



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Maestría y Doctorado en Ciencias Bioquímicas

“CARACTERIZACIÓN GENÓMICA, POBLACIONAL Y FUNCIONAL DE
ENTEROCOCCUS AISLADOS DE UN ALIMENTO ARTESANAL”

TESIS

QUE PARA OPTAR POR EL GRADO DE:

Doctor en Ciencias

PRESENTA:

M. en C. MYRNA ELENA OLVERA GARCÍA

TUTOR PRINCIPAL

DRA. MARICARMEN QUIRASCO BARUCH
[Facultad de Química](#)

MIEMBROS DEL COMITÉ TUTOR

DRA. ROSA MARÍA GUTIÉRREZ RÍOS
[Instituto de Biotecnología](#)

DRA. ROMINA RODRÍGUEZ SANOJA
[Instituto de Investigaciones Biomédicas](#)

Ciudad de México. Septiembre, 2018



Universidad Nacional
Autónoma de México



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.

“Caracterización genómica, poblacional y funcional de *Enterococcus* aislados de un alimento artesanal”

Reconocimientos

Esta tesis de doctorado se realizó bajo la tutoría de la Dra. Maricarmen Quirasco Baruch en el laboratorio 312 del conjunto E, Departamento de Alimentos y Biotecnología de la Facultad de Química en la Universidad Nacional Autónoma de México.

Se reconocen las siguientes colaboraciones:

Al Dr. Alejandro Sánchez Flores por la dirección y tutoría en el análisis bioinformático y el apoyo técnico de la Dra. Alejandra Escobar, MC Verónica Jiménez Jacinto, MC Jérôme Verleyen y MC Karel Estrada Guerra de la Unidad de Secuenciación Masiva y Bioinformática del Instituto de Biotecnología de la UNAM.

Al Q.A. Andrés Guzmán Ovalle por la colaboración en el análisis poblacional de cepas de *E. faecium* aisladas de queso Cotija.

El Comité Tutorial que asesoró el desarrollo de esta tesis estuvo formado por:

Dra. Rosa María Gutiérrez Ríos Instituto de Biotecnología, UNAM

Dra. Romina Rodríguez Sanoja Instituto de Investigaciones Biomédicas, UNAM

Durante la realización de este trabajo, se contó con el apoyo financiero de los siguientes proyectos: DGAPA-UNAM-PAPIIT IN222115 y IN222717, PAIP-FQ 5000-9102, CONACyT (Beca de doctorado), Programa de Apoyo a los Estudios del Posgrado (asistencia a congreso nacional e internacional).

El jurado asignado:

Presidente Dra. Esperanza Martínez Romero Centro de Ciencias Genómicas, UNAM

Vocal Dra. Clarita Olvera Carranza Instituto de Biotecnología, UNAM

Vocal Dra. Yolanda López Vidal Facultad de Medicina, UNAM

Vocal Dra. Amelia María de Guadalupe
Farrés González Sarabia Facultad de Química, UNAM

Secretario Dr. Francisco Ruíz Terán Facultad de Química, UNAM

Contenido

1. RESUMEN	7
2. MARCO TEÓRICO	9
2.1 <i>Enterococcus</i> spp.	10
2.2 <i>Enterococcus</i> en alimentos y su potencial biotecnológico	10
2.2.1 Metabolismo de citrato por enterococos	11
2.2.2 Proteólisis	11
2.2.3 Lipólisis.	12
2.2.4 Compuestos con actividad antibacteriana.	12
2.3 Potencial de patogenicidad del género <i>Enterococcus</i>	14
2.4 Era de la genómica en el estudio de <i>Enterococcus</i>	16
3. ANTECEDENTES	20
3.1 Microbiología del queso Cotija	20
4. JUSTIFICACIÓN	22
5. HIPÓTESIS	23
6. OBJETIVO GENERAL	23
5.1 OBJETIVOS PARTICULARES	23
7. METODOLOGÍA GENERAL	24
7.1 Determinación de estructura poblacional (MLST)	24
7.2 Análisis Genómicos	26
7.2.1 Extracción de ADN	26
7.2.2 Secuenciación de genomas y análisis genómico	27
7.2.3 Análisis de ortología	28
7.2.4 Anotación funcional	28
7.2.5 Detección de genes relacionados con factores de virulencia y con síntesis de aminas biógenas	29
7.2.6 Identificación de genes asociados al desarrollo de aromas y sistemas proteolítico y lipolítico	29
7.3 Pruebas de susceptibilidad a antibióticos	29
7.4 Elección y evaluación de posibles biomarcadores	30
7.5 Rearmado de genomas a partir del metagenoma y búsqueda de marcadores	31
8. RESULTADOS Y DISCUSIÓN	32
8.1 ¿En qué se diferencian las cepas de enterococos aisladas de queso Cotija de cepas de otros ambientes?	32
Estructura poblacional de <i>E. faecalis</i> y <i>E. faecium</i> aisladas de queso Cotija	32
8.1.1 MLST y estructura poblacional de <i>E. faecalis</i>	32

8.1.2 MLST y estructura poblacional de <i>E. faecium</i>	39
8.1.3 Secuenciación, ensamble y anotación de los genomas	45
8.1.4 Genómica comparativa	46
8.1.5 Descripción del potencial de patogenicidad en las cepas del QC	51
8.2 Potencial funcional del género <i>Enterococcus</i> en el queso Cotija	54
8.2.1 Sistema proteolítico, lipolítico y desarrollo de aromas y sabores	54
8.2.2 Producción de compuestos con actividad antibacteriana	60
8.3 Propuesta de genes marcadores de adaptación a alimentos	63
8.3.1 Estrategia de evaluación de biomarcadores usando el ensamble de metagenoma	67
9. CONCLUSIONES	68
10. PERSPECTIVAS	68
11. BIBLIOGRAFIA	69
12. Anexos	78
12.1 Anexo I	78
12.2 Anexo II	79
12.3 Anexo III	82
12.4 Anexo IV	83

FIGURAS

Fig 1. Modelo de sinergia para causar una infección por el género <i>Enterococcus</i> .	15
Fig 2. Tecnologías de nueva generación de secuenciación.	17
Fig 3. Queso Cotija artesanal madurado y zona de producción en la sierra de Jalmich.	20
Fig 4. Anotación taxonómica de las principales fila bacterianas identificadas a través del marcador filogenético 16S ADNr.	22
Fig 5. Metodología general para el desarrollo del proyecto	24
Fig 6. Esquema de metodología de MLST.	26
Fig 7. Filogenia de los genes <i>gdh</i> y <i>gyd</i> de <i>E. faecalis</i> y <i>E. faecium</i> .	33
Fig 8. Alineamiento de las secuencias del gen <i>gdh</i> <i>E. faecium</i> y <i>E. faecalis</i> .	34
Fig 9. Análisis de estructura poblacional de <i>E. faecalis</i> .	36
Fig 10. Árbol de agrupación de cepas de <i>E. faecalis</i> .	38
Fig 11. Análisis de estructura poblacional de <i>E. faecium</i> .	41
Fig 12. Árbol de agrupación de cepas de <i>E. faecium</i> .	43
Fig 13. Identificación de genes que codifican para factores de virulencia.	45
Fig 14. Alineamiento de genomas de <i>E. faecalis</i> provenientes de diferentes ambientes.	48
Fig 15. Alineamiento de genomas de <i>E. faecium</i> provenientes de diferentes ambientes.	49
Fig 16. Análisis de familias de proteínas expandidas en ambas especies.	51
Fig 17. Propuesta de sistema proteolítico en las cepas de <i>Enterococcus</i> spp.	58
Fig 18. Genes codificantes para bacteriocinas.	61
Fig 19. Identificación de proteínas con actividad antibacteriana.	62
Fig 20. Identificación enterocina A en cepas aisladas del queso Cotija.	65
Fig 21. Análisis de regiones CRISPR y CRISPR-cas.	66
Fig AI. Reordenamiento del ensamble de <i>E. faecalis</i> usando como referencia el genoma de <i>E. faecalis</i> Symbioflor 1	82
Fig AII. Reordenamiento del ensamble de <i>E. faecium</i> D usando la referencia NRRL B-2354	82

TABLAS

Tabla 1. Cebadores y condiciones para realizar PCR de biomarcadores	30
Tabla 2. Resultados de número de alelo y secuencia tipo para las cepas de <i>E. faecalis</i> del queso Cotija	35
Tabla 3. Resultados de número de alelo y secuencia tipo para las cepas de <i>E. faecium</i> del queso Cotija	39
Tabla 4. Comparación de características genómicas de las cepas <i>E. faecalis</i> F y <i>E. faecium</i> D contra las referencias	46
Tabla 5. Genes involucrados en el sistema proteolítico de las cepas <i>E. faecium</i> D y <i>E. faecalis</i> F	54
Tabla 6. Genes con anotación de actividad PGH	62
Tabla 7. Rutas metabólicas asociadas a los genes únicos (KEEG)	63
Tabla 8. Anotación de los genes únicos encontrados en los ensamblajes generados a partir del metagenoma	67
Tabla AI. Lista de cepas utilizadas en este trabajo	78
Tabla AII. Información de cepas utilizadas para construir el árbol filogenético de <i>E. faecalis</i>	79
Tabla AIII. Información de cepas utilizadas para construir el árbol filogenético de <i>E. faecium</i>	80
Tabla AIV. Anotación funcional de los genes codificantes de proteína presentes en los plásmidos de <i>E. faecalis</i> F	83

ABREVIATURAS

BAL: Bacteria Ácido Láctica

BLAST: Basic Local Alignment Search Tool

BRIG: BLAST Ring Image Generator

CC: Complejo Clonal

COG: Clusters of Orthologous Groups

CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats

EGM: Elementos Genéticos Móviles

GRAS: Generalmente Reconocido como Seguro

HPLC: Cromatografía de Líquidos de Alta Eficacia

MLST: Tipificación de secuencias multilocus

PGHs: Peptidoglucano hidrolasas

ST: Secuencia Tipo

QC: queso Cotija

I+D: Investigación y desarrollo

GLOSARIO

Alelo: es cada una de las formas alternativas que puede tener un mismo gen y que se diferencian en su secuencia y pueden manifestar modificaciones concretas de la función de ese gen.

Clado: se denomina en la biología a cada una de las ramificaciones que se obtiene después de hacer un único corte en el árbol filogenético.

Cluster: un conjunto de genes homólogos dentro un organismo.

Contig: son segmentos de ADN superpuestos, que juntos representan una región consenso de ADN.

Core: genes presentes en todos los organismos que se están comparando.

Ensamble: se denomina al alineamiento y mezcla de múltiples fragmentos de una secuencia de ADN mucho mayor para reconstruir la secuencia original.

Fila (plural de filum): es una categoría en taxonomía situada entre el reino y la clase, y usada en los reinos animales (que contiene 35 filos), fungi (que contiene 6), protistas y dominio bacterias.

Mapeo: reordenamiento de secuencias cortas o largas teniendo un genoma de referencia.

Scaffolding: unión de una serie no contigua de secuencias genómicas en un andamio, que consiste en secuencias separadas por huecos de longitud conocida. Las secuencias que están unidas son comunmente secuencias contiguas que corresponden a superposiciones.

*Fuente: https://en.wikipedia.org/wiki/Main_Page

1. RESUMEN

Los *Enterococcus* están presentes en la microbiota nativa de muchos alimentos fermentados tradicionales distribuidos alrededor del mundo. Se caracterizan por producir compuestos con actividad antibacteriana, principalmente activos contra *Listeria monocytogenes*, lo cual ha incrementado el interés en estudiar este género. A pesar de su constante aparición en los alimentos artesanales, es poca la información disponible acerca de sus actividades proteolíticas, lipolíticas, así como su potencial aplicación biotecnológica, probablemente debido a su reciente asociación con brotes de infecciones nosocomiales, que en la última década lo ubica como el tercer Gram-positivo causante de infecciones intrahospitalarias.

En este trabajo se realizaron diferentes estrategias genómicas para poder generar información que diferenciara cepas de *E. faecalis* y *E. faecium* aisladas de queso Cotija artesanal madurado de cepas nosocomiales. A través de la tipificación de secuencias multilocus, se definió la no pertenencia de las cepas de enterococos aisladas del queso Cotija con los Complejos Clonales de alto riesgo. Así mismo, se propuso al CC 72 como un complejo enriquecido por cepas de *E. faecalis* aisladas de alimentos. El análisis de genómica comparativa, empleando los genomas completos (secuenciados y ensamblados en este trabajo), de dos cepas, una de *E. faecalis* y una de *E. faecium*, nos permitió evidenciar las discrepancias entre cepas nosocomiales y las presentes en otros nichos ecológicos, principalmente en alimentos. Además, se propusieron regiones genómicas características de cepas adaptadas a un alimento: genes que codifican para enzimas relacionadas con el sistema proteolítico, un sistema de inmunidad bacteriano (CRISPR-cas) y enterocinas.

Por otro lado, el armado de los genomas, generando drafts de alta calidad y su anotación funcional, ayudó a plantear un sistema proteolítico hipotético de las especies *faecalis* y *faecium*, detectar genes codificantes de actividades lipolíticas (lipasas y esterases) así como enzimas relacionadas a la producción de metabolitos que contribuyen a la formación de aromas y sabores característicos del queso Cotija. Además, se identificaron genes codificantes para bacteriocinas y peptidoglucano hidrolasas, compuestos que pueden tener actividad antimicrobiana. Por otro lado, se hizo una descripción del potencial de patogenicidad de ambas cepas, concluyendo que carecen de las islas de patogenicidad que previamente han sido reportadas en cepas nosocomiales. *E. faecalis* y *E. faecium* aisladas del queso Cotija; presentan únicamente la resistencia a antibióticos intrínseca del género, así como genes

relacionados con la capacidad de colonizar el tracto gastrointestinal, la formación de biopelículas y la agregación celular.

En conclusión, las cepas de enterococos presentes en el queso Cotija Región de Origen presentan regiones genómicas frecuentemente ausentes en los aislados nosocomiales. Dichas regiones, se componen de genes que codifican para bacteriocinas, un sistema de inmunidad bacteriano que regula la entrada de ADN exógeno y algunos genes relacionados con el metabolismo de carbohidratos. Sin embargo, la principal diferencia radica en la ausencia de factores de virulencia que puedan dañar al hospedero y resistencia reducida a antibióticos, intrínseca y/o adquirida.

2. MARCO TEÓRICO

El género *Enterococcus* está formado, hasta el momento, por 49 especies (Zhong et al., 2017). A pesar de que a este género se le considera ubicuo, su principal nicho ecológico es el tracto gastrointestinal de los mamíferos, por lo que algunas cepas han sido empleadas como probióticos para mejorar o promover la salud de los humanos y animales (Franz et al., 2011).

Enterococcus también se puede encontrar, con alta frecuencia, formando parte de otros ambientes como plantas, suelo, agua y alimentos (Giraffa, 2003). Pertenece al grupo de las bacterias ácido lácticas (BAL), y está presente durante la producción de quesos artesanales, principalmente del Mediterráneo, elaborados con leche bronca o pasteurizada (Feta, Cheddar, Manchego, Mozzarella y Venaco), y se ha propuesto que son clave en el proceso de maduración, debido a sus actividades de proteólisis, lipólisis y ruptura del citrato, que son indispensables para el desarrollo de aromas y sabores (Franz et al, 2011). Además, cepas de enterococos aisladas de alimentos han resultado de gran interés biotecnológico debido a su capacidad productora de compuestos con actividad antibacteriana (Giraffa, 2002). A pesar de ello, su papel durante el proceso de maduración, relacionado a las actividades enzimáticas mencionadas previamente, ha sido escasamente descrito. En la actualidad, la secuenciación de genomas de bacterias y su análisis bioinformático es una herramienta muy importante, ya que se pueden conocer todas las secuencias codificantes y con ello inferir las rutas metabólicas completas, y así tener un panorama más amplio de su participación en el proceso de fermentación de un alimento (van Kranenburg et al., 2002).

Cepas de este género han sido asociadas a infecciones nosocomiales causando enfermedades como sepsis, endocarditis e infecciones del tracto urinario. Aunado a esto, algunas cepas son resistentes a una amplia gama de antibióticos (Ogier y Serror, 2008; Franz et al., 2001). Sin embargo, se ha planteado que la incidencia de dichos factores de virulencia es menor en aislados de alimentos, agua y vegetales, comparada con la presentada por aislados provenientes de infecciones clínicas (Eaton y Gasson, 2001). Estudios basados en las características genómicas han mostrado que *E. faecium* y *E. faecalis* exhiben una gran diversidad inter-cepa asociada, principalmente, a la presencia de fagos, plásmidos, islas de patogenicidad y elementos conjugativos (Kim y Marco, 2014). Por lo tanto, la divergencia entre cepas de origen nosocomial y de otros ambientes podría ser el resultado de una adaptación nicho-específica (van Schaik y Willems, 2010).

2.1 *Enterococcus* spp.

Los enterococos son cocos Gram-positivos, que pueden encontrarse solos, en pares o cadenas cortas, son anaerobios facultativos y pueden crecer en medio sólido expuesto al aire, son no esporulados, generalmente no móviles, catalasa, reductasa y oxidasa negativos, pero algunas cepas pueden presentar actividad de pseudocatalasa (Salminen, 2004). El género *Enterococcus* forma parte del grupo de las BAL, las cuales se caracterizan por producir ácido láctico, como componente mayoritario, de la fermentación de carbohidratos (Giraffa, 2003). En la clasificación de BAL, por su capacidad de fermentación de azúcares, los enterococos se encuentran dentro del grupo III junto con los pediococos, estreptococos, tetragenococos, algunos lactobacilos y vagococos. Se consideran heterofermentativos facultativos ya que en presencia de hexosas se comportan como homofermentativos, sin embargo, cuando la fuente de carbono está constituida por pentosas o compuestos parecidos (gluconato), se induce la síntesis de fosfocetolasas, dando como resultado la producción de ácido acético, CO₂, etc., por una fermentación heteroláctica (Salminen et al., 2004). Tienen una capacidad biosintética muy limitada, por lo que requieren medios de cultivo ricos en vitaminas del complejo B, diversos aminoácidos, bases púricas y pirimídicas para crecer adecuadamente (Garg y Mital, 1991). A pesar de pertenecer a las BAL, sólo seis especies del género *Enterococcus*, incluyendo *E. faecium* y excluyendo a *E. faecalis* son reconocidas como “generalmente reconocido como seguro” (GRAS) y se encuentran dentro del manual de la AAFCO (The Association of American Feed Control Officials) (Giraffa, 2003).

2.2 *Enterococcus* en alimentos y su potencial biotecnológico

El género *Enterococcus* puede colonizar diversos nichos ecológicos y, por ello, está presente en vegetales y diversos alimentos fermentados, especialmente en los que son de origen animal como son salchichas y quesos (Foulquié-Moreno et al., 2006). En la industria de cárnicos, los enterococos generalmente no son deseables porque se les asocia con la descomposición del alimento y la formación de aminas biógenas (Franz et al., 2003). Sin embargo, su papel en la producción de alimentos fermentados artesanales tanto de origen lácteo como cárnicos y algunos vegetales, se basa en su contribución en el desarrollo de los sabores y aromas característicos del alimento, debido principalmente a sus actividades proteolíticas, lipolíticas y de oxidación de citrato (Foulquié-Moreno, 2006; Giraffa, 2002). Diversos estudios han demostrado que estas bacterias contribuyen al desarrollo de sabor, aroma, color y estructura de los quesos debido a la producción de diversos metabolitos como:

concentración de nitrógeno soluble, aminoácidos libres, ácidos grasos volátiles, ácidos grasos de cadena media y larga, etc. (Giraffa, 2003). Su presencia en los alimentos se atribuye a su excepcional aptitud para sobrevivir y persistir en ambientes desfavorables: resisten estrés oxidativo, desecación, son capaces de crecer en un amplio rango de pH, que va de 4.6 a 9.5, y de temperaturas, que abarca de 10 a 45°C e incluso sobreviven a una pasteurización clásica de 60 °C por 30 min.; también crecen en presencia de un alto contenido de NaCl y de sales biliares, además de algunos detergentes y antimicrobianos (Paulsen et al., 2003; Orgier y Serror, 2008). La presencia de éstos en quesos artesanales es muy común y han sido ampliamente reportados en los elaborados en el Mediterráneo, por ejemplo: Manchego, Armada, Cebreiro, Picante, Majoero, Feta, Teleme, Mozzarella, Monte Veronese, Caprino, Serra, Venaco y Comté (Giraffa, 2003).

2.2.1 Metabolismo de citrato por enterococos: El citrato está presente en muchas materias primas usadas para la elaboración de alimentos fermentados, tales como frutas, vegetales, leche, y en algunos casos es usado como aditivo para la producción de salchichas fermentadas (Foulquié-Moreno, 2006). La capacidad de metabolizar el citrato está invariablemente vinculada con la presencia de plásmidos endógenos que contienen los genes que codifican para un transportador, el cual es responsable de la captación del citrato (Hugenholtz, 1993). Dado que el citrato es un sustrato altamente oxidable, no hay equivalentes reductores, como el NADH, producidos durante su degradación, lo cual resulta en la formación de productos metabólicos finales diferentes del ácido láctico. Estos incluyen diacetilo, acetaldehído, acetoína y 2,3 butanodiol, que son los responsables de distintos aromas y de ahí su importancia en la calidad de los alimentos fermentados (Foulquié-Moreno, 2006; Martino et al., 2016).

2.2.2 Proteólisis. La degradación de caseína juega un papel importante en el desarrollo de la textura de los quesos. Además, algunos péptidos contribuyen a la formación de aromas importantes en el perfil sensorial del alimento, mientras que otro, con sabor amargo, no son deseables (van Kranenburg et al., 2002). La degradación secundaria de los aminoácidos tiene mucho impacto en el desarrollo de los aromas a través de la conversión de los aminoácidos libres en alcoholes, ácidos y aldehídos. Sin embargo, se sabe que esta actividad es baja para el género

Enterococcus, a excepción de algunas cepas de *E. faecalis* (Sarantinopoulus et al., 2001; Foulquié-Moreno, 2006).

2.2.3 Lipólisis. Los lípidos tienen un mayor efecto en el sabor y textura de los quesos, contribuyendo principalmente por tres formas diferentes: 1) son una fuente de ácidos grasos de cadena corta que después son convertidos a compuestos aromáticos, especialmente metilcetonas, lactonas y ésteres, 2) oxidan ácidos grasos y producen aldehídos insaturados, los cuales están fuertemente relacionados en el enranciamiento oxidativo y 3) finalmente, actúan como disolventes para compuestos aromáticos, producidos no sólo a partir de lípidos, sino también desde proteínas y lactosa (Sarantinopoulus et al., 2001). A pesar de que el género *Enterococcus* no es conocido por tener gran actividad lipolítica, ya se han reportado y caracterizado algunas lipasas y esterases principalmente en *E. faecalis*, *E. faecium* y *E. durans*. También se ha establecido que *E. faecalis* es el que ha presentado mayor actividad esterolítica (Foulquié-Moreno, 2006; Serio et al., 2010).

2.2.4 Compuestos con actividad antibacteriana. Especies de *Enterococcus* tienen la capacidad de producir bacteriocinas, llamadas enterocinas, las cuales son péptidos, sintetizados ribosomalmente, de bajo peso molecular con actividad antibacteriana contra bacterias Gram positivas relacionadas filogenéticamente incluyendo bacterias asociadas a descomposición de alimentos y patógenas como lo son *Listeria monocytogenes*, *Staphylococcus aureus*, *Bacillus cereus* y *Clostridium botulinum* (Franz et al., 2007; Martín-Platero, 2009). Además, se han reportado otras enzimas llamadas peptidoglucano hidrolasas (PGHs), las cuales se encargan de hidrolizar el peptidoglucano, componente principal de la pared celular de las bacterias, y con ello efectúan una actividad antibacteriana (García-Cano et al., 2014).

2.2.4.1 Bacteriocinas: Las cepas de enterococos productoras de bacteriocinas son de gran interés biotecnológico, ya que se han caracterizado por ser activas contra microorganismos importantes en la inocuidad alimentaria, como son: *Listeria monocytogenes*, *Bacillus cereus*, *Staphylococcus* spp. y *Clostridium* spp., por lo que son propuestas para su uso como bioconservadores de alimentos (Foulquié Moreno et al. 2003a y 2006). Específicamente, las enterocinas son clasificadas en cuatro grupos: la clase I son enterocinas con un residuo lantibiótico, la clase II, las

cuales son enterocinas pequeñas y no lantibióticas; la clase III incluye a las enterocinas cíclicas y la clase IV son enterocinas grandes, es decir de un peso molecular de más de 10 kDa. Además, la clase II se divide en tres subclases: II.1 son enterocinas que pertenecen a la familia de las pediocinas (con un motivo YGNGV), II.2 son enterocinas sintetizadas sin péptido líder; y finalmente las II.3 son lineales y no pertenecen a la familia de las pediocinas (Franz et al., 2007).

2.2.4.2 Peptidoglucano hidrolasas (PGHs): La pared celular de las bacterias está compuesta de peptidoglucano y durante todo su crecimiento, las células necesitan romper de manera controlada los enlaces de dicha estructura. Las enzimas responsables de este proceso son conocidas como peptidoglucano hidrolasas o PGHs (Eckert et al., 2006). Por lo tanto, las PGHs están involucradas en un gran número de funciones que requieren la modificación de la red rígida del peptidoglucano durante el crecimiento y la división celular (Lepeuple et al., 1998). Las PGHs se han clasificado de acuerdo con el tipo de enlace que son capaces de romper en el peptidoglucano. Las N-acetilmuramoyl-L-alaninamidases rompen el enlace amida entre MurNAc y la L-alanina del péptido. Las peptidasas son capaces de hidrolizar el último aminoácido del extremo carboxilo de los péptidos (carboxipeptidasas), o de romper completamente los puentes formados por los péptidos (endopeptidasas). Las N-acetilglucosaminidasas y las N-acetilmuramidases hidrolizan los enlaces β -1,4 de la cadena de glicanos. Las primeras lo hacen dejando un extremo GlcNAc reductor, mientras que las segundas hidrolizan el enlace entre MurNAc y GlcNAc, pudiendo dejar libre un extremo MurNAc reductor, llamadas lisozimas, o formando un anillo 1,6-anhidro en MurNAc, conocidas como transglicosilasas líticas (Vollmer et al., 2008). El uso de estas enzimas líticas en el desarrollo de la inocuidad de alimentos ya ha sido ampliamente reportado, por ejemplo: la lisostafina ha demostrado tener un fuerte efecto contra *S. aureus* en diversos alimentos como leche, mayonesa y cerdo; la lisozima suprime el crecimiento de *Listeria monocytogenes* y *Salmonella anatum* en salmón ahumado y un coctel de enzimas líticas producidas por *Streptomyces* sp. es muy efecto para el control de bacterias ácido lácticas y acéticas en la producción de vinos (Callewaert et al. 2011).

2.3 Potencial de patogenicidad del género *Enterococcus*

A pesar de la importancia de los enterococos en alimentos fermentados, este género ha tenido recientemente mucha relevancia en la microbiología clínica. *Enterococcus* ha sido tradicionalmente catalogado como patógeno oportunista o de bajo nivel (Franz et al., 2003). Sin embargo, no hay una hipótesis clara sobre el impacto en la inocuidad debido a su presencia en productos alimentarios. Su reciente y acentuada ambigüedad en la relación con los seres humanos, se debe a su hábitat entérico, su participación en la cadena de alimentos, la resistencia a antibióticos y su implicación en enfermedades adquiridas en los hospitales (Giraffa, 2003; Orgier y Serror, 2008).

Así mismo el papel de los enterococos como probiótico es controversial debido al incremento de infecciones nosocomiales (infecciones de tracto urinario, sepsis, endocarditis, etc.) atribuidas a los mismos y a su múltiple resistencia a antibióticos (Franz et al, 2003). El desarrollo de virulencia se da por las siguientes características: la adherencia al tejido del hospedero, invasión y formación de abscesos, capacidad de modulación de la respuesta inflamatoria del hospedero y secreción de sustancias tóxicas (Semedo et al, 2003). Sin embargo, la posible patogenicidad de las diferentes especies y cepas dentro de este género sigue en estudio.

Hasta este momento, casi una docena de factores de virulencia han sido reportados a través del estudio en varios modelos animales. Dichos factores están involucrados tanto en interacciones entre las células hospederas y proteínas extracelulares (*asa1*, *ace*, *esp*, *efaA*), en resistencia a macrófagos (*asa*, *hypR*), en daño a tejido y células (*cyl*, *gelE*, *sprE*), así como en evasión del sistema inmune (cápsula: operón *cps*) (Ogier J. y Serror P., 2008). Por lo tanto, los enterococos cuentan con los factores necesarios para considerarse patógenos intrahospitalarios, como son la persistencia al ambiente, colonización del tracto gastrointestinal, mecanismos de dispersión (translocación a través del epitelio intestinal), capacidad de crecimiento en condiciones adversas, producción de toxinas y capacidad de dañar un tejido (Fig 1).

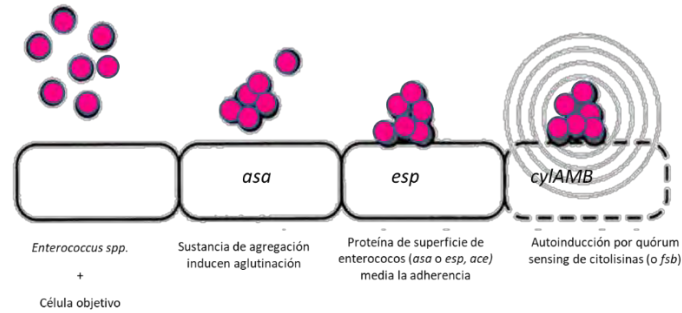


Fig 1. Modelo de sinergia para causar una infección por el género *Enterococcus*. Los enterococos que tengan la capacidad de adherirse entre ellos (gen *asa*), llegar y colonizar las células del hospedero (gen *esp*) y finalmente causar un daño al hospedero debido a la producción de enzimas que hidrolizan las células eucarióticas (operón citolisinas, proteasas, etc) pueden aumentar hasta 8 veces su patogenicidad (Gilmore, 2002).

Enterococcus presenta dos tipos de resistencia a los antibióticos: la resistencia intrínseca y la adquirida o extrínseca (Orgier y Serror, 2008). La mayoría de las cepas de *E. faecium* y *E. faecalis*, son intrínsecamente resistente a bajos niveles de diversas clases de antibióticos, incluidos algunos agentes activos contra la pared celular como los betalactámicos, sulfonamidas, aminoglucósidos, clindamicina y lincomicina y dentro de la resistencia adquirida se pueden encontrar el cloranfenicol, eritromicina, altos niveles de β -lactámicos, fluoroquinolonas, tetraciclinas y glicopéptidos, tales como la vancomicina (Willems, 2010). Esta característica agrava el tratamiento terapéutico cuando se encuentran causando una infección nosocomial ya que el uso de antibióticos se ve limitado. Por lo anterior, es que en febrero de 2017 la Organización Mundial de la Salud urgió a que “se establezcan políticas que incentiven la investigación científica básica y la I+D avanzada tanto a través de organismos financiados con fondos públicos como del sector privado que inviertan en el descubrimiento de nuevos antibióticos.” (Lawe-Davies y Bennett, 2017/ <http://www.who.int/mediacentre/news/releases/2017/bacteria-antibiotics-needed/es/>). Sin embargo, la presencia de este género en otros ambientes, diferentes al nosocomial, continuamente se ha reportado con una menor incidencia de multirresistencia a antibióticos, limitándose a su resistencia intrínseca (Franz et al., 2001).

A finales de la década de 1970, Sanger y colaboradores desarrollaron un método para la secuenciación de ADN mediante la técnica llamada “terminación de cadena”. Esta técnica transformó la biología y la forma de ser estudiada, ya que se convirtió en una herramienta poderosa para descifrar la secuencia de genes y años más tarde incluso genomas. La

tecnología de Sanger con los años se fue mejorando hasta hacerla paralelizable y acoplada a electroforesis capilar con detectores de fluorescencia que después son transformados en cromatogramas. Hoy en día es ampliamente utilizada para secuenciación de fragmentos genómicos, genes o para corroborar datos obtenidos de secuenciación de nueva generación (van Dijk et al., 2014).

El término “Nueva-generación” hace referencia a las diversas implementaciones de secuenciación basadas en “arreglos cíclicos”, por ejemplo: 454 sequencing (454 Genome Sequencers, Roche Applied Science), Solexa technology (Illumina, Genome Analyzer), plataforma SOLiD (Applied Biosystems), Polonator (Dover/Harvard), HeliScope SingleMolecule Sequencer technology (Helicos; Cambridge, MA, USA). El concepto de secuenciación basada en “arreglo-cíclico” puede ser resumido como la secuenciación de un arreglo denso de fragmentos de ADN mediante ciclos iterativos de manipulación enzimática y recolección de datos basada en imágenes (Shendure y Ji, 2008). Hoy en día se encuentran tomando fuerza tecnologías que conocemos por “tercera generación” dentro de las cuales se encuentran: PacBio (SMRT, Pacific Biosystem) y Nanopore (Oxford Nanopore Technologies) las cuales llevan a cabo secuenciación en tiempo real (Fig 2).

2.4 Era de la genómica en el estudio de *Enterococcus*

A finales de la década de 1970, Sanger y colaboradores desarrollaron un método para la secuenciación de ADN mediante la técnica llamada “terminación de cadena”. Esta técnica transformó la biología y la forma de ser estudiada, ya que se convirtió en una herramienta poderosa para descifrar la secuencia de genes y años más tarde incluso genomas. La tecnología de Sanger con los años se fue mejorando hasta hacerla paralelizable y acoplada a electroforesis capilar con detectores de fluorescencia que después son transformados en cromatogramas. Hoy en día es ampliamente utilizada para secuenciación de fragmentos genómicos, genes o para corroborar datos obtenidos de secuenciación de nueva generación (van Dijk et al., 2014).

El término “Nueva-generación” hace referencia a las diversas implementaciones de secuenciación basadas en “arreglos cíclicos”, por ejemplo: 454 sequencing (454 Genome Sequencers, Roche Applied Science), Solexa technology (Illumina, Genome Analyzer), plataforma SOLiD (Applied Biosystems), Polonator (Dover/Harvard), HeliScope

SingleMolecule Sequencer technology (Helicos; Cambridge, MA, USA). El concepto de secuenciación basada en “arreglo-cíclico” puede ser resumido como la secuenciación de un arreglo denso de fragmentos de ADN mediante ciclos iterativos de manipulación enzimática y recolección de datos basada en imágenes (Shendure y Ji, 2008). Hoy en día se encuentran tomando fuerza tecnologías que conocemos por “tercera generación” dentro de las cuales se encuentran: PacBio (SMRT, Pacific Biosystem) y Nanopore (Oxford Nanopore Technologies) las cuales llevan a cabo secuenciación en tiempo real (Fig 2).

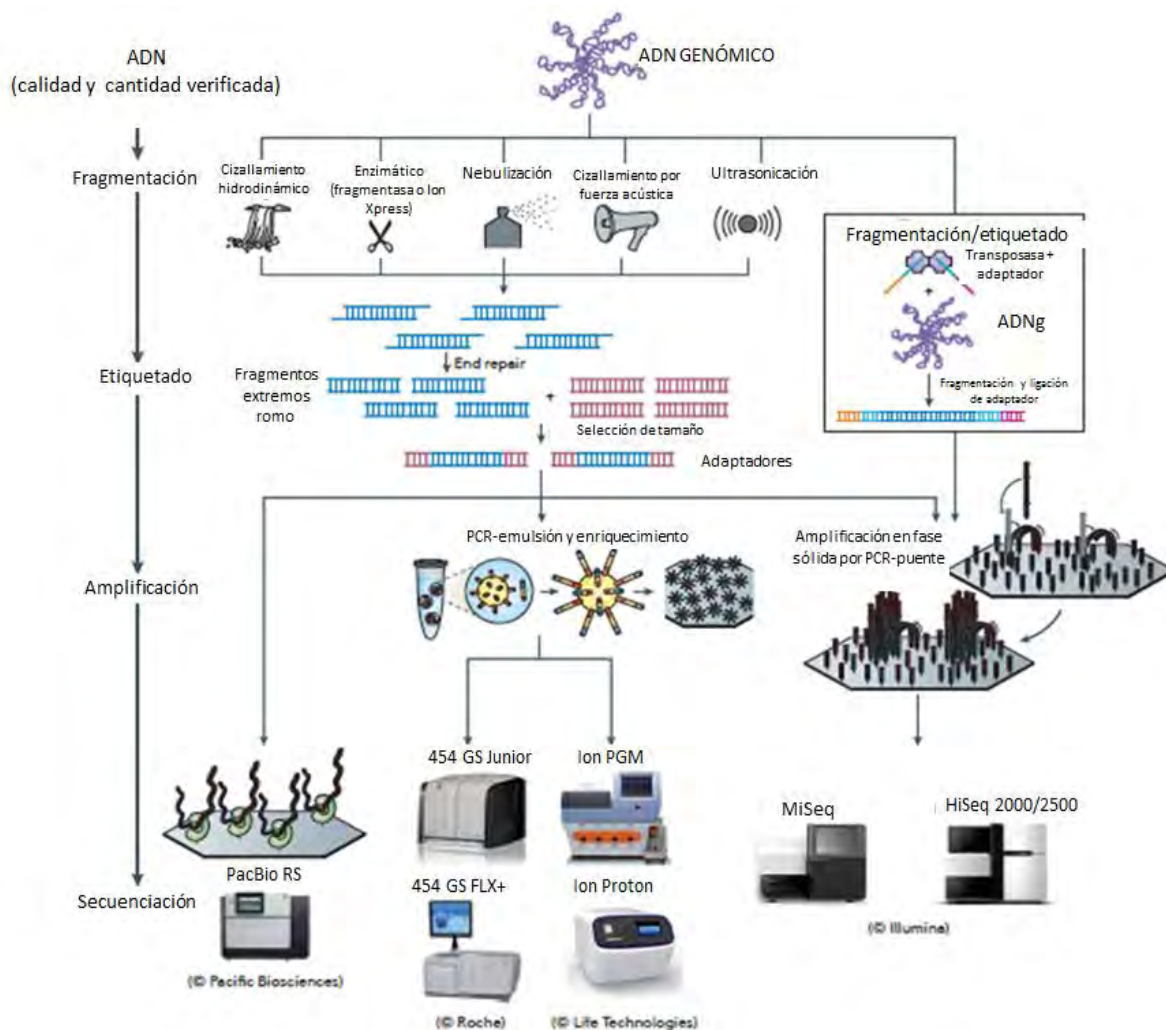


Fig 2. Tecnologías de nueva generación. El esquema muestra las principales plataformas de secuenciación de alto rendimiento disponibles y los procedimientos asociados de preparación de muestras y amplificación de plantillas (Modificada de Loman et al., 2012).

El estudio de genomas de *Enterococcus* ha crecido considerablemente en los últimos años con particular atención en el análisis de genómica comparativa entre aislados clínicos. En ese mismo sentido, varios estudios se han enfocado en la comparación de aislados clínicos y no clínicos en cepas de *E. faecalis* y *E. faecium* con el fin de evaluar aspectos de seguridad (Kim y Marco, 2014). Sin embargo, las cepas aisladas de alimentos han sido escasamente investigadas a este nivel de detalle, lo cual es sorprendente ya que este microorganismo tiene una amplia importancia, abundancia y distribución en alimentos fermentados de manera no intencionada (Bonacina et al., 2016).

Varios autores señalan que la información genética del género *Enterococcus* ha sufrido cambios dependiendo del nicho ecológico en el que se encuentre. Se ha observado que la cantidad de marcadores genéticos de patogenicidad es mayor en los aislados clínicos, seguido de aislados de alimentos fermentados tradicionales y por último los que se pueden encontrar en cepa utilizadas como cultivo iniciador en alimentos y probióticos (Eaton y Gason, 2001; Aubriouel et al, 2008). Se ha demostrado, mediante técnicas de tipificación por PCR de longitud de heterogénea (LH-PCR), que los aislados clínicos se pueden concentrar en un mismo grupo, el cual es diferente del que asocia a los aislados de matrices alimentarias, los cuales en su mayoría forman un grupo separado que es compartido con cepas aisladas de suelo y agua (Aubriouel et al, 2008).

van Schaik y Willems (2010), describen la evolución de los enterococos con base en su genoma, y se plantea, a través de técnicas como tipificación de secuencias multilocus (MLST), la existencia de subpoblaciones específicas de *E. faecalis* y *E. faecium* asociadas con infecciones nosocomiales, formando lo que se ha denominado complejo(s) clonal(es) de alto riesgo (CC). Además, mediante técnicas de genómica comparativa por hibridación (HCG) se ha buscado identificar genes que estén presentes únicamente en aislados clínicos y ausentes en cepas de alimentos (productos lácteos), con el objetivo de estudiar la diversidad genética de este género y específicamente para determinar la variación en sus islas de patogenicidad (Palmer, 2014; Kim and Marco, 2014). Gracias a estos y otros estudios, se ha propuesto que la variación en el contenido genómico de diferentes cepas de *E. faecalis*, se basa en la presencia o ausencia de elementos genéticos móviles (EGM), tales como fagos y elementos transponibles conjugativos y variaciones en sus islas de patogenicidad (Kopit et al., 2014) La presencia de genes relacionados con patogenicidad es común en aislados no clínicos, lo que sugiere que dichos genes podrían tener papeles desconocidos hasta ahora en la colonización

de individuos sanos (por ejemplo: sustancia de agregación y proteína de superficie) y que la adquisición de genes de virulencia por sí solo es insuficiente para conferir un fenotipo infeccioso de *E. faecalis*. Los estudios antes mencionados sugieren que las infecciones en humanos están asociadas con subpoblaciones específicas, que probablemente sea incorrecto considerar estas subpoblaciones estrictamente grupos clonales, ya que pueden presentar grandes diferencias en el contenido genético y que la variabilidad en la presencia de los EGM entre los aislados de una especie provenientes de diferentes nichos podría ser el factor más importante en la determinación de sus características fenotípicas (van Schaik y Willems, 2010; Palmer et al., 2014).

Es importante recordar que varios factores que actúan para conferir virulencia son los mismos que podrían tener una acción simbiótica con el hospedero. Tal es el caso de los factores para persistencia al ambiente, colonización del tracto gastrointestinal y mecanismos de dispersión, los cuales se han propuesto en la cepa probiótica *E. faecalis* Symbioflor1 (Domman et al., 2007). Por lo tanto, lo que podría determinar cuál es la función de algunos de dichos factores es el entorno ambiental en el que se encuentra la bacteria. Así ocurre con la sustancia de agregación, la proteína de adhesión de colágeno, la resistencia contra aniones de oxígeno y la formación de una cápsula (Domman et al., 2007). Una revisión de los elementos genéticos y regulatorios pertenecientes a una asociación patogénica o de simbiosis revela que no hay principios generales que justifiquen una distinción entre ambas. Por el contrario, las estrategias subyacentes de las interacciones bacteria-hospedero son similares en patógenos y simbiosis, aunque con propiedades y funciones modificadas para adaptarse a sus necesidades (Hentschel et al., 2000).

3. ANTECEDENTES

3.1 Microbiología del queso Cotija

El queso Cotija es un alimento artesanal, madurado que se elabora a partir de leche bronca y durante su elaboración no se incluye ningún proceso térmico que disminuya la carga microbiana presente. Sin embargo, pasados al menos tres meses de maduración, en el lugar donde se elabora, la inocuidad del producto es apropiada (sin riesgo microbiológico) para el consumidor (NMX-F-735-COFOCALEC-2011). Este queso se elabora en la sierra conocida como “Jalmich”, ya que se ubica entre los estados de Jalisco y Michoacán y cada año es vendido en la feria de Cotija, de donde toma su nombre. Su proceso de elaboración se ha conservado por más de 400 años (Cervantes et al., 2008). Este alimento es de gran formato, con piezas de entre 20-22 kg, de pasta desmoronable, sabor y olor pronunciado y color amarillo-crema (Fig 3A). En los últimos años se han llevado a cabo diversos esfuerzos para ayudar a los productores a obtener la denominación de origen, sin embargo, hasta ahora sólo se cuenta con la marca colectiva: “Queso Cotija Región de Origen” (Fig 3B) (Escobar-Zepeda et al., 2016).



Fig 3. Queso Cotija artesanal madurado y zona de producción en la sierra de Jalmich. Tomando de Álvarez et al., 2005.

A partir del análisis taxonómico del metagenoma del queso Cotija se observó que la diversidad bacteriana está formada por 574 géneros, donde tres especies representan más del 80% de la abundancia total: *Lactobacillus plantarum*, *Leuconostoc mesenteroides* y *Weisella paramesenteroides*. Mientras que cuatro géneros son parte de una población subdominante con una abundancia de entre 1-10%: *Aerococcus*, *Enterococcus*, *Lactococcus* y *Staphylococcus* (Fig 4, cuadro verde). Dentro del género *Enterococcus* se identificaron las especies *E. faecium*, *E. faecalis*, *E. casseliflavus*, *E. gallinarum* y *E. avium* (Escobar-Zepeda et al., 2016).

A pesar de la subdominancia de enterococos en el queso Cotija (QC), frecuentemente y con mucha facilidad, se han aislado cepas de este género pertenecientes principalmente a las especies de *E. faecalis* y *E. faecium*. Algunas de ellas demostraron tener actividad lítica, por técnicas de zimografía, contra *Micrococcus lysodeikticus*, *S. aureus* y *Escherichia coli* con proteínas de pesos moleculares entre los 60 y 86 kDa (García Cano et al., 2014). Por otra parte, el análisis del potencial de patogenicidad asociado a este género indicó que las cepas presentaban genes relacionados con su capacidad para adherirse y colonizar el tracto gastrointestinal y carecían de un gen involucrado en la síntesis de una proteína capaz de hidrolizar células eucarióticas (operón de citolisina) (Olvera, 2013). En otro estudio se determinó que la cepa de *E. faecalis* F tiene un amplio espectro de actividad antibacteriana contra microorganismos tanto Gram-positivos como Gram-negativos de importancia en seguridad alimentaria, atribuida a la producción de peptidoglucano hidrolasas (PGHs) (García-Cano et al., 2014). A su vez, la cepa de *E. faecium* D presenta actividad bacteriocinogénica debido a la producción de la enterocina A, la cual es activa contra *Listeria* spp., así como actividad lítica contra *M. lysodeikticus* presumiblemente por la presencia de una PGH de 60 kDa (Olvera, 2013).

