

GENERACIÓN DE MÉTODOS BASADOS EN GENÓMICA PARA DETECTAR MICROORGANISMOS NOCIVOS EN SISTEMAS ACUÁTICOS

DCV. ALVARO GONZÁLEZ REVELLO

Orientador: Dr. José Sotelo-Silveira
Departamento de Genómica, Instituto de
Investigaciones Biológicas Clemente Estable

Co-Orientadora: Dra. Claudia Piccini
Departamento de Microbiología, Instituto de
Investigaciones Biológicas Clemente Estable



UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY



FACULTAD DE
CIENCIAS
UDELAR | fcien.edu.uy



**TESIS DE
MAESTRÍA EN
BIOTECNOLOGÍA**

Montevideo, Diciembre de 2020

PÁGINA DE APROBACIÓN

Director de Tesis: Dr. José Sotelo-Silveira

Codirectora de Tesis: Dra. Claudia Piccini

Tribunal: Dra. Claudia Etchebehere (Presidente), Dr. Martín Fraga, Dr. Andrés Iriarte (Vocales)

Fecha: 15 de diciembre de 2020

Calificación: Aprobado con mención

Autor: Álvaro González Revello

AGRADECIMIENTOS

A mis tutores ambos, Coya y Claudia, por aceptarme como estudiante y permitirme trabajar en este proyecto. Por las enseñanzas transmitidas, la confianza brindada y por abrirme las puertas al IIBCE y a sus líneas de investigación.

A Karina Cal por pensar en mí para esta tesis y recomendarme con el Coya.

A la ANII y al Posgrado de Biotecnología (FCien) por el apoyo económico brindado.

A los colaboradores de las distintas instituciones que me facilitaron cepas o muestras: Ana Umpiérrez y Susana Deus (IIBCE, Microbiología), Valeria Braga (Instituto de Higiene), Alejandro Perreta (Instituto de Investigaciones Pesqueras) y Rosario de los Santos (Facultad de Veterinaria) y que muy amablemente me brindaron su tiempo en las distintas consultas realizadas.

A Pablo Zunino y Andrés Iriarte por su colaboración y el apoyo en la etapa de diseño y análisis bioinformático.

A Eliana, mi compañera durante gran parte de este trabajo, por su colaboración, entusiasmo y amistad.

A los integrantes del Departamento de Microbiología (IIBCE) y especialmente a mis compañeros del grupo de ecología acuática (Flor, Gaby, Nacho y Paula) con quienes compartí horas de laboratorio/escritorio y siempre estuvieron dispuestos a darme una mano.

A los integrantes del Departamento de Genómica (IIBCE), especialmente con los que compartí laboratorio en la plataforma de secuenciación, Rafa, Valeria, Carlitos y Joaquina, gracias por la ayuda continua.

A mis compañeros docentes del Departamento de Ciencia y Tecnología de la Leche (FVet) por el apoyo y comprensión en los momentos que se me dificultó estar presente.

Finalmente, a mi familia (padres y hermano) por el aliento, confianza y apoyo brindado. Y a los amigos de siempre, Lucía, Diana, Rodrigo, Natasha, Bettina y Jose quienes estuvieron a mi lado, sosteniendo y haciendo el camino más fácil.

Muchas gracias a tod@s...

FINANCIACION

Este trabajo fue financiado por la Agencia Nacional de Investigación e Innovación (ANII) a través de la beca nacional de maestría (POS_FMV_2015_1_1005288).

También recibió apoyo económico del Posgrado de Biotecnología (Facultad de Ciencias, UdelAR) a través de las alcúotas para apoyo a tesis.

El trabajo se enmarca en el proyecto de investigación “Generación de un método basado en secuenciación masiva para detectar microorganismos nocivos en muestras de agua”. ANII-Fondo María Viñas (FMV_1_2014_1_104673).

DIFUSION DE RESULTADOS

Living together: the heterotrophic bacterial community associated to *Microcystis* spp. blooms (2017). Piccini, C., Nervi, E., González Revello, A., Sotelo-Silveira, J., Martinez de la Escalera, G., Segura, A., Kruk, C. 18th Workshop of the International Association of Phytoplankton Taxonomy & Ecology (IAP). Natal, Brasil.

Análisis de comunidades bacterianas en muestras de agua: puesta a punto de un protocolo in-house basado en secuenciación masiva del gen ribosomal 16S (2017). Nervi, E., González Revello, A., Bertoglio, F., Iriarte, A., Sotelo-Silveira, J., Piccini, C. Congreso Nacional de Biociencias Ciudad. Montevideo, Uruguay.

Selección de elementos conservados en genes de virulencia bacterianos para la implementación de una herramienta de detección de patógenos mediante secuenciación masiva (2017). González Revello, A., Nervi, E., Iriarte, A., Zunino, P., Piccini, C., Sotelo-Silveira, J. Congreso Nacional de Biociencias Ciudad. Montevideo, Uruguay.

Potencial patogénico del agua: diversidad filogenética vs. diversidad de ecotipos (2017). González Revello, A., Piccini, C., Nervi, E., Iriarte, A., Zunino, P., Sotelo-Silveira, J. 1^{er} Taller de la Red colaborativa en Ecología Microbiana Acuática en América Latina. Rocha, Uruguay.

Detección de elementos de patogenicidad y toxicidad bacterianos mediante AmpliSeq™ en sistemas acuáticos (2018). González Revello, A., Nervi, E., Fort, R., Iriarte, A., Zunino, P., Piccini, C., Sotelo Silveira, J. 1^{er} Encuentro Bienal de la Sociedad de Bioquímica y Biología Molecular. Montevideo, Uruguay.

Multiplex NGS detection of conserved regions of bacterial toxins in environmental water (2018). González Revello, A., Nervi, E., Fort, R., Iriarte, A., Zunino, P., Piccini, C., Sotelo Silveira, J. ASM Conference on Rapid Applied Microbial Next-Generation Sequencing and Bioinformatic Pipeline. Washington, USA.

Neutral and niche-based mechanisms shape bacterial community composition from freshwater to marine realm (2019). Piccini, C., Segura, A.M., Kruk, C., Nervi E., González Revello, A., Bertoglio, F., Sotelo-Silveira, J. 1st.Congress International Society for Microbial Ecology, ISME-Latin America Ciudad. Valparaiso, Chile.

