Solon, A. J., Porazinska, D. & Dorador, C. (2017). A *Naganishia* in high places: functioning populations or dormant cells from the atmosphere?. *Mycology*, 8(3), 153-163.
- Schuster-Böckler, B., Conrad, D., & Bateman, A. (2010). Dosage sensitivity shapes the evolution of copy-number varied regions. *PloS one*, 5(3), e9474.
- Scorzetti, G., Petrescu, I., Yarrow, D., & Fell, J. W. (2000). *Cryptococcus adeliensis* sp. nov., a xylanase producing basidiomycetous yeast from Antarctica. *Antonie van Leeuwenhoek*, 77, 153-157.
- Scorzetti, G., Fell, J. W., Fonseca, A., & Statzell-Tallman, A. (2002). Systematics of basidiomycetous yeasts: a comparison of large subunit D1/D2 and internal transcribed spacer rDNA regions. *FEMS yeast research*, 2(4), 495-517.
- Sedmak, J. J., Weerasinghe, D. K., & Jolly, S. O. (1990). Extraction and quantitation of astaxanthin from *Phaffia rhodozyma*. *Biotechnology techniques*, 4, 107-112.
- Segal-Kischinevzky, C., Romero-Aguilar, L., Alcaraz, L. D., López-Ortiz, G., Martínez-Castillo, B., Torres-Ramírez, N., Sandoval, G. & González, J. (2022). Yeasts inhabiting extreme environments and their biotechnological applications. *Microorganisms*, 10(4), 794.
- Selbmann, L., Zucconi, L., Onofri, S., Cecchini, C., Isola, D., Turchetti, B., & Buzzini, P. (2014). Taxonomic and phenotypic characterization of yeasts isolated from worldwide cold rock-associated habitats. *Fungal Biology*, 118(1), 61-71.
- Sharma, A., Parashar, D., & Satyanarayana, T. (2016). Acidophilic microbes: biology and applications. In *Biotechnology of extremophiles: advances and challenges* (pp. 215-241). Springer.
- Shelford, V. E. (1913). Animal Communities in Temperate America: As Illustrated in the Chicago Region; a Study in Animal Ecology (No. 5). *Geographic Society of Chicago*.

- Shen, X. X., Opulente, D. A., Kominek, J., Zhou, X., Steenwyk, J. L., Buh, K. V., Haase, M. A., Wisecaver, J. H., Wang, M., Doering, D. T. & Rokas, A. (2018). Tempo and mode of genome evolution in the budding yeast subphylum. *Cell*, 175(6), 1533-1545.
- Shi, J., Cao, Y., Fan, X., Li, M., Wang, Y., & Ming, F. (2012). A rice microsomal delta-12 fatty acid desaturase can enhance resistance to cold stress in yeast and *Oryza sativa*. *Molecular Breeding*, 29, 743-757.
- Siddiqui, K. S., & Cavicchioli, R. (2006). Cold-adapted enzymes. *Annu. Rev. Biochem.*, 75, 403-433.
- Singh, A., Krishnan, K. P., Prabakaran, D., & Sinha, R. K. (2017). Lipid membrane modulation and pigmentation: a cryoprotection mechanism in Arctic pigmented bacteria. *Journal of basic microbiology*, 57(9), 770-780.
- Sliwa, P., & Korona, R. (2005). Loss of dispensable genes is not adaptive in yeast. *Proceedings of the National Academy of Sciences*, 102(49), 17670-17674.
- Spear, J. R., Walker, J. J., McCollom, T. M., & Pace, N. R. (2005). Hydrogen and bioenergetics in the Yellowstone geothermal ecosystem. *Proceedings of the National Academy of Sciences*, 102(7), 2555-2560.
- Spielman, S. J., & Wilke, C. O. (2015). The relationship between dN/dS and scaled selection coefficients. *Molecular biology and evolution*, 32(4), 1097-1108.
- Spiwok, V., Lipovová, P., Skálová, T., Dušková, J., Dohnálek, J., Hašek, J., Russell, N. J. & Králová, B. (2007). Cold-active enzymes studied by comparative molecular dynamics simulation. *Journal of Molecular Modeling*, 13, 485-497.
- Stajich, J. E., Harris, T., Brunk, B. P., Brestelli, J., Fischer, S., Harb, O. S., Kissinger, J. C., Li, W., Nayak, V., Pinney, D. F. & Roos, D. S. (2012). FungiDB: an integrated functional genomics database for fungi. *Nucleic acids research*, 40(D1), D675-D681.
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9), 1312-1313.
- Steinway, S. N., Dannenfelser, R., Laucius, C. D., Hayes, J. E., & Nayak, S. (2010). JCoDA: a tool for detecting evolutionary selection. *BMC bioinformatics*, 11, 1-9.
- Straw, M. L., Chaplin, A. K., Hough, M. A., Paps, J., Bavro, V. N., Wilson, M. T., Vijgenboom, E. & Worrall, J. A. (2018). A cytosolic copper storage protein provides a second level of copper tolerance in *Streptomyces lividans*. *Metallomics*, 10(1), 180-193.
- Su, Y., Jiang, X., Wu, W., Wang, M., Hamid, M. I., Xiang, M., & Liu, X. (2016). Genomic, transcriptomic, and proteomic analysis provide insights into the cold adaptation mechanism of the obligate psychrophilic fungus *Mrakia psychrophila*. *G3: Genes, Genomes, Genetics*, 6(11), 3603-3613.
- Suh, J. K., Poulsen, L. L., Ziegler, D. M., & Robertus, J. D. (1999). Yeast flavin-containing monooxygenase generates oxidizing equivalents that control protein folding in the endoplasmic reticulum. *Proceedings of the National Academy of Sciences*, 96(6), 2687-2691.
- Suyama, M., Torrents, D., & Bork, P. (2006). PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic acids research*, 34(suppl_2), W609-W612.
- Talke, I. N., Hanikenne, M., & Krämer, U. (2006). Zinc-dependent global transcriptional control, transcriptional deregulation, and higher gene copy number for genes in metal homeostasis of the hyperaccumulator *Arabidopsis halleri*. *Plant physiology*, 142(1), 148-167.
- Tamás, M. J., & Wysocki, R. (2001). Mechanisms involved in metalloloid transport and tolerance acquisition. *Current genetics*, 40, 2-12.

- Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: molecular evolutionary genetics analysis version 11. *Molecular biology and evolution*, 38(7), 3022-3027.
- Tan, Y. S., Zhang, R. K., Liu, Z. H., Li, B. Z., & Yuan, Y. J. (2022). Microbial adaptation to enhance stress tolerance. *Frontiers in Microbiology*, 13, 888746.
- Tapia, H., Young, L., Fox, D., Bertozzi, C. R., & Koshland, D. (2015). Increasing intracellular trehalose is sufficient to confer desiccation tolerance to *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences*, 112(19), 6122-6127.
- Tatusov, R. L., Galperin, M. Y., Natale, D. A., & Koonin, E. V. (2000). The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic acids research*, 28(1), 33-36.
- Teitzel, G. M., & Parsek, M. R. (2003). Heavy metal resistance of biofilm and planktonic *Pseudomonas aeruginosa*. *Applied and environmental microbiology*, 69(4), 2313-2320.
- Ter-Hovhannisyan, V., Lomsadze, A., Chernoff, Y. O., & Borodovsky, M. (2008). Gene prediction in novel fungal genomes using an *ab initio* algorithm with unsupervised training. *Genome research*, 18(12), 1979-1990.
- Thatcher, J. W., Shaw, J. M., & Dickinson, W. J. (1998). Marginal fitness contributions of nonessential genes in yeast. *Proceedings of the National Academy of Sciences*, 95(1), 253-257.
- Thieringer, H. A., Jones, P. G., & Inouye, M. (1998). Cold shock and adaptation. *Bioessays*, 20(1), 49-57.
- Tornos, F., López Pamo, E., & Sánchez España, F. (2008). *Una Aproximación Al Patrimonio Geológico de Relevancia Internacional* (pp. 57-64). Instituto Geológico Y Minero De España.
- Tran, K., & Green, E. M. (2019). SET domains and stress: uncovering new functions for yeast *Set4*. *Current genetics*, 65(3), 643-648.
- Tronelli, D., Maugini, E., Bossa, F., & Pascarella, S. (2007). Structural adaptation to low temperatures— analysis of the subunit interface of oligomeric psychrophilic enzymes. *The FEBS journal*, 274(17), 4595-4608.
- Tshisekedi, K. A. (2021). Genome assembly and comparative genome analysis of the basidiomycetous yeast *Naganishia randhawae* isolated from avian guano in South Africa (Doctoral dissertation).
- Tulha, J., Lima, A., Lucas, C., & Ferreira, C. (2010). *Saccharomyces cerevisiae* glycerol/H⁺ symporter Stl1p is essential for cold/near-freeze and freeze stress adaptation. A simple recipe with high biotechnological potential is given. *Microbial cell factories*, 9, 1-7.
- Turchetti, B., Selbmann, L., Blanchette, R. A., Di Mauro, S., Marchegiani, E., Zucconi, L., Arenz B. & Buzzini, P. (2015). *Cryptococcus vaughanmartinae* sp. nov. and *Cryptococcus onofrii* sp. nov.: two new species isolated from worldwide cold environments. *Extremophiles*, 19, 149-159.
- Ullrich, S. R., González, C., Poehlein, A., Tischler, J. S., Daniel, R., Schlömann, M., Holmes, D.S., & Mühling, M. (2016). Gene loss and horizontal gene transfer contributed to the genome evolution of the extreme acidophile “Ferrovum”. *Frontiers in microbiology*, 7, 797.
- Vacic, V., Uversky, V. N., Dunker, A. K., & Lonardi, S. (2007). Composition Profiler: a tool for discovery and visualization of amino acid composition differences. *BMC bioinformatics*, 8(1), 1-7.
- Vajpeyi, S., & Chandran, K. (2016). Draft genome sequence of the oleaginous yeast *Cryptococcus albidus* var. *albidus*. *Genome Announcements*, 4(3), 10-1128.
- van Gorcom, R. F., van den Hondel, C. A., & Punt, P. J. (1998). Cytochrome P450 enzyme systems in fungi. *Fungal Genetics and Biology*, 23(1), 1-17.

- Varekamp, J. C., Pasternack, G. B., & Rowe Jr, G. L. (2000). Volcanic lake systematics II. Chemical constraints. *Journal of volcanology and geothermal research*, 97(1-4), 161-179.
- Vergara, E., Neira, G., González, C., Cortez, D., Dopson, M., & Holmes, D. S. (2020). Evolution of predicted acid resistance mechanisms in the extremely acidophilic *Leptospirillum* genus. *Genes*, 11(4), 389.
- Vido, K., Spector, D., Lagniel, G., Lopez, S., Toledano, M. B., & Labarre, J. (2001). A proteome analysis of the cadmium response in *Saccharomyces cerevisiae*. *Journal of Biological Chemistry*, 276(11), 8469-8474.
- Villegas, L. B., Amoroso, M. J. D. R., & Castellanos, L. I. (2011). Interaction of copper or chromium with yeasts: potential application on polluted environmental clean up. In *Bioremediation: Biotechnology, Engineering and Environmental Management* (pp. 177-206). Nova Science Publishers.
- Vincent, W. F. (2000). Evolutionary origins of Antarctic microbiota: invasion, selection and endemism. *Antarctic Science*, 12(3), 374-385.
- Vishniac, H. S., & Hempfling, W. P. (1979a). *Cryptococcus vishniacii* sp. nov., an Antarctic yeast. *International Journal of Systematic and Evolutionary Microbiology*, 29(2), 153-158.
- Vishniac, H. S., & Hempfling, W. P. (1979b). Evidence of an indigenous microbiota (yeast) in the dry valleys of Antarctica. *Microbiology*, 112(2), 301-314.
- Vishniac, H. S. (1985). *Cryptococcus friedmannii*, a new species of yeast from the Antarctic. *Mycologia*, 77(1), 149-153.
- Vishniac, H. S. (2006). Yeast biodiversity in the Antarctic. In *Biodiversity and ecophysiology of yeasts* (pp. 419-440). Springer.
- Walker, G. M. (1998). Yeast Cytology. In *Yeast physiology and biotechnology* (pp.11-48). John Wiley & Sons.
- Wang, J., & Chen, C. (2009). Biosorbents for heavy metals removal and their future. *Biotechnology advances*, 27(2), 195-226.
- Wang, D., Zhang, Y., Zhang, Z., Zhu, J., & Yu, J. (2010). KaKs_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. *Genomics, proteomics & bioinformatics*, 8(1), 77-80.
- Wang, C., Pakhomova, S., Newcomer, M. E., Christner, B. C., & Luo, B. H. (2017). Structural basis of antifreeze activity of a bacterial multi-domain antifreeze protein. *PLoS One*, 12(11), e0187169.
- Waterhouse, R. M., Seppey, M., Simão, F. A., Manni, M., Ioannidis, P., Klioutchnikov, G., Kriventseva, E. V. & Zdobnov, E. M. (2018). BUSCO applications from quality assessments to gene prediction and phylogenomics. *Molecular biology and evolution*, 35(3), 543-548.
- Weisman, C. M., Murray, A. W., & Eddy, S. R. (2020). Many, but not all, lineage-specific genes can be explained by homology detection failure. *PLoS biology*, 18(11), e3000862.
- Westerdijk Fungal Biodiversity Institute. (n.d.). *Naganishia vishniacii*. Retrieved November 8, 2023, from https://wi.knaw.nl/fungal_table
- White, T. J., Bruns, T., Lee, S. J. W. T., & Taylor, J. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR protocols: a guide to methods and applications*, 18(1), 315-322.
- Whitman, W. B., Coleman, D. C., & Wiebe, W. J. (1998). Prokaryotes: the unseen majority. *Proceedings of the National Academy of Sciences*, 95(12), 6578-6583.
- Wiemken, A. (1990). Trehalose in yeast, stress protectant rather than reserve carbohydrate. *Antonie van Leeuwenhoek*, 58, 209-217.