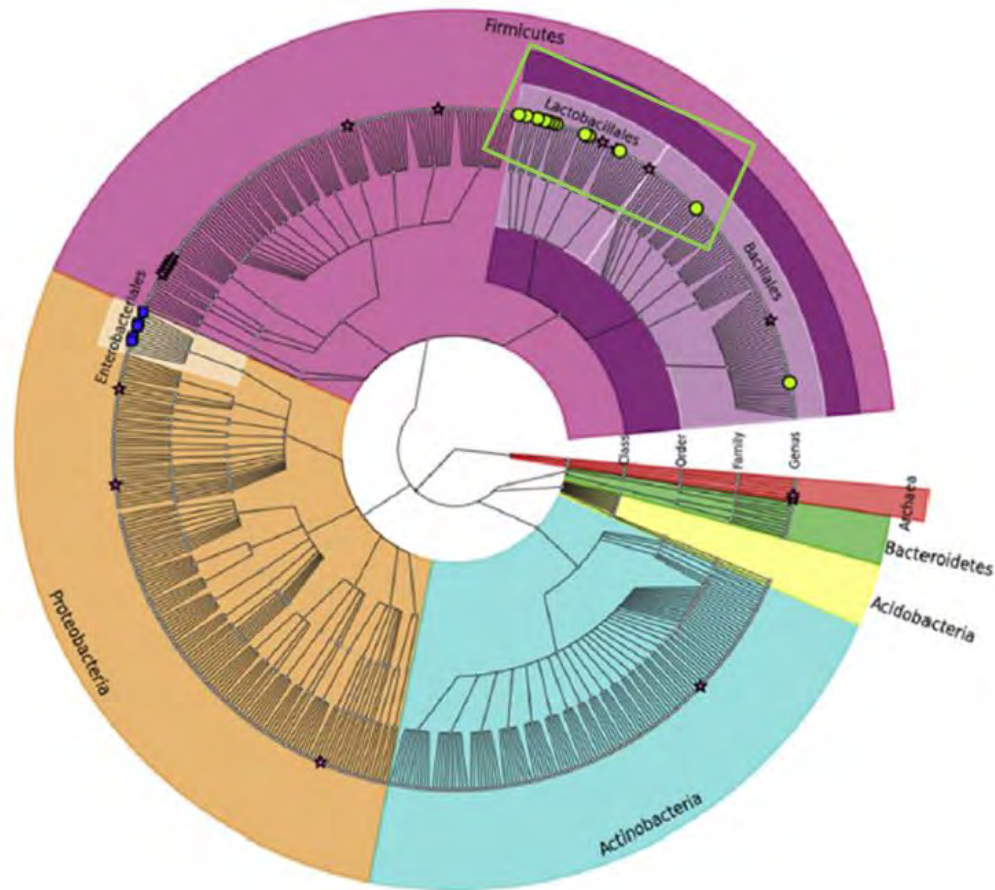


Fig 4. Anotación taxonómica de las principales fila bacterianas identificadas a través del marcador filogénico 16S ADNr. Los círculos verdes, dentro de recuadro verde, marcan los géneros dominantes y subdominantes (Modificado de Escobar-Zepeda et al., 2016).

4. JUSTIFICACIÓN

Los *Enterococcus* se caracterizan por ser ubicuos, al tener la capacidad de sobrevivir a diferentes ambientes, por lo cual se han aislado de alimentos, suelo, agua, humanos y animales y, desafortunadamente, de infecciones nosocomiales. Por lo anterior, existe una alta preocupación por la presencia de este género en alimentos. A pesar, de la alta interconexión que existe entre los diferentes nichos en donde habita el género *Enterococcus* y aunado al bajo control de residuos sanitarios en los hospitales, las cepas de *E. faecalis* y *E. faecium* aisladas de un nicho ecológico diferente al nosocomial se han caracterizado por tener una baja incidencia de genes relacionados con factores de virulencia, deleciones en sistemas de

regulación de éstos, baja o nula resistencia a antibióticos, así como la baja presencia de elementos móviles y sitios de inserción. Sin embargo, la controversia sobre el efecto negativo de su presencia en alimentos continúa siendo de gran relevancia. Por lo tanto, es importante generar mayor evidencia genómica de cepas de *E. faecalis* aisladas de diversos ambientes, principalmente de origen alimentario, que ayude a contar con información que permita diferenciar cepas inocuas de aquéllas presentes en hospitales.

Además, aunque se considera de gran importancia al género *Enterococcus* en el proceso de maduración de diversos alimentos lácteos poco se ha reportado acerca de las actividades que podría estar desempeñando. Por lo que resulta de interés conocer el potencial proteolítico y lipolítico que presentan estas especies.

5. HIPÓTESIS

Las cepas de *E. faecalis* y *E. faecium* aisladas del queso Cotija presentarán diferencias genómicas en comparación con los aislados de otros nichos ecológicos, principalmente el nosocomial, las que estarán relacionadas a su adaptación a los alimentos.

6. OBJETIVO GENERAL

Caracterización genómica de cepas pertenecientes a las dos principales especies del género *Enterococcus* con el fin de establecer evidencia genómica que diferencie cepas de *E. faecalis* y *E. faecium* aisladas de alimentos de un aislado de un nicho ecológico diferente, así como determinar su funcionalidad en los alimentos.

5.1 OBJETIVOS PARTICULARES

- Definir la estructura poblacional de cepas de *Enterococcus* aisladas de queso Cotija y su relación con complejos clonales de alto riesgo a través del uso de MLST.
- A través de la genómica comparativa evaluar la relación que existe entre las cepas de *Enterococcus* aisladas del queso Cotija con las de origen nosocomial.
- Establecer secuencias genómicas relacionadas a la adaptación de los enterococos a los alimentos lácteos fermentados
- Dilucidar el potencial metabólico de *E. faecalis* F y *E. faecium* D relacionado al proceso de maduración y desarrollo de aromas y sabores en el queso Cotija, así como la producción de compuestos con actividad antibacteriana.

7. METODOLOGÍA GENERAL

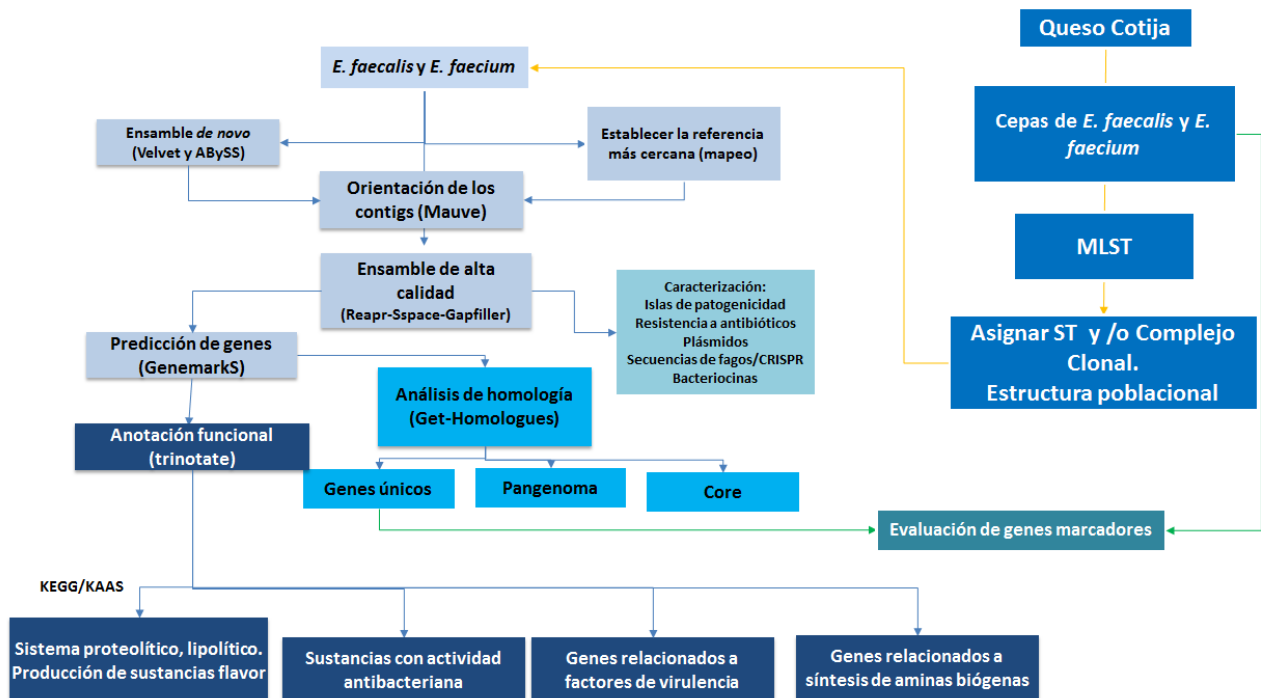


Fig 5. Metodología general para el desarrollo del proyecto

7.1 Determinación de estructura poblacional (MLST)

La estructura poblacional se determinó a través de la metodología de MLST para cada especie de *Enterococcus*. Para este análisis se utilizaron 3 cepas de *E. faecalis* aisladas de queso Cotija en 2008 (Bravo, 2008) y 4 de una colecta del 2012 (Anexo I y Guzmán, 2015). La estructura poblacional de *E. faecium* incluyó 7 cepas aisladas en 2015, el análisis fue realizado por Guzmán en el 2015 y en el apartado de discusión se retomarán sus resultados. De cada cepa se amplificaron fragmentos de los 7 genes de mantenimiento descritos en el protocolo (www.mlst.net), para lo cual se utilizó la Phusion High-Fidelity PCR Master Mix, la cual tiene una ADN polimerasa con muy baja tasa de error (Life Technologies Inc. California, EUA). Los genes amplificados fueron: *gdh* (enzima glucosa-6-fosfato deshidrogenasa), *gyd* (gliceraldehido-3-fosfato deshidrogenasa), *pstS* (transportador del casete de unión de ATP-fosfato), *gki* (glucoquinasa), *aroE* (shikimato 5- deshidrogenasa), *xpt* (xantina fosforribosiltransferasa) y *yiqL* (acetil-CoA acetiltransferasa) (Ruíz-Garbajosa, 2006). Los amplicones se purificaron con el uso del kit DNA Clean & Concentrator™ (Zymo Research, Carlifornia, EUA) y se secuenciaron por método de Sanger (IBT-UNAM, Morelos, México y Macrogen, Seúl, Corea del Sur). Finalmente, los resultados se analizaron a través de la base

de datos (MLST) con el fin de obtener el número de alelo correspondiente (<https://pubmlst.org/>), siguiendo el protocolo de análisis reportado por Guzmán (2015). Con el análisis de las secuencias se obtuvieron los perfiles alélicos (combinación) y con ello las secuencias tipo (ST) (Fig 6). Por último, se realizó el análisis de agrupamiento de las secuencias tipo obtenidas, utilizando la base de datos de MLST (<http://pubmlst.org/>) y el algoritmo e-BURST versión.3 disponible en línea (<http://eburst.mlst.net/default.asp>), en el cual se estableció que un complejo clonal estaba definido por aquellas ST que comparte al menos 5 de los 7 genes de la combinación alélica.

Con el fin de describir los complejos clonales con mayor detalle (analizando los cambios puntales a nivel nucleótido), se elaboraron árboles de agrupación usando la concatenación de las secuencias nucleotídicas de los 7 genes de cada una de las cepas. Para llevar a cabo ese análisis se tomaron secuencias alélicas reportadas para aislados clínicos pertenecientes a los complejos clonales de alto riesgo y cepas no nosocomiales, principalmente aisladas de alimentos (el listado de cepas se encuentra en el Anexo II). Las secuencias se alinearon y se pusieron en fase a través del uso de BioEdit v7.0.5 (Hall, 1999), después se alinearon usando clustalW, disponible en Mega 7. Posteriormente, se calculó el mejor método de sustitución y finalmente se construyeron los árboles con el método de Neighbor Joining usando Mega 7.0.26 con un Bootstrap de 1000 (Tamura et al., 2013).

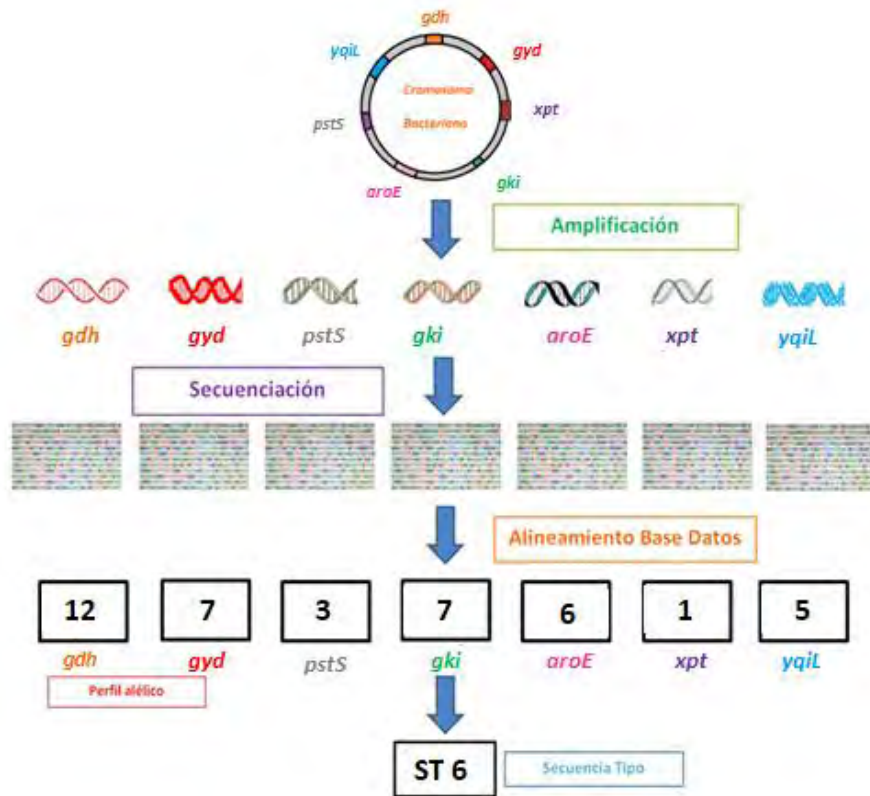


Fig 6. Esquema de metodología de MLST (Modificado de Guzmán, 2015)

7.2 Análisis Genómicos

7.2.1 Extracción de ADN

Se continuó el análisis genómico con las cepas *E. faecium* D y *E. faecalis* F. Ambas cepas se aislaron del queso Cotija en 2008 y se encontraban almacenadas bajo condiciones criogénicas, por lo que se reactivaron en 10 mL de caldo MRS (de Man, Rogosa and Sharp, Oxoid, Basingstoke, UK). Posteriormente, 1 mL de ese cultivo se inoculó en 100 mL caldo MRS y se incubaron 12 h a 37 °C y 250 rpm. La extracción de ADN se realizó con el kit FAST-ID siguiendo las indicaciones del proveedor (Genetic ID NA, Inc., Fairfield IA, EUA). Del mismo cultivo se tomaron 10 mL para realizar la extracción de plásmidos usando el GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific, Massachusetts, EUA). La concentración de ADN obtenida se cuantificó por medio del equipo take3-Epoch (BioTeK, Vermont, EUA). La presencia de plásmidos se visualizó por electroforesis en gel de agarosa al 1% (m/v) teñido con bromuro de etidio al 0.1% (m/v).

7.2.2 Secuenciación de genomas y análisis genómico

Se preparó una biblioteca de un tamaño de inserto de 500 pb usando el kit TruSeq v2 (Illumina, Inc. San Diego CA, EUA). La secuenciación se llevó a cabo en una plataforma de HiSeq 2000 (Illumina, Inc. San Diego CA, EUA) en ciclos de extremos pareados de 100 pb en el Laboratorio Nacional de Genómica y Biodiversidad (LANGEBIO, Irapuato, Guanajuato, México). La calidad de las lecturas obtenidas para cada cepa fue verificada empleando el software FastQC y teniendo como base un score Phred ≥ 20 para cada nucleótido (Andrews, 2016). Previo a realizar el ensamble *de novo*, se normalizó la cobertura de las lecturas eliminando aquellas que no fueran útiles para el ensamble con el software Khmer toolkit (Brown et al., 2012). El ensamble de cada genoma se realizó usando los programas ABySS v1.5.2 (Simpson et al., 2009) y Velvet v1.2.09 (Zerbino y Birney, 2008) y para encontrar el mejor ensamble se construyó un script en Perl que calculó las estadísticas N50/N90, para *E. faecalis* el mejor fue con un k-mero de 61 (ABySS) y para *E. faecium* con k-mero 89 (Velvet). La mejora de los ensamblados comenzó con un mapeo de las lecturas sobre los genomas de las referencias completamente cerradas y anotadas, y que han sido caracterizados ampliamente en otros trabajos, con el mapeador Smalt (`> smalt map -r 0 -y 0.8 -f samsoft -o`) (Ponstingl and Ning, 2010, <ftp://ftp.sanger.ac.uk/pub/resources/software/smalt/>; ftp://ftp.ncbi.nlm.nih.gov/genomes/archive/old_refseq/Bacteria/). Una vez seleccionada la referencia más cercana (mayor número de secuencias mapeadas sobre la referencia) se realizó un reordenamiento de los genomas usando Mauve 2.3.1 (Darling et al., 2010). La evaluación de la calidad del nuevo ensamble se llevó a cabo re-mapeando las lecturas sobre el ensamble utilizando REAPR v3 con el fin de colapsar zonas de baja cobertura (Hunt et al., 2013). Finalmente, para la corrección de errores de ensamblado y scaffolding, y relleno de gaps se utilizaron los programas SSPACE v3 y Gapfiller v1.10, respectivamente (Boetzer et al., 2011; Nadalin et al., 2013).

A partir de los ensamblados mejorados se hizo la predicción de genes usando GeneMarkS.hmm Prokaryotic 3.25 con los parámetros establecidos para bacteria (Besemer et al., 2001). Una anotación preliminar fue hecha a través del servidor online Rapid Annotation using Subsystem Technology (RAST-SEED) (Aziz et al., 2008; Overbeek et al., 2014). Al mismo tiempo se realizaron anotaciones funcionales por búsquedas específicas: islas de patogenicidad (PAIDB v2.0) (Yoon et al., 2014), resistencia a antibióticos (ResFinder v2.1) (Zankari et al., 2012), predicción de regiones CRISPR-cas (CRISPRfinder) (Grissa et al., 2007) plásmidos

(PlasmidFinder v1.3) (Carattoli et al. 2014) y secuencias correspondientes a profagos y elementos móviles (PHAge Search Tool-PHAST) (Zhou et al., 2011).

7.2.3 Análisis de ortología

La genómica comparativa de cada especie se realizó utilizando el software Get-Homologues versión 20140627, el cual clusteriza los genes de familias homólogas usando los algoritmos bidireccional best-hit, COGtriangles y OrthoMCL (Contreras-Moreira y Vinuesa, 2013). Los grupos de proteínas obtenidos son analizados con scripts en Perl. Finalmente, se identifican los genes que forman el pangenoma (todos los genes de la especie), el core del genoma (genes presentes en todos los genomas) y genes exclusivos (genes presentes en un solo genoma) y estos son resumidos en una matriz de ausencia/presencia.

La representación visual de los alineamientos de genomas se realizó con el software BRIG (BLAST Ring Image Generator) v 0.95 (Alikhan et al., 2011). El genoma de mayor tamaño se seleccionó como la referencia para hacer el alineamiento (círculo concéntrico).

7.2.4 Anotación funcional

Para obtener la anotación funcional de cada genoma, se partió de la predicción de genes traducidos a aminoácidos y fue realizada con una adaptación del pipeline de Trinote de la siguiente manera: los genes en proteínas se compararon contra las bases de datos UniProtKB/Swiss-Prot a través del algoritmo de Blast v2.2.28+ (Camacho et al., 2009). Además, se hizo una anotación de dominios funcionales a través de hmmscan HMMER v3.1b1 (Eddy, 2011) y los modelos obtenidos se compararon contra la base de datos de Pfam-A. Las predicciones de péptido señales y dominios transmembranales se obtuvieron usando SignalP y tmHMM, respectivamente (<http://trinotate.github.io/>). A través de la base de datos Bagel3 (Bagel automated bacteriocin mining) se identificaron genes codificantes para bacteriocinas (van Heel et al., 2013) y su búsqueda se complementó con los resultados obtenidos de la anotación funcional de Trinotate. La búsqueda de PGHs se llevó a cabo a través de la búsqueda de dominios funcionales de Pfam.

7.2.5 Detección de genes relacionados con factores de virulencia y con síntesis de aminas biógenas

La identificación de genes que codifican para factores de virulencia y enzimas para la producción de aminas biógenas se realizó mediante una búsqueda exhaustiva en los resultados de la anotación funcional con respecto a lo que se ha reportado en la bibliografía (Orgier y Serror, 2008; Jiménez et al., 2013).

Para el análisis de identificación y cuantificación de aminas biógenas (putrescina, tiramina, histamina y cadaverina) en quesos Cotija se partió de 4 quesos Cotija “Región de Origen” colecta 2008 y 4 quesos de una colecta del 2012. De cada queso (triturado) se tomaron 250 g y se analizaron por medio de Cromatografía de Líquidos de Alta Eficacia (HPLC por sus siglas en inglés) (Saarka Nutrición y Tecnología, S.A., Ciudad de México, México). El análisis fue realizado siguiendo el método oficial 977.13 de la AOAC, con las modificaciones necesarias para ser detectados por derivados fluorescentes en HPLC.

7.2.6 Identificación de genes asociados al desarrollo de aromas y sistemas proteolítico y lipolítico

El análisis del potencial metabólico de los genes codificantes (traducidos a aminoácidos) de cada cepa se realizó a través del servidor KEEG Automatic Annotation para la asignación de ortólogos y mapeo de rutas metabólicas (KAAS) (Moriya et al., 2007). Dicha anotación en conjunto con la obtenida por Trinotate se utilizaron para complementar la identificación y predicción de moléculas relacionadas a la producción de aromas y sabores, así como la reconstrucción de un sistema proteolítico y lipolítico de ambas cepas a través del reconocimiento de proteinasas, peptidasas, transportadores ABC y enzimas con actividad lipolítica reportadas en otras BAL. La búsqueda se llevó a cabo teniendo como base lo reportado en la bibliografía (Park et al., 2015; Lui et al., 2010; Savijoki et al., 2006).

7.3 Pruebas de susceptibilidad a antibióticos

La susceptibilidad a antibióticos se evaluó por medio de pruebas de difusión en disco. Círculos de papel filtro de 1 mm de diámetro conteniendo: 16 µg/mL de tetraciclina, 20 µg/mL de kanamicina, 8 µg/mL de eritromicina, 1 mg/mL estreptomina (alto grado de resistencia), 32

µg/mL de vancomicina, 500 µg/mL de gentamicina (alto grado de resistencia) y 16 µg/mL de ampicilina (Sigma, St. Louis MO, EUA) (Franz et al., 2001). Primero se colocó una capa de agar BHI al 1 % y sobre ella una segunda capa de agar BHI al 0.8 %, en la cual previamente se habían inoculado 10⁷ UFC/mL de cada cepa de enterococos. Las cajas Petri se incubaron a 37°C por 24 h. Zonas de inhibición alrededor de los círculos con antibiótico se consideraban como un resultado positivo de susceptibilidad.

7.4 Elección y evaluación de posibles biomarcadores

Basados en la bibliografía se propusieron genes codificantes para bacteriocinas de la clase II.1 (Enterocina A, P y SE-K4) y los codificantes para enzimas de restricción tipo I o III (Nes et al., 2014; O’Sullivan et al., 2009; Franz et al., 2007) como posibles biomarcadores de adaptación a alimentos. Por lo anterior, se diseñaron los cebadores para cada gen y se sintetizaron en la empresa IDT (Coralville, EUA). Las condiciones de PCR se optimizaron para las cepas D y F (controles positivos) (Tabla 1). Para este ensayo se utilizó la enzima Taq polimerasa recombinante EP0402 (Life Technologies Inc. California, EUA). La presencia de los genes se evaluó en 14 cepas aisladas de queso Cotija (Sección 5.1), así como 5 cepas de *E. faecium* aisladas de queso artesanal Bola de Ocosingo (Urrieta, 2018). Los amplicones de cada PCR se analizaron mediante una electroforesis en gel de agarosa al 1% teñido con bromuro de etidio (0.1% m/v) y se visualizaron en un transiluminador UVTM (Hoefer). Cepas reportadas como aisladas de infecciones nosocomiales (V583, ATCC 29212 y NCIMB 700585) se emplearon como controles negativos.

Tabla 1. Cebadores y condiciones para realizar PCR de biomarcadores

Especie	Gen	Cebadores	Condiciones PCR
<i>E. faecalis</i>	Enterocina P	D: TAT TAG GAA CAA ATG TAG AAG C R: CTA TCT ACT TAT ACT AGC CAA A	95°C, 3 min, 35 (95°C 30 s, 45 °C 30 s, 72 °C 1 min) 72 °C, 10 min.
	Enzima de restricción I	D: CAA CGA ACT AAA CTA CAA TCA CA R: TGT AAA AAC ATA TTC ATA CGC GC	
<i>E. faecium</i>	Enterocina SE-K4	D: GAT TTT AGG AAT TGT TTT ATT GAG TG R: GCA CTT TCG TAA AAC CAT TAA CG	
	Enzima de restricción III	D: TTG CAG AAA CTG TTC CCT AAA AA	

	R: AAC TTT GCA ACA GAA CCA GTG
Enterocina A (Du Toit et al., 2000)	D:AAATATTATGGAAATGGAGTGTAT R:GCACTTCCCTGGAATTGCTC

7.5 Rearmado de genomas a partir del metagenoma y búsqueda de marcadores

A partir del ensamble de metagenoma (Escobar-Zepeda et al., 2016), se realizó la reconstrucción de los genomas de *E. faecalis* y *E. faecium* tomando como molde los genomas armados en este proyecto (Sección 5.3), utilizando el software Abacas (abacas.1.03.pl; Assefa et al., 2009). En estos genomas rearmados se hizo la búsqueda de los genes del sistema proteolítico, lipolítico y de enzimas relacionadas a la síntesis aromas y sabores, los cuales son propuestos como biomarcadores de adaptación a alimentos (en este trabajo), así como aquellos genes que codifican para bacteriocinas (genes únicos).

8. RESULTADOS Y DISCUSIÓN

8.1 ¿En qué se diferencian las cepas de enterococos aisladas de queso Cotija de cepas de otros ambientes?

Estructura poblacional de *E. faecalis* y *E. faecium* aisladas de queso Cotija

En el análisis del metagenoma del queso Cotija Región de Origen se concluyó que el género *Enterococcus* forma parte de la población subdominante y está representado por diferentes especies que incluyen a *E. faecium* y *E. faecalis* (Escobar-Zepeda et al., 2016). Sin embargo, durante los diferentes estudios poblacionales, dependientes de cultivo, del queso Cotija es frecuente su presencia, y resalta la facilidad con que se le aísla, por encima de los microorganismos reportados como dominantes en el alimento (*Lactobacillus*, *Leuconostoc* y *Weissella*). Por lo anterior y dada su controversia como patógenos oportunistas para el cuidado de la salud (intra-hospitalaria), se decidió seguir analizando estas especies con el fin de dilucidar su relación con complejos clonales de alto riesgo y así tener un panorama del potencial de riesgo por la presencia de este género en el queso Cotija.

A partir de las cepas identificadas y caracterizadas por Guzmán (2015), se realizó la descripción de la estructura poblacional de los *Enterococcus* del queso Cotija. En este trabajo se describe la estructura poblacional de *E. faecalis* y se analiza los hallazgos reportados por Guzmán (2015) y Urrieta (2018), para la especie *E. faecium*.

8.1.1 MLST y estructura poblacional de *E. faecalis*

Para determinar la estructura poblacional de las cepas de *E. faecalis* aisladas del queso Cotija nos basamos en el análisis de secuencias multilocus (MLST), el cual resulta muy útil para determinar la asociación de cepas con los denominados CC de alto riesgo, que se caracterizan por estar compuestos de cepas causantes de infecciones hospitalarias. Para llevar a cabo esta metodología se utilizaron cepas de *E. faecalis* que ya han sido caracterizadas en el grupo de trabajo, identificadas como: A, E, F, QA1, QB2, QC4, A51 (García-Cano et al., 2014; Guzmán, 2015, Anexo I). Para cada cepa se amplificaron y secuenciaron los siete genes requeridos por la metodología reportada en la base de datos (*gdh*, *gyd*, *pstS*, *gki*, *aroE*, *xpt* y *yigL*); posteriormente se realizó el análisis de esas secuencias interrogando la base de datos

PubMLST y de esta manera se obtuvo el número alélico de cada gen, y finalmente la combinación alélica nos dio como resultado una secuencia tipo (ST).

Como parte del análisis, las secuencias nucleotídicas obtenidas para cada gen se identificaron en la base de datos no redundante (nr) usando el algoritmo Blastn (nucleótidos) con el fin de corroborar la correcta amplificación y purificación de cada uno de los genes. De dicho escrutinio se determinó que los genes *gyd* y *gdh*, de tres de las cepas en este estudio apuntaban a pertenecer a la especie *E. faecium* con un porcentaje de identidad mayor al 98% y con más del 90% de cobertura de la secuencia. Dado lo anterior, se llevaron a cabo diferentes estrategias para corroborar la identidad de las cepas en cuestión pues se tenía una identificación incorrecta de nuestras cepas. Para ello primero, se verificó que no existiera una contaminación en las cepas usando PCR especie-específica (metodología descrita por Guzmán, 2015), y los resultados arrojaron una identificación consistente como *E. faecalis*. A continuación, se reamplificaron los genes en cuestión y los amplicones fueron re-secuenciados en dos diferentes instituciones (IBT y Macrogen Corea). Los resultados nuevamente indicaban una identidad como *E. faecium*. Finalmente, el análisis bibliográfico no dice que recientemente se ha reportado que existe transferencia horizontal de los genes de mantenimiento entre las especies *E. faecalis* y *E. faecium*, los cuales son empleados para realizar la técnica MLST, dicha transferencia se presenta principalmente en los genes *gdh* y *gyd* que codifican para las enzimas glucosa-6-fosfato deshidrogenasa y gliceraldehido-3-fosfato deshidrogenasa, respectivamente. Además, estos genes pueden llegar a compartir un porcentaje de identidad >73% (Fig. 7). En conclusión, una transferencia horizontal de estos genes podría explicar los resultados observados en este estudio.

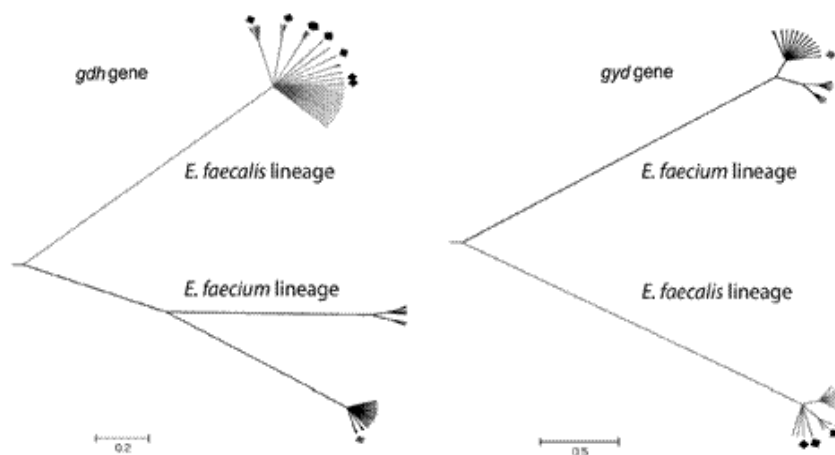


Fig 7. Tomada de Willems (2010). Filogenia de los genes *gdh* y *gyd*. Los diamantes presentes en las ramas pertenecen a los genes de aislados de *E. faecalis* y *E. faecium* encontrados en los grupos a los que no pertenecen, lo que indica una transferencia horizontal de dichos genes.

Para corroborar si esta coincidencia sucedía a menudo, se realizó un alineamiento de nucleótidos seleccionado aleatoriamente el alelo 1 del gen *gdh* de *E. faecium* y el alelo 11 del mismo gen de *E. faecalis*, los cuales son alelos diferentes a los encontrados en nuestros resultados. Como se puede observar en la Figura 8, estos genes comparten un 100% de similitud y cobertura, con lo que se sugiere un cambio en la metodología pues con este tipo de eventos la discriminación entre cepas de la misma especie podría ser de forma incorrecta. Aprovechando la era de la secuenciación masiva, algunos grupos de investigación han propuesto el uso de lo que se conoce como “core-genome” MLST (cg-MLST), donde se analizan todos los genes pertenecientes al core del genoma de cada especie y de esta manera se logra tener un alto poder discriminatorio entre diferentes cepas (de Been et al. 2015). Por ahora, el MLST convencional nos da un buen panorama o diagnóstico rápido del posible origen de las cepas en estudio, además de que es sigue siendo una técnica de menor costo comparado con la secuenciación masiva, además la cantidad de datos disponible por esta metodología es cada vez en mayor haciendo posible un mejor análisis poblacional.

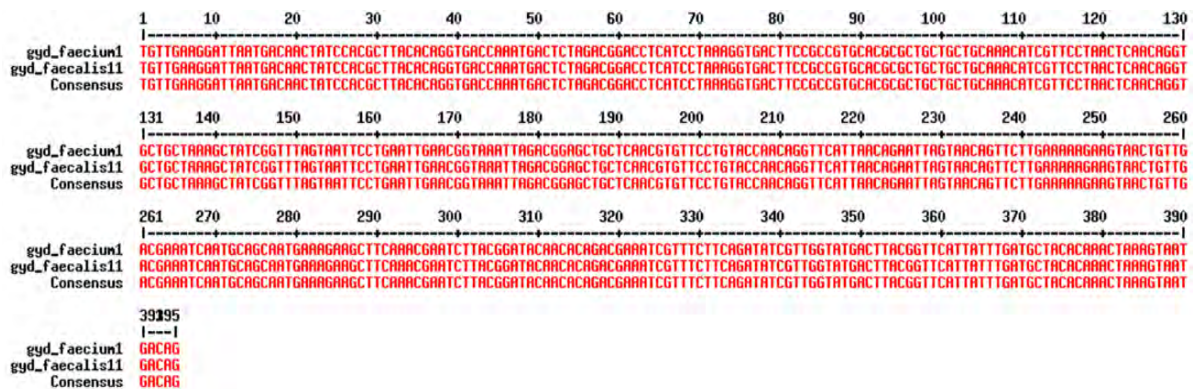


Fig 8. Alineamiento de las secuencias del gen *gdh* de alelo 1 de *E. faecium* y el alelo 11 de *E. faecalis* utilizando el software/plataforma online MultAlin (Corpet F. 1988). Los alelos comparten un 100% de similitud.

En la Tabla 2 se muestran los resultados de los perfiles alélicos del MLST para *E. faecalis*. Se puede observar que tenemos alelos nuevos (no reportados previamente), para tres genes diferentes, los cuales se muestran sombreados en naranja. Después de someter los nuevos alelos a revisión por el curador de la base de datos, se asignaron los números de alelos para *gdh* 93 y 92, *gyd* 33 y *xpt* 80, y de esta manera se asignaron las nuevas ST: 712 para E y F, 757 corresponde a QB-3 y finalmente 758 para QC-4 y A5-1. Las cepas A y QA1 presenta la

ST 21, reportada previamente en una amplia gama de aislados de diferentes nichos ecológicos.

Tabla 2. Resultados de número de alelo y secuencia tipo para las cepas de *E. faecalis* del queso Cotija

Cepa	ST	Año	Estado	<i>gdh</i>	<i>gyd</i>	<i>pstS</i>	<i>gki</i>	<i>aroE</i>	<i>xpt</i>	<i>yiqL</i>
A	21	2008	Michoacán	1	7	9	1	1	1	1
E	712	2008	Michoacán	20	3	7	28	23	80	2
F	712	2008	Jalisco	20	3	7	28	23	80	2
QA1	21	2012	Michoacán	1	7	9	1	1	1	1
QB3	757	2012	Jalisco	93	33	16	1	1	1	1
QC4	758	2012	Jalisco	92	11	9	1	1	1	1
A51	758	2008	Jalisco	92	11	9	1	1	1	1
CC2	2			15	7	3	2	6	1	5
CC2	6			12	7	3	7	6	1	5
CC9	9			4	6	16	4	1	1	4

Los ST y alelos en amarillo son nuevos y han sido asignadas por el curador de la base de datos.

Complejos Clonales (CC) 2, 6 y 9 han sido establecidos como complejos clonales de alto riesgo (Ruiz-Garbajosa, 2006).

En el análisis de la estructura poblacional, se compararon los aislados del queso Cotija con toda la base de datos a través del algoritmo de eBurst, el cual sirve para hacer una representación de evolución genética y agrupamiento de poblaciones a partir de los datos obtenidos por MLST. Éste se basa en identificar el número máximo de alelos compartidos entre las diferentes cepas; cabe señalar que el número mínimo de alelos que deben ser compartidos entre los aislados es definido por el usuario. Para este análisis, en un inicio, se empleó un grado de astringencia alto en el cual se pueden definir complejos clonales, es decir grupos de cepas que comparten al menos 5 de los 7 genes analizados, así como genotipos fundadores que son aquellos que en algún momento se diferenciaron en los 7 alelos y que comienzan a tener otras cepas compartiendo al menos 5 genes, formando de esta manera un nuevo CC (Martínez-Castilla, 2007).

Un análisis astringente no permitía definir la agrupación de las cepas por lo que se cambiaron los parámetros y para la construcción de este diagrama se estableció que las cepas compartieran al menos 4 de los 7 genes, lo cual podría implicar un análisis laxo, pero fue la

única estrategia con la cual se logró la visualización de los CC de alto riesgo. En la Figura 9 se puede observar que las cepas de este trabajo están fuera de los CC de alto riesgo 2 y 9; los cuales engloban a las ST 2, 6, 49, 51 y 9, 17, 18, 42 y 52, respectivamente (Ruíz-Garbajosa et al., 2006). Las cepas A, QA1, QC4 y A51 se encuentran formando parte del complejo clonal 21 (ST 21, 22, 5, 46, 50, 157, 202 y 70), el cual ya ha sido reportado por estar enriquecido de cepas aisladas de animales, humanos sanos, quesos artesanales de Francia y Portugal y algunos casos de heces de humanos hospitalizados. Palmer y colaboradores (2012) reportan que las cepas causantes de infecciones son diferentes de las cepas colonizadoras del tracto gastrointestinal, ya que las primeras están muy adaptadas y especializadas al ambiente nosocomial, por lo que su probable presencia en las heces humanas no necesariamente indica un origen intra-hospitalario. Así mismo, las cepas E y F, que comparten la misma secuencia tipo, a su vez comparten al menos 5 de 7 alelos con cepas aisladas de otros quesos artesanales franceses y portugueses, y de hortalizas (cebollas y aceitunas), las cuales tendrían a la ST 72 como fundadora, por lo que se propone la formación del CC72 el que podría estar agrupando cepas de alimentos. Para corroborar esta hipótesis se propone seguir caracterizando cepas aisladas de alimentos a través de esta metodología.

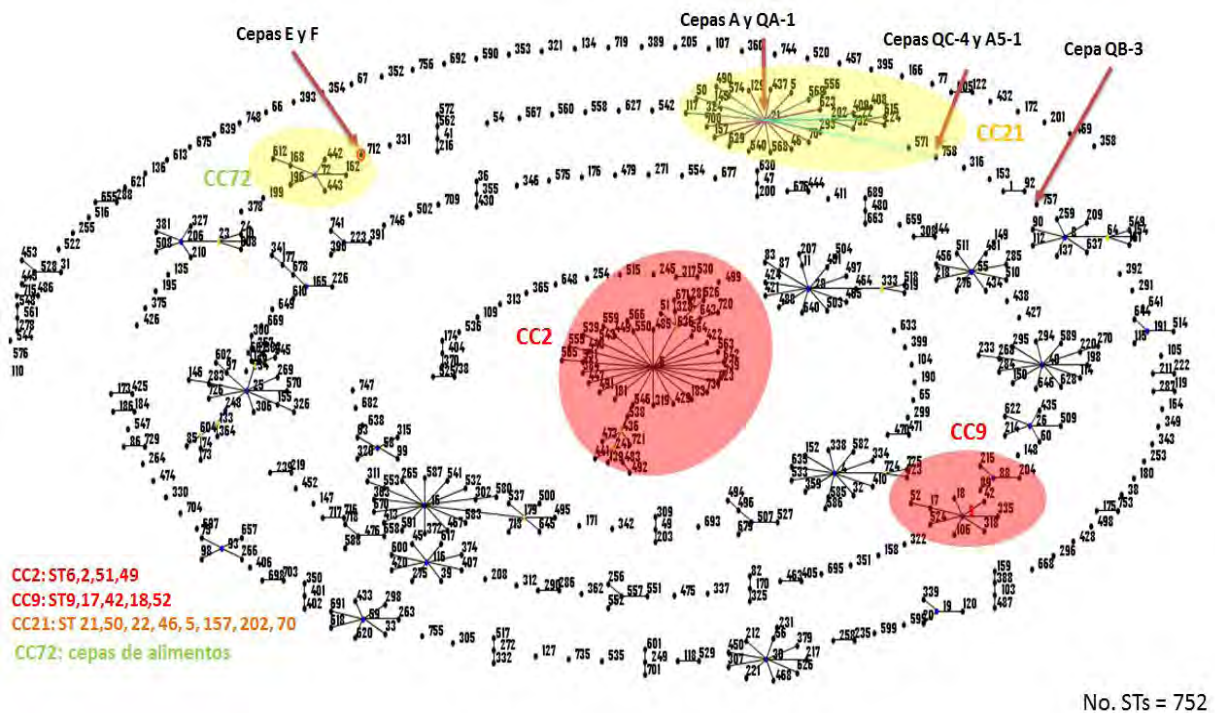


Fig 9. Análisis de estructura poblacional de *E. faecalis* utilizando el programa eBurst v3 y la base de datos completa de esta especie (<http://pubmlst.org/efaecalis/>). En rojo se muestran los CC de alto riesgo, en amarillo el CC 21 y en naranja el CC 72 de alimentos. El diagrama fue construido estableciendo que las cepas compartieran al menos 4 de los 7 genes de mantenimiento.

Con el objetivo de tener un análisis más estricto de los cambios puntuales y su efecto en la asociación de cepas de diferentes ambientes, se construyó un árbol filogenético usando las secuencias nucleotídicas concatenadas de los 7 genes, usados para el MLST, como una única secuencia. Para comparar se tomaron las ST reportadas para cepas de otros nichos ecológicos: alimentos, agua, probióticos e infecciones hospitalarias; con el fin de corroborar la no asociación de los aislados del queso Cotija con un ambiente nosocomial. En el Anexo II se presenta la información de todas las cepas incluidas en este análisis.

En la Figura 10 se muestra el árbol de agrupamiento. En él se puede ver que las cepas aisladas del queso Cotija E y F se agrupan con cepas aisladas de otros alimentos (con círculos amarillos), así como con cepas de origen no nosocomial. En ese mismo clado se encuentran cepas aisladas de vegetales, donde la ST 168 ha sido reportada para 14 cepas de diferente origen como lechugas, aceitunas, apio y cerezas (Burgos et al., 2009). Por otro lado, las cepas del CC21 al que pertenecen las cepas A y QA1 quedaron en un clado separado (sostenido por un bootstrap de 76), formado por cepas nosocomiales (círculos rojos) y no nosocomiales. Las cepas QC4 y A5-1 quedaron fuera de cualquier grupo, indicando su no pertenencia a los complejos nosocomiales, los cuales se observan en la parte inferior del árbol. Para enraizar el árbol se eligió a una cepa nosocomial de *E. faecium* con el fin de corroborar que a través de esta metodología hay una diferenciación entre ambas especies.

Con el empleo de esta estrategia se pudo corroborar el resultado obtenido por el análisis de eBurst, donde concluimos que las cepas aisladas del queso Cotija no se agrupan por cepas reportadas como aislados de infecciones nosocomiales (CC2 y 9). De la misma forma, este análisis nos deja ver la agrupación de cepas aisladas de alimentos en lo que hemos denominado el “Complejo Clonal 72”, por ser la secuencia tipo 72 la fundadora. Sin embargo, es importante mencionar que al menos una cepa aislada de un ambiente nosocomial se agrupa con las de otros ambientes, lo cual nos habla de la dispersión de cepas patógenas a un ambiente diferente al hospitalario.