TABLA DE CONTENIDO

PÁGINA DE APROBACIÓN.....	2
AGRADECIMIENTOS	3
FINANCIACION	4
DIFUSION DE RESULTADOS.....	4
LISTA DE TABLAS	7
LISTA DE FIGURAS	8
LISTA DE ABREVIATURAS	10
RESUMEN	11
ABSTRACT	13
1. INTRODUCCIÓN	15
1.1. Contaminación de Los Recursos Hídricos Por Agentes Patógenos.....	15
1.1.1. Impacto Global de Las Enfermedades Transmitidas Por el Agua	15
1.1.2. Enfermedades Microbianas Asociadas al Agua	17
1.2. Bacterias Transmitidas Por el Agua.....	19
1.2.1. Persistencia y Crecimiento en Diferentes Sistemas Acuáticos	19
1.2.2. Bacterias Patógenas Transmitidas Por el Agua Potable	20
1.2.3. Eutrofización, Floraciones Tóxicas y Cianobacterias	22
1.2.4. Factores de Virulencia Bacterianos	25
1.3. Métodos Para la Detección de Patógenos en Agua	27
1.3.1. Métodos clásicos.....	27
1.3.2. Aplicación de Métodos de Secuenciación Masiva al Estudio de Patógenos.....	28
2. OBJETIVOS Y ESTRATEGIA DE TRABAJO	31
2.1. Objetivo General	31
2.2. Objetivos Específicos.....	31
2.3. Estrategia de Investigación	32
3. MATERIALES Y MÉTODOS.....	33
3.1. Diseño Del Sistema de Amplificación de Genes de Virulencia Bacterianos	33
3.1.1. Selección de Agentes Patógenos y Genes Candidatos	33
3.1.2. Evaluación de la Conservación en Los Genes de Virulencia	33
3.1.3. Construcción del Panel de <i>Primers</i> AmpliSeq	34
3.1.4. Análisis <i>in Silico</i> Del Panel de <i>Primers</i> AmpliSeq.....	35
3.2. Evaluación Del Panel en Muestras Controles Con Potencial Patogénico	36
3.2.1. Colección de Cepas y Extracción de ADN.....	36
3.2.2. Diseño Experimental	36
3.2.3. Construcción de Pooles de ADN.....	37
3.2.4. Amplificación y Construcción de Librerías	38
3.2.5. Amplificación Clonal, Enriquecimiento y Secuenciación.....	41
3.2.6. Análisis Bioinformático	41
3.2.7. Análisis Estadístico	42
3.3. Evaluación Del Panel en Una Muestra Ambiental	43
3.3.1. Sitio Del Muestreo	43

3.3.2.	Procesamiento de la Muestra y Extracción de ADN	44
3.3.3.	Amplificación y Secuenciación de la Muestra.....	44
3.4.	Evaluación de la Estructura Poblacional.....	45
3.4.1.	Amplificación Por PCR.....	45
3.4.2.	Construcción de Librerías y Secuenciación	45
3.4.3.	Análisis Bioinformático	46
4.	RESULTADOS	47
4.1.	Diseño Del Sistema de Amplificación de Genes de Virulencia Bacterianos	47
4.1.1.	Genes de Virulencia Candidatos.....	47
4.1.2.	Evaluación de la Conservación en Los Genes de Virulencia	51
4.1.3.	Panel de <i>Primers</i> Fvir: Características Generales.....	54
4.1.4.	Evaluación <i>in Silico</i> Del Panel.....	57
4.2.	Evaluación Del Panel en Muestras Controles con Potencial Patogénico.....	60
4.2.1.	Cuantificación de Librerías y Secuenciación	60
4.2.2.	Control Con Potencial Patogénico (Cepas Únicas)	62
4.2.3.	Control Con Potencial Patogénico (Cepas Múltiples)	67
4.3.	Evaluación Del Panel en Una Muestra Ambiental	75
4.3.1.	Amplificación y Secuenciación.....	75
4.3.2.	Genes de Virulencia Amplificados	75
4.3.3.	Funcionamiento Del Panel	77
4.3.4.	Control de la Estructura Poblacional.....	79
4.4.	Funcionamiento General Del Panel Fvir en Las Muestras Analizadas.....	81
5.	DISCUSIÓN	83
5.1.	Diseño del Panel	83
5.1.1.	Genes Candidatos.....	83
5.1.2.	Evaluación de la conservación en los genes de virulencia.....	84
5.1.3.	Panel FVIR, Características Generales y Evaluación <i>in Silico</i>	85
5.2.	Evaluación Del panel en Controles Con Potencial Patogénico	87
5.2.1.	Control Con Potencial Patogénico (Cepas Únicas)	87
5.2.2.	Control Con Potencial Patogénico (Cepas Múltiples)	89
5.2.3.	Control de la estructura poblacional.....	90
5.3.	Evaluación Del Panel en una Muestra Ambiental.....	91
5.3.1.	Genes de Virulencia Detectados.....	91
5.3.2.	Control de la Estructura Poblacional.....	92
6.	CONCLUSIONES	93
7.	PERSPECTIVAS.....	94
8.	REFERENCIAS BIBLIOGRÁFICAS	95
9.	ANEXO I. MATERIAL ADICIONAL	110
10.	ANEXO II. TRABAJOS SUPLEMENTARIOS	115

LISTA DE TABLAS

TABLAS:

Tabla 1. Bacterias patógenas transmitidas por el agua potable (OMS)	21
Tabla 2. Bacterias patógenas con transmisión propuesta a través del agua potable (OMS)	22
Tabla 3. Genes de virulencia de bacterias asociadas al agua, clasificación funcional	26
Tabla 4. Ensayo experimental I, control con cepas únicas.....	36
Tabla 5. Ensayo experimental II, control con cepas múltiples.....	37
Tabla 6. Factores de virulencia candidatos para panel de <i>primers</i> de AmpliSeq™.....	48
Tabla 7. Número de Reads obtenidos durante el análisis de las muestras control.....	61
Tabla 8. Antecedentes de la muestra de Gualeguaycito (Deus, 2018).	75
Tabla 9. Funcionamiento general de los <i>primers</i> del panel Fvir en las muestras analizadas	82

TABLAS ANEXOS:

Tabla S1. Aislamientos bacterianos utilizados en los ensayos experimentales.....	110
Tabla S2. Referencias genómica de genes de virulencia con región conservada.....	111
Tabla S3. Referencias genómicas de genes de virulencia sin región conservada	113
Tabla S4. Genes seleccionados para estudio de comparativo con ARNr16S.	115

LISTA DE FIGURAS

FIGURAS:

Figura 1. Principales riesgos generados por el cambio climático sobre las enfermedades transmitidas por el agua.	16
Figura 2. Floraciones de cianobacterias en la costas de Uruguay (verano de 2019)	24
Figura 3. Estrategia general del presente trabajo, flujograma del proceso.	32
Figura 4. Pasos secuenciales para el estudio de conservación de genes de virulencia.....	33
Figura 5. Esquema de la localización de <i>primers</i> en un gen seleccionado	34
Figura 6. Diseño de <i>primers</i> mediante AmpliSeq™ designer.....	35
Figura 7. Flujograma general para la secuenciación utilizando la plataforma de Ion Torrent PGM™.	38
Figura 8. Amplificación y construcción de librerías con AmpliSeq™.....	39
Figura 9. Purificación de amplicones con perlas magnéticas, pasos secuenciales (AMPure™ XP). ..	40
Figura 10. Muestra ambiental, localización de la estación Gualeguaycito.	43
Figura 11. Flujograma del proceso de análisis con QIIME..	46
Figura 12. Distribución filogenética de los microorganismos seleccionados como blanco.	47
Figura 13. Análisis general de la base de datos construida a partir de secuencias de GenBank.....	51
Figura 14. Genes Conservados. Tamaño y porcentaje de conservación de la secuencia.	52
Figura 15. Genes no conservados. Tamaño y porcentaje de conservación de la secuencia.	53
Figura 16. Descriptores de la base de datos generada, genes conservados vs no conservados..	54
Figura 17. Cobertura del Panel de <i>primers</i> Fvir en la secuencia de los genes candidatos	55
Figura 18. Número de juegos de <i>primers</i> del panel Fvir, filos <i>Cyanobacteria</i> y <i>Actinobacteria</i>	55
Figura 19. Número de juegos de <i>primers</i> del panel Fvir, filo <i>Firmicutes</i>	56
Figura 20. Número de juegos de <i>primers</i> del panel Fvir, filo <i>Proteobacteria</i>	56
Figura 21. Características generales de los <i>primers</i> del panel Fvir.....	57
Figura 22. Análisis de la estructura interna de los <i>primers</i> del panel Fvir	58
Figura 23. Histograma de frecuencias del tamaño de insertos del Panel Fvir.	58
Figura 24. Ejemplo de localización de <i>primers</i> fuera de la región de conservación.....	59
Figura 25. Evaluación cualitativa y cuantitativa de librerías mediante <i>Bioanalyzer 2100 HS</i>	60
Figura 26. Mapeo general de <i>reads</i> del control con cepas únicas	62
Figura 27. Reads mapeados del control de cepas únicas, comparación entre tratamientos	63

Figura 28. Reads mapeados del control de cepas únicas, compasión entre especies y genes.	64
Figura 29. Reads mapeados del control de cepas únicas, comparación entre <i>primers</i> por especie. ...	65
Figura 30. Reads mapeados del control de cepas únicas, comparación entre <i>sets</i> de <i>primers</i>	66
Figura 31. Resumen de la performance del panel en el control con cepas únicas	67
Figura 32. Mapeo general de <i>reads</i> del control con cepas múltiples.	68
Figura 33. Reads mapeados del control de cepas múltiples, comparación tratamiento/especie.....	69
Figura 34. Reads mapeados del control de cepas múltiples, comparación tratamiento/genes.	70
Figura 35. Porcentaje de <i>primers</i> con amplificación del control de cepas múltiples	71
Figura 36. Resumen del funcionamiento de <i>primers</i> en el control con cepas múltiples	72
Figura 37. Control de la estructura poblacional en el control C50. Estimadores de diversidad alfa	72
Figura 38. Estructura de la comunidad (ARNr16S) vs genes de virulencia detectados (panel Fvir) en el control C50	73
Figura 39. Correlaciones entre la aproximación del ARNr 16S y las copias de genomas iniciales o los <i>reads</i> mapeados en genes de virulencia en el tratamiento C50.....	74
Figura 40. Genes de virulencia bacterianos detectados en Gualaguaycito. Mapeo por gen	76
Figura 41. Genes de virulencia bacterianos detectados en Gualaguaycito. Mapeo por gen y funcionamiento de <i>primers</i>	77
Figura 42. Cobertura de los genes <i>mcv</i> por el panel Fvir en muestra Gualaguaycito	78
Figura 43. Control de la estructura poblacional en la muestra de Gualaguaycito. Estimadores de diversidad alfa.	79
Figura 44. Estructura poblacional en la muestra de Gualaguaycito. Principales Filos. Distribución de OTUs del Filo Cyanobacteria	79
Figura 45. Estructura de la comunidad (ARNr16S) vs genes de virulencia detectados (panel Fvir) en la muestra de Gualaguaycito.	80
Figura 46. Estructura de la comunidad (ARNr16S) en la muestra de Gualaguaycito géneros sin detección de genes de virulencia.....	81

FIGURAS ANEXOS:

Figura S1. Relación entre variabilidad genética del gen ARNr 16S y genes de virulencia.	116
Figura S2. Variabilidad del gen ARNr 16S vs genes de virulencia	117
Figura S3. Variabilidad genética del ARNr 16S vs genes de toxinas en cianobacterias.....	117
Figura S4. Variabilidad del gen ARNr 16S vs genes de virulencia de <i>E. coli</i>	118
Figura S5. Variabilidad en genes de virulencia de <i>E. coli</i> según su localización	118
Figura S6. Árbol filogenético construido a partir del gen <i>eaeA</i> para todos los genomas analizados.	119

LISTA DE ABREVIATURAS

- ADN: Ácido desoxirribonucleico
- ARN: Ácido ribonucleico
- Barcode: Código de barras
- CMA: Complejo *Microcystis aeruginosa*
- Chip: circuito integrado
- dNTPs: deoxinucleótidos tri fosfato
- G+C: referido al contenido en guanina y citosina
- ISP: perlas, del inglés *Ion Sphere™ Particle*
- MAC: Complejo *Mycobacterium avium*, del inglés *Mycobacterium avium complex*
- ND: no detectable
- NGS: Secuenciación de nueva generación, del inglés *Next Generation Sequencing*
- Panel: Conjunto de todos los pares de *primers*
- pb: Pares de bases
- PCR: Reacción en cadena de la polimerasa, del inglés *Polymerase Chain Reaction*
- PGM™: en referencia a plataforma de secuenciación Ion Torrent™, del inglés *Personal Genome Machine*
- Pool: Conjunto
- *Primers*: Cebadores u oligonucleótidos
- qPCR: PCR cuantitativo
- Read/s: Lectura/s de secuencia de ADN
- SNP: Polimorfismo de nucleótido simple, del inglés *Single Nucleotide Polymorphism*
- Targeted-NGS: Secuenciación dirigida
- Tm: Temperatura de disociación, del inglés *melting temperatura*
- *Trimming*: referido al control de calidad de secuencias filtrado o “recorte” de las lecturas
- UN: Organización de las Naciones Unidas (ONU), del inglés *United Nations*
- UNICEF: Fondo de las Naciones Unidas para la Infancia, del inglés *United Nations International Children's Emergency Fund*
- WGS: Secuenciación del genoma completo, del inglés *Whole Genome Sequencing*
- WHO: Organización Mundial de la Salud (OMS), del inglés *World Health Organization*