- Wilson, G. A., Bertrand, N., Patel, Y., Hughes, J. B., Feil, E. J., & Field, D. (2005). Orphans as taxonomically restricted and ecologically important genes. *Microbiology*, 151(8), 2499-2501.
- Wolf, Y. I., & Koonin, E. V. (2013). Genome reduction as the dominant mode of evolution. *Bioessays*, 35(9), 829-837.
- Wolfe, K. H. (2006). Comparative genomics and genome evolution in yeasts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 361(1467), 403-412.
- Wolfe, K. H. (2015). Origin of the yeast whole-genome duplication. *PLoS biology*, 13(8), e1002221.
- Wynn-Williams, D. D. (1990). Ecological aspects of Antarctic microbiology. In *Advances in microbial ecology* (pp. 71-146). Springer.
- Wysocki, R., & Tamás, M. J. (2010). How *Saccharomyces cerevisiae* copes with toxic metals and metalloids. *FEMS microbiology reviews*, 34(6), 925-951.
- Yadav, A. N., Verma, P., Sachan, S. G., Kaushik, R., & Saxena, A. K. (2018). Psychrotrophic microbiomes: molecular diversity and beneficial role in plant growth promotion and soil health. *Microorganisms for Green Revolution: Volume 2: Microbes for Sustainable Agro-ecosystem*, 197-240.
- Yang, Z., & Nielsen, R. (2000). Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Molecular biology and evolution*, 17(1), 32-43.
- Yong, X., Chen, Y., Liu, W., Xu, L., Zhou, J., Wang, S., Chen, P., Ouyang, P. & Zheng, T. (2014). Enhanced cadmium resistance and accumulation in *Pseudomonas putida* KT2440 expressing the phytochelatin synthase gene of *Schizosaccharomyces pombe*. *Letters in applied microbiology*, 58(3), 255-261.
- Yong, X., Yan, Z., Xu, L., Zhou, J., Wu, X., Wu, Y., Li, Y., Chen, Z., Zhou, H., Wei, P. & Jia, H. (2016). Genome sequence of a microbial lipid producing fungus *Cryptococcus albidus* NT2002. *Journal of Biotechnology*, 223, 6-7.
- Yu, H., Li, Q., Shen, X., Zhang, L., Liu, J., Tan, Q., Li, Y., Lv, B. & Shang, X. (2020). Transcriptomic Analysis of Two *Lentinula edodes* Genotypes With Different Cadmium Accumulation Ability. *Frontiers in microbiology*, 11, 558104
- Yue, J. X., Li, J., Aigrain, L., Hallin, J., Persson, K., Oliver, K., Bergström, A., Coupland, P., Warringer, J., Lagomarsino, M. C. & Liti, G. (2017). Contrasting evolutionary genome dynamics between domesticated and wild yeasts. *Nature genetics*, 49(6), 913-924.
- Yurkov, A. M., Kachalkin, A. V., Daniel, H. M., Groenewald, M., Libkind, D., de Garcia, V., Zalar, P., Gouliamova, D. E., Boekhout, T. & Begerer, D. (2015). Two yeast species *Cystobasidium psychroaquaticum* f.a. sp. nov. and *Cystobasidium rietchieii* f.a. sp. nov. isolated from natural environments, and the transfer of *Rhodotorula minuta* clade members to the genus *Cystobasidium*. *Antonie Van Leeuwenhoek*, 107, 173-185.
- Zalar, P., & Gunde-Cimerman, N. (2014). Cold-adapted yeasts in Arctic habitats. In *Cold-adapted yeasts: Biodiversity, adaptation strategies and biotechnological significance*, (pp. 49-74). Springer.
- Zarroca, M., Roqué, C., Linares, R., Salminci, J. G., & Gutiérrez, F. (2021). Natural acid rock drainage in alpine catchments: A side effect of climate warming. *Science of the total environment*, 778, 146070.
- Zhang, Y., & Rao, R. (2010). Beyond ergosterol: linking pH to antifungal mechanisms. *Virulence*, 1(6), 551-554.
- Zhang, X., Liu, X., Liang, Y., Guo, X., Xiao, Y., Ma, L., Miao, B., Liu, H., Peng, D., Huang, W. & Yin, H. (2017). Adaptive evolution of extreme acidophile *Sulfobacillus thermosulfidooxidans* potentially driven by horizontal gene transfer and gene loss. *Applied and environmental microbiology*, 83(7), e03098-16.
- Zhang, X., Liu, C. G., Yang, S. H., Wang, X., Bai, F. W., & Wang, Z. (2022). Benchmarking of long-read sequencing, assemblers and polishers for yeast genome. *Briefings in Bioinformatics*, 23(3), bbac146.

- Zhao, X., Zhou, L., Riaz Rajoka, M. S., Yan, L., Jiang, C., Shao, D., Zhu, J., Shi, J., Huang, Q., Yang, H. & Jin, M. (2018). Fungal silver nanoparticles: synthesis, application and challenges. *Critical reviews in biotechnology*, 38(6), 817-835.
- Zhou, Y., Jia, B. S., Zhou, Y. G., Li, A. H., & Xue, L. (2020). *Naganishia floricola* sp. nov., a novel basidiomycetous yeast species isolated from flowers of *Sorbaria sorbifolia*. *International Journal of Systematic and Evolutionary Microbiology*, 70(8), 4496-4501.
- Zhu, J., Dong, C. H., & Zhu, J. K. (2007). Interplay between cold-responsive gene regulation, metabolism and RNA processing during plant cold acclimation. *Current opinion in plant biology*, 10(3), 290-295.
- Zmasek, C. M., & Eddy, S. R. (2001). A simple algorithm to infer gene duplication and speciation events on a gene tree. *Bioinformatics*, 17(9), 821-828.
- Zomorodian, K., Pourshahid, S., Sadatsharifi, A., Mehryar, P., Pakshir, K., Rahimi, M. J., & Arabi Monfared, A. (2016). Biosynthesis and characterization of silver nanoparticles by *Aspergillus* species. *BioMed research international* 2016.



RESEARCH ARTICLE

Unique genomic traits for cold adaptation in *Naganishia vishniacii*, a polyextremophile yeast isolated from Antarctica

Paula Nizovoy¹, Nicolás Bellora¹, Sajeet Haridas², Hui Sun², Chris Daum², Kerrie Barry², Igor V. Grigoriev^{2,3,†}, Diego Libkind¹, Laurie B. Connell⁴ and Martín Moliné^{1,*‡}

¹Centro de Referencia en Levaduras y Tecnología Cervecera (CRELTEC), Instituto Andino Patagónico de Tecnologías Biológicas y Geoambientales (IPATEC) – CONICET / Universidad Nacional del Comahue, San Carlos de Bariloche, Río Negro 8400, Argentina, ²U.S. Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA 94598, USA, ³Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720, USA and ⁴School of Marine Sciences, University of Maine, Orono, ME 04469, USA

*Corresponding author: Centro de Referencia en Levaduras y Tecnología Cervecera (CRELTEC), Instituto Andino Patagónico de Tecnologías Biológicas y Geoambientales (IPATEC) – CONICET / Universidad Nacional del Comahue, San Carlos de Bariloche, Río Negro 8400, Argentina. Tel.: +54 294 442 3374 (#102); E-mail: mmoline@comahue-conicet.gob.ar

One sentence summary: Through a comparative genomic approach, this study reveals a series of characteristics possibly involved in the evolution of *N. vishniacii* in the context of its extreme environment and its tolerance of extremely cold conditions.

Editor: Prof. Isak Pretorius

[†]Igor V. Grigoriev, <http://orcid.org/0000-0002-3136-8903>

[‡]Martín Moliné, <http://orcid.org/0000-0001-9211-1143>

ABSTRACT

Cold environments impose challenges to organisms. Polyextremophile microorganisms can survive in these conditions thanks to an array of counteracting mechanisms. *Naganishia vishniacii*, a yeast species hitherto only isolated from McMurdo Dry Valleys, Antarctica, is an example of a polyextremophile. Here we present the first draft genomic sequence of *N. vishniacii*. Using comparative genomics, we unraveled unique characteristics of cold associated adaptations. 336 putative genes (total: 6183) encoding solute transfers and chaperones, among others, were absent in sister species. Among genes shared by *N. vishniacii* and its closest related species we found orthologs encompassing possible evidence of positive selection ($dN/dS > 1$). Genes associated with photoprotection were found in agreement with high solar irradiation exposure. Also genes coding for desaturases and genomic features associated with cold tolerance (i.e. trehalose synthesis and lipid metabolism) were explored. Finally, biases in amino acid usage (namely an enrichment of glutamine and a trend in proline reduction) were observed, possibly conferring increased protein flexibility. To the best of our knowledge, such a combination of mechanisms for cold tolerance has not been previously reported in fungi, making *N. vishniacii* a unique model for the study of the genetic basis and evolution of cold adaptation strategies.

Received: 18 June 2020; **Accepted:** 15 October 2020

© The Author(s) 2020. Published by Oxford University Press on behalf of FEMS. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

Keywords: genomics; extremophile; cold adaptation; yeast; *Naganishia*; Antarctica

INTRODUCTION

A large proportion of the Earth's surface area is covered by permanently low temperature environments. Polar regions, Antarctica and areas within the Arctic circle, constitute 20% of the world's land surface area with less than 0.5% seasonally ice- or snow-free. Nevertheless several ecosystems are represented, with a native biota that has been adapted to the extreme conditions of these environments (Cowan *et al.* 2007; Convey and Peck 2019). The combination of low temperature, low liquid water availability, periodic freeze-thaw cycles, increased osmotic and oxidative stresses and high solar irradiation, *inter alia*, make these regions extremely inhospitable to most forms of life. However, such places have been successfully colonized by psychrophilic and psychrotolerant organisms. Current knowledge of microbial biodiversity has shown that cold habitats harbor a wide diversity of adapted prokaryotic and eukaryotic microbial life, including archaea, bacteria, cyanobacteria, protists, microalgae, virus, filamentous fungi and yeasts (Margesin and Miteva 2011; Connell *et al.* 2014; Zalar and Gunde-Cimerman 2014; Margesin and Collins 2019).

To successfully overcome the negative effects of low temperatures, organisms harbor a complex range of structural and functional adaptations. Changes in membrane fluidity by restructuring lipid composition thus counteracting membrane rigidifying, production of pigments and metabolites that act as reactive species scavengers and photoprotective compounds, synthesis of anti freezing proteins enabling survival to freeze-thaw cycles and accumulation of osmolytes that avoid freezing injury, have been reported as common mechanisms in a wide variety of cold tolerant organisms from vertebrates to microorganisms (Clark and Worland 2008; Singh *et al.* 2017; Collins and Margesin 2019). However, many of these responses and mechanisms are not fully characterized in cold tolerant yeasts. Understanding these mechanisms has been a challenge and a matter of interest, not only because of their biotechnological potential, but also because of their importance in biogeochemical nutrient cycling, ecological role and also to gain an understanding of the origins of life on our planet (Hébraud and Potier 1999, Albanese *et al.* 2020).

Yeasts are ubiquitous eukaryotic organisms belonging to the fungi and even though they are considered less tough than prokaryotic microorganisms, a great diversity of them are able to overcome a plethora of extreme conditions (Buzzini and Margesin 2014). *Naganishia vishniacii*, a basidiomycetous yeast (class Tremellomycetes, order Filobasidiales), was first isolated from the Beacon supergroup sandstone, Antarctica, and its discovery in the Ross Desert provided the first indication that indigenous microorganisms do inhabit the soils of the unglaciated high valleys (Vishniac and Hempfing 1979a). This species is apparently unique to the McMurdo Dry Valleys and even though its autochthony cannot be inferred based on the impossibility of finding it in similar soils, its ability to grow at 4°C and even at -3°C, but not at or above 21°C, and the fact that many other species found in these regions have also been isolated from other places, makes *N. vishniacii* an interesting candidate for indigenous microbiota. Little is known about cold-adapted yeast inhabiting these environments (Alcaíno, Cifuentes and Baeza 2015). The advent of 'omics' technologies with whole genome sequences of microbes inhabiting cold places are helping us understand their particularities and genes for cold adaptation,

cold shock proteins, pigments biosynthesis pathways, chaperones and DNA repair systems (Yadav *et al.* 2018). In this sense, the more sequenced genomes that become available, the more we will be able to increase and deepen our knowledge.