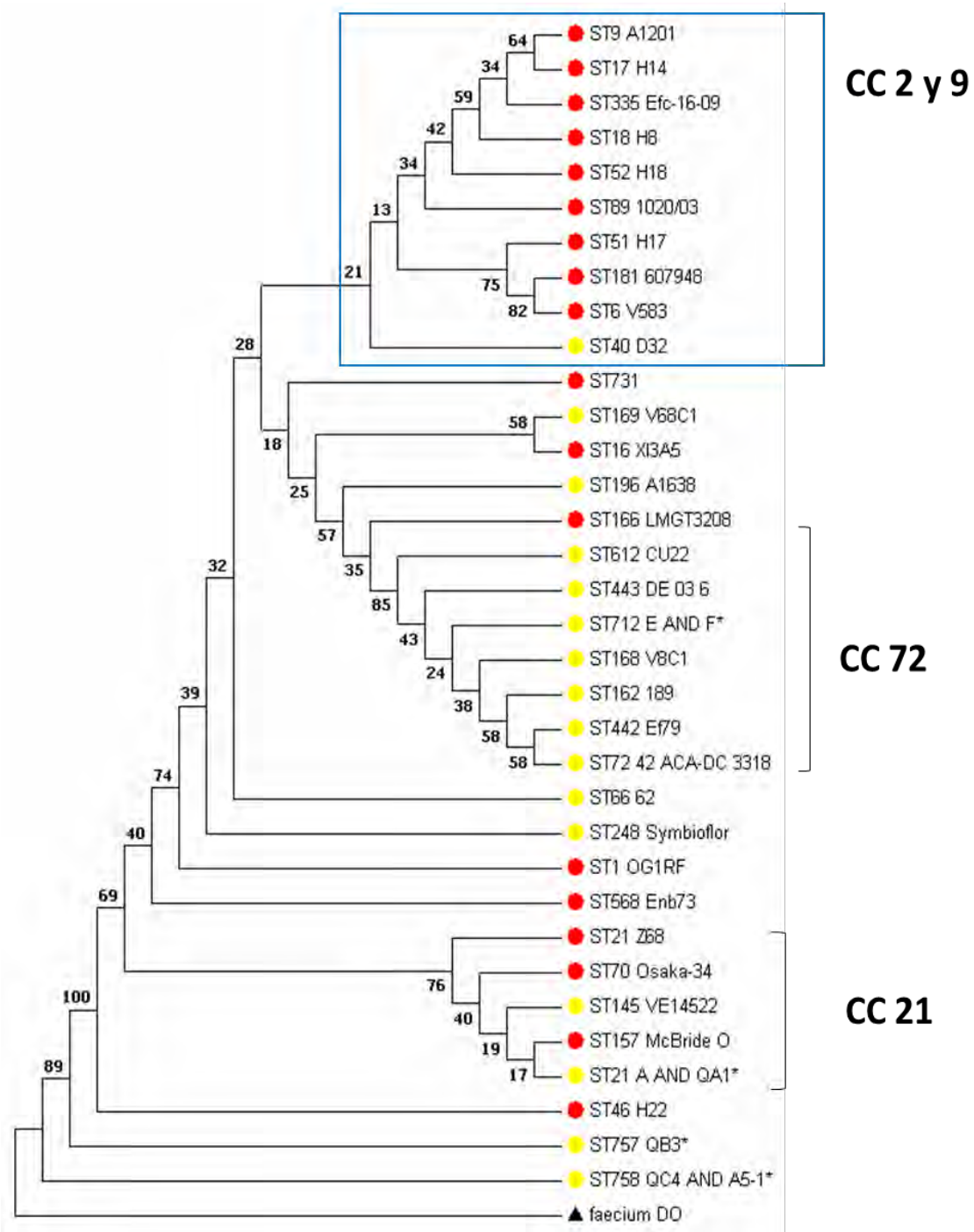


Fig 10. Árbol de agrupación de cepas de *E. faecalis* construido con las secuencias concatenadas de los genes *gdh*, *gyd*, *pstS*, *gki*, *aroE*, *xpt* y *yiqL*. El análisis se realizó con el método de Neighbor joining. En círculos rojos se resaltan las cepas que pertenecen al CC 2 y 9 (recuadro azul), y en círculos amarillos cepas reportadas para alimentos y no nosocomiales (CC 21 y 72). Las cepas aisladas del queso Cotija se muestran con asteriscos. Los mismos 7 genes conservados para *E. faecium* DO (TX16) se utilizaron para enraizar el árbol (triángulo negro).

8.1.2 MLST y estructura poblacional de *E. faecium*

Esta parte del proyecto, como se ha mencionado previamente, fue realizada por Guzmán en 2015, quien reportó la presencia de un nuevo alelo (74) para el gen *ddl* y 4 ST nuevas, las cuales fueron asignadas por el curador de la base de datos (<http://pubmlst.org/efaecium/>). A las cepas B y QD2 se les asignó la ST 820 y 135, respectivamente y estas comparten 6 de los 7 alelos con la ST de cepas aisladas de quesos italianos y noruegos (A0037 y A0038), de la misma manera el resto de las cepas del queso Cotija comparten al menos 4 alelos con dichas cepas. Las cepas D y B por otro lado, comparten 6 de los 7 alelos con cepas aisladas de heces de animales y de humanos sanos, mientras que las cepas C y G pueden llegar a compartir 5 alelos con ST asignadas a cepas aisladas de infecciones. Para complementar este análisis se contó con los resultados obtenidos por Urrieta (2018), quien caracterizó cepas de *E. faecium* aisladas de otro queso artesanal mexicano llamado “queso Bola de Ocosingo”, Chiapas. De sus resultados se resalta que nuevamente se encontraron y reportaron alelos nuevos para los genes *gyd* y *pstS* y con ello se dio lugar a la asignación de tres nuevas secuencias tipo (Tabla 3).

Tabla 3. Resultados de número de alelo y secuencia tipo para las cepas de *E. faecium* del queso Cotija (Guzmán, 2015 y Urrieta, 2018).

Cepa	ST	Región	Año	<i>atpA</i>	<i>ddl</i>	<i>gdh</i>	<i>purK</i>	<i>gyd</i>	<i>pstS</i>	<i>adk</i>
B	820	La Tinaja (Michoacán)	2008	5	7	5	7	1	1	1
C	1045	S. Ma. del oro (Jalisco)	2008	51	2	12	9	1	1	5
D	1046	S. Ma. del oro (Jalisco)	2008	3	7	5	7	1	1	1
G	1045	La Tinaja (Michoacán)	2008	51	2	12	9	1	1	5
QD-2	135	Quitupán (Jalisco)	2015	27	3	1	2	1	1	1
QE-2	1047	Quitupán (Jalisco)	2015	27	3	25	2	1	12	1
QG-5	1048	S. Ma del oro (Jalisco)	2015	8	74	12	6	1	1	5
QO 9 y 10	130	Ocosingo (Chiapas)	2016	8	4	12	9	1	1	5
QO 6	1293	Ocosingo (Chiapas)	2016	8	4	12	9	1	124	5
QO 3	1294	Ocosingo (Chiapas)	2016	8	4	12	9	53	123	5
QO 2	1295	Ocosingo (Chiapas)	2016	8	4	12	6	53	1	5
CC17	17	Complejo clonal de alto riesgo (CC 17)		1	1	1	1	1	1	1
CC17	18			7	1	1	1	1	1	1
CC17	78			5	1	1	1	1	1	1

*Los ST y alelos en amarillo son nuevos y han sido asignadas por el curador de la base de datos.

CC 17: complejo clonal de alto riesgo para la salud pública (intrahospitalaria).

Leavis y colaboradores (2006) determinaron que existe un complejo clonal de alto riesgo para la salud pública denominado CC 17 (ST 17, 18, 71, 16, 204 y 412) y se concluyó que dichas ST se caracterizan por presentar el alelo 1 para el gen *purk*. De esta manera se han catalogado e identificado cepas de origen nosocomial (Leavis et al., 2006; Freitas et al., 2009). En la Tabla 3 se observa que ninguna de las cepas aisladas de los quesos evaluados presenta el alelo 1 para el gen *purk*, el cual es característico de cepas de aislados clínicos, por lo que de esta manera se puede presumir que su origen no es nosocomial (Freitas et al, 2009). Además, a través del análisis de estructura poblacional se determinó que las cepas del queso Cotija y Ocosingo no pertenecen al CC de alto riesgo 17. Por el contrario, se agrupó a las cepas (C, G, 9 y 10) en el CC5 (fundado por las ST5, ST6, 150 y ST185), el cual ha sido descrito por estar enriquecido por cepas aisladas de porcinos, así como de heces de humanos u otros animales. Se ha planteado que el CC5 con el tiempo se ha ido esparciendo y ha comenzado a agrupar cepas de alto riesgo, por lo que se resalta el papel de los huéspedes no hospitalarios como posibles depósitos de enterococos patógenos y multirresistentes a antibióticos (Freitas et al., 2009). Por su parte, las cepas B, D y QD2, se encuentran formando parte del CC 22 (ST32, ST392, ST435 y ST436), el cual está enriquecido de cepas aisladas de heces de humanos y muestras ambientales, que se caracterizan por ser resistentes a tetraciclina, ciprofloxacino y a altas concentraciones de kanamicina (Fig 11).

Aunque se esperaba que las cepas estuvieran estrechamente relacionadas entre sí, dado su origen alimentario, dicha asociación no pudo ser establecida, sin embargo, se estableció que no pertenecen al CC de alto riesgo (17). A pesar de estos resultados, se sugiere que las cepas aisladas de alimentos siempre sean caracterizadas a profundidad, incluyendo la evaluación de resistencia a antibióticos, ya que se ha reportado una expansión y diseminación del CC17 a otros ambientes fuera del hospitalario por lo que ahora podemos encontrar cepas de origen animal, humano y alimentos que presentan ST pertenecientes a dicho complejo.

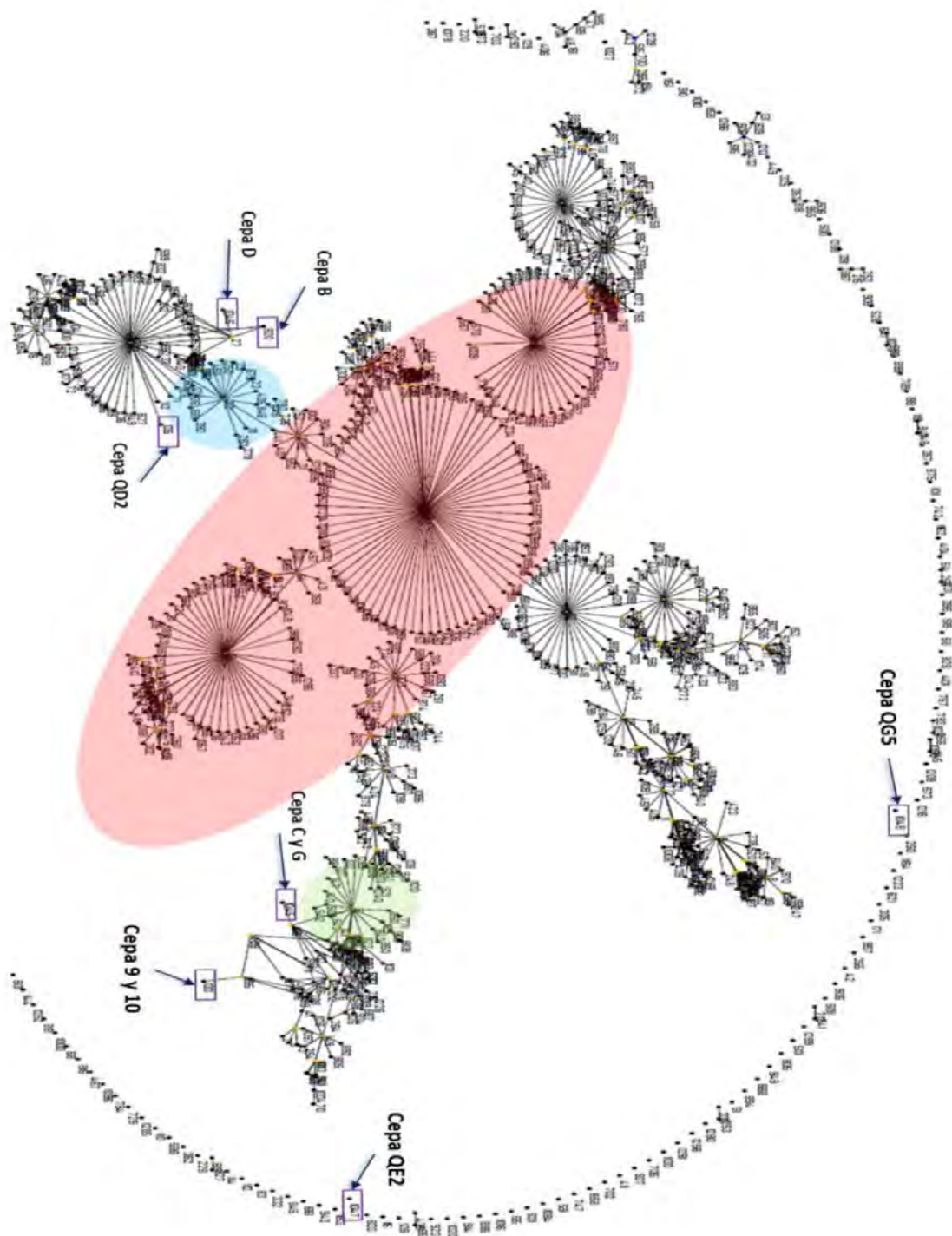


Fig 11. Análisis de estructura poblacional de *E. faecium* utilizado el programa eBurst v3 y la base de datos completa de esta especie (<http://pubmlst.org/efaecium/>). Sombreado con rojo se resalta el CC de alto riesgo 17, en verde el CC5 y en azul el CC22. El diagrama fue construido estableciendo que las cepas compartieran al menos 5 de los 7 genes de mantenimiento.

En el caso del estudio de estructura poblacional de *E. faecium* también se incluyó un análisis de alineamiento de secuencias, con el fin de corroborar lo observado mediante el diagrama generado por eBurst. En la Figura 12 se muestra el árbol de agrupamiento, el cual nos muestra que las cepas del queso Cotija (con círculos amarillos y asteriscos) se encuentran fuera del CC17 de alto riesgo. Las cepas B, QD2 y QE2 se agruparon juntas en lo que nombramos clado "A", cabe mencionar que estos aislados pertenecen a quesos Cotija de diferentes colectas y zonas productoras (2008 y 2012). Mientras tanto, otras cepas se encuentran en un clado enriquecido de las cepas aisladas de diferentes quesos artesanales mexicanos, a este grupo lo denominamos "clado B". Como caso extraordinario se tiene a la cepa D agrupada en un clado diferente pero muy cercano al de la cepa NRRL B-2354. Dicha cepa, en los años 70, era usada como cultivo iniciador para la elaboración de quesos y actualmente se utiliza como indicador positivo de tratamientos térmicos en plantas de alimentos. Finalmente, se observa la formación del clado C, conformado por cepas de origen no nosocomial (Fig 12).

A través de este gráfico se puede confirmar la formación del complejo clonal de alto riesgo 17, aquí identificado con las secuencias tipo fundadoras correspondientes: 17, 18 y 78. Además, se identificaron los clados A, B y C conformados por cepas aisladas de diferentes ambientes, es de particular interés el clado B por contener a la mayoría de las cepas analizadas en este trabajo. Finalmente, con la ayuda de esta metodología se resuelve la no asociación con patógenos por parte de las cepas de *E. faecium* aisladas del queso Cotija.

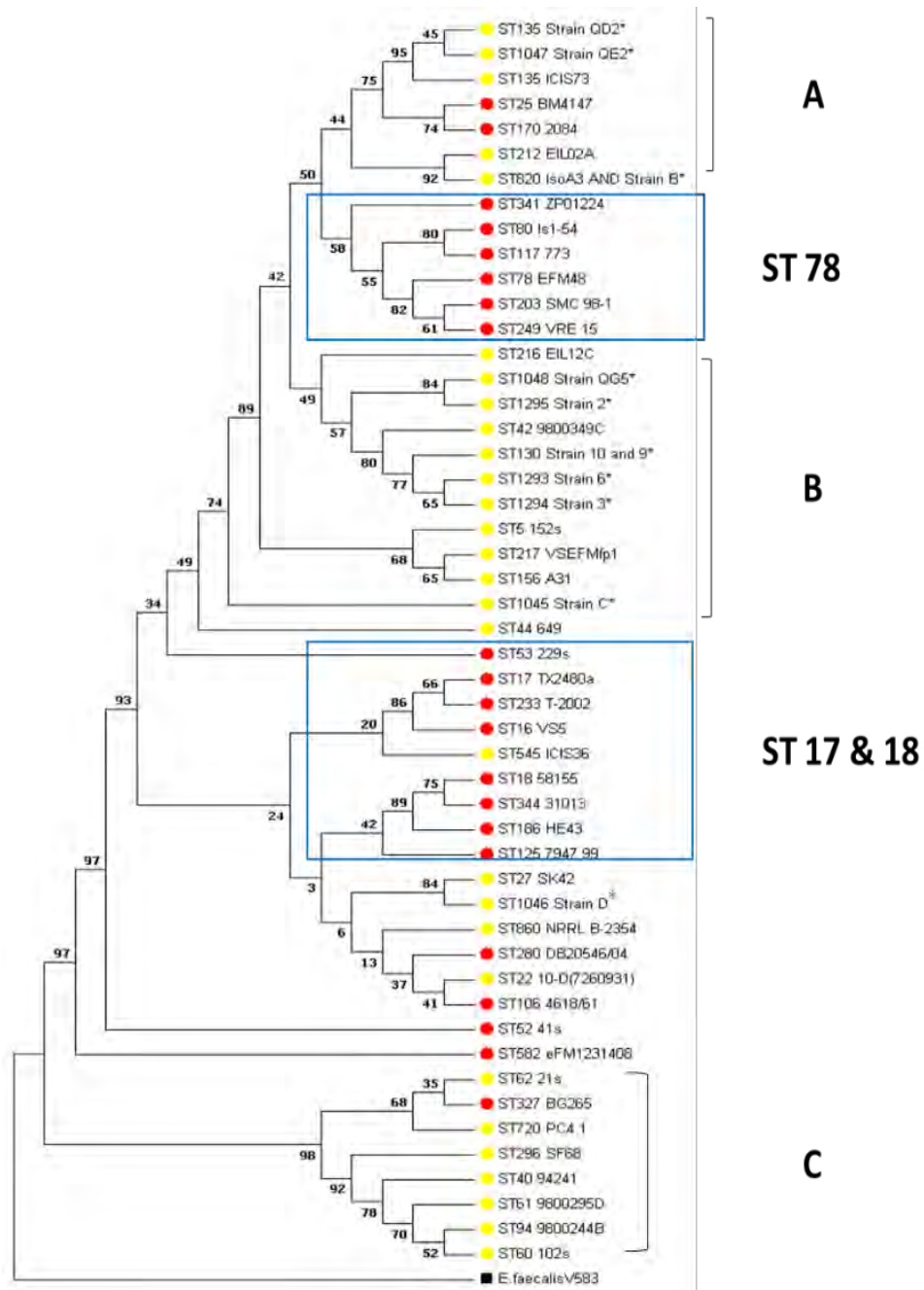


Fig 12. Árbol de agrupación de cepas de *E. faecium* construido con las secuencias concatenadas de los genes *atpA*, *adk*, *ddl*, *gdh*, *purK*, *gyd* y *pstS*. El análisis se realizó con el método de Neighbor joining. En círculos rojos se resaltan las cepas que pertenecen al CC de alto riesgo 17. En círculos amarillos cepas reportadas para alimentos y no nosocomiales. Las cepas aisladas del queso Cotija y queso Bola de Ocosingo se muestran con asterisco. Los mismos 7 genes conservados para *E. faecalis* V583 se utilizaron para enraizar el árbol (cuadro negro). Los clados A, B y C son agrupaciones donde se incluye a las cepas de los quesos artesanales mexicanos.

Además, las cepas QA1, QB3, QC4, A51, QD2, QE2 y QG5 se analizaron para determinar si presentaban genes relacionados factores de virulencia, siguiendo la metodología descrita por Olvera en el 2013. Para todas las cepas fue común la presencia del gen *asa1* el cual codifica para la proteína de agregación, en 6 de ellas se encontró el gen codificante de la proteína de superficie enterococcica *Esp* (ejemplo Fig 13), mientras que 2 cepas de *E. faecium* presentan el gen codificante de la citolisina A (*cylA*), por lo que se propone realizar pruebas de hemolisis que corroboren este genotipo. Con respecto a la susceptibilidad a antibióticos, se determinó que todas eran sensibles a tetraciclina, ampicilina, eritromicina y vancomicina, mientras que fueron resistentes a kanamicina, estreptomina, espectinomicina y en dos casos se observó resistencia a penicilina y gentamicina. Urrieta reportó que las de *E. faecium* aisladas de queso Bola de Ocosingo Chiapas, e incluidas en este análisis, presentaban resistencia kanamicina, espectinomicina y estreptomina, mientras que algunas también eran resistentes a eritromicina, penicilina, vancomicina y tetraciclina, sin embargo, todas fueron negativas para la presencia de genes de virulencia (Urrieta, 2018). Dichos resultados concuerdan con su agrupación en los CC 5 y 22, los cuales se caracterizan por presentar resistencia a diferentes antibióticos. En este punto, hay que recordar que todos los enterococos poseen una resistencia intrínseca de bajo nivel a los aminoglucósidos y macrólidos, a los primeros debido un transporte deficiente del aminoglucósido al interior de la bacteria y éste se caracteriza por presentar valores de concentración mínima inhibitoria (CMI) que oscilan entre 4 y 64 mg/L de gentamicina y entre 16 y 256 mg/L de estreptomina. Por otro lado, la resistencia a macrólidos puede ser dada por varios factores, pero el más ampliamente descrito se basa en el gen *msr(A)*, el cual confiere resistencia a este tipo de antibióticos y a las estreptograminas B (Cercenado, 2011). A través de este análisis podemos decir que en las cepas de *E. faecalis* aisladas de alimentos es más frecuente la presencia de genes que codifican para factores de virulencia, mientras que las cepas de *E. faecium* se caracterizan por albergar genes de resistencia a antibióticos, lo cual ya ha sido ampliamente reportado (Franz et al., 2001).

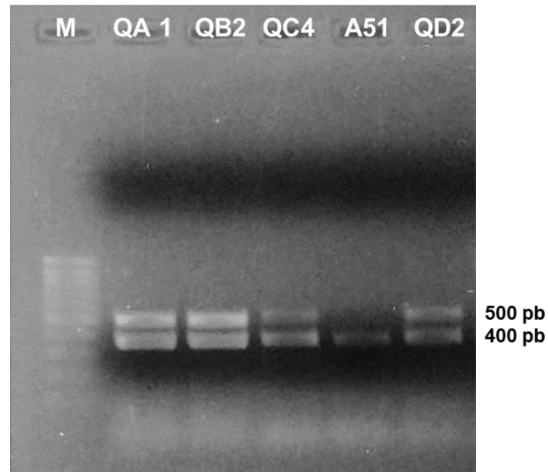


Fig 13. Electroforesis para visualizar amplicones de genes relacionados a factores de virulencia, *asa1* (375pb) y *esp* (510 pb). M: marcador low range (Thermo Fisher Scientific). Gel de agarosa 1%, teñido con bromuro de etidio al 0.1 %.

8.1.3 Secuenciación, ensamble y anotación de los genomas

Para el análisis genómico, se seleccionaron las cepas *E. faecium* D y *E. faecalis* F, ya que son de las que más información previa se tiene, además de que no pertenecen a los CC de alto riesgo. Del proceso de secuenciación se obtuvieron un total de 11,247,214 lecturas pareadas para *E. faecalis* y de 9,082,672 para *E. faecium*. Las lecturas fueron mapeadas contra las cepas de referencia completamente secuenciadas y cerradas (Tabla 4), donde el mayor porcentaje de lecturas ($\geq 75\%$) se obtuvo sobre la cepa Symbioflor1 para *E. faecalis* cepa F y NRRL B-2354 para *E. faecium* cepa D (Anexo III Fig AI). Como parte de la estrategia, se llevó a cabo un ensamble *de novo*, donde se obtuvieron 70 contigs para el genoma de *E. faecalis* F (lecturas totales 3,494,182 pb, N50/N90 de 91,252/ 28,483 pb, y 37.6%GC) y 82 contigs para *E. faecium* D (lecturas totales 2,779,320 pb, N50/N90 de 88,852 /21,856 pb y 37.9% GC). Los resultados estadísticos sugieren que el proceso de cerrado de ensamblado fue el adecuado, obteniendo ensambles mejorados y de alta calidad (Chain et al., 2009). En la Tabla 4 se muestra un resumen de la comparación de los genomas de las cepas F y D contra los genomas de referencia, donde se puede resaltar que la cepa F tiene un genoma de mayor tamaño que las referencias. Experimentalmente, se identificaron 2 plásmidos: uno de ~4,000 pb y otro > 10,000 pb, y a partir de la secuenciación se ensamblaron dos plásmidos de 4,718 y 9,882 pb (anotación en anexo IV, Tabla AIV). Por su parte, la cepa D tiene un tamaño de genoma parecido al de las referencias; sin embargo, se diferencia de las demás cepas debido a la ausencia de plásmidos.

Tabla 4. Comparación de características genómicas de las cepas *E. faecalis* F y *E. faecium* D contra las referencias

Número de acceso	Cepa	Origen	Tamaño de genoma (Mb)	%GC	Genes	Plásmidos	Referencia
CP002621.1	OG1RF	Cepa virulenta de laboratorio	2.74	37.8	2,579	-	Burgogne y Garsin, 2008
HF558530.1	Symbioflor1	Probiótico	2.81	37.7	2,808	-	Fritzenwanker y Kuenne, 2012
CP003726.1	D32	Heces de porcino	3.06	37.5	2,965	2	Zischka y Kuenne, 2012
CP002491.1	62	Heces de infantes sanos	3.13	37.2	3,157	3	Dag y Lars, 2011
AE016830.1	V583	Nosocomial	3.36	37.5	3,264	3	Paulsen et al., 2003
MEGW00000000	Cepa F	Queso Cotija	3.49	37.6	3,271	2	En este trabajo
Número de acceso	Cepa	Origen	Tamaño de genoma (Mb)	%GC	Genes	Plásmidos	Referencia
CP003583.1	DO	Nosocomial	2.7	38.15	2,703	3	Qin et al., 2012
CP004063.1	NRRL B-2354	Utilizada en la industria alimentaria	2.84	37.8	2,639	1	Kopit et al., 2014
CP006620.1	Aus0085	Nosocomial	2.99	38.2	2,922	6	Lam et al., 2013
CP003351.1	Aus004	Nosocomial	3.02	38.31	2,932	3	Lam et al., 2012
MEGX00000000	Cepa D	Queso Cotija	2.77	37.9	2,804	.*	En este trabajo

*Se corroboró experimentalmente a través de dos metodologías

8.1.4 Genómica comparativa

Como primera estrategia se llevó a cabo un alineamiento entre las cepas aisladas del queso Cotija contra sus respectivas referencias. En la Figura 14A se pueden observar grandes bloques de sintenia entre la cepa de *E. faecalis* F (círculo interno negro) y las cepas pertenecientes a diferentes nichos ecológicos, incluido el nosocomial. Dentro del alineamiento circular, las zonas en blanco corresponden a regiones genómicas presentes únicamente en la cepa F. En el mismo análisis para la cepa *E. faecium* D (círculo rojo), se puede ver que hay

varias regiones genómicas que no comparte con las cepas patógenas y, por el contrario, presenta un patrón de bloques de sintenia muy parecido al de la cepa NRRL B-2354 (Fig 15A).

A partir de la predicción de proteínas de los genomas de cada cepa, se realizó un análisis de homología a través de la clusterización de familias de genes codificantes usando una intersección de los resultados obtenidos con las metodologías de triángulos de COG y el algoritmo OrthoMCL (Get-homologues), lo cual nos permitió dilucidar el core del genoma, el pangenoma y genes linaje específicos, así como familias de genes expandidas o comprimidas. La expansión o compresión de familias de genes dentro de un linaje dado puede ser el resultado de selección natural o adaptación al ambiente y el análisis de dicha información nos llevó a conclusiones interesantes.

El análisis del core del genoma de cada especie, en este estudio, se llevó a cabo analizando únicamente los genomas cerrados y anotados reportados en las bases de datos, y los drafts generados para las cepas D y F. Para la especie *E. faecalis* se formaron 4363 clusters de proteínas, donde 2050 eran parte del “core” del genoma (presentes en todos los genomas), 2211 clusters formando el “soft-core” (en el 95% de los genomas analizados), 1820 clusters en el “cloud” (en al menos 2 de los genomas analizados) y 332 clusters referidos al “shell” (sólo en uno de los genomas). Para la especie *E. faecium* se identificaron un total de 3908, de los cuales 1844 clusters están formando el “core” del genoma, 2063 clusters integrando el “soft-core”, 1864 clusters como parte del “cloud”, así como 376 clusters en el “shell”. A través de un script auxiliar (compare_clusters.pl) se generó una matriz de ausencia presencia de proteínas del pangenoma, los cuales fueron inferidos a partir de los cluster generados por el script principal. El análisis de la matriz nos permitió obtener los genes únicos para una de las cepas D y F (531 y 338 respectivamente).

Teniendo como base todos los clusters de proteínas obtenidos se generó un árbol de parsimonia con el objetivo de visualizar con cuál de las cepas comparadas se tenía mayor similitud. El árbol obtenido deja ver que la cepa F se agrupa en un mismo clado con la cepa probiótica Symbioflor1 y en un clado diferente está agrupada la cepa patógena V583 (Fig 14B). Empleando la misma metodología, la cepa D se agrupó con la cepa no patógena NRRL-B2354, dejando en un clado diferente a las cepas patógenas (Fig 15B).

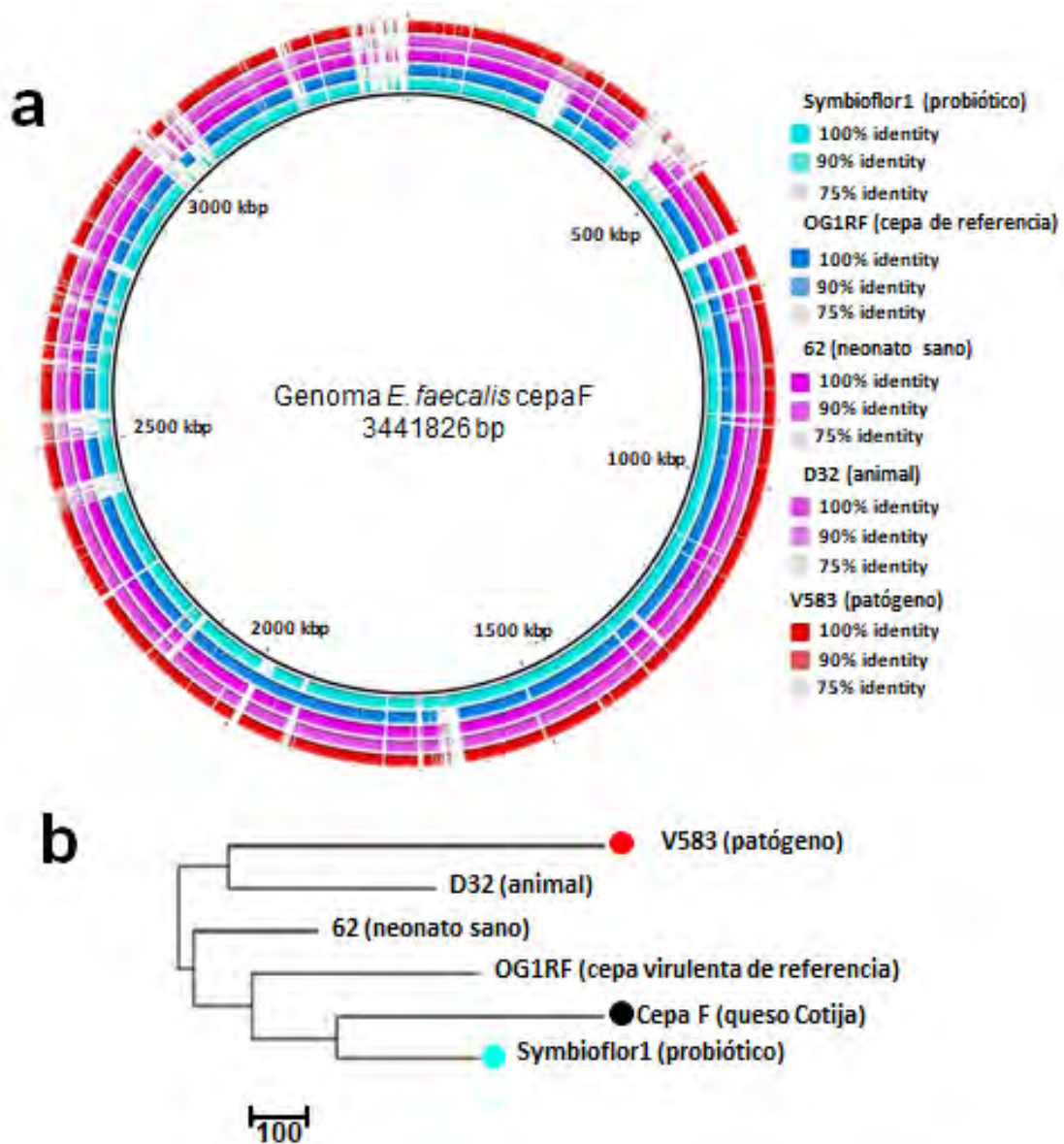


Fig 14. Alineamiento de genomas de *E. faecalis* provenientes de diferentes ambientes. (A) Alineamiento circular, espacios vacíos (blanco) representan genes presentes únicamente en la cepa F. (B) Árbol de parsimonia del pangenoma de las seis cepas, derivado de la matriz de ausencia/presencia (OMCL y COGtriangles). El árbol realizado con PARS desde PHYLIP package y visualizado con MEGA6.

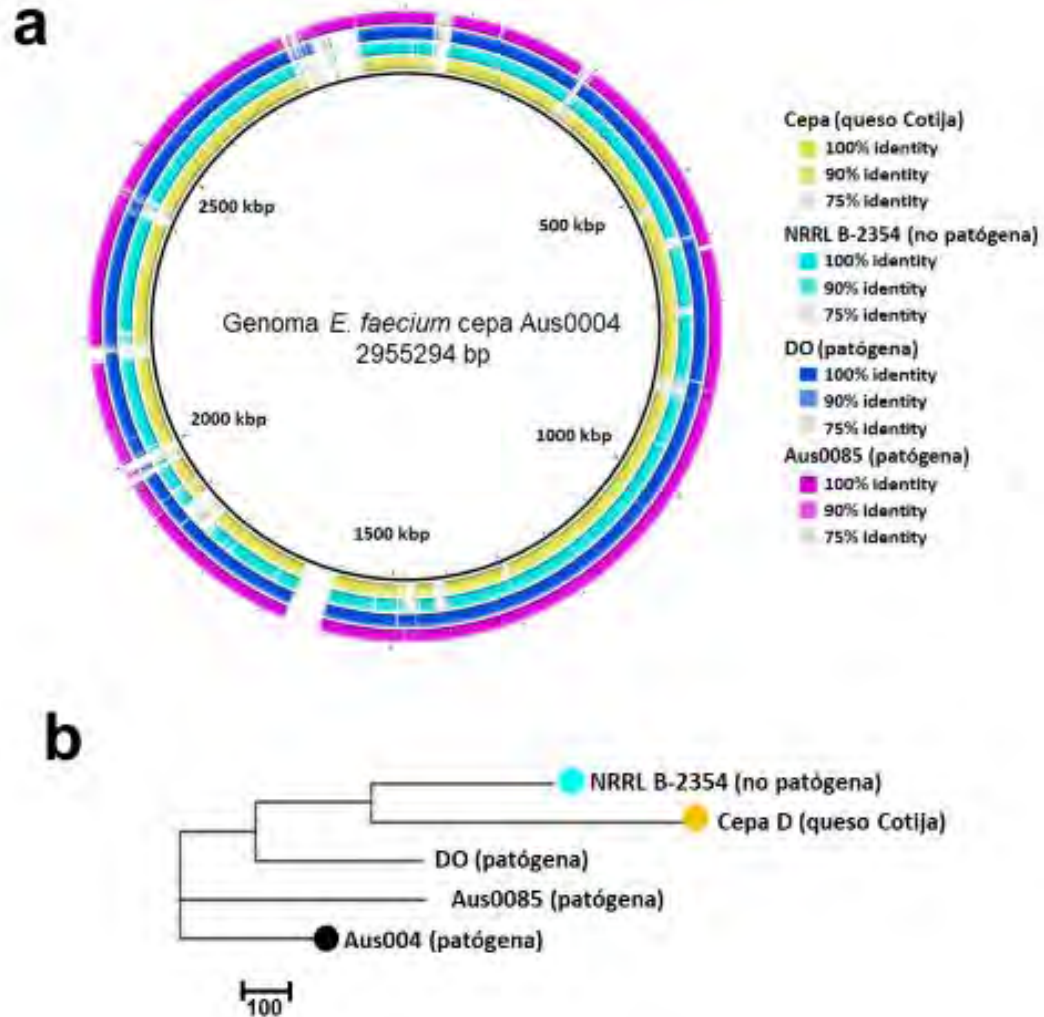
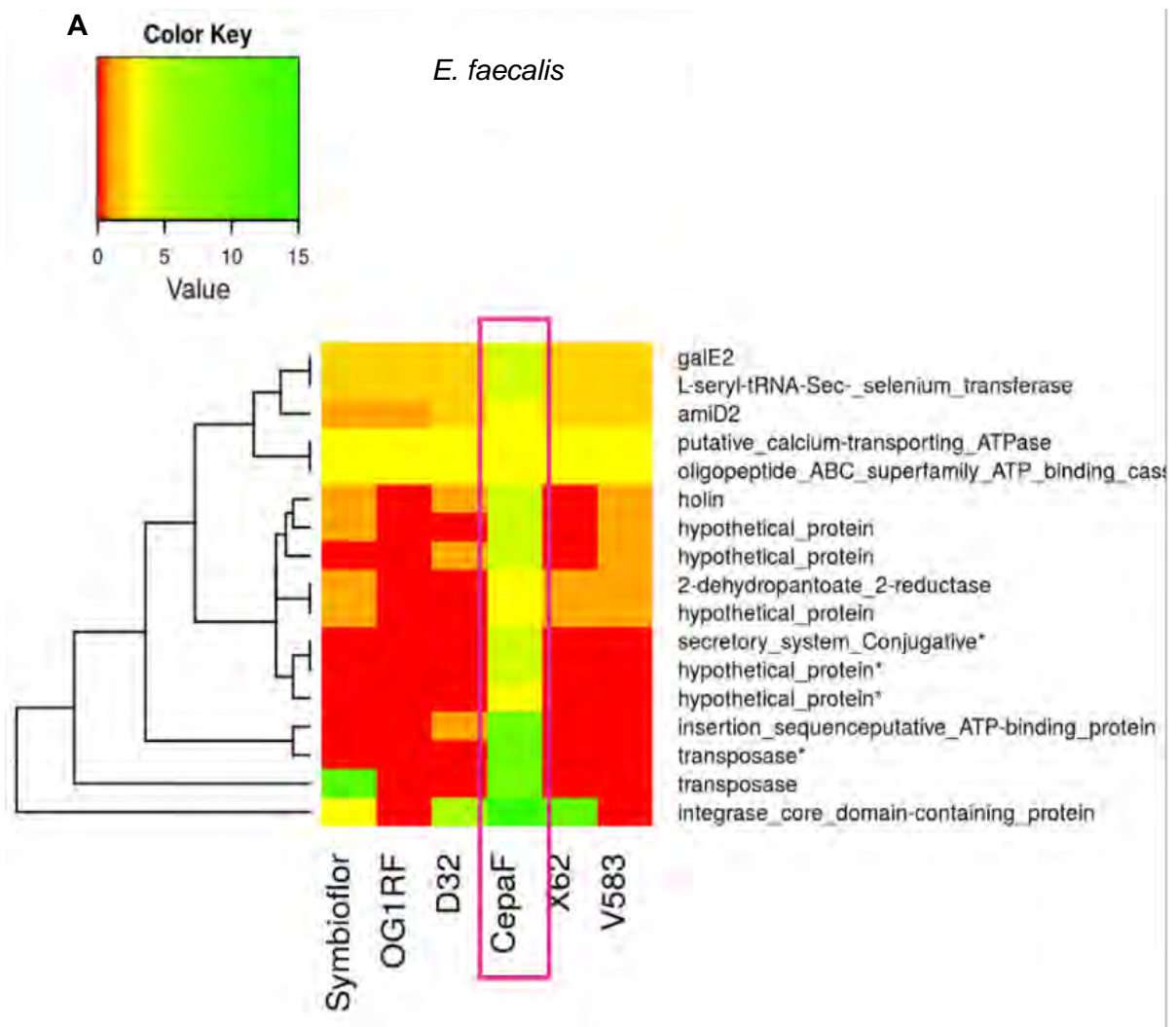


Fig 15. Alineamiento de genomas de *E. faecium* provenientes de diferentes ambientes. (A) Alineamiento circular, espacios vacíos (blanco) representan genes presentes únicamente en la cepa Aus0004 (patógeno hospitalario). (B) Árbol de parsimonia del pangenoma de las seis cepas, derivado de la matriz de ausencia/presencia (OMCL y COGtriangles). El árbol realizado con PARS desde PHYLIP package y visualizado con MEGA6.

Con la ayuda de la matriz obtenida del análisis de ortología previamente descrito, se pudieron identificar familias de genes codificantes de proteínas enriquecidas en las cepas D y F, es decir, que presentaban varias copias para algunos genes en comparación con las referencias. En ambos casos se puede notar que las familias relacionadas con la utilización de azúcares como la galactosa (*galE2*), actividades relacionadas con la autólisis por ruptura de la pared celular (*amiD2*), así como otras pertenecientes al sistema proteolítico (Clp B y OppB) están

expandidas (Fig 16 A y B), probablemente por la necesidad de utilización de nutrientes presentes en el queso Cotija donde en mayor proporción se encuentra la caseína, grasa butírica y como fuente de carbono se encuentra la lactosa. Resulta interesante, la presencia de varias copias del gen *amiD2* ya que esta actividad de autólisis de la célula, la cual podría estar relacionada con la liberación del complejo Clp al espacio extracelular, y de esta manera, dicho complejo podría iniciar la hidrólisis de caseína, principal nutriente proteínico del queso. Sin embargo, genes altamente transferibles como transposasas y elementos de inserción también están sobre representados, lo cual nos habla de la naturaleza de este género a generar sitios de recambio e inserción.



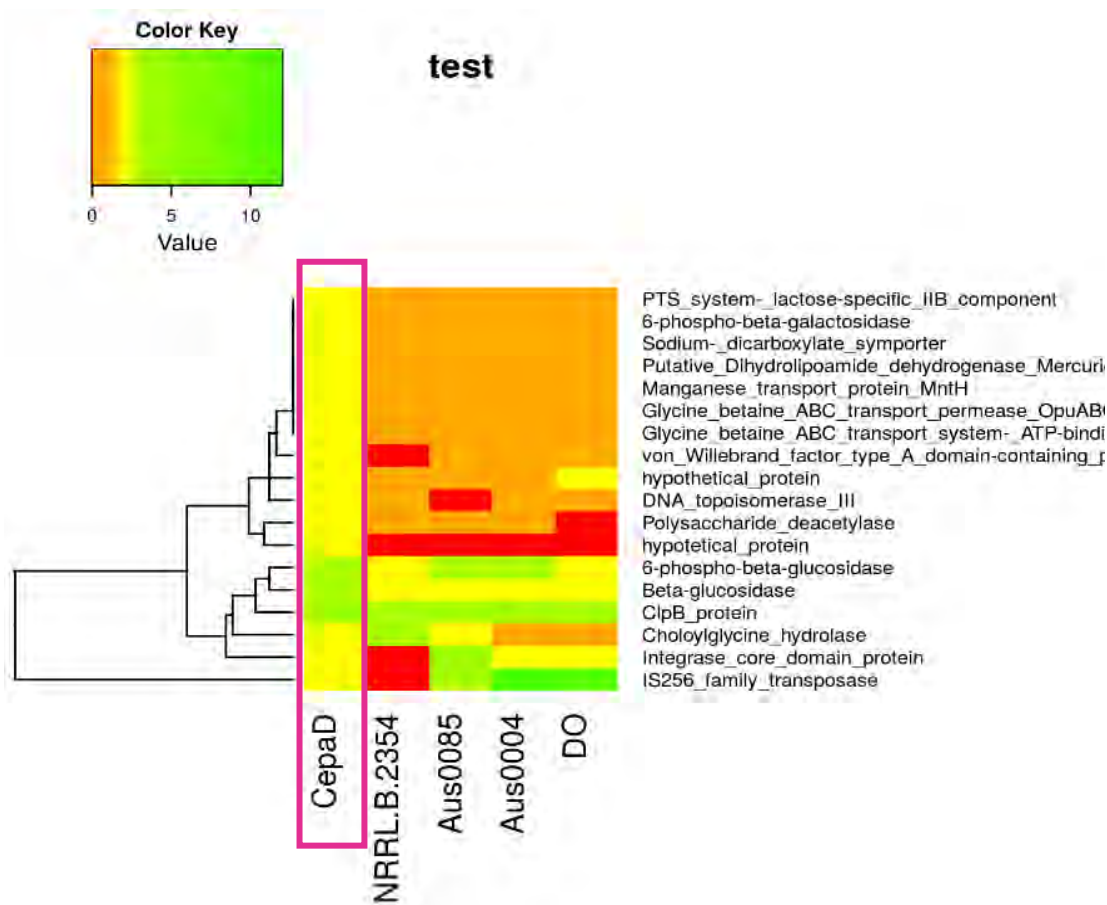


Fig 16. (A) Heatmap de familias de proteínas expandidas en *E. faecalis* F, donde las relacionadas a uso de galactosa y peptidoglucano hidrolasas se ven enriquecidas (más de una copia del gen), así como algunas transposasas y elementos de inserción. (B) Heatmap de familias de proteínas expandidas en *E. faecium* D, donde el uso de galactosa (lactosa) y genes del sistema proteolítico están en varias secuencias con respecto a las referencias. La escala va de 0 a 15 copias del gene.

8.1.5 Descripción del potencial de patogenicidad en las cepas del queso Cotija

Dado que se contaba con la anotación funcional para cada gen, se decidió hacer una búsqueda exhaustiva de todos los genes, reportados, que codifican para factores de virulencia y producción de aminas biógenas con el fin de dilucidar el potencial de patogenicidad de ambas cepas. Como antecedentes, en estudios preliminares de la evaluación de la patogenicidad de *Enterococcus* spp. aislados del queso Cotija, se buscaron por medio PCR los genes relacionados a la síntesis de una sustancia de agregación (*asa1*), citolisina (*cytA*) y una proteína de superficie enterococcica (*esp*), y se concluyó que la cepa *E. faecalis* F era

positiva para la presencia del gen *asa1* y pruebas fenotípicas de adherencia a células de colon corroboran la síntesis de proteínas involucradas en la colonización de células de mamíferos (Olvera, 2013).