RESUMEN

Asegurar la calidad microbiológica del agua exige el uso de herramientas diagnósticas sensibles y específicas que garanticen la identificación de patógenos y sus capacidades patogénicas o tóxicas. Hasta la fecha los métodos de detección de bacterias patógenas en el agua se basan principalmente en microscopía y/o cultivos, contando con claras limitaciones. Los avances en la detección de patógenos, particularmente mediante técnicas genómicas como la NGS, que permiten un análisis masivo de ácidos nucleicos/amplicones ambientales ofrecen mejores perspectivas para la detección de estos microorganismos. A su vez el enriquecimiento selectivo mediante multiplex-PCR previo a la secuenciación (NGS-dirigida) permite que las regiones blanco sean secuenciadas con mayor profundidad, independientemente del “fondo” complejo de la muestra.

El objetivo de este trabajo fue contribuir a la mejora del diagnóstico de patógenos bacterianos en sistemas acuáticos utilizando secuenciación dirigida a genes de virulencia. Para ello, seleccionamos 72 genes de virulencia de más de 50 especies bacterianas de relevancia para los sistemas acuáticos. La diversidad de los genes candidatos es vasta e incluye varios mecanismos de virulencia, múltiples genes de un clúster, variantes más relevantes descritas, así como elementos codificados en islas de patogenicidad, plásmidos y fagos. A partir de un análisis comparativo/evolutivo de las secuencias disponibles en la base de datos establecimos regiones de conservación en 39 de los 72 genes candidatos. Los restantes 33 genes presentaron un grado de variabilidad genética alta y como alternativa para el diseño de *primers* se utilizó el genoma de una cepa de referencia. Nuestra estrategia en ambos casos apuntó a la generación de múltiples *primers* dirigidos a la cobertura completa de cada gen candidato. Mediante el software *Ampliseq Designer*[™] diseñamos un panel de *primers* Ampliseq[™], el panel Fvir, alcanzando una alta cobertura sobre la secuencia de los genes candidatos (97,3 %). Nuestro panel, consta de 548 juegos de *primers*, que cubren 71 genes de virulencia bacterianos, compatibles en dos reacciones de PCR-multiplex (284 y 264 juegos respectivamente) y asociados al sistema de secuenciación Ion Torrent (PGM). Para verificar que los *primers* sintetizados estén dentro de los parámetros razonables para la PCR-multiplex se realizó la evaluación *in silico* del panel. No existieron diferencias en la longitud, contenido en G+C y Tm entre los dos pooles, indicando un correcto balance entre ellos. Los *primers* del panel presentaron, baja formación de dímeros de/entre *primers* y horquillas, así como Tm adecuadas para la PCR. En relación con el mapeo, el panel mostró niveles elevados de sensibilidad (100 %) y especificidad (99,5 %) sobre la secuencia de los genes candidatos. La capacidad de detección de los *primers* se evaluó utilizando ADN de cepas control a través de dos ensayos. El primer ensayo se basó en un número reducido de cepas (*E. coli*, *L. monocytogenes*, *P. aeruginosa* y *S. enterica* Typhimurium), variando la concentración de ADN molde y los ciclos de PCR iniciales. El segundo ensayo se basó en una mezcla de 92 aislamientos pertenecientes a los géneros *Escherichia*, *Salmonella*, *Listeria*, *Pseudomonas*, *Enterococcus*, *Staphylococcus*, *Aeromonas* y *Cylindrospermopsis*. En el primer ensayo el panel logró detectar los 8 genes candidatos referidos a las cepas utilizadas en un rango estimado entre 30-30.000 copias de genomas/μL, amplificando el 87,5 % de los *primers* diseñados y alcanzando niveles altos de sensibilidad (95,2 %) y especificidad (99,5 %). En el segundo ensayo el panel Fvir detectó 25 genes asociados a los géneros que conformaron el pool de ADN, 16 de ellos caracterizados anteriormente y 9 genes

cuya presencia en las cepas era desconocida pero factible. Asimismo, el porcentaje de *primers* que amplificaron en las muestras de este ensayo osciló entre 69 y 80 %. Por último, en este ensayo se determinó que el enriquecimiento de un microorganismo en la muestra (50 % del ADN total) afecta la detección de otros menos representados.

Por otra parte, se evaluó la utilidad diagnóstica del panel en una muestra del Río Uruguay durante un evento de floración del Complejo *Microcystis aeruginosa* (CMA). La muestra, previamente caracterizada, era abundante en genotipos tóxicos productores de microcistinas (gen *mcy*). Los resultados mostraron que más del 99,8 % del total de *reads* correspondieron a los genes *mcyE* y *mcyJ*, mostrando amplificación con todos los juegos de *primers* diseñados para estos genes (*mcyE*= 37 juegos, *mcyJ*= 4 juegos). Además, se detectaron 16 genes de virulencia que presentaron muy baja abundancia (mapeo inferior al 1 % de los *reads*) e incluyeron genes que codifican para otras cianotoxinas, así como también otros genes de virulencia de grupos de bacterias heterótrofas.

Paralelamente a la estrategia de Ampliseq™, se puso a punto y validó el control de la estructura poblacional mediante secuenciación masiva del ARNr 16S. Ésta se utilizó para contrastar los resultados obtenidos por el panel en las muestras control y ambiental. Todos los genes de virulencia detectados (panel Fvir) se pudieron relacionar a posibles taxones presentes en la muestra (ARNr16 S) la mayoría a nivel de género. En comparación con el ARNr 16S la secuenciación dirigida a genes de virulencia presentó un mayor valor diagnóstico de la muestra, ya que fue posible determinar el potencial patogénico de algunos géneros presente.