In this work, we report the genome of *N. vishniacii*. Using comparative genomics with model species and with phylogenetically closely related genomes, we were able to identify genome-wide characteristics of cold adaptation. Particularly, characterization of unique genes linked with protein stability, trends in amino acid composition and identification of specific pathways already associated with cold tolerance in other organisms (photoprotection, trehalose synthesis and lipid metabolism) were achieved. All these traits constitute valuable targets for further functional studies that may account for cold-adaptation in *N. vishniacii*.

MATERIALS AND METHODS

Isolate source

Naganishia vishniacii (ANT03-052; CBS 10616) was isolated from a soil sample collected in the Labyrinths, Wright Valley, South Victoria Land, Antarctica (160.702E, -77.562S). The soil was collected aseptically during the 2003–2004 Antarctic field campaign and isolates were cultured and identified as previously described (Connell *et al.* 2008).

Genome sequencing

The *N. vishniacii* genome was sequenced using two Illumina libraries, Regular Fragments (300bp insert-size) and Long Mate Pairs (LMP, 4kbp). For the Fragment library, 100 ng of DNA was sheared to 300 bp using the Covaris LE220 (Covaris, Woburn, MA, USA) and size selected using SPRI beads (Beckman Coulter, Indianapolis, IN, USA). The fragments were treated with end-repair, A-tailing and ligation of Illumina compatible adapters (IDT, Inc, Coralville, IA, USA) using the KAPA-Illumina library creation kit (KAPA Biosystems, Wilmington, MA, USA).

For the LMP library, 5 µg of DNA was sheared using the Covaris g-TUBE (Covaris) and gel size selected for 4kb. The sheared DNA was treated with end repair and ligated with biotinylated adapters containing loxP. The adapter ligated DNA fragments were circularized via recombination by a Cre excision reaction (NEB). The circularized DNA templates were then randomly sheared using the Covaris LE220 (Covaris). The sheared fragments were treated with end repair and A-tailing using the KAPA-Illumina library creation kit (KAPA Biosystems) followed by immobilization of mate pair fragments on streptavidin beads (Invitrogen, Carlsbad, CA, USA). Illumina compatible adapters (IDT, Inc) were ligated to the mate pair fragments and eight cycles of PCR were used to enrich for the final library (KAPA Biosystems).

All libraries were quantified using KAPA Biosystem's next-generation sequencing library qPCR kit and run on a Roche LightCycler 480 real-time PCR instrument, multiplexed with other libraries and prepared for sequencing on the Illumina HiSeq sequencing platform utilizing a TruSeq paired-end cluster kit, v3 and Illumina's cBot instrument to generate clustered flowcells for sequencing. Sequencing of the flowcells was performed on the Illumina HiSeq2000 sequencer using a TruSeq SBS

sequencing kit, v3, following a 2×150 bp (2×100 bp for LMP) indexed run recipes.

Genome assembly, gene predictions and annotations

Genomic reads from both Fragment and LMP libraries were initially assembled using AllPathsLG release version R41043, (Gnerre *et al.* 2011). Genome assembly was made available via the JGI fungal portal MycoCosm (mycocosm.jgi.doe.gov /Nagvi1; Grigoriev *et al.* 2014). Reads were also assembled using SPAdes v3.6.0 (Bankevich *et al.* 2012) including adapter removal, trimming, quality filtering and error correction. Both assemblies were merged with the program NGAP4 (Nizovoy *et al.* 2018). Briefly, regions flanking undefined nucleotides in the initial AllPaths assembly were recovered and scanned for in SPAdes versions in order to correct nucleotide identities with those included in flanked fragments when possible. This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession JABEVT000000000. The version described in this paper is version JABEVT010000000.

Ab initio gene prediction with GeneMark-ES v2.3e (Ter-Hovhannisyanyan *et al.* 2008) was self-trained on the genome scaffolds (parameters: `-min_contig 8000 -max_nnn 1000`). These predictions as well as genome scaffolds were used to feed the Funannotate Pipeline (Palmer 2016). By means of this pipeline, Pfam domains, Basidiomycota BUSCO groups and functional categories were also annotated when possible. For genome comparisons, predictions and annotations were also performed for most closely related genomes available in NCBI portal, pertaining to the Filobasidiales order, namely *Naganishia albida* Y-1402 and NT2002 (DDBJ/EMBL/GenBank accession numbers PRJNA291988 and PRJNA298369; Vajpeyi and Chandran 2016; Yong *et al.* 2016), *Naganishia adliensis* IF15W-F1 (PRJNA623412; Bijlani *et al.* 2020), *Filobasidium wieringae* UCDFST 05-544 (PRJDB3683), *Solicozozyma phenolicus* JCM11743 (PRJDB3680) and *Solicozozyma terricola* JCM24523 (PRJNA327103). Strains selection of available assemblies was based on phylogenetic closeness to *N. vishniacii*, prioritizing already published genomes (for species with no publication available, assemblies were included upon permission of authors). Genomic sequences of *Cryptococcus neoformans* JEC21 (FungiDB-42) (Stajich *et al.* 2012) were processed equally in order to compare results with the model yeasts species.

Phylogenetic tree reconstruction

Set of *single copy shared* genes among Filobasidiales species and *Cryptococcus neoformans* JEC21 as determined by eggNOG-Mapper v4.5.1 were aligned by means of MUSCLE with default parameters (Edgar 2004) and converted to codon alignment with PAL2NAL (Suyama, Torrents and Bork 2006). Alignments were concatenated, gaps eliminated and the resulting length was 3082 974 columns. Rooted phylogenetic trees were constructed using RAxML v. 8.2.12 (Stamatakis 2014) with the PROTGAMMALG4M model of amino acid substitutions. Branch support values were determined using 1000 rapid bootstrap replicates. Evolutionary model was selected using ModelTest-NG with default parameters (Darriba *et al.* 2020). RAxML and ModelTest-NG were run on CIPRES science gateway (Miller, Pfeiffer and Schwartz 2010). The level of genomic sequence divergence between closely related species was estimated using de Kr value, an alignment-free pairwise distance measure calculated with Genome Tools (Gremme, Steinbiss and Kurtz 2013).

Orthologous genes and functional assignment

To find specific genes that may be involved in cold tolerance, species genes were categorized as *shared* or *unique* by means of fast orthology assignment using precomputed clusters. Clustering was performed with eggNOG v4.5.1 (Huerta-Cepas *et al.* 2017), employing the fungal database. This process enabled functional assignment of groups of genes.

Predicted CDS of *N. vishniacii* not matching with eggNOG database were classified into *shared* or *unique* genes by performing tblastn (e-value < $1e-05$, identity $\geq 50\%$) against the six Filobasidiales genomes. Predictions that did not group with any of those in the rest of the species were considered unique to *N. vishniacii*. Putative functions were inferred by homology search against non-redundant NCBI databases.

dN/dS ratio analysis

Ratio of non-synonymous (*dN*) and synonymous (*dS*) substitutions were computed for all the genes common to *N. vishniacii* and both *N. albida* strains. Analysis between single copy orthologs was performed by means of KaKs Calculator Version 2.0 (Wang *et al.* 2010). Predicted common proteins were aligned with MUSCLE with default parameters (Edgar 2004) and converted into gap-free codon alignments with PAL2NAL (Suyama, Torrents and Bork 2006). Derived **axt* files were split into 102 nucleotide-length chunks and analysed by a sliding window strategy (24 nucleotide steps) using YN model (Yang and Nielsen 2000). Analysis was also run for the unsplit alignments (whole sequences). All predictions with positive regions (*dN/dS* ratio > 1, e-value < $1e-02$) were subsequently compared between *N. vishniacii* and *F. wieringae*. Genes with positive regions for both pairs of species were evaluated with JCoDa (Steinway *et al.* 2010). Signals of positive selection were considered only if *dN/dS* > 1 was shared for the same region on the sequence among *N. vishniacii* vs. any or both *N. albida* strains and *N. vishniacii* vs. *F. wieringae* but not *N. albida* strains vs. *F. wieringae*.

Survey of genes involved in cold tolerance

We used tblastn and blastp to identify genes already known to be involved in pathways normally related to cold tolerance. A survey was made on Filobasidiales genomes and derived proteomes. Antifreeze proteins (AFPs) were screened using sequences obtained from NCBI database by using 'antifreeze' as keyword. Fungal (166) and bacterial (3832) query sequences were used for blastp suite search (e-val < $1e-5$) (available as Supplementary material: **antifreeze.faa*). Genes involved in the synthesis of photoprotection responses were identified in the predicted proteins using blastp (e-val < $1e-10$). Queries used are listed: genes involved in carotenoid synthesis: isopentenyl diphosphate isomerase (acc: BAA33979.1), farnesyltransferase (acc: AAY33922.1), farnesyl pyrophosphatase synthetase (acc: AHW57996.1), phytoene dehydrogenase (G01847.P), phytoene-beta carotene synthase (acc: AA047570.1), NADP-cytochrome P450 reductase (acc: ACI43098.1), beta-carotene 15,15-monooxygenase (G04735.P). Genes involved in mycosporine synthesis: 2-epi-5-epi-valiolone synthase (G05217.P), catechol o-methyltransferase (G05218.P) and carbamoylphosphate synthase (G05219.P) previously characterized in *Phaffia rhodozyma* (Bellora *et al.* 2016). Backbone enzymes involved in the formation and degradation of trehalose, namely

trehalose-6-phosphate-synthase (acc: XP.02 554 8824.1), trehalose 6-phosphate phosphatase (acc: NP.594 975) and trehalases (acc: NP.0 09555.1, NP.01 0284.1 and NP.01 5351.1) were prospected in proteomes using fungal homologues queries. Identity was confirmed by reciprocal-blastp (e-values $\leq 2e-48$), specific Pfam domains presence (PF02358, PF0982/PF07492, PF1204) and grouping into common clusters of orthologous genes. Fatty acid desaturases genes were screened in all predicted proteomes by searching specific Pfam domains (PF00487 and PF08557). Identity of putative targets was confirmed by reciprocal blastp (e-values $\leq 4e-27$). Also differences in copy numbers of enzymes involved in lipid metabolism determined by eggNOG clustering were analyzed among Filobasidiales species.

Analysis of Amino Acids usage

The set of 248 core eukaryotic genes (CEGs; Parra *et al.* 2008) was scanned with HMMER v 3.1b1 (Eddy 2008) in all proteomes from Filobasidiales species as well as in *C. neoformans*. Screening was accomplished by applying the specific models with a threshold defined by Parra *et al.* (2008). Sets of common CEGs among each of the Filobasidiales species and *C. neoformans* (among 239 and 243) were pairwise aligned using MUSCLE (Edgar 2004) with default parameters and converted to codon alignment with PAL2NAL (Suyama, Torrents and Bork 2006). Translated alignments were concatenated and gaps eliminated. Concatenated sequences of each species, as well as those resulting from pairwise alignment of 2130 genes common to all, as determined by eggNOG-Mapper v4.5.1, were analyzed in order to study if there was an amino acid composition bias. Comparisons of amino acid composition were made over both sets of resulting sequences using Composition Profiler with default parameters (Vacic *et al.* 2007) and the analysis was referred to *C. neoformans*.

Experimental confirmation of carotenoids and mycosporines production in *N. vishniacii*

Yeasts were cultured in modified MMS agar (Moliné *et al.* 2011) under three different treatments: (a) radiation, (b) oxidative stress and (c) low temperature. Plates were inoculated with 100 μL of a 10^6 cell suspension with a Drigalski-spatula until dry and cultured: (a) at 19°C under photosynthetically active irradiation in an environmental test chamber with light provided by 10 white light fluorescent tubes (Sanyo, 40 W); (b) at 19°C adding 1 mL of H_2O_2 0.01 mM and (c) at 0.5°C. Carotenoid extraction was performed as described by Sedmak (Sedmak, Weerasinghe and Jolly 1990), with modifications proposed by Moliné *et al.* (2009). Samples were suspended in n-hexane and the absorbance was measured at 473 nm in a UV-visible spectrophotometer (Shimadzu UV1800). Mycosporines extraction was performed as described before (Moliné *et al.* 2011) and absorbance was measured at 310 nm.

RESULTS

Genome assembly, annotations and phylogenetic reconstruction

The 19.7-Mbp genome of *N. vishniacii* (ANT03-052; CBS 10616) was assembled into 38 scaffolds, ranging from 2.1 kb to 1.8 Mb, with a 146.7-fold depth of coverage and a very low number of undefined nucleotides (25 706 Ns, 0.13%). It was 90.9% complete using the Basidiomycota BUSCOs and 99.2% complete

based on CEGs genes. Gene predictions yielded 6183 putative coding genes accounting for 53.4% of the genome sequence. From these models 4956 putative genes (80.3%) were assigned to 23 different eggNOG functional categories and classified into 4609 orthogroups. 2613 (42.3%) gene predictions were annotated with a putative function. G + C content among this species was considerably higher than the model species *Cryptococcus neoformans* JEC21 as shown in Fig. 1.

