

UNIVERSIDAD NACIONAL DE SAN ANTONIO ABAD DEL CUSCO

FACULTAD DE CIENCIAS

ESCUELA PROFESIONAL DE BIOLOGIA



Descripción de enzimas proteolíticas y obtención de una proteasa recombinante a partir del análisis de datos del genoma de *Chromohalobacter salexigens* MP25462.

TESIS PARA OPTAR AL TÍTULO PROFESIONAL DE BIÓLOGO

PRESENTADA POR:

Bach. Jhomira Ilución Mamani Morales

ASESORA:

Dra. María Antonieta Quispe Ricalde

CO-ASESOR:

Blgo. José Luis Sierra Herrera

TESIS FINANCIADA POR LA UNSAAC

CUSCO – PERÚ

2022

A Dios, que iluminó mi camino cuando sentía que no podía continuar. Me dio luz, fuerza y me presentó personas que han sido mi soporte y compañía durante todo este periodo.

A mis queridos padres Policarpo y Estela, por su gran amor y su guía constante en cada uno de mis pasos.

A mis hermanas Jhemsin, Maristel y Yesenia, por sus palabras de aliento y su apoyo incondicional.

A mi hermano Polisman por la fuerza que genera en mí a pesar de ya no estar presente en esta vida.

A mi hermana Calipso y su esposo Miguel, que son unos padres para mí y los principales promotores de mis sueños, me alegra llegar a término y festejar este logro con ustedes. Gracias por su apoyo y comprensión, los quiero infinitamente.

A mí querida sobrina Lucerito por ser la amiga que me impulsa en los proyectos de mi vida.

A Princesa, mi compañera fiel y mi motor para luchar por nuestros sueños. De igual manera a Copito, que se unió a nosotros en esta travesía.

La realización del trabajo presentado en esta tesis de investigación ha sido posible gracias al Proyecto de investigación ***“Inhibición de proteasas de Leishmania a través de péptidos obtenidos de proteínas de Lupinus mutabilis por acción enzimática de proteasas de bacterias halófilas: Nuevas perspectivas de búsqueda de moléculas antileishmanicidas”*** subvencionado con fondos CANON de la Universidad Nacional San Antonio Abad del Cusco, aprobado por resolución N° R-0392-2018-UNSAAC.

Por otro lado, agradecer al proyecto ***“Secuenciación del metagenoma de bacterias halambientes salinos del Departamento de Cusco”*** con resolución N° R-005-2016-UNSAAC, *por ser la base de este trabajo.*

AGRADECIMIENTOS

Expresar en primer lugar mi más profundo agradecimiento a la Dra. María Antonieta Quispe Ricalde, por la formación, la orientación y la confianza depositada en mi persona.

A Blgo. Jose Luis Sierra Herrera por sus importantes sugerencias y aportes durante el desarrollo del presente trabajo.

A todos mis profesores de la escuela profesional de Biología por los consejos otorgados que atesoró y la formación brindada durante mi etapa de estudiante.

Al Dr. Basilio Balladares por aceptar mi pasantía en el Instituto Universitario de Enfermedades Tropicales y Salud Pública de Canarias de la Universidad de La Laguna (ULL) – España.

Al Dr. Enrique Carretero, director del Instituto Universitario de Enfermedades Tropicales y Salud Pública de Canarias de la Universidad de La Laguna – España, por autorizar mi pasantía en dicha institución.

A la Dra. Pilar Foronda y al Dr. Jose Antonio, por abrirme las puertas de sus respectivos laboratorios, por apoyarme durante el desarrollo del presente trabajo de investigación, por el buen ánimo y las palabras de apoyo que nunca olvidaré.

Al MsCs. Hugo Castellán Sanchez, Investigador del Centro de Investigación en Dinámica Celular de la Universidad Autónoma del Estado de Morelos (Morelos, México), por la capacitación y formación en bioinformática, por haber estado siempre ávido para colaborar con sus enseñanzas, y sobre todo por su amistad.

A mis amigos, Sharmely, Joss, Pame, Duly, Leito, Fercho y Bryan, quiénes con su cariño y apoyo incondicional en muchos momentos difíciles y gratos, contribuyeron a la culminación de este trabajo de investigación. Un agradecimiento singular merece la amistad y apoyo de mi compañero Kevin. Así mismo, he sido privilegiada en conocer a muchos otros amigos, que a pesar que no los nombro, a cada uno de ustedes les tengo presente.

Un agradecimiento muy especial merece la comprensión, paciencia y el ánimo recibido de parte de mi familia durante todo este tiempo.

A todos ustedes muchas gracias.

“Los antiguos, para conocer su destino, consultaban a los dioses.

Nosotros, modernos, aprendemos de los genes”.

F. Jacob

Premio Nobel de Medicina

TABLA DE CONTENIDO

RESUMEN	<i>i</i>
INTRODUCCIÓN	<i>ii</i>
PLANTEAMIENTO DEL PROBLEMA	<i>iii</i>
JUSTIFICACIÓN	<i>iv</i>
OBJETIVOS	<i>v</i>
OBJETIVO GENERAL	
OBJETIVOS ESPECIFICOS	
HIPÓTESIS	<i>vi</i>
VARIABLES	<i>vii</i>
VARIABLES INDEPENDIENTES	
VARIABLES DEPENDIENTES	
CAPITULO I	<i>1</i>
MARCO TEÓRICO	<i>1</i>
1.1 MARCO REFERENCIAL	<i>2</i>
1.1.1 ESTUDIOS DE ENZIMAS PROTEOLÍTICAS EN EL MUNDO	<i>2</i>
1.1.2 ESTUDIOS DE ENZIMAS PROTEOLITICAS EN EL PERÚ	<i>5</i>
1.1.3 ESTUDIOS DE ENZIMAS PROTEOLÍTICAS EN LA REGION DE CUSCO	<i>6</i>
1.2 MARCO CONCEPTUAL	<i>8</i>
1.2.1 BACTERIAS HALÓFILAS	<i>8</i>
1.2.2 GÉNERO <i>Chromohalobacter salexigens</i>	<i>9</i>
1.2.2.1 Posición sistemática	<i>9</i>
1.2.3 ÓMICAS	<i>10</i>
1.2.4 SECUENCIACIÓN	<i>11</i>
1.2.4.1 Secuenciación de primera generación	<i>13</i>
1.2.4.2 Secuenciación de próxima generación (NGS)	<i>14</i>
1.2.4.3 Secuenciación de tercera generación (TGS)	<i>14</i>
1.2.5 BIOINFORMÁTICA	<i>15</i>
1.2.5.1 Ensamblaje	<i>15</i>
1.2.5.2 Anotación	<i>16</i>
1.2.6 BASES DE DATOS Y PROGRAMAS BIOINFORMATICOS	<i>16</i>
1.2.6.1 FastQC	<i>16</i>
1.2.6.2 Trimmomatic	<i>17</i>
1.2.6.3 SPAdes	<i>18</i>
1.2.6.4 Quast	<i>18</i>

1.1.1.1.	RAST (Rapid Annotation using Subsystem Technology)	19
1.2.6.5	MEGA 7 (Molecular Evolutionary Genetic Analysis)	19
1.2.6.6	GENE RUNNER	19
1.2.6.7	I-Tasser	20
1.2.6.8	PDB (Protein Data Bank)	20
1.2.6.9	Blast (Basic Local Alignment Search Tool)	21
1.2.7	PROTEÍNAS	21
1.2.7.1	Composición de las proteínas	21
1.2.7.2	Niveles de organización de las proteínas	24
1.2.8	ENZIMAS	25
1.2.8.1	Clasificación y nomenclatura de las enzimas	25
1.2.8.2	Origen y uso de enzimas	26
1.2.9	PROTEASAS	28
1.2.9.1	Sistema EC de clasificación de proteasas	28
1.2.9.2	El sistema MEROPS de clasificación de proteasas	29
1.2.9.3	Fuentes de proteasas	30
1.2.10	HALOENZIMAS PROTEOLITICAS Y SU APLICACIÓN INDUSTRIAL	32
1.2.11	TÉCNICAS DE BIOLOGIA MOLECULAR	33
1.2.11.1	Extracción de ADN	33
1.2.11.2	Primers, cebadores u oligos	34
1.2.11.3	PCR - Amplificación del ADN mediante la técnica reacción en cadena de la polimerasa	35
1.2.11.4	Cuantificación del ADN	37
1.2.11.5	Electroforesis	37
1.2.11.6	Comprobación de la integridad del ADN por electroforesis	38
1.2.12	TECNOLOGÍA DEL ADN RECOMBINANTE	38
1.2.12.1	Clonación	38
1.2.12.2	Vectores de clonación	40
1.2.12.3	Transformación de células no competentes	41
1.2.12.4	PCR de colonias	44
1.2.12.5	Criopreservación de células con glicerol	44
CAPÍTULO II		46
MATERIALES Y MÉTODOS		46
2.1	LUGAR DE ESTUDIO	47
2.2	MATERIALES	47
2.2.1	MATERIAL BIOLÓGICO	47
2.2.2	MATERIAL DE LABORATORIO	47
2.2.2.1	Equipos de laboratorio	47
2.2.2.2	Material fungible y otros	48
2.2.3	MATERIAL BIOINFORMÁTICO	49
2.2.3.1	Datos primarios	49
2.2.3.2	Bases de datos	49

2.2.3.3	Softwares y programas informáticos	49
2.3	METODOLOGÍA	50
2.3.1	TIPO DE INVESTIGACIÓN	50
2.3.2	FLUJOGRAMA DEL TRABAJO DE INVESTIGACIÓN	50
2.4	MÉTODOS Y FUNDAMENTOS	52
2.4.1	TRABAJO IN SILICO	52
2.4.1.1	Control de calidad de las secuencias crudas	52
2.4.1.2	Recorte de las secuencias	53
2.4.1.3	Ensamble de las secuencias crudas de adn de <i>C. salexigens</i> .	53
2.4.1.4	Evaluación de calidad del ensamble	53
2.4.1.5	Anotación del genoma <i>C. salexigens</i> MP25462	54
2.4.1.6	Descripción de proteasas del genoma de <i>C. salexigens</i> MP25462.	54
2.4.1.7	Modelamiento estructural de la proteasa Amidohidrolasa MP25462	54
2.4.1.7	Diseño de cebadores	55
2.4.2	TRABAJO EN LABORATORIO	56
2.4.2.1	Recuperación a partir de la criopreservación de <i>C. salexigens</i> MP25462.	56
2.4.2.2	Preparación de células competentes <i>E. coli</i> JM109	56
2.4.2.3	Aislamiento y purificación del ADN genómico de <i>C. salexigens</i> MP25462	58
2.4.2.4	Reacción en cadena de la polimerasa (PCR)	60
2.4.2.5	Estandarización de las reacciones de PCR.	60
2.4.2.6	PCR para amplificar el gen Amidohidrolasa	62
2.4.2.7	Electroforesis	63
2.4.2.8	Purificación de la Amidohidrolasa	64
2.4.2.9	Secuenciación de la Amidohidrolasa	66
2.4.2.10	Análisis bioinformático del gen Amidohidrolasa	66
2.4.2.11	Clonación	66
2.4.2.12	PCR de colonias clonadas	69
2.4.2.13	Extracción del ADN plasmídico por lisis alcalina	70
2.4.2.14	Secuenciación de la Amidohidrolasa clonado	72
2.4.2.15	Criopreservación de células clonadas	72
CAPITULO III		73
RESULTADOS Y DISCUSIÓN		73
3.1	RESULTADOS	74
3.1.1	ORGANIZACIÓN DE LAS SECUENCIAS CRUDAS DE ADN DE <i>C. salexigens</i> .	74
3.1.2	CALIDAD DEL ENSAMBLE DE LAS SECUENCIAS DEL GENOMA <i>C. salexigens</i> MP25462	82
3.1.3	ANOTACIÓN DEL GENOMA <i>C. salexigens</i> MP25462 CON EL SERVIDOR RAST.	83
3.1.4	DESCRIPCIÓN DE PROTEASAS DEL GENOMA DE <i>C. salexigens</i> MP25462.	85
3.1.4.1	Endopeptidasas	86
3.1.4.2	Exopeptidasas	88
3.1.5	DESCRIPCIÓN DE LA PROTEASA AMIDOHIDROLASA	90
3.1.5.1	Modelamiento estructural de la proteasa Amidohidrolasa MP25462	92

3.1.6 DISEÑO DE CEBADORES	96
3.1.7 RECUPERACIÓN A PARTIR DE LA CRIOPRESERVACIÓN DE <i>C. salexigens</i> MP25462	98
3.1.8 AISLAMIENTO DEL ADN GENÓMICO DE <i>C. salexigens</i> MP25462.	98
3.1.9 ESTANDARIZACIÓN DE LA PCR PARA LA AMPLIFICACIÓN DEL GEN AMIDOHIDROLASA	99
3.1.10 AMPLIFICACIÓN DEL GEN AMIDOHIDROLASA POR PCR	100
3.1.11 PURIFICACIÓN Y SECUENCIACIÓN DEL GEN AMIDOHIDROLASA.	101
3.1.12 CLONACIÓN EN EL VECTOR pGEM-T.	105
3.1.13 SECUENCIACIÓN DE LA PROTEASA AMIDOHIDROLASA CLONADO	106
3.1.14 CRIOPRESERVACIÓN DE CÉLULAS CLONADAS	107
3.2 DISCUSIÓN	108
3.2.1 DE LAS CARACTERÍSTICAS GENERALES DEL GENOMA DE <i>Chromohalobacter salexigens</i> .	108
3.2.2 DE LAS PROTEASAS DEL GENOMA DE <i>Chromohalobacter salexigens</i> MP25462	110
3.2.3 DE LA OBTENCIÓN DE UNA PROTEASA RECOMBINANTE	113
CONCLUSIONES	114
RECOMENDACIONES	115
BIBLIOGRAFÍA	116

ANEXOS

ANEXO 1. _____	144
<i>Fichas que describen a las 36 enzimas proteolíticas encontradas en C. salexigens MP25462.</i> _____	144
ANEXO 2. _____	180
<i>Preparación de dNTPs a 2mM</i> _____	180
ANEXO 3. _____	180
<i>Preparación de tampón de caraga para electroforesis</i> _____	180
ANEXO 4. _____	181
<i>Preparación de gel de Agarosa al 1%</i> _____	181
ANEXO 5. _____	181
<i>Preparación de medio SW al 15%</i> _____	181
ANEXO 6. _____	182
<i>Preparación de medio LB con Ampicilina 100ug/ml</i> _____	182
ANEXO 7. _____	182
<i>Preparación de medio de recuperación SOC</i> _____	182
ANEXO 8. _____	183
<i>Informe de préstamo de servicio para secuenciación del SEGAI (Servicios Generales de Apoyo a la Investigación de la Universidad de la Laguna)</i> _____	183

INDICE DE FIGURAS

Figura 1: Bacilos de <i>Chomohalobacter salexigens</i> . _____	9
Figura 2 Ciencias Ómicas _____	10
Figura 3: Desarrollo histórico de las generaciones de tecnologías de secuenciación en función de la longitud de lectura. _____	12
Figura 4: Pasos básicos en la identificación de la secuencia de un pequeño fragmento hipotético mediante la técnica de Sanger (didesoxi). _____	13
Figura 5: Estructura básica de un aminoácido. _____	22
Figura 6: Fuente de proteasas y ejemplos de sus diversos usos _____	32
Figura 7: Principio de la amplificación de una secuencia de ADN mediante PCR. _____	36
Figura 8: Panorama general de la clonación de genes con un plasmido bacteriano que muestra los diversos usos de los genes clonados. _____	39
Figura 9: Topoisformas de los vectores plasmídicos. A. ADN circular cerrado B. ADN enrollado C. ADN lineal de longitud completa. _____	40
Figura 10: El vector pGEM®-T (número de acceso GenBank® X65308). _____	41
Figura 11 Criopreservación y su efecto sobre la célula _____	45
Figura 12: Flujo de secuencia de actividades del trabajo de investigación in silico. _____	50
Figura 13: Flujo de secuencia de actividades del trabajo de investigación en laboratorio. ____	51
Figura 14: Parámetros de calidad que evalúa el software FastQC. _____	52
Figura 15: Proceso de obtención de células competentes para la etapa de la transformación (clonación). _____	57
Figura 16: Proceso de aislamiento y purificación del ADN genómico _____	59
Figura 17: Programación de la PCR en gradiente con el termociclador T Professional Gradient (Biometra) _____	62
Figura 18: Proceso de la purificación de la Amidohidrolasa _____	65
Figura 19: Proceso de la transformación en la clonación del gen amidohidrolasa. _____	68
Figura 20: Procedimiento de la extracción del ADN plasmídico por lisis alcalina _____	71
Figura 21: Calidad de las secuencias crudas de Forward y Reverse. El eje Y es el intervalo de calidad Phred, el eje X son los pb de la secuencia. Fuente: FastQC _____	76
Figura 22: Calidad de las secuencias crudas por celda de flujo. _____	77

Figura 23: Contenido de bases nitrogenadas en las secuencias crudas Forward y Reverse. ___	78
Figura 24: Contenido de G+C en la secuencias crudas de Forward y Reverse. _____	79
Figura 25: Calidad de las secuencias, después del Trimming según el índice de Phred. ____	81
Figura 26: Anotación del genoma <i>C. salexigens</i> MP25462 con el servidor RAST. _____	85
Figura 27: Endoproteasas de tipo serina encontradas en el genoma de <i>C. salexigens</i> MP25462	86
<hr/>	
Figura 28: Endoproteasas de tipo aspartil encontradas en el genoma de <i>C. salexigens</i> MP25462	86
<hr/>	
Figura 29: Endoproteasas de tipo metalo encontradas en el genoma de <i>C. salexigens</i> MP25462	87
<hr/>	
Figura 30: Endoproteasas de tipo treonina encontradas en el genoma de <i>C. salexigens</i> MP25462 _____	87
Figura 31: Exoproteasas de tipo amino encontradas en el genoma de <i>C. salexigens</i> MP25462	88
Figura 32: Exopeptidasas de tipo carboxi encontradas en el genoma de <i>C. salegigens</i> MP25462	89
<hr/>	
Figura 33: Exopeptidasas de tipo omega encontradas en el genoma de <i>C. salegigens</i> MP25462	89
<hr/>	
Figura 34: Ubicación en el genoma de la proteína Amidohidrolasa de <i>C. salexigens</i> MP25462	90
Figura 35: Secuencia de aminoácidos en formato FASTA de la Amidohidrolasa de <i>C. salexigens</i> MP25462. _____	90
Figura 36: Secuencia de nucleótidos en formato FASTA del gen Amidohidrolasa de <i>C. salexigens</i> MP25462. _____	91
Figura 37: Modelamiento de la estructura de la proteína Amidohidrolasa MP25462 con el programa I-TASSER _____	94
Figura 38: Visualización de los sitios de unión a los ligandos de Manganeso y sitio activo de la Amidohidrolasa MP25462 con el programa I-Tasser. _____	95
Figura 39: Secuencia del gen amidohidrolasa, donde se señalan los cebadores AMD1F y AMD1R _____	97
Figura 40: A. Cultivo fresco de <i>C. salexigens</i> MP25462 en medio líquido SW al 15% de NaCl. B. Colonias de <i>C. salexigens</i> MP25462 en Agar SW al 15% de NaCl. _____	98

Figura 41: Cuantificación en gel del ADN genómico de <i>C. salexigens</i> , Carril 1: Estándar Lambda HindIII 23kb (Bioline), Carril 2: Muestra de ADN genómico, Carril 3: Muestra blanco	99
<hr/>	
Figura 42: Estandarización de la temperatura de hibridación en PCR en gradiente, para la amplificación de la proteína Amidohidrolasa. Carril 1: Marcador molecular Hyper Lader II 2Kb (Bioline), Carril 2: Amplicón a 52°C, Carril 3: Amplicón a 54.9°C, Carril 4: Amplicón a 58.1°C, Carril 5: Amplicón a 61°C y Carril 6: Muestra blanco.	100
Figura 43: Amplificación del gen Amidohidrolasa a 53°C. Carril 1: Marcador molecular Hyper Lader II (Bioline), Carril 2, 3,4 y 5: Bandas del gen Amidohidrolasa, Carril 6: Muestra blanco.	101
<hr/>	
Figura 44: Cuantificación en gel de la proteasa Amidohidrolasa, Carril 1: Marcador molecular Hyper Lader II (Bioline) Carril 2: Muestra del amplicon de Amidohidrolasa purificado Carril 3: Muestra blanco	102
Figura 45: Imagen del cromatograma de la secuenciación de la proteasa amidohidrolasa amplificado	103
Figura 46: Alineación de secuencias. 1. Secuencia de la proteasa Amidohidrolasa, producto de la anotación 2. Secuencia de la proteasa Amidohidrolasa purificado, producto de la PCR. El segmento amarillo es la secuencia del cebador AmdF1.	103
Figura 47: Porcentaje de identidad del amplicón de Amidohidrolasa secuenciado en la base de datos del Blast- NCBI.	104
Figura 48: Secuencias de la base de datos alineadas con la secuencia de consulta (gen de Amidohidrolasa secuenciado)	104
Figura 49: Crecimiento de colonias transformadas con inserto Amd- vector pGEM-T	105
Figura 50: PCR colonia. Carril 1: Marcador molecular Hyper Lader II (Bioline) Carril 2- 16: colonias aisladas del 1 a 15. Último carril: blanco de reacción.	106
Figura 51: Alineamiento entre secuenciación del plásmido clonado, Amidohidrolasa y cebadores AmdIF/AmdIR.	107
Figura 52: Distribución de las endopeptidasas del genoma de <i>C. salexigens</i> MP25462.	110
Figura 53: Distribución de las exopeptidasas del genoma de <i>C. salexigens</i> MP25462	112

INDICE DE TABLAS

Tabla 1: Clasificación de los organismos halófilos en base a la concentración óptima de NaCl para su crecimiento. _____	8
Tabla 2: Índice phred: Probabilidad de que la asignación de una base sea incorrecta. _____	17
Tabla 3: Clasificación y abreviatura de los 20 α -aminoácidos. _____	23
Tabla 4: Enzimas industriales según su uso _____	27
Tabla 5: Clasificación de las proteasas. _____	29
Tabla 6: Distribución de familias de peptídicas de organismos de diferentes reinos en el sistema propuesto por MEROPS (http://merops.sanger.ac.uk , consultado 12/11/2021). _____	30
Tabla 7: Marcadores genéticos en <i>E. coli</i> JM109 _____	43
Tabla 8: Sistema de dilución para cada componente de la PCR. _____	60
Tabla 9: Sistema del mix de PCR para un volumen final de siete reacciones. _____	61
Tabla 10: Sistema del mix de PCR para un volumen final de 5 reacciones _____	63
Tabla 11: Componentes de la reacción de ligación _____	67
Tabla 12: Sistema de reacción para la PCR de colonias clonadas _____	69
Tabla 13: Parámetros de calidad de las secuencias Forward y Reverse con el programa FastQC _____	74
Tabla 14: Estadísticas básicas del archivo forward _____	75
Tabla 15: Parámetros de calidad de las secuencias Forward y Reverse después del Trimming _____	80
Tabla 16: Parámetros de calidad del genoma ensamblado de <i>C. salexigens</i> MP25462 _____	82
Tabla 17: Número de genes agrupados en 27 funciones de los subsistemas. _____	84
Tabla 18: Proteínas estructuralmente cercanas al modelado de la Amidohidrolasa MP25462 _____	93
Tabla 19: Especificaciones del diseño de los cebadores. _____	96
Tabla 20: Valores de la cuantificación del ADN genómico de <i>Chromohalobacter salexigens</i> . _____	98
Tabla 21: Valores de la cuantificación del ADN del gen Amidohidrolasa purificado. _____	101
Tabla 22: Valores de la cuantificación del ADN del gen Amidohidrolasa purificado _____	106
Tabla 23: Características generales de reportes de genomas de <i>Chromohalobacter salexigens</i> ensamblados y anotados. _____	109

ABREVIATURAS

ADN:	Acido desoxirribonucleico
ARN:	Acido ribonucleico
ARNm:	ARN mensajero
ADNsa:	Desoxirribonucleasa
A, T, C, G y U:	Adenina, Timina, Citocina, Guanina y Uracilo.
D.O:	Densidad óptica
dNTP:	Desoxirribonucleotidos trifosfato
EDTA:	Acido eilendiaminotetracético.
kb:	Kilobases
kDa:	Kilodaltons
pb:	Pares de bases
ORF:	“open reading frame” pauta de lectura abierta
M:	Molar
μl:	Microlitro
ng:	Nanogramos
nm:	Nanometros
U:	Unidad
Tris:	Trihidroximetil aminometano
pI:	Punto isoeléctrico
Tm:	Temperatura de <i>melting</i> o también llamada temperatura de fusión del ADN

GLOSARIO

Lectura: o *read*, pequeño segmento de bases consecutivas de ADN identificado mediante un instrumento de secuenciación.

Genoteca (*library*): Es el conjunto de fragmentos de ADN amplificados durante el proceso de secuenciación donde cada uno contiene la fuente original de ADN, es decir, debería contener muestras de todo el ADN del organismo.

Sintenia: Conservación del orden génico y la orientación a lo largo del cromosoma; un elevado grado de sintenia es indicativo de proximidad filogenética, es decir, es la alineación de las secuencias de ADN entre dos organismos que muestran alta similaridad.

Contig: Son segmentos de ADN superpuestos, que juntos representan una región consenso de ADN, es decir, define una secuencia de ADN reconstruido a partir de un conjunto de lecturas. Agrupación de secuencias o lecturas

Fasta: En bioinformática, el formato FASTA (.fasta, .fna, .ffn, .faa, .frn), es un formato de fichero informático basado en texto, utilizado para representar secuencias bien de ácidos nucleicos, bien de péptido, y en el que los pares de bases o los aminoácidos se representan usando códigos de una única letra.

Sybr Safe: Es un agente intercalante (se intercala entre las bases nitrogenadas) que se usa como colorante fluorescente para la visualización de ácidos nucleicos en geles de agarosa.

Dímero: Refiere a dos subunidades estructuralmente similares denominadas monómeros unidas por enlaces que pueden ser fuertes o débiles. Es uno de los casos de daño más generales en el ADN, el dímero puede ocurrir entre pares adyacentes de bases pirimidínicas (C, T, U) entre dos residuos de timina, entre 2 citosinas o una citosina y un uracilo, muchas veces catalizado por la radiación ultravioleta o por agentes químicos mutagénicos, puede ser producto de la acción de un clastógeno.

Cebadores: primers, oligos u iniciadores, oligonucleótidos que son, cada uno, complementarios a una de las dos hebras del ADN.

Iones divalentes: Se suele usar magnesio (Mg^{+2}), agregado comúnmente como cloruro de magnesio ($MgCl_2$), o algún otro catión divalente.

Buffer: También llamados disoluciones amortiguadoras o sistemas tampón, son aquellas disoluciones cuya concentración de protones tienden a impedir la variación del pH al añadir ácidos o bases fuertes.

ADN molde: que contiene la región de ADN que se va a amplificar.

RESUMEN

Las enzimas tienen un importante rol dentro de la biotecnología, gracias a su mecanismo de acción se realizan diferentes procesos que son importantes para la ciencia y el desarrollo industrial, lo cual ha generado una amplia información acerca de estos biocatalizadores (Fasim, *et. al.*, 2021), sin embargo, las investigaciones disminuyen cuando se trata de exhibir a la totalidad de genes que codifican enzimas contenidas en el genoma de un microorganismo, es así que la presente investigación se enfocó en la descripción de todas las enzimas proteolíticas que codifican los genes del genoma de la bacteria halófila *Chromohalobacter salexigens* MP25465 y a su vez la obtención de una proteasa recombinante a partir del análisis de datos del genoma de dicha bacteria.

Este trabajo se divide en dos fases: *in silico* y en laboratorio. Para la primera fase, se organizó las secuencias crudas de ADN, realizando control de calidad con los programas bioinformáticas FastQC y Trimmomatic, ensamblaje de *novo* con el software Spades, la anotación y la ubicación con el servidor RAST y a partir de estos datos se describió las secuencias que expresan para enzimas proteolíticas y se diseñó una pareja de cebadores para aislar la proteasa Amidohidrolasa para el trabajo de la segunda etapa experimental que consistió en amplificar la secuencia de la Amidohidrolasa con la pareja de cebadores diseñados para luego clonarlo en un sistema heterólogo.

Como resultado se obtuvo el genoma ensamblado de *C. salexigens* MP25462 con una longitud de 3'712,216 pb, la anotación predijo 263 genes implicados en el metabolismo de proteínas de los cuales se identificó 36 enzimas proteolíticas, 21 endopeptidasas y 15 exopeptidasas. Las endopeptidasas son de tipo serina, aspartil, metalo y treonina y las exopeptidasas son de tipo aminopeptidasa, carboxipeptidasa y omegapeptidasa; así mismo se logró aislar y clonar *in vitro* el gen completo de la proteasa Amidohidrolasa con una longitud de 1013 pb, utilizando la pareja de cebadores diseñados AMD1F 5'GGTCTGGGTGGTCATGGC 3' y AMD1R 5'CATCCGACTGATGGAGCCTC 3'.

Palabras clave: secuenciación de próxima generación, ensamblaje, anotación, bioinformática, enzimas proteolíticas, amidohidrolasa, clonación.

INTRODUCCIÓN

En el departamento de Cusco existen ecosistemas salinos que tiene su origen debido a la formación de la cordillera de los Andes que causo la retención de masas de agua oceánica en la superficie continental y que por evaporación llegaron a formar los depósitos de halita (Marocco & Garcia-Zabaleta, 1974), que hoy en día son ambientes para el estudio de bacterias halófilas ya que son una fuente de enzimas proteolíticas por su importancia en la biomedicina y en procesos biotecnológicos debido a su actividad y estabilidad a altas concentraciones de salinidad (Mokashe *et al.*, 2018).

La secuenciación ha permitido adquirir información de genes de cualquier ácido nucleico que existe en un genoma o metagenoma mediante diversas tecnologías presentes en plataformas como Sanger, PacBio, Ion Torrent e Illumina, producto de este proceso se obtienen las lecturas de la secuenciación o secuencias crudas, cuya única forma de análisis es mediante la bioinformática, que de forma rápida y versátil identifica y caracteriza *in silico* las secuencias de genes de interés de un organismo u organismos (Hernández, *et al.*, 2020). Posterior a este proceso se amplifica el gen diana mediante PCR y por medio de la técnica de ADN recombinante se obtiene miles de copias del amplificado para su mejor estudio.

En la última década los estudios de proteasas aumentaron exponencialmente a nivel global (Uritskiy & DiRuggiero, 2019), así mismo, en la región del Cusco realizaron la evaluación de actividad enzimática hidrolítica de bacterias halofilas (Gárate, 2016; Romoacca, 2018), la purificación de metaloproteasas (Huaihua, 2019) y el análisis *in silico* y clonación de una haloproteasa (Pacheco, 2021).

El presente trabajo trata de la descripción de enzimas proteolíticas codificadas por los genes del genoma del halófilo moderado *Chromohalobacter salexigens* MP25462 que fue aislado previamente del salar de Maras, Urubamba; y secuenciado por tecnología MiSeq de la plataforma Illumina, dentro del proyecto “Secuenciación del metagenoma de ambientes salinos del departamento de Cusco” (2015-2019). A partir de sus datos NGS (*Next Generation Sequencing*) se planteó efectuar el correcto análisis bioinformático de su genoma para describir todos los genes proteolíticos; esta etapa bioinformática se corroboró amplificando un gen proteolítico e insertándolo en un vector de clonación.

PLANTEAMIENTO DEL PROBLEMA

El rápido desarrollo de la tecnología de secuenciación de próxima generación (NGS), la bioinformática, el análisis e interpretación de datos, ha cobrado impulso en los últimos años a medida que, en todo el mundo, los científicos se han unido en una búsqueda común para secuenciar, almacenar y analizar genomas completos con la finalidad de dilucidar y caracterizar los genes de muchos genomas (Manzoni, *et al.*, 2018), siendo así los genes de las proteasas de microorganismos extremófilos los que han cobrado mayor interés por la importancia biotecnológica que poseen en diversos campos de la biomedicina, industria y biorremediación; así por ejemplo una Serin-proteasa alcalina termoactiva en detergentes, disolventes y sales, de *Bacillus sp.* que promete su aplicación en las industrias de lavandería y farmacéutica (Ibrahim *et al.*, 2015), otro organismo halófilo es la *Idiomarina sp.* C9-1 que posee una proteasa alcalina con aplicación potencial para depilación enzimática ecológica en la industria del cuero (Zhou *et al.*, 2018), y demás proteasas implicadas en la producción de papel, la degradación de almidón, las fábricas de procesamiento de alimentos, etc. Por lo tanto, las proteasas están creando un punto de referencia en cada campo con su amplio potencial y es una de las áreas de exploración que se expande más rápidamente (Solanki, *et al.*, 2021), por ello es imprescindible el análisis bioinformático ya que estos datos son la base para la obtención y producción de proteasas recombinantes, lo que da lugar a la ingeniería de genes de proteasas para mejorar la calidad y cantidad del producto. Sin embargo, hasta el momento no se tiene antecedentes del total de proteasas codificados por genes del genoma de un microorganismo, particularmente un halófilo, es así que se evidencia la necesidad de datos al respecto. Por lo cual surge la interrogante:

¿Qué información sobre enzimas proteolíticas podemos encontrar en los datos del genoma de *Chromohalobacter salexigens*? y, a partir de estos datos ¿Se puede obtener una proteasa recombinante?

JUSTIFICACIÓN

Actualmente, se observa que la sociedad esta optando por consumir productos amigables con el medio ambiente dando lugar a que inversores de diversas industrias innoven en sus tecnologías de producción y descarte de agentes corrosivos y dañinos, como la mezcla de metales tóxicos, los plásticos, los productos químicos manufacturados, el petróleo, los desechos urbanos e industriales, los pesticidas, fertilizantes, los productos químicos farmacéuticos, etc. (Landrigan *et al.*, 2020). Por ello, las enzimas como biocatalizadores, específicamente, las proteasas microbianas (Banerjee & Ray, 2017; Razzaq *et al.*, 2019), son una alternativa al uso de productos químicos peligrosos, ya que generan una producción mas limpia que ayuda a mitigar el riesgo de contaminación, por lo tanto ascendió la búsqueda de proteasas en el mercado mundial generando una creciente demanda en el mercado enzimático (Dumorné *et al.*, 2017; Flores *et al.*, 2010).

Asi mismo, surge el interés por las proteasas halófilas debido a las características que poseen estos microorganismos a partir del estrés salino, como son: estabilidad y actividad a altas concentraciones de sal (Madigan & Oren, 1999), son tolerantes a pH alcalino, a altas temperaturas, a medios salinos con baja actividad hídrica, debido a que tienen suficiente agua para mantener una distribución de carga adecuada en el sitio activo manteniendo la conformación de la enzima (Dumorné *et al.*, 2017; Zaccai, 2004), y son estables en presencia de solventes orgánicos (Oren, 2010; de Lourdes Moreno *et al.*, 2013).

Por estas razones se justifica el propósito de describir todas las proteasas que alberga una bacteria halófila, motivando así, ensayos experimentales a partir de un repertorio de proteasas donde cada una será de interés para realizar investigaciones en biomedicina y biotecnología. De igual manera, se justifica la obtención de una proteasa recombinante a partir de un correcto diseño de primers, lo que asevera la posición fidedigna del gen dentro del genoma de *C. salexigens*, además que la técnica de clonación otorga una ventaja sobre otros métodos convencionales de síntesis de proteínas.

OBJETIVOS

OBJETIVO GENERAL

Describir las enzimas proteolíticas codificados por los genes de *Chromohalobacter salexigens* MP25462 y a partir de estos datos obtener una proteasa recombinante.

OBJETIVOS ESPECIFICOS

1. Anotar el genoma de *Chromohalobacter salexigens* MP25462.
2. Buscar las secuencias de los genes que codifiquen a enzimas proteolíticas en el genoma de *Chromohalobacter salexigens* MP25462.
3. Caracterizar *in silico* a los genes de las enzimas proteolíticas del genoma de *Chromohalobacter salexigens* MP25462.
4. Diseñar el clonaje de una enzima proteolítica de *Chromohalobacter salexigens* MP25462.
5. Clonar la enzima proteolítica en un sistema heterólogo.

HIPÓTESIS

Utilizando el análisis de datos del genoma de *Chromohalobacter salexigens* MP25462 se puede describir a las enzimas proteolíticas y conseguir una proteasa recombinante a través de la clonación.

VARIABLES

VARIABLES INDEPENDIENTES

- Programas bioinformáticos de recorte, ensamblaje y anotación del genoma.
- Cebadores para amplificar el gen de la proteasa Amidohidrolasa.
- Temperatura de hibridación de los cebadores en la reacción de PCR.
- Plásmidos para clonación.

VARIABLES DEPENDIENTES

- Genoma de *Chromohalobacter salexigens* cepa MP25462.
- Secuencia de la proteasa Amidohidrolasa.
- Clonación del gen de la Amidohidrolasa.

CAPITULO I
MARCO TEÓRICO

1.1 MARCO REFERENCIAL

1.1.1 ESTUDIOS DE ENZIMAS PROTEOLÍTICAS EN EL MUNDO

Lee *et al.*, (1994), en Japón revelaron que la carboxipeptidasa Taq es un nuevo tipo de metalocarboxipeptidasa dependiente de zinc, esto debido a estudios de clonación y secuenciación de este gen carboxipeptidasa Taq, que vendría a ser una metalocarboxipeptidasa termoestable de *Thermus aquaticus* YT-1. Además dieron a conocer que el gen comprendía un marco de lectura abierto de 1536 pares de bases con un codón de iniciación GTG y un codón de terminación TGA, que codifica una proteína de 56,210 Da que consta de 511 residuos de aminoácidos. El análisis de iones metálicos unidos a la enzima encontró que una molécula de la enzima contiene un ión de zinc fuertemente unido y la comparación de la secuencia completa mostró que la enzima no tiene similitud de secuencia obvia con otras metalopeptidasas. Sin embargo, se encontró una secuencia His-Glu-XX-His, que es una secuencia conservada en el sitio activo de endopeptidasas y aminopeptidasas dependientes de zinc, en las posiciones 276 a 280 de la enzima.

Oren *et al.*, (2005), en Israel analizaron la composición de aminoácidos de diferentes categorías de proteínas de la bacteria moderadamente halófila *Chromohalobacter salexigens*, y la compararon con proteínas de representantes no halófilos de las γ -proteobacterias (*Escherichia coli* , *Pseudomonas aeruginosa* , *Vibrio cholerae*) dando a conocer que existe una diferencia importante entre las proteínas periplásmicas de *C. salexigens* halófilo moderado y las bacterias no halófilas, por ejemplo las proteínas de unión al sustrato de los sistemas de transporte.

Rashid & Kamran Azim, (2011), en Pakistán realizaron bioinformática estructural de la proteína NMB1620 utilizando modelos de homología y acoplamiento de ligando, esta proteína fue anotada como LD-carboxipeptidasa de *Neisseria meningitidis* involucrada en el reciclaje de peptidoglicanos, y dieron a conocer que su genoma posee 2160 regiones codificantes predichas, incluidos 1000 genes hipotéticos y en una segunda anotación se identificó nueve peptidasas putativas. La comparación estructural del sitio de unión al sustrato de LD-carboxipeptidasa se realizó basándose en la unión del sustrato tetrapéptido 'L-alanil-D-glutamil-meso-diaminopimelil-D-alanina'. La inspección de los diferentes residuos formadores de subsitios mostró una variabilidad en el subsitio S1 a través de diferentes especies bacterianas.

Yamaguchi et al., (2012) en Japon clonaron una proteína periplasmática de unión a metales de un halófilo moderado, *Chromohalobacter salexigens* caracterizada por un alto contenido de histidina. La proteína, denominada proteína de unión a metales rica en histidina (HP), se expresó y se purificó a partir de *E. coli* como forma nativa. HP se unió a columnas de quelato cargadas con Ni^{2+} y Cu^{2+} con alta afinidad, y columnas de Co^{2+} y Zn^{2+} con afinidad moderada.

Jamdar et al., (2015) en la India, caracterizaron tres miembros de la familia de peptidasas M20D de *Burkholderia cepacia* (BcepM20D), *Deinococcus radiodurans* R1 (DradM20D) y *Staphylococcus aureus* (HmrA), para varios sustratos. Por tanto, los resultados revelan que todas las enzimas, incluida la HmrA, carecen de actividades endopeptidasa y aminopeptidasa y poseen una fuerte actividad carboxipeptidasa. Además, la actividad amidohidrolasa ejercida sobre otros sustratos como N-Acetil-Aminoácidos, N-Carbobenzoxil-Aminoácidos y Ácido indol acético (IAA) -Aminoácidos se debe a la capacidad de estas enzimas para acomodar diferentes tipos de grupos químicos distintos del aminoácido en el bolsillo S1. La evidencia de apoyo para las especificidades del sustrato también se proporcionó con los estudios de acoplamiento molecular llevados a cabo utilizando la estructura de SACOL0085 y la estructura modelada por homología de BcepM20D por lo cual discuten la preferencia por diferentes sustratos, su unión en el sitio activo de la enzima y el posible papel de estas enzimas en el reciclaje de metionina.

Lylloff et al., (2016), en Dinamarca revelaron una abundancia de proteasas capaces de secretar una amplia gama de homólogos de subtilisina y otras oligo-, di- y exopeptidasas de bacterias adaptadas al frío y pH alcalino, todo esto mediante análisis genómicos y exoproteómicos, donde identificaron dos aislados relacionados del género *Arsukibacterium* como principales productores de proteasa extracelular de dos ambientes fríos y alcalinos y es probable que la actividad extracelular observada sea el resultado de varias actividades endoproteolíticas.

Surachat et al., (2017), en Tailandia realizan el análisis bioinformático integral y la caracterización de las toxinas proteicas producidas por *Lactobacillus paracasei* SD1, identifican y analizan las toxinas proteicas y las proteínas antimicrobianas utilizando recursos confiables de bases de datos en línea y establecen su relación filogenética con otros genomas bacterianos. Su investigación sugiere que esta cepa es segura para uso humano y contiene varias bacteriocinas que confieren beneficios para la salud del huésped.

Da Silva R. (2017), en Brasil, dan a conocer la importancia biotecnológica de cada subfamilia de las enzimas proteolíticas microbianas y fúngicas, revelando sus propiedades bioquímicas y sus potenciales aplicaciones en productos farmacéuticos, alimentos, síntesis de péptidos bioactivos, estudios de investigación básica, tratamiento de cueros e industria de detergentes, entre otros.

Contesini et al., (2018), en Brasil, describieron las enzimas proteolíticas que presentan las especies del género *Bacillus*. Esta investigación, identificó con la base de datos PDB (protein data bank) 34% serina proteasas, 28% metaloproteasas, 22% cistein proteasas, 7% aspartil proteasa, 4% treonina proteasas y concluyeron que la mayoría de las proteasas extracelulares producidas por *Bacillus* son serina proteasas (*B. subtilis*, *B. pumilus*), cisteína proteasas (*B. licheniformis*) y metaloproteasas (*B. stearothermophilus*).

Srivastava et al., (2019), en la India, aislaron la cepa *Chromohalobacter salexigens* ANJ207 y analizaron *in silico* las métricas de su genoma, que se secuenció con una plataforma Illumina HiSeq. Utilizaron una puntuación Phred de ≥ 20 , obtuvieron 26,75 millones de lecturas de alta calidad. El ensamblaje se realizó con Velvet v.1.2.10, generando un total de 116 contigs en 33 andamios con un tamaño de genoma de 3.664.372 pb (contenido G+C, 63.71%). La anotación del genoma se realizó mediante Anotación de subsistemas (RAST) v.2.0, que predijo un número total de 3406 genes. De los 3406 genes predichos, 3344 eran genes codificadores de proteínas y 62 eran genes no codificantes de proteínas.

Jutras et al., (2020), en Reino Unido analizaron el repertorio de proteasas de *Nicotiana benthamiana* que puede afectar la producción de proteínas recombinantes y los avances recientes en las estrategias de depleción de proteasas para aumentar la producción de proteínas recombinantes en *Nicotiana benthamiana*. Las últimas anotaciones del genoma indican que *N. benthamiana* codifica al menos 1243 proteasas putativas que probablemente actúan de forma redundante y consecutiva sobre sustratos en diferentes compartimentos subcelulares. Las proteasas putativas centrales incluyen 165 aspártico proteasas, 307 cisteína proteasas, 66 treonina proteasas, 207 metaloproteasas y 498 serina proteasas.

1.1.2 ESTUDIOS DE ENZIMAS PROTEOLITICAS EN EL PERÚ

Chavez-Gamarra & Sandoval (2014), estudiaron bioinformáticamente a la proteasa ClpP de *Streptococcus mutans*, responsable de la formación de la caries dental humana; para ello descargaron la secuencia de aminoácidos de la proteasa a partir de la base de datos del GenBank y determinaron sus parámetros bioquímicos, dominios conservados, predicción de estructuras secundarias. Las herramientas bioinformáticas que utilizaron son ProtParam, Prosite, PHD Secondary Structure Prediction, SWISS-MODEL y Pymol, este último para la visualización de los modelos tridimensionales de la proteasa. Como resultado, la proteasa ClpP de *S. mutans* es una molécula ácida y de mediano peso molecular y además presenta dos dominios altamente conservados relacionados con endopeptidasas bacterianas, por lo que resulta importante para el proceso infectivo de esta bacteria lo cual permitirá generar una variedad de estrategias para la inhibición de la colonización en la cavidad bucal.

Chamorro et al. (2015), realizaron el análisis bioinformático de la proteína SAG29 de *Arabidopsis thaliana* involucrada en la respuesta al estrés salino, estableciendo así, sus parámetros bioquímicos, los dominios conservados, predijeron la estructura secundaria y su modelamiento tridimensional. Dando a conocer que dicha proteína posee 292 aminoácidos y un peso molecular de 32.9 kDa, así como dominios conservados correspondientes a proteínas involucradas en el transporte de azúcares a través de la membrana plasmática de células vegetales. Además, tiene importancia en la regulación de la viabilidad celular y respuesta al estrés salino.

Punil & Sandoval (2016), estudiaron *in silico* a la proteína endolisina LysB4 empleada en el biocontrol de *Bacillus cereus* en la industria alimentaria; determinaron sus parámetros bioquímicos, dominios conservados, estructuras secundarias y modelamiento tridimensional, empleando herramientas bioinformáticas. Estableciendo así, que la proteína endolisina LysB4 posee 262 aminoácidos, un peso molecular de 27.8 kDa y un pI de 9.21, así como dominios conservados correspondientes a las superfamilias VanY y SH3, involucradas en su actividad catalítica, así como en la interacción con sustrato. Con respecto a su estructura secundaria, la endolisina LysB4 presenta 21.4% de hélices α , 28.6% de láminas β y 50% de *loops* internos.

1.1.3 ESTUDIOS DE ENZIMAS PROTEOLÍTICAS EN LA REGION DE CUSCO

Gárate (2016), estudió las enzimas proteolíticas producidas por cepas bacterianas halotolerantes aisladas de las salineras de Maras-Cusco. Las cepas N9 y N10 resaltaron por su actividad proteolítica en medios de gelatina y caseína, siendo así el rendimiento enzimático de las proteasas excretadas por la cepa N9 y N10 fue de 49.09 U/mL/min y 140.45 U/mL/min respectivamente, además al realizar los ensayos de inhibición de la actividad en sustratos copolimerizados, concluyó que la cepa N9 produce 3 proteasas, metalo proteasas, serín y aspártico proteasas; y la cepa N10, produce metaloproteasas y serín - cisteín - aspártico proteasas.

Romoacca (2018), seleccionó 3 ambientes salinos acuosos: las salineras de Maras, el afluyente salino de Huanoquite y el afluyente salino de Acos, cada uno de ellos en distintas provincias del Cusco, donde hizo la colecta de microorganismos, llegando a aislar 35 cepas bacterianas en el que determinó sus características morfológicas mediante la tinción de Gram teniendo como resultado 31 cepas Gram negativas y 4 cepas Gram positivas, seguidamente determinó su identidad mediante el análisis de las secuencias del gen ribosomal 16S, así mismo evaluó la actividad de proteasas en sustrato gelatina, lipasas con el sustrato tween80 y amilasas con el sustrato almidón en donde llegó a la conclusión que las bacterias halófilas aisladas, presentan actividad enzimática hidrolítica (proteolítica, lipolítica y amilolítica) positiva.

Leon (2018), determinó la actividad de proteasas intracelulares y extracelulares de las bacterias halófilas aisladas de los manantiales salinos de los distritos de Acos y Huanoquite. Para conseguirlo, primero realizó una búsqueda de aquellas cepas bacterianas aisladas con actividad proteolítica, una vez encontradas, las cepas correspondientes crecieron en medio líquido SW (Medio de cultivo Sea Water) a partir del cual se obtuvieron las proteínas extracelulares y luego las proteínas intracelulares. Luego, determinó la actividad semicuantitativa de las proteasas extracelulares e intracelulares y finalmente cuantificó la concentración de proteínas de cada uno de los extractos crudos. Concluyó que las cepas M3H159 y M3H1027 fueron las que mostraron actividad proteolítica cualitativa formando halos de actividad al hidrolizar la gelatina, en cuanto a su análisis semicuantitativo los extractos extracelulares e intracelulares hidrolizaron la gelatina contenida en el medio de cultivo. La cepa M3H159 mostró mayor diámetro de halo y fue la que obtuvo actividad proteasa extracelular e intracelular con 4.65mm/μg de proteína y 18.9mm/μg de proteína respectivamente

Mamani et al. (2019), aislaron la cepa del halófilo moderado *Halomonas elongata* MH25661 de un arroyo salino en Huanquite- Cusco, Peru, y realizaron la secuenciación genómica de dicha bacteria, para luego analizar bioinformáticamente las métricas de la secuenciación genómica resultando 375 660 secuencias de lectura y un tamaño entre 2 a 231 pb, así mismo, el ensamblaje de dichas secuencias se realizó con el programa A5 Miseq Linux, resultando el genoma de un tamaño de 3'552,403 pb agrupados en 117 contigs. Además, realizaron la anotación del genoma ensamblado que reveló genes que responden al estrés osmótico y al estrés oxidativo, genes involucrados en la esporulación y los genes que participan en los sistemas de transporte y resistencia a antibióticos y compuestos tóxicos.

Huaihua (2019), estudió la purificación, caracterización e inmovilización de proteasas de excreción de *Staphylococcus sp.* de las salineras de Maras obtenidos por fermentación sumergida que se concentraron por liofilización, pasaron por sistemas de cromatografía de afinidad, intercambio aniónico y filtración en gel. Los extractos de *Staphylococcus sp.* N10, obtuvieron una actividad enzimática de 29.25 U/mL/min y 28.2 U/mL/min en los medios de cultivo MH (Agar Mueller Hinton) y TSB (Caldo de Trypticase y soya), con un peso de 60 kDa, un grado de pureza de 13.56 veces más frente al extracto inicial y un rendimiento de 7.1 % de la actividad total, la enzima se determinó como una metaloproteasa, con temperatura y pH óptimo de 40°C y 7.0 respectivamente.

Pacheco (2021), caracterizó *in silico* y clonó el gen aspartil aminopeptidasa a partir de datos de secuenciación de próxima generación (NGS) del halófilo moderado *Chromohalobacter salexigens* MP25462. La secuencia nucleotídica del gen aspartil aminopeptidasa presentó una cobertura de 20.7 X y una longitud de 1299 pares de bases traducidas a una secuencia de 432 aminoácidos, dando por resultado su estructura tridimensional similar a la aspartil aminopeptidasa de *Homo sapiens* (DNPEP) y su funcionalidad idéntica a la proteína de la arquea *Thaumarchaeota archaeon* (TET) ubicando su sitio activo en el aminoácido 263 dependiendo de dos iones de cobalto para su actividad.

1.2 MARCO CONCEPTUAL

1.2.1 BACTERIAS HALÓFILAS

El término halófilo se origina del griego, *halo* es “sal” y *philo* es “amante de”. Los halófilos bacterianos son abundantes en ambientes tales como lagos salados, suelos salinos y productos alimenticios salados, los mismos que están adaptados a altas concentraciones de sal y a la alta presión osmótica de su entorno. Oren en 2008, revela que las bacterias halofilas usan dos estrategias para equilibrar su citoplasma osmóticamente con su medio:

- La primera estrategia implica la acumulación de concentraciones molares de KCl, esta estrategia requiere la adaptación de la maquinaria enzimática intracelular, ya que las proteínas deben mantener su conformación y actividad adecuadas a concentraciones de sal casi saturantes.
- La segunda estrategia es excluir la sal del citoplasma y sintetizar y / o acumular solutos orgánicos "compatibles" que no interfieran con la actividad enzimática.

Kushner en 1978 clasificó a los organismos halófilos en base a la concentración óptima de NaCl para su crecimiento como se muestra en la tabla 1:

Tabla 1: Clasificación de los organismos halófilos en base a la concentración óptima de NaCl para su crecimiento.

CATEGORÍA	DESCRIPCIÓN	EJEMPLO
No halófilos	Crecimiento óptimo: NaCl inferior a 0.2 M (1 % p/v).	La mayoría de las bacterias de agua fresca
Halófilos débiles	Crecimiento óptimo: NaCl de 0.2 -0.5 M (1-3 % p/v).	Generalmente bacterias marinas
Halófilos moderados	Crecimiento óptimo: NaCl 0.5 – 2.5 M (3-15 % p/v).	<i>Salinivibrio costicola</i>
Halófilos extremos	Crecimiento óptimo: NaCl 2.5 y 5.2 M (15 – 32 % p/v).	<i>Halorhodospira halophila</i>
Halotolerantes	Crecen en presencia y en ausencia de altas concentraciones de NaCl.	<i>Staphylococcus aureus</i>

Fuente: Kushner, 1978.

1.2.2 GÉNERO *Chromohalobacter salexigens*

Las bacterias *C. salexigens*, son bacilos móviles gramnegativos que no forman esporas, se presentan individualmente o en parejas. El crecimiento óptimo ocurre a 7.5-10% (p/v) de sales, con pH 7.5 y temperatura de 37 °C. En un medio sólido, las colonias son de color crema, opacas y circulares y menos de 2 mm de diámetro; en un medio líquido, se produce turbidez homogénea. Las bacterias de *C. salexigens* son estrictamente aeróbicos, con catalasa positiva y oxidasa negativa; producen ácido a partir de L-arabinosa, D-fructuosa, D-galactosa, glicerol, D-glucosa, lactosa, maltosa, D-manosa, sacarosa y D-xilosa pero no de trehalosa; reducen el nitrato a nitrito pero no se reduce el nitrito, citrato positivo; no hidrolizan la gelatina, el almidón, la aesculina, el ADN y el Tween 80, hidrolizan la caseína; no producen fenilalanina deaminasa, lisina descarboxilasa y ornitina descarboxilasa (Arahal *et al.*, 2001).

1.2.2.1 Posición sistemática

Dominio: Bacteria
Filo: Proteobacteria
Clase: Gammaproteobacteria
Orden: Oceanospirillales
Familia: Halomonadaceae
Género: *Chromohalobacter*
(Ventosa *et al.*, 1989)
Especie: *Chromohalobacter salexigens*
(Arahal *et al.*, 2001)

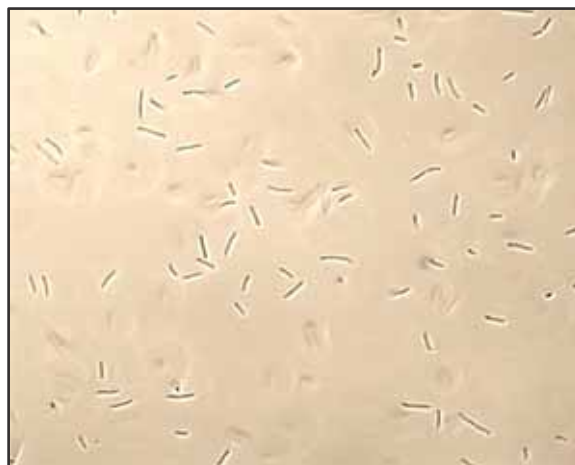


Figura 1: Bacilos de *Chomohalobacter salexigens*.

Fuente: Microbewiki, 2012.

1.2.3 ÓMICAS

Ómica es un neologismo derivado del alemán *genom Om* = oma, que se refiere al estudio de la totalidad o del conjunto de algo. Las ómicas, se definen como un grupo de disciplinas que tienen como objetivo el análisis de un gran número de moléculas a partir de muestras biológicas involucradas en la función de un organismo, con ayuda de las nuevas tecnologías y la formación de equipos multidisciplinarios que ayudan a la interpretación de los datos. Cada una de estas tecnologías tiene por objeto conocer el conjunto de todos los potenciales genes (genómica), de ARN expresados por los genes (transcriptómica), de proteínas (proteómica) o de sustancias metabólicas (metabolómica) que en un organismo puede contener en todas las condiciones (Frigolet & Gutiérrez-Aguilar, 2017)

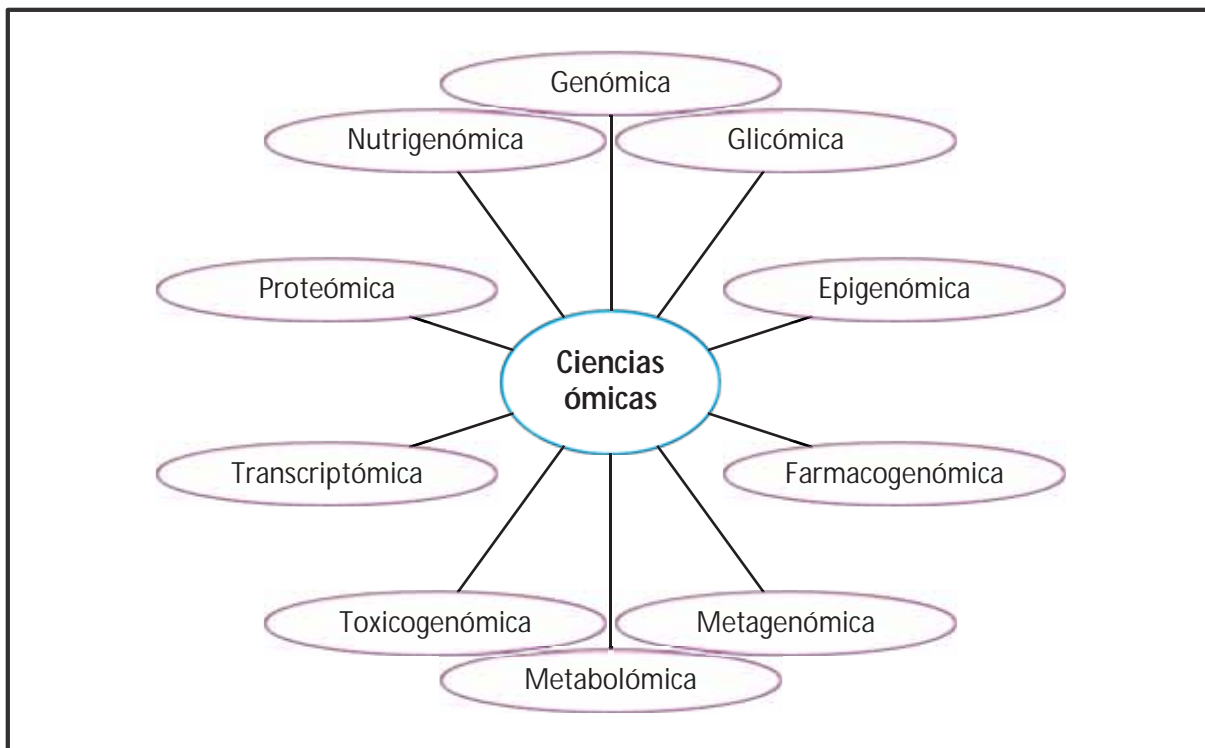


Figura 2: Ciencias Ómicas

✚ Genómica:

Estudia el genoma o ADN en una célula o tejido.

✚ Epigenómica:

Se refiere al conjunto de procesos por medio de los cuales se regula la transcripción de los genes sin afectar la secuencia del ADN.

✚ Transcriptómica:

Estudia todos los genes expresados en una célula o tejido.

✚ Proteómica:

Estudia todas las proteínas en una célula o tejido.

✚ Metabolómica:

Analiza proteínas y vías enzimáticas implicadas en el metabolismo celular.

✚ Toxicogenómica:

Analiza los efectos de los químicos tóxicos en los genes, incluidas las mutaciones creadas por las toxinas y los cambios en la expresión génica causados por las toxinas.

✚ Metagenómica:

El análisis de genomas de organismos recogidos del medio ambiente.

✚ Farmacogenómica:

Estudia el desarrollo de medicamentos personalizados basados en el perfil genético de una persona para una condición particular.

✚ Glicómica:

Analiza los carbohidratos de una célula o tejido.

✚ Nutrigenómica:

Estudia los cambios en la expresión de genes en respuesta al consumo de un nutrimento, alimento o dieta.

1.2.4 SECUENCIACIÓN

La secuenciación del ADN es cómo medimos una de las principales propiedades por las cuales las formas de vida terrestre pueden definirse y diferenciarse entre sí (Heather & Chain, 2016), mediante la determinación del orden de los cuatro componentes básicos químicos (A, T o U, C y G), llamados "bases", que forman la molécula de ADN o ARN. Las utilidades principales de la secuenciación son ampliar el conocimiento de la biología evolutiva, la biología forense, la biología de sistemas, la metagenómica y la medicina con sus aplicaciones clínicas. Es así, que durante las últimas décadas muchos investigadores han invertido mucho tiempo desarrollando diferentes tecnologías que mejoren aspectos de tiempo, costos, precisión y tamaño de lectura en pares de bases (Figura 3).

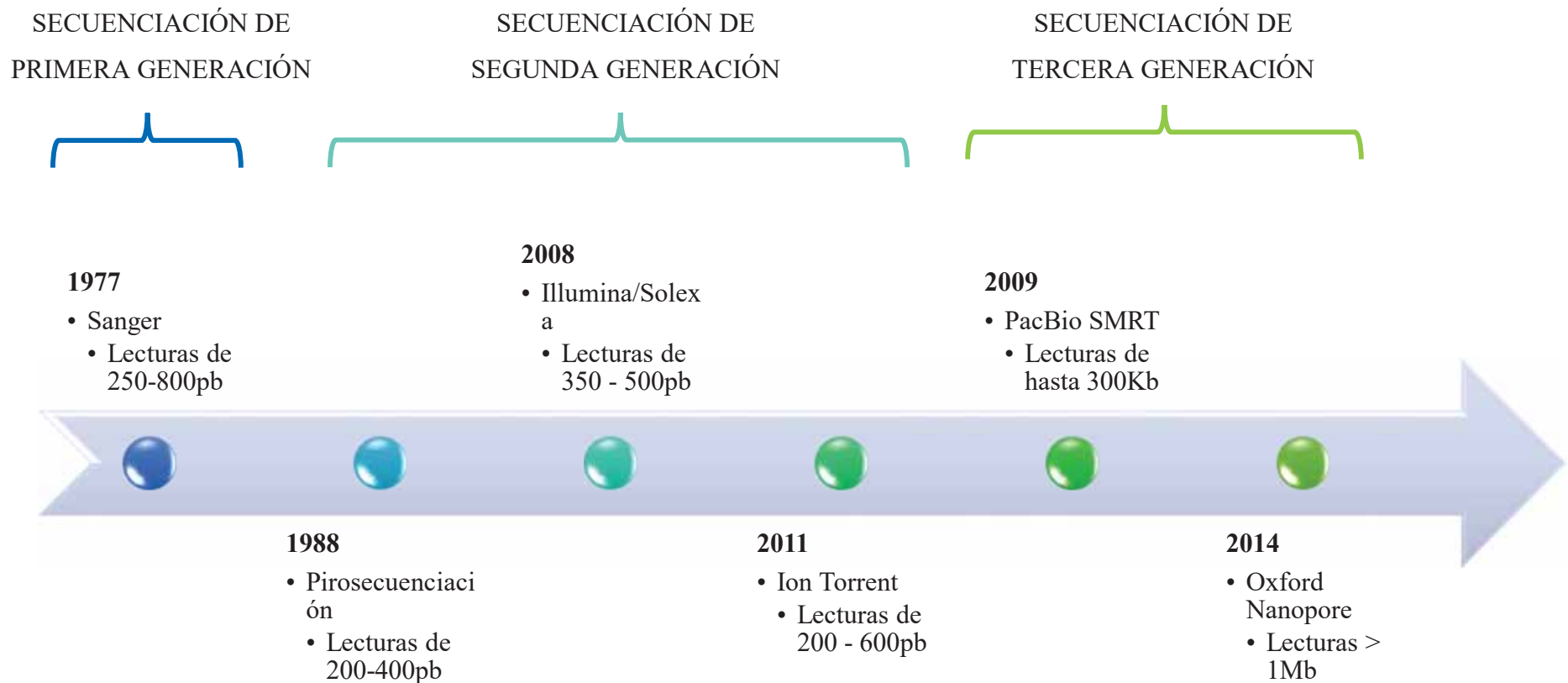


Figura 3: Desarrollo histórico de las generaciones de tecnologías de secuenciación en función de la longitud de lectura.

1.2.4.1 Secuenciación de primera generación

🚩 Método Sanger

El método de Sanger o método de los dideoxinucleótidos, es una secuenciación por síntesis ya que hace uso de la ADN polimerasa para sintetizar cadenas de ADN con una terminación específica, generando lecturas de 700-800 pb. Cada fragmento acabado en dideoxinucleótido está marcado fluorescentemente, es decir, que cada dideoxinucleótido (ddATP, ddCTP, ddGTP, ddTTP) están diseñados para que carezcan del grupo 3'-OH, lo que permite la adición del nucleótido consecutivo con un color de fluoróforo diferente, de este modo, se produce una "escalera" de ADN de fragmentos que difieren en una base de longitud estos son ordenados por tamaño en electroforesis capilar, con detección automatizada de los fragmentos de ADN marcados fluorescentemente, proporcionando la secuencia ordenada de los fragmentos en cromatogramas (Sanger *et al.*, 1977).

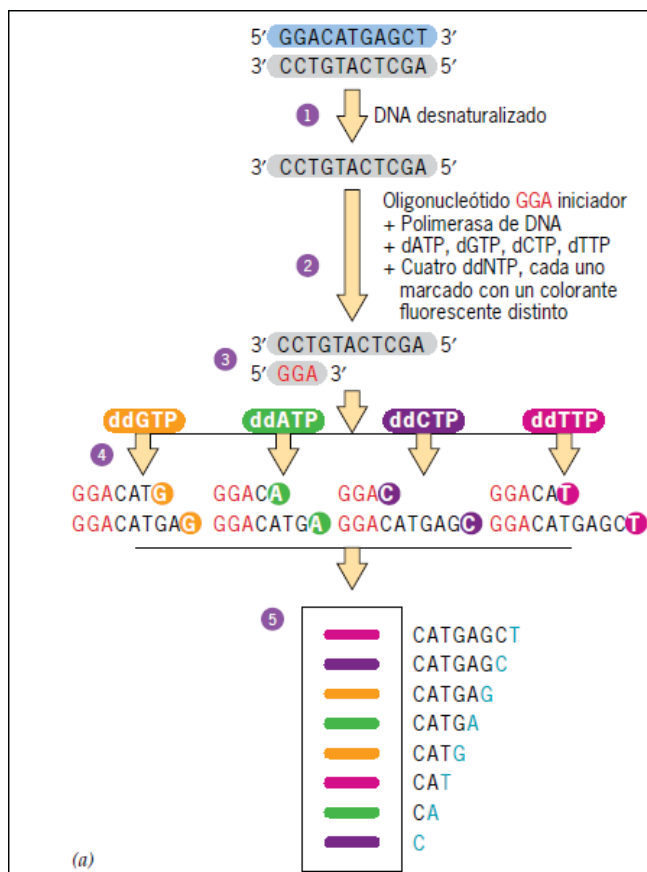


Figura 4: Pasos básicos en la identificación de la secuencia de un pequeño fragmento hipotético mediante la técnica de Sanger (didesoxi).