En suma, en este trabajo se diseñó un panel Ampliseq™ dirigido a 71 genes de virulencia bacterianos con alta capacidad de detección de los genes candidatos presentes. La utilidad diagnóstica del panel fue comprobada en una muestra ambiental, brindando información más profunda sobre los genotipos tóxicos presentes en una floración del CMA. El funcionamiento general del panel presentó resultados alentadores en cuanto a su capacidad de detección, sensibilidad y especificidad. La amplificación dirigida a genes de virulencia bacterianos seguida de NGS se presenta como una estrategia novedosa e innovadora que contribuye a los actuales sistemas de diagnóstico de bacterias patógenas en sistemas acuáticos.

Palabras Clave:

Secuenciación dirigida, Bacterias Patógenas, Genes de virulencia, Sistemas acuáticos, Herramientas diagnósticas.

ABSTRACT

The microbiological quality of water requires sensitive and specific diagnostic tools to identify pathogens and their pathogenic or toxic capacities. Nowadays, detecting pathogenic bacteria in the water is based mainly on microscopy and cultures, having clear limitations. Advances in detecting pathogens, particularly by genomic techniques, like next generation sequencing (NGS), allow a massive analysis of PCR amplicons to offer the best prospects for detecting these microorganisms. Moreover, the selective enrichment by multiplex-PCR before sequencing (targeted-NGS) allows the target regions to be sequenced in greater depth, regardless of the sample's complex "background."

This work aimed to improve the diagnosis of bacterial pathogens in aquatic systems using NGS targeting virulence genes. Thereby, we selected 72 virulence genes across fifty bacterial relevant species for aquatic systems. The diversity of target genes is extensive and includes several virulence mechanisms, multiple genes in a cluster, more relevant variants described, as well as elements encoded in pathogenicity islands, plasmids, and phages. From a comparative/evolutionary analysis of the sequences available in the database, we established conserved regions in 39 of the 72 target genes. The remaining 33 genes presented a high level of genetic variability. In this case, for primers design, the genome of a reference strain was used. For both cases, the strategy was to generate several primers aimed to obtain complete coverage of each target gene. Using the Ampliseq Designer™ software, we designed a panel of Ampliseq™ primers, the Fvir panel, achieving high coverage on the sequence of target genes (97.3 %). Our panel consisted of 548 sets of primers, covering 71 bacterial virulence genes, compatible in two multiplex-PCR reactions (284 and 264 sets, respectively) and associated with the Ion Torrent sequencing system (PGM). Using the Ampliseq Designer™ software, we designed a panel of Ampliseq™ primers, the Fvir panel, achieving high coverage on the sequence of target genes (97.3 %). To verify reasonable parameters for multiplex-PCR, the panel was submitted for in silico evaluation. There were no differences in the length, G+C content, and Tm between the two pools, indicating a correct balance. The panel showed low levels of dimers in/between primers and hairpins and adequate Tm for the PCR. The panel showed high levels of sensitive (100 %) and specificity (99.5 %) on the target genes' sequences concerning mapping. The detection capacity of the primers was evaluated using DNA from control strains by two assays. The first assay was performed on a reduced number of strains (*E. coli*, *L. monocytogenes*, *P. aeruginosa*, and *S. enterica* Typhimurium), varying the template DNA concentration and initial PCR cycles. The second assay consisted of 92 isolates belonging to the genus *Escherichia*, *Salmonella*, *Listeria*, *Pseudomonas*, *Enterococcus*, *Staphylococcus*, *Aeromonas*, and *Cylindrospermopsis*. In the first experiment, the panel managed to detect the eight target genes referred to the strains, used in an estimated range between 30-30,000 genome copies/μL, amplified 87.5 % of the designed primers. It reached high levels of sensitivity (95.2 %) and specificity (99.5 %). In the second experiment, the Fvir panel detected twenty-five genes associated with the genus that made up the DNA pool, sixteen of them previously characterized, and nine genes whose presence in the strains was unknown but feasible. Besides, the percentage of primers amplifying products in the samples of this test ranged between 69 and 80 %. Finally, in this test, it was determined that the enrichment of a microorganism in the sample (50 % of the total DNA) affects the detection of other less represented ones.

On the other hand, the panel's diagnostic utility was evaluated in a sample from the Uruguay River during a *Microcystis aeruginosa* complex (MAC) bloom event. The sample, previously characterized, was abundant in microcystin-producing toxic genotypes (*mcy* genes). The results showed that more than 99.8 % of the total *reads* corresponded to the *mcyE* and *mcyJ* genes, showing amplification with all primers sets designed for these genes (*mcyE* = 37, *mcyJ* = 4). Also, 16 virulence genes were detected in low abundance (mapping less than 1 % of *reads*). They included genes codifying for other cyanotoxins and virulence genes from heterotrophic bacteria.

Parallel to the Ampliseq™ strategy, the bacterial structure community was evaluated by massive 16S rRNA sequencing and was used to contrast the panel results in the control and environmental samples. All the virulence genes detected (panel Fvir) could be related to possible taxa present in the sample (rRNA16 S), most of them at the genus level. Compared with 16S rRNA, virulence gene sequencing presented a more accurate diagnostic value of the sample since it was possible to determine some genera's pathogenic potential.

In summary, in this work, an Ampliseq™ panel was designed targeting 71 bacterial virulence genes with a high detection capacity of the genes present in the analyzed samples. The panel's diagnostic utility was verified in an environmental sample, providing more information on the toxic genotypes present in a MAC bloom. The panel's performance showed encouraging results in terms of its detection capacity, sensitivity, and specificity. Amplification targeting virulence genes followed by NGS is presented as a novel and innovative strategy that contributes to the current diagnostic systems for pathogenic bacteria in aquatic systems.

Keywords:

Targeted-NGS, Pathogenic bacteria, Virulence genes, Aquatic environments, Diagnostic tools.