The phylogenetic tree for the eight species constructed based on common single copy orthologs clusters (described below) grouped species by genera (Fig. 1).

Orthologous genes and functional assignment

Orthology inference analysis of the 4609 different groups revealed that among them, 4593 (~99.7%) were shared between *N. vishniacii* and one or more of the related species. 2926 orthogroups were common to the seven Filobasidiales species, whereas 16 groups were only found in the *N. vishniacii* genome. Almost 18% of predicted genes in *N. vishniacii* ($n = 1073$) did not match any orthogroup of the eggNOG database. Tblastn of derived proteins performed against the rest of the species genomes for this set of genes yielded 753 (70%) predictions with positive matches in any or several species. For the rest of the 320 predictions without significant matches and no orthology assignment, the search against NCBI non redundant database resulted in 297 of these annotations (almost 93%) not matching any deposited query, while 23 (~7%) were linked to hypothetical proteins or partially matched with a deposited sequence and six of them (~2%) were linked with putative functions. Complete list is available as Table S1 (Supporting Information; Fig. 2).

Among 16 orthogroups found only in the *N. vishniacii* genome, the eggNOG functional categories most represented included unknown function (43%), signal transduction mechanisms (13%), secondary metabolites biosynthesis transport and catabolism (13%) and transcription (13%). Also carbohydrate transport and metabolism and RNA processing and modification were present. Functional description resulted, among others, in putative transferases (Nv.005274), solute transporters (NV_002329) and chaperones (Nv.005012). A complete list is available as Supplementary material (Table S2, Supporting Information).

In order to explore possible gene redundancy in *N. vishniacii*, we compared orthogroups copy number. Analysis revealed that 2366 shared groups (~81%) encompassed one copy candidate by species, 19 (~1%), two copies and 1, three copies, while 540 groups (18%) showed differences in copy numbers among species. When comparing the copy numbers between *N. vishniacii* and *N. albida* strains we found that in 12 orthogroups *N. vishniacii* gene redundancy was higher than that in both *N. albida* strains (Fig. 3). Genes involved in RNA processing and modification, amino acids and solute transport and metabolism and energy production were more abundant in *N. vishniacii* regarding both *N. albida* strains. For groups ENOG410PJ40, ENOG410PFJG, ENOG410PH2P, ENOG410PFKH and ENOG410PF97 (green highlighted), even though *N. vishniacii* surpassed *N. albida* copy numbers, it showed equal or even lower number of genes than those in the rest of the Filobasidiales species belonging to non *Naganishia* genera.

dN/dS analysis

Distribution of *dN/dS* mean values for the 3942/3991 single copy common genes between *N. vishniacii* and *N. albida* strains

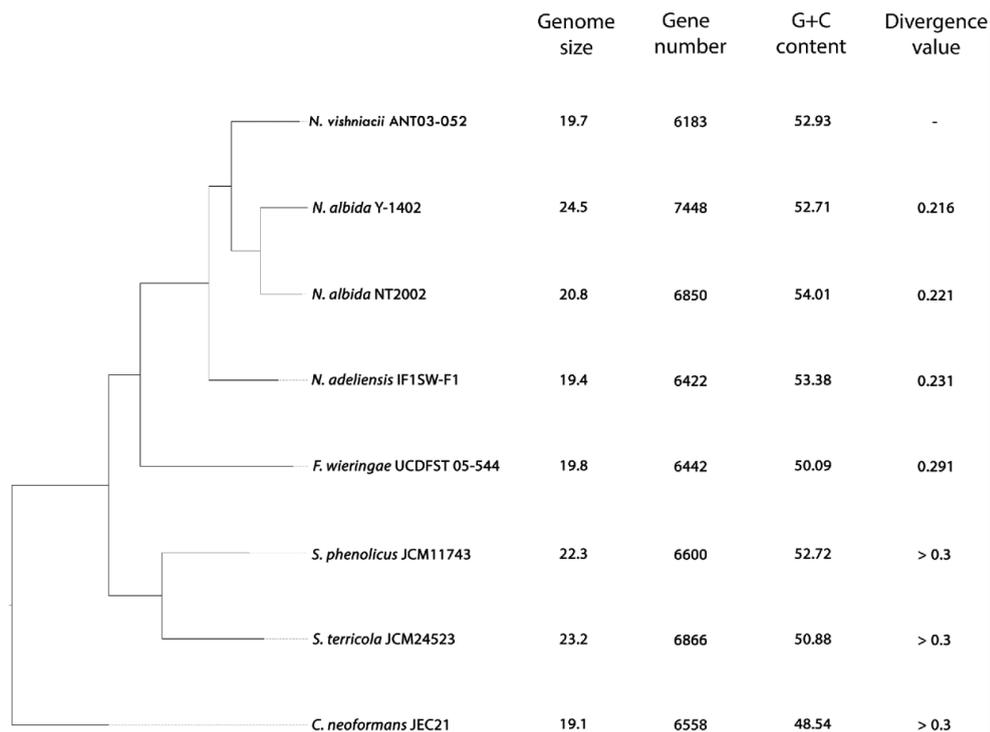


Figure 1. Phylogenetic tree of the organisms used in this study, constructed from 2130 single copy ortholog protein sequences using RAxML v. 8.2.12 with the PROTGAM-MALG4M model of amino acid substitution. All nodes showed 100% bootstrap support.

(Y-1402/NT2002 respectively) are shown in Fig. 4A. The dN/dS ranged from 0 to 0.526/0.590 although the 98% of the genes had $dN/dS < 0.2$ (81/79% with $dN/dS < 0.1$, respectively). Even though no gene showed dN/dS mean value > 1 , by means of a sliding windows strategy we could detect that 65/77 annotations (1.9/1.6% respectively) overcame $dN/dS > 1$ in any region between *N. vishniacii* and *N. albida* strains and 49/23 of those also showed a $dN/dS > 1$ between *N. vishniacii* vs. *F. wieringae* pairwise comparisons. One by one comparison of these annotations among the three species using JCoDA resulted in 31/13 (0.79/0.30%) orthologs encompassing possible evidence of positive selection, with $dN/dS > 1$ in shared regions between *N. vishniacii* vs. *N. albida* (YN-102/NT2002 strains) or *F. wieringae*, but not between *F. wieringae* vs. *N. albida* (numbers shown in brackets in Fig. 4A). Among them, nine showed evidence of positive selection in both *N. albida* strains. Orthogroups were classified by means of eggNOG functional categories (Fig. 4B). Function of most of the predicted genes (42.9%) could not be inferred (grouped as Function unknown). Post-translational modification, protein turnover and chaperones (20%), RNA processing and modification (8.6%) and Replication, recombination and repair (8.6%) were the most represented categories. A total of ten of these 35 genes also matched BUSCO database (EOG nomenclature in Fig. 4B).

Survey of genes involved in cold tolerance

A survey of specific genes related to cold adaptations demonstrated that *N. vishniacii* possesses genes involved in photoprotection and antioxidant responses as well as in the synthe-

sis of trehalose. Furthermore, we were able to identify genes involved in lipid transport and metabolism in *N. vishniacii*. Homology search of antifreeze proteins yielded positive results in *N. adeliensis* genome only. This species possesses a gene with a high homology ($< 1e-11$) to a characterized AFP glycoprotein from *Rhodotorula toruloides*. For the rest of the Filobasidiales species, we could not find any homology evidence employing the 3998 AFPs queried. Regarding genes involved in photoprotection and antioxidant responses, a cluster of three genes previously shown to be involved in mycosporine-glutaminol-glucoside (MGG) synthesis (Bellora et al. 2016), consisting of *DHQS* homolog (2-epi-5-epi-valiolone synthase, EEVS-like; Nv.000059), O-methyltransferase (*O-MT*; Nv.000060) and an ATP-grasp (Nv.000058) was found in all the *Naganishia* genomes as well as in *F. wieringae* and *S. terricola*. Queried sequences were not found in *S. phenolicus* nor in the outgroup species. MGG biosynthesis was experimentally confirmed for all positive species: *F. wieringae* (Muñoz 2010, Brandão et al. 2011), *S. terricola* (Libkind et al. 2009) and *Naganishia* spp. (Guzmán 2016 and in all three conditions tested in this work). The complete set of genes involved in the synthesis of carotenoids were characterized in *N. vishniacii* genome (Nv.003146: isopentenyl diphosphate isomerase; Nv.004897: farnesyltransferase; Nv.000648: farnesyl pyrophosphatase synthetase; Nv.002341: phytoene dehydrogenase; Nv.005201: phytoene-beta carotene synthase; Nv.004862: NADP-cytochrome P450 reductase; Nv.000692: beta-carotene 15,15-monooxygenase) as well as in both *N. albida* strains, *N. adeliensis* and *F. wieringae*, but not in the other studied species (Table S3, Supporting Information). *Naganishia vishniacii* colonies are creamy to light pink colored when cultured under different

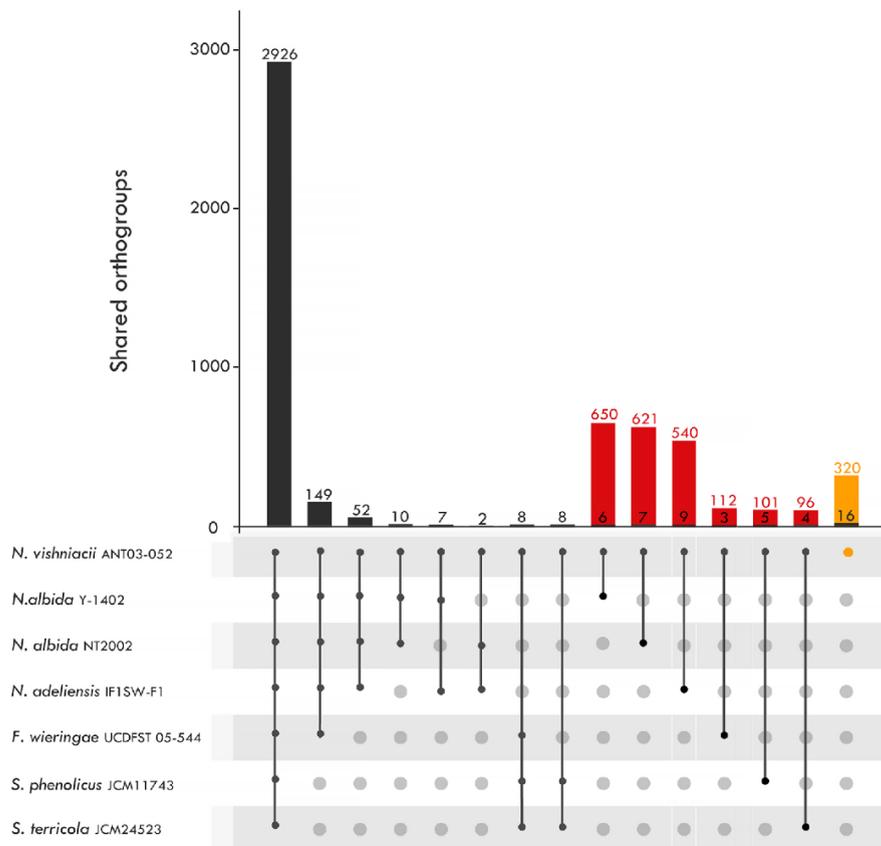


Figure 2. Orthology analysis of seven Filobasidiales species. Black bars represent the number of genes present in shared orthogroups. Filled dark dots denote presence and empty dots (light gray) indicate absence of orthogroups in each species. Only main intersections/overlap relative to *N. vishniacii* are shown. Predicted proteins pertaining to *N. vishniacii* that did not cluster in orthogroups but with positive blastn hits against the other genomes are represented by red bars above each pairwise intersection. Yellow vertical bar represents genes in *N. vishniacii* not clustered by eggNOG and not matching the rest of the Filobasidiales genomes used in this study. Plot visualized with UpSetR (Conway, Lex and Gehlenborg 2017).

conditions, however experimental confirmation of carotenoid production was negative in all tested conditions used in this work.