El análisis de la anotación funcional de la cepa *E. faecalis* F nos permitió dilucidar otros genes, diferentes a los ya identificados, que codifican para factores de virulencia, como son: gelatinasa (*gelE*), proteína de unión a colágeno (*ace*), hemolisina (*hly III*) y genes involucrados en la biosíntesis de cápsula (Orgier y Serror, 2008). Previamente, se reportó que un fenotipo negativo asociado a la expresión de la hemolisina, corroborado en pruebas de hidrolisis en agar sangre, así como el fenotipo asociado a la capacidad de formación de biopelículas por presencia de la proteína Esp, evaluado en placas de poliestireno (Olvera, 2013). Por otra parte, la síntesis del gen que codifica para la gelatinasa está regulada por el sistema Frs, el cual consta de los genes *frsA-D*, el cual se encuentra incompleto en el genoma de la cepa F, ya que sólo se identificó el gen *frsB*, que por sí solo es insuficiente para llevar a cabo la regulación (Garsin et al., 2014).

Finalmente, la presencia de los genes *asa1*, *ace* y de síntesis de cápsula en la cepa F también han sido reportando para la cepa *E. faecalis* Symbioflor1 (probiótico), atribuyéndole a ellos sus características probióticas (Domann et al., 2007). De este análisis concluimos que la cepa de *E. faecalis* cuenta con una batería de genes relacionado a su capacidad de colonización y carece de aquellos que han sido reportados por tener un efecto negativo en el hospedero. Por su parte, dentro del genoma de la *E. faecium* D no se encontró ninguno de los genes relacionados a virulencia.

Además de los factores de virulencia, la presencia de genes involucrados en la producción de aminas biógenas es de gran relevancia en la BAL presentes en alimentos, ya que éstas podrían ser la causa de alergias e intoxicaciones en alimentos como queso y vino. En ambos genomas identificamos las enzimas ornitina carbamoiltransferasa, putrescina carbamoiltransferasa (agmatina desaminasa) y tirosina decarboxilasa, para la síntesis de ornitina, putrescina y tiramina, respectivamente. Esta última es de gran interés ya que se ha reportado que la intoxicación por alimentos conteniendo tiramina pueden provocar síntomas como: náuseas, hipertensión, dolor de cabeza y reacciones alérgicas (Gardini et al., 2001).

Con el fin de conocer la concentración de aminas biógenas y con ello el riesgo por consumo, se seleccionaron 8 quesos Cotija 4 de la colecta 2008 y 4 del 2012 (2 quesos de Michoacán y 2 quesos Jalisco) y se analizaron mediante la técnica descrita por la AOAC modificada y acoplada a HPCL (ver metodología). Como resultados se obtuvo que el máximo de concentración de tiramina era de 11.2 ppm y sólo se presentó en uno de los quesos analizados. En tres de ellos se obtuvo un promedio de 59.9 ppm de histamina, mientras que en otros dos se observó menos de 0.1 ppm de cadaverina y de putrescina. Cabe señalar que, en dos de los ocho casos, los quesos presentaban valores por debajo de 0.1 ppm para todas las aminas biógenas. Los resultados encontrados en este análisis concuerdan con lo reportado en otros estudios, por ejemplo, en el análisis de detección y cuantificación de aminas biógenas en el queso Feta se concluyó que las bajas concentraciones de estos compuestos podrían estar relacionadas a las altas concentraciones de NaCl presentes en el alimento y responsables de la inhibición de producción de las aminas (Wunderlichová et al., 2014). Este mismo fenómeno podría estar sucediendo en el queso Cotija, el cual se caracteriza por tener concentraciones de NaCl cercanas al 5%. Las concentraciones detectadas en este análisis nos dicen que no existe un riesgo de intoxicación debido al consumo de este alimento, ya se ha establecido que las concentraciones necesarias para provocar una intoxicación en personas susceptibles son de 50-100 mg en total.

8.2 Potencial funcional del género *Enterococcus* en el queso Cotija

8.2.1 Sistema proteolítico, lipolítico y desarrollo de aromas y sabores

El sistema proteolítico de las BAL convierte las proteínas de la matriz alimentaria en péptidos y estos en aminoácidos libres, los cuales son esenciales para el crecimiento de la bacteria y también contribuyen significativamente a las propiedades organolépticas finales del alimento (Smit et al., 2005). Debido a que la información acerca de las capacidades: proteolítica, lipolítica y de formación de aromas por parte del género *Enterococcus* es escasa, en este trabajo se realizó un análisis comparativo de proteínas reportadas en otras BAL, en particular las descritas para *Lactococcus lactis* (van Kranenburg et al. 2002 y Liu et al., 2010).

En cada genoma se identificaron un conjunto de genes codificantes para enzimas capaces de hidrolizar las moléculas de caseína, para transportadores de péptidos y para algunas peptidasas (Tabla 5). En la Figura 17 se presenta una reconstrucción hipotética del sistema proteolítico del género *Enterococcus* el cual consiste en un complejo Clp que podría hidrolizar la caseína en oligopéptidos, éstos serían transportados dentro de la célula bacteriana por medio del sistema Opp, y di-tri péptidos por transportadores tipo ABC (Dpp) y otros transportadores que actúan bajo fuerza motriz de protones (DptT). Una vez, dentro de la célula, diferentes peptidasas (Pep O, F, D, C, A, Q, T, B y S) que pueden hidrolizar los péptidos y generar aminoácidos libres quienes son los responsables de formar precursores de aromas y sabores. Es importante acotar que se corroboró que los transportadores de péptidos se localizan en la membrana celular basados en las predicciones de hélices transmembranales (Tmhmm).

Tabla 5. Genes involucrados en el sistema proteolítico de las cepas *E. faecium* D y *E. faecalis* F

Sistema proteolítico	Gen	Actividad ^a	<i>E. faecium</i> cepa D ^b	<i>E. faecalis</i> cepa F ^b
Complejo Clp (actividad caseinolítica)^{1,2}	Clp P	Subunidad proteolítica. Proteasa Clp ATP- dependiente	1	1
	Clp E	Proteasa Clp ATP- dependiente	1	1
	Clp B	Proteasa chaperona	2	2
	Clp L	Proteasa Clp ATP- dependiente	1	1
	Clp X	Proteasa Clp ATP- dependiente	1	1

Sistema de transporte de oligopéptidos ^{1,3}	Opp A	Proteína extracelular de unión a proteínas	1	3
	Opp B	Sistema de transporte de la membrana celular. Unión a proteínas	2	3
	Opp C	Permeasa. Transporte transmembranal	1	2
	Opp D	Transportador ABC/Transportador de oligopéptidos y dipéptidos	2	3
	Opp F	Transportador ABC/Transportador de oligopéptidos y dipéptidos	2	3
	Sistema de transporte de Di/tripeptidos ^{1,3}	Dpp A/P	Proteína de unión a ATP. Transportador periplásmico de dipéptidos	-
Dpp B		Permeasa. Sistema de transporte de dipéptidos	-	-
Dpp C		Permeasa. Sistema de transporte de dipéptidos	1	1
Dpp D		Proteína de unión a ATP. Transporte de dipéptidos (acoplamiento de energía)	-	-
DppF		Proteína de unión a ATP. Transporte de dipéptidos.	-	-
Dpp E		Proteína de unión a ATP. Transporte de dipéptidos.	1	9
DptT		Sistema de transporte de Di/tripeptidos. Fuerza protón motriz	1	2
Endopeptidasas ^{1,3}		Pep E/G	(X)m (X)n	-
	Pep O	(X)m (X)n	1	-
	Pep F	(X)m (X)n	1	1
Dipeptidasas ^{1,3}	Pep D	X X	-	-
	Pep V	X X	2	2
Aminopeptidasas ^{1,3}	Pep C	X(X) n	1	1
	Pep N	X(X) n	-	-
	Pep M	Met (X)n	-	-
	PepA	Glu/Asp (X)n	1	1
	Pcp	PyroGlu (X)n	-	-
Prolinepeptidasas ^{1,3}	Pep X	X-Pro (X)n	-	-
	Pep I	Pro X-(X)n	-	--
	Pep R	Pro X	-	-
	Pep L	Leu (X)n	-	-
	Pep P	X Pro (X)n	-	-
	Pep Q	X Pro	1	1
Otras ¹	PepT	Tripeptide aminopeptidase	1	2
	PepB	Oligoendopeptidase	1	1
	PepS	Aminopeptidase	1	1

^a X representa cualquier aminoácido

^b Representa el número de copias de cada gen, identificado en el análisis funcional.

¹ Sistema proteolítico identificado en este trabajo.

² Park et al., 2015.

³ Lui et al., 2010.

A continuación, se describirán los resultados de anotación funcional, ejemplificados en la Figura 17, los cuales revelaron que los enterococos del queso Cotija contienen un complejo con actividad caseinolítica denominado Clp. Se ha reportado que durante el proceso de maduración, las BAL se encuentran bajo condiciones de estrés, lo cual las lleva a la acumulación de proteínas mal plegadas y como propia respuesta se acciona el sistema proteolítico Clp, el cual ha sido reportado por tener un papel importante en el control de la calidad de las proteínas durante un proceso de estrés, a través de la hidrólisis de proteínas dañadas o mal plegadas, actuando como chaperonas en el proceso de replegamiento en una estructura terciaria adecuada. Bajo las condiciones que se encuentran en un queso, las bacterias podrían estar bajo un proceso de autólisis, lo cual podría liberar el complejo Clp fuera de la célula (Savijoki et al, 2006). El complejo Clp es dependiente de ATP e incluye proteínas encargadas de la hidrólisis (ClpP y ClpL), otras responsables del reconocimiento (ClpE y ClpX), desplegamiento (ClpB) y la translocación dentro de la estructura multimérica que conforma el complejo (Park et al., 2015; Frees et al., 2007). Se tiene la evidencia de que varias de estas proteasas de bacterias Gram-positivas son inducibles por condiciones de estrés y capaces de degradar caseína. Las cepas de *Enterococcus*, de este trabajo, albergan todas las unidades necesarias para que el complejo Clp pueda ser funcional, por lo que se propone que la hidrólisis de la caseína por este complejo podría ser el primer paso durante la maduración del queso.

Los oligopéptidos y péptidos generados por el complejo Clp son capturados e internalizados por transportadores adheridos a la célula, llamados sistemas Opp y Dpp y por otro, anclado a la pared celular denominado DptT, los cuales han sido ampliamente descritos en BAL. El sistema Opp pertenece a la superfamilia de transportadores ABC de unión a ATP y este lleva a cabo la captación de los oligopéptidos de 8 a 18 residuos de aminoácidos. En ambas genomas de enterococos el sistema Opp está completo, e incluye un gen codificante de una proteína de unión a oligopéptidos (OppA), dos proteínas de membrana (OppB y OppC) y dos

transportadores de proteínas dependientes de ATP (OppD y OppF) (Kunji et al., 1996; Savijoki et al., 2006). En ambos genomas se identificaron genes de otro sistema de transportadores homólogos del sistema Opp, llamado Dpp, del que únicamente se encontraron los genes codificando para las subunidades DppA, DppP, DppC y DppE. Estos tienen las funciones de unión de péptidos, permeasa y transportador dependiente de ATP, respectivamente. Sin embargo, se ha reportado que los sistemas Dpp y Opp pueden tener una actividad complementaria por lo que, aunque el sistema Dpp no está completo, puede verse complementado por el sistema Opp y viceversa (Doeven et al., 2005; Lamarque et al., 2004). Además, nuestras cepas contienen otro transportador de dipéptidos y tripéptidos, que trabaja bajo el efecto de fuerza motriz de protones (DptT) y presenta alta afinidad por tripéptidos, pero también es capaz de transportar dipéptidos y tetrapéptidos siempre y cuando contengan un aminoácido de cadena hidrofóbica (Savijoki et al., 2006) (Fig 17).

Una vez dentro de la célula, los péptidos son degradados en aminoácidos por peptidasas. Las endopeptidasas intracelulares PepC, PepF y PepO, han sido propuestas como las primeras enzimas en actuar. Los di/tripéptidos generados a continuación son degradados por tripeptidasas como la PepT y la dipeptidasa PepV. Finalmente, exopeptidasas como la PepA y PepB hidrolizan desde el extremo-N. Además, existen otras peptidasas de corte más específico como la PepQ y la PepS, donde la primera hidroliza dipéptidos con una prolina y la segunda tiene preferencia por péptidos que tienen de dos a cinco residuos de arginina o aminoácidos aromáticos en el N-terminal (Savijoki et al., 2006).

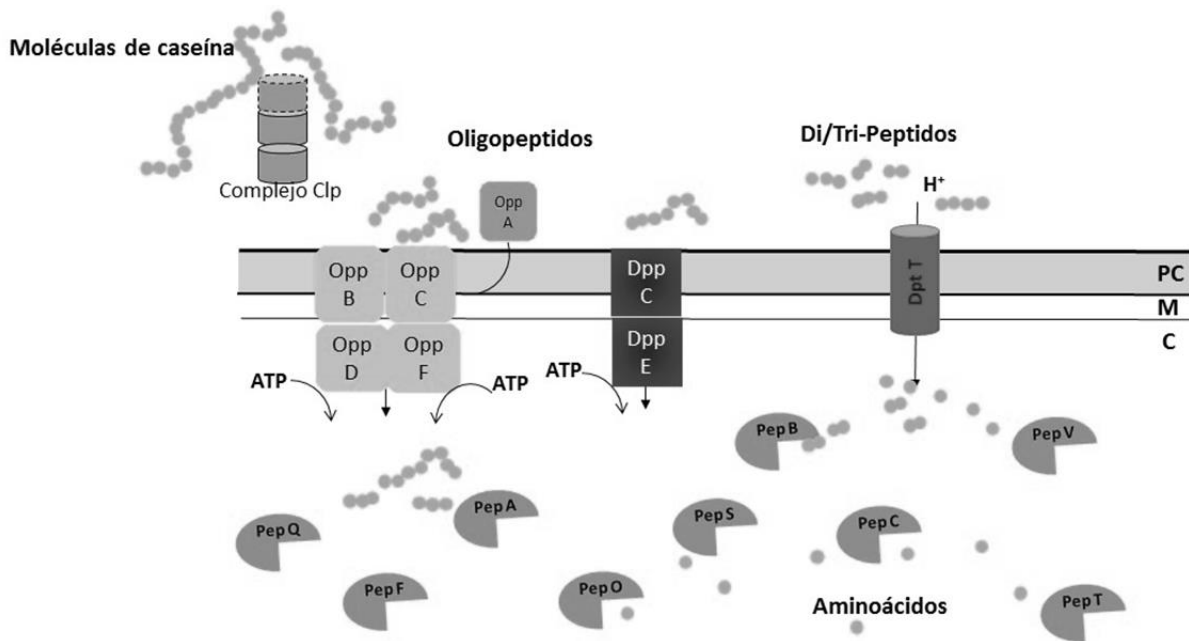


Fig 17. Propuesta de sistema proteolítico en las cepas de *Enterococcus* spp. aisladas de queso Cotija. PC: pared celular, M: membrana celular, C: citoplasma.

En 1970, Dovat y colaboradores observaron que el género *Enterococcus* presentaba baja actividad proteolítica, con excepción de *E. faecalis*. Después de algunos años, la primera proteinasa fue reportada en *E. faecalis* var. *Liquefaciens*. Esta enzima mostraba alta actividad contra caseína y baja actividad contra lactoglobulina y α -lactoalbúmina. En estudios donde se llevaron a cabo ensayo usando como sustrato la leche, *E. faecalis* demostró tener la mayor actividad proteolítica comparado con otras especies de enterococos, y se concluyó que esa actividad era dependiente de las concentraciones de NaCl y calcio presentes (Foulquié Moreno et al., 2003b). Otras investigaciones de enterococos aislados de quesos, por ejemplo, de Pecorino Abruzzese, han demostrado la importancia de la actividad aminopeptidasa del género; debido principalmente a que la producción de aminoácidos hidrofóbicos está relacionada con el efecto de eliminación de amargor y también está asociada con el desarrollo de sabores en el alimento (Serio et al., 2010).

Por otro lado, la lipólisis es una actividad importante en la maduración de los quesos, por lo tanto, también realizamos una búsqueda de genes codificantes para esterases en ambos genomas. En el genoma de *E. faecalis* cepa F, se identificaron cinco esterases putativas, dos acetilesterasas (una tributirina esterasa y una lipasa) y dos carboxilesterasas, mientras que en el genoma de *E. faecium* cepa D se encontraron únicamente dos acetilesterasas (una

tributirina esterasa y una lipasa) y una carboxilestera. Resultados experimentales preliminares han demostrado la capacidad lipolítica de ambas cepas en ensayos con agar tributirina (Alberto, 2016).

La lipólisis es un proceso que lleva a la formación de ácidos grasos libres, los cuales contribuyen al gusto y aromas finales de un producto alimentario. Durante los procesos de elaboración y maduración de quesos algunas lipasas endógenas de la leche son las responsables de la actividad lipolítica, así como enzimas pertenecientes a las bacterias iniciadoras, no iniciadoras y psicotróficas y enzimas exógenas (Sarantinopoulos et al., 2001). Otros grupos de trabajo han demostrado que algunas cepas de enterococos son capaces de hidrolizar triacilgliceroles, principalmente cepas aisladas de alimentos, dentro de las cuales, nuevamente *E. faecalis* mostró mayor actividad comparando con *E. durans* y *E. faecium*; y por el contrario *E. faecium* mostraba alta actividad esterolítica (Sarantinopoulos et al., 2001; Foulquié Moreno et al., 2006). Los ácidos grasos libres pueden ser transformados en metilcetonas, alcoholes secundarios, ésteres y lactonas, los cuales han sido reportados para contribuir al perfil sensorial de diversos productos lácteos (van Kranenburg et al., 2002).

A pesar de la existencia de pocos reportes, ha sido demostrado los enterococos son capaces de metabolizar el citrato y la lactosa con el fin de transformarlas en compuestos volátiles como son el acetaldehído, diacetilo, acetona, acetoína y etanol (Sarantinopoulos et al., 2002). En este estudio, pudimos determinar que ambas cepas presentan los genes que codifican para enzimas involucradas en la formación de compuestos volátiles: acetoína (EC 4.1.1.5), diacetilo (descarboxilación oxidativa) y acetaldehído (EC 1.2.1.2), los cuales contribuyen con aromas y sabores a mantequilla, yogurt y mantequilla rancia, así como acetaldehído y acetato (EC 2.7.2.1) quienes contribuyen con aromas pungentes, además de etanol (EC 1.1.1.1) el cual podría impartir notas frutales y de nuez a un queso madurado como es el queso Cheddar y el Roquefort (Smit et al., 2005; Engels et al., 1997).

La presencia de enterococos en altas cantidades es una característica común en diferentes quesos madurados, por ejemplo, en queso Feta (griego elaborado a partir de leche sin pasteurizar), Manchego, Mozzarella, Serra, Kefalotyru, Crebreiro y Comté (Sarantinopoulos et al., 2002), en donde este género es predominante incluso hasta el final de la maduración. Sin embargo, si la presencia de enterococos en un queso madurado contribuye al desarrollo de sabores y aromas buenos y aceptables o no, aún está en debate, ya que altas cantidades de

este microorganismo podría llevar al deterioro y desarrollo de características organolépticas no deseadas. No obstante, el papel benéfico de enterococos en la formación de aromas en quesos ha llevado a la inclusión de ciertas cepas como cultivos iniciadores, tal es el caso de *E. faecium* K77D, la cual ha sido considerada como aceptable para su uso en alimentos lácteos por la UK Advisory Committee on Novel Foods and Processes (ACNFP, 1996).

8.2.2 Producción de compuestos con actividad antibacteriana

Las bacteriocinas producidas por las BAL son péptidos antimicrobianos sintetizados ribosomalmente, son en su mayoría, catiónicos, anfifílicos, producidos por la mayoría de las bacterias, las cuales varían en el espectro y modo de actividad, estructura y masa molecular, termoestabilidad, actividad a diferentes pHs y determinantes genéticos (Moreno et al., 2002; de Vuyst y Vandamme, 1994). Previamente se identificó en *E. faecium* D, el gen que codifica para la enterocina A, la cual era activa contra *Listeria monocytogenes*.

El análisis del genoma permitió la identificación de 1 operón para la síntesis de la enterocina P en el genoma de *E. faecalis* F, así como 4 genes que codifican para las enterocinas A, B, P, SE-K4 y lactacina F (subunidad LacFX) en *E. faecium* D (Fig 17). Basados en esta clasificación de Franz y colaboradores (2007) (sección 2.2.4.1), las bacteriocinas identificadas en los genomas de las cepas del queso Cotija pertenecen principalmente a la clase II.1 (enterocina A, P, SE-K4), mientras que la enterocina B se agrupa en la clase II.3 (Fig 18). En ambas cepas, los genes que codifican para la bacteriocina están flanqueados por aquellos que codifican para proteínas de inmunidad, las cuales protegen a las cepas productoras de ser susceptibles a la bacteriocina. Es importante mencionar que las enterocinas pertenecientes a la clase II han sido principalmente identificadas en cepas aisladas de alimentos, por el contrario, enterocinas de la clase I (citolisina y bacteriocinas 32 y 41) frecuentemente se reportan en cepas aisladas de infecciones (Nes et al., 2014).

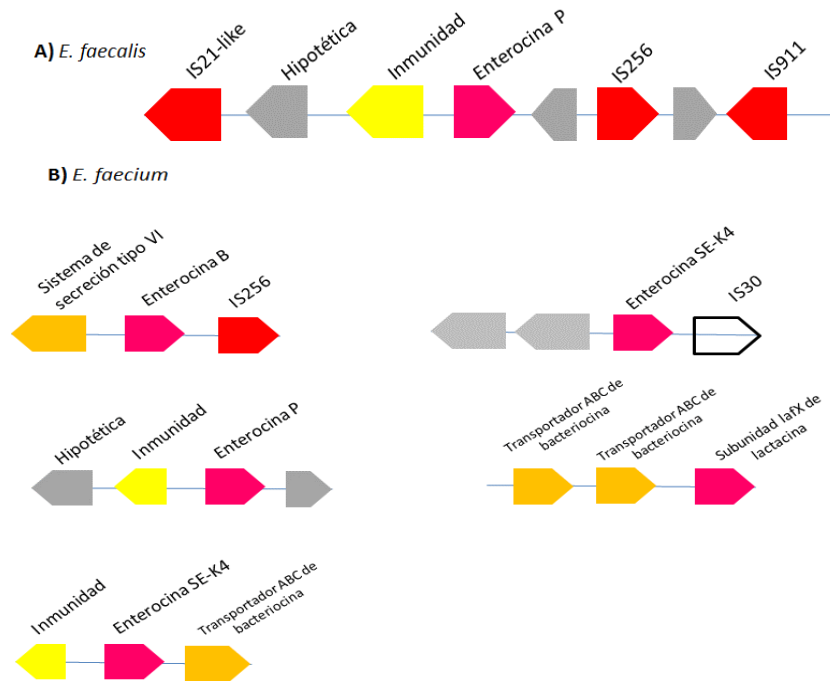


Fig 18. Genes que flanquean a los codificantes para bacteriocinas. A) Presentes en el genoma de *E. faecalis* F. B) Identificados en genoma de *E. faecium* D.

Además, ambas cepas presentan más de 10 genes que codifican para peptidoglucano hidrolasas (PGHs) principalmente con actividad de amidasa y glucosaminidasa con uno o varios dominios de LysM para reconocimiento de sustrato (Tabla 6). Previamente en el grupo de trabajo se reportó que la cepa de *E. faecalis* F, producía PGHs con pesos moleculares entre 60-90 kDa (Fig 19 B2), con actividad lítica en contra de bacterias como *E. coli*, *Salmonella sp*, *Pseudomonas aeruginosa*, *Streptococcus pyogenes*, *L. monocytogenes*, *S. aureus*, entre otras, lo que podría estar influenciando la calidad microbiológica del queso Cotija (García-Cano et al., 2014). Adicionalmente, se reportó una proteína de 74 kDa con actividad lítica producida por la cepa de *E. faecium* D (Fig 19 B1) (Olvera, 2013). Ambos experimentos se repitieron con el fin de corroborar la presencia de las bandas de actividad. Las PGHs se han reportado como enzimas con actividad lítica contra bacterias Gram-positivas y Gram-negativas (Callewaert et al., 2011). Después de la secuenciación de varios genomas bacterianos se ha descrito que una bacteria puede albergar entre 10 y 30 genes codificantes para alguna peptidoglucano hidrolasa y dicha redundancia se explica con el hecho de que este tipo de actividades son vitales para el crecimiento de la bacteria, ya que las PGHs están involucradas en el recambio de unidades de peptidoglucano para la síntesis de la pared celular, en la autólisis celular, división celular y en el anclaje de algunas estructura como puede

ser el pili o flagelo (Mesnage et al., 2008). Como se mencionó previamente, experimentalmente sólo se visualizaron 1 y 3 bandas con actividad lítica en zimografía contra *M. lysodeikticus*, lo cual se puede deber a que la producción de PGH depende de la fase de crecimiento en la que se encuentre la bacteria, además de que no todas las enzimas son igualmente activas contra este microorganismo (Lepeuple et al., 1998). Además, en nuestra experiencia, después de un tratamiento con ácido tricloroacético, no todas las enzimas pueden renaturalizarse de manera eficiente y por lo tanto no son activas.

Tabla 6. Genes con anotación funcional de PGH en las cepas de estudio

<i>E. faecium</i> D	<i>E. faecalis</i> F
(6) endo- β -acetilglucosaminidasas	(5) N-Acilmuramoil-L-alanina amidasa
(6) N-acilmuramoil-L-alanina amidasas	(3) endo- β -acetilglucosaminidasas+ dominio lysM
(3) amidasas con regiones NlpC/p60	(3) amidasas
NlpC/p60+amidasa+ dominio endo- β -acetilglucosaminidasa	(2) 72 y 54 kDa reportadas previamente

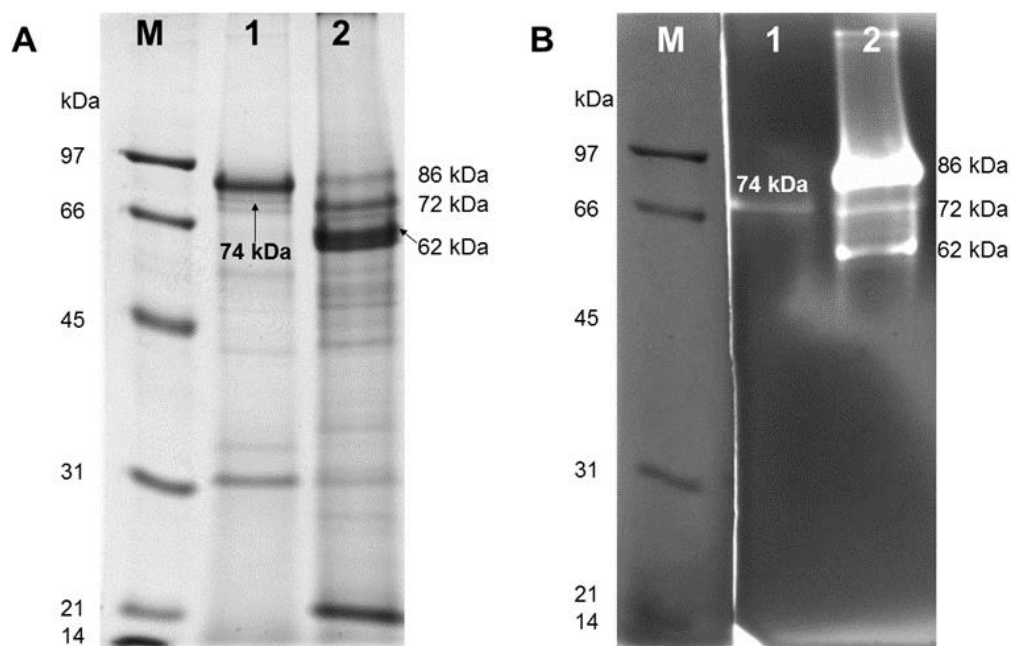


Fig 19. A) SDS-PAGE de extractos semipurificados de una fermentación por precipitación con TCA. M: Marcador de peso molecular, 1: *E. faecium* D y 2: *E. faecalis* F. B) Zimografía contra *M. lysodeikticus*

8.3 Propuesta de genes marcadores de adaptación a alimentos

El análisis de ortología nos permitió identificar un conjunto de genes únicos de las cepas F y D, los cuales podrían ayudarnos a generar biomarcadores de cepas asociadas a alimentos (Tabla 7). En color naranja se muestran algunos de los genes de *E. faecalis* que han sido planteados como marcadores de adaptación al tracto gastrointestinal, así como genes únicos de la cepa *E. faecium* NRRL B-2354, también identificados en el conjunto de genes únicos de las cepas analizadas en este trabajo (Weigand et al., 2014; Kopit et al., 2014).

Tabla 7. Rutas metabólicas asociadas a los genes únicos (KEEG)

<i>E. faecalis</i> cepa F	<i>E. faecium</i> cepa D	<i>E. faecium</i> NRRL-B2354 (Kopit L. et al., 2014)
338 genes únicos	531 genes únicos	45 genes únicos
Biosíntesis de amino ácidos Vía de las pentosas fosfato Metabolismo de arginina y prolina Biosíntesis de terpenoides (esqueleto) Exportación de proteínas Reparación de ADN Enzimas de restricción Transportadores ABC Sistema fosfotransferasa de azúcares (PTS) Sistemas de dos componentes Bacteriocinas y PGHs	Metabolismo de carbono Glicolisis/gluconeogénesis Biosíntesis de amino ácidos Metabolismo de fructosa y manosa Metabolismo de galactosa Metabolismo de amino azúcares Transportadores ABC Enzimas de restricción Sistemas fosfotransferasa de azúcares (PTS) Sistemas de dos componentes Bacteriocinas	20 proteínas hipotéticas 8 proteínas de fagos 4 reparación de ADN 2 glicosiltransferasa

En otros estudios, se ha comprobado que la adaptación de algunas cepas de *Lactobacillus* spp., *Lactococcus* spp. y *Streptococcus thermophilus*, a alimentos lácteos y al tracto gastrointestinal, se basa en la presencia de genes codificantes para actividades relacionadas a su capacidad de sobrevivir bajo las condiciones de ambos ambientes, así como el fenómeno de multicopia de genes codificantes para la utilización de los sustratos disponibles en ambos nichos ecológicos. Las cepas que están completamente adaptadas al tracto gastrointestinal se caracterizan por tener genes que codifican para actividades asociadas a la capacidad de hidrolizar las sales biliares, así como de permeasas para la degradación de ciertos azúcares, por ejemplo, para el uso metabólico de manosa. Por otro lado, las cepas adaptadas a alimentos lácteos presentan ciertos genes como proteasas y enzimas de restricción tipo I y III,

las cuales ayudan a la bacteria a persistir en presencia de la microbiota acompañante (O'Sullivan et al., 2009).

Además, se ha propuesto que cepas de *Enterococcus* aisladas de ambientes nosocomiales se caracterizan por presentar tipos específicos de bacteriocinas. Las bacteriocinas Bac41, Bac32, citolisina L y citolisina S, se encuentran con mayor frecuencia en cepas nosocomiales, mientras que las enterocinas A, B, P, Q y L50 son características de cepas aisladas de alimentos y de heces de animales y humanos sanos. Estos hallazgos sugieren que algunos tipos de bacteriocinas son dominantes en enterococos de origen clínico las cuales difieren de las predominantes en los enterococos aislados de alimentos fermentados y cepas comensales, y que el tipo de bacteriocina presente en una cepa está influenciado por factores nutricionales y microbianos. Por lo tanto, se sugiere que las enterocinas podrían proveer una ventaja de competencia y promover la predominancia y colonización de ciertas cepas a diferentes ambientes, tales como el tracto gastrointestinal (Nes et al., 2014).

Weigand y colaboradores (2014) obtuvieron el genoma de varias cepas de *E. faecalis* aislados de diferentes ambientes, principalmente de aguas residuales. El análisis de los genomas reveló que existen genes que codifican para transportadores de azúcares y de metales (Cu o Mn), característicos de cada cepa, lo cual se plantea está ampliamente relacionado con su adaptación al tracto gastrointestinal o su sobrevivencia en un ambiente acuático, respectivamente. Además, proponen al operón de captación de níquel como un biomarcador de cepas pertenecientes a un ambiente no entéricos, mientras que genes que codifican para enzimas encargadas de la utilización de azúcares (xilosa) son propuestos como indicadores de cepas de origen gastrointestinal.

Con todo lo anterior y basados en el análisis de los genes únicos identificados en ambos genomas se propuso a los genes codificantes para enterocinas de la clase II.1 y enzimas de restricción como biomarcadores de adaptación a un alimento lácteo (queso Cotija).

Los genes seleccionados fueron analizados, por PCR punto final, en las cepas de enterococos empleadas para describir la estructura poblacional por MLST (Anexo I). En los resultados se observó que el gen que codifica para la enterocina A está presente en todas las cepas de *E. faecium* y en 3 cepas de *E. faecalis* y está ausente en las cepas patógenas V583 y ATCC 29212 (Fig 20). La presencia del gen que codifica para la enterocina P únicamente se encontró

en una cepa de *E. faecalis* (QA1) y la enterocina SE-K4 sólo está presente una cepa de *E. faecium* (QD2). Por otro lado, el análisis de la presencia de las enzimas de restricción fue positivo sólo en una cepa de cada especie por lo que se descartan como posibles biomarcadores de adaptación dada su baja frecuencia de detección. Finalmente, se propone a las enterocinas de la clase II.1 como biomarcadores de adaptación de cepas a los alimentos fermentados.

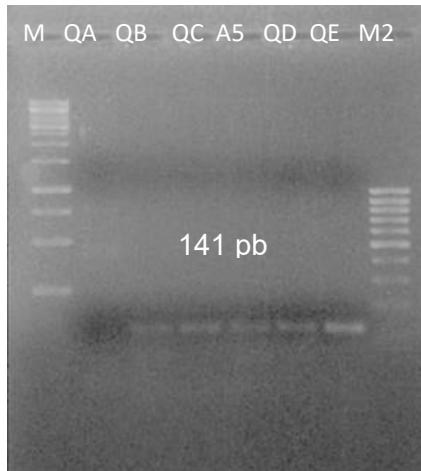


Fig 20. Identificación de amplicón correspondiente a la enterocina A en cepas aisladas del queso Cotija. M: marcador 1kb (Thermo Scientific), QA1, QB2, QC4, A5-1: *E. faecalis*, QD2 y QE2: *E. faecium*, M2: marcador low range (Thermo Scientific). Gel de agarosa al 1%.

Los CRISPR-cas son un sistema de defensa procariótico de secuencia específica que proporciona un tipo de inmunidad adquirida. En general, funcionan integrando un segmento pequeño de un elemento móvil invasor dentro del arreglo del CRISPR, entre regiones palíndromas repetidas y finalmente son degradadas por las enzimas nucleasas Cas (Palmer et al., 2010). Se ha reportado que la presencia de regiones CRISPR asociadas a nucleasas Cas, podrían ser un diferenciador entre cepas de origen nosocomial, ya que éstas últimas con el paso del tiempo se ha ido deshaciendo de los genes que codifican para la actividad de nucleasas (Fig 21A). Con base en lo anterior, Palmer y colaboradores (2010), han propuesto que hay tres tipos de regiones CRISPR en la especie *E. faecalis*. El tipo 1 es característico de cepas aisladas en años anteriores a la aparición de la cepa multirresistente a antibióticos *E. faecalis* V853 en el año de 1987, estas secuencias repetidas se encuentran localizadas entre los genes Ef0672 y Ef0673. A partir de la aparición de la cepa patógena se desarrollan los tipos CRISPR 2 y 3, siendo el tipo 2 característico en cepas causantes de infecciones, mientras que el tipo 3 se ha encontrado en cepas distintas a las nosocomiales. La principal

diferencia entre ellas es la presencia de genes codificantes para nucleasas cas 1 y cas 2. Como se observa en la Figura 21B, en el genoma de la cepa F se pudo reconstruir la región de CRISPR que la clasifica dentro del tipo 3. Sin embargo, hace falta más información de este tipo de regiones en cepas aisladas de alimentos que ayuden a sostener la hipótesis de que esta clasificación podría discernir entre cepas nosocomiales y las que no lo son.

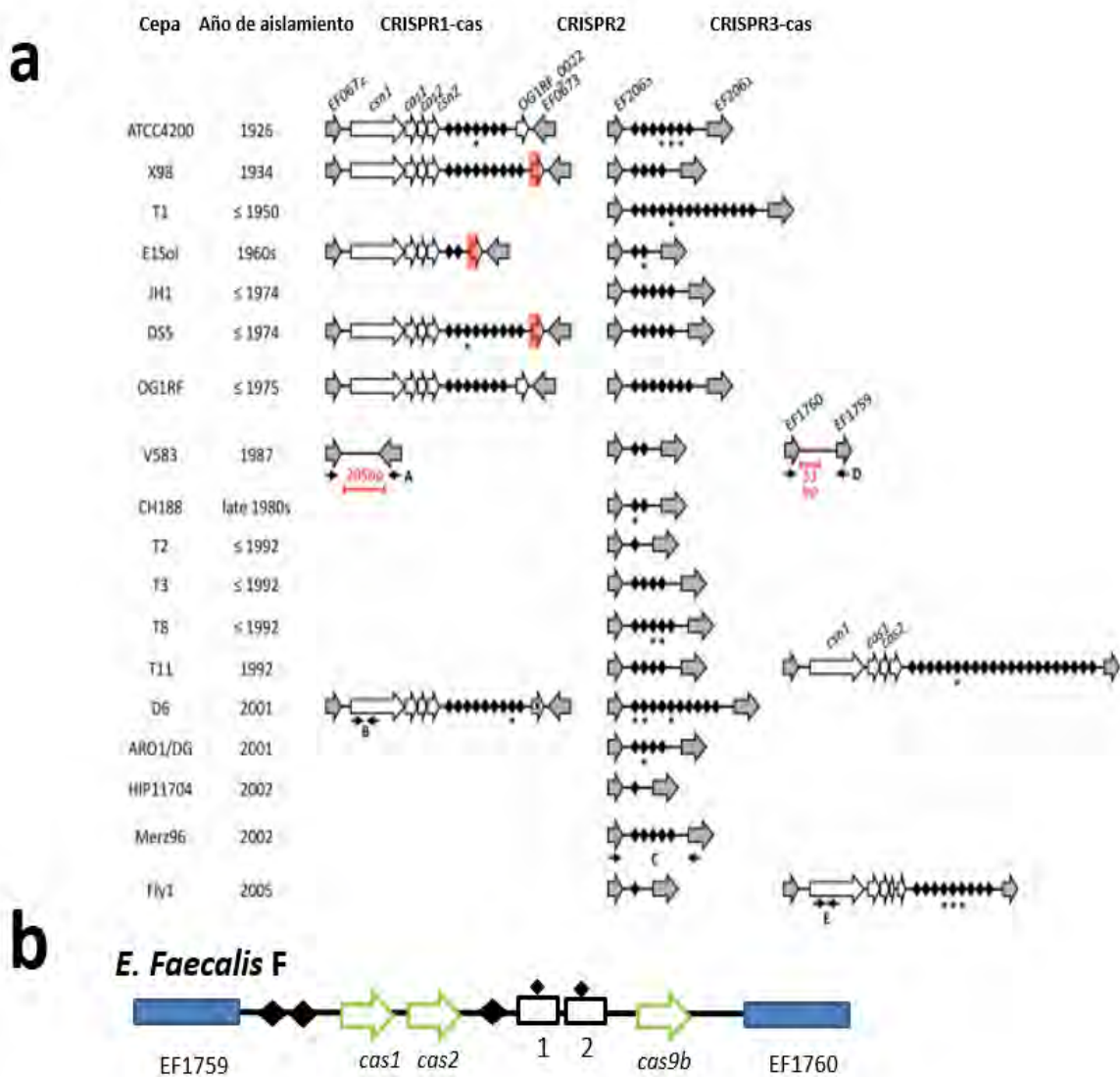


Fig 21. Análisis de regiones CRISPR y CRISPR-cas. A) Homólogos de *E. faecalis* V583 se muestran en flechas negras, mientras que los genes codificantes a nucleasas cas se representan como flechas blancas. Las regiones rojas indican las deleciones características de la cepa patógena. CRISPR1-cas característico de cepa anteriores a la cepa patógena. CRISPR2 presente en cepas nosocomiales.

CRISPR3-cas diferentes a las patógenas y de reciente aparición. Modificada de Palmer y colaboradores (2010). B) CRISPR3-cas encontrada en la cepa F.

8.3.1 Estrategia de evaluación de biomarcadores usando el ensamble de metagenoma

A partir del ensamble obtenido del metagenoma bacteriano del queso Cotija (Escobar-Zepeda, 2016) se realizó la reconstrucción de genomas de *E. faecalis* y *E. faecium* tomando como referencia los genomas secuenciados en el presente trabajo y con la ayuda del software ABACAS el cual busca bloques de sintenia entre la referencia y el ensamble dado.

Como resultado de la comparación y búsqueda de los genes únicos de las cepas D y F en los ensamblajes reconstruidos pudimos encontrar sólo 188 de 338 de los genes únicos de *E. faecalis* F y 376 de los 531 genes únicos de *E. faecium* D con un 99% de identidad y 95-100% de cobertura. En la Tabla 8 se muestra el resultado del análisis de la anotación funcional que correspondía a dichos genes. Con los estudios previamente descritos de genes de adaptación de alimentos lácteos y tracto gastrointestinal podemos inferir que las cepas presentes en el queso Cotija podrían tener un origen gastrointestinal al presentar genes característicos de adaptación a ese ambiente, de acuerdo con lo establecido por Weigand y colaboradores (2014). Además, se corrobora la presencia de las enterocinas de la clase II.1 como comunes en cepas aisladas de alimentos.

Tabla 8. Anotación de los genes únicos encontrados en los ensamblajes generados a partir del metagenoma

Genes de <i>E. faecalis</i>	Genes de <i>E. faecium</i>
Elementos de inserción	Transposasas
Transposasas	Elementos DUF
Enzimas de restricción tipo I y III	Enzimas de restricción tipo III
Enterocina P	Peptidoglucano hidrolasas
Reguladores transcripcionales	Genes del sistema fosfotranferasa de azúcares (PTS)
Transportadores de metales (Cu, Zn y Mn)	Permeasa de lactosa, manosa y celobiosa
Peptidoglucano hidrolasas	Reguladores transcripcionales
Genes del sistema fosfotranferasa de azúcares (PTS)	
Elementos DUF	

9. CONCLUSIONES

- Los enterococos (*E. faecalis* y *E. faecium*) del queso Cotija (diferentes colectas) no pertenecen a los complejos clonales de alto riesgo en salud pública, descritos para ambas especies, y se plantea que su origen sea relacionado a animales y/o humanos sanos.
- A través del análisis de genómica comparativa se logró establecer evidencias genéticas que diferencian a *E. faecalis* y *E. faecium* aislados del queso Cotija con respecto de aislados nosocomiales. Además, se demostró que dichas cepas que tienen gran similitud con cepas utilizadas como probióticos o cultivos iniciadores, dando un indicio de su origen.
- Las cepas de enterococos aisladas del queso Cotija se caracterizan por tener baja incidencia de genes relacionados con factores de virulencia, así como de genes que confieran resistencia a antibióticos. Sin embargo, se caracterizan por albergar genes relacionados a la colonización y proliferación en el tracto gastrointestinal, lo que habla de su habilidad para sobrevivir a dicho ambiente.
- La anotación funcional de los genomas permitió establecer el papel potencial de especies de *Enterococcus* en la producción el queso Cotija, a través de probables sistemas proteolítico (Complejo Clp, transportadores de péptidos y peptidasas), lipolítico (esterasas y lipasas) y de enzimas relacionadas la producción de metabolitos implicados en el desarrollo de aromas y sabores.
- Las bacteriocinas de la clase II y las regiones CRISPR-cas (Tipo3) podrían funcionar como biomarcadores de adaptación de enterococos a los alimentos, ya que se pudieron identificar en otras cepas aisladas de quesos artesanales.

10. PERSPECTIVAS

- Secuenciar más cepas de ambas especies aisladas de otros alimentos fermentados para ampliar el análisis de genómica comparativa, el cual ayude a fortalecer las conclusiones planteadas en este proyecto.
- Evaluar la presencia de los genes propuestos como biomarcadores de adaptación a alimentos (enterocinas tipo II y CRISPR-cas tipo3) en cepas aisladas de otros alimentos fermentados.
- Realizar estudios de transcriptómica en condiciones similares en las que se encuentran los enterococos en el queso Cotija, que nos permitan dilucidar si los genes propuestos como marcadores (enterocinas, CRISPR-cas, los sistemas proteolíticos y lipolíticos) se expresan y en qué niveles.

11. BIBLIOGRAFIA

- Abriouel, H., Ben, O.N., Cobo, A.M., Lucas, R.L., Grande, M.J., Martínez-Viedma, P., Ortega, E., Martínez-Cañamero, M., Gálvez, A. 2008. Comparative analysis of genetic diversity and incidence of virulence factors and antibiotic resistance among enterococcal populations from raw fruit and vegetable foods, water and soil, and clinical samples. *International Journal of Food Microbiology* 123, 38-49.
- Alberto Ríos Claudia, 2016. Potencial biotecnológico de cepas de *Enterococcus* spp. aisladas de queso Cotija. Tesis de licenciatura. Facultad de Química, UNAM. Ciudad de México, México.
- Alikhan, N.F., Petty, N.K., Ben Zakour, N.L., Beatson, S.A., 2011. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics*, 12:402.
- Álvarez, R., Barragán E., Chombo P., 2005. Reglas de Uso Marca Colectiva Queso Cotija Región de Origen. Colmich/ Ciatej. México.
- Andrews, S., 2015. Babraham Bioinformatics-FastQC a Quality Control Tool for High Throughput Sequence Data. URL. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (accessed 4.13.16.) AOAC. 2012. Official methods of analysis of Association of Official Analytical Chemists. Official Method 977.13. Colorimetric Method, Chapter 35. In: G. W. Latimer, (Ed.) Robert L. Bradley Jr, 19 ed. AOAC Inc., Arlington, pp. 17-18.
- Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., Meyer, F., Olsen, G.J., Olson, R., Osterman, A.L., Overbeek, R.A., McNeil, L.K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A., Zagnitko, O., 2008. The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 9,75.
- Besemer, J., Lomsadze, A., Borodovsky, M., 2001. Prokaryotic GeneMark.hmm version 2 GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Research* 29, 2607-2618.
- Boetzer, M., Henkel, C.V., Jansen, H.J., Butler, D., Pirovano, W. 2011. Scaffolding pre-assembled contigs using SSPACE. 15;27(4), 578-9.
- Bonacina, J., Suárez, N., Hormigo, R., Fadda, S., Lechner, M., Saavedra, L., 2016. A genomic view of food-related and probiotic *Enterococcus* strains. *DNA research* 24, 11-24.
- Bourgogne, A. y Garsin, D., 2008. Large scale variation in *Enterococcus faecalis* illustrated by the genome analysis of strain OG1RF. *Genome Biology* 9:R110.
- Brown, C.T., Howe, A., Zhang, Q., Pyrkosz, A. B., Brom, T.H., 2012. A Reference-Free Algorithm for Computational Normalization of Shotgun Sequencing Data. <http://arxiv.org/abs/1203.4802> (accessed 4.13.16).
- Burgos, M.J., López R.L., Abriouel, H., Omar N.B., Galvez A., 2009. Multilocus sequence typing of *Enterococcus faecalis* from vegetable foods reveals two new sequence types. *Foodborne Pathogens and Disease* 6, 321-327.