Fuente: Karp, 2009.

1.2.4.2 Secuenciación de próxima generación (NGS)

El término NGS (*Next generation Sequence*), hace referencia a las nuevas tecnologías diseñadas para analizar gran cantidad de bases de forma masiva en un menor lapso de tiempo y a menor costo, también conocida como secuenciación de segunda generación.

✚ Pirosecuenciación:

La pirosecuenciación fue la primera NGS en el mercado, ofrece lecturas alrededor de 700 pb, se basa en la detección quimioluminiscente del pirofosfato liberado durante la elongación de la cadena complementaria de ADN, lo que permite la rápida determinación de secuencias a tiempo real. A pesar de que es una tecnología capaz de ofrecer mayor cantidad de lecturas y datos de mejor calidad, el coste por secuenciación es relativamente caro y además, poco sensible en la detección de homopolímeros (Heather & Chain, 2016; Hyman, 1988; Nyrén & Lundin, 1985; Ronaghi, 1998).

✚ Secuenciación illumina:

Se caracteriza por el uso de nucleótidos marcados con fluoróforos que bloquean de forma reversible la elongación de la cadena. De este modo, tras la detección de la incorporación del fluoróforo, y la eliminación del mismo, es posible continuar con un nuevo ciclo de adición de un nuevo nucleótido. Genera lecturas de aproximadamente 700 bases (Bentley *et al.*, 2008).

✚ Secuenciación de Ion Torrent :

Determina la composición de la secuencia de ADN por cambios en la medición de pH, debido a la liberación de iones H^+ por incorporación de nucleótidos durante la síntesis de la hebra. El uso de circuitos integrados para medir los cambios de pH e identificar la incorporación de bases elimina la necesidad de sistemas de detección de luz, más caros, generando 250 bases en 2 horas reduciendo sustancialmente los costos y los tiempos de secuenciación (Rothberg *et al.*, 2011; Whiteley *et al.*, 2012).

1.2.4.3 Secuenciación de tercera generación (TGS)

También conocida como secuenciación de lecturas largas. Son aquellas capaces de secuenciar moléculas individuales, anulando la fase de amplificación de ADN compartido por todas las tecnologías anteriores y puede generar secuencias > 10Kb a partir de ADN nativo. Actualmente se

tiene dos tecnologías principales de lectura larga, *Pacific Biosciences* (PacBio) y *Oxford Nanopore Technology* (ONT).

✚ Secuenciación de Pacbio o SMRT

SMRT, por sus siglas en inglés *Single Molecule, RealTime* o molécula única, en tiempo real, genera longitudes promedio de 15kb. Es una secuenciación por síntesis, opera en tiempo real en lugar de incorporar una sola base (o tipo de base) a la vez. Esta secuenciación captura la información de la secuencia durante el proceso de replicación de la molécula de ADN objetivo. La plantilla, llamada SMRTbell, es un ADN circular monocatenario cerrado que se crea ligando adaptadores de horquilla a ambos extremos de una molécula de ADN bicatenario (dsDNA) objetivo (Rhoads & Au, 2015)

✚ Secuenciación Oxford Nanopore u ONT

ONT, por sus siglas en inglés, *Oxford Nanopore Technology*. Actualmente ONT desarrolló tecnologías como GridION y MinION, los cuales se caracterizan por que generan lecturas superiores a 1 Mb y su pequeño tamaño es comparado con un teléfono móvil. La secuenciación se basa en la conducción del ácido nucleico monocatenario a través de una bicapa lipídica con poros iónicos de α -hemolisina mediante electroforesis, la ligadura del adaptador al ADN de doble cadena facilita su captura por el poro de la proteína (Heather & Chain, 2016; T. Hu *et al.*, 2021).

1.2.5 BIOINFORMÁTICA

El término "bioinformática" se originó en 1970 con Paulien Hogeweg, biólogo teórico holandés (Hogeweg, 2011). Actualmente se define a la bioinformática como un campo interdisciplinario que se ocupa del desarrollo y la aplicación de algoritmos que analizan datos biológicos para investigar la estructura y función de los polímeros biológicos y sus relaciones con los sistemas vivos (Tapprich *et al.*, 2021).

1.2.5.1 Ensamblaje

La secuenciación masiva genera varias secuencias cortas de ADN no ordenadas (*reads*), este gran volumen de datos (*big data*) exige análisis computacionales para evaluar la calidad de la secuenciación, posteriormente la *big data* requiere un ensamble, que es el proceso de alinear y ensamblar todas las *reads* similares según su tamaño con el objetivo de reconstruir la secuencia de

ADN original del genoma. Para organizar dichas secuencias de ADN se tiene dos categorías de ensamblaje: ensamblaje por comparación, en el que se utiliza un genoma como referencia; y ensamblaje de novo, en el cual se utiliza solo la información obtenida de la secuenciación para reconstruir el genoma en cuestión, sin conocimiento a priori de la organización del mismo (Wajid & Serpedin, 2012). Actualmente, se han desarrollado diversos algoritmos computacionales con tecnologías más eficientes en cuanto a precisión y memoria, ya que el ensamble es un proceso vital para un posterior análisis biológico del genoma en el campo de la investigación y los sistemas de salud, de hecho ya se espera que los sistemas de salud, tengan su propia dirección IP que se pueda integrar con las secuencias del genoma del patógeno (Pallen, 2016), de igual manera, se menciona el beneficio de metagenomas personalizados para usos clínicos en el futuro y hacer realidad la medicina personalizada y precisa (Sohn & Nam, 2018).

1.2.5.2 Anotación

La anotación se basa en utilizar la información de la secuencia de un genoma para interpretarlo con datos biológicos relevantes. Para ello es necesaria la identificación de las principales características del genoma, proceso conocido como anotación. La anotación de genomas comprende dos etapas fundamentales: la anotación estructural (predicción de regiones codificantes) y la anotación funcional (asignación de información biológica a los genes previamente predichos) (Aguilar-Bultet & Falquet, 2015). La anotación genómica no termina con la identificación de los genes codificadores de proteínas. Además de los genes tradicionales, varias otras características de relevancia biológica caracterizan un genoma, como los ARN no codificantes (ARNnc), promotores, elementos reguladores o aspectos estructurales intrínsecos en la arquitectura de los genomas (es decir, repeticiones de secuencias). Todas estas características representan anotaciones que se pueden adjuntar a una secuencia (Del Giacco & Cattaneo, 2012). Por lo tanto, la anotación ofrece el sentido biológico al material genético.

1.2.6 BASES DE DATOS Y PROGRAMAS BIOINFORMATICOS

1.2.6.1 FastQC

FastQC es un software que facilita el control de calidad en archivos con secuencias sin procesar que provienen de tuberías de secuenciación de alto rendimiento, esta herramienta puede importar datos de archivos BAM, SAM o FastQ (cualquier variante). FastQC, realiza un protocolo de control

de calidad, analizando los resultados y agregando métricas de calidad en un panel interactivo diseñado para resumir de forma rica las ejecuciones de secuenciación individuales. El tablero agrupa muestras en menús desplegables para la navegación entre los conjuntos de datos, como tablas, gráfico de líneas, gráfico de barras, gráfico de rango de área, mapa de calor o mapa de calor de placas (Brown *et al.*, 2017) Además, exporta los resultados a un informe permanente en HTML y puede operar sin conexión para permitir la generación automatizada de informes (Andrews, 2009).

Un valor importante en FastQC es Phred. Las puntuaciones de calidad de Phred se definen como una propiedad que está relacionada logarítmicamente con las probabilidades de error de llamadas de base. Por ejemplo, si Phred asigna un puntaje de calidad de 30 a una base (Tabla 2), las posibilidades de que esta base se llame incorrectamente son 1 en 1000 (Ewing & Green, 1998).

Tabla 2: Índice phred: Probabilidad de que la asignación de una base sea incorrecta.

Nivel de calidad de Phred	Probabilidad de llamada base incorrecta	Precisión de llamada base
10	1 en 10	90%
20	1 en 100	99%
30	1 en 1000	99,9%
40	1 en 10,000	99,99%
50	1 en 100,000	99,999%
60	1 en 1,000,000	99,9999%

Fuente: (Ewing *et al.*, 1998)

1.2.6.2 Trimmomatic

Dado que la presencia de secuencias técnicas o de baja calidad, como adaptadores en los datos de secuenciación de próxima generación (NGS) pueden resultar subóptimos en análisis posteriores, Trimmomatic es una herramienta de preprocesamiento de recorte de secuencias más flexible, eficiente y optimizada para los datos Illumina NGS, incluye una variedad de pasos de procesamiento para el recorte y filtrado de lectura, sus parámetros asociados son suministrados en la línea de comandos (Bolger *et al.*, 2014). Los parámetros comunes utilizados, son:

- ILLUMINACLIP: Corta el adaptador y otras secuencias específicas de la lectura.
- SLIDINGWINDOW: Realiza un enfoque de recorte de ventana deslizante. Comienza escaneando en el extremo 5' y recorta la lectura una vez que la calidad promedio dentro de la ventana cae por debajo de un umbral.
- LEADING: Corta las bases al inicio de una lectura, si está por debajo del umbral de calidad.
- TRAILING: Corta las bases al final de una lectura, si está por debajo del umbral de calidad.
- CROP: Corta la lectura a una longitud especificada quitando las bases del extremo.
- HEADCROP: Corta el número especificado de bases desde el comienzo de la lectura.
- MINLEN: Descarta la lectura si está por debajo de una longitud especificada.
- TOPHRED33: Convierte puntajes de calidad a Phred-33

1.2.6.3 SPAdes

Es uno de los ensambladores de gráficos Eulerian de Bruijn, y fue diseñado para la secuenciación de una sola célula. Este programa utiliza el gráfico de Bruijn emparejado, que es una especie de gráfico de Bruijn de doble capa. Los k -mers de las lecturas de fragmentos de ADN construyen el gráfico interno de Bruijn, que se utiliza para el ensamblaje contig. Por otro lado, los k -mers emparejados con un tamaño de inserción grande construyen el gráfico externo de Bruijn, que se utiliza para la resolución repetida o el andamiaje (Bankevich *et al.*, 2012) (Medvedev *et al.*, 2011)

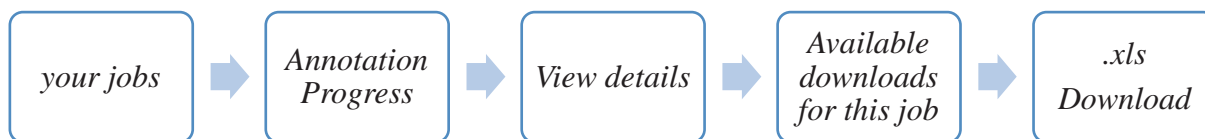
1.2.6.4 Quast

Es una herramienta de evaluación de la calidad para evaluar y comparar conjuntos de genomas. Esta herramienta mejora el *software* líder de comparación de ensamblajes con nuevas ideas y métricas de calidad. Quast puede evaluar ensamblajes tanto con un genoma de referencia como sin una referencia, además, produce muchos informes, tablas de resumen y gráficos, donde se enumera las métricas que evalúan la representación del genoma en contigs y el número de elementos funcionales ensamblados. Así se muestra el número *de contigs*, el número total de contigs en el ensamblaje, el *contig más grande*, la longitud total del genoma ensamblado, el porcentaje de Guanina mas Citocina, etc., (Gurevich *et al.*, 2013).

1.1.1.1. RAST (Rapid Annotation using Subsystem Technology)

El servicio RAST es completamente automatizado para anotar genomas bacterianos y de arqueas. El servicio identifica la codificación de proteínas, los genes de ARNr y ARNt, asigna funciones a los genes, predice qué subsistemas están representados en el genoma, utiliza esta información para reconstruir la red metabólica y hace que la salida sea fácilmente descargable para el usuario. Además, el genoma anotado se puede explorar en un entorno que admite el análisis comparativo con los genomas anotados mantenidos en el entorno SEED (Aziz *et al.*, 2008).

Para utilizar el servicio RAST se debe registrar un usuario y subir el archivo con el genoma procariótico de extensión fasta. La anotación tarda dependiendo de cuántos trabajos tiene en cola, por lo general son 24 - 48hrs. Después de la anotación, el usuario puede seguir su trabajo desde la página inicial, siguiendo este esquema:



1.2.6.5 MEGA 7 (Molecular Evolutionary Genetic Analysis)

El Análisis Genético Evolutivo Molecular (MEGA) es un software bio-computacional para llenar el vacío entre el desarrollo y el análisis de datos, el cual está disponible gratuitamente en www.megasoftware.net. La finalidad de este *software* es interpretar esta información biológica de un extenso repertorio de información de secuencias sin procesar (Sohpal *et al.*, 2010).

La nueva versión Mega 7 se ha optimizado para su uso en sistemas informáticos de 64 bits para analizar conjuntos de datos más grandes, el cual proporciona un asistente avanzado para construir árboles del tiempo e incluye una nueva funcionalidad para predecir automáticamente eventos de duplicación de genes en árboles genealógicos (S. Kumar *et al.*, 2016).

1.2.6.6 GENE RUNNER

El *software* Gene Runner versión 6.5.51 Beta, es un programa de biología molecular que dentro de su interfaz tiene la opción para el diseño de cebadores. Se sube las secuencias de ADN en archivo .txt o .fasta al programa *Gene Runner*, y conociendo la localización del fragmento a

amplificar se selecciona una secuencia de nucleótidos y se da a la opción *Oligo Analysis*, seguidamente, se eliminan algunos nucleótidos de ambos extremos hasta que el oligonucleótido muestre características de un buen cebador como el contenido de GC de 35–60% y una pequeña diferencia de 0 – 3 °C en la T_m de 55 – 65 °C entre las imprimaciones, evitando la formación de dímeros de imprimación, horquillas y bucles. Al usar el botón de cambio de oligo, es posible completar el recuadro de oligo para el cebador antisentido y realizar los mismos pasos que para el cebador sentido o directo. Finalmente, las secuencias y características del cebador sentido y antisentido deben guardarse (A. Kumar & Chordia, 2015).

1.2.6.7 I-Tasser

El servidor I-TASSER es un recurso en línea para la predicción automatizada de la estructura de proteínas y la anotación de funciones basadas en la estructura, las plantillas estructurales se reconocen primero desde el PDB (*Protein Data Bank*) utilizando múltiples enfoques de alineación de subprocesos. A continuación, se construyen modelos de estructura de longitud completa mediante simulaciones interactivas de ensamblaje de fragmentos. Los conocimientos funcionales se obtienen finalmente al hacer coincidir los modelos de estructura predichos con proteínas conocidas en las bases de datos de funciones. Los algoritmos involucrados se han evaluado rigurosamente en experimentos ciegos en toda la comunidad y han demostrado ventajas considerables en comparación con los métodos de pares en la predicción de la estructura y función de las proteínas (Yang & Zhang, 2015).

1.2.6.8 PDB (Protein Data Bank)

El Banco de datos de proteínas (PDB), un servidor de acceso libre, contiene en su base de datos estructuras 3D de macromoléculas biológicas y sus complejos determinadas experimentalmente, se estableció en 1971 y se convirtió en el primer recurso digital de acceso abierto en las ciencias biológicas. El archivo de la PDB alberga actualmente ~130.000 entradas (mayo de 2017). Es administrado por la organización *Worldwide Protein Data Bank* (wwPDB; wwpdb.org), que incluye el *RCSB Protein Data Bank* (RCSB PDB; rcsb.org), el *Protein Data Bank Japan* (PDBj; pd bj.org), el *Protein Data Bank in Europe* (PDBe; pd be.org) y *BioMagResBank* (BMRB; www.bmr b.wisc.edu). En la actualidad, el PDB archiva datos experimentales, metadatos asociados y modelos estructurales de nivel atómico 3D derivados de tres métodos bien establecidos:

cristalografía, espectroscopía de resonancia magnética nuclear (RMN) y microscopía electrónica (3DEM) (Burley *et al.*, 2017).

1.2.6.9 Blast (Basic Local Alignment Search Tool)

La herramienta básica de búsqueda de alineación local (BLAST) es un programa de búsqueda de similitud de secuencias que se puede utilizar a través de una interfaz web o como una herramienta independiente para comparar la consulta de un usuario con una base de datos de secuencias. Varias variantes de BLAST comparan todas las combinaciones de consultas de nucleótidos o proteínas con bases de datos de nucleótidos o proteínas. BLAST es una heurística que encuentra coincidencias cortas entre dos secuencias e intenta iniciar alineaciones desde estos "puntos calientes". Además de realizar alineaciones, BLAST proporciona información estadística sobre una alineación; este es el valor "esperado" o la tasa de falsos positivos (Johnson *et al.*, 2008).

1.2.7 PROTEÍNAS

Las proteínas son los principales polímeros estructurales y funcionales en los seres vivos. Cumplen un amplio abanico de funciones, incluida la catálisis de reacciones metabólicas y el transporte de vitaminas, minerales, oxígeno y combustibles. Esta información proviene del ADN, pasa al ARN mensajero y se utiliza para elaborar una proteína específica. Las proteínas son sintetizadas en los ribosomas como una secuencia de aminoácidos unidos en una estructura lineal, pero adoptan estructuras tridimensionales complejas al realizar sus funciones (Baynes & Dominiczak, 2019).

1.2.7.1 Composición de las proteínas

Las proteínas son polímeros formados por cadenas lineales de aminoácidos unidos covalentemente. Existen alrededor de 300 aminoácidos en varios sistemas animales, vegetales y microbianos, pero solamente 20 aminoácidos están codificados por el ADN para aparecer en las proteínas (Baynes & Dominiczak, 2019), como se observa en la tabla 3. Muchas proteínas también contienen aminoácidos modificados y componentes accesorios, denominados grupos prostéticos, por lo que la variedad real de proteínas aisladas de células o tejidos es algo mayor (Watson *et al.*, 2013).

Cada aminoácido tiene un carbono central, denominado carbono α , al cual se enlazan cuatro grupos diferentes (Figura 5):

- Un grupo amino básico (—NH_2).
- Un grupo carboxilo ácido (—COOH).
- Un átomo de hidrógeno (—H).
- Una cadena lateral distintiva (—R).

Las unidades de aminoácidos de una cadena de péptidos se denominan residuos de aminoácidos. El residuo aminoácido que tiene un grupo amino libre al final del péptido se denomina el aminoácido N-terminal (amino terminal), mientras que el residuo que tiene un grupo carboxilo libre en el otro extremo se denomina aminoácido C-terminal (carboxilo terminal).

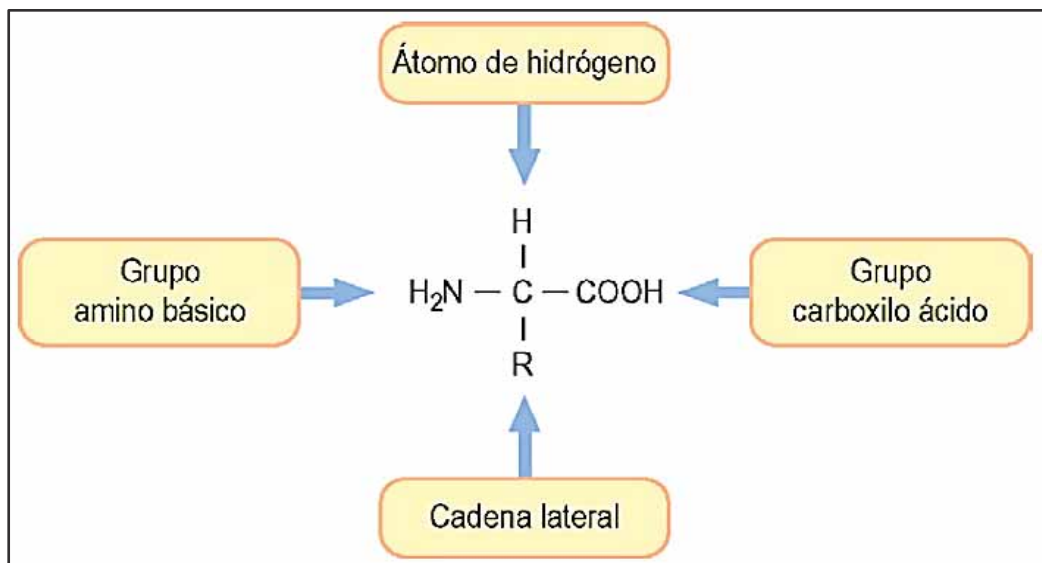


Figura 5: Estructura básica de un aminoácido.

Fuente: Baynes & Dominiczak, 2019.

Tabla 3: Clasificación y abreviatura de los 20 α -aminoácidos.

	Abreviatura de tres letras	Abreviatura de una letra
Aminoácidos alifáticos		
Glicina	Gly	G
Alanina	Ala	A
Valina	Val	V
Leucina	Leu	L
Isoleucina	Leu	I
Aminoácidos que contienen azufre		
Cisteína	Cys	C
Metionina	Met	M
Aminoácidos aromáticos		
Fenilalanina	Phe	F
Tirosina	Tyr	Y
Triptófano	Trp	W
Iminoácido		
Prolina	Pro	P
Aminoácidos neutros		
Serina	Ser	S
Treonina	Thr	T
Asparagina	Asn	N
Glutamina	Gln	Q
Aminoácidos básicos		
Histidina	His	H
Lisina	Lys	K
Arginina	Arg	R
Aminoácidos ácidos		
Ácido aspártico	Asp	D
Ácido glutámico	Glu	E

Fuente: Ellington & Cherry, 1996.

1.2.7.2 Niveles de organización de las proteínas

- Estructura primaria: Es la secuencia lineal de aminoácidos, es decir, la unión del grupo carboxilo de un aminoácido con el grupo amino del aminoácido siguiente, formando un puente amida (péptido), en la reacción se elimina agua (Baynes & Dominiczak, 2019).
- Estructura secundaria: Esta estructura está determinada por las interacciones mediante enlaces o puentes de hidrógeno entre el oxígeno del grupo carbonilo de una cadena peptídica y el hidrógeno de amida de otra cadena peptídica próxima. Existen dos tipos de estructura secundaria: la hélice α y la hoja plegada β (Baynes & Dominiczak, 2019).

La hélice α es una estructura en forma de varilla con la cadena peptídica fuertemente enrollada y con las cadenas laterales de los residuos aminoácidos extendiéndose fuera del eje de la espiral. Cada grupo carbonilo amídico está unido mediante un puente de hidrógeno al hidrógeno del grupo amida de una cadena peptídica que está alejado cuatro residuos a lo largo de la misma cadena. Hay un promedio de 3,6 residuos de aminoácidos por vuelta de hélice y la hélice gira hacia la derecha (en sentido de las agujas del reloj) en la mayoría de las proteínas naturales (Baynes & Dominiczak, 2019).

La hoja β plegada es una estructura extendida, a diferencia de la hélice α , que está enrollada. Está plegada porque los enlaces carbono carbono (C—C) son tetraédricos y no pueden existir en una configuración plana. Si la cadena polipeptídica discurre en la misma dirección, forma una hoja β paralela, pero si sigue la dirección opuesta, forma una estructura antiparalela. El giro β se refiere al segmento en el que el polipéptido gira y cambia abruptamente de dirección (Baynes & Dominiczak, 2019).

- Estructura terciaria: La conformación tridimensional, plegada y biológicamente activa de una proteína es propia de su estructura terciaria, esta estructura refleja la forma global de la molécula siendo estabilizada por interacciones entre grupos funcionales de las cadenas laterales: puentes disulfuro covalentes, enlaces de hidrógeno, puentes salinos e interacciones hidrofóbicas. La estructura terciaria de las proteínas se ha determinado mediante cristalografía de rayos X y espectroscopia de resonancia magnética (Baynes & Dominiczak, 2019).
- Estructura cuaternaria: Corresponde a un complejo o un ensamblaje de dos o más cadenas de péptidos que se mantienen unidas por interacciones no covalentes o, en algunos casos,

covalentes. En general, la mayoría de proteínas mayores de 50 kDa constan de más de una cadena y se conocen como proteínas diméricas, triméricas o multiméricas. Muchas proteínas que contienen varias subunidades están compuestas de diferentes tipos de subunidades funcionales, como las subunidades reguladoras y las catalíticas (Baynes & Dominiczak, 2019).

1.2.8 ENZIMAS

En 1878, Wilhelm Kühne acuñó el término enzima que proviene del griego: *en* = en, *zima* = levadura (Devlin, 1999). Los enzimas son proteínas especializadas que tienen la función de acelerar reacciones químicas (actividad catalítica) debida a su capacidad de activación específica (Battaner Arias, 2013).

1.2.8.1 Clasificación y nomenclatura de las enzimas

La Unión Internacional de Bioquímica y Biología Molecular - IUBMB (*International Union of Biochemistry and Molecular Biology*), constituyó una comisión (*Enzyme Commission*, E.C.) que se encargó de establecer un sistema que contiene a todas las enzimas descritas y de desarrollar una nomenclatura para identificarlos, los denominados números EC; este sistema se basa en el mecanismo de reacción química que cataliza cada enzima, clasificando en grupos cuando catalizan procesos semejantes, y en subgrupos cuando especifican con mayor exactitud la reacción particular considerada (Ludeña, 2009). A continuación se nombran los principios generales del sistema IUBMB:

- Las reacciones y las enzimas que las catalizan se dividen en 6 grupos principales, cada una con 4 a 13 subgrupos.
- El nombre de la enzima tiene dos partes: la primera es el nombre del o los sustratos; la segunda, con terminación –asa, indica el tipo de reacción catalizada.
- Información adicional, si es necesario aclarar la reacción, puede seguir un paréntesis.
- Cada enzima tiene un número clave (E.C.) que caracteriza al tipo de reacción según la clase (primer dígito), subclase (segundo dígito), sub-subclase (tercer dígito). El cuarto dígito es para la enzima específica.

La clasificación de las enzimas se hace distribuyéndolos en seis grupos conforme a la naturaleza de la reacción catalizada, donde cada una se subdivide en subgrupos. El número de grupo (1-6) es el que aparece como primer elemento en el número sistemático de la enzima. Estos grupos son los siguientes:

1. Oxidorreductasas, 2. Transferasas, 3. Hidrolasas, 4. Liasas, 5. Isomerasas y 6. Ligasas.

1.2.8.2 Obtención y uso de enzimas

Las enzimas se obtienen de diversas fuentes: animales, plantas, hongos y bacterias, sobresaliendo los microorganismos extremófilos debido a sus propiedades extraordinarias (Sandhya *et al.*, 2005 & Dumorné *et al.*, 2017), y aunque muchos procesos industriales, como la fabricación de queso, han utilizado tradicionalmente fuentes de enzimas impuras, a menudo de animales o plantas, el desarrollo de gran parte de la enzimología industrial moderna ha ido de la mano con la explotación comercial de enzimas microbianas debido a su sencilla manipulación genética, destacando a las hidrolasas en el mercado global de enzimas. Actualmente, las enzimas se utilizan en cuatro campos distintos de comercio y tecnología: como catalizadores industriales, agentes terapéuticos, reactivos analíticos y como herramientas de manipulación genética (Robinson, 2015). En la tabla 4 se detalla los usos de algunas enzimas:

Tabla 4: Enzimas industriales según su uso

Enzima	Reacción	Fuente	Uso
Catalizadores industriales			
Proteasas ácidas	Digestión de proteínas	<i>A. niger</i> , <i>K. lactis</i>	Coagulación de la leche en la fabricación de queso
Proteasas alcalinas	Digestión de proteínas	Especies de <i>Bacilos</i>	Detergentes
Aminoacilasa	Hidrólisis L- aminoácidos acilados	Especies de <i>Aspergillus</i>	Producción de L- aminoácidos
α -amilasa	Hidrólisis de almidón	Especies de <i>Bacilos</i>	Producción de glucosa en la industria alimentaria.
Amiloglucosidasa	Hidrólisis de dextrina	Especies de <i>Aspergillus</i>	Producción de glucosa
β -galactosidasa	Hidrólisis de lactosa	Especies de <i>Aspergillus</i>	Hidrólisis de lactosa en leche o suero
Isomerasa glucosa	Conversión de glucosa a fructosa	Especies <i>Streptomyces</i>	Producción de jarabe de alta fructosa
Penicilina acilasa	Escisión cadena lateral penicilina	<i>E. coli</i>	Producción de penicilinas semisintéticas
Agentes terapéuticos			
L- asparaginasa	Eliminación de L- asparagina	<i>E. coli</i>	Quimioterapia contra el cáncer, la leucemia.
Uroquinasa	Activación de plasminógeno	Humano	Eliminación de coágulos de fibrina.
Reactivos analíticos			
Glucosa oxidasa	Oxidación de glucosa	<i>Aspergillus niger</i>	Detección de glucosa en sangre.
Luciferasa	Bioluminiscencia	Bacterias marinas o luciérnagas.	Ensayos bioluminiscentes con ATP
Peroxidasa	Colorante de oxidación de H ₂ O ₂	Rábano picante	Cuantificación de hormonas y anticuerpos.
Ureasa	Hidrólisis de urea a CO ₂ y NH ₃	Jack Bean	Cuantificación de urea en fluidos corporales
Herramientas manipuladoras			
Lisozima	Hidrólisis de enlaces glucosídicos	Clara de huevo	Interrupción del mucopeptido en bacterias.
Nucleasas	Hidrólisis de enlaces fosfodiéster	Varias bacterias	Enzimas de restricción
ADN polimerasas	Síntesis de ADN	<i>Thermus aquaticus</i>	En la reacción en cadena de la polimerasa

Fuente: Robinson, 2015.

1.2.9 PROTEASAS

Las proteasas o peptidasas son enzimas que catalizan la hidrólisis específica de uno o varios enlaces peptídicos dentro de una proteína eliminando una molécula de agua. Las proteasas son universales en todos los organismos y se requieren para las actividades metabólicas de la célula (Baynes & Dominiczak, 2019)..

1.2.9.1 Sistema EC de clasificación de proteasas

De acuerdo con la UICBBM, las proteasas pertenecen al sub-grupo 3.4 de las hidrolasas y se clasifican sobre la base de tres criterios principales: Tipo de reacción catalizada, naturaleza química del sitio catalítico y la relación evolutiva con referencia a la estructura.

Dependiendo de su sitio de acción las proteasas se subdividen en dos grupos principales (Rao *et al.*, 1998):

- ✚ **Las exopeptidasas**, hidrolizan un enlace peptídico próximo al extremo amino (N-terminal) o carboxi terminal (C-terminal) del sustrato.
- ✚ **Las endopeptidasas**, hidrolizan las cadenas peptídicas en las regiones internas alejadas de los extremos carboxi y amino terminal.

Según el grupo funcional presente en el sitio activo, las proteasas se clasifican en cuatro grupos prominentes (Hartley, 1960; Rao *et al.*, 1998):

- ✚ **Serín proteasas:** Presentan una serina en el centro activo, acompañada de una histidina y un residuo dicarboxílico (aspartato o glutamato), constituyendo lo que se llama una tríada catalítica, realiza un ataque nucleofílico en el carbonilo del enlace peptídico.
- ✚ **Aspártil proteasas:** Poseen 2 residuos de Aspartato conservados en su sitio activo. La molécula de agua se desprotona al ser atacado por un residuo aspártico del sitio activo.
- ✚ **Tiol o Cisteín proteasas:** Poseen un residuo de Cisteína (Cis) en su sitio activo, siendo el grupo tiol de la Cisteína quien realiza el ataque nucleofílico.
- ✚ **Metalo proteasas:** Poseen un ión catiónico de origen metálico en su sitio activo, generalmente Zn^{2-} , que activa la molécula del agua para transformarla en nucleofílico.

Hay unas pocas proteasas misceláneas que no encajan de forma precisa en la clasificación estándar, por ejemplo, proteasas dependientes de ATP que requieren ATP para la actividad. (Rao *et al.*, 1998).

Tabla 5: Clasificación de las proteasas.

	Sub-subclases	Descripción
Exopeptidasas	EC 3.4.11	Aminopeptidasas
	EC 3.4.13	Dipeptidasas
	EC 3.4.14	Di- y tri-peptidil-peptidasas
	EC 3.4.15	Peptidil-dipeptidasas
	EC 3.4.16	Carboxipeptidasas Tipo Serina
	EC 3.4.17	Metalocarboxipeptidasas
	EC 3.4.18	Carboxipeptidasas Tipo Cisteína
	EC 3.4.19	Omega peptidasas
Endopeptidasas	EC 3.4.21	Endopeptidasas Serínicas
	EC 3.4.22	Endopeptidasas Cisteínicas
	EC 3.4.23	Endopeptidasas Aspárticas
	EC 3.4.24	Metaloendopeptidasas
	EC 3.4.25	Endopeptidasas Treonínicas
	EC 3.4.99	Endopeptidasas de mecanismo desconocido

Fuente: Rao *et al.*, 1998.

1.2.9.2 El sistema MEROPS de clasificación de proteasas

De acuerdo a la base datos MEROPS (<http://merops.sanger.ac.uk>), las proteasas son asignadas dentro de una Familia en base a similitudes significativas en la secuencia de aminoácidos, y dichas familias que pueden ser homólogas se agrupan en clanes (Rawlings *et al.*, 2014). A continuación se definen los tres niveles de clasificación:

- Clan. Es un conjunto de familias cuyas peptidasas tienen un origen evolutivo común que, si bien no llega a evidenciarse por comparación de sus secuencias de aminoácidos, sí por las similitudes de sus estructuras terciarias. Se construye por comparación de las estructuras terciarias y cuando no están disponibles se comparan los ordenamientos de aminoácidos catalíticos en la cadena polipeptídica o las secuencias aminoacídicas alrededor de los aminoácidos catalíticos.

- Familia. Se construye por comparación de la estructura primaria. Está constituida por peptidasas cuyas secuencia de aminoácidos, que debe incluir la secuencia de la “unidad peptidasa” (región de la proteína responsable de la actividad catalítica) presentan similitudes estadísticamente significativas.
- Peptidasa. Cada proteína con actividad peptidasa recibe un nombre y un identificador (ID MEROPS). Se diferencian por su estructura, función biológica, sustrato, mecanismo catalítico y origen genético.

Tabla 6: Distribución de familias de peptídicas de organismos de diferentes reinos en el sistema propuesto por MEROPS (<http://merops.sanger.ac.uk>, consultado 12/11/2021).

Tipo de peptidasas	REINO						
	Bacteria	Archaea	Protozoa	Fungi	Planta	Animalia	Virus
Aspárticas	9	5	6	6	9	7	4
Cisteínicas	27	11	20	16	20	22	30
Glutámicas	2	1	-	1	-	-	1
Metalopeptidasas	62	34	32	34	34	33	14
Serínicas	25	18	20	16	22	25	18
Treonínicas	1	1	1	1	1	1	-
Asparagínicas	3	-	-	-	-	-	6

1.2.9.3 Fuentes de proteasas

Las proteasas son fisiológicamente necesarias para los organismos vivos y se encuentran en una gran diversidad de fuentes, como plantas, animales y microorganismos.

Las enzimas proteolíticas derivadas de plantas están encontrando cada vez más aplicaciones clínicas e industriales (Banik *et al.*, 2018). La papaína (Amri & Mamboya, 2012), la bromelina (Arshad *et al.*, 2014), las queratinasas y la ficina son algunas de las ilustres proteasas de origen vegetal, siendo las más predominantes las endopeptidasas de cisteína y rara vez se observan otras proteasas (Badgajar & Mahajan, 2009), las plantas carnívoras segregan proteasas de gran importancia biotecnológica y farmacéutica, *Drosera capensis* es una de esas plantas que tiene una gran cantidad de genes que codifican las cisteína proteasas, y que por comparación de secuencias mostraron homología con varias proteasas de plantas (Butts *et al.*, 2016). Sin embargo, la

producción de proteasas a partir de plantas, es un proceso lento por varios factores, como la disponibilidad de tierra para el cultivo y la idoneidad de las condiciones climáticas para el crecimiento (Rao et al., 1998; Gurumalles *et al.*, 2019).

Por otra parte, las proteasas de origen animal se preparan en forma pura en grandes cantidades pero su producción depende de la disponibilidad de ganado para el sacrificio que se rige por varias políticas clasificadas en sectores políticos y agrícolas; las proteasas a partir de animales más familiares son la tripsina pancreática, la quimotripsina, la pepsina y las reninas (Rao et al., 1998).

En cambio, las proteasas de origen microbiano, son preferidas a las enzimas de origen vegetal y animal por su gran diversidad, rango específico de acción y propiedad de ser activos en un rango muy amplio de temperatura, pH, mayor producción a través de la fermentación, disponibilidad, diversidad, estabilidad y menor costo, atrayendo la atención de los biotecnólogos de todo el mundo (de Souza *et al.*, 2015) (Banerjee & Ray, 2017). Las proteasas microbianas son originadas por las bacterias y los hongos que son buenas fuentes de diferentes tipos de enzimas como proteasa, amilasa, lipasa, celulasa, pectinasa, quitinasa, fitasa, β -galactosidasa, etc. (Banerjee & Ray, 2017). Los microorganismos productores de proteasas se aíslan de diversas fuentes, como el suelo, los desechos vegetales, las plantas de procesamiento de leche y suelos contaminados con desechos de aves de corral, etc. (Chu, 2007; Gaur et al., 2010), es importante conocer las características del ambiente de origen ya que para la producción cada microorganismo necesita un medio de crecimiento optimizado y condiciones ambientales para secretar la enzima proteasa.

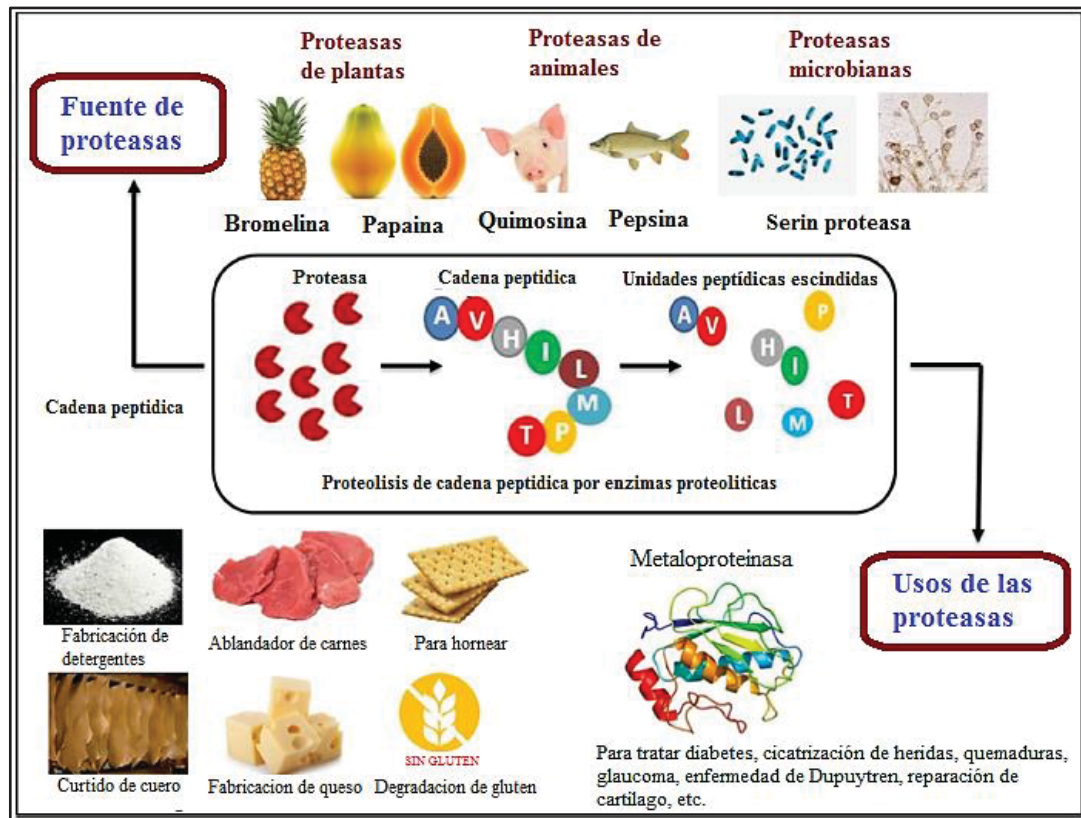


Figura 6: Fuente de proteasas y ejemplos de sus diversos usos

Fuente: Gurumallesh et al., 2019.

1.2.10 HALOENZIMAS PROTEOLITICAS Y SU APLICACIÓN INDUSTRIAL

Las haloenzimas son las enzimas proteolíticas provenientes de bacterias halófilas. El primer estrés químico encontrado durante la evolución de la vida en la tierra puede haber sido el estrés salino, así, desde el principio, los organismos deben haber desarrollado estrategias y mecanismos efectivos para la estabilización de las estructuras protoplasmáticas y la regulación de iones (Sarwar et al., 2015); tales estrategias y mecanismos les confiere propiedades importantes en aplicaciones industriales, como el procesamiento de alimentos, la biorremediación ambiental y los procesos biosintéticos. En este sentido, el hallazgo de nuevas enzimas que muestran actividades óptimas en varios rangos de concentraciones de sal, temperaturas y valores de pH es de gran importancia (Gomes & Steiner, 2004).

Los microorganismos halófilos tiene la capacidad de tolerar el estrés causado por la salinidad mediante diversas adaptaciones. Las halobacterias y algunas bacterias extremadamente halófilas acumulan iones inorgánicos (K^+ , Na^+ , Cl^-) en el citoplasma para equilibrar la presión osmótica del

medio, y han desarrollado proteínas específicas que son estables y activas en presencia de sales. En contraste, los halófilos moderados acumulan en el citoplasma grandes cantidades de osmolitos orgánicos específicos, que funcionan como osmoprotectores, proporcionando un equilibrio osmótico sin interferir con el metabolismo normal de la célula (Cira-Chávez *et al.*, 2018; de Lourdes Moreno *et al.*, 2013).

Por otro lado, la explotación biotecnológica de las enzimas de microorganismos halófilos no se limita a su estabilidad y actividad a altas concentraciones de sal (Madigan & Oren, 1999), ya que también son tolerantes a pH alcalino, altas temperaturas, medios con baja actividad de agua, debido a que tienen suficiente agua para mantener una distribución de carga adecuada en el sitio activo manteniendo la conformación de la enzima (Dumorné *et al.*, 2017; Zaccai, 2004), y en presencia de solventes orgánicos (Oren, 2010) que generalmente se consideran que causan desnaturalización enzimática, produciendo la distorsión de las interacciones agua-biocatalizador, sin embargo, se han informado varios estudios sobre hidrolasas producidas por extremófilos aislados de ambientes salinos que son muy estables en soluciones que contienen solventes orgánicos (de Lourdes Moreno *et al.*, 2013).

Los microorganismos halofílicos tienen varias aplicaciones biotecnológicas, se están utilizando varias enzimas de hidrólisis de microorganismos halófilos en la degradación de la biomasa para la consiguiente producción de bioetanol o biodiesel. Además, existe la posibilidad de encontrar halófilos con actividad de celulasa y actividad de eliminación de fenol para disminuir la concentración de la demanda química de oxígeno y la contaminación ambiental. Además, los microorganismos halófilos con sus habilidades especiales tienen un papel innegable en la producción de biohidrógeno y biogás (Amoozegar *et al.*, 2019).

1.2.11 TÉCNICAS DE BIOLOGIA MOLECULAR

1.2.11.1 Extracción de ADN

La extracción y purificación de ácidos nucleicos constituye la primera etapa de la mayoría de los estudios de biología molecular y de todas las técnicas de recombinación de ADN; siendo la calidad y la pureza los dos elementos más importantes en ese tipo de análisis.

La extracción consiste en el aislamiento y purificación de moléculas de ADN y se basa en las características fisicoquímicas de la molécula. El ADN está constituido por dos cadenas de nucleótidos unidas entre sí formando una doble hélice, la unión de los nucleótidos ocurre entre el grupo fosfato y el azúcar, mediante enlaces fosfodiéster, dando lugar al esqueleto de la molécula. Las bases de cadenas opuestas se unen mediante puentes de hidrógeno y mantienen estable la estructura helicoidal (Figura 10). Los grupos fosfato están cargados negativamente y son polares, lo que le confiere al ADN una carga negativa neta y lo hace altamente polar, características que son aprovechadas para su extracción. Los grupos fosfato tienen una fuerte tendencia a repelerse, debido a su carga negativa, lo que permite disolver al ADN en soluciones acuosas y formar una capa hidratante alrededor de la molécula pero, en presencia de etanol, se rompe la capa hidratante y quedan expuestos los grupos fosfato, la carga neta negativa del ADN le permite unirse a moléculas y matrices inorgánicas cargadas positivamente (Cornejo Romero *et al.*, 2014; Sambrook *et al.*, 1989).

Actualmente existen diferentes protocolos para la extracción de ADN que se han ido desarrollando a lo largo del tiempo con la finalidad de obtener una cantidad y calidad de ADN adecuados. Esta la técnica tradicional, que requiere preparar varias soluciones y la extracción puede tomar varias horas por los numerosos pasos que deben realizarse pero es más económica, en esta técnica se utilizan solventes orgánicos para separar a las proteínas del ADN y, una vez suspendido en la fase acuosa, aislarlo por precipitación con etanol (Sambrook *et al.*, 1989). Por otro lado están los *kits* de extracción que utilizan matrices inorgánicas compactas cargadas positivamente capaces de retener varios microgramos de ADN y separarlos del resto de las biomoléculas, permitiendo obtener un extracto libre de inhibidores, de esta manera disminuyendo el tiempo de extracción. Las matrices inorgánicas pueden ser membranas de sílice que están formadas por una resina insertadas dentro de un tubo de polipropileno o perlas magnéticas con un centro de hierro recubierto por resina (microesferas), que están suspendidas en una solución amortiguadora (Dundas *et al.*, 2008).

1.2.11.2 *Primers, cebadores u oligos*

Los cebadores u oligos son cortas secuencias de ADN que se utilizan para la amplificación de determinados genes, estos oligos deben estar enfrentados y no a mucha distancia ya que delimitan la zona de ADN a amplificar, es decir, corresponden a los nucleótidos que definen los extremos

de la secuencia que se desea replicar. El diseño de primers es uno de los aspectos más importantes de la PCR. Los primers mal diseñados pueden amplificar otros fragmentos de ADN distintos a los buscados (amplificación inespecífica) (A. Kumar & Chordia, 2015).

1.2.11.3 PCR - Amplificación del ADN mediante la técnica reacción en cadena de la polimerasa

Se define como la amplificación *in vitro* de uno o varios fragmentos de ADN específicos utilizando una ADN polimerasa que puede trabajar a temperaturas muy elevadas, ya que proviene de la bacteria *Thermus aquaticus* que vive a altas temperaturas, de ahí su nombre comercial Taq polimerasa (Eguiarte *et al.*, 2007).

La reacción en cadena de la polimerasa, PCR por sus siglas en inglés de *Polymerase Chain Reaction*, fue desarrollada por Kary Mullis a principio de los años 80 (Mullis *et al.*, 1986). Mullis basó su técnica en el proceso de replicación del ADN de doble cadena que realiza la ADN polimerasa. La función que realiza esta enzima es sintetizar la cadena complementaria de ADN en sentido 5' → 3' utilizando como molde la cadena sencilla pero a partir de una región que tiene doble cadena. De esta manera se crea la región de doble cadena usando los cebadores u oligos, estos son sintetizados de una forma en que sean complementarios a cada uno de los extremos 3' del fragmento de ADN que se quiere amplificar (Mullis *et al.*, 1986).

Si el proceso de duplicación del ADN bicatenario se repite de manera cíclica, es decir, utilizando como molde las moléculas sintetizadas en la replicación anterior, se producirá un aumento en la cantidad de ADN. De acuerdo a este principio la técnica de Reacción en Cadena de la Polimerasa se basa en la repetición de un ciclo formado por tres etapas.

- **Desnaturalización:** Inicia la separación de las dos cadenas de ADN, siendo el calentamiento de 94-95 °C, la forma más habitual.
- **Hibridación de los cebadores:** Es cuando el cebador se unirá a su secuencia complementaria en el ADN molde. Para ello es necesario bajar la temperatura a 40-68 °C durante 20-40 segundos, permitiendo así el alineamiento. La polimerasa une el híbrido de la cadena molde y el cebador, y empieza a sintetizar ADN. Los cebadores actuarán como límites de la región de la molécula que va a ser amplificada.

- Extensión o elongación de la cadena: La polimerasa sintetiza una nueva hebra de ADN complementaria a la hebra molde añadiendo los dNTP complementarios en dirección 5'→ 3', uniendo el grupo 5'-fosfato de los dNTP con el grupo 3'-hidroxilo del final de la hebra de ADN creciente (la cual se extiende). La temperatura de la Taq polimerasa está en 75-80 °C.

Además, existe una etapa final de la extensión, que se lleva a cabo a una temperatura de 70-74 °C durante 5-15 minutos tras el último ciclo de PCR. Con ella se asegura que cualquier ADN de cadena simple restante sea totalmente ampliado.

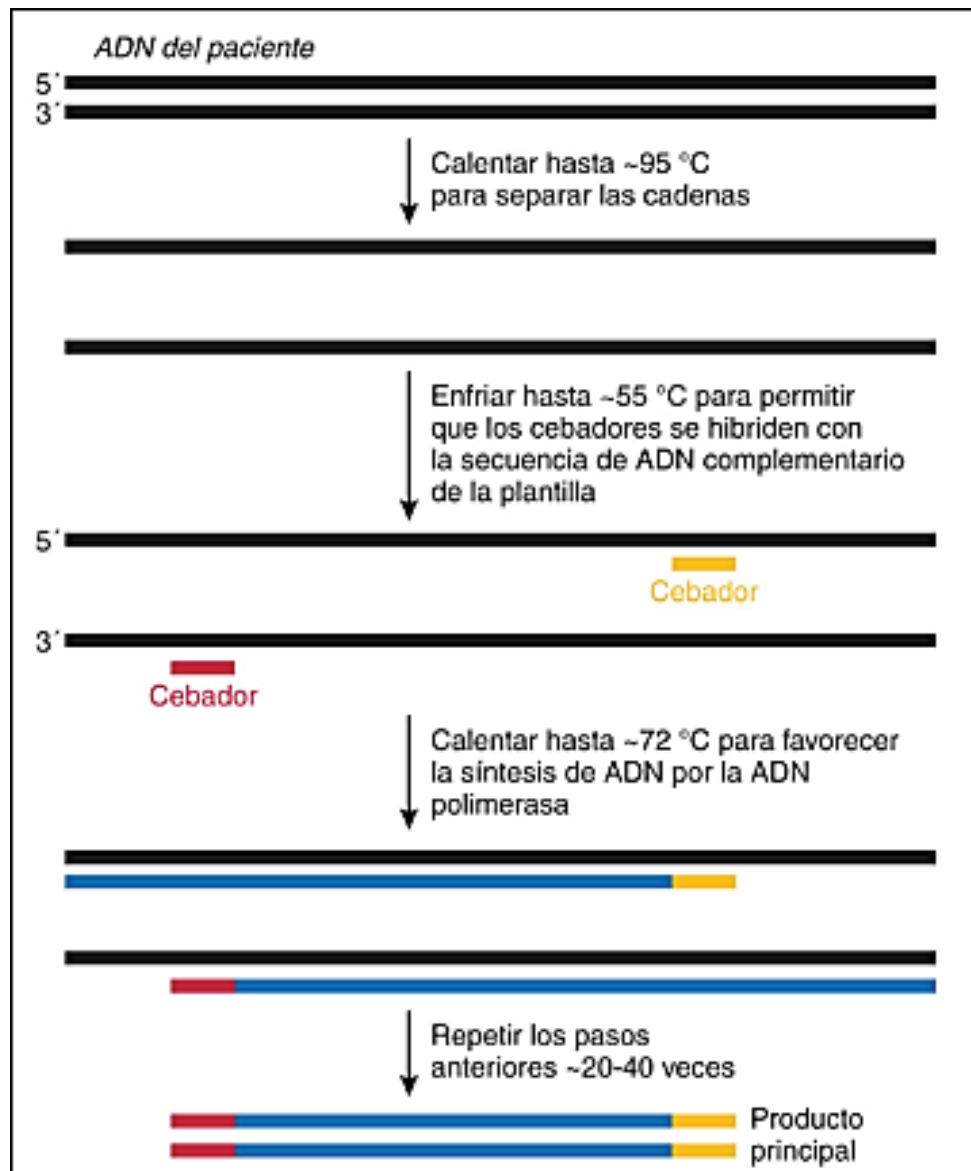


Figura 7: Principio de la amplificación de una secuencia de ADN mediante PCR.

Fuente: Ronner, 2019.

1.2.11.4 *Cuantificación del ADN*

Consiste en cuantificar el índice de pureza del material genético purificado utilizando un nanoespectrofotometro y un lul de volumen de muestra. La ecuación de Beer-Lambert indica que la concentración molar de la muestra se obtiene al dividir la absorbancia medida de una solución de muestra por su coeficiente de extinción molar, es decir, que la concentración de una molécula en solución depende de la cantidad de luz absorbida de las moléculas disueltas (Mayerhöfer *et al.*, 2020). Una característica del ADN es que absorbe la luz ultravioleta (UV) a 260 nm y permite estimar su concentración mediante espectrofotometría (Cornejo *et al.*, 2014). La interferencia de contaminantes puede determinarse calculando un “cociente”. Para estimar la pureza del ADN se emplea el cociente A260/A280. Los cocientes respectivos del ADN puro son aproximadamente de 1.8 y 2.0, un valor entre 1.6 – 1.8 se considera aceptable, proporciones menores a este valor indican la presencia de proteínas. Una segunda valoración de la pureza de ácidos nucleicos es la proporción 260/230, los valores aceptados de pureza se encuentran en el rango de 2.0 a 2.2 si la relación es menor indican la presencia de contaminantes como hidratos de carbono, péptidos, fenoles, compuestos aromáticos u otras sustancias (Cornejo *et al.*, 2014).

1.2.11.5 *Electroforesis*

La electroforesis es una técnica que se utiliza para separar las moléculas biológicas, ya sean ADN, proteínas o ARN en función de si son más grandes o más pequeñas. Las biomoléculas se someten a un campo eléctrico a travez de una matriz de gel, ya sea geles de agarosa o de acrilamida, que forman una especie de red con poros por donde pasan las moléculas biológicas con carga negativa hacia el polo positivo cuando se somete a un campo eléctrico durante un tiempo determinado, ya que la carga de las moléculas de ADN es negativa por la presencia de los grupos fosfato (P⁻). Los fragmentos más pequeños migraran mas rápido y terminarán en la parte inferior del gel, porque han ido más lejos, y las más grandes terminarán quedándose en la parte superior; de esta manera se juntaran los fragmentos de tamaños similares por que han migrado al mismo ritmo y podremos verlos formando una banda en el gel (Cornejo *et al.*, 2014) (Slater *et al.*, 2002).

Para realizar esta tecnica es necesario tener un transiluminador de luz ultravioleta, una cámara de electroforesis, una fuente de poder, un tampon de electroforesis el cual tendrá el pH entre 8 – 8.3, y los iones necesarios para que fluya la corriente y pueda migrar el ADN (Strutz & Stellwagen,

1998). Para la preparación de geles de agarosa se debe tener en cuenta el tamaño de los fragmentos que se espera obtener y utilizar una concentración de agarosa mayor o menor (Rickwood & Hames, 1990), según ese dato obtener una matriz en gel de agarosa con poros mas o menos grandes y una mejor resolución de las bandas. Por el contrario, si no se conoce el tamaño de los fragmentos, se puede empezar con agarosa al 1%.

1.2.11.6 Comprobación de la integridad del ADN por electroforesis

Además de conocer la cantidad y calidad del ADN por espectrofotometría, es importante conocer si el ADN obtenido está integro. La integridad del ADN se puede observar mediante electroforesis en gel de agarosa. Si el ADN está integro, se debe observar una banda estrecha cercana al pozo en que se colocó la mezcla de ADN. Si está fragmentado, se observará una banda de más de un cm de ancho o un sendero luminoso en el carril de la muestra. El ADN fragmentado dificulta la amplificación de productos de PCR de alto peso molecular y afecta la reproducibilidad de las técnicas (Cornejo *et al.*, 2014).

1.2.12 TECNOLOGÍA DEL ADN RECOMBINANTE

La tecnología del ADN recombinante es un método que depende de la realización artificial de eventos de recombinación genética. Es posible obtener cualquier gen ordenado o su producto con este método. Antes del paso de producción, el gen ordenado se deriva del cromosoma original mediante la enzima endonucleasa y se integra a un vector como un plásmido o un fago, después, este vector se transforma en una bacteria o una levadura. Luego, el gen o la proteína ordenados se producen en las cantidades deseadas cultivando estos microorganismos (Koçoğlu & Yalçinkaya, 1992).

1.2.12.1 Clonación

La clonación del ADN es una técnica molecular importante para la obtención de grandes cantidades de un gen de interés para diversas aplicaciones, esta secuencia de ADN es ligado a un vector, que es un vehículo para transportar ADN recombinante a una célula huésped, generalmente las células de *E. coli*, quienes son sometidas a un tratamiento físico o químico para recibir al vector con el ADN recombinante, convirtiéndose en bacterias recombinantes o bacterias clonadas con el gen de interés. Durante el proceso de división celular, las bacterias replican el vector recombinado

y lo transfieren a sus descendientes y el gen de interés es clonado al mismo tiempo, obteniendo muchas copias del gen (Campbell & Reece, 2006)

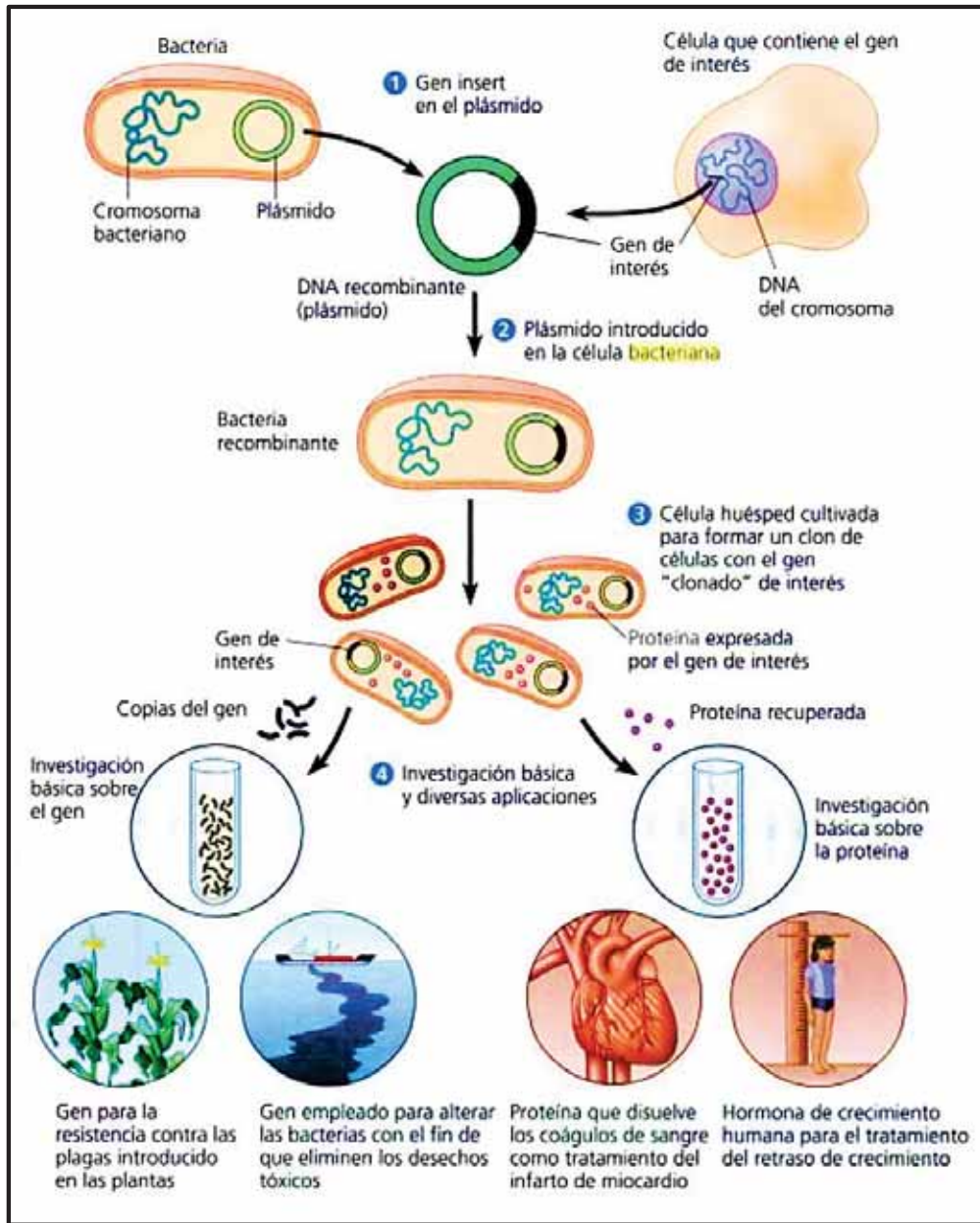


Figura 8: Panorama general de la clonación de genes con un plásmido bacteriano que muestra los diversos usos de los genes clonados.

Fuente: Campbell & Reece, 2006.

1.2.12.2 Vectores de clonación

Los vectores de clonación son plásmidos, término que fue presentado por primera vez por el biólogo molecular norteamericano Joshua Lederberg en 1952, se definen como moléculas de ADN extracromosómico generalmente circular, que se replican y transmiten independientes del ADN cromosómico. Dentro del área de biología molecular los vectores de clonación llevan insertados fragmentos de ADN de hasta 4 kb, típicamente un gen, que se desea introducir o expresar en el hospedador y para poder ser insertado es necesario cortar el vector con enzimas de restricción y una vez introducido pasa a denominarse, vector recombinante (T. Brown, 2008).

Todos los vectores de clonación deben al menos contener (Campbell & Reece, 2006):

- Un origen de replicación para poder tener más de una copia del mismo en la célula infectada.
- Dos genes que confieran resistencia a diferentes antibióticos, (cloranfenicol y ampicilina) lo que permite la identificación de las células que portan a dicho vector.

Por otro lado, los plasmidos presentan tres topoisofomas, que son: moléculas enrolladas, moléculas lineales de longitud completa, y moléculas circulares cerradas (Watson *et al.*, 2013).

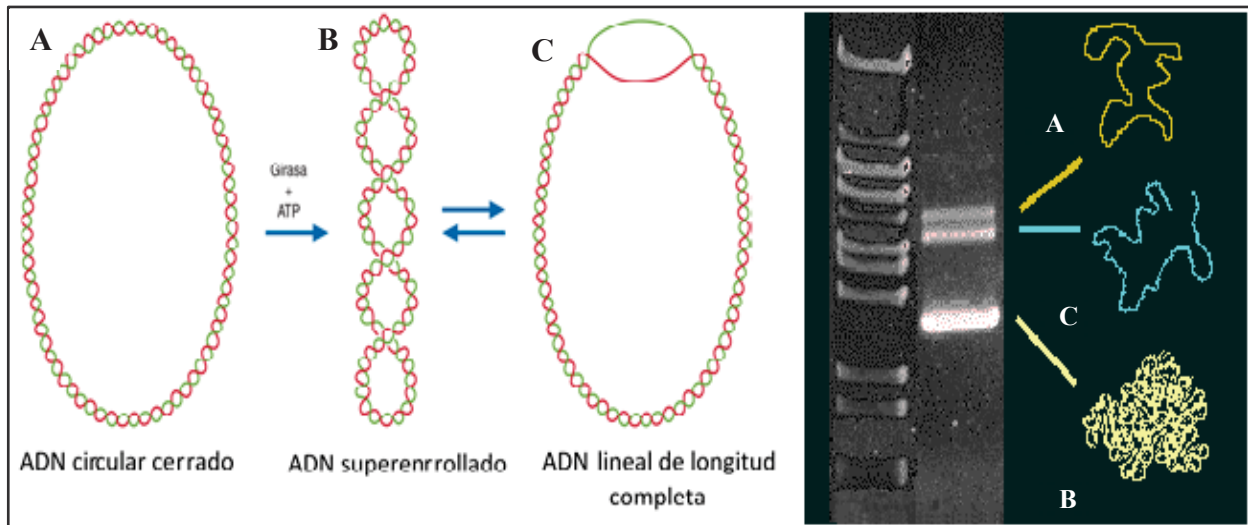


Figura 9: Topoisofomas de los vectores plasmídicos. A. ADN circular cerrado B. ADN enrollado C. ADN lineal de longitud completa.

Fuente: Watson *et al.*, 2013.

Vector PGEM-T

Los sistemas de vectores pGEM-T son sistemas convenientes para clonar productos de PCR que hayan sido amplificados con polimerasas termoestables que a menudo agregan una sola desoxiadenosina, de forma independiente de la plantilla, a los extremos 3' de los fragmentos amplificados proporcionando extremos compatibles para Adenina. La característica principal del vector pGEM-T se centra en mejorar en gran medida la eficiencia de la ligadura evitando la recircularización del vector, debido a que presenta brazos de Timina individuales en el sitio de inserción, asimismo, los promotores de ARN polimerasa T7 y SP6 flanquean una región de clonación múltiple que abarca diferentes enzimas de restricción (Promega Corporation, 2007).

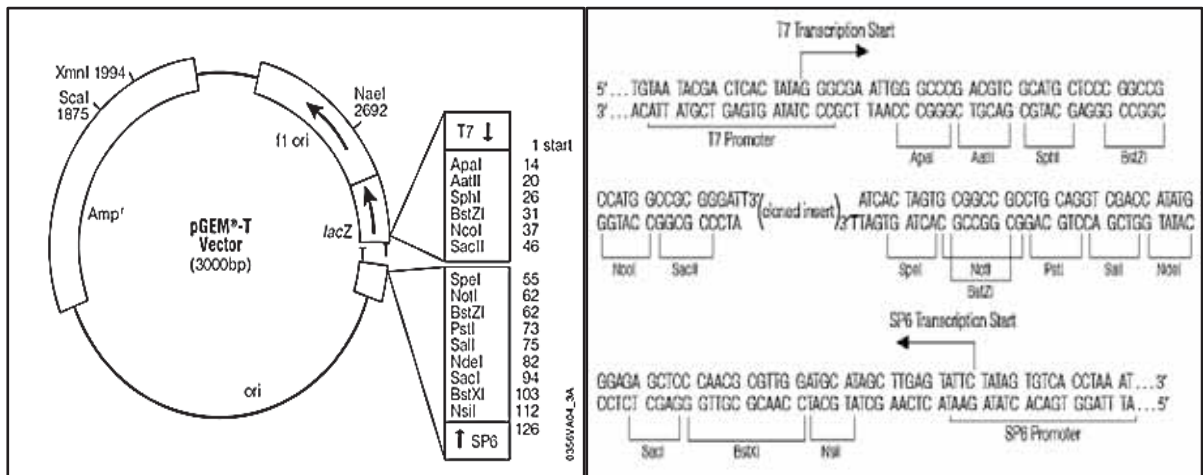


Figura 10: El vector pGEM®-T (número de acceso GenBank® X65308).