1. INTRODUCCIÓN

1.1. Contaminación de Los Recursos Hídricos Por Agentes Patógenos

1.1.1. Impacto Global de Las Enfermedades Transmitidas Por el Agua

El agua, recurso natural imprescindible, tiene una importancia vital para todos los sectores socioeconómicos y la pérdida de calidad de ésta da lugar a problemas de salud, económicos y ambientales en todo el mundo (World Health Organization, 2017; Magana-Arachchi & Wanigatunge, 2020). Los cuerpos de agua dulce superficiales, incluyendo lagos, ríos y arroyos, constituyen importantes ecosistemas acuáticos y son una fuente de agua potable en muchos países incluido Uruguay (Tan *et al.*, 2015). El agua potable es definida como aquella “adecuada para el consumo humano y para todo uso doméstico habitual, incluida la higiene personal” siendo libre de microorganismos causantes de enfermedades (World Health Organization, 2017). En el año 2010, la Asamblea General de las Naciones Unidas declara unívocamente el derecho humano al agua y al saneamiento, reconociendo que estos son esenciales para la realización plena de todos los derechos humanos (UNGA, 2010). Sin embargo, debido a la insuficiencia, inaccesibilidad, indisponibilidad o la no seguridad del agua, la mayoría de las personas a nivel mundial se ven privadas de este derecho universal (Magna-Arachchi & Wanigatunge, 2020). Según el informe del Programa de Monitoreo Conjunto OMS/UNICEF, aproximadamente 3 de cada 10 personas en todo el mundo (2.1 mil millones), no tienen acceso a agua segura o esta es de difícil disponibilidad en su hogar y 6 de cada 10 (4.5 mil millones), carecen de saneamiento administrado de manera segura (WHO/UNICEF, 2017). Es así como, en 2015, los estados miembros de las Naciones Unidas adoptaron en la agenda 2030 para el desarrollo sostenible como uno de los objetivos el “Garantizar la disponibilidad y la gestión sostenible del agua y el saneamiento para todos” (UNSD, 2015).

Sumado a lo anterior, el saneamiento deficiente y el agua contaminada están vinculados a la transmisión de enfermedades vehiculizadas por el agua, como el cólera, la disentería, la hepatitis A y la fiebre tifoidea (WHO/UNICEF, 2017). En este marco, sólo en 2016, el agua, el saneamiento y la higiene fueron responsables de 829.000 muertes por enfermedad diarreica y se estima que cada año 361.000 niños menores de 5 años mueren a causa de diarrea en todo el mundo (WHO/UNICEF, 2017). Estas enfermedades constituyen una preocupación importante de la salud pública a nivel mundial, no sólo por la morbilidad y la mortalidad que causan, sino por el alto costo que representa su prevención y tratamiento (Ramírez-Castillo *et al.*, 2015). Se estima que las enfermedades transmitidas por el agua tienen un costo económico asociado de mil millones de dólares anuales sólo en los Estados Unidos (Collier *et al.*, 2012), mientras que a nivel mundial, la pérdida económica ronda los 12 mil millones de dólares por año (Alhamlan, Al-Qahtani & Al-Ahdal, 2015).

En este contexto, el aumento del tamaño de la población humana y la urbanización en las últimas décadas han ejercido una inmensa presión sobre el uso de los recursos hídricos para el consumo y la recreación. El mal uso de estos recursos ha tenido como consecuencia el agotamiento de las existencias en algunos lugares, la disminución de las capas freáticas y la reducción de los flujos de corriente (Ward & Trimble, 2004; UNESCO, 2016). En este sentido, la presencia o el aumento de bacterias, parásitos, virus y hongos en el agua, surge usualmente por efecto directo o indirecto de cambios

en el medio ambiente y en la población. Entre estos cambios se destacan: la urbanización no controlada, el crecimiento industrial, la pobreza, la ocupación de regiones antes deshabitadas, y la disposición inadecuada de excretas humanas y animales. Estos cambios relacionados con las actividades antropogénicas se ven reflejados directamente en el entorno y, por consiguiente, en el recurso hídrico (Ramírez-Castillo *et al.*, 2015; Ríos, Agudelo & Gutiérrez, 2017).

El clima es otro factor clave que contribuye a la aparición de brotes de estas enfermedades, por ejemplo, introduce contaminantes en las fuentes de agua por la escorrentía de una fuerte lluvia o inundación (Ramírez-Castillo *et al.*, 2015). En este sentido, el aumento de las precipitaciones como consecuencia del cambio climático puede desencadenar una secuencia de eventos en cascada de magnitud significativa con consecuencias directas sobre las enfermedades transmitidas por el agua (Semenza, 2020). En la Figura 1 se esquematiza los posibles riesgos y rutas desencadenadas por el cambio climático en relacionan a las enfermedades transmitidas por el agua.



Figura 1. Principales riesgos generados por el cambio climático y que afectan a las enfermedades transmitidas por el agua (efecto cascada) (Adaptado de Semenza, 2020). El aumento de la temperatura media global de la superficie debido al cambio climático acelera la evaporación y aumenta el contenido de agua de la atmósfera (1). Como resultado, la precipitación promedio mundial ha aumentado (2). Los eventos de lluvia extrema pueden movilizar materia fecal de la vida silvestre o de los animales domésticos, que pueden ingresar y superar la capacidad de las plantas de tratamiento de agua (3). Los microorganismos patógenos de origen fecal animal (o humano) pueden infiltrarse en el sistema de distribución de agua y causar un brote transmitido por el agua (4). Durante eventos de lluvia tan extremos, las instalaciones de tratamiento de alcantarillado aguas abajo también pueden verse abrumadas, lo que da como resultado un desbordamiento combinado de alcantarillado que descarga las aguas residuales sin tratar y tratadas directamente en los ríos y arroyos cercanos (5). Las inundaciones pueden dañar las plantas de alcantarillado o contaminar los pozos de agua subterránea con patógenos transmitidos por el agua, y los contaminantes de la cría de animales pueden contaminar el acuífero (6). Los incendios forestales durante las olas de calor pueden degradar la calidad del agua del río e impedir el tratamiento del agua (7).

Los cambios en la temperatura pueden alterar la dinámica de los microorganismos patógenos determinando su crecimiento en forma planctónica o formando *biofilms* (Ingerson-Mahar & Reid, 2012). Las fuertes lluvias, las inundaciones y el clima cálido están asociados con estas enfermedades. Las vías de transmisión de las enfermedades asociadas al agua, no sólo se ven afectadas por los cambios en la temperatura media y la precipitación, sino también por fenómenos meteorológicos repentinos y atípicos (Semenza, 2020).

Además, la contaminación por aguas residuales, sustancias químicas tóxicas, nutrientes y las proliferaciones de algas nocivas resultantes pueden hacer que el agua no sea apta para el consumo humano o para actividades recreativas (Tan *et al.*, 2015). Dado que algunos microorganismos son causantes de toxicidad o patologías severas, el monitoreo de este tipo de microorganismos es fundamental para la detección temprana, el control, el manejo paliativo y el rastreo de la fuente de la contaminación.