Callewaert, L., Walmagh, M., Michiels, C.W., Lavigne, R., 2011. Food applications of bacterial cell wall hydrolases. *Current Opinion in Biotechnology* 22,164–171.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 15; 10-421.

Carattoli, A., Zankari, E., Garcia-Fernandez, A., Volby Larsen, M., Lund, O., Villa, L., Aarestrup, F.M., Hasman, H., 2014. PlasmidFinder and pMLST: in silico detection and typing of plasmids. *Antimicrobial Agents Chemotherapy* 58, 3895-3903.

Cercenado, E., 2011. *Enterococcus*: resistencias fenotípicas y genotípicas y epidemiología en España. *Enfermedades Infecciosas y Microbiología Clínica*, 29, 59–65.

Cervantes-Escoto, F., Villegas-de Gante, A., Cesin-Vargas, A., Espinoza- Ortega, A., 2008. Los Quesos Mexicanos Genuinos –Patrimonio cultural que debe rescatarse-. 1ª edición. Mundi-Prensa México. pp. 149-162.

Chain, P.S., Grafham, D.V., Fulton, R.S., Fitzgerald, M.G., Hostetler, J., Muzny, D., Ali, J., Birren, B., Bruce, D.C., Buhay, C., Cole, J.R., Ding, Y., Dugan, S., Field, D., Garrity, G.M., Gibbs, R., Graves, T., Han, C.S., Harrison, S.H., Highlander, S., Hugenholtz, P., Khouri, H.M., Kodira, C.D., Kolker, E., Kyripides, N.C., Lang, D., Lapidus, A., Malfatti, S.A., Markowitz, V., Metha, T., Nelson, K.E., Parkhill, J., Pitluck, S., Qin, X., Read, T.D., Schmutz, J., Sozhamannan, S., Sterk, P., Strausberg, R.L., Sutton, G., Thomson, N.R., Tiedje, J.M., Weinstock, G., Wollam, A., Genomic Standards Consortium Human Microbiome Project Jumpstart Consortium, Detter, J.C., 2009. Genomics. Genome project standards in a new era of sequencing. *Science* 326(5950), 236-237.

Contreras-Moreira, B., Vinuesa, P., 2013. GET_HOMOLOGUES, a Versatile Software Package for Scalable and Robust Microbial Pangenome Analysis. *Applied and Environmental Microbiology* 79, 7696 –7701.

Corpet, F., 1988. Multiple sequence alignment with hierarchical clustering. *Nucleic acids research* 16(22), 10881-10890.

Dag Anders, B. y Lars, G., 2011. Complete Genome Sequence of the Commensal *Enterococcus faecalis* 62, Isolated from a Healthy Norwegian Infant. *Journal of Bacteriology* 193, 2377-2378.

Darling, A.E., Mau, B., Perna, N.T., 2010. progressiveMauve: Multiple Genome Alignment with Gene Gain, Loss and Rearrangement. *PLoS ONE* 5(6), e11147.

de Been M, Pinholt M, Top J, Bletz S, Mellmann A, van Schaik W, Brouwer E, Rogers M, Kraat Y, Bonten M, Corander J, Westh H, Harmsen D, Willems RJL., 2015. Core genome multilocus sequence typing scheme for high-resolution typing of *Enterococcus faecium*. *J Clin Microbiol* 53, 3788 –3797.

de Vuyst, L. y Vandamme, E.J., 1994. Antimicrobial potential of lactic acid bacteria. In: De Vuyst, L., Vandamme, E.J. (Eds.), *Bacteriocins of Lactic Acid Bacteria: microbiology, genetics and applications*. Blackie Academic & Professional, London, United Kingdom, pp. 91–142.

Darling, A.E., Mau, B., Perna, N.T., 2010. Progressive Mauve: Multiple Genome Alignment with Gene Gain, Loss and Rearrangement. PLoS ONE 5, e11147.

Doeven, M.K., Kok, J., Poolman B., 2005. Specificity and selectivity determinants of peptide transport in *Lactococcus lactis* and other microorganisms. Molecular Microbiology 57(3), 640–649.

Domann, E., Hain, T., Ghai, R., Billion, A., Kuenne, C., Zimmermann, K., Chakraborty, T., 2007. Comparative genomic analysis for the presence of potential enterococcal virulence factors in the probiotic *Enterococcus faecalis* strain Symbioflor 1. International Journal of Medical Microbiology, 297, 533-539.

Dovat, A. M., Reinbold, G. W., Hammond, E. G., Vedamuthu, E. R., 1970. Lipolytic and proteolytic activity of enterococci and lactic acid group streptococci isolated from young Cheddar cheese. Journal of Milk and Food Technology 9, 365-423.

Eaton, T.J. y Gasson, M.J., 2001. Molecular screening of *Enterococcus* virulence determinants and potential for genetic exchange between food and medical isolates. Applied and Environmental Microbiology 67, 1628-1635.

Eddy, S.R., 2011. Accelerated profile HMM searches. PLoS Computational Biology 7, e1002195

Ecker, C., Lecerf, M., Dubost, L., Arthur, M., Mesnage, S., 2006. Funcional Analysis of AtlaA, the Major N-Acetylglucosaminidase of *Enterococcus faecalis*. Journal of Bacteriology. 188, 8513-8519

Engels, W. J. M., Dekker, R., de Jong C., Neeter R., Visser, S., 1997. Comparative Study of Volatile Compounds in the Water-soluble Fraction of Various Types of Ripened Cheese. International Dairy Journal 7, 255-263

Escobar-Zepeda, A., Sanchez-Flores, A., Quirasco, B., 2016. Metagenomic analysis of a Mexican ripened cheese reveals a unique complex microbiota. Food Microbiology 57, 116-127.

Foulquié, M.R.M., Callewaert, R., Devreese, B., Beeumen J.V. y De Vuyst, L. 2003a. Isolation and biochemical characterization of enterocins produced by enterococci from different sources. Journal of Applied Microbiology. 94, 214-229.

Foulquié, M.R.M., Rea, M.C., Cogan, T.M., De Vuyst, L., 2003b. Applicability of a bacteriocin-producing *Enterococcus faecium* as a co-culture in Cheddar cheese manufacture. International Journal of Food Microbiology 81, 73– 84

Foulquié, M.R.M., Sarantinopoulos, P., Tsakalidou, E., De Vuyst L., 2006. The role and application of enterococci in food and health. International Journal of Food Microbiology 106, 1 – 24.

Franz, C.M., Muscholl-Silberhorn, A.B., Yousif, N.M.K., Vancanneyt, M., Swings, J., Holzapfel W.H., 2001. Incidence of Virulence Factors and Antibiotic Resistance among *Enterococci* Isolated from Food. Applied Environmental Microbiology 67, 4385–4389.

Franz, C.M., van Belkum, M.J., Holzapfel, W.H., Abriouel, H., Gálvez, A., 2007. Diversity of enterococcal bacteriocins and their grouping in a new classification scheme. *FEMS Microbiology Reviews* 31, 293–310.

Franz, C.M., Huch, M., Abriouel, H., Holzapfel, W., Gálvez, A., 2011. Enterococci as probiotics and their implications in food safety. *International Journal of Food Microbiology* 151, 125–40.

Frees, D., Savijoki, K., Varmanen, P., Ingmer, H., 2007. Clp ATPases and ClpP proteolytic complexes regulate vital biological processes in low GC, Gram-positive bacteria. *Molecular Microbiology* 63(5), 1285–1295.

Freitas, A. R., Novais, C., Ruiz-Garbajosa, P., Coque, T. M., & Peixe, L., 2009. Dispersion of Multidrug-Resistant *Enterococcus faecium* Isolates Belonging to Major Clonal Complexes in Different Portuguese Settings. *Applied and Environmental Microbiology*, 75(14), 4904–4908.

Fritzenwanker, M. y Kuenne, C., 2013. Complete Genome Sequence of the Probiotic *Enterococcus faecalis* Symbioflor 1 Clone DSM 16431. *Genome Announcements* 1, Issue 1 e00165-12.

Garg, S.K. y Mittal, B.K., 1991. Enterococci in milk and milk products. *Critical Reviews in Microbiology* 18:1, 15-45.

Garsin, D.A., Frank, K.L., Silanpää, J., Ausubel, F.M., Hartke, A., Shankar, N., Murray, B.E., 2014. Pathogenesis and Models of Enterococcal Infection. *Enterococci From Commensals to Leading Causes of Drug Resistant Infection*. Massachusetts Eye and Ear Infirmary Boston pp 204-206.

Gardini, F., Martuscelli, M., Caruso, M.C., Galgano, F., Crudele, M.A., Favati, F., Guerzoni, M.E., Suzzi G., 2001. Effects of pH, temperature and NaCl concentration on the growth kinetics, proteolytic activity and biogenic amine production of *Enterococcus faecalis*. *International Journal of Food Microbiology* 64, 105–117.

García-Cano, I., Serrano-Maldonado, C.E., Olvera-García, M., Delgado Arciniega, E., Peña-Montes, C., Quirasco, M., 2014. Antibacterial activity produced by *Enterococcus* spp. isolated from artisanal Mexican dairy product, Cotija cheese. *LWT Food Science and Technology* 59, 26-34.

Gilmore, M., 2002. *The Enterococci: Pathogenesis, Molecular Biology and Antibiotic Resistance*. Washington, DC: American Society for Microbiology.

Giraffa, G., 2002. *Enterococci* from foods. *FEMS Microbiology Reviews* 26, 163-171.

Giraffa, G., 2003. Funcionalidad de enterococci in dairy products. *International Journal of Food Microbiology* 88, 215-222.

Grissa, I., Vergnaud, G., Pourcel, C., 2007. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Research* 35, W52-W57.

Guzmán Ovalle Andrés, 2015. Estudio genético poblacional de cepas de *Enterococcus faecium* aisladas del queso Cotija artesanal madurado. Tesis de licenciatura. Facultad de Química. UNAM. Ciudad de México, México.

- Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41,95-98.
- Hentschel, U., Steinert, M, Hacker, J. 2005. Common molecular mechanisms of symbiosis and pathogenesis. Trends in Microbiology 8, 226-230.
- Hugenholtz, J., 1993. Citrate metabolism in lactic acid bacteria. FEMS Microbiology Reviews 12, 165 – 178.
- Hunt, M., Kikuchi, T., Sanders, M., Newbold, C., Berriman, M., y Otto, T.D., 2013. REAPR: a universal tool for genome assembly evaluation. Genome Biology 14: R47.
- Jiménez, E., Ladero, V., Chico, I., Maldonado-Barragán, A., López, M., Martín, V., Fernández, L., Fernández, M., Álvarez, MA., Torres, C., y Rodríguez J.M., 2013. Antibiotic resistance, virulence determinants and production of biogenic amines among enterococci from ovine, feline, canine, porcine and human milk. BMC Microbiology 13-288.
- Kim, E.B., Marco, M.L., 2014. Nonclinical and clinical *Enterococcus faecium* strains, but not *Enterococcus faecalis* strains, have distinct structural and functional genomic features. Applied Environmental Microbiology 80,154-165.
- Kopit, L.M., Kim, E.B., Siezen, R.J., Harris, L.J., Marco, M.L., 2014. Safety of the surrogate microorganism *Enterococcus faecium* NRRL B-2354 for use in thermal process validation. Applied Environmental Microbiology 80, 1899-1909.
- Kunji, E.R., Mierau, I., Hagting, A., Poolman, B., Konings, W.N., 1996. The proteolytic systems of lactic acid bacteria. Antonie van Leeuwenhoek 70,187-221.
- Lam MM, Seemann T, Bulach DM, Gladman SL, Chen H, Haring V, Moore RJ, Ballard S, Grayson ML, Johnson PD, Howden BP, Stinear TP. 2012. Comparative analysis of the first complete *Enterococcus faecium* genome. Journal of Bacteriology 194, 2334-2341.
- Lam, M.M., Seeman, T., Tobias, N.J., Chen, H., Haring, V., Moore, R.J., Ballard, S., Grayson, L.M., Johnson, P.D., Howden, B.P., Stinear, T.P., 2013. Comparative analysis of the complete genome of an epidemic hospital sequence type 203 clone of vancomycin-resistant *Enterococcus faecium*. BMC Genomics 2013, 14-595.
- Lamarque, M., Charbonnel, P., Aubel, D., Piard, J.C., Atlan, D., Juillard, V., 2004. A multifunction ABC transporter (Opt) contributes to diversity of peptide uptake specificity within the genus *Lactococcus*. Journal of Bacteriology 186, 6492–6500.
- Leavis, H.L., Bonten, M.J.M., Willems, R.J.L., 2006. Identification of high-risk enterococcal clonal complexes: global dispersion and antibiotic resistance. Current Opinion in Microbiology 9, 454-460.
- Lepeuple, A.S., van Gemert, E., Chapot-Chartier, M.P.,1998. Analysis of the Bacteriolytic Enzymes of the Autolytic *Lactococcus lactis* subsp. cremoris Strain AM2 by Renaturing Polyacrylamide Gel Electrophoresis: Identification of a Prophage-Encoded Enzyme. Applied and Environmental Microbiology 64,4142–4148.

Loman, N.J., Constantinidou, C., Chan, J.Z.M., Halachev, M., Sergeant, M., Penn, C.W., Robinson, E. R., and Pallen, J.M., 2012. High-throughput bacterial genome sequencing: an embarrassment of choice, a world of opportunity. *Nature Reviews Microbiology* 10, 599-606.

Liu, M., Bayjanov, J.R., Renckens, B., Nauta, A., Siezen, R.J., 2010. The proteolytic system of lactic acid bacteria revisited: a genomic comparison. *BMC Genomics*,11:36.

Martín-Platero, A., Valdivia, E., Maqueda, M., Martínez-Bueno, M., 2009. Characterization and safety evaluation of enterococci isolated from Spanish goats' milk cheeses. *International Journal of Food Microbiology* 132, 24–32.

Martínez-Castilla, L.P., 2007. Cap 4 "Reconstrucción de la historia de cambio de los caracteres". En: Eguiarte LE, Souza V y Aguirre X. (Ed.) *Ecología molecular*. Instituto Nacional de Ecología, Semarnat, Ciudad de México, México. pp. 152.

Martino, G. P., Quintana, I. M., Espariz, M., Blancato, V.S., Magni C. Aroma compounds generation in citrate metabolism of *Enterococcus faecium*: Genetic characterization of type I citrate gene cluster. *International Journal of Food Microbiology* 218, 27-37.

Mesnage, S., Chau, F., Dubost, L., Arthur, M., 2008. Role of N-acetylglucosaminidase and N-acetylmuramidase activities in *Enterococcus faecalis* peptidoglycan metabolism. *The Journal of Biological Chemistry* 283, 19845-19853.

Moriya, Y., Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. 2007. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Research* 35 (Web Server issue): W182-5.

Moreno, M. R, Leisner, J.J, Tee, L.K., Ley, C., Radu, S., Rusul, G., Vancanneyt, M. y De Vuyst, L., 2002. Microbial analysis of Malaysian tempeh, and characterization of two bacteriocins produced by isolates of *Enterococcus faecium*. *Journal of Applied Microbiology* 92, 147-157.

Nadalin, F., Vezzi, F., Policriti, A., 2012; 2013. GapFiller: a de novo assembly approach to fill the gap within paired reads. *BMC Bioinformatics*. doi: 10.1186/1471-2105-13-S14-S8.

Nes, I.F., Diep, D.B., Ike, Y., 2014. Enterococcal Bacteriocins and Antimicrobial Proteins that Contribute to Niche Control. *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*. Massachusetts Eye and Ear Infirmary Boston, pp 446-482.

Olvera García Myrna Elena, 2013. Evaluación de la inocuidad de *Enterococcus* spp. aislado de queso Cotija. Tesis de Maestría. Facultad de Química. UNAM. Ciudad de México, México.

O'Sullivan, O., Callaghan, J.O, Sangrador-Vegas, A., McAuliffe, O., Slattery, L, Kaleta, P., Callanan, M., Fitzgerald, G.F., Ross, R.P. y Beresford, T., 2009. Comparative genomics of lactic acid bacteria reveals a niche-specific gene set. *BMC Microbiology* 9:50.

Orgier, J.C., y Serror, P., 2008. Safety assessment of dairy microorganisms: The *Enterococcus* genus. *International Journal of Food Microbiology* 126, 291–301.

Overbeek, R., Olson, R., Pusch, G.D., Olsen, G.J., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., Parrello, B., Shukla, M., Vonstein, V., Wattam, A.R., Xia, F., Stevens, R., 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Research* 42, D206–D214.

Palmer, K.L., and M.S. Gilmore. 2010. Multidrug-resistant enterococci lack CRISPR-cas. *mBio* 1(4):e00227-10.

Palmer, K.L., et al. 2012. Comparative genomics of enterococci: variation in *Enterococcus faecalis*, clade structure in *E. faecium*, and defining characteristics of *E. gallinarum* and *E. casseliflavus*. *mBio* 3(1):e00318-11.

Palmer, K.L., Godfrey, P., Griggs, A., Kos, V.N., Zucker, J., Desjardins, C., Cerqueira, G., Gevers, D., Walker, S., Wortman, J., Feldgarden, M., Haas, B., Birren, B., Gilmore, M.S., 2012. Comparative genomics of enterococci: variation in *Enterococcus faecalis*, clade structure in *E. faecium*, and defining characteristics of *E. gallinarum* and *E. casseliflavus*. *MBio* 3, e00318-11.

Palmer, K.L., van Schaik, W., Willems, R.J.L., Gilmore, M.S., 2014. Enterococcal Genomics. In: Gilmore, M.S., Clewell, D.B., Ike, Y., Shankar, N. (Ed.) *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection* [Internet]. Eye and Ear Infirmary, Boston, Massachusetts.

Park, S.S., Kwon, H.Y., Tran, T.D., Choi, M.H., Jung, S.H., Lee, S., Briles, D.E., Rhee, D.K., 2015. ClpL is a chaperone without auxiliary factors. *FEBS Journal* 282, 1352–1367.

Paulsen, I.T., Banerjee, L., Myers, G.S., Nelson, K.E., Seshadri, R., Read, T.D., Fouts, D.E., Eisen, J.A., Gill, S.R., Heidelberg, J.F., Tettelin, H., Dodson, R.J., Umayam, L., Brinkac, L., Beanan, M., Daugherty, S., DeBoy, R.T., Durkin, S., Kolonay, J., Madupu, R., Nelson, W., Vamathevan, J., Tran, B., Upton, J., Hansen, T., Shetty, J., Khouri, H., Utterback, T., Radune, D., Ketchum, K.A., Dougherty, B.A., Fraser, C.M., 2003. Role of Mobile DNA in the Evolution of Vancomycin-Resistant *Enterococcus faecalis*. *Science* 299, 2071-2074.

Ponstingl, H., Ning Z., 2010. SMALT. <http://www.sanger.ac.uk/science/tools/smalt-0> (accessed 4.13.16.).

Qin, X., Galloway-Pena, J.R., Sillanpaa, J., Roh, J.H., Nallapareddy, S.R., Chowdhury, S., Bourgogne, A., Choudhury, T., Muzny, D.M., Buhay, C.J., 2012 Complete genome sequence of *Enterococcus faecium* strain TX16 and comparative genomic analysis of *Enterococcus faecium* genomes. *BMC Microbiology* 12:135.

Ruiz-Garbajosa, P., Bonten, M.J.M., Robinson, D.A., Top, J., Nallapareddy, S. R., Torres, C., Coque, T.M., Catón, R., Baquero F., Murray, B.E., del Campo, R., Willems, R.J.L., 2006. Multilocus Sequence Typing Scheme for *Enterococcus faecalis* Reveals Hospital-Adapted Genetic Complexes in a Background of High Rates of Recombination. *Journal of Clinical Microbiology*, 44(6), 2220-2228.

Salminen, S. von Wright, A., Ouwehand, A., 2004. Lactic Acid Bacteria. *Microbiological and Functional Aspects*. 3ª edición. Marcel Dekker Inc. Estados Unidos, pp 8-10.

- Savijoki, K., Ingmer, H., Varmanen, P., 2006. Proteolytic systems of lactic acid bacteria. *Applied Microbiology Biotechnology* 71, 394–406.
- Sarantinopoulos, P., Andrighetto, C., Georgalakis, M.D., Rea, M.C., Lombardi, A., Coganc, T.M., Kalantzopoulou, G., Tsakalidou E., 2001. Biochemical properties of enterococci relevant to their technological performance. *International Dairy Journal* 11, 621–647.
- Sarantinopoulos, P., Kalantzopoulos, G., Tsakalidou, E., 2002. Effect of *Enterococcus faecium* on microbiological, physicochemical and sensory characteristics of Greek Feta cheese. *International Journal of Food Microbiology* 76, 93–105.
- Serio, A., Chaves-López, C., Paparella, A., Suzzi, G., 2010. Evaluation of metabolic activities of enterococci isolated from Pecorino Abruzzese cheese. *International Dairy Journal* 20, 459–464.
- Semedo, T., Almeida, M.S., Martins, P., Silva, L.M.F., Figueiredo, M.J.J., Tereiro, R. y Barreto, C.M.T. 2003. Comparative study using type strains and clinical and food isolates to examine hemolytic activity and occurrence of the *cyl* operon in *Enterococci*. *Journal of Clinical Microbiology* 41, 2569-2576.
- Shendure, J. y Ji, H., 2008. Next-generation DNA sequencing. *Nature Biotechnology* 26, 1135-1145.
- Simpson, J.T., Wong, K., Jackman, S. D., Schein, J. E., Jones, S. J., Birol, I., 2009. ABySS: a parallel assembler for short read sequence data. *Genome Research* 19, 1117-1123.
- Smit, G., Smit, B.A., Engels, W.J.M., 2005. Flavour formation by lactic acid bacteria and biochemical flavour profiling of cheese products. *FEMS Microbiology Reviews* 29, 591–610.
- Tamura, K., Stecher, G., Peterson, D., Filipinski, A., Kumar, S., 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* 12, 2725-2729.
- Urrieta Vázquez Jazmín, 2018. Análisis multilocus de *Enterococcus faecium* aislados del queso de Ocosingo. Tesis de licenciatura. Facultad de Química, UNAM. Ciudad de México, México.
- Vollmer, W., Joris, B., Charlier, P., Foster, S., 2008. Bacterial peptidoglycan (murein) hydrolases. *FEMS Microbiology Reviews* 32, 259–286.
- van Dijk, E.L., Auger, H., Jaszczyszyn, Y., Thermes, C., 2014. Ten years of next-generation sequencing technology 30 (9), 418-426.
- van Heel, A.J., de Jong, A., Montalbán-López, M., Kok, J., Kuipers, O.P., 2013. BAGEL3: automated identification of genes encoding bacteriocins and (non-)bactericidal posttranslationally modified peptides. *Nucleic Acids Research* 41, W448eW453.
- van Kranenburg, R., Kleerebezem, M., van Hylckama Vlieg J., Ursing, B.M., Boekhorst, J., Smit, B.A., Ayad, E.H.E, Smit, G., Siezen, R.J., 2002. Flavour formation from amino acids by lactic acid bacteria: predictions from genome sequence analysis. *International Dairy Journal* 12, 111–121.

van Schaik W. y R. J. L. Willems, 2010. Genome-based insights into the evolution of enterococci. *Clinical Microbiology and Infection*, 16, 527–532.

Weigand, M., Ashbolt, J., Konstantinidis, K.T., Santo Domingo, J.W., 2014. Genome Sequencing Reveals the Environmental Origin of *Enterococci* and Potencial Biomarkers for Water Quality Monitoring. *Environmental Science & Technology* 48, 3707-3714.

Willems, 2010. Population Genetics of *Enterococcus*. In: Robinson A.D., Falush D., Feil E. J. Bacterial Population Genetics in Infectious Disease. John Wiley & Sons, Inc. and Edward J. Feil, pp 195-216.

Wunderlichová, L., Bunková, L., Koutný, M., Jancová, P., Bunka F., 2014. Formation, Degradation, and Detoxification of Putrescine by Foodborne Bacteria: A Review. *Comprehensive Reviews in Food Science and Food Safety* 13, 1012-1030.

Yoon SH, Park YK, and Kim JF. 2014. PAIDB v2.0: exploration and analysis of pathogenicity and resistance islands. *Nucleic Acids Research*. doi: 10.1093/nar/gku985

Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV. 2012. Identification of acquired antimicrobial resistance genes. *Journal antimicrobial chemotherapy* 67(11), 2640-4.

Zerbino, D.R. y Birney E., 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research* 18, 821-829.

Zischka M. y Kuenne C. 2012. Complete Genome Sequence of the Porcine Isolate *Enterococcus faecalis* D32. *Journal of Bacteriology*.194:5490-5491.

Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. 2011. "PHAST: A Fast Phage Search Tool". *Nucleic Acid Research*. 39 (Web Server issue): W347-52.

Sitios de internet consultados:

- Base de datos MLST
<http://efaecalis.mlst.net/>
<http://efaecium.mlst.net/>
- Base de datos PubMLST
<https://pubmlst.org/efaecalis/>
<https://pubmlst.org/efaecium/>
- e-Burst
<http://efaecalis.mlst.net/eburst/>
<http://efaecium.mlst.net/eburst/>
- Genebank (genomas de bacterias)
ftp://ftp.ncbi.nlm.nih.gov/genomes/archive/old_refseq/Bacteria/
- Illumina
<http://www.illumina.com/>
- MultAlin
<http://multalin.toulouse.inra.fr/multalin/>
- OligoCalc
<http://simgene.com/OligoCalc>
- Organización Mundial de la Salud (OMS) (Lawe-Davis y Bennet, 2017)
<http://www.who.int/mediacentre/news/releases/2017/bacteria-antibiotics-needed/es/>

- NCBI (Blastn)
https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LO C=blasthome
- Trinotate
<http://trinotate.github.io/>
- Bagel3
<http://bagel.molgenrug.nl/>

12. Anexos

12.1 Anexo I

Tabla AI. Lista de cepas utilizadas en este trabajo

Abreviatura	Especie	Origen	Referencia
Cepa A	<i>E. faecalis</i>	Queso Cotija, Región de Origen	Bravo, 2008
Cepa B	<i>E. faecium</i>		
Cepa C	<i>E. faecium</i>		
Cepa D	<i>E. faecium</i>		
Cepa E	<i>E. faecalis</i>		
Cepa F	<i>E. faecalis</i>		
Cepa G	<i>E. faecium</i>		
Cepa QA-1	<i>E. faecalis</i>		Guzmán, 2015
Cepa QB-3	<i>E. faecalis</i>		
Cepa QC-4	<i>E. faecalis</i>		
Cepa QD-2	<i>E. faecium</i>		
Cepa QE-2	<i>E. faecalis</i>		
Cepa QG-5	<i>E. faecium</i>		
Cepa A5-1	<i>E. faecalis</i>		
Cepa 10	<i>E. faecium</i>		
Cepa 9	<i>E. faecium</i>		
Cepa 6	<i>E. faecium</i>		
Cepa 3	<i>E. faecium</i>		
Cepa 2	<i>E. faecium</i>		

12.2 Anexo II

Tabla AII. Información de cepas utilizadas para construir el árbol filogenético de *E. faecalis*

ST	Cepa	Epidemiología
1	OG1RF	Laboratorio
6	V583	Brote hospitalario
16	XI3A5	Aislado clínico animal
17	H14	Aislado clínico
18	H8	Aislado clínico
21	Z68	Aislado de animal
24	H4	Aislado clínico
335	Efc-16-09	Aislado clínico
89	1020/03	Aislado clínico
51	H17	Aislado clínico
181	607948	Aislado clínico
731	TRE157002	Roedor
566	Enb67	Aislado clínico
52	H18	Aislado clínico
46	H22	Paciente hospitalizado
70	Osaka-34	Paciente hospitalizado
157	McBride O	Desconocida
568	Enb73	Paciente hospitalizado
9	A1201	Paciente hospitalizado
40	D32	Animal
248	Symbioflor1	Comensal de una personal
21	A*	Alimento-origen animal
712	E*	Alimento-origen animal
712	F*	Alimento-origen animal
21	QA1*	Alimento-origen animal
757	QB3*	Alimento-origen animal
758	QC4*	Alimento-origen animal
758	A51*	Alimento-origen animal
5	BM5	Pollo
21	QCB1	Alimento-origen animal
72	42 ACA-DC 3318	Alimento-origen animal
168	V8C1	Alimento-origen vegetal
145	VE14522-	Alimento-origen animal (queso francés)
166	LMGT3208-	Alimento-origen animal (queso griego)
169	V68C1	Alimento-origen vegetal

443	DE 03_6	Ambiental
442	Ef79	Desconocido
196	A1638	Comensal
73	QA46	Alimento-origen animal
612	CU22	Alimento-origen animal
162	189	Comensal
	DO	<i>E. faecium</i> aislado nosocomial

*Cepas reportadas en este estudio

Tabla AIII. Información de cepas utilizadas para construir el árbol filogenético de *E. faecium*

ST	Cepa	Epidemiología
5	152s	Agricultura
27	SK42	Pollos
212	EIL02A	Persona no hospitalizada
44	649	Becerro
217	VSEFMfp1	Cerdo
135	ICIS73	Persona no hospitalizada
216	EIL12C	Persona no hospitalizada
545	ICIS36	Persona no hospitalizada
156	A31	Aguas residuales
296	SF-68	Probiótico
94	9800244B	Persona no hospitalizada
42	9800349C	Persona no hospitalizada
720	PC4.1	Alimento
860	NRRL B-2354	Agricultura
61	9800295D	Persona no hospitalizada
40	9424.1	Ambiental
60	102s	Persona no hospitalizada
820	IsoA3	Ambiental
1045	C*	Alimento-origen animal
1046	D*	Alimento-origen animal
135	ICIS73	Persona no hospitalizada
1047	QE2*	Alimento-origen animal
1048	QG5*	Alimento-origen animal
1295	6*	Alimento-origen animal
1293	2*	Alimento-origen animal
1294	3*	Alimento-origen animal
130	10* y 9*	Alimento-origen animal

53	229s	Agricultura
125	7947_99	Ambiental
22	GRAVanE2	Alimento
18	58155	Paciente hospitalizado /Brote hospitalario
582	Efm1231408	Paciente hospitalizado
17	TX2480a	Paciente hospitalizado
52	41s	Paciente hospitalizado
203	SMC 98-1	Paciente hospitalizado
16	VS5	Paciente hospitalizado
25	BM4147	Paciente hospitalizado
106	4618/61	Paciente hospitalizado
78	EFM48	Paciente hospitalizado
280	DB20546/04	Paciente hospitalizado
125	7947_99	Ambiental
80	Is1-54	Paciente hospitalizado
170	2084	Paciente hospitalizado
344	31013	Paciente hospitalizado
249	VRE 15	Paciente hospitalizado
186	HE43	Paciente hospitalizado
341	ZP01224	Paciente hospitalizado
117	773	Paciente hospitalizado
233	T-2002	Paciente hospitalizado
62	21s	Paciente hospitalizado
	V583	<i>E. faecalis</i> aislado nosocomial

*Cepas reportadas en este estudio

12.3 Anexo III

Alineamiento contra las referencias más cercanas usando Mauve.

Fig AI. Reordenamiento del ensamble de *E. faecalis* usando como referencia el genoma de *E. faecalis* Symbioflor 1

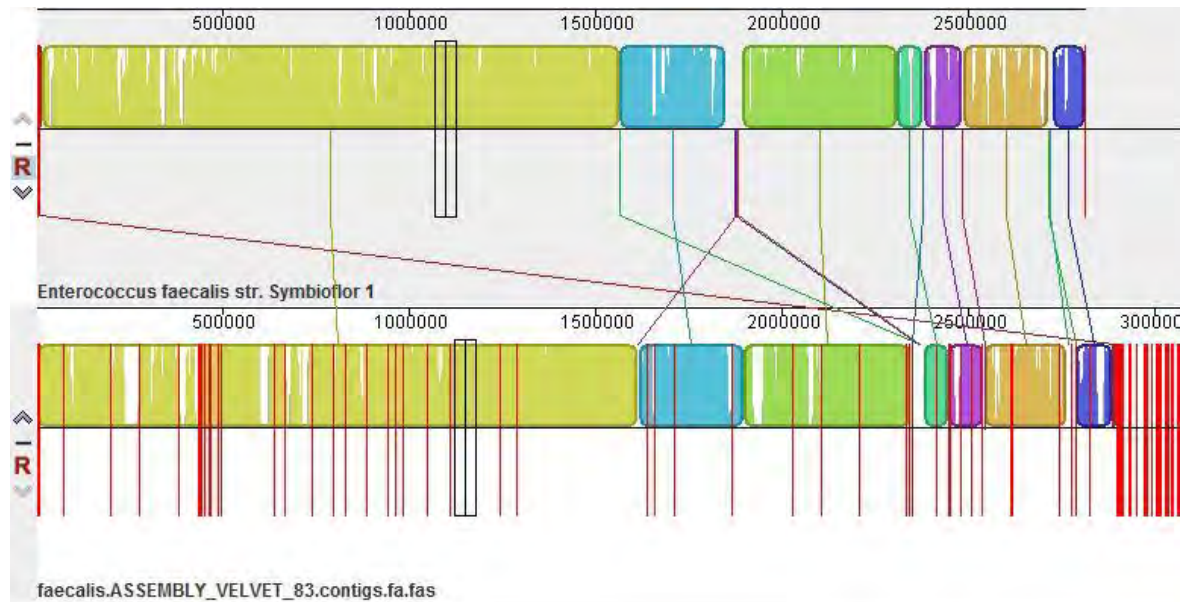
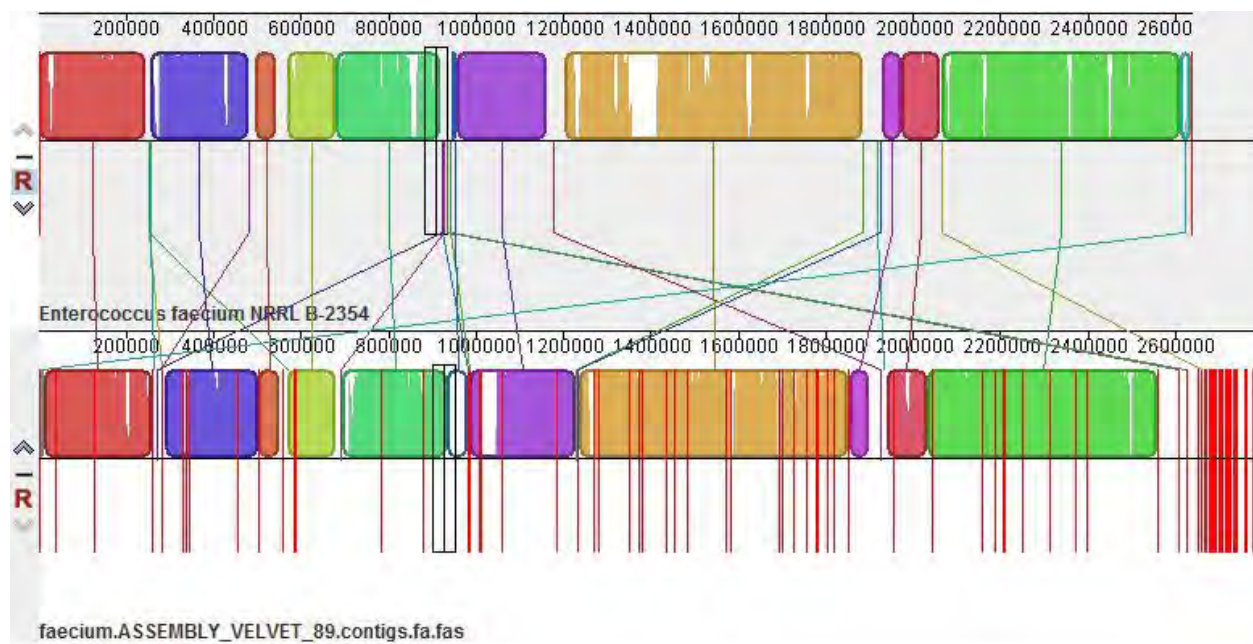


Fig AII. Reordenamiento del ensamble de *E. faecium* D usando la referencia NRRL B-2354



12.4 Anexo IV

Tabla AIV. Anotación funcional de los genes codificantes de proteína presentes en los plásmidos de *E. faecalis* F

Plásmido	Tamaño (pb)	Anotación funcional
pstrainF1	9,882	Integrasa(2), subunidad PTS transportado de azúcar IIB (2), proteína hipotética (5), MATE Transportador de salida, fosfolglicerol mutasa, fosfoglicolato fosfatasa y proteína TraC de conjugación.
pstrainF2	4,718	Proteína hipotética (3), proteínas de movilización <i>Mob E</i> , proteína iniciadora <i>rep E</i> .

Entre paréntesis se identifica el número de genes para esa anotación.



Genomic and functional characterisation of two *Enterococcus* strains isolated from Cotija cheese and their potential role in ripening

Myrna Olvera-García¹ · Alejandro Sanchez-Flores² · Maricarmen Quirasco Baruch¹

Received: 9 August 2017 / Revised: 18 December 2017 / Accepted: 26 December 2017
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

Enterococcus spp. are present in the native microbiota of many traditional fermented foods. Their ability to produce antibacterial compounds, mainly against *Listeria monocytogenes*, has raised interest recently. However, there is scarce information about their proteolytic and lipolytic potential, and their biotechnological application is currently limited because enterococcal strains have been related to nosocomial infections. In this work, next-generation sequencing and optimised bioinformatic pipelines were used to annotate the genomes of two *Enterococcus* strains—one *E. faecium* and one *E. faecalis*—isolated from the Mexican artisanal ripened Cotija cheese. A battery of genes involved in their proteolytic system was annotated. Genes coding for lipases, esterases and other enzymes whose final products contribute to cheese aroma and flavour were identified as well. As for the production of antibacterial compounds, several peptidoglycan hydrolase- and bacteriocin-coding genes were identified in both genomes experimentally and by bioinformatic analyses. *E. faecalis* showed resistance to aminoglycosides and *E. faecium* to aminoglycosides and macrolides, as predicted by the genome functional annotation. No pathogenicity islands were found in any of the strains, although traits such as the ability of biofilm formation and cell aggregation were observed. Finally, a comparative genomic analysis was able to discriminate between the food strains isolated and nosocomial strains. In summary, pathogenic strains are resistant to a wide range of antibiotics and contain virulence factors that cause host damage; in contrast, food strains display less antibiotic resistance, include genes that encode class II bacteriocins and express virulence factors associated with host colonisation rather than invasion.

Keywords Comparative genomics · Lactic acid bacteria proteolytic system · Antibacterial compounds · Bacteriocins · Bioinformatic tools

Introduction

Enterococci are ubiquitous bacteria found predominantly in the gastrointestinal tract of humans and animals, where they play a beneficial role in the balance of the microbial

population, as implied by the use of some strains as probiotics. It has recently been proposed that the ancestral habitat of *Enterococcus faecalis* was related to an extra-enteric environment and that their adaptation to the animal gastrointestinal tract took place relatively recently (Weigand et al. 2014). Additionally, they are non-starter lactic acid bacteria (NSLAB) in a variety of cheeses, especially artisanal ones manufactured with either raw or pasteurised milk (Cebrián et al. 2012). The most common *Enterococcus* species in food products are *E. faecium*, *E. faecalis* and *E. durans* (Orgier and Serror 2008). *Enterococcus* strains may function as starter cultures and co-cultures in cheeses such as Feta, Cheddar, Mozzarella and Venaco (Foulquié-Moreno et al. 2006). Moreover, enterococci could play a key role in the ripening process, probably through proteolysis, lipolysis and citrate breakdown. In a fermented dairy product, caseins are partially degraded into peptides, which in turn are hydrolysed into amino acids; finally, these could be transformed into acids,

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00253-018-8765-3>) contains supplementary material, which is available to authorized users.

✉ Maricarmen Quirasco Baruch
quirabma@unam.mx

¹ Fac. de Química, Departamento de Alimentos y Biotecnología, Universidad Nacional Autónoma de México, Av. Universidad 3000, C. U, 04510 Ciudad de México, Mexico

² Instituto de Biotecnología, Unidad de Secuenciación Masiva y Bioinformática, Av. Universidad 2001, Chamilpa, 62210 Cuernavaca, Morelos, Mexico

alcohols and aldehydes. In particular, the conversion of methionine, as well as aromatic and branched-chain amino acids, is crucial to flavour and aroma development in cheese. The role of lipolysis in the synthesis of flavour precursors is also important, through the hydrolysis of triacylglycerides into free fatty acids, which are later converted into methyl ketones, secondary alcohols, esters and lactones. In addition, citrate catabolism results in the production of compounds such as diacetyl, acetoin, acetaldehyde and 2,3-butene-diol, which are also responsible for the final flavour and odour features of a given dairy product (Smit et al. 2005). Besides, the production of carbon dioxide can contribute to texture in some cheeses (Foulquié-Moreno et al. 2006). Enterococci isolated from food have been studied because of their useful biotechnological traits, such as bacteriocin production and probiotic behaviour; however, their role in cheese ripening, related to enzymatic activities such as proteolysis and lipolysis, is still not fully elucidated. Nowadays, bacterial genome sequencing and bioinformatic analysis are powerful tools to improve the knowledge of all coding sequences and putative complete metabolic pathways; these could yield information to better understand the process involved in cheese ripening (van Kranenburg et al. 2002).

On the other hand, enterococci are the most controversial group among lactic acid bacteria (LAB). They have been associated with human nosocomial infections and can harbour some genes related to virulence factors, such as those coding for haemolysin, adhesins and invasins; additionally, some strains are resistant to antibiotics (e.g. vancomycin) commonly used in clinical practice (Franz et al. 2001; Orgier and Serror 2008). The incidence of virulence factors in enterococci isolated from food and other natural sources (e.g. water and vegetables) has shown to be low compared to the one of clinical strains (Franz et al. 2001). Moreover, genetic analyses have revealed that *E. faecium* and *E. faecalis* exhibit considerable inter-strain genomic diversity, which is mainly associated to a variable presence of phage DNA, plasmids, pathogenicity islands and conjugative elements (Kim and Marco 2014; Palmer et al. 2014). Therefore, the divergent traits among nosocomial and food strains could be the result of broader niche-specific adaptations that can be explored through a comparative genomic analysis.

E. faecalis and *E. faecium* strains isolated from Cotija cheese showed an antibacterial activity that is potentially useful in food preservation technology. These strains produce extracellular peptidoglycan hydrolases active against Gram-positive and Gram-negative pathogenic bacteria such as *Staphylococcus aureus* and *Salmonella enterica* (García-Cano et al. 2014). However, aspects still uncertain are the role of these strains in the Cotija cheese ripening process and, more importantly, whether there is any potential human health risk related to their presence in this artisanal food.

Therefore, the aim of this work was to provide a comprehensive outlook of the metabolic potential of one *E. faecalis* and one *E. faecium* strain, both isolated from Cotija cheese, related to ripening and flavour development, as well as to their antibacterial compound production. On the other hand, a comparative genomic analysis was carried out in order to differentiate food strains from nosocomial ones and find biomarkers associated with fermented food strains.

Materials and methods

Strain selection, culture conditions, DNA and plasmid extraction

The strains used in this study, *E. faecalis* F (CFQ-B-254) and *E. faecium* D (CFQ-B-255), were isolated from 'Region of Origin' handmade Cotija cheese, manufactured and ripened following the process established in the Mexican standard (NMX-F-735-COFOCALEC 2011), and are deposited in the Culture Collection of the School of Chemistry (CFQ) at the National Autonomous University of Mexico (WDCM No. 100). They were identified by 16S rRNA sequencing (García-Cano et al. 2014). *E. faecalis* F was previously reported as a peptidoglycan hydrolase producer with antibacterial activity against pathogenic bacteria (García-Cano et al. 2014). The strains were stored in cryogenic conditions ($-70\text{ }^{\circ}\text{C}$) and were reactivated in MRS broth (de Man, Rogosa and Sharp; Oxoid, Basingstoke, UK). High-quality genomic DNA was extracted with the Fast ID Genomic DNA extraction kit (Genetic ID NA, Inc., Fairfield, IA, USA). Plasmid extraction was carried out using the GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific, Vilnius, Lithuania), according to the manufacturer's protocol, with some modifications: the volume of bacterial culture used was 20 mL and the incubation time with the lysis solution for plasmid extraction was 7 min.