Fuente: Manual Pgem-T

1.2.12.3 Transformación de células no competentes

La transformación bacteriana es un proceso por el cual las células captan ADN libre presente en el medio, para lo cual las células deben estar en un estado fisiológico de competencia, por lo que se les denomina células competentes, en este estado la bacteria presenta alteraciones en su pared y membranas celulares que permiten la entrada de ácidos nucleicos. Este efecto se desarrolla en los procariontes como resultado de una transformación natural, en el que pueden tomar material genético de su entorno, absorberlos y recombinarlos en su propio genoma. Como consecuencia, la competencia natural para la transformación se considera un modo primario de transferencia horizontal de genes (Blokesch, 2016). Sin embargo, los científicos han encontrado maneras de

inducir artificialmente la competencia celular y mejorar la capacidad de una célula bacteriana, ya que es importante para el desarrollo de la biología molecular en procedimientos de clonación y subclonación. Se puede tratar a la célula de manera química o física para introducir ADN plasmídico:

Las células químicamente competentes que se tratan con cloruro de calcio y cloruro de magnesio para facilitar la unión del ADN plasmídico a la membrana celular competente. La célula competente se calienta alternativamente en un baño de agua, esto abre los poros de la membrana celular permitiendo la entrada del plásmido (Bergmans *et al.*, 1981).

Las células electrocompetentes funcionan utilizando el proceso de electroporación, que induce la competencia mediante la aplicación de un pulso eléctrico breve e intenso que crea poros que permiten que el material genético permee la membrana bacteriana (Drury, 1994).

Células competentes de *Escherichia coli* JM109

Las células competentes de *E. coli* cepa JM109 están disponibles en *Promega Corporation* <https://worldwide.promega.com/>. Su genotipo es *endA1, recA1, gyrA96, thi-1 hsdR17 (r_k⁻ m_k⁺), supE44 relA1 Δ(lac-proAB), [F' traD36 proAB laqIq ZΔM15]*, la descripción de los genes se observa en la tabla 7, estas células son utilizadas para la transformación y propagación de plásmidos de clonaje de productos de PCR, es por eso que para una transformación conveniente en dos eficiencias existe la alta eficiencia a más de 10⁸ UFC/μg para clonación y uso único, y la eficiencia de subclonación a más de 10⁷ UFC / μg.

Tabla 7: Marcadores genéticos en *E. coli* JM109

SÍMBOLO	DESCRIPCIÓN	EFEECTO DE LA MUTACIÓN
<i>endA1</i>	Mutación de endonucleasa I específica de ADN.	Mejora la calidad de los aislamientos de ADN plasmídico.
<i>recA1</i>	Mutación en recombinación.	Minimiza la recombinación del ADN introducido con el ADN del huésped, aumentando la estabilidad de los insertos.
<i>gyrA96</i>	Mutación de la ADN girasa	Confiere resistencia al ácido nalidíxico.
<i>thi-1</i>	Mutación en el metabolismo de la tiamina.	Los mutantes requieren tiamina para crecer en medios mínimos.
<i>hsdR17</i>(r_k^- m_k^+)	Restricción del ADN del huésped y mutación del sistema de metilación.	Permite la clonación sin escisión del ADN transformado por endonucleasas de restricción endógenas.
<i>relA1</i>	Mutación de ppGpp sintetasa I	Permite la síntesis de ARN en ausencia de síntesis de proteínas. La ppGpp sintetasa I asociada a ribosomas puede detectar ribosomas con una ARNt sin carga en el sitio A, un signo de inanición.
<i>proAB</i>	Mutaciones en el metabolismo de la prolina.	Requiere prolina para crecer en medios mínimos
<i>supE44</i>	Mutaciones supresoras	Suprime las mutaciones ámbar (UAG).
<i>traD36</i>	Mutación del factor de transferencia	Previene la transferencia del episoma F'.
<i>lacIq</i>	Sobreproducción de la proteína represora lac	Conduce a altos niveles de la proteína represora lac, inhibiendo la transcripción del promotor lac
<i>lacZΔM15</i>	Supresión parcial del gen de la β-d-galactosidasa	Permite la complementación de la actividad de β-galactosidasa por la secuencia de complementación α en los vectores pGEM [®] -Z. Permite la selección azul / blanco para colonias recombinantes cuando se colocan en placas sobre X-Gal.

Fuente: Manual de promega – Marcadores genéticos de *E. coli*

1.2.12.4 *PCR de colonias*

La PCR de colonias es un método conveniente de alto rendimiento para determinar la presencia o ausencia de inserto de ADN en construcciones de plásmidos, también se puede usar para determinar la orientación del inserto, por lo tanto, es una herramienta poderosa para la detección rápida y fácil a través de cantidades potencialmente grandes de colonias para distinguir los verdaderos positivos de los falsos positivos (Bergkessel & Guthrie, 2013). Los transformantes individuales se pueden lisar en agua con un breve paso de calentamiento o agregarse directamente a la reacción de PCR y lisarse durante el paso de calentamiento inicial. Este paso de calentamiento inicial provoca la liberación del ADN plasmídico de la célula, por lo que puede servir como plantilla para la reacción de amplificación. Los cebadores específicos de inserto pueden usarse para determinar si la construcción contiene el fragmento de ADN de interés, además de proporcionar información sobre la especificidad y el tamaño del ADN de inserto, alternativamente, los cebadores que se dirigen al ADN del vector que flanquea el inserto se pueden usar para determinar si el inserto tiene o no el tamaño molecular correcto. En todos los diseños experimentales, la presencia o ausencia de un amplicón de PCR y el tamaño del producto se determinan por electroforesis junto con un marcador de tamaño de ADN en un gel de agarosa (*PCR de colonias / NEBRASKA, s. f.*)

1.2.12.5 *Criopreservación de células con glicerol*

La criopreservación (*crio* = frío) es el uso de temperaturas muy bajas para preservar células y tejidos vivos estructuralmente intactos. Los crioprotectores son soluciones que protegen del frío actuando sinérgicamente en la deshidratación de la célula extrayendo o desplazando el agua del citoplasma y así evitar que durante la congelación se formen cristales de hielo en el interior de la célula, pero para ser biológicamente aceptables el compuesto debe ser altamente soluble en agua y permanecer así a bajas temperaturas para producir una profunda depresión de la temperatura de congelación; debe poder penetrar en las células; y debe tener una baja toxicidad para que pueda usarse en las altas concentraciones que se requieren para producir estos efectos. Muchos compuestos tienen tales propiedades, que incluyen glicerol, dimetilsulfóxido, etanodiol y propanodiol (Pegg, 2007).

Los crioprotectores se dividen en permeables y no permeables, entre los permeables se encuentra el glicerol, sus propiedades como crioprotector fueron descubiertas en 1949 por los científicos Polge, Smith y Parkes (Polge *et al.*, 1949), el glicerol forma uniones de hidrógeno con el agua a través de sus grupos -OH, mejorando la estabilidad de las células al reducir al mínimo el contenido de agua intracelular para evitar la vitrificación y proteger las macromoléculas en el ambiente interno de la célula (Ávila-Portillo *et al.*, 2006).

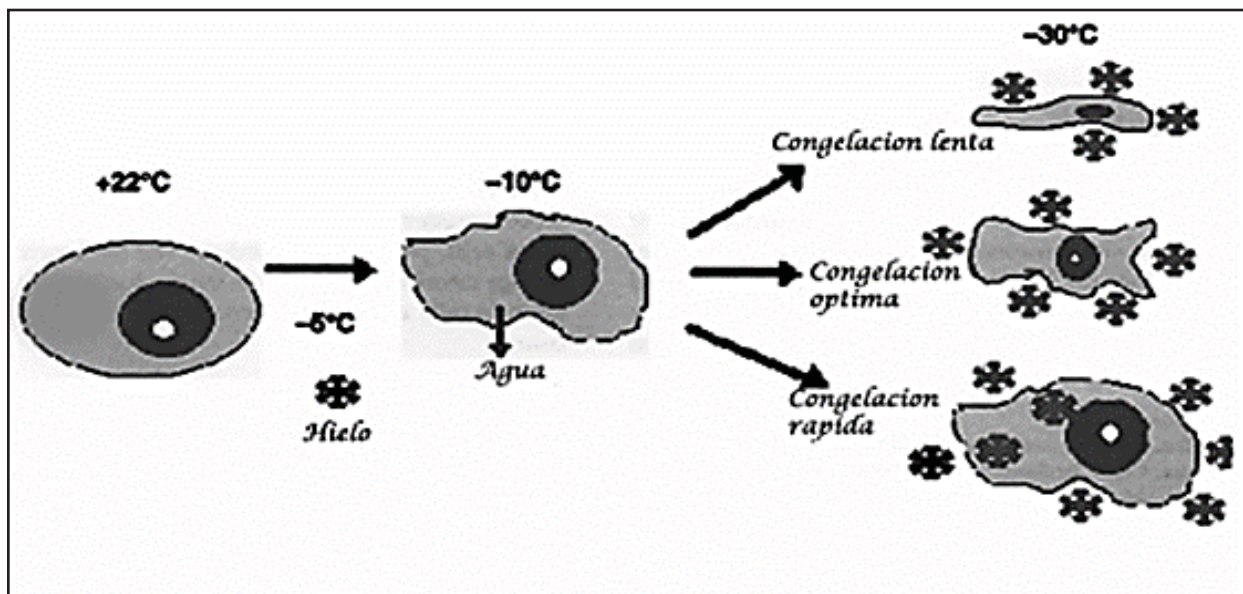


Figura 11 Criopreservación y su efecto sobre la célula

Fuente: Mazur, 1977.

CAPÍTULO II
MATERIALES Y MÉTODOS

2.1 LUGAR DE ESTUDIO

El análisis *in silico* se realizó en el aula virtual C-234 de la Facultad de Ciencias de la Universidad Nacional San Antonio Abad del Cusco y la parte experimental se realizó en los laboratorios N°9 y N°10 del Instituto Universitario de Enfermedades Tropicales y Salud Pública de Canarias de la Universidad de La Laguna – España, en el 2018.

2.2 MATERIALES

2.2.1 MATERIAL BIOLÓGICO

- ✓ Cepa de *Chromohalobacter salexigens* MP25462.
- ✓ Cepas bacterianas de *Escherichia coli* JM109 (Promega).
- ✓ Plásmido pGEM-T (Promega).

2.2.2 MATERIAL DE LABORATORIO

2.2.2.1 Equipos de laboratorio

- ✓ Micropipetas de 0.5-10 µl, 10-100 µl y 100-1000 µl (Eppendorf).
- ✓ Agitador de tubos (Vórtex)
- ✓ Multi-Rotator PTR-35 (Grant-bio)
- ✓ Microcentrifuga 5418 (Eppendorf)
- ✓ Baño maria (Univeba)
- ✓ Lector de Imágenes ChemiDoc XRS+ (Bio-Rad).
- ✓ Centrifuga Megafuge 16R (Thermo Scientific).
- ✓ Cámara de flujo Laminar Bio-II-A (TelStar).
- ✓ Congeladora (Corbero).
- ✓ Microondas.
- ✓ Thermo Mixer C (Eppendorf).
- ✓ Termociclador T Professional Basic (Biometra).
- ✓ Termociclador T Professional Gradient (Biometra).
- ✓ Cámara de electroforesis Sub-Cell GT Basic (Bio-Rad).
- ✓ Fuente de alimentación eléctrica (Bio-Rad).
- ✓ Cubeta de electroforesis Wide Mini-Sub Cell GT Basic (Bio-Rad).

- ✓ Balanza electrónica JT-1200 (Cobos).
- ✓ Transiluminador UV.
- ✓ Espectrofotometro DS-11 + (DeNovix).
- ✓ Congeladora -20°C (Innova).
- ✓ Autoclave (Raypa).
- ✓ Destilador de agua MiliQ Paranity TU (VWR).
- ✓ Reactivos y kits PureYield™ Plasmid Miniprep System (Promega).
- ✓ Ilustra™ Tissue & Cells genomic Prep Mini Spin Kit (GE Healthcare).
- ✓ Ilustra™ GFX PCR DNA and Gel Band Purification (GE Healthcare).
- ✓ Agarosa de grado molecular (BIOLINE).
- ✓ T4 DNA ligasa (Thermo Scientific).
- ✓ Marcador molecular Lambda Hind III (BIOLINE).
- ✓ Marcador molecular Hyper Ladder II (BIOLINE).
- ✓ Tampon de reacción 2X (Thermo Scientific).
- ✓ Cloruro de Sodio (NaCl)
- ✓ Ácido etilendiaminotetraacético (EDTA)
- ✓ Taq DNA polimerasa (Promega)
- ✓ Cloruro de magnesio (MgCl₂)

2.2.2.2 Material fungible y otros

- ✓ Probetas graduadas, de 250 mL, y 1000 mL.
- ✓ Puntas para micropipetas de 1-20 µL, 1-200 µL, 100-1000 µL.
- ✓ Tubos de microcentrífuga de 1.5 mL
- ✓ Tubos de PCR de 0.2 mL.
- ✓ Tubos de criopreservación de 1.8 mL.
- ✓ Tubos de centrífuga 15 mL y 50 mL.
- ✓ Placas Petri de 100mm x 20mm.
- ✓ Matraces de 50 mL y 500 mL.
- ✓ Pipetas Pasteur de 3 mL
- ✓ Jeringas de 10 mL.
- ✓ Filtros de jeringa de 0.22 µm.

- ✓ Gradillas para tubos de 1.5 ml.
- ✓ Asas de siembra descartables.
- ✓ Palitos de mondadientes estériles.
- ✓ Mechero de alcohol.
- ✓ Encendedor.
- ✓ Papel toalla.
- ✓ Equipo de protección personal (barbijo, mandil, guantes).

2.2.3 MATERIAL BIOINFORMÁTICO

2.2.3.1 Datos primarios

- ✓ Secuencias crudas pe-1_S15_L001_R1_001.fastq
- ✓ Secuencias crudas pe-1_S15_L001_R2_001.fastq

2.2.3.2 Bases de datos

- ✓ RAST: *Rapid Annotation using Subsystem Technology* (Anotación rápida utilizando tecnología de subsistema) – Servicio de anotación del genoma procariótico.
<http://rast.nmpdr.org/rast.cgi>.
- ✓ NCBI: *National Center Biotechnology Information* – Centro nacional de información biotecnológica.
<https://www.ncbi.nlm.nih.gov/>
- ✓ MEROPS: *Peptidase Database* Es un recurso de información para las peptidasas.
<https://www.ebi.ac.uk/merops/index.shtml>.

2.2.3.3 Softwares y programas informáticos

- ✓ Sistema operativo Linux, Ubuntu (Versión 16.04LTS)
- ✓ Trimmomatic v. 0.36
- ✓ SPAdes v.3.13.1
- ✓ Quast v. 4.6.1
- ✓ I-Tasser
- ✓ Gene Runner
- ✓ Mega VII
- ✓ ImageLab
- ✓ Microsoft 2010

2.3 METODOLOGÍA

2.3.1 TIPO DE INVESTIGACIÓN

Descriptiva – experimental

2.3.2 FLUJOGRAMA DEL TRABAJO DE INVESTIGACIÓN

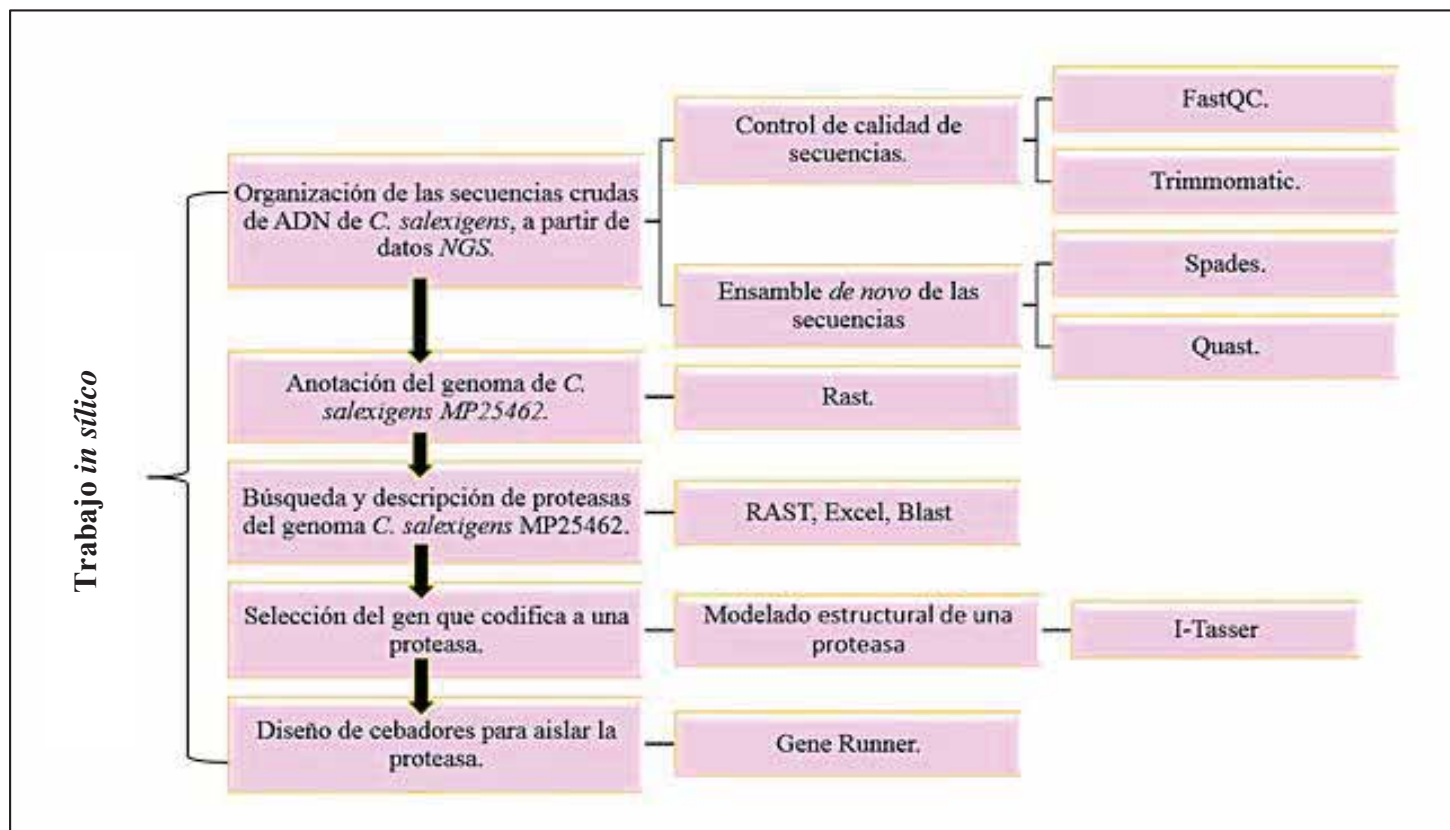


Figura 12: Flujo de secuencia de actividades del trabajo de investigación in silico.

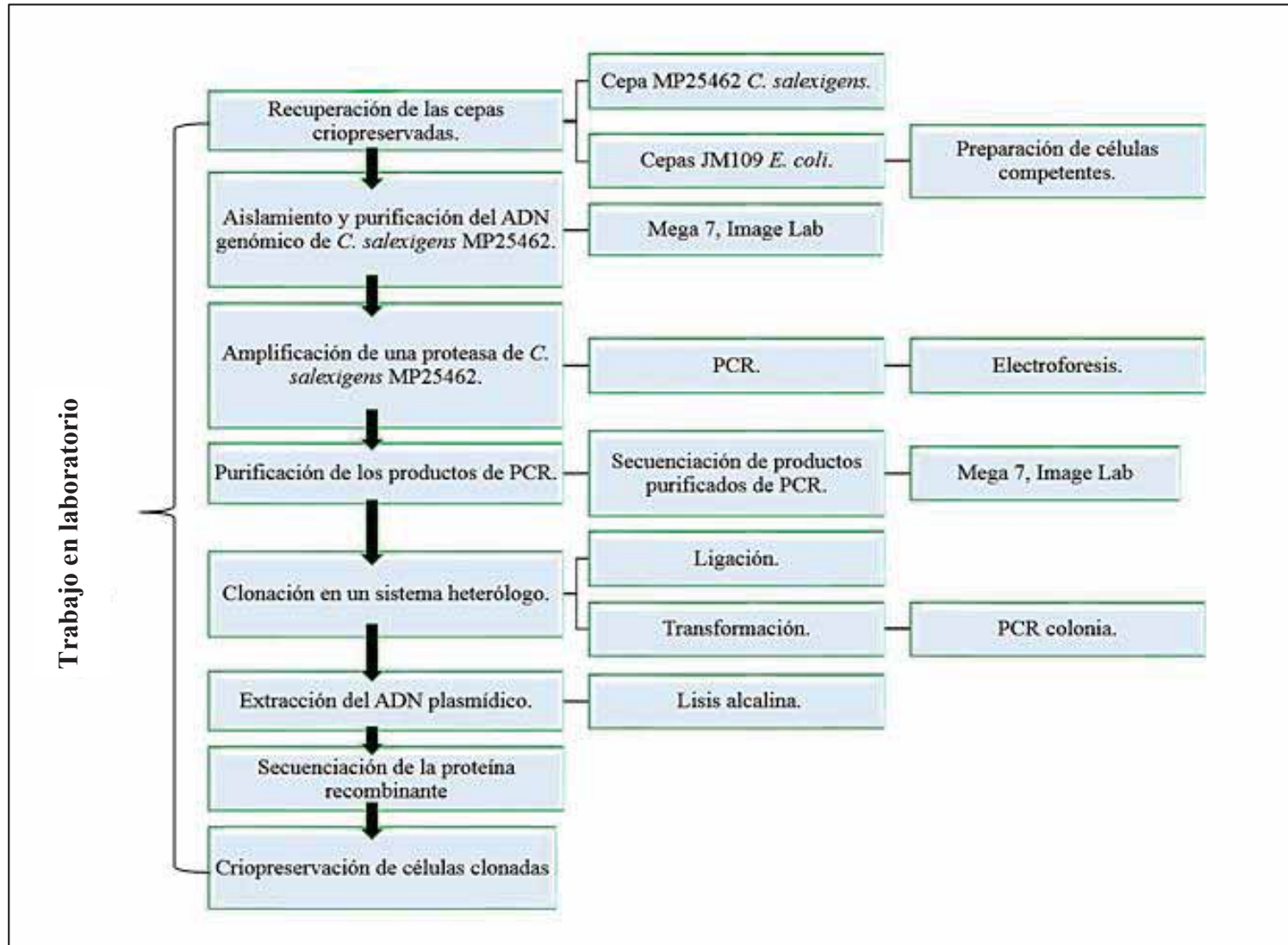


Figura 13: Flujo de secuencia de actividades del trabajo de investigación en laboratorio.

2.4 MÉTODOS Y FUNDAMENTOS

La presente investigación tiene como datos primarios las secuencias crudas de ADN de la bacteria *Chromohalobacter salexigens* MP25462 resultado del Proyecto: “Secuenciación del metagenoma de bacterias halófilas en ambientes salinos del departamento del Cusco” con resolución N° R-005-2016-UNSAAC.

2.4.1 TRABAJO IN SILICO

2.4.1.1 Control de calidad de las secuencias crudas

Se utilizó el software FastQC v. 0.11.7, que se ejecutó desde la terminal del sistema operativo de Linux, Ubuntu, los archivos de entrada fueron las dos secuencias crudas *.fastq Forward* y *.fastq Reverse*, se obtuvo un formato “html” como archivo de salida en cada secuencia respectivamente, en donde se puede visualizar de forma general las calidades de las lecturas que se obtuvieron del proceso de secuenciación y que fueron consideradas para sus correcciones.

En la Figura 14, se describe un conjunto modular de análisis que se considera dentro de la interfaz de este programa.

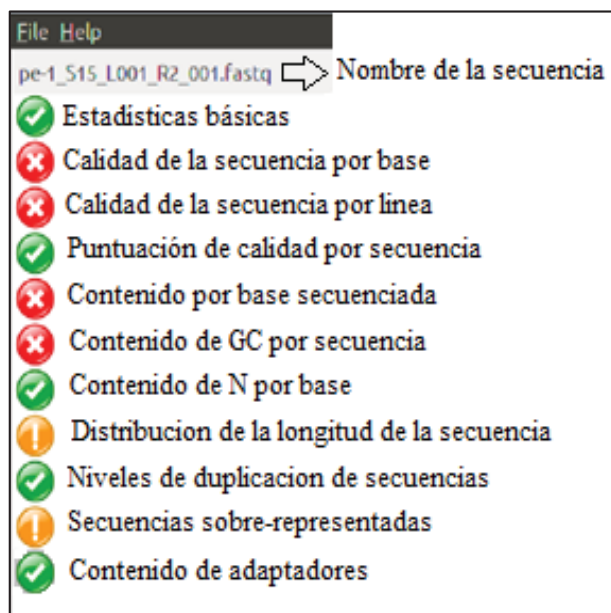


Figura 14: Parámetros de calidad que evalúa el software FastQC.

Fuente: Traducido de la interfaz de FastQC

2.4.1.2 Recorte de las secuencias

Se utilizó el software Trimmomatic v. 0.36 encargado de varios recortes útiles para mejorar la calidad de las secuencias, este programa se ejecutó desde la terminal del sistema operativo de Linux, Ubuntu, los archivos de entrada son las dos secuencias crudas en formato fastq y como producto del análisis resultaron 4 archivos de salida (2 Paired-End y 2 Single-End). La línea de comandos es la siguiente:

```
→ /Usuario/Downloads/trimmomatic java -jar trimmomatic.0.36.jar PE -phred33
   'archivo_1.fastq' 'archivo_2.fastq'          'archivo_1_PE.fastq' 'archivo_1_UP.fastq'
   'archivo_2_PE.fastq'          'archivo_2_UP.fastq' ILLUMINACLIP:/home/usuario/
   Aplicaciones/Trimmomatic.0.36/NexteraPE-PE.fa:2:30:10 LEADING:3TRAILING:28
   HEADCROP:20 MINLEN:100 SLIDINGWINDOW:4:15
```

2.4.1.3 Ensamble de las secuencias crudas de adn de *C. salexigens*.

El producto de la secuenciación masiva con IlluminaMiseq fue ensamblado con el programa SPAdes v.3.13.1, este es un ensamblador automático que genera un archivo de salida en formato FASTA, *contigs.fa*, este archivo alberga los datos de las secuencias ensambladas. El programa se ejecuto desde la terminal de Linux-Ubuntu con la siguiente línea de comando:

```
→ /Usuario/Downloads/SPAdes-3.13.1-Linux/bin/ spades.py -1 archivo_1.fastq -2
   archivo_2.fastq -o archivodesalida
```

2.4.1.4 Evaluación de calidad del ensamble

La calidad del ensamble se analizó con el programa Quast v. 4.6.1 el cual generó un archivo en formato *.html*, donde se observó diversas métricas como la longitud del genoma, las variaciones del N50 basados en el alineamiento de los bloques, el contenido de GC (guanina, citosina), la cantidad de Ns (bases no identificadas), la longitud y número de *contigs* que conforman el genoma ensamblado. Esta herramienta se ejecuto desde la terminal de Linux-Ubuntu, utilizando el archivo *contigs.fa* resultado del ensamble con SPAdes; la línea de comando es la siguiente:

```
→ /Usuario/Downloads/quast-4.6.1-Linux/ python quast.py archivocontigs.fa
```

2.4.1.5 Anotación del genoma *C. salexigens* MP25462

La anotación del genoma de *C. salexigens* cepa MP25462 se realizó con el programa RAST (*Rapid Annotation using Subsystem Technology*) el cual solicitó subir el archivo *contigs.fa* producto del ensamble y completar datos de taxonomía del microorganismo. El resultado se obtuvo a las 24 horas y en la página de este programa se logró visualizar diferentes grupos de genes divididos en subsistemas, además de diferentes opciones para descargar archivos de interés, entre ellos, Excel (archivo de extensión .xls), dentro del cual se pudo hallar una lista de todos los genes que posee el genoma de esta bacteria, su respectiva descripción, secuencia de nucleótidos y secuencia de aminoácidos.

2.4.1.6 Descripción de proteasas del genoma de *C. salexigens* MP25462.

Para la descripción de las enzimas proteolíticas que alberga el genoma de *C. salexigens*, se descargó el archivo Excel del programa de anotación RAST, en el cual se halla una lista de todos los genes codificantes organizados por nombre, función, cobertura de la secuencia en el genoma, secuencia de aminoácidos, secuencia de nucleótidos, sentido de la secuencia dentro del genoma, ubicación respecto al genoma. Así mismo, para encontrar los genes que transcriben solo a proteasas, se utilizó todas las denominaciones que representan a este grupo de enzimas según la clasificación UIBBM, es así que se utilizó la opción *Buscar* en el Excel, en donde se escribió las siguientes palabras: *peptidase*, *aminopeptidase*, *carboxipeptidase*, *dipeptidil peptidase*, *serinepeptidase*, *cysteinepeptidase*, *aspartylpeptidase*, *metallopeptidase*, de esta manera se fue reuniendo las secuencias de cada enzima proteolítica, las cuales a su vez fueron filtradas utilizando la secuencia de aminoácidos en la base de datos del Blast-NCBI, esto para corroborar la identidad de cada proteasa hallada. Seguidamente, se procedió a agruparlas según su sitio de acción catalítica, ya sean endopeptidasas o exopeptidasas, cada proteasa con su respectiva descripción según sus antecedentes, además, se indicó su número de nucleótidos, su número de aminoácidos, y se definió su ubicación intergenica así como los genes próximos de cada proteasa descrita.

2.4.1.7 Modelamiento estructural de la proteasa Amidohidrolasa MP25462

Posterior a la realización de la búsqueda de la secuencia nucleotídica del gen que expresa la Amidohidrolasa se comenzó con la descripción de la misma con el servidor I-TASSER (*Iterative Threading ASSEmbly Refinement*). Este servidor generó predicciones estructurales automatizadas

de la proteína Amidohidrolasa de longitud completa en 3D así como la ubicación del sitio activo, el ion metálico y los ligandos. Para lo cual, se inició sesión en la plataforma de I-Tasser con una cuenta institucional, seguidamente se subió la secuencia de aminoácidos en formato FASTA en la opción “*Or upload the sequence from your local computer*”, indicando la identificación o nombre de la proteína en estudio. Finalmente se seleccionó el botón “*Run I-TASSER*” para terminar de subir la secuencia y se ponga en cola debido a la demanda de trabajo que tiene el servidor. El proceso de identificación estructural, se basa en que el programa I - Tasser utiliza una serie de algoritmos que se van alineando con la plantilla estructural más similar del PDB (*Protein Data Base*), este proceso tardó aproximadamente 96 horas. Al término, se obtuvo el modelado 3D de la Amidohidrolasa la cual tuvo alta similitud con una Amidohidrolasa cristalizada del PDB (*Protein Data Base*). El servidor I-TASSER mediante su base de datos otorgó la predicción de funcionalidad de la proteína Amidohidrolasa de *C.salexigens* MP25462 realizando una comparación estructural con las proteínas cristalizadas que abarca el PDB tomando en cuenta el C-score, la cobertura y comparaciones de ligando, es así, que el servidor otorgó el archivo PDB de la enzima cristalizada con mayor probabilidad de ubicación de sitio activo así como ligandos implicados, con lo cual se realizó la descarga del archivo de extensión *.pdb*.

Con la información obtenida del artículo de la Amidohidrolasa cristalizada se realizó la descripción de la proteína Amidohidrolasa de *C. salexigens* cepa MP25462, donde se ubicó los sitios de unión de ligando y el sitio activo, ya que, con esta similaridad obtenida se consiguen los aminoácidos comprometidos en el sitio activo así como el hallazgo del ion metálico de la proteína para luego realizar la simulación de su estructura proteica.

2.4.1.7 Diseño de cebadores

Para amplificar la enzima proteolítica Amidohidrolasa presente en el genoma de *C. salexigens* se diseñó un par de cebadores con el programa *Gene Runner v. 6.5.51 Beta*, para ejecutar el programa, primero se cargó la secuencia de nucleótidos del gen, se identificó la posición del inicio y final del gen, y con la opción *Oligo Analysis*, se realizó el análisis seleccionando previamente entre 18-20pb de ambos oligos y luego se consideró el porcentaje de GC entre 50-60%, con una $T_m < 5^\circ\text{C}$ entre la pareja de cebadores y se evitó que las terminaciones 3' sean complementarias, ya que esto puede llevar a la formación de dímeros.

2.4.2 TRABAJO EN LABORATORIO

2.4.2.1 Recuperación a partir de la criopreservación de *C. salexigens* MP25462.

Para recuperar la bacteria halófila *Chromohalobacter salexigens* cepa MP25462 de la criopreservación a -80°C se procedió de la siguiente manera:

- Se preparó un medio líquido SW al 15% de NaCl (Anexo 5).
- Se descongeló en hielo la cepa MP25462 y se cultivó 200 μl en 5ml del medio líquido SW al 15% de NaCl.
- Se incubó a 37°C con movimiento constante a 300 RPM durante 18 horas.
- A partir del cultivo fresco de la cepa MP25462 se sembró en placas con medio SW al 15% de NaCl y se dejó incubar a 37°C durante 48 horas .

2.4.2.2 Preparación de células competentes *E. coli* JM109

Se preparó células competentes de *E. coli* para ser utilizadas durante la fase de la transformación según el protocolo modificado por (Tang *et al.*, 1994). La cepa no competente JM109, fue tratada químicamente para su competencia de la siguiente manera (Figura15):

- Inicialmente se preparó medio LB (Anexo 6) donde se dejó incubando la cepa en medio LB líquido (5ml) durante toda la noche a 37°C .
- A partir del cultivo fresco se tomó 1ml y se transfirió a 50ml de LB líquido, este último se dejó a 37°C , y se leyó a una longitud de onda de 450 nm hasta una D.O. = 0.6.
- Luego, se dejó enfriar el matraz en hielo por 10min y se trasvaso a un falcón de 50ml que se centrifugo 4000 RPM durante 17min a 4°C , se descartó el sobrenadante y el pellet se re suspendió con 50ml de MgCl_2 0.1M, luego se dejó la muestra en hielo por 15min.
- La segunda centrifugación fue a 4000 RPM durante 7min a 4°C , se descartó el sobrenadante y el pellet se re suspendió en 25ml de CaCl_2 0.1M, luego se dejó la muestra en hielo por 1 hora.
- La tercera centrifugación fue a 4000 RPM durante 5min a 4°C , se descartó el sobrenadante y se re suspendio el pellet en 5ml de CaCl_2 0.1M.
- Finalmente, a este volumen de la muestra se le añadió glicerol al 20% para su almacenamiento a -80°C , en tubos de 1.5ml.

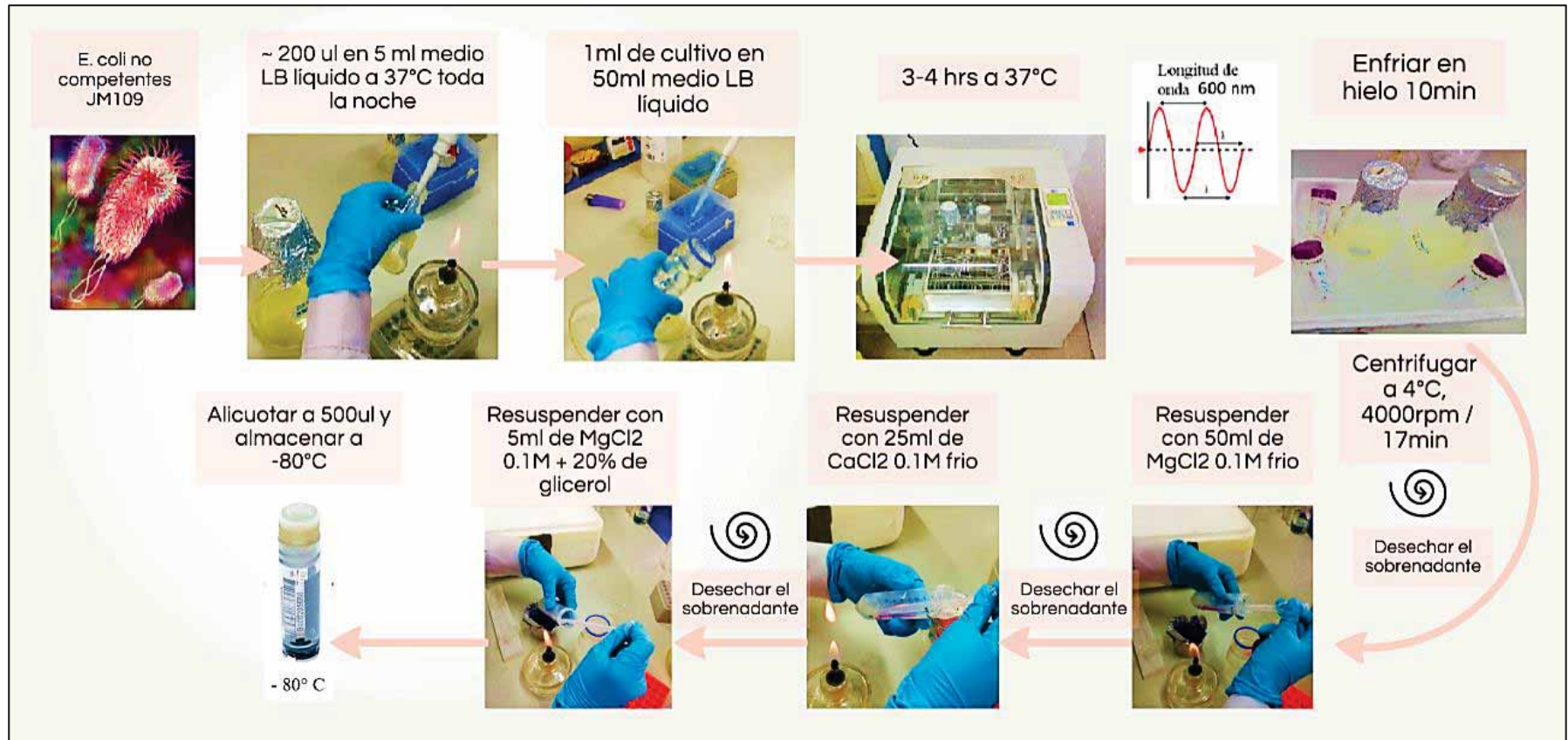


Figura 15: Proceso de obtención de células competentes para la etapa de la transformación (clonación).

2.4.2.3 Aislamiento y purificación del ADN genómico de *C. salexigens* MP25462

Para el aislamiento y purificación de ADN genómico de *Chromohalobacter salexigens* MP25462 se utilizó el *Kit GE Healthcare Illustra Tissue & cells genomicPrep Mini Spin* (Figura 16), se describe a continuación:

- Descongelar en hielo la cepa MP25462 y se cultivó 200µl en 5ml de medio líquido SW al 15% de NaCl, durante 18 horas a 37°C.
- Separar 1ml del cultivo fresco en un tubo de 1.5ml y se centrifugó a 16000g por 3 minutos, este paso se realizó tres veces, descartando el sobrenadante al final de cada centrifugación, con el objetivo de obtener un pellet concentrado en el tubo de 1.5ml.
- Llevar a lisis las células añadiendo 100µl de tampón de lisis Tipo I en la muestra y se dejó incubar durante 15min a 56°C.
- Agregar a la muestra 5µl de ARNasa (10mg/ml) para la remoción del ARN y se dejó incubar durante 15 minutos a temperatura ambiente.
- Purificar la muestra añadiéndole 500µl de tampón de lisis Tipo IV y se dejó incubar durante 10 minutos a temperatura ambiente, luego se centrifugó a 11000g durante 10 segundos.
- Luego se lavó y dejó secar la columna añadiendo 500µl de tampón TipoIV y 500µl de tampón TipoVI con centrifugaciones a 11000g durante 3 minutos cada uno respectivamente.
- Finalmente, el ADN genómico purificado se eluyó con 200µl de tampón Tipo V precalentado en el centro de la columna y se dejó incubar durante 1 minuto a temperatura ambiente, luego se centrifugó a 11000g durante 1minuto y se almacenó a -20°C.

2.4.2.4 Reacción en cadena de la polimerasa (PCR)

Para empezar, se calculó las diluciones de los 4 dNTPs, los dos cebadores, el cloruro de magnesio ($MgCl_2$), la solución tampón, el ADN polimerasa y el ADN molde. Para esto, se tomó en cuenta la concentración inicial de cada componente de la PCR, con el fin de obtener la misma cantidad de moles al final de la disolución. La ecuación es la siguiente:

$$C_{inicial} \times V_{inicial} = C_{final} \times V_{final}$$

Donde:

$C_{inicial}$: Concentración inicial en molares

$V_{inicial}$: Volumen inicial de la disolución

C_{final} : Concentración final en molares

V_{final} : Concentración final de la disolución

Tabla 8: Sistema de dilución para cada componente de la PCR.

Componente	$C_{inicial}$	C_{final}	$V_{inicial}$
Agua	-	-	8.75 μ l
Tampón	5X	1X	4 μ l
MgCl ₂	25Mm	1.5mM	1.2 μ l
dNTPs	2mM	0.2mM	2 μ l
Primer Forward	10mM	10 μ mol	1 μ l
Primer Reverse	10mM	10 μ mol	1 μ l
Taq Polimerasa	5U/ μ l	0.25U/ μ l	0.05 μ l
ADN	-	5ng/ μ l	2 μ l
V_{final}			20ul

Para la preparación de las reacciones, se irradia con luz ultravioleta por 5 minutos la cámara de flujo laminar, y los materiales que se utilizarán.

2.4.2.5 Estandarización de las reacciones de PCR.

Previamente a la amplificación del gen, se estandarizó la temperatura de fusión (T_m) de los cebadores en el termociclador *T Professional Gradient (Biometra)* (Tabla 9 y Figura 17).

- Separar el mix de PCR en un tubo de 1.5ml para siete reacciones con un volumen final de 140µl, considerando cinco reacciones para la gradiente de temperatura y una reacción para el blanco.
- Transferir 18µl del tubo mix a cada tubo para PCR de 0.2ml.
- Añadir 2µl de ADN a cada tubo para PCR de 0.2ml fuera de la cámara de bioseguridad excepto al tubo que vendría a ser el blanco.
- Programar el termociclador *T Professional Gradient (Biometra)*, con la gradiente de temperatura de 55°C, 57.5°C, 59.8°C, 65°C y 68°C.
- El ciclo de amplificación fue de 4 minutos a 94°C, seguido de 30 ciclos de 94°C durante 30 segundos, 55-68°C durante 30 segundos, y 72°C durante 30 segundos, finalizando a 72°C durante 5min.

Tabla 9: Sistema del mix de PCR para un volumen final de siete reacciones.

Componente	C _{inicial}	C _{final}	V _{inicial}	V ₇
Agua	-	-	8.75µl	61.25µl
Tampón	5X	1X	4µl	28µl
MgCl ₂	25Mm	1.5mM	1.2µl	8.4µl
dNTPs	2mM	0.2mM	2µl	14µl
Primer Forward	10mM	10pmol	1µl	7µl
Primer Reverse	10mM	10pmol	1µl	7µl
Taq Polimerasa	5U/µl	0.25U	0.05µl	0.35µl
ADN	17.08ng/µl	5ng/µl	2µl	-
V_{final}			20µl	140µl

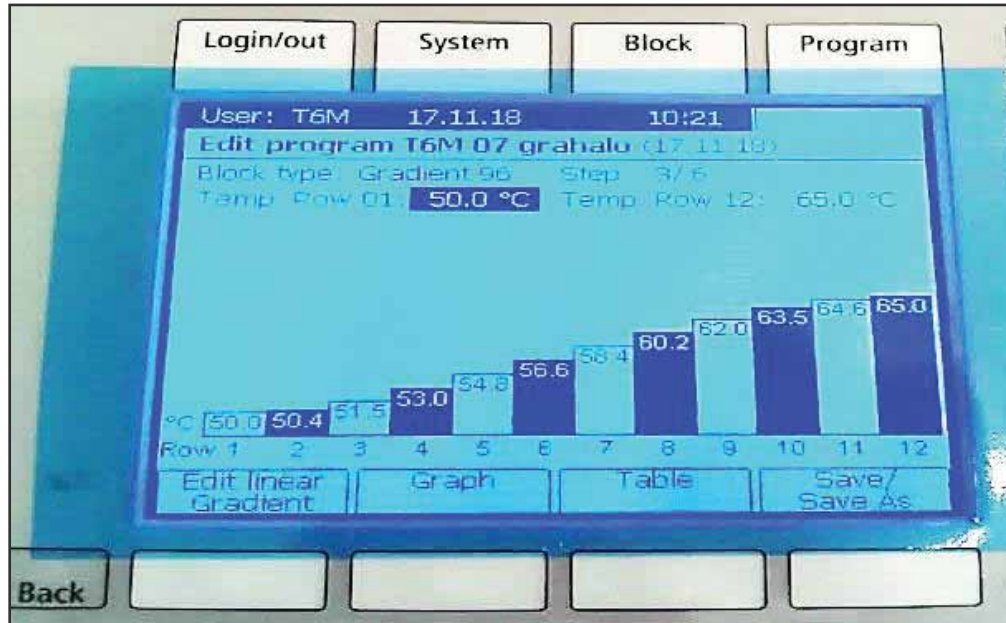


Figura 17: Programación de la PCR en gradiente con el termociclador *T Professional Gradient* (Biometra)

2.4.2.6 PCR para amplificar el gen Amidohidrolasa

La amplificación del fragmento de la proteasa amidohidrolasa fue realizada en el termociclador *T Professional Basic* (Biometra).

- Esterilizar la cámara de flujo laminar durante 5 minutos con luz ultravioleta.
- Preparar el mix de PCR en un tubo de 1.5ml para 5 reacciones donde una reacción es para el blanco y tres reacciones para la amplificación del gen con el fin de obtener una alta cantidad y concentración del gen aislado.
- Transferir 18µl del tubo mix a cada tubo para PCR de 0.2ml.
- Añadir 2µl de ADN a cada tubo para PCR de 0.2ml fuera de la cámara de bioseguridad excepto al tubo que vendría a ser el blanco.
- Programar el termociclador *T Professional Basic* (Biometra), los ciclos de amplificación fueron: 4 minutos a 94°C, seguida de 30 ciclos a 94°C durante 30 segundos, temperatura de fusión de los cebadores durante 30 segundos y 72 °C durante 30 segundos, finalizando a 72°C durante 5 minutos.

Tabla 10: Sistema del mix de PCR para un volumen final de 5 reacciones

Componente	C _{inicial}	C _{final}	V _{inicial}	V ₅
Agua	-	-	8.75µl	43.75
Tampón	5X	1X	4µl	20
MgCl ₂	25Mm	1.5mM	1.2µl	6
dNTPs	2mM	0.2mM	2µl	10
Primer Forward	10mM	10pmol	1µl	5
Primer Reverse	10mM	10pmol	1µl	5
Taq Polimerasa	5U/µl	0.25U	0.05µl	0.25
ADN	17.08ng/µl	5ng/µl	2µl	-
V _{final}			20µl	100µl

2.4.2.7 Electroforesis

Para visibilizar el tamaño del gen amidohidrolasa se realizó la técnica de electroforesis en gel de agarosa, para lo cual primero se alistó el equipo de electroforesis, las muestras de ADN producto de la PCR, el marcador de peso molecular (*standard*), el tampón de carga 6X, el tampón TAE 1X (Anexo 3), la solución *Sybr Safe* que es un componente que se intercala con las bases nitrogenadas emitiendo una fluorescencia bajo la luz ultravioleta. Luego se desarrolló el procedimiento de la siguiente manera:

- Se preparó el gel de agarosa al 1% con 1µl de *Sybr Safe* (Anexo 4)
- Luego se vertió la mezcla del gel con mucho cuidado en el molde para electroforesis con el peine bien puesto y se dejó solidificar.
- Se sumergió el gel en la cámara electroforética que contenía tampón TAE 1X y se retiró cuidadosamente el peine para que queden libres los pocillos para las muestras.
- Luego se preparó las muestras mezclando tanto las muestras de ADN como el marcador de peso molecular con 8µl de agua destilada y 2µl de tampón para un volumen final de 12µl.
- Se conectó los cables a la fuente de poder y se aplicó un voltaje de 90V durante una hora.
- Para la visualización del ADN se utilizó el lector de imágenes *ChemiDoc XRS+* (*Bio-Rad*) donde se colocó el gel en el transiluminador y se encendió la lámpara de luz ultravioleta (λ

≈ 300 nm), la imagen se observa mediante el programa *Imagen Lab TM* (Bio-Rad Laboratories, Inc.).

2.4.2.8 Purificación de la Amidohidrolasa

Para purificar el gen Amidohidrolasa se amplificó el gen y se realizó la electroforesis respectiva. Para empezar, se colocó el gel de agarosa en un transiluminador de luz ultravioleta con una longitud de onda a 365nm para visualizar las bandas con el ADN de la amplificación del gen y se recortó rápidamente el gel de agarosa con el ADN, los cuales fueron purificados con el *Kit Illustra GFX PCR DNA and Gel Band Purification* (GE Healthcare). El desarrollo del procedimiento se observa en la figura 18 y se detalla a continuación:

- Primero se sumergió los trozos de gel de agarosa con el ADN en 310µl de tampón de captura Tipo 3 y se incubó a 60°C hasta que el gel se disolvió.
- Luego se preparó la columna *GFX Micro Spin* en un tubo de recolección.
- Seguidamente se transfirió 610µl de la muestra de captura con los geles disueltos a la columna *GFX Micro Spin* y se dejó a temperatura ambiente durante 1 minuto.
- Se centrifugó la columna ensamblada a 16000g durante 30 segundos, descartando el eluido del tubo de recogida.
- Se adicionó 500µl de tampón de lavado Tipo 1 a la columna ensamblada y se centrifugó a 16000g durante 30 segundos, descartando el eluido del tubo de recolección.
- Se traslado la columna *GFX Micro Spin* a un nuevo tubo de 1.5ml y se adicionó 30µl de tampón de elución Tipo 4.
- Se centrifugó a 16000g durante 1 minuto y se colectó el eluido que contiene el ADN de la Amidohidrolasa purificado en un volumen de 30µl.

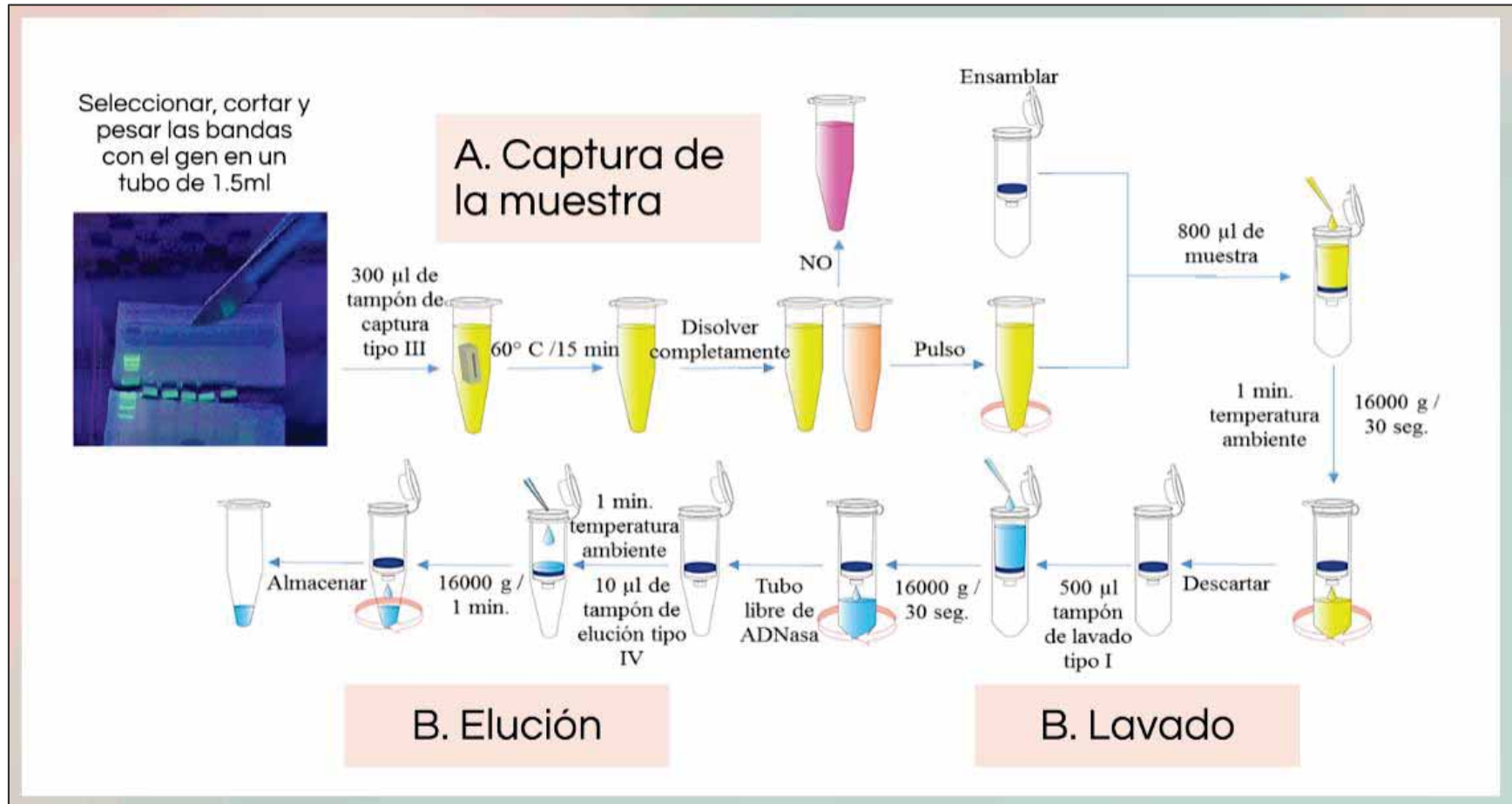


Figura 18: Proceso de la purificación del gen Amidohidrolasa.

2.4.2.9 Secuenciación de la Amidohidrolasa

La secuenciación de los productos purificados del gen Amidohidrolasa se realizó mediante la técnica de Sanger en el Servicio General de Apoyo a la Investigación (SEGAI) de la Universidad de La Laguna, Tenerife.

2.4.2.10 Análisis bioinformático del gen Amidohidrolasa

Se utilizó el programa Mega VII para revisar el cromatograma de la secuenciación y corroborar la similitud de nucleótidos del gen Amidohidrolasa entre los resultados de la secuenciación proporcionado por el SEGAI y la secuencia de nucleótidos proporcionado por la anotación RAST.

2.4.2.11 Clonación

El procedimiento de clonación del gen Amidohidrolasa se realizó en el sistema vector de clonación *pGEM-T Vector Sistem I*, el cual cuenta con un protocolo que se divide en dos pasos:

- Ligación: Inserto-Vector
- Transformación.

Ligación: Inserto-Vector

Se refiere a la ligación entre el inserto que viene a ser el gen Amidohidrolasa purificado y el vector *pGEM-T Vector Sistem I*. El desarrollo del procedimiento fue el siguiente:

- Se calculó la cantidad de inserto requerido para una proporción de 1:3 (vector:inserto) según la siguiente formula:

$$\frac{\eta g \text{ de vector} \times \text{tamaño del inserto pb}}{\text{tamaño del vector pb}} \times \text{proporcion} \frac{\text{inserto}}{\text{vector}} = \eta g \text{ de inserto}$$

- Se centrifugó brevemente el vector pGEM-T y el inserto del ADN control.
- Se vortexeó vigorosamente el tampón de ligación 2X.
- Se preparó la reacción de ligación como se muestra en la Tabla 11.
- Se incubó la reacción de ligación a 4°C durante toda la noche.

Tabla 11: Componentes de la reacción de ligación

Componentes de reacción	Reacción Estándar	Control Positivo	Control de Fondo
Tampón de ligación rápida al ADN T4 ligasa 2X	5 μ l	5 μ l	5 μ l
Vector pGEM-T (50 η g)	1 μ l	1 μ l	1 μ l
Productos de PCR	X μ l	-	-
Inserto de ADN control	-	2 μ l	-
ADN T4 ligasa 3U/ μ l	1 μ l	1 μ l	1 μ l
Agua libre de nucleasas para un volumen final	10 μ l	10 μ l	10 μ l

Transformación

Para esta fase de la clonación previamente se preparó el medio SOC (Anexo 7) y las placas de Luria Bertani (LB) que contengan ampicilina a una concentración de 100 μ g/ml, como se describe en el Anexo 6. El desarrollo del procedimiento para la transformación se observa en la figura 19 y se describe a continuación:

- Colocar las células competentes *E. coli* JM109 en un recipiente con hielo para su descongelación.
- Combinar 2 μ l de cada reacción de ligación con 50 μ l de células competentes JM109 mezclando suavemente en un tubo de 1.5ml y se incubó durante 20 minutos en hielo.
- Realizar la transformación mediante *shock* térmico sometiendo la mezcla de la reacción de ligación con las células a 42°C durante 1 minuto seguido de 2 minutos en hielo.
- Realizar la recuperación de las células añadiendo 950 μ l de medio SOC a temperatura ambiente en los tubos con las transformaciones de reacción de ligación y se dejó incubando a 37°C durante 90 minutos con agitación a 150 RPM.
- Realizar la siembra de las células transformadas, diseminando 100 μ l en placas con LB más Ampicilina, y se dejó a 37°C durante 24 horas.

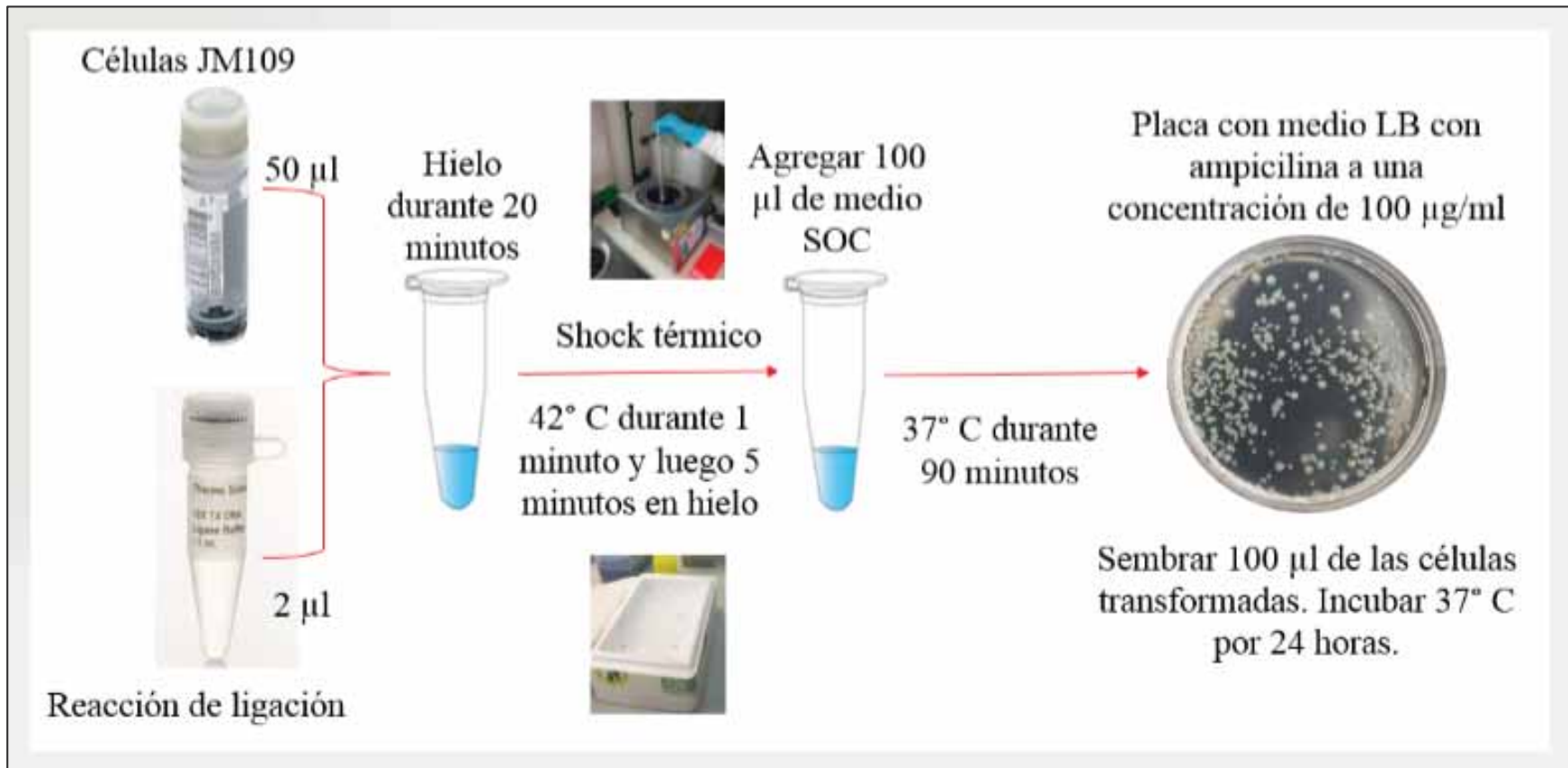


Figura 19: Proceso de la transformación en la clonación del gen amidohidrolasa.

2.4.2.12 *PCR de colonias clonadas*

Se seleccionarán 5 colonias al azar de la placa con medio LB más ampicilina, estas colonias fueron individualmente inoculadas en 700µl de caldo LB más ampicilina y se incubaron a 37°C durante 18 horas. El procedimiento para la PCR de cada colonia fue el siguiente:

- A partir del cultivo fresco de cada colonia se separó 100µl a nuevos tubos de 1.5ml respectivamente y se centrifugó a 6000 RPM durante 10 minutos, descartando el sobrenadante.
- Se añadió 50µl de agua ultrapura al pellet y se llevó a un baño de agua a 100°C durante 5 minutos luego se centrifugó a 13000 RPM por 5 minutos. Se conservó el sobrenadante y se descartó el pellet.
- Se utilizó 2µl del sobrenadante como ADN molde para cada reacción de PCR, utilizando los cebadores SP6 y T7 del vector pGEM-T para comprobar la transformación en las colonias. El sistema de reacción utilizado se describe en la Tabla 12.
- El ciclo de amplificación fue de 4 minutos de denaturación inicial a 94°C, seguido de 31 ciclos de 30 segundos a 94°C, 30 segundos de hibridación a 50°C y 30 segundos de extensión a 72°C; y una extensión final de 5min.a 72°C. La electroforesis de las muestras se realizó en gel de agarosa al 1% con un voltaje de 90V constante durante 1 hora.

Tabla 12: Sistema de reacción para la PCR de colonias clonadas

Componente	C _{inicial}	C _{final}	V _{inicial}	V ₇
Agua	-	-	8.75µl	61.25µl
Tampón	5X	1X	4µl	28µl
MgCl ₂	25Mm	1.5mM	1.2µl	8.4µl
dNTPs	2mM	0.2mM	2µl	14µl
Cebador T7	10mM	10pmol	1µl	7µl
Cebador SP6	10mM	10pmol	1µl	7µl
Taq Polimerasa	5U/µl	0.25U	0.05µl	0.35µl
ADN	-	-	2µl	-
V _{final}			20µl	140µl

2.4.2.13 Extracción del ADN plasmídico por lisis alcalina

Para extraer al plásmido pGEM-T que contiene al segmento del gen en las células JM109, se utilizó el *Kit PureYield™ Plasmid Miniprep System (Promega)*. Para empezar, se seleccionó una única colonia clonada que fue inoculado en 5ml de caldo LB (con ampicilina 100 µg/ml) y se cultivó a 37°C durante 16 horas con agitación vigorosa de 300 RPM. El desarrollo del procedimiento para la extracción del ADN plasmídico se observa en la figura 20 y se detalla a continuación:

- Obtener el pellet de 3ml del cultivo fresco mediante centrifugación a 16000g durante 30 segundos en un tubo de 1.5ml.
- Agregar 600µl de tampón TE al pellet y se resuspendió.
- Agregar 100µl de *Cell Lysis Buffer* y se mezcló invirtiendo el tubo seis veces hasta que la solución viró de color opaco a un azul claro, indicando que la lisis estaba completa.
- Añadir 350µl de *Neutralization Solution* fría (4°C) y se mezcló invirtiendo el tubo hasta que la solución viró a color amarillo.
- Centrifugar a 16000g durante 3 minutos y se alistó la minicolumna *Pure Yield* con su tubo de colección.
- Transferir 800µl del sobrenadante a la minicolumna y se centrifugó a 16000g durante 15 segundos.
- Descartar el eluido y añadir 200µl *Endotoxin Removal Wash* a la minicolumna y se centrifugó a 16000g durante 15 segundos.
- Descartar el eluido y añadir 400µl *Column Wash Solution* a la minicolumna y se centrifugó a 16000g durante 15 segundos.
- Transferir la minicolumna a un tubo de 1.5ml y agregar 30µl de *Elution Buffer* directamente al centro de la matriz de la minicolumna e incubar a temperatura ambiente durante un minuto.
- Centrifugar a 16000g durante 15 segundos para eluir el ADN plasmídico.
- Almacenar el ADN plasmídico a -20°C hasta su uso en posteriores procedimientos.

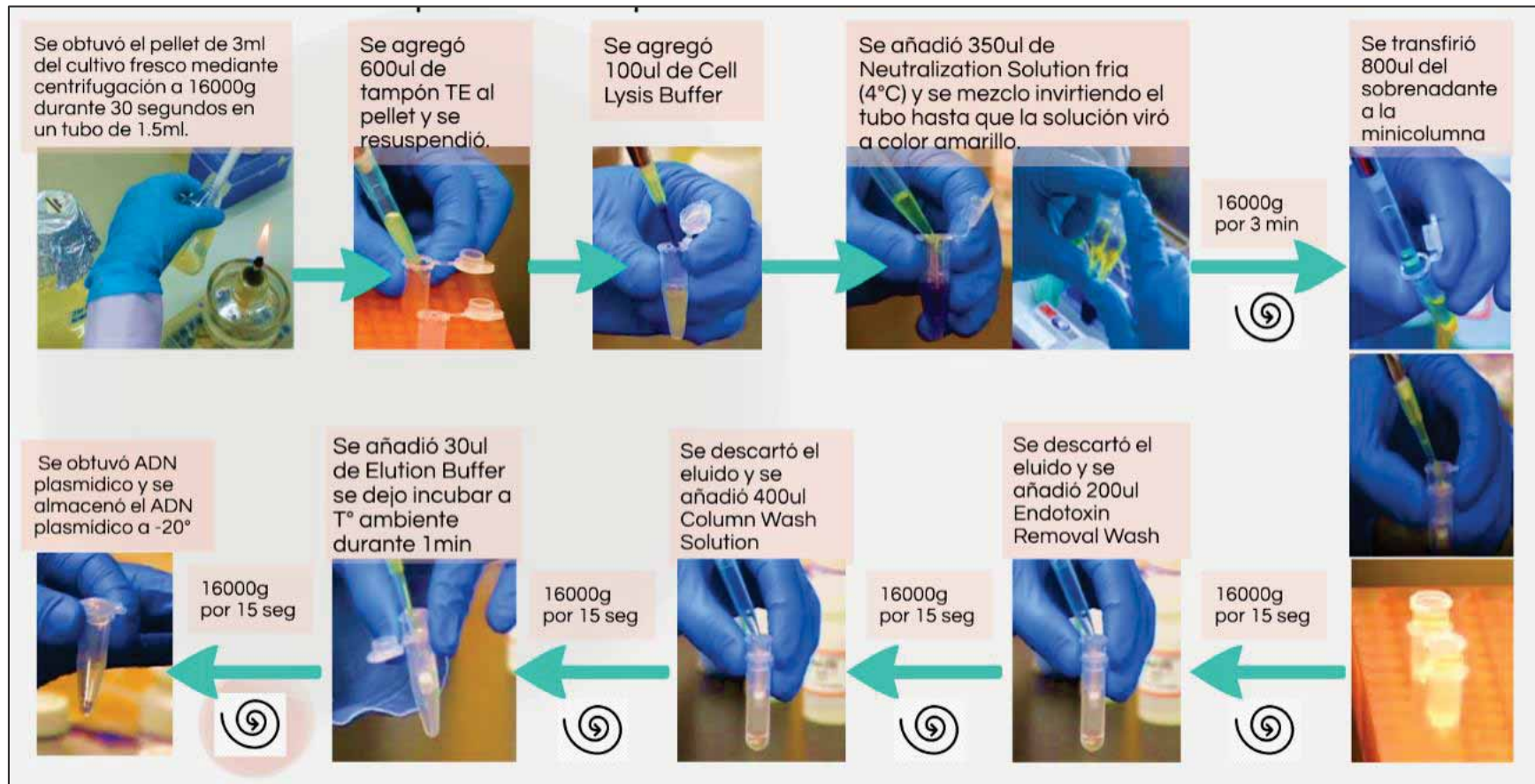


Figura 20: Procedimiento de la extracción del ADN plasmídico por lisis alcalina

2.4.2.14 *Secuenciación de la Amidohidrolasa clonado*

Para asegurar la clonación del gen amidohidrolasa y conocer en que orientación ingreso el gen en el plásmido pGEM-T, se envió un volumen de 10µl a una concentración de 5ng/µl del producto de la extracción del ADN plasmídico clonado al Servicio General de Apoyo a la Investigación de la Universidad de La Laguna (SEGAI) para su respectiva secuenciación. Los resultados de este paso fueron analizados con el programa MEGA 7 donde se alinearon las cadenas nucleotídicas sentido y antisentido del gen de la amidohidrolasa con los resultados de secuenciación del plásmido con el cebador T7 y el cebador SP6 del vector.

