



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Maestría y Doctorado en Ciencias Bioquímicas

ANÁLISIS DEL METAGENOMA DE LEVADURA
DE UN QUESO ARTESANAL MEXICANO

TESIS

QUE PARA OPTAR POR EL GRADO DE:
Maestra en Ciencias

PRESENTA:

Q.A. Blanca Daniela Martínez Aviña

TUTOR PRINCIPAL

Dra. Maricarmen Quirasco Baruch
[Facultad de Química, UNAM](#)

MIEMBROS DEL COMITÉ TUTOR

Dr. Francisco Ruiz Terán
[Facultad de Química, UNAM](#)

Dr. Luis David Alcaraz Peraza
[Facultad de Ciencias, UNAM](#)

Ciudad de México. Junio, 2023



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

ANÁLISIS DEL METAGENOMA DE LEVADURA DE UN QUESO ARTESANAL MEXICANO

Jurado Asignado

Presidente	Dra. María Amanda Gálvez Mariscal
Secretario	Dr. Adrián Ochoa Leyva
Vocal	Dr. José Adelfo Escalante Lozada
Vocal	Dra. Rosa María Gutiérrez Ríos
Vocal	Dra. Esperanza Martínez Romero

Este trabajo fue realizado en el Laboratorio 312 del Departamento de Alimentos y Biotecnología del Conjunto E de la Facultad de Química de la Universidad Nacional Autónoma de México, bajo la tutoría de la Dra. Maricarmen Quirasco Baruch. El análisis bioinformático se llevó a cabo con el apoyo del Dr. Luis Fernando Lozano Aguirre de la Unidad de Análisis Bioinformáticos del Centro de Ciencias Genómicas de la Universidad Nacional Autónoma de México, campus Cuernavaca, Morelos.

Este trabajo fue financiado por el PAIP FQ con la clave 5000-9102. Se agradece a CONACyT por la beca otorgada.

El Comité Tutor estuvo conformado por:

Dr. Francisco Ruiz Terán	Facultad de Química, UNAM
Dr. Luis David Alcaraz Peraza	Facultad de Ciencias, UNAM

ÍNDICE

Resumen

Glosario

1 Marco Teórico	9
1.1 Quesos y su maduración.....	9
1.2 Importancia de las levaduras en quesos madurados	9
1.2.1 Levaduras aisladas de quesos	11
1.3 Análisis metagenómicos.....	13
1.3.1 Metagenómica de escopeta y dirigida	14
1.3.2 Marcadores para la identificación de microorganismos fúngicos.....	15
1.4 Análisis bioinformático de los datos	17
2 Antecedentes	22
2.1 Queso Cotija	22
2.1.1 Proceso de elaboración	23
2.1.2 Microorganismos presentes en el queso	24
3 Justificación	28
4 Hipótesis	28
5 Objetivos	28
5.1 Objetivo general	28
5.2 Objetivos particulares.....	28
6 Diagrama experimental	29
7 Metodología	29
7.1 Selección de muestras	29
7.2 Tratamiento de las muestras.....	29
7.3 Obtención del paquete celular.....	29
7.4 Extracción y purificación del ADN metagenómico	30
7.5 Secuenciación del ADN metagenómico	31
7.6 Análisis bioinformático de las secuencias	32
7.6.1 Calidad de las secuencias	32
7.6.2 Datos de secuenciación dirigida a la región ITS1	32

7.6.3	Datos de secuenciación <i>shotgun</i>	33
8	Resultados y Discusión	36
8.1	Extracción y purificación de ADN	36
8.2	Análisis bioinformático de las secuencias	36
8.2.1	Calidad de las secuencias	36
8.2.2	Secuenciación dirigida a la región ITS1.....	39
8.2.3	Datos metagenómicos de <i>shotgun</i>	56
8.2.4	Análisis metagenómico dirigido y <i>shotgun</i>	62
9	Conclusiones	64
10	Perspectivas	65
	Referencias	66
	Anexo	76
I.	Resultados obtenidos por Escobar (2016) con el programa Parallel-meta.....	76
II.	Cebadores utilizados en la secuenciación de la región ITS1	79
III.	Soluciones	79
IV.	Asignación taxonómica de las secuencias de la región ITS1	80
V.	Asignación taxonómica de las secuencias metagenómicas de <i>shotgun</i>	86

Resumen

El queso Cotija es un producto lácteo artesanal mexicano que se elabora a partir de leche cruda de vaca, sin la adición de cultivos iniciadores y es madurado por al menos 3 meses en la región de origen. La microbiota incorporada de forma no intencional durante su manufactura, así como la propia de los ingredientes utilizados, interaccionan con los componentes del alimento y es clave para el desarrollo de las propiedades y características distintivas de este queso.

Mediante estudios de microbiología clásica, métodos moleculares dependientes de PCR y por secuenciación, se ha identificado que su población microbiana está conformada mayoritariamente por bacterias *Firmicutes*. Mientras que solo algunas especies de levaduras han sido descritas, las cuales tienen un papel importante en el proceso de maduración al contribuir en el desarrollo de compuestos de sabor y olor gracias a sus actividades proteolítica y lipolítica.

Con el avance en las tecnologías de secuenciación y de las técnicas moleculares independientes de cultivo, ha sido posible obtener más información sobre la diversidad microbiológica en alimentos tradicionales fermentados de manera natural y, se podría obtener un perfil más completo sobre las levaduras.

El objetivo principal de este trabajo fue identificar la diversidad de levaduras presentes en tres muestras de Queso Cotija Región de Origen mediante análisis a la región ITS1 y contrastar con lo descrito según el análisis de datos de secuenciación de *shotgun*. Se encontró que, mediante dos estrategias distintas de análisis de los datos de ITS1, la comunidad de levaduras está dominada por las familias *Phaffomycetaceae*, *Debaryomycetaceae* y *Saccharomycetaceae*. Los géneros dominantes con más del 18% de abundancia, fueron *Komagataella*, *Starmarella* y *Debaryomyces*, el primero de ellos fue común entre las muestras. Con la primera estrategia de análisis basada en unidades taxonómicas operacionales (OTUs, por sus siglas en inglés) se identificaron 60 géneros, mientras que con la segunda estrategia basada en variantes de secuencia de amplicones (ASVs, por sus siglas en inglés), se asignaron 18 géneros. Ambas estrategias identificaron los mismos géneros dominantes y subdominantes entre las muestras. Con respecto a los datos de secuenciación de *shotgun*, se corroboró que el grupo microbiano dominante en el queso Cotija son bacterias y solo una fracción menor de levaduras, dentro de las cuales *Debaryomyces hansenii* es la especie dominante.

GLOSARIO

Actividad de agua (aw): Es definida como la cantidad de agua disponible o libre en un alimento que puede ser utilizado por los microorganismos para su crecimiento o para llevar a cabo diferentes reacciones químicas. Matemáticamente, se representa como la relación entre la presión de vapor de agua del alimento (P) y la presión de vapor de agua pura (P₀), ambos permaneciendo a una temperatura constante.

$$a_w = (P) / (P_0)$$

Ancestro común más bajo (LCA, siglas en inglés): Algoritmo basado en la teoría de grafos y utilizado para estimar o inferir la clasificación taxonómica de una secuencia dada. Corresponde al taxón más cercano compartido entre dos o más taxones encontrados por una búsqueda para una secuencia. Por ejemplo, si una secuencia de consulta tiene dos asignaciones o hits pertenecientes a dos especies diferentes, una de *Lactobacillus acidophilus* y la otra de *Lactobacillus casei*, el algoritmo LCA asigna la secuencia de consulta al género *Lactobacillus*, que es el nivel taxonómico común más bajo de estas dos especies.

Brecha de código de barras (barcoding gap): Se refiere a la diferencia que existe entre la variación interespecífica (entre especies cercanamente relacionadas) e intraespecífica (dentro de la misma especie) dentro de un grupo de organismos. Cuanto mayor sea la diferencia entre estas, la discriminación específica será más acertada (Meyer y Paulay, 2005).

Células somáticas: Células blancas propias del organismo que le sirven como defensa a la glándula mamaria de la vaca contra organismos patógenos.

Código de barras: Secuencia corta estandarizada de ADN de entre 400 y 800 bp de largo que puede ser fácilmente aislada y caracterizada para todas las especies en el planeta. El uso de tales secuencias cortas de ADN para identificaciones biológicas fue propuesto por Hebert *et al.* (2003) con el objetivo final de una identificación rápida y fiable a nivel de especies en todas las formas de vida, incluyendo animales, plantas y microorganismos.

Contig: Es un conjunto de segmentos o secuencias de ADN que se superponen de manera que proporciona una secuencia contigua de una región genómica. Dentro de la secuencia ensamblada no debe de haber huecos (gaps) o bases ambiguas diferentes de A, C, G o T.

Desreplicación: Etapa durante el preprocesamiento de los datos secuenciados donde se condensa los datos colapsando todas las lecturas que presentan la misma secuencia, lo que reduce significativamente los tiempos de procesamiento posteriores.

Firma genómica: Secuencia de nucleótidos con longitud dada presente en un genoma y taxón específico, utilizada para detectar ADN transferido horizontalmente, reconstruir relaciones filogenéticas e inferir estilos de vida de bacteriófagos, así como para asignar fragmentos de secuencias metagenómicas a taxones microbianos.

k-mero: Subsecuencia de longitud “k” obtenida de una secuencia dada ya sea de ADN, ARN o proteína.

Lectura: Secuencia de ADN que se obtiene de la secuenciación de un genoma o metagenoma.

Lecturas pareadas: Par de secuencias de ADN obtenidas del proceso de secuenciación desde ambos extremos de cada fragmento, proveen información sobre localización relativa y orientación de cada par.

Marca colectiva: signo distintivo que sirve para diferenciar en el mercado los productos o servicios de los miembros de las asociaciones o sociedades de productores, fabricantes, comerciantes o prestadores de servicios legalmente constituidos, quienes están obligados a cumplir con las reglas de uso establecidas.

Perl: Lenguaje de programación diseñado por Larry Wall en 1987, ha sido ampliamente adoptado por su destreza en el procesamiento de texto y no tener ninguna de las limitaciones de los otros lenguajes.

Puntuación Phred (*Phred quality score*, Q): Métrica utilizada para evaluar la precisión de una plataforma de secuenciación. Indica la probabilidad de error de que una base dada sea llamada incorrectamente por el secuenciador (P), toma valores que varían de 2 a 40, puntajes más altos indican una mayor confianza en la llamada. Está definida matemáticamente por:

$$Q = -10 \log_{10} (P)$$

Región de Origen: Región geográfica del país que sirve para designar un producto originario de la misma, y cuya calidad o características se deban exclusivamente al medio geográfico.

Unidades formadoras de colonias (UFC): Unidad de medida para cuantificar el número de células viables en una muestra sólida o líquida, es el número mínimo de células separables en un medio de cultivo y que da lugar al desarrollo de una colonia.

Secuencias adaptadoras: Secuencias cortas de ADN de doble cadena que se incorporan a los extremos de los fragmentos cuya secuencia se desea conocer, son necesarios para llevar a cabo la secuenciación y su patrón de nucleótidos varía de acuerdo a la tecnología utilizada.

Potencial metabólico: Conjunto de predicciones de función obtenidas a partir de secuencias metagenómicas que no se asocian generalmente a un género o especie en particular, sino al consorcio como un sistema completo.

Modelos ocultos de Markov (HMM, por sus siglas en inglés): Herramienta del aprendizaje automático que permite explorar las secuencias de ADN o de aminoácidos y detectar componentes estructurales, localizar genes y otros elementos funcionales. Se basa en modelos estadísticos y de probabilidad para representar procesos aleatorios que consisten de estados de un sistema en etapas sucesivas; el sistema puede realizar transiciones desde un estado a otro y la probabilidad de estos cambios depende solo del estado anterior.

1 Marco Teórico

1.1 Quesos y su maduración

De acuerdo con la norma NOM-243-SSA1-2010, el queso es un derivado de la leche obtenido por la coagulación de la caseína por acción del cuajo, al cual se le pueden incorporar ingredientes comestibles dando lugar a distintas variedades.

En México se estima que alrededor del 70% de los quesos son elaborados de forma artesanal y, a pesar de que los quesos frescos como el Panela o el Oaxaca son los más consumidos, existen otras variedades semi-maduras y maduras, y entre los cuales se encuentra el queso Cotija, del cual se hablará en un apartado posterior.

1.2 Importancia de las levaduras en quesos madurados

El proceso de maduración de los quesos se caracteriza por cambios en los componentes del alimento y en los parámetros fisicoquímicos. Es resultado de la actividad metabólica de la microbiota presente y de la acción de las enzimas de los ingredientes utilizados, como las de la leche o del cuajo, así como por las sintetizadas por los microorganismos. Estos cambios incluyen procesos fermentativos, proteolíticos y lipolíticos, que generan moléculas que intensifican el aroma y el sabor, disminuyen la actividad de agua (a_w) y el producto final adquiere características sensoriales distintivas (Beresford *et al.*, 2001, Weimer, 2007).

Se han llevado a cabo múltiples estudios con el objetivo de caracterizar la microbiota presente en los quesos producidos artesanalmente y con éstos se ha demostrado que son un alimento microbiológicamente dinámico que alberga distintos grupos microbianos que son metabólicamente activos (Ndoye, 2011).

El grupo más estudiado es el bacteriano, en particular las bacterias ácido lácticas (LAB, por sus siglas en inglés). Éstas se pueden clasificar en dos grupos de acuerdo a su capacidad para modificar las propiedades fisicoquímicas del queso, mismas que se relacionan con sus características metabólicas. Al primero se le denomina como LAB iniciadoras (SLAB, por sus siglas en inglés) y el segundo como LAB no iniciadoras (NSLAB). Las bacterias SLAB son homofermentadoras cuyo metabolito principal de la fermentación de carbohidratos es el ácido láctico y por tanto son responsables de la disminución drástica del pH en la primera etapa de fermentación. Mientras que el segundo grupo tiene importancia en el proceso de maduración al producir diferentes metabolitos relacionados con olor y sabor en el queso, y cuyo metabolismo es principalmente heterofermentativo. Algunos ejemplos de bacterias SLAB son *Lactobacillus acidophilus*, *Lactobacillus helveticus* y *Lactococcus lactis*; y para el grupo NSLAB se encuentran los géneros *Lactobacillus* (*Lactobacillus*

plantarum y *Lactobacillus casei*), *Leuconostoc*, *Pediococcus*, *Streptococcus*, entre otros (Escobar, 2016; Bintsis, 2018).

Las bacterias LAB también contribuyen a la maduración a través de su actividad proteolítica y de conversión de aminoácidos en compuestos que otorgan sabor al producto lácteo. Así mismo, se ha asociado a las bacterias LAB con la inocuidad de los quesos artesanales debido a su capacidad para producir diferentes sustancias antagónicas como bacteriocinas y péptidos antimicrobianos, peróxido de hidrógeno, diacetilo, ácidos orgánicos y moléculas de bajo peso molecular como heterociclos (Bassi *et al.*, 2020).

Aunque las bacterias LAB son responsables de la producción de ácidos orgánicos y ayudan al proceso de maduración, otros microorganismos presentes como las levaduras tienen una aportación importante en el proceso de maduración del queso, sin embargo, la contribución de éstas al desarrollo de compuestos de sabor y olor durante dicha etapa generalmente se subestima.

Algunas especies de levaduras se encuentran naturalmente en muchos quesos, pero particularmente en los elaborados con leche cruda y pueden crecer durante la fabricación, así como en las primeras etapas de maduración. Pueden desarrollarse tanto en la superficie, alcanzando valores de entre 5 a 6 log UFC g⁻¹, como en el interior de numerosos tipos de quesos blandos o semiblandos, llegando hasta 9 log UFC g⁻¹ (Beresford y Williams, 2004; Jacques y Casaregola, 2008).

La diferencia en la proporción y tipo de levaduras en los productos lácteos y quesos puede estar relacionada con la diferencia en los procesos de manufactura, la ubicación geográfica, las diferentes condiciones climáticas y los ecosistemas de los pastizales o vegetación que se ve reflejada en la calidad de la leche obtenida durante la ordeña (Sessou *et al.*, 2019).

La prevalencia de las levaduras en los productos lácteos y en particular en quesos se atribuye a su tolerancia a pH ácido, a una *a_w* reducida y a concentraciones elevadas de sal. Otras características que pueden explicar su crecimiento en los quesos y su contribución al perfil sensorial del producto final son la capacidad de metabolizar la lactosa residual y el ácido láctico producido por las bacterias LAB, formación de metabolitos alcalinos y aromáticos, actividades lipolítica y proteolítica.

Además, se encuentran ampliamente distribuidas en los entornos de producción de lácteos y aparecen como contaminantes naturales tanto en la leche cruda, el aire y la salmuera, como en los utensilios y el agua empleados durante la manufactura de quesos. También se ha informado de la capacidad de ciertas especies de levaduras aisladas de productos lácteos para sobrevivir a los tratamientos térmicos, a los agentes desinfectantes y sanitizantes, por ejemplo *Debaryomyces hansenii* resiste

al dióxido de cloro, *Rhodotorula mucilaginosa* a detergentes alcalinos, *Diklera custersiana* a detergentes ácidos y *Torulaspota delbrueckii* a solución comercial de yodo (van den Tempel y Jakobsen, 2000; Ferreira y Viljoen, 2003; Jacques y Casaregola, 2008).

Por otro lado, las levaduras también pueden actuar como organismos de descomposición causando defectos en los quesos como mal sabor, pérdida de la calidad de la textura, formación excesiva de gas, aumento de la acidez debido a los efectos estimulantes sobre las bacterias LAB y coloración indeseable en la superficie (Gardini *et al.*, 2006). Por ejemplo, Tofalo *et al.* (2014) asocian a *Yarrowia lipolytica*, *Candida catenulata* y a *D. hansenii* con la pigmentación en quesos al presentar actividad de tirosinasa y llevar a cabo la oxidación de la tirosina a melanina. Mientras que las levaduras *Kluyveromyces lactis*, *Pichia guilliermondii* y *Saccharomyces cerevisiae* se identificaron como productoras de gas en quesos españoles como el Arzúa-Ulloa, Tetilla y Cebeiro (Atanassova *et al.*, 2016).

1.2.1 Levaduras aisladas de quesos

Recientemente se ha incrementado el interés por conocer las levaduras que forman parte de la microbiota de varios alimentos y se han reportado diferentes especies que han sido aisladas no solo de productos lácteos, sino también de otros alimentos fermentados. El crecimiento de diferentes especies de levaduras en algunos quesos se ve favorecido por el microambiente propio del alimento y por las características metabólicas del microorganismo que le permiten crecer en mayor o menor proporción.

Se han reportado a las especies *Kluyveromyces marxianus*, *K. lactis*, *D. hansenii*, *Y. lipolytica*, entre otras como productoras de enzimas lipolíticas y/o proteolíticas que logran metabolizar las proteínas y la grasa presente en la leche o en el queso, por lo cual estas especies tienen un papel significativo en el desarrollo de compuestos de aroma y sabor durante la maduración de los quesos (Bolumar *et al.*, 2008)

En la tabla 1 se enlistan algunas especies aisladas e identificadas mediante técnicas moleculares en diferentes quesos. Por ejemplo, Gardini *et al.* (2006) estudiaron las levaduras presentes en el queso Pecorino en diferentes etapas de su producción y encontraron que la mayor cuenta se obtenía durante las primeras horas de salado y se mantenía durante la maduración. La cuenta más alta fue de $5.11 \log \text{ UFC g}^{-1}$ a los 15 días de madurado e identificaron a 7 levaduras de las cuales *D. hansenii* fue el organismo predominante.

Tabla 1. Levaduras aisladas de diversos quesos semi-madurados y madurados

Muestra	Especies	Técnica de identificación	Referencia
Tipo parmesano	<i>Debaryomyces hansenii</i> <i>Galactomyces candidus</i>	Aislamiento en medio YGCA; amplificación y secuenciación ITS	Banjara <i>et al.</i> , 2015
Pecorino crotonese	<i>Debaryomyces hansenii</i> <i>Candida intermedia</i> <i>Candida inconspicua</i> <i>Kluyveromyces lactis</i> <i>Yarrowia lipolytica</i> <i>Saccharomyces cerevisiae</i> <i>Pichia carsonii</i>	Aislamiento en medio agar Sabouraud dextrosa; RFLP-PCR región 5.8S-ITS	Gardini <i>et al.</i> , 2006
De cuajada agria	<i>Debaryomyces hansenii</i> <i>Kluyveromyces marxianus</i> <i>Candida caterulata</i> <i>Geotrichum candidum</i>	Aislamiento en medio YGCA; amplificación región D1/D2 del gen ribosomal 26S	Lopandic <i>et al.</i> , 2006
Azules	<i>Debaryomyces hansenii</i> <i>Saccharomyces cerevisiae</i> <i>Yarrowia lipolytica</i> <i>Candida sake</i>	Aislamiento en medio YEPD; RFLP-PCR región 5.8S-ITS	Lopandic <i>et al.</i> , 2006
Kashkaval	<i>Kluyveromyces marxianus</i> <i>Debaryomyces hansenii</i> <i>Candida albicans</i> <i>Rhodotorula mucilaginosa</i> <i>Meyerozyma guilliermondii</i> <i>Candida parapsilosis</i>	Aislamiento en medio YGCA y PDA; RAPD-PCR región D1/D2 del gen ribosomal 26S	Zmen <i>et al.</i> , 2020
De cabra marroquí	<i>Kluyveromyces lactis</i> <i>Saccharomyces cerevisiae</i> <i>Yarrowia lipolytica</i> <i>Candida parapsilosis</i> <i>Pichia fermentans</i> <i>Kazachstania unispora</i> <i>Kluyveromyces marxianus</i>	Aislamiento en medio agar Sabouraud dextrosa y PDA; Pruebas bioquímicas API ID 32C; amplificación región D1/D2 del gen ribosomal 26S, región ITS-5.8S, gen ACT1 y gen COX2	Ouadghiri <i>et al.</i> , 2014
Tradicional turco	<i>Yarrowia lipolytica</i> <i>Kluyveromyces marxianus</i> <i>Pichia membranifaciens</i> <i>Debaryomyces hansenii</i> <i>Pichia fermentans</i> <i>Candida zeylanoides</i>	Aislamiento en medio YGCA y PDA; RAPD-PCR región D1/D2 del gen ribosomal 26S	Melike <i>et al.</i> , 2021
Surk	<i>Debaryomyces hansenii</i> <i>Candida zeylanoides</i> <i>Yarrowia lipolytica</i> <i>Kluyveromyces lactis</i>	Aislamiento en medio PDA; RAPD-PCR región D1/D2 del gen ribosomal 26S	Yusuf y Bülent, 2021

YGCA – Agar extracto de levadura, glucosa, cloranfenicol

YEPD – Agar extracto de levadura, peptona, dextrosa

RFLP-PCR (siglas en inglés) - Polimorfismo de longitud de fragmentos de restricción.

RAPD-PCR (siglas en inglés) - Amplificación aleatoria de ADN polimórfico

En otro estudio realizado sobre productos lácteos como yogurt, queso fresco, queso madurado, cuajada, entre otros, Lopandic *et al.* (2006) reportaron que las especies más comunes fueron *Candida zeylanoides*, *Geotrichum candidum*, *D. hansenii*, *K. marxianus* y *Y. lipolytica*.

D. hansenii es una levadura halotolerante que, al tener capacidad para metabolizar el lactato y otras fuentes de carbono como lactosa, la convierten en una especie idónea para habitar en la superficie del queso y, por lo tanto, es una de las especies de levadura frecuentemente aislada en este tipo de alimentos. La especie es capaz de producir metabolitos alcalinos como el amoníaco a partir de la desaminación de aminoácidos y recientemente se demostró su potencial para producir compuestos volátiles que contribuyen al sabor del queso, como aldehídos y alcoholes de cadena ramificada, aunque otras especies como *G. candidum* y *Y. lipolytica* tienen mayor capacidad metabólica (Fröhlich-Wyder *et al.*, 2019).

En el caso de *K. marxianus* es una de las principales especies de levadura de este género que se encuentra en los productos lácteos. Es capaz de influir en la maduración del queso a través de su actividad proteolítica y lipolítica, y la producción de compuestos aromáticos volátiles. Además, produce etanol y ácidos carboxílicos por la fermentación de lactosa y, a partir de ellos, producen ésteres y acetaldehído. Su crecimiento se ha asociado con sabores ácidos, de sidra, alcohólicos, fermentados y afrutados (Belloch *et al.*, 2011; Fröhlich-Wyder *et al.*, 2019).

Por su parte, *Y. lipolytica* pertenece a las levaduras dimórficas capaces, según las condiciones ambientales, de crecer en forma de levadura o en forma de micelio. Se sabe que posee una excelente capacidad para degradar proteínas y lípidos, y por ende la especie se encuentra frecuentemente en alimentos con alto contenido de proteínas y/o grasas. Debido a su peculiar metabolismo, se conoce que contribuye considerablemente a la maduración del queso al producir grandes cantidades de compuestos volátiles, como ácidos orgánicos, sulfuros, furanos y cetonas de cadena corta (Sørensen *et al.*, 2011). Además, la especie es capaz de degradar la caseína, hidrolizar triglicéridos, catabolizar ácidos grasos a metilcetonas y aminoácidos a compuestos de azufre volátiles y amoníaco (Fröhlich-Wyder *et al.*, 2019).

1.3 Análisis metagenómicos

Los microorganismos son parte importante de la diversidad en el planeta y se encuentran en diversos nichos desde suelos, ríos y océanos, hasta ambientes extremos como desiertos, aguas termales o glaciales. Sin embargo, la gran mayoría aún permanecen sin anotación y solo un pequeño porcentaje pueden cultivarse, se

estima que se ha identificado menos del 1% de la diversidad (Wooley *et al.*, 2010; Baltoumas *et al.*, 2023).

Un enfoque para explorar dicha diversidad es a través de la metagenómica que permite estudiar el material genético (genomas) de una comunidad mixta de microorganismos mediante la detección de genes funcionales o análisis por secuenciación (Zhang *et al.*, 2021). Se basa en extraer del ADN total de la muestra y usar las tecnologías de secuenciación de nueva generación (NGS, por sus siglas en inglés).

Con el análisis metagenómico y las tecnologías NGS se ha realizado la identificación más precisa de los grupos microbianos, incluidos aquellos que son difíciles de cultivar y/o están presentes en poca abundancia, proporcionando así un análisis más completo de la diversidad (Ercolini *et al.*, 2012). Además, este tipo de análisis proporciona a los investigadores e industrias información valiosa sobre la diversidad bacteriana, fúngica y viral en muchos alimentos producidos tradicionalmente. Por ejemplo, en granos de kéfir, Nalbantoglu *et al.* (2014) identificaron nuevas especies que se encontraban en su mayoría en baja abundancia, lo que indica una mayor resolución y precisión en el estudio de las comunidades microbianas presentes en el alimento.

En el caso particular de los quesos elaborados tradicionalmente se ha incrementado su estudio mediante metagenómica. Por mencionar algunos, Wolfe *et al.* (2014) identificaron la diversidad microbiana de 137 muestras de superficie de tres quesos madurados naturalmente y buscaron funciones relacionadas con la producción de compuestos volátiles a partir de datos metagenómicos. En otro estudio, Dugat-Bony *et al.* (2015) emplearon a la metagenómica para monitorear la dinámica de especies y funciones durante la maduración de quesos inoculados superficialmente con nueve géneros de hongos y bacterias.

Debido al avance de las tecnologías de secuenciación, el crecimiento de las bases de datos y el desarrollo de herramientas computacionales, la metagenómica se ha posicionado como una técnica confiable de análisis.

1.3.1 Metagenómica de escopeta y dirigida

Existen dos vías de análisis del ADN extraído de una muestra: mediante metagenómica o secuenciación de escopeta (*metagenomic shotgun sequencing*, en inglés) y mediante análisis dirigido a secuencias particulares de ADN, mejor denominado como *metaprofiling* (Sharpton *et al.*, 2014; Shikha *et al.*, 2021).

En metagenómica de *shotgun* se extrae, se fracciona y se secuencia de forma independiente todos los genomas existentes en la muestra, lo que permite analizar

tanto la biodiversidad microbiana como las capacidades funcionales. Dentro de los datos generados por este enfoque se encuentran lecturas de genes marcadores o de regiones características que pueden ser utilizadas para describir la composición microbiana, pero también pueden ser alineadas a múltiples genomas de referencias conocidos y anotados, y recuperar la información sobre genes, proteínas y funciones disponibles de los mismos e incluso se puede ensamblar y recuperar genomas (*metagenome-assembled genomes*, MAGs, por sus siglas en inglés).

En cambio, *metaprofiling* es una de las formas más sencillas y computacionalmente eficientes de caracterizar la diversidad microbiana a partir de un metagenoma. Este enfoque se puede utilizar para la clasificación taxonómica y filogenética en muestras que son grandes y/o complejas, como los alimentos. En este procedimiento se amplifica y analiza un solo gen marcador taxonómicamente informativo para uno o varios grupos microbianos y se comparan las lecturas con una base de datos de referencia, se identifican aquellas lecturas que son homólogas y se asignan taxonómicamente. Los genes marcadores utilizados con más frecuencia incluyen al gen ribosomal 16S para analizar la presencia de arqueas y bacterias, y a los genes ribosomales 18S y 28S para microorganismos eucariontes.

1.3.2 Marcadores para la identificación de microorganismos fúngicos

Los hongos y levaduras se encuentran entre los organismos eucariontes más abundantes de la tierra que, desafortunadamente, no han sido completamente identificados. Se estima que tan solo 100 000 especies han sido descritas de una diversidad aproximada de entre 1.5 a 5.1 millones (Badotti *et al.*, 2017).

El uso de los marcadores de códigos de barras para estos organismos ha permitido un avance en su identificación. El grupo del Dr. Paul Hebert de la Universidad Guelph, Canadá, desarrolló la herramienta del código de barras genético con el objetivo de identificar especies de manera más rápida y precisa. Esta región corresponde a una secuencia de ADN estandarizada que debe de cumplir tres criterios: i) contener variabilidad y divergencia genética significativa a nivel de especie, (ii) poseer sitios flanqueantes conservados para desarrollar cebadores de PCR universales para una amplia aplicación taxonómica, y (iii) tener una longitud de secuencia corta (400 – 800 pb) para facilitar las capacidades actuales de extracción y amplificación de ADN (Hebert *et al.*, 2003; Kress y Erickson, 2008; Stielow *et al.*, 2015).

Con el estudio de Schoch *et al.* (2012) se propuso a la región del espaciador transcrito interno, o ITS por sus siglas en inglés, como marcador o código de barras universal para hongos y levaduras. Los autores compararon cuatro regiones de ADN como posibles códigos de barras, tres regiones ribosómicas nucleares (ITS, LSU y SSU) y un gen codificante representativo, el gen de la subunidad grande de la ARN

polimerasa II o *rpb1*. Analizaron entre otros parámetros, su capacidad de amplificación por PCR y la brecha de códigos de barras (*barcoding gap*, en inglés). Observaron que las regiones ribosómicas tenían una mejor tasa de amplificación a comparación de la región codificante, así como la región ITS presentaba una mayor diferencia entre la variación intraespecífica e interespecífica en distintos linajes de hongos al obtener valores promedios entre 0.02 y 0.37, respectivamente (figura 1). Se espera que la variación interespecífica sea mayor que la variación intraespecífica en al menos 10 veces como se observó en el estudio (Čandek y Kuntner, 2015).

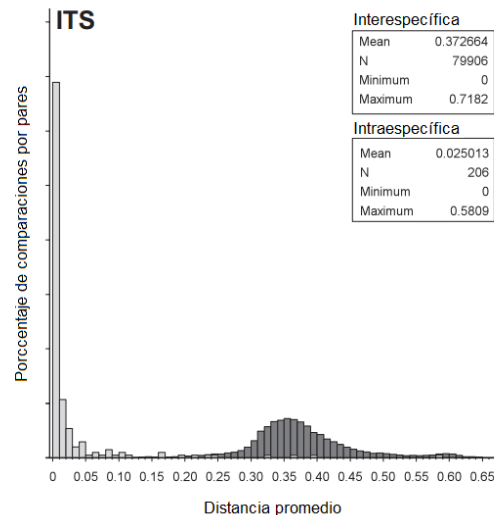


Figura 1. Histograma del análisis de brechas de códigos de barras para el marcador ITS. Frecuencia de valores promedio para las variaciones intraespecífica e interespecífica entre 17 linajes fúngicos. Gris claro - variación intraespecífica; gris oscuro - variación interespecífica. Tomado y modificado de Schoch *et al.*, 2012.

1.3.2.1 Región ITS

La región ITS es una unidad de repetición ribosomal nuclear compuesta de tres partes (ITS1, 5.8S que es una secuencia muy conservada e ITS2), con tamaño que oscila los 1100 pb (figura 2) (Bridge *et al*, 2005). Se transcribe junto con los genes que codifican las subunidades ribosomales, pero éstos no son incorporados en los ribosomas maduros, y son eliminados del ARNr precursor.

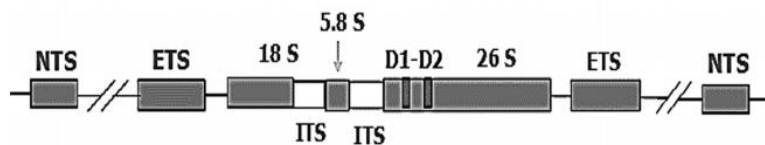


Figura 2. Estructura del espaciador transcrito interno (ITS).

Tomado de Kawahata *et al.*, 2007.

Las tecnologías de secuenciación de alto rendimiento más utilizadas tienen una limitación de longitud de lectura máxima y por tanto se usa sólo una de las dos subregiones (ITS1 o ITS2). Ambas se han evaluado como código de barras para hongos y levaduras (Monard *et al.*, 2013; Mbareche *et al.*, 2020), así como para otros organismos como insectos, peces y distintos tipos de plantas (Wang *et al.*, 2015).

Han observado que la región ITS1 tiene un mayor poder de discriminación y regiones flanqueantes más conservadas para la amplificación por PCR que ITS2 en varios grupos taxonómicos y, en general, representa un mejor código de barras de ADN que ITS2 en eucariontes (Wang *et al.*, 2015). Con respecto a los hongos y levaduras, la diversidad fúngica con ITS1 es mayor que la obtenida con ITS2, de acuerdo con lo reportado por Nilsson *et al.* (2008) y Monard *et al.* (2013). Además, permite identificar tanto a levaduras ascomicetas como basidiomicetas (Vásquez *et al.*, 2016).

Sin embargo, Badotti *et al.* (2017) mencionan que la región ITS no es tan variable para algunos géneros *Ascomycota* entre los cuales se encuentran *Alternaria*, *Aspergillus*, *Cladosporium*, *Penicillium* y *Fusarium*. Y se debe considerar que, en los estudios realizados hasta el momento, ni la región ITS1 ni ITS2 por sí solos son perfectos en términos de capturar la diversidad total de las muestras.