Genome sequencing and in silico analysis

Sequencing and annotation

A 500-bp insert library was prepared using the TruSeq kit v2 (Illumina, Inc., San Diego, CA, USA). Sequencing was performed on the HiSeq 2000 (Illumina) platform with a 200-cycle paired-end configuration at the National Laboratory of Genomics for Biodiversity (LANGEBIO), Irapuato, Guanajuato, Mexico. The reads for each strain were obtained with a high quality (Phred score ≥ 20 at each nucleotide position), assessed with the FastQC software (Andrews 2016). Prior to the de novo read assembly, the khmer tool kit was used to normalise the read coverage and to exclude reads that were not suitable for assembling because of their size ($< 100\text{ bp}$) (Brown et al. 2012). For each strain, read assemblies

were carried out using the ABySS v1.5.2 and Velvet v1.2.09 software, using default parameters, in a k-mer range 21–91 (Simpson et al. 2009; Zerbino and Birney 2008). Assemblies with the best statistics based on N_{50}/N_{90} were selected: for *E. faecalis* k-mer 61 (ABySS) and for *E. faecium* k-mer 89 (Velvet). In order to find the closest reference genome for each species (*faecalis* and *faecium*), reads were mapped against the available GenBank's complete genomes using the mapper Smalt (> smalt map -r 0 -y 0.8 -f samsoft -o) (Ponstingl and Ning 2010, <ftp://ftp.sanger.ac.uk/pub/resources/software/smalt/>; ftp://ftp.ncbi.nlm.nih.gov/genomes/archive/old_refseq/Bacteria/). Genomes selected as reference are shown in Table 1. The resulting assemblies were oriented using progressive alignments with the Mauve 2.3.1 software with default parameters (Darling et al. 2010). To identify and correct assembling errors, each genome assembly was evaluated using the REAPR v3 software (Hunt et al. 2013). The SSPACE v3 and Gapfiller v1.10 programs were used for scaffolding and scaffold gap filling, respectively (Boetzer et al. 2011; Nadalin et al. 2012). GenBank accession numbers of the genome assemblies generated in this work are in Table 1. Gene prediction was carried out with GeneMark.hmm Prokaryotic 3.25 for each species, with default parameters (Besemer et al. 2001). Preliminary functional annotation was performed for each genome with the online server Rapid Annotation using Subsystem Technology (RAST-SEED)

(Aziz et al. 2008; Overbeek et al. 2014). Additional databases were searched for genes associated with pathogenicity islands (PAIDB v2.0) (Yoon et al. 2014), CRISPR-cas region prediction (CRISPRfinder) (Grissa et al. 2007), prophage sequences and features (PHAge Search Tool-PHAST) (Zhou et al. 2011), and plasmid prediction (plasmidFinder) (Carattoli et al. 2014).

Functional annotation

The annotation for gene prediction was performed by adapting the Trinotate pipeline (<http://trinotate.github.io/>) as follows: the predicted amino acids were compared against the UniProtKB/Swiss-Prot database using the Blast v2.2.28+ algorithm (Camacho et al. 2009). A functional domain annotation was performed using the hmmscan from suiteHMMER v3.1b1 (Eddy 2011) to search against the Pfam-A database. Signal peptide and transmembrane domain prediction were performed using SignalP (Petersen et al. 2011) and tmHMM (Krogh et al. 2001), respectively. The in silico detection of bacteriocin genes was performed using the Bagel3 (Bagel automated bacteriocin mining) web software (van Heel et al. 2013) and was completed with the Trinotate functional annotation. Considering the virulence traits and genes related to biogenic amine formation reported by Orgier and Serror (2008), a comprehensive search for genes with such annotation was carried out. A search of Pfam domains and Swiss-

Table 1 General genomic features of *E. faecalis* strain F, *E. faecium* strain D and other related genomes from *Enterococcus* strains

Accession number (GenBank)	Strain	Source	Genome size (Mbp)	%GC	CDS	Plasmid	Reference
CP002621.1	OG1RF ^a	Virulent laboratory strain	2.74	37.8	2579	–	Bourgogne and Garsin (2008)
HF558530.1	Symbioflor1 ^a	Probiotic	2.81	37.7	2808	–	Fritzenwanker and Kuenne (2013)
CP003726.1	D32 ^a	Porcine faeces	3.06	37.5	2965	2	Zischka and Kuenne (2012)
CP002491.1	62 ^a	Healthy infant	3.13	37.2	3157	3	Dag Anders and Lars (2011)
AE016830.1	V583 ^a	Nosocomial	3.36	37.5	3264	3	Paulsen et al. (2003)
CP014949.1	LD33	Traditional dairy product	2.8	37.6	2695	–	Yuehua et al. (2016)
CP015410.2	KB1	Mammalian intestinal microbiota	3.03	37.2	2815	–	Uchimura et al. (2016)
GCA_002006705.1 ^b	Strain F	Cotija cheese	3.4	37.6	3252	2	This study
CP003583.1	DO ^a	Nosocomial	2.7	38.15	2703	3	Qin et al. (2012)
CP004063.1	NRRL B-2354 ^a	Milk and dairy utensils	2.84	37.8	2639	1	Kopit et al. (2014)
CP006620.1	Aus0085 ^a	Nosocomial	2.99	38.2	2922	6	Lam et al. (2013)
CP003351.1	Aus004 ^a	Nosocomial	3.02	38.31	2932	3	Lam et al. (2012)
CP006030.1	T110	Probiotic	2.69	38.5	2503	1	Natarajan and Parani
CP012522	64/3	Hospitalised patient	2.57	38.2	2382	–	Bender et al. (2015)
GCA_002006745.1 ^b	Strain D	Cotija cheese	2.77	37.9	2804	–	This study

CDS coding DNA sequences

^a Used as references in this study

^b Draft genomes assemblies

Prot products for peptidoglycan hydrolases activity annotation was also performed.

Identification of genes associated with development of flavours and proteolytic and lipolytic systems

The metabolic analysis of complete sequences was carried out with the KEEG Automatic Annotation Server for orthologue assignment and pathway mapping (KAAS), using gene prediction in amino acids (Moriya et al. 2007). The production of flavour and odour molecules was predicted through the identification of genes corresponding to the complete metabolic pathway, as well as searching for genes that encode enzymes reported for specific reactions. The functional annotation described before also allowed the identification of proteinases, peptidases and amino acid transporters related to proteolytic systems and enzymes related to lipolytic systems in LAB.

Comparative genomic analysis

For the comparison within different enterococci species (*faecalis* or *faecium*), the Get-Homologues v20140627 was used. Homologous genes were clustered using the bidirectional best-hit, COGtriangles or/and OrthoMCL algorithms (Contreras-Moreira and Vinuesa 2013). The protein clusters obtained from the Get-Homologues software were parsed using Perl scripts. The auxiliary script *compare_clusters.pl* was used to generate a pangenomic matrix of the presence/absence of protein clusters, which was inferred from the output generated by the main script *get_homologues.pl*. For each species, the pangenome (gene repertoire), the genome core (common genes, mutually conserved) and the exclusive genes (only found in one genome) were analysed using Perl scripts and summarised in an absence/presence matrix, taking the smallest genome as a reference. A parsimony pangenome tree was built for each species using the presence/absence matrix. Finally, to perform a circular comparison between reference genomes and the genomes from our strains, the BRIG 0.95 software was employed (Alikhan et al. 2011).

Phenotypic assays

Bacteriocin analysis

For bacteriocin production, each strain was cultured in MRS broth at 37 °C for 10 h and the concentration of extracellular bacteriocins was carried out by a pH-mediated adsorption–desorption method (Alvarez-Cisneros et al. 2010). Afterwards, protein bands with lytic activity were detected by zymography using Tricine-SDS-PAGE (Schägger and von Jagow 1987) against *Listeria innocua*. *E. faecium* MXVK29 was used as positive control (Alvarez-Cisneros et al. 2010).

Peptidoglycan hydrolase identification and activity test

Peptidoglycan hydrolase (PGH) activity was tested as described by García-Cano et al. (2014). *Enterococcus* strains were cultured in MRS broth and extracellular PGH was prepared by trichloroacetic acid (10% w/v) (TCA) precipitation of the neutralised fermentation supernatant, followed by testing the concentrated protein preparation (CPP). Protein bands with bacteriolytic activity were identified through SDS-PAGE and zymograms against *Micrococcus lysodeikticus*.

Production of cytolysin (haemolysin)

The phenotype associated with cytolysin gene expression was determined by the ability to hydrolyse erythrocyte membrane. The test was considered positive when clear zones were apparent around the *Enterococcus* colonies after growth at 37 °C for 24 h in Blood Agar Base (Oxoid, Basingstoke, UK) supplemented with 5% (v/v) of defibrinated ram blood (DIBICO, Mexico City, Mexico).

Adherence to cellular lines Caco-2 and HT-29

Cell aggregation and adherence to cell hosts were tested following the method described by Nueno-Palop and Narbad (2011) with some modifications. For cell line culture conditions, HT-29 cells were cultured in Minimal Essential Medium, whereas Enterocyte-like Caco-2 (Caco-2) cells were grown in Dulbecco's modified Eagle's Minimal Essential Medium (Sigma, St. Louis MO, USA), both supplemented with 10% (v/v) foetal bovine serum (FBS) (Sigma, St. Louis MO, USA) and 1% (v/v) penicillin–streptomycin (Sigma, St. Louis MO, USA). Cells were grown at 37 °C in a humidified atmosphere of 5% CO₂ and 95% air for 24 h (for HT-29) and 48 h (for Caco-2), or until a confluent monolayer was microscopically observed (×40). Afterwards, cells were rinsed with sterile PUKS (0.14 M NaCl, 5 mM KCl, 4 mM NaHCO₃ and 1.15 mM EDTA) and transferred to 24-well tissue culture plates (Corning Costar; Sigma Aldrich, St. Louis MO, USA) with the corresponding medium, without FBS and antibiotics. Finally, plates were incubated at the conditions previously described for 24 h. In parallel, bacterial strains were cultivated in MRS with 10% (w/v) D-mannose for 24 h at 37 °C. A bacterial culture with OD₆₀₀ of 1.5 units approximately was added to each well containing the mammal tissue cultures, and plates were incubated for 3 h at 37 °C. After incubation, each well was rinsed three times with a sterile pre-heated (37 °C) phosphate buffer, pH 7, then the bacterial cells were fixed using methanol and stained with the Giemsa stain. Bacterial adhesion was observed under light microscopy (×100) and described qualitatively. Assays were performed in triplicate. *Escherichia coli* strain 2348/69 (enteropathogenic, EPEC) (Jerse et al. 1990) was used as positive control for localised

adherence; *E. coli* 049766 (enteroaggregative, EAEC), for aggregative adherence (Navarro-García et al. 1998).

Biofilm formation on polystyrene

The capability of cells to adhere to polystyrene plates was evaluated using a previously described method by Cebrián et al. (2012) with some modifications. Each strain was grown in MRS broth at 37 °C, for 10 h; afterwards, 200 µL of a 1:100 dilution of the cultures was added to a sterile 96-well polystyrene microtitre plate (Corning Costar; Sigma Aldrich, St. Louis MO, USA). After incubating for 14 h at 37 °C, wells were emptied, dried in an inverted position for 15 min and stained with 1% (*m/v*) crystal violet for 15 min. Afterwards, wells were washed with 200 µL of 70% (*v/v*) ethanol. OD₆₂₀ was read in a microplate spectrophotometer (Epoch BioTek, Winooski VT, USA). *E. faecalis* ATCC 29212 was used as a positive control.

Antibiotic susceptibility testing

Antibiotic susceptibility was assessed by means of a diffusion test. Circular filters were supplemented with the following antibiotic concentrations: 16 µg/mL tetracycline, 8 µg/mL penicillin, 20 µg/mL kanamycin, 8 µg/mL erythromycin,

1 mg/mL streptomycin (high level), 32 µg/mL vancomycin, 500 µg/mL gentamicin (high level) and 16 µg/mL ampicillin (Sigma, St. Louis MO, USA) (Clinical and Laboratory Standards Institute 2012; Franz et al. 2001). These were placed on top of a bacterial lawn of 10⁷ CFU/mL. Plates were incubated at 37 °C for 24 h and evaluated. Inhibition zones around discs were considered as a positive result.

Results

Genome sequencing, assembly and annotation

A total of 5.6 and 4.5 Mbp, in 100-bp paired-end reads, were obtained for strains D and F, respectively. They showed an average Phred score ≥ 20 , indicating a high-quality sequencing. Results are summarised in Table 2 and Supplemental Table S1 for plasmid information. Reads were mapped against all the available fully closed and well-characterised *Enterococcus* genomes at the time (Table 1). The highest percent of mapped reads and synteny were obtained with the genome of an *Enterococcus* used in the food industry (*E. faecium* NRRL B-2354) and the one belonging to a probiotic (*E. faecalis* Symbioflor1) for *E. faecium* strain D and *E. faecalis* strain F, respectively.

Table 2 Results for sequencing and principal characteristics of each strain in this study

	<i>E. faecalis</i> F	<i>E. faecium</i> D
Total number of reads sequenced	5.6 Mbp	4.5 Mbp
Assembly		
Contigs	70 (ABYSS)	81 (Velvet)
Total of reads assembled	3,441,804 bp	2,777,228 bp
N50/90	91,252/28,483 bp	88,852/21,856 bp
Average contig length	49,168 bp	33,874 bp
%GC	37.59	37.92
Chromosome		
Size	3,427,226 bp	2,779,320 bp
Protein-coding ORFs	3252	2804
tRNA-CDS	65	63
rRNA-CDS	3 copies of 5S 3 copies of 23S 4 copies of 16S	3 copies of 5S 1 copy of 23S 1 copy of 16S
Plasmids ^a		
Number	2	Although two CDS for <i>repA</i> and one for <i>repE</i> were predicted, experimentally the absence of plasmids was assessed
Size	9882 and 4718 bp	

The software used for assembling is shown in parentheses

CDS coding DNA sequences

^a The protein-coding ORFs detected for each plasmid are shown in Supplemental Table S1

Potential roles of enterococci in Cotija cheese ripening and safety

Metabolic activities associated with ripening in cheese

Due to the importance of casein hydrolysis during ripening, the reconstructed genome and its annotation were analysed in order to identify proteolytic systems similar to the ones reported for other LAB used as starter cultures (Liu et al. 2010). In each genome, we detected genes coding for enzymes that could hydrolyse casein molecules, others coding for peptide transporters and some peptidases (Table 3). A hypothetical enterococcal proteolytic system is depicted in Fig. 1. The Clp complex could break casein molecules into oligopeptides, which would then be transported into the cell by the Opp system, the dipeptide transporter (Dpp) and the tripeptide transporter (DptT). Inside the cell, different peptidases would hydrolyse peptides into amino acids. The peptide transporter systems would be located in the cell membrane; this location was predicted through a Tmhmm analysis. Once inside the cell, amino acids would be used as a nitrogen source for growth and as precursors of flavour molecules.

Since lipolysis is also critical for cheese ripening, we searched for esterase genes in each genome functional annotation. Genes for five putative esterases, two acetyl esterases (a tributyrin esterase and a lipase) and two carboxylesterases were found in the *E. faecalis* F genome. As for the *E. faecium* D genome, only two acetyl esterases (a tributyrin esterase and a lipase) and a carboxylesterase were annotated. Preliminary experimental results support proteolytic and lipolytic activities in these strains.

Genes coding for enzymes associated with flavour generation were detected using a KEGG metabolic pathway analysis. Based on the description of flavour formation by LAB in cheese products by Smit et al. (2005), we found that *E. faecalis* F and *E. faecium* D may play a role in flavour development. We found genes related to metabolic pathways for the production of acetoin and diacetyl, acetaldehyde, acetate and ethanol, mediated by acetolactate decarboxylase (EC 4.1.1.5), acetaldehyde dehydrogenase (EC 1.2.1.10), acetate kinase (EC 2.7.2.1) and alcohol dehydrogenase (EC 1.1.1.1), respectively. Diacetyl may also be produced by a non-enzymatic oxidative decarboxylation of acetolactate.

Production of antibacterial compounds

Genes coding for antagonistic substances against pathogenic bacteria were identified in both genomes. According to the functional annotation and Bagel3 analyses, only one operon coding for bacteriocins was identified in the *E. faecalis* strain F chromosome. This operon includes six genes encoding the following proteins: enterocin P, which is flanked by two immunity proteins; an ABC transporter; a modification enzyme

and a transposase. The first four proteins are essential for bacteriocin activity. On the other hand, five genes were detected in *E. faecium* strain D coding for bacteriocins: enterocins A, B, P and two genes for enterocin SE-K4 (when these were translated in silico, the percent identity in the amino acid sequence was >90%). In addition, a sixth gene encoding lactacin F, subunit LafX, was identified, yielding 63% identity in the amino acid sequence. Also, six proteins related to enterocin A immunity (*entAI*) were found in this strain. Experimentally, two peptides with lytic activity against *L. innocua* were identified in *E. faecium* strain D, with molecular masses of approximately 4 and 5 kDa, respectively (Fig. 2). No lytic activity was observed in *E. faecalis* strain F.

The functional annotation allowed identifying 15 PGH-coding genes in the *E. faecalis* genome annotated as genes coding for three *N*-acetylmuramoyl-L-alanine amidases, four *N*-acetylmuramoyl-L-alanine amidases + SH3 domain, two *N*-acetylmuramoyl-L-alanine amidases + CHAP domain, three endo- β -acetylglucosaminidases + lysM domains and one amidase. As for the PGH-coding genes annotated in the *E. faecium* D genome, six genes coding for endo- β -acetylglucosaminidases (some with a lysM domain), four for *N*-acetylmuramoyl-L-alanine amidases, two for amidases with NlpC/p60 regions and one for NlpC/p60 + amidase + endo- β -acetylglucosaminidase domains were identified. Experimentally, several protein bands with bacteriolytic activity against *M. lysodeikticus* were detected in CPPs; three in *E. faecalis* F, with molecular masses of 86, 72 and 62 kDa, and a 74-kDa protein from *E. faecium* D (Fig. 3).

Analysis of virulence potential

A preliminary screening was performed in search of pathogenicity genes in the *Enterococcus* strains characterised. Phenotypic tests related to virulence were also carried out. Although the haemolysin III gene was detected in the genome analysis, β -haemolysis activity was negative on agar blood tests. Biofilm formation, associated with the expression of gene *esp*, was negative ($OD_{620} 0.09$)—a well-structured biofilm should have an $OD_{620} \geq 0.22$, as described by Cebrián et al. (2012). A qualitative evaluation of bacterial aggregation was performed through the adherence to human intestinal HT-29 and Caco-2 cell lines. Microscopic examinations revealed neither bacterial internalisation into the mammalian cell nor cell damage. In contrast, bacterial adherence (localised and aggregative) was observed, similar to both *E. coli* strains used as reference (Fig. 4).

On the other hand, *E. faecium* strain D tested negative for the amplification of genes related to virulence, and none of the genes reported by Orgier and Serror (2008) were found in the genome functional analysis. Nevertheless, some phenotypic tests were performed. This strain tested negative for β -haemolysis and was able to form a non-structured biofilm

Table 3 Genes related to proteolysis in the genomes of *E. faecium* strain D and *E. faecalis* strain F

Proteolytic system	Gene	Activity ^a	<i>E. faecium</i> strain D ^b	<i>E. faecalis</i> strain F ^b
Clp complex (caseinolytic activity) ^{1, 2}	Clp P	ATP-dependent Clp protease proteolytic subunit	1	1
	Clp E	ATP-dependent Clp protease	1	1
	Clp B	Chaperone protein	2	2
	Clp L	ATP-dependent Clp protease	1	1
	Clp X	ATP-dependent Clp protease	1	1
Oligopeptides ABC transport system ^{1, 3}	Opp A	Bacterial extracellular solute-binding proteins	1	3
	Opp B	Binding-protein-dependent transport system inner membrane component	2	3
	Opp C	Transmembrane domain of oligopeptide transport permease	1	2
	Opp D	ABC transporter/oligopeptide and dipeptide transport	2	3
	Opp F	ABC transporter/oligopeptide and dipeptide transport	2	3
Di/tri peptide transport system ^{1, 3}	Dpp A/P	ATP-binding protein. Periplasmic dipeptide transport	–	1
	Dpp B	Dipeptide transport system, permease	–	–
	Dpp C	Dipeptide transport system, permease	1	1
	Dpp D	ATP-binding protein, dipeptide transport (energy coupling)	–	–
	Dpp F	ATP-binding protein, dipeptide transporter	–	–
	Dpp E	ATP-binding protein, dipeptide transporter	1	9
	Dpt T	Di-/tripeptide transporter. Major Facilitator Superfamily	1	2
Endopeptidases ^{1, 3}	Pep E/G	(X)m (X)n	–	–
	Pep O	(X)m (X)n	1	–
	Pep F	(X)m (X)n	1	1
Dipeptidases ^{1, 3}	Pep D	X X	–	–
	Pep V	X X	2	2
Aminopeptidases ^{1, 3}	Pep C	X(X)n	1	1
	Pep N	X(X)n	–	–
	Pep M	Met (X)n	–	–
	Pep A	Glu/Asp (X)n	1	1
	Pcp	PyroGlu (X)n	–	–
Prolinepeptidases ^{1, 3}	Pep X	X-Pro (X)n	–	–
	Pep I	Pro X-(X)n	–	–
	Pep R	Pro X	–	–
	Pep L	Leu (X)n	–	–
	Pep P	X Pro/(X)n	–	–
	Pep Q	X Pro	1	1
Others ¹	Pep T	Tripeptide aminopeptidase	1	2
	Pep B	Oligoendopeptidase	1	1
	Pep S	Aminopeptidase	1	1

^a X represents any amino acid

^b Number of copies of each gene, as identified by functional analysis

¹ Proteolytic systems identified in this work by functional annotation

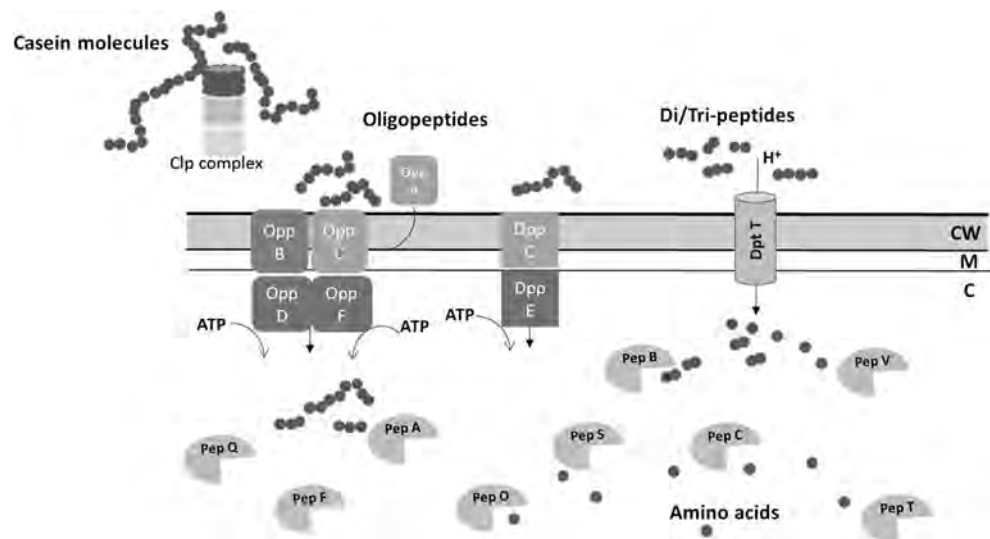
² Park et al. (2015)

³ Liu et al. (2010)

(OD₆₂₀ = 0.21). To note, the strain used as positive control yielded an OD₆₂₀ of 0.274. Cell adhesion tests were not performed with this strain.

In addition to the virulence factors already mentioned, genes involved in the production of biogenic amines, such as ornithine carbamoyltransferase, putrescine carbamoyltransferase

Fig. 1 Hypothetical enterococcal proteolytic system. Based on the scheme of Savijoki et al. (2006) for *Lactococcus lactis*. CW, cell wall; M, membrane; C, cytoplasm



(agmatine deiminase) and tyrosine decarboxylase, were identified in both *Enterococcus* strains.

Assessment of antibiotic resistance, pathogenicity islands, phage sequences and CRISPR regions

According to the protocols established by the Clinical and Laboratory Standards Institute to evaluate antibiotic susceptibility, *E. faecalis* strain F was sensitive to ampicillin, vancomycin, tetracycline and penicillin. However, it was not susceptible to erythromycin, gentamicin, streptomycin and kanamycin, probably because the *macA*, *macB*, *macB2* and *msrA* genes (related to macrolide resistance) and the *scp* gene (associated with aminoglycoside resistance) were detected in its genome. On the other hand, pathogenicity islands were not observed when genomic sequences were compared to the PAIDB. The analysis of prophage sequences revealed five regions with genes encoding proteins such as integrases, transposases, phage tail and tRNA coding regions. Finally, we found two confirmed CRISPR regions plus two that are incomplete; in one of the confirmed CRISPR regions, there were three *cas*-associated genes, namely *cas2*, *cas1* and *cas9*.

E. faecium strain D was sensitive to ampicillin, vancomycin, tetracycline and penicillin. The genomic analysis yielded two antibiotic resistance genes, i.e. *aac* and *msrA*, related to aminoglycoside and macrolide resistance, respectively. Phenotypic tests corroborated *E. faecium* strain D resistance to gentamicin, kanamycin, streptomycin and erythromycin. Although genes related to resistance to other antibiotics were also found, these belonged to complex resistance systems and were incomplete (see Supplemental Table S2 for details), in accordance to the antibiotic sensitivity observed. No pathogenicity islands reported in the PAIDB were found, in agreement with RAST results and as confirmed by a search in our gene functional annotation. In this strain, four prophage regions

with genes coding for transposases, integrases and a lysis protein were identified. The analysis of CRISPR regions showed only one confirmed and three incomplete regions. No *cas* genes were found.

Comparative genomics

The core genome for enterococci species included in this study was established by the comparative analysis of our draft

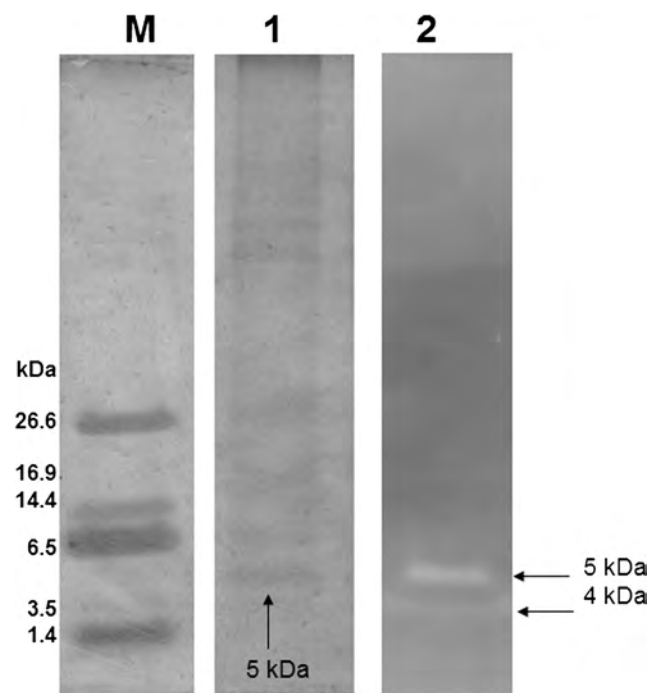
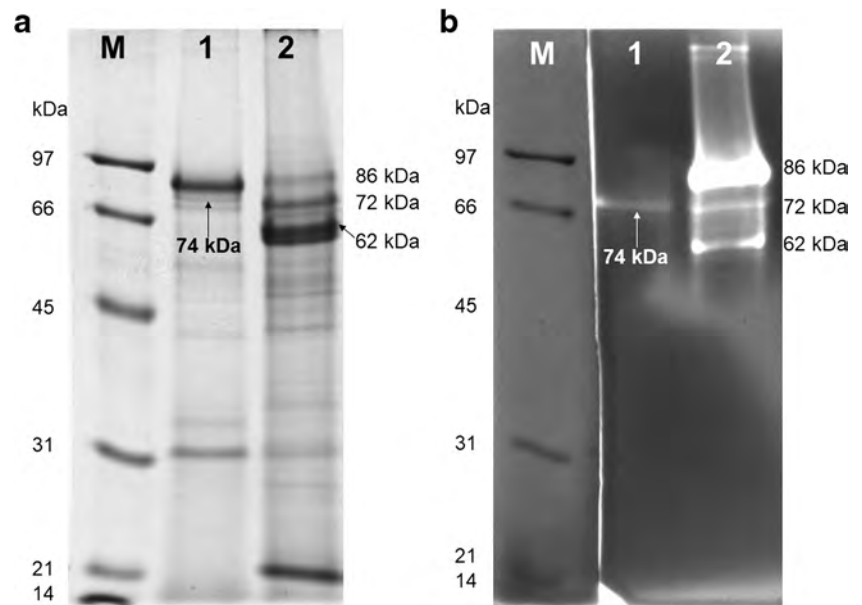


Fig. 2 Protein profile (Tricine-SDS-PAGE 10%) of bacteriocin preparation for *E. faecium* strain D. Lanes: M, molecular weight marker; 1, Coomassie stain; 2, zymogram against *L. innocua*. Total protein per well, 20 mg. Images are in grey scale to improve band visualisation

Fig. 3 **a** Protein profile (SDS-PAGE 10%) of CPPs. M, molecular weight marker. **b** Zymograms against *M. lysodeikticus*. 1, *E. faecium* strain D; 2, *E. faecalis* strain F. Total protein per well 25 mg. Images are in grey scale to improve band visualisation



genomes against complete reference genomes (Table 1), and these have been used commonly as references because they are well characterised, e.g. source, genome functional annotation, epidemiological analysis etc. A total of 4363 protein clusters were obtained for *E. faecalis* F; 2050 were part of the core genome (genes present in all genomes); 2211 belonged to the soft core (present in at least 95% of the input genomes); 1820, to cloud genes (present in at least two genomes); and 336, to shell clusters (present in only the

E. faecalis F strain genome), calculated relative to the smallest genome. The pangenomic matrix analysis allowed constructing a cladogram and obtaining the genes that were unique to each strain. Figure 5a shows the alignment of five *E. faecalis* genomes; the genomes used as reference are listed in Table 1. The circular alignment reflects the differences between strains from various environments: nosocomial, probiotic, healthy infants, animal and Cotija cheese (Fig. 5a, b). The largest genome, belonging to *E. faecalis* strain F, was used as the

Fig. 4 **a** Adherence to Caco-2. **b** Adherence to HT-29. 1, Cells with no bacterial adherence; 2, cells with *E. coli* adherence; and 3, cells with *E. faecalis* strain F adherence

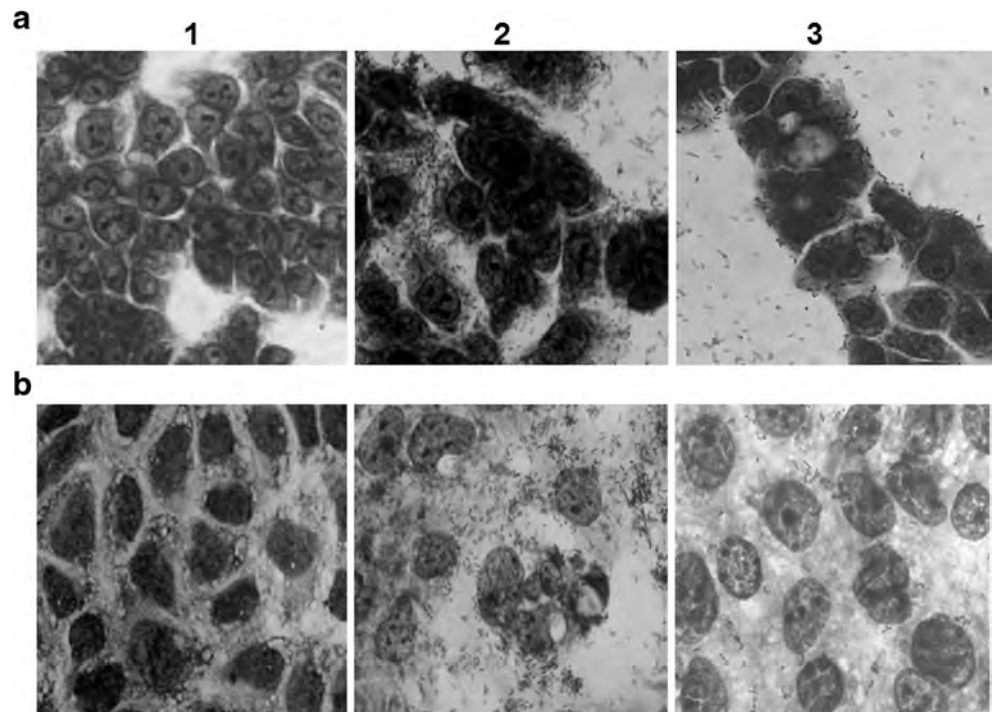
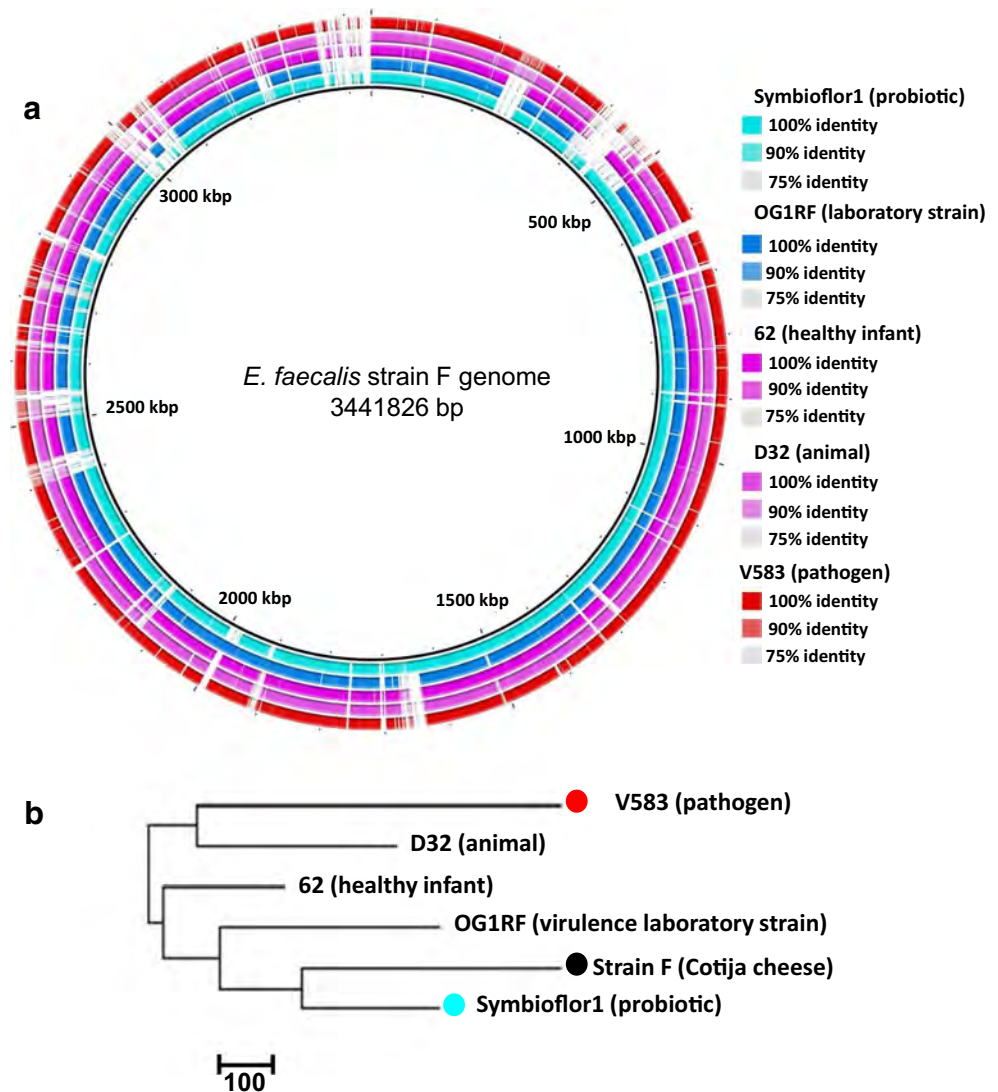


Fig. 5 Alignment of six *E. faecalis* genomes. **a** Circular alignment of *E. faecalis* genomes from different environments against the strain F genome. Blank spaces are genomic zones present only in the strain F genome. The colour scale represents different identity percentages. **b** Parsimony pangenome tree for six *E. faecalis* proteomes derived from presence/absence data in a consensus (OMCL and COGtriangles) pangenome matrix computed from the OMA data set. Tree search performed with PARS from PHYLIP and visualised with MEGA6 (Tamura et al. 2013)



reference and it is represented in the inner ring. The comparison of all protein clusters was used to build a midpoint tree, and this led to obtaining a single clade grouping together Cotija cheese strain F and the probiotic strain Symbioflor1. It is interesting that the pathogenic strain V583 was allocated to a different group (Fig. 5b). The absence/presence matrix revealed 335 genes unique to *E. faecalis* strain F, which were grouped in 62 systems related to the following functions according to KASS and RAST: biosynthesis of amino acids, pentose phosphate pathway, arginine and proline metabolism, terpenoid backbone biosynthesis, protein export, aminoacyl tRNA biosynthesis, DNA repair, ABC transporters, phosphotransferase systems (PTS), two-component system, homologous recombination and defence. The genes with a defence annotation include those encoding the production of ribosomally synthesised antibacterial peptides (bacteriocins).

The *E. faecium* D orthology analysis was performed against four reference closed genomes, several with a virulent phenotype and one belonging to a strain used as a starter

culture in the 1970s, when it was named *Micrococcus freudenreichii*. It was subsequently classified as *E. faecium* and is currently used as a surrogate culture to validate thermal processes in the food industry, being recognised as a non-pathogenic strain (Kopit et al. 2014). The analysis output led to the identification of 1844 protein clusters in the core genome, 2063 in the soft core, 1864 as part of the cloud and 530 in shell clusters. In the genome alignment shown in Fig. 6a, the reference genome in the inner circle belongs to the Aus0004 pathogenic strain because it is the largest of all. The same RAST and KASS platform yielded 368 protein clusters unique to *E. faecium* D. These were represented in 37 systems with the following main functions: glycolysis and gluconeogenesis, carbon, fructose, mannose and galactose metabolisms, amino acid biosynthesis, ABC transporters, PTS, two-component systems and defence (synthesis of bacteriocins). The pangenome comparison showed that both non-pathogenic strains were clustered together in a different clade relative to the pathogenic ones (Fig. 6b).

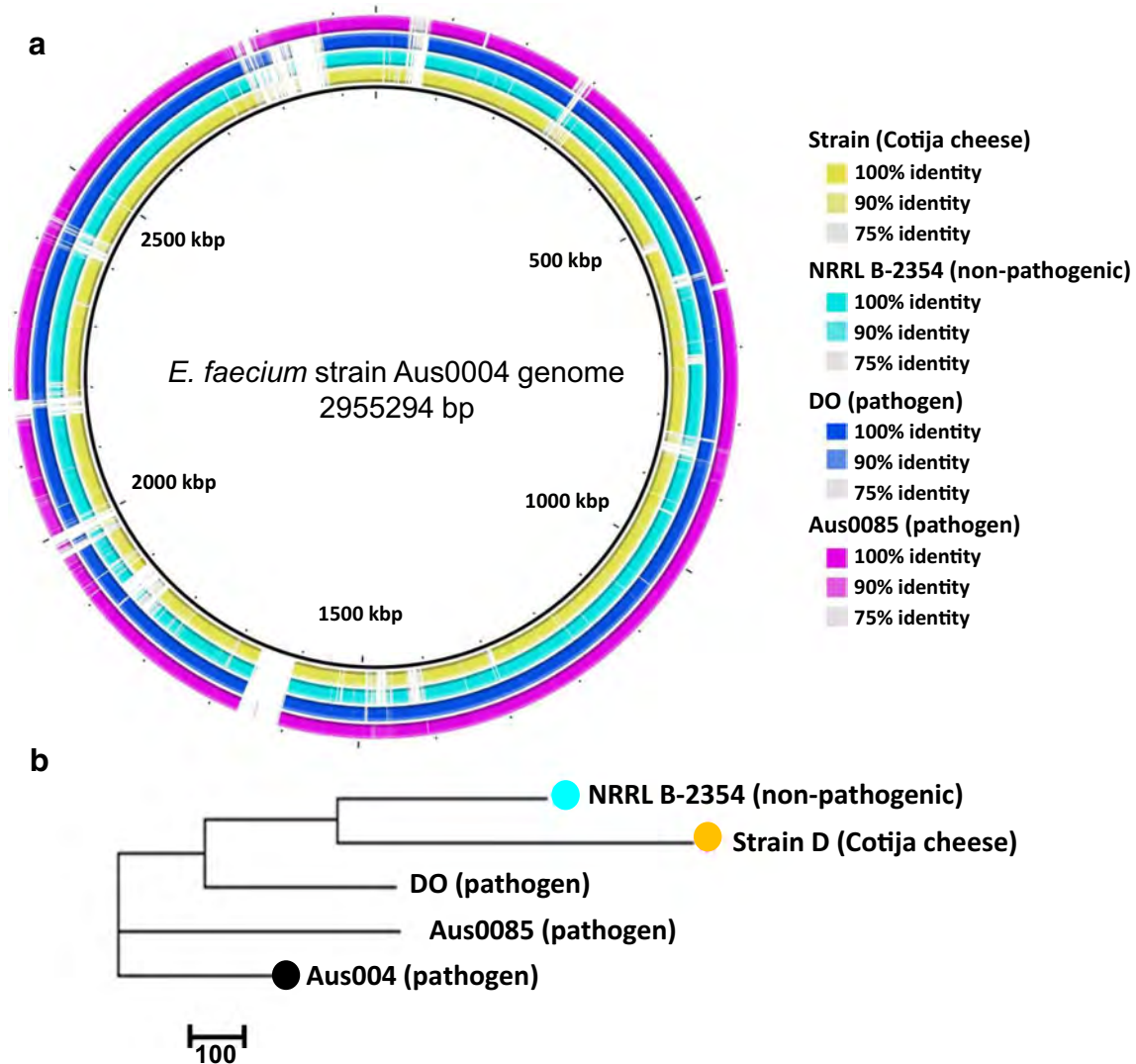


Fig. 6 Alignment of four *E. faecium* genomes. **a** Circular alignment of *E. faecium* genomes from different environments against Aus0004 strain genome. Blank spaces are genomic zones present only in the reference genome. The colour scale represents different identity percentages; the circular alignment was drawn using BRIG 0.95. **b** Parsimony pangenome

tree for five *E. faecium* proteomes derived from presence/absence data in a consensus (OMCL and COGtriangles) pangenome matrix computed from the OMA data set. Tree search performed with PARS from PHYLIP and visualised with MEGA6 (Tamura et al. 2013)

Discussion

Genome sequencing, assembly and annotation

For both strains, *E. faecalis* F and *E. faecium* D, we enhanced the assembly strategy by using several bioinformatic tools to generate an improved high-quality draft genome, as per the criteria established by Chain et al. (2009), and also because 90% of each genome was represented in 38 and 36 contigs for *E. faecalis* and *E. faecium*, respectively. The analysis of the complete genomes of each strain revealed that *E. faecalis* strain F has the largest chromosome of all, and that it harbours two plasmids. In contrast, *E. faecium* strain D has a smaller genome than the pathogenic strains, and it lacks any plasmids,

contrary to the other references (virulent and non-virulent) which have up to six plasmids.

Metabolic activities associated with ripening in cheese

Our interest in the *Enterococcus* genus is related to its role in dairy fermentations. The study of European artisanal matured cheeses has evidenced that enterococci may play a key role in the ripening process. These NSLAB could also contribute to the microbial safety of the product, given that some enterococcal strains produce bacteriocins. However, their role in cheese ripening has not been fully elucidated yet due to the scarce interest in enterococci as a useful culture in food

fermentations. Accordingly, in this work, we searched for genes related to the enzymatic activities associated with aroma, texture and flavour development.

The results of the genome functional annotation revealed that the enterococci strains studied have the caseinolytic protease (Clp) complex (Table 3). It has been reported that during cheese ripening, LABs are under stress conditions and could also be undergoing an autolysis process that, in turn, could release the Clp complex outside the cell (Savijoki et al. 2006). Clp includes proteolytic units (ClpP and ClpL), plus others responsible for recognition (ClpE, ClpX), unfolding (ClpB) and translocation into the multimeric protein structure, which resembles a hollow chamber with a central pore (Frees et al. 2007; Park et al. 2015). Given this evidence, we hereby propose that the Clp complex could be the first step in casein breakdown during cheese ripening. Then, the oligopeptides and peptides produced are needed to be captured and internalised into the cell by transporters, namely the Opp and Dpp systems, and by the cell wall-bound proteinase DtpT, which has been described for LAB. This system is complete in the *Enterococcus* strains studied and includes genes coding for an oligopeptide-binding protein (OppA), two membrane-bound proteins (OppB and OppC) and intracellular ATP-dependent transporter proteins (OppD and OppF) (Savijoki et al. 2006). These strains harbour other dipeptide/tripeptide transporters that function through a proton-motive force (DtpT) and the ATP-driven Dpp system. Although the Dpp system is not complete in the strains studied, it has been reported that Opp can completely replace Dpp and vice versa (Lamarque et al. 2004). Thus, we can assume that the Opp system may perform the functions that the Dpp system is lacking. Intracellular peptidases, like the ones found (Table 3) have been proposed to be the first enzymes to break down oligopeptides into amino acids. The first proteinase from *E. faecalis* var. *liquefaciens* showed a high activity against casein and low against β -lactoglobulin and α -lactalbumin. Assays carried out with milk as substrate demonstrated that *E. faecalis* displayed the highest proteolytic activity relative to other enterococcal species; this activity was dependent on NaCl and calcium concentrations (Foulquié-Moreno et al. 2006). Studies on *Enterococcus* spp. isolated from cheese, e.g. Pecorino Abruzzese, have demonstrated the importance of their aminopeptidase activity because the production of hydrophobic amino acids is related to a debittering effect and it is associated to the development of cheese flavour (Serio et al. 2010).