2.4.2.15 *Criopreservación de células clonadas*

Las cepas clonadas con el inserto amidohidrolasa fueron criopreservados para un uso posterior en tubos de 2ml de capacidad. El procedimiento fue el siguiente:

- Realizar un cultivo fresco de las cepas bacterianas clonadas en 5ml de caldo LB con ampicilina durante 16 horas a 37°C con agitación vigorosa a 300 RPM.
- Añadir 160 µl del cultivo fresco en cada tubo de 2ml.
- Añadir 40µl de glicerol estéril al 20%.
- Invertir los tubos para homogenizar la mezcla.
- Finalmente, almacenar los tubos a -80 °C.

CAPITULO III
RESULTADOS Y DISCUSIÓN


























3.1 RESULTADOS

3.1.1 ORGANIZACIÓN DE LAS SECUENCIAS CRUDAS DE ADN DE *C. salexigens*.

De la organización de las secuencias se reporta los informes elaborados por las herramientas bioinformáticas FastQC, Trimmomatic y Quast.

En la Tabla 13, se muestra el resultado del estado de las secuencias crudas producto de la secuenciación masiva, donde el programa FastQC, analizó los archivos *.fastq Forward* y *.fastq Reverse*, revelando el estado de las secuencias en once parámetros de calidad mediante tres diferentes iconos de color: el Rojo representa un estado dañado de la secuencia, el Amarillo representa un estado alarmante de la secuencia y el Verde representa un estado bueno de la secuencia.

Tabla 13: Parámetros de calidad de las secuencias Forward y Reverse con el programa FastQC

N°	Parámetros de calidad	Estado	
		Forward	Reverse
1	Estadísticas básicas		
2	Calidad de la secuencia por base		
3	Calidad de la secuencia por línea		
4	Puntuación de calidad por secuencia		
5	Contenido por base secuenciada		
6	Contenido de GC en la secuencia		
7	Contenido de N por base		
8	Distribución de la longitud de la secuencia		
9	Niveles de duplicación de secuencias		
10	Secuencias sobre-representadas		
11	Contenido de adaptadores		
Leyenda:	 Estado bueno de la secuencia		
	 Estado alarmante de la secuencia		
	 Estado dañado o peligroso de la secuencia		

✚ **Los parámetros de análisis con icono verde son cinco**, las cuales no requerirán de alguna corrección ya que su estado es bueno, estas son:

- 1) Las estadísticas básicas, donde el programa describió detalles generales de la secuenciación como el número total de secuencias, el porcentaje de GC, el nombre del fichero original de secuenciación, (Tabla 14).

Tabla 14: Estadísticas básicas del archivo forward

Características	Valores
Nombre del archivo <i>forward</i>	pe_1_S15_L001_R1_001.fastq
Encoding	Sanger / Illumina 1.9
Total de secuencias	594952 lecturas
Secuencias marcadas como de baja calidad	0
Longitud de secuencias	35-301 pb
%GC	62

- 2) La puntuación de calidad por secuencia, que permitió ver si algún conjunto pequeño de secuencias tiene baja calidad.
- 3) El contenido de N por base, que representa la indeterminación de bases a lo largo de las lecturas.
- 4) Los niveles de duplicación de secuencias, este parámetro está relacionado con la redundancia de las lecturas.
- 5) El contenido de adaptadores, que revela la presencia de una posible contaminación ya sea durante el aislamiento o durante el procesamiento del genoma en estudio.

✚ **Los parámetros de análisis con icono amarillo son dos**, las cuales no requerirán de alguna corrección ya que su estado es aceptable, estas son:

- 1) La distribución de la longitud de la secuencia, representa la distribución en la longitud de las secuencias, en este caso todas las secuencias no tienen el mismo tamaño, sin embargo, el porcentaje es aceptable.
- 2) Las secuencias sobre-representadas, son las secuencias de adaptadores o cebadores que suelen restar de la secuenciación masiva, no siempre indica presencia de agentes externos a la secuencia.

✚ **Los parámetros que presentan iconos rojos son cuatro**, los cuales fueron tratados con el software Trimmomatic con la finalidad de mejorar la calidad de las secuencias.

1) Calidad de la secuencia por base: Da a conocer la calidad de las secuencias mediante el Intervalo de calidad Phred que va de 0-40, donde cifras menores a 20 expresan una mala calidad, entre 20-28 es aceptable y superiores a 28 indica buena calidad. En la figura 21, se observa la calidad de la secuencia forward y reverse producto de la secuenciación con la plataforma Miseq-Illumina. La grafica de la secuencia *Forward* revela que a partir de 270pb el intervalo de calidad decae de 20 – 11 Phred, y la gráfica de la secuencia *Reverse* revela que a partir de 230pb el intervalo de calidad es de 10 - 6 Phred.

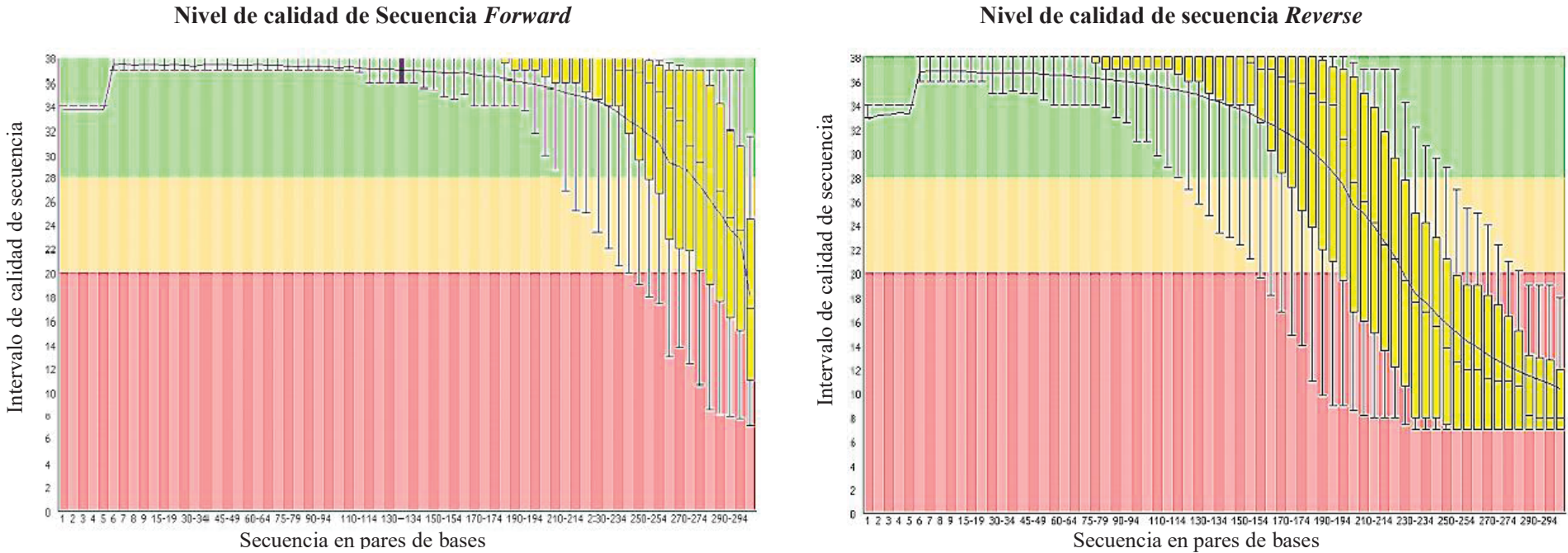


Figura 21: Calidad de las secuencias crudas de Forward y Reverse. El eje Y es el intervalo de calidad Phred, el eje X son los pb de la secuencia. Fuente: FastQC

2) Calidad de la secuencia por azulejo: Este parámetro se evidencia cuando la plataforma de secuenciación es Illumina, los gráficos de la figura 22 revelan pérdidas de calidad de las secuencias en algunas celdas donde se observa tonos cálidos como rojo, naranja y amarillo, lo ideal es que las celdas se muestren de un color azul uniforme.

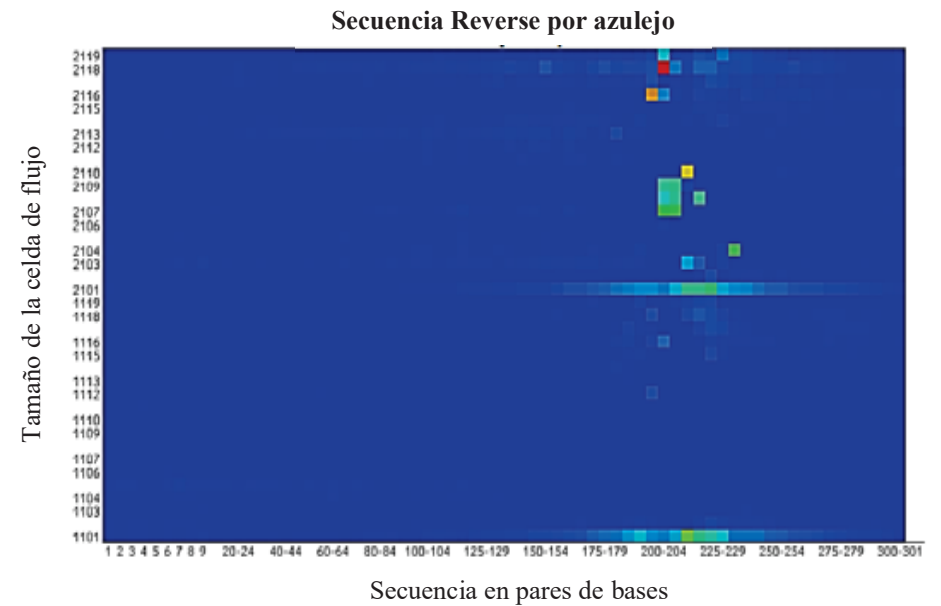
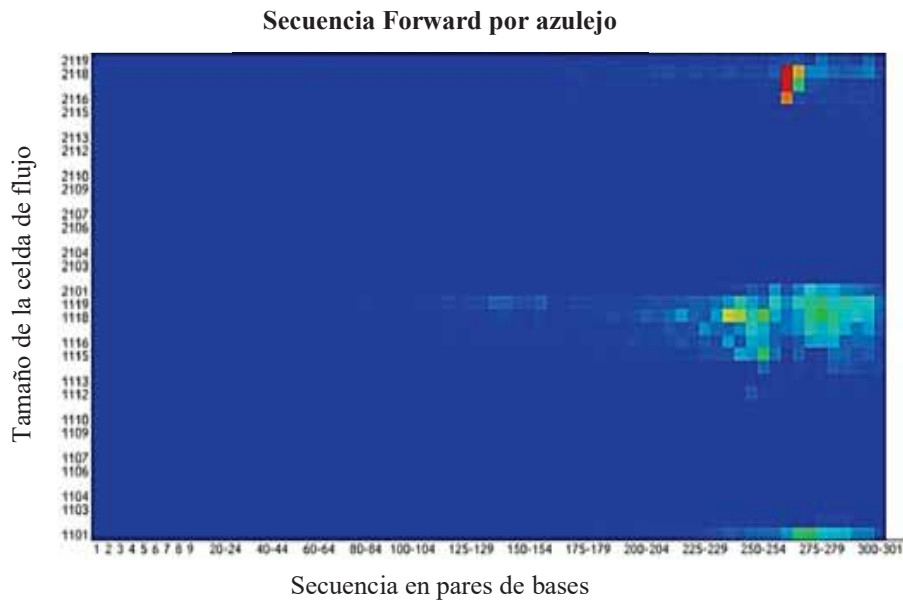
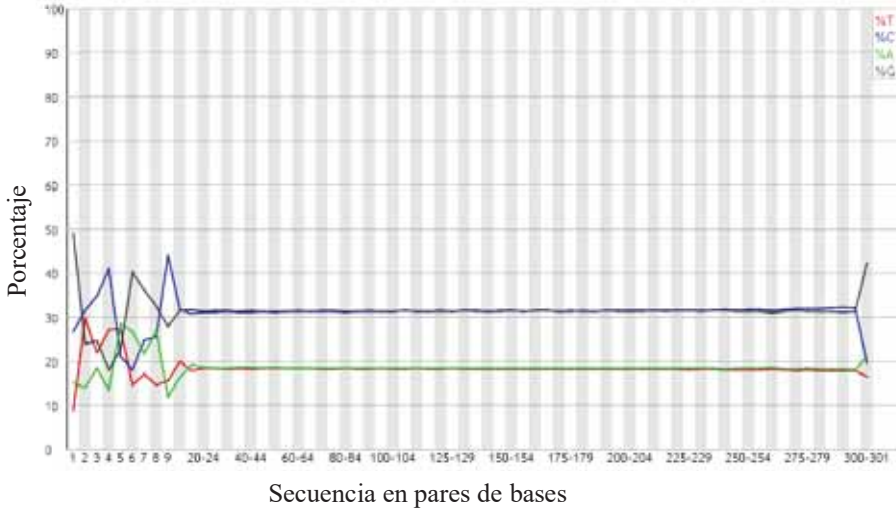


Figura 22: Calidad de las secuencias crudas por celda de flujo.

3) Contenido de bases en la secuencia: Representa el porcentaje de las 4 bases nitrogenadas a lo largo de las secuencias. En la figura 23 se observa que las líneas azul (C), negro (G), rojo (T) y verde (A) se presentan como un ruido al inicio de ambas secuencias, esto puede deberse a la presencia de adaptadores como parte del procedimiento de la secuenciación, recién a partir de 20pb las líneas de ambas secuencias se mantienen constantes hasta el final, en el caso de la secuencia forward se separan a partir de 300pb, de la secuencia reverse se observa separación de las líneas azul y negro desde los 200pb y las líneas verde y rojo desde las 275 pb.

Contenido de bases en la secuencia Forward



Contenido de bases en la secuencia Reverse

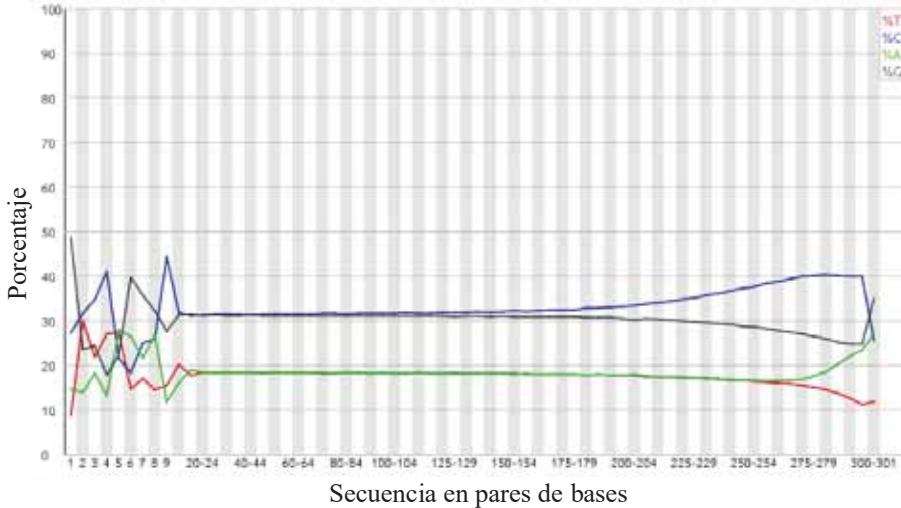


Figura 23: Contenido de bases nitrogenadas en las secuencias crudas Forward y Reverse.

4) Contenido de GC en la secuencia: Representa la proporción de guanina mas citocina en las secuencias. En la figura 24 se observa dos líneas en forma de campana, una roja y otra azul que muestran gran similitud para ambas secuencias. La línea roja representa al valor de las lecturas trabajadas y la línea azul representa al valor teórico, ambos valores coinciden en $G+C = 63\%$.

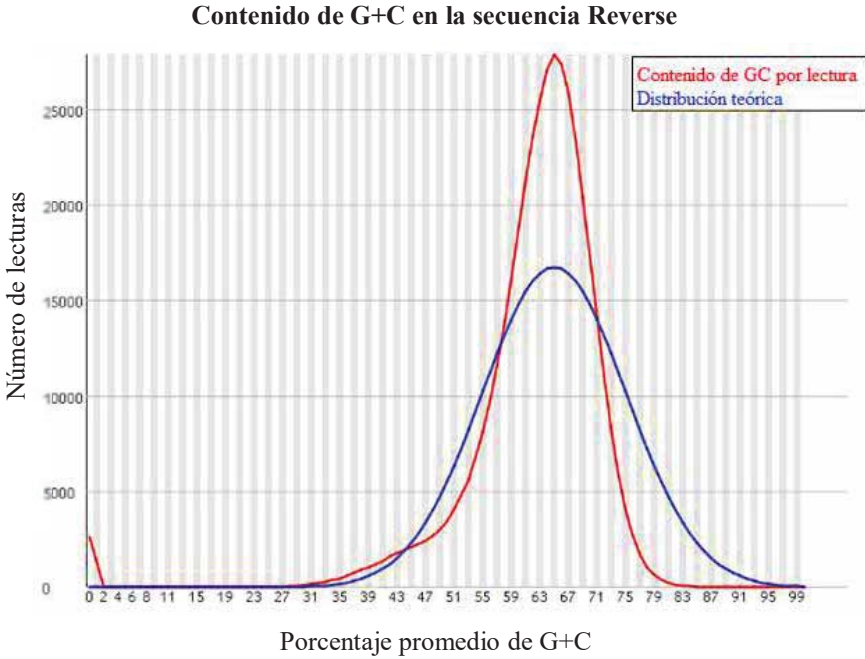
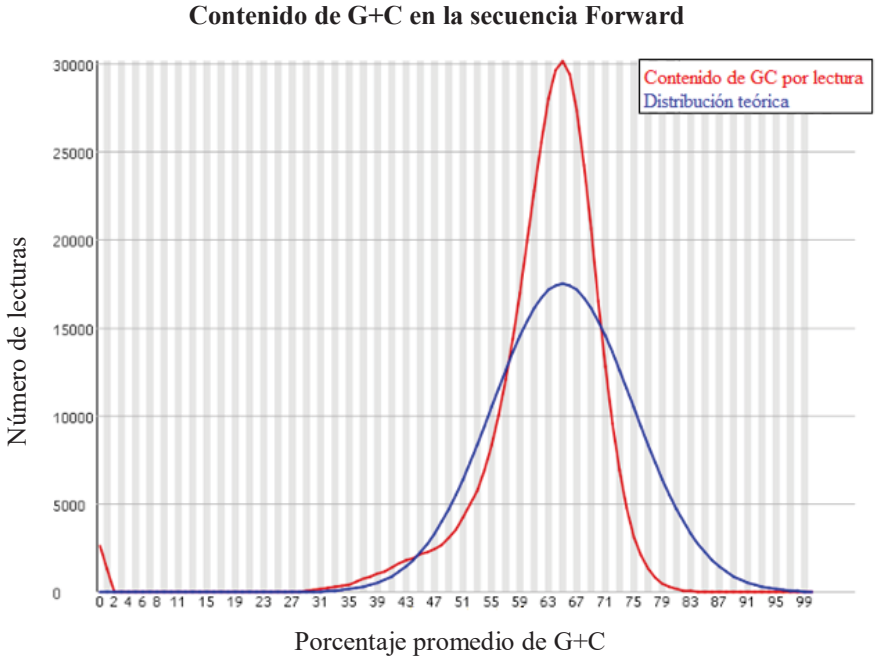















Figura 24: Contenido de G+C en la secuencias crudas de Forward y Reverse.

De acuerdo a estos resultados del análisis de las secuencias crudas analizadas por el programa FastQC, se realizaron las correcciones de las secuencias mediante el *Trimming* (recorte de secuencias), reportándose una mejora considerable en los parámetros de: calidad de la secuencia por línea, contenido por base secuenciada y contenido de GC por secuencia, (Tabla 15), cabe mencionar que al momento de realizar el *trimming* se tuvo severo cuidado con los cortes para evitar la pérdida de información. El contenido de Guanina y Citocina (GC) en la secuencia, se observó que el valor teórico es compatible con el valor de las lecturas trabajadas, no ha de variar con los recortes; la calidad de secuencia por línea mantiene el porcentaje de error, esto demuestra que el origen de contaminación es de origen específico, ya que este parámetro analiza los problemas que han podido suceder durante la reacción de la secuenciación (Por ejemplo, déficit en la manipulación del reactivo), el quinto parámetro revela que los cortes no fueron demasiados y que aún se mantienen ligeras superposiciones de bases, pero que ahora se encuentran en los extremos, el parámetro 8 indica que el tamaño de las secuencias aun difieren de tamaños, pero en proporción aceptable. Por otro lado, los demás parámetros presentan estados buenos de la secuencia.

Tabla 15: Parámetros de calidad de las secuencias Forward y Reverse después del Trimming

N°	Parámetros de calidad	Estado	
		Forward	Reverse
1	Estadísticas básicas		
2	Calidad de la secuencia por base		
3	Calidad de la secuencia por línea		
4	Puntuación de calidad por secuencia		
5	Contenido por base secuenciada		
6	Contenido de GC por secuencia		
7	Contenido de N por base		
8	Distribución de la longitud de la secuencia		
9	Niveles de duplicación de secuencias		
10	Secuencias sobre-representadas		
11	Contenido de adaptadores		

En la figura 25 se revela la calidad de secuencias después del *Trimming*, el primer gráfico con la calidad de secuencia *Forward* muestra una calidad constante de izquierda a derecha, de igual manera para el segundo gráfico con la calidad de secuencia *Reverse*, ambos presentan un índice de Phred por encima de 20.

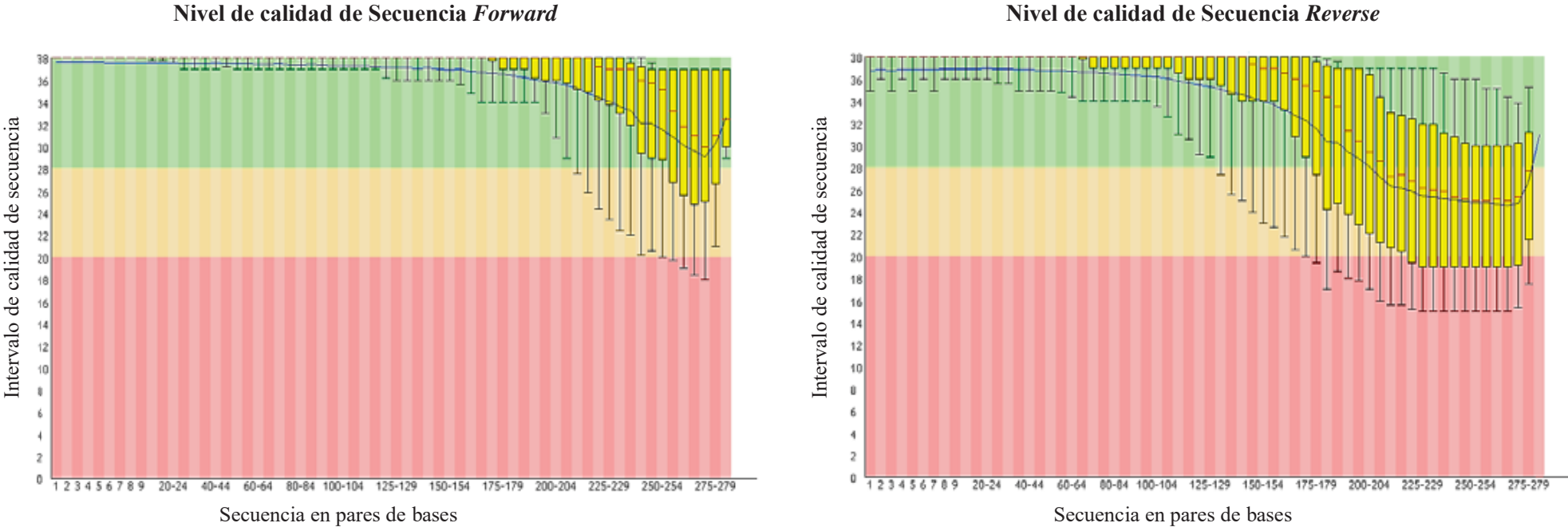


Figura 25: Calidad de las secuencias, después del *Trimming* según el índice de Phred.

Con estas correcciones en las secuencias crudas, se garantiza una mejor calidad de resultados para los siguientes procesos, como el ensamblaje y la posterior anotación. Para ello, el *software Trimmomatic* proporcionó un archivo de salida importante para cada secuencia:

- Secuencia Forward - pe-1_S15_L001_R1_PE.fastq
- Secuencia Reverse - pe-1_S15_L001_R2_PE.fastq

3.1.2 CALIDAD DEL ENSAMBLE DE LAS SECUENCIAS DEL GENOMA *C. salexigens* MP25462

Los resultados que se obtuvo del programa de calidad de ensamblajes QUAST se presentan en la tabla 16, estos datos dan a conocer las características del genoma ahora ensamblado, siendo la longitud del genoma de 3 712 216 pb, con un contenido de guanina mas citocina (G + C) del 64 %, asimismo comprende 42 *contigs* con los siguientes parámetros: el contig mas grande mide 491 015 pb, el **N50** es la longitud del *contig* que se encuentra en la mitad del total genómico y mide 447 080 pb, el **N75** es la longitud del *contig* que se encuentra en el 75% del tamaño total del genoma y mide 169 102 pb, el **L50** es 4 y se refiere al número de *contigs* cuya longitud sumada es el N50, **L75** es 8 y se refiere al número de *contigs* cuya longitud sumada es el N75, el número de **Ns** (enes) es cero, se refiere a la sustitución de un nucleótido indeterminado por una llamada N, ya que carece de confianza de asignación de base.

La evaluación del ensamblado genómico revela que el programa SPAdes es destacado para ensamblajes genómicos ya que presenta una baja cantidad de contigs y además no contiene Ns que es un parámetro exigente de confianza.

Tabla 16: Parámetros de calidad del genoma ensamblado de *C. salexigens* MP25462

Parámetro de calidad	Pares de bases (pb)
Largo total del genoma	3 712 216
Número de contigs	42
Contig más grande	491 015
N50	447 018
GC (%)	63,99
Desajustes	
# N's	0
# N por 100 kpb	0

3.1.3 ANOTACIÓN DEL GENOMA *C. salexigens* MP25462 CON EL SERVIDOR RAST.

El análisis funcional del genoma completo de la cepa MP25461 se obtuvo a las 48 horas de haber subido las secuencias organizadas en archivo *fasta* al servidor RAST, el cual predice un contenido de G + C del 64%, 3400 genes codificantes incluyendo 69 secuencias que codifican ARN.

Los 3400 genes codificantes fueron organizados en 486 subsistemas. Los subsistemas son agrupaciones de genes que realiza el servidor RAST en función de sus propiedades biológicas, por ejemplo, el número de genes que codifica a proteínas de la pared celular, genes de virulencia, genes implicados en la respiración, la respuesta al estrés, la motilidad, división celular, el metabolismo de diferentes proteínas, etc., (Tabla 17) (Figura 26).

A cerca de la cobertura de los genes, el 60% esta dentro de los subsistemas y el 40% no está dentro de los subsistemas. Los genes que estan dentro de los subsistemas son 2025, de los cuales 1909 son genes no hipotéticos y 116 genes hipotéticos. Los genes que no estan presentes en los subsistemas son 1375, de ellos 779 son genes no hipotéticos y 596 son genes hipotéticos.

Se encontró una alta proporción de genes que codifican para el metabolismo de proteínas, 263 genes implicados, de los cuales 25 genes se asocian a proteínas plegables, 4 genes a selenoproteínas, 173 genes implicados en la biosíntesis de proteínas, 24 genes en el procesamiento y modificación de proteínas y 37 genes en la degradación de proteínas. Además, se encontró que una gran cantidad de genes están involucrados en el mecanismo de transporte de membranas y el metabolismo de carbohidratos, lípidos y ácidos grasos. Mientras que los genes involucrados en el metabolismo de compuestos aromáticos, azufre, fosforo, nitrógeno, hierro y potasio, fueron menos abundantes.

Tabla 17: Número de genes agrupados en 27 funciones de los subsistemas.

Nº	Funciones de los subsistemas	Número de genes
01	Cofactores, vitaminas, grupos prostéticos, pigmentos	268
02	Pared celular y cápsula	133
03	Virulencia, enfermedad y defensa	62
04	Metabolismo del potasio	23
05	Fotosíntesis	0
06	Misceláneo	79
07	Fagos, profagos, elementos transponibles, plásmidos	20
08	Transporte de membrana	203
09	Adquisición de hierro y metabolismo	17
10	Metabolismo de ARN	199
11	Nucleósidos y nucleótidos	107
12	Metabolismo de proteínas	263
13	División celular y ciclo celular	32
14	Motilidad y quimiotaxis	83
15	Regulación y señalización celular	93
16	Metabolismo secundario	4
17	Metabolismo del ADN	95
18	Ácidos grasos, lípidos e isoprenoides	141
19	Metabolismo de nitrógeno	26
20	Dormancia y esporulación	2
21	Respiración	131
22	Respuesta al estrés	163
23	Metabolismo de los compuestos aromáticos	39
24	Aminoácidos y derivados	364
25	Metabolismo de azufre	33
26	Metabolismo del fósforo	42
27	Carbohidratos	395

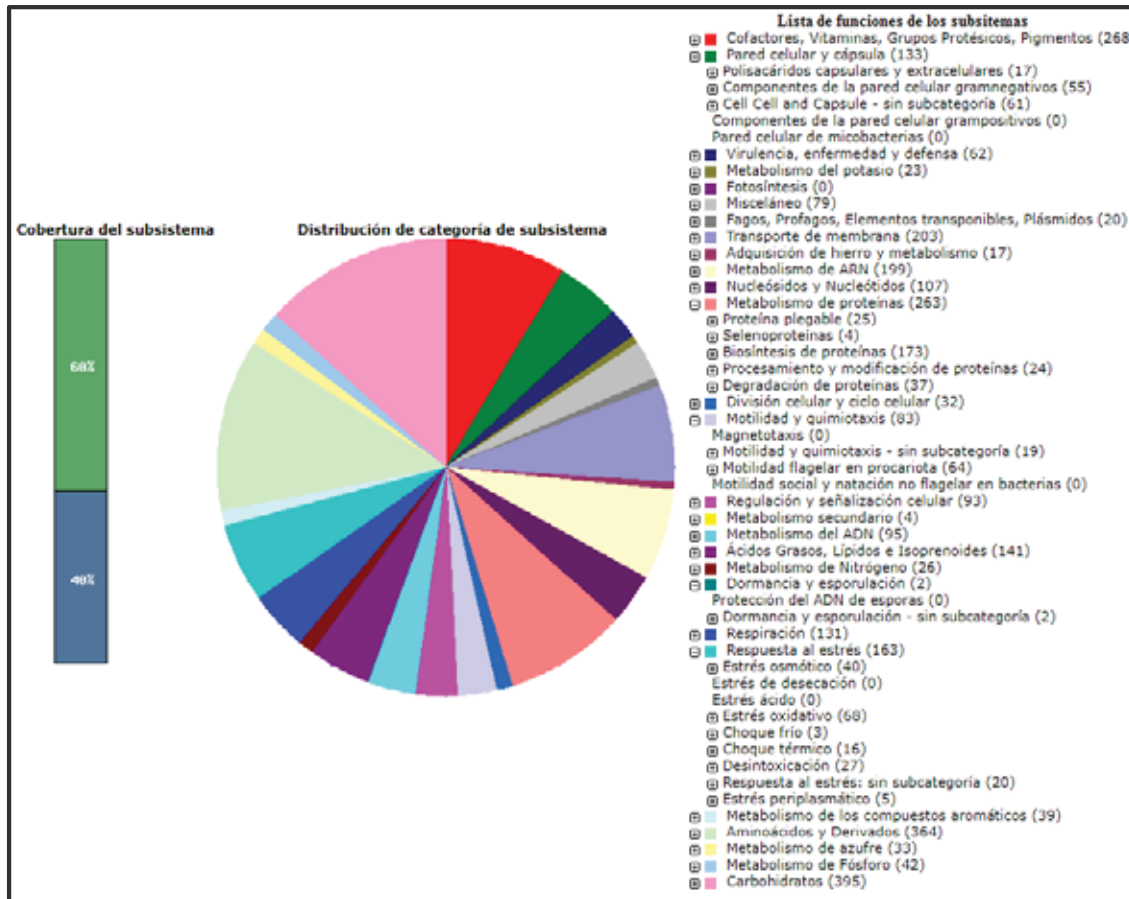


Figura 26: Anotación del genoma *C. salexigens* MP25462 con el servidor RAST.

3.1.4 DESCRIPCIÓN DE PROTEASAS DEL GENOMA DE *C. salexigens* MP25462.

El servidor RAST predijo los genes utilizando diversas tuberías de información en su base de datos, que se basa en alinear mediante algoritmos el genoma ensamblado con otros genomas ya descritos, considera la longitud de los genes, sitios conservados, secuencias palindrómicas de algunas secuencias, hasta obtener una imagen más resuelta de la anotación, es así que predijo 263 genes que codifican para el metabolismo de proteínas de *C. saligens* MP25462 y mediante la exploración del archivo de descarga disponible Excel .xls, se logró identificar manualmente y usando la clasificación enzimática, los 36 genes que codifican varios tipos de proteasas en el genoma de *C. salexigens*, según la clasificación de la *Enzyme Comision* EC, se subdividió en endopeptidasa y exopeptidasas (Anexo 1)

3.1.4.1 Endopeptidasas

Dentro de la clasificación del grupo de las endopeptidasas, se obtuvo 9 proteasas de tipo serina (Figura 27), 2 proteasas de tipo aspartil (Figura 28), 9 proteasas de tipo metalo (Figura 29) y 1 proteasa de tipo treonina (Figura 30).

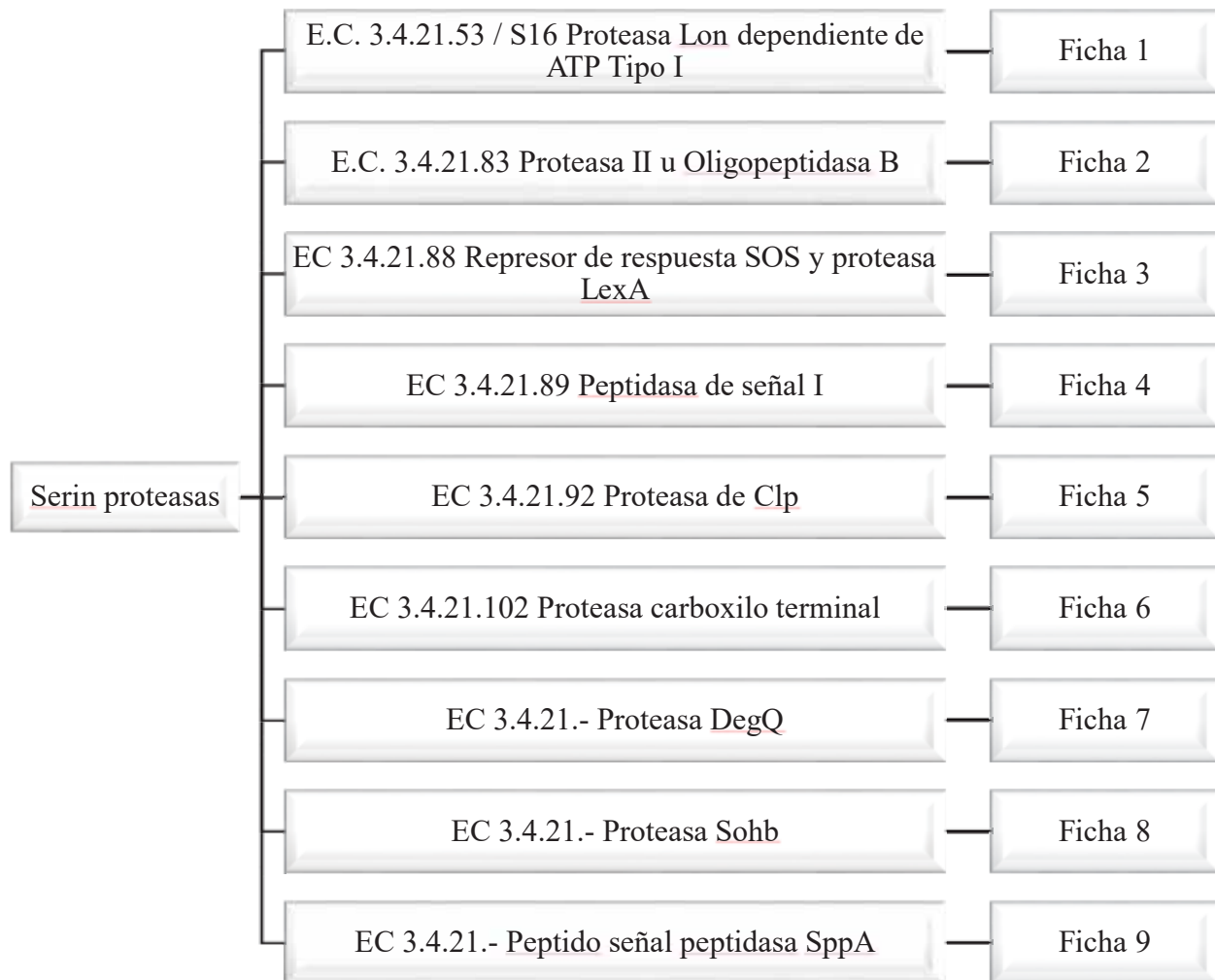


Figura 28: Endopeptidasas de tipo serina encontradas en el genoma de *C. salexigens* MP25462

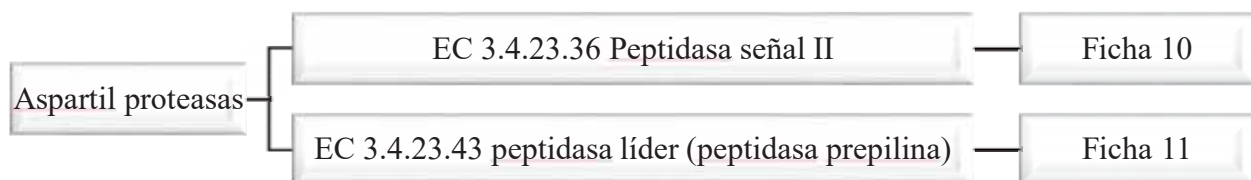


Figura 27: Endopeptidasas de tipo aspartil encontradas en el genoma de *C. salexigens* MP25462

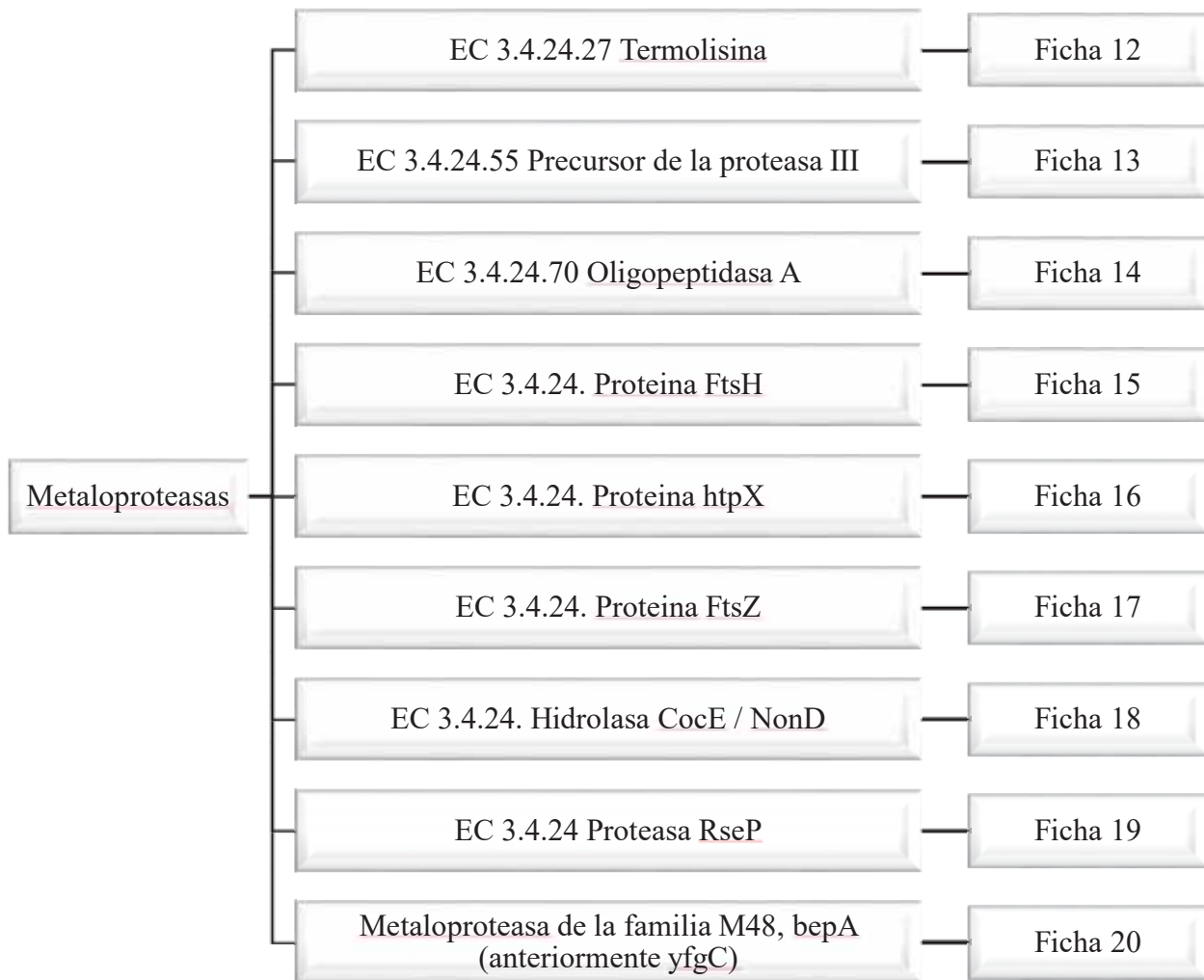


Figura 29: Endoproteasas de tipo metalo encontradas en el genoma de *C. salexigens* MP25462



Figura 30: Endoproteasas de tipo treonina encontradas en el genoma de *C. salexigens* MP25462

3.1.4.2 Exopeptidasas

Dentro de la clasificación del grupo de las exopeptidasas, se obtuvo 7 proteasas de tipo aminopeptidasa (Figura 31), 7 proteasas de tipo carboxipeptidasa (Figura 32) y 1 proteasa de tipo omegapeptidasa (Figura 33), dentro del genoma de *C. salexigens*. Diversas amino y carboxipeptidasas se encuentran presentes en este genoma, responsables de la escisión de cualquier aminoácido N- terminal o C-terminal adyacente a un residuo de Prolina, Metionina, Aspartil, Histidina, Serina. En su mayoría están descritas en mamíferos, plantas, hongos, parásitos, bacterias, etc.

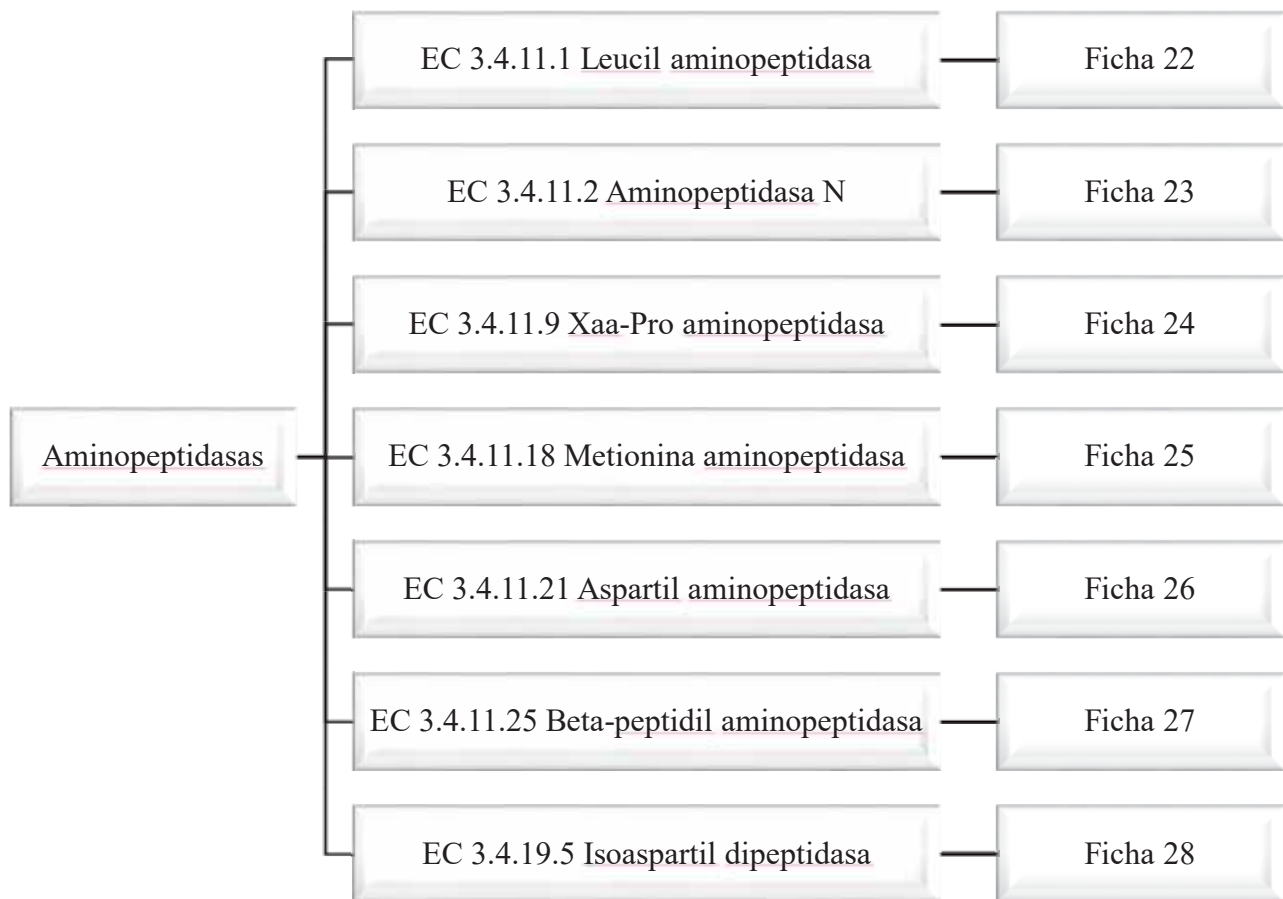


Figura 31: Exoproteasas de tipo amino encontradas en el genoma de *C. salexigens* MP25462

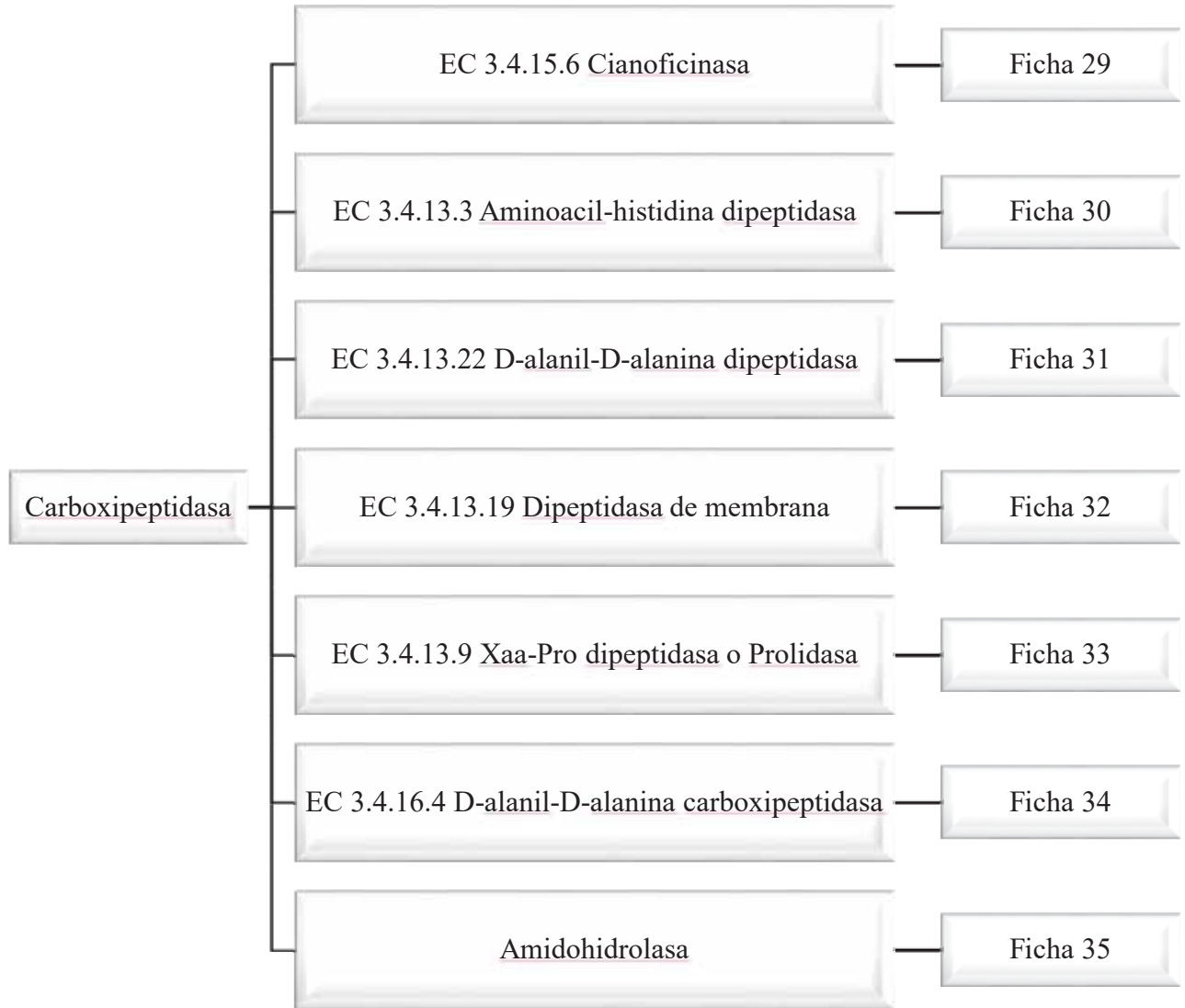


Figura 32: Exopeptidasas de tipo carboxi encontradas en el genoma de *C. salegigens* MP25462



Figura 33: Exopeptidasas de tipo omega encontradas en el genoma de *C. salegigens* MP25462

3.1.5 DESCRIPCIÓN DE LA PROTEASA AMIDOHIDROLASA

A partir de la descripción de las proteasas obtenidas se aisló por técnicas moleculares la proteasa Amidohidrolasa de *C. salexigens* MP25462, esta proteasa se ubica en la hebra sentido del ADN, inicia en el nucleótido 134947 y termina en el nucleótido 136098 teniendo una longitud 1151pb (Figura 36) que se traducen en 383 aminoácidos (Figura 35). Esta molécula tiene un peso de 41.15 Kda. La cobertura de la secuencia del gen amidohidrolasa en el genoma resultó de 19.6x. El gen amidohidrolasa está flanqueado por el extremo 5' por la proteína que contiene el dominio DUF2489 (naranja) y por el extremo 3' por la proteína OmpA/MotB (verde). La proteína que contiene el dominio DUF2489 tiene un tamaño de 462pb que se traduce en 134 aminoácidos, el espacio intergenómico entre esta proteína y la proteína amidohidrolasa comprende 23pb. La proteína OmpA/MotB tiene un tamaño de 885pb que se traduce en 295 aminoácidos, el espacio intergenómico entre esta proteína y la proteína amidohidrolasa comprende 264pb (Figura 34)

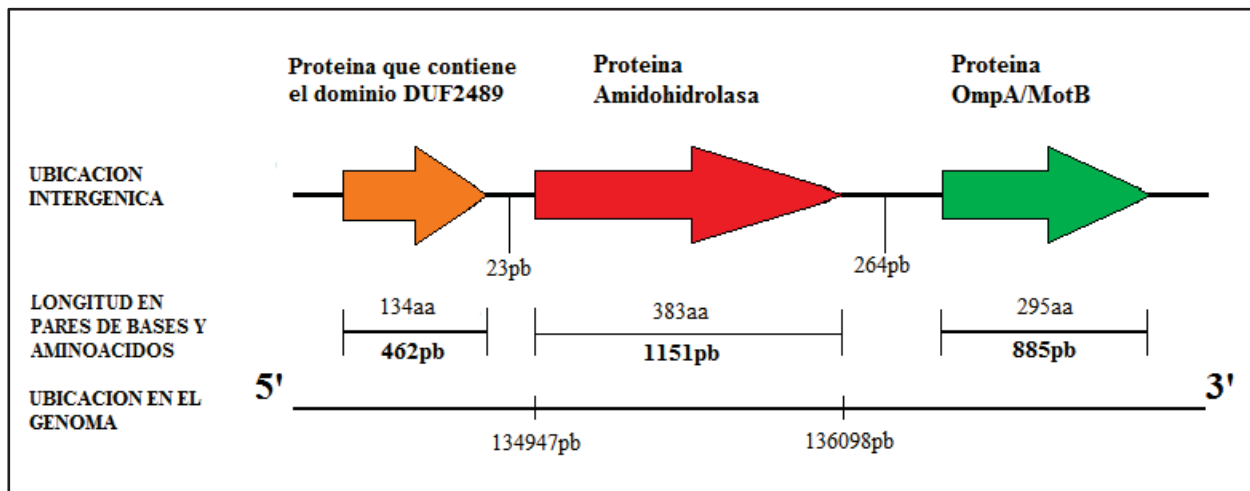


Figura 34: Ubicación en el genoma de la proteína Amidohidrolasa de *C. salexigens* MP25462

```
> fig | 6666666.597456.peg.3228 Peptidasa M20D, amidohidrolasa
[Chromohalobacter salexigens MP25462]
MNEAFLSLTPAQIRHRLHAAPELSGQETETARRVAGWLEQAGATRIVTGLGGHGVAGVFAGT
LPGPRVMFRAELDGLPIDETAELIAWKSTCEHVAHSCGHDGHMAVLLGLAQLAEQPPTHGET
VLLFQPSEETGRGAREVVESSRFEEIAPDHVYALHNLPGKPLGEVSVREGAMTCASRGLIIR
LNGLCAHAAHPEQGRNPALAMCRILQGLKRLPQQLPDDQGLVMVTLIHARLGEVAFGTSPGQ
AEVMATLRTEDDAMMRALAEASVALARAEAEEDDLGIEIDWCDVFQATRNPAAVASIREAA
RSLDMPLTALPEAHRWSEDFGWLAAQGGALLTLGAGERCLPLHNPYDFPDALLDTGIALF
ERLVRLLHQSDG
```

Figura 35: Secuencia de aminoácidos en formato FASTA de la Amidohidrolasa de *C. salexigens* MP25462.

```
>fig|6666666.597456.peg.3228 Peptidasa M20D, amidohidrolasa
[Chromohalobacter salexigens MP25462]
atgaacgaggcatttttgtcgtgacgcccggcccagattcgtcatcggctgcatgccgcgccg
gagctgtcagggcaggaaccgaaacggcccgggggtggcgggctggctcgagcagggcggg
gacacgcatcgtcaccggtctgggtgggtcatggcgtggccgggggtgttcgcgggcacctg
ccggggccgcgcgtcatgttccgtgccgaactcgatggccttcccacgcagacgagaccgcgaa
atcgcctggaaatcgacctgtgagcacgtggcgcacagtgttgccatgacgggcacatggcg
gtattgctcggattggcgcagcgcctggcggagcagccgcgacgcatggcgaaacggtgctg
ctgtttcagccttcggaggaaaccggggcgtggcgcgccgcgaggtgggtggaatcatcgcgttc
gaggaaatcgctcccgatcatgtatatgcaactgcacaatctgcccggcaagccgtgggagag
gtgtcgggtgagagagggcgccatgacctgcgcctcgcgagggctgatcatccggctcaacggc
ttgtgcgcgcatgcggcacaccggagcagggccgcaaccggcgctggcgatgtgccggatt
ctgcaggggtctcaagcggttgccccagcaactgcccgacgatcaagggctgggtcatggtcacg
ctgatccacgcgcgattggggcagggtcgccttcggtagctgcggggggcaggccgaggtcatg
gccacgttgcgcacccgaggatgacgccatgatgcgcgcgctggcggaggcctcggtgccgctt
gcgcgggccgaagcggaggaggacgacctggggatcgagatcgactgggtgcatgtctttcag
gccacgcgcaatcatcccgcgcgcttgccctcgattcgcgaaagccgcccggctcgtggacatg
cccttgacggcgctgcccgaggcccaccgctggctcggaagacttcggctggctggccgcgcag
gggcaggggtgcgctgttgacctggggcgcggcgagcgggtgcctgccgctgcacaaccccggc
tacgactttccggatgccctgctggacacgggaatcgcgttgttcgagcgcctggtgaggctc
catcagtcggatggctga
```

Figura 36: Secuencia de nucleótidos en formato FASTA del gen Amidohidrolasa de *C. salexigens* MP25462.

Según la base de datos de Merops la proteasa Amidohidrolasa pertenece al grupo de las peptidasas M20D, miembros de la familia M20, estas enzimas requieren dos iones metal para su actividad catalítica, además se describe a la cisteína como residuo en el sitio activo, lo cual las diferencia de otras peptidasas M20D (Bitto *et al.*, 2009).

Las amidohidrolasas tienen fama de tener actividades promiscuas debido a cierta afinidad por varias especificidades de sustrato convirtiéndola en una enzima generalista multifuncional., altamente evolutivo y con una alta diversidad conformacional (Barba *et al.*, 2013), para ello, se demostró que la pérdida del ion metálico y el reordenamiento del sitio activo permite a la enzima cambiar drásticamente su mecanismo catalítico por lo cual puede degradar nuevos compuestos por ejemplo molinatos y las diuron hidrolasas, de esta manera confirmaron que las alteraciones del metal nativo de esta metaloenzima impulsa actividad promiscua (Sugrue *et al.*, 2015). Debido a ello, estas enzimas están emergiendo como un activo popular y valioso en el campo de la enzimología para estudiar las relaciones estructura-función y la evolución (Liu & Huo, 2014).

3.1.5.1 Modelamiento estructural de la proteasa Amidohidrolasa MP25462

Mediante el *software* I-Tasser se predijo el modelado estructural de la proteína en estudio, el programa basó su diseño según la base de datos que tiene de todas las plantillas de cristalografías que existen en el PDB. El resultado obtenido fue un modelado estructural con alta similaridad con la proteína de código PDB 4ewtA que corresponde a la estructura cristalina de una Aminohidrolasa de *Staphylococcus aureus* COL traducida por el SACOL0085, con un índice de cobertura 0.984 y TM-score de 0.980 (valores igual a 0.96 ± 0.05 lo cual indica una topología correcta), donde la cristalografía del gen amidohidrolasa reveló que dicha molécula tiene dos iones Mn^{+2} unidos en el sitio activo. Además, posee una característica distintiva que es la disposición cuaternaria homotetramérica en forma de mariposa con extensas interacciones entre subunidades. La comparación de estas cuatro unidades monoméricas del tetramero mostró variaciones en la orientación entre dominios, lo cual podría ayudar a los movimientos de dominio y podría desempeñar un papel en la regulación de la actividad catalítica (Girish *et al.*, 2013).

Las 10 proteínas estructuralmente cercanas al modelado de la proteasa Amidohidrolasa MP25462 se muestran en la Tabla 18. Entre ellos se encuentran proteínas cristalizadas de *Staphylococcus aureus subsp. aureus* COL (Girish *et al.*, 2013), *Arabidopsis thaliana* (Levin *et al.*, 2007), *Bacillus subtilis* (Minasov *et al.*, 2005), *Staphylococcus aureus subsp. aureus* Mu50 (Joint Center for Structural Genomics, 2011), *Pseudomonas sp. RS-16* (Rowell *et al.*, 1997), *Corynebacterium striatum* (Natsch & Emter, 2020), *Shigella flexneri 2a str. 301* (Osipiuk *et al.*, 2021), *Bacillus cereus* ATCC 10987 (Joint Center for Structural Genomics, 2009), *Escherichia coli* (Badger *et al.*, 2005) y *Legionella pneumophila subsp. pneumophila str. Filadelfia 1* (Joint Center for Structural Genomics, 2010).

Tabla 18: Proteínas estructuralmente cercanas al modelado de la Amidohidrolasa MP25462

	PDB ENCONTRADO	ORGANISMO	COV	Puntuación TM
1	4ewtA	<i>Staphylococcus aureus subsp. aureus COL</i>	0.984	0.980
2	2q43A	<i>Arabidopsis thaliana</i>	0.937	0.867
3	1ysjA	<i>Bacillus subtilis</i>	0.911	0.861
4	3rzaA	<i>Staphylococcus aureus subsp. aureus Mu50</i>	0.911	0.809
5	1cg2A	<i>Pseudomonas sp. RS-16</i>	0.935	0.805
6	6slfA	<i>Corynebacterium striatum</i>	0.977	0.800
7	7lgpA	<i>Shigella flexneri</i>	0.927	0.799
8	3gb0A	<i>Bacillus cereus ATCC 10987</i>	0.901	0.794
9	1vixA	<i>Escherichia coli</i>	0.909	0.790
10	3pfeA	<i>Legionella pneumophila subsp. pneumophila</i>	0.953	0.787

* Puntuación TM: clasifica a las proteínas alineando la estructura de consulta y las estructuras conocidas en la biblioteca de PDB. Puntuación TM igual a 0.96 ± 0.05 mide la similitud estructural.

* Cov: representa la cobertura de la alineación por Puntuación TM y es igual al número de residuos alineados estructuralmente dividido por la longitud de la proteína de consulta.

Las características de la estructura proteica de Amidohidrolasa MP25462 se describen a continuación, contiene trece cadenas con disposición β -plegada intercaladas entre diez α -hélice. La secuencia comienza a partir del extremo N-terminal con disposición α -hélice desde el Aac. 3-19 ; 28-40, avanzando hasta la disposición β -plegada con el Aac 45-47 ; 54-60 ; 67 -71, continuando hasta la α -hélice desde el 101 – 116 seguido de una β -plegada 123-128 continua α -hélice 137- -143 avanzando hasta la β -plegada 154 -158 ; 167 – 171 ; 175 – 177 seguido de α -hélice 182 – 190 ; 204 – 219 seguida de β -plegada 228 – 236 ; 248 – 256 continua la α -hélice desde

260 – 281, luego sigue la β -plegada 284 – 291 ; 297 – 298 continua la α -helice 300 – 314 luego la β -plegada 318 -319, seguido de una α -helice 327 – 336, y ya llegando al extremo C- terminal esta la β -plegada 339 – 344 y por ultimo la α -helice desde 362 – 381. En la figura 37 se aprecia el modelamiento final de la estructura otorgado para la proteasa Amidohidrolasa, las estructuras secundarias se visualizan en dos colores, esta la α -helice con el color rosado y la β -plegada de color naranja. El modelado de la proteasa Amidohidrolasa obtuvo puntajes muy altos en cuanto la confianza, similitud y precisión en el alineamiento estructural, demostrando que la proteína existe.



Figura 37: Modelamiento de la estructura de la proteína Amidohidrolasa MP25462 con el programa I-TASSER

Dentro de la interfaz de I-TASSER se obtuvo la predicción de funcionalidad de la proteasa Amidohidrolasa MP25462 comparándola estructuralmente con las proteínas cristalizadas contenidas en PDB. La peptidasa 4ewtD de *Staphylococcus aureus subsp. aureus COL* resultó ser muy similar en cuanto a la conformación de los sitios de unión al ligando, ubicando a los residuos del sitio de unión al ligando en C98 (Cisteína), E133 (ácido glutámico), H354 (Histidina) y E132 (ácido glutámico) como sitio activo, además se revela un ion de Mn^{+2} , centro metálico mononuclear, donde el sustrato se activará mediante transferencia de protones desde el sitio activo, y el agua se activará mediante ligadura de metales y catálisis de bases generales. Entonces se define la topología del sitio activo y la composición de ligandos de la proteasa Amidohidrolasa *C. salexigens* MP25462 similar a la de *Staphylococcus aureus subsp. aureus COL*.