1.4 Análisis bioinformático de los datos

Con el desarrollo de mejores y más eficientes técnicas de secuenciación, e incremento en su aplicación en nuevas áreas, se ha impulsado el desarrollo de herramientas estadísticas y computacionales que permiten el análisis e interpretación de la enorme cantidad de datos generados.

Independientemente de la tecnología empleada, las secuencias de ADN son almacenadas en archivos de salida con distintos formatos, siendo *fastq* el más utilizado. Es un archivo de texto plano que contiene información de cada una de las secuencias con su identificador y la calidad asociada a cada nucleótido según su puntuación Phred (Q), la cual es una medida de la probabilidad de error e indica si el nucleótido ha sido llamado correctamente durante la secuenciación.

De los primeros pasos de análisis de los datos están la limpieza y análisis de calidad, mediante el cual se eliminan los adaptadores y las secuencias de ADN contaminante perteneciente a otros organismos o del hospedero. Así mismo, se pueden recortar aquellas bases con baja puntuación de calidad Phred, comúnmente aquellas con puntuación inferior a 20. En la leche de vaca es común encontrar células somáticas en baja concentración ($<1 \times 10^5$ células/mL) y su conteo es indicador de la salud de la glándula mamaria y de la calidad de la leche obtenida (Kochoski *et al.*, 2011; Pinto, 2013; PROY-NMX-F-700-COFOCALEC-2012). Dichas

células somáticas pueden pasar a los productos lácteos y ser una fuente de ADN contaminante en investigaciones que utilizan la secuenciación.

En los estudios realizados sobre el queso Cotija, de interés para el proyecto, se ha reportado que el paquete celular acarrea ADN de vaca y Escobar (2016) lo detectó por PCR en tiempo real a partir de tres paquetes celulares provenientes de extracciones independientes de muestras del queso. Por lo cual, aun cuando se espera la presencia de ADN bovino en la extracción del material genético, este se considera material contaminante al no ser de interés para el trabajo y se elimina de los datos de secuenciación.

Después del preprocesamiento de las lecturas, el análisis bioinformático puede tomar dos rutas diferentes dependiendo del tipo de datos de secuenciación (*shotgun* o *metaprofiling*) así como del enfoque del proyecto (figura 3). En ambos casos, se puede extraer información taxonómica y conocer la diversidad microbiana; sin embargo, con datos metagenómicos de *shotgun* además se puede conocer el potencial metabólico a través de la predicción y anotación de genes o secuencias codificantes. Cabe destacar que el flujo de análisis propuesto es general y, se pueden omitir y/o añadirse otros pasos según el objetivo de la investigación.



Figura 3. Análisis bioinformático general de datos metagenómicos
Adaptado de Oulas *et al.*, 2015 y Cameron *et al.*, 2021

La asignación taxonómica a partir de las secuencias derivadas de *metaprofiling* implica un paso de agrupación o *clustering* de las lecturas en unidades taxonómicas

operacionales (OTUs, por sus siglas en inglés) o en variantes de secuencia exacta o de amplicones (ASVs, por sus siglas en inglés).

Las OTUs se definen como un grupo de individuos estrechamente relacionados que se agrupan en función de su similitud de secuencia, a menudo establecido en 97% para inferir la identidad a nivel de género. Esta agrupación reduce el tamaño del conjunto de los datos sin procesar, así como disminuye la capacidad computacional para el análisis y reduce el impacto de los errores de secuenciación en las estimaciones de diversidad (Kozich *et al.*, 2013; Oulas *et al.*, 2015; Siegwald *et al.*, 2017). Mientras que la generación de ASVs, es un método de reducción de ruido en el que se genera un modelo de error basado en la calidad de la ejecución de secuenciación y se utiliza este modelo para distinguir la variación biológica “verdadera” y la probablemente generada por el error de secuenciación. Y es entonces que las secuencias “verdaderas” que varían desde un solo nucleótido las que se definen como ASV separadas, permitiendo una mayor precisión que el 97% de identidad en las OTUs (Glassman y Martiny, 2018; Nearing *et al.*, 2018, Chiarello *et al.*, 2022). El programa más utilizado para inferir las ASVs es dada2 (Callahan *et al.*, 2016), el cual convierte los archivos *fastq* pareados en secuencias de muestra inferidas, sin ruido y sin quimeras.

Posteriormente, de cada OTU o ASV se extrae una secuencia representativa y se alinea con cada una de las secuencias de una base de datos de referencia. La secuencia representativa y la OTU/ASV a la que pertenece se asignan finalmente a un grupo taxonómico inspeccionando las mejores alineaciones (Siegwald *et al.*, 2017). Algunas bases de datos comúnmente utilizadas para los hongos y levaduras se enlistan en la tabla 2, de las cuales UNITE es la más empleada.

Tabla 2. Bases de datos para el grupo fúngico y levaduriforme.

Base de datos	Región	Versión	Fecha actualización	Total de secuencias
UNITE	ITS completa	9.0	2022	6 441 764
BOLD		4.0	2023	236 100
ITS Refseq (NCBI)		---	2023	15 725
ITSoneBD	ITS1	1.141	2020	1 218 745
ITS2	ITS2	3.013	2011	288 044

Con respecto a los datos metagenómicos de *shotgun*, la agrupación de lecturas y su asignación taxonómica se puede llevar a cabo por dos estrategias dependiendo de la información utilizada, la composición o la similitud u homología.

El método según la composición se basa en la observación de que los genomas individuales de referencia tienen una distribución única de secuencias k-meros, subcadenas cortas de coincidencia exacta de longitud fija k, también denotadas como firmas genómicas. Al hacer uso de estas firmas genómicas se agrupan las secuencias en sus respectivos genomas y se compara su frecuencia en las lecturas metagenómicas. Un ejemplo de programa que emplea este enfoque para realizar la asignación taxonómica es Kaiju y Kraken2.

Por otro lado, el método basado en similitud implica la extracción de secuencias de genes marcadores ya sea mediante uso de algoritmos de alineación como BLAST o su predicción con Modelos Ocultos de Markov (HMM, por sus siglas en inglés) para obtener información de similitud sobre secuencias o genes específicos de bases de datos disponibles públicamente. A partir de la cual las secuencias se agrupan en OTUS o ASVs y se asigna la taxonomía por su similitud con el mejor hit de la referencia o por identificación de su último ancestro común (LCA por sus siglas en inglés) [Oulas *et al.*, 2015; Dröge y McHardy, 2012].

Continuando con el análisis de datos de *metaprofiling*, antes de analizar la diversidad es necesario estandarizar el tamaño de los datos de secuenciación o de la biblioteca mediante un submuestreo aleatorio que recibe el nombre de rarefacción. Esto permite la comparación entre muestras sin tener un sesgo por las diferencias en el tamaño de los datos (número de lecturas).

El tamaño adecuado de la biblioteca estandarizada se determina mediante las curvas de acumulación de especies o también llamadas curvas de rarefacción (figura 4). Se selecciona un tamaño de biblioteca que abarque la parte de aplanamiento de la curva para cada muestra y se asume que capturará adecuadamente la diversidad a pesar de la exclusión de algunas lecturas durante el proceso de rarefacción (Cameron *et al.*, 2021). Además, permite conocer si el esfuerzo de muestreo y secuenciación fueron lo suficientemente altos para capturar la diversidad de la muestra al observar que la curva alcanza el aplanamiento o una asíntota horizontal.

Finalmente, se realiza el análisis de la diversidad. La diversidad alfa se define como la diversidad de especies en diferentes sitios o hábitats dentro de una escala local. Este término fue acuñado por Robert Harding Whittaker junto con otros conceptos relacionados como la diversidad beta y gamma. Este tipo de diversidad considera la riqueza (número de especies por muestra) y la uniformidad de distribución o la abundancia de las especies presentes. Hay distintas métricas o índices para su determinación, entre los que se encuentra diversidad observada, Chao1, Shannon-Waever y Gini-Simpson.

La diversidad observada es la medida más simple de la diversidad alfa y corresponde al recuento del número de especies (OTUs o ASVs) presentes en la muestra o comunidad. Mientras que el índice de Chao1 es un estimador basado en la abundancia; por lo tanto, requiere datos de abundancia y, considera el número de singletons (especies observadas una vez) y de dobletones (especies observadas dos veces) en la muestra. Por su parte, el índice de Shannon-Weaver, también denominado Shannon-Wiener o entropía de Shannon, es uno de los índices de diversidad populares en los estudios ecológicos. La idea de esta métrica es que cuantas más especies diferentes estén presentes, y cuanto más similares sean sus abundancias en la muestra, más difícil será predecir qué especie será la siguiente en “tomar”. En otras palabras, cuantifica la incertidumbre en la predicción de la identidad de especie que se toma al azar del conjunto de datos y así, al tener una mayor diversidad, se tiene una mayor incertidumbre en elegir de manera aleatoria a una especie en específico. Por último, el índice de Gini-Simpson fue introducido por el estadístico Edward H. Simpson en 1949, pero ya había sido introducido parcialmente en 1945 por el economista Albert O. Hirschman. Este índice mide el grado de concentración cuando los individuos se clasifican en tipos, y tiene en cuenta el número de especies presentes, incluyendo su abundancia relativa en una muestra. Lo anterior puede interpretarse como la probabilidad de que dos individuos cualesquiera de la muestra, elegidos al azar, pertenezcan a la misma especie (Chao y Chiu, 2016; Thukral, 2017; Fedor y Zvaríková, 2019).

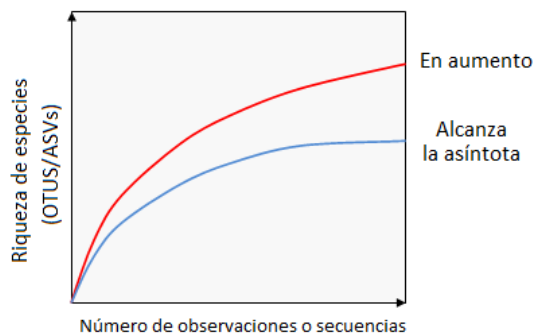


Figura 4. Curva de rarefacción
Tomado y modificado de Edgar, s.f.

Para analizar la diversidad entre muestras, o el grado en que éstas son diferentes, se utiliza la diversidad beta. Se puede estimar a partir de distintas métricas como disimilitud de Bray-Curtis, índice Jaccard o UniFrac ponderado y no ponderado. El primero considera las abundancias de las secuencias mientras que el índice de Jaccard se basa en la presencia/ausencia de organismos. Las métricas de UniFrac se basan en la relación filogenética o distancias, el no ponderado mide la longitud relativa de aquellas ramas que conducen exclusivamente a especies presentes en una sola de las dos muestras con respecto a la longitud total de todas las ramas del

árbol mientras que el ponderado además considera las abundancias relativas (Goodrich *et al.*, 2014; Calle, 2019; Cameron *et al.*, 2021).

2 Antecedentes

2.1 Queso Cotija

El queso Cotija es un producto lácteo mexicano genuino originario de la región de JalMich, ubicada entre los estados de Jalisco y Michoacán (figura 5). Actualmente, es elaborado de manera artesanal en la misma región bajo la Marca Colectiva de “Queso Cotija Región de Origen” y su producción se limita a los meses de julio a octubre para comercializarlo a partir de diciembre.

Se produce con leche entera de vaca o ligeramente descremada y sin pasteurizar. Se clasifica como un queso madurado, de pasta dura y de gran formato con un peso que oscila entre los 20 y 30 kg. Posee una pasta dura friable o desmoronable, ácida y con un porcentaje elevado de sal (aproximadamente 6%) que cuando está bien madurado presenta color blanco amarillento agradable, aroma pronunciado y sabor fuerte (Villegas, 2004).



Figura 5. Región de origen del queso Cotija (Tomado de Álvarez *et al.*, 2005).

La zona productora, definida por las reglas de uso de la marca colectiva, comprende seis municipios: Quitupan, Santa María del Oro y Jilotlán, ubicados en el estado de Jalisco, y Cotija, Tocumbo y Buena Vista, en Michoacán (figura 5). La región es una ladera templada, de transición climática, muy plegada y con escalonamiento altitudinal, que va desde los pies de monte que se elevan a partir del Valle de Tierra Caliente, formando a partir de allí una figura de herradura (con el arco orientado al norte) que llega hasta un poco antes de las cumbres frías del eje volcánico. La vegetación es selva baja caducifolia, con vegetación secundaria irregular, limitada con bosque mixto de pino-encino en el perímetro externo de esa forma de herradura que tiene la zona. Dicha región se ubica entre los 700 y 1 700 msnm,

aproximadamente; con una temperatura media anual de entre 25 °C al sur, y los 20 °C en los otros tres puntos cardinales que forman el arco de la herradura, y se registra una precipitación pluvial media anual de entre 900 mm a 1200 mm.

Las vacas productoras de la leche destinada a la elaboración del queso son de raza criolla o cebú de doble propósito (Pardo Suizo-Cebú, Holstein-Cebú) y son alimentadas por libre pastoreo en la región. Los factores como el clima, altura y suelo son los que originan la vegetación típica del lugar descrita anteriormente, misma que se refleja en la composición de la leche producida. Mientras que la humedad relativa específica, relacionada a la temperatura, las lluvias y la altura, se relacionan directamente con las características particulares del queso Cotija elaborado y añejado en la región (Álvarez *et al.*, 2005).

2.1.1 Proceso de elaboración

El queso Cotija se obtiene a partir de leche cruda de vaca y sal de grano de la laguna de Cuyutlán, Colima, sin adición de cultivo iniciador, ni tratamiento térmico, por lo que el proceso de maduración es llevado a cabo por la microbiota incorporada de forma natural o no intencional durante su proceso de manufactura.

El proceso de elaboración del auténtico queso Cotija consta de 9 etapas principales (figura 6) que se describen a continuación (García, 2001; Álvarez *et al.*, 2005; Torres, 2016):

1.- Ordeña: En esta etapa se obtiene la leche fresca, la cual debe ser filtrada a través de cedazos limpios para remover cualquier materia extraña antes de su utilización. La leche obtenida se deja reposar durante 2 a 4 h a temperatura ambiente en un recipiente de acero inoxidable limpio.

2.- Cuajado: Una vez que la leche alcanza la temperatura óptima (34 °C), se incorpora el cuajo. Se mezcla rápidamente y se deja cuajar por una hora aproximadamente. El cuajo utilizado debe provenir del extracto bovino del estómago de rumiantes, el producto comercial es Cuamix M.R., Cuajo líquido estandarizado de CHR Hansen®.

3.- Cortado: Una vez que la cuajada alcanza el punto deseado, se corta hasta obtener grumos equivalentes en tamaño a un grano de maíz. Se deja reposar nuevamente hasta que la cuajada se deposite en el fondo.

4.- Drenado: Se elimina el suero mediante decantación y colado de la cuajada en un trapo limpio estéril. La cuajada se comprime con ayuda del trapo para contribuir al desuerado y se deja escurrir en las mesas de trabajo.

5.- Amasado y salado: La cuajada se amasa de forma manual y al mismo tiempo se incorporan aproximadamente 140 g de sal por cada 20 litros de leche empleada (la cantidad de sal añadida varía según el productor).

6.- Moldeado: La cuajada salada se deposita en un molde de acero inoxidable cubierto con dos piezas de yute o ixtle (fibra de maguey). De esta forma se logran piezas cilíndricas de 40 cm de diámetro y 18 cm de altura con un peso aproximado de 20 kg.

7.- Prensado: La pasta se prensa de 18 a 24 h empleando piedras de 50 a 90 kg o prensas rústicas de tornillo. Se voltea sobre su propia cara al concluir el periodo.

8.- Oreado: La pieza se faja con un cilindro y se voltea diariamente durante mínimo 15 días o hasta que deje de escurrir suero para que adquiera la firmeza necesaria para ser manipulada.

9.- Maduración: El queso se desfaja al adquirir la firmeza adecuada y se continúa volteando para alternar la cara expuesta al medio ambiente durante los siguientes tres meses. Además, durante esta etapa es importante limpiar la superficie de la pieza con un lienzo limpio y suave, y se coloca una malla de tela limpia sobre el queso para evitar su contaminación. Se mantiene en un lugar seco, ventilado y cerrado para evitar la presencia de insectos.

De los pasos descritos anteriormente, los que son importantes para la obtención de un producto de calidad serían la ordeña, amasado y salado, oreado y maduración. Mientras que las etapas que posiblemente son punto de inoculación de microorganismos ya sea por el proceso, utensilios empleados o materia prima, serían ordeña, amasado y salado, prensado y maduración (Torres, 2016).

2.1.2 Microorganismos presentes en el queso

El queso Cotija es una matriz alimentaria compleja cuyas características sensoriales se ven influenciadas por varios factores, entre los que destacan el proceso y las condiciones de elaboración, la zona geográfica, las materias primas y su microbiota, siendo este último el que más impacto tiene sobre el alimento.

Los primeros estudios sobre este queso se enfocaron en su caracterización fisicoquímica a través de análisis proximales, determinación del a_w , pH y acidez. Esto con el objetivo de establecer algunos parámetros para tener una mayor estandarización del producto lácteo y que ayudaron a la obtención de la marca colectiva Región de Origen.

Con respecto a la caracterización microbiológica del alimento, ha sido objeto de estudio desde distintos enfoques demostrando que es abundante, diversa y

compleja al aislarse varios grupos microbianos como bacterias LAB, bacterias halófilas, bacterias con actividad proteolítica y lipolítica, y algunas levaduras.

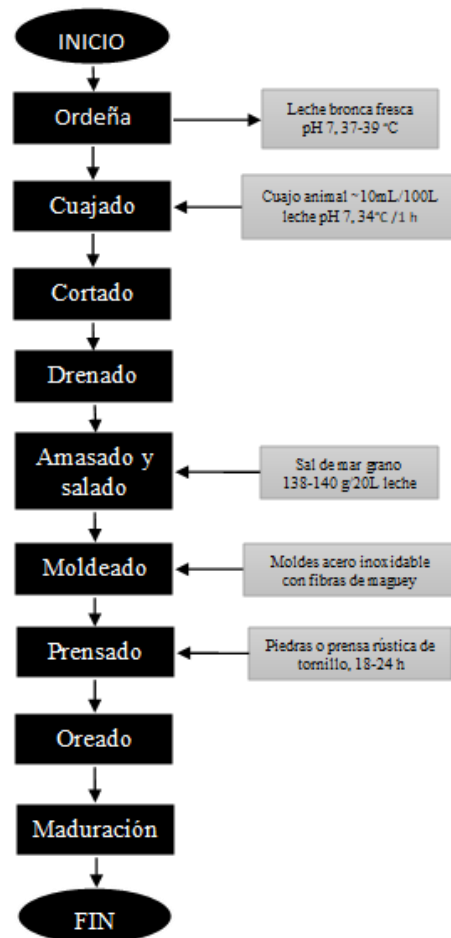


Figura 6. Proceso de elaboración del auténtico queso Cotija Región de Origen

Al menos se han aislado 15 géneros bacterianos, por ejemplo, Hernández (2007) logró aislar especies de los géneros *Bacillus*, *Staphylococcus* y *Enterococcus* empleando métodos bioquímicos tradicionales y ARDRA. También se llevaron a cabo estudios para la detección de diversos microorganismos patógenos en el producto lácteo y, con ellos se demostró que es un alimento inocuo y su consumo no representa un riesgo para la salud al no identificarse en ninguna de las muestras analizadas a *Listeria monocytogenes*, *Escherichia coli* ETEC O EHEC, *Brucella spp.*, *Staphylococcus aureus*, miembros del complejo *Mycobacterium* ni *Salmonella enterica*.

Por otra parte, García (2006) además de identificar bacterias, realizó la cuantificación en placa de hongos y levaduras (tabla 3), lo cual derivó en el

aislamiento e identificación de dos levaduras, *Y. lipolytica* y *C. zeylanoides*, a través del análisis y secuenciación del gen 18S del ADNr.

Tabla 3. Resultados de cuenta en placa para hongos y levaduras.

Muestra	García, 2006		Martínez, 2009	
	Levaduras (UFC/g)	Hongos (UFC/g)	Levaduras (UFC/g)	Hongos (UFC/log)
1	1.3x10 ⁸	<10	11.13x10 ²	<67
2	6.4x10 ⁶	<10	13.33x10 ²	<67
3	5.6x10 ⁵	<10	3.0x10 ²	<67
4	2.7x10 ⁷	4.6x10 ⁴	6.67x10 ³	<67
5	3.5x10 ²	<10	2.50x10 ⁴	<67
6	<10	<10	2.28x10 ⁴	<67
7	----	----	60.80x10 ³	<67
8	----	----	6.67x10 ²	<67

Mientras que Martínez (2009) estudió la población levaduriforme a través de técnicas dependientes e independientes de cultivo. Analizó ocho muestras de queso provenientes de varios municipios pertenecientes al colectivo de la Región de Origen. Identificó a las levaduras *Candida parapsilosis*, *C. zeylanoides* y *K. lactis*, de las cuales las dos últimas se encontraron en la mayoría de las muestras analizadas.

Se han realizado algunos estudios metagenómicos en el alimento que además de identificar las especies microbianas, se enfocaron en la búsqueda de genes codificantes para bacteriocinas. Estos análisis revelaron que la microbiota bacteriana consiste en aproximadamente 98% de *Firmicutes*, con *Lactobacillus*, *Leuconostoc* y *Weissella* como los géneros dominantes, y más de 500 géneros no dominantes agrupados en 31 *phyla* de bacterias y arqueas (Escobar, 2016).

Además, Escobar (2016) logró identificar una fracción de eucariontes mediante la anotación del gen ribosomal 18S a pesar de que su análisis se enfocó en la caracterización bacteriana (Anexo I). Las especies de levaduras anotadas se muestran en la figura 7 y tabla 4, mientras que las familias *Saccharomycetaceae* y *Debaryomycetaceae* son las más abundantes, que al formar parte de la microbiota contribuyen al perfil sensorial del alimento durante su proceso de maduración.

Considerando los estudios realizados sobre esta muestra y conociendo su importancia en el proceso de maduración, podemos inferir que el grupo de levaduras presentes en el queso también son parte importante de la microbiota y pueden estar relacionadas con el desarrollo de compuestos de sabor y aroma tan característicos de este producto.

Tabla 4. Anotación taxonómica para eucariontes del ADNr 18S de muestras de Queso Cotija de distintos municipios de la Región de Origen. (Tomado y modificado de Escobar, 2016)

	Género ^a	Especie ^a
Dominantes	<i>Debaryomyces</i> (35.20%)	<i>D. udonii</i> (14.15%)
		<i>D. hansenii</i> var. <i>hansenii</i> (10.00%)
		<i>D. hansenii</i> (7.75%)
		<i>D. sp.MPSR02</i> (3.25%)
	<i>Kazachstania</i> (16.65%)	<i>K. unispora</i> (16.50%)
Subdominantes	<i>Schwanniomyces</i> (9.80%)	<i>S. capriottii</i> (9.80%)
	<i>Naumovia</i> (7.25%)	<i>N. castellii</i> (7.25%)
	<i>Williopsis</i> (4.05%)	<i>W. salicorniae</i> (3.25%)
	<i>Candida</i> (3.50%)	<i>C. dubliniensis</i> (2.40%)
	<i>Dothidea</i> (2.80%)	<i>D. berberidis</i> (2.80%)
	<i>Teratosphaeria</i> (2.15%)	<i>T. molleriana</i> (2.15%)
	<i>Saccharomyces</i> (1.55%)	<i>S. cerevisiae</i> (0.90%)
	<i>Bagnisiella</i> (1.35%)	<i>B. examinans</i> (1.35%)
	<i>Kabatiella</i> (1.10%)	<i>K. microsticta</i> (1.10%)
	<i>Mycosphaerella</i> (1.10%)	<i>M. podagrariae</i> (1.10%)

^a En paréntesis la abundancia relativa con respecto a la abundancia total

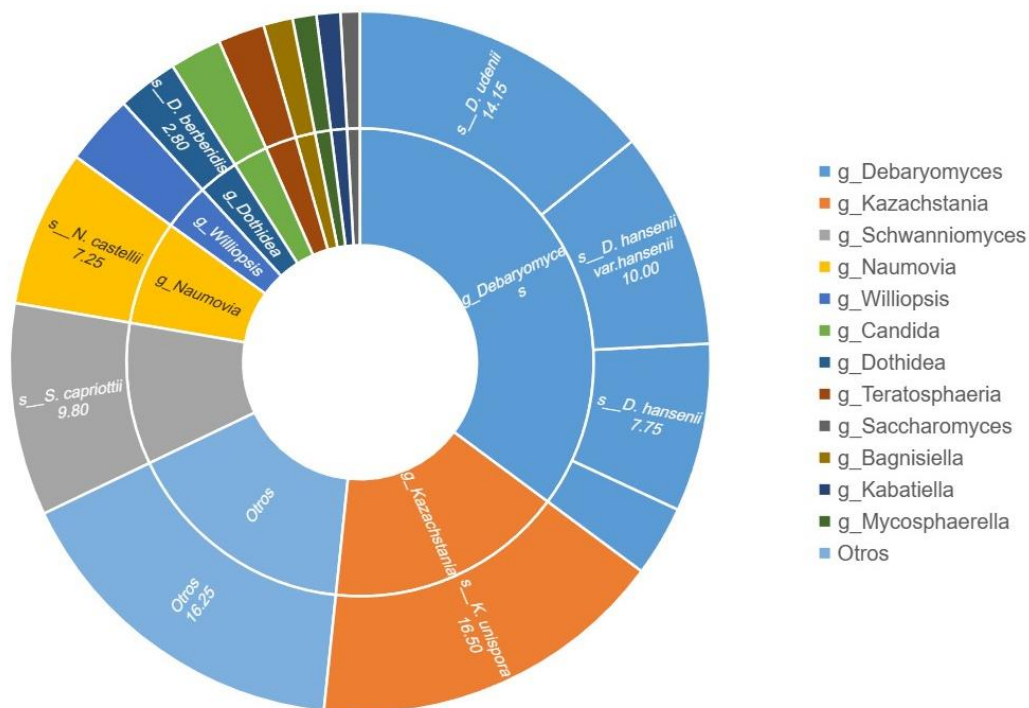


Figura 7. Abundancia relativa de las lecturas metagenómicas de *shotgun* de muestras del Queso Cotija. El número seguido a la especie corresponde a la abundancia relativa con respecto a la abundancia total en porcentaje (Tomado y modificado de Escobar, 2016).

3 Justificación

Los quesos madurados elaborados de forma artesanal son un reservorio de microorganismos que poseen actividades metabólicas interesantes desde el punto de vista alimentario y biotecnológico. Durante su proceso de maduración, la microbiota participa en el desarrollo de compuestos de sabor y aroma, generando un producto con características únicas. La población de levaduras presente posee actividades metabólicas particulares que están directamente relacionadas con dicho proceso. No obstante, su contribución a la maduración y al perfil sensorial generalmente se subestima. Por lo que, al analizar el metagenoma de levadura en el queso Cotija Región de Origen, permitirá caracterizar el grupo levaduriforme y complementar los estudios que se han realizado sobre este.

4 Hipótesis

Se espera que el perfil taxonómico de levaduras de tres muestras de Queso Cotija Región de Origen sea semejante y que las familias dominantes sean *Saccharomycetaceae* y *Debaryomycetaceae*.

5 Objetivos

5.1 Objetivo general

Identificar la diversidad de levaduras presentes en tres muestras de Queso Cotija Región de Origen mediante secuenciación dirigida a la región ITS1 y contrastar con el perfil taxonómico descrito según el análisis de datos de secuenciación de *shotgun*.

5.2 Objetivos particulares

- Identificar las levaduras presentes en tres muestras de queso Cotija mediante la amplificación y secuenciación de la región ITS1 del ADN metagenómico.
- Interrogar los datos metagenómicos de *shotgun*, obtenidos en un estudio previo, para la búsqueda de lecturas pertenecientes a levaduras y posibles secuencias de ITS.
- Analizar la diversidad de levaduras identificada por ambos enfoques.

6 Diagrama experimental

Para lograr los objetivos planteados, se propone la metodología ilustrada en la figura 8. Cada etapa se detalla en la siguiente sección.

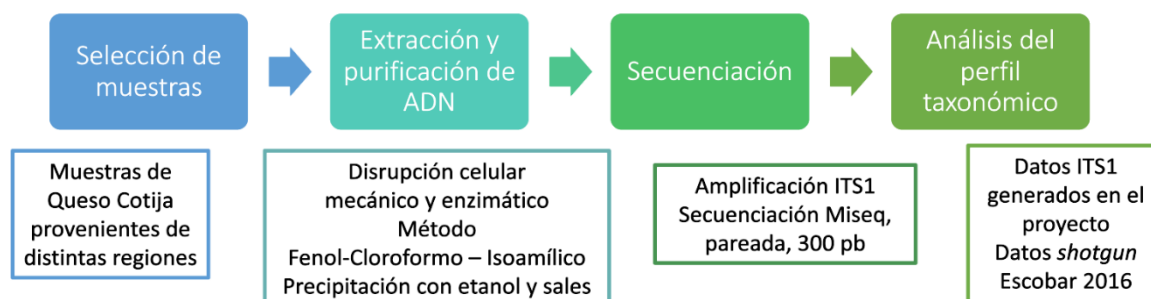


Figura 8. Diagrama de la metodología seguida

7 Metodología

7.1 Selección de muestras

Las muestras por analizar fueron seleccionadas con base en los resultados obtenidos previamente de cuenta en placa para distintos grupos microbianos realizado por Chombo-Morales (datos no publicados, tabla 5). Se seleccionaron tres muestras que presentaron mayor número de levaduras y que cumplieron con los parámetros fisicoquímicos y microbiológicos indicados por la norma NMX-F-735-COFOCALEC-2018 (tabla 6).

Tabla 5. Muestras a analizar

Muestra	Procedencia	Levaduras (log UFC/g)
QC-22	Cotija	5.02x10 ⁶
QC-57	Quitupán	2.88X10 ⁶
QC-95	Santa María del Oro	6.47X10 ⁶

7.2 Tratamiento de las muestras

Se realizó un muestreo de cada uno de los quesos, se retiró la corteza, se cortaron en trozos pequeños y se trituraron 500 g de cada uno con un procesador. Se almacenaron en una bolsa plástica de cierre hermético en congelación a -70°C hasta su uso.

7.3 Obtención del paquete celular

La metodología que a continuación se describe se ha empleado anteriormente en el grupo de trabajo. Sin embargo, se optimizó para obtener un ADN de mejor calidad e integridad, principalmente se adecuó la velocidad de centrifugado durante la obtención del paquete celular; la velocidad y tiempo de agitación con las perlas de

vidrio en la disrupción de la pared celular y se controló la temperatura durante la disrupción mecánica. Se describe a más detalle en los apartados siguientes.

Tabla 6. Parámetros establecidos por la norma NMX-F-735-COFOCALEC-2018

Parámetros fisicoquímicos	a_w		≤ 0.90
	Humedad		Máx. 36%
	Grasa (BH)		Min. 23%
	Proteína (BH)		Mín. 25%
Parámetros microbiológicos	Cuenta:		
	<i>Staphylococcus aureus</i>		< 100 UFC/g
	<i>Salmonella sp.</i>		Ausencia
	Bacterias mesófilas aerobias		10 ⁵ – 10 ⁶ UFC/g
	Coliformes totales		100 – 300 UFC/g
	Por qPCR:		Ausencia
	<i>Listeria monocytogenes</i>	<i>Salmonella enterica</i>	
	<i>Staphylococcus aureus</i>	<i>Escherichia coli</i>	
	<i>E. coli</i> ETEC	<i>E. coli</i> O157:H7	
	Complejo <i>Mycobacterium</i> (<i>Mycobacterium bovis</i>)		
	<i>Brucella spp.</i>		

Se descongelaron las muestras en refrigeración por al menos 12 h. Se pesaron 25 g de cada queso en bolsas para *Stomacher* (Seward) en la campana de flujo laminar y se adicionaron 40 mL de buffer de fosfatos 50 mM pH 7 previamente estéril. La mezcla se homogeneizó en un *Stomacher 400 Circulator* (Seward) por 2 min a velocidad alta para romper mecánicamente la matriz del queso. Se incubó 5 min a 37°C y se adicionó 1 mL de solución de proteasas de *Bacillus amyloliquefaciens* (Sigma) con el fin de romper la red proteínica y permitir la separación de las células microbianas de los restos del alimento, y se mezcló a alta velocidad en el *Stomacher* durante 5 min. Posteriormente, la bolsa se incubó a 45°C durante 1 h en agitación a 100 rpm y al término el contenido se dividió en dos tubos Falcon estériles de 50 mL. Se centrifugó a 121 x g (1 000 rpm) durante 10 min a 4°C para separar los restos del queso y la grasa de las células de los microorganismos que se encontraban suspendidas en la fase acuosa.

La fase intermedia se transfirió a tubos de fondo redondo de 50 mL estériles y se centrifugó a 20 817 x g por 10 min a 4°C. El sobrenadante se desechó y el pellet se lavó 3 veces con solución salina al 0.85% pH 7 estéril. El pellet se resuspendió en 1 mL de solución salina y se colocó en tubo Eppendorf de 2 mL estéril, se centrifugó a 20 817 x g por 10 min a 4°C. El sobrenadante se eliminó y cada paquete celular se resuspendió en 300 µL de buffer TE pH 8.

7.4 Extracción y purificación del ADN metagenómico

Para la extracción y purificación del ADN, a cada uno de los tubos con los paquetes celulares resuspendidos en buffer TE se les adicionaron 70 µL de una solución de

lítica 1 000 U/mL. Se agitaron las muestras en vortex y se incubaron a 37°C por 1 h. Posteriormente, se centrifugó a 20 817 x g para eliminar el sobrenadante. Se añadieron 500 µL buffer de lisis (Anexo II) y se adicionaron 200 µL de perlas de vidrio 0.5 mm (Sigma). Se agitó en vortex por 30 s a velocidad media y se colocaron en un baño de hielo por 2 min. Se agitó nuevamente en vortex por 30 s a velocidad media y se incubó a 65°C por 30 min. Al término de la incubación, se adicionaron 20 µL de una solución de RNasa A de páncreas bovino a 20 µg/mL y se incubó por 30 min a 37°C. Después se adicionaron 20 µL de una solución de proteinasa K 20 mg/mL y se incubó durante 30 min a 65°C con agitación a 600 rpm.

La mezcla se dejó enfriar hasta 37°C, se adicionaron 10 µL de una solución de SDS 10% y se incubó a 37°C por 40 min. Terminado este tiempo, se centrifugó a 20 817 x g por 10 min a 4°C y el sobrenadante se transfirió a un tubo Eppendorf de 2 mL estéril.