Lipolysis produces free fatty acids that contribute to the final flavour of food. In both strains, the genome annotation obtained here revealed esterases, lipases and carboxylesterases, the activities of which were experimentally corroborated. Other research groups have demonstrated that enterococci strains can hydrolyse triglycerides, mainly those isolated from food matrices, from

which *E. faecalis* displayed the highest lipolytic activity relative to *E. durans* and *E. faecium*. *E. faecium* showed the highest esterolytic activity among these species, with no difference regarding the environment from which they were isolated (Foulquié-Moreno et al. 2006; Sarantinopoulos et al. 2001; Serio et al. 2010). Free fatty acids are later transformed to methyl ketones, secondary alcohols, esters and lactones, which contribute to the sensory profile of dairy products (van Kranenburg et al. 2002). In cheese fermentations, enterococci have been shown to metabolise citrate and lactose into volatile compounds such as acetaldehyde, diacetyl, acetone, ethanol and acetoin (Sarantinopoulos et al. 2002). In both strains, the annotation obtained revealed some enzymes involved in the production of volatile compounds. These include acetoin, diacetyl and acetaldehyde, which contribute to buttery and rancid butter odour; acetate and acetaldehyde, to pungent odour; and ethanol, which confers fruity and nutty notes, and which has been reported in different ripened cheeses such as Cheddar and Roquefort (Engels et al. 1997; Smit et al. 2005).

Antibacterial activity and safety

As for the production of antibacterial compounds, many enterocins (bacteriocins produced by *Enterococcus*) are well characterised because they are of interest as food biopreservatives and are usually active towards food-borne pathogens such as *Listeria* spp., *Staphylococcus* spp., *Bacillus cereus* and *Clostridium* spp. (Foulquié-Moreno et al. 2006). Based on the classification proposed by Franz et al. (2007), the bacteriocins identified by the annotations obtained here belong to class II (small, non-lantibiotic peptides), mainly subclass II.1 (pediocin family). Furthermore, the YGNGV motif, characteristic of subclass II.1, was identified in genes encoding enterocins A, SE-K4 and P. The gene coding for enterocin B was also found; it belongs to subclass II.3, which include linear, non-pediocin enterocins. Two active bacteriocins were detected in *E. faecium* by zymography. According to their molecular mass, they could be enterocins A and B (Fig. 2), in accordance to the genomic annotation. As for *E. faecalis*, the experimental approach used failed to detect any active peptide; alternative methods for peptide concentration or analysis should be assayed. Enterocins belonging to class II have been identified mainly in strains isolated from food.

Additionally, PGHs have been reported as antibacterial extracellular enzymes with activity against Gram-positive and Gram-negative bacteria (García-Cano et al. 2014). These are required in vital bacterial functions such as septum digestion after cell division, autolysis and peptidoglycan subunit exchange. Bacterial genome sequence analyses have revealed that one genome could harbour as far as 10–30 PGH-

encoding genes with a redundancy that can be explained by the dependence of bacterial growth on the activity thereof (Mesnage et al. 2008). In our strains, we identified up to 15 genes coding for this activity, including two PGHs with molecular masses of 72 and 54 kDa produced by *E. faecalis* strain F as previously described by García-Cano et al. (2014). Zymography allowed the detection of only one or three bands with lytic activity against *M. lysodeikticus*. This finding could be explained by the fact that PGH production depends on the bacterial growth phase and because not all PGHs are equally active against the *M. lysodeikticus* cell wall, even though this bacterium is used as the substrate in the lytic activity detection standard methodology (Lepeuple et al. 1998). In our experience, not all enzymes will be efficiently renatured after TCA precipitation, and thus their activity may not be detected afterwards. The enterococcal strains analysed here were isolated from Cotija cheese, which is made with raw milk using a process that does not involve a heating process. After 3 months of ripening, its microbiota does not include pathogenic bacteria (Escobar-Zepeda et al. 2016) and its microbiological quality is acceptable according to the Mexican official standards. Thus, we hypothesise that enterococci strains could play a key role in product microbial safety resulting from the potential production of various antibacterial compounds.

Although the genus *Enterococcus* belongs to the LAB group, it is not considered as safe (GRAS). Unlike other bacterial pathogens, enterococci may not be intrinsically virulent; instead, they are efficient in collecting and exchanging genes encoding antibiotic resistances. For this reason, this genus has emerged as a nosocomial pathogen. The pathogenic potential in *Enterococcus* is related to the expression of proteins such as cytolysin, adhesion proteins, aggregation substance and the enterococcal surface protein (Esp). These are encoded in the *Enterococcus* pathogenicity islands or in its core genome. Infection models have shown that cytolysins and the aggregation substance or protein Esp are all necessary to cause infection and mortality in rabbits (Garsin et al. 2014). Accordingly, the genes coding for these proteins were searched in this study. Our findings showed that none of these genes were detected in *E. faecium*, and only the gene encoding for the aggregation substance was present in *E. faecalis*. Furthermore, other genes related to virulence were found in this strain by functional annotation, such as the ones coding for gelatinase (*gelE*), adhesion proteins (*ace*, *asa1*) and enzymes related to capsule synthesis. However, the presence of virulence genes does not imply that the harbouring strain should be pathogenic (Garsin et al. 2014). Genes coding for the aggregation substance, the collagen adhesion protein and the proteins needed for capsule formation were detected in the probiotic Symbioflor1 strain. These genes are considered to be advantageous because they confer the capability to colonise and proliferate in the gastrointestinal tract, which provides probiotic attributes (Domann et al. 2007). We observed that

E. faecalis strain F was able to adhere to mammal cells without damaging them; also, both strains were able to form non-structured biofilms, a feature that could be correlated to capsule formation and a potential probiotic behaviour. Although the detection of virulence genes may indicate a likely virulent phenotype in food strains, food-borne enterococcal infections have never been reported (Orgier and Serror 2008). Besides, the incidence of virulence factors is higher in enterococcal isolates from clinical environments than in isolates from food or in starter strains; this suggests that strains from environments other than clinical have a lower pathogenicity potential (Franz et al. 2001).

The production of biogenic amines (BA) by LAB is also a concern in the food industry. Both genome annotations revealed enzymes related to the production of biogenic amines such as ornithine, putrescine and tyramine (Wunderlichová et al. 2014). However, BA tests in several Cotija cheeses in our laboratory have shown concentrations of putrescine, cadaverine and tyramide below 0.1 ppm, a finding that suggests that BA production by enterococcal strains in this particular cheese is likely low.

Although antibiotic resistance is not considered as a virulence trait in enterococci, it is associated with their ability to proliferate in an infection. It has been reported that strains of *E. faecium* and *E. faecalis* isolated from food are commonly resistant to ciprofloxacin, penicillin, kanamycin, erythromycin, chloramphenicol, tetracycline, streptomycin, gentamicin, vancomycin and/or ampicillin (Franz et al. 2001). In contrast, experimental results involving both strains in study revealed that they were resistant to only four of the antibiotics mentioned; these results were consistent with the genes identified by our annotation. It has been proposed that antibiotic resistance is widespread in bacteria proliferating in fermented food ecosystems due to gene exchange between them in these environments. Based on this assumption, fermented foods could be potential reservoirs of genes associated with antibiotic resistance (Orgier and Serror 2008). It is worth considering that virulence is not the result of the expression of a single gene but of multiple genomic determinants, such as genes coding for virulence factors, antibiotic resistance and mobile genetic elements (transposons, phages, insertion sequences, pathogenicity islands and plasmids). These determinants have been proposed as the bacterial adaptation and specialisation strategy to proliferate in different niches (Franz et al. 2011). It has been suggested that CRISPR-cas could provide an immune defence system against viral infections and mobile-element integration in the prokaryotic chromosome (Palmer et al. 2014). At least one CRISPR region was identified in both strains studied; however, only in *E. faecalis* strain F the CRISPR regions were associated with *cas* genes, which in turn are responsible for the degradation of exogenous DNA through nuclease activity. The presence of the CRISPR-cas systems and/or genes encoding restriction enzymes could

explain the absence of plasmids and pathogenicity islands, as well as the low abundance of genes conferring antibiotic resistance. Palmer and Gilmore (2010) have demonstrated that in multi-resistant enterococci strains there is a lack of the CRISPR-cas system, accordingly to our results.

Comparative genomics

Some typing strategies such as multilocus sequence typing (MLST) have been used in order to differentiate enterococcal strains from diverse environments from the ones isolated from clinical infections. However, the wide intra-specific genome diversity had made it difficult to differentiate between them (Palmer et al. 2012). Nevertheless, the comparison of all coding DNA sequences in the genome comparison with the strains selected from different representative environments allowed to group together non-pathogenic strains in the same clade (Figs. 5b and 6b). Strategies based on complete genome sequencing and comparison have differentiated enteric from non-enteric species based on the identification of specific genes associated with adaptation to different environments. For example, operons related to manganese uptake and genes required for sugar utilisation (xylosides and lactose) were over-represented in enteric strains, whereas nickel-uptake genes were representative in *E. faecalis* strains isolated from freshwater environments (Weigand 2014). In addition, class II bacteriocins are usually reported in enterococcal strains from non-clinical environments. On the other hand, class I enterocins, including cytolysin produced by *E. faecalis* and *E. faecium*, are frequently reported in strains isolated from human infections (Nes et al. 2014). Enterococcal genomes were interrogated for genes unique to food strains, and some that could function as biomarkers of adaptation to cheese were noted, mainly genes coding for class II bacteriocins. Clearly, a few genes are not sufficient to determine the niche of a certain enterococcal strain; however, biomarkers could be a useful tool to differentiate between pathogenic and non-pathogenic strains.

Through the study of the metabolic potential obtained by genomic analysis, this work proposes a proteolytic system in *E. faecium* and *E. faecalis* that broadens the knowledge about their role in food fermentation, which may be relevant for the development of aromas and flavours, combined with the lipolytic system that they also possess. In addition, it was demonstrated that these species have the capacity to harbour genes and produce different compounds with antibacterial activity, hence contributing to the microbial safety of the final product. On the other hand, new sequencing technologies and the subsequent bioinformatic analyses are useful tools for unveiling the differences between non-pathogenic and pathogenic strains. Furthermore, these supplement other methods that together could assist in the detailed characterisation of the strains engaged in food fermentation. Because of the ease of

gene transfer between strains and species within the genus *Enterococcus*, their continuous evaluation is required, particularly as regards the presence and extent of antibiotic resistance.

Acknowledgments This work was supported by UNAM-PAPIIT IN222115 and IN222717 grants. We thank Alejandra Escobar-Zepeda, Jérôme Verleyen and Veronica Jiménez-Jacinto for their support in bioinformatic analyses, which were performed in the ‘Unidad de Secuenciación Masiva y Bioinformática’ at the ‘Laboratorio Nacional de Apoyo Tecnológico a las Ciencias Genómicas’, CONACyT #260481, Instituto de Biotecnología/UNAM; to Dr. Edith Ponce Alquicira, from Universidad Autónoma Metropolitana-Iztapalapa (UAM-I) for providing the *Enterococcus faecium* strain MXVK29 and *Enterococcus faecalis* NCIMB-700585; and to Dr. Carlos Eslava-Campos, for providing mammal cell lines and for his assistance in performing the adherence experiments. M.O.-G. was granted a doctoral fellowship from Consejo Nacional de Ciencia y Tecnología (CONACyT). We also thank Dr. Romina Rodríguez Sanoja and Dr. Rosa María Gutierrez for their valuable comments and support during the development of this project. María Elena Sánchez-Salazar edited the English manuscript.

Compliance with ethical standards

Ethical statement This manuscript is in compliance with ethical standards. This manuscript does not contain any studies with human participants or animals performed by any of the authors.

Conflict of interest The authors declare that they have no conflict of interest.

References

- Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA (2011) BLAST ring image generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* 12(1):402. <https://doi.org/10.1186/1471-2164-12-402>
- Alvarez-Cisneros YM, Fernández FJ, Wachter-Rodarte C, Aguilar MB, Sáinz Espuñes T del R, Ponce-Alquicira E (2010) Biochemical characterization of a bacteriocin-like inhibitory substance produced by *Enterococcus faecium* MXVK29, isolated from Mexican traditional sausage. *J Sci Food Agric* 90:2475–2481. doi: <https://doi.org/10.1002/jsfa.4109>, 14
- Andrews S (2016) Babraham Bioinformatics e FastQC a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>. Accessed 9 Sep 2016
- Aziz RK, Bartels D, Best AA, De Jongh M, Disz T, Edwards RA, Formisano K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O (2008) The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9(1):75. <https://doi.org/10.1186/1471-2164-9-75>
- Bender JK, Fiedler S, Klare I, Werner G (2015) Complete genome sequence of the gut commensal and laboratory strain *Enterococcus faecium* 64/3. *Genome Announc* 3(6):e01275–e01215. <https://doi.org/10.1128/genomeA.01275-15>
- Besemer J, Lomsadze A, Borodovsky M (2001) Prokaryotic GeneMark.Hmm version 2 GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for

- finding sequence motifs in regulatory regions. *Nucleic Acids Res* 29(12):2607–2618. <https://doi.org/10.1093/nar/29.12.2607>
- Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W (2011) Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27(4):578–579. <https://doi.org/10.1093/bioinformatics/btq683>
- Bourgogne A, Garsin D (2008) Large scale variation in *Enterococcus faecalis* illustrated by the genome analysis of strain OG1RF. *Genome Biol* 9(7):R110. <https://doi.org/10.1186/gb-2008-9-7-r110>
- Brown CT, Howe A, Zhang Q, Pyrkosz AB, Brom TH (2012) A reference-free algorithm for computational normalization of shotgun sequencing data. <http://arxiv.org/abs/1203.4802>. Accessed 13 April 2016; arXiv:1203.4802
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. *BMC Bioinform* 10(1):421. <https://doi.org/10.1186/1471-2105-10-421>
- Carattoli A, Zankari E, Garcia-Fernandez A, Volby Larsen M, Lund O, Villa L, Aarestrup FM, Hasman H (2014) In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob Agents Chemother* 58(7):3895–3903. <https://doi.org/10.1128/AAC.02412-14>
- Cebrián R, Baños A, Valdivia E, Pérez-Pulido R, Martínez-Bueno M, Maqueda M (2012) Characterization of functional, safety, and probiotic properties of *Enterococcus faecalis* UGRA10, a new AS-48-producer strain. *Food Microbiol* 30(1):59–67. <https://doi.org/10.1016/j.fm.2011.12.002>
- Chain PS, Grafham DV, Fulton RS, Fitzgerald MG, Hostetler J, Muzny D, Ali J, Birren B, Bruce DC, Buhay C, Cole JR, Ding Y, Dugan S, Field D, Garrity GM, Gibbs R, Graves T, Han CS, Harrison SH, Highlander S, Hugenholtz P, Khouri HM, Kodira CD, Kolker E, Kyrpides NC, Lang D, Lapidus A, Malfatti SA, Markowitz V, Metha T, Nelson KE, Parkhill J, Pitluck S, Qin X, Read TD, Schmutz J, Sozhamannan S, Sterk P, Strausberg RL, Sutton G, Thomson NR, Tiedje JM, Weinstock G, Wollam A, Genomic Standards Consortium Human Microbiome Project Jumpstart Consortium, Dettler JC (2009) Genomics. Genome project standards in a new era of sequencing. *Science* 326(5950):236–237. <https://doi.org/10.1126/science.1180614>
- Clinical and Laboratory Standards Institute (2012) Performance standards for antimicrobial susceptibility testing; twenty-second informational supplement. CLSI document M100-S22. Clinical and Laboratory Standards Institute, Wayne
- Contreras-Moreira B, Vinuesa P (2013) GET_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. *Appl Environ Microbiol* 79(24):7696–7701. <https://doi.org/10.1128/AEM.02411-13>
- Dag Anders B, Lars G (2011) Complete genome sequence of the commensal *Enterococcus faecalis* 62, isolated from a healthy Norwegian infant. *J Bacteriol* 193(9):2377–2378. <https://doi.org/10.1128/JB.00183-11>
- Darling AE, Mau B, Perna NT (2010) Progressive mauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 5(6):e11147. <https://doi.org/10.1371/journal.pone.0011147>
- Domann E, Hain T, Ghai R, Billion A, Kuenne C, Zimmermann K, Chakraborty T (2007) Comparative genomic analysis for the presence of potential enterococcal virulence factors in the probiotic *Enterococcus faecalis* strain Symbioflor 1. *Int J Med Microbiol* 297(7-8):533–539. <https://doi.org/10.1016/j.ijmm.2007.02.008>
- Eddy SR (2011) Accelerated profile HMM searches. *PLoS Comput Biol* 7(10):e1002195. <https://doi.org/10.1371/journal.pcbi.1002195>
- Engels WJM, Dekker R, de Jong C, Neeter R, Visser S (1997) A comparative study of volatile compounds in the water-soluble fraction of various types of ripened cheese. *Int Dairy J* 7(4):255–263. [https://doi.org/10.1016/S0958-6946\(97\)00003-4](https://doi.org/10.1016/S0958-6946(97)00003-4)
- Escobar-Zepeda A, Sanchez-Flores A, Quirasco Baruch M (2016) Metagenomic analysis of a Mexican ripened cheese reveals a unique complex microbiota. *Food Microbiol* 57:116–127. <https://doi.org/10.1016/j.fm.2016.02.004>
- Foulquié-Moreno MR, Sarantinopoulos P, Tsakalidou E, De Vuyst L (2006) The role and application of enterococci in food and health. *Int J Food Microbiol* 106(1):1–24. <https://doi.org/10.1016/j.ijfoodmicro.2005.06.026>
- Franz CM, Muscholl-Silberhorn AB, Yousif NMK, Vancanneyt M, Swings J, Holzapfel WH (2001) Incidence of virulent factors and antibiotic resistance among enterococci isolated from food. *Appl Environ Microbiol* 67(9):4385–4389. <https://doi.org/10.1128/AEM.67.9.4385-4389.2001>
- Franz CM, van Belkum MJ, Holzapfel WH, Abriouel H, Gálvez A (2007) Diversity of enterococcal bacteriocins and their grouping in a new classification scheme. *FEMS Microbiol Rev* 31(3):293–310. <https://doi.org/10.1111/j.1574-6976.2007.00064.x>
- Franz CM, Huch M, Abriouel H, Holzapfel W, Gálvez A (2011) Enterococci as probiotics and their implications in food safety. *Int J Food Microbiol* 151(2):125–140. <https://doi.org/10.1016/j.ijfoodmicro.2011.08.014>
- Frees D, Savijoki K, Varmanen P, Ingmer H (2007) Clp ATPases and ClpP proteolytic complexes regulate vital biological processes in low GC, gram-positive bacteria. *Mol Microbiol* 63(5):1285–1295. <https://doi.org/10.1111/j.1365-2958.2007.05598.x>
- Fritzenwanker M, Kuenne C (2013) Complete genome sequence of the probiotic *Enterococcus faecalis* Symbioflor 1 clone DSM 16431. *Genome Announc* 1(1):e00165–e00112. <https://doi.org/10.1128/genomeA.00165-12>
- García-Cano I, Serrano-Maldonado CE, Olvera-García M, Delgado Arciniega E, Peña-Montes C, Quirasco M (2014) Antibacterial activity produced by *Enterococcus* spp. isolated from artisanal Mexican dairy product, Cotija cheese. *LWT Food Sci Technol* 59(1):26–34. <https://doi.org/10.1016/j.lwt.2014.04.059>
- Garsin DA, Frank KL, Silanpää J, Ausubel FM, Hartke A, Shankar N, Murray BE (2014) Pathogenesis and models of enterococcal infection. In: Gilmore MS, Clewell DB, Ike Y, Shankar N (eds) *Enterococci: from commensals to leading causes of drug resistant infection* [internet]. Eye and Ear Infirmary, Boston, pp 135–169
- Grissa I, Vergnaud G, Pourcel C (2007) CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res* 35(Web Server):W52–W57. <https://doi.org/10.1093/nar/gkm360>
- Hunt M, Kikuchi T, Sanders M, Newbold C, Berriman M, Otto TD (2013) REAPR: a universal tool for genome assembly evaluation. *Genome Biol* 14(5):R47. <https://doi.org/10.1186/gb-2013-14-5-r47>
- Jerse AE, Yu J, Tall BD, Kaper JB (1990) A genetic locus of enteropathogenic *Escherichia coli* necessary for the production of attaching and effacing lesions on tissue culture cells. *Proc Natl Acad Sci U S A* 87(20):7839–7843. <https://doi.org/10.1073/pnas.87.20.7839>
- Kim EB, Marco ML (2014) Nonclinical and clinical *Enterococcus faecium* strains, but not *Enterococcus faecalis* strains, have distinct structural and functional genomic features. *Appl Environ Microbiol* 80(1):154–165. <https://doi.org/10.1128/AEM.03108-13>
- Kopit LM, Kim EB, Siezen RJ, Harris LJ, Marco ML (2014) Safety of the surrogate microorganism *Enterococcus faecium* NRRL B-2354 for use in thermal process validation. *Appl Environ Microbiol* 80(6):1899–1909. <https://doi.org/10.1128/AEM.03859-13>
- Krogh A, Larsson B, von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305(3):567–580. <https://doi.org/10.1006/jmbi.2000.4315>
- Lam MM, Seemann T, Bulach DM, Gladman SL, Chen H, Haring V, Moore RJ, Ballard S, Grayson ML, Johnson PD, Howden BP, Stinear TP (2012) Comparative analysis of the first complete *Enterococcus faecium* genome. *J Bacteriol* 194(9):2334–2341. <https://doi.org/10.1128/JB.00259-12>

- Lam MM, Seeman T, Tobias NJ, Chen H, Haring V, Moore RJ, Ballard S, Grayson LM, Johnson PD, Howden BP, Stinear TP (2013) Comparative analysis of the complete genome of an epidemic hospital sequence type 203 clone of vancomycin-resistant *Enterococcus faecium*. *BMC Genomics* 14(1):595. <https://doi.org/10.1186/1471-2164-14-595>
- Lamarque M, Charbonnel P, Aubel D, Piard JC, Atlan D, Juillard V (2004) A multifunction ABC transporter (opt) contributes to diversity of peptide uptake specificity within the genus *Lactococcus*. *J Bacteriol* 186(19):6492–6500. <https://doi.org/10.1128/JB.186.19.6492-6500.2004>
- Lepeuple AS, van Gemert E, Chapot-Chartier MP (1998) Analysis of the bacteriolytic enzymes of the autolytic *Lactococcus lactis* subsp. *cremoris* strain AM2 by renaturing polyacrylamide gel electrophoresis: identification of a prophage-encoded enzyme. *Appl Environ Microbiol* 64:4142–4148
- Liu M, Bayjanov JR, Renckens B, Nauta A, Siezen RJ (2010) The proteolytic system of lactic acid bacteria revisited: a genomic comparison. *BMC Genomics* 11(1):36. <https://doi.org/10.1186/1471-2164-11-36>
- Mesnager S, Chau F, Dubost L, Arthur M (2008) Role of N-acetylglucosaminidase and N-acetylmuramidase activities in *Enterococcus faecalis* peptidoglycan metabolism. *J Biol Chem* 283(28):19845–19853. <https://doi.org/10.1074/jbc.M802323200>
- Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M (2007) KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res* 35(Web Server):W182–W185. <https://doi.org/10.1093/nar/gkm321>
- Nadalin F, Vezzi F, Policriti A (2012) GapFiller: a de novo assembly approach to fill the gap within paired reads. *BMC Bioinf* 14(Suppl 14):S8. <https://doi.org/10.1186/1471-2105-13-S14-S8>
- Natarajan P, Parani M (2015) First complete genome sequence of a probiotic *Enterococcus faecium* strain T-110 and its comparative genomic analysis with pathogenic and non-pathogenic *Enterococcus faecium* genomes. *J Genet Genomics* 42(1):43–46. <https://doi.org/10.1016/j.jgg.2014.07.002>
- Navarro-García F, Eslava C, Villaseca JM, López-Revilla R, Czczulin JR, Srinivas S, Nataro JP, Cravioto A (1998) In vitro effects of a high-molecular-weight heat-labile enterotoxin from enteroaggregative *Escherichia coli*. *Infect Immun* 66(7):3149–3154
- Nes IF, Diep DB, Ike Y (2014) Enterococcal bacteriocins and antimicrobial proteins that contribute to niche control. In: Gilmore MS, Clewell DB, Ike Y, Shankar N (eds) *Enterococci: from commensals to leading causes of drug resistant infection* [internet]. Eye and Ear Infirmary, Boston, pp 446–482
- NMX-F-735-COFOCALEC (2011) Mexican Official Standard. Sistema productivo leche-alimentos-lácteos-alimento lácteo regional-queso Cotija artesanal madurado-denominación, especificaciones y métodos de prueba. DOF: 24/06/2011, Diario Oficial de la Federación, Mexico City
- Nueno-Palop C, Narbad A (2011) Probiotic assessment of *Enterococcus faecalis* CP58 isolated from human gut. *Int J Food Microbiol* 145(2-3):390–394. <https://doi.org/10.1016/j.ijfoodmicro.2010.12.029>
- Orgier JC, Serror P (2008) Safety assessment of dairy microorganisms: the *Enterococcus* genus. *Int J Food Microbiol* 126(3):291–301. <https://doi.org/10.1016/j.ijfoodmicro.2007.08.017>
- Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R (2014) The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). *Nucleic Acids Res* 42(D1):D206–D214. <https://doi.org/10.1093/nar/gkt1226>
- Palmer KL, Gilmore MS (2010) Multidrug-resistant enterococci lack CRISPR-cas. *mBio* 1(4):e00227-10. <https://doi.org/10.1128/mBio.00227-10>
- Palmer KL, Godfrey P, Griggs A, Kos VN, Zucker J, Desjardins C, Cerqueira G, Gevers D, Walker S, Wortman J, Feldgarden M, Haas B, Birren B, Gilmore MS (2012) Comparative genomics of enterococci: variation in *Enterococcus faecalis*, clade structure in *E. faecium*, and defining characteristics of *E. gallinarum* and *E. casseliflavus*. *mBio* 3:e00318-11. <https://doi.org/10.1128/mBio.00318-11>
- Palmer KL, van Schaik W, Willems RJL, Gilmore MS (2014) Enterococcal genomics. In: Gilmore MS, Clewell DB, Ike Y, Shankar N (eds) *Enterococci: from commensals to leading causes of drug resistant infection* [internet]. Eye and Ear Infirmary, Boston, pp 189–216
- Park SS, Kwon HY, Tran TD, Choi MH, Jung SH, Lee S, Briles DE, Rhee DK (2015) ClpL is a chaperone without auxiliary factors. *FEBS J* 282(8):1352–1367. <https://doi.org/10.1111/febs.13228>
- Paulsen IT, Banerjee L, Myers GS, Nelson KE, Seshadri R, Read TD, Fouts DE, Eisen JA, Gill SR, Heidelberg JF, Tettelin H, Dodson RJ, Umayam L, Brinkac L, Beanan M, Daugherty S, DeBoy RT, Durkin S, Kolonay J, Madupu R, Nelson W, Vamathevan J, Tran B, Upton J, Hansen T, Shetty J, Khouri H, Utterback T, Radune D, Ketchum KA, Dougherty BA, Fraser CM (2003) Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. *Science* 299(5615):2071–2074. <https://doi.org/10.1126/science.1080613>
- Petersen TN, Brunak S, von Heijne G, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 8(10):785–786. <https://doi.org/10.1038/nmeth.1701>
- Ponstingl H, Ning Z (2010) SMALT. <http://www.sanger.ac.uk/science/tools/smalt-0>. Accessed 13 April 2016
- Qin X, Galloway-Peña JR, Sillanpää J, Roh JH, Nallapareddy SR, Chowdhury S, Bourgogne A, Choudhury T, Muzny DM, Buhay CJ, Ding Y, Dugan-Rocha S, Liu W, Kovar C, Sodergren E, Highlander S, Petrosino JF, Worley KC, Gibbs RA, Weinstock GM, Murray BE (2012) Complete genome sequence of *Enterococcus faecium* strain TX16 and comparative genomic analysis of *Enterococcus faecium* genomes. *BMC Microbiol* 12(1):135. <https://doi.org/10.1186/1471-2180-12-135>
- Sarantinopoulos P, Andrighetto C, Georgalakis MD, Rea MC, Lombardi A, Coganc TM, Kalantzopoulou G, Tsakalidou E (2001) Biochemical properties of enterococci relevant to their technological performance. *Int Dairy J* 11(8):621–647. [https://doi.org/10.1016/S0958-6946\(01\)00087-5](https://doi.org/10.1016/S0958-6946(01)00087-5)
- Sarantinopoulos P, Kalantzopoulos G, Tsakalidou E (2002) Effect of *Enterococcus faecium* on microbiological, physicochemical and sensory characteristics of Greek feta cheese. *Int J Food Microbiol* 76(1-2):93–105. [https://doi.org/10.1016/S0168-1605\(02\)00021-1](https://doi.org/10.1016/S0168-1605(02)00021-1)
- Savijoki K, Ingmer H, Varmanen P (2006) Proteolytic systems of lactic acid bacteria. *Appl Microbiol Biotechnol* 71(4):394–406. <https://doi.org/10.1007/s00253-006-0427-1>
- Schägger H, von Jagow G (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in range from 1 to 100 kDa. *Anal Biochem* 166(2):368–379. [https://doi.org/10.1016/0003-2697\(87\)90587-2](https://doi.org/10.1016/0003-2697(87)90587-2)
- Serio A, Chaves-López C, Paparella A, Suzzi G (2010) Evaluation of metabolic activities of enterococci isolated from pecorino Abruzzese cheese. *Int Dairy J* 20(7):459–464. <https://doi.org/10.1016/j.idairyj.2010.02.005>
- Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJ, Birol I (2009) ABySS: a parallel assembler for short read sequence data. *Genome Res* 19(6):1117–1123. <https://doi.org/10.1101/gr.089532.108>
- Smit G, Smit BA, Engels WJM (2005) Flavour formation by lactic acid bacteria and biochemical flavour profiling of cheese products. *FEMS Microbiol Rev* 29(3):591–610. <https://doi.org/10.1016/j.femsre.2005.04.002>
- Tamura K, Stecher G, Peterson D, Filipowski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 12(12):2725–2729. <https://doi.org/10.1093/molbev/mst197>
- Uchimura Y, Wyss M, Brugiroux S, Limenitakis JP, Stecher B, McCoy KD, Macpherson AJ (2016) Complete genome sequences of 12

- species of stable defined moderately diverse mouse microbiota 2. *Genome Announc* 4(5):e00951–e00916. <https://doi.org/10.1128/genomeA.00951-16>
- van Heel AJ, de Jong A, Montalbán-López M, Kok J, Kuipers OP (2013) BAGEL3: automated identification of genes encoding bacteriocins and (non-)bactericidal post translationally modified peptides. *Nucleic Acids Res* 41(W1):W448–W453. <https://doi.org/10.1093/nar/gkt391>
- van Kranenburg R, Kleerebezem M, van HylckamaVlieg J, Ursing BM, Boekhorst J, Smit BA, Ayad EHE, Smit G, Siezen RJ (2002) Flavour formation from amino acids by lactic acid bacteria: predictions from genome sequence analysis. *Int Dairy J* 12(2-3):111–121. [https://doi.org/10.1016/S0958-6946\(01\)00132-7](https://doi.org/10.1016/S0958-6946(01)00132-7)
- Weigand MR, Ashbolt NJ, Konstantinidis KT, Santo Domingo JW (2014) Genome sequencing reveals the environmental origin of enterococci and potential biomarkers for water quality monitoring. *Environ Sci Technol* 48(7):3707–3714. <https://doi.org/10.1021/es4054835>
- Wunderlichová L, Bunková L, Koutný M, Jancová P, Bunka F (2014) Formation, degradation, and detoxification of putrescine by foodborne bacteria: a review. *Compr Rev Food Sci Food Saf* 13(5):1012–1030. <https://doi.org/10.1111/1541-4337.12099>
- Yoon SH, Park YK, Kim JF (2014) PAIDB v2.0: exploration and analysis of pathogenicity and resistance islands. *Nucleic Acids Res* 43(D1):D624–D630. <https://doi.org/10.1093/nar/gku985>
- Yuehua J, Lanwei Z, Fei L, Huaxi Y, Xue H (2016) Complete genome sequence of *Enterococcus faecalis* LD33, a bacteriocin-producing strain. *J Biotechnol* 227:79–80. <https://doi.org/10.1016/j.jbiotec.2016.04.030>
- Zerbino DR, Birney E (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 18(5):821–829. <https://doi.org/10.1101/gr.074492.107>
- Zischka M, Kuenne C (2012) Complete genome sequence of the porcine isolate *Enterococcus faecalis* D32. *J Bacteriol* 194(19):5490–5491. <https://doi.org/10.1128/JB.01298-12>
- Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS (2011) PHAST: a fast phage search tool. *Nucleic Acid Res* 39(suppl):W347–W352. <https://doi.org/10.1093/nar/gkr485>

Detección de Proteínas con Actividad Antibacteriana Producidas por Bacterias Ácido Lácticas

Myrna Olvera-García, Carlos Eduardo Serrano-Maldonado y Maricarmen Quirasco*
*L-312, Conjunto E, Departamento de Alimentos y Biotecnología, Facultad de Química,
Universidad Nacional Autónoma de México, 04510 Coyoacán, D.F., México.*

E-mail: quirabma@unam.mx

RESUMEN

Las bacterias ácido lácticas se han empleado históricamente en la producción de alimentos y bebidas fermentadas, como agentes de bioconservación. Tal función se atribuye a la producción de compuestos antibacterianos, como pueden ser los ácidos orgánicos o incluso compuestos de naturaleza proteínica, como las bacteriocinas y las enzimas que degradan la pared celular de bacterias (p. ej., peptidoglucano hidrolasas). Dichas proteínas pueden ser producidas por la bacteria en el alimento directamente durante la fermentación, o se pueden adicionar en su forma purificada. Su potencial biotecnológico ha ocasionado que en varios grupos de investigación exista el interés de caracterizar a las proteínas responsables de dicho efecto antagónico. En este trabajo se revisarán algunas de las estrategias para su purificación, identificación y caracterización bioquímica, ya que frecuentemente se producen en baja concentración, por lo que son difíciles de detectar y de analizar. También es común que la actividad antibacteriana sólo se atribuya al alguno de los dos tipos de moléculas, cuando existe la posibilidad de que la bacteria ácido láctica presente ambos tipos de actividades.

Palabras clave: *bacterias ácido lácticas, actividad antibacteriana, peptidoglucano hidrolasas, bacteriocinas, zimogramas.*

ABSTRACT

Lactic acid bacteria have been historically used in the production of fermented foods and beverages, due to their function as biopreservative agents. That role is attributed to the production of antibacterial compounds, such as organic acids or even to proteinaceous compounds, like bacteriocins and bacterial cell wall-degrading enzymes (i.e. peptidoglycan hydrolases). Those proteins could be produced directly in the food matrix during the fermentation process, or they can be added in a purified form. Several research groups are interested in characterizing the proteins responsible of this antagonistic effect, because of their biotechnological potential. In this work, some strategies for their purification, identification and biochemical characterization will be reviewed, considering that frequently they are produced in low concentrations, which makes their

detection and analysis difficult. It is also common that the antibacterial activity would be attributed only to one of those sort of molecules, when there is the possibility that the lactic acid bacteria would show both kind of activities.

Key words: lactic acid bacteria, antibacterial activity, peptidoglycan hydrolases, bacteriocins, zymograms.

INTRODUCCIÓN

Las bacterias ácido lácticas (BAL) son un grupo conformado por cocos o bacilos Gram positivos, no esporulados, catalasa negativos, anaerobios facultativos, que producen ácido láctico como producto final mayoritario de la fermentación de carbohidratos. Los principales géneros que conforman este grupo son: *Aerococcus*, *Carnobacterium*, *Enterococcus*, *Lactobacillus*, *Lactococcus*, *Leuconostoc*, *Oenococcus*, *Pediococcus*, *Streptococcus*, *Tetragenococcus*, *Vagococcus* y *Weissella* (Salminen *et al.*, 2004).

Las BAL se han utilizado en la elaboración de alimentos fermentados, dentro de los que se encuentran: productos vegetales (aceitunas, pepinillos), lácteos (yogurt, algunos quesos) y cárnicos (salami, ciertos tipos de salchichas y de chorizos). Algunas cepas de BAL aportan beneficios a la salud, en su uso como probióticos, en la producción de péptidos bioactivos y como vehículos transportadores de moléculas con valor terapéutico (Nouaille *et al.*, 2003). Pero sin duda, la aplicación más importante es su uso como cultivos iniciadores en varios productos lácteos fermentados (Savijoki *et al.*, 2006), ya que la utilización de los azúcares presentes en el alimento lleva a la producción de ácidos orgánicos y a un descenso en el pH, importante en el fenómeno de la

coagulación de las proteínas y en la reducción o prevención del crecimiento de la microbiota acompañante. Por otra parte, su presencia puede tener un efecto en el desarrollo de características sensoriales deseables. Por ejemplo, participan en la hidrólisis de proteínas, contribuyendo a la textura y sabor final de diferentes alimentos fermentados; sintetizan compuestos que dan sabor y aroma, como ácidos orgánicos, acetaldehído, diacetilo, acetoina, etc.; y pueden producir agentes texturizantes, como exopolisacáridos, los cuales aportan consistencia al producto. Adicionalmente, pueden producir componentes inhibitorios de otros microorganismos, dentro de los que se pueden encontrar patógenos y de descomposición (Salminen *et al.*, 2004).

La mayoría de las BAL no representan un riesgo para los seres humanos. Han sido utilizadas consciente e inconscientemente desde hace siglos en procesos de elaboración de alimentos sin causar daños a los consumidores. Poseen el estatus de GRAS (Generally Recognised As Safe) por la FDA (Food and Drug Administration) en Estados Unidos y de QPS (Qualified Presumption of Safety) por la Unión Europea (Peterbauer *et al.*, 2011). Como se mencionó, este grupo de bacterias tiene la capacidad de producir una gran variedad de sustancias que pueden inhibir

inhibir a otros microorganismos y, aunado a su grado alimenticio, se han utilizado como agentes de bioconservación con el fin de contribuir a la inocuidad de los alimentos y/o de incrementar la vida de anaquel de los mismos. Su aplicación puede hacerse inoculando la bacteria para que produzca el agente antibacteriano directamente en el alimento; o usando dicho agente, previamente purificado, como aditivo alimentario (Elsser-Gravesen & Elsser-Gravesen, 2014).

PROTEÍNAS Y PÉPTIDOS CON ACTIVIDAD ANTIBACTERIANA PRODUCIDOS POR LAS BAL

Además de la producción de ácido láctico u otros ácidos orgánicos (acético y propiónico), las BAL pueden producir compuestos de diferente naturaleza con actividad antibacteriana. Algunos de esos compuestos son: peróxido de hidrógeno, dióxido de carbono, diacetilo, reuterina, reuteriicina, bacteriocinas, BLIS (Bacteriocin Like Inhibitor Substances) y peptidoglucano hidrolasas (Moreno *et al.*, 2002; Kang *et al.*, 2003; Salminen *et al.*, 2004). De los cuales, los últimos tres son de naturaleza proteínica.

Las peptidoglucano hidrolasas (PGHs) son enzimas que hidrolizan el peptidoglucano, principal componente de la pared celular de bacterias. El peptidoglucano consiste en cadenas de residuos alternados de N-acetilglucosamina (GlcNAc) y ácido N-acetilmurámico (MurNAc) unidos mediante enlaces β -1,4, que se entrecruzan por péptidos cortos formados por L- y D-aminoácidos. Las PGHs están involucradas en un gran número de funciones que requieren la modificación de la

red rígida del peptidoglucano durante el crecimiento y la división celular. Algunas de estas funciones incluyen la regulación del crecimiento de la pared celular, el intercambio de unidades de peptidoglucano durante el crecimiento, la separación de las células hijas durante la división y la autólisis, que generalmente se induce en condiciones adversas como la falta de nutrientes (Lortal & Chapot-Chartier, 2005; Eckert *et al.*, 2006; Vollmer *et al.*, 2008). Tienen pesos moleculares que pueden ir desde los 27 kDa de la lisostafina, hasta los 137 kDa de la proteína Atl, la principal autolisina de *Staphylococcus aureus* y *S. epidermidis* (Kumar, 2008; Götz *et al.*, 2014). Las PGHs se han clasificado de acuerdo con el tipo de enlace que son capaces de romper en el peptidoglucano (Fig. 1). Las N-acetilmuramoyl-L-alanina amidasas rompen el enlace amida entre MurNAc y la L-alanina del péptido. Las peptidasas son capaces de hidrolizar el último aminoácido del extremo carboxilo de los péptidos (carboxipeptidasas), o de romper completamente los puentes formados por los péptidos (endopeptidasas). Las N-acetilglucosaminidasas y las N-acetilmuramididasas hidrolizan los enlaces β -1,4 de la cadena de glicanos. Las primeras lo hacen dejando un extremo GlcNAc reductor, mientras que las segundas hidrolizan el enlace entre MurNAc y GlcNAc, pudiendo dejar libre un extremo MurNAc reductor, llamadas lisozimas, o formando un anillo 1,6-anhidro en MurNAc, conocidas como transglicosilasas líticas (Vollmer *et al.*, 2008; Layec *et al.*, 2008). Algunas PGHs han sido utilizadas como agentes antibacterianos por ejemplo, la lisostafina que es

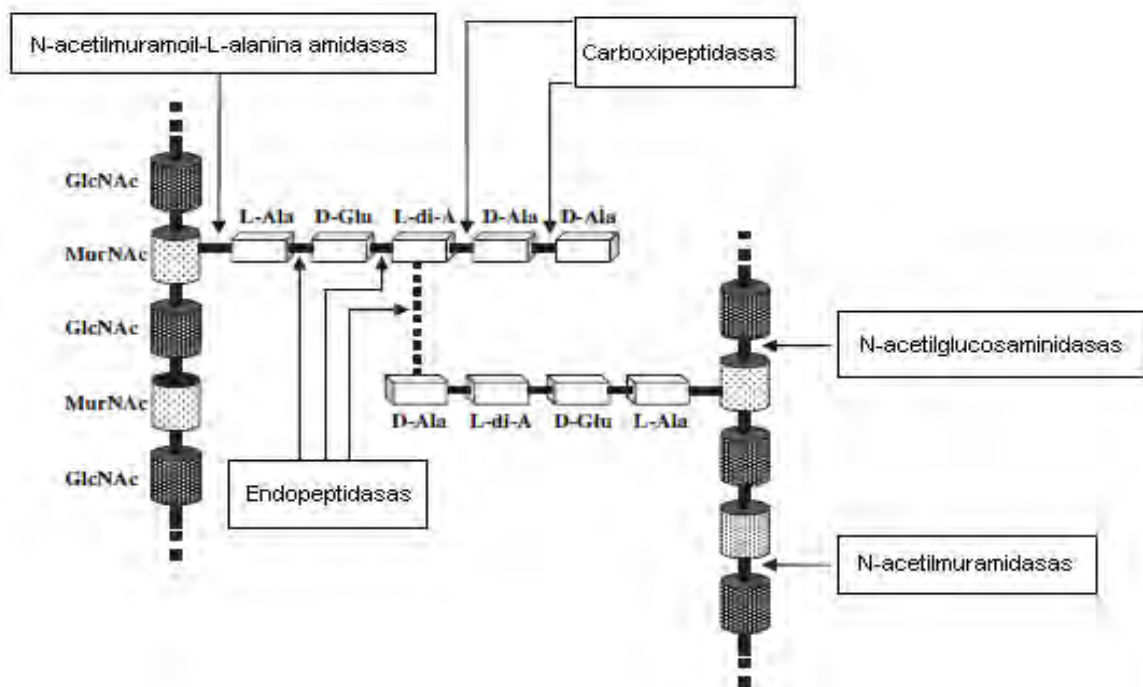


Fig. 1. Estructura del peptidoglucano y especificidad por sustrato de las PGHs (adaptado de Layec *et al.*, 2008).

una endopeptidasa que hidroliza los enlaces glicina-glicina específicos de la pared celular de *S. aureus*, tiene aplicaciones terapéuticas para tratar infecciones causadas por este microorganismo (Kumar, 2008). Otro ejemplo es la lisozima de huevo de gallina, que es utilizada como conservador en quesos y vinos (Callewaert *et al.*, 2011).

Las bacteriocinas son péptidos antibacterianos de bajo peso molecular (por lo regular, de menos de 10 kDa), sintetizadas ribosomalmente, son estables en rangos amplios de pH y son resistentes a tratamientos térmicos. Pueden ser activas en contra de bacterias de la misma especie (espectro reducido), o de diferente género (de amplio espectro); por otra parte, las cepas de bacterias que las producen tienen un mecanismo

específico que les otorga autoinmunidad. Una clasificación propuesta por Cotter *et al.* (2005), comprende tres clases de bacteriocinas (Fig. 2). La **clase I** incluye a los lantibióticos, son pequeños péptidos de entre 19 y 38 aminoácidos de longitud, que contienen residuos de lantionina o β-metil-lantionina, los que son producto de modificaciones post-traduccionales. La lantionina se obtiene por la condensación de una serina deshidratada (dehidroalanina) con el grupo sulfhidrilo de una cisteína vecina, con lo que se forma un puente entre los dos residuos. Cuando los aminoácidos que se unen son una treonina deshidratada (deshidrobutilina) y cisteína, lo que se forma es la β-metil-lantionina. La formación del puente entre residuos del péptido es catalizada enzimáticamente forma un anillo en la estructura

Artículos

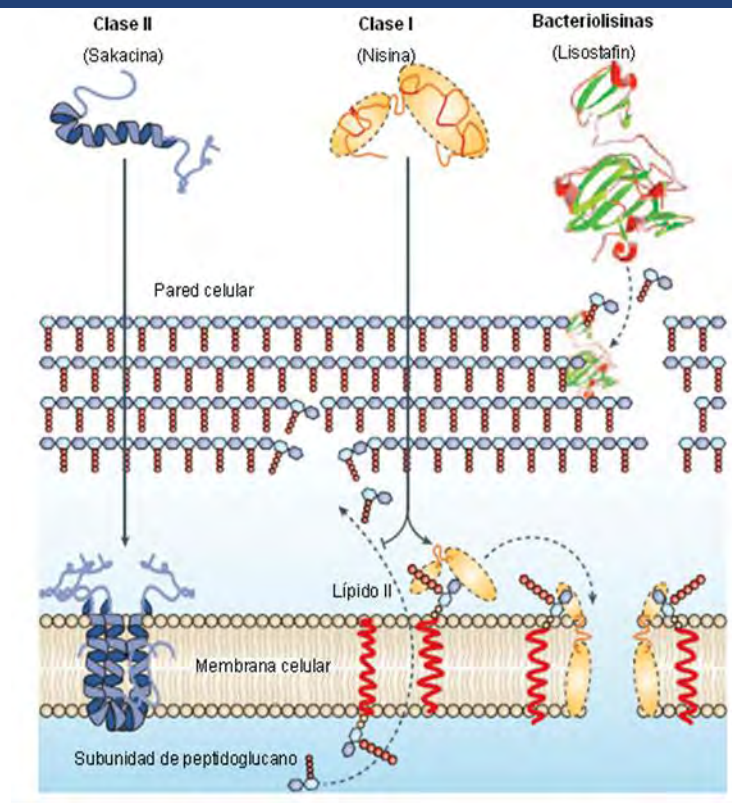


Fig. 2. Modo de acción de las bacteriocinas (Adaptado de Cotter *et al.*, 2005).

del lantibiótico, característica estructural importante de las bacteriocinas clase I.