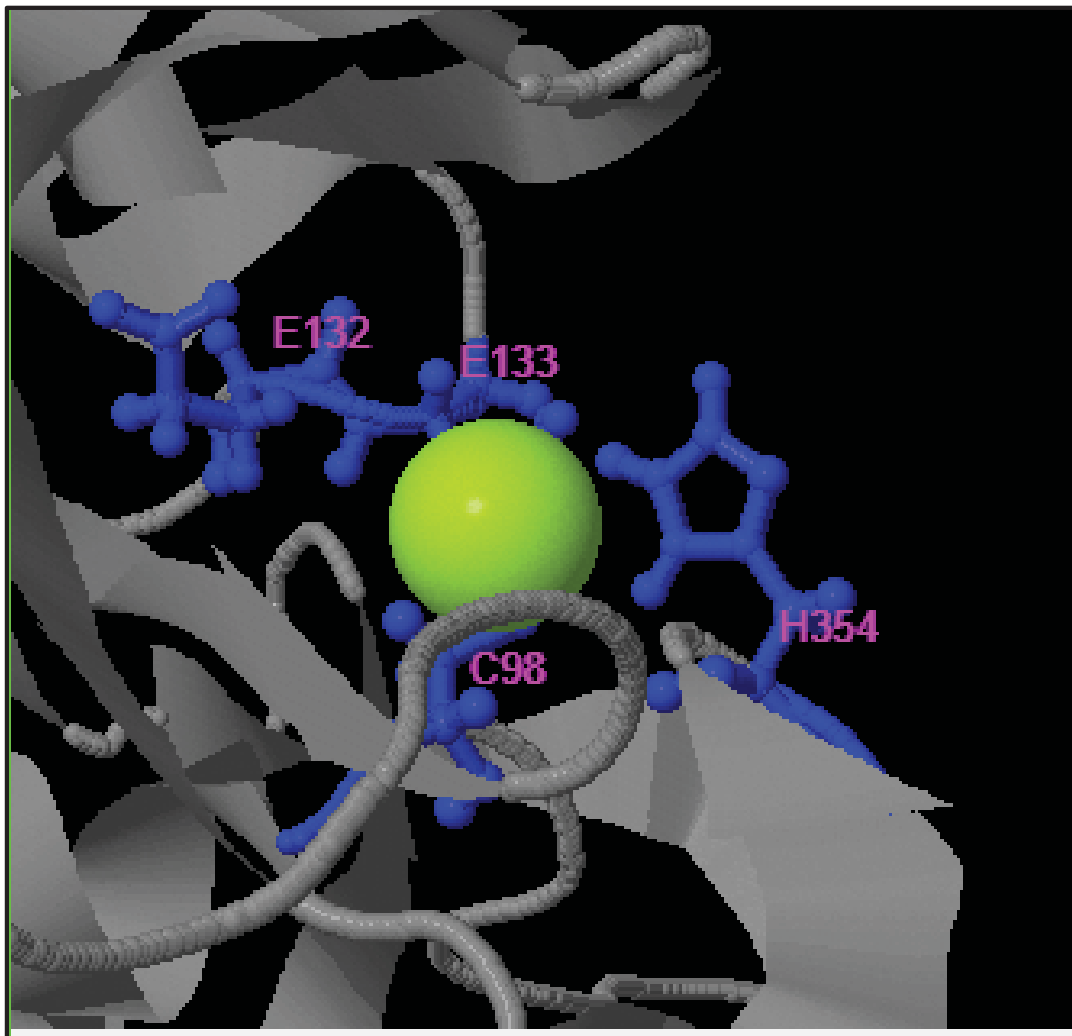


Figura 38: Visualización de los sitios de unión a los ligandos de Manganeso y sitio activo de la Amidohidrolasa MP25462 con el programa I-Tasser.

3.1.6 DISEÑO DE CEBADORES

Las secuencias diseñadas para el cebador sentido fue 5' GGTCTGGGTGGTCATGGC 3' que inicia en el nucleótido 145 hasta el nucleótido 162 y su Tm es de 62°C denominado AMD1F y para el cebador antisentido fue 5'CATCCGACTGATGGAGCCTC 3' que inicia en el nucleótido 1147 hasta el nucleótido 1128 y su Tm es de 60°C denominado AMD1R (Tabla 19), esta pareja de cebadores amplificó un fragmento de ADN de 1003pb que pertenece al gen amidohidrolasa (Figura 39).

Tabla 19: Especificaciones del diseño de los cebadores.

ID	POSICIÓN	SECUENCIA DE CEBADORES	Tm (C°)
AMD1F	145-162	5'GGTCTGGGTGGTCATGGC 3'	62°C
AMD1R	1147-1128	5'CATCCGACTGATGGAGCCTC 3'	60°C

5'	1	21	41	61	81
	ATGAACGAGGCATTTTTGTC	GCTGACGCCGGCCAGATTC	GTCATCGGCTGCATGCCGCG	CCGGAGCTGTGAGGGCAGGA	AACCGAAACGGCCCGCCGGG
	TACTTGCTCCGTAAAAACAG	CGACTGCGGCCGGGTCTAAG	CAGTAGCCGACGTACGGCGC	GGCCTCGACAGTCCCCTCCT	TTGGCTTTGCCGGGCGGCC
3'	101	121	141	161	181
	TGGCGGGCTGGCTCGAGCAG	GCGGGGGCGACACGCATCGT	CACGGTCTGGGTGGTCATG	GCGTGGCCGGGGTGTTCGCG	GGCACCTTGCCGGGGCCGCG
	ACCGCCCGACCGAGCTCGTC	CGCCCCGCTGTGCGTAGCA	GTGSCCAGACCCACCAGTAC	CCACCCGGCCCCACAAGCGC	CCGTGGAACGGCCCCGGGCGC
	201	221	241	261	281
	CGTCATGTTCCGTGCCGAAC	TCGATGGCTTGCCCATCGAC	GAGACCGCGAAATCGCCTG	GAAATCGACCTGTGAGCAGC	TGGCGCACAGTTGTGGCCAT
	GCAGTACAAGGCACGGCTTG	AGCTACCGAACGGGTAGCTG	CTCTGGCGCCTTTAGCGGAC	CTTTAGCTGGACACTCGTGC	ACCGCGTGTCAACACCGGTA
	301	321	341	361	381
	GACGGGCACATGGCGGTATT	GCTCGGATTGGCGCAGCGCC	TGGCGGAGCAGCCGCCGACG	CATGGCGAAACGGTGTCTGCT	GTTTCAGCCTTCGGAGGAAA
	CTGCCCGTGTACCGCCATAA	CGAGCCTAACC GCGTCGCGG	ACCGCCTCGTCGGCGGCTGC	GTACCGCTTTGCCACGACGA	CAAAGTCGGAAGCCTCCTTT
	401	421	441	461	481
	CCGGGCGTGGCGCCCGCGAG	GTGGTGAATCATCGCGTTT	CGAGGAAATCGCTCCCGATC	ATGTATATGCACTGCACAAT	CTGCCCGGCAAGCCGTTGGG
	GGCCCGACCGCGGGGCGCTC	CACCACCTTAGTAGCGCAA	GCTCCTTTAGCGAGGGCTAG	TACATATACGTGACGTGTTA	GACGGGCGGTTTCGGCAACCC
	501	521	541	561	581
	AGAGGTGTCGGTGAGAGAGG	GCGCCATGACCTGCGCCTCG	CGAGGGCTGATCATCCGGCT	CAACGGCTTGTGCGCGCATG	CGGCACACCCGGAGCAGGGC
	TCTCCACAGCCACTCTCTCC	CGCGGTACTGGACGCGGAGC	GCTCCCGACTAGTAGGCCGA	GTTGCCGAACACGCGCGTAC	GCCGTGTGGGCTCGTCCCG
	601	621	641	661	681
	CGCAACCCGGCGCTGGCGAT	GTGCCGATTCTGCAGGGTC	TCAAGCGTTGCCAGCAA	CTGCCGACGATCAAGGGCT	GGTCATGGTCACGCTGATCC
	GCGTTGGGCGCGACCGCTA	CACGGCCTAAGACGTCCAG	AGTTCCCAACGGGGTCGTT	GACGGGCTGCTAGTTCCCGA	CCAGTACAGTGCAGCTAGG
	701	721	741	761	781
	ACGCGCGATTGGGCGAGGTC	GCCTTCGGTACGTCGCCGGG	GCAGGCCGAGGTCATGGCCA	CGTTGCGCACCGAGGATGAC	GCCATGATGCGCGCGCTGGC
	TGCGCGCTAACCGCTCCAG	CGGAAGCCATGCAGCGGCC	CGTCCGGCTCCAGTACCGGT	GCAACGCGTGGCTCCTACTG	CGGTACTACGCGCGGACCG
	801	821	841	861	881
	GGAGGCCTCGGTGGCGCTTG	CGCGGGCCGAAGCGGAGGAG	GACGACCTGGGGATCGAGAT	CGACTGGTGCATGTCTTTC	AGGCCACGCGCAATCATCCC
	CCTCCGGAGCCACCGCGAAC	GCGCCCGGCTTCGCCTCCTC	CTGCTGGACCCCTAGCTCTA	GCTGACCACGCTACAGAAAAG	TCCGGTGC GCGTTAGTAGGG
	901	921	941	961	981
	GCCGCCGTTGCCTCGATTTCG	CGAAGCCGCCCGGTCGCTGG	ACATGCCCTTGACGGCGCTG	CCCAGGGCCCACCGCTGGTC	GGAAGACTTCGGCTGGCTGG
	CGGCGGCAACGGAGCTAAGC	GCTTCGGCGGGCCAGCGACC	TGTACGGGAAGTCCCGCGAC	GGGCTCCGGGTGGCGACCAG	CCTTCTGAAGCCGACCGACC
	1001	1021	1041	1061	1081
	CCGCGCAGGGGCAAGGGTGGC	CTGTTGACCCTGGGCGCCGG	CGAGCGGTGCCTGCCGCTGC	ACAACCCCGGCTACGACTTT	CCGGATGCCCTGCTGGACAC
	GGCGCGTCCCCGTCCACGC	GACAACCTGGGACCCGCGGCC	GCTCGCCACGGACGGCGACG	TGTTGGGGCCGATGCTGAAA	GGCCTACGGGACGACCTGTG
	1101	1121	1141	1161	1181
	GGGAATCGCGTTGTTTCGAGC	GCCTGGTGAGGCTCCATCAG	TCGGATCGCTGA		
	CCCTTAGCGCAACAAGCTCG	CGGACCACTCCGAGGTAGTC	AGCCTACCGACT		

Figura 39: Secuencia del gen amidohidrolasa, donde se señalan los cebadores AMD1F y AMD1R

3.1.7 RECUPERACIÓN A PARTIR DE LA CRIOPRESERVACIÓN DE *C. salexigens* MP25462

La cepa MP25462 se recuperó en medio líquido SW al 15% de NaCl al cabo de las 18 horas observándose turbidez de color ámbar (Figura 40-A), el cultivo fresco fue utilizado para la extracción del ADN genómico y para sembrar en agar con SW al 15% de NaCl, observándose a las 48 horas las colonias blanquecinas, cremosas, convexas y circulares (Figura 40-B).

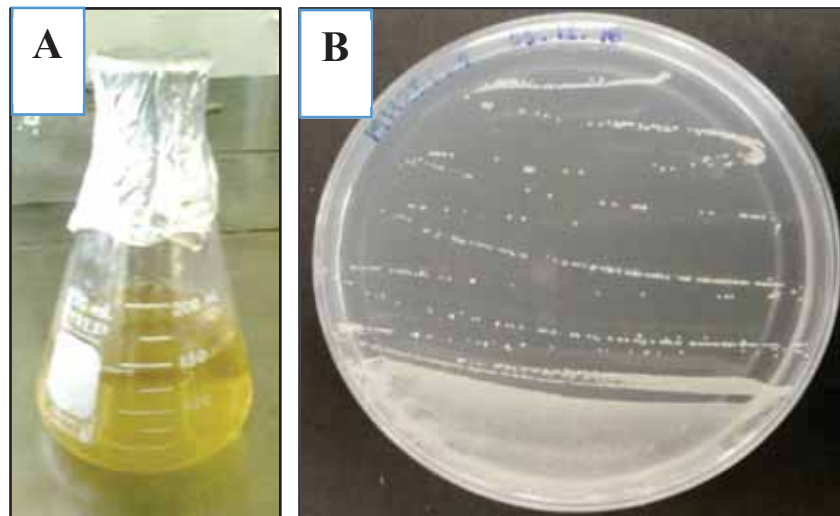


Figura 40: A. Cultivo fresco de *C. salexigens* MP25462 en medio líquido SW al 15% de NaCl. B. Colonias de *C. salexigens* MP25462 en Agar SW al 15% de NaCl.

3.1.8 AISLAMIENTO DEL ADN GENÓMICO DE *C. salexigens* MP25462.

El ADN aislado y purificado de *Chromohalobacter salexigens* se obtuvo en un volumen de 30 μ l y se cuantificó en el nanoespectrofotómetro y en gel de agarosa. Para la cuantificación al nanoespectrofotómetro se utilizó un volumen de muestra de 2 μ l de ADN genómico, los parámetros de cuantificación se aprecian en la tabla 20.

Tabla 20: Valores de la cuantificación del ADN genómico de *Chromohalobacter salexigens*.

Concentración	17.08 η g/ μ l
A260	0.342
260/230	0.66
260/280	2.21

Para la cuantificación en gel se preparó la agarosa al 1%, se uso un volumén de muestra de 2 μ l del ADN purificado y se uso el marcador molecular Lambda HindIII (Estándar o Std), con ayuda del programa *Image Lab* se visualizó y determinó la concentración del ADN purificado el cual fue de 9.45 η g/ μ l, además, el fabricante recomienda que la banda del ADN genómico este por encima de un tamaño de 20000pb, como se evidencia en la figura 41.

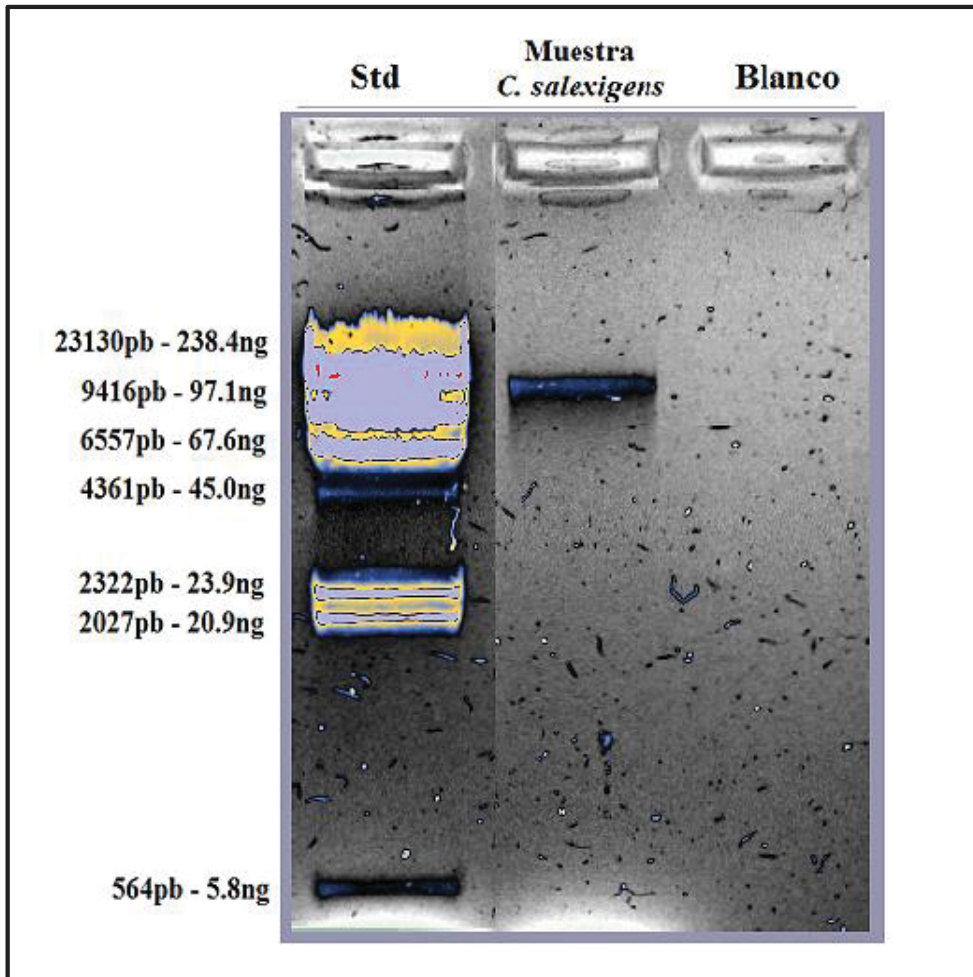


Figura 41: Cuantificación en gel del ADN genómico de *C. salexigens*, **Carril 1:** Estándar Lambda HindIII 23kb (Bioline), **Carril 2:** Muestra de ADN genómico, **Carril 3:** Muestra blanco

3.1.9 ESTANDARIZACIÓN DE LA PCR PARA LA AMPLIFICACIÓN DEL GEN AMIDOHIDROLASA

Para la amplificación del gen amidohidrolasa se realizó la estandarización de la temperatura de hibridación en la PCR utilizando los cebadores AMDF1 y AMDR1, se tomó en cuenta las temperaturas de hibridación de cada cebador al presentar AMDF1 62°C Y AMDR1 60°C y se opto por realizar un gradiente térmico de hibridación con las temperaturas que van desde 52, 54.9, 58.1

hasta 61°C. Como resultado, la PCR en gradiente amplificó a 52°C y 54.9°C una sola banda de aproximadamente 1000pb que pertenece al gen en estudio (Figura 42), revelando que la temperatura optima de hibridación es de 53°C.

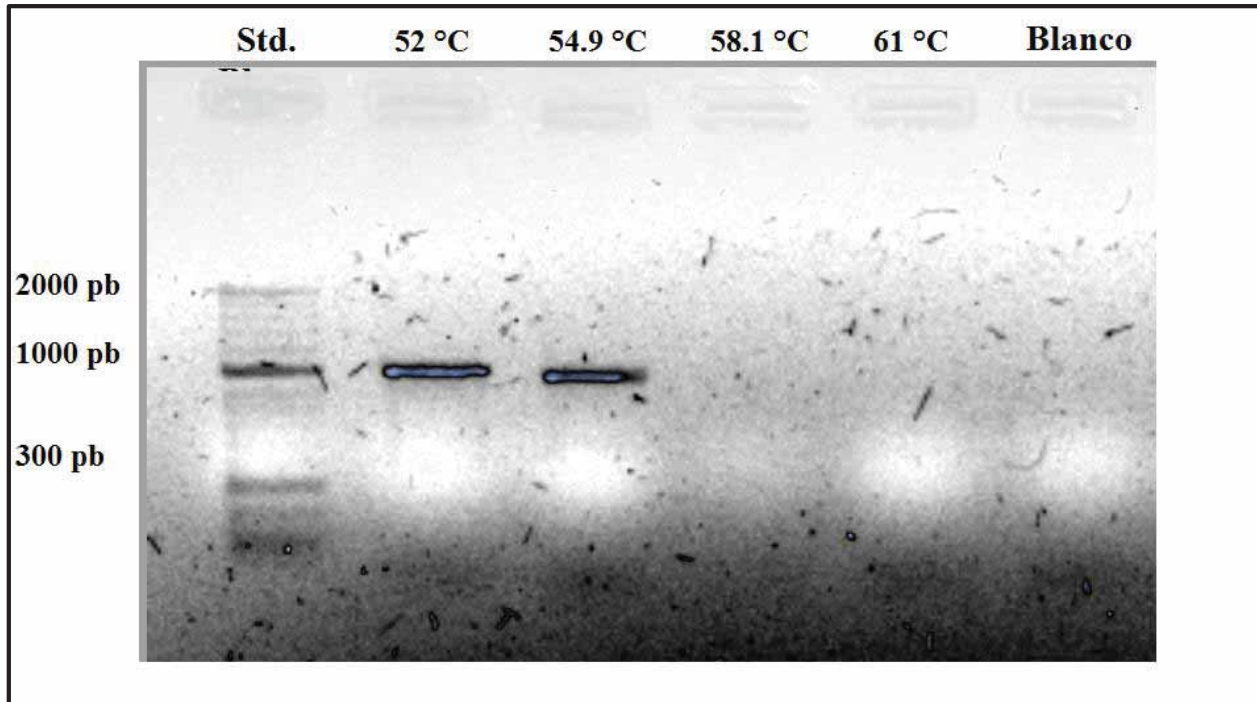


Figura 42: Estandarización de la temperatura de hibridación en PCR en gradiente, para la amplificación de la proteína Amidohidrolasa. Carril 1: Marcador molecular Hyper Lader II 2Kb (Bioline), Carril 2: Amplicón a 52°C, Carril 3: Amplicón a 54.9°C, Carril 4: Amplicón a 58.1°C, Carril 5: Amplicón a 61°C y Carril 6: Muestra blanco.

3.1.10 AMPLIFICACIÓN DEL GEN AMIDOHIDROLASA POR PCR

La temperatura de hibridación escogida para la amplificación del gen amidohidrolasa fue de 53°C. La finalidad de esta PCR es obtener la suficiente cantidad de muestra para llevar a cabo la purificación, secuenciación y clonación de la misma.

Para visualizar el producto de amplificación, se cargó 20µl de las reacciones de PCR en un gel de agarosa al 1%, se llevo a electroforesis durante 1 hora a 90V constantes (Figura 43)

Con estos resultados se entiende que los cebadores diseñados tienen su temperatura ideal de hibridación a 53°C, además, se realiza la amplificación del gen en cuatro carriles para su posterior uso en la purificación del ADN a partir de geles.

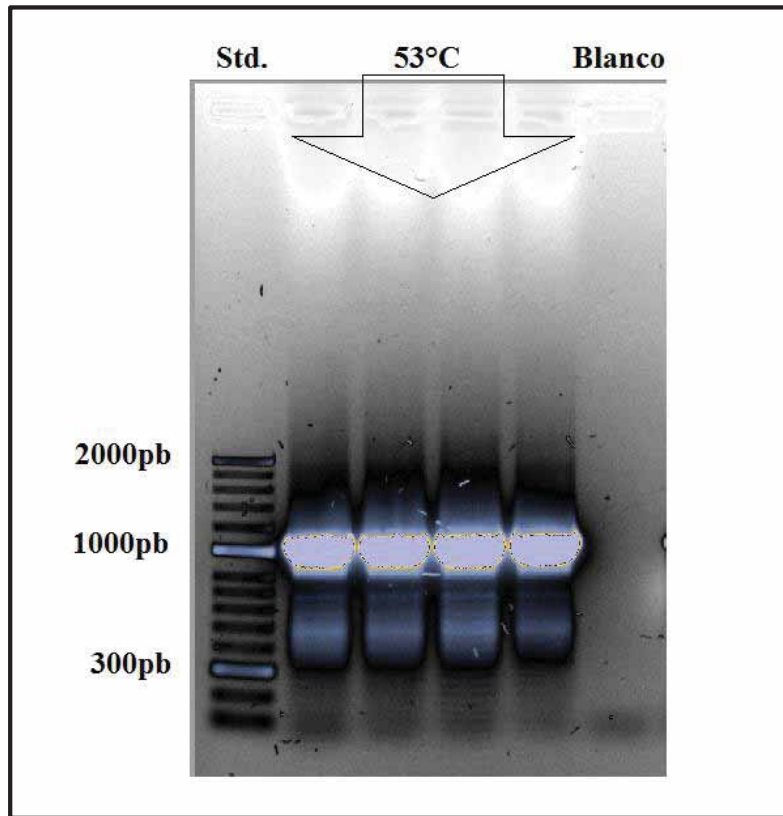


Figura 43: Amplificación del gen Amidohidrolasa a 53°C. **Carril 1:** Marcador molecular Hyper Lader II (Bioline), **Carril 2, 3,4 y 5:** Bandas del gen Amidohidrolasa, **Carril 6:** Muestra blanco.

3.1.11 PURIFICACIÓN Y SECUENCIACIÓN DEL GEN AMIDOHIDROLASA.

Para obtener una mayor concentración de la proteasa amidohidrolasa, se purificó las 4 bandas amplificadas (Figura 44) mediante el protocolo del *Kit Illustra™ GFX PCR DNA and Gel Band Purification* como se explicó en el capítulo 2.4.11 de materiales y métodos. La cuantificación se realizó en gel y al espectrofotómetro. La lectura al espectrofotómetro del ADN purificado resultó 117.30 ng/μl (Tabla 21).

Tabla 21: Valores de la cuantificación del ADN del gen Amidohidrolasa purificado.

Concentración	117.30 ng/μl
A260	2.346
260/230	2.08
260/280	1.94

La cuantificación en gel y su análisis se realizó con *Image Lab* teniendo en cuenta la concentración de cada banda del marcador *Hyper Lader II* (Figura 44). El resultado de esta cuantificación es de aproximadamente 114ng/μl.

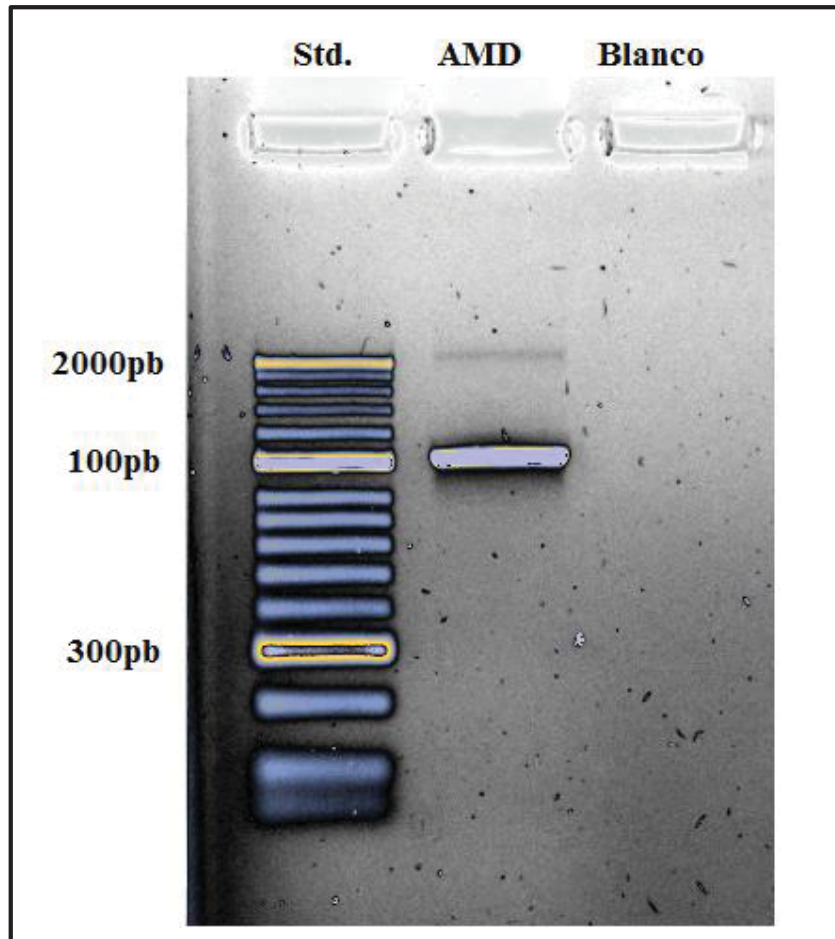


Figura 44: Cuantificación en gel de la proteasa Amidohidrolasa, **Carril 1:** Marcador molecular *Hyper Lader II* (Bioline) **Carril 2:** Muestra del amplicon de Amidohidrolasa purificado **Carril 3:** Muestra blanco

Los resultados del fragmento secuenciado por el SEGAI, fueron cargados al programa Mega 7 donde se revisó la asignación de bases nucleotídicas visualizando el electroferograma, el cual revela la asignación de bases mediante picos de colores, y cuanto mas definidos estén los picos de estos colores, se interpreta como una buena reacción de secuenciación, en este caso, llego hasta aproximadamente los 560pb (Figura 45). Además, utilizando el mismo programa, se alineo la secuencia teórica de la proteasa Amidohidrolasa producto de la anotación con el fragmento secuenciado, en el cual se observó similitud de nucleótidos entre las secuencias (Figura 47), lo cual demostró que se aisló el gen de la proteasa Amidohidrolasa.

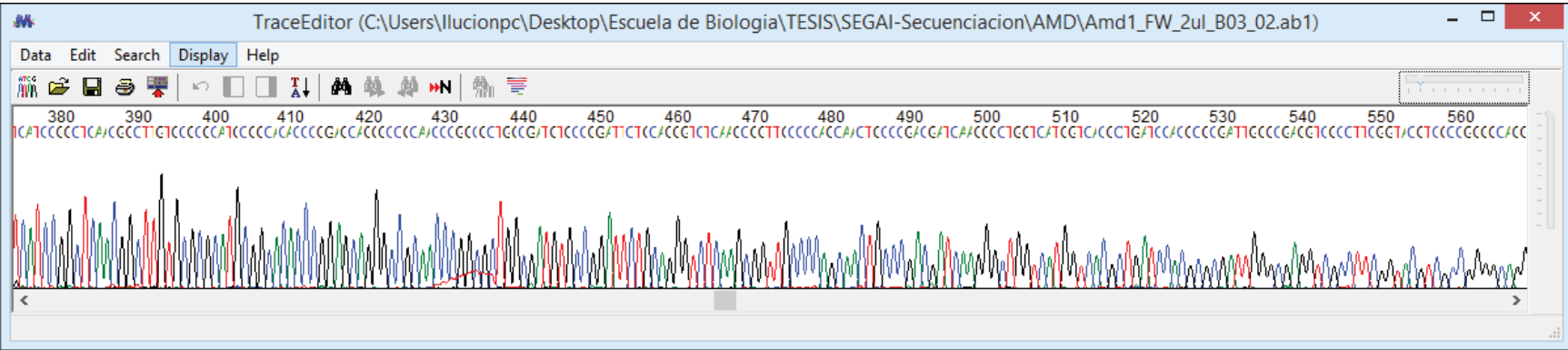


Figura 45: Imagen del cromatograma de la secuenciación de la proteasa amidohidrolasa amplificado

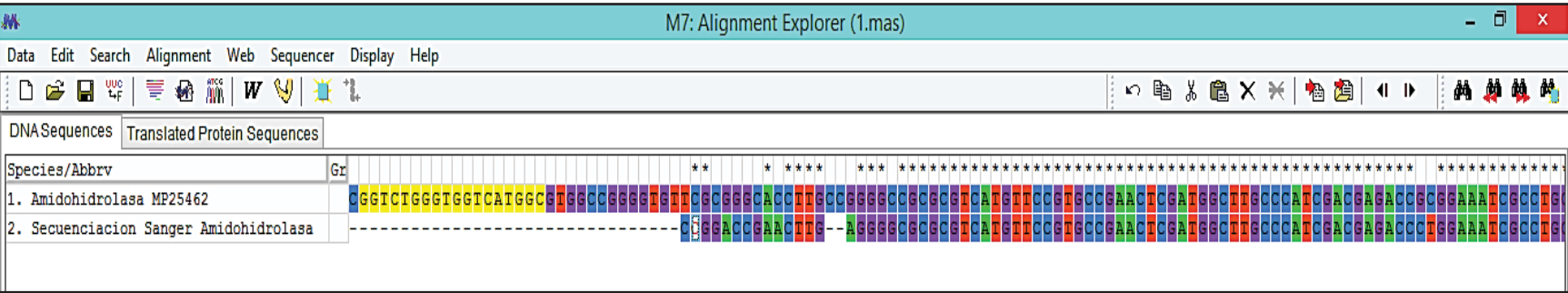


Figura 46: Alineación de secuencias. 1. Secuencia de la proteasa Amidohidrolasa, producto de la anotación 2. Secuencia de la proteasa Amidohidrolasa purificado, producto de la PCR. El segmento amarillo es la secuencia del cebador AmdF1.

De igual manera, el fragmento secuenciado fue analizado mediante BLAST para conocer la identidad del fragmento amplificado. Dicho análisis otorgo 100% de identidad con la proteasa amidohidrolasa de *Chromohalobacter salexigens*, como se observa en la figura 47 y figura 48. Tomando en cuenta estos datos se procedieron con seguridad los siguientes experimentos.

	Descripción	Nombre científico	Cubierta de consulta	Por. Identificador	Cuenta Len	Adhesión
✓	amidohidrolasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100,00%	383	WP_179153462.1
✓	amidohidrolasa [Chromohalobacter sp.]	Chromohalobacter sp.	100%	99,74%	383	NQY44242.1
✓	amidohidrolasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,48%	383	WP_110062444.1
✓	amidohidrolasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	99%	99,48%	383	WP_011507095.1
✓	amidohidrolasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,22%	383	WP_128690692.1
✓	amidohidrolasa [Chromohalobacter marismortui]	Chromohalobacter marismortui	100%	90,60%	384	WP_133694381.1

Figura 47: Porcentaje de identidad del amplicón de Amidohidrolasa secuenciado en la base de datos del Blast- NCBI.

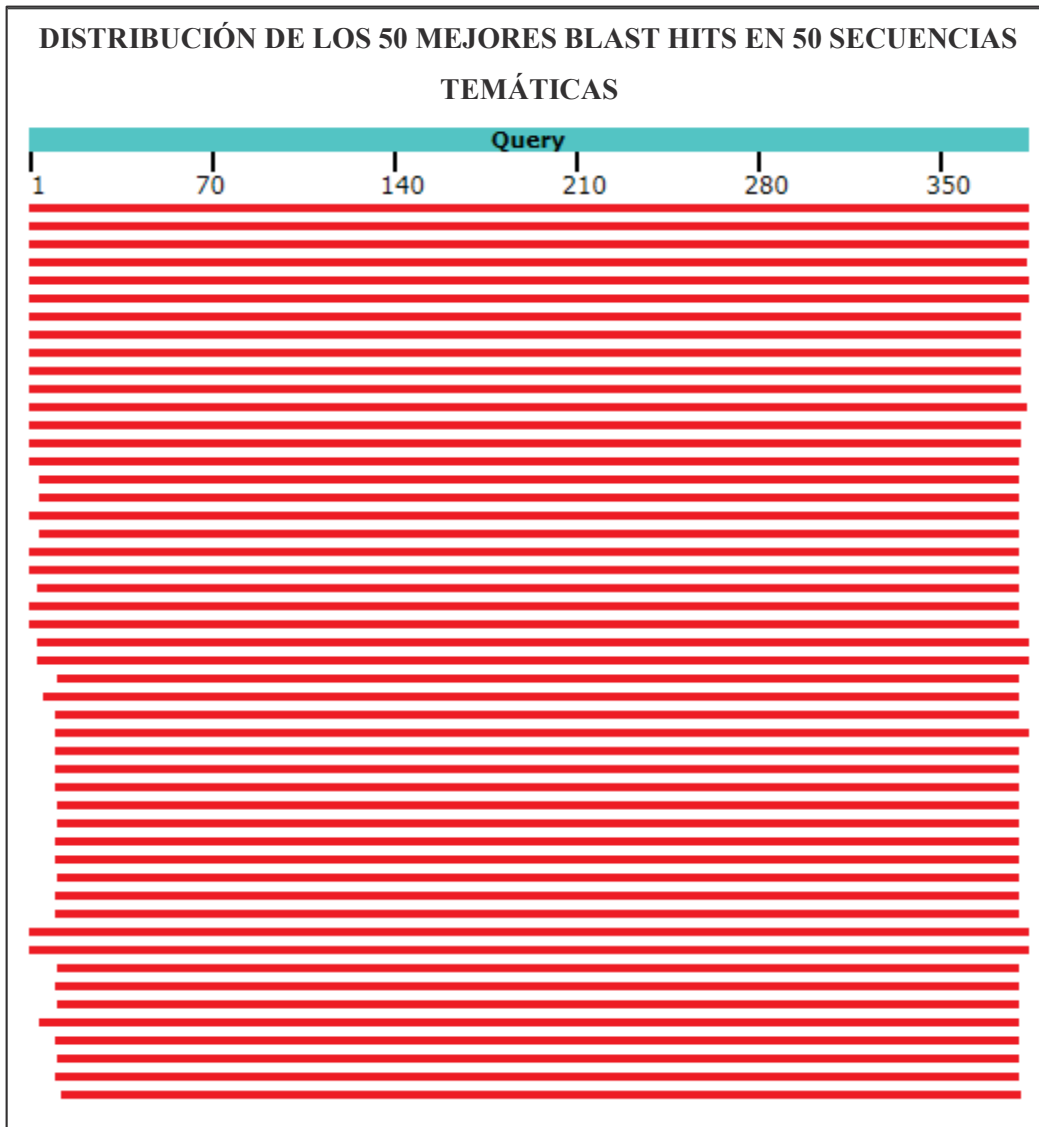


Figura 48: Secuencias de la base de datos alineadas con la secuencia de consulta (gen de Amidohidrolasa secuenciado)

3.1.12 CLONACIÓN EN EL VECTOR pGEM-T.

El valor de cuantificación del gen amidohidrolasa fué 117.3 ng/ μ l, a partir de este dato se hizo los cálculos correspondientes. Como resultado de la transformación del inserto-vector en las células competentes JM109 se obtuvo el crecimiento de colonias en el medio LB/ampicilina al cabo de las 24 horas (Figura 49), las cuales contienen al vector pGEM-T mas el inserto Amidohidrolasa. Posteriormente, a partir del crecimiento en el medio LB/ampicilina se seleccionaron 15 colonias transformadas para la PCR colonia con los primers del vector pGEM-T.

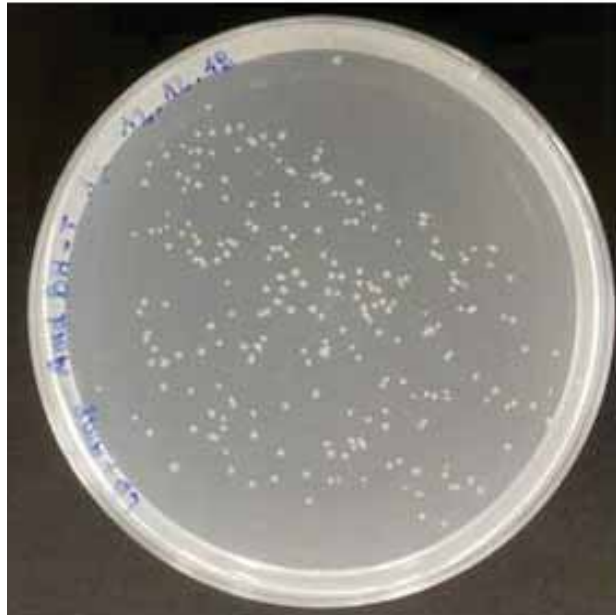


Figura 49: Crecimiento de colonias transformadas con inserto Amd- vector pGEM-T

La visualización de la PCR colonia se muestra en la figura 50, esta imagen demuestra los tamaños de las bandas en pares de bases y siendo el inserto, el gen amidohidrolasa de 1003 pb ligado a la región de policlonaje del vector pGEM-T de 163 pb, genera una banda con una longitud de 1166 pb. Resultando que las bandas de 14 colonias están próximas a la banda de 1200pb del marcador molecular evidenciando que estas células contienen el vector mas el inserto. El cuarto carril que corresponde a la colonia N°3 donde la banda esta por debajo de los 300pb indica que no se produjo la ligación inserto-vector y su crecimiento en el medio con antibiótico demuestra que estas células contienen al vector mas no al inserto, vector recircularizado. A partir de esta visualización se seleccionó la primera colonia denominada AMD-T1

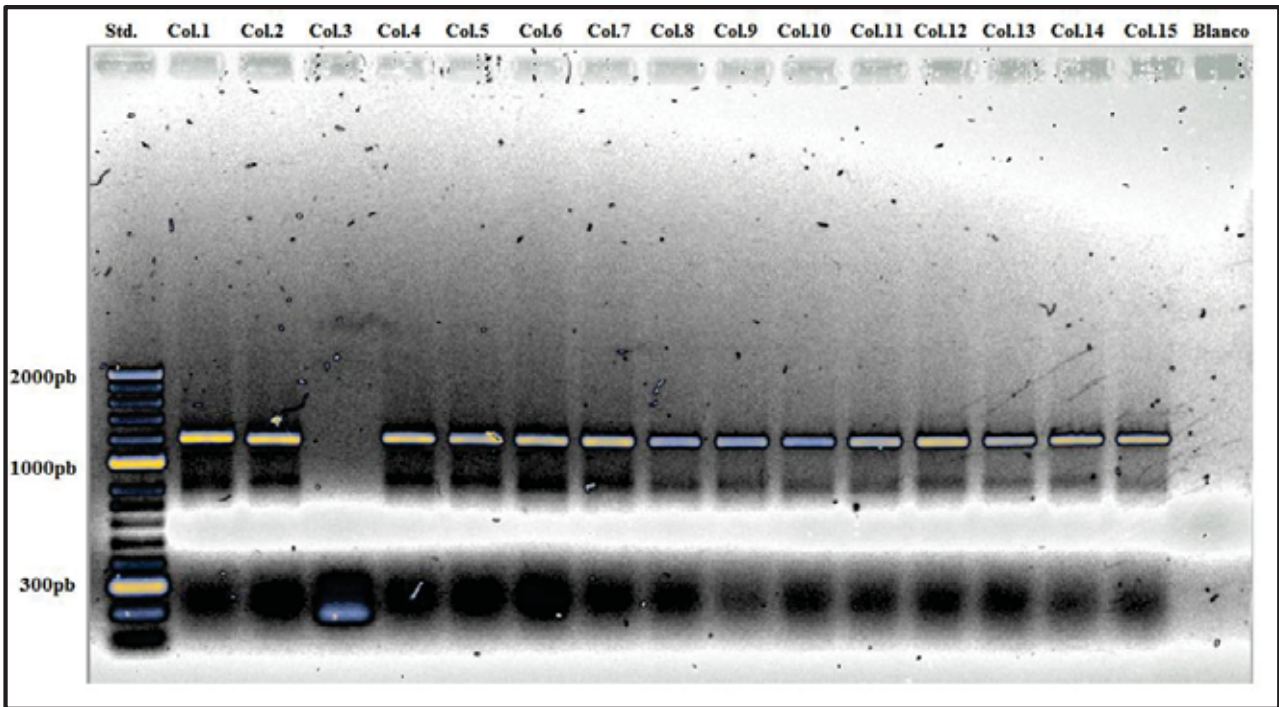


Figura 50: PCR colonia. Carril 1: Marcador molecular Hyper Lader II (Bioline) Carril 2- 16: colonias aisladas del 1 a 15. Último carril: blanco de reacción.

3.1.13 SECUENCIACIÓN DE LA PROTEASA AMIDOHIDROLASA CLONADO

A partir de la colonia AMD-T1 se purificó el plásmido y se cuantificó en el nanoespectrofotómetro resultando una concentración de 181.06 ng/ul. En la tabla 22, se observa el parámetro A260 que representa la longitud de onda a la cual se lee los ácidos nucleicos, el cociente 260/230 que es mayor a 1.8 lo que indica que no hay contaminación por componentes orgánicos y el cociente 260/280 que es mayor a 1.8, menor a este valor indicaría la presencia de proteínas. Es así que se determina que la cuantificación del ADN plasmídico es puro.

Tabla 22: Valores de la cuantificación del ADN del gen Amidohidrolasa purificado

Concentración AMDT1	181.06 ng/μl
A260	3.621
260/230	1.97
260/280	1.88

En la figura 51 se observa el alineamiento de nucleótidos de la secuencia antisentido del gen amidohidrolasa con la secuencia del plásmido, resultado de la secuenciación con el cebador T7, el cual amplifica en la cadena sentido del plásmido pGEM-T, también, se observa alineamiento de nucleótidos de la secuencia sentido del gen amidohidrolasa con la secuencia del plasmido, resultado de la secuenciación con el cebador SP6, que amplifica en la cadena antisentido del vector, lo que sugiere que el inserto ingreso al sitio de multiclonaje del vector en forma invertida.

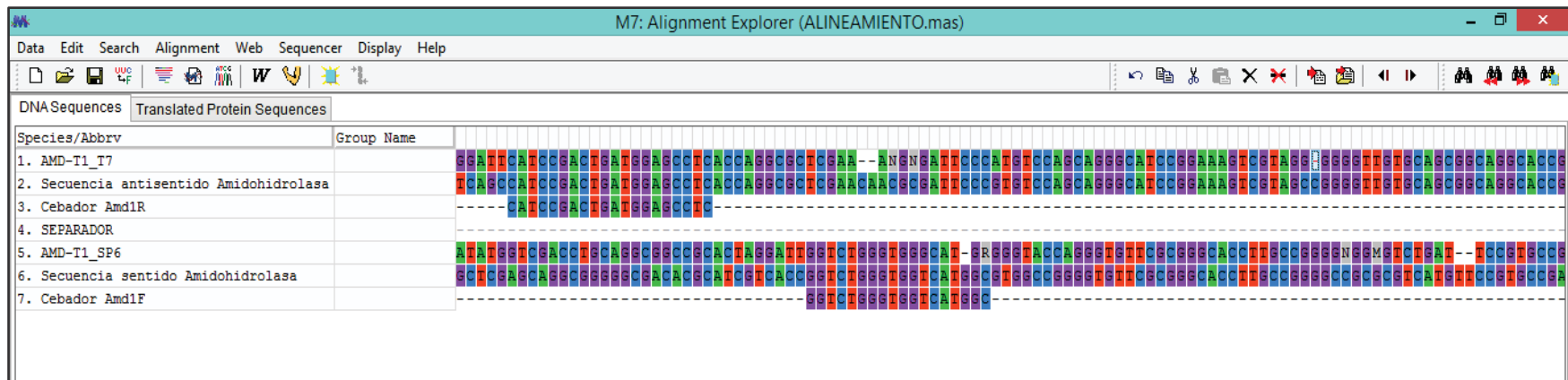


Figura 51: Alineamiento entre secuenciación del plásmido clonado, Amidohidrolasa y cebadores Amd1F/Amd1R. **Fila 1:** secuenciación del plásmido con cebador T7, **Fila 2:** secuencia anti sentido del gen Amidohidrolasa, **Fila 3:** Cebador anti sentido Amd1R, **Fila 5:** secuenciación del plásmido con cebador SP6, **Fila 6:** secuencia sentido del gen Amidohidrolasa, **Fila 7:** Cebador sentido Amd1F.

3.1.14 CRIOPRESERVACIÓN DE CÉLULAS CLONADAS

Las colonias positivas secuenciadas que contienen el plásmido con el inserto amidohidrolasa (AMDT1) fueron criopreservadas con fecha 9 de Diciembre del 2018.

3.2 DISCUSIÓN

3.2.1 DE LAS CARACTERÍSTICAS GENERALES DEL GENOMA DE *Chromohalobacter salexigens*.

La anotación del genoma de *C. salexigens*, actualmente tiene 5 reportes que datan desde el 2006 (Tabla 23), el primer genoma completo secuenciado y anotado fue la cepa de *C. salexigens* DSM 3043 actualmente *C. salexigens* cepa 1H11 (Arahal *et al.*, 2001), dicha cepa fue ensamblado en un solo andamio llegando a poseer 63.9% de G+C y un cromosoma largo de 3'696,649 pb con 3.319 genes (Copeland *et al.*, 2011). Posteriormente, se tiene registros de la cepa *C. salexigens* 40a_TX, ensamblado con Spades, llegando a obtener 63.9 % de G+C y una longitud de 3,600,000pb, por otro lado, estudios recientes reportaron el proyecto de secuencia del genoma de *C. salexigens* ANJ207 que tiene un cromosoma largo de 3'660.000 pb y un número total de 3,406 genes (Srivastava *et al.*, 2019) y el proyecto de secuencia del genoma de *C. salexigens* MP25462 del año 2020, en samblado con A5-miseq v.20160825, que tiene un cromosoma de 3,690,000pb y un total de 3437 genes.

En la tabla 23, se debe tener en cuenta que la cepa modelo es la DSM3043, debido a su cuidadoso análisis bioinformático manual. Contrastando la cepa modelo con la cepa en estudio se tiene una semejanza en cuanto a su longitud, porcentaje de G+C y el número de genes. Sin embargo comparando la cepa en estudio con las cepas 40a_TX y KG13, se observa diferencias en cuando a la longitud del genoma, el número de genes y el número de G+C, esto puede deberse a la diferencia de andamios ya que a mayor número de andamiaje menor garantía presta la anotación, en este caso la cepa en estudio tiene menor número de andamios en relación a estas cepas que usaron el mismo ensamblador.

Tabla 23: Características generales de reportes de genomas de *Chromohalobacter salexigens* ensamblados y anotados.

	<i>C. salexigens</i> DSM 3043	<i>C. salexigens</i> 40a_TX	<i>C. salexigens</i> ANJ 207	<i>C. salexigens</i> KG13	<i>C. salexigens</i> MP25462	<i>C. salexigens</i> MP25462*
Longitud	3.70	3,60	3,67	3,53	3,69	3.73
Porcentaje de G+C	63.90	63.90	63.70	61.00	64.00	63.90
Nº de genes	3402	3380	3432	3351	3437	3400
Nº de proteínas	3287	3239	3337	3207	3347	
Nº de andamios	1	81	5	69	17	42
contigs	-	82	59	72	17	89
ensamble	-	SPAdes v.3.10.1	Velvet v.1.2.10	SPAdes v.3.12.0	A5-miseq v.20160825	SPAdes v.3.13.1
Código de acceso	GCA_00005578 5.1	GCA_003182525. 1	GCA_004102695.1	GCA_01339340 5.1	GCA_013395975.1	-
Fecha de publicación	11-Abr-2006 Instituto Conjunto del Genoma del DOE	31-May-2018 Instituto Conjunto del Genoma del DOE	22-Ene-2019 Consejo Indio de Investigación Agrícola (ICAR-NBAIM)	08-Jul-2020 Universidad de Chung-Ang	10-Jul-2020 Universidad Nacional de San Antonio Abad	-
Referencia	(Copeland <i>et al.</i> , 2011) (Oren <i>et al.</i> , 2005)	-		(Chun <i>et al.</i> , 2021)	-	

* Cepa en estudio

3.2.2 DE LAS PROTEASAS DEL GENOMA DE *Chromohalobacter salexigens* MP25462

En este estudio se logró describir de forma manual 36 proteasas para la cepa *C. salexigens* MP25462 que codifican varios tipos de proteasas, según la clasificación de la *Enzyme Comision* (EC), se subdividido en endopeptidasa y exopeptidasas.

Dentro de las endopeptidasas se revelo 9 serin proteasas, 2 aspartil proteasa, 9 metalo proteasa, 1 treonin proteasa y ninguna cistein proteasa. (Figura 52)

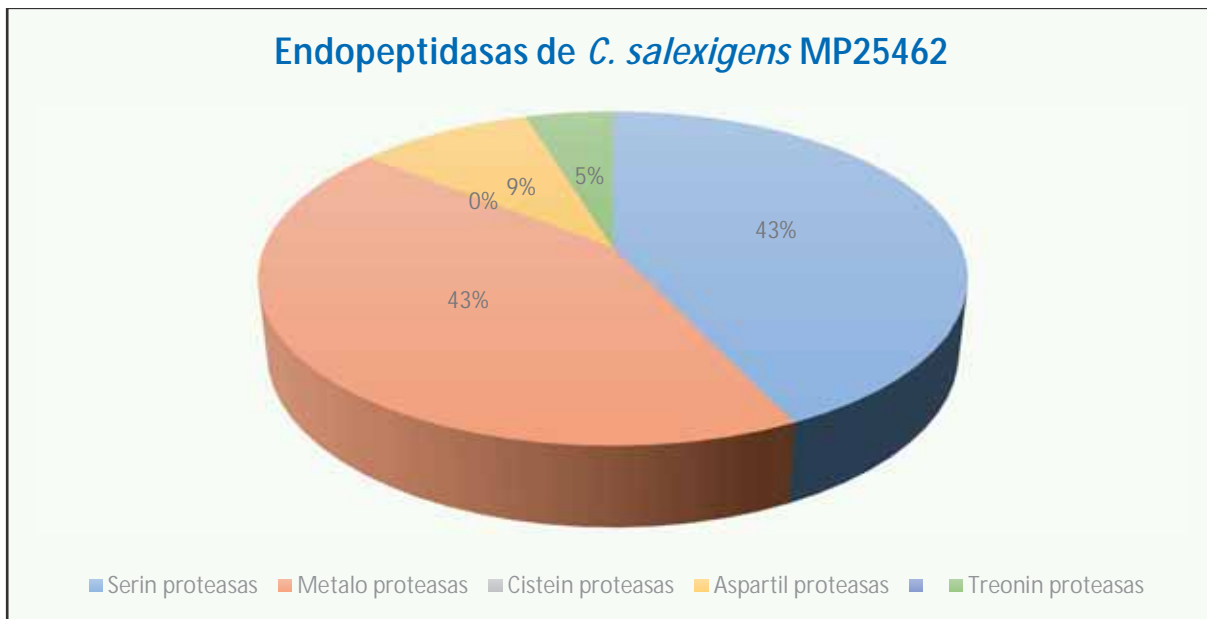


Figura 52: Distribución de las endopeptidasas del genoma de *C. salexigens* MP25462.

Las serin proteasas están descritas en organismos eucariotas y procariotas y han surgido durante la evolución como el grupo más abundante y funcionalmente diverso. Inicialmente se encontraron en *Neisseria*, *Shigella* y *Escherichia coli* patógena, pero ahora también se han identificado en especies de *Citrobacter rodentium*, *Salmonella* y *Edwardsiella*, en este estudio se describió serin proteasas de una bacteria halófila gram negativa, estas son: La proteasa Lon dependiente de ATP Tipo I (Gibellini *et al.*, 2020), proteasa II u oligopeptidasa B (Coetzer *et al.*, 2008; Motta *et al.*, 2019), represor de respuesta SOS y proteasa LexA (Mo *et al.*, 2018), peptidasa de señal I (Paetzel, 2019; Tuteja, 2005), proteasa de Clp (Ishizawa *et al.*, 2019), proteasa carboxilo terminal (Chueh *et al.*, 2019), proteasa DegQ (Wruse *et al.*, 2011), proteasa SohB (Baird *et al.*, 1991) y peptido señal peptidasa SppA (Cai *et al.*, 2017). En las bacterias gramnegativas, existe una familia

creciente de serina proteasas de alto peso molecular secretadas al medio externo por un mecanismo de secreción bacteriana fascinante y ampliamente utilizado, conocido como la vía del autotransportador. Se tiene a las serina proteasas para albergar sustratos proteicos intracelulares o extracelulares involucrados en numerosas funciones biológicas, como los implicados en la estabilidad del citoesqueleto, la autofagia o la inmunidad innata y adaptativa. (Ruiz-Perez & Nataro, 2014).

Las aspartil proteasas (EC 3.4.23) son un grupo de enzimas proteolíticas que se pueden encontrar en diferentes organismos, desde humanos hasta plantas y retrovirus. Las fuentes más conocidas de proteasas aspárticas son el estómago de los mamíferos, las levaduras y los hongos, (Szecsi, 1992). En este estudio las aspartil proteasas como peptidasa de señal II (EC 3.4.23.36) y Peptidasa líder (peptidasa prepilina) (EC 3.4.23.43) juegan un rol importante en la fisiología bacteriana, patogenicidad y resistencia a los antibióticos según (Lory & Strom, 1997; Vogeley *et al.*, 2016).

Las metaloproteasas extracelulares que contienen zinc están ampliamente distribuidas en el mundo bacteriano. Los más estudiados son los que están asociados con bacterias patógenas o bacterias que tienen importancia industrial. En este estudio se describen las siguientes metaloproteasas: Termolisina (Inouye *et al.*, 2007), precursor de la proteasa III (Cheng & Zipser, 1979), oligopeptidasa A (Lorenzon *et al.*, 2010) Proteína FtsH (Okuno & Ogura, 2013), Proteína htpX (Yoshitani *et al.*, 2019), Proteína FtsZ (Camberg *et al.*, 2009), Hidrolasa CocE / NonD (Ayarza *et al.*, 2015), Proteasa RseP (Hizukuri *et al.*, 2017), Metaloproteasa de la familia M48, bepA (anteriormente yfgC) (Bryant *et al.*, 2020; Narita *et al.*, 2013), muchas de ellas participan en el control de calidad de las proteínas de membrana, división celular y dentro de la industria como edulcorantes, las metaloproteasas se encuentran tanto en microorganismos gramnegativos como grampositivos, ya sean aeróbicos o anaeróbicos, dentro de la gran familia de metaloendopeptidasas bacterianas que contienen zinc, se observan unidades familiares más pequeñas, como metaloproteasas de tipo termolisina, elastasa. (Häse & Finkelstein, 1993)

Las treonin proteasas no han sido reportadas en organismos procariotas, sin embargo en este estudio se tiene a la Proteasa de ClpX dependiente de ATP que fue expresada en *E. coli* (Rohrwild *et al.*, 1996), por otro lado se tiene referencias de una treonin proteasa PfHs IV que es un objetivo farmacológico atractivo en la malaria y que se requiere más investigación para aclarar su papel funcional en el parásito. (Tschan *et al.*, 2011)

Las cistein proteasas descritas en plantas, animales y humanos, desempeñan múltiples funciones en básicamente todos los aspectos de la fisiología y el desarrollo. En las plantas, participan en el crecimiento, desarrollo, en la acumulación y movilización de proteínas de almacenamiento, en los animales y humanos, son responsables de la senescencia y la apoptosis, el procesamiento de prohormonas. (Wilkesman, 2017), en este estudio no se logro identificar alguna cistein proteasa para la bacteria halófila *C. salexigens* MP2462, esto puede deberse a que las cistein proteasas tienen actividad enzimática en un rango de pH 5-6, y se entiende que *C. salexigens* se encuentran en medios halo-alcalinos llegando a no expresar este grupo de proteasas (Wilkesman, 2017). Dentro de las exopeptidasas se revelo 7 aminopeptidasas, 7 carboxipeptidasas y 1 Omegapeptidasa (Figura 53)

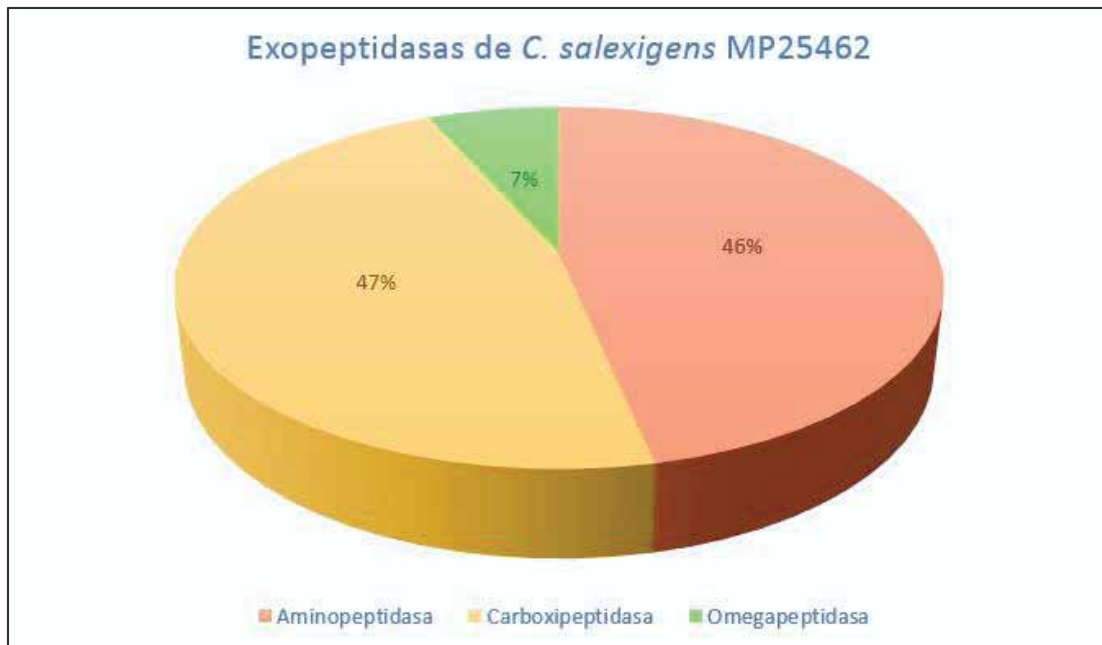


Figura 53: Distribución de las exopeptidasas del genoma de *C. salexigens* MP25462

Las aminopeptidasas, liberan selectivamente residuos de aminoácidos N-terminales de polipéptidos y proteínas. Las bacterias muestran varias actividades aminopeptidásicas que pueden estar localizadas en el citoplasma, en las membranas, asociadas con la envoltura celular o secretadas al medio extracelular. Los estudios sobre el sistema de aminopéptidos bacterianos se han llevado a cabo durante las últimas tres décadas y son importantes en dominios fundamentales y biotecnológicos (Gonzales & Robert-Baudouy, 1996). En este estudio se tiene las siguientes aminopeptidasas: Leucil aminopeptidasa (Zhang *et al.*, 2018), Aminopeptidasa N (Su *et al.*, 2011), Xaa-Pro aminopeptidasa (Yaron & Mlynar, 1968), Metionina aminopeptidasa (Helgren *et al.*,

2016), Aspartil aminopeptidasa (Yokoyama *et al.*, 2006), Beta-peptidil aminopeptidasa (Geueke & Kohler, 2007), Isoaspartil dipeptidasa (Park *et al.*, 2017).

Dentro de las carboxipeptidasas de la bacteria halófila en estudio se tiene a 7 integrantes, son: Cianoficinasa (Law *et al.*, 2009), Aminoacil-histidina dipeptidasa (Aoki *et al.*, 2012), D-alanil-D-alanina dipeptidasa (Lessard & Walsh, 1999), Dipeptidasa de membrana (Keynan *et al.*, 1997), Xaa-Pro dipeptidasa o Prolidasa (Wilk *et al.*, 2021), D-alanil-D-alanina carboxipeptidasa (Hu *et al.*, 2019), Amidohidrolasa (Seibert & Raushel, 2005), implicadas en el almacenamiento de nitrógeno, en la eliminación de radicales libres, en la estabilización de la capa del peptidoglicano. Estas carboxipeptidasas hidrolizan un enlace peptídico situado en el extremo carboxi-terminal de una proteína o polipéptido, liberando de esta forma el aminoácido situado al final de la cadena.

Dentro de las exopeptidasas se encuentra 1 omegapeptidasa Isoaspartil dipeptidasa (Park *et al.*, 2017), Esta enzima cataliza la escisión hidrolítica de dipéptidos beta-aspartilo. Los perfiles de tasa de pH para la hidrólisis de beta-Asp-Leu indican que la catálisis depende de la ionización de dos grupos; uno que se ioniza a un pH de aproximadamente 6 y el otro de aproximadamente 9. El grupo que debe ionizarse para la catálisis depende directamente de la identidad del ión metálico unido al sitio activo. (Martí-Arbona *et al.*, 2005)

3.2.3 DE LA OBTENCIÓN DE UNA PROTEASA RECOMBINANTE

La obtención de la proteasa recombinante se desarrollo mediante la técnica de clonación del inserto (Amidohidrolasa) en el plasmido p-GEMT, esta técnica es el paso previo para la expresión y purificación de proteínas y se ha venido desarrollandose a lo largo de los últimos años con mayor influencia en diversas partes del mundo así se tiene referencias de Yamaguchi *et al.*, (2012), que clonó una proteína periplasmica de unión a metales de un halófilo moderado en células de *E. coli*. Así mismo, con la obtención de la proteasa recombinante se demuestra que la base de estos trabajos es el analisis bioinformático del genoma, el cual sirve de pie para el diseño de primers que logra amplificar el gen esperado y no otro segmento del genoma. Así mismo, este proceso de obtención de proteasas recombinantes sirve de base para futuros proyectos de investigación, que ahora tendran conocimiento del repertorio de todas las proteasas que contiene el genoma en estudio con su respectiva posición intergenica y al demostrarse el aislamiento de una de las proteasas (gen Amidohidrolasa) se garantiza la ubicación y el futuro aislamiento de los demas genes proteolíticos.

CONCLUSIONES

1. Se anotó el genoma de *Chromohalobacter salexigens* MP25462, a partir del ensamble de novo de las secuencias corregidas, con el *software* Spades, caracterizando al genoma con una longitud de 3'712,216 pb, 42 número de contigs, estadístico N50 con 447018 pb, y contenido de G + C del 63.99%.
2. Se localizó 36 secuencias de genes que codifican a enzimas proteolíticas en el genoma de *Chromohalobacter salexigens* MP25462.
3. Se describió *in silico* las proteasas localizadas, agrupando en endopeptidasas y exopeptidasas, obteniendo dentro de las endopeptidasas 9 proteasas de tipo serina, 2 proteasas de tipo aspartil, 9 proteasas de tipo metalo y 1 proteasa de tipo treonina; y dentro de las exopeptidasas, 7 proteasas de tipo aminopeptidasa, 7 proteasas de tipo carboxipeptidasa y 1 proteasa de tipo omegapeptidasa, donde se destacó la caracterización *in silico* de la proteasa Amidohidrolasa, una carboxipeptidasa que tiene afinidad por varias especificidades de sustrato, altamente evolutivo y con una alta diversidad conformacional.
4. Se logró aislar *in vitro* el gen completo de la proteasa Amidohidrolasa de *C. salexigens* MP25462 con la pareja de cebadores diseñados AMD1F 5'GGTCTGGGTGGTCATGGC 3'y AMD1R 5'CATCCGACTGATGGAGCCTC 3' con Tm de 62°C y 60°C respectivamente, con una longitud de 1013pb.
5. Se clonó el gen de la proteasa Amidohidrolasa en el vector de clonación pGEM-T y se seleccionó la primera colonia denominada AMDT1, la cual fue criopreservada.

RECOMENDACIONES

1. Es importante desarrollar nuevas técnicas y herramientas bioinformáticas para la descripción de enzimas proteolíticas, que favorezcan la identificación y determinación de sitios activos en bacterias halófilas para generar mayor información sobre su modo de acción enzimática.
2. Continuar con la identificación y descripción bioinformática de otros grupos de proteínas de interés biotecnológico de la bacteria halófila *C. salexigens*, aislada de la salineras de Maras - Cusco.

BIBLIOGRAFÍA

- Aguilar-Bultet, L., & Falquet, L. (2015). Secuenciación y ensamblaje de novo de genomas bacterianos: Una alternativa para el estudio de nuevos patógenos. *Revista de Salud Animal*, 37(2), 125-132. http://scielo.sld.cu/scielo.php?script=sci_abstract&pid=S0253-570X2015000200008&lng=es&nrm=iso&tlng=es
- Amoozegar, M. A., Safarpour, A., Noghabi, K. A., Bakhtiary, T., & Ventosa, A. (2019). Halophiles and Their Vast Potential in Biofuel Production. *Frontiers in Microbiology*, 10. <https://doi.org/10.3389/fmicb.2019.01895>
- Amri, E., & Mamboya, F. (2012). Papain, a Plant Enzyme of Biological Importance: A Review. *American Journal of Biochemistry and Biotechnology*, 8(2), 99-104. <https://doi.org/10.3844/ajbbsp.2012.99.104>
- Andrews, S. (2009). *FastQC: a quality control tool for high throughput sequence data*. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Aoki, A., Shibata, Y., Okano, S., Maruyama, F., Amano, A., Nakagawa, I., & Abiko, Y. (2012). Transition metal ions induce carnosinase activity in PepD-homologous protein from *Porphyromonas gingivalis*. *Microbial Pathogenesis*, 52(1), 17-24. <https://doi.org/10.1016/j.micpath.2011.09.003>
- Arahal, D. R., García, M. T., Vargas, C., Cánovas, D., Nieto, J. J., & Ventosa, A. (2001). *Chromohalobacter salexigens* sp. Nov., a moderately halophilic species that includes *Halomonas elongata* DSM 3043 and ATCC 33174. *International Journal of Systematic and Evolutionary Microbiology*, 51(Pt 4), 1457-1462. <https://doi.org/10.1099/00207713-51-4-1457>

- Arakawa, T., Yamaguchi, R., & Tokunaga, H. T. and M. (2016). Unique Features of Halophilic Proteins. *Current Protein & Peptide Science*, 18(1), 65-71. [https://doi.org/10.2174 / 1389203717666160617111140](https://doi.org/10.2174/1389203717666160617111140)
- Arshad, Z. I. M., Amid, A., Yusof, F., Jaswir, I., Ahmad, K., & Loke, S. P. (2014). Bromelain: An overview of industrial application and purification strategies. *Applied Microbiology and Biotechnology*, 98(17), 7283-7297. <https://doi.org/10.1007/s00253-014-5889-y>
- Ávila-Portillo, L. M., Madero, J. I., López, C., León, M. F., Acosta, L., Gómez, C., Delgado, L. G., Gómez, C., Lozano, J. M., & Reguero, M. T. (2006). Fundamentos de criopreservación. *Revista Colombiana de Obstetricia y Ginecología*, 57(4), 291-300. <https://doi.org/10.18597/rcog.468>
- Ayarza, J. M., Mazzella, M. A., & Erijman, L. (2015). Expression of stress-related proteins in *Sediminibacterium* sp. Growing under planktonic conditions. *Journal of Basic Microbiology*, 55(9), 1134-1140. <https://doi.org/10.1002/jobm.201400725>
- 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., ... Zagnitko, O. (2008). The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics*, 9, 75. <https://doi.org/10.1186/1471-2164-9-75>
- Badger, J., Sauder, J. m., Adams, J. m., Antonysamy, S., Bain, K., Bergseid, M. g., Buchanan, S. g., Buchanan, M. d., Batiyenko, Y., Christopher, J. a., Emtage, S., Eroshkina, A., Feil, I., Furlong, E. b., Gajiwala, K. s., Gao, X., He, D., Hendle, J., Huber, A., ... Harris, T. j. r. (2005). Structural analysis of a set of proteins resulting from a bacterial genomics project.