Genes involved in trehalose synthesis were identified in all of the eight species in this work. Backbone enzymes involved in the formation (*TPS1* and *TPS2*) and degradation (*NTH1/2*, *ATH1*) of these two glucose molecules disaccharide were screened in the proteomes. One putative copy of *TPS1* and *TPS2* enzymes were found in all Filobasidiales species (*TPS1* was found in consecutive predictions on scaffold 2: Nv_002644 and Nv_002645 and *TPS2* was found on scaffold 17: Nv_002256). Trehalase gene *NTH1* (Nv_002469) was found as a single copy in all the genomes (Table S4, Supporting Information). Regarding lipid metabolism genes, 4 (Nv_004586), 8 (Nv_004525) and 12 (Nv_000815) desaturases were found in all the genomes from Filobasidiales species by looking for specific protein domains and the presence of 9 (Nv_003331) desaturase was confirmed in all analyzed genomes except for *S. terricola*. Orthology clustering of enzymes involved in lipid metabolism yielded between 119 and 136 groups composed of unidentified hypothetical proteins in the seven Filobasidiales species, with 78 groups common to all. In addition, 49 groups were annotated with associated functions, with 37 groups shared among all the species. Even though the copy numbers of enzymes pertaining to each group were similar among species, an enzyme related to putative enoyl-CoA-hydratase was

not found in any of the *Naganishia* species and sterol 3-beta-glucosyltransferase was not found in *Naganishia* or in *F. wieringae* genomes (Table S5, Supporting Information).

Analysis of amino acids usage

Differences in amino acid usage was computed for the common-to-all set of single copy orthologous genes (ScCOg) among the seven Filobasidiales species and differences in usage were compared with *C. neoformans*, and also between *N. vishniacii* and the rest of the *Naganishia* species. Enrichment in six amino acid usage (Ala, Arg, Asp, Gln, Met and Thr), and depletion in other five amino acids (Glu, Lys, Pro, Ser and Trp) was observed in all the *Naganishia* species compared with *C. neoformans* ($p < 0.001250$; Fig. 5).

Comparison among *Naganishia* species showed a lower frequency of Ala usage in *N. vishniacii* with the other studied species belonging to the genera. A depletion in Pro and an enrichment in Ser was also confirmed when comparison was made against both *N. albida* strains.

In contrast, when comparison was made over CEG dataset significant differences were found in one amino acid only, with an enrichment of Gln (p -value = 0.001250) for *N. vishniacii* and *N. adeliensis* species with respect to *C. neoformans*, but no significant differences in amino acid composition were observed between

Orthogroup	<i>N. vishniacii</i> ID	COG functional category	Functional description	#Nv	#NaY	#NaN	#Nad	#roF
ENOG410PGGB	Nv_000525 Nv_001541 Nv_003488	RNA processing & modification	Dicer-like endonuclease	3	2	2	1	2
ENOG410PGKM	Nv_002463 Nv_003321 Nv_003922	AA transport & metab. post-translational modif. prot. turnover & chaperones	Aminopeptidase	4	2	1	1	1-2
ENOG410PMCP	Nv_003280 Nv_003328 Nv_004462 Nv_005845 Nv_005846	Energy production & conversion	Oxidoreductase	5	4	3	3	2-4
ENOG410PII4	Nv_000220 Nv_003422 Nv_005291 Nv_005418	Carbohydrate transport & metabolism	alpha-1-6-mannosyltransferase	4	2	1	2	1-3
ENOG410PNF7	Nv_003563 Nv_004049 Nv_004279		Glycosyl hydrolase family 71	3	1	1	3	1-2
ENOG410PNRT	Nv_002066 Nv_002068 Nv_002408 Nv_004073	Function unknown	SUR7/Pall family	4	3	3	3	1-2
ENOG410PWG1	Nv_001317 Nv_001333 Nv_003778		NA	3	2	2	5	2
ENOG410PJ40	Nv_001387 Nv_001586 Nv_003277	AA transport & metab. post-translational modif. prot. turnover & chaperones	Amidohydrolase	3	2	1	2	1-4
ENOG410PFJG	Nv_000658 Nv_003680 Nv_003681		MFS transporter	3	1	2	2	1-3
ENOG410PH2P	Nv_000862 Nv_001298 Nv_005799			3	2	2	2	2-3
ENOG410PFKH	Nv_000275 Nv_000284 Nv_002412 Nv_005830	Carbohydrate transport & metabolism	Quinate permease	4	3	3	3	3-7
ENOG410PF97	Nv_001371 Nv_002166 Nv_002181 Nv_002182 Nv_005515		Maltose permease, trehalose transport-related protein	5	3	1	4	5-16

Figure 3. Orthogroups in which *N. vishniacii* copy numbers were higher than *N. albida*. Copy numbers in *N. vishniacii* (#Nv), *N. albida* Y-1402 (#NaY), *N. albida* NT2002 (#NaN) and *N. adeliensis* (#Nad) and observed range in the rest of the Filobasidiales species (#roF) are shown. Green highlighted: groups in which the rest of the Filobasidiales showed higher or equal copy numbers than *N. vishniacii*.

DISCUSSION

Genome assembly, gene predictions and annotations

Using a combinatorial approach as described in the methods, we were able to generate a high quality genome assembly for *N. vishniacii* (ANT03-052; CBS 10616) with a low number of scaffolds (38), a very low number of Ns (0.13%) and a high number of identified complete BUSCO (90.9%) and CEGs genes (99.2%). The genome size of *N. vishniacii* was smaller than those of *N. albida* strains but similar to *F. wieringae* and *N. adeliensis*. The predicted gene number was also considerably lower than that observed for all the other species in this study. Gene loss is a pervasive phenomenon in all organisms with positive effects (Albalat and Cañestro 2016) given that in free living microorganisms gene loss can be beneficial when they are inhabiting oligotrophic or

extreme environments (García-Fernández, de Marsac and Diez 2004; Ullrich *et al.* 2016; Zhang *et al.* 2017) or are in complex microbial communities, particularly in homogeneous environments (Morris, Lenski and Zinser 2012; Mas *et al.* 2016). Although the impact of gene loss in yeast fitness is still a matter of debate (Thatcher, Shaw and Dickinson 1998; Sliwa and Korona 2005; Lang, Murray and Botstein 2009), the considerably low number of genes in *N. vishniacii* with respect to the other related species suggest that this can be an adaptive feature to an environment with low resources and many stressing factors but relatively constant conditions such as low temperature, low water activity and high solar irradiation (Selbman *et al.* 2014).

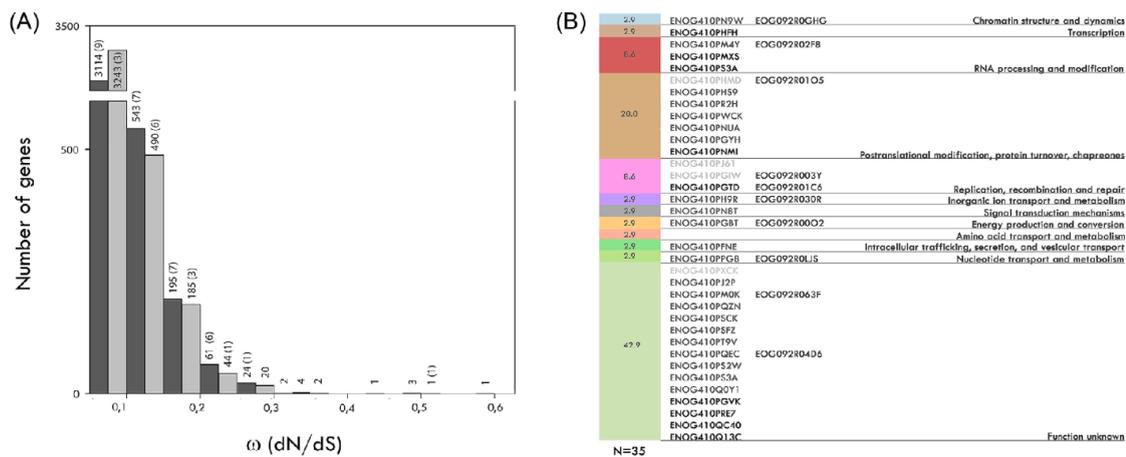


Figure 4. dN/dS analysis for single copy shared genes. **(A)** Distribution of dN/dS mean values for 3942/3991 orthogroups shared among *N. vishniacii* and *N. albida* Y-1402 (dark bars) and NT2002 (light bars) strains. Shown in brackets: numbers of genes with $dN/dS > 1$ considered under positive selection in any region ($n = 35$). **(B)** eggNOG functional categories of genes under positive selection. Groups with positive signal among *N. vishniacii* and both strains (bold), Y-1402 (dark color) only and NT2 (light color) only. Genes belonging to the BUSCO database are shown (EOG).

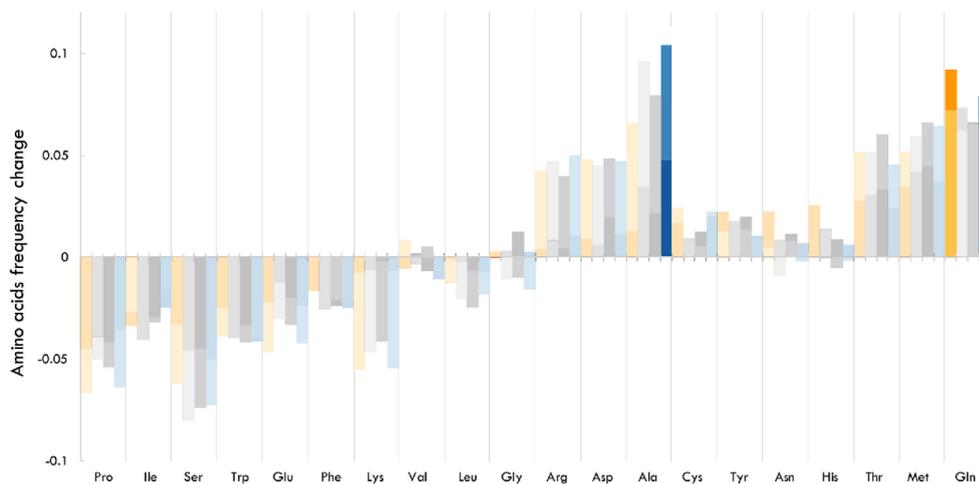


Figure 5. Amino acid frequency changes in *N. vishniacii* (orange), *N. albida* strains Y-1402 and NT2 (dark and light gray respectively) and *N. adeliensis* (blue) with respect to *C. neoformans*, tested in conserved CEGs genes (dark colors) and SccOg (light colors). Highlighted bars show amino acids with significant differences for both sets of genes.

with low resources and many stressing factors but relatively constant conditions such as low temperature, low water activity and high solar irradiation (Selbman *et al.* 2014).

Orthologous genes and functional assignment

Comparison between orthologous groups of genes among all the Filobasidiales genomes available to date as well as *C. neoformans* allowed us to identify that most of the genes present in *N. vishniacii* were also present in related species. Interestingly, orthology inference resulted in a set of 44 candidate genes corresponding to 12 orthogroups (Fig. 3) with increased copy numbers (gene redundancy) in *N. vishniacii* genome when compared with *N. albida* strains and, in most cases, also with the rest of the Filobasidiales species, and 16 genes (corresponding to 16 orthogroups) unique to *N. vishniacii*.

Gene redundancy is usually related to an increase in dosage of encoded genes (Schuster-Böckler, Conrad and Bateman 2010) and a mechanism of adaptation to stressful or novel environmental conditions (Kondrashov 2012). Among the functional groups with greater gene redundancy in *N. vishniacii*, the major facilitator superfamily (MFS) reported here has been previously proposed as part of an important mechanism for cold adaptation in the obligate psychrophilic yeast *Mrakia psychrophila* (Su *et al.* 2016), probably contributing to the accumulation of nutrients from the environment in order to give rise to all the other metabolic and rearrangement modifications needed to counteract cold imposed conditions. For other functional groups, the relationship with cold is not clear. Different organisms counteract shifts in temperature by overexpression of genes involved in modulating RNA processing in different ways (Phadtare, Inouye and Severinov 2004; Zhu, Dong and Zhu 2007) such as by the

production and accumulation of reserve carbohydrates as cold protectant metabolites (Schade *et al.* 2004), and by modification of membrane lipid composition to maintain an appropriate physical state and functionality (discussed below). Gene redundancy in Dicer-like endonucleases, oxidoreductases (enoyl reductases), and in trehalose transport proteins could be related to those general responses.

By means of our comparative approach we were also able to unravel 16 orthogroups that appeared to be present only in the *N. vishniacii* genome (with respect to the other six genomes used for comparison in this work). Most of them were associated with signal transduction mechanisms, secondary metabolites biosynthesis/transport/catabolism and transcription. Also, groups related to carbohydrate transport and metabolism, RNA processing and modification and cell wall/membrane biogenesis were present. Among them we found one gene containing the MFS domain (Nv_002329) that provides additional evidence for the role of this group in cold adaptation; one transferase (Nv_005274), a functional group that was previously suggested as having important role in cold tolerance of Antarctic bacteria for being involved in membrane biogenesis, carbohydrate, proteins or lipids modifications (Grzyski *et al.* 2006) and one gene related to chaperones (Nv_005012), enzymes required for protein folding, important for protein conformation maintenance in the cold (Ferrer *et al.* 2003) and reported to be involved in a synergistic effect together with trehalose during yeast adaptation to near-freezing temperatures (Kandror *et al.* 2004).