A continuación, se realizó la purificación del ADN. Se adicionaron 200 µL de una solución de NaCl 6M y se mezcló lentamente por inversión. Se adicionaron 400 µL de una solución de fenol-cloroformo-alcohol isoamílico (F:Cl:Al) 25:24:1 en la campana de extracción, se agitó en vortex por 30 s y se centrifugó a 20 817 x g por 10 min a 4°C. La mayoría de las proteínas se extraen con fenol debido a que este solvente permite su desnaturalización y concentración en la fase orgánica. En tanto que el ADN se recuperó en la fase acuosa y se transfirió a un tubo Eppendorf de 2 mL estéril. Se dio un lavado con un volumen de cloroformo para eliminar las proteínas residuales y el fenol, se agitó en vortex por 30 s y se centrifugó a 20 817 x g por 10 min a 4°C. Nuevamente la fase acuosa se recuperó y transfirió a un Eppendorf de 1.5 mL estéril, se adicionó 0.1 V de una solución de acetato de amonio 3M y se adicionó 1 mL de etanol 100% frío, se agitó por inversión de 2 a 5 veces. Se incubó a -20°C toda la noche. La mezcla se centrifugó a 20 817 x g por 10 min a 4°C y el pellet se lavó dos veces con 1 mL de una solución de etanol 75% frío centrifugando 10 min en las mismas condiciones.

Se desechó el etanol y se evaporó el remanente en Thermomixer durante 5 min a 37°C. El pellet de ADN se resuspendió en 30 µL de agua grado biología molecular y se incubó a 37°C por 15 min para su total hidratación. Se midió la concentración y pureza del ADN total por método espectrofotométrico (NanoDrop, ThermoScientific) y se evaluó la integridad con electroforesis en gel de agarosa al 1%. Los ADN totales extraídos se almacenaron en congelación a -20°C.

7.5 Secuenciación del ADN metagenómico

La construcción y secuenciación de la librería de amplicones de la región ITS1 se llevó a cabo en la Unidad Universitaria de Secuenciación Masiva y Bioinformática (UUSMB) del Instituto de Biotecnología, UNAM en Cuernavaca, Morelos, de acuerdo con el protocolo "*Fungal Metagenomic Sequencing Demonstrated Protocol*" (Illumina, 2021). De manera breve, se amplificó la región ITS1 utilizando 15 pares

de cebadores (Illumina) específicos para la región de interés (Anexo III) y se verificó el tamaño de los amplicones en un gel de agarosa al 1.5%. Se eliminaron los cebadores y sus dímeros de la reacción anterior con un volumen de Ampure XP Beads y con un lavado de etanol al 80%, y se resuspendió en un volumen de buffer EB. Posteriormente se llevó a cabo el marcaje de cada muestra mediante una segunda reacción de amplificación, para ello se tomó 5 µL de la primera reacción purificada y se utilizaron los cebadores Nextera XT v2 index 1 Primer y Nextera XT v2 index 2 Primer. Al término del marcaje, las reacciones se mezclaron en un solo tubo para purificarse nuevamente con Ampure XP Beads y eluirse en un volumen de buffer EB. La librería construida se normalizó a una concentración de 4 nM y se cargó un volumen en la celda para obtener aproximadamente 100 mil lecturas por muestra en la secuenciación final. Se realizó la secuenciación en un equipo MiSeq de Illumina, con configuración pareada y con una longitud de lectura promedio de 300 pb.

7.6 Análisis bioinformático de las secuencias

El análisis bioinformático de los datos de secuenciación se realizó con el apoyo y asesoría del Dr. Luis Fernando Lozano Aguirre de la Unidad de Análisis Bioinformáticos del Centro de Ciencias Genómicas de la UNAM y se utilizó el servidor Pangolín.

7.6.1 Calidad de las secuencias

Previo al tratamiento de los datos, se evaluó la calidad de las lecturas crudas obtenidas por secuenciación de la región ITS1 y por secuenciación *shotgun* (Escobar, 2016). Se empleó el programa *fastQC* v.0.11.9 de Babraham Bioinformatics, de la Universidad John Hopkins (disponible en <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). La media de calidad de las secuencias se esperaba por arriba de una puntuación Phred de 28.

7.6.2 Datos de secuenciación dirigida a la región ITS1

7.6.2.1 Anotación taxonómica

Para la asignación taxonómica de las lecturas de la región ITS1 se siguieron dos estrategias de análisis. La primera de ellas se basó en el *pipeline* propuesto por Bálint *et al.* (2014), que está enfocado en el análisis de lecturas obtenidas por la tecnología de Illumina, pero se hicieron algunas modificaciones. El primer paso consistió en la eliminación de adaptadores y secuencias cortas con el programa Cutadapt v3.5 indicando la secuencia del adaptador (Nextera transposasa, 5'-CTGTCTCTTATACACATCT-3'). Después se ensamblaron las lecturas pareadas con PANDAseq v2.11 (Masella *et al.*, 2012) seguida de la búsqueda y extracción de la región ITS1 con el programa ITSx v1.1.3 del *Department of Neuroscience and Physiology de University of Gothenburg*, se especificó el grupo de interés (fungi). Con este programa se pretendió descartar aquellas lecturas que no coincidieran con

el modelo estructural de la región ITS, así como las posibles regiones conservadas que se pudieron recuperar en la secuenciación y que derivarían en un sesgo al hacer la asignación taxonómica con BLAST, ya que podrían aumentar la similitud entre las secuencias (Nilsson *et al.*, 2010).

Posteriormente, empleando USEARCH se hizo la desreplicación y eliminación de secuencias únicas (singletons) y, por último, la clusterización en OTUs con un porcentaje de similitud de 97%. Finalmente, las OTUs generadas se asignaron con la herramienta BLAST v11.0.6 empleando la base de datos de UNITE+INSD v9.0 (Abarenkov *et al.*, 2022a).

Para la segunda estrategia se utilizó el programa dada2 (Callahan *et al.*, 2016) y el flujo de análisis propuesto por los autores para la región ITS (disponible en https://benjjneb.github.io/dada2/ITS_workflow.html). El filtrado y recorte de las lecturas se realizó con los parámetros establecidos en el programa (filtrado de lecturas con nucleótidos ambiguos, MaxN=0; filtrado de lecturas en función de errores esperados, maxEE=2; filtrado de lecturas con longitud menor de la indicada, truncLen=50; filtrado de secuencias en las que hay al menos una base con la alta probabilidad de llamada errónea; truncQ=2) y se eliminaron aquellas lecturas que pertenecieran al control positivo utilizado en la secuenciación (fago PhiX). Para la asignación taxonómica el programa emplea el clasificador Bayesiano ingenuo (*the naive Bayesian classifier*), que es un algoritmo de aprendizaje supervisado que se basa en la probabilidad de que una secuencia de consulta desconocida sea miembro de un nivel taxonómico en particular. Se utilizó la base de datos general de UNITE en formato fasta que incluye secuencias singletons (globales y 97% similitud) v9.0 (Abarenkov *et al.*, 2022b).

7.6.2.2 Análisis de diversidad

A partir de la tabla de ASVs obtenida durante la anotación, se generó la curva de rarefacción (parámetros empleados: establecer una semilla y obtener resultados repetibles, set.seed=2021; lecturas submuestreadas o profundidad, sample.size=43320; iteraciones 10) y se calcularon los índices o estimadores de la diversidad alfa con los paquetes *phyloseq* v1.42.0 (McMurdie y Holmes, 2013) y *microbiome* v1.20.0 (Lahti, 2017) disponibles por *Bioconductor* en R.

7.6.3 Datos de secuenciación *shotgun*

Debido a que parte del proyecto se realizó aun estando en pandemia por COVID-19 (2021 - 2022) y se implementaron restricciones en el trabajo experimental en los laboratorios, se decidió complementar el estudio analizando los datos metagenómicos de *shotgun* obtenidos en un estudio previo con el fin de identificar posibles lecturas pertenecientes a levaduras. Esto gracias a que, con dichos datos, Escobar (2016) anotó una fracción de eucariontes mediante análisis del gen ribosomal 18S usando el programa Parallel-meta y la base de datos de SILVA. Para ello, los archivos (proyecto PRJNA286900, disponible en

<https://www.ebi.ac.uk/ena/browser/view/PRJNA286900>) fueron descargados del European Nucleotide Archive (ENA) del Instituto Europeo de Bioinformática (EMBL-EBI),

7.6.3.1 Limpieza de adaptadores y secuencias contaminantes

El programa FastQC v0.11.9 se utilizó para visualizar las calidades de las secuencias antes y después del paso de limpieza. Para la limpieza se empleó el programa Trim Galore v06.5 (disponible en https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/), con el cual se eliminaron los adaptadores, las bases de baja calidad del extremo 3' y se recortaron las primeras bases del extremo 5'. Después, las lecturas fueron mapeadas contra el genoma de *Bos taurus* (ARS-LIC_NZ_Hostein-Friesian_1, NCBI) con bowtie2 v2.4.5 (disponible en <http://bowtie-bio.sourceforge.net/bowtie2/index.shtml>) con el fin de eliminar las lecturas contaminantes del ADN de vaca acarreado durante el proceso de extracción.

7.6.3.2 Anotación taxonómica

La asignación taxonómica de las lecturas obtenidas después de la limpieza del ADN contaminante se llevó a cabo mediante dos programas. El primero de ellos fue Kaiju v1.8.2 que realiza la clasificación taxonómica por comparación de secuencias de aminoácidos contra una base de datos de referencia de proteínas. Se corrió bajo los parámetros establecidos y se utilizó la base de datos nr_euk (disponible en <https://kaiju.binf.ku.dk/server>; 24-02-2021, 52 GB) que comprende tanto a bacterias, arqueas, virus, hongos, levaduras, y otros eucariontes.

El segundo programa utilizado fue Kraken2 v2.0.8, también se corrió bajo los parámetros establecidos y, se descargó y utilizó la base de datos PlusPF (disponible en <https://benlangmead.github.io/aws-indexes/k2>; 5/17/2021, 41 GB) elaborada por Ben Langmead de la Universidad John Hopkins, abarca los grupos siguientes: bacteria, arquea, virus, humano, protozoarios, hongos y levaduras.

7.6.3.3 Extracción de lecturas de ITS e identificación

Ya que el objetivo principal del proyecto es identificar las levaduras del queso Cotija mediante análisis del marcador ITS, se buscaron y extrajeron las lecturas correspondientes a dicha región a partir de los datos metagenómicos de *shotgun* con el programa ITSx v1.1.3 (Bengtsson-Palme *et al.*, 2013), el cual permite extraer la región completa ITS o las subregiones ITS1 e ITS2. La extracción de las secuencias se basa en las posiciones predichas de los genes ribosomales (LSU/28S y SSU/18S) en las secuencias mediante los modelos ocultos de Markov (HMM, por sus siglas en inglés) calculados a partir de alineaciones de un total de 20 grupos de eucariontes que incluye hongos, levaduras, metazoos y plantas.

Debido a que las lecturas tenían una longitud corta (20-95 pb) se realizó el ensamblaje en secuencias largas o también denominadas *contigs* con el programa SPAdes

v3.15.4 (Bankevich *et al.*, 2012; Prjibelski *et al.*, 2020). Es un ensamblador iterativo *de novo* de secuencias cortas que se basa en la construcción de gráficas de Bruijn a partir de k-meros con distintas longitudes k (Bankevich *et al.*, 2012). El programa se ejecutó bajo los parámetros establecidos, se indicó que eran datos metagenómicos (--meta) y se probaron los k-meros 21, 33, 55 y 77 pb. Al no aceptar lecturas con longitudes mayores a 100 kbp el programa ITSx, se fragmentaron los contigs con un script interno en Perl. Finalmente, se ejecutó el programa ITSx con los parámetros establecidos.

A partir de los contigs identificados como secuencias de ITS, se hizo un alineamiento contra la base de datos de GenBank usando el programa BLAST para su identificación.

8 Resultados y Discusión

8.1 Extracción y purificación de ADN

En total se realizaron 18 extracciones independientes de cada muestra de queso Cotija, para posteriormente mezclar los ADN de una misma muestra según su calidad e integridad. Después, se seleccionó los ADN que presentaron mejor relación A260/280 e integridad de cada muestra y fueron secuenciados por la UUSMB del Instituto de Biotecnología UNAM (tabla 7). En la figura 9 se muestran las bandas definidas del ADN secuenciado, así como los perfiles de absorbancia en los cuales se ilustra un pico máximo en 280 nm indicando que la calidad era adecuada.

Tabla 7. Cantidades totales de ADN extraído de tres quesos Cotija y las relaciones de absorbancias

Muestra	Cantidad total (μg)	A260/A280
QC-22	6.12	1.79
QC-57	16.94	1.76
QC-95	3.31	1.77

8.2 Análisis bioinformático de las secuencias

8.2.1 Calidad de las secuencias

Para verificar la calidad de la secuenciación, tanto de las lecturas de la región ITS1 como de la secuenciación *shotgun*, se corrió el programa fastQC. Este programa despliega distintos gráficos que ilustran varios aspectos de la calidad. En la figura 9, se muestran los gráficos de calidad de acuerdo con la puntuación Phred para cada posición de las lecturas pareadas. La media de calidad (línea azul) se espera que se encuentre por arriba de una puntuación Phred de 28 y que el propio programa indica en color verde.

Para las lecturas de la región ITS1 (figura 10A – C) se observa que mantienen una buena calidad a pesar de que en las últimas posiciones hay un decaimiento. Este decaimiento se ha reportado como un efecto de la secuenciación puesto que la exposición del ADN al láser ocasiona su degradación. Así mismo, el proceso de síntesis durante la secuenciación no es homogéneo derivando en una desincronización y acumulación en la tasa de error (Fuller *et al.*, 2009).

Por otro lado, la calidad de las lecturas derivadas de la secuenciación de *shotgun*, presentaron una calidad aceptable a pesar de que las lecturas directas en sus primeras bases tienen un decaimiento atípico de su calidad, la puntuación Phred media se encontró entre 10 y 28 (figura 10G y H).

En cuanto a los gráficos de contenido de adaptadores (figura 11), el programa solo detecta secuencias adaptadoras en las lecturas de la región ITS1 y que

corresponden al adaptador Nextera transposasa utilizado en la construcción de la librería, que tiene una longitud de 13 pb (5'-CTGTCTCTTATACACATCT-3').

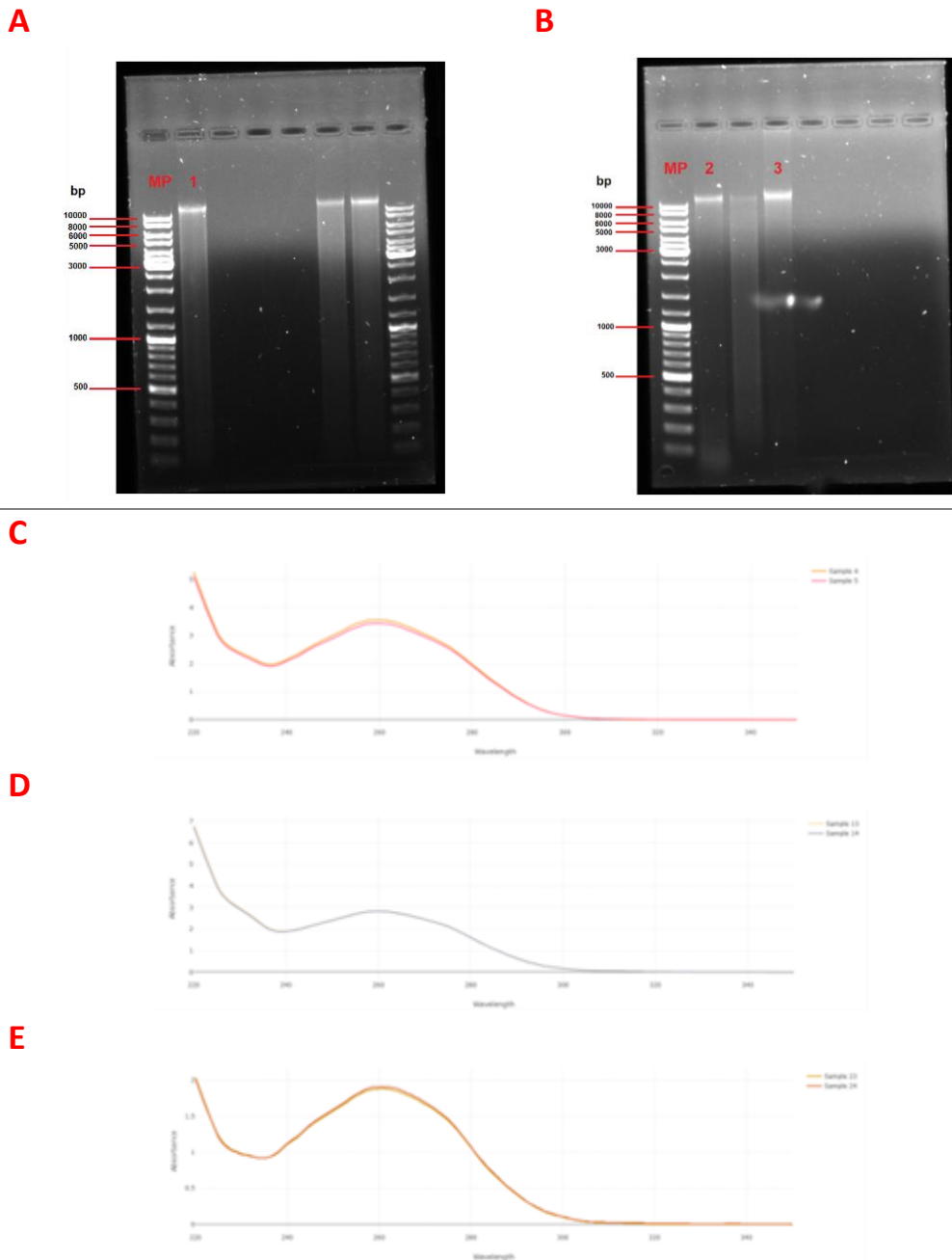


Figura 9. ADN extraído de tres muestras de queso Cotija. **A y B** - Geles de agarosa 1% teñido con bromuro de etidio. **Gel A:** MP: marcador de peso molecular *GeneRuler DNA Mix*; 1: QC-22; **Gel B:** MP: marcador de peso molecular *GeneRuler DNA Mix*; 2: QC-57; 3: QC-95. **C-E** – Espectros de absorbancia de las muestras que se mandaron a secuenciar. C, QC-22; D, QC-57 y E, QC-95.

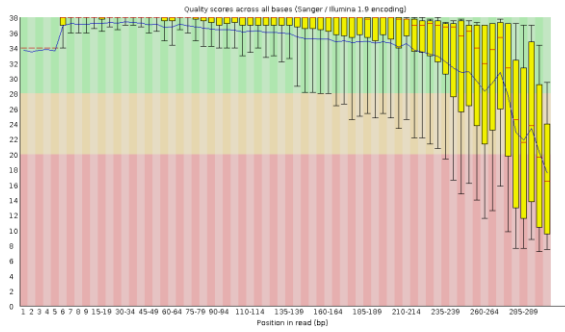
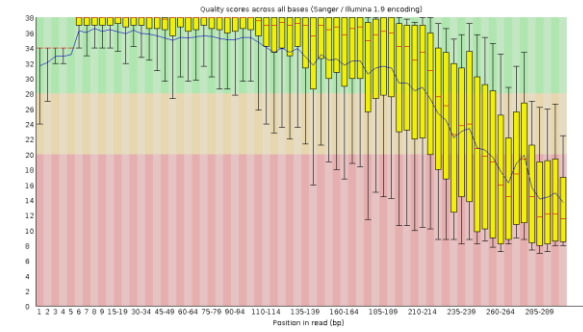
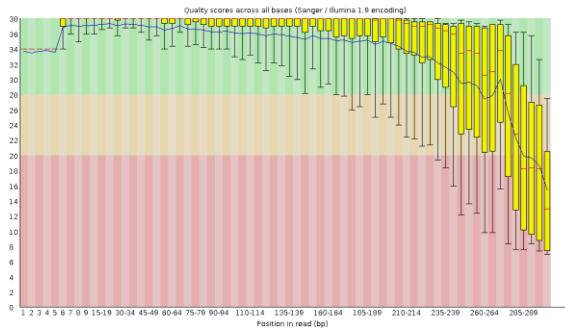
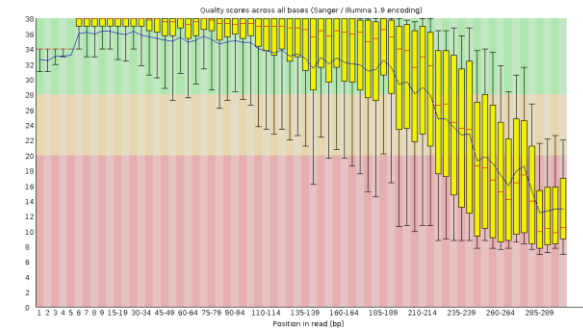
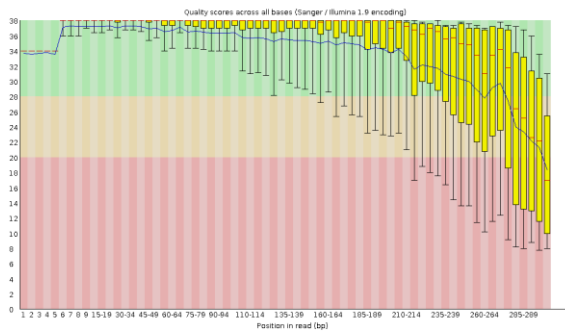
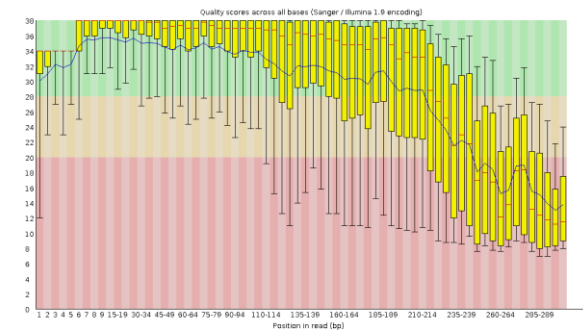
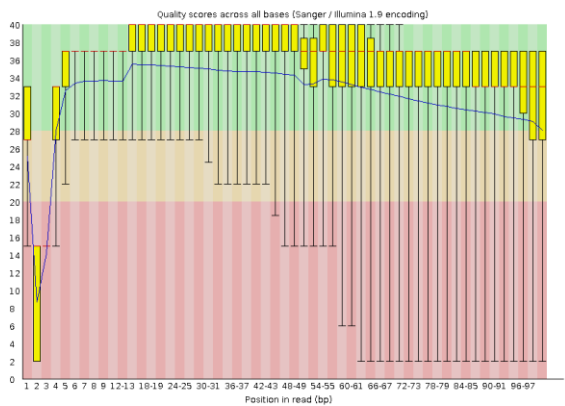
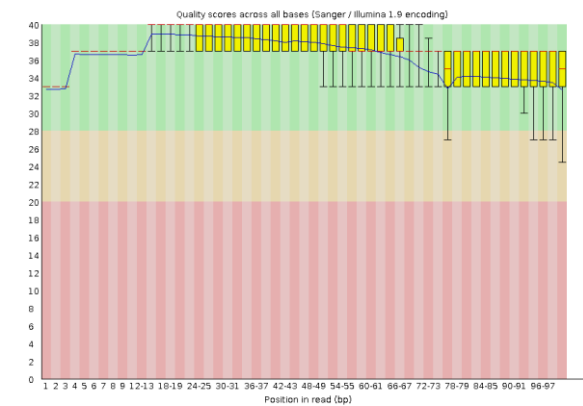
A1**A2****B1****B2****C1****C2****D1****D2**

Figura 10. Gráficos de caja para las lecturas obtenidas por secuenciación en sentido 5' (1) y en sentido 3' (2). A – C, secuenciación de la región ITS1; D, secuenciación *shotgun*. A, QC-22; B, QC-57, y C, QC-95.

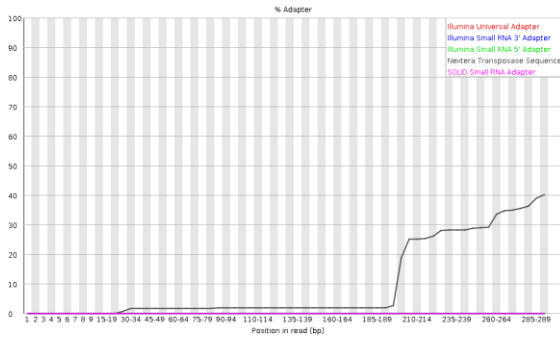
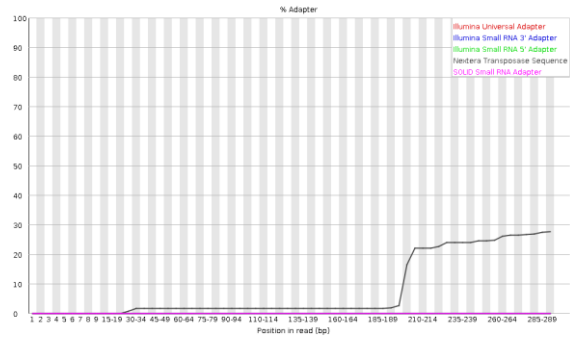
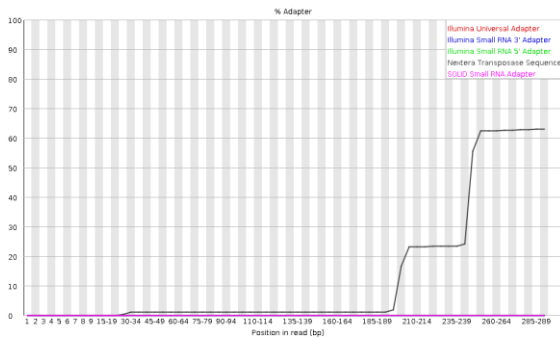
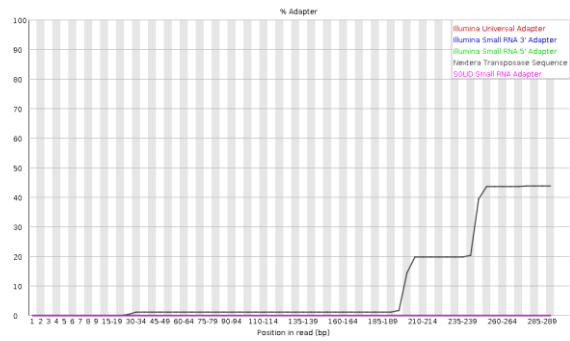
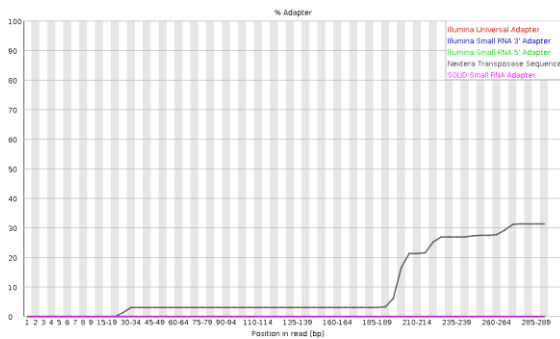
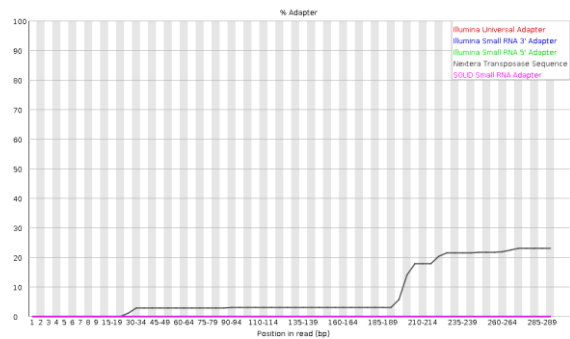
A1**A2****B1****B2****C1****C2**

Figura 11. Gráficos de contenido de adaptadores de los datos de secuenciación de la región ITS1 en sentido 5' (1) y en sentido 3' (2). A, QC-22; B, QC-57, y C, QC-95.

8.2.2 Secuenciación dirigida a la región ITS1

8.2.2.1 Análisis taxonómico

Como resultado de la secuenciación de la región ITS1 se obtuvieron alrededor de 237 mil a 369 mil lecturas para cada muestra (tabla 8) con una longitud promedio de 301 pb.

Con el análisis de los datos según lo propuesto por Bálint *et al.* (2014) basado en OTUs, se encontró que entre el 60% y 71% de las lecturas crudas pasaron el filtro de calidad, las cuales fueron agrupadas en 75, 45 y 64 OTUs para la muestra QC-22, QC-57 y QC-95, respectivamente (tabla 8). Además, se encontró que entre el

7% y 19% de las lecturas representaban secuencias quiméricas que son secuencias “artefactos” formadas por dos o más secuencias biológicas unidas incorrectamente y al no eliminarse, pueden formar OTUs únicas y sobreestimar la diversidad y riqueza de la muestra. En el estudio conducido por Bjørnsgaard *et al.* (2017) reportan menor proporción de lecturas quiméricas cuando la secuenciación se realiza en las subregiones del ITS, en comparación con los estudios dirigidos a los genes ribosomales 16S y 18S donde entre 30% y 60% de las secuencias totales son quiméricas (Edgar *et al.*, 2011; Schloss *et al.*, 2011; Shin *et al.*, 2014; Bjørnsgaard *et al.*, 2017), concordando con lo obtenido.

Tabla 8. Rendimiento de los datos de secuenciación de la región ITS1 durante el análisis siguiendo la estrategia de Bálint *et al.*, 2014.

Etapa	QC-22	QC-57	QC-95
Secuencias crudas	319 173	237 623	369 143
Secuencias con adaptador ^a	139 988 (43.9 %) 104 670 (32.8 %)	161 109 (67.8 %) 124 490 (52.4 %)	124 460 (33.7 %) 104 265 (28.3 %)
Secuencias sin adaptadores, con calidad y longitud >50pb ^a	312 754 (98.0%)	234 522 (98.7%)	356 980 (96.7%)
Secuencias ensambladas ^b	264 384 (84.3%)	187 552 (80.0%)	314 977 (88.2%)
Secuencias identificadas ITS1 ^b	264 185 (99.9%)	187 403 (99.9%)	314 415 (99.8%)
Secuencias únicas (singletones) ^b	12 561 (4.8%)	7 284 (3.9%)	15 703 (5.0%)
Secuencias quiméricas ^b	18 259 (7.3%)	3 436 (19.1%)	33 775 (11.3%)
OTUs	75	46	64
Secuencias agrupadas en OTUs ^a	233 365 (73.1%)	145 754 (61.3%)	264 937 (71.8%)
OTUs asignadas	73	45	63
OTUs NO asignadas	2	1	1

^a Porcentaje respecto al número de lecturas crudas o iniciales de cada muestra

^b Porcentaje respecto al número de lecturas del paso anterior

La asignación taxonómica de las OTUs generadas con BLASTn permitió identificar una riqueza de 60 géneros fúngicos en los que se encontraban contenidas 56 especies distintas, la mayoría pertenecientes al *phylum Ascomycota* (96-99%) con solo algunas pertenecientes a *Basidiomycota* (0.2-3%). A nivel de familia, aquellas con más del 20% de abundancia fueron *Phaffomycetaceae*, *Saccharomycetaceae* y *Debaryomycetaceae*. Los géneros *Candida* y *Starmerella* presentes en algunas muestras no se asignó a nivel de familia (*Saccharomycetales fam Incertae sedis*) según la base de datos de UNITE+INSD. Inicialmente *Candida* se asignó en la familia *Saccharomycetaceae* (Schauer y Hanschke, 1999), sin embargo, acorde con la base de datos taxonómicos del NCBI, pertenece a *Debaryomycetaceae* (Schoch *et al.*, 2020). Por su parte, la relación filogenética del género *Starmerella* con respecto a otras levaduras no se ha establecido satisfactoriamente a partir del análisis de un solo gen (Santos *et al.*, 2018). Kurtzman y Robnett (2013) sugieren

que, basándose en análisis filogenéticos multigénicos, *Starmerella* puede ser un miembro de la familia *Trichomonascaceae*.

Los géneros asignados se clasificaron en tres grupos claramente definidos de acuerdo con su abundancia. Se consideró como dominantes a aquellos que representaran más del 10% de abundancia; subdominantes si estaban entre el 1% a 10%, y como no dominantes a los que tuvieran una abundancia inferior al 1%. Considerando la clasificación anterior, la población dominante entre las tres muestras de queso Cotija estuvo comprendida por *Kluyveromyces*, *Debaryomyces*, *Starmerella* y *Komagataella* en orden decreciente de abundancia (tabla 9, figura 12), siendo el último común entre las muestras. Dentro de la población subdominante se encontraron a los géneros *Aspergillus*, *Yamadazyma*, *Candida*, *Moniliella*, *Kodamaea*, *Issatchenkia* y *Whickerhamiella* cuya presencia y abundancia variaba según la muestra, pero se encontraban en el rango definido. El resto de los géneros detectados formaban parte de la población pobremente representada (Anexo IV).