1.1.2. Enfermedades Microbianas Asociadas al Agua

Uno de los mayores problemas enfrentados a nivel mundial, en términos de calidad de agua, es la contaminación de cuerpos de agua ambientales con patógenos transmitidos por el agua y las enfermedades vinculadas a su presencia (World Health Organization, 2017). A pesar de los continuos esfuerzos para mantener la seguridad del agua, todavía se informan brotes de enfermedades transmitidas por el agua en todo el mundo (Ramírez-Castillo *et al.*, 2015). La contaminación por patógenos es un problema para casi todos los tipos de cuerpos de agua ambientales, por lo que su reconocimiento y comprensión son esenciales (US-EPA, 2012).

Se entiende como patógeno al agente causal de enfermedad a un huésped (generalmente organismos vivos), siendo los patógenos transmitidos por el agua aquellos causantes de enfermedades que se transmiten a través de ésta (Magana-Arachchi & Wanigatunge, 2020). Se estima que hay 1407 especies de patógenos que infectan a los humanos, que incluyen bacterias (538 especies), virus (208 tipos), protozoarios parásitos (57 especies) y varias especies de hongos y helmintos (Woolhouse, 2006; Bitton, 2014).

Clásicamente los patógenos asociados con el agua han sido categorizados en cuatro grupos de acuerdo a la forma de transmisión de la enfermedad (White *et al.*, 1972; Bartram *et al.*, 2015; Forstinus *et al.*, 2016; Gerba & Pepper, 2019; Oleiwi, 2020):

Enfermedades propagadas por el agua (Water-borne diseases): son aquellas transmitidas por la ingestión de agentes patógenos en el agua contaminada. En estos casos el agua sirve como transporte pasivo del agente infeccioso. Históricamente, el cólera y la fiebre tifoidea son las enfermedades propagadas por el agua con mayor mortalidad y con frecuencia asolaron áreas densamente pobladas a lo largo de la historia humana (Bartram *et al.*, 2015; Forstinus *et al.*, 2016). Se transmiten a través de la ruta fecal-oral, de humano a humano o de animal a humano, de modo que el agua es sólo una de varias fuentes posibles de infección (Oleiwi, 2020). El agua contaminada puede albergar una variedad de patógenos, como *Escherichia coli*, *Salmonella*, *Shigella*, *Campylobacter*, *Yersinia*, *Vibrio*, enterovirus, norovirus, adenovirus, coxsackievirus, echovirus y virus de hepatitis A y E (HAV, HEV), *Squitosoma*, *Dracunculus medinensis*, *Echinococcus granulosus*, *Entamoeba*

histolytica, *Cryptosporidium parvum* y *Giardia lamblia* (Lipp, Farrah & Rose, 2001; Symonds, Griffin & Breitbart, 2009). La mayoría de estos microorganismos son agentes infecciosos causantes de enfermedades intestinales (entéricas) y se transmiten a través de los desechos fecales. Estas enfermedades son más frecuentes en áreas con malas condiciones sanitarias, ya que estos patógenos viajan a través de fuentes de agua y se propagan a través de personas que manejan alimentos y agua. Las fuentes de estos patógenos infecciosos se originan principalmente de una fuente puntual, como la descarga de aguas residuales, o de fuentes no puntuales, como la agricultura, la vida silvestre y la escorrentía urbana (Alhamlan, Al-Qahtani & Al-Ahdal, 2015). Una descarga puntual se puede manejar fácilmente tratando la fuente (es decir, las aguas residuales); sin embargo, las fuentes no puntuales se consideran una amenaza debido a su amplia difusión (Stewart *et al.*, 2008).

Enfermedades “lavadas” (por contacto o escasez) con el agua (Water-washed diseases): son aquellas estrechamente relacionadas con la falta de higiene y prosperan en condiciones de escasez de agua dulce y/o saneamiento deficiente. En este caso, la disponibilidad de agua es un factor determinante y el control de la enfermedad depende más de la cantidad de agua que de la calidad de esta (Forstinus *et al.*, 2016). A diferencia de las enfermedades propagadas por el agua, el rol de esta es evitar la difusión de la enfermedad, más que actuar como un vehículo de transporte de patógenos (Gerba & Pepper, 2019). Ejemplos de estas enfermedades incluyen: sarna, tifus, pian, fiebre recurrente, impétigo, tracoma, conjuntivitis y úlceras cutáneas. Estas enfermedades representan una de las principales causas de morbilidad y mortalidad infantil en los países en desarrollo (Forstinus *et al.*, 2016).

Enfermedades con base en el agua (Water-based diseases): son causadas por agentes patógenos que pasan toda o parte esencial de su vida en el agua o dependen de organismos acuáticos para completar sus ciclos de vida. La infección puede ocurrir por ingestión, inhalación o el microorganismo puede atravesar la piel intacta o dañada (percutánea). Ejemplos de estos microorganismos son parásitos helmintos como varias especies de *Schistosoma* y bacterianos como *Legionella spp.* y *Leptospira sp.* (Bartram *et al.*, 2015; Kulinkina *et al.*, 2016).

Enfermedades relacionadas con el agua (Water-related diseases): son aquellas enfermedades causadas por insectos vectores que se reproducen dentro o alrededor de los cuerpos de agua. Enfermedades como la fiebre amarilla (*Flavivirus*), el dengue (*Flavivirus*), la filariasis (nemátodos superfamilia *Filarioidea*), la malaria (*Plasmodium*), la oncocercosis (*Onchocerca volvulus*) y la enfermedad del sueño (*Trypanosoma*) (Gerba & Pepper, 2019). Estas enfermedades no están directamente relacionadas con la calidad del agua potable, sin embargo, la consideración del control de vectores durante el diseño, construcción y operación de reservorios y canales de agua superficial (para agua potable o para riego) puede reducir el potencial de transmisión (Forstinus *et al.*, 2016).

Un quinto grupo ha sido propuesto por Bradley, (2009) y constituye a las *Enfermedades transmitidas por inhalación*. Este grupo también se ha denominado como “asociado a la ingeniería del sistema de agua”. El principal peligro son las bacterias del género *Legionella spp.*, patógenos causales de la legionelosis, o enfermedad del legionario. Estos microorganismos proliferan formando *biofilms*, particularmente en los sistemas de ingeniería vinculados al procesamiento y

distribución de agua donde las condiciones de temperatura y nutrientes contribuyen a su crecimiento (Oleiwi, 2020).

1.2. Bacterias Transmitidas Por el Agua

Los patógenos bacterianos son agentes etiológicos clásicos de las enfermedades propagadas por el agua a nivel mundial y son responsables de más del 50 % de las enfermedades infecciosas emergentes (Magana-Arachchi & Wanigatunge, 2020). Estos organismos pueden aparecer de manera ubicua en muchos hábitats acuáticos y sistemas de potabilización y/o distribución. En este trabajo nos centraremos en aquellas bacterias patógenas ubicuas asociados a los distintos sistemas de agua.