Finally, for the 320 unique genes of *N. vishniacii*, not grouped by orthology clustering nor found in the other species in this work, the vast majority (297) did not yield significant matches in the GenBank non-redundant database. Similar numbers of genes without matches were observed in yeasts belonging to the Basidiomycetes and the Ascomycetes (Bellora *et al.* 2016; Levasseur *et al.* 2014; Blevins *et al.* 2019) and are usually categorized as putative orphan or lineage-specific genes. These are in some cases novel genes but they mostly have undetectable homology with other related sequences (Weisman, Murray and Eddy 2020) and are often related to ecological adaptations (Wilson *et al.* 2005). Further analyses are needed to understand if any of these predictions are related to the extreme habitat preference of *N. vishniacii*.

dN/dS analysis

Distribution of *dN/dS* mean values among *N. vishniacii* and *N. albida* strains' single copy shared genes revealed that most genes are under negative selection. However, selective pressure on genes generally occurs during some lapses of time on the evolution affecting only a few amino acids on specific regions of domains and *dN/dS* is usually lower than one in divergent species, moreover values lower or higher than one can occur under both negative and positive selection (Kryazhimskiy and Plotkin 2008; Spielman and Wilke 2015). For those reasons the averaged *dN/dS* ratio is tricky and rarely exceeds values greater than one (Gillespie 1994). By means of a sliding windows strategy we could find 35 genes with regions exceeding *dN/dS* > 1 thus encompassing possible evidence of positive selection. Since our search strategy was highly conservative and we only kept genes with selection signals present in *N. vishniacii* only, this set of 35 genes is highly significant in terms of evolutionary history.

Among them, Almost 40% of the genes were functionally related with translational modification, protein turnover and chaperones, RNA processing and modification and replication, recombination and repair, as the main represented groups. In

this sense, it is important to highlight that all these genes are involved in genome expression and protein folding and, as mentioned before, the role of these functional categories in cold tolerance has already been reported for psychrotolerant organisms (Phadtare, Inouye and Severinov 2004). In total, two genes are remarkably important considering the environment from which *N. vishniacii* was isolated. The gene Nv_005203 (ENOG410PSCK) showed the highest mean value for those whose function could be inferred. HMMER search of Nv_005203 reveals that this gene is related to the Uniprot target A0A1E3HK01 of *C. depauperatus*, and Q5KHT7 of *C. neoformans*, both related with the function of sensing extracellular osmolarity to initiate a change in cell activity, and to span the membrane of the cell. *N. vishniacii* was described not only as a psychrophilic yeast, but also as a xerophilic species (Vishniac 2006). Considering that the McMurdo Dry Valleys are the most extreme cold deserts known (Vishniac and Hempfing, 1979b) and lack of water of its xeric soils is one of the main factors limiting microbial growth there (Klingler and Vishniac 1988), this selection pressure on Nv_005203 gene is most probably related with the need to survive in environments with low water activity. One of the most studied responses in xerophilic fungi to survive and grow at reduced water availability is the intracellular accumulation of compatible solutes like glycerol, to maintain balance between water availability inside and outside the cell (Vinnere and Leong 2001). The membrane osmosensor Nv_005203 found in this analysis is highly similar to the *Sho1* gene described for *S. cerevisiae* and *Aspergillus* spp., involved in the synthesis and accumulation of glycerol (Salazar *et al.* 2009). In *S. cerevisiae* this protein is part of a transduction cascade that, upon different stresses (including cold, hyperosmotic and oxidative stresses), trigger transcriptional responses (Rodríguez Peña *et al.* 2010). Also remarkable was the positive signal in gene Nv_003572 (ENOG410PGIW). In this case, HMMER search yielded a putative serine threonine protein kinase related to A0A1E3IMU4 Uniprot target, from *C. depauperatus*. This protein is a DNA damage response kinase that belongs to the phosphatidylinositol 3-kinase-like (PIKK) family. Members of this family are activated upon DNA damage further amplifying the signal by activating downstream kinases (Durocher and Jackson 2001). In cells with damaged DNA due to genotoxic stresses such as ionizing radiation, ultraviolet light or DNA replication stalling (Kozhina, Kozhin and Korolev 2011), propagation of the signal results in cell-cycle arrest and stimulation of DNA repair (Bøe *et al.* 2018). Selection of this gene, thus, may be related to the high incidence of solar radiation characteristic of this harsh environment (Cary *et al.* 2010) as will be discussed below.

The function of the remaining genes with strong signals of positive selection could not be inferred, and were annotated as hypothetical proteins. Functional relevance of these predictions in the response of *N. vishniacii* to its habitat's extreme conditions should be the target for further studies.

Survey of genes involved in cold tolerance

Genome mining in the genome of *N. vishniacii* for genes previously known to be related to cold tolerance in yeasts and other organisms suggested the presence of most of these traits but surprisingly the lack of others, for example antifreeze-proteins (AFPs). AFPs have been identified in various psychrophilic organisms (Ewart *et al.* 1999), including yeasts such as basidiomycetous species of the genera *Glaciozyma*, *Leucosporidium* and *Rhodotorula* (Kim *et al.* 2014). Although AFPs are thought to prevent cell damage from recrystallization of extracellular ice

sequences in *N. vishniacii*. Because this species lives in Antarctic soils, a strategy to avoid the formation of ice crystals was expected. However, it is important to highlight that AFPs possess a large, relatively flat and hydrophobic plane on the protein surface and commonly have repeating motifs (Wang *et al.* 2017), and since this characteristic is common to several proteins, AFPs evolved several times in different organisms from different peptides (Cheng 1998). For those reasons, we cannot eliminate the possibility of other still unknown AFP in *N. vishniacii*, different from those already described in other organisms.

We were able to detect genes involved in the synthesis of antioxidant and UV sunscreen molecules (carotenoid pigments and mycosporines). Production of carotenoids and UV-screening compounds is a wide-spread trait present in several yeasts species (Libkind *et al.* 2011; Moliné *et al.* 2014) and both compounds play an important role in photoprotection and possess antioxidant properties (Moliné *et al.* 2009; 2011). In cold rock-associated ecosystems, such as the ones of the McMurdo Dry Valleys, high solar irradiation is a limiting stress factor for life (Onofri *et al.* 2007). In these environments yeasts species from the orders Filobasidiales and Tremellales are usually isolated (Selbmann *et al.* 2014). The presence of photoprotective compounds in yeasts from such environments was expected. All known genes responsible for the synthesis of both carotenoid pigments and mycosporines (particularly MGG) were unambiguously found in *N. vishniacii*, as well as in the rest of the *Naganishia* species and *F. wieringae* but not in the other species in this study. The presence of genes involved in the production of carotenoids and MGG make sense considering that Antarctica is a region of high UV radiation incidence as a result of the long solar radiation periods in summer and the ozone reduction (Madronich *et al.* 1998; McKenzie *et al.* 2011) and it has been reported that *Naganishia* species are among the more resistant organisms to UV radiation (Schmidt *et al.* 2017). However, we could not experimentally confirm carotenoid synthesis in *N. vishniacii* even under photostimulation and oxidative conditions. Moreover, the three *Naganishia* species and *F. wieringae* usually grow in culture as white to creamy colored colonies (Viviani and Tortorano 2009; Kurtzman, Fell and Boekhout 2011). It is still not clear whether the lack of carotenoids in culture colonies of *N. vishniacii* is due to a failure to meet the necessary conditions to stimulate their synthesis or a possible inhibition at the level of gene expression. On the contrary, the accumulation of MGG was experimentally confirmed for *N. vishniacii*, suggesting that this molecule is enough to cope with extreme UV radiation conditions in Antarctic environments.

The presence of the primary enzymes involved in the synthesis and degradation of trehalose were also confirmed in all the Filobasidiales genomes. Trehalose is a well-known protective metabolite against harsh environmental conditions such as freezing temperatures (Wiemken 1990) that helps to maintain membrane integrity and function (Rudolph, Crowe and Crowe 1986). In yeasts, increasing intracellular trehalose accumulation is sufficient to confer high tolerance to desiccation (Tapia *et al.* 2015). The presence of trehalose genes together with the expanded trehalose transport proteins described above, suggests their capability to synthesize this important compound that probably helps to withstand cold conditions and low water availability of the McMurdo Dry Valleys.

In the same way, the presence of a complete set of desaturases genes in *N. vishniacii* is in agreement with the capability to respond to cold. Membrane fluidity can be regulated by changing the lipid composition (Ernst *et al.* 2016), and the number of double bonds in phospholipid fatty acyl chains are increased at

low temperature to maintain the appropriate physical state of the membrane (Chintalapati, Kiran and Shivaji 2004). Recombinant expression of Δ -12 fatty acid desaturase in *S. cerevisiae* enhances their survival at low temperatures (Shi *et al.* 2012). The complete set of desaturases described for Filobasidiales, including the presence of Δ -12 fatty acid desaturases, indicate the ability of these species to regulate the membrane fluidity under extreme cold conditions. Moreover, when we individually compared gene redundancy of a variety of enzymes involved in lipid metabolism, we determined that all the *Naganishia* species used in this work, as well as the closely related *F. wieringae* genomes lack sterol 3-beta-glucosyltransferase. This enzyme catalyzes the synthesis of sterol glycosides and acyl sterol glycosides and is ubiquitous in fungi (Grille *et al.* 2010). The attachment of a sugar moiety to the 3-hydroxyl group of a sterol increases the size of the hydrophilic head of the lipid changing its biophysical properties and it has been observed that a decrease in sterol glycosides and acylated sterol glycosides content was correlated with an increase in resistance to freezing stress (Thieringer *et al.* 1998). Lack of this enzyme probably serves as a complementary trait contributing to the cold tolerance of *N. vishniacii*.

Analysis of amino acids usage

When comparing differences in the amino acid usage, we found a considerable enrichment of glutamine for *N. vishniacii*, a phenomenon particularly evident in the group of highly conserved CEGs and higher in magnitude than on its psychrophilic cognate *N. adeliensis* (Scorzetti *et al.* 2000). Increased flexibility has been proposed to be the main structural feature of cold-adapted enzymes to compensate for the reduction in chemical reaction rate inherent to low temperatures (Gerday *et al.* 2000; Gianese, Argos and Pascarella 2001). For this reason, a higher usage of small residues is proposed to be preferred by these enzymes (Adekoya *et al.* 2006). In this way, the energetic cost of conformational changes required to interact with substrates could be reduced. Increased catalytic efficiency is a key feature for low-temperature enzymes that have to compensate for slow metabolic rates that would occur at low temperatures due to the low kinetic energy available to overcome reaction barriers (Georgette *et al.* 2004).

The relationship between glutamine usage in enzymes and temperature has been previously described in Archaeas, following an almost linear trend over the complete range of optimum growth temperatures from psychrophiles to hyperthermophiles (Saunders *et al.* 2003). Notably, *N. vishniacii* glutamine percentage was the highest among all the compared species.

Also, a marked trend in reducing proline content was observed in *N. vishniacii* when compared with *C. neoformans* and also with both *N. albida* strains, but not with *N. adeliensis*. Although this change was not statistically significant, it is remarkable as a similar trend has already been reported as an attribute of cold-adapted enzymes (Feller and Gerday 1997). Proline usage in peptides restricts backbone rotations and reduces flexibility (Siddiqui and Cavicchioli 2006), and for that reason proline is a more common amino acid in proteins adapted to high temperatures (Aghajari *et al.* 1998; Spiwok *et al.* 2007; Sælensminde, Halskau and Jonassen 2009; Tronelli *et al.* 2007). Depletion of Pro in *N. vishniacii* and in *N. adeliensis* may be in line with adapting to their freezing environments. To our knowledge, this is the first time that changes in amino acids usage related to cold adaptation are reported for yeasts.

FINAL REMARKS

Understanding adaptations to the environment has been one of the main questions in biology from the time of Darwin to the present. Scientists seek to understand how evolution shapes these characteristics in all organisms and what are the particular changes that make organisms more suited to their environments. Comparative genomics is shedding light on the understanding not only of the phylogenetic relationships between species, but also of the molecular mechanisms involved in sequence evolution, speciation, genome rearrangements, the origin of new genes as well as the evolution of those involved in different adaptations (Koonin, Aravind and Kondrashov 2000; Miller *et al.* 2004; Wolfe 2006). Availability of an increasing range of yeasts genomes is a powerful source of information for determining the genetic basis for phenotypic variation, the mechanisms of genomic rearrangements and those involved in speciation and isolation. They also help to elucidate processes that take place on short timescales, such as recent adaptation of species to their environments (Wolfe 2006; Marsit *et al.* 2016; Riley *et al.* 2016; Libkind *et al.* 2020). In this work we explored the genome of the extremophile yeast *N. vishniacii* and were able to reveal a plethora of genomic traits that may be the underlying basis of its capability to grow in the extreme environments of Antarctic dry soils.