Tabla 9. Anotación taxonómica de la fracción que representa más del 1% de la riqueza total relativa realizada con BLAST

Muestra	Abundancia	Género ^a	Especie ^b
QC-22	Dominante	<i>Debaryomyces</i> (54.4%) <i>Komagataella</i> (27.2%)	<i>D. hansenii</i> (100%) <i>K. phaffii</i> (31.0%)
	Subdominante	<i>Aspergillus</i> (6.3%) <i>Yamadazyma</i> (3.3%) <i>Moniliella</i> (3.0%) <i>Candida</i> (2.3%)	<i>A. westerdijkiae</i> (51.2%) <i>C. parapsilosis</i> (6.2%)
QC-57	Dominante	<i>Starmerella</i> (45.2%) <i>Komagataella</i> (32.7%) <i>Debaryomyces</i> (12.5%)	<i>K. phaffii</i> (31.8%) <i>D. hansenii</i> (99.9%)
	Subdominante	<i>Kluyveromyces</i> (4.7%) <i>Yamadazyma</i> (4.1%)	<i>K. lactis</i> (100%)
QC-95	Dominante	<i>Kluyveromyces</i> (68.0%) <i>Komagataella</i> (17.9%)	<i>K. lactis</i> (100%) <i>K. phaffii</i> (34.2%)
	Subdominante	<i>Issatchenkia</i> (3.2%) <i>Kodamaea</i> (3.4%) <i>Candida</i> (2.4%) <i>Wickerhamiella</i> (1.2%)	<i>I. orientalis</i> (100%) <i>K. ohmeri</i> (100%) <i>C. parapsilosis</i> (89.5%) <i>W. pararugosa</i> (100%)

^a En paréntesis la abundancia relativa con respecto a la abundancia total

^b En paréntesis la abundancia relativa con respecto al 100% de su género

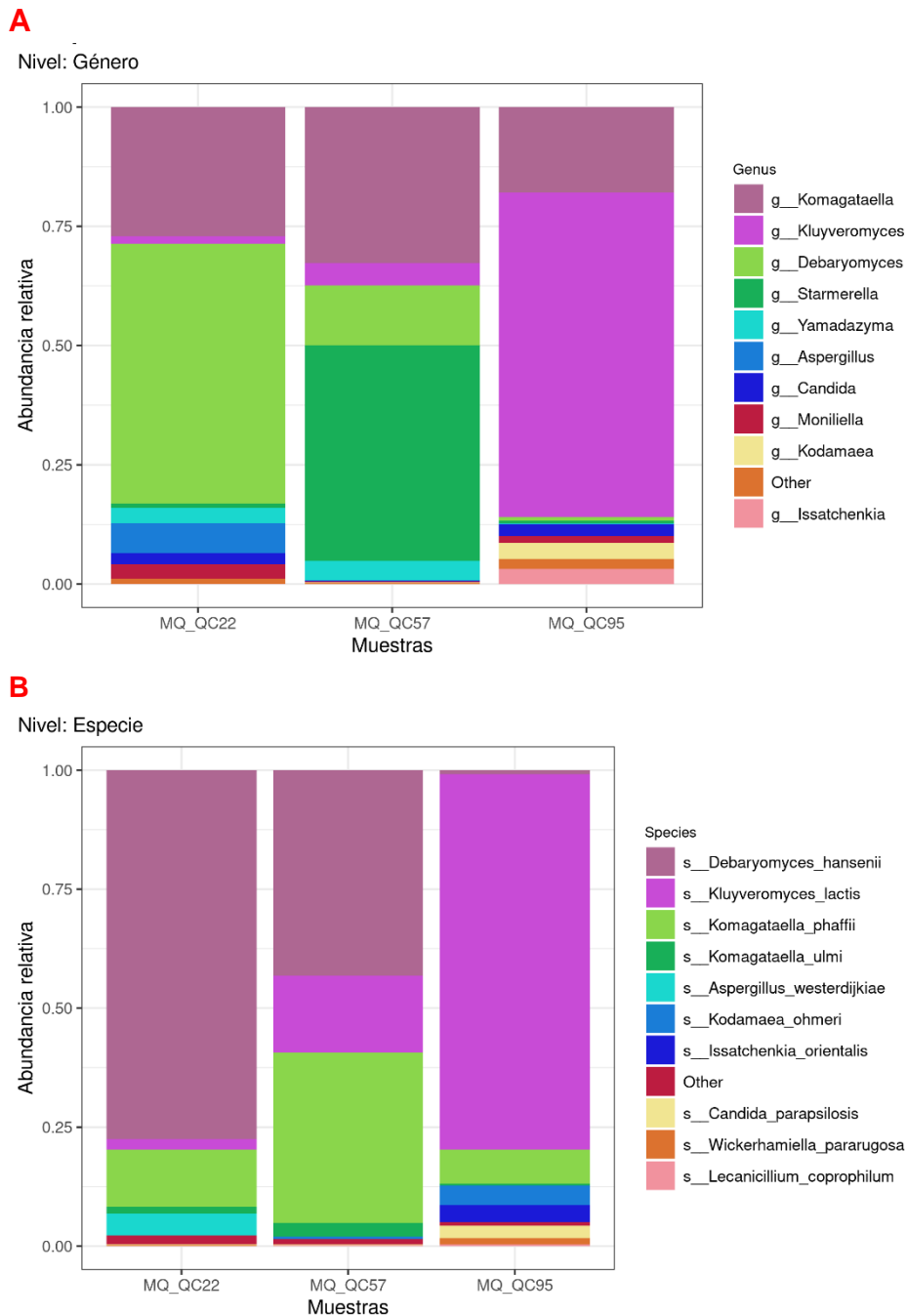


Figura 12. Abundancia relativa de las OTUs asignadas en las muestras de Queso Cotija según la estrategia propuesta por Bálint *et al.*, 2014. **A** – A nivel género; **B** – a nivel especie.

Por otra parte, el programa dada2 (basado en ASVs) realiza el filtrado de las lecturas considerando tanto la ausencia de nucleótidos ambiguos (N) ya que el programa no admite ninguna secuencia con este tipo de nucleótidos y es eliminada (maxN), la calidad global de la secuencia (truncQ) y las lecturas con errores superiores a los esperados son eliminadas (maxEE). Este paso de filtrado se corrió bajo los

parámetros establecidos, se obtuvo que entre el 24 y 46% de las lecturas cumplieron con lo anterior (tabla 10). De estas lecturas entre el 35 y 49% se identificaron como secuencias quiméricas y el resto se agruparon en ASVs. Para la muestra QC-22 se generaron 77 ASVs mientras que para la muestra QC-57 y QC-95 fueron 63 y 69 ASVs respectivamente.

Tabla 10. Rendimiento de los datos de secuenciación de la región ITS1 durante el análisis con dada2

Etapa	QC-22	QC-57	QC-95
Secuencias crudas	319 173	237 623	369 143
Secuencias con adaptador ^a	139 988 (43.9 %) 104 670 (32.8 %)	161 109 (67.8 %) 124 490 (52.4 %)	124 460 (33.7 %) 104 265 (28.3 %)
Secuencias sin adaptadores, con calidad y longitud >50pb ^a	113527 (35.6%)	109 942 (46.3%)	90 502 (24.5%)
Secuencias ensambladas ^b	99 229 (87.4%)	101 794 (92.6%)	84 860 (93.8%)
Secuencias quiméricas ^b	35 189 (35.5%)	42 403 (41.7%)	41 540 (49.0%)
Secuencias agrupadas en ASVs ^a	64 040 (20.1%)	59 391 (25%)	43 320 (11.7%)
ASVs	77	63	69
ASVs asignadas	77	62	68
ASVs NO asignadas	----	1	2

^a Porcentaje respecto al número de lecturas crudas o iniciales de cada muestra

^b Porcentaje respecto al número de lecturas del paso anterior

Para la asignación taxonómica dada2 emplea el clasificador Bayesiano ingenuo. Genera el perfil de k-meros de cada secuencia (8 bases) y los compara con los perfiles de k-meros de las secuencias de referencia de la base de datos utilizada. La secuencia de referencia con el perfil más similar se utiliza para asignar la taxonomía a la secuencia de consulta y evalúa la confianza de la asignación en cada nivel taxonómico (Wang *et al.*, 2007).

El perfil taxonómico basado en ASVs estuvo representado por 18 géneros y 22 especies, que al igual que con el perfil generado con OTUs, la mayoría pertenecían al *phylum Ascomycota* (94 – 99%) con una menor proporción asignados a *Basidiomycota* (0.4 - 5.6%). Las familias dominantes entre las tres muestras fueron *Phaffomycetaceae*, *Debaryomycetaceae* y *Saccharomycetaceae*; la primera de ellas con una abundancia mayor al 49% (tabla 11, figura 16). Al igual que con la base de datos UNITE+INSD, la utilizada en esta estrategia (base general de UNITE) tampoco asignaba la familia para los géneros *Candida* y *Starmerella*.

El género dominante compartido en las muestras de queso Cotija fue *Komagataella* con abundancias de entre 49 – 66%, coincidiendo con lo asignado por BLAST en el análisis por OTUs (tabla 11, figura 16). *Debaryomyces* y *Starmerella* también fueron

parte de los géneros dominantes en la muestra QC-22 y QC-57 respectivamente, con abundancias superiores al 18%. Los géneros identificados como subdominantes en la estrategia por OTUS, fueron los mismos que se asignaron con dada2 pero con abundancias distintas según la muestra. Los 7 géneros restantes presentaron una abundancia menor al 1% (Anexo IV).

Tabla 11. Anotación taxonómica de la fracción que representa más del 1% de la riqueza total relativa realizada con dada2

Muestra	Abundancia	Género ^a	Especie ^b
QC-22	Dominante	<i>Komagataella</i> (66.2%) <i>Debaryomyces</i> (18.9%)	<i>K. phaffii</i> (95.3%) <i>K. ulmi</i> (2.7%) <i>D. coudertii</i> (5.4%)
	Subdominante	<i>Moniliella</i> (5.6%) <i>Aspergillus</i> (3.2%) <i>Starmerella</i> (1.6%) <i>Yamadazyma</i> (1.1%)	<i>A. ruber</i> (78.8%) <i>S. etchellsii</i> (100%)
QC-57	Dominante	<i>Starmerella</i> (44.9%) <i>Komagataella</i> (49.6%)	<i>S. etchellsii</i> (100%) <i>K. phaffii</i> (95.9%) <i>K. ulmi</i> (2.1%)
	Subdominante	<i>Debaryomyces</i> (3.1%)	
QC-95	Dominante	<i>Komagataella</i> (66.6%)	<i>K. phaffii</i> (96.2%) <i>K. ulmi</i> (1.7%)
	Subdominante	<i>Issatchenkia</i> (6.5%) <i>Kluyveromyces</i> (5.0%) <i>Moniliella</i> (4.5%) <i>Candida</i> (4.3%) <i>Kodamaea</i> (2.6%) <i>Wickerhamiella</i> (2.6%) <i>Starmerella</i> (1.4%)	<i>I. orientalis</i> (100%) <i>K. lactis</i> (100%) <i>C. parapsilosis</i> (96.1%) <i>K. ohmeri</i> (100%) <i>W. pararugosa</i> (100%) <i>S. etchellsii</i> (100%)

^a En paréntesis la abundancia relativa con respecto a la abundancia total

^b En paréntesis la abundancia relativa con respecto al 100% de su género

Al comparar los perfiles taxonómicos de las dos estrategias o enfoques, se observa una diferencia en los géneros identificados a pesar de la similitud en el número de OTUs y ASVs. Con la generación de OTUs el perfil estuvo integrado por una mayor riqueza, 60 géneros entre las tres muestras; mientras que con ASVs sólo por 18 (figura 13). Los géneros dominantes y subdominantes se compartieron entre las estrategias, aunque dos de ellos difirieron en su proporción (figura 15). Las diferencias más marcadas fueron en los géneros pobremente representados que se identificaron por OTUs.

Durante el preprocesamiento de los datos se percató de una diferencia en el número de lecturas que se emplearon para la inferencia de OTUs y ASVs. Mientras que basado en OTUs se clusterizaron entre 145 y 264 mil lecturas, en la estrategia por

ASVs se agruparon entre 43 y 64 mil para su análisis e inferencia. Esto explicaría, en parte, la identificación de géneros con baja abundancia por OTUs, así como la menor riqueza de géneros asignada por ASVs.

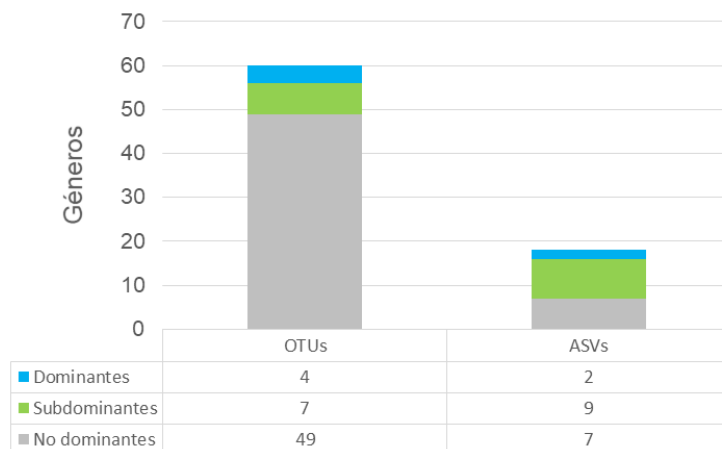


Figura 13. Diferencia en los géneros identificados según la estrategia de análisis.

De acuerdo con Callahan *et al.* (2016) el filtrado de las lecturas con dada2 considera tanto la calidad global de la lectura como las secuencias quiméricas y los errores causados por el proceso de secuenciación (modelo de error), lo cual favorece la detección de la variación biológica “verdadera” y se alcanza una mayor precisión en comparación con el 97% utilizado en OTUs (Glassman y Martiny, 2018; Nearing *et al.*, 2018, Chiarello *et al.*, 2022). Sin embargo, este filtrado de lecturas elimina gran parte de los datos y podría ser un factor en la disimilitud en los géneros de levaduras identificados en el queso Cotija.

El filtrado de las lecturas según ASVs se llevó a cabo de acuerdo a los parámetros recomendados por los autores y el modelo de error generado durante el análisis mostró el comportamiento esperado para las lecturas secuenciadas (figura 14). Este modelo representa la frecuencia observada de cada posible transición (cambio de una purina a otra purina, adenina a guanina; o de una pirimidina a otra pirimidina, citosina a timina) en función de la puntuación de calidad Phred asociada a las lecturas. Los puntos representan las tasas de error o las frecuencias de transición observadas según los datos y la línea negra representan las tasas de error que estima dada2. Se espera un ajuste de lo observado con lo esperado y que disminuya el error al aumentar la calidad Phred de la secuencia obtenida. Se probó un filtrado más laxo (modificando los parámetros maxEE y truncQ, datos no mostrados) pero se seguía eliminando gran parte de los datos así que se prefirió trabajar con los recomendados. Podría ser interesante evaluar el efecto del filtrado sobre las lecturas recuperadas y el perfil taxonómico inferido.

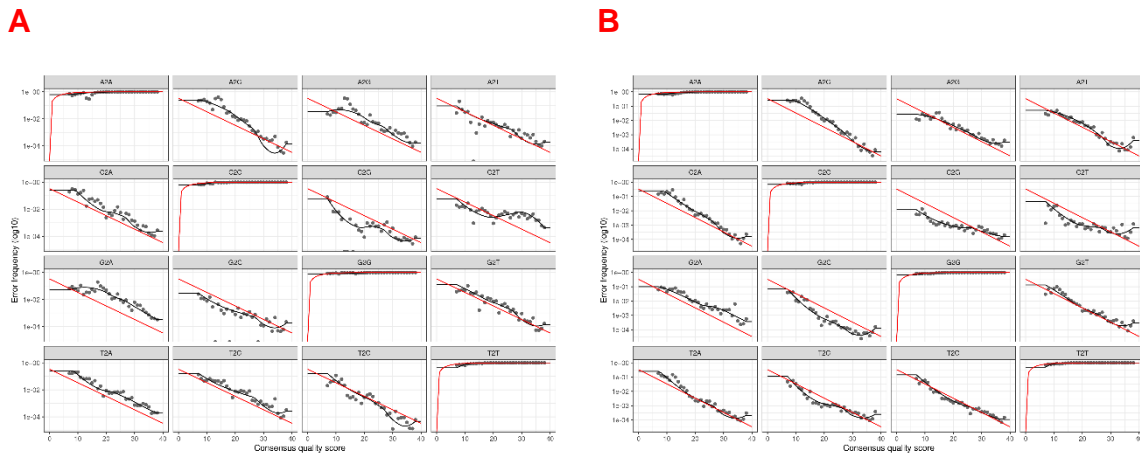


Figura 14. Modelo de error generado con el programa dada2. **A** – Lecturas directas, **B** - Lecturas reversas.

Cabe destacar que se ha reportado por varios estudios que el enfoque basado en OTUs presenta cierta variación en la riqueza y suele sobreestimar la diversidad (Prodan *et al.*, 2020; Chiarello *et al.*, 2022). Por lo que es necesario analizar con más detalle las OTUs con abundancia menor al 1% puesto que es más propenso a retener secuencias raras y a detectar OTUs espurias (Callahan *et al.*, 2016; Edgar, 2017). Así mismo, Reitmeier *et al.* (2021) proponen un filtrado según la abundancia relativa de 0.25% como un punto de corte, debido a que la agrupación en OTUS y eliminación de singletons, enfoque comúnmente utilizado durante el análisis y empleado en este proyecto, mantiene entre el 50 y 80% de las OTUs espurias, por lo que podría ser interesante llevarlo a cabo en los datos de ITS1.

Con respecto a los géneros con diferencia en su proporción según la estrategia de análisis, encontramos a *Debaryomyces* y *Kluyveromyces* (figura 15). El primer género se detectó en las muestras QC-22 y QC-57 donde formaba parte de la población dominante según OTUs, pero al evaluar con ASVs, éste se encontraba en menor proporción pasando a la población subdominante. *Kluyveromyces* se encontró en la muestra QC-95 y tuvo el mismo comportamiento que se comentó anteriormente. Esto se debe a que hay menos variación de secuencia, y por tanto menos ASVs detectadas a estos géneros. Además de que *Komagataella* resulta ser más representativa por ASVs al contar con una comunidad más compleja y con más variantes de secuencia con abundancia sustancial en las muestras que se analizaron del queso Cotija. No es posible detectar esta variación por la estrategia de OTUs ya que estas variantes observadas difieren en solo uno o dos nucleótidos sobre la región secuenciada. Lo anteriormente descrito destaca a la estrategia de análisis por ASVs, lo que le confiere una mayor precisión en la inferencia.

De acuerdo con lo anterior es posible concluir que ambas estrategias de análisis son buenas aproximaciones para conocer la diversidad de levaduras en una muestra a partir de lecturas de amplicones, dado que se identificaron los mismos géneros dominantes y subdominantes en el queso Cotija. Mientras que con OTUs se analiza una mayor proporción de los datos secuenciados y se detectan géneros que están en baja abundancia (menor al 1%), con ASVs se pueden encontrar variantes de secuencias que difieren en tan solo un nucleótido y tener más precisión. Sin embargo, considerando los reportes de la sobreestimación de la diversidad basado en OTUS y que puede ser complicada la instalación de programas, el análisis por ASVs con dada2 resulta más práctico y confiable. Además, los estudios basados en ASVs tienen mayor reproducibilidad y facilidad de comparación con otras investigaciones. Por lo cual, en las siguientes dos secciones se analizarán a más detalle los perfiles taxonómicos de cada queso y la diversidad alfa obtenidos únicamente con el programa dada2.

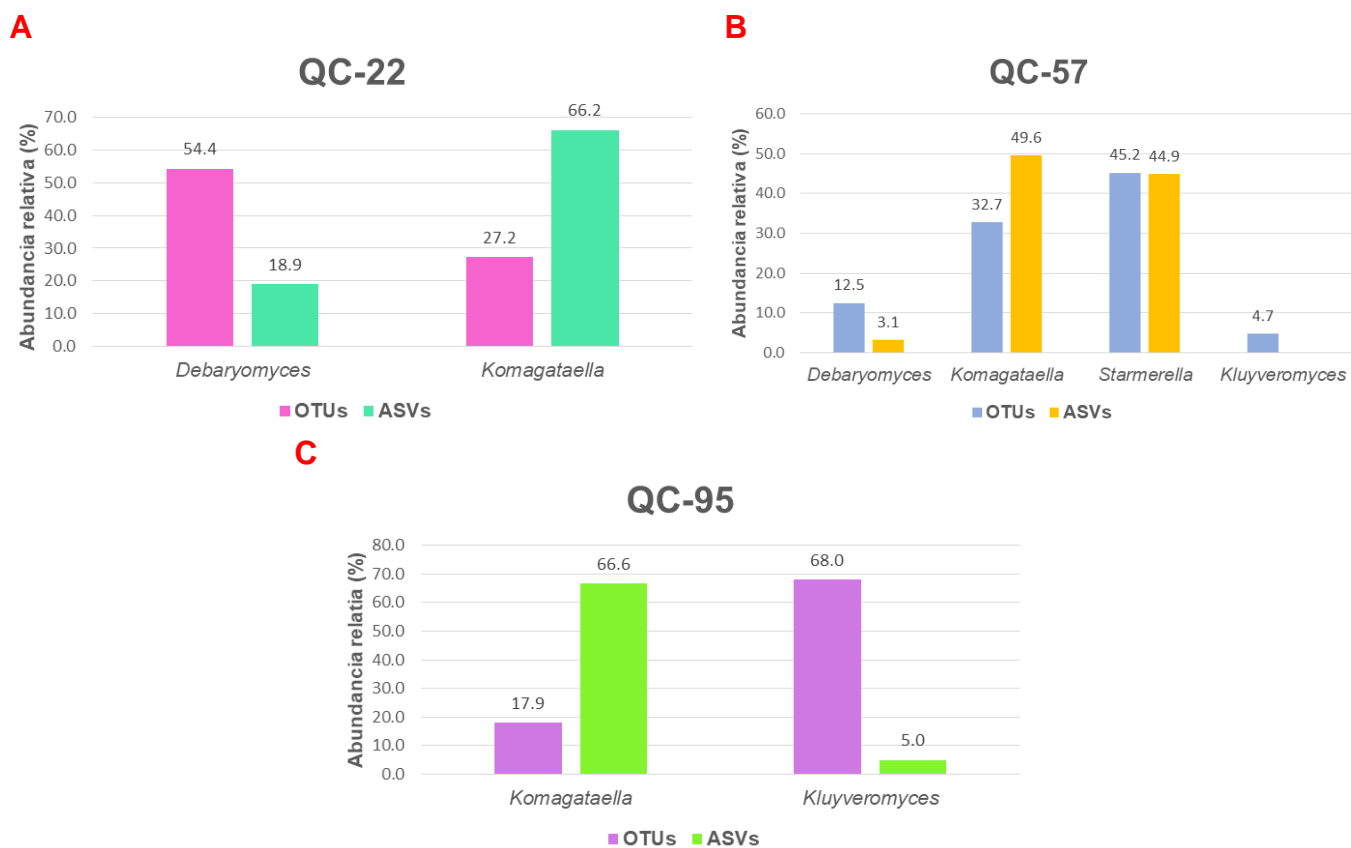


Figura 15. Diferencia entre los géneros dominantes y subdominantes según la estrategia de análisis OTUs o ASVs. **A** – QC-22; **B** – QC-57; **C** – QC-95

Otro factor a considerar en el análisis de los datos es la base de referencia utilizada, ya que con éstas se determina la asignación filogenética de las lecturas obtenidas por secuenciación. Las incluidas en cada estrategia son de diferente tamaño,

UNITE+INDS tiene aproximadamente 6 millones de secuencias, mientras que la base general UNITE tiene 188 070 entradas. Sin embargo, con los dos enfoques de análisis se recuperaron la misma población de levaduras dominantes y subdominantes, aunque con abundancias variables, y su influencia no fue evidente.

8.2.2.2 Perfil taxonómico de cada muestra de queso Cotija

Las muestras de queso Cotija analizadas presentaron géneros y abundancias distintas según la estrategia por ASVs (figura16), es decir, un perfil taxonómico diferente.

La muestra QC-22 presentó como géneros dominantes a *Komagataella* y *Debaryomyces* con 66.2% y 18.9% de abundancia respectivamente. La población subdominante estuvo conformada por *Aspergillus*, *Yamadazyma*, *Moniliella* y *Starmerella* con distintas abundancias.

Para la muestra QC-57 la comunidad de levaduras dominantes fueron *Starmerella* y *Komagataella* con abundancias superiores al 30%. *Debaryomyces* fue el único subdominante con 3.1%. Fue la muestra con menor diversidad.

Finalmente, la muestra QC-95 exhibió la mayor diversidad al tener 8 géneros formando parte de su comunidad dominante y subdominante, en comparación con las dos otras muestras analizadas. Presentó a *Komagataella* como género dominante. Los géneros *Kluyveromyces*, *Issatchenkia*, *Kodamaea*, *Candida*, *Wickerhamiella*, *Moniliella* y *Starmerella* presentaron abundancias de entre 1 y 5%.

Las diferencias en los géneros dominantes y subdominantes entre las muestras pueden atribuirse, entre otros factores, al lugar de procedencia recordando que las tres muestras analizadas se elaboraron en distintas regiones de la sierra de Jalisco. Dos de las muestras, QC-57 y QC-95, se elaboraron en regiones distantes del estado de Jalisco (Quitupan y Sta. Ma. del Oro) y solo una muestra del estado de Michoacán de la región de Cotija (QC-22). Se ha reportado que las comunidades microbianas presentes en quesos se ven influenciadas por diversos factores entre los que se encuentra la ubicación geográfica, las diferentes condiciones climáticas y los ecosistemas de los pastizales que se ve reflejada en la calidad de la leche obtenida durante la ordeña (Sessou *et al.*, 2019). Así mismo, a pesar de que las muestras de queso Cotija se elaboraron siguiendo las Reglas de uso, la interacción del quesero con la cuajada y los utensilios empleados durante la manufactura, pueden también estar contribuyendo a la disimilitud del perfil taxonómico. Considerando la diversidad y los géneros dominantes entre las muestras, se puede considerar que *Komagataella* es la firma del queso Cotija, ya que se encontró en abundancias superiores al 18% y presente en las tres muestras.

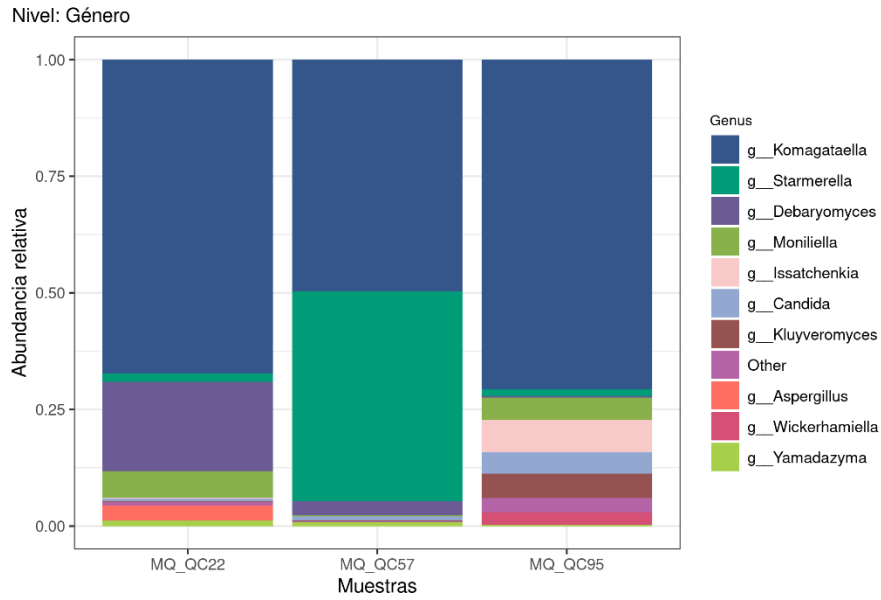
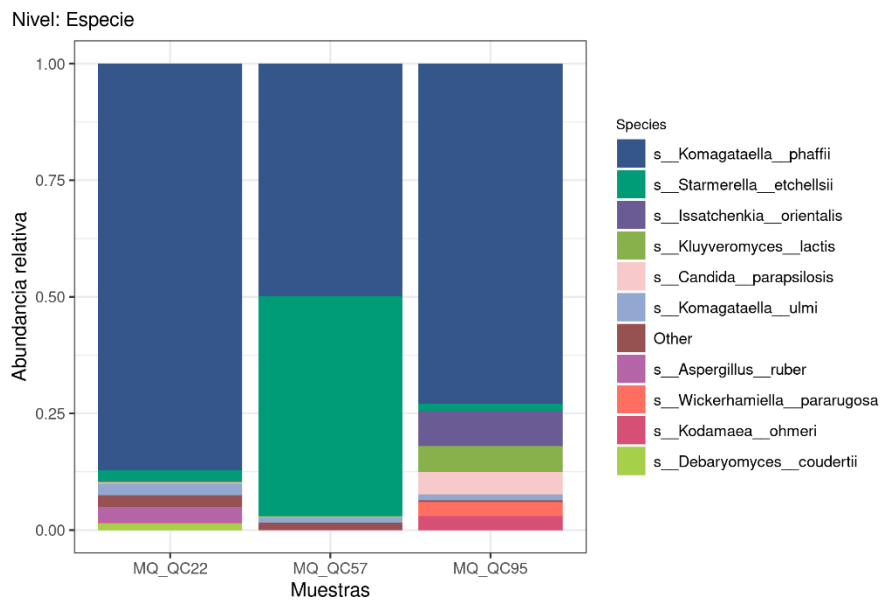
A**B**

Figura 16. Abundancia relativa de las ASVs asignadas en las muestras de Queso Cotija con dada2. **A** – Nivel género; **B** – Nivel especie.

Finalmente, un factor que no se consideró durante la extracción del ADN a partir de las muestras del queso fue la posible contaminación con ADN fúngico presente en los reactivos utilizados. Hay reportes sobre presencia de ADN de hongos y levaduras en las enzimas líticas utilizadas comúnmente en la extracción de ADN, como zimolasa y liticasa, y en reactivos para PCR (Miyajima *et al.*, 2009; Czurda *et al.*, 2016). Miyajima *et al.* (2009) estimaron por PCR en tiempo real la contaminación por ADN ribosomal fúngico en dos enzimas comerciales (Zimolasa-100T de

Seikagaku Biobusiness y Liticasa de Sigma-Aldrich, L2524 grado biología molecular), para ello amplificaron y detectaron la región D1/D2 del gen ribosomal 28S. Cuantificaron una concentración de 0.0323 copias/unidad de actividad enzimática para la liticasa, que también se empleó en el proyecto, de tal forma que no estaría contribuyendo de manera importante al ADN extraído y al perfil taxonómico descrito. Así mismo, se tomaron precauciones a la hora de realizar la extracción y purificación de ADN para evitar su contaminación con ADN ambiental como por ejemplo uso de guantes, cubrebocas y campana de flujo laminar. Sin embargo, sería interesante evaluar la presencia de ADN fúngico en la enzima liticasa utilizada y en los reactivos de PCR.

8.2.2.3 Diversidad alfa y esfuerzo de muestreo

A partir de la tabla de ASVs se evaluó el esfuerzo de muestreo y se calcularon algunos índices de diversidad alfa para analizar la comunidad de levaduras presentes. En la figura 17 se muestran las curvas de acumulación de especies o de rarefacción para cada una de las muestras, las cuales presentaron el comportamiento asintótico esperado y que representa un buen esfuerzo de muestreo.

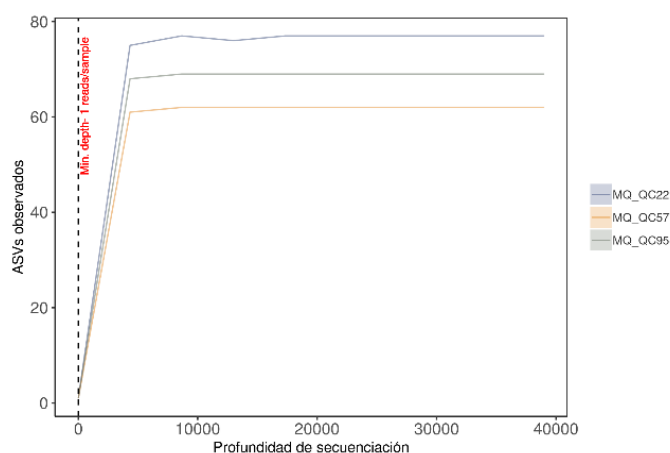


Figura 17. Curva de rarefacción basada en ASVs.

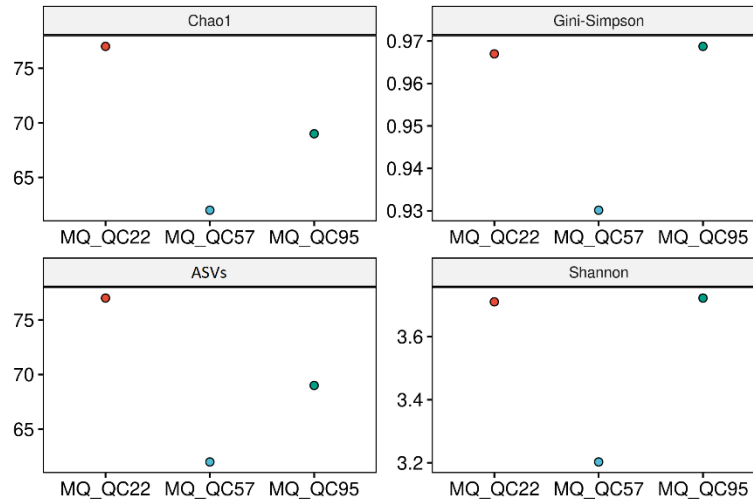
El número de especies observadas (ASVs) es el descriptor más básico de la diversidad. Se estima mayor número especies en la muestra QC-22 y QC-95 que en la muestra QC-57 (figura 18). Respecto al estimador no paramétrico Chao1 este considera la riqueza de las especies añadiendo un factor de corrección para las especies singletones y dobletones presentes. Es sensible al tamaño de la muestra y subestima la riqueza real cuando el muestreo es pobre (Escalante, 2007). Los valores de este índice son cercanos al observado lo que concuerda con el comportamiento descrito en la curva de rarefacción.

El siguiente parámetro es el índice de Gini-Simpson [complemento del índice de Simpson, (1-D)], que puede tomar valores desde cero para comunidades de baja diversidad, hasta valores cercanos a 1 (Krebs, 2014). Éste índice alcanzó valores cercanos a uno indicando una mayor diversidad en las muestras analizadas manteniendo que las más diversas son QC-22 y QC-95.

Por otro lado, el índice de Shannon es una medida de la entropía que muestra la heterogeneidad de la comunidad tomando en cuenta el número de especies presentes y su abundancia relativa. Para comunidades con una sola especie toma un valor de cero y, para comunidades con mayor diversidad la incertidumbre incrementa y adquiere valores más altos. De acuerdo con Krebs (2014) en muestras biológicas el índice suele ser menor a 5.0, tal como ocurre en este caso. Murugesan *et al.* (2018) reportan valores de entre 2.9 y 3.5 en quesos elaborados con leche no pasteurizada. Valores bajos para este índice se han reportado en la cuajada y sueros cuyos límites van de 0.47 a 2.60 (Filippis *et al.*, 2014). Considerando los valores anteriores y el tipo de muestra, la comunidad de levaduras en el queso Cotija es de baja a media según este parámetro y la muestra QC-57 es la de menor diversidad.

Tomando en cuenta los tres índices podemos inferir que las muestras QC-22 y QC-95 son más diversas con respecto a QC-57. Ambas muestras presentaron valores más altos y cercanos en cada índice. Por su parte, la muestra QC-57 estuvo representada casi en su totalidad por dos especies que sumaban una abundancia relativa de aproximadamente 95%, por lo que las métricas de diversidad fueron más bajas. Además, en la población subdominante solo estaba presente un género. En cambio, las otras dos muestras, a pesar de contar con una a dos especies dominantes que alcanzaba una abundancia de entre 60% a 80%, su población subdominante contaba con más géneros, contribuyendo a los índices de diversidad.

Diversidad alfa



	Chao1	ASVs observados	Gini-Simpson	Shannon
QC-22	77.0	77	0.97	3.71
QC-57	62.0	62	0.93	3.20
QC-95	69.0	69	0.97	3.72

Figura 18. Índices de diversidad alfa basada en ASVs.

8.2.2.4 Población dominante

La población de levaduras en el queso Cotija está caracterizada principalmente por tres géneros y que corresponden a *Komagataella*, *Starmarella* y *Debaryomyces*.