- Proteins: Structure, Function, and Bioinformatics*, 60(4), 787-796.
<https://doi.org/10.1002/prot.20541>
- Badgajar, S. B., & Mahajan, R. T. (2009). *Proteolytic Enzymes of Some Laticiferous Plants Belonging to Khandesh Region of Maharashtra, India*. 2(9), 1434-1437.
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.736.2723&rep=rep1&type=pdf>
- Baird, L., Lipinska, B., Raina, S., & Georgopoulos, C. (1991). Identification of the *Escherichia coli* sohB gene, a multicopy suppressor of the HtrA (DegP) null phenotype. *Journal of Bacteriology*, 173(18), 5763-5770. <https://doi.org/10.1128/jb.173.18.5763-5770.1991>
- Banerjee, G., & Ray, A. K. (2017). Impact of microbial proteases on biotechnological industries. *Biotechnology & Genetic Engineering Reviews*, 33(2), 119-143.
<https://doi.org/10.1080/02648725.2017.1408256>
- Banik, S., Biswas, S., & Karmakar, S. (2018). Extraction, purification, and activity of protease from the leaves of *Moringa oleifera*. *F1000Research*, 7.
<https://doi.org/10.12688/f1000research.15642.1>
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., & Pevzner, P. A. (2012). SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology*, 19(5), 455-477. <https://doi.org/10.1089/cmb.2012.0021>
- Barba, M., Glansdorff, N., & Labedan, B. (2013). Evolution of Cyclic Amidohydrolases: A Highly Diversified Superfamily. *Journal of Molecular Evolution*, 77(3), 70-80.
<https://doi.org/10.1007/s00239-013-9580-1>

- Battaner Arias, E. (2013). En *Compendio de Enzimologia* (p. 48,66). Ediciones Universidad de Salamanca y Enrique Battaner Arias.
- Baynes, J., & Dominiczak, M. (2019). *Bioquímica médica—5th Edition*. Elsevier.
<https://www.elsevier.com/books/bioquimica-medica/baynes/978-84-9113-406-0>
- Bentley, D. R., Balasubramanian, S., Swerdlow, H. P., Smith, G. P., Milton, J., Brown, C. G., Hall, K. P., Evers, D. J., Barnes, C. L., Bignell, H. R., Boutell, J. M., Bryant, J., Carter, R. J., Keira Cheetham, R., Cox, A. J., Ellis, D. J., Flatbush, M. R., Gormley, N. A., Humphray, S. J., ... Smith, A. J. (2008). Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*, *456*(7218), 53-59. <https://doi.org/10.1038/nature07517>
- Bergkessel, M., & Guthrie, C. (2013). Chapter Twenty Five—Colony PCR. En J. Lorsch (Ed.), *Methods in Enzymology* (Vol. 529, pp. 299-309). Academic Press.
<https://doi.org/10.1016/B978-0-12-418687-3.00025-2>
- Bergmans, H. E., van Die, I. M., & Hoekstra, W. P. (1981). Transformation in *Escherichia coli*: Stages in the process. *Journal of Bacteriology*, *146*(2), 564-570.
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC216999/>
- Bitto, E., Bingman, C. A., Bittova, L., Houston, N. L., Boston, R. S., Fox, B. G., & Phillips Jr., G. N. (2009). X-ray structure of ILL2, an auxin-conjugate amidohydrolase from *Arabidopsis thaliana*. *Proteins: Structure, Function, and Bioinformatics*, *74*(1), 61-71.
<https://doi.org/10.1002/prot.22124>
- Blokesch, M. (2016). Natural competence for transformation. *Current Biology*, *26*(21), R1126-R1130. <https://doi.org/10.1016/j.cub.2016.08.058>

- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114-2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Brown, J., Pirrung, M., & McCue, L. A. (2017). FQC Dashboard: Integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool. *Bioinformatics*, 33(19), 3137-3139. <https://doi.org/10.1093/bioinformatics/btx373>
- Brown, T. (2008). *Genomas/ Genome*. Ed. Médica Panamericana.
- Bryant, J. A., Cadby, I. T., Chong, Z.-S., Boelter, G., Sevastyanovich, Y. R., Morris, F. C., Cunningham, A. F., Kritikos, G., Meek, R. W., Banzhaf, M., Chng, S.-S., Lovering, A. L., & Henderson, I. R. (2020). Structure-Function Characterization of the Conserved Regulatory Mechanism of the Escherichia coli M48 Metalloprotease BepA. *Journal of Bacteriology*, 203(2), e00434-20. <https://doi.org/10.1128/JB.00434-20>
- Burley, S. K., Berman, H. M., Kleywegt, G. J., Markley, J. L., Nakamura, H., & Velankar, S. (2017). Protein Data Bank (PDB): The Single Global Macromolecular Structure Archive. *Methods in Molecular Biology (Clifton, N.J.)*, 1607, 627-641. https://doi.org/10.1007/978-1-4939-7000-1_26
- Butts, C. T., Zhang, X., Kelly, J. E., Roskamp, K. W., Unhelkar, M. H., Freitas, J. A., Tahir, S., & Martin, R. W. (2016). Sequence comparison, molecular modeling, and network analysis predict structural diversity in cysteine proteases from the Cape sundew, *Drosera capensis*. *Computational and Structural Biotechnology Journal*, 14, 271-282. <https://doi.org/10.1016/j.csbj.2016.05.003>
- Cai, D., Wang, H., He, P., Zhu, C., Wang, Q., Wei, X., Nomura, C. T., & Chen, S. (2017). A novel strategy to improve protein secretion via overexpression of the SppA signal peptide

- peptidase in *Bacillus licheniformis*. *Microbial Cell Factories*, 16, 70.
<https://doi.org/10.1186/s12934-017-0688-7>
- Camberg, J. L., Hoskins, J. R., & Wickner, S. (2009). ClpXP protease degrades the cytoskeletal protein, FtsZ, and modulates FtsZ polymer dynamics. *Proceedings of the National Academy of Sciences of the United States of America*, 106(26), 10614-10619.
<https://doi.org/10.1073/pnas.0904886106>
- Campbell, N. A., & Reece, J. B. (2006). *Biología*. Ed. Médica Panamericana.
- Chamorro, M., Sandoval Peña, G. A., Chávez Gamarra, M., & La Rosa, R. (2015). Análisis bioinformático de la proteína SAG29 de *Arabidopsis thaliana* involucrada en la respuesta al estrés salino. *Ágora Revista Científica*, 2, 158. <https://doi.org/10.21679/arc.v2i2.34>
- Chavez-Gamarra, M., & Sandoval, G. A. (2014). Análisis bioinformático de la proteasa ClpP de *Streptococcus mutans* asociada a la formación de caries dental humana. *Ágora Revista Científica*, 1(2), 31-35. <https://doi.org/10.21679/arc.v1i2.29>
- Cheng, Y. S., & Zipser, D. (1979). Purification and characterization of protease III from *Escherichia coli*. *The Journal of Biological Chemistry*, 254(11), 4698-4706.
- Chromohalobacter* *Salexigens*—*Microbewiki*. (2012).
https://microbewiki.kenyon.edu/index.php/Chromohalobacter_Salexigens
- Chu, W.-H. (2007). Optimization of extracellular alkaline protease production from species of *Bacillus*. *Journal of Industrial Microbiology & Biotechnology*, 34(3), 241-245.
<https://doi.org/10.1007/s10295-006-0192-2>
- Chueh, C.-K., Som, N., Ke, L.-C., Ho, M.-R., Reddy, M., & Chang, C.-I. (2019). Structural Basis for the Differential Regulatory Roles of the PDZ Domain in C-Terminal Processing Proteases. *MBio*, 10(4), e01129-19. <https://doi.org/10.1128/mBio.01129-19>

- Chun, B. H., Han, D. M., Kim, H. M., Park, D., Jeong, D. M., Kang, H. A., & Jeon, C. O. (2021). Metabolic Features of Ganjang (a Korean Traditional Soy Sauce) Fermentation Revealed by Genome-Centered Metatranscriptomics. *MSystems*, 6(4), e0044121. <https://doi.org/10.1128/mSystems.00441-21>
- Cira-Chávez, L. A., Guevara-Luna, J., YadiraSoto-Padilla, M., Román-Ponce, B., Vásquez-Murrieta, M. S., & Estrada-Alvarado, M. I. (2018). Kinetics of Halophilic Enzymes. *Kinetics of Enzymatic Synthesis*. <https://doi.org/10.5772/intechopen.81100>
- Coetzer, T. H. T., Goldring, J. P. D., & Huson, L. E. J. (2008). Oligopeptidase B: A processing peptidase involved in pathogenesis. *Biochimie*, 90(2), 336-344. <https://doi.org/10.1016/j.biochi.2007.10.011>
- Contesini, F. J., Melo, R. R., & Sato, H. H. (2018). An overview of Bacillus proteases: from production to application. *Critical reviews in biotechnology*, 38(3), 321–334. <https://doi.org/10.1080/07388551.2017.1354354>
- Copeland, A., O'Connor, K., Lucas, S., Lapidus, A., Berry, K. W., Detter, J. C., Del Rio, T. G., Hammon, N., Dalin, E., Tice, H., Pitluck, S., Bruce, D., Goodwin, L., Han, C., Tapia, R., Saunders, E., Schmutz, J., Brettin, T., Larimer, F., ... Woyke, T. (2011). Complete genome sequence of the halophilic and highly halotolerant Chromohalobacter salexigens type strain (1H11(T)). *Standards in Genomic Sciences*, 5(3), 379-388. <https://doi.org/10.4056/sigs.2285059>
- Cornejo Romero, A., Serrato Diaz, A., Rendon Aguilar, B., Rocha Munive, M. G., & (editoras). (2014). *Herramientas moleculares aplicadas en ecología: Aspectos teóricos y prácticos* (Primera edición). Instituto Nacional de Ecología y Cambio Climático(INECC-SEMARNAT).

[http://www2.congreso.gob.pe/sicr/cendocbib/con4_uibd.nsf/770DBBBBD5ADF759505257D4900580FE6/\\$FILE/HerramientasMolecularesAplicadasEcolog%C3%ADa.pdf](http://www2.congreso.gob.pe/sicr/cendocbib/con4_uibd.nsf/770DBBBBD5ADF759505257D4900580FE6/$FILE/HerramientasMolecularesAplicadasEcolog%C3%ADa.pdf)

de Lourdes Moreno, M., Pérez, D., García, M. T., & Mellado, E. (2013). Halophilic Bacteria as a Source of Novel Hydrolytic Enzymes. *Life: Open Access Journal*, 3(1), 38-51. <https://doi.org/10.3390/life3010038>

Da Silva R. R. (2017). Bacterial and Fungal Proteolytic Enzymes: Production, Catalysis and Potential Applications. *Applied biochemistry and biotechnology*, 183(1), 1–19. <https://doi.org/10.1007/s12010-017-2427-2>

De Souza, P. M., Bittencourt, M. L. de A., Caprara, C. C., de Freitas, M., de Almeida, R. P. C., Silveira, D., Fonseca, Y. M., Ferreira, E. X., Pessoa, A., & Magalhães, P. O. (2015). A biotechnology perspective of fungal proteases. *Brazilian Journal of Microbiology*, 46(2), 337-346. <https://doi.org/10.1590/S1517-838246220140359>

Del Giacco, L., & Cattaneo, C. (2012). Introduction to Genomics. En V. Espina & L. A. Liotta (Eds.), *Molecular Profiling: Methods and Protocols* (pp. 79-88). Humana Press. https://doi.org/10.1007/978-1-60327-216-2_6

Devlin, T. M. (1999). En *BIOQUIMICA, Libro de Texto con Aplicaciones Clínicas* (Tercera edición, p. 128). Editorial Reverté.

Drury, L. (1994). Transformation of Bacteria by Electroporation. En A. J. Harwood (Ed.), *Protocols for Gene Analysis* (pp. 1-8). Humana Press. <https://doi.org/10.1385/0-89603-258-2:1>

Dumorné, K., Córdova, D. C., & Renganathan, M. A.-E. and P. (2017). *Extremozymes: A Potential Source for Industrial Applications*. 27(4), 649-659. <https://doi.org/10.4014/jmb.1611.11006>

- Dundas, N., Leos, N. K., Mitui, M., Revell, P., & Rogers, B. B. (2008). Comparison of Automated Nucleic Acid Extraction Methods with Manual Extraction. *The Journal of Molecular Diagnostics : JMD*, 10(4), 311-316. <https://doi.org/10.2353/jmoldx.2008.070149>
- Eguiarte, L. E., Souza, V., & Aguirre, X. (2007). *Ecología Molecular* (primera edición). https://www.researchgate.net/publication/258129643_Ecologia_Molecular
- Ellington, A., & Cherry, J. M. (1996). Characteristics of Amino Acids. *Current Protocols in Molecular Biology*, 33(1), A.1C.1-A.1C.12. <https://doi.org/10.1002/0471142727.mba01cs33>
- Escobar, C. A. M., Murillo, L. V. R., & Soto, J. F. (2011). Tecnologías bioinformáticas para el análisis de secuencias de ADN. *Scientia et Technica*, 3(49), 116-121. <https://doi.org/10.22517/23447214.1491>
- Ewing, B., & Green, P. (1998). Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Research*, 8(3), 186-194.
- Ewing, B., Hillier, L., Wendl, M. C., & Green, P. (1998). Base-Calling of Automated Sequencer Traces Using Phred. I. Accuracy Assessment. *Genome Research*, 8(3), 175-185. <https://doi.org/10.1101/gr.8.3.175>
- Fasim, A., More, V. S., & More, S. S. (2021). Large-scale production of enzymes for biotechnology uses. *Current opinion in biotechnology*, 69, 68-76. <https://doi.org/10.1016/j.copbio.2020.12.002>
- Flores, M. L., Zavaleta, A. I., Zambrano, Y., Cervantes, L., & Izaguirre, V. (2010). Bacterias halófilas moderadas productoras de hidrolasas de interés biotecnológico. *Ciencia e Investigación*, 13(1), 42-46. <http://revistasinvestigación.unmsm.edu.pe/index.php/farma/article/view/3224>

- Frigolet, M. E., & Gutiérrez-Aguilar, R. (2017). Ciencias “ómicas”, ¿cómo ayudan a las ciencias de la salud? *Revista Digital Universitaria*, 18(7).
<https://doi.org/10.22201/codeic.16076079e.2017.v18n7.a3>
- Gárate Palomino, A. W. (2016). *Obtención y Caracterización de proteasas de excreción a partir de bacterias halófilas aisladas de las salineras de Maras*. Universidad Nacional De San Antonio Abad Del Cusco Facultad De Ciencias.
- Gaur, S., Agrahari, S., & Wadhwa, N. (2010). Purification of Protease from *Pseudomonas thermaerum* GW1 Isolated from Poultry Waste Site. *The Open Microbiology Journal*, 4(1).
<https://doi.org/10.2174/1874285801004010067>
- Geueke, B., & Kohler, H.-P. E. (2007). Bacterial β -peptidyl aminopeptidases: On the hydrolytic degradation of β -peptides. *Applied Microbiology and Biotechnology*, 74(6), 1197-1204.
<https://doi.org/10.1007/s00253-007-0872-5>
- Gibellini, L., De Gaetano, A., Mandrioli, M., Van Tongeren, E., Bortolotti, C. A., Cossarizza, A., & Pinti, M. (2020). The biology of Lonp1: More than a mitochondrial protease. *International Review of Cell and Molecular Biology*, 354, 1-61.
<https://doi.org/10.1016/bs.ircmb.2020.02.005>
- Girish, T. S., B, V., Colaco, M., Misquith, S., & Gopal, B. (2013). Structure of an amidohydrolase, SACOL0085, from methicillin-resistant *Staphylococcus aureus* COL. *Acta Crystallographica Section F*, 69(2), 103-108.
<https://doi.org/10.1107/S1744309112049822>
- Gomes, J., & Steiner, W. (2004). The Biocatalytic Potential of Extremophiles and Extremozymes. *Food Technology and Biotechnology*, 42(4), 223-225. <https://hrcak.srce.hr/110869>

- Gonzales, T., & Robert-Baudouy, J. (1996). Bacterial aminopeptidases: Properties and functions. *FEMS Microbiology Reviews*, 18(4), 319-344. <https://doi.org/10.1111/j.1574-6976.1996.tb00247.x>
- Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUASt: Quality assessment tool for genome assemblies. *Bioinformatics*, 29(8), 1072-1075. <https://doi.org/10.1093/bioinformatics/btt086>
- Gurumalles, P., Alagu, K., Ramakrishnan, B., & Muthusamy, S. (2019). A systematic reconsideration on proteases. *International Journal of Biological Macromolecules*, 128, 254-267. <https://doi.org/10.1016/j.ijbiomac.2019.01.081>
- Hartley, B. S. (1960). Proteolytic enzymes. *Annual Review of Biochemistry*, 29, 45-72. <https://doi.org/10.1146/annurev.bi.29.070160.000401>
- Häse, C. C., & Finkelstein, R. A. (1993). Bacterial extracellular zinc-containing metalloproteases. *Microbiological Reviews*, 57(4), 823-837.
- Heather, J. M., & Chain, B. (2016). The sequence of sequencers: The history of sequencing DNA. *Genomics*, 107(1), 1-8. <https://doi.org/10.1016/j.ygeno.2015.11.003>
- Helgren, T. R., Wangtrakuldee, P., Staker, B. L., & Hagen, T. J. (2016). Advances in Bacterial Methionine Aminopeptidase Inhibition. *Current topics in medicinal chemistry*, 16(4), 397-414. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5009474/>
- Hernández, Marta, Quijada, Narciso M., Lázaro, David Rodríguez, & Eiros, José María. (2020). Aplicación de la secuenciación masiva y la bioinformática al diagnóstico microbiológico clínico. *Revista argentina de microbiología*, 52(2), 101-110. <https://dx.doi.org/10.1016/j.ram.2019.06.003>

- Hizukuri, Y., Akiyama, K., & Akiyama, Y. (2017). Biochemical Characterization of Function and Structure of RseP, an Escherichia coli S2P Protease. *Methods in Enzymology*, 584, 1-33. <https://doi.org/10.1016/bs.mie.2016.09.044>
- Hogeweg, P. (2011). The Roots of Bioinformatics in Theoretical Biology. *PLoS Computational Biology*, 7(3). [https://doi. urnal.pcbi.1002021org/10.1371/jo](https://doi.urnal.pcbi.1002021org/10.1371/jo)
- Hu, J., Lu, X., Wang, H., Wang, F., Zhao, Y., Shen, W., Yang, H., & Chen, X. (2019). Enhancing extracellular protein production in Escherichia coli by deleting the d-alanyl-d-alanine carboxypeptidase gene dacC. *Engineering in Life Sciences*, 19(4), 270-278. <https://doi.org/10.1002/elsc.201800199>
- Hu, T., Chitnis, N., Monos, D., & Dinh, A. (2021). Next-generation sequencing technologies: An overview. *Human Immunology*, 82(11), 801-811. <https://doi.org/10.1016/j.humimm.2021.02.012>
- Huaihua Puma, P. (2019). *Purificación, caracterización e inmovilización de proteasas de excreción de Staphylococcus sp. De las salineras de Maras—Cusco*. Universidad Nacional De San Antonio Abad Del Cusco Facultad De Ciencias.
- Hyman, E. D. (1988). A new method of sequencing DNA. *Analytical Biochemistry*, 174(2), 423-436. [https://doi.org/10.1016/0003-2697\(88\)90041-3](https://doi.org/10.1016/0003-2697(88)90041-3)
- Ibrahim, A. S. S., Al-Salamah, A. A., El-Badawi, Y. B., El-Tayeb, M. A., & Antranikian, G. (2015). Detergent-, solvent- and salt-compatible thermoactive alkaline serine protease from halotolerant alkaliphilic Bacillus sp. NPST-AK15: Purification and characterization. *Extremophiles: Life Under Extreme Conditions*, 19(5), 961-971. <https://doi.org/10.1007/s00792-015-0771-0>

- Inouye, K., Kusano, M., Hashida, Y., Minoda, M., & Yasukawa, K. (2007). Engineering, expression, purification, and production of recombinant thermolysin. *Biotechnology Annual Review*, 13, 43-64. [https://doi.org/10.1016/S1387-2656\(07\)13003-9](https://doi.org/10.1016/S1387-2656(07)13003-9)
- Ishizawa, J., Zarabi, S. F., Davis, R. E., Halgas, O., Nii, T., Jitkova, Y., Zhao, R., St-Germain, J., Heese, L. E., Egan, G., Ruvolo, V. R., Barghout, S. H., Nishida, Y., Hurren, R., Ma, W., Gronda, M., Link, T., Wong, K., Mabanglo, M., ... Andreeff, M. (2019). Mitochondrial ClpP-Mediated Proteolysis Induces Selective Cancer Cell Lethality. *Cancer Cell*, 35(5), 721-737.e9. <https://doi.org/10.1016/j.ccell.2019.03.014>
- Jamdar, S. N., Are, V. N., Navamani, M., Kumar, S., Nagar, V., & Makde, R. D. (2015). The members of M20D peptidase subfamily from Burkholderia cepacia, Deinococcus radiodurans and Staphylococcus aureus (HmrA) are carboxydipeptidases, primarily specific for Met-X dipeptides. *Archives of Biochemistry and Biophysics*, 587, 18-30. <https://doi.org/10.1016/j.abb.2015.10.003>
- Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., & Madden, T. L. (2008). NCBI BLAST: A better web interface. *Nucleic Acids Research*, 36(Web Server issue), W5-W9. <https://doi.org/10.1093/nar/gkn201>
- Joint Center for Structural Genomics, (JCSG). (2009). *Crystal structure of aminopeptidase PepT (NP_980509.1) from Bacillus cereus ATCC 10987 at 2.04 Å resolution*. 10(12), 980-980. <https://doi.org/10.1038/nsb1203-980>
- Joint Center for Structural Genomics, (JCSG). (2010). *Crystal structure of a M20A metallo peptidase (dapE, lpg0809) from Legionella pneumophila subsp. Pneumophila str. Philadelphia 1 at 1.50 Å resolution*. *Nature Structural & Molecular Biology*, 10(12), 980-980. <https://doi.org/10.2210/pdb3pfe/pdb>

- Joint Center for Structural Genomics, (JCSG). (2011). Crystal structure of a tripeptidase (SAV1512) from staphylococcus aureus subsp. Aureus mu50 at 2.10 Å resolution. *Nature Structural & Molecular Biology*, 10(12). <https://doi.org/10.2210/pdb3rza/pdb>
- Jutras, P. V., Dodds, I., & van der Hoorn, R. A. (2020). Proteases of *Nicotiana benthamiana*: An emerging battle for molecular farming. *Current Opinion in Biotechnology*, 61, 60-65. <https://doi.org/10.1016/j.copbio.2019.10.006>
- Karp, G. (2009). *BIOLOGÍA CELULAR Y MOLECULAR Conceptos y experimentos* (Quinta edición). McGraw Hill.
- Keynan, S., Hooper, N. M., & Turner, A. J. (1997). Identification by site-directed mutagenesis of three essential histidine residues in membrane dipeptidase, a novel mammalian zinc peptidase. *The Biochemical Journal*, 326 (Pt 1), 47-51. <https://doi.org/10.1042/bj3260047>
- Koçoğlu, T., & Yalçinkaya, T. (1992). [Recombinant DNA technology]. *Mikrobiyoloji Bulteni*, 26(2), 177-188.
- Kumar, A., & Chordia, N. (2015). In silico PCR primer designing and validation. *Methods in Molecular Biology (Clifton, N.J.)*, 1275, 143-151. https://doi.org/10.1007/978-1-4939-2365-6_10
- Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, 33(7), 1870-1874. <https://doi.org/10.1093/molbev/msw054>
- Kushner D. J. (1978). *Microbial life in extreme environments*. Academic Press. <http://agris.fao.org/agris-search/search.do?recordID=US201300558777>
- Landrigan, P. J., Stegeman, J. J., Fleming, L. E., Allemand, D., Anderson, D. M., Backer, L. C., Brucker-Davis, F., Chevalier, N., Corra, L., Czerucka, D., Bottein, M.-Y. D., Demeneix,

- B., Depledge, M., Deheyn, D. D., Dorman, C. J., Fénichel, P., Fisher, S., Gaill, F., Galgani, F., ... Rampal, P. (2020). Human Health and Ocean Pollution. *Annals of Global Health*, 86(1), 151. <https://doi.org/10.5334/aogh.2831>
- Law, A. M., Lai, S. W. S., Tavares, J., & Kimber, M. S. (2009). The structural basis of beta-peptide-specific cleavage by the serine protease cyanophycinase. *Journal of Molecular Biology*, 392(2), 393-404. <https://doi.org/10.1016/j.jmb.2009.07.001>
- Lee, S. H., Taguchi, H., Yoshimura, E., Minagawa, E., Kaminogawa, S., Ohta, T., & Matsuzawa, H. (1994). Carboxypeptidase Taq, a thermostable zinc enzyme, from *Thermus aquaticus* YT-1: Molecular cloning, sequencing, and expression of the encoding gene in *Escherichia coli*. *Bioscience, Biotechnology, and Biochemistry*, 58(8), 1490-1495. <https://doi.org/10.1271/bbb.58.1490>
- Leon Escobar, M. (2018). *Determinación de proteasas extracelulares e intracelulares de bacterias halófilas obtenidas de Huanoquite y Acos, provincias de Paruro y Acomayo, Cusco*. Universidad Nacional De San Antonio Abad Del Cusco Facultad De Ciencias.
- Lessard, I. A., & Walsh, C. T. (1999). VanX, a bacterial D-alanyl-D-alanine dipeptidase: Resistance, immunity, or survival function? *Proceedings of the National Academy of Sciences of the United States of America*, 96(20), 11028-11032. <https://doi.org/10.1073/pnas.96.20.11028>
- Levin, E. J., Kondrashov, D. A., Wesenberg, G. E., & Phillips, G. N. (2007). Ensemble Refinement of Protein Crystal Structures: Validation and Application. *Structure*, 15(9), 1040-1052. <https://doi.org/10.1016/j.str.2007.06.019>
- Liu, A., & Huo, L. (2014). Amidohydrolase Superfamily. *ELS*. John Wiley & Sons,. <https://doi.org/10.1002/9780470015902.a0020546.pub2>

- Lorenzon, R. Z., Cunha, C. E. L., Marcondes, M. F., Machado, M. F. M., Juliano, M. A., Oliveira, V., Travassos, L. R., Paschoalin, T., & Carmona, A. K. (2010). Kinetic characterization of the Escherichia coli oligopeptidase A (OpdA) and the role of the Tyr(607) residue. *Archives of Biochemistry and Biophysics*, 500(2), 131-136. <https://doi.org/10.1016/j.abb.2010.05.025>
- Lory, S., & Strom, M. S. (1997). Structure-function relationship of type-IV prepilin peptidase of Pseudomonas aeruginosa—A review. *Gene*, 192(1), 117-121. [https://doi.org/10.1016/s0378-1119\(96\)00830-x](https://doi.org/10.1016/s0378-1119(96)00830-x)
- Ludeña, M. (2009). *La enzimología*. El Cid Editor | apuntes. <https://books.google.com.pe/books?id=WgeLAQAACAAJ>
- Lylloff, J. E., Hansen, L. B. S., Jepsen, M., Sanggaard, K. W., Vester, J. K., Enghild, J. J., Sørensen, S. J., Stougaard, P., & Glaring, M. A. (2016). Genomic and exoproteomic analyses of cold- and alkaline-adapted bacteria reveal an abundance of secreted subtilisin-like proteases. *Microbial Biotechnology*, 9(2), 245-256. <https://doi.org/10.1111/1751-7915.12343>
- Madigan, M. T., & Oren, A. (1999). Thermophilic and halophilic extremophiles. *Current Opinion in Microbiology*, 2(3), 265-269.
- Mamani, J. I., Pacheco, K. B., Elorrieta, P., Romoacca, P., Castelan, H., Davila, S., Sierra, J. L., & Quispe-Ricalde, M. A. (2019). Draft Genome Sequence of Halomonas elongata MH25661 Isolated from a Saline Creek in the Andes of Peru. *Microbiology Resource Announcements*, 8(1). <https://doi.org/10.1128/MRA.00934-18>
- Manzoni, C., Kia, D. A., Vandrovцова, J., Hardy, J., Wood, N. W., Lewis, P. A., & Ferrari, R. (2018). Genome, transcriptome and proteome: the rise of omics data and their integration

- in biomedical sciences. *Briefings in bioinformatics*, 19(2), 286–302. <https://doi.org/10.1093/bib/bbw114>
- Martí-Arbona, R., Thoden, J. B., Holden, H. M., & Raushel, F. M. (2005). Functional significance of Glu-77 and Tyr-137 within the active site of isoaspartyl dipeptidase. *Bioorganic Chemistry*, 33(6), 448-458. <https://doi.org/10.1016/j.bioorg.2005.10.002>
- Mayerhöfer, T. G., Pahlow, S., & Popp, J. (2020). The Bouguer-Beer-Lambert law: Shining light on the obscure. *ChemPhysChem*, n/a(n/a). <https://doi.org/10.1002/cphc.202000464>
- Mazur, P. (1977). The role of intracellular freezing in the death of cells cooled at supraoptimal rates. *Cryobiology*, 14(3), 251-272. [https://doi.org/10.1016/0011-2240\(77\)90175-4](https://doi.org/10.1016/0011-2240(77)90175-4)
- Medvedev, P., Pham, S., Chaisson, M., Tesler, G., & Pevzner, P. (2011). Paired de Bruijn Graphs: A Novel Approach for Incorporating Mate Pair Information into Genome Assemblers. *Journal of Computational Biology*, 18(11), 1625-1634. <https://doi.org/10.1089/cmb.2011.0151>
- Minasov, G., HShuvalovaenrick, L., Brunzelle, J., Collart, F., & Anderson, W. (2005). Estructura cristalina de la proteína Bacillus Subtilis YXEP (APC1829), una peptidasa de unión a metales dinucleares de la familia M20. *Nature Structural & Molecular Biology*, 10(12), 980-980. <https://doi.org/10.2210/pdb1ysj/pdb>
- Mo, C. Y., Culyba, M. J., Selwood, T., Kubiak, J. M., Hostetler, Z. M., Jurewicz, A. J., Keller, P. M., Pope, A. J., Quinn, A., Schneck, J., Widdowson, K. L., & Kohli, R. M. (2018). Inhibitors of LexA Autoproteolysis and the Bacterial SOS Response Discovered by an Academic-Industry Partnership. *ACS Infectious Diseases*, 4(3), 349-359. <https://doi.org/10.1021/acsinfecdis.7b00122>

- Mokashe, N., Chaudhari, B., & Patil, U. (2018). Operative utility of salt-stable proteases of halophilic and halotolerant bacteria in the biotechnology sector. *International Journal of Biological Macromolecules*, *117*, 493-522. <https://doi.org/10.1016/j.ijbiomac.2018.05.217>
- Motta, F. N., Azevedo, C. D. S., Neves, B. P., Araújo, C. N. de, Grellier, P., Santana, J. M. de, & Bastos, I. M. D. (2019). Oligopeptidase B, a missing enzyme in mammals and a potential drug target for trypanosomatid diseases. *Biochimie*, *167*, 207-216. <https://doi.org/10.1016/j.biochi.2019.10.006>
- Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G., & Erlich, H. (1986). Specific Enzymatic Amplification of DNA In Vitro: The Polymerase Chain Reaction. *Cold Spring Harbor Symposia on Quantitative Biology*, *51*, 263-273. <https://doi.org/10.1101/SQB.1986.051.01.032>
- Narita, S., Masui, C., Suzuki, T., Dohmae, N., & Akiyama, Y. (2013). Protease homolog BepA (YfgC) promotes assembly and degradation of β -barrel membrane proteins in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(38), E3612-3621. <https://doi.org/10.1073/pnas.1312012110>
- Natsch, A., & Emter, R. (2020). The specific biochemistry of human axilla odour formation viewed in an evolutionary context. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1800), 20190269. <https://doi.org/10.2210/pdb6slf/pdb>
- Nyrén, P., & Lundin, A. (1985). Enzymatic method for continuous monitoring of inorganic pyrophosphate synthesis. *Analytical Biochemistry*, *151*(2), 504-509. [https://doi.org/10.1016/0003-2697\(85\)90211-8](https://doi.org/10.1016/0003-2697(85)90211-8)
- Okuno, T., & Ogura, T. (2013). FtsH protease-mediated regulation of various cellular functions. *Sub-Cellular Biochemistry*, *66*, 53-69. https://doi.org/10.1007/978-94-007-5940-4_3

- Oren, A. (2008). Microbial life at high salt concentrations: Phylogenetic and metabolic diversity. *Saline Systems*, 4(1), 2. <https://doi.org/10.1186/1746-1448-4-2>
- Oren, A. (2010). Industrial and environmental applications of halophilic microorganisms. *Environmental Technology*, 31(8-9), 825-834. <https://doi.org/10.1080/09593330903370026>
- Oren, A., Larimer, F., Richardson, P., Lapidus, A., & Csonka, L. N. (2005). How to be moderately halophilic with broad salt tolerance: Clues from the genome of *Chromohalobacter salexigens*. *Extremophiles*, 9(4), 275-279. <https://doi.org/10.1007/s00792-005-0442-7>
- Osiipiuk, J., Endres, M., Becker, D. P., Joachimiak, A., & Center for Structural Genomics of Infectious Diseases, (CSGID). (2021). DapE enzyme from *Shigella flexneri*. *Nature Structural & Molecular Biology*, 10(12), 980-980. <https://doi.org/10.2210/pdb7lgp/pdb>
- Pacheco Capcha, K. B. (2021). *Caracterización In silico y clonación del gen Aspartil Aminopeptidasa a partir de datos NGS (next generation sequencing) del halófilo moderado Chromohalobacter salexigens MP25462* [Universidad Nacional de San Antonio Abad del Cusco]. <https://repositorio.unsaac.edu.pe/handle/20.500.12918/5647>
- Paetzel, M. (2019). Bacterial Signal Peptidases. *Sub-Cellular Biochemistry*, 92, 187-219. https://doi.org/10.1007/978-3-030-18768-2_7
- Pallen, M. J. (2016). Microbial bioinformatics 2020. *Microbial Biotechnology*, 9(5), 681-686. <https://doi.org/10.1111/1751-7915.12389>
- Park, S.-H., Lee, C. W., Lee, S. G., Shin, S. C., Kim, H. J., Park, H., & Lee, J. H. (2017). Crystal structure and functional characterization of an isoaspartyl dipeptidase (CpsIadA) from *Colwellia psychrerythraea* strain 34H. *PloS One*, 12(7), e0181705. <https://doi.org/10.1371/journal.pone.0181705>

- PCR de colonias* / NEBRASKA. (s. f.). Recuperado 9 de octubre de 2019, de <https://www.neb.com/applications/cloning-and-synthetic-biology/dna-analysis/colony-pcr>
- Pegg, D. E. (2007). Principles of Cryopreservation. En J. G. Day & G. N. Stacey (Eds.), *Cryopreservation and Freeze-Drying Protocols* (pp. 39-57). Humana Press. https://doi.org/10.1007/978-1-59745-362-2_3
- Polge, C., Smith, A. U., & Parkes, A. S. (1949). Revival of Spermatozoa after Vitrification and Dehydration at Low Temperatures. *Nature*, 164(4172), 666-666. <https://doi.org/10.1038/164666a0>
- Promega Corporation. (2007). *PGEM-T and pGEM-T Easy Vector Systems*. www.promega.com/tbs/
- Punil, R., & Sandoval, G. A. (2016). Análisis bioinformático de la endolisina LysB4: Un agente antimicrobiano utilizado en el biocontrol de *Bacillus cereus* en alimentos. *Ágora Revista Científica*, 3(2), 375-380. <http://revistaagora.com/index.php/cieUMA/article/view/71>
- Rao, M. B., Tanksale, A. M., Ghatge, M. S., & Deshpande, V. V. (1998). Molecular and Biotechnological Aspects of Microbial Proteases. *Microbiology and Molecular Biology Reviews*, 62(3), 597-635. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC98927/>
- Rashid, Y., & Kamran Azim, M. (2011). Structural bioinformatics of *Neisseria meningitidis* LD-carboxypeptidase: Implications for substrate binding and specificity. *The Protein Journal*, 30(8), 558-565. <https://doi.org/10.1007/s10930-011-9364-7>
- Rawlings, N. D., Waller, M., Barrett, A. J., & Bateman, A. (2014). MEROPS: The database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Research*, 42(Database issue), D503-D509. <https://doi.org/10.1093/nar/gkt953>

- Razzaq, A., Shamsi, S., Ali, A., Ali, Q., Sajjad, M., Malik, A., & Ashraf, M. (2019). Microbial Proteases Applications. *Frontiers in Bioengineering and Biotechnology*, 7. <https://doi.org/10.3389/fbioe.2019.00110>
- Rhoads, A., & Au, K. F. (2015). PacBio Sequencing and Its Applications. *Genomics, Proteomics & Bioinformatics*, 13(5), 278-289. <https://doi.org/10.1016/j.gpb.2015.08.002>
- Rickwood, D., & Hames, B. D. (Eds.). (1990). *Gel Electrophoresis of Nucleic Acids: A Practical Approach* (Edición: 2). Oxford University Press.
- Robinson, P. K. (2015). Enzymes: Principles and biotechnological applications. *Essays in Biochemistry*, 59, 1-41. <https://doi.org/10.1042/bse0590001>
- Rohrwild, M., Coux, O., Huang, H. C., Moerschell, R. P., Yoo, S. J., Seol, J. H., Chung, C. H., & Goldberg, A. L. (1996). HslV-HslU: A novel ATP-dependent protease complex in *Escherichia coli* related to the eukaryotic proteasome. *Proceedings of the National Academy of Sciences of the United States of America*, 93(12), 5808-5813. <https://doi.org/10.1073/pnas.93.12.5808>
- Romoacca Hanco, P. (2018). *Identificación molecular mediante el gen ribosomal 16s y evaluación de las capacidades enzimáticas de bacterias halófilas aisladas en zonas salinas del departamento del Cusco*. Universidad Nacional De San Antonio Abad Del Cusco Facultad De Ciencias.
- Ronaghi, M. (1998). DNA SEQUENCING: A Sequencing Method Based on Real-Time Pyrophosphate. *Science*, 281(5375), 363-365. <https://doi.org/10.1126/science.281.5375.363>
- Ronner, P. (2019). *Netter. Bioquímica Esencial*. Elsevier Health Sciences.

- Rothberg, J. M., Hinz, W., Rearick, T. M., Schultz, J., Mileski, W., Davey, M., Leamon, J. H., Johnson, K., Milgrew, M. J., Edwards, M., Hoon, J., Simons, J. F., Marran, D., Myers, J. W., Davidson, J. F., Branting, A., Nobile, J. R., Puc, B. P., Light, D., ... Bustillo, J. (2011). An integrated semiconductor device enabling non-optical genome sequencing. *Nature*, *475*(7356), 348-352. <https://doi.org/10.1038/nature10242>
- Rowse, S., Paupit, R. A., Tucker, A. D., Melton, R. G., Blow, D. M., & Brick, P. (1997). Crystal structure of carboxypeptidase G2, a bacterial enzyme with applications in cancer therapy. *Structure*, *5*(3), 337-347. [https://doi.org/10.1016/S0969-2126\(97\)00191-3](https://doi.org/10.1016/S0969-2126(97)00191-3)
- Ruiz-Perez, F., & Nataro, J. P. (2014). Bacterial serine proteases secreted by the autotransporter pathway: Classification, specificity, and role in virulence. *Cellular and Molecular Life Sciences: CMLS*, *71*(5), 745-770. <https://doi.org/10.1007/s00018-013-1355-8>
- Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory.
- Sandhya, C., Alagarsamy, S., Szakacs, G., & Pandey, A. (2005). Comparative evaluation of neutral protease production by *Aspergillus oryzae* in submerged and solid-state fermentation. *PROCESS BIOCHEMISTRY*, *40*(8), 2689-2694. <https://doi.org/10.1016/j.procbio.2004.12.001>
- Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences*, *74*(12), 5463-5467. <https://doi.org/10.1073/pnas.74.12.5463>
- Sarwar, M. K., Azam, I., & Iqbal, T. (2015). Biology and Applications of Halophilic Bacteria and Archaea: A Review. *Electronic Journal of Biology*, *11*(3).

<https://ejbio.imedpub.com/abstract/biology-and-applications-of-halophilic-bacteria-and-archaea-a-review-7370.html>

- Seibert, C. M., & Raushel, F. M. (2005). Structural and Catalytic Diversity within the Amidohydrolase Superfamily. *Biochemistry*, 44(17), 6383-6391. <https://doi.org/10.1021/bi047326v>
- Slater, G. W., Guillouzie, S., Gauthier, M. G., Mercier, J.-F., Kenward, M., McCormick, L. C., & Tessier, F. (2002). Theory of DNA electrophoresis (~ 1999 –2002 ½). *ELECTROPHORESIS*, 23(22-23), 3791-3816. <https://doi.org/10.1002/elps.200290002>
- Sohn, J., & Nam, J.-W. (2018). The present and future of de novo whole-genome assembly. *Briefings in Bioinformatics*, 19(1), 23-40. <https://doi.org/10.1093/bib/bbw096>
- Sohpal, V. K., Dey, A., & Singh, A. (2010). MEGA biocentric software for sequence and phylogenetic analysis: A review. *International Journal of Bioinformatics Research and Applications*, 6(3), 230-240. <https://doi.org/10.1504/IJBRA.2010.034072>
- Solanki, P., Putatunda, C., Kumar, A., Bhatia, R., & Walia, A. (2021). Microbial proteases: ubiquitous enzymes with innumerable uses. *3 Biotech*, 11(10), 428. <https://doi.org/10.1007/s13205-021-02928-z>
- Srivastava, A. K., Sharma, A., Srivastava, R., Tiwari, P. K., Singh, A. K., Yadav, J., Jamali, H., Bharati, A. P., Srivastava, A. K., Kashyap, P. L., Chakdar, H., Kumar, M., & Saxena, A. K. (2019). Draft Genome Sequence of Halotolerant Bacterium *Chromohalobacter salexigens* ANJ207, Isolated from Salt Crystal Deposits in Pipelines. *Microbiology Resource Announcements*, 8(15). <https://doi.org/10.1128/MRA.00049-19>
- Strutz, K., & Stellwagen, N. C. (1998). Do DNA gel electrophoretic mobilities extrapolate to the free-solution mobility of DNA at zero gel concentration? *ELECTROPHORESIS*, 19(5), 635-642. <https://doi.org/10.1002/elps.1150190504>

- Su, L., Fang, H., & Xu, W. (2011). Aminopeptidase N (EC 3.4.11.2) inhibitors (2006 - 2010): A patent review. *Expert Opinion on Therapeutic Patents*, 21(8), 1241-1265. <https://doi.org/10.1517/13543776.2011.587002>
- Sugrue, E., Fraser, N. J., Hopkins, D. H., Carr, P. D., Khurana, J. L., Oakeshott, J. G., Scott, C., & Jackson, C. J. (2015). Evolutionary Expansion of the Amidohydrolase Superfamily in Bacteria in Response to the Synthetic Compounds Molinate and Diuron. *Applied and Environmental Microbiology*, 81(7), 2612-2624. <https://doi.org/10.1128/AEM.04016-14>
- Surachat, K., Sangket, U., Deachamag, P., & Chotigeat, W. (2017). In silico analysis of protein toxin and bacteriocins from *Lactobacillus paracasei* SD1 genome and available online databases. *PLoS ONE*, 12(8). <https://doi.org/10.1371/journal.pone.0183548>
- Szececi, P. B. (1992). The aspartic proteases. *Scandinavian Journal of Clinical and Laboratory Investigation. Supplementum*, 210, 5-22.
- Tang, X., Nakata, Y., Li, H. O., Zhang, M., Gao, H., Fujita, A., Sakatsume, O., Ohta, T., & Yokoyama, K. (1994). The optimization of preparations of competent cells for transformation of *E. coli*. *Nucleic Acids Research*, 22(14), 2857-2858. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC308259/>
- Tapprich, W. E., Reichart, L., Simon, D. M., Duncan, G., McClung, W., Grandgenett, N., & Pauley, M. A. (2021). An instructional definition and assessment rubric for bioinformatics instruction. *Biochemistry and molecular biology education : a bimonthly publication of the International Union of Biochemistry and Molecular Biology*, 49(1), 38-45. <https://doi.org/10.1002/bmb.21361>
- Thermo Fisher Scientific. (2016). *NanoDrop One User Guide-v1.3-sw.pdf*. <https://tools.thermofisher.com/content/sfs/manuals/3091-NanoDrop-One-User-Guide-v1.3-sw-SPANISH.pdf>

- Tschan, S., Mordmüller, B., & Kun, J. F. J. (2011). Threonine peptidases as drug targets against malaria. *Expert Opinion on Therapeutic Targets*, 15(4), 365-378. <https://doi.org/10.1517/14728222.2011.555399>
- Tuteja, R. (2005). Type I signal peptidase: An overview. *Archives of Biochemistry and Biophysics*, 441(2), 107-111. <https://doi.org/10.1016/j.abb.2005.07.013>
- Uritskiy, G., & DiRuggiero, J. (2019). Applying Genome-Resolved Metagenomics to Deconvolute the Halophilic Microbiome. *Genes*, 10(3), E220. <https://doi.org/10.3390/genes10030220>
- Ventossa, A., Gutierrez, M. C., Garcia, M. T., & Ruiz-Berraquero, F. (1989). Classification of “*Chromobacterium marismortui*” in a New Genus, *Chromohalobacter* gen. Nov., as *Chromohalobacter marismortui* comb. Nov., nom. Rev. *International Journal of Systematic and Evolutionary Microbiology*, 39(4), 382-386. <https://doi.org/10.1099/00207713-39-4-382>
- Vogele, L., El Arnaout, T., Bailey, J., Stansfeld, P. J., Boland, C., & Caffrey, M. (2016). Structural basis of lipoprotein signal peptidase II action and inhibition by the antibiotic globomycin. *Science (New York, N.Y.)*, 351(6275), 876-880. <https://doi.org/10.1126/science.aad3747>
- Wajid, B., & Serpedin, E. (2012). Review of general algorithmic features for genome assemblers for next generation sequencers. *Genomics, Proteomics & Bioinformatics*, 10(2), 58-73. <https://doi.org/10.1016/j.gpb.2012.05.006>
- Watson, J. D., Baker, T. A., Bell, S. P., Gann, A., Levine, M., & Losick, R. (2013). *Molecular Biology of the Gene* (Edición: 7). Pearson.
- Whiteley, A. S., Jenkins, S., Waite, I., Kresoje, N., Payne, H., Mullan, B., Allcock, R., & O'Donnell, A. (2012). Microbial 16S rRNA Ion Tag and community metagenome

- sequencing using the Ion Torrent (PGM) Platform. *Journal of Microbiological Methods*, 91(1), 80-88. <https://doi.org/Escobar>
- Wilk, P., Wałtor, E., & Weiss, M. S. (2021). Prolidase—A protein with many faces. *Biochimie*, 183, 3-12. <https://doi.org/10.1016/j.biochi.2020.09.017>
- Wilkesman, J. (2017). Cysteine Protease Zymography: Brief Review. *Methods in Molecular Biology (Clifton, N.J.)*, 1626, 25-31. https://doi.org/10.1007/978-1-4939-7111-4_3
- Wrase, R., Scott, H., Hilgenfeld, R., & Hansen, G. (2011). The Legionella HtrA homologue DegQ is a self-compartmentizing protease that forms large 12-meric assemblies. *Proceedings of the National Academy of Sciences of the United States of America*, 108(26), 10490-10495. <https://doi.org/10.1073/pnas.1101084108>
- Yamaguchi, R., Arakawa, T., Tokunaga, H., Ishibashi, M., & Tokunaga, M. (2012). Halophilic properties of metal binding protein characterized by high histidine content from *Chromohalobacter salexigens* DSM3043. *The Protein Journal*, 31(2), 175-183. <https://doi.org/10.1007/s10930-011-9390-5>
- Yang, J., & Zhang, Y. (2015). I-TASSER server: New development for protein structure and function predictions. *Nucleic Acids Research*, 43(Web Server issue), W174-W181. <https://doi.org/10.1093/nar/gkv342>
- Yaron, A., & Mlynar, D. (1968). Aminopeptidase-P. *Biochemical and Biophysical Research Communications*, 32(4), 658-663. [https://doi.org/10.1016/0006-291X\(68\)90289-1](https://doi.org/10.1016/0006-291X(68)90289-1)
- Yokoyama, R., Kawasaki, H., & Hirano, H. (2006). Identification of yeast aspartyl aminopeptidase gene by purifying and characterizing its product from yeast cells. *The FEBS Journal*, 273(1), 192-198. <https://doi.org/10.1111/j.1742-4658.2005.05057.x>

- Yoshitani, K., Hizukuri, Y., & Akiyama, Y. (2019). An in vivo protease activity assay for investigating the functions of the Escherichia coli membrane protease HtpX. *FEBS Letters*, 593(8), 842-851. <https://doi.org/10.1002/1873-3468.13368>
- Zaccai, G. (2004). The effect of water on protein dynamics. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 359(1448), 1269-1328. <https://doi.org/10.1098/rstb.2004.1503>
- Zhang, S., Cai, X., Luo, X., Wang, S., Guo, A., Hou, J., & Wu, R. (2018). Molecular cloning and characterization of leucine aminopeptidase gene from Taenia pisiformis. *Experimental Parasitology*, 186, 1-9. <https://doi.org/10.1016/j.exppara.2018.01.012>
- Zhou, C., Qin, H., Chen, X., Zhang, Y., Xue, Y., & Ma, Y. (2018). A novel alkaline protease from alkaliphilic Idiomarina sp. C9-1 with potential application for eco-friendly enzymatic dehairing in the leather industry. *Scientific Reports*, 8(1), 16467. <https://doi.org/10.1038/s41598-018-34416-5>

ANEXOS

ANEXO 1.

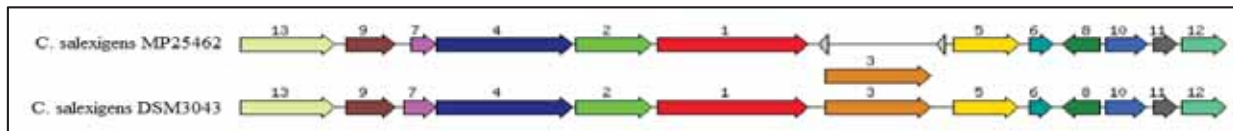
Fichas que describen a las 36 enzimas proteolíticas encontradas en *C. salexigens* MP25462.

FICHA 1

EC 3.4.21.53 - Proteasa Lon dependiente de ATP Tipo I

Es una enzima multifacética, ya que cumple diferentes funciones, en humanos participa en la regulación de la respuesta al estrés oxidativo, al choque térmico y en la regulación de la mitofagia, en el mantenimiento del ADN mitocondrial. Además, su actividad proteolítica puede regular varias vías bioquímicas como el ciclo del ácido tricarboxílico, la fosforilación oxidativa, la biosíntesis de esteroides y hemo y la producción de glutamina. Es por ello, que esta enzima es bien conservada a lo largo de la evolución y cualquier mutación en su gen desencadena enfermedades graves como síndromes poco común caracterizado por anomalías cerebrales, oculares, dentales, auriculares y esqueléticas (Gibellini *et al.*, 2020)

POSICIÓN INTERGENICA:



Donde: **1:** Proteasa Lon dependiente de ATP, **2,6,7:** Proteína hipotética, **3:** Dihidroxiácido deshidratasa. **4:** Histidina quinasa híbrida multisensor, **5:** Oxidoreductasa no caracterizada **8:** Proteína Nudix, **9:** Flavodoxina reductasa, **10:** Hidrolasa putativa, **11:** Proteína de choque térmico, **12:** Chaperona de 33Kda, **13:** sistema de transporte no caracterizado.

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

Descripción	Nombre científico	Cubierta de consulta	Por. Ident	Acc. Len	Adhesión
<input checked="" type="checkbox"/> endocapósida La [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100,00%	815	WP_179152879.1
<input checked="" type="checkbox"/> endocapósida La [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,88%	815	WP_011507958.1
<input checked="" type="checkbox"/> endocapósida La [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,51%	815	WP_110052216.1
<input checked="" type="checkbox"/> endocapósida La [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,39%	815	WP_128689943.1
<input checked="" type="checkbox"/> endocapósida La [Chromohalobacter canadensis]	Chromohalobacter canadensis	99%	91,98%	818	WP_097021895.1

SECUENCIA DE AMINOÁCIDOS:

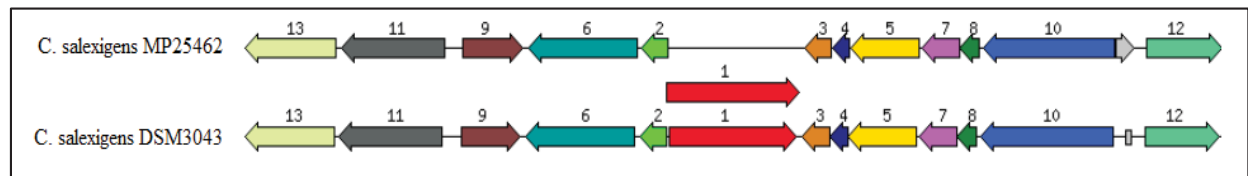
```
> fig | 6666666.597456.peg.1906 Proteasa La dependiente de ATP (EC 3.4.21.53) Tipo I [Chromohalobacter salexigens MP25462]
MSEHDQDQPRDFDAPRGGSPRNTDDTLSELDGSTSHGGAVVPTQEYLPRIYLLPIHNRPFPPAQVQPLVIHRERW
QETMDRVDNTPHHSVGVAVYVG DAGVDELGPDDFPEIGTAVKVHRTQVEDQQIQFIAQGVERRFRIVRWLSKPPY
LVEVSYPKIEIASDEEARAYAMAMINGIKELLPINPLYGEELKHYLNRFSPHEPSPLTDFAAAITSAKGGELQEV
ETLPVMSRMQKVLPLLRKEIEVAQLQSEISEQVNAQMQRQREFFLREQLKVIQRELGISKDDRENDVDTFRERLE
NLOVPEHVMERIEDELGKLSVLETGSPEYGTTRNYLDWLTSLPWGITSEDOLDLPRARKVLDRDHDGLTDVKERI
```

FICHA 2

EC 3.4.21.83 - Proteasa II u oligopeptidasa B

No se encuentra en ninguno de los genomas de mamíferos disponibles hasta la fecha, pero si esta presente en bacterias Gram negativas, protozoos y plantas. El papel fisiológico de esta enzima aún no se ha dilucidado totalmente, sin embargo, la proteasa se ha identificado como un importante factor de virulencia y agente terapéutico en la tripanosomosis animal, donde se pretende desarrollar fármacos contra la enfermedad de Chagas, la leishmaniasis y la tripanosomiasis africana. (Coetzer *et al.*, 2008; Motta *et al.*, 2019)

POSICIÓN INTERGENICA EN COMPARACIÓN CON *C. salexigens* DSM3045:



Donde: **1**: Proteasa II **2,12**: Proteína hipotética, **3**:Proteína YcgN, **4**:Proteína YcgL, **5**: Ribonucleasa, **6**: Factor de Von Willebrand Tipo A, **7**:Proteína de recombinación RecR, **8**: Proteína hipotética que coexiste con RecR **9**:Cisteina sintasa, **10**:Subunidades de ADN polimerasa III gamma y tau, **11**:Acido graso de cadena larga – CoA ligasa, **13**: ARN Helicasa HrpA dependiente de ATP

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

Descripción	Nombre científico	Cubierta de consulta	Por. Ident	Acc. Len	Adhesión
<input checked="" type="checkbox"/> trnll oligopeptidasa serina peptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	96%	100.00%	693	WP_179152193.1
<input checked="" type="checkbox"/> trnll oligopeptidasa serina peptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	96%	99.99%	693	WP_110062374.1
<input checked="" type="checkbox"/> trnll oligopeptidasa serina peptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	96%	98.55%	693	MR25877041.1
<input checked="" type="checkbox"/> trnll oligopeptidasa serina peptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	96%	98.27%	693	WP_120690470.1
<input checked="" type="checkbox"/> trnll oligopeptidasa serina peptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	96%	98.12%	693	WP_011506763.1

SECUENCIA DE AMINOÁCIDOS:

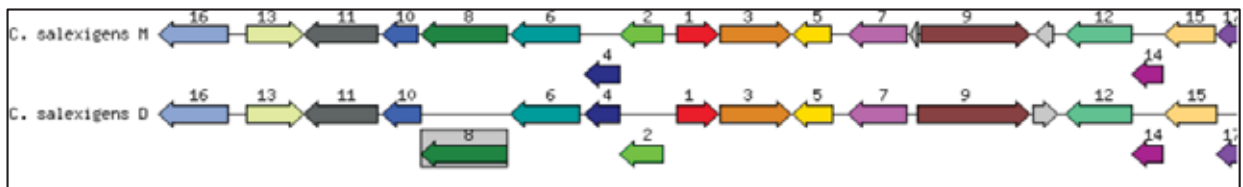
```
> fig | 6666666.597456.peg.778 Proteasa II (EC 3.4.21.83) [Chromohalobacter salexigens MP25462]
MWQGIRNQHRVLSRSIFPRHGTPLHMKAQPHIPALHDATQLSPVARFQRKEDPDWHWLENRDDPEVVAFLQANQEFQSQWSAP
LAPLVEALYSGLHARRELAVEGLGTPLDHYTYWSETARDADYPVWWRHPNQVPSRRECVDLDQARASEQAFMELGDIAPDEHW
LAWTEDTSGDEYFTLFRHALPDGTPRRLLDIDGPELCSAEDNRTLFFTRYDDTQRPDVSVWRLDIESGETTCVVFRETDPFEWVG
VGKTRSREWLVLLETASKDTSCHLVPAAPHVPPRCVREIRKIGIEYALEHRPGHFYILHNQDAPHFRDLVADETAPDTWHPLV
GHDAQLTLESIDAFAGLVI TERDHREAQVHLSVLDLDAVPVTRRRRLPLPEAPCSLMLGDI PDFRTRRLRLHHEESFTLPSRWI
EHDLDSDARRLLKVQPVYGDLPPELVCRRVWAEAHGGERIPVSVVARDLWEQGPMTLLYGYGAYGKVLDPWFSVARLELL
SRGVAFVAHVRRGGDRGEPWYLAGKLEHKENSFRDFLAARHALVEHGIADGERIAAYGASAGLLVSASLNLDPTAFCAAVL
DVPFVDVLRMTMENPDLPLTTAEYSEWGNPSEPEAHRIRDYSPDLNLVARPYPTLFLQGSWHDSDRVPYWEPAKLYARLTELVA
QLPAAERRPIMLRDMDAAGHGASGRFKAWHDNARQDAFILWALGLAETAAP
```


FICHA 3

EC 3.4.21.88 Represor de respuesta SOS y proteasa LexA

En la mayoría de las bacterias, el eje RecA / LexA es el gen de la respuesta al daño del ADN bacteriano (SOS), por ello es un objetivo farmacológico prometedor, aunque no tradicional. La respuesta SOS se inicia tras el estrés genotóxico, cuando RecA, un sensor de daño del ADN, induce a LexA, el represor SOS, a someterse a autoproteólisis, desreprimiendo así los genes posteriores que pueden mediar la reparación del ADN y acelerar la mutagénesis . (Mo *et al.*, 2018)

POSICIÓN INTERGENICA:



Donde: **1:** Represor de respuesta SOS y proteasa LexA, **2:** regulador transcripcional previsto para degradación de acidos grasos FadQ, familia TetR, **3:** ADN polimerasa IV (EC 2.7.7.7) **4:** Glutamato sintasa NADPH de cadena grande EC 1.4.1.13 **5:** Proteína de función desconocida **6:** beta N-acetilglucosaminidasa (EC 3.2.1.52) **7:** formiltetrahidrofolato deformilasa (EC 3.5.1.10) **8:** canal mecanosensible de pequeña conductancia **9:** proteína similar a la betalactamasa **10:** Hipoxantina – guanina fosforribosiltransferasa (EC 2.4.2.8). **11:** enoil-CoA hridratasa/isomerasa **12:** proteína de biosíntesis del cofactor de molibdeno MoaA. **13:** Proteína Hipotetica **14:** proteína de biosíntesis del cofactor de molibdeno MoaE **15:** proteína de biosíntesis del cofactor de molibdeno MoaC / proteína de biosíntesis del cofactor de molibdeno MoaD **16:** Fosfoenolpiruvato sintasa (EC 2.7.9.2)

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

Descripción	Nombre científico	Cubierta de consulta	Por. Ident.	Acc. Len	Adhesión
<input checked="" type="checkbox"/> represor transcripcional LexA [Chromohalobacter salexigens]	Chromohalobacter salexigens	90%	100,00%	212	WP_035412251.1
<input checked="" type="checkbox"/> represor transcripcional LexA [Chromohalobacter salexigens]	Chromohalobacter salexigens	90%	99,48%	212	WP_011507307.1
<input checked="" type="checkbox"/> represor transcripcional LexA [Chromohalobacter sp.]	Chromohalobacter sp.	90%	98,96%	212	NQY44231.1
<input checked="" type="checkbox"/> represor transcripcional LexA [Chromohalobacter salexigens]	Chromohalobacter salexigens	90%	98,96%	212	WP_110061178.1
<input checked="" type="checkbox"/> represor transcripcional LexA [Chromohalobacter japonicus]	Chromohalobacter japonicus	90%	93,78%	212	WP_040241548.1

SECUENCIA DE AMINOÁCIDOS:

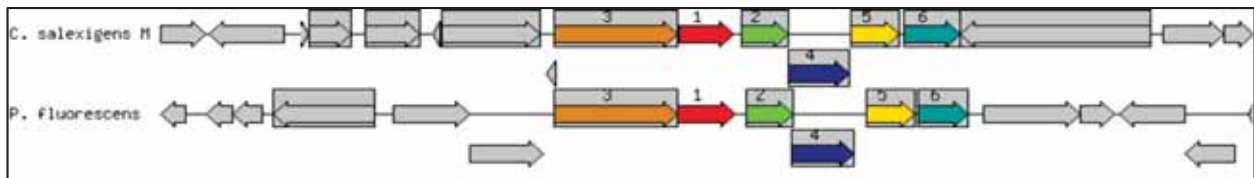
```
> fig | 6666666.597456.peg.2155 Represor de respuesta SOS y proteasa LexA (EC 3.4.21.88) [Chromohalobacter salexigens MP25462]
MSRSLTSRQQHVDFIVKTMNELGYPPTRAEIARALGFRSPNAAEEHLRALDRKGVIRMIPGTSRGIRLTSEEDAAATSTDGL
PVVGEVAAGSPIIAAAHIDRHCLAPDYFTPRADYLLRVRGLSMKDAGILDGDL LAVHRTQQIRDGQIVVARLDDEVTVKRFK
RQGHVWLVAENADFAPIEVDLRHQDLEIEGLGVGVIRGGGNGLH
```

FICHA 4

EC 3.4.21.89 Peptidasa de señal I

Se han encontrado peptidasas señal de tipo I en bacterias, arqueas, hongos, plantas y animales. Las peptidasas señal de tipo I utiliza un mecanismo de díada catalítica de serina / lisina y es la peptidasa señal principal en la mayoría de las bacterias, son enzimas indispensables, que catalizan la escisión de las secuencias del péptido señal aminoterminal de las preproteínas, que se traslocan a través de las membranas biológicas. (Paetzel, 2019; Tuteja, 2005)

POSICIÓN INTERGENICA:



Donde: **1:** Peptidasa señal I. **2:** Ribonucleasa III. **3:** Factor de elongación de translocación LepA. **4:** Era de las proteínas de unión a GTP. **5:** Recombinación y reparación de ADN proteína RecO. **6:** Pridoxina 5-fosfato sintasa (EC 2.6.99.2)

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

Descripción	Nombre científico	Cubierta de consulta	Por. Ident.	Acc. Len.	Adhesión
<input checked="" type="checkbox"/> peptidasa señal I [Chromohalobacter salexigens]	Chromohalobacter salexigens	99%	100.00%	267	WP_011509929.1
<input checked="" type="checkbox"/> peptidasa señal I [Chromohalobacter salexigens]	Chromohalobacter salexigens	99%	99.63%	267	WP_110061695.1
<input checked="" type="checkbox"/> peptidasa señal I [Chromohalobacter salexigens]	Chromohalobacter salexigens	99%	87.64%	267	WP_178996214.1
<input checked="" type="checkbox"/> peptidasa señal I [Chromohalobacter utronica]	Chromohalobacter utronica	99%	96.99%	267	WP_040243620.1
<input checked="" type="checkbox"/> peptidasa señal I [Chromohalobacter canadensis]	Chromohalobacter canadensis	99%	87.64%	267	WP_097032284.1

SECUENCIA DE AMINOÁCIDOS:

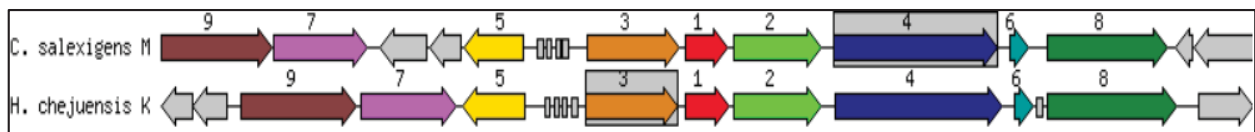
```
> fig | 6666666.597456.peg.399 Peptidasa de señal I (EC 3.4.21.89)
[Chromohalobacter salexigens MP25462]
MDFSLLLVLAVAITGIVWLLDIVWRRARRARLATAESQVDGELDAATRGLAKEPWPVDYARSFFPVLIVLVLRSFVVEPF
QIPSGSMRPTLKVGDIFLVNKFYGLRPLVINTEILDGEPERGDIMVFRFPDDPSVNF I KRVVGLPGDRIRYEDKQLYVNGQ
PVAKSVTDDDADDAPGERQFEERLGDVAHA I YNNPQDPGPQMREVVVDPDGQYFMMGDNRDHSNDSRYWGFVPEENIVGEAFV
WMHWDGGLPSFSSVRLLE
```

FICHA 5

EC 3.4.21.92 Proteasa de Clp dependiente de ATP

La proteasa caseinolítica mitocondrial P (ClpP) juega un papel central en el control de la calidad de la proteína mitocondrial al degradar las proteínas mal plegadas. A demás se demostró, que la hiperactivación de la proteasa mata selectivamente las células cancerosas, mediante la degradación selectiva de sus sustratos de proteínas de la cadena respiratoria y altera la estructura y función mitocondrial, mientras que no afecta a las células no malignas. (Ishizawa *et al.*, 2019)

POSICIÓN INTERGENICA:



Donde: **1:** Subunidad proteolítica de proteasas Clp dependiente de ATP (EC 3.4.21.92) **2:** Proteasas de Clp dependiente de ATP subunidad de unión a ATP ClpX **3:** Factor desencadenante de la división celular (EC 5.2.1.8) **4:** PROTEASAS La dependiente de ATP (EC 3.4.21.53) Tipo I **5:** Metilentetrahidrofolato deshidrogenasa (NADP+) / Metilentetrahidrofolato ciclohidrolasa (EC 3.5.4.9) **6:** Proteina de unión al ADN HU-beta **7:** Cisteinil-t RNA sintetasa (EC6.1.1.16) **8:** Peptidil-propil cis-trans isomerasa PpiD (EC 5.2.1.8) **9:** GLUTAMIL-tRNA sintetasa (ec 6.1.1.18)

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

	Descripción	Nombre científico	Cubierta de consulta	Por. Ident.	Acc. Len.	Adhesión
<input checked="" type="checkbox"/>	ClpP subunidad proteolítica de endopeptidasa Clp dependiente de ATP [Chromohalobacter salexigens]	Chromohalobacter salexigens	99%	100.00%	206	WP_011507343.1
<input checked="" type="checkbox"/>	ClpP subunidad proteolítica de endopeptidasa Clp dependiente de ATP [Chromohalobacter salexigens]	Chromohalobacter salexigens	99%	99.51%	206	WP_224113267.1
<input checked="" type="checkbox"/>	ClpP subunidad proteolítica de endopeptidasa Clp dependiente de ATP [Chromohalobacter]	Chromohalobacter	99%	96.60%	206	WP_040241459.1
<input checked="" type="checkbox"/>	ClpP subunidad proteolítica de endopeptidasa Clp dependiente de ATP [Chromohalobacter marismortui]	Chromohalobacter marismortui	99%	97.00%	206	WP_133696496.1
<input checked="" type="checkbox"/>	Subunidad proteolítica de endopeptidasa Clp dependiente de ATP ClpP [Chromohalobacter canadensis]	Chromohalobacter canadensis	99%	96.12%	206	WP_097024142.1

SECUENCIA DE AMINOÁCIDOS:

```
> fig | 6666666.597456.peg.2181 Subunidad proteolítica de proteasa Clp dependiente de ATP (EC 3.4.21.92) [Chromohalobacter salexigens MP25462]
```

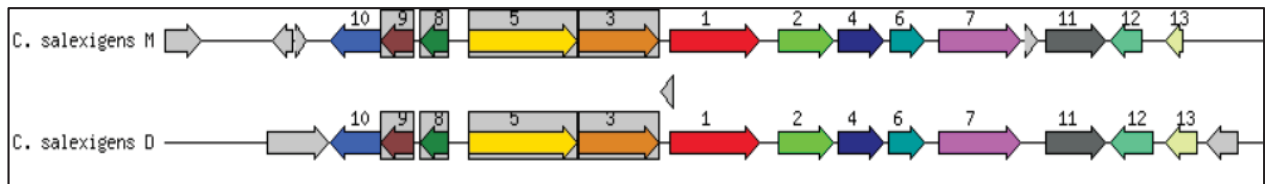
```
MGNEFDISNAGGLVPMVVEQNARGERAYDIYSRLKERVIFLIGPVEDYMANLVVAQMLFLESENPKDIHLYINSPGG  
SVTAGMSIYDTMQFIKPDVSTVCVGGQAASMGALLLAGGAAGKRYCLPHSRMMIHQPLGGYQQAADIEIHTKEILNIRQ  
QLNEILAKHTGQDAETVARDTDRDNFMNGTQAVEYGLIDAMLDKRPVS
```

FICHA 6

EC 3.4.21.102 Proteasa carboxilo-terminal

Las proteasas de procesamiento de carboxilo (C) -terminal (CTP) participan en la proteólisis protectora y reguladora en bacterias. El dominio PDZ es fundamental para la actividad de los CTP, pero desempeña funciones reguladoras inherentemente diferentes. Por ejemplo, el dominio PDZ inhibe la actividad de la proteasa de señalización CtpB bloqueando el sitio activo, pero es necesario para la activación de Prc (o Tsp), una proteasa específica de la cola que degrada las proteínas marcadas con SsrA. (Chueh *et al.*, 2019).