To the best of our knowledge, since this species was first isolated from the Beacon supergroup sandstone (Vishniac and Hempfing 1979a), no report of its presence has been made outside this zone in Antarctica (Buzzini and Margesin 2014). Even though many authors have proposed much of the microbiota of Antarctica as potentially originating elsewhere and colonizing Antarctica by aerial dispersion (Vincent 2000; Schmidt *et al.* 2017). The ability of *N. vishniacii* to withstand daily and yearly environmental extremes of its peculiar habitat and its repeated occurrence in this place only, suggest it may constitute an endemic species, since emigrant recolonization of vacant sites, or sites with unstable biodiversity, is expected to be random with respect to the taxon (Vishniac and Hempfing 1979b). Even though lack of systematic and culture-independent studies in this region, make biogeography, ecological relevance and patterns of spatial distribution of fungi an unrevealed matter (Dreesens *et al.* 2014), presence of specific traits of adaptation in *N. vishniacii*, much of them postulated in this work for the first time, supports a tight association of this species to this place.

Comparative genomics allowed the identification of a set of candidate genes probably involved in cold adaptation and also different trends during the evolution of *N. vishniacii* genome in relation to its capability to cope with extreme cold conditions, including genes that are unique, in higher copy number, carrying signatures of positive selective pressure, involved in specific metabolic pathways, and with unique changes in amino acids composition. To the best of our knowledge, such a combination of genomic features for cold tolerance has not been previously reported in fungi. While the genome sequence analyses are the first steps to understand the adaptation of a given organism to life under extreme conditions, further integrative *omic* studies are required to infer the possible underlying mechanisms.

AUTHOR CONTRIBUTIONS

P.N. carried out experiments, assembled and merged genomes, performed annotation and bioinformatics analysis, analyzed data and wrote the manuscript. N.B. assisted, performed and supervised bioinformatics analysis. S.H. annotated genome and

submitted genomic data to GenBank. H.S. performed the original genome assembly. C.D. sequenced genome. K.B. and I.G. coordinated genome sequencing, annotations and data submission. L.C. collected the samples and isolated the yeast. M.M., L.C., D.L. and N.B. conceived the study and contributed with the writing and revision of the manuscript. M.M. also assisted with data analysis, bioinformatics analysis and experiments. D.L. and M.M. provided the funding.

ACKNOWLEDGMENTS

The authors thank the following researchers for permission to use genome data prior to publications: Rikiya Endoh and the Genome Information Upgrading Program of the National BioResource Project of the MEXT, Japan for *S. phenolicus*, Dan Close for *S. terricola*, and Kyria Boundy-Mills for *F. wieringae*. The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy under Contract number DE-AC02-05CH11231.

SUPPLEMENTARY DATA

Supplementary data are available at [FEMSYR](https://www.ncbi.nlm.nih.gov/bioproject/112201) online.

FUNDING

This work was partially supported by Consejo Nacional de Investigaciones Científicas y Técnicas [PIP11220130100392CO], Fondo para la Investigación Científica y Tecnológica [PICT 2542 to D.L., PICT 0565 to M.M.] and Universidad Nacional del Comahue [B199].

Conflicts of interest. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

REFERENCES

- Adekoya OA, Helland R, Willassen NP *et al.* Comparative sequence and structure analysis reveal features of cold adaptation of an enzyme in the thermolysin family. *Proteins Struct Funct Bioinf* 2006;**62**:435–49.
- Aghajari N, Feller G, Gerday C *et al.* Structures of the psychrophilic *Alteromonas haloplanctis* α -amylase give insights into cold adaptation at a molecular level. *Structure* 1998;**6**:1503–16.
- Albalat R, Cañestro C. Evolution by gene loss. *Nat Rev Genet* 2016;**17**:379.
- Albanese D, Coleine C, Rota-Stabelli O *et al.* Antarctic cryptoendolithic bacterial lineages of pre-Cambrian origin as proxy for Mars colonization. *bioRxiv*. 2020, DOI: 10.1101/2020.02.27.967604.
- Alcaño J, Cifuentes V, Baeza M. Physiological adaptations of yeasts living in cold environments and their potential applications. *World J Microbiol Biotechnol* 2015;**31**:1467–73.
- Bankevich A, Nurk S, Antipov D *et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J of Comput Biol* 2012;**19**:455–77.
- Bellora N, Moliné M, David-Palma M *et al.* Comparative genomics provides new insights into the diversity, physiology, and sexuality of the only industrially exploited tremellomycete: *Phaffia rhodozyma*. *BMC Genomics* 2016;**17**:901.

- Bijlani S, Singh NK, Mason CE *et al.* Draft genome sequences of tremellomycetes strains isolated from the International Space Station. *Microbiol Resour Announc* 2020;**9**:e00504–20.
- Blevins WR, Ruiz-Orera J, Messegueur X *et al.* Frequent birth of de novo genes in the compact yeast genome. *bioRxiv*, 575837 2019. DOI: 10.1101/575837.
- Bøe CA, Håland TW, Boye E *et al.* A novel role for ATR/Rad3 in G1 phase. *Sci Rep* 2018;**8**:1–12.
- Brandão LR, Libkind D, Vaz AB *et al.* Yeasts from an oligotrophic lake in Patagonia (Argentina): diversity, distribution and synthesis of photoprotective compounds and extracellular enzymes. *FEMS Microbiol Ecol* 2011;**76**:1–13.
- Buzzini P, Margesin R. *Cold-Adapted Yeasts. Miscellaneous Cold-Active Yeast Enzymes of Industrial Importance* New York: Springer, 2014.
- Cary SC, McDonald IR, Barrett JE *et al.* On the rocks: the microbiology of Antarctic Dry Valley soils. *Nat Rev Microbiol* 2010;**8**:129–38.
- Cheng CHC. Evolution of the diverse antifreeze proteins. *Curr Opin Genet Dev* 1998;**8**:715–20.
- Chintalapati S, Kiran MD, Shivaji S. Role of membrane lipid fatty acids in cold adaptation. *Cell Mol Biol (Noisy-le-grand)* 2004;**50**:631–42.
- Clark MS, Worland MR. How insects survive the cold: molecular mechanisms — a review. *J Compar Physiol B* 2008;**178**:917–33.
- Collins T, Margesin R. Psychrophilic lifestyles: mechanisms of adaptation and biotechnological tools. *Appl Microbiol Biotechnol* 2019;**103**:2857–71.
- Connell L, Redman R, Craig S *et al.* Diversity of soil yeasts isolated from South Victoria Land, Antarctica. *Microb Ecol* 2008;**56**:448–59.
- Connell LB, Rodriguez RR, Redman RS *et al.* Cold-adapted yeasts in Antarctic deserts. In *Cold-Adapted Yeasts* (pp. 75–98). Springer, Berlin, Heidelberg, 2014.
- Convey P, Peck LS. Antarctic environmental change and biological responses. *Sci Adv* 2019;**5**:eaaz0888.
- Conway JR, Lex A, Gehlenborg N. UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics* 2017;**33**:2938–40.
- Cowan DA, Casanueva A, Stafford W. Ecology and biodiversity of cold-adapted microorganisms. In *Physiology and Biochemistry of Extremophiles* (pp. 119–132). American Society of Microbiology, 2007. DOI: 10.1128/9781555815813.ch9.
- Darriba D, Posada D, Kozlov AM *et al.* ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. *Mol Biol Evol* 2020;**37**:291–4.
- Durocher D, Jackson SP. DNA-PK, ATM and ATR as sensors of DNA damage: variations on a theme? *Curr Opin Cell Biol* 2001;**13**:225–31.
- Eddy SR. A probabilistic model of local sequence alignment that simplifies statistical significance estimation. *PLoS Comput Biol* 2008;**4**:e1000069.
- Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004;**32**:1792–7.
- Ernst R, Ejsing CS, Antonny B. Homeoviscous adaptation and the regulation of membrane lipids. *J Mol Biol* 2016;**428**:4776–91.
- Ewart KV, Lin Q, Hew CL. Structure, function and evolution of antifreeze proteins. *Cell Mol Life Sci CMLS* 1999;**55**:271–83.
- Feller G, Gerday C. Psychrophilic enzymes: molecular basis of cold adaptation. *Cell Mol Life Sci CMLS* 1997;**53**:830–41.
- Ferrer M, Chernikova TN, Yakimov MM *et al.* Chaperonins govern growth of *Escherichia coli* at low temperatures. *Nat Biotechnol* 2003;**21**:1266–7.
- García-Fernández JM, de Marsac NT, Diez J. Streamlined regulation and gene loss as adaptive mechanisms in *Prochlorococcus* for optimized nitrogen utilization in oligotrophic environments. *Microbiol Mol Biol Rev* 2004;**68**:630–8.
- Gerday C, Aittaleb M, Bentahir M *et al.* Cold-adapted enzymes: from fundamentals to biotechnology. *Trends Biotechnol* 2000;**18**:103–7.
- Georlette D, Blaise V, Collins T *et al.* Some like it cold: biocatalysis at low temperatures. *FEMS Microbiol Rev* 2004;**28**:25–42.
- Gianese G, Argos P, Pascarella S. Structural adaptation of enzymes to low temperatures. *Protein Eng* 2001;**14**:141–8.
- Gillespie JH. *The Causes of Molecular Evolution (Vol. 2)*. Oxford University Press On Demand, 1994. ISBN-13: 978-0195092714.
- Gnerre S, MacCallum I, Przybylski D *et al.* High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc Natl Acad Sci* 2011;**108**:1513–8.
- Gremme G, Steinbiss S, Kurtz S. GenomeTools: a comprehensive software library for efficient processing of structured genome annotations. *IEEE/ACM Trans Comput Biol Bioinf* 2013;**10**:645–56.
- Grigoriev IV, Nikitin R, Haridas S *et al.* MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Res* 2014;**42**:D699–704.
- Grille S, Zaslowski A, Thiele S *et al.* The functions of steryl glycosides come to those who wait: Recent advances in plants, fungi, bacteria and animals. *Prog Lipid Res* 2010;**49**:262–88.
- Grzymalski JJ, Carter BJ, DeLong EF *et al.* Comparative genomics of DNA fragments from six Antarctic marine planktonic bacteria. *Appl Environ Microbiol* 2006;**72**:1532–41.
- Guzmán L. *Distribución filogenética de la producción de micosporinas en los phyla Ascomycota y Basidiomycota. [master's thesis]. [SC de Bariloche (Río Negro)]*, Universidad Nacional del Comahue, 2016.
- Hashim NHF, Sulaiman S, Bakar FDA *et al.* Molecular cloning, expression and characterisation of Afp4, an antifreeze protein from *Glaciozyma antarctica*. *Pol Biol* 2014;**37**:1495–505.
- Hébraud M, Potier P. Cold shock response and low temperature adaptation in psychrotrophic bacteria. *J Mol Microbiol Biotechnol* 1999;**1**:211–9.
- Huerta-Cepas J, Forslund K, Coelho LP *et al.* Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. *Mol Biol Evol* 2017;**34**:2115–22.
- Kandror O, Bretschneider N, Kreydin E *et al.* Yeast adapt to near-freezing temperatures by STRE/Msn2:4-dependent induction of trehalose synthesis and certain molecular chaperones. *Mol Cell* 2004;**13**:771–81.
- Kim HJ, Lee JH, Do H *et al.* Production of antifreeze proteins by cold-adapted yeasts. In *Cold-Adapted Yeasts* (pp. 259–280). Springer, Berlin, Heidelberg, 2014.
- Klingler JM, Vishniac HS. 3.4 Water potential of Antarctic soils. *Polarforschung* 1988;**58**, 231–8.
- Kondrashov FA. Gene duplication as a mechanism of genomic adaptation to a changing environment. *Proc R Soc B Biol Sci* 2012;**279**:5048–57.
- Koonin EV, Aravind L, Kondrashov AS. The impact of comparative genomics on our understanding of evolution. *Cell* 2000;**101**:573–6.
- Kozhina TN, Kozhin SA, Korolev VG. Gene RAD31 is identical to gene MEC1 of yeast *Saccharomyces cerevisiae*. *Russ J Genet* 2011;**47**:533–7.
- Kryazhimskiy S, Plotkin JB. The population genetics of dN/dS. *PLoS Genet* 2008;**4**:e1000304.
- Kurtzman C, Fell JW, Boekhout T (Eds.). *The Yeasts: A Taxonomic Study*. Elsevier. ISBN: 9780080931272, 2011.