La presencia del género *Debaryomyces* es común en quesos madurados y semi-madurados. Comprende a levaduras halotolerantes y psicrótrofas capaces de crecer en ambientes con condiciones extremas. En total son 18 especies y la más representativa es *D. hansenii* dadas sus características metabólicas y que se describirán en un apartado siguiente. Son ubicuas y pueden aislarse de diversos hábitats naturales desde el suelo, el polen, exudados de árboles y plantas, hasta en alimentos salados y fermentados con una actividad de agua reducida (Jacques *et al.*, 2015). Las levaduras de este género pueden asimilar distintas fuentes de carbono, no son capaces de realizar fermentación y por tanto su crecimiento en ausencia de oxígeno es bajo (Rossi y Torriani, 2018). Considerando que el queso Cotija cuenta con una pasta fiable o desmoronable, es probable que se generen microcanales que favorezcan la entrada de aire en él (Escobar, 2012) y se permita el crecimiento de este género además de que la concentración de cloruro de sodio no sería una limitante. Se ha descrito que, en quesos, las levaduras pueden metabolizar el ácido láctico producido por bacterias LAB y exhiben actividad lipolítica

que contribuye al desarrollo de compuestos que dan sabor característico. La especie *Debaryomyces coudertii* anotada con la estrategia basada en ASVs, no se ha identificado en alimentos sino en excrementos de ave. Se ha reportado que, a nivel de secuencia, la región 5.8-ITS tiene una alta similitud (mayor al 99%) con otras especies del género *Debaryomyces* (*D. castellii*, *D. hansenii*, *D. polymorphus*, *D. pseudopolymorphus*, *D. robertsiae*, *D. udonii* y *D. vanriijiae*) en especial con *Debaryomyces nepalensis* (Martorell *et al.*, 2005).

El género *Starmerella* incluye aproximadamente 50 especies, la mayoría de las cuales se han descrito asociadas con abejas e insectos y sus nichos florales. No obstante, algunas especies se han reportado en alimentos fermentados y bebidas, por ejemplo, en la producción de vinos son capaces de reducir la acidez y metabolizar el etanol, producir glicerol y algunos compuestos volátiles (Ciani *et al.*, 2022). Ciertas especies que pertenecían al género *Candida* han sido anotadas a este género, tal es el caso de la especie *Candida etchellsii* (*Starmerella etchellsii*). Esta especie se identificó previamente como parte de la microbiota del queso Cotija de acuerdo con Chombo-Morales *et al.* (2016), encontrándose dispersa tanto en la superficie como en la matriz principal. Es una levadura halotolerante capaz de producir compuestos que dan sabor y aroma como alcoholes, ésteres y aminoácidos libres. Bajo estrés salino (NaCl 5%), muestra una capacidad fuerte para producir ésteres como hexanoato de etilo y acetato de hexilo, que otorgan sabor afrutado, y es necesario para que sintetice ésteres (Feng *et al.*, 2012; Lu *et al.*, 2022). Considerando lo anterior, es probable que *S. etchellsii* contribuya de manera importante al perfil sensorial del producto final.

El otro género dominante, *Komagataella*, comprende a especies metilotróficas que son capaces de utilizar como única fuente de carbono el metanol y muestran un rápido crecimiento a altas densidades celulares en fuentes de carbono como glucosa, glicerol, etanol y metanol en condiciones aeróbicas. Varias especies de *Pichia* se han reanotado en este género. Actualmente el género consiste en siete especies aisladas de ambientes arbóreos. Las cepas bien caracterizadas son *Komagataella phaffii* y *Komagataella pastoris* (ambas anteriormente clasificadas como *Pichia pastoris*) por su importancia en la producción de proteínas recombinantes (Heisting *et al.*, 2022). No se ha descrito con anterioridad este género en productos lácteos ni quesos madurados y semi-madurados, pero su presencia podría deberse a los pastizales en los que el ganado es alimentado por libre pastoreo y tiene influencia en la calidad de la leche obtenida durante la ordeña y en su microbiota.

8.2.2.5 Población subdominante

El análisis taxonómico de la región ITS1 muestra la presencia de una variedad de levaduras como parte de la población subdominante en las muestras de queso Cotija con abundancias de entre 1 a 10%. Fue la población más dispareja de entre

las muestras ya que no todos los géneros se compartieron. Entre los géneros identificados se encuentran *Kluyveromyces*, *Aspergillus*, *Candida*, *Kodamaea*, *Yamadazyma*, *Moniliella*, *Issatchenkia* y *Wickerhamiella*.

El género *Kluyveromyces* fue un caso especial debido a que presentó diferencias en la abundancia dependiendo la estrategia de análisis empleada. Es común identificarlo como parte de la microbiota de productos lácteos y en estudios previos se identificó en el queso Cotija. Comprende al menos 17 especies y pueden dividirse según la fuente de carbono que pueden metabolizar. En quesos madurados y semi-madurados, la presencia de las especies *K. lactis* y *K. marxianus* se justifica principalmente por su capacidad de fermentar la lactosa residual en el queso y gracias a esta característica es utilizada en la industria biotecnológica para la producción de la enzima β -galactosidasa (Johnson y Echavarrí-Erasun, 2011). También poseen actividades proteolíticas y lipolíticas, pero más débiles que otras levaduras presentes en el queso como *Y. lipolytica* y *D. hansenii*. Sin embargo, estas dos especies al metabolizar los componentes de la leche (lactosa, proteínas y grasas) las hace importantes para la maduración y contribuyen en la formación de compuestos que dan aroma como por ejemplo alcoholes, aldehídos, ésteres y terpenos (Leclercq-Perlat *et al.*, 2004; Belloch *et al.*, 2011).

La presencia del género *Aspergillus* llama la atención ya que es la primera vez que se reporta en el queso Cotija y, además, durante los análisis microbiológicos de cuenta en placa para hongos filamentosos realizados en proyectos anteriores, no sobrepasan el límite establecido (tabla 3). Los miembros de este género se conocen por ser osmo-, xero- y/o halotolerantes. Tienen una amplia distribución y se han aislado de alimentos con baja humedad como los cereales, así como en productos que contienen altas concentraciones de azúcar, como jarabes, mermeladas y jaleas, y productos cárnicos salados. También, algunas especies son utilizadas en la industria biotecnológica para la producción de ácidos orgánicos, antibióticos o enzimas, entre otros. Generalmente la presencia de especies de este género en quesos se ha relacionado con defectos en los productos según lo reportado por Garnier *et al.* (2017), así mismo pueden ser de preocupación clínica ya que algunas especies pueden producir micotoxinas. Se ha reportado a *Aspergillus westerdijkiae* como una de las especies importantes para la producción de ocratoxina A en café, arroz, bebidas y otros alimentos (Visagie *et al.*, 2014). Por lo cual, será importante investigar la presencia de este género en el queso Cotija.

Por su parte, *Aspergillus ruber* se ha aislado de granos de café, té, suelo y de muestras ambientales y no se ha detectado la producción de micotoxinas. Es una especie importante en la investigación al ser capaz de producir una cantidad sustancial de metabolitos secundarios pertenecientes a varias clases en las que algunos de ellos mostraron actividades biológicas prometedoras como anticancerígeno (Chen *et al.*, 2017; Youssef y Singab, 2021).

Ciertas especies del género *Candida* se han descrito no solo en quesos madurados sino también en alimentos fermentados. Es un género heterogéneo que comprende microorganismos con diferentes características fisiológicas y metabólicas que les permite crecer en diversos ambientes. Abarca alrededor de 200 especies de las cuales aproximadamente 40 son patógenas para los humanos. La especie *C. parapsilosis* ya había sido reportada como parte de la microbiota del queso Cotija y se confirma su presencia en este estudio. Posee actividad lipolítica, contribuyendo en la producción de compuesto volátiles y por tanto al perfil sensorial del alimento. En algunos estudios, proponen a la especie como un agente de biocontrol contra hongos como *Fusarium* y *Aspergillus* (Niknejad *et al.*, 2012; Fallah *et al.*, 2016), por lo que es posible que la presencia de esta levadura en el queso Cotija inhiba el crecimiento de hongos filamentosos y por ello su cuenta en placa sea baja.

En el estudio realizado por Chombo-Morales *et al.* (2016) sobre una muestra de queso Cotija, describieron a dos de los géneros subdominantes identificados en este estudio, *Kodamaea* y *Moniliella*, presentes tanto en la superficie del queso como en la matriz principal. Estos géneros se han reportado en alimentos fermentados como el cacao y alimentos en escabeche, así como en muestras ambientales e insectos, pero no en productos lácteos. *Kodamaea* se ha estudiado y aislado principalmente de flores e insectos y consta de 22 especies (Gao *et al.*, 2017). Por su parte el género *Moniliella* comprende alrededor de 11 especies basidiomicetas, algunas de las cuales se han utilizado para la producción de eritritol mientras que otras son consideradas patógenas en humanos (Thanh *et al.*, 2013; Joint Genome Institute, 2022).

El siguiente género, *Yamadazyma* es uno de los géneros más grandes de la familia *Debaryomycetaceae*, sus especies se puede encontrar en regiones tropicales, subtropicales y templadas de diferentes continentes. Se han aislado principalmente de madera e insectos, pero también se recuperan de flores, hojas, frutas, corteza de árboles, agua de mar, minerales y la atmósfera. Actualmente consta de 24 especies, pero 38 especies más aún se encuentran asignadas al género *Candida* (Gao *et al.*, 2021).

Por último, *Wickerhamiella* e *Issatchenkia* no se han descrito en ambientes lácteos ni en quesos. El primero está constituido por 35 especies, la mayoría con un alto grado de especialización nutricional y ecológica, un perfil estrecho de asimilación de carbono y una fuerte asociación con flores e insectos que las visitan, especialmente moscas y escarabajos (de Vega *et al.*, 2017). Sin embargo, algunas especies de este género también se aislaron de bebidas alcohólicas, agua, rizosfera y de muestras de humanos. Algunos estudios revelaron que la planta de caña de azúcar podría ser un nuevo hábitat para algunas levaduras asociadas con el género (Avchar *et al.*, 2019). La especie *Wickerhamiella pararugosa* antes pertenecía al género *Candida*, puede asimilar glicerol y requiere tiamina y biotina para su crecimiento, diferenciándose de la especie *Candida rugosa* que no asimila el glicerol.

El género *Issatchenkia* pertenece a la familia *Pichiaceae* y la especie más estudiada es *Issatchenkia orientalis*, la cual se utiliza en la producción de vinos. Producen etanol y tiene mayor tolerancia a temperaturas altas, a concentraciones de sal y ácidos que *S. cerevisiae*. Debido a su resistencia a múltiples factores de estrés, *I. orientalis* tiene una aplicación potencial en la producción de bioetanol y la producción de ácido succínico (Miao *et al.*, 2018). Dado que el queso Cotija cuenta con una concentración de cloruro de sodio de aproximadamente 5% y actividad de agua baja, es probable que el género resista estas condiciones y formen parte de la microbiota.

En general, la mayoría de los géneros subdominantes asignados en las muestras de queso Cotija no habían sido descritos en productos lácteos y en particular en quesos, por lo que es la primera vez que se describen como parte de la microbiota de este queso artesanal. Además, resalta que el ambiente en que se elabora el producto tiene un impacto tanto en la diversidad microbiana de leche con la que se elabora y en el producto final ya que, como se mencionó anteriormente, los géneros como *Kodamaea*, *Moniliella*, *Yamadazyma*, *Wickerhamiella* e *Issatchenkia* se han asociado a ambientes arbóreos, flores e insectos, que forman parte del ecosistema y el cual puede ser una fuente directa de levaduras en la leche y el queso.

8.2.3 Datos metagenómicos de *shotgun*

8.2.3.1 Limpieza de adaptadores y secuencias contaminantes

El total de lecturas pareadas obtenidas por secuenciación *shotgun* fue de 133 229 485, que presentaron una longitud promedio de 101 pb. Se detectaron secuencias adaptadoras de la tecnología Illumina (5'- AGATCGGAAGAGC-3') tanto en las lecturas directas como en las inversas en el 31.6% y 38.1% de las lecturas, respectivamente (tabla 12), los cuales fueron eliminados y, dado que algunas bases presentaron baja calidad Phred, estas se recortaron utilizando el programa Trim Galore.

Después se realizó un alineamiento contra el genoma de *Bos taurus* pues, como se describió anteriormente, durante la extracción de ADN de las muestras de queso se acarrea también el ADN de vaca. Aproximadamente mil lecturas se alinearon contra el genoma de referencia, que representa solo el 0.001% (tabla 12). El resto de las secuencias corresponderían a la microbiota presente en la muestra. Probablemente la baja cantidad de lecturas alineadas contra el genoma de *Bos taurus* se debe a que los datos que se compartieron y subieron a ENA fueron los limpios y no los crudos.

Tabla 12. Rendimiento de los datos de secuenciación de *shotgun* durante el preprocesamiento

	Lecturas directas (5')	Lecturas inversas (3')
Lecturas totales	133 229 485	133 229 485
Lecturas con adaptadores	42 139 170 (31.6%)	50 800 793 (38.1%)
Lecturas filtradas sin adaptadores	122 364 037 (91.84%)	
Lecturas alineadas	1 713 (0.0014%)	
Lecturas NO alineadas	122 362 324 (99.99%)	

8.2.3.2 Análisis taxonómico

Continuando con el procesamiento de los datos de *shotgun*, se realizó la anotación taxonómica de las secuencias del queso Cotija empleando dos programas diferentes, Kraken2 y Kaiju, cuyos resultados se presentan en las tablas 13 a 15. Los dos programas tienen la misma aproximación de análisis de las lecturas, pero difieren en su algoritmo.

El programa Kaiju realiza la clasificación taxonómica basada en la comparación de secuencias de aminoácidos contra una base de datos de referencia de proteínas. Antes de la clasificación traduce las lecturas de nucleótidos a secuencias de aminoácidos y después busca las coincidencias exactas máximas (MEM) en una base de datos dada de proteínas anotadas a genomas de referencia. El algoritmo de comparación de secuencias utiliza la transformada de Burrows-Wheeler (BWT) para lograr una alta velocidad de clasificación (Menzel *et al.*, 2016). Mientras que Kraken2, trabaja con las secuencias de nucleótidos obtenidas por secuenciación. Primero construye un índice a partir de todos los k-meros que se encuentran en los genomas de referencia y asigna cada k-mero al ancestro común más bajo (LCA, por sus siglas en inglés) de todas las especies que tienen ese k-mero. Luego, hace coincidir los k-meros encontrados en las lecturas de secuenciación con este índice y, finalmente, asigna la lectura al taxón con la mayoría de los k-meros coincidentes siguiendo un camino desde la raíz del árbol (Wood *et al.*, 2019).

Ambos programas clasificaron una mayor proporción de lecturas a bacterias en comparación con las clasificadas a hongos y levaduras (tabla 13). Escobar (2016) reportó que la microbiota del queso Cotija estaba representada casi en su totalidad por bacterias al analizar genes marcadores de copia única y al gen ribosomal 16S, siendo los géneros dominantes *Lactobacillus*, *Weisella* y *Leuconostoc*, concordando con lo obtenido.

Tabla 13. Lecturas anotadas con Kraken2 y Kaiju

Lecturas anotadas a:	Kraken2 v2.0.8 ^a	Kaiju v1.8.2 ^a
Totales	122 362 324	122 362 324
Bacterias	111 491 103 (91.11%)	112 206 288 (91.70%)
Arqueas	2 477 (0.002%)	4 957 (0.004%)
Eucariontes	88 448 (0.072%)	103 780 (0.085%)
Hongos y levaduras	78 608 (0.064%)	100 501 (0.082%)
Virus	10 205 (0.008%)	16 138 (0.013%)
Sin clasificar	10 770 091 (8.80%)	10 031 161 (8.20%)

^a En paréntesis la abundancia relativa de cada grupo con respecto a la abundancia total.

Con respecto al grupo fúngico, la familia dominante fue *Debaryomycetaceae* con el 80.85% con Kaiju y 96.92% con Kraken2 (tabla 14). Esta familia comprende los siguientes 11 géneros *Debaryomyces*, *Kurtzmaniella*, *Lodderomyces*, *Meyerozyma*, *Millerozyma*, *Priceomyces*, *Scheffersomyces*, *Schwanniomyces*, *Spathaspora*, *Wickerhamia* y *Yamadazyma*. En la actualidad, más de 100 especies de levaduras pertenecen a esta familia que han sido aislados de una amplia variedad de muestras, como suelo, agua, alimentos, sustratos vegetales y muestras asociadas a animales (Hui *et al.*, 2014).

En quesos en particular se ha descrito a *Debaryomycetaceae* como familia dominante seguida de *Saccharomycetaceae*, *Dipodascaceae* y *Pichiaceae* (Michailidou *et al.*, 2021). Una característica importante de las especies de esta familia es su capacidad de tolerar ambientes de alta salinidad, otorgándole una ventaja con respecto a otras especies. Por lo cual, no es sorprendente la identificación de especies de esta familia como parte de la microbiota de quesos madurados y semi-madurados. El queso Cotija tiene una concentración de cloruro de sodio de 5% y anteriormente, se había descrito a las familias *Debaryomycetaceae* y *Saccharomycetaceae* como familias dominantes (Escobar, 2016).

A nivel de especie, ambos programas, identificaron a *Debaryomyces hansenii* como dominante con el 47.44% (Kaiju) y 93.84% (Kraken2) [figura 19]. Esta levadura junto con *Kluyveromyces spp.*, son consideradas como las especies más comunes y abundantes que se encuentran casi en todos los tipos de quesos y ambientes lácteos (Fox *et al.*, 2017). Es una levadura marina osmotolerante y halotolerante capaz de crecer en ambientes con una salinidad de hasta 24% y baja temperatura, metaboliza lactosa, ácido láctico y cítrico (Sacerdot *et al.*, 2008).

Se debe destacar que *D. hansenii* produce aldehídos y alcoholes de cadena ramificada que aceleran el desarrollo de sabor a queso (Sørensen *et al.*, 2011). Además, muestra actividad proteolítica y lipolítica, pero la intensidad es específica de la cepa por lo que la influencia de *D. hansenii* en la textura y el sabor del queso pueden variar. Sin embargo, se ha reportado que tiene un papel importante durante

el proceso de maduración en especial en la etapa inicial debido a su metabolismo (Mansour *et al.*, 2008; Fröhlich-Wyder *et al.*, 2019).

Tabla 14. Principales familias anotadas a partir de los datos metagenómicos de *shotgun* del queso Cotija.

Kraken2 v2.0.8			Kaiju v1.8.2		
Especie	Conteo	Abundancia relativa (%) ^a	Especie	Conteo	Abundancia relativa (%) ^a
<i>Debaryomycetaceae</i>	76184	96.916	<i>Debaryomycetaceae</i>	66090	80.848
<i>Saccharomycetaceae</i>	1004	1.277	<i>Glomeraceae</i>	11235	13.744
<i>Nectriaceae</i>	190	0.242	<i>Mucoraceae</i>	3326	4.069
<i>Pichiaceae</i>	109	0.139	<i>Saccharomycetaceae</i>	1818	2.224
<i>Trichomonascaceae</i>	88	0.112	<i>Harpellaceae</i>	806	0.986
<i>Glomerellaceae</i>	86	0.109	<i>Aspergillaceae</i>	949	1.161
<i>Malasseziaceae</i>	79	0.100	<i>Pleosporaceae</i>	679	0.831
<i>Ustilaginaceae</i>	68	0.087	<i>Saccharomycodaceae</i>	609	0.745
<i>Pyriculariaceae</i>	66	0.084	<i>Trichomonascaceae</i>	474	0.580
<i>Phaffomycetaceae</i>	44	0.056	<i>Punctulariaceae</i>	423	0.517

^a Abundancia relativa con respecto al 100% de las lecturas asignadas a hongos y levaduras

Al comparar con las especies asignadas previamente por Escobar (2016) con el gen ribosomal 18S (Figura 7, tabla 4), ella identificó distintas especies de *Debaromyces* las cuales se encontraban entre las dominantes (*D. udenii*, 14.2%) y subdominantes (*D. hansenii var. hansenii*, *D. hansenii* y *D. sp. MPSR02*) al contar con abundancias >10% y 1-10%, respectivamente, contribuyendo a la idea de que la especie *D. hansenii* forma parte de la microbiota del queso Cotija.

Siguiendo con las especies fúngicas asignadas con Kaiju, en total asignó 676 especies de las cuales dos fueron dominante con más de 10% de abundancia. *Rhizophagus sp.* MUCL 43196 contó con una abundancia de 11.09% siendo parte de la población dominante. Sin embargo, este hongo no se ha reportado previamente en productos lácteos sino se ha asociado a muestras vegetales y ambientales ya que forma parte del filo *Glomeromycota* de hongos micorrízicos. Este se asocia por simbiosis a las raíces de plantas terrestres y, facilita la absorción de nutrientes esenciales al actuar como extensiones de la raíz y promueve el crecimiento de las plantas (Morin *et al.*, 2019). La población subdominante estuvo representada por tres especies, *Debaryomyces fabryi*, *Mucor ambiguus* y *Lachancea nothofagi* con el 5.25%, 3.30 y 1.22% de abundancia respectivamente. El resto de las especies se encontraron pobremente representados al tener abundancias inferiores al 1% (Anexo V).

Por su parte, Kraken2 asignó menor cantidad de especies con un total de 58. De las cuales solo una especie conformó la población dominante (*D. hansenii*). La especie

Scheffersomyces stipitis, también denominada como *Pichia stipitis*, fue la única especie subdominante al contar con 2.16% de abundancia. Esta levadura ha sido muy estudiada debido a su capacidad de metabolizar azúcares de cinco carbonos como la D-xilosa y la L-arabinosa. Se ha utilizado en la industria para producir bioetanol a partir de residuos de cultivos como del maíz y de la caña de azúcar, y para la producción de ácido fumárico (Yang *et al.*, 2019). También se ha empleado como fuente de genes para la ingeniería del metabolismo de la xilosa en otros microorganismos (Papini *et al.*, 2012). A pesar de su metabolismo, se ha relacionado a quesos. La especie fue reportada por Murugesan *et al.* (2018) como abundante en distintos quesos mexicanos entre los que se encuentra un queso tipo Cotija siendo coherente con lo obtenido; sin embargo, su papel en los quesos aún no ha sido investigada. Esta especie también fue identificada por Kaiju pero con una abundancia inferior al 1%.

El resto de las especies identificadas con Kraken2 presentaron una abundancia inferior al 1% (Anexo V), entre estas se encontró a *K. lactis* con una abundancia de 0.80%, la cual también se ha reportado como parte de la microbiota de quesos y ambientes lácteos. Dentro del grupo de trabajo, esta especie fue identificada como parte de la población fúngica del queso Cotija (Martínez, 2009), por lo que la asignación de lecturas a esta especie concuerda con lo que se ha reportado anteriormente dentro del grupo.

Si analizamos las levaduras identificadas con Kraken2 y Kaiju, en el primero de ellos se describen más géneros que se han descrito como parte de la microbiota de quesos tradicionales, tal es el caso de *Kluyveromyces*, *Candida* y *Saccharomyces*. Mientras que con Kaiju, la mayoría de ellos han sido relacionados con muestras ambientales de suelo y árboles.

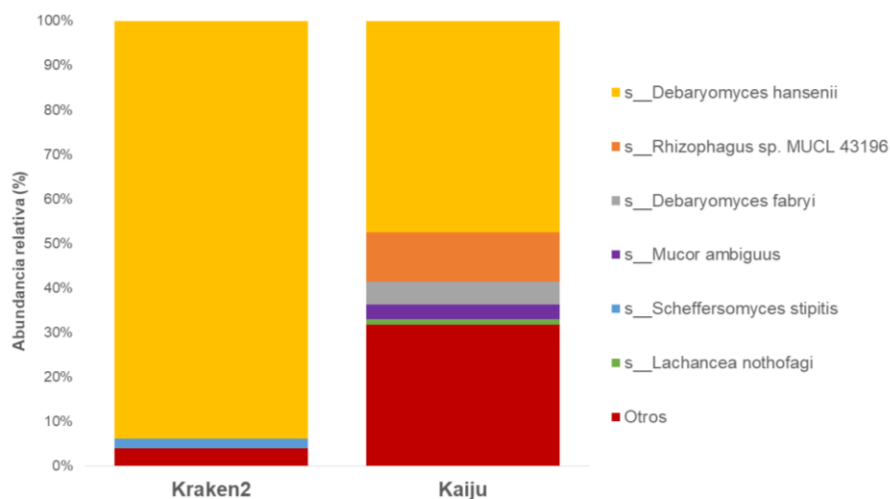


Figura 19. Principales especies anotadas a partir de los datos metagenómicos de *shotgun* del queso Cotija.

8.2.3.3 Ensamble de las lecturas con SPAdes

Las lecturas de *shotgun* filtradas por calidad se ensamblaron con el programa SPAdes cuyas estadísticas se muestran en la tabla 15. El número total de contigs ensamblados fue de 89 479 con una N50 de 3 903 pb y un L50 de 6 169. El N50 y L50 son dos parámetros utilizados para evaluar la calidad del ensamble, el primero hace referencia a la longitud del *contig* más corto con el que se alcanza el 50% del tamaño total del genoma. Por otro lado, L50 es el número de *contigs* con el que se alcanza el 50% del tamaño total del genoma. Se espera que la longitud del *contig* de N50 sea larga y el menor número de *contigs* para L50, garantizando que se ensamble y recupera la mayor proporción de los genomas presentes en la muestra (Gurevich *et al.*, 2013).

Tabla 15. Estadísticas del ensamble del metagenoma del queso Cotija con SPAdes

Número total de bases ensambladas (pb)	125 477 222
Número total de contigs	89 479
Tamaño del contig más largo (bp)	271 050
N50 (bp)	3 903
L50	6 169

8.2.3.4 Extracción de secuencias ITS e identificación

El programa ITSx extrajo solo una secuencia de las lecturas metagenómicas y la asignó al grupo de hongos y levaduras. Esta secuencia corresponde a la región ITS completa (tabla 16) con una longitud de 550 pb. Al alinearla con BLASTn se identificó como *D. hansenii* (figura 20), lo cual concuerda con la especie dominante asignada con Kraken2 y Kaiju.

Tabla 16. Secuencias identificadas como ITS

Región	Identificador	Posición (longitud)
ITS completa		---- (550 bp)
ITS1	>NODE_2247_length_8396_cov_1	3971 - 4168 (198 bp)
ITS2	4.663060 F	4327 - 4519 (193 bp)

Score	Expect	Identities	Gaps	Strand
1016 bits(550)	0.0	550/550(100%)	0/550(0%)	Plus/Plus
Query 1	CAGTATTC	TTTTGGCCAGCGCTTAATTGCGCGCGAAAAACCTTACACACAGTGT	TTTT	60
Sbjct 35	CAGTATTC	TTTTGGCCAGCGCTTAATTGCGCGCGAAAAACCTTACACACAGTGT	TTTT	94
Query 61	TGTATTACAAGAACTTTT	TGGCTGGACTAGAAATAGTTTGGGCCAGAGGTTTAC		120
Sbjct 95	TGTATTACAAGAACTTTT	TGGCTGGACTAGAAATAGTTTGGGCCAGAGGTTTAC		154
Query 121	TGAAC	TAACTTCAATATTTATATTGAATTGTTATTTAATTGTC	CAATTTGTTGATT	180
Sbjct 155	TGAAC	TAACTTCAATATTTATATTGAATTGTTATTTAATTGTC	CAATTTGTTGATT	214
Query 181	AAATTC	AAAAATCTTCAAACTTTCAACAACGGATCTCTGGTTCTCGCATCGATGAAG		240
Sbjct 215	AAATTC	AAAAATCTTCAAACTTTCAACAACGGATCTCTGGTTCTCGCATCGATGAAG		274
Query 241	AACGCAGCGAAATGCGATAAGTAATATGAATTGCAGATTTTCGTGAATCATCGAATCTTT			300
Sbjct 275	AACGCAGCGAAATGCGATAAGTAATATGAATTGCAGATTTTCGTGAATCATCGAATCTTT			334
Query 301	GAACGCACATTGCGCCCTCTGGTATTCAGAGGGCATGCCTGTTT	GAGCGTCATTTCTCT		360
Sbjct 335	GAACGCACATTGCGCCCTCTGGTATTCAGAGGGCATGCCTGTTT	GAGCGTCATTTCTCT		394
Query 361	CTCAAACCTTCGGGTTTGGTATTGAGTGATACTCTTAGTCGAACTAGGC	GTTTGCTTGAA		420
Sbjct 395	CTCAAACCTTCGGGTTTGGTATTGAGTGATACTCTTAGTCGAACTAGGC	GTTTGCTTGAA		454
Query 421	ATGTATTGGCATGAGTGGTACTGGATAGTGTATATGACTTTCAATGTATTAGGTTATC			480
Sbjct 455	ATGTATTGGCATGAGTGGTACTGGATAGTGTATATGACTTTCAATGTATTAGGTTATC			514
Query 481	CAACTCGTTGAATAGTTTAAATGGTATATTTCTCGGTATTCTAGGCTCGGCCTTACAATAT			540
Sbjct 515	CAACTCGTTGAATAGTTTAAATGGTATATTTCTCGGTATTCTAGGCTCGGCCTTACAATAT			574
Query 541	AACAACAAG		550	
Sbjct 575	AACAACAAG		584	

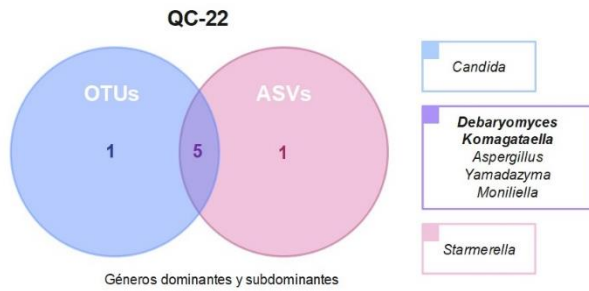
Figura 20. Secuencia ITS (Node 2247) alineada con *D. hansenii* con el número de acceso MT321273

8.2.4 Análisis metagenómico dirigido y *shotgun*

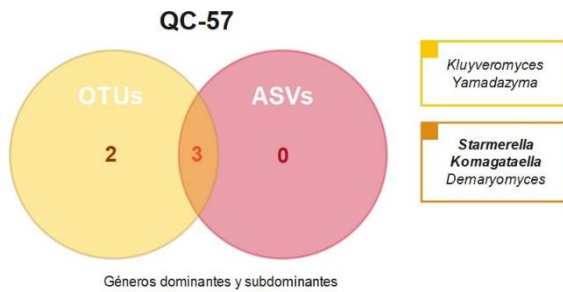
Al analizar las asignaciones de los dos tipos de datos metagenómicos, observamos que a pesar de la diferencia en la profundidad de secuenciación se logra recuperar un género en común, *Debaryomyces*, el cual como se menciona forma parte de la comunidad microbiana de quesos. Así mismo, *Komagataella*, otro de los géneros con mayor abundancia según los datos de amplicones, se rescató con Kaiju pero como parte de la comunidad pobremente representada. Los dos enfoques otorgan un buen acercamiento sobre la posible diversidad de levaduras presentes en el queso Cotija, sin embargo, no se pudo aprovechar la ventaja que tiene la metagenómica de *shotgun* debido a que originalmente el objetivo del proyecto que realizó la secuenciación era caracterizar la comunidad bacteriana y por lo cual se optó por una metodología de extracción de ADN diferente (utilizando enzimas que degradan la pared celular de bacterias, lisozima y mutanolisina), con lo que la proporción original de las levaduras en el queso no se pudo ver representada correctamente. Por lo cual, el enfoque basado en amplicones resulta una estrategia más eficiente para describir al grupo de levaduras en el queso Cotija al enfocarse en un gen marcador y, logró describir más géneros, aunque no se pudo discriminar entre especies dada la diferencia observada y el poder de resolución del marcador. Así mismo, el análisis de amplicones con los enfoques de OTUs o ASV identificaron los mismos géneros dominantes y subdominantes en el queso Cotija (figura 21) y por tanto son buenas estrategias de análisis para conocer la diversidad de levaduras en una muestra.

Datos metagenómicos de ITS1

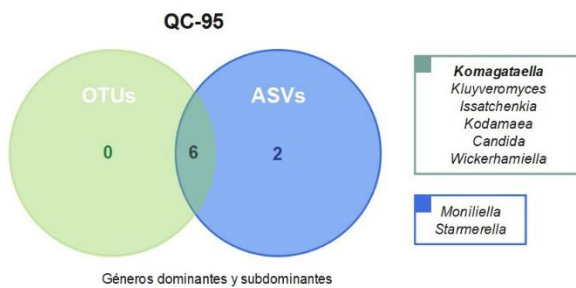
A



B

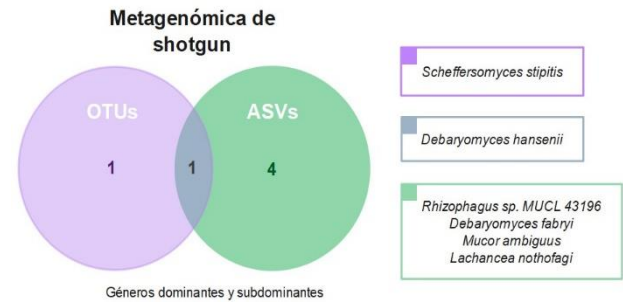


C



Datos metagenómicos de *shotgun*

D



Datos de ITS1 y *shotgun*

E

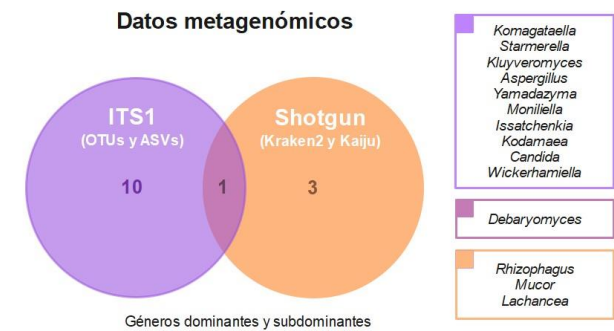


Figura 21. Diagramas de Venn de los géneros dominantes y subdominantes asignados por las estrategias de análisis sobre los datos metagenómicos de ITS1 y *shotgun*. **A** – QC-22; **B** - QC-57; **C** – QC-95.

9 Conclusiones

- La amplificación y secuenciación de la región ITS1 permitió identificar a las levaduras que forman parte de la microbiota del Queso Cotija Región de Origen, encontrando tres familias dominantes, *Phaffomycetaceae*, *Debaryomycetaceae* y *Saccharomycetaceae*, a nivel de género destacaron *Komagataella*, *Starmerella* y *Debaryomyces* al contar con más del 18% de abundancia. El primer de ellos se presentó como género en común entre las tres muestras analizadas.
- Los perfiles taxonómicos obtenidos por el análisis de las secuencias de ITS1 fueron distintos para cada una de las muestras analizadas, siendo más similares QC-22 y QC-95 que la muestra QC-57.
- La aplicación de dos estrategias de análisis bioinformático en los datos de ITS1, propuestas en este trabajo, permitió la anotación del mismo género de levaduras para cada una de las tres muestras analizadas, aunque a nivel de especie fueron distintas.
- Al analizar los datos metagenómicos de *shotgun* obtenidos en un proyecto anterior, se corroboró que la microbiota del Queso Cotija está dominada por bacterias *Firmicutes* y en menor proporción se encuentran las levaduras. Dentro de la fracción de levaduras identificadas por los programas Kraken2 y Kaiju, la familia *Debaryomycetaceae* fue la dominante; mientras que, a nivel de especie, *Debaryomyces hansenii* obtuvo el mayor porcentaje de abundancia con el 93.8% y 47.4%, respectivamente.
- La búsqueda y extracción de posibles secuencias de ITS, con el programa ITSx a partir de los *contigs* ensamblados, condujo a la identificación de una secuencia completa de ITS asignada a la levadura *D. hansenii* con BLAST.
- Al comparar los perfiles taxonómicos obtenidos por el análisis de las secuencias de ITS1 con ambas estrategias y la fracción de levaduras identificadas de los datos de *shotgun*, observamos que coinciden en la presencia del género *Debaryomyces* como parte de dicho grupo en el Queso Cotija Región de Origen. Así mismo, la metagenómica con secuenciación dirigida a la región ITS1 permitió conocer más géneros de levaduras, aunque no se pudo discriminar entre especies dada la diferencia observada por las estrategias utilizadas.