POSICIÓN INTERGENICA:



Donde: **1:** proteasas carboxi-terminal (EC 3.4.21.102) **2:** regulador transcripcional, familia AraC **3:** peptidasa M25B **4:** proteína exportadora de lisina (LYSE / YGGA) **5:** Fosfoglicerato mutasa independiente de 2,3-bisfosfoglicerato (EC 5.4.2.1) **6:** Regulador transcripcional, Familia TetR **7:** Superfamilia facilitadora principal MFS_1 **8:** Sulfurtransferasa relacionada con Rhodanesa **9:** Proteína de exportación de proteína chaperona del citoplasma (SecB, mantiene la proteína para ser exportada en estado desplegado) **10:** Metiltransferasa E de subunidad pequeña de ARN ribosómico (Ec 2.1.1.-) SSU ARNr m3u1498 **11:** permeasa de la superfamilia de transportadores de fármacos / metabolitos (DMT) **12:** proteína exportadora de lisina (LYSE / YGGA) **13:** Proteína de la familia OSmC / Ohr

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

Descripción	Nombre científico	Cubierta de consulta	Por. Ident	Acc. Len	Adhesión
<input checked="" type="checkbox"/> Peptidasa de la familia S41 [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100,00%	440	WP_179152435.1
<input checked="" type="checkbox"/> proteasa carboxilo-terminal [Chromohalobacter salexigens DSM 3043]	Chromohalobacter salexigens DSM 3043	100%	99,77%	440	ABE57409.1
<input checked="" type="checkbox"/> Peptidasa de la familia S41 [Chromohalobacter sp.]	Chromohalobacter sp.	100%	99,77%	440	NGY45556.1
<input checked="" type="checkbox"/> Peptidasa de la familia S41 [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,09%	440	WP_224114096.1
<input checked="" type="checkbox"/> Peptidasa de la familia S41 [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	98,86%	440	WP_110090985.1

SECUENCIA DE AMINOÁCIDOS:

```
> fig | 6666666.597456.pcg.2541 Proteasa carboxilo-terminal (EC 3.4.21.102)
[Chromohalobacter salexigens MP25462]
```

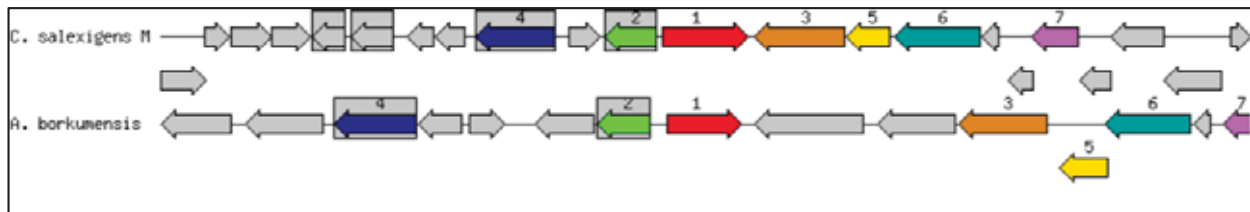
```
MTASQSRSPRLVVRHCLHLGMALAI GALTLP LPAHAQQPAGDDALPVEDVQTFAEVFERIKRAYVDEVDTTLMRNAMRG
MLGELDPHSAYLDAESFEALRETTEGEFSGVGIIEVGMQEGQLTIIAPIDDS PAARAGLQAQDAILRIDDTPTESLSLQEA
VEMMRGDEGEEIRLTILREGE EAPREVTLTRETIRTDSVKHEMLSPGYGLRISQFQSRTGEQARDAIAALREEGDGNLK
GLVLDLRNNPGGVLD SAVDVADLFLDSGLIVYTEGR LADSDMRFSASPQTSAPDVP MVVLINGGSASAAEIVAGALQDQQ
RAVLMTGESFGKGSVQQLPLNNGDGLKLT TALYTPDGRS IQAQGIAPDVEVVRGRLEVAEATGLSIRESDLENHLRNI
NGERERTERESSLAESDYQLGEALNLLKALNVL PRAQSGN
```

FICHA 7

EC 3.4.21 Familia DegQ

Se sabe poco sobre DegQ. Una identidad de secuencia de aminoácidos del 59% con DegP y la presencia de una secuencia señal típica indican que DegQ es una segunda chaperona-proteasa principal del periplasma, ya que por sí solo parece ser responsable de mantener la homeostasis de las proteínas en el periplasma de muchos procariontes. (Wrase *et al.*, 2011)

POSICIÓN INTERGENICA:



Donde: **1:** sensor de tensión de la membrana externa de proteasa DegQ, Serina proteasa **2:** Proteína hipotética Ybgl **3:** histidinol deshidrogenasa (EC 1.1.1.23) **4:** ATPasa, familia AFG1 **5:** ATP fosforribosiltransferasa (EC 2.4.2.17) **6:** UDP-N-acetilglucosamina 1-carboxiviniltransferasa (EC 2.5.1.7) **7:** Transportador ABC no caracterizado, componente auxiliar YrbC.

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

Descripción	Nombre científico	Cubierta de consulta	Por. Ident	Acc. Len	Adhesión
<input checked="" type="checkbox"/> ¿Familia serina endopeptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100.00%	412	WP_011507508.1
<input checked="" type="checkbox"/> ¿Familia serina endopeptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99.76%	412	WP_224113217.1
<input checked="" type="checkbox"/> ¿Familia serina endopeptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99.51%	412	WP_128880923.1
<input checked="" type="checkbox"/> ¿La familia serina endopeptidasa [Chromohalobacter sp.]	Chromohalobacter sp.	100%	98.70%	412	NGY46089.1
<input checked="" type="checkbox"/> ¿Familia serina endopeptidasa [Chromohalobacter japonicus]	Chromohalobacter japonicus	98%	89.63%	405	WP_075369828.1

SECUENCIA DE AMINOÁCIDOS:

```
> fig | 6666666.597456.peg.2010 Proteasa DegQ del sensor de tensión de la membrana externa, serina proteasa [Chromohalobacter salexigens MP25462]
```

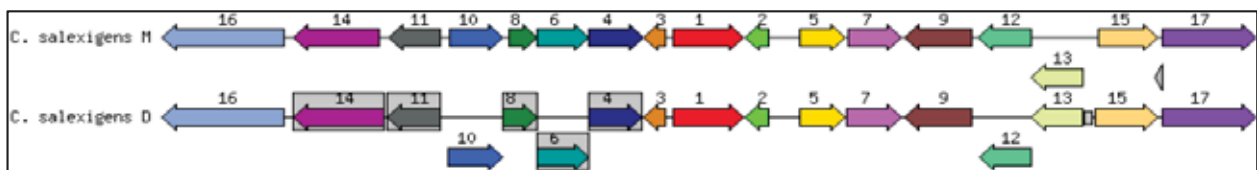
```
MRRSLMSLVWPTISGILLAIIVLLNAFPQLLGGRNADDTVTPVPDTSAVRAIAERTDTAARAPEVVEAAPLDRAQGPASAYA  
TAVDKAAPAVVNIYSSRMVDPSEHPLMSDPFFEQFFGKDMPPQRQMLSSLGSGVIVSPEGYVLTNNHVIRNADQIQVALR  
DGRETLAEVVGTDPESDLAVLKIIPVDNLPVIELSDSEQVAVGDVSLAIGNPFGVGQTVTMGIISATGRNHLGLNAYEDFI  
QTDAAINPGNSGGALVNAEGALVGINTAIFSRSGSQGIGFAIPANLAHQVLDQIVAHGRVIRGWLIDVQAMTPDLATS  
FGLKTLKGVVIANVVPGGPGEKAGLQPGDVLMSVNGKIVDAREAMADIAEISPGTSLPVTIVRDGEKREVTLTVGERPQ  
AAQRQPTAPSSSE
```

FICHA 8

EC 3.4.21 Proteasa Sohb, Familia U7

Se identifico en *E. coli*, aseverando que es posible que el gen sohB codifique una proteasa periplásmica no descubierta previamente en *E. coli* que, cuando se sobreexpresa, puede compensar parcialmente la función de la proteína HtrA faltante (gen htrA codifica una proteasa periplásmica requerida para la viabilidad bacteriana solo a alta temperatura, es decir, por encima de 39 °C) (Baird *et al.*, 1991)

POSICIÓN INTERGENICA:



Donde: **1:** Peptidasa Sohb familia U7, **2:** regulador transcripcional familia ArsR **3:** Proteina de la familia de transferencia de esteroides SCP-2 **4:** pirofosfatasa NADH (EC 3.6.1.22) **5:** Tnrasportador putativo preQ0 **6:** Subunidad épsilon de la ADN polimerasa III (EC 2.7.7.7) **7:** HIDROLASA ALFA / BETA **8:** Ribonucleasa HI (EC 3.1.26.4) **9:** transportador de hidroximetilpirimidina **10:** metiltransferasa dependiente de SAM (EC 2.1.1.-) **11:** Hidroxiacilglutation hidrolasa (EC 3.1.2.6) **12:** transportador de hidroximetilperimidina ABC, componente transmembrana **13:** Transportador ABC de hierro férrico, proteína de unión a ATP **14:** Precursor de transglicosilasa D de mureina lítica unida a membrana (EC 3.2.1.-) **15:** proteína hipotética **16:** proteína extracelular de unión a solutos, familia 5 **17:** intercambiador de sodio / hidrogeno

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

Descripción	Nombre científico	Cubierta de consulta	Por. Ident	Acc. Len	Adhesión
<input checked="" type="checkbox"/> proteasa Sohb [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100.00%	349	WP_128690773.1
<input checked="" type="checkbox"/> proteasa Sohb [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99.71%	349	WP_011507238.1
<input checked="" type="checkbox"/> proteasa Sohb [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99.71%	367	WP_218781831.1
<input checked="" type="checkbox"/> proteasa Sohb [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99.43%	349	WP_224114342.1
<input checked="" type="checkbox"/> proteasa Sohb [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99.43%	349	WP_110061367.1

SECUENCIA DE AMINOÁCIDOS:

```
> fig | 6666666.597456.peg.2289 Peptidasa, familia U7 [Chromohalobacter salexigens MP25462]
```

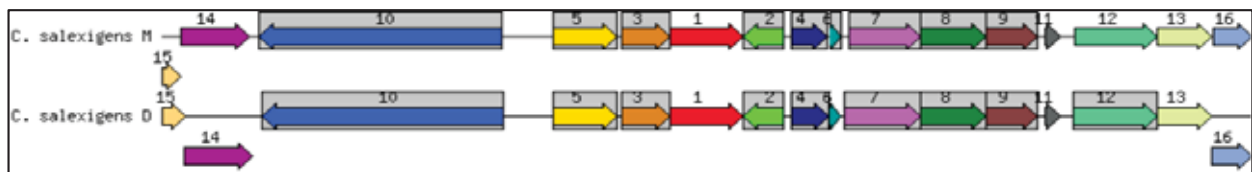
```
MTEWIAEYALFAAKLATLVVAVGVVLVGLVVTLKQRASGQVEGNRLRIVALHERQRRRRESRLAGLPAAQRRRAQ  
KALRKEEKARRKKGKGEAQEAARPRVWVLDGDLKASRTPQLAEQVSLLLGELQDDDEVVVRTSGGGLVHAYG  
LASAQLDRLRDAGVRLTVCVDKVAASGGYMMACCAHRLIAAPFAVIGSIGVVAQVPNVHRLKHKHDIDVELLTAG  
RYKRTLTVLGENTEEGRAKFLDLRETHDLFKRYVGERPALDIEKVATGEIWIYGRQALDDGLIDEVNTSDGYLAA  
RMEQARVLQVVLEPRHSLMQRIGVVGASQGIERLSDRVMERLEESGWQRR
```


FICHA 9

EC 3.4.21 Peptido señal peptidasa SppA

La peptidasa de péptido señal (SPP) es una enzima unida a la membrana que utiliza un mecanismo de díada catalítica de serina / lisina para escindir los péptidos señal remanentes en la membrana celular y ayuda en la secreción de proteínas. A pesar del papel central de SppA en la secreción de proteínas, no se han realizado estudios centrados en la mejora de la secreción de proteínas a través de la sobreexpresión de SPP. (Cai *et al.*, 2017).

POSICIÓN INTERGENICA:



Donde: 1: serina proteasa periplasmática (clase ClpP) 2: Proteína de la familia Maf/ YceF/ YHdE 3: similar a la fosfoglicolato fosfatasa, agrupada con la subunidad ribosómica grande pseudouridina sintasa C 4: Proteína COG1399, agrupada con proteína ribosómica L32p 5: Seudouridina sintasa C de subunidad grande ribosómica (EC4.2.1.70) 6: proteína ribosómica LSU L32p 7: Fosfato acil-ACP acetiltransferasa PlsX 8: Malonil CoA-acil proteína transportadora transacilasa 9: 3-oxoacil-(proteína transportadora de acilo) reductasa (EC 1.1.1.100) 10: ribonucleasa E (EC 3.1.26.12) 11: proteína de acilo 12: 3-oxoacil-sintasa, KASII (EC 2.3.1.41) 13: aminodesoxicorismato liasa (EC 4.1.3.38) 14: UDP N acetilenoilpiruvilglucosamina reductasa (EC 1.1.1.158) 15: proteína tirosina fosfatasa de bajo peso molecular (EC 3.1.3.48) 16: proteína YceG similar

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

Descripción	Nombre científico	Cubierta de consulta	Por. Ident	Acc. Len	Adhesión
<input checked="" type="checkbox"/> péptido señal peptidasa SppA [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100,00%	357	WP_179153622.1
<input checked="" type="checkbox"/> péptido señal peptidasa SppA [Chromohalobacter sp.]	Chromohalobacter sp.	100%	99,72%	357	NQY44445.1
<input checked="" type="checkbox"/> péptido señal peptidasa SppA [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,44%	357	WP_011506084.1
<input checked="" type="checkbox"/> proteasa-4 [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,16%	358	PWW36273.1
<input checked="" type="checkbox"/> péptido señal peptidasa SppA [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,16%	357	WP_110061690.1

SECUENCIA DE AMINOÁCIDOS:

```
> fig | 6666666.597456.peg.440 Serina proteasas periplasmáticas (clase ClpP) [Chromohalobacter salexigens MP25462]
```

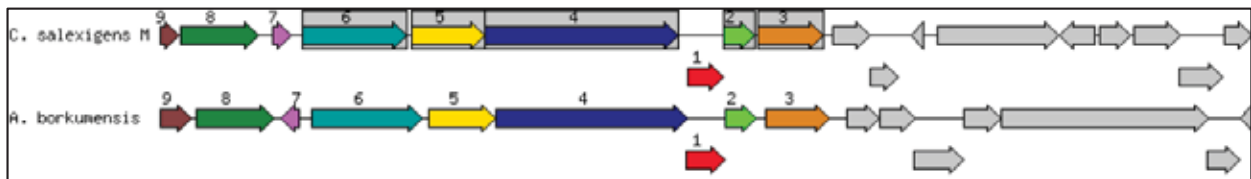
```
MGDDKRDDAWTQGPEVGRGRDEDRRAEEAAAGDPEQLRELRLREQHMLDRWIDGVLVEQRRTRRWKLFRRFVFAALIA  
ASLAMTTYAVLFAGRDMTGAAGPHLGVVEVQGVDAQGEASAERIIIEGLRDAWKAPSSRGVVLHINSPGGSPVQSQRVYDE  
VRRLLAAQGDKPVVAVIEDIGASGAYMASAADDIYASPASLVGSI GVIYSGFGLQEAIKVGVERRVFTSGDNKAFLDPFS  
PIRDEQRAFWQTVLDTTTRQFIDDVKAGRGDRLDDDPRFLFSGLIWTGQALGLGLVDGLSSLDALSRRERFGGVPRRDYTPA  
LDPLSRLSRQLGRVAAEWTGMAHHASPVRYQVQ
```

FICHA 10

EC 3.4.23.36 Peptidasa señal II

Las lipoproteínas se sintetizan con una péptido señal que las fija a la membrana citoplasmática con el dominio de lipoproteínas en el periplasma o fuera de la célula. El procesamiento postraduccional requiere una señal peptidasa II (LspA) que elimina el péptido señal. Se sintetizan en una forma precursora que requiere procesamiento por enzimas que son esenciales en muchas bacterias patógenas pero no tienen equivalente en los humanos, lo que las convierte en posibles dianas farmacológicas. (Vogeley *et al.*, 2016)

POSICIÓN INTERGENICA:



Donde: **1:** peptidasa señal de lipoproteína **2:** peptidil-propil cis-trans isomerasa SlpA de tipo FKBP (EC 5.2.1.8) **3:** 4-hidroxi-3-metilbut-2-enil difosfato reductasa (EC 1.17.1.2) **4:** isoleucil-tRNA sintetasa **5:** robodflavina quinasa (EC 2.7.1.26) / FMN adenililtransferasa (EC 2.7.7.2) **6:** propuesto peptidoglicano lípido II flippasa MurJ **7:** Proteína ribosómica SSU S20p **8:** Glutamato 5-quinasa (EC 2.7.2.11) / PUA de dominio C-terminal de unión a ARN **9:** Proteína de unión a GTP Ogb

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

Descripción	Nombre científico	Cubierta de consulta	Por Ident	Acc Len	Adhesión
<input checked="" type="checkbox"/> peptidasa señal de lipoproteínas [Chromohalobacter salexigens DSM 3043]	Chromohalobacter salexigens DSM 3043	98%	100,00%	181	AME57844.1
<input checked="" type="checkbox"/> peptidasa de señal II [Chromohalobacter salexigens]	Chromohalobacter salexigens	98%	99,44%	181	WP_110062095.1
<input checked="" type="checkbox"/> peptidasa de señal II [Chromohalobacter salexigens]	Chromohalobacter salexigens	90%	100,00%	166	WP_148266850.1
<input checked="" type="checkbox"/> Peptidasa señal de lipoproteína [Vriobacillus halodentificans]	Vriobacillus halodentificans	97%	83,05%	181	CDQ37523.1
<input checked="" type="checkbox"/> peptidasa de señal II [Chromohalobacter marismortui]	Chromohalobacter marismortui	97%	83,05%	180	WP_133688373.1

SECUENCIA DE AMINOÁCIDOS:

```
> fig | 6666666.597456.peg.528 Peptidasa señal de lipoproteína (EC 3.4.23.36) [Chromohalobacter salexigens MP25462]
```

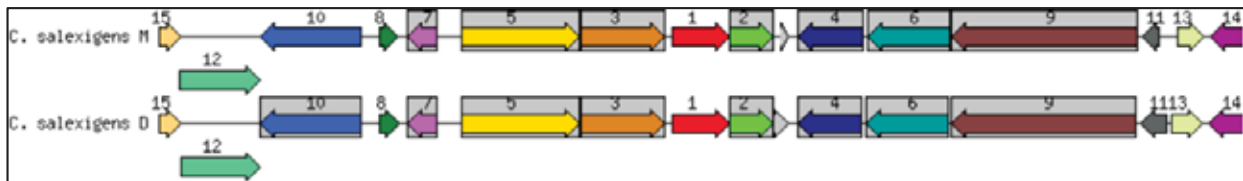
```
MNDASTSEKIAAAPAMRRPLRWLWALLVVVLDLGSKFVASAYLVYGPVEVLPVFNLTLLHNTGAAFSFLAEHA  
GWQRWLFAAIAIGVSIGLTVWLRKHAHERLSAIAIVLVIGGALGNLYDRLVHGYVVDFLSFHWQDYYPFANLAD  
TAITLGAMGLILASLRDDRRNTTQKETPH
```


FICHA 11

EC 3.4.23.43 Peptidasa líder (peptidasa prepilina)

Es responsable de la escisión endopeptídica de los péptidos líderes únicos que caracterizan a los precursores de pilina de tipo IV, así como de proteínas con secuencias líderes homólogas que son componentes esenciales de la vía de secreción general que se encuentra en una variedad de patógenos gramnegativos. (Lory & Strom, 1997).

POSICIÓN INTERGENICA:



Donde: **1**: peptidasa líder (prepilin peptidasa) (EC 3.4.23.43) / N-metiltransferasa (EC 2.1.1.-) **2**: defosfo-CoA quinasa (EC 2.7.1.24) **3**: proteína de ensamblaje fimbrial de tipo IV PilC **4**: proteína 7-mutT mutante **5**: conjunto fimbrial tipo IV, ATPase PilB **6**: Glutamato N-acetiltransferasa (EC 2.3.1.35) / N-acetilglutamato sintasa (EC 2.3.1.1) **7**: Pilin Pila tipo IV **8**: Proteína hipotética **9**: exportación de proteínas del citoplasma SecA ATPasa AN helicasa **10**: proteína de absorción de potasio TrkH **11**: Proteína hipotética **12**: Fosfolípido / glicerol aciltransferasa **13**: proteína y derivados truncados que contienen cinta de Zn, posiblemente de unión a ARN **14**: UDP-3-O-N-acetilglucosamina desacetilasa (EC 3.5.1.108) **15**: Factor de crecimiento C dependiente de la recombinación de ADN.

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

Descripción	Nombre científico	Cubierta de consulta	Por. Ident	Acc. Len	Adhesión
<input checked="" type="checkbox"/> peptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100,00%	280	NWC55929.1
<input checked="" type="checkbox"/> peptidasa 1 de prepilina tipo 4, peptidasa asimétrica, familia MEROPS A24 [Chromohalobacter salexigens DS]	Chromohalobacter salexigens	100%	98,21%	280	ABE59527.1
<input checked="" type="checkbox"/> Peptidasa de la familia A24 [Chromohalobacter salexigens]	Chromohalobacter salexigens	95%	100,00%	268	WP_179153232.1
<input checked="" type="checkbox"/> prepilina peptidasa [Chromohalobacter sp.]	Chromohalobacter sp.	96%	98,88%	268	NQY46124.1
<input checked="" type="checkbox"/> Peptidasa de la familia A24 [Chromohalobacter salexigens]	Chromohalobacter salexigens	95%	98,13%	268	WP_224113310.1

SECUENCIA DE AMINOÁCIDOS:

```
> fig | 6666666.597456.peg.2046 Peptidasa líder (Prepilin peptidasa) (EC 3.4.23.43) / N-metiltransferasa (EC 2.1.1.-) [Chromohalobacter salexigens MP25462]
```

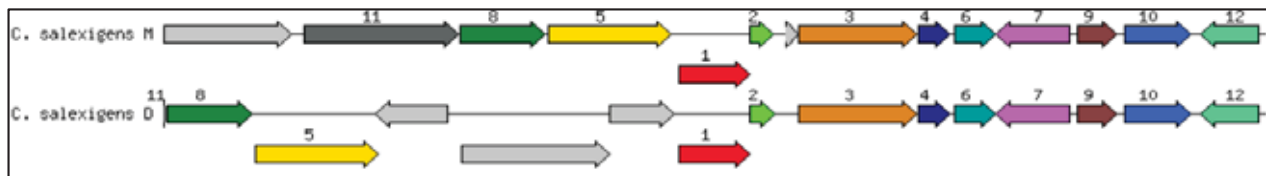
```
MLADIPLPWLAVMAFVVGTTLLGSFLNVVIVRLPVMLMQQWRQEARESLALETEDMPAYNLCTPRSQPCRAPITWH  
DNLPLIGWLKRRGRCAHCQGRISAQYPLVELTAGLLLSVVMVHGATWSALAWGGLCLALLVVAVIDLRTQLLPDI I  
TLPLLWAGLIYQLVFPYWLGNVIGALAGYLLWGFYWLFKWI TGKEGMGYGDFKLLAALGAWFGWTMLPILLILLS  
AGLGAIVIGLLQLTHPRLRGMPLPFGPFLALAGWVAVLFGAPLRLTGL
```

FICHA 12

EC 3.4.24.27 Termolisina, Metalopeptidasa putativa, familia M4

Es una metaloproteína de zinc neutra termoestable identificada originalmente en el caldo de cultivo de *Bacillus thermoproteolyticus*. La termolisina es un representante de la metaloproteína de zinc y un objetivo atractivo en la ingeniería de proteínas para comprender el mecanismo catalítico, la termoestabilidad y la halofilia. La termolisina se utiliza en la industria, especialmente para la síntesis enzimática del éster metílico de N-carbobenzoil L-Asp-L-Phe (ZDFM), un precursor de un edulcorante artificial, el aspartamo. (Inouye *et al.*, 2007)

POSICIÓN INTERGENICA:



Donde: 1: Metaloproteasa putativa 2: proteína hipotética 3: 2-isopropilmalato sintasa (EC 2.3.3.13) 4: proteína hipotética 5: Fosfohidrolasa dependiente de metales 6: proteína hipotética 7: lisil-lisina 2,3-aminomutasa 8: fosfoserina serina (EC 3.1.3.3) 9: factor de elongación de traslación P 10: factor de elongación de traducción P Lys34 lisina transferasa 11: topoisomerasa IV subunidad A (EC 5.99.1.-) 12: fosfatidilserina descarboxilasa (EC 4.1.1.65)

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

	Descripción	Nombre científico	Cubierta de consulta	Por. Ident	Acc. Len	Adhesión
<input checked="" type="checkbox"/>	Metalopeptidasa de la familia M4 (Chromohalobacter salexigens)	Chromohalobacter salexigens	100%	100,00%	349	WP_128681117.1
<input checked="" type="checkbox"/>	Metalopeptidasa de la familia M4 (Chromohalobacter salexigens)	Chromohalobacter salexigens	100%	99,71%	347	WP_224114500.1
<input checked="" type="checkbox"/>	proteólisis de proteolisis (Chromohalobacter salexigens)	Chromohalobacter salexigens	100%	98,56%	347	EW035713.1
<input checked="" type="checkbox"/>	Metalopeptidasa de la familia M4 (Chromohalobacter salexigens)	Chromohalobacter salexigens	100%	98,56%	349	WP_110061998.1
<input checked="" type="checkbox"/>	Metalopeptidasa de la familia M4 (Chromohalobacter salexigens)	Chromohalobacter salexigens	98%	98,25%	348	WP_011507830.1

SECUENCIA DE AMINOÁCIDOS:

```
> fig | 6666666.597456.peg.121 metaloproteasa putativa [Chromohalobacter salexigens MP25462]
```

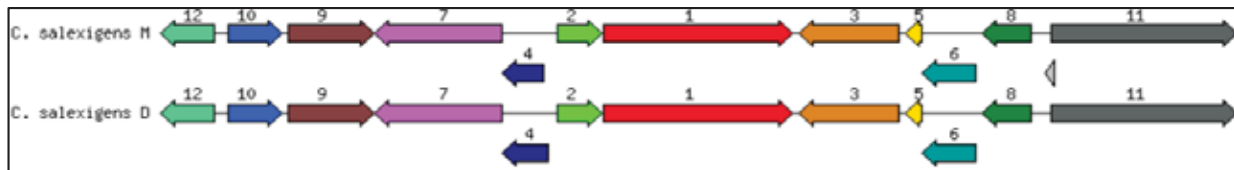
```
MVFPKSPLRGFMPPHVLDRIALQGTERQRRCAQQTLLQADQWFLRLASPPPARDAARAVAGRPDRRIHSADHEQTLPGRLVREEG  
QAAHGDAVDEAYEWLGATYRFYWEVFGGRSDIDDRGMPLIGTVHYGRDYDNFAWNGAQMVFGDGDGLFRRFTAAPVVAHEL  
T  
HGVIERDVGVLVYAGQSGALNESLADVFGVVVKQYHAGQTAQEADWLI GAALLTDRVQGRALRSMEAPGTAYDDPVLGRDPQPGH  
MRDFVDTQADNGGVHINSI P N RAFYLA AVALEAPAWESVAPVWYAAMRDDALSRESDFAAFAALTVAHARRQHGEGRSREARAV  
DDAWREVGVS
```

FICHA 13

EC 3.4.24.55 Precursor de la proteasa III

La proteasa III carece de actividad frente a sustratos de aminopeptidasa, carboxipeptidasa o esterasa, pero degrada rápidamente las proteínas pequeñas. Cuando se utilizan fragmentos de beta-galactosidasa como sustratos para la proteasa III, la enzima degrada preferentemente proteínas con pesos moleculares inferiores a 7.000. (Cheng & Zipser, 1979)

POSICIÓN INTERGENICA:



Donde: 1: precursor de la proteasas III 2: proteína de función desconocida DUF159 3: Exodesoxirribonucleasa I (EC 3.1.11.1) 4: regulador transcripcional 5: Factor de especificidad topológica de división celular MinE 6: Proteína que determina el sitio del tabique MinD 7: 1-acil-sn-glicerol-3-fosfato aciltransferasa (EC 2.3.1.51) 8: Proteína MinC que determina el sitio del tabique 9: hemolisinas y proteínas relacionadas que contienen dominios CBS 10: proteína relacionada con CcsA 11: proteína hipotética 12: partícula de reconocimiento de señal, subunidad Fth SRP54 (TC 3.A.5.1.1)

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

Descripción	Nombre científico	Cubierta de consulta	Porcent. Ident.	Acc. Len.	Adhesión
<input checked="" type="checkbox"/> peptidasa M16 [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100,00%	941	NW056555.1
<input checked="" type="checkbox"/> proteína similar a la peptidasa M16 [Chromohalobacter salexigens DSM 3043]	Chromohalobacter salexigens DSM 3043	100%	99,88%	941	ABE60367.1
<input checked="" type="checkbox"/> proteína de la familia de las insulinasas [Chromohalobacter sp.]	Chromohalobacter sp.	100%	99,26%	941	NQY45333.1
<input checked="" type="checkbox"/> proteína de la familia de las insulinasas [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,15%	941	WP_110062316.1
<input checked="" type="checkbox"/> proteína de la familia de las insulinasas [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,04%	941	WP_128921454.1

SECUENCIA DE AMINOÁCIDOS:

```
> fig | 6666666.597456.pcg.1545 Precursor de la proteasa III (EC 3.4.24.55)
[Chromohalobacter salexigens MP25462]
```

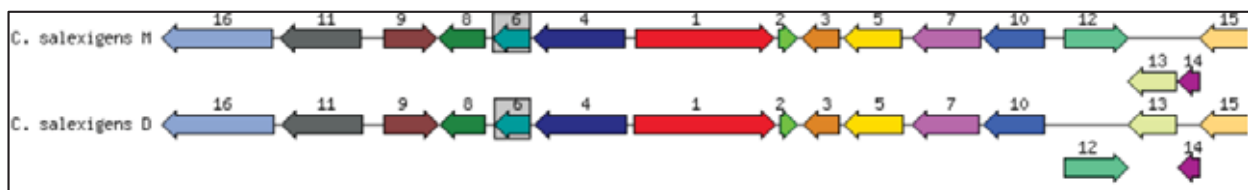
```
MLSSRPRAGRWLLGMCLFLLAVSQTAYASDPVAKVVDPIASPNSRDYRALTLDNGLIILLVSDPEADEAAAAMNVDVGS
SDDPDATPGLAHFLEHMLFLGTDRYPEADAYQNFISAHGGDHNAFTASRDTNYYFDIEPTALPEALDRFSRFFVAPRFNP
EYVERERNAVHSEYQARLRDDGRRINEATDRALNPEHPATRFVAVGSLFTLQGGERSLREKLIDFYESHYGANVMHLTVIG
PQSLDTLES MVRDRFAEIPDRGLTRTPLETPLVTDAELPARLAVKSLSRDREVRFLFPI PDPQQDYRTKPAEYLANLLGH
```

FICHA 14

EC 3.4.24.70 Oligopeptidasa A

La oligopeptidasa A (OpdA) pertenece a la subfamilia M3A de peptidasas bacterianas con propiedades catalíticas y estructurales similares a la timetoligopeptidasa (TOP) y la neurolisina (NEL) de mamíferos. Las tres enzimas tienen cuatro residuos de Tyr conservados en un bucle flexible muy cerca del sitio catalítico. Se afirma que el residuo OpdA Tyr (607) juega un papel importante en la interacción enzima-sustrato y en la actividad hidrolítica (Lorenzon, *et.al.*, 2010)

POSICIÓN INTERGENICA:



Donde: 1: oligopeptidasa A (EC 3.4.24.70) 2: Proteína hipotética 3: precursor de superóxido dismutasa Cu-Zn (EC 1.15.1.1) 4: Simportador de neurotransmisor 5: 5,10-metilentetrahidrofolato reductasa (EC 1.5.1.20) 6: anhidrasa carbonica, familia 3 7: proteína de resistencia de cobalto-zinc-cadmio CzcD 8: proteína hipotética 9: regulador LldR sensible al lactato en enterobacterias, familia GntR 10: Permeasa de la familia de transportadores de fármacos / metabolitos (DMT) 11: L-lactato deshidrogenasa (EC 1.1.2.3) 12: portador de eflujo de auxina 13: regulador transcripcional, familia AraC 14: Transporte de aminoácidos de cadena ramificada 15: Transporte de aminoácidos de cadena ramificada, AzlC 16: D-lactato deshidrogenasa prevista, proteína Fe- S, que contiene FAd /FMN

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

Descripción	Nombre científico	Cubierta de consulta	Por Ident	Acc. Len	Adhesión
<input checked="" type="checkbox"/> oligopeptidasa A [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100,00%	683	WP_179352616.1
<input checked="" type="checkbox"/> oligopeptidasa A [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,71%	683	WP_043557395.1
<input checked="" type="checkbox"/> oligopeptidasa A [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,71%	683	WP_129889770.1
<input checked="" type="checkbox"/> oligopeptidasa A, metaloexopeptidasa, familia MEROPS M03A [Chromohalobacter salexigens DSM 3]	Chromohalobacter salexigens DSM	100%	99,71%	684	ABE57761.1
<input checked="" type="checkbox"/> oligopeptidasa A [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,71%	683	WP_224113898.1

SECUENCIA DE AMINOÁCIDOS:

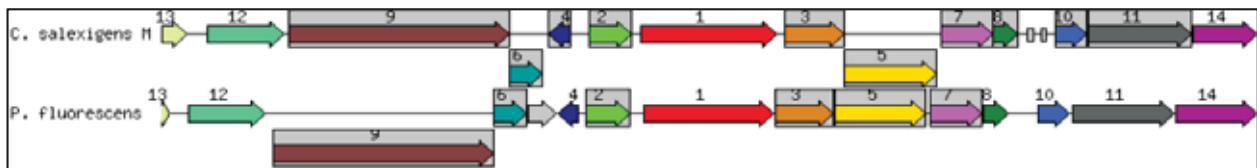
```
> fig | 6666666.597456.peg.2750 Oligopeptidasa A (EC 3.4.24.70)
[Chromohalobacter salexigens
MP25462]MSRNPLLESHVLPFFDDIQPEHVVPAIEQLLAENRRDIEALAQQSQISWESLAAPLEALNDRLSQAWSPVS
HLNSTMNNEALREAYQACLAMLSDYSTWLGQHQLFEAFTRLKESDEYARLEGGQORSIDNTLRDFRLAGVDLPEDRKR
RYGEIQARLSELANTFSNHVLDATQAWHLDLTDDTRLGGLPDSALATLKANA EAKGVDGYRITLDFPSPFYVLSFADDR
ALREEVYTAFVTRASDKGPHAGRFDNAPIIEETLRLRRELAEELGFDTYADYSLATKMADSPQQVGLGFLGLADRAHPQ
AQREFNELEAFARESLGLETLPKWDIGYASEKLRERYAISOEQLRPYFPAPRVIEGLFOVTGTLYGIDFAERDDVPRY
```

FICHA 15

EC 3.4.24 Proteína FtsH

FtsH, un miembro de la familia de proteínas AAA (ATPasas asociadas con una variedad de actividades celulares), es una proteasa dependiente de ATP de 71 kDa anclada a la membrana interna. Es responsable de la degradación tanto de la membrana como de las proteínas del sustrato citoplasmático. Las proteínas del sustrato se despliegan y se traslocan a través del poro central del dominio ATPasa hacia la cámara proteolítica, donde las cadenas polipeptídicas se degradan procesivamente en péptidos cortos. (Okuno & Ogura, 2013)

POSICIÓN INTERGENICA:



Donde: 1: proteína de división celular FtsH 2: proteína de división celular FtssJ /ARN ribosómico subunidad grande metiltransferasa E (EC 2.1.1.-) LSU ARNr Um2552 3: Dihidropteroato sintasa (EC 2.5.1.15) 4: proteína de unión a ARN 5: Fosfoglucoamina mutasa (EC 5.4.2.10) 6: Factor de elongación de la transcripción GreA 7: triosafosfato isomerasa (EC 5.3.1.1) 8: Subunidad de translocasas de preproteínas (TC 3.A.5.1.1) 9: Carbamoilfosfato sintasa de cadena larga (EC 6.3.5.5) 10: FIG000325 agrupado con la proteína de terminación de la transcripción NusA 11: proteína de terminación de la transcripción NusA 12: Carbamoilfosfato sintasa de cadena pequeña (EC 6.3.5.5) 13: 4-hidroxi-tetrahidropicolinato reductasa (EC 1.17.1.8)

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

Descripción	Nombre científico	Cubierta de consulta	Por. Ident.	Acc. Len.	Adhesión
<input checked="" type="checkbox"/> Metaloproteasa de zinc dependiente de ATP FtsH [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100.00%	655	WP_011500371.1
<input checked="" type="checkbox"/> Metaloproteasa de zinc dependiente de ATP FtsH [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99.69%	655	WP_110061568.1
<input checked="" type="checkbox"/> Metaloproteasa de zinc dependiente de ATP FtsH [Vibrio parvulus]	Vibrio parvulus	100%	95.11%	654	CDQ34902.1
<input checked="" type="checkbox"/> Metaloproteasa de zinc dependiente de ATP FtsH [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	94.81%	654	WP_178998992.1
<input checked="" type="checkbox"/> Metaloproteasa de zinc dependiente de ATP FtsH [Chromohalobacter canadensis]	Chromohalobacter canadensis	100%	92.37%	654	WP_097021322.1

SECUENCIA DE AMINOÁCIDOS:

```
> fig | 6666666.597456.peg.320 Proteína de división celular FtsH (EC 3.4.24.-) [Chromohalobacter salexigens MP25462]
```

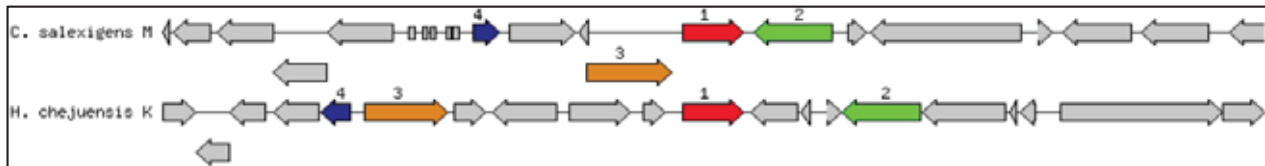
```
MNDMAKNLILWLIVIAAVLLTFVFNNSVDSSPQAMSYSQFVQVQNDQIESVTIEGYTINGEREDGTQFQTIRPAEDPKLMD  
DLLAHDVSVIGKKPEEQSLWTRLLVASFPILILAIFFMRQMQGGGGKGGPMSFGKSKAKLLTQDQIKTTFADVAGCDE  
AKEEVEELVDFLKDPKQFRLGGQIPRGVLMVGPPTGKTLAKAISGEAKVFFFTISGSDVFVEMFVGVGASRVDRDMFEQAK  
KQAPCIIFIDEIDAAGRHRGSGMGGGHDEREQTLNQLLVEMDGFANDGIIVIAATNRPDVLDPALLRPRGFRDQVTVGLPD  
IRGREHILGVHLRKVPLADDVQPSFIARGTSGADLANLVNEAALFAARRNKRLVGMDELEMAKDKILMGEKRSVMVSE  
KEKSNTAYHESGHAIIGLLMPEHDPVYKVTIIPRGRALGVTMFLPEEDRYSLSROOIIISOICSLFGGRLAEEMTLGPNQVTT
```

FICHA 16

EC 3.4.24 Proteína HtpX

Escherichia coli HtpX es una metaloproteínasa de zinc de la familia M48 ubicada en la membrana citoplasmática. Estudios anteriores sugirieron que participa en el control de calidad de las proteínas de membrana (Sakoh, Ito & Akiyama 2005). Sin embargo, su función proteolítica in vivo no se ha caracterizado en detalle, principalmente porque no se han identificado los sustratos fisiológicos y no se dispone de un sustrato modelo que permita la detección sensible de la actividad de la proteasa. (Yoshitani *et al.*, 2019)

POSICIÓN INTERGENICA:



Donde: 1: probable homólogo de proteasahtpX (EC 3.4.24.-) 2: Eritronato-4-fosfato deshidrogenasa (EC 1.1.1.290) 3: Aspartato aminopeptidasa (EC 2.6.1.1) 4: Peptido metionina sulfoxido reductasa MsrB (EC 1.8.4.12)

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

Descripción	Nombre científico	Cubierta de consulta	Por Ident	Acc Len	Adhesión
<input checked="" type="checkbox"/> proteasa HtpX [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100.00%	300	WP_011506648.1
<input checked="" type="checkbox"/> proteasa HtpX [Chromohalobacter japonicus]	Chromohalobacter japonicus	100%	96.67%	300	WP_075270529.1
<input checked="" type="checkbox"/> proteasa HtpX [Chromohalobacter canadensis]	Chromohalobacter canadensis	100%	96.67%	300	WP_097022300.1
<input checked="" type="checkbox"/> proteasa HtpX [Salinicota sp.]	Salinicota sp.	100%	91.33%	300	NRR56271.1
<input checked="" type="checkbox"/> proteasa HtpX [Salinicota sp.]	Salinicota sp.	100%	91.33%	300	MAM58725.1

SECUENCIA DE AMINOÁCIDOS:

```
> fig | 6666666.597456.peg.842 Probable homólogo de proteasa htpX (EC 3.4.24.-) [Chromohalobacter salexigens MP25462]
```

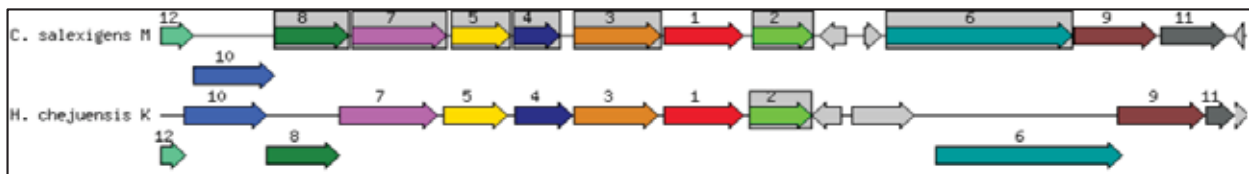
```
MMRILLFLATNLAVVLVASITLRLGLGVEPYLNANGLNMNSLLIFCFVIGMAGSLVSLFISKWMAKMKSTKAKVIEQPG  
NATERWLLDTVGELARDAGIKMPEVAIFPAQQSNFATGWKNKDALVAVSAGLLERMRPEEIRAVLAHEIGHVANGD  
MVTLALIQGVNLNTFVMFFARIVAQLVDSFLRRDDEGGGLGFFGYMAVVI VAEIVFGLVASMVVAVWFSRFREYRADAA  
GAKLAGSGAMINALARLKAETQMPDQMPDTLTAFAITTGQTRKLMERL FASHPPLDDRIRALKESAYRE
```


FICHA 17

EC 3.4.24 Proteína FtsZ

FtsZ es la principal proteína citoesquelética en bacterias y un homólogo de tubulina. Encontramos que FtsZ es degradado por *E. coli* ClpXP, una proteasa dependiente de ATP. In vitro, ClpXP degrada tanto los protómeros como los polímeros de FtsZ; sin embargo, la FtsZ polimerizada se degrada más rápidamente que el monómero. El análisis de delección muestra que el dominio N-terminal de ClpX es importante para el reconocimiento del polímero y que el terminal FtsZ C contiene una señal de reconocimiento de ClpX. (Camberg *et al.*, 2009)

POSICIÓN INTERGENICA:



Donde: 1: Proteína de división celular FtsZ (EC 3.4.24.-) 2: UDP-3-O-N-acetilglucosamina desacetilasa (EC 3.5.1.108) 3: Proteína de división celular FtsA 4: Proteína de división celular FtsQ 5: D-alanina-D-alanina ligasa (EC 6.3.2.4) 6: Exportación de proteínas, proteína del citoplasma SecA ATPasa ARN helicasa (TC 3.A.5.1.1) 7: UDP-N-acetilmuramato-alanina ligasa (EC 6.3.2.8) 8: UDP-N-acetilglucosamina-N-acetilmuramyl-(pentapéptido) pirofosforilundecaprenol N-acetilglucosamina transferasa (EC 2.4.1.227) 9: Glutamato N-acetiltransferasa (EC 2.3.1.35) / N-acetilglutamato sintasa (EC 2.3.1.1) 10: proteína de división celular FtsW 11: proteína mutT mutante (7,8-dihidro-8-oxoguanina-trifosfatasa) (EC 3.6.1.-) / proteína similar a la pirofosforilasa de tiamina-fosfato 12: UDP-N-acetilmuramoilalanina-D-glutamato ligasa (EC 6.3.2.9)

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

Descripción	Nombre científico	Cubierta de consulta	Por. Ident.	Acc. Len.	Adhesión
<input checked="" type="checkbox"/> proteína de división celular FtsZ [Chromohalobacter salexigens]	<i>Chromohalobacter salexigens</i>	100%	100.00%	394	WP_011507482.1
<input checked="" type="checkbox"/> proteína de división celular FtsZ [Chromohalobacter canadensis]	<i>Chromohalobacter canadensis</i>	100%	97.21%	394	WP_097022964.1
<input checked="" type="checkbox"/> proteína de división celular FtsZ [Chromohalobacter japonicus]	<i>Chromohalobacter japonicus</i>	100%	97.21%	394	WP_075369906.1
<input checked="" type="checkbox"/> proteína de división celular FtsZ [Chromohalobacter marismortui]	<i>Chromohalobacter marismortui</i>	100%	96.95%	394	WP_133696794.1
<input checked="" type="checkbox"/> proteína de división celular FtsZ [Chromohalobacter salexigens]	<i>Chromohalobacter salexigens</i>	100%	96.95%	394	WP_178997189.1

SECUENCIA DE AMINOÁCIDOS:

```
> fig | 6666666.597456.peg.2037 Proteína de división celular FtsZ (EC 3.4.24.-) [Chromohalobacter salexigens MP25462]
```

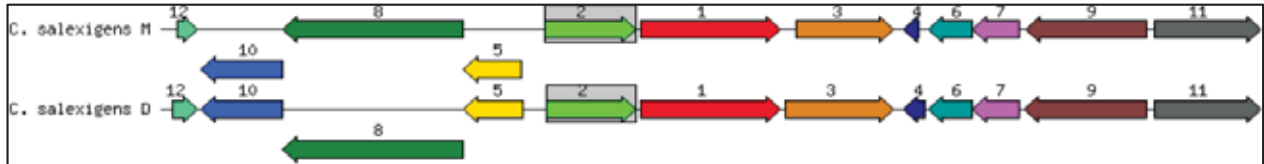
```
MFELVDNAPSSSAVIKIVIGVGGGGGNAVNHMVESNIEGVFICANTDAQALKRVAAKTVLQLGSEITKGLGAGANPEVGRQAA  
MEDRERVAELLQGADVMVITAGMGGGTGTGGAPVVAQVAKELGILTVAVVTRPFPFEGPKRMRAAEEGMASLSEYVDSLITIP  
NEKLLAVLGKNASLLSAFSAANDVLLGAVQGI AELITSPGIINVDFAVVRTVMSEMGMAMMGTGGATGENRAREAAEKAIRSP  
LLEDIDLHGARGILVNITAGPDLSIGEFNDVGTATVQEFASQDATIVVGTSIDMEMSDELRTVVAAAGLEGLKEKAAVSTPQRE
```

FICHA 18

EC 3.4.24 Hidrolasa Coc/NonD

Es una proteína que protege a las bacterias contra el estrés oxidativo, pertenece a la familia del estrés universal –UspA, (Ayarza, Mazzella & Erijman 2015)

POSICIÓN INTERGENICA:



Donde: 1: Proteína M115128 2: Proteína hipotética 3: ATPasa predicha relacionada con la proteína PhoH inducible por inanición de fosfato 4: proteína hipotética 5: Cianofinasa (EC 3.4.15.6) 6: Metiltransferasa dependientes de SAM 7: Carboxilesterasa 8: Cianoficina sintasa (EC 6.3.2.29) 9: Permeasa de sulfato, Tipo Trk 10: isoaspartil dipeptidasa (EC3.4.19.5) Asp-X dipeptidasa 11: proteína hipotética 12: Nitrito reductasa (NAD (P) H) subunidad grande, NirD

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

	Descripción	Nombre científico	Cubierta de consulta	Poc. Ident	Acc. Len	Adhesión
<input checked="" type="checkbox"/>	Hidrolasa de la familia CocE / NonD [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100,00%	677	WP_1791152290.1
<input checked="" type="checkbox"/>	Hidrolasa de la familia CocE / NonD [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,70%	677	WP_110002136.1
<input checked="" type="checkbox"/>	Hidrolasa de la familia CocE / NonD [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,56%	677	WP_224113547.1
<input checked="" type="checkbox"/>	Hidrolasa de la familia CocE / NonD [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	98,97%	677	WP_011506413.1
<input checked="" type="checkbox"/>	Hidrolasa de la familia CocE / NonD [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	98,23%	677	WP_129990272.1

SECUENCIA DE AMINOÁCIDOS:

```
> fig | 6666666.597456.peg.187 Proteína M115128 [Chromohalobacter salexigens MP25462]
```

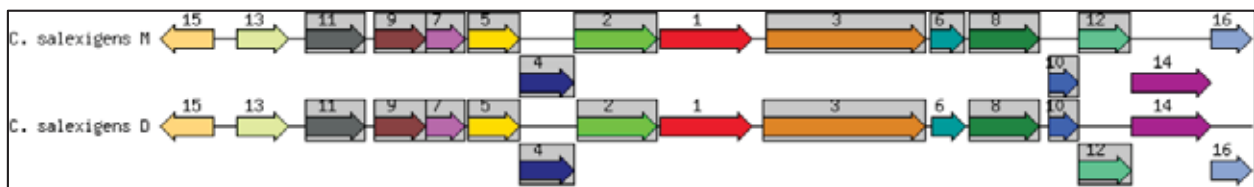
```
MRIVSHFPRQIHEEEDWITLADGCR LAVRIWRPVDAERDPVPAILEYLPYRKRDLTAMRDAQSHAYWAGHGYAGIRVDMRGSGE  
SDGVL RDEY LQQELDDGVEILQWLGRQPWCTGDVGMIGISWGGFNGLQIAALQPPELKAVITLCSTDDRYADDVHHMGGCLLGD  
NLSWASTMFDANACPPDPLLVGDRWRDMWQERLEDSGLWLAKWLQHQRDAYWKHGSVCEDFSRIRCPVYAVSGWADGYCNAV  
RLLAGLDVPRKGLVGPWAHKYPHLGVPGPAIGFLQESLRWWDHHLKGGKATGIMDEPMLRVVMQESVPPSARYEARPGRWVSEPC  
WSPRIDTRPFRRLTSGHDLLPEAQEGAAPDDTPLTVRSPSLVGLYAGKWCSSYNAPPDLPHDQRDEGGALVVFQTARLDRDIEIC  
GQPVVELELEADQPVAMVALRLSDIAEDDKATRISYGLLNLTHRDSDEHPQPLEPGKRYRVRVLLKHIAQQFPAGNAIRLSVST  
SYWPLAWPAPAPVKLTLYPAGSRLLPCRPQPREEADLPADFADPEGAPPLAVTRIQPNQETWRVIRDLANDQTMLEVINDEGV  
IRLDDIDLELSVHVIERITYAYAGNYDSLSGWTQWERRFRRGDWEVRSVTRTLMTSSAESFRLRATLDAYENDSRVFAKSWDEDI  
PRDLV
```


FICHA 19

EC 3.4.24 Proteasa RseP

Escherichia coli RseP, que pertenece a la familia de metaloproteasas de zinc S2P de I-CLiP, desempeña un papel esencial en la activación de una respuesta al estrés de la envoltura celular a través de la escisión de anti- σ Eproteína RseA, una proteína transmembrana de un solo tramo. Un estudio reciente mostró que también escinde los péptidos señal remanentes generados por la translocación de proteínas secretoras en la membrana (Hizukuri *et al.*, 2017)

POSICIÓN INTERGENICA:



Donde: 1: proteasa intermembrana RasP / YluC, implicada en la división celular basada en la escisión de FtsL 2: 1-desoxi-D-xilulosa 5-fosfato reductoisomerasa (EC 1.1.1.267) 3: precursor de YaeT del factor de ensamblaje de proteínas de la membrana externa 4: fosfatidato citidililtransferasa (EC 2.7.7.41) 5: undecaprenil difosfato sintasa (EC 2.5.1.31) 6: precursor de la proteína H de la membrana externa 7: factor de reciclaje de ribosomas 8: UDP-3-O-(3-hidroximiristoil) glucosamina N-acetiltransferasa (EC 2.3.1.-) 9: urudina monofosfato quinasa (EC 2.7.4.22) 10: 3-hidroxiacil-(proteína transportadora de acilo) deshidratasa, forma FabZ (EC 4.2.1.59) 11: Factor de elongación de traslación Ts 12: Acil-(Proteína transportadora de acilo)-UDP-N-acetilglucosamina O-aciltransferasa (EC 2.3.1.129) 13: Proteína ribosómica SSU S2p (SAe) 14: Lipido-A-disacarido sintasa (EC 2.4.1.182) 15: metionina aminopeptidasa (EC 3.4.11.18) 16: Ribonucleasa HII (EC 3.1.26.4)

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

	Descripción	Nombre científico	Cubierta de consulta	Por Ident	Acc Len	Adhesión
<input checked="" type="checkbox"/>	regulador de proteasa sigma E RseP [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100,00%	451	WP_179153335.1
<input checked="" type="checkbox"/>	regulador de proteasa sigma E RseP [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,56%	451	WP_12862673.1
<input checked="" type="checkbox"/>	regulador de proteasa sigma E RseP [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,78%	451	WP_011505678.1
<input checked="" type="checkbox"/>	regulador de proteasa sigma E RseP [Chromohalobacter sp.]	Chromohalobacter sp.	100%	99,56%	451	NQY46817.1
<input checked="" type="checkbox"/>	regulador de proteasa sigma E RseP [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,11%	451	WP_224113851.1

SECUENCIA DE AMINOÁCIDOS:

```
> fig | 6666666.597456.peg.3026 Proteasa intramembrana RasP / YluC, implicada en la división celular basada en la escisión de FtsL [Chromohalobacter salexigens MP25462]
```

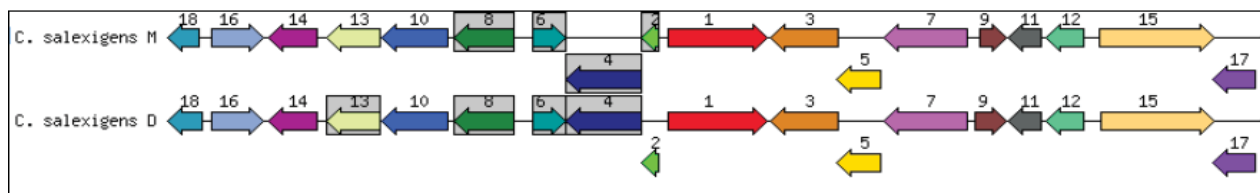
```
MGVIENLLAVIVVLGLLITFHEFGHFVWARRCGVKVLRFSVGF GKPLWSRFD RHGTEFAVA AIP LGGYVKMLDEREGPVAPEE  
QAHAFNRKSVWARIAIVSAGPLANFLLA FVAYWALFIYGTATVAPVIGDVT PDSPAAQGG LQRGQEIVAVQGEPTPSWGEVNL  
KLVA AIGADGTL DVTT RTSEDATPTQHRVPVNDWLV RQNPQPLASLGVT PWRPELPAVLG EVLDDGRAHQAGLQRGDRIVSV  
DGVAVDDWMAFVERVRANPETPLTLQVTRDGERREITLTPAVREQEDGSTIGYIGAGVQPSEWPERYRREIRYGPLDAVGEAV  
AKTGEMSL LTLDSIRKMLVGLISPSNLSGPVTIARIAGDSARDGVESFISFLAYLSISLGVNLNLLPIPVLDGGHLLVYIIIEAV  
RGRPVPEAVQAFGLRVGIALVGS LMLMALYFDLMRL
```

FICHA 20

Metaloproteasa de la familia M48, bepA (anteriormente yfgC)

BepA posee actividad proteasa y es responsable de la degradación de LptD plegada incorrectamente, la alteración de bepA (anteriormente yfgC), un gen regulado por σ (E) que codifica una metaloproteasa periplásmica putativa, sensibiliza a las células a múltiples fármacos, lo que sugiere que puede estar involucrado en el mantenimiento de la integridad de la OM. . (Bryant *et al.*, 2020; Narita *et al.*, 2013)

POSICIÓN INTERGENICA:



Donde: **1.** Precursor exportado de metaloproteasa de Zinc YfgC **2.** Proteína similar a SirA **3.** Proteína reguladora de la vía de la arginina ArgR **4.** Permeasa putativa PerM **5,9.** Proteína hipotética **6.** Tiol peroxidasa **7.** Alcohol deshidrogenasa de cadena corta **8.** 4-hidroxitetrahidropicolinato **10.** Proteína de membrana externa NlpB **11.** Proteína de la glutatión peroxidasa **12.** Hidrolasa putativa de Nudex **13.** Proteína similar a la betalactamasa **14.** Fosforribosilaminoimidazol- succinocarboxamida sintasa **15.** Acil-CoA deshidrogenasa **16.** COG1720, proteína conservada no caracterizada **17.** Proteína exportadora de lisina **18.** N-acetiltransferasa relacionada con GCN5

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

	Descripción	Nombre científico	Cubierta de consulta	Por. Ident	Acc. Len	Adhesión
<input checked="" type="checkbox"/>	Metaloproteasa de la familia M48 [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100,00%	481	WP_179152971.1
<input checked="" type="checkbox"/>	Metaloproteasa de la familia M48 [Chromohalobacter sp.]	Chromohalobacter sp.	100%	99,79%	481	NQY45132.1
<input checked="" type="checkbox"/>	Metaloproteasa de la familia M48 [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,58%	481	WP_110061108.1
<input checked="" type="checkbox"/>	peptidasa M48_Site24p [Chromohalobacter salexigens DSM 3043]	Chromohalobacter salexigens DSM 3043	100%	99,58%	481	ABE60166.1
<input checked="" type="checkbox"/>	Metaloproteasa de la familia M48 [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,58%	483	WP_148266835.1

SECUENCIA DE AMINOÁCIDOS:

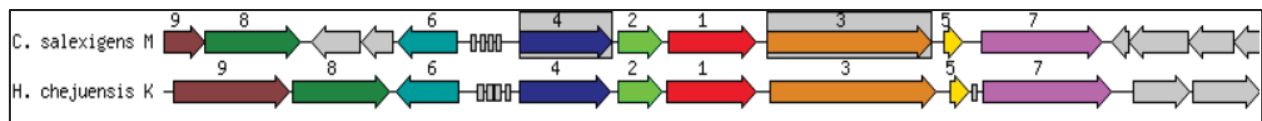
```
> fig | 6666666.597456.peg.1739 Precursor exportado de metaloproteasa de zinc YfgC
[Chromohalobacter salexigens MP25462]
MPRRWNLLPVLCLALAACVATPVVAMAQDDYGLPELGGASASVSSGEEYRLGRAWLRQFRAQTDEWEDPIARDYIHGLVARLLPY
TDVHDPVIVTMVDSRLLNAFAVPGGVMGVNSGLTFADSEDTLASVIAHELGHLSQHHYARRMQRVEETQLPTMAAMLGAMVLA
AGGAGDAGLATMVGSAAFIQDQLAYSRRFEQEADRIGLDAMADAGFDPQAMPEMFRAMQHLASLQGGNPPEFLLTHPVTESRI
SDTQTRANQLPSPAPHTSKVFMIRGRALLSLHRSDPEQAMTRLRQDDPPEAAVRYLQALIDAQRGNTAKALATLDALSEAQPD
LSMLPASAAEVALDAGQRDDALRRARRILRLQPDYYPQRIEAEVLLQAPDQAFNVLDRDMSDQYPEDPHVFALLAEAGRSR
DLWGMLARAHLQLTGHIDRAIKQIDIAEQTARDRGDFAMASRLGERRQAYLGYRQTLRNF
```

FICHA 21

Proteasa de ClpX dependiente de ATP - Proteasa hslU - HslV

Se aisló un nuevo tipo de proteasa dependiente de ATP de *Escherichia coli*. Es el producto del locus de choque térmico hslVU que codifica dos proteínas: HslV, una proteína de 19 kDa similar a las subunidades beta del proteasoma, y HslU, una proteína de 50 kDa relacionada con la ATPasa ClpX. En presencia de ATP, la proteasa hidroliza rápidamente el péptido fluorogénico Z-Gly-Gly-Leu-AMC y muy lentamente algunos otros sustratos de quimotripsina. (Rohrwild *et al.*, 1996)

POSICIÓN INTERGENICA:



Donde: 1: Proteína de Clp dependiente de ATP subunidad de unión a ATP ClpX 2: Subunidad proteolítica de proteasa Clp dependiente de ATP (EC 3.4.21.92) 3: proteasas La dependiente de ATP (EC 3.4.21.53) Tipo I 4: Factor desencadenante de la división celular (EC 5.2.1.8) 5: proteína de unión al ADN HU-beta 6: metilentetrahidrofolato deshidrogenasa (NADP +) (EC 1.5.1.5) / meteniltetrahidrofolato ciclohidrolasa (EC 3.5.4.9) 7: Peptidil-propil cis-trans isomerasa PpiD (EC 5.2.1.8) 8: cisteinil-tRNA sintetasa (EC 6.1.1.16) 9: glutaminil- tRNA sintetasa (EC 6.1.1.18)

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

	Descripción	Nombre científico	Cubierta de consulta	Por. Ident.	Acc. Len.	Adhesión
<input checked="" type="checkbox"/>	Proteasa de Clp dependiente de ATP Subunidad de unión a ATP ClpX [Chromohalobacter salexigens]	<i>Chromohalobacter salexigens</i>	100%	100.00%	426	WP_011507342.1
<input checked="" type="checkbox"/>	Proteasa Clp dependiente de ATP Subunidad de unión a ATP ClpX [Chromohalobacter marismortui]	<i>Chromohalobacter marismortui</i>	100%	97.89%	426	WP_133696494.1
<input checked="" type="checkbox"/>	Proteasa Clp dependiente de ATP Subunidad de unión a ATP ClpX [Chromohalobacter]	<i>Chromohalobacter</i>	100%	97.42%	426	WP_040241497.1
<input checked="" type="checkbox"/>	Proteasa Clp dependiente de ATP Subunidad de unión a ATP ClpX [Halomonas sp. 1513]	<i>Halomonas sp. 1513</i>	100%	91.55%	426	WP_076748092.1
<input checked="" type="checkbox"/>	Proteasa de Clp dependiente de ATP Subunidad de unión a ATP ClpX [Halomonas socii]	<i>Halomonas socii</i>	100%	91.78%	426	WP_163503154.1

SECUENCIA DE AMINOÁCIDOS:

```
> fig | 6666666.597456.peg.2182 Proteasa Clp dependiente de ATP Subunidad de unión a ATP ClpX [Chromohalobacter salexigens MP25462]
```