- Lang GI, Murray AW, Botstein D. The cost of gene expression underlies a fitness trade-off in yeast. *Proc Natl Acad Sci* 2009;**106**:5755–60.
- Lee JK, Park KS, Park S et al. An extracellular ice-binding glycoprotein from an Arctic psychrophilic yeast. *Cryobiology* 2010;**60**:222–8.
- Levasseur A, Lomascolo A, Chabrol O et al. The genome of the white-rot fungus *Pycnoporus cinnabarinus*: a basidiomycete model with a versatile arsenal for lignocellulosic biomass breakdown. *BMC Genomics* 2014;**15**:486.
- Libkind D, Moliné M, Sampaio JP et al. Yeasts from high-altitude lakes: influence of UV radiation. *FEMS Microbiol Ecol* 2009;**69**:353–62.
- Libkind D, Moliné M, Sommaruga R et al. Phylogenetic distribution of fungal mycosporines within the Pucciniomycotina (Basidiomycota). *Yeast* 2011;**28**:619–27.
- Libkind D, Peris D, Cubillos FA et al. Into the wild: new yeast genomes from natural environments and new tools for their analysis. *FEMS Yeast Res* 2020;**20**:foaa008.
- Madronich S, McKenzie RL, Björn LO et al. Changes in biologically active ultraviolet radiation reaching the Earth's surface. *J Photochem Photobiol B* 1998;**46**:5–19.
- Margesin R, Miteva V. Diversity and ecology of psychrophilic microorganisms. *Res Microbiol* 2011;**162**:346–61.
- Margesin R, Collins T. Microbial ecology of the cryosphere (glacial and permafrost habitats): current knowledge. *Appl Microbiol Biotechnol* 2019;**103**:2537–49.
- Marsit S, Sanchez I, Galeote V et al. Horizontally acquired oligopeptide transporters favour adaptation of *Saccharomyces cerevisiae* wine yeast to oenological environment. *Environ Microbiol* 2016;**18**:1148–61.
- Mas A, Jamshidi S, Lagadeuc Y et al. Beyond the black queen hypothesis. *ISME J* 2016;**10**:2085–91.
- McKenzie RL, Aucamp PJ, Bais AF et al. Ozone depletion and climate change: impacts on UV radiation. *Photochem Photobiol Sci* 2011;**10**:182–98.
- Miller W, Makova KD, Nekrutenko A et al. Comparative genomics. *Annu Rev Genomics Hum Genet* 2004;**5**:15–56.
- Miller MA, Pfeiffer W, Schwartz T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In *2010 Gateway Computing Environments Workshop (GCE)* (pp. 1–8). IEEE, 2010. DOI: 10.1109/GCE.2010.5676129.
- Moliné M, Libkind D, del Carmen Diéguez M et al. Photoprotective role of carotenoids in yeasts: response to UV-B of pigmented and naturally-occurring albino strains. *J Photochem Photobiol B* 2009;**95**:156–61.
- Moliné M, Arbeloa EM, Flores MR et al. UVB photoprotective role of mycosporines in yeast: photostability and antioxidant activity of mycosporine-glutaminol-glucoside. *Radiat Res* 2011;**175**:44–50.
- Moliné M, Libkind D, de Garcia V et al. Production of pigments and photo-protective compounds by cold-adapted yeasts. In *Cold-Adapted Yeasts* (pp. 193–224). Springer, Berlin, Heidelberg, 2014. DOI: 10.1007/978-3-642-39681-6_9.
- Morris JJ, Lenski RE, Zinser ER. The Black Queen Hypothesis: evolution of dependencies through adaptive gene loss. *MBio* 2012;**3**:e00036–12.
- Muñoz M, *Levaduras y hongos dimórficos del floglano de Nothofagus pumilio y el papel de la exposición solar en su distribución y producción de metabolitos fotoprotectores. [master's thesis]. [SC de Bariloche (Río Negro)]*, Universidad Nacional del Comahue, 2010.
- Nizovoy P, Moliné M, Libkind D et al. Cleaning 'Nwanted regions: a novel approach to getting rid of N content by assemblies combination. PeerJ Preprints [Preprint]. Available at: <https://peerj.com/preprints/3504.pdf>, 2018.
- Onofri S, Zucconi L, Selbmann L et al. Fungal associations at the cold edge of life. In *Algae and Cyanobacteria in Extreme Environments* (pp. 735–757). Springer, Dordrecht, 2007. DOI: 10.1007/978-1-4020-6112-7_40.
- Palmer JM. Funannotate: a fungal genome annotation and comparative genomics pipeline. Available at: <https://github.com/nextgenusfs/funannotate>, 2016.
- Parra G, Bradnam K, Ning Z et al. Assessing the gene space in draft genomes. *Nucleic Acids Res* 2008;**37**:289–97.
- Phadtare S, Inouye M, Severinov K. The mechanism of nucleic acid melting by a CspA family protein. *J Mol Biol* 2004;**337**:147–55.
- Riley R, Haridas S, Wolfe KH et al. Comparative genomics of biotechnologically important yeasts. *Proc Natl Acad Sci* 2016;**113**:9882–7.
- Rodríguez-Peña JM, García R, Nombela C et al. The high-osmolarity glycerol (HOG) and cell wall integrity (CWI) signalling pathways interplay: a yeast dialogue between MAPK routes. *Yeast* 2010;**27**:495–502.
- Rudolph AS, Crowe JH, Crowe LM. Effects of three stabilizing agents—proline, betaine, and trehalose—on membrane phospholipids. *Arch Biochem Biophys* 1986;**245**:134–43.
- Sælensminde G, Halskau Ø, Jonassen I. Amino acid contacts in proteins adapted to different temperatures: hydrophobic interactions and surface charges play a key role. *Extremophiles* 2009;**13**:11.
- Salazar M, Vongsangnak W, Panagiotou G et al. Uncovering transcriptional regulation of glycerol metabolism in *Aspergilli* through genome-wide gene expression data analysis. *Mol Genet Genomics* 2009;**282**:571.
- Saunders NF, Thomas T, Curmi PM et al. Mechanisms of thermal adaptation revealed from the genomes of the Antarctic Archaea *Methanogenium frigidum* and *Methanococcoides burtonii*. *Genome Res* 2003;**13**:1580–8.
- Schade B, Jansen G, Whiteway M et al. Cold adaptation in budding yeast. *Mol Biol Cell* 2004;**15**:5492–502.
- Schmidt SK, Vimercati L, Darcy JL et al. A *Naganishia* in high places: functioning populations or dormant cells from the atmosphere? *Mycology* 2017;**8**:153–63.
- Schuster-Böckler B, Conrad D, Bateman A. Dosage sensitivity shapes the evolution of copy-number varied regions. *PLoS One* 2010;**5**:e9474.
- Scorzetti G, Petrescu I, Yarrow D et al. *Cryptococcus adeliensis* sp. nov., a xylanase producing basidiomycetous yeast from Antarctica. *Antonie Van Leeuwenhoek*, 2000;**77**:153–7.
- Sedmak JJ, Weerasinghe DK, Jolly SO. Extraction and quantitation of astaxanthin from *Phaffa rhodozyma*. *Biotechnol Tech* 1990;**4**:107–12.
- Selbmann L, Zucconi L, Onofri S et al. Taxonomic and phenotypic characterization of yeasts isolated from worldwide cold rock-associated habitats. *Fung Biol* 2014;**118**:61–71.
- Shi J, Cao Y, Fan X et al. A rice microsomal delta-12 fatty acid desaturase can enhance resistance to cold stress in yeast and *Oryza sativa*. *Mol Breed* 2012;**29**:743–57.
- Siddiqui KS, Cavicchioli R. Cold-adapted enzymes. *Annu Rev Biochem*, 2006;**75**:403–33.
- Singh A, Krishnan KP, Prabakaran D et al. Lipid membrane modulation and pigmentation: a cryoprotection mechanism in Arctic pigmented bacteria. *J Basic Microbiol* 2017;**57**:770–80.
- Sliwa P, Korona R. Loss of dispensable genes is not adaptive in yeast. *Proc Natl Acad Sci* 2005;**102**:17670–4.

- Spielman SJ, Wilke CO. The relationship between dN/dS and scaled selection coefficients. *Mol Biol Evol* 2015;**32**:1097–108.
- Spiwok V, Lipovová P, Skálová T *et al.* Cold-active enzymes studied by comparative molecular dynamics simulation. *J Mol Model* 2007;**13**:485–97.
- Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014;**30**:1312–3.
- Stajich JE, Harris T, Brunk BP *et al.* FungiDB: an integrated functional genomics database for fungi. *Nucleic Acids Res* 2012;**40**:D675–81.
- Steinway SN, Dannenfeller R, Laucus CD *et al.* JCoDA: a tool for detecting evolutionary selection. *BMC Bioinformatics* 2010;**11**:284.
- Su Y, Jiang X, Wu W *et al.* Genomic, transcriptomic, and proteomic analysis provide insights into the cold adaptation mechanism of the obligate psychrophilic fungus *Mrakia psychrophila*. *G3 Genes Genomes Genet* 2016;**6**:3603–13.
- Suyama M, Torrents D, Bork P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res* 2006;**34**:W609–12.
- Tapia H, Young L, Fox D *et al.* Increasing intracellular trehalose is sufficient to confer desiccation tolerance to *Saccharomyces cerevisiae*. *Proc Natl Acad Sci* 2015;**112**:6122–7.
- Ter-Hovhannisyan V, Lomsadze A, Chernoff YO *et al.* Gene prediction in novel fungal genomes using an *ab initio* algorithm with unsupervised training. *Genome Res* 2008;**18**:1979–90.
- Thatcher JW, Shaw JM, Dickinson WJ. Marginal fitness contributions of nonessential genes in yeast. *Proc Natl Acad Sci* 1998;**95**:253–7.
- Thieringer HA, Jones PG, Inouye M. Cold shock and adaptation. *Bioessays* 1998;**20**:49–57.
- Tronelli D, Maugini E, Bossa F *et al.* Structural adaptation to low temperatures— analysis of the subunit interface of oligomeric psychrophilic enzymes. *FEBS J* 2007;**274**:4595–608.
- Ullrich SR, González C, Poehlein A *et al.* Gene loss and horizontal gene transfer contributed to the genome evolution of the extreme acidophile 'Ferrofum'. *Front Microbiol* 2016;**7**:797.
- Vacic V, Uversky VN, Dunker AK *et al.* Composition Profiler: a tool for discovery and visualization of amino acid composition differences. *BMC Bioinformatics* 2007;**8**:211.
- Vajpeyi S, Chandran K. Draft genome sequence of the oleaginous yeast *Cryptococcus albidus* var. *albidus*. *Genome Announc* 2016;**4**:e00390–16.
- Vinnere Pettersson O, Leong SLL. *Fungal xerophiles (osmophiles)* 2001. DOI: 10.1186/1471-2105-8-211.
- Vishniac HS, Hempfing WP. *Cryptococcus vishniacii* sp. nov., an Antarctic yeast. *Int J Syst Evol Microbiol* 1979a;**29**:153–8.
- Vishniac HS, Hempfing WP. Evidence of an indigenous microbiota (yeast) in the dry valleys of Antarctica. *Microbiology* 1979b;**112**:301–14.
- Vishniac H. Yeast biodiversity in the Antarctic. Biodiversity and Ecophysiology of Yeasts, *The Yeast Handbook*, Rosa CA, Peter G (Eds). Heidelberg: Springer, **25**:419–40, 2006.
- Viviani MA, Tortorano AM. *Cryptococcus*. In *Clinical Mycology* (pp. 231–249). Churchill Livingstone, 2009. ISBN: 978-1-4160-5680-5.
- Wang D, Zhang Y, Zhang Z *et al.* KaKs-Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. *Genom Proteom Bioinformatics* 2010;**8**:77–80.
- Wang C, Pakhomova S, Newcomer ME *et al.* Structural basis of antifreeze activity of a bacterial multi-domain antifreeze protein. *PLoS One* 2017;**12**:e0187169.
- Weisman CM, Murray AW, Eddy SR. Many but not all lineage-specific genes can be explained by homology detection failure. BioRxiv. DOI: 10.1101/2020.02.27.968420, 2020.
- Wiemken A. Trehalose in yeast, stress protectant rather than reserve carbohydrate. *Antonie Van Leeuwenhoek* 1990;**58**:209–17.
- Wilson GA, Bertrand N, Patel Y *et al.* Orphans as taxonomically restricted and ecologically important genes. *Microbiology* 2005;**151**:2499–501.
- Wolfe KH. Comparative genomics and genome evolution in yeasts. *Philos Trans R Soc B Biol Sci* 2006;**361**:403–12.
- Yadav AN, Verma P, Sachan SG *et al.* Psychrotrophic microbiomes: molecular diversity and beneficial role in plant growth promotion and soil health. In *Microorganisms for Green Revolution* (pp. 197–240). Springer, Singapore, 2018. DOI: 10.1007/978-981-10-7146-1_11.
- Yang Z, Nielsen R. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Mol Biol Evol* 2000;**17**:32–43.
- Yong X, Yan Z, Xu L *et al.* Genome sequence of a microbial lipid producing fungus *Cryptococcus albidus* NT2002. *J Biotechnol* 2016;**223**:6–7.
- Zalar P, Gunde-Cimerman N. Cold-adapted yeasts in Arctic habitats. In *Cold-Adapted Yeasts* (pp. 49–74). Springer, Berlin, Heidelberg, 2014. DOI:10.1007/978-3-642-39681-6.
- Zhang X, Liu X, Liang Y *et al.* Adaptive evolution of extreme acidophile *Sulfobacillus thermosulfidooxidans* potentially driven by horizontal gene transfer and gene loss. *Appl Environ Microbiol* 2017;**83**:e03098–16.
- Zhu J, Dong CH, Zhu JK. Interplay between cold-responsive gene regulation, metabolism and RNA processing during plant cold acclimation. *Curr Opin Plant Biol* 2007;**10**:290–5.