10 Perspectivas

Algunas de las actividades derivadas de este análisis y que quedaron fuera de los alcances de este estudio pero que resultan de interés son:

- Enfocarse en el aislamiento de los géneros identificados con las lecturas de ITS1, especialmente los dominantes. Probar medios de cultivo específicos para levaduras halófilas debido a que los géneros identificados tenían esta característica.
- Evaluar la presencia de hongos filamentosos, en particular del género *Aspergillus* dada su importancia en alimentos y clínica.
- Evaluar el efecto del filtrado sobre las lecturas recuperadas (parámetros maxEE y truncQ) y sobre el perfil taxonómico inferido según la estrategia por ASVs.
- Evaluar la contribución de la contaminación presente en los reactivos en la diversidad microbiana descrita.

Referencias

- Abarenkov, K., Zirk, A., Piirmann, T., Pöhönen, R., Ivanov, F., Nilsson, R.H. and Kõljalg, U. (2022a). Full UNITE+INSD dataset for Fungi. Version 16.10.2022. UNITE Community. <https://unite.ut.ee/repository.php> (último acceso 28 febrero 2023).
- Abarenkov, K., Zirk, A., Piirmann, T., Pöhönen, R., Ivanov, F., Nilsson, R. H. and Kõljalg, U. (2022b). UNITE general FASTA release for Fungi 2. Version 16.10.2022. UNITE Community. <https://unite.ut.ee/repository.php> (último acceso 28 febrero 2023).
- Álvarez, B., Barragán, L. and Chombo, M. (2005). Reglas de uso. Marca colectiva (Queso Cotija Región de Origen). México. INDAUTOR: 03-2004-090812435800-01.
- Atanassova, M. R., Fernández-Otero, C., Rodríguez-Alonso, P., Fernández-No, I. C., Garabal, J. I., and Centeno, J. A. (2016). Characterization of yeasts isolated from artisanal short-ripened cows' cheeses produced in Galicia (NW Spain). *Food microbiology*, 53(Pt B), 172–181. <https://doi.org/10.1016/j.fm.2015.09.012>.
- Avchar, R., Groenewald, M., and Baghela, A. (2019). *Wickerhamiella shivajii* sp. nov., a thermotolerant yeast isolated from distillery effluent. *International Journal of Systematic and Evolutionary Microbiology*, 69(10), 3262–3267. <https://doi.org/10.1099/ijsem.0.003616>.
- Badotti, F., de Oliveira, F. S., García, C. F., Vaz, A. B., Fonseca, P. L., Nahum, L. A., Oliveira, G., and Góes-Neto, A. (2017). Effectiveness of ITS and sub-regions as DNA barcode markers for the identification of Basidiomycota (Fungi). *BioMed Central Microbiology*, 17(1), 42. <https://doi.org/10.1186/s12866-017-0958-x>.
- Baltoumas, F. A., Karatzas, E., Paez-Espino, D., Venetsianou, N. K., Aplakidou, E., Oulas, A., Finn, R. D., Ovchinnikov, S., Pafilis, E., Kyrpides, N. C., and Pavlopoulos, G. A. (2023). Exploring microbial functional biodiversity at the protein family level-From metagenomic sequence reads to annotated protein clusters. *Frontiers in Bioinformatics*, 3, 1157956. <https://doi.org/10.3389/fbinf.2023.1157956>.
- Banjara, N., Suhr, M.J. and Hallen-Adams H.E. (2015). Diversity of yeast and mold species from a variety of cheese types. *Current Microbiol.*, 70(6): 792-800. <https://doi.org/10.1007/s00284-015-0790-1>.
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Pribelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., and 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. <https://doi.org/10.1089/cmb.2012.0021>.
- Bálint, M., Schmidt, P. A., Sharma, R., Thines, M., and Schmitt, I. (2014). An Illumina metabarcoding pipeline for fungi. *Ecology and Evolution*, 4(13), 2642–2653. <https://doi.org/10.1002/ece3.1107>.
- Bassi, D., Gazzola, S., Sattin, E., Dal Bello, F., Simionati, B., and Cocconcelli, P. S. (2020). Lactic Acid Bacteria Adjunct Cultures Exert a Mitigation Effect against Spoilage Microbiota in Fresh Cheese. *Microorganisms*, 8(8), 1199. <https://doi.org/10.3390/microorganisms8081199>.
- Belloch, C., Querol, A. and E. Barrio (2011). Yeasts and Molds | *Kluyveromyces* spp. In Fuquay, J. W. (Ed.), *Encyclopedia of Dairy Sciences* (2nd Ed.), Academic Press, 754-764. <https://doi.org/10.1016/B978-0-12-374407-4.00499-4>.
- Bengtsson-Palme, J., Veldre, V., Ryberg, M., Hartmann, M., Branco, S., Wang, Z., Godhe, A., Bertrand, Y., De Wit, P., Sanchez, M., Ebersberger, I., Sanli, K., de Souza, F., Kristiansson, E., Abarenkov, K., Eriksson, K.M. and Nilsson R.H. (2013). Improved software detection and extraction of ITS1 and ITS2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data. *Methods in Ecology and Evolution*, 4: 914-919. <https://doi.org/10.1111/2041-210X.12073>.

- Beresford, T., Fitsimons, N., Brennan, N. and Cogan, T. (2001). Recent advances in cheese microbiology. *International Dairy Journal*, 11: 259-274. [https://doi.org/10.1016/S0958-6946\(01\)00056-5](https://doi.org/10.1016/S0958-6946(01)00056-5).
- Beresford, T and Williams, A. (2004). The Microbiology of Cheese Ripening. In P.F. Fox, P.L.H. McSweeney, T. M. Cogan, T. P. Guinee (Ed), *Cheese: Chemistry, Physics and Microbiology*, Academic Press, 287-317. [https://doi.org/10.1016/S1874-558X\(04\)80071-X](https://doi.org/10.1016/S1874-558X(04)80071-X).
- Bintsis, T. (2018). Lactic acid bacteria as starter cultures: An update in their metabolism and genetics. *AIMS Microbiology*, 4(4), 665–684. <https://doi.org/10.3934/microbiol.2018.4.665>.
- Bjørnsgaard, A., Davey, M. L., and Kauserud, H. (2017). ITS all right mama: investigating the formation of chimeric sequences in the ITS2 region by DNA metabarcoding analyses of fungal mock communities of different complexities. *Molecular Ecology Resources*, 17(4), 730–741. <https://doi.org/10.1111/1755-0998.12622>.
- Bridge, P.D, Spooner, B.M and Roberts P.J (2005). The impact of molecular data in fungal systematics. *Advances in Botanical Research*, 42: 33-67. [https://doi.org/10.1016/S0065-2296\(05\)42002-9](https://doi.org/10.1016/S0065-2296(05)42002-9).
- Bolumar, T., Sanz, Y., Aristoy, M. C., and Toldrá, F. (2008). Purification and characterization of proteases A and D from *Debaryomyces hansenii*. *International Journal of Food Microbiology*, 124(2), 135–141. <https://doi.org/10.1016/j.ijfoodmicro.2008.03.001>.
- Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J., and Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, 13(7), 581–583. <https://doi.org/10.1038/nmeth.3869>.
- Calle, M. L. (2019). Statistical Analysis of Metagenomics Data. *Genomics & Informatics*, 17(1), e6. <https://doi.org/10.5808/GI.2019.17.1.e6>.
- Cameron, E. S., Schmidt, P. J., Tremblay, B. J., Emelko, M. B., and Müller, K. M. (2021). Enhancing diversity analysis by repeatedly rarefying next generation sequencing data describing microbial communities. *Scientific Reports*, 11(1), 22302. <https://doi.org/10.1038/s41598-021-01636-1>.
- Čandek, K., and Kuntner, M. (2015). DNA barcoding gap: reliable species identification over morphological and geographical scales. *Molecular Ecology Resources*, 15(2), 268–277. <https://doi.org/10.1111/1755-0998.12304>.
- Chao, A., and Chiu, C. (2016). Species Richness: Estimation and Comparison. *Wiley Statistics Reference Online*, 1-26. <https://doi.org/10.1002/9781118445112.stat03432.pub2>.
- Chen, A. J., Hubka, V., Frisvad, J. C., Visagie, C. M., Houbraeken, J., Meijer, M., Varga, J., Demirel, R., Jurjević, Ž., Kubátová, A., Sklenář, F., Zhou, Y. G., and Samson, R. A. (2017). Polyphasic taxonomy of *Aspergillus* section *Aspergillus* (formerly *Eurotium*), and its occurrence in indoor environments and food. *Studies in Mycology*, 88, 37–135. <https://doi.org/10.1016/j.simyco.2017.07.001>.
- Chiarello, M., McCauley, M., Villéger, S., and Jackson, C. R. (2022). Ranking the biases: The choice of OTUs vs. ASVs in 16S rRNA amplicon data analysis has stronger effects on diversity measures than rarefaction and OTU identity threshold. *PloS One*, 17(2), e0264443. <https://doi.org/10.1371/journal.pone.0264443>.
- Chombo-Morales, P., Kirchmayr, M.R., Gschaedler, A., Lugo-Cervantes, E., and Villanueva-Rodríguez, S.J. (2016). Effects of controlling ripening conditions on the dynamics of the native microbial population of Mexican artisanal Cotija cheese assessed by PCR-DGGE. *LWT Food Science and Technology*, 65, 1153-1161. <https://doi.org/10.1016/j.lwt.2015.09.044>.
- Ciani, M., Canonico, L., and Comitini, F. (2022). Chapter 10 - Improving white wine aroma and structure by non-*Saccharomyces* yeasts. In Morata, A. (Ed), *White Wine Technology*, Academic Press, 117-130. <https://doi.org/10.1016/B978-0-12-823497-6.00020-X>.

- de Vega, C., Albaladejo, R. G., Guzmán, B., Steenhuisen, S. L., Johnson, S. D., Herrera, C. M., and Lachance, M. A. (2017). Flowers as a reservoir of yeast diversity: description of *Wickerhamiella nectarea* f.a. sp. nov., and *Wickerhamiella natalensis* f.a. sp. nov. from South African flowers and pollinators, and transfer of related *Candida* species to the genus *Wickerhamiella* as new combinations. *FEMS Yeast Research*, 17(5). <https://doi.org/10.1093/femsyr/fox054>.
- Czurda, S., Smelik, S., Preuner-Stix, S., Nogueira, F., & Lion, T. (2016). Occurrence of Fungal DNA Contamination in PCR Reagents: Approaches to Control and Decontamination. *Journal of Clinical Microbiology*, 54(1), 148–152. <https://doi.org/10.1128/JCM.02112-15>.
- Dröge, J., and McHardy, A. C. (2012). Taxonomic binning of metagenome samples generated by next-generation sequencing technologies. *Briefings in Bioinformatics*, 13(6), 646–655. <https://doi.org/10.1093/bib/bbs031>.
- Dugat-Bony, E., Straub, C., Teissandier, A., Onésime, D., Loux, V., Monnet, C., Irlinger, F., Landaud, S., Leclercq-Perlat, M. N., Bento, P., Fraud, S., Gibrat, J. F., Aubert, J., Fer, F., Guédon, E., Pons, N., Kennedy, S., Beckerich, J. M., Swennen, D., & Bonnarne, P. (2015). Overview of a surface-ripened cheese community functioning by meta-omics analyses. *PLOS One*, 10(4), e0124360. <https://doi.org/10.1371/journal.pone.0124360>.
- Edgar, R. C. (2017). Accuracy of microbial community diversity estimated by closed- and open-reference OTUs. *PeerJ Journals*, 5, e3889. <https://doi.org/10.7717/peerj.3889>.
- Edgar, R. C. (s.f.) Rarefaction. <http://drive5.com/usearch/manual/rare.html> [Último acceso 20 agosto 2023].
- Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*, 27(16), 2194–2200. <https://doi.org/10.1093/bioinformatics/btr381>.
- Ercolini, D., De Filippis, F., La Storia, A., and Iacono, M. (2012). "Remake" by high-throughput sequencing of the microbiota involved in the production of water buffalo mozzarella cheese. *Applied and Environmental Microbiology*, 78(22), 8142–8145. <https://doi.org/10.1128/AEM.02218-12>.
- Escalante, A. (2007). La ecología molecular en el estudio de las comunidades bacterianas. In: Eguiarte, L.E., Souza, V., Aguirre, X. (Eds.), *Ecología molecular*. Secretaría de Medio Ambiente y Recursos Naturales, Mexico, pp. 393–425.
- Escobar, G.A. (2012). Estudio de la distribución espacial de la microbiota bacteriana en queso Cotija por la técnica de FISH. Tesis de maestría. Programa de Maestría y Doctorado en Ciencias Bioquímicas. UNAM.
- Escobar, G.A. (2016). Análisis taxonómico y del potencial metabólico del metagenoma de queso Cotija. Tesis de doctorado. Programa de Maestría y Doctorado en Ciencias Bioquímicas. UNAM.
- Fallah, B., Zaini, F., Daei Ghazvini, R., Kachuei, R., Kordbacheh, P., Safara, M., and Mahmoudi, S. (2016). The antagonistic effects of *Candida parapsilosis* on the growth of *Fusarium* species and fumonisin production. *Current Medical Mycology*, 2(1), 1–6. <https://doi.org/10.18869/acadpub.cmm.2.1.1>.
- Ferreira, A. D., and Viljoen, B. C. (2003). Yeasts as adjunct starters in matured Cheddar cheese. *International Journal of Food Microbiology*, 86(1-2), 131–140. [https://doi.org/10.1016/s0168-1605\(03\)00252-6](https://doi.org/10.1016/s0168-1605(03)00252-6).
- Feng, J., Zhan, X. B., Wang, D., Zhang, L. M., and Lin, C. C. (2012). Identification and analysis of the metabolic functions of a high-salt-tolerant halophilic aromatic yeast *Candida etchellsii* for soy sauce production. *World Journal of Microbiology and Biotechnology*, 28(4), 1451–1458. <https://doi.org/10.1007/s11274-011-0945-6>.

- Filippis, F.D., Stora, A.L., Stellato, G., Gatti, M. and Ercolini, D. (2014). A Selected Core Microbiome Drives the Early Stages of Three Popular Italian cheese Manufactures. *PLOS ONE*, 9, e89680. <https://doi.org/10.1371/journal.pone.0089680>.
- Fox, P. F., Guinee, T. P., Cogan, T. M., and McSweeney, P. L. (2017). *Fundamentals of cheese science*, Springer US., 121-183. <https://doi.org/10.1007/978-1-4899-7681-9>.
- Fedor, P and M. Zvaríková (2019). Biodiversity Indices. In Fath, B (Ed), *Encyclopedia of Ecology* (2nd Ed.), 1: 337-346. <https://doi.org/10.1016/B978-0-12-409548-9.10558-5>.
- Fröhlich-Wyder, M. T., Arias-Roth, E., & Jakob, E. (2019). Cheese yeasts. *Yeast*, 36(3), 129–141. <https://doi.org/10.1002/yea.3368>.
- Fuller, C. W., Middendorf, L. R., Benner, S. A., Church, G. M., Harris, T., Huang, X., Jovanovich, S. B., Nelson, J. R., Schloss, J. A., Schwartz, D. C., and Vezenov, D. V. (2009). The challenges of sequencing by synthesis. *Nature Biotechnology*, 27(11), 1013–1023. <https://doi.org/10.1038/nbt.1585>.
- Gao, W. L., Li, Y., Chai, C. Y., Yan, Z. L., and Hui, F. L. (2021). New species of *Yamadazyma* from rotting wood in China. *MycKeys*, 83, 69–84. <https://doi.org/10.3897/mycokeys.83.71156>.
- Gao, W. L., Liu, T. T., Zheng, J., & Hui, F. L. (2017). *Kodamaea neixiangensis* f.a., sp. nov. and *Kodamaea jinghongensis* f.a., sp. nov., two yeast species isolated from rotting wood. *International Journal of Systematic and Evolutionary Microbiology*, 67(9), 3358–3362. <https://doi.org/10.1099/ijsem.0.002117>.
- Gardini, F., Tofalo, R., Belletti, N., Iucci, L., Suzzi, G., Torriani, S., Guerzoni, M. E., and Lanciotti, R. (2006). Characterization of yeasts involved in the ripening of Pecorino Crotonese cheese. *Food Microbiology*, 23(7), 641–648. <https://doi.org/10.1016/j.fm.2005.12.005>.
- García, S.V. (2001). Detección de microorganismos lipolíticos durante las etapas de elaboración de un queso artesanal mediante técnicas moleculares. Tesis de maestría. Programa de Maestría y Doctorado en Ciencias Bioquímicas. UNAM.
- García, S.V. (2006). Aislamiento de microorganismos con mayor actividad lipolítica del queso Cotija. Tesis de licenciatura. Facultad de Química. UNAM.
- Garnier, L., Valence, F., and Mounier, J. (2017). Diversity and Control of Spoilage Fungi in Dairy Products: An Update. *Microorganisms*, 5(3), 42. <https://doi.org/10.3390/microorganisms5030042>.
- Glassman, S. I., and Martiny, J. B. H. (2018). Broad-scale Ecological Patterns Are Robust to Use of Exact Sequence Variants versus Operational Taxonomic Units. *mSphere*, 3(4), e00148-18. <https://doi.org/10.1128/mSphere.00148-18>.
- Goodrich, J. K., Di Rienzi, S. C., Poole, A. C., Koren, O., Walters, W. A., Caporaso, J. G., Knight, R., and Ley, R. E. (2014). Conducting a microbiome study. *Cell*, 158(2), 250–262. <https://doi.org/10.1016/j.cell.2014.06.037>.
- Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics*, 29(8), 1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>.
- Hebert P. D. N., Cywinska A., Ball S. L., deWaard J. R. (2003). Biological identifications through DNA barcodes. *Proceedings for the Royal Society B* 270: 313–321. <https://doi.org/10.1098/rspb.2002.2218>.
- Heisteringer, L., Dohm, J. C., Paes, B. G., Koizar, D., Troyer, C., Ata, Ö., Steininger-Mairinger, T., and Mattanovich, D. (2022). Genotypic and phenotypic diversity among *Komagataella* species reveals a hidden pathway for xylose utilization. *Microbial Cell Factories*, 21(1), 70. <https://doi.org/10.1186/s12934-022-01796-3>.
- Hernández, B. (2007). Queso Cotija: Estudio del análisis fisicoquímico, proximal y actividad antioxidante. Tesis de licenciatura. Facultad de Química. UNAM.

- Hui, F., Ren, Y., Chen, L., Li, Y., Zhang, L., and Niu, Q. (2014). Molecular phylogenetic analysis reveals the new genus *Hemisphaericaspora* of the family Debaryomycetaceae. *PLOS One*, 9(7), e103737. <https://doi.org/10.1371/journal.pone.0103737>.
- Illumina (2021). Fungal Metagenomic Sequencing Demonstrated Protocol. <https://support.illumina.com/downloads/fungal-metagenomic-sequencing-demonstrated-protocol-1000000064940.html> [Último acceso 17 mayo 2021].
- Jacques, N., and Casaregola, S. (2008). Safety assessment of dairy microorganisms: the hemiascomycetous yeasts. *International Journal of Food Microbiology*, 126(3), 321–326. <https://doi.org/10.1016/j.ijfoodmicro.2007.08.020>.
- Jacques, N., Zenouche, A., Gunde-Cimerman, N., and Casaregola, S. (2015). Increased diversity in the genus *Debaryomyces* from Arctic glacier samples. *Antonie van Leeuwenhoek*, 107(2), 487–501. <https://doi.org/10.1007/s10482-014-0345-7>.
- Joint Genome Institute (2022). *Moniliella* sp. MCA 3643 v1.0. <https://mycocosm.jgi.doe.gov/Monili2/Monili2.home.html> [Último acceso 18 noviembre 2022].
- Johnson, E. and Echavarri-Erasun, C. (2011). Chapter 3 – Yeast Biotechnology. In Kurtzman, C. P., Fell, J. W. y T. Boekhout (Eds), *The Yeasts* (5 ed), Elsevier, 21-44. <https://doi.org/10.1016/B978-0-444-52149-1.00003-3>.
- Kawahata, M., Fujii, T., and Iefuji, H. (2007). Intraspecies diversity of the industrial yeast strains *Saccharomyces cerevisiae* and *Saccharomyces pastorianus* based on analysis of the sequences of the internal transcribed spacer (ITS) regions and the D1/D2 region of 26S rDNA. *Bioscience, Biotechnology, and Biochemistry*, 71(7), 1616–1620. <https://doi.org/10.1271/bbb.60673>.
- Krebs, C. (2014). Chapter 13 – Species Diversity Measures. In: *Ecological Methodology*. Addison-Wesley Educational Publishers, Inc., Boston, 532 – 593.
- Kress, W. J., and Erickson, D. L. (2008). DNA barcodes: genes, genomics, and bioinformatics. *Proceedings of the National Academy of Sciences*, 105(8), 2761–2762. <https://doi.org/10.1073/pnas.0800476105>.
- Kochoski, L.J., Kalevska, T. and Trajkovska, B. (2011). Influence of somatic cell count in raw milk on cheese production. *Agrár- és Vidékfejlesztési Szemle*, 6 (1). pp. 18-23.
- Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., and Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and Environmental Microbiology*, 79(17), 5112–5120. <https://doi.org/10.1128/AEM.01043-13>.
- Kurtzman, C. P. and Robnett, C. J. (2013). Relationships among genera of the *Saccharomycotina* (Ascomycota) from multigene phylogenetic analysis of type species. *FEMS Yeast Research*, 13(1), 23–33. <https://doi.org/10.1111/1567-1364.12006>.
- Lahti, L. (2017). Tools for microbiome analysis in R. *Microbiome* package version. <http://microbiome.github.io/microbiome> [Último acceso 20 agosto 2023].
- Leclercq-Perlat, M. N., Corrieu, G., and Spinnler, H. E. (2004). Comparison of volatile compounds produced in model cheese medium deacidified by *Debaryomyces hansenii* or *Kluyveromyces marxianus*. *Journal of Dairy Science*, 87(5), 1545–1550. [https://doi.org/10.3168/jds.S0022-0302\(04\)73306-8](https://doi.org/10.3168/jds.S0022-0302(04)73306-8).
- Lopandic, K., Zelger, S., Bánszky, L. K., Eliskases-Lechner, F., and Prillinger, H. (2006). Identification of yeasts associated with milk products using traditional and molecular techniques. *Food Microbiology*, 23(4), 341–350. <https://doi.org/10.1016/j.fm.2005.05.001>.
- Lu, Y., Yang, L., An, Y., Liu, D., Yang, G., and He, Q. (2022). Salt tolerance and ester production mechanisms of *Candida etchellsii* in Chinese horse bean-chili-paste. *LWT Food Science and Technology*. 167, 113806. <https://doi.org/10.1016/j.lwt.2022.113806>.

- Mansour, S., Beckerich, J. M., and Bonnarne, P. (2008). Lactate and amino acid catabolism in the cheese-ripening yeast *Yarrowia lipolytica*. *Applied and Environmental Microbiology*, 74(21), 6505–6512. <https://doi.org/10.1128/AEM.01519-08>.
- Martínez, P.D. (2009). Identificación de levaduras en el queso cotija por métodos moleculares dependiente e independiente de cultivo (ARDRA, RFLP y DGGE). Tesis de maestría. Programa de Maestría y Doctorado en Ciencias Bioquímicas. UNAM.
- Martorell, P., Fernández-Espinar, M. T., and Querol, A. (2005). Sequence-based identification of species belonging to the genus *Debaryomyces*. *FEMS Yeast Research*, 5(12), 1157–1165. <https://doi.org/10.1016/j.femsyr.2005.05.002>.
- Masella, A.P., Bartram, A.K., Truszkowski, J.M., Brown, D.G. and Neufeld, J.D. (2012). PANDAseq: paired-end assembler for illumina sequences. *BMC Bioinformatics* 13, 31. <https://doi.org/10.1186/1471-2105-13-31>.
- Meyer, C. P. and Paulay, G. (2005). DNA barcoding: error rates based on comprehensive sampling. *PLOS Biology*, 3(12), e422. <http://doi.org/10.1371/journal.pbio.0030422>.
- Mbareche, H., Veillette, M., Bilodeau, G., and Duchaine, C. (2020). Comparison of the performance of ITS1 and ITS2 as barcodes in amplicon-based sequencing of bioaerosols. *PeerJ Journals*, 8, e8523. <https://doi.org/10.7717/peerj.8523>.
- McMurdie, PJ and Holmes, S. (2013). phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLOS ONE*, 8(4), e61217. <http://dx.plos.org/10.1371/journal.pone.0061217>.
- Melike, Y., Tamer, T., Bulent, C. and Zulal, K. (2021). Microbiological characteristics and identification of yeast microbiota of traditional mouldy civil cheese, *International Dairy Journal*, 116. <https://doi.org/10.1016/B978-0-12-818766-1.00272-5>.
- Menzel, P., Ng, K. and Krogh, A. (2016). Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nature Communications* 7, 11257. <https://doi.org/10.1038/ncomms11257>.
- Miao, Y., Xiong, G., Li, R., Wu, Z., Zhang, X., and Weng, P. (2018). Transcriptome profiling of *Issatchenkia orientalis* under ethanol stress. *AMB Express*, 8(1), 39. <https://doi.org/10.1186/s13568-018-0568-5>.
- Michailidou, S., Pavlou, E., Pasentsis, K., Rhoades, J., Likotrafiti, E., and Argiriou, A. (2021). Microbial profiles of Greek PDO cheeses assessed with amplicon metabarcoding. *Food Microbiology*, 99, 103836. <https://doi.org/10.1016/j.fm.2021.103836>.
- Miyajima, Y., Satoh, K., Umeda, Y., and Makimura, K. (2009). Quantitation of fungal DNA contamination in commercial zymolyase and lyticase used in the preparation of fungi. *Japanese Journal of Medical Mycology*, 50(4), 259–262. <https://doi.org/10.3314/jjmm.50.259>.
- Monard, C., Gantner, S., and Stenlid, J. (2013). Utilizing ITS1 and ITS2 to study environmental fungal diversity using pyrosequencing. *FEMS Microbiology Ecology*, 84(1), 165–175. <https://doi.org/10.1111/1574-6941.12046>.
- Morin, E., Miyauchi, S., San Clemente, H., Chen, E. C. H., Pelin, A., de la Providencia, I., Ndikumana, S., Beaudet, D., Hainaut, M., Drula, E., Kuo, A., Tang, N., Roy, S., Viala, J., Henrissat, B., Grigoriev, I. V., Corradi, N., Roux, C., and Martin, F. M. (2019). Comparative genomics of *Rhizophagus irregularis*, *R. cerebriforme*, *R. diaphanus* and *Gigaspora rosea* highlights specific genetic features in Glomeromycotina. *The New Phytologist*, 222(3), 1584–1598. <https://doi.org/10.1111/nph.15687>.
- Murugesan, S., Reyes-Mata, M. P., Nirmalkar, K., Chavez-Carbajal, A., Juárez-Hernández, J. I., Torres-Gómez, R. E., Piña-Escobedo, A., Maya, O., Hoyo-Vadillo, C., Ramos-Ramírez, E. G., Salazar-Montoya, J. A., and García-Mena, J. (2018). Profiling of bacterial and fungal communities of Mexican cheeses by high throughput DNA sequencing. *Food Research International*, 113, 371–381. <https://doi.org/10.1016/j.foodres.2018.07.023>.

- Nalbantoglu, U., Cakar, A., Dogan, H., Abaci, N., Ustek, D. and Sayood, K. (2014). Metagenomic analysis of the microbial community in kefir grains. *Food Microbiology*, 41: 42-51. <https://doi.org/10.1016/j.fm.2014.01.014>.
- Nearing, J. T., Douglas, G. M., Comeau, A. M., and Langille, M. G. I. (2018). Denoising the Denoisers: an independent evaluation of microbiome sequence error-correction approaches. *PeerJ Journals*, 6, e5364. <https://doi.org/10.7717/peerj.5364>.
- Niknejad, F., Zaini, F., Faramarzi, M., Amini, M., Kordbacheh, P., Mahmoudi, M., and Safara, M. (2012). *Candida parapsilosis* as a Potent Biocontrol Agent against Growth and Aflatoxin Production by *Aspergillus* Species. *Iranian Journal of Public Health*, 41(10), 72–80. PMID: PMC3494234.
- Nilsson, R. H., Kristiansson, E., Ryberg, M., Hallenberg, N. and Larsson, K-H (2008). Intraspecific ITS variability in the kingdom fungi as expressed in the international sequence databases and its implications for molecular species identification. *Evolutionary Bioinformatics*, 4: 193 – 201. <https://doi.org/10.4137/EBO.S653>.
- Nilsson, R. H., V. Veldre, M. Hartmann, M. Unterseher, A. Amend, J. Bergsten, Kristiansson, E., Ryberg, M., Jumpponen, A. and Abarenkov, K. (2010). An open source software package for automated extraction of ITS1 and ITS2 from fungal ITS sequences for use in high-throughput community assays and molecular ecology. *Fungal Ecology*, 3: 284– 287. <https://doi.org/10.1016/j.funeco.2010.05.002>.
- Ndoye, B., Rasolofoa, E., LaPointe, G. and Roy, D. (2011). A review of the molecular approaches to investigate the diversity and activity of cheese microbiota. *Dairy Science and Technology*, 91:495–524. <https://doi.org/10.1007/s13594-011-0031-8>.
- NMX-F-735-COFOCALEC, 2011. Norma Mexicana. Sistema producto Leche-Alimentos-Lácteos-Alimento Lácteo Regional - Queso Cotija Artesanal Madurado- Denominación, especificaciones y Métodos de Prueba. 24 de junio del 2011. DOF.
- NOM-243-SSA1, 2010. NORMA Oficial Mexicana. Productos y servicios. Leche, fórmula láctea, producto lácteo combinado y derivados lácteos. Disposiciones y especificaciones sanitarias. Métodos de prueba. 27 de septiembre de 2010. DOF.
- Ouadghiri, M., Amar, M., Heide-Marie, D., Swings, J. and Vandamme, P. (2014). Yeasts in traditional Moroccan goat cheese. *African Journal of Microbiology Research*. 8 (32): 3029-3037. <https://doi.org/10.5897/AJMR2013.6277>.
- Oulas, A., Pavludi, C., Polymenakou, P., Pavlopoulos, G. A., Papanikolaou, N., Kotoulas, G., Arvanitidis, C., and Iliopoulos, I. (2015). Metagenomics: tools and insights for analyzing next-generation sequencing data derived from biodiversity studies. *Bioinformatics and Biology Insights*, 9, 75–88. <https://doi.org/10.4137/BBI.S12462>.
- Papini, M., Nookaew, I., Uhlén, M., and Nielsen, J. (2012). *Scheffersomyces stipitis*: a comparative systems biology study with the Crabtree positive yeast *Saccharomyces cerevisiae*. *Microbial Cell Factories*, 11, 136. <https://doi.org/10.1186/1475-2859-11-136>.
- Pinto, R. (2013). UF1284: Yogures, leches fermentadas y pastas untadas. Ed. 1. España, IC Editorial, 214.
- Prijbelski, A., Antipov, D., Meleshko, D., Lapidus, A., and Korobeynikov, A. (2020). Using SPAdes De Novo Assembler. *Current Protocols in Bioinformatics*, 70(1), e102. <https://doi.org/10.1002/cpbi.102>.
- Prodan, A., Tremaroli, V., Brolin, H., Zwinderman, A. H., Nieuwdorp, M., and Levin, E. (2020). Comparing bioinformatic pipelines for microbial 16S rRNA amplicon sequencing. *PLOS One*, 15(1), e0227434. <https://doi.org/10.1371/journal.pone.0227434>.
- PROY-NMX-F-700-COFOCALEC-2012, Sistema producto leche – Alimento - lácteo - leche cruda de vaca – Especificaciones fisicoquímicas, Sanitarias y métodos de prueba. 04 de mayo del 2019. DOF.