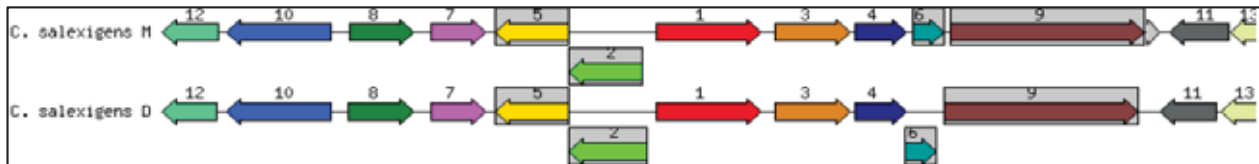
```
MADGKGDSESGKLLYCSFCGKNQNEVRKLIAGPSVYICDECVDLCNDI IREEVLEADAEGDEERLPAPREIRNTLDDYVIGQDR  
AKRVLSSVAVYNHYKRLRADVKSEDELGKSNILLIGPTGSGKTLLEAETMARLLNVPTIADATTLTEAGYVGEDVENI IQKLLQ  
KCDYDVKAKQRGIVYIDEIDKISRKSDNPSITRDVSGEGVQALLKLI EGTTASIPPQGRKHPQQEFLOVDTSNMLFIVGGAF  
AGLEKVI RDRVEKGGIGFNAEVKSKDSEKAMGDILAGVEPEDLVKFLIPEFVGRVPIATLTELSEALI QILTEPKNSLIKQ  
YVHLFAMEGVELEFREDALRAVAHKAMARNTGARGLSILESVLDDTMYEVP SLEDVTKVVIDASVITGDSEPLLIYSQQEESK  
VAGKDG
```

FICHA 22

EC 3.4.11.1 Leucil aminopeptidasa

La leucina aminopeptidasa (LAP, EC: 3.4.11.1) es una metaloexopeptidasa importante que cataliza la hidrólisis de residuos de leucina aminoterminales de polipéptidos y proteínas. (Zhang *et al.*, 2018)

POSICIÓN INTERGENICA:



Donde: 1: Citosol aminopeptidasa PepA (EC 3.4.11.1) 2: FIG000988 permeasa prevista 3: aminotransferasa de aminoácidos de cadena ramificada (EC 2.6.1.42) 4: 3-oxoacil-proteína transportadora de acilo reductasa (EC 1.1.1.100) 5: FIG000906 permeasa prevista 6: subunidad chi de la ADN polimerasa III (EC 2.7.7.7) 7: proteína hipotética 8: regulador transcripcional, familia AraC 9: valil-tRNA sintetasa (EC 6.1.1.9) 10: aldehído deshidrogenasa 11: sistema de transporte K⁺ tipo Kef dependiente de cAMP 12: alcohol deshidrogenasa (EC 1.1.1.1) 13: Xaa-Pro aminopeptidasa

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

	Descripción	Nombre científico	Cubierta de consulta	Por Ident	Acc. Len	Adhesión
<input checked="" type="checkbox"/>	leucil aminopeptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100.00%	507	WP_128991388.1
<input checked="" type="checkbox"/>	leucil aminopeptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99.80%	507	WP_011508209.1
<input checked="" type="checkbox"/>	leucil aminopeptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99.80%	507	WP_035413351.1
<input checked="" type="checkbox"/>	leucil aminopeptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99.61%	507	WP_110062030.1
<input checked="" type="checkbox"/>	leucil aminopeptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	95.07%	507	WP_178099219.1

SECUENCIA DE AMINOÁCIDOS:

```
> fig | 6666666.597456.peg.1651 Citosol aminopeptidasa PepA (EC 3.4.11.1)
[Chromohalobacter salexigens MP25462]
```

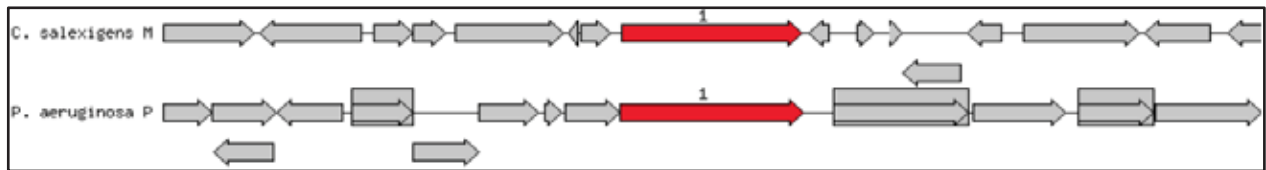
```
MEFPVLTVPNAKIETACLVPVFKDGDLLPAADKLDASERLIGQLLDGRDGFEPVSLGNVQLIPFAPGLGAERLLLVGLGERA
KCGESQLIKALDAAFGAI AKLPLDDVAATFTDVPVGERDTAWKARVTLEAAHRACYRFDEFKSEPPAPRLSSLTLLVGDED
QAAQAE LGARVGA AVGEGVSLTRTLGNLPGNVCTPRYLADQAQLADGADGALNVEILDEDALEALGANALLAVGQGS AEPS
RLIVMEYRGAEDPEEAPHILVKGKITFDSSGGISLKPAGMDEMCFMCGAASVFGTMKTVLALKPRINVI GVVASAENMPDG
RAIKPGDIVKTLKGLNVEILNTDAEGRMVLCDALSYVERFKPASVVDIATLTGA AIIALGHHASGLLSNDDDLALD LLDL DAGE
NAWDRAWHLPLWDEYQSQLDSNFADLAHIGRPA GTITACFLSRFADAYPWAHLDIAGTAWASGD KKGGA
```

FICHA 23

EC 3.4.11.2 Aminopeptidasa N

La aminopeptidasa N (APN / CD13) actúa como exopeptidasa y se ha estudiado durante décadas. En los últimos años, se ha visto no solo como un objetivo relacionado con el tumor, sino también como una proteína funcional potencial en varios otros procesos fisiológicos o patológicos, como la analgesia, la infección por virus y la inflamación. (Su *et al.*, 2011)

POSICIÓN INTERGENICA:



Donde: 1. membrana alanina aminopeptidasa N (EC 3.4.11.2)

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

Descripción	Nombre científico	Cubierta de consulta	Por Ident	Acc. Len	Adhesión
<input checked="" type="checkbox"/> aminopeptidasa N [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100.00%	879	WP_011507814.1
<input checked="" type="checkbox"/> aminopeptidasa N [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99.89%	879	WP_224114492.1
<input checked="" type="checkbox"/> aminopeptidasa N [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99.89%	879	WP_110061700.1
<input checked="" type="checkbox"/> aminopeptidasa N [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99.88%	879	WP_129891107.1
<input checked="" type="checkbox"/> Aminopeptidasa N [Vriobacillus halodentificans]	Vriobacillus halodentificans	100%	94.77%	879	CDQ35182.1

SECUENCIA DE AMINOÁCIDOS:

```
> fig | 6666666.597456.peg.103 Membrana alanina aminopeptidasa N (EC 3.4.11.2)
[Chromohalobacter salexigens MP25462]
```

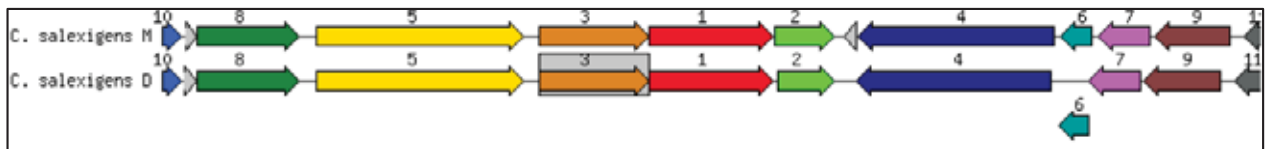
```
MSEPQATHLKDYLPAYRVVTHTELTFDLAPSATTVTARLHVERHPEREAGLPLRFAGDKLSLERIAVDGQTLQADEYQVDDEGL
TIPTVPERFLDQTQS IDPAANTALEGLYLSNGMFTQCEPEGFRITFYPRPDVMTFSTTVGDKETLPVLLSNGNPVERG
ELPGERHFVTTWEDPHPKPSYLFALVAGKLEKVEDHFTTMSGRDVTLQIWWPENLDKTDHAMASLKRAMQWDEETYGREYDLDL
FMIVAVNDFNMGAMENKGLNIFNSAAVLTHPQTATDAAFQAVEGIVAHAYFHNWGNRVTCRDWFQLSLKEGFTVFRDQTFSD
TNSAPVKRIEDVSFFRTAQFAEDAGPTAHPVVRPDHYIEIISNFYTLTIYEKGAEIVRMLRNLLGWETFRQGS DLYFARFDGQAVT
IEDFVDCMAEVSGLDLDQFMRWYSQAGTPEIDAHGEYDYAKCEYHLRLSQRTPTTPGQAEKAPLHIPVRMGLVGTKSGRDLSLT
LDGEALGTETVLHLRDSEQTYVFTGIDEAPTPSLLRGRFSAPVYLRYYPYSREDLSFLLTHDADDFNRWDAGQRLTMLALDDLIAA
HRNGVEKVM DGRVIDAYRRLLTTE TDDKAVLAEMLTLPSEAYIAEQPMVDVDAIHAARDFVKQSLATALRDEFLLALYEAHRSD
APYAPEPEQIAQRRVKNVALEYLMSIEDEQGIALANAQVEAEDNMTDVRAALTMLVHSSRTDLAEPALKAFFGEKWAHDPLVMNQ
WFTIQVTRPQADVLERVKFLMAHPLFSLKNPNKVRALIGAFAAQNRVNFHRLDGEYRLLADVVIELNRLNPEIAARIITPLTR
WQRFDEQRQALMKAELERIRAEELSPNVFEMVERALADA
```

FICHA 24

EC 3.4.11.9 Xaa-Pro aminopeptidasa

Las aminopeptidasas Xaa-Pro son peptidasas que están generalmente unidas a la membrana, dependiente de Mn^{2+} , presente tanto en células de mamíferos como en células bacterianas, y son responsables de la escisión de cualquier aminoácido N-terminal adyacente a un residuo de prolina. (Yaron & Mlynar, 1968)

POSICIÓN INTERGENICA:



Donde: 1: Xaa-Pro aminopeptidasa (EC 3.4.11.9) 2: sistema de transporte K^+ tipo Kef dependiente cAMP 3: transportador de di/tripeptidos 4: Valil-tRNA sintetasa (6.1.1.9) 5: Glutamato-amoniaco-ligasa adeniltransferasa (EC 2.7.7.42) 6: subunidad chi de la ADN polimerasa III 7: 3-oxoacil-proteína transportadora de acilo reductasa (EC 1.1.1.100) 8: acetato permeasa ActP (simportador /acetato) 9: aminotransferasa de aminoácidos de cadena ramificada (EC 2.6.1.42) 10: diacilglicerol quinasa (EC 2.7.1.107) 11: citosol aminopeptidasa PepA (EC 3.4.11.1)

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

Descripción	Nombre científico	Cubierta de consulta	Por. Ident	Acc. Len	Adhesión
<input checked="" type="checkbox"/> proteína de la familia de la aminopeptidasa P [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100,00%	605	WP_179153018.1
<input checked="" type="checkbox"/> proteína de la familia de la aminopeptidasa P [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,67%	605	WP_011508215.1
<input checked="" type="checkbox"/> proteína de la familia de la aminopeptidasa P [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,34%	605	WP_224114117.1
<input checked="" type="checkbox"/> proteína de la familia de la aminopeptidasa P [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,50%	605	WP_110082024.1
<input checked="" type="checkbox"/> proteína de la familia de la aminopeptidasa P [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,01%	605	WP_128691392.1

SECUENCIA DE AMINOÁCIDOS:

```
> fig | 6666666.597456.peg.1644 Xaa-Pro aminopeptidasa (EC 3.4.11.9)
[Chromohalobacter salexigens MP25462]
```

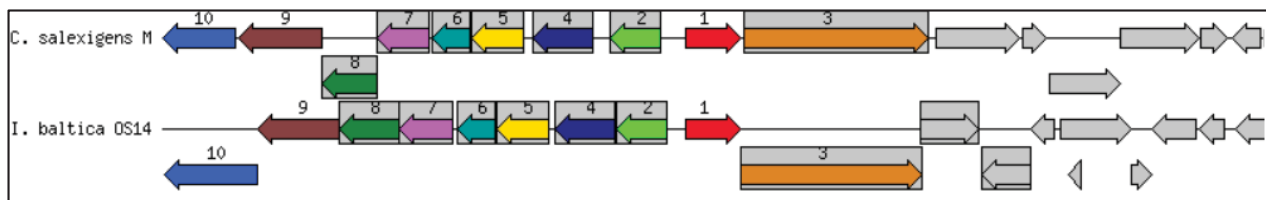
```
MSLSEHPRTPAERLALRETMRNAVDAWWLPSSDPHSSEYLPEHWQGRAWLSGFDGSGVGLTVVTQQAGLWVDSRYVWVQAEQQ
LAGSGIELMKLQPGQAQRPMAWLVEQLAPGATVGFDAVVSLATVRQLQAHLPADIRWEGHRDLLDAIWPNRPALPEAPVRAH
PSDYVDTARREKLAALREKMEEQGADTHLVSTLDDVAWLTNLRGADVDFNPVFLAHLLEQARATLFFVAPDKLGTALIKALAE
GIEVADYTEVASALALPHDARLLIDPARVSLALTEAVPAGVSFVEAMQPSTLAKSRKSDRDIEHVRHAMEEDGAALCAFFAWL
EAALADGETVTELTVDERLTAERARDGFVRSFATIAAFNANGALPHYHATPAAHSVIKGDGLLLIDSGAQYLGTTDITRVV
PVGQIDAAHRRDFTRVLKGTIALSRARFPRGIPSPQLDAIARAPLWAAGLDYGHTGHGVGYFLNVHEGPQVIWYAPVTPQTA
MQPGMITSIIEPGVYRPGQWGVRIENLVVNRDPDEASDFGDFLRFETLTLCPIDTRALDMSLLDAAEIAWLDAYHDEVRRLLPRV
EGPARDWLEQRTAPLAR
```


FICHA 25

EC 3.4.11.18 Metionina aminopeptidasa

Las metionina aminopeptidasas (MetAP) son metaloenzimas que escinden la metionina N-terminal de péptidos y proteínas recién sintetizados. Estas enzimas MetAP están presentes en bacterias, y los experimentos de knockout han demostrado que la actividad de MetAP es esencial para la vida celular, lo que sugiere que las MetAP son buenas dianas de fármacos antibacterianos. Las enzimas MetAP también están presentes en el huésped humano y la selectividad es esencial. (Helgren *et al.*, 2016)

POSICIÓN INTERGENICA:



Donde: 1: metionina aminopeptidasa (EC 3.4.11.18) 2: proteína ribosómica SSU S2p (SAe) 3: proteína-PII uridiltransferasa (EC 2.7.7.59) 4: factor de elongación de traslación Ts 5: Uridina monofosfato quinasa (EC 2.7.4.22) 6: factor de reciclaje de ribosomas 7: undecaprenil difosfato sintasa (EC 2.5.1.31) 8: fosfatidato citidiltransferasa (EC 2.7.7.41) 9: 1-desoxi-D-xilulosa 5-fosfato reductoisomerasa (EC 1.1.1.267) 10: Proteasa intermembrana RasP/YluC, implicada en la división celular basada en la escisión de FtsL

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

Descripción	Nombre científico	Cubierta de consulta	Por. Ident	Acc. Len	Adhesión
<input checked="" type="checkbox"/> metionil aminopeptidasa de tipo I [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100,00%	263	WP_011505870.1
<input checked="" type="checkbox"/> metionil aminopeptidasa de tipo I [Chromohalobacter japonicus]	Chromohalobacter japonicus	100%	90,11%	263	WP_075367959.1
<input checked="" type="checkbox"/> metionil aminopeptidasa de tipo I [Chromohalobacter]	Chromohalobacter	99%	89,27%	263	WP_040239548.1
<input checked="" type="checkbox"/> metionil aminopeptidasa de tipo I [Chromohalobacter canadensis]	Chromohalobacter canadensis	100%	89,73%	263	WP_097023450.1
<input checked="" type="checkbox"/> metionil aminopeptidasa de tipo I [Chromohalobacter marismortui]	Chromohalobacter marismortui	99%	88,93%	263	WP_133695764.1

SECUENCIA DE AMINOÁCIDOS:

```
> fig | 6666666.597456.peg.3034 Metionina aminopeptidasa (EC 3.4.11.18)
[Chromohalobacter salexigens MP25462]
```

```
MNIPIKTADAEIEKMREAGRQAASVLEMIAPHVAAGVTTGEIDRLCHEYIVNELGSTPAPLNYHGFPNATCTSINHVVCHGIPDM
GKKLKKGDIMNLDITVKTADGYHGDSSVMFVIGESIQQGRLSRVTQECLYKSIAEVRPGVRLSELARTIQQHAERHGYSVVRDF
CGHGIGAGFHEDPQFLHYDGYAPEADAVLAEGMCFTEPMLNAGDYRTKVLRDGTAVTKDKLSLSAQWEHTLLVTANGVEVLTA
RQDEDFSLAA
```

FICHA 26

EC 3.4.11.21 Aspartil aminopeptidasa

La aspartil aminopeptidasa (EC 3.4.11.21) escinde solo los residuos de aminoácidos ácidos N-terminales no bloqueados. Hasta la fecha, solo se ha encontrado en mamíferos (Yokoyama, *et al.* . 2006).

POSICIÓN INTERGENICA:



Donde: 1: Aspartil aminopeptidasa 2: proteasa peri plasmática

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

Descripción	Nombre científico	Cubierta de consulta	Por Ident	Acc Len	Adhesión
<input checked="" type="checkbox"/> Aminopeptidasa de la familia M18 [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100,00%	432	WP_011506819.1
<input checked="" type="checkbox"/> Aminopeptidasa de la familia M18 [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,77%	432	WP_035410067.1
<input checked="" type="checkbox"/> Aminopeptidasa de la familia M18 [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,77%	433	WP_110061933.1
<input checked="" type="checkbox"/> Aminopeptidasa de la familia M18 [Chromohalobacter salexigens]	Chromohalobacter salexigens	99%	92,34%	432	WP_178998123.1
<input checked="" type="checkbox"/> Aminopeptidasa de la familia M18 [Chromohalobacter japonicus]	Chromohalobacter japonicus	99%	92,11%	432	WP_040243768.1

SECUENCIA DE AMINOÁCIDOS:

```
> fig | 6666666.597456.peg.707 Aspartil aminopeptidasa [Chromohalobacter salexigens MP25462]
```

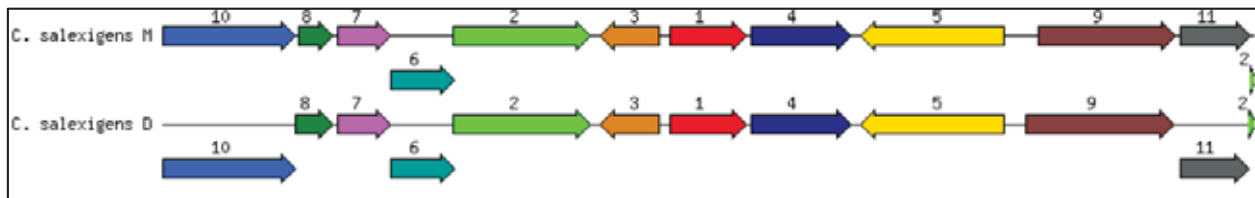
```
MAHAPTLDRLLHFLERSPTPWHAVDNMARRLEQAGYRRLEETEAWQLAPGDRFYVTRNASSLIAMQVPTDPLSGLRMIGAHTDSP  
GLRLKPQPVVAKKDWLQLSVEVYGGALLAPWFDRDLGLAGRIHVRRREDGRLQGVLLHVDRPVAIIPSLAIHLDREANNGRALNAQ  
TQMLPVVLQGGGEADLERWLRWLYEQHGLNIQLLDYELSLYDMQRPSRVGIEGELIASARLDNLLSCTFTGIEALLAGDGRQGA  
LFVANDHEEVGSASACGAQGFPLGDVLRVHAQLGEGGEDGWVRLIQGSRMISCDNAHAVHPNFPEKHDEHHGPAINGGPVIKVN  
ANQRYATNSATAAMFRDICEAGTPVQTFVTRADMGCSTIGPITATELGVPTLDVGI PQWGMHSIRETAGSRDADYLIRALTAF  
VNRTELD
```


FICHA 27

EC 3.4.11.25 Beta-peptidil aminopeptidasa

Las beta-peptidil aminopeptidasas forman la primera familia de enzimas que hidrolizan una variedad de péptidos beta cortos y péptidos que contienen beta-aminoácidos. Todos los representantes de esta familia fueron aislados de bacterias Gram-negativas. Las especificidades de sustrato de las peptidasas varían mucho, pero las enzimas tienen propiedades estructurales comunes y se puede esperar un mecanismo de reacción similar (Geueke & Kohler, 2007),

POSICIÓN INTERGENICA:



Donde: 1: beta-peptidil aminotransferasa (EC 3.4.11.25) 2: Transportador ABC de hidroxamato férrico (TC 3.A.1.14.3) componente permeasa FhuB 3: regulador transcripcional, familia RpiR 4: región asociada a la permeasa de aminoácidos 5: receptor de membrana externa de hidroxamato férrico 6: Transportador ABC de hidroxamato férrico (TC 3.A.1.14.3), proteína de unión al sustrato periplasmico FhuD 7: Transportador ABC de hidroxamato férrico (TC 3.A.1.14.3), proteína de unión a ATP FhuC 8: sigma-24 (similar a Fecl) 9: proteínas receptoras de la membrana externa, principalmentetransporte de Fe 10: proteína de biosíntesis de ferrioxamina E DesD superfamilia de sideroforo sintetasa, grupo C componente de sideroforo sintetasa 11: sistema de transporte de sideroforo Fe³ + tipo ABC, componente periplasmatico de unión a hierro

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

Descripción	Nombre científico	Cubierta de consulta	Por. Ident	Acc Len	Adhesión
<input checked="" type="checkbox"/> Peptidasa de la familia P1 [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100.00%	370	WP_179152642.1
<input checked="" type="checkbox"/> Peptidasa de la familia P1 [Chromohalobacter sp.]	Chromohalobacter sp.	100%	99.19%	370	NGY47470.1
<input checked="" type="checkbox"/> Peptidasa de la familia P1 [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99.19%	370	WP_011506350.1
<input checked="" type="checkbox"/> Peptidasa de la familia P1 [Halomonas campaniensis]	Halomonas campaniensis	98%	83.84%	373	WP_038477495.1
<input checked="" type="checkbox"/> Peptidasa de la familia P1 [Halomonas venusta]	Halomonas venusta	98%	83.29%	373	WP_125750705.1

SECUENCIA DE AMINOÁCIDOS:

```
> fig | 6666666.597456.peg.1501 beta-peptidil aminopeptidasa (EC 3.4.11.25)
[Chromohalobacter salexigens MP25462]
```

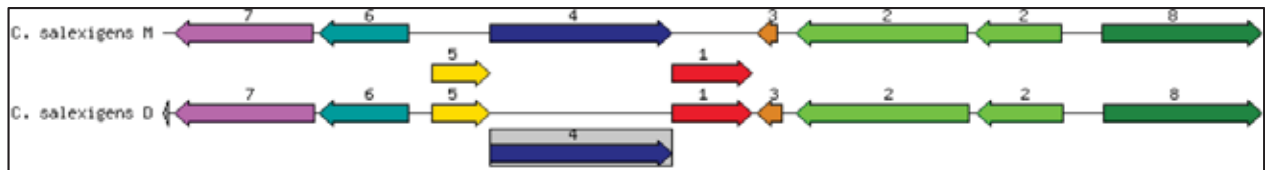
```
MQRMRAREMGLPLPGEPLNNAITDVPGLVGVYETIDGTAENGRPIKGTGVTAILPRARSNTAPVWAGFHALNGNGEMTGTHW
IEQGGYFVGPICLTNSHSVGI VVHHAATRWMLDTYAQTFDAHHLWAMPVVAETYDGLVNDINGQHVEAAHVHAALASASGGAIK
EGNVGGNGMICYGFKGGTGTASRRVGLDQDYTLGVLVQANHGKRDWLNVLGVPVGEALHDADLPEELNRERGSIIAVIATD
APMLPHQLKRLAQRAGLGITRSGSPGGNDSGDMFLAFSTANEGPLPQLGPARQQMHHMNDYFDDFYMAVVQATDEAVLNAMC
MARGAPMAKPEGWCPALDPERLEPLLRAGISIGERN
```

FICHA 28

EC 3.4.19.5 Isoaspartil dipeptidasa

La isoaspartil dipeptidasa (IadA) es una enzima que cataliza la hidrólisis de un resto similar a un dipéptido de isoaspartilo, que puede formarse inapropiadamente en proteínas, entre la cadena lateral del grupo β -carboxilo de Asp y el grupo amino del siguiente aminoácido. (Park, *et al.*, 2017). Es decir, escinde un residuo de Asp ligado a β del extremo N de un polipéptido.

POSICIÓN INTERGENICA:



Donde: 1: isoaspartil dipeptidasa (EC 3.4.19.5) 2: nitrito reductasa (NAD (P) H) subunidad grande, NirB 3: nitrito reductasa (NAD (P) H) subunidad grande, NirD 4: cianoficina sintasa (EC 6.3.2.29) (EC 6.3.2.30) 5: Cianoficina (EC 3.4.15.6) 6: FIG 01057880 proteína hipotética 7: proteína MII5128 8: Subunidad grande de nitrato reductasa asimiladora (CE 1.7.99.4)

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

Descripción	Nombre científico	Cubierta de consulta	Por Ident	Acc. Len	Adhesión
<input checked="" type="checkbox"/> beta-aspartil ceptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100,00%	393	WP_170152288.1
<input checked="" type="checkbox"/> beta-aspartil ceptidasa [Chromohalobacter sp.]	Chromohalobacter sp.	100%	99,24%	393	NGY44879.1
<input checked="" type="checkbox"/> beta-aspartil ceptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	98,98%	393	WP_011506417.1
<input checked="" type="checkbox"/> beta-aspartil ceptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	98,73%	393	WP_224113543.1
<input checked="" type="checkbox"/> beta-aspartil ceptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	98,47%	393	WP_035410783.1

SECUENCIA DE AMINOÁCIDOS:

```
> fig | 6666666.597456.peg.183 Isoaspartil dipeptidasa (EC 3.4.19.5) @ Asp-X dipeptidasa [Chromohalobacter salexigens MP25462]
```

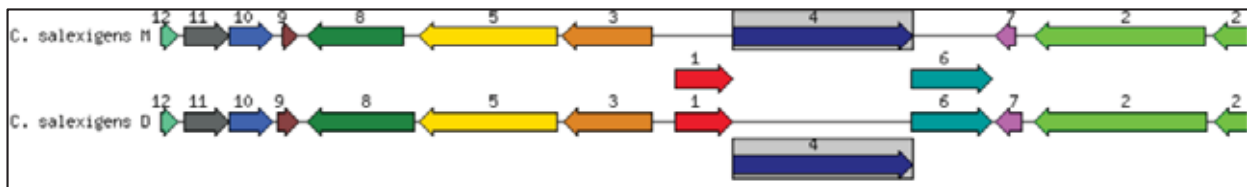
```
MMSGKARQPAPMPTLLRVAEAFAPERLAATDILMAGGRIVALGQGLAVPQGWVQVVDARHLI AVPAFIDQHVHVTGGG  
GEGGCGTRCPPITTRDIVAMGIGTVVGVLTDSISRSPADLLAAVRGLAADGLAAYMYTGAYRVPAPTLTGD IQRD LA  
WIPEVIGVGEIAISDHRSSQPRQDELERLVSDARV GAMLAKRGICHFHLGDGKRGLEPLRRLLTETET EIPADQVIPT H  
VNRRTELLEEAAYALAFDASVDVTA FEDAGDGLSAFD AVSRLLARGVSPARITLSSDCNGSLPEFDADGAYVGMQVA  
RNTTLIADWRRRLVHARVLP LESALGLLAGNVARV LGLADKGR LAVGSGADITLLDKALQPQRTFVAGRCLYGAVDHHE  
TAR
```

FICHA 29

EC 3.4.15.6 Cianofinasa

La cianoficina, o poli-L-Asp-multi-L-Arg, es un polímero peptídico no sintetizado ribosómicamente que se utiliza para el almacenamiento de nitrógeno por cianobacterias y otras eubacterias seleccionadas. Tras la síntesis, se autoasocia para formar gránulos insolubles, cuya degradación es catalizada únicamente por una proteasa carboxi-terminal específica, cianofinasa. (Law *et al.* 2009)

POSICIÓN INTERGENICA:



Donde: 1: cianofinasa (EC 3.4.15.6) 2: Nitroreductasa (NAD (P) H) subunidad grande (EC 1.7.1.4) 3: FIG 01057880 Proteína hipotética 4:cianoficina sintasa (EC 6.3.2.29) (EC 6.3.2.30) 5: proteína MII5128 6: isoaspartil dipeptidasa (EC 3.4.19.5) Asp-X dipeptidasa 7: nitro reductasa (NAD (P) H) subunidad pequeña (EC 1.7.1.4) 8: ATPasa predicha relacionada con la proteína PhoH inducible por inanición de fosfato 9: proteína hipotética 10: metiltransferasas dependientes de SAM 11: carboxilesterasa 12: permeasa de sulfato, tipo Trk.

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

Descripción	Nombre científico	Cubierta de consulta	Por Ident	Acc. Len	Adhesión
<input checked="" type="checkbox"/> cianofinasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100,00%	284	WP_179152288.1
<input checked="" type="checkbox"/> cianofinasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,65%	284	WP_110062133.1
<input checked="" type="checkbox"/> cianofinasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,30%	284	WP_128690274.1
<input checked="" type="checkbox"/> cianofinasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	98,94%	284	WP_224113545.1
<input checked="" type="checkbox"/> cianofinasa [Chromohalobacter sp.]	Chromohalobacter sp.	100%	95,77%	284	NQY44881.1

SECUENCIA DE AMINOÁCIDOS:

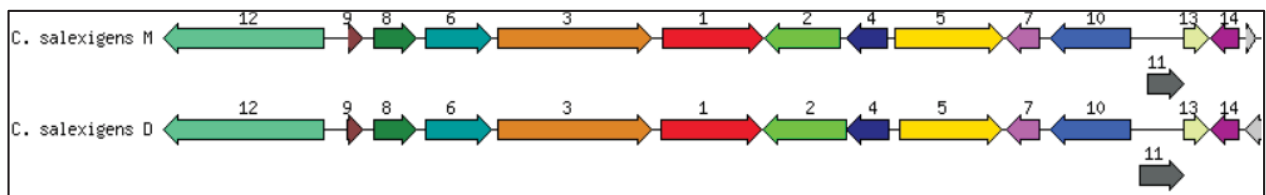
```
> fig | 6666666.597456.peg.185 Cianofinasa (EC 3.4.15.6) [Chromohalobacter salexigens MP25462]
MASRTPTVCGHIVAIGGAEDKTSSELAILRRVAFALAPQGNDDEVAVIATASSIPEQLLPSYEAAPKRLGASRVHALNIQDRQQA
ADTDNVLRIQRSGVIFFTGGDQLRLTTVLGGSDTLRAVRERLREGAVVAGTSAGAAAMPSTMIYNGAAADALRKGAVNMTFG
LGFVRGMVIDSHFLERGRFTRLMEVVGASNPEQLGVGLGEDAAVI IHPGRILETIGPGHVIIIDSRDLASSNIAELEMGAPVA
VENMILHAMVSGHGYDVDARRYLVAAEELEAVLAGRRQE
```

FICHA 30

EC 3.4.13.20 Aminoacil-histidina dipeptidasa

La aminoacilhistidina dipeptidasa, también Xaa-His dipeptidasa, carnosinasa o PepD, cataliza la escisión y liberación de un aminoácido N-terminal, que suele ser un residuo neutro o hidrófobo, a partir de un dipéptido Xaa-His. Durante las infecciones bacterianas, la degradación de la l-carnosina a través de la carnosinasa o enzimas similares a PepD puede aumentar el potencial destructivo de las bacterias, lo que resulta en un impacto patológico. (Aoki, *et al.* 2012)

POSICIÓN INTERGENICA:



Donde: 1: aminoacil-histidina dipeptidasa (peptidasa D) (EC 3.4.13.3) 2: proteína hipotética 3: ATPasa P-loop prevista fusionada con una acetiltransferasa COG1444 4: Proteína de la familia ThiJ/ PfpI 5: proteína de la recombinación de ADN RmuC 6: tiorredoxina reductasa (EC 1.8.1.9) 7: proteína de unión a ADN no específica Dps/ proteína antioxidante de tipo ferritina de unión a hierro/ferroxidasa (EC 1.16.3.1) 8: proteína similar a la subunidad C de la alquil hidropéroxido reductasa 9: proteína hipotética 10: glucerato quinasa 11: proteína hipotética 12: aconitate hidratasa (EC 4.2.1.3) 2-metilisocitrato deshidratasa (EC 4.2.1.99) 13: proteína hipotética 14: proteína de la triada histidina (HIT)

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

Descripción	Nombre científico	Cubierta de consulta	Por Ident	Acc Len	Adhesión
<input checked="" type="checkbox"/> aminoacil-histidina dipeptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100,00%	487	WP_128891417.1
<input checked="" type="checkbox"/> aminoacil-histidina dipeptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,79%	487	WP_011508248.1
<input checked="" type="checkbox"/> aminoacil-histidina dipeptidasa [Chromohalobacter sp.]	Chromohalobacter sp.	100%	99,59%	487	NCY45267.1
<input checked="" type="checkbox"/> aminoacil-histidina dipeptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,38%	487	WP_110062009.1
<input checked="" type="checkbox"/> aminoacil-histidina dipeptidasa [Chromohalobacter canadensis]	Chromohalobacter canadensis	100%	90,76%	487	WP_097022182.1

SECUENCIA DE AMINOÁCIDOS:

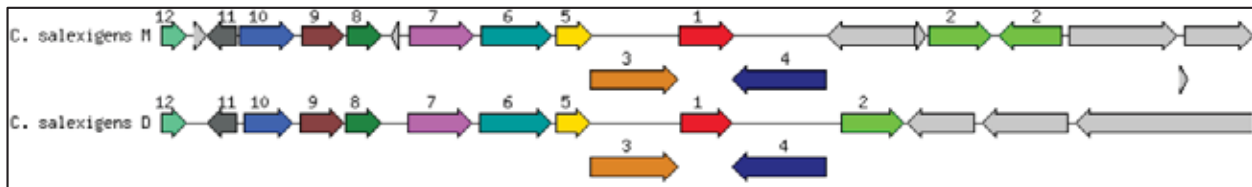
```
> fig | 6666666.597456.peg.1611 Aminoacil-histidina dipeptidasa (Peptidasa D) (EC 3.4.13.3) [Chromohalobacter salexigens MP25462]
MNDHLLQLSPEPLWRHFRTL CNTPRPSGHEAPLIATLTAWADAHGLTHERDAAGNLLIRKAATPGCEALPGVILQGHLDMVAQ
ANA EHPHDFTRDPIQTYVEEGWLHAHETTLGADNGIGVAAALAI L EDETLRHGPLEALFTLEEEASMRGAL ELDTDWLQGRYL
LNL DSEDRGQVYIGCAGGADVVEARLPTRDLGEHEVVRQVALTGLVGGHSGIDIHKGRGNANRLLVRVLRAL E PFGACL VAY
QGGTLRNALPREAFASIALPADEEAAVVARLAELQAELAAELAGVDEDVQLALEPAESAAPLTGAASRL LIAALHVAPCGVER
MSVEVPGVVE TSNNLGVVKLEEGRFHLCALVRSRDSATQDLADRFRALFSLIGAETRVENAYPGWTPAPESTLLARFRLHL
ERTGVDPDVKVIHAGLECGILGGKYPQLEMISFGPLIRGAHSPAERVEIDSVAEFWQVLR A I I E D
L.ASARAA
```

FICHA 31

EC 3.4.13.22 D-alanil-D-alanina dipeptidasa

La D-alanil-D-alanina (D-Ala-D-Ala) dipeptidasa VanX que contiene zinc se ha detectado en bacterias Gram-positivas y Gram-negativas, donde parece haberse adaptado a al menos tres funciones fisiológicas distintas. (Lessard & Walsh, 1999)

POSICIÓN INTERGENICA:



Donde: 1: D-alanil-D-alanina dipeptidasa (EC 3.4.13.22) 2: regulador transcripcional, familia LysR 3: sistema de transporte de dicarboxilato C4 de tipo TRAP, gran componente de permeasa 4: FIG00960493 proteína hipotética 5: sistema de transporte de dicarboxilato C4 tipo TRAP pequeño componente de permeasa 6: sistema de transporte de dicarboxilato C4 tipo TRAP pequeño componente periplasmático 7: regulador transcripcional, familia LysR 8: proteína hipotética 9: proteína fijadora de fosfolípidos 10: proteína similar a la extensina 11: proteína hipotética 12 transportador de glicinabetaína OpuD

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

Descripción	Nombre científico	Cubierta de consulta	Por. Ident.	Acc. Len	Adhesión
<input checked="" type="checkbox"/> D-alanil-D-alanina dipeptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100.00%	259	EWW06253.1
<input checked="" type="checkbox"/> Metalopeptidasa de la familia M15 [Chromohalobacter salexigens]	Chromohalobacter salexigens	99%	99.61%	258	WP_218781802.1
<input checked="" type="checkbox"/> Metalopeptidasa de la familia M15 [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	98.84%	259	WP_222832148.1
<input checked="" type="checkbox"/> Metalopeptidasa de la familia M15 [Chromohalobacter salexigens]	Chromohalobacter salexigens	97%	99.60%	253	WP_084075934.1
<input checked="" type="checkbox"/> Metalopeptidasa de la familia M15 [Chromohalobacter sp.]	Chromohalobacter sp.	96%	99.60%	251	NQY47136.1

SECUENCIA DE AMINOÁCIDOS:

```
> fig | 6666666.597456.peg.1097 D-alanil-D-alanina dipeptidasa (EC 3.4.13.22)
[Chromohalobacter salexigens MP25462]
```

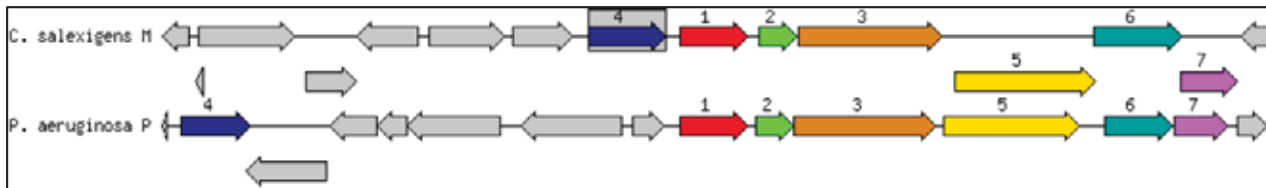
```
MLETRNVNVTDSLRLPIDPIAEPWETIAAIPVDEGGALHPVSLAAPMLRTLPLYHALGVP GAVPECWAREAVYRRLLEVAR
HLPDGLGLVLD AWRP YV VQRHLYDTLLEMVETRYADLDES DRRAMTRAFVSP PSE RVD RPS PHLTGGAVDVALCDADGLLLD
MGS AFDEARAASH TVHF EHHAVDADTRTYRDRRRILYHAMVDAGFANLPSE WWHYSYGDQMWAWYRGEPRALFGPSG PDSLEA
RWRRSLGQDA
```

FICHA 32

EC 3.4.13.19 Dipeptidasa de membrana

La dipeptidasa de membrana (EC 3.4.13.19) es una peptidasa de zinc de la membrana plasmática que participa en el metabolismo renal del glutatión y sus conjugados, como el leucotrieno D4. La enzima carece de las firmas clásicas de otras hidrolasas dependientes de zinc y no muestra homología con ninguna otra proteína de mamífero. (Keynan, Hooper & Turner, 1997)

POSICIÓN INTERGENICA:



Donde: 1: dipeptidasa microsomal (EC 3.4.13.19) 2: FIG00535603: proteína hipotética 3: subunidad A de dimetilglicina desmetilasa de DgcA 4: regulador transcripcional que contiene un dominio amidasa y un dominio HTH de unión a ADN de tipo AraC 5: DgcB subunidad B de bdimetilglicina desmetilasa 6: flavoproteína de transferencia de electrones, subunidad alfa 7: flavoproteína de transferencia de electrones, subunidad beta

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

Descripción	Nombre científico	Cubierta de consulta	Por. Ident	Acc. Lien	Adhesión
<input checked="" type="checkbox"/> dipeptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100.00%	325	WP_011505291.1
<input checked="" type="checkbox"/> dipeptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99.68%	325	WP_128660183.1
<input checked="" type="checkbox"/> dipeptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99.38%	325	WP_224113629.1
<input checked="" type="checkbox"/> dipeptidasa [Chromohalobacter marismortui]	Chromohalobacter marismortui	100%	96.31%	325	WP_133694828.1
<input checked="" type="checkbox"/> Dipeptidasa de membrana (familia de las peptidasas M19) [Vriobacillus halodentificans]	Vriobacillus halodentificans	100%	96.31%	325	CDQ33488.1

SECUENCIA DE AMINOÁCIDOS:

```
> fig | 6666666.597456.peg.1442 Dipeptidasa microsomal (EC 3.4.13.19)
[Chromohalobacter salexigens MP25462]
```

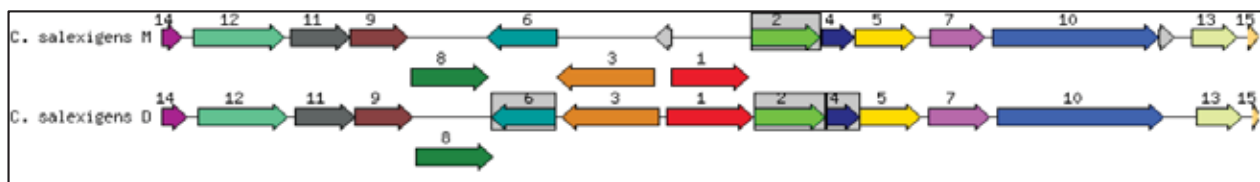
```
MTAQSLHDDAIVIDGLIIAKWGRDLDLDMRRGGLTAANCTVSVWEGFQATVDNIVETNALLAACDDLVRPVHTTADII
RAKEEGKTGIIYGFQNAHAFEDQIGYVEVFKKLGVGIVQMCYNTQNLVGTGCIYERDGGLSGFGREIVAEMNRVGMCD
LSHVGETTSREVI EASEKPVCSHCLPSGLKEHPRNKSDEELRFIADHGGFVGVTMFTFPFLRAGVDATVDDYVEAIEY
VMNIVGEDAIGIGTDFDTQGQDAFFEWLTHDKGYARRLTNFGTIINPEGIRITIGFEPNLTEALLRRGMSEQRVKIMG
R.NW.MR.VI.K.F.V.W.G.A
```


FICHA 33

EC 3.4.13.9 Xaa-Pro dipeptidasa

La prolidasa es una peptidasa dependiente de metales especializada en la escisión de dipéptidos que contienen prolina o hidroxiprolina en sus extremos C-terminales. Un número creciente de estudios destaca la importancia de la prolidasa en varias otras afecciones humanas, incluido el cáncer. Algunos estudios recientes relacionan el papel regulador independiente de la actividad de la prolidasa con la tumorigénesis. (Wilk *et al.*, 2021).

POSICIÓN INTERGENICA:



Donde: 1: Xaa-Pro aminopeptidasa, 2:Subunidad succinato deshidrogenasa, 3:Regulador transcripcional GabR de la utilización de GABA (familia GntR con dominio similar a la aminotransferasa), 4:Proteína reguladora que responde a la leucina, regulador del regulon de leucina (o *lpr*) y sistema de transporte de aminoácidos de cadena ramificada de alta afinidad, 5: Activador de genes inducibles por peróxido de hidrogeno, 6: D-3-fosfoglicerato deshidrogenasa, 7:Cadena D de formiato deshidrogenasa, 8:Transportador ABC de glicerol-3-fosfato, proteína de unión a ATP UgpC, 9: Transportador ABC de glicerol-3-fosfato, proteína permeasa UgpE, 10:Proteína putativa de formiato deshidrogenasa oxidorreductasa, 11: Transportador ABC de glicerol-3-fosfato, proteína permeasa UgpA, 12: Transportador ABC de glicerol-3-fosfato, proteína de unión a glicerol-3-fosfato periplasmica, 13:Proteína hipotética, 14: 4-hidroxi-2-oxoglutarato aldolasa, 15:Aldehido deshidrogenasa B.

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

Descripción	Nombre científico	Cubierta de consulta	Por Ident	Acc Len	Adhesión
<input checked="" type="checkbox"/> peptidasa M24 [Chromohalobacter salexigens DSM 3043]	Chromohalobacter salexigens DSM 3043	100%	100.00%	419	ABE60079.1
<input checked="" type="checkbox"/> ectoína hidrolasa DaeA [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100.00%	399	WP_036413636.1
<input checked="" type="checkbox"/> ectoína hidrolasa DaeA [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99.73%	399	WP_224113745.1
<input checked="" type="checkbox"/> ectoína hidrolasa DaeA [Chromohalobacter japonicus]	Chromohalobacter japonicus	100%	95.74%	399	WP_040242726.1
<input checked="" type="checkbox"/> ectoína hidrolasa DaeA [Chromohalobacter]	Chromohalobacter	100%	95.48%	399	WP_075368193.1

SECUENCIA DE AMINOÁCIDOS:

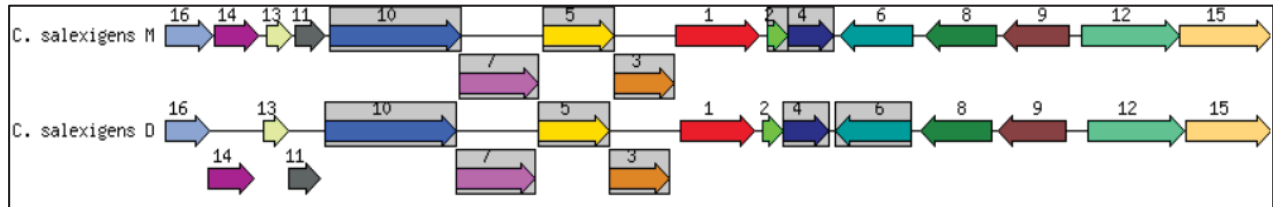
```
> fig | 6666666.597456.peg.1829 Xaa-Pro aminopeptidasa (EC 3.4.13.9)
[Chromohalobacter salexigens MP25462]
MAHRGIDVLIVSDPSNMAWLTYGDGWSFYVHQCVLLGLEGEPEVWYGRMDANGALRTCWIHPDNITYYPDYVQNPDMHPMEYL
AQSILPDRGWHNGIIGMEMDNYFSAKAYLSLVRELPHARFEDANSLVNWCRAIKSPQIEIYMRIAARIVEGMHSRILEVIQPG
LPKSKLVSEIYRVGIEGWTDPDGRVFGGDYPAIVPMLPTGKDAAAPHLTWDDTPFRQGEFTFFFEIAGVYKRYHAPMSRTVFLGT
PPSAFIRAESALLEGIENGLAVAKPGNRTADIAMALGAAMDKYGFDRGGARGCYSIGISYPPDWGERMNSLRPSDDTTLEPGMT
FHFMPGLWEDDWGLEITESILITETGCETLANFPRLFVC
```

FICHA 34

EC 3.4.16.4 D-alanil-D-alanina carboxipeptidasa

La d-Alanil-d-alanina carboxipeptidasa DacC es importante para la síntesis y estabilización de la capa de peptidoglicano de *Escherichia coli*. (Hu *et al.*, 2019)

POSICIÓN INTERGENICA:



Donde: 1: D-alanil-D-alanina carboxipeptidasa (EC 3.4.16.4) 2: proteína reguladora de lipoatos propuesta YbeD 3: precursor de la lipoproteína A poco común 4: Octanoato-proteína transportadora de acilo-prteina-N-octanoiltransferasa 5: precursor de transglicosilasa B de mureina lítica unida a membrana (EC 3.2.1.-) 6: lipoato sintasa 7: proteína que determina la forma de la barra RodA 8: proteínatipo GroES de alcohol deshidrogenasa 9: regulador transcripcional del operon RuBisCO CbbR 10: proteína 2 de unión a penicilina (PBP-2) 11: LSU m3Psi1915 metiltransferasa RImH 12: subunidad A de piruvato carboxilasa (EC 6.4.1.1) 13: proteína lojap 14: nicotinato-nucleotido adeniltransferasa (EC2.7.7.18) 15: subunidad B de piruvato carboxilasa (que contiene biotina) (EC 6.4.1.1) 16: Gamma-glutamil fosfato reductasa (EC 1.2.1.41)

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

Descripción	Nombre científico	Cubierta de consulta	Porcent. Ident.	Acc. Len.	Adhesión
<input checked="" type="checkbox"/> D-alanil-D-alanina carboxipeptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100,00%	400	WP_043558204.1
<input checked="" type="checkbox"/> D-alanil-D-alanina carboxipeptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,25%	400	WP_181396304.1
<input checked="" type="checkbox"/> serina hidrolasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	97,25%	400	WP_035410003.1
<input checked="" type="checkbox"/> proteína de unión a penicilina 9, serina peptidasa, familia MEROPS S11 [Chromohalobacter salexigens D-]	Chromohalobacter salexigen	89%	100,00%	359	ABE50902.1
<input checked="" type="checkbox"/> Precursor de D-alanil-D-alanina carboxipeptidasa DacC [Vergibacillus halodentificans]	Vergibacillus halodentificans	100%	90,25%	400	CDG32091.1

SECUENCIA DE AMINOÁCIDOS:

```
> fig | 6666666.597456.peg.676 D-alanil-D-alanina carboxipeptidasa (EC 3.4.16.4) [Chromohalobacter salexigens MP25462]
```

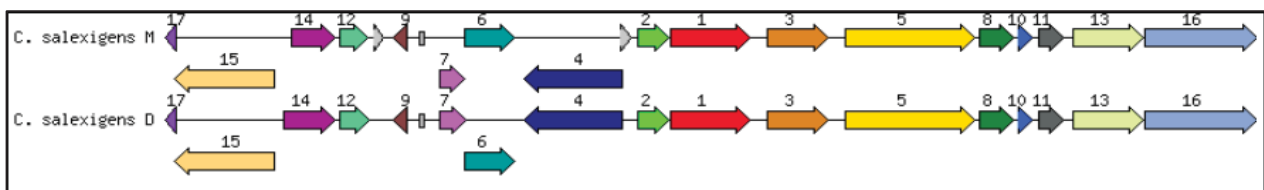
```
MRVFAQSSLLRVVLLVTLICIGVLAPTAHAQNPPPEPSPVPSSMVPASPELAASAWVLMADDSGKILVEHNADKELPPASLTKMM  
TAYLVEREIDAGNISGDDEATISENAWRTGGSRMFLNPGSKVSDNLLRGVVIQSGNDASVALAEHIAGSEDAFADLMNQQA  
RMKLEHHTFMNPTGLPHEEHYSAHDLAVIAQHIKDYDPDHYEVYSEKYFTWNDIRQPNRNRLLRDPNVDGLKTGHTTEAAGY  
CLVASAKKDGTTRLVAVVMGTDSDEARAQETQKLLSYGFRFFETORLYESGAVLNTPRIWGGKDELRLGVADDVYLVTVPRGRS
```


FICHA 35

Proteasa Amidohidrolasa

Catalizan la hidrólisis de una amplia gama de sustratos que llevan grupos funcionales amida o éster en los centros de carbono y fósforo. El centro metálico de esta superfamilia de enzimas tiene una doble funcionalidad en la expresión de la actividad catalítica global. En todos los casos, la molécula de agua nucleofílica se activa mediante la complejación con un centro metálico mononuclear o binuclear. (Seibert & Raushel, 2005)

POSICIÓN INTERGENICA:



Donde: 1. Peptidasa M20D, Amidohidrolasa, 2,9,12. Proteína hipotética 3. OmpA/MotB 4. Transportador de aniones 5. Treonil-tRNA sintetasa 6. Proteína de membrana asociada aLrgA LrgB 7. Proteína de tipo antiholin LrgA 8. Factor de iniciación de la traducción 3, 10. Proteína ribosómica LSU L35p 11. Proteína ribosómica 13. Cadena alfa de fenilalanil -tRNA sintetasa 14. Regulador transcripcional, familia TetR 15. Subunidad de hierro-azufre de xantina deshidrogenasa 16. Cadena beta de fenilalanil -tRNA sintetasa 17. Xantina deshidrogenasa

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

Descripción	Nombre científico	Cubierta de consulta	Por. Ident	Acc. Len	Adhesión
<input checked="" type="checkbox"/> amidohidrolasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100,00%	383	WP_179153462.1
<input checked="" type="checkbox"/> amidohidrolasa [Chromohalobacter sp.]	Chromohalobacter sp.	100%	99,74%	383	NQY44242.1
<input checked="" type="checkbox"/> amidohidrolasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,48%	383	WP_110062444.1
<input checked="" type="checkbox"/> amidohidrolasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	99%	99,48%	383	WP_011507095.1
<input checked="" type="checkbox"/> amidohidrolasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	99,22%	383	WP_128690692.1

SECUENCIA DE AMINOÁCIDOS:

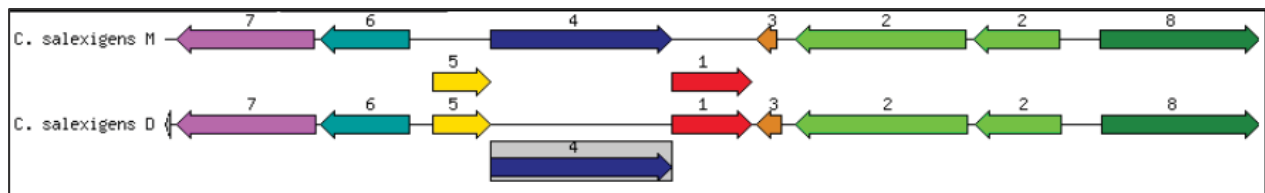
```
> fig | 6666666.597456.peg.3228 Peptidasa M20D, amidohidrolasa [Chromohalobacter salexigens MP25462]
MNEAFSLSLTPAQIRHRLHAAPELSGQETETARRVAGWLEQAGATRIIVTGLGGHVAGVVFAGTLPGPRVMFRAELDGLPIDETA
IAWKSTCEHVAHSCGHGDMVALLGLAQRLAEQPPHGETVLLFQPSEETGRGAREVVESRFEIAPDHVYALHNLPGKPLGE
VSVREGAMTCASRLIIRLNLCAHAAHPEQGRNPALAMCRILQGLKRLPQQLPDDQGLVMVTLIHARLGEVAFGTSPGQAEVM
ATLRTEDDAMMRALAEASVALARAEAEEDDLGIEIDWCDVVFQATRNHPAAVASIREAARSLDMLPTALPEAHRWSEDFGWLAAQ
GQGALLTLGAGERCLPLHNPYDFPDALLDTGIALFERLVRLLHQSDG
```

FICHA 36

Isoaspartil dipeptidasa EC 3.4.19.5

La isoaspartil dipeptidasa (IadA) es una enzima que cataliza la hidrólisis de un resto similar a un dipéptido de isoaspartilo, que puede formarse inapropiadamente en proteínas, entre la cadena lateral del grupo β -carboxilo de Asp y el grupo amino del siguiente aminoácido. (Park, *et al.*, 2017)

POSICIÓN INTERGENICA:



Donde: 1. Isoaspartil dipeptidasa (EC 3.4.19.5), 2. Nitrito reductasa [NAD(P)H] subunidad grande (EC 1.7.1.4), 3. Nitrito reductasa [NAD(P)H] subunidad pequeña (EC 1.7.1.4), 4. Cianoficina sintasa (EC 6.3.2.29)(EC 6.3.2.30), 5. Cianoficina (EC 3.4.15.6), 6. Proteína hipotética, 7. Proteína Mll5128, 8. Subunidad grande de nitrato reductasa asimilatoria (EC:1.7.99.4).

ALINEACIÓN DE SECUENCIAS CON BASE DE DATOS DEL NCBI:

	Descripción	Nombre científico	Cubierta de consulta	Por Identificador	Cuenta Len	Adhesión
<input checked="" type="checkbox"/>	beta-aspartil-peptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	100.00%	393	WP_179152286.1
<input checked="" type="checkbox"/>	beta-aspartil-peptidasa [Chromohalobacter sp.]	Chromohalobacter sp.	100%	99.24%	393	NOY44879.1
<input checked="" type="checkbox"/>	beta-aspartil-peptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	98.98%	393	WP_011506417.1
<input checked="" type="checkbox"/>	beta-aspartil-peptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	98.73%	393	WP_224113543.1
<input checked="" type="checkbox"/>	beta-aspartil-peptidasa [Chromohalobacter salexigens]	Chromohalobacter salexigens	100%	98.47%	393	WP_035410783.1

SECUENCIA DE AMINOÁCIDOS:

```
>fig|66666666.597456.peg.183 Isoaspartil dipeptidasa (EC 3.4.19.5) @ Asp-X
dipeptidasa [Chromohalobacter salexigens MP25462]
MMSGKARQPAPMTLLRVEAFAPERLAATDILMAGGRIVALGQGLAVPQGWVQVVDARHLIAVPAFIDQHVHVTGGGGEGGCGT
RCPPITTRDIVAMGIGTVVGLGTDSISRSPADLLAAVRGLAADGLAAYMYTGAYRVPAPTLTGDIQRDLAWIPEVIGVEIAI
SDHRSSQPRQDELERLVS DARV GAML AGKRG ICH FHLGDGKRGLEPLRRLLTETEIPADQVIPTHVNRRETELLEEA EYALAFD
ASVDVTAFEDAGDGLSAF DAVSRLLARGVSPARITLSSDCNGSLPEFDADGAYVGMQVARNTT LIADWRRLVHARVLP LESALG
LLAGNVARV LGLADKGR LAVGSGADITLLDKALQPQRTFVAGRCLYGAVDHHETAR
```

ANEXO 2.

Preparación de dNTPs a 2mM

Componentes:

- dGTP 4 μ l (200 μ M)
- dCTP 4 μ l (200 μ M)
- dTTP 4 μ l (200 μ M)
- dATP 4 μ l (200 μ M)

Preparación:

- Preparar la cámara de flujo laminar e irradiar por 5min con UV.
- Colocar en un tubo de microcentrifuga estéril, 4 μ l de cada dNTP y añadir 184 μ l de agua bidestilada, homogenizar.
- Almacenar a -20°C hasta el momento de su uso.

ANEXO 3.

Preparación de tampón de caraga para electroforesis

Buffer TAE (50x) 1 L

Tris base, 242 g; ácido acético glacial, 57.1 ml; 0.5 M EDTA pH 8, 100 ml.

Ajustar hasta 1 litro con agua destilada o desionizada.

Buffer TBE (5x) 1 L

Tris base, 54 g; ácido bórico, 27.5 g; 0.5 M EDTA pH 8, 20 ml. Ajustar hasta 1 litro con agua destilada o desionizada. El EDTA debe estar preparado previamente: Se añade al agua destilada o desionizada (un volumen 20% inferior al volumen final deseado) la cantidad adecuada de EDTA para obtener una concentración final de 0.5 M. Se ajusta el pH con NaOH hasta que el EDTA se disuelva y llegue la solución a un pH de 8. Se ajusta el volumen, se esteriliza y se conserva a temperatura ambiente.

Buffer TE 1X

(20 mM Tris-HCl, 0.1mM EDTA, pH 8.0)

ANEXO 4.

Preparación de gel de Agarosa al 1%

Componentes:

- 1g Agarosa (Bioline)
- 100ml Agua bidestilada
- 1ul Syber Safe

Preparación:

- Pesar 1g de agarosa y mezclar con 100ml de tampón TAE 1X, homogenizar.
- Disolver y llevar a ebullición.
- Añadir 1µl de Saber Safe a la mezcla de agarosa y homogenizar.
- Verter la mezcla de agarosa a la cubeta con el peine puesto y dejar gelificar aproximadamente de 15 a 30 minutos
- Colocar el gel en la cámara electroforética y cubrir con TAE 1X, después quitar con suavidad el peine.

ANEXO 5.

Preparación de medio SW al 15%

Componentes:

- NaCl 15%
- 9.6g MgSO₄.7H₂O
- 7.0g MgCl₂.6H₂O
- 2g KCl
- 0.36g CaCl₂
- 0.06g NaHCO₃
- 10g Extracto de levadura
- 5g Peptona
- 1g Glucosa
- 20g Agar
- 1000ml Agua destilada
- Para un medio liquido SW, excluir el agar de los componentes.

Preparación:

- Pesar todos los componentes, y añadir 1000ml, homogenizar.
- Autoclavar a 121°C por 15 minutos.
- Esperar a que se atempere y con cuidado apresurar el plaqueado.

ANEXO 6.

Preparación de medio LB con Ampicilina 100ug/ml

Componentes:

- 100 µg/ml Ampicilina
- 1% NaCl
- 0.5% Extracto de levadura
- 1% Triptona
- 2% Agar

Preparación:

- Pesarse todos los componentes excepto la ampicilina, añadir 100ml de agua destilada y autoclavar a 121°C por 15 minutos.
- Para la Ampicilina preparar una solución concentrada de 100 mg/ml con agua destilada, se esterilizó por filtración y se guardó a -20°C. Se utiliza a una concentración final de 100 µg/ml que se añade antes de realizar el plaqueo.

ANEXO 7.

Preparación de medio de recuperación SOC

Componentes:

- 2g Triptona
- 0.5g Extracto de levadura NaCl (Panreac) 0'58
- 0.25ml KCl 1M
- 1 ml NaCl 1M
- 1ml Mg 2M
- 1ml Glucosa 2M

Preparación:

- Pesarse triptona, extracto de levadura, NaCl, KCl a 97ml de agua destilada
- Mezclar y autoclavar a 121°C por 15 minutos. Dejar enfriar a temperatura ambiente.
- Agregar 2M de Mg y 2M de glucosa, cada uno a una concentración final de 20mM.
- Llevar a 100ml con destilada esteril.

ANEXO 8.

Informe de préstamo de servicio para secuenciación del SEGAI (Servicios Generales de Apoyo a la Investigación de la Universidad de la Laguna)



Universidad de La Laguna
Instituto de Investigación



SEGAI

Página 1 de 1

INFORME DE PRESTACIÓN DE SERVICIO

Referencia:	R.09-PG-SG
Rev.:	05
ID:	47305

Resumen de datos Técnicos:

TECNICAS	PROCEDIMIENTOS	CANTIDAD
<input type="checkbox"/> Purificación de ADN + electroforesis	Kit de extracción de ADN: <input type="checkbox"/> DNeasy Blood & Tissue Kit (Qiagen) <input type="checkbox"/> DNeasy Plant Mini Kit (Qiagen). <input type="checkbox"/> Otros:	
<input type="checkbox"/> PCR estándar + electroforesis	Kit de PCR utilizado: <input type="checkbox"/> AmpONE taq DNA Polymerase (GeneAII) <input type="checkbox"/> Otros:	
<input type="checkbox"/> Electroforesis ADN	Marcador de peso molecular: <input type="checkbox"/> 1 Kb DNA Ladder (Panzac) <input type="checkbox"/> Otros: 1 Kb Ladder (New England Biolabs)	
<input type="checkbox"/> Purificación de producto de PCR	Kit de purificación utilizado: <input type="checkbox"/> ExoSAP-IT (Affimetrix) <input type="checkbox"/> ZymoClean™ Gel DNA Recovery Kit (Zymo) <input type="checkbox"/> Otros:	
<input type="checkbox"/> PCR a tiempo real	Kit de PCR a tiempo real utilizado: <input type="checkbox"/> iQ™ SYBR® Green Supermix (BIORAD) <input type="checkbox"/> Otros:	
<input type="checkbox"/> Síntesis de cDNA	Kit de síntesis de cDNA utilizado: <input type="checkbox"/> iScript™ cDNA Synthesis Kit (BIORAD) <input type="checkbox"/> Otros:	
<input checked="" type="checkbox"/> Secuenciación de ADN	Kit de reacción de secuenciación: <input checked="" type="checkbox"/> Big Dye Terminator v3.1 (Applied Biosystems) <input type="checkbox"/> Otros: Purificación reacción de secuenciación: <input checked="" type="checkbox"/> Etanol + acetato amónico + EDTA <input type="checkbox"/> Otros: Polímero en Genetic Analyzer AB 3500: <input checked="" type="checkbox"/> POP-7 <input type="checkbox"/> Otros:	4
<input type="checkbox"/> Análisis de fragmentos de ADN	Estándar de peso molecular: <input type="checkbox"/> GeneScan 600 LIZ Size Standard <input type="checkbox"/> Otros: Polímero en Genetic Analyzer AB 3500: <input type="checkbox"/> POP-7 <input type="checkbox"/> Otros:	
<input type="checkbox"/> Otro tipo de análisis	<input type="checkbox"/> Otros:	

Observaciones a los resultados:	Personal del Servicio:
Se ha procedido al análisis de 4 secuencias, cuyo pago ha sido realizado a través del sistema UZI (hoja de encargo PS-A18110009-003) Quedan pendientes por realizar otras 79 secuencias. NOTA: Se ha repetido una secuencia de la solicitud ID 47241 para mejorar la calidad de los resultados	Fecha 13/11/2018  Fdo: María Andrés González Carracedo

Universidad de La Laguna, Vicerrectorado de Investigación, SEGAI Apartado 456 CP: 38200 San Cristóbal de La Laguna/ Santa Cruz de Tenerife. Tel: 922319523 - 922310940. E-mail: suplives@ull.es. CIF: O3818001D





“CUSCO CAPITAL HISTÓRICA DEL PERÚ”

Año Del Fortalecimiento De La Soberanía Nacional

CUSCO, 03 de Agosto del 2022

CARTA

*Dr. Olintho Aguilar Condemayta
Decano De La Facultad De Ciencias
Universidad Nacional San Antonio Abad del Cusco*

Presente.

Asunto: LEVANTAMIENTO DE OBSERVACIONES DE TESIS SUSTENTADA.

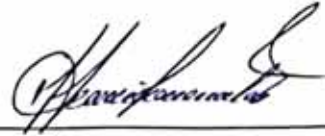
De nuestra mayor consideración, es grato dirigirnos a usted para saludarlo y a la vez para poner en su conocimiento que en mérito al acta de sustentación de tesis de fecha 15 de Julio del 2022 a horas 11:10 a.m., en el SALON VIRTUAL DEL SALON DE GRADOS, se resuelve señalar que las observaciones realizadas durante la sustentación oral de la tesis intitulada “**Descripción de enzimas proteolíticas y obtención de una proteasa recombinante a partir del análisis de datos del genoma de Chromohalobacter salexigens MP25462**”, presentada por la Bachiller de la Escuela Profesional de Biología **Jhomira Ilucion Mamani Morales** para optar al título profesional de Biólogo, han sido levantadas apropiadamente. En fe de lo cual, firmamos el presente.

Sin otro particular, aprovechamos la oportunidad para expresarle las consideraciones más distinguidas.

Atentamente



**Dra. JULIA GRISELDA MUÑIZ
DURÁN**
PRIMER DICTAMINANTE



Mgt. JORGE ACURIO SAAVEDRA
SEGUNDO DICTAMINANTE



**Mgt. RAMON GUSTAVO QUISPE
MONTAYA**
PRIMER REPLICANTE



**Blga. OLGA LIBIA CJUNO
HUANCA**
SEGUNDO REPLICANTE