Éstas pueden actuar de dos formas: se pueden enlazar al lípido II, el mayor del transportador de subunidades de peptidoglucano del citoplasma a la pared celular, lo que evita la síntesis correcta de la pared, con la consecuente muerte celular; además, pueden usar al lípido II como molécula de acoplamiento para insertarse en la membrana e iniciar la formación de poros, lo que lleva a una muerte rápida. La **clase II**, o bacteriocinas que no contienen lantionina, son péptidos termoestables pequeños (menos de 10 kDa), inducen la permeabilización de la membrana y, consecuentemente, la salida de moléculas del interior de la bacteria. Esta clase

está formada por un grupo heterogéneo de moléculas, dentro de las que destacan: las que tienen alta especificidad contra *Listeria monocytogenes* - tipo pediocinas - (clase IIa), las que están compuestas por dos péptidos (clase IIb) y las que tienen una unión covalente entre el extremo N- y el C- terminal, esto es, que tienen una estructura circular (clase IIc). Y en la **clase III** se encuentran las bacteriolisinas o proteínas bacteriolíticas, son termolábiles, de alto peso molecular y su modo de acción es mediante la lisis de la pared celular. En general, poseen el dominio catalítico (homólogo a las endopeptidasas) en el extremo N-terminal y el de reconocimiento al sustrato, en el amino terminal. A diferencia de las otras clases de bacteriocinas, no tienen relacionado un gen que

confiera inmunidad a la bacteria productora (Cotter *et al.*, 2005).

La única bacteriocina aprobada para su utilización como conservador en alimentos por la FDA (denominada tipo GRAS) y aceptada como aditivo alimentario en otros 45 países es la nisina. Esta bacteriocina es producida por *Lactococcus lactis* subsp. *lactis* y tiene un peso molecular de 3.5 kDa. La marca comercial más usada es Nisaplin (Danisco); es una preparación que contiene 2.5 % de nisina, NaCl (77.5%) y leche en polvo descremada. Otra bacteriocina comercial es la pediocina PA-1, denominada Atla® 2341, que inhibe el desarrollo de *Listeria monocytogenes* en productos cárnicos. Pese a que son pocas las bacteriocinas de las cuales se permite su aplicación en forma purificada, éstas pueden ser producidas directamente en alimentos por cepas iniciadoras o no iniciadoras o en humanos y animales por cepas probióticas (Cotter *et al.*, 2005; Yang *et al.*, 2014).

BLIS es un término aplicado a sustancias antimicrobianas de naturaleza proteínica de bajo peso molecular, cuyo mecanismo de acción no está completamente definido o que no cumplen con alguno de los criterios de la definición de bacteriocinas, por lo que se clasifican en un grupo aparte. Las BLIS por lo general tienen un espectro más amplio de actividad que las propias bacteriocinas, ya que son capaces de inhibir el crecimiento de bacterias Gram positivas y Gram negativas, así como hongos y levaduras (Atanassova *et al.*, 2003).

¿CÓMO DETECTAR LA ACTIVIDAD ANTIBACTERIANA?

Las BAL pueden producir una gran variedad de compuestos inhibitorios, como ya se mencionó, por lo que tienen un gran potencial de aplicación biotecnológica. Uno de los aspectos a evaluar cuando se está analizando la posibilidad de aplicación de alguna BAL, es su capacidad de producir compuestos que inhiben el crecimiento de otros microorganismos y ubicar la localización de dichos compuestos, ya sean extracelulares, intracelulares, o adheridos a la membrana o a la pared celular.

Uno de los métodos más utilizados para detectar actividad inhibitoria es el de difusión en agar (Liu *et al.*, 2014). Éste consiste en inocular el microorganismo contra el cual se quiera detectar la actividad en un medio con agar suave (0.7 - 0.8% p/v) a una concentración tal que, al crecer, forme un césped homogéneo. Previamente se puede poner una primera capa más sólida con mayor concentración de agar (1.0 - 1.5% p/v). Sobre la capa de agar suave se forman pozos dentro de los cuales se carga la muestra a la que se le quiera analizar la actividad antibacteriana. Las placas se incuban para permitir el crecimiento del microorganismo control y la difusión de la muestra. Si presenta actividad antibacteriana, se observará un halo de inhibición alrededor del pozo (Fig. 3).

Los compuestos antibacterianos que pueden producir las BAL, que hay que considerar primeramente, son los ácidos orgánicos. La influencia de éstos sobre la actividad inhibitoria se elimina al ajustar el pH de la muestra a valores entre 6 y 7, lo cual se puede hacer empleando una solución de NaOH. La forma no disociada es la forma tóxica de la molécula, ya

que ésta difunde a través de la membrana celular por ser liposoluble y, una vez dentro,

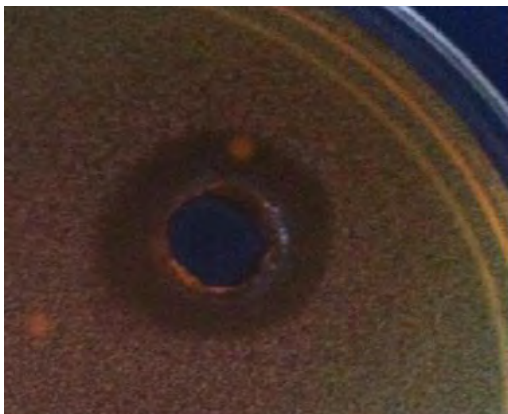


Fig. 3. Halo de inhibición del sobrenadante de la fermentación de una BAL en contra de *Listeria innocua*. Tomado de Olvera-García, 2013.

dependiendo del pH intracelular, puede disociarse en mayor o menor medida liberando iones H^+ que acidifican el citoplasma. Además del efecto del pH, la forma no disociada de la molécula contribuye al efecto antibacteriano mediante el colapso del gradiente electroquímico de protones, causando bacteriostasis y la eventual muerte de la bacteria susceptible. El efecto es más pronunciado a valores de pH por debajo del pKa del ácido, ya que una mayor proporción de éste está en su forma no disociada (Magnusson, 2003; Salminen, 2004).

Otro compuesto a considerar, por su capacidad inhibitoria del crecimiento bacteriano, es el peróxido de hidrógeno. La actividad antibacteriana se atribuye a un efecto fuertemente oxidante sobre la pared celular y a la modificación de las estructuras moleculares de proteínas celulares. Se puede eliminar su acción mediante la neutralización de la muestra

y con el uso de catalasa, ya que esta enzima cataliza la descomposición del peróxido de hidrógeno en oxígeno y agua (Liu *et al.*, 2014).

Mediante la neutralización de la muestra también se puede reducir el efecto de otros compuestos como el diacetilo, ya que su actividad antibacteriana es más efectiva a valores de pH menores a 7. Además, las cantidades necesarias para que se presente la actividad antibacteriana son muy elevadas. También se ve disminuida la actividad de la reuteriicina, ya que es más efectiva a valores de pH menores a 4.5 (Magnusson, 2003; Salminen, 2004).

Tras haber realizado estos pasos, si se observa actividad antibacteriana, ésta puede estar dada por la presencia de bacteriocinas, PGHs o ambas. A continuación se mencionan algunas estrategias para la detección y purificación de dichas proteínas con actividad antibacteriana, que pueden ser útiles para diferenciarlas.

ESTRATEGIAS PARA LA DETECCIÓN Y PURIFICACIÓN DE PEPTIDOGLUCANO HIDROLASAS

Las PGHs son enzimas extracitoplasmáticas debido a que su sustrato es el peptidoglucano, una de las capas más externas de las bacterias (Navarre & Schneewind, 1999; Vollmer *et al.*, 2008). Las PGHs pueden ser secretadas mediante la vía dependiente de Sec o mediante el sistema TAT, por lo que la mayoría presentan un péptido señal. Para efectos de la purificación, el que su localización sea extracelular representa una gran ventaja, ya que trabajar con el sobrenadante de la fermentación es más

sencillo que con células completas. Sin embargo, las PGHs también se pueden encontrar adheridas a la célula ya que la interacción entre las enzimas y la pared celular es crucial para su actividad. Algunas PGHs se unen de manera covalente al sustrato, aunque no es muy común; la mayoría presenta dominios conservados de interacción iónica con el peptidoglucano u otros componentes de la pared celular (dominios de unión a peptidoglucano o CBD: cell wall binding domains). Muchas enzimas presentan carga positiva a pH neutro, lo que facilita su unión a los componentes cargados negativamente de la pared celular (Vollmer *et al.*, 2008). De encontrarse de esta manera, la purificación requeriría un paso más para separar a la enzima de su sustrato. Algunas PGHs presentan mayor capacidad de unión al peptidoglucano a valores de pH cercanos a su punto isoeléctrico, tal es el caso de la AcmD de *Lactococcus lactis*, que en medio ácido presenta mayor capacidad de unión que en un medio neutro (Visweswaran *et al.*, 2013), por lo que al modificar el pH podrían liberarse al medio.

Para detectar actividad de PGH, no siempre es necesario que estén puras. Es posible identificarlas a partir de extractos crudos o semipuros del medio extracelular. Como se mencionó anteriormente, las PGHs desempeñan diferentes funciones en la célula, por lo que es posible encontrarlas en cualquier fase de crecimiento; sin embargo, una función importante que desempeñan es la autólisis, la cual se presenta después de la fase exponencial de crecimiento, en donde se puede encontrar la mayor actividad. Una vez que se

tiene un cultivo en este punto, deben separarse las células del sobrenadante.

Concentración y semipurificación de la preparación enzimática

La precipitación diferencial de proteínas puede utilizarse como primera operación en la estrategia de purificación; los métodos de precipitación de proteínas se basan en la disminución de la solubilidad, reparto selectivo a una fase en un sistema de dos fases o precipitación por afinidad. De tal modo que la precipitación proteínica puede ser realizada por la adición de sales, solventes o polímeros orgánicos, por alteraciones de pH y temperatura o por la adición de ligandos.

La solubilidad de las proteínas se reduce generalmente a altas concentraciones de sal; cuando una sal es adicionada al sistema, el agua solvata los iones de sal y a medida que la concentración salina aumenta, el agua que rodea a la proteína es secuestrada, exponiendo eventualmente las porciones hidrofóbicas. Es entonces cuando los segmentos hidrofóbicos de una molécula proteínica pueden interactuar con los de otra, resultando en la agregación de éstas. Por lo que las proteínas con segmentos hidrofóbicos largos tienden a agregarse y precipitar antes que aquellas con pequeños o pocos segmentos hidrofóbicos. La precipitación ocurre por la neutralización de cargas de la superficie de la proteína por la sal, por la reducción de la actividad acuosa de la proteína y por la disminución efectiva de la concentración de agua en el medio, este efecto es conocido como "salting out".

Por otra parte, la adición de ciertas sales en la cantidad correcta puede precipitar selectivamente algunas proteínas, mientras que otras permanecen en solución. El sulfato de amonio ((NH₄)₂SO₄) es particularmente efectivo y se usa generalmente para precipitar proteínas. Las proteínas precipitadas pueden ser removidas de las que permanecen solubles mediante una centrifugación a baja velocidad. Las cantidades residuales de sulfato de amonio pueden interferir con otras mediciones y métodos de purificación y tiene que ser removido posteriormente (Prado-Barragán *et al.*, 1999; Nelson & Cox, 2008).

También, es posible utilizar ácido tricloroacético (TCA) para la precipitación de proteínas. Con este método se concentra la muestra, y es efectivo para remover sales, polisacáridos y muchos detergentes (como el SDS) de las proteínas (Matsudaira, 1993; Nandakumar *et al.*, 2002). Los residuos de TCA deben ser removidos con lavados exhaustivos con acetona o etanol. La principal desventaja de este método es que se somete a las enzimas a un tratamiento drástico y no todas mantienen la actividad o son capaces de recuperarla con tratamientos renaturalizantes.

Por lo general, las PGHs se encuentran por arriba de los 20 kDa y se pueden separar de moléculas pequeñas mediante la diálisis a través de una membrana semipermeable, como puede serlo una membrana porosa de celulosa. Las moléculas de dimensiones significativamente mayores que el diámetro del poro se retienen dentro de la bolsa de diálisis, mientras que las moléculas más pequeñas y los iones atraviesan los poros de esta membrana.

Esta técnica es útil para retirar las sales u otras moléculas pequeñas como el sulfato de amonio después de una precipitación (Berg *et al.*, 2008).

Otra técnica que se puede emplear para la semipurificación y concentración de PGHs es la ultrafiltración. Esta técnica involucra la aplicación de una fuerza (presión) sobre el fluido a través membranas semipermeables. Los componentes que son más pequeños que los poros de la membrana pasan a través de ella (sales, azúcares, proteínas más pequeñas, etc.) y los más grandes son retenidos. La ultrafiltración tiene un umbral de exclusión en el rango de 1 a 200 kDa, de acuerdo al poro de la membrana que se utilice. Se puede utilizar una con tamaño de corte de 10 kDa, con la que se podrían concentrar las PGHs (Prado-Barragán *et al.*, 1999).

Purificación de la preparación enzimática

Uno de los métodos de utilidad general para la purificación de proteínas es la cromatografía, la cual aprovecha las diferencias en carga, tamaño, afinidad de enlaces, etc. que hubiera en una mezcla de proteínas. Un material poroso sólido con propiedades químicas apropiadas (fase estacionaria) se coloca dentro de una columna, y una solución amortiguada (fase móvil) pasa a través de ésta. La solución que contiene las proteínas se vierte por la parte superior de la columna y pasa a través de la matriz sólida. Las proteínas migran más rápido o más lento a través de la columna dependiendo de su interacción con la fase estacionaria, con lo que se logra su separación. Hay diferentes tipos de matrices para

cromatografía. Las de intercambio iónico constan de pequeñas esferas que contienen cargas positivas o negativas, de modo que las proteínas se separan en función de la disposición de cargas de la superficie. Las columnas hidrofóbicas están empacadas con esferas que poseen cadenas laterales hidrofóbicas que sobresalen de ellas, de modo que las proteínas que tengan regiones hidrofóbicas al descubierto quedan retardadas en su tránsito por la columna. Las columnas de filtración en gel, que separan proteínas en función de su tamaño, contienen diminutas esferas porosas: a medida que van descendiendo por la columna, las moléculas suficientemente pequeñas para penetrar en los poros pasan sucesivamente por el interior de estas esferas, mientras que las moléculas más grandes permanecen en la solución fluyendo entre las esferas y, por lo tanto, eluyen más rápidamente a través de la columna y salen primero. Existe también la cromatografía de afinidad, donde la fase estacionaria interactúa específicamente con la proteína de interés. Ejemplos de estas matrices son las que se utilizan para separar a proteínas que tengan una etiqueta de histidinas en el extremo N-terminal, por la presencia de algún metal como Ni, Co o Cu en la fase estacionaria. O por la interacción de la proteína de interés con un anticuerpo específico que se encuentra unido a la fase estacionaria.

También existen resinas cromatográficas en forma de diminutas esferas (de 3 a 10 μm de diámetro) que pueden ser empaquetadas mediante un aparato especial formando un lecho uniforme en la columna. Con estas

columnas de cromatografía líquida de alta resolución (HPLC, de High Performance Liquid Chromatography) se alcanza un alto grado de separación. Como las columnas de HPLC contienen partículas empaquetadas de forma muy compacta, la velocidad de flujo a través de ellas es prácticamente nula a menos que se apliquen altas presiones (Alberts *et al.*, 2002).

A través de la aplicación de uno o varios de los métodos descritos anteriormente, se puede obtener la PGH de interés con diferentes grados de pureza. Así, por ejemplo, para la purificación de la AtIB de *Enterococcus faecalis* JH2-2, Mesnage *et al.* (2008) emplearon la siguiente metodología: la cepa fue reactivada en caldo BHI hasta una $\text{DO}_{600\text{nm}}$ de 0.7. Las células fueron recolectadas por centrifugación y se usaron para inocular un medio sintético. Tras la fermentación, se retiraron las células y al sobrenadante se le agregó gradualmente sulfato de amonio hasta una concentración final de 0.7 g/mL, con agitación a 4 °C. Las proteínas que precipitaron se recolectaron por centrifugación y se resuspendieron en Tris-HCl (pH 8.5). Posteriormente, la enzima fue purificada mediante dos cromatografías de intercambio aniónico (HiTrap Q-Sepharose fast flow y MonoQHR5) usando gradientes de NaCl en Tris-HCl.

La metodología a seguir depende de las características de cada proteína, como peso molecular, punto isoelectrico, carga, hidrofobicidad, resistencia a diferentes tratamientos, etc. y se pueden combinar uno o varios métodos de los ya descritos.

Evaluación de la actividad enzimática

Artículos

Una vez que se tiene la muestra concentrada o con una mayor pureza, se debe verificar la actividad lítica. Esto se puede hacer espectrofotométricamente mediante el monitoreo de la disminución de la DO_{600nm} de una suspensión de células de la cepa productora de la enzima, debido a que las PGHs son activas en contra de la misma. También se puede hacer un ensayo con el peptidoglucano en lugar de las células completas, esta vez monitoreando la DO a 450 nm (Leclerc & Asselin, 1989). Hay un método para la medición de la actividad de muramidasa, el cual es descrito por Sigma-Aldrich como ensayo enzimático de lisozima (EC 3.2.1.17). El cual se basa en que la lisozima es capaz de lisar células intactas de *Micrococcus lysodeikticus*, reduciendo la densidad óptica a 450 nm de una suspensión de este microorganismo. Esta misma bacteria se puede utilizar como sustrato para la realización de zimogramas (Sigma-Aldrich).

Los métodos zimográficos se basan en la separación de enzimas por electroforesis en gel de poliacrilamida seguida de un paso de renaturalización; al recobrar su estructura funcional, la actividad de la enzima se detecta por su efecto sobre el sustrato que se ha embebido, esparcido o sobrepuesto en el gel. Los zimogramas se han utilizado para la detección de PGHs en numerosos estudios. Generalmente se utiliza como sustrato células de *M. lysodeikticus* (también llamado *M. luteus*), ya sea liofilizadas, de cultivo fresco o inactivadas por calor, las cuales se embeben dentro de la red de la poliacrilamida

polimerizada. Este microorganismo es utilizado debido a que posee una pared celular delgada y susceptible a la degradación; sin embargo, es posible utilizar como sustrato cualquier bacteria contra la que se desee probar la actividad lítica. Las enzimas se separan por electroforesis en condiciones desnaturalizantes, para evitar interacción prematura con el sustrato. Posteriormente, se lleva a cabo una renaturalización (con una solución de Tritón X-100 1 % v/v) para que las enzimas recuperen la actividad e hidrolicen el sustrato embebido, ésta se observa como una banda clara en un fondo opaco. Para aumentar el contraste, el gel puede ser teñido con una solución de azul de metileno (0.1 % p/v) en KOH (0.01 % p/v). Este método permite una detección simultánea de la actividad enzimática y una estimación del peso molecular, además de que posee una elevada sensibilidad (Hardt *et al.*, 2003). En la figura 4 se muestra un zimograma donde se pueden observar varias bandas de pesos moleculares entre 45 y 95 kDa que tienen actividad lítica contra varias bacterias indicadoras. También es posible identificar a la proteína responsable de la actividad y conocer parcialmente su secuencia de aminoácidos si de manera paralela al zimograma se corre un gel para observar el patrón de bandas tiñendo con azul de Coomassie, posteriormente se comparan ambos geles y la banda con actividad se corta para analizarla. La identificación de la proteína se puede realizar por el método de cromatografía líquida acoplada a espectrometría de masas, capaz de producir espectros en tándem (LC-MS/MS).

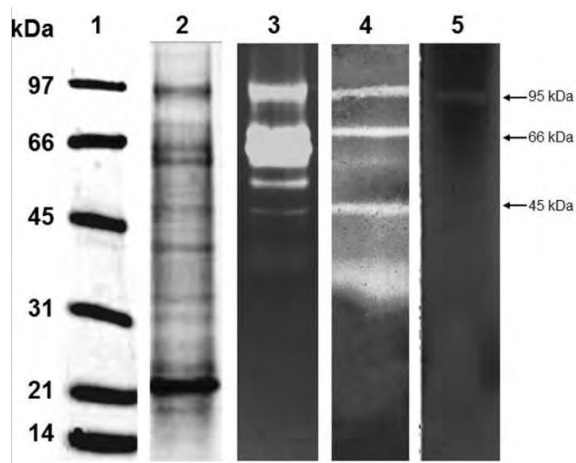


Fig. 4. Perfil de proteínas (carril 2) y zimograma de actividad lítica contra *Micrococcus lysodeikticus*, *Staphylococcus aureus* y *Escherichia coli*, carriles 3, 4 y 5. Modificado de García-Cano *et al.*, 2014.

En este método, la proteína es digerida (generalmente con tripsina) y el análisis de espectrometría de masas arroja la secuencia de los fragmentos generados. Al someter los espectros de fragmentación o las secuencias de estos péptidos a los bancos de datos se puede identificar una determinada proteína (Unidad de Proteómica, Instituto de Biotecnología, UNAM).

ESTRATEGIAS PARA LA DETECCIÓN Y PURIFICACIÓN DE BACTERIOCINAS

Como se mencionó anteriormente, las bacteriocinas son péptidos antimicrobianos, catiónicos y con tendencia a adoptar estructuras helicoidales. Son producidas por diferentes microorganismos con actividad antagónica contra bacterias relacionadas filogenéticamente.

Concentración y purificación de bacteriocinas

Los métodos de extracción y purificación de bacteriocinas, como el resto de las proteínas, depende de su naturaleza. Por lo tanto para obtenerlas, en la mayoría de las ocasiones se deben combinar diferentes estrategias. Cabe resaltar que estos péptidos con actividad antibacteriana se caracterizan por producirse durante la fase logarítmica de crecimiento, y su mayor presencia se observa durante la fase logarítmica tardía.

A continuación se enuncian algunas estrategias para la extracción y purificación de bacteriocinas:

1. Adsorción y desorción.

Se parte de un cultivo que se encuentra en fase logarítmica tardía, el cual se somete a un tratamiento térmico durante 30 min a 70 °C para la inactivación de proteasas que pudieran estar presentes en el medio (Katla *et al.*, 2003). Posteriormente, se ajusta el medio a pH 5 para permitir la adsorción del compuesto antimicrobiano a la membrana celular bacteriana (Dündar *et al.*, 2014). A continuación se obtiene el paquete celular por centrifugación, el cual se lava con buffer de fosfatos. Finalmente, las células se resuspenden en una solución 100 mM de NaCl a pH 2 para favorecer la desorción y liberación del compuesto de interés de la membrana celular de la bacteria hacia el medio. Así mismo, el NaCl previene la aglutinación de las bacteriocinas, lo cual evita que éstas no sean removidas de la membrana (Yang *et al.*, 1992; Álvarez-Cisneros, 2010). Los sobrenadantes se dializan empleando membranas de celulosa con un tamaño de poro de 1 kDa y, finalmente, se liofilizan para

concentrar. La actividad antibacteriana se observa en geles SDS-Tris-tricina (que se describirán más adelante) y por prueba de difusión en agar (Olvera-García, 2013).

2. Precipitación con sulfato de amonio.

La siguiente metodología fue descrita para la purificación de la enterocina 62-6, producida por *Enterococcus faecium* 62-6. El sobrenadante libre de células, obtenido de cultivo en fase logarítmica tardía, se somete a una precipitación de proteínas usando sulfato de amonio al 60% p/v (se recomienda probar un gradiente) a 4 °C y es separado por centrifugación. El precipitado se resuspende en buffer de citratos (pH 5), y se dializa usando una membrana de celulosa (corte de 12 kDa) durante 3 días. La muestra dializada se pasa a través de una columna de intercambio catiónico, cargando 20 mL en 10 mL de la columna de Carboximetil Sepharosa Fast-Flow que ha sido previamente equilibrada con buffer de citratos pH 5 a un flujo de 8 mL/h. Las proteínas son eluidas de la columna empleando un gradiente salino de 0.1 - 0.7 M de NaCl. Se colectan fracciones de 2 mL y éstas se emplean para el ensayo de actividad inhibitoria a través de la prueba de difusión en agar. Las fracciones que fueron positivas en el ensayo anterior se analizaron en un gel SDS-PAGE al 15 % (p/v) de acrilamida, teñido con azul de Coomassie R-250 o plata. El peso molecular se determinó usando un patrón de referencia y por análisis de espectrometría de masas (Dezwaan *et al.*, 2007).

3. Precipitación con acetona.

A continuación se cita la metodología empleada para la purificación de la lactococina 972. Se realizó un cultivo en medio M17 hasta

llegar a la fase logarítmica tardía y se obtuvo el sobrenadante por centrifugación (fracción I). El sobrenadante se precipitó con 5 volúmenes de acetona fría, dejando reposar durante 30 min a -20 °C. Posteriormente, se centrifugó y el precipitado obtenido, una vez seco, se solubilizó en buffer de fosfato sódico 50 mM, pH 7 (fracción II). A continuación, se tomaron muestras de 20 mL de la fracción II y se cargaron en una columna de intercambio catiónico High-S de 5 mL, equilibrada con buffer de fosfatos, tras haberla lavado con NaCl 0.1 M. La bacteriocina se eluyó con 30 mL de NaCl 0.4 M con un flujo de 1 mL/min, obteniendo fracciones de 1 mL (fracción III). La actividad inhibitoria de las tres fracciones se evaluó por la prueba de difusión en agar (Martínez-Fernández, 1996).

4. Adsorción y desorción /Cromatografía de intercambio catiónico en SP-Sepharosa y Cromatografía líquida de alta resolución en fase reversa (RP-HLPC).

Dündar *et al.* (2014) reportan la siguiente metodología para la purificación de las enterocinas A, B, P, I, L50A y L50B, producidas por *E. faecium* W3. Se parte de un cultivo en medio MRS de 16 h de incubación. Esta técnica es empleada cuando la bacteriocina se encuentra asociada a la superficie celular, por lo que el paquete celular del cultivo es sometido al tratamiento de adsorción y desorción descrito anteriormente. Las muestras liofilizadas son resuspendidas en ácido trifluoroacético 0.1 % (p/v) para, posteriormente, separar los péptidos por RP-HLPC inyectando alícuotas de 100 µL. Las fracciones peptídicas son colectadas y ensayadas para probar su actividad de

bacteriocinas. El análisis de secuencia de aminoácidos se realizó a través de espectrometría de masas MALDI-TOF.

La siguiente estrategia general se ha empleado en la producción de diversas bacteriocinas, como lo es la pediocina PA-1 por *Pediococcus pentosaceus*, lo cual resalta que las técnicas descritas anteriormente pueden ser aplicadas para cualquier BAL (Fig. 5).

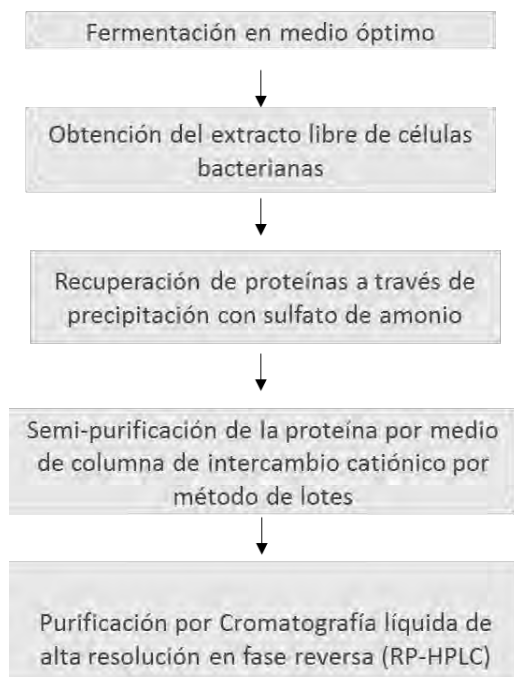


Fig. 5. Esquema de purificación rápida utilizada para la producción de la pediocina PA1. Modificado de Vijay *et al.*, 2012.

Detección de la actividad bacteriolítica

Una vez que se han purificado las bacteriocinas, se debe utilizar una metodología para su detección. A continuación se mencionan algunos métodos, reportados en la bibliografía, para este fin:

1. Geles SDS-Tris-tricina y zimografía.

Para identificar las bandas con actividad antibacteriana se pueden realizar geles de SDS-Tris-tricina al 10 % (p/v) con células embebidas con un microorganismo indicador (ej.: *Listeria innocua*). Este método, que está basado en la introducción de tricina como ión acarreador, da una resolución mayor que el de SDS-Tris-glicina, especialmente en el rango de 5 a 30 kDa (Schägger, 2006).

Se parte del paquete celular de un cultivo del microorganismo indicador, el cual se lava y centrifuga dos veces con buffer de fosfatos a pH 7. El pellet limpio se resuspende en solución salina al 0.85 % (p/v), pH7 y se mide D.O. a 600 nm. Una D.O. de 7 a 8 se considera óptima para *L. innocua* o *Micrococcus lysodeikticus*, como microorganismos indicadores (Olvera-García, 2013). En la tabla 1 se presentan los reactivos y el orden en que se deben agregar al paquete celular que corresponde a la D.O. mencionada. Las muestras obtenidas por el método de adsorción y desorción son ideales para la visualización de la actividad por zimografía ya que se encuentran libres de impurezas. Los liofilizados se resuspenden en 30 μ L de buffer de carga del gel (Tris-HCl 3M/SDS 0.3 % p/v, pH 8.45) y se calientan 5 min a 95 °C para después añadirse en los pozos del gel. Como patrón de referencia se utiliza un marcador de peso molecular de polipéptido (de 1.4 a 26.6 kDa). Las condiciones de la electroforesis son 80 V durante 5 h, a 4 °C. Posteriormente, se incuban a 37°C y 50 rpm, durante 18 h en una solución renaturalizante (Tris-HCl 100 mM y Tritón X-100 1 % v/v, pH 8) para que las proteínas puedan recobrar su estructura y actividad, la cual se verá reflejada en el gel por

Artículos

Tabla 1. Reactivos para la preparación del zimograma para bacteriocinas (Mini-PROTEAN III, Bio-rad).

Orden de incorporación	Reactivo	Gel	Gel
		separador	concentrador
1	Agua	1.93 mL	1.56 mL
2	Glicerol 80 % (v/v)	0.89 mL	-
3	Acrilamida Tris-tricina	2.83 mL	324 µL
4	Buffer de gel	2.83 mL	620 µL
5	Temed	5.66 µL	1.5 µL
6	Persulfato de amonio 10 % (p/v)	42.5 µL	13.3 µL
	Volumen final	8.5 mL	2.5 mL

(Olvera-García, 2013)

la presencia de una banda clara o translúcida sobre un fondo opaco. Para aumentar el contraste, el gel es teñido con una solución de azul de metileno, durante 20 minutos y posteriormente se destiñe con agua destilada hasta observar la aparición de las bandas. En la figura 6 se pueden visualizar bandas de actividad antibacteriana de aproximadamente 5 kDa, correspondiente a la enterocina A, producida por diferentes cepas de de *E. faecium* C, D y G) y *E. faecalis* (E) aisladas del queso Cotija Región de Origen (Olvera-García, 2013). Se recomienda hacer un gel SDS-PAGE Tris-tricina sin microorganismo indicador al mismo tiempo, con el fin de visualizar la presencia de la banda que corresponde a la proteína con actividad inhibitoria.

2. SDS-PAGE acoplado a ensayo en placa.

Geles de SDS-PAGE-Tris-tricina se preparan y corren tal y como se mencionó previamente,

pero sin adicionar las células del microorganismo a inhibir. El primer gel es teñido, ya sea usando azul de Coomassie o tinción con plata para visualizar las bandas pertenecientes a

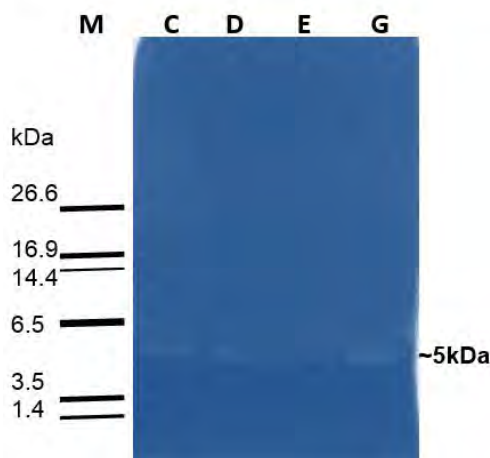


Fig. 6. Zimograma contra el microorganismo indicador *Micrococcus lysodeikticus*.

M: Marcador de peso molecular de polipéptido Bio-Rad. C, D, E y G, cepas de *Enterococcus faecalis* (Olvera-García, 2013).

Artículos

proteínas. A su vez, el segundo gel es sometido a la técnica de superposición. En ésta se tiene una placa de agar donde se inocula al microorganismo indicador a inhibir, por vertido. El gel se coloca sobre el agar y se incuba a la temperatura óptima de crecimiento de la bacteria. Una zona translúcida en el tapete del cultivo, indica que la prueba es positiva para actividad antibacteriana (Fig. 7). Ambos geles pueden ser comparados para poder determinar el peso molecular de la(s) bacteriocina(s). Sin embargo, esta técnica es menos sensible que los zimogramas con células embebidas, además de que es más difícil distinguir la zona de inhibición del microorganismo. En la figura 7 se observa la detección directa de la enterocina E-760, contra *Campylobacter jejuni*. Los autores identificaron una banda de 5.5 kDa, como responsable de la actividad antibacteriana (Line *et al.*, 2008).

Este ensayo no es tan sensible como la zimografía, descrita en el punto anterior, y en nuestra experiencia de laboratorio, la zona translúcida no se obtiene tan clara como lo que reportan Line *et al.* Adicionalmente, el crecimiento del césped del microorganismo indicador puede afectarse por otros factores, no necesariamente a la actividad de una determinada proteína del gel.

CONCLUSIONES

En los últimos años, proteínas como las PGHs o las bacteriocinas han cobrado gran relevancia, como alternativa al uso de antibióticos, al provenir de organismos que no representan riesgo alguno para el consumidor; ya que pueden ser utilizadas para evitar la

presencia de microorganismos indeseables en alimentos, así como en ambientes clínicos o áreas que deban permanecer en condiciones de asepsia. Para su análisis bioquímico es necesario contar con una preparación suficientemente concentrada y libre de otros compuestos, como carbohidratos y otras proteínas, que dificultan la aplicación exitosa de herramientas analíticas tales como la electroforesis y la zimografía. Es frecuente que cuando se observa que una bacteria tiene una actividad antagonica hacia otras (por ejemplo, por pruebas de difusión en agar) se piense inmediatamente en la posibilidad de que esté produciendo bacteriocinas, sin considerar el

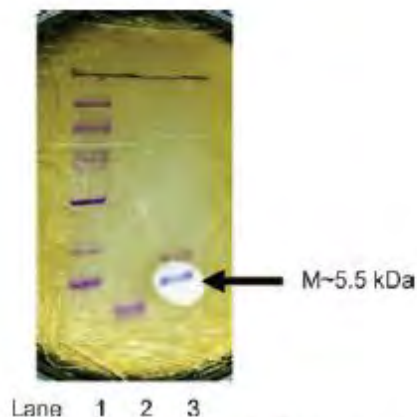


Fig. 7. Técnica de superposición de gel en placa. Tomado de Line *et al.*, 2008.

efecto que pueden producir otros metabolitos, como los ácidos orgánicos o el peróxido de hidrógeno. Generalmente se recurre al análisis de las preparaciones de proteínas a través de SDS-PAGE-Tris-tricina, pero no se toma en cuenta la fracción de mayor peso molecular, por lo que se excluye del análisis a un grupo de enzimas que también puede ser

responsable de la lisis bacteriana. En esta revisión se proporcionan herramientas que han tenido resultados exitosos en diferentes casos para ampliar el espectro de posibilidades que puede presentar una bacteria con actividad antagonista hacia otras.

REFERENCIAS

- Alberts B, Bray D, Lewis J, Raff M, Roberts K & Watson JD (2002) *Biología molecular de la célula*. 3ª edición. Ediciones Omega. Barcelona, España. pp. 176-181.
- Álvarez-Cisneros YM, Fernández FJ, Wachter-Rodarte C, Aguilar MB, Sainz ETR & Ponce-Alquicira E (2010) Biochemical characterization of a bacteriocin-like inhibitory substance produced by *Enterococcus faecium* MXVK29, isolated from Mexican traditional sausage. *J. Sci. Food Agric.* 90: 2475–2481.
- Atanassova M, Choiset Y, Dalgalarondo M, Chobert JM, Dousset X, Ivanova I & Haertlé T (2003) Isolation and partial biochemical characterization of a proteinaceous anti-bacteria and anti-yeast compound produced by *Lactobacillus paracasei* subsp. *paracasei* strain M3. *Int. J. Food Microbiol.* 83: 63-73.
- Berg JM, Stryer L & Tymoczko JL (2008) *Bioquímica*. 6ª edición. Editorial Reverté. Barcelona, España. p. 69.
- Callewaert L, Walmagh M, Michiels CW & Lavigne R (2011) Food applications of bacterial cell wall hydrolases. *Curr. Opin. Biotech.* 22: 164–171.
- Cotter PD, Hill C. & Ross RP (2005) Bacteriocins: Developing innate immunity for food. *Nat. Rev. Microbiol.* 3: 777-788.
- Dezwaan DC, Mequio MJ, Littell JS, Allen JP, Rossbach S & Pybus V (2007) Purification and characterization of enterocin 62-6, a two-peptide bacteriocin produced by a vaginal strain of *Enterococcus faecium*: Potential significance in bacterial vaginosis. *Microb. Ecol. Health Dis.* 19(4): 241–250.
- Dündar H, Atakay M, Çeliklebiçak Ö, Salih B & Bozoğlu F (2014) Comparison of two methods for purification of enterocin B, a bacteriocin produced by *Enterococcus faecium* W3. *Prep. Biochem. Biotech.* 45(8): 796-809.
- Eckert C, Lecerf M, Dubost L, Arthur M & Mesnage S (2006) Functional analysis of AtlA, the major *N*-acetylglucosaminidase of *Enterococcus faecalis*. *J. Bacteriol.* 188: 8513-8513.
- Elsser-Gravesen D & Elsser-Gravesen A (2014) Biopreservatives. *Adv. Biochem. Eng. Biot.* 143: 29-49.
- García-Cano I, Serrano-Maldonado CE, Olvera-García M, Delgado-Arciniega E, Peña-Montes C, Mendoza-Hernández G & Quirasco M (2014) Antibacterial activity produced by *Enterococcus* spp. isolated from an artisanal Mexican dairy product, Cotija cheese. *LWT – Food Sci. Technol.* 59:26-34.
- Götz F, Heilmann C & Stehle T (2014) Functional and structural analysis of the major amidase (Atl) in *Staphylococcus*. *Int. J. Med. Microbiol.* 304(2): 156-163.
- Hardt M, Guo Y, Henderson G & Laine RA (2003) Zymogram with Remazol brilliant blue-labeled *Micrococcus lysodeikticus* cells for the detection of lysozymes: example of a

Artículos

- new lysozyme activity in Formosan termite defense secretions. *Anal. Biochem.* 312: 73–76.
- Kang OJ, Laberge S & Simard RE (2003) Detection and localization of a peptidoglycan hydrolase in *Lactobacillus delbrueckii* subsp. *bulgaricus*. *J. Dairy Sci.* 86: 96-104.
- Katla T, Naterstad K, Vancanneyt M, Swings J & Axelsson L (2003) Differences in nisin. *Appl. Environ. Microb.* 69(8): 4431-4437. Kumar JK (2008) Lysostaphin: an antistaphylococcal agent. *Appl. Microbiol. Biot.* 80(4): 555-561.
- Layec S, Decaris B & Leblond-Bourget N (2008) Diversity of Firmicutes peptidoglycan hydrolases and specificities of those involved in daughter cell separation. *Res. Microbiol.* 159: 507-515.
- Leclerc D & Asselin A (1989) Detection of bacterial cell wall hydrolases after denaturing polyacrylamide gel electrophoresis *Can. J. Microbiol.* 35(8): 749-753.
- Line JE, Svetoch EA, Eruslanov BV, Perelygin VV, Mitsevich EV, Mitsevich IP, Levchuk VP, Svetoch OE, Seal BS, Siragusa GR & Stern NJ (2008) Isolation and purification of enterocin E-760 with broad antimicrobial activity against Gram-positive and Gram-negative bacteria. *Antimicrob. Agents Chemother.* 52(3): 1094–1100.
- Liu W, Zhang L, Yi H, Shi J, Xue C, Li H, Jiao Y, Shigwedha N, Du M & Han X (2014) Qualitative detection of class IIa bacteriocinogenic lactic acid bacteria from traditional Chinese fermented food using a YGNGV-motif-based assay. *J. Microbiol. Meth.* 100:121-127.
- Lortal S & Chapot-Chartier MP (2005) Role, mechanisms and control of lactic acid bacteria lysis in cheese. *Int. Dairy J.* 15: 857-871.
- Magnusson, J. 2003. Antifungal activity of lactic acid bacteria. Ph. D. Dissertation. Swedish University of Agricultural Sciences. Upsala, Suecia. pp. 3-10.
- Martínez-Fernández B (1996) Bacteriocinas de *Lactococcus lactis* aislados de quesos asturianos: nisina z y lactococina 972. Tesis de Doctorado en Biología. Universidad de Oviedo. Oviedo, España. pp. 30-31.
- Matsudaira PT (1993) A Practical guide to protein and peptide purification for microsequencing. 2ª edición. Academic Press, Inc. San Diego, CA. p. 48.
- Mesnage S, Chau F, Dubost L & Arthur M (2008) Role of *N*-acetylglucosaminidase and *N*-acetylmuramidase activities in *Enterococcus faecalis* peptidoglycan metabolism. *J. Biol. Chem.* 283: 19845-19853.
- Moreno MRF, Leisner JJ, Tee LK, Ley C, Radu S, Rusul G, Vancanneyt M & De Vuyst L (2002) Microbial analysis of Malaysian tempeh, and characterization of two bacteriocins produced by isolates of *Enterococcus faecium*. *J. Appl. Microbiol.* 92: 147-157.
- Nandakumar MP, Shen J, Raman B & Marten MR (2002) Solubilization of trichloroacetic acid (TCA) precipitated microbial proteins via NaOH for two-dimensional electrophoresis. *J. Proteome Res.* 2:89-93.

Artículos

- Navarre WW & Schneewind O (1999) Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. *Microbiol. Mol. Biol. Rev.* 63(1):174-229.
- Nelson DL & Cox MM (2008) Lehninger. Principles of Biochemistry. 5ª edición. W. H. Freeman and Company. Nueva York, NY. pp. 85-91.
- Nouaille S, Ribeiro LA, Miyoshi A, Pontes D, Le Loir Y, Costa S, Langella P & Azevedo V (2003) Heterologous protein production and delivery systems for *Lactococcus lactis*. *Genet. Mol. Res.* 2:102-111.
- Olvera-García ME (2013) Evaluación de la inocuidad de *Enterococcus* spp. aislados del queso Cotija. Tesis de Maestría en Ciencias Bioquímicas. Universidad Nacional Autónoma de México. México, D.F. pp. 40-41, 67-68.
- Peterbauer C, Maischberger T & Haltrich D (2011) Food-grade gene expression in lactic acid bacteria. *Biotechnol. J.* 6: 1147-1161.
- Prado-Barragán LA, Huerta-Ochoa S, Rodríguez-Serrano G & Saucedo-Castañeda G (1999) Avances en purificación y aplicación de enzimas en biotecnología. Colección Tópicos en Biotecnología. 1ª edición. Universidad Autónoma Metropolitana. Unidad Iztapalapa. México, DF. pp. 33-87.
- Salminen S, von Wright A & Ouwehand A (2004) Lactic acid bacteria. microbiological and Functional Aspects. 3ª edición. Marcel Dekker Inc. Nueva York, NY. pp. 375-398.
- Savijoki K, Ingmer H & Varmanen P (2006) Proteolytic systems of lactic acid bacteria. *Appl. Microbiol. Biotechnol.* 71: 394-406.
- Schägger H (2006). Tricine-SDS-PAGE. *Nat. Protoc.* 1(1):16-22.
- Sigma-Aldrich. Enzymatic Assay of Lysozyme (EC 3.2.1.17). Disponible en: <http://www.sigmaaldrich.com/technical-documents/protocols/biology/enzymatic-assay-of-lysozyme.html> Unidad de Proteómica, Instituto de Biotecnología, UNAM. Disponible en: <http://www.ibt.unam.mx/proteomica/informaciones.pdf>
- Vijay SB, Sood SK, Kumariya R & Garsa AK (2012) Simple and rapid purification of pediocin PA-1 from *Pediococcus pentosaceus* NCDC 273 suitable for industrial application. *Microbiol. Res.* 167(9): 544-549.
- Visweswaran GR, Steen A, Leenhouts K, Szeliga M, Ruban B, Hesseling-Meinders A, Dijkstra BW, Kuipers OP, Kok J & Buist G (2013) AcmD, a homolog of the major autolysin AcmA of *Lactococcus lactis*, binds to the cell wall and contributes to cell separation and autolysis. *PLoS One.* 8(8): e72167. doi: 10.1371/journal.pone.0072167.
- Vollmer W, Joris B, Charlier P & Foster S (2008) Bacterial peptidoglycan (murein) hydrolases. *FEMS Microbiol. Rev.* 32(2): 259-86.