- Reitmeier, S., Hitch, T.C.A., Treichel, N., Fikas, N., Hausmann, B., Ramer-Tair, A.E., Neuhaus, K., Berry, D., Haller, D. y Lagkouvardos, I. (2021) Handling of spurious sequences affects the outcome of high-throughput 16S rRNA gene amplicon profiling. *ISME Communications*, 1, 31. <https://doi.org/10.1038/s43705-021-00033-z>.
- Rossi, F. and S. Torriani (2018). *Debaryomyces*. In D. Lui (Ed.), *Handbook of Foodborne Diseases*, CRC Press Taylor and Francis Group. 505-512
- Sacerdot, C., Casaregola, S., Lafontaine, I., Tekaiia, F., Dujon, B., and Ozier-Kalogeropoulos, O. (2008). Promiscuous DNA in the nuclear genomes of hemiascomycetous yeasts. *FEMS Yeast Research*, 8(6), 846–857. <https://doi.org/10.1111/j.1567-1364.2008.00409.x>.
- Santos, A. R. O., Leon, M. P., Barros, K. O., Freitas, L. F. D., Hughes, A. F. S., Morais, P. B., Lachance, M. A., and Rosa, C. A. (2018). *Starmerella camargoi* f.a., sp. nov., *Starmerella ilheusensis* f.a., sp. nov., *Starmerella litoralis* f.a., sp. nov., *Starmerella opuntiae* f.a., sp. nov., *Starmerella roubikii* f.a., sp. nov. and *Starmerella vitae* f.a., sp. nov., isolated from flowers and bees, and transfer of related *Candida* species to the genus *Starmerella* as new combinations. *International Journal of Systematic and Evolutionary Microbiology*, 68(4), 1333–1343. <https://doi.org/10.1099/ijsem.0.002675>.
- Sessou, P., Keisam, S., Tuikhar, N., Gagara, M., Farougou, S., and Jeyaram, K. (2019). High-Throughput Illumina MiSeq Amplicon Sequencing of Yeast Communities Associated With Indigenous Dairy Products From Republics of Benin and Niger. *Frontiers in Microbiology*, 10, 594. <https://doi.org/10.3389/fmicb.2019.00594>.
- Schauer, F., and Hanschke, R. (1999). Zur Taxonomie und Ökologie der Gattung *Candida* [Taxonomy and ecology of the genus *Candida*]. *Mycoses*, 42 Suppl 1, 12–21. <https://doi.org/10.1111/j.1439-0507.1999.tb04521.x>.
- Schloss, P.D., Gevers, D. and Westcott, S.L. (2011). Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. *PLOS One*, 6, e27310. <https://doi.org/10.1371/journal.pone.0027310>.
- Schoch, C. L., Ciufu, S., Domrachev, M., Hottton, C. L., Kannan, S., Khovanskaya, R., Leipe, D., Mcveigh, R., O'Neill, K., Robbertse, B., Sharma, S., Soussov, V., Sullivan, J. P., Sun, L., Turner, S., and Karsch-Mizrachi, I. (2020). NCBI Taxonomy: a comprehensive update on curation, resources and tools. *Database*, 2020, baaa062. <https://doi.org/10.1093/database/baaa062>.
- Schoch, C. L., Seifert, K. A., Huhndorf, S., Robert, V., Spouge, J. L., Levesque, C. A., Chen, W., Fungal Barcoding Consortium, & Fungal Barcoding Consortium Author List (2012). Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences*, 109(16), 6241–6246. <https://doi.org/10.1073/pnas.1117018109>.
- Sharpton, T. J. (2014). An introduction to the analysis of shotgun metagenomic data. *Frontiers in Plant Science*, 5, 209. <https://doi.org/10.3389/fpls.2014.00209>.
- Shikha, Singh, S. and S. Shankar (2021). Chapter 8 – Microbial metagenomics: potential and challenges. In Mondal, S and Singh, R.L (Ed), *Advances in Animal Genomics*, Academic Press, 109-122. <https://doi.org/10.1016/B978-0-12-820595-2.00008-4>.
- Shin, S., Lee, T. K., Han, J. M., and Park, J. (2014). Regional effects on chimera formation in 454 pyrosequenced amplicons from a mock community. *Journal of Microbiology*, 52(7), 566–573. <https://doi.org/10.1007/s12275-014-3485-6>.
- Siegwald, L., Touzet, H., Lemoine, Y., Hot, D., Audebert, C., and Caboche, S. (2017). Assessment of Common and Emerging Bioinformatics Pipelines for Targeted Metagenomics. *PLOS one*, 12(1), e0169563. <https://doi.org/10.1371/journal.pone.0169563>.
- Sørensen, L.M., Gori, K., Petersen, M.A., Jespersen, L. and Arneborg, N. (2011). Flavour compound production by *Yarrowia lipolytica*, *Saccharomyces cerevisiae* and *Debaryomyces hansenii* in

- a cheese-surface model *International Dairy Journal*, (12):970-978. <https://doi.org/10.1016/j.idairyj.2011.06.005>.
- Stielow, J. B., Lévesque C. A., Seifert K. A., Meyer W., Iriny L., Smits D., Renfurm, R., Verkley, G. J., Groenewald, M., Chaduli, D., Lomascolo, A., Welti, S., Lesage-Meessen, L., Favel, A., Al-Hatmi, A. M., Damm, U., Yilmaz, N., Houbraeken, J., Lombard, L., Quaedvlieg, W. and Robert, V. (2015). One fungus, which genes? Development and assessment of universal primers for potential secondary fungal DNA barcodes. *Persoonia*, 35, 242–263. <https://doi.org/10.3767/003158515X689135>.
- Thanh, V. N., Hien, D. D., and Thom, T. T. (2013). *Moniliella byzovii* sp. nov., a chlamydospore-forming black yeast isolated from flowers. *International Journal of Systematic and Evolutionary Microbiology*, 63(3), 1192–1196. <https://doi.org/10.1099/ijs.0.049767-0>.
- Tofalo, R., Fasoli, G., Schirone, M., Perpetuini, G., Pepe, A., Corsetti, A., and Suzzi, G. (2014). The predominance, biodiversity and biotechnological properties of *Kluyveromyces marxianus* in the production of Pecorino di Farindola cheese. *International Journal of Food Microbiology*, 187, 41–49. <https://doi.org/10.1016/j.ijfoodmicro.2014.06.029>.
- Torres, I (2016). Elaboración de un banco metagenómico del queso Cotija e identificación funcional de esterasas. Tesis de maestría. Facultad de Química, UNAM.
- Thukral, A. K. (2017). A review on measurement of Alpha diversity in biology. *Agricultural Research Journal*, 2017, 54(1): 1-10. <https://doi.org/10.5958/2395-146X.2017.00001.1>.
- van den Tempel, T and Jakobsen M (2000). The technological characteristics of *Debaryomyces hansenii* and *Yarrowia lipolytica* and their potential as starter cultures for production of Danablu. *International Dairy Journal*, 10: 263–270. [https://doi.org/10.1016/S0958-6946\(00\)00053-4](https://doi.org/10.1016/S0958-6946(00)00053-4).
- Vásquez J.A., Ramírez, M. y Monsalve, Z. I. (2016). Actualización en caracterización molecular de levaduras de interés industrial. *Revista Colombiana de Biotecnología*, 18(2), 129-139. <https://doi.org/http://dx.doi.org/10.15446/rev.colomb.biote.v18n2.61530>.
- Villegas, A (2004). Tecnología quesera. México: Trillas, pág. 316-321. ISBN: 978-607-17-0994-3 .
- Visagie, C.M., Varga, J., Houbraeken, J., Meijer, M., Kocsubé, S., Yilmaz, N., Fotedar, R., Seifert, K. A., Frisvad, J. C., and Samson, R. A. (2014). Ochratoxin production and taxonomy of the yellow aspergilli (*Aspergillus* section *Circumdati*). *Studies in Mycology*, 78, 1–61. <https://doi.org/10.1016/j.simyco.2014.07.001>.
- Wang, Q., Garrity, G. M., Tiedje, J. M., and Cole, J. R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology*, 73(16), 5261–5267. <https://doi.org/10.1128/AEM.00062-07>.
- Wang, X.-C., Liu, C., Huang, L., Bengtsson-Palme, J., Chen, H., Zhang, J.-H., Cai, D. and Li, J.-Q. (2015). ITS1: a DNA barcode better than ITS2 in eukaryotes? *Molecular Ecology Resources* 15: 573-586. <https://doi.org/10.1111/1755-0998.12325>.
- Weimer, B. (2007). Improving the flavour of cheese. First edición. Ed. Woodhead Publishing Limited, 144-156. <https://doi.org/10.1111/j.1471-0307.2008.00378.x>.
- Wood, D. E., Lu, J., and Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. *Genome Biology*, 20(1), 257. <https://doi.org/10.1186/s13059-019-1891-0>.
- Wolfe, B. E., Button, J. E., Santarelli, M., and Dutton, R. J. (2014). Cheese rind communities provide tractable systems for in situ and in vitro studies of microbial diversity. *Cell*, 158(2), 422–433. <https://doi.org/10.1016/j.cell.2014.05.041>.
- Wooley, J. C., Godzik, A., and Friedberg, I. (2010). A primer on metagenomics. *PLOS Computational Biology*, 6(2), e1000667. <https://doi.org/10.1371/journal.pcbi.1000667>.
- Yang, S.T., Zhang, K., Zhang, B. and H. Huang (2019). 3.16 - Fumaric Acid. In Moo-Young, M. (Ed.) *Comprehensive Biotechnology* (3ra Ed.), Pergamon, 188-207. <https://doi.org/10.1016/B978-0-12-809633-8.09141-X>.

- Youssef, F. S., and Singab, A. N. B. (2021). An Updated Review on the Secondary Metabolites and Biological Activities of *Aspergillus ruber* and *Aspergillus flavus* and Exploring the Cytotoxic Potential of Their Isolated Compounds Using Virtual Screening. *Evidence-based complementary and alternative medicine*, 2021, 8860784. <https://doi.org/10.1155/2021/8860784>.
- Yusuf, E. and Bülent, Ç. (2021). Bacterial and yeast microbial diversity of the ripened traditional Middle East surk cheese, *International Dairy Journal*, 117, 105004. <https://doi.org/10.1016/j.idairyj.2021.105004>.
- Zhang, L., Chen, F., Zeng, Z., Xu, M., Sun, F., Yang, L., Bi, X., Lin, Y., Gao, Y., Hao, H., Yi, W., Li, M., & Xie, Y. (2021). Advances in Metagenomics and Its Application in Environmental Microorganisms. *Frontiers in Microbiology*, 12, 766364. <https://doi.org/10.3389/fmicb.2021.766364>.
- Zmen, S., Capece, A., Siesto, G., Aksu, H., Sandikci, S., Yilmaz, F., Romano, P. and Karagul, Y. (2020). Molecular characterization of yeasts isolated from traditional Turkish cheeses. *Food Science and Technology*, 40(4), 871-876. <https://doi.org/10.1590/fst.24319>.

Anexo

I. Resultados obtenidos por Escobar (2016) con el programa Parallel-meta

Tabla 1. Anotación taxonómica para eucariontes del ADNr 18S de muestras de Queso Cotija de distintos municipios de la Región de Origen. (Tomado y modificado de Escobar, 2016).

Dominio	Reino	Filum	Familia	Género	Especie	Conteo	Abundancia Relativa (%)			
							Familia	Género	Especie	
Eukaryota	Fungi	Ascomycota	Mitosporic Dothioraceae	Kabatiella	Kabatiella microsticta	22	1.10	1.10	1.10	
Eukaryota	Fungi	Ascomycota	Dothideaceae	Bagnisiella	Bagnisiella examinans	27	4.15	1.35	1.35	
Eukaryota	Fungi	Ascomycota	Dothideaceae	Dothidea	Dothidea berberidis	56		2.80	2.80	
Eukaryota	Fungi	Ascomycota	Mycosphaerellaceae	Mycosphaerella	Mycosphaerella podagrariae	22	1.10	1.10	1.10	
Eukaryota	Fungi	Ascomycota	Teratosphaeriaceae	Teratosphaeria	Teratosphaeria molleriana	43	2.20	2.15	2.15	
Eukaryota	Fungi	Ascomycota	Teratosphaeriaceae	Teratosphaeriaceae	Teratosphaeriaceae sp. D007_09	1		---	0.05	
Eukaryota	Fungi	Ascomycota	Delitschiaceae	Delitschia	Delitschia winteri	1	0.05	0.05	0.05	
Eukaryota	Fungi	Ascomycota	Ajellomycetaceae	Ajellomyces	Ajellomyces capsulatus	4	0.20	0.20	0.20	
Eukaryota	Fungi	Ascomycota	Sclerotiniaceae	Botryotinia	Botryotinia fuckeliana B05.10	1	0.05	0.05	0.05	
Eukaryota	Fungi	Ascomycota	Sarcoscyphaceae	Sarcoscypha	Sarcoscypha austriaca	1	0.88	0.05	0.05	
Eukaryota	Fungi	Ascomycota	Sarcosmataceae	Urnula	Urnula hiemalis	16	0.80	0.80	0.80	
Eukaryota	Fungi	Ascomycota	Debaryomycetaceae	Debaryomyces	Debaryomyces udonii	283	45.70	35.20	14.15	
Eukaryota	Fungi	Ascomycota	Debaryomycetaceae	Debaryomyces	Debaryomyces hansenii var.hansenii	200			10.00	
Eukaryota	Fungi	Ascomycota	Debaryomycetaceae	Debaryomyces	Debaryomyces hansenii	155			7.75	
Eukaryota	Fungi	Ascomycota	Debaryomycetaceae	Debaryomyces	Debaryomyces sp.MPSR02	65			3.25	
Eukaryota	Fungi	Ascomycota	Debaryomycetaceae	Debaryomyces	Debaryomyces maramus	1			0.05	
Eukaryota	Fungi	Ascomycota	Debaryomycetaceae	Lodderomyces	Lodderomyces elongisporus	10			0.50	0.50
Eukaryota	Fungi	Ascomycota	Debaryomycetaceae	Meyerozyma	Meyerozyma guilliermondii ATCC6260	2			0.20	0.10
Eukaryota	Fungi	Ascomycota	Debaryomycetaceae	Meyerozyma	Meyerozyma guilliermondii	2				0.10
Eukaryota	Fungi	Ascomycota	Debaryomycetaceae	Schwannomyces	Schwannomyces capriottii	196			9.80	9.80
Eukaryota	Fungi	Ascomycota	Endomycetaceae	Endomyces	Endomyces sp.	1			0.05	0.05
Eukaryota	Fungi	Ascomycota	Metschnikowiaceae	Metschnikowia	Metschnikowia sp.UWOPS04-218.3	2	0.10	0.10	0.10	

Eukaryota	Fungi	Ascomycota	Mitosporic Saccharomycetales	Blastobotrys	Blastobotrys adenivorans	3	3.70	0.15	0.15
Eukaryota	Fungi	Ascomycota	Mitosporic Saccharomycetales	Brettanomyces	Brettanomyces naardenensis	1		0.05	0.05
Eukaryota	Fungi	Ascomycota	Mitosporic Saccharomycetales	Candida	Candida dubliniensis	48		3.50	2.40
Eukaryota	Fungi	Ascomycota	Mitosporic Saccharomycetales	Candida	Candida lignicola	7			0.35
Eukaryota	Fungi	Ascomycota	Mitosporic Saccharomycetales	Candida	Candida sp.SK75	3			0.15
Eukaryota	Fungi	Ascomycota	Mitosporic Saccharomycetales	Candida	Candida lyxosophila	3			0.15
Eukaryota	Fungi	Ascomycota	Mitosporic Saccharomycetales	Candida	Candida sp.BG02-7-20-001A-2-1	2			0.10
Eukaryota	Fungi	Ascomycota	Mitosporic Saccharomycetales	Candida	Candida silvanorum	2			0.10
Eukaryota	Fungi	Ascomycota	Mitosporic Saccharomycetales	Candida	Candida sp.NRRLY-27170	1			0.05
Eukaryota	Fungi	Ascomycota	Mitosporic Saccharomycetales	Candida	Candida psychrophila	1			0.05
Eukaryota	Fungi	Ascomycota	Mitosporic Saccharomycetales	Candida	Candida palmioleophila	1			0.05
Eukaryota	Fungi	Ascomycota	Mitosporic Saccharomycetales	Candida	Candida lassenensis	1			0.05
Eukaryota	Fungi	Ascomycota	Mitosporic Saccharomycetales	Candida	Candida fennica	1		0.05	
Eukaryota	Fungi	Ascomycota	Saccharomycetaceae	Cyberlindnera	Cyberlindnera saturnus	3		0.15	0.15
Eukaryota	Fungi	Ascomycota	Saccharomycetaceae	Dekkera	Dekkera anomala	4	0.25	0.20	
Eukaryota	Fungi	Ascomycota	Saccharomycetaceae	Dekkera	Dekkera bruxellensis	1		0.05	
Eukaryota	Fungi	Ascomycota	Saccharomycetaceae	Kazachstania	Kazachstania unispora	330	16.65	16.50	
Eukaryota	Fungi	Ascomycota	Saccharomycetaceae	Kazachstania	Kazachstania sp.QMW-2009a	1		0.05	
Eukaryota	Fungi	Ascomycota	Saccharomycetaceae	Kazachstania	Kazachstania exigua	1		0.05	
Eukaryota	Fungi	Ascomycota	Saccharomycetaceae	Kazachstania	Kazachstania telluris	1		0.05	
Eukaryota	Fungi	Ascomycota	Saccharomycetaceae	Kluyveromyces	Kluyveromyces wickerhamii	1		0.05	0.05
Eukaryota	Fungi	Ascomycota	Saccharomycetaceae	Kodamaea	Kodamaea ohmeri	1		0.05	0.05
Eukaryota	Fungi	Ascomycota	Saccharomycetaceae	Lachancea	Lachancea kluyveri	8		0.40	0.40
Eukaryota	Fungi	Ascomycota	Saccharomycetaceae	Naumovia	Naumovia castelli	145		7.25	7.25
Eukaryota	Fungi	Ascomycota	Saccharomycetaceae	Pichia	Pichia sp.	4		0.40	0.20
Eukaryota	Fungi	Ascomycota	Saccharomycetaceae	Pichia	Pichia triangularis	1			0.05
Eukaryota	Fungi	Ascomycota	Saccharomycetaceae	Pichia	Pichia sp.SG3L01	1	0.05		
Eukaryota	Fungi	Ascomycota	Saccharomycetaceae	Pichia	Pichia pijperi	1	0.05		
Eukaryota	Fungi	Ascomycota	Saccharomycetaceae	Pichia	Pichia mexicana	1	0.05		
Eukaryota	Fungi	Ascomycota	Saccharomycetaceae	Saccharomyces	Saccharomyces cerevisiae	18	1.55	0.90	
Eukaryota	Fungi	Ascomycota	Saccharomycetaceae	Saccharomyces	Saccharomyces spencerorum	8		0.40	
Eukaryota	Fungi	Ascomycota	Saccharomycetaceae	Saccharomyces	Saccharomyces servazzii	2		0.10	
Eukaryota	Fungi	Ascomycota	Saccharomycetaceae	Saccharomyces	Saccharomyces bayanus	2		0.10	
Eukaryota	Fungi	Ascomycota	Saccharomycetaceae	Saccharomyces	Saccharomyces cerevisiae AWRI796	1		0.05	
Eukaryota	Fungi	Ascomycota	Saccharomycetaceae	Williopsis	Williopsis salicorniae	65		3.25	
Eukaryota	Fungi	Ascomycota	Saccharomycetaceae	Williopsis	Williopsis pratensis	16	4.05	0.80	
Eukaryota	Fungi	Ascomycota	Saccharomycetaceae	Zygosaccharomyces	Zygosaccharomyces lentus	17	0.85	0.85	
Eukaryota	Fungi	Ascomycota	Saccharomycopsidaceae	Saccharomycopsis	Saccharomycopsis fibuligera	1	0.05	0.05	0.05
Eukaryota	Fungi	Ascomycota	Wickerhamomycetaceae	Wickerhamomyces	Wickerhamomyces sydowiorum	1	0.05	0.05	0.05
Eukaryota	Fungi	Ascomycota	Cordycipitaceae	Cordyceps	Cordyceps bifusispora	1	0.05	0.05	0.05
Eukaryota	Fungi	Ascomycota	Hypocreaceae	Hypocrea	Hypocrea lixi	2	0.10	0.10	0.10
Eukaryota	Fungi	Ascomycota	Mitosporic Halosphaeriaceae	Monodictys	Monodictys castaneae	2	0.10	0.10	0.10
Eukaryota	Fungi	Ascomycota	Lasiosphaeriaceae	Podospora	Podospora anserina	2	0.10	0.10	0.10
Eukaryota	Fungi	Ascomycota	Schizosaccharomycetaceae	Schizosaccharomyces	Schizosaccharomyces japonicus	1	0.05	0.05	0.05

Eukaryota	Fungi	Ascomycota	Taphrinaceae	Taphrina	Taphrina deformans	2			0.10
Eukaryota	Fungi	Ascomycota	Taphrinaceae	Taphrina	Taphrina tosquinetii	1	0.20		0.05
Eukaryota	Fungi	Ascomycota	Taphrinaceae	Taphrina	Taphrina farlowii	1		0.20	0.05
Eukaryota	Fungi	Basidiomycota	Agaricales	Uncultured Russulales	Uncultured Russulales	1	0.05	0.05	0.05
Eukaryota	Fungi	Basidiomycota	Cystofilobasidiaceae	Cystofilobasidium	Cystofilobasidium infirmominiatum	2	0.15		0.10
Eukaryota	Fungi	Basidiomycota	Cystofilobasidiaceae	Cystofilobasidium	Cystofilobasidium ferigula	1		0.15	0.05
Eukaryota	Fungi	Basidiomycota	Dacrymycetaceae	Calocera	Calocera viscosa	12	0.60	0.60	0.60
Eukaryota	Fungi	Basidiomycota	Hymenochaetales incertae sedis	Resinicium	Resinicium saccharicola	2	0.10	0.10	0.10
Eukaryota	Fungi	Basidiomycota	Sparassidaceae	Sparassis	Sparassis sp.TENN50289/ss31	2	0.10	0.10	0.10
Eukaryota	Fungi	Basidiomycota	Sebacinaceae	Craterocolla	Craterocolla cerasi	2	0.95	0.10	0.10
Eukaryota	Fungi	Basidiomycota	Sebacinaceae	Sebacina	Sebacina vermifera	17		0.85	0.85
Eukaryota	Fungi	Basidiomycota	Mitosporic Entylomataceae	Tilletiopsis	Tilletiopsis oryzicola	4	0.20	0.20	0.20
Eukaryota	Fungi	Basidiomycota	Mitosporic Agaricostilbomycetidae	Bensingtonia	Bensingtonia musae	1	0.05	0.05	0.05
Eukaryota	Fungi	Chytridiomycota	Unclassified Lobulomycetales	Lobulomycetales	Lobulomycetales sp.AF011	3	0.15	0.15	0.15
Eukaryota	Fungi	Chytridiomycota	Unclassified	----	----	1	0.05	----	0.05
Eukaryota	Fungi	Dikarya	Taiwanofungus	Taiwanofungus	Taiwanofungus camphoratus	1	0.05	0.05	0.05
Eukaryota	Fungi	Dikarya	Bispora	Bispora	Bispora christiansenii	18	0.90	0.90	0.90
Eukaryota	Fungi	Dikarya	Graphium	Graphium	Graphium rubrum	2	0.10	0.10	0.10
Eukaryota	Fungi	Dikarya	Moniliella	Moniliella	Moniliella suaveolens	17	1.20		0.85
Eukaryota	Fungi	Dikarya	Moniliella	Moniliella	Moniliella acetobutans	7		1.20	0.35
Eukaryota	Fungi	Dikarya	Phaeomoniella	Phaeomoniella	Phaeomoniella prunicola	2	0.10	0.10	0.10
Eukaryota	Fungi	Dikarya	Spicellum	Spicellum	Spicellum roseum	1	0.05	0.05	0.05
Eukaryota	Fungi	Dikarya	Tetracladium	Tetracladium	Tetracladium marchalianum	1	0.05	0.05	0.05
Eukaryota	Fungi	Dikarya	Neolecta	Neolecta	Neolecta irregularis	1	0.05	0.05	0.05
Eukaryota	Fungi	Dikarya	Unclassified	----	----	1	0.05	----	----
Eukaryota	Fungi	Environmental	----	----	----	4	----	----	----
Eukaryota	Fungi	Fungi incertae sedis	Zygomycete sp.AM-2008a	Unclassified	Unclassified	20	1.00	1.00	1.00
Eukaryota	Fungi	Fungi incertae sedis	Mucor	Mucor	Mucor mucedo	2	0.10	0.10	0.10
Eukaryota	Fungi	Fungi incertae sedis	Zygorhynchus	Zygorhynchus	Zygorhynchus heterogamus	1	0.05	0.05	0.05
Eukaryota	Fungi	Glomeromycota	Environmental samples	Uncultured Glomus	Uncultured Glomus	1	0.05	0.05	0.05
Eukaryota	Alveolata	Apicomplexa	Babesia	Babesia	Babesia odocoilei	2	0.10	0.10	0.10
Eukaryota	Alveolata	Ciliophora	Leptopharynx	Leptopharynx	Leptopharynx costatus	1	0.05	0.05	0.05
Eukaryota	Alveolata	Dinophyceae	Amphidinium	Amphidinium	Amphidinium belauense	1	0.05	0.05	0.05
Eukaryota	Alveolata	Dinophyceae	Uncultured syndiniales	Unclassified	Unclassified	1	0.05	0.05	0.05
Eukaryota	Amoebozoa	Centramoebida	Protacanthamoeba bohemica	Unclassified	Unclassified	5	0.25	0.25	0.25
Eukaryota	Fonticula	Fonticula	----	----	----	2	----	----	----
Eukaryota	Metazoa	Craniata	Ophichthidae	Echiophis	Echiophis punctifer	1	0.05	0.05	0.05
Eukaryota	Metazoa	Malacostraca	Homoloidea	Latreilla	Latreilla valida	1	0.05	0.05	0.05
Eukaryota	Metazoa	Porifera	Crambe	Crambe	Crambe crambe	2	0.10	0.10	0.10
Eukaryota	Metazoa	Vertebrata	Murinae	Mus	Mus musculus	4	0.20	0.20	0.20
Eukaryota	Viridiplantae	Tracheophyta	Acanthaceae	Avicennia	Avicennia marina	2	0.10	0.10	0.10
Eukaryota	Environmental	Uncultured eukaryote	----	----	----	19	----	---	----

II. Cebadores utilizados en la secuenciación de la región ITS1

Tabla 2. Cebadores utilizados para la amplificación de la región ITS1.
Tomado de Illumina, 2021.

Cebadores directos		Cebadores inversos	
Nombre	Secuencia (5')	Nombre	Secuencia (5')
ITS_fwd_1	CTTGGTCATTTAGAGGAAGTAA	ITS_rev_1	GCTGCGTTCTTCATCGATGC
ITS_fwd_2	CTCGGTCATTTAGAGGAAGTAA	ITS_rev_2	GCTGCGTTCTTCATCGATGG
ITS_fwd_3	CTTGGTCATTTAGAGGAACTAA	ITS_rev_3	GCTACGTTCTTCATCGATGC
ITS_fwd_4	CCCGGTCATTTAGAGGAAGTAA	ITS_rev_4	GCTGCGTTCTTCATCGATGT
ITS_fwd_5	CTAGGCTATTTAGAGGAAGTAA	ITS_rev_5	ACTGTGTTCTTCATCGATGT
ITS_fwd_6	CTTAGTTATTTAGAGGAAGTAA	ITS_rev_6	GCTGCGTTCTTCATCGTTGC
ITS_fwd_7	CTACGTCATTTAGAGGAAGTAA	ITS_rev_7	GCGTTCTTCATCGATGC
ITS_fwd_8	CTTGGTCATTTAGAGGTCGTAA	----	----

III. Soluciones

Tabla 3. Buffer de lisis

Reactivo / Solución	Cantidad	Concentración
Tris-HCl	2 mL	20 mM
NaCl 5M	5 mL	250 mM
SDS 20%	5 mL	25 mM
EDTA•Na ₂ 0.5 M pH 8	2.5 mL	0.5%
H ₂ O destilada	cbp 100 mL	---

IV. Asignación taxonómica de las secuencias de la región ITS1

Tabla 4. Hongos y levaduras asignadas con BLASTn según Bálint *et al.* (2014).

Phylum	Clase	Orden	Familia	Género	Especie	Conteo	QC-22		QC-57		QC-95				
							Abundancia		Conteo	Abundancia		Conteo	Abundancia		
							Género	Especie		Género	Especie		Género	Especie	
NA	NA	NA	NA	NA	NA	44									
NA	NA	NA	NA	NA	NA	11									
Ascomycota	NA	NA	NA	NA	NA	5									
Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium_oxysporum	16	0.0099	0.0069	16	0.0110	0.011	59	0.0230		
Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium_halotolerans	4		0.0017							
Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium_sphaerospermum	3		0.0013							
Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium_dominicanum							2	0.0008		
Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Toxicocladosporium	Toxicocladosporium_protearum							8	0.0030	0.0030	
Ascomycota	Dothideomycetes	Mycosphaerellales	Teratosphaeriaceae	NA	NA							4			
Ascomycota	Dothideomycetes	Mycosphaerellales	Teratosphaeriaceae	Pseudotaeniolina	Pseudotaeniolina_globosa							7	0.0026	0.0026	
Ascomycota	Dothideomycetes	Pleosporales	Didymellaceae	Ascochyta	Ascochyta_medicaginicola							76	0.0287	0.0287	
Ascomycota	Dothideomycetes	Pleosporales	Didymellaceae	Didymella	Didymella_pomorum				3	0.0021	0.0021				
Ascomycota	Dothideomycetes	Pleosporales	Didymellaceae	Epicoccum	NA							148	0.0559		
Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	Phaeosphaeria	NA	6	0.0026								
Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	Alternaria_alstroemeriae	48	0.0206	0.0206							
Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Curvularia	NA							2	0.0008		
Ascomycota	Eurotiomycetes	Chaetothyriales	Chaetothyriales_fam_Incertae_sedis	Neophaeococcomyces	Neophaeococcomyces_aloes							5	0.0019	0.0019	
Ascomycota	Eurotiomycetes	Chaetothyriales	Herpotrichiellaceae	Coniosporium	NA							6	0.0023		
Ascomycota	Eurotiomycetes	Chaetothyriales	Herpotrichiellaceae	Exophiala	Exophiala_oligosperma				3	0.0021	0.0021				
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Aspergillus	NA	6437	6.2777					52	0.0223		
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Aspergillus	NA	11									
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Aspergillus	Aspergillus_candidus	557		0.2387						7	0.0026
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Aspergillus	Aspergillus_jensenii	144		0.0617							
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Aspergillus	Aspergillus_westerdijkiae	7496									
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Aspergillus	Aspergillus_westerdijkiae	3		3.2143							
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Aspergillus	Aspergillus_westerdijkiae	2									
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	NA	67						4	0.0042		
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium_chrysogenum	32	0.0441	0.0137				7		58.3333	
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium_citreonigrum										
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium_citrinum	4		0.0017							
Ascomycota	Leotiomycetes	NA	NA	NA	NA							5			
Ascomycota	Leotiomycetes	Erysiphales	Erysiphaceae	Erysiphe	Erysiphe_vignae				3	0.0021	0.0021				
Ascomycota	Leotiomycetes	Helotiales	Sclerotiniaceae	Monilia	Monilia_medoacensis	2	0.0009	0.0009							

Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Debaryomyces	NA	2												
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Debaryomyces	NA		54.3535		2	12.5136							0.8644	
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Debaryomyces	Debaryomyces_hansenii	126840		54.3526	18237		12.51	1825						0.6888
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Debaryomyces	Debaryomyces_delbrueckii							465						0.1755
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Kurtzmaniella	NA	55	0.0236		3	0.0021		577					0.2178	
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Meyerozyma	Meyerozyma_guilliermondii	115	0.0493	0.0493	5	0.0034	0.0034	80					0.0302	0.030
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Yamadazyma	NA	5472			3752			508						
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Yamadazyma	NA	1178			736		4.1227	47					0.2204	
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Yamadazyma	NA	957			1521			29						
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Yamadazyma	NA	2												
Ascomycota	Saccharomycetes	Saccharomycetales	Dipodasaceae	Geotrichum	NA	5												
Ascomycota	Saccharomycetes	Saccharomycetales	Dipodasaceae	Geotrichum	Geotrichum_silvicola	3	0.0034	0.0013									0.0045	
Ascomycota	Saccharomycetes	Saccharomycetales	Dipodasaceae	Geotrichum	Geotrichum_candidum							12						0.0045
Ascomycota	Saccharomycetes	Saccharomycetales	Dipodasaceae	Yarrowia	Yarrowia_alimentaria							20					0.0075	0.0075
Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Clavispora	NA	12												
Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Clavispora	Clavispora_lusitanae		0.0051		2	0.0014								
Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Kodamaea	Kodamaea_ohmeri	9						9020						
Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Kodamaea	Kodamaea_ohmeri		0.0039	0.0039	184	0.1276	0.1276						3.4065	3.4065
Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Kodamaea	Kodamaea_ohmeri				2									
Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Kodamaea	Kodamaea_ohmeri							5						
Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Pichia	Pichia_kluyveri	4		0.0017										
Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Pichia	Pichia_mandschurica	3	0.0030	0.0013										
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Barnettozyma	Barnettozyma_californica							12					0.0045	0.0045
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	NA	41320			31252			30050						
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_phaffii	19644			15125			16251						
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_phaffii	3	27.1476	8.4190		32.6715	10.3771						17.9122	6.1347
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_phaffii							2						
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_ulmi	2386		1.0224	1243			0.8528	1153					0.4352
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Issatchenkia	Issatchenkia_orientalis	28	0.0120	0.0120		0.0027	0.0027						3.2057	3.2057
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Issatchenkia	Issatchenkia_orientalis				4			8493						
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Kluyveromyces	Kluyveromyces_lactis	3627	1.5542	1.5542	6868	4.7120	4.7120	180014					67.95	67.9460
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	NA	4841			7									
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	NA	3												
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida_parapsilosis	337		0.1444	89			0.0611	5733					2.1639
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida_metapsilosis	258						125						
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida_metapsilosis	12		0.1157	160			0.1585	58					0.0691
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida_metapsilosis		2.3375		71		0.2490							2.4179
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida_orthopsilosis	2		0.0009										
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida_oceani	2		0.0009	36			0.0247	38					0.0143
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida_spenceri								430					0.1623
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida_intermedia								22					0.0083

Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Diuina	Diuina_catenulata	232	0.0994	0.0994	6	0.0041	0.0041	74	0.0279	0.0279
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Starmerella	NA	2064	0.8845		65899	45.2303		1326	0.5005	
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Starmerella	Starmerella_apicola				26		0.0178			
Ascomycota	Saccharomycetes	Saccharomycetales	Trichomonasaceae	Wickerhamiella	Wickerhamiella_pararugosa	251	0.1084	0.1084				3185	1.2044	1.2044
Ascomycota	Saccharomycetes	Saccharomycetales	Trichomonasaceae	Wickerhamiella	Wickerhamiella_pararugosa	2						6		
Ascomycota	Sordariomycetes	Hypocreales	Bionectriaceae	Clionastix	NA	4	0.0017							
Ascomycota	Sordariomycetes	Hypocreales	Bionectriaceae	Clonostachys	NA	3	0.0013							
Ascomycota	Sordariomycetes	Chaetosphaeriales	Chaetosphaeriaceae	Pseudodinemasprium	NA	3	0.0013							
Ascomycota	Sordariomycetes	Hypocreales	Cordycipitaceae	Lecanicillium	Lecanicillium_coprophilum	103	0.0441	0.0441	46	0.0316	0.0316	738	0.2786	0.2786
Ascomycota	Sordariomycetes	Hypocreales	Cordycipitaceae	Simplicillium	NA							43	0.0192	
Ascomycota	Sordariomycetes	Hypocreales	Cordycipitaceae	Simplicillium	Simplicillium_lanosoniveum							8		0.0030
Ascomycota	Sordariomycetes	Hypocreales	Hypocreales_fam_Incertae_sedis	Acremonium	Acremonium_charticola	579	0.4701	0.4701						
Ascomycota	Sordariomycetes	Hypocreales	Hypocreales_fam_Incertae_sedis	Acremonium	Acremonium_charticola	518								
Ascomycota	Sordariomycetes	Hypocreales	Hypocreales_fam_Incertae_sedis	Sarocladium	NA				2	0.0014				
Ascomycota	Sordariomycetes	Microascales	Microasaceae	Scopulariopsis	Scopulariopsis_flava	3	0.0013	0.0013				4	0.0015	0.0015
Ascomycota	Sordariomycetes	Hypocreales	Nectriaceae	Fusarium	NA	41								
Ascomycota	Sordariomycetes	Hypocreales	Nectriaceae	Fusarium	Fusarium_oxysporum	346	0.1658	0.1483				2	0.0008	0.0008
Ascomycota	Sordariomycetes	Xylariales	Diatrypaceae	Eutypella	NA	5	0.0021							
Ascomycota	Taphrinomycetes	Taphrinales	Taphrinales	Taphrina	NA	3	0.0013							
Basidiomycota	Agaricomycetes	Boletales	Pisolithaceae	Pisolithus	NA	3	0.0013							
Basidiomycota	Agaricomycetes	Polyporales	Ganodermataceae	Ganoderma	NA				5	0.0034				
Basidiomycota	Agaricomycetes	Polyporales	Steccherinaceae	Junghuhnia	NA							18	0.006794	
Basidiomycota	Agaricomycetes	Russulales	Peniophoraceae	NA	NA	8	0.0034							
Basidiomycota	Agaricomycetes	Russulales	Peniophoraceae	NA	NA							127		
Basidiomycota	Agaricostilbomycetes	Agaricostilbales	Agaricostilbaceae	Sterigmatomyces	NA				7	0.1063		16	0.0506	
Basidiomycota	Agaricostilbomycetes	Agaricostilbales	Agaricostilbaceae	Sterigmatomyces	Sterigmatomyces_halophilus	37	0.0159	0.0159	148		0.1015	118		0.0445
Basidiomycota	Malasseziomycetes	Malasseziales	NA	NA	NA							3		
Basidiomycota	Malasseziomycetes	Malasseziales	Malasseziaceae	Malassezia	NA	20	0.0086		4	0.0027		63	0.0287	
Basidiomycota	Malasseziomycetes	Malasseziales	Malasseziaceae	Malassezia	NA							13		
Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodotorula	NA				3					
Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodotorula	NA					0.0055		4	0.0015	
Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodotorula	Rhodotorula_mucilaginosa				5		0.0034			
Basidiomycota	Moniliellomycetes	Moniliellales	Moniliellaceae	Moniliella	NA	5139			123			1970		
Basidiomycota	Moniliellomycetes	Moniliellales	Moniliellaceae	Moniliella	NA	1897	3.0236		67		0.1578	1534	1.40	
Basidiomycota	Moniliellomycetes	Moniliellales	Moniliellaceae	Moniliella	NA	20			15			169		
Basidiomycota	Moniliellomycetes	Moniliellales	Moniliellaceae	Moniliella	NA				25			35		
Basidiomycota	Tremellomycetes	Filobasidiales	Filobasidiaceae	Naganishia	Naganishia_diffluens	3	0.0013	0.0013						
Basidiomycota	Tremellomycetes	Filobasidiales	Filobasidiaceae	Filobasidium	Filobasidium_magnum	3	0.0013	0.0013						
Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Bulleribasidium	Bulleribasidium_oberjochense				4	0.0027	0.0027			
Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Dioszegia	NA				2	0.0014				
Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma_victoriae	21	0.0090	0.0090						
Basidiomycota	Tremellomycetes	Trichosporonales	Trichosporonaceae	Trichosporon	Trichosporon_coremiiforme	6	0.0026	0.0026	6	0.0041	0.004	61	0.0230	0.0230

Basidiomycota	Ustilaginomycetes	Ustilaginales	Ustilaginaceae	Dirkmeia	NA	5	0.0021							
Basidiomycota	Ustilaginomycetes	Ustilaginales	Ustilaginaceae	Pseudozyma	NA				3	0.0021				
Basidiomycota	Wallemiomycetes	Wallemiales	Wallemiaceae	Wallemia	NA	22								
Basidiomycota	Wallemiomycetes	Wallemiales	Wallemiaceae	Wallemia	NA				3					
Basidiomycota	Wallemiomycetes	Wallemiales	Wallemiaceae	Wallemia	NA		0.0146		2	0.0034				0.0030
Basidiomycota	Wallemiomycetes	Wallemiales	Wallemiaceae	Wallemia	NA							8		
Basidiomycota	Wallemiomycetes	Wallemiales	Wallemiaceae	Wallemia	Wallemia_muriae	12		0.0051						
Mucoromycota	Mucoromycetes	Mucorales	Mucoraceae	Mucor	NA							39	0.0147	
NA	NA	NA	NA	NA	NA				29					
NA	NA	NA	NA	NA	NA							12		
					Total	233365	---	----	145754	---	----	264937	---	----

Tabla 5. Hongos y levaduras asignadas con dada2

Phylum	Clase	Orden	Familia	Género	Especie	QC-22		QC-57		QC-95				
						Conteo	Abundancia		Conteo	Abundancia		Conteo	Abundancia	
							Género	Especie		Género	Especie		Género	Especie
Ascomycota	NA	NA	NA	NA	NA						16			
Ascomycota	NA	NA	NA	NA	NA	10								
Ascomycota	NA	NA	NA	NA	NA	1101								
Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	NA	17	0.0265							
Ascomycota	Dothideomycetes	Pleosporales	Didymellaceae	NA	NA							37		
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Aspergillus	Aspergilluostianus	214								
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Aspergillus	Aspergilluostianus	76								
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Aspergillus	Aspergilluostianus	74								
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Aspergillus	Aspergilluostianus	67	3.1793							
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Aspergillus	Aspergilluruber	559								
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Aspergillus	Aspergilluruber	555		2.5062						
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Aspergillus	Aspergilluruber	491								
Ascomycota	Sordariomycetes	Hypocreales	Hypocreales_fam_Incertae_sedis	Acremonium	Acremonium_charticola	159	0.2483	0.2483						
Ascomycota	Saccharomycetes	Saccharomycetales	NA	NA	NA							1164		
Ascomycota	Saccharomycetes	Saccharomycetales	NA	NA	NA							1158		
Ascomycota	Saccharomycetes	Saccharomycetales	NA	NA	NA				29					
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Debaryomyces	NA	3720			475			86		
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Debaryomyces	NA	3286			461			42		
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Debaryomyces	NA	3276			476					
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Debaryomyces	NA	634								
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Debaryomyces	NA	267								
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Debaryomyces	NA	258	18.8913			3.0459				
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Debaryomyces	Debaryomyces_coudertii	657		1.0259						
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Debaryomyces	Debaryomyces_nepalensis				117					
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Debaryomyces	Debaryomyces_nepalensis				114					
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Debaryomyces	Debaryomyces_nepalensis				113					
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Debaryomyces	Debaryomyces_nepalensis				53					
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Meyerozyma	Meyerozyma_carpophila	64	0.0999	0.0999						
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Yamadazyma	NA	242			166			42		
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Yamadazyma	NA	217			155			47		
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Yamadazyma	NA	247	1.1134		171	0.8284			0.2054	
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Yamadazyma	Yamadazyma_triangularis	7		0.0109						
Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Kodamaea	Kodamaea_ohmeri							1147		
Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Kodamaea	Kodamaea_ohmeri				61	0.1027	0.1027		2.6477	

Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Barnettozyma	Barnettozyma_californica							6	0.0139	0.0139
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	NA	438						306		320
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	NA	408						292		310
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_phaffii	3797						2410		2529
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_phaffii	3460						2308		2336
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_phaffii	3363						2234		2186
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_phaffii	2587						1856		1976
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_phaffii	2598						1805		1766
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_phaffii	2475						1681		1929
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_phaffii	2324						1612		1763
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_phaffii	2440						1740		1505
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_phaffii	2253						1619		1442
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_phaffii	2221						1658		1421
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_phaffii	2046						1465		1490
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_phaffii	1956						1454		1274
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_phaffii	1085		63.08			47.53	751		671
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_phaffii	1022						700		593
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_phaffii	942	66.1993			49.56		660	66.6090	681
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_phaffii	879						703		599
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_phaffii	844						556		565
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_phaffii	717						538		591
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_phaffii	641						505		477
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_phaffii	569						449		394
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_phaffii	526						358		423
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_phaffii	437						271		297
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_phaffii	416						307		276
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_phaffii	408						305		279
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_phaffii	388						285		283
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_ulmi	237						111		99
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_ulmi	221						104		103
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_ulmi	208						102		88
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_ulmi	185		1.8020			1.02	106		96
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_ulmi	142						79		93
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella_ulmi	161						103		
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_1ncertae_sedis	Candida	Candida_atlantica	76						95		
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_1ncertae_sedis	Candida	Candida_atlantica	85		0.3716			0.4176	83		
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_1ncertae_sedis	Candida	Candida_atlantica	77						70		
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_1ncertae_sedis	Candida	Candida_intermedia								21	0.0485
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_1ncertae_sedis	Candida	Candida_metapsilosis							168		
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_1ncertae_sedis	Candida	Candida_metapsilosis	67	0.6496	0.1046		0.8082	0.2829		4.3236	0.1200
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_1ncertae_sedis	Candida	Candida_metapsilosis							52		
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_1ncertae_sedis	Candida	Candida_parapsilosis	90						64		647
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_1ncertae_sedis	Candida	Candida_parapsilosis			0.1405			0.1078		609	4.1551
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_1ncertae_sedis	Candida	Candida_parapsilosis							544		
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_1ncertae_sedis	Candida	Candida_zeilanoidea	21		0.0328						

Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Diutina	Diutina_catenuolata	119	0.1858	0.1858				72	0.1662	0.1662						
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Starmerella	Starmerella_etchellsii	339	1.6334	1.6334	8620	44.9209	44.9209	205	1.4151	1.4151						
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Starmerella	Starmerella_etchellsii	375			8089			211								
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Starmerella	Starmerella_etchellsii	332			7535			197								
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Starmerella	Starmerella_etchellsii				838											
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Starmerella	Starmerella_etchellsii				804											
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Starmerella	Starmerella_etchellsii				793											
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Issatchenka	Issatchenka_orientalis		0.1202	0.1202				2121	6.5328	6.5328						
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Issatchenka	Issatchenka_orientalis											395				
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Issatchenka	Issatchenka_orientalis	77										314				
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Kluyveromyces	Kluyveromycelactis	121	0.1889	0.1889	127	0.2138	0.2138	1915	4.9723	4.9723						
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Kluyveromyces	Kluyveromycelactis													135		
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Kluyveromyces	Kluyveromycelactis													104		
Ascomycota	Saccharomycetes	Saccharomycetales	Trichomonascaceae	Wickerhamiella	Wickerhamiella_pararugosa	73	0.1140	0.1140				307	2.6270	2.6270						
Ascomycota	Saccharomycetes	Saccharomycetales	Trichomonascaceae	Wickerhamiella	Wickerhamiella_pararugosa											343				
Ascomycota	Saccharomycetes	Saccharomycetales	Trichomonascaceae	Wickerhamiella	Wickerhamiella_pararugosa											258				
Ascomycota	Saccharomycetes	Saccharomycetales	Trichomonascaceae	Wickerhamiella	Wickerhamiella_pararugosa											230				
Basidiomycota	NA	NA	NA	NA	NA				27											
Basidiomycota	Agaricostilbomycetes	Agaricostilbales	Agaricostilbaceae	Sterigmatomyces	Sterigmatomycehalophilus		0.0156	0.0156	47	0.0791	0.0791									
Basidiomycota	Agaricostilbomycetes	Agaricostilbales	Agaricostilbaceae	Sterigmatomyces	Sterigmatomycehalophilus	10														
Basidiomycota	Moniliellomycetes	Moniellales	Moniellaceae	Moniella	NA	816	5.5731		128	0.3199		330	4.5175							
Basidiomycota	Moniliellomycetes	Moniellales	Moniellaceae	Moniella	NA	893									368					
Basidiomycota	Moniliellomycetes	Moniellales	Moniellaceae	Moniella	NA	783									357					
Basidiomycota	Moniliellomycetes	Moniellales	Moniellaceae	Moniella	NA	371									336					
Basidiomycota	Moniliellomycetes	Moniellales	Moniellaceae	Moniella	NA	374									300					
Basidiomycota	Moniliellomycetes	Moniellales	Moniellaceae	Moniella	NA	332					62				266					
Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma_victoriae	17	0.0265	0.0265												
NA	NA	NA	NA	NA	NA							48								
NA	NA	NA	NA	NA	NA							35								
NA	NA	NA	NA	NA	NA				17											
				Total		64040	----	----	59391	----	----	43320	----	----						

V. Asignación taxonómica de las secuencias metagenómicas de *shotgun*

Tabla 6. Hongos y levaduras asignadas con Kraken2

Phylum	Clase	Orden	Familia	Género	Especie	Conteo	Abundancia (%)
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Debaryomyces	Debaryomyces hanseni	73764	93.8378
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Scheffersomyces	Scheffersomyces stipitis	1696	2.1575
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Candida	Candida orthopsilosis	91	0.1158
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Candida	Candida dubliniensis	31	0.0394
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Candida	Candida albicans	17	0.0216
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Kluyveromyces	Kluyveromyces lactis	629	0.8002
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Kluyveromyces	Kluyveromyces marxianus	16	0.0204
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Saccharomyces	Saccharomyces cerevisiae	23	0.0293
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Saccharomyces	Saccharomyces paradoxus	20	0.0254
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Saccharomyces	Saccharomyces eubayanus	8	0.0102
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Kazachstania	Kazachstania africana	22	0.0280
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Kazachstania	Kazachstania naganishii	16	0.0204
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Eremothecium	Eremothecium gossypii	21	0.0267
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Eremothecium	Eremothecium sinecaudum	4	0.0051
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Eremothecium	Eremothecium cymbalariae	1	0.0013
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Naumovozyma	Naumovozyma castellii	18	0.0229
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Naumovozyma	Naumovozyma dairenensis	7	0.0089
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Tetrapisispora	Tetrapisispora phaffii	13	0.0165
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Tetrapisispora	Tetrapisispora blattae	10	0.0127
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Lachancea	Lachancea thermotolerans	14	0.0178
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Zygosaccharomyces	Zygosaccharomyces rouxii	13	0.0165
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Torulaspora	Torulaspora globosa	7	0.0089
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Torulaspora	Torulaspora delbrueckii	4	0.0051
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Nakaseomyces	[Candida] glabrata	5	0.0064
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Zygotorulaspora	Zygotorulaspora mrakii	3	0.0038
Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Ogataea	Ogataea parapolyomorpha	70	0.0890
Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Pichia	Pichia kudriavzevii	35	0.0445
Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Brettanomyces	Brettanomyces nanus	4	0.0051
Ascomycota	Saccharomycetes	Saccharomycetales	Trichomonascaceae	Sugiyamaella	Sugiyamaella lignohabitans	88	0.1119
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella phaffii	44	0.0560
Ascomycota	Saccharomycetes	Saccharomycetales	Dipodascaceae	Yarrowia	Yarrowia lipolytica	8	0.0102

Ascomycota	Sordariomycetes	Hypocreales	Nectriaceae	Fusarium	Fusarium pseudograminearum	72	0.0916
Ascomycota	Sordariomycetes	Hypocreales	Nectriaceae	Fusarium	Fusarium venenatum	10	0.0127
Ascomycota	Sordariomycetes	Hypocreales	Nectriaceae	Fusarium	Fusarium graminearum	5	0.0064
Ascomycota	Sordariomycetes	Hypocreales	Nectriaceae	Fusarium	Fusarium verticillioides	69	0.0878
Ascomycota	Sordariomycetes	Hypocreales	Nectriaceae	Fusarium	Fusarium fujikuroi	5	0.0064
Ascomycota	Sordariomycetes	Hypocreales	Nectriaceae	Fusarium	Fusarium oxysporum	19	0.0242
Ascomycota	Sordariomycetes	Hypocreales	Ophiocordycipitaceae	Drechmeria	Drechmeria coniospora	23	0.0293
Ascomycota	Sordariomycetes	Hypocreales	Clavicipitaceae	Pochonia	Pochonia chlamydosporia	3	0.0038
Ascomycota	Sordariomycetes	Glomerellales	Glomerellaceae	Colletotrichum	Colletotrichum higginsianum	86	0.1094
Ascomycota	Sordariomycetes	Magnaporthales	Pyriculariaceae	Pyricularia	Pyricularia pennisetigena	50	0.0636
Ascomycota	Sordariomycetes	Magnaporthales	Pyriculariaceae	Pyricularia	Pyricularia grisea	9	0.0114
Ascomycota	Sordariomycetes	Magnaporthales	Pyriculariaceae	Pyricularia	Pyricularia oryzae	3	0.0038
Ascomycota	Sordariomycetes	Sordariales	Sordariaceae	Neurospora	Neurospora crassa	24	0.0305
Ascomycota	Sordariomycetes	Sordariales	Chaetomiaceae	Thermothelomyces	Thermothelomyces thermophilus	10	0.0127
Ascomycota	Sordariomycetes	Sordariales	Chaetomiaceae	Thermothielavioides	Thermothielavioides terrestris	9	0.0114
Ascomycota	Leotiomycetes	Helotiales	Sclerotiniaceae	Botrytis	Botrytis cinerea	13	0.0165
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Aspergillus	Aspergillus fumigatus	18	0.0229
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Aspergillus	Aspergillus oryzae	5	0.0064
Ascomycota	Eurotiomycetes	Eurotiales	Trichocomaceae	Talaromyces	Talaromyces rugulosus	21	0.0267
Ascomycota	Dothideomycetes	Mycosphaerellales	Mycosphaerellaceae	Cercospora	Cercospora beticola	13	0.0165
Ascomycota	Dothideomycetes	Mycosphaerellales	Mycosphaerellaceae	Zymoseptoria	Zymoseptoria tritici	3	0.0038
Ascomycota	Schizosaccharomycetes	Schizosaccharomycetales	Schizosaccharomycetaceae	Schizosaccharomyces	Schizosaccharomyces pombe	32	0.0407
Basidiomycota	Malasseziomycetes	Malasseziales	Malasseziaceae	Malassezia	Malassezia restricta	79	0.1005
Basidiomycota	Ustilaginomycetes	Ustilaginales	Ustilaginaceae	Sporisorium	Sporisorium graminicola	23	0.0293
Basidiomycota	Ustilaginomycetes	Ustilaginales	Ustilaginaceae	Ustilago	Ustilago maydis	12	0.0153
Basidiomycota	Tremellomycetes	Tremellales	Cryptococcaceae	Cryptococcus	Cryptococcus neoformans	12	0.0153
Basidiomycota	Tremellomycetes	Tremellales	Cryptococcaceae	Cryptococcus	Cryptococcus gattii VGI	6	0.0076
Microsporidia	Unikaryonidae	Encephalitozoon	Encephalitozoon romaleae			1	0.0013
					Total	78608	----

Tabla 7. Hongos y levaduras asignadas con Kaiju

Phylum	Clase	Orden	Familia	Género	Especie	Conteo	Abundancia
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Debaryomyces	Debaryomyces hansenii	47683	47.4453
Mucoromycota	Glomeromycetes	Glomerales	Glomeraceae	Rhizophagus	Rhizophagus sp. MUCL 43196	11148	11.0924
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Debaryomyces	Debaryomyces fabryi	5279	5.2527
Mucoromycota	Mucoromycetes	Mucorales	Mucoraceae	Mucor	Mucor ambiguus	3317	3.3005
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Lachancea	Lachancea nothofagi	1222	1.2159
Zoopagomycota	Harpellomycetes	Harpellales	Harpellaceae	Furculomyces	Furculomyces boomerangus	806	0.8020
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycodaceae	Hanseniaspora	Hanseniaspora guilliermondii	594	0.5910
Ascomycota	Saccharomycetes	Saccharomycetales	Trichomonascaceae	Sugiyamaella	Sugiyamaella lignohabitans	455	0.4527
Basidiomycota	Agaricomycetes	Corticiales	Punctulariaceae	Punctularia	Punctularia strigosozonata	423	0.4209
Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Bipolaris	Bipolaris oryzae	371	0.3692
Basidiomycota	Exobasidiomycetes	Entylomatales	Entylomatales incertae sedis	Tilletiopsis	Tilletiopsis washingtonensis	363	0.3612
Ascomycota	Sordariomycetes	Coniochaetales	Coniochaetaceae	Coniochaeta	Coniochaeta ligniaria	260	0.2587
Basidiomycota	Exobasidiomycetes	Exobasidiales	Cryptobasidiaceae	Acaromyces	Acaromyces ingoldii	214	0.2129
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Saccharomyces	Saccharomyces cerevisiae	199	0.1980
Ascomycota	Leotiomycetes	Helotiales	Sclerotiniaceae	Sclerotinia	Sclerotinia borealis	192	0.1910
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Scheffersomyces	Scheffersomyces stipitis	179	0.1781
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Wickerhamomyces	Wickerhamomyces anomalus	175	0.1741
Ascomycota	Sordariomycetes	Ophiostomatales	Ophiostomataceae	Sporothrix	Sporothrix schenckii	163	0.1622
Ascomycota	Eurotiomycetes	Chaetothyriales	unclassified Chaetothyriales	Chaetothyriales	Chaetothyriales sp. CBS 134920	157	0.1562
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Aspergillus	Aspergillus fumigatus	138	0.1373
Basidiomycota	Exobasidiomycetes	Georgefischeriales	Tilletiariaceae	Tilletiaria	Tilletiaria anomala	133	0.1323
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Phialosimplex	Phialosimplex sp. HF37	119	0.1184
Basidiomycota	Pucciniomycetes	Pucciniales	Pucciniaceae	Puccinia	Puccinia coronata	114	0.1134
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Millerozyma	Millerozyma farinosa	110	0.1095
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Spathaspora	Spathaspora sp. JA1	106	0.1055
Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Ogataea	Ogataea polymorpha	98	0.0975
Ascomycota	Dothideomycetes	Capnodiales	Teratosphaeriaceae	Baudoinia	Baudoinia panamericana	96	0.0955
Microsporidia			Enterocytozoonidae	Enterocytozoon	Enterocytozoon hepatopenaei	94	0.0935
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Debaryomyces	[Candida] psychrophila	92	0.0915
Basidiomycota	Agaricomycetes	Boletales	Coniophoraceae	Coniophora	Coniophora puteana	90	0.0896
Ascomycota	Eurotiomycetes	Onygenales	Ajellomycetaceae	Blastomyces	Blastomyces gilchristii	87	0.0866

Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Kluyveromyces	Kluyveromyces lactis	86	0.0856
Ascomycota	Dothideomycetes	Capnodiales	Teratosphaeriaceae	Hortaea	Hortaea werneckii	85	0.0846
Basidiomycota	Agaricomycetes	Hymenochaetales	Schizoporaceae	Schizopora	Schizopora paradoxa	84	0.0836
Basidiomycota	Agaricomycetes	Agaricales	Tricholomataceae	Laccaria	Laccaria bicolor	84	0.0836
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Spathaspora	Spathaspora passalidarum	83	0.0826
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Hyphopichia	Hyphopichia burtonii	81	0.0806
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Pachysolen	Pachysolen tannophilus	79	0.0786
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Aspergillus	Aspergillus steynii	78	0.0776
Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Pichia	Pichia membranifaciens	74	0.0736
Basidiomycota	Agaricomycetes	Polyporales	Polyporaceae	Polyporus	Polyporus arcularius	73	0.0726
Ascomycota	Taphrinomycotina incertae sedis			Saitoella	Saitoella complicata	72	0.0716
Ascomycota	Eurotiomycetes	Chaetothyriales	Herpotrichiellaceae	Capronia	Capronia epimyces	72	0.0716
Basidiomycota	Agaricomycetes	Polyporales	Meruliaceae	Phlebia	Phlebia centrifuga	72	0.0716
Ascomycota	Sordariomycetes	Glomerellales	Glomerellaceae	Colletotrichum	Colletotrichum salicis	72	0.0716
Basidiomycota	Tremellomycetes	Cystofilobasidiales	Mrkakiaceae	Xanthophyllomyces	Xanthophyllomyces dendrorhous	71	0.0706
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Meyerozyma	Meyerozyma guilliermondii	70	0.0697
Basidiomycota	Malasseziomycetes	Malasseziales	Malasseziaceae	Malassezia	Malassezia restricta	69	0.0687
Ascomycota	Saccharomycetes	Saccharomycetales	Lipomycetaceae	Lipomyces	Lipomyces starkeyi	67	0.0667
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Suhomyces	Suhomyces tanzawaensis	66	0.0657
Chytridiomycota	Neocallimastigomycetes	Neocallimastigales	Neocallimastigaceae	Piromyces	Piromyces sp. E2	66	0.0657
Ascomycota	Dothideomycetes	Pleosporomycetidae incertae sedis	Gloniaceae	Glonium	Glonium stellatum	63	0.0627
Ascomycota	Sordariomycetes	Hypocreales	Clavicipitaceae	Claviceps	Claviceps purpurea	62	0.0617
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Candida	Candida albicans	62	0.0617
Mucoromycota	Glomeromycetes	Glomerales	Glomeraceae	Rhizophagus	Rhizophagus irregularis	62	0.0617
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Babjeviella	Babjeviella inositovora	60	0.0597
Basidiomycota	Dacrymycetes	Dacrymycetales	Dacrymycetaceae	Calocera	Calocera viscosa	60	0.0597
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Aspergillus	Aspergillus aculeatinus	59	0.0587
Basidiomycota	Agaricomycetes	Agaricales	Lyophyllaceae	Termitomyces	Termitomyces sp. J132	58	0.0577
Ascomycota	Sordariomycetes	Xylariales	Xylariales incertae sedis	Monosporascus	Monosporascus sp. GIB2	58	0.0577
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Nakaseomyces	[Candida] glabrata	56	0.0557
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium italicum	54	0.0537
Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Clavispora	[Candida] auris	54	0.0537
Basidiomycota	Tremellales	Cryptococcaceae	Cryptococcus	Cryptococcus neoformans species complex	Cryptococcus neoformans	52	0.0517
Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodotorula	Rhodotorula toruloides	52	0.0517
Ascomycota	Lecanoromycetes	Teloschistales	Teloschistaceae	Xanthoria	Xanthoria polycarpa	52	0.0517
Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Metschnikowia	Metschnikowia bicuspidata	51	0.0507
Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Clavispora	Clavispora lusitaniae	51	0.0507
Ascomycota	Sordariomycetes	Hypocreales	Hypocreaceae	Trichoderma	Trichoderma atroviride	51	0.0507

Mucoromycota	Glomeromycetes	Diversisporales	Gigasporaceae	Gigaspora	Gigaspora rosea	50	0.0498
Ascomycota	Sordariomycetes	Glomerellales	Plectosphaerellaceae	Verticillium	Verticillium longisporum	50	0.0498
Ascomycota	Sordariomycetes	Glomerellales	Glomerellaceae	Colletotrichum	Colletotrichum tofieldiae	50	0.0498
Ascomycota	Sordariomycetes	Hypocreales	Ophiocordycipitaceae	Purpureocillium	Purpureocillium lilacinum	49	0.0488
Basidiomycota	Agaricomycetes	Cantharellales	Ceratobasidiaceae	Rhizoctonia	Rhizoctonia solani	47	0.0468
Ascomycota	Sordariomycetes	Hypocreales	Hypocreaceae	Trichoderma	Trichoderma parareesei	44	0.0438
Ascomycota	Sordariomycetes	Glomerellales	Glomerellaceae	Colletotrichum	Colletotrichum sublineola	43	0.0428
Chytridiomycota	Neocallimastigomycetes	Neocallimastigales	Neocallimastigaceae	Neocallimastix	Neocallimastix californiae	42	0.0418
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Yamadazyma	Yamadazyma tenuis	42	0.0418
Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Clavispora	[Candida] intermedia	40	0.0398
Basidiomycota	Agaricomycetes	Sebacinales	Serendipitaceae	Serendipita	Serendipita vermifera	40	0.0398
Ascomycota	Sordariomycetes	Coniochaetales	Coniochaetaceae	Coniochaeta	Coniochaeta pulveracea	40	0.0398
Ascomycota	Pezizomycetes	Pezizales	Tuberaceae	Tuber	Tuber magnatum	39	0.0388
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Aspergillus	Aspergillus nidulans	39	0.0388
Chytridiomycota	Chytridiomycetes	Chytridiales	Chytriomycetaceae	Rhizoclosmatium	Rhizoclosmatium globosum	38	0.0378
Mucoromycota	Glomeromycetes	Diversisporales	Diversisporaceae	Diversispora	Diversispora epigaea	38	0.0378
Ascomycota	Sordariomycetes	Glomerellales	Glomerellaceae	Colletotrichum	Colletotrichum graminicola	36	0.0358
Basidiomycota	Agaricomycetes	Polyporales	Polyporaceae	Ganoderma	Ganoderma sinense	36	0.0358
Ascomycota	Eurotiomycetes	Eurotiales	Trichocomaceae	Talaromyces	Talaromyces amestolkiae	36	0.0358
Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Pichia	Pichia kudriavzevii	33	0.0328
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Aspergillus	Aspergillus terreus	33	0.0328
Ascomycota	Sordariomycetes	Xylariales	Xylariaceae	Xylaria	Xylaria longipes	33	0.0328
Ascomycota	Sordariomycetes	Sordariales	Chaetomiaceae	Thermothelomyces	Thermothelomyces thermophilus	32	0.0318
Chytridiomycota	Chytridiomycetes	Spizellomycetales	Spizellomycetaceae	Spizellomyces	Spizellomyces punctatus	32	0.0318
Ascomycota	Leotiomycetes	Helotiales	Dermateaceae	Marssonina	Marssonina brunnea	32	0.0318
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella phaffii	31	0.0308
Basidiomycota	Agaricomycetes	Agaricales	Psathyrellaceae	Psathyrella	Psathyrella aberdarensis	30	0.0299
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Aspergillus	Aspergillus niger	29	0.0289
Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Ogataea	[Candida] boidinii	29	0.0289
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Aspergillus	Aspergillus heteromorphus	29	0.0289
Ascomycota	Sordariomycetes	Ophiostomatales	Ophiostomataceae	Sporothrix	Sporothrix insectorum	29	0.0289
Cryptomycota				Paramicrosporidium	Paramicrosporidium saccamoebae	29	0.0289
Ascomycota	Dothideomycetes	Botryosphaeriales	Botryosphaeriaceae	Diplodia	Diplodia seriata	28	0.0279
Ascomycota	Sordariomycetes	Hypocreales	Ophiocordycipitaceae	Ophiocordyceps	Ophiocordyceps sp. 'camponoti-saundersi'	28	0.0279
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Candida	Candida parapsilosis	27	0.0269
Mucoromycota	Mucoromycetes	Mucorales	Cunninghamellaceae	Absidia	Absidia glauca	26	0.0259
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales incertae sedis	Candida	Candida sake	25	0.0249

Basidiomycota	Exobasidiomycetes	Microstromatales	Microstromatales incertae sedis	Jaminaea	Jaminaea rosea	25	0.0249
Ascomycota	Eurotiomycetes	Eurotiales	Elaphomycetaceae	Elaphomyces	Elaphomyces granulatus	24	0.0239
Ascomycota	Eurotiomycetes	Onygenales	Ajellomycetaceae	Blastomyces	Blastomyces dermatitidis	23	0.0229
Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	Parastagonospora	Parastagonospora nodorum	23	0.0229
Mucoromycota	Mucoromycetes	Mucorales	Lichtheimiaceae	Lichtheimia	Lichtheimia corymbifera	23	0.0229
Ascomycota	Sordariomycetes	Sordariales	Lasiosphaeriaceae	Podospora	Podospora comata	23	0.0229
Ascomycota	Sordariomycetes	Hypocreales	Nectriaceae	Neonectria	Neonectria ditissima	23	0.0229
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium antarcticum	23	0.0229
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Aspergillus	Aspergillus taichungensis	23	0.0229
Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Komagataella	Komagataella pastoris	22	0.0219
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Candida	Candida maltosa	22	0.0219
Ascomycota	Leotiomycetes	Helotiales	Helotiaceae	Glarea	Glarea lozoyensis	22	0.0219
Ascomycota	Sordariomycetes	Hypocreales	Cordycipitaceae	Beauveria	Beauveria brongniartii	22	0.0219
Ascomycota	Sordariomycetes	Hypocreales	Ophiocordycipitaceae	Ophiocordyceps	Ophiocordyceps polyrhachis-furcata	22	0.0219
Basidiomycota	Tremellomycetes	Tremellales	Tremellaceae	Tremella	Tremella mesenterica	21	0.0209
Blastocladiomycota	Blastocladiomycetes	Blastocladales	Blastocladiaceae	Allomyces	Allomyces macrognus	21	0.0209
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium freii	21	0.0209
Basidiomycota	Tremellomycetes	Trichosporonales	Trichosporonaceae	Apiotrichum	Apiotrichum porosum	21	0.0209
Basidiomycota	Agaricomycetes	Agaricales	Omphalotaceae	Gymnopus	Gymnopus luxurians	21	0.0209
Ascomycota	Leotiomycetes	Helotiales	Dermateaceae	Diplocarpon	Diplocarpon rosae	21	0.0209
Basidiomycota	Ustilaginomycetes	Violaceomycetales	Violaceomycetaceae	Violaceomyces	Violaceomyces palustris	21	0.0209
Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Meyerozyma	Meyerozyma sp. JA9	21	0.0209
Basidiomycota	Agaricomycetes	Auriculariales	Exidiaceae	Exidia	Exidia glandulosa	20	0.0199
Ascomycota	Eurotiomycetes	Onygenales	Onygenales incertae sedis	Coccidioides	Coccidioides immitis	20	0.0199
Ascomycota	Eurotiomycetes	Eurotiales	Trichocomaceae	Talaromyces	Talaromyces marneffeii	20	0.0199
Basidiomycota	Agaricomycetes	Agaricales	Cortinariaceae	Hebeloma	Hebeloma cylindrosporum	20	0.0199
Ascomycota	Sordariomycetes	Sordariales	Sordariales incertae sedis	Madurella	Madurella mycetomatis	20	0.0199
Basidiomycota	Agaricomycetes	Trechisporales	Hydnodontaceae	Sistotremastrum	Sistotremastrum niveocreum	20	0.0199
Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales incertae sedis	Kuraishia	Kuraishia capsulata	20	0.0199
Mucoromycota	Glomeromycetes	Glomerales	Glomeraceae	Glomus	Glomus cerebriforme	20	0.0199
Basidiomycota	Agaricomycetes	Auriculariales	Auriculariaceae	Auricularia	Auricularia subglabra	20	0.0199
Ascomycota	Sordariomycetes	Xylariales	Xylariales incertae sedis	Monosporascus	Monosporascus sp. 5C6A	20	0.0199
Ascomycota	Sordariomycetes	Hypocreales	Nectriaceae	Fusarium	Fusarium oxysporum	19	0.0189
Ascomycota	Sordariomycetes	Magnaporthales	Magnaporthaceae	Gaeumannomyces	Gaeumannomyces tritici	19	0.0189
Mucoromycota	Mortierellomycetes	Mortierellales	Mortierellaceae	Mortierella	Mortierella verticillata	19	0.0189
Basidiomycota	Microbotryomycetes	Leucosporidiales	Leucosporidiaceae	Leucosporidium	Leucosporidium creatinivorum	19	0.0189
Basidiomycota	Agaricomycetes	Russulales	Bondarzewiaceae	Heterobasidion	Heterobasidion irregulare	19	0.0189
Ascomycota	Xylonomycetes	Xylonales	Xylonaceae	Xylona	Xylona heveae	19	0.0189
Ascomycota	Leotiomycetes	Leotiomycetes incertae sedis	Leotiomycetes incertae sedis	Phialophora	Phialophora cf. hyalina BP 5553	19	0.0189