1.2.1. Persistencia y Crecimiento en Diferentes Sistemas Acuáticos

Como se mencionó anteriormente las enfermedades asociadas con el agua incluyen una diversidad de patógenos. Desde un punto de vista ecológico, algunos de estos patógenos pasan la mayor parte de sus ciclos de vida en el ambiente acuático y sólo accidentalmente encuentran un hospedero. Por ello, están típicamente bien adaptados a concentraciones bajas de nutrientes y a las condiciones fisicoquímicas que encuentran en el agua. Estos denominados “patógenos ambientales” son bacterias como *Legionella* spp., *Pseudomonas aeruginosa* y algunas especies de micobacterias no tuberculosas. Estos patógenos se caracterizan por su independencia del hospedero, para el cual son facultativos. Otros patógenos transmitidos por el agua son los denominados obligados, ya que su propagación sólo puede ocurrir dentro de un hospedero infectado, típicamente se replican dentro del intestino del individuo infectado. Ejemplos de éstos son los patógenos entéricos *Campylobacter*, *Salmonella* o los virus entéricos. Los patógenos obligados dependen del hospedero para ser diseminados en el ambiente y alcanzar así otros hospederos, estos patógenos transmitidos ambientalmente tienen en general dos estilos de vida, uno dentro del hospedero y otro en el ambiente. Entre los nichos que estos patógenos pueden ocupar cuando están fuera del hospedero, el agua juega un rol importantísimo para muchos de ellos (Nocker, Burr & Camper, 2013).

En cuanto a los sistemas de potabilización y de distribución, si bien los patógenos típicos propagados por el agua pueden persistir en el agua potable, la mayoría no crece ni prolifera. Los microorganismos como *E. coli* y *Campylobacter* pueden acumularse en sedimentos y se movilizan cuando aumenta el flujo de agua. Después de abandonar al huésped, la mayoría de los patógenos pierden gradualmente la viabilidad y la capacidad de infectar. La tasa de degradación suele ser exponencial, y un patógeno se volverá indetectable después de un cierto período de tiempo (World Health Organization, 2017). La persistencia se ve afectada por varios factores, de los cuales la temperatura es la más importante. La degradación suele ser más rápida a temperaturas más altas y puede estar mediada por los efectos de la radiación UV solar que actúa en la superficie del agua. Los patógenos y parásitos transmitidos por el agua más comunes son aquellos que tienen una alta infectividad y pueden proliferar en el agua o poseer una alta resistencia a la descomposición fuera del cuerpo del huésped. Por el contrario, cantidades relativamente altas de carbono orgánico biodegradable, junto con temperaturas cálidas y bajas concentraciones residuales de cloro pueden permitir el crecimiento de *Legionella*, *V. cholerae* y otros

microorganismos en algunas aguas superficiales y durante la distribución del agua (World Health Organization, 2017).

Otro aspecto importante es la capacidad de formación de *biofilms* de estos microorganismos en los sistemas de agua potable. Los patógenos que resisten el proceso de desinfección tienden a persistir en las tuberías de distribución de agua formando comunidades microbianas en *biofilms*, lo que conduce a su difusión hacia los usuarios finales por un proceso de desprendimiento posterior (Figueras & Borrego 2010). Esto constituye un reservorio ambiental de patógenos y es una fuente persistente de contaminación del agua, resultando en un riesgo potencial para la salud de no ser detectado (Wingender & Flemming, 2011). Entre los microorganismos con este potencial se incluyen bacterias indicadoras fecales (*Escherichia coli*), patógenos bacterianos obligados de origen fecal (*Campylobacter spp.*) y bacterias oportunistas de origen ambiental (*Legionella spp.*, *Pseudomonas aeruginosa*) entre otros (Wingender & Flemming, 2011). El agua de pozo no tratada también representa un gran riesgo, patógenos como *Legionella* y *Campylobacter* han sido descritos como los principales agentes etiológicos transmitidos por el agua en los Estados Unidos (Yoder *et al.* 2011).

1.2.2. Bacterias Patógenas Transmitidas Por el Agua Potable

La OMS establece en sus directrices para la calidad del agua potable, una lista de agentes etiológicos para una cantidad sustancial de enfermedades transmitidas por el agua. En ella, incluye patógenos "clásicos" para los cuales considera hay alguna evidencia de importancia para la salud, por ejemplo: la gravedad del impacto y su asociación con brotes de enfermedad, persistencia en los sistemas acuáticos, resistencia al tratamiento con cloro o la tasa de infectividad entre otros (World Health Organization, 2017). En la Tabla 1 se resumen las principales bacterias patógenas relacionadas con los sistemas de agua potable y sus enfermedades relacionadas de acuerdo con la OMS.

Sin embargo, en la actualidad, nuevos organismos o nuevas cepas de agentes patógenos ya conocidos han sido identificadas y representan importantes desafíos adicionales para los sectores vinculados al agua y a la salud pública. Entre las bacterias patógenas, se incluyen *Helicobacter pylori*, *Tsukamurella*, especies de *Bacillus* y cianobacterias tóxicas a las que se debe prestar especial atención para lograr suministrar agua limpia, prevenir percances y proteger y mejorar aspectos relacionados a la salud pública (Magana-Arachchi & Wanigatunge, 2020). En la Tabla 2 se enumeran algunas de las especies bacterianas patógenas para los cuales se ha propuesto la transmisión a través del agua potable, pero para los cuales la evidencia no es concluyente según la OMS (World Health Organization, 2017).

Tabla 1. Lista de bacterias patógenas transmitidas por el agua potable más relevantes según la Organización Mundial de la Salud (OMS)¹

Patógeno ²	Enfermedad asociada	Importancia salud ³	Persistencia agua ⁴	Resistencia al Cl ⁵	Infectividad ⁶	Fuente Animal
<i>Burkholderia pseudomallei</i>	Melioidosis	Alta	Puede multiplicarse	Baja	Baja	No
<i>Campylobacter</i> (<i>C. jejuni</i> , <i>C. coli</i>)	Gastroenteritis	Alta	Moderada	Baja	Moderada	Si
<i>E. coli</i> diarreogénicas ⁷	Gastroenteritis	Alta	Moderada	Baja	Baja	Si
<i>E. coli</i> enterohemorrágica (O157:H7)	Gastroenteritis Síndrome Urémico hemolítico	Alta	Moderada	Baja	Alta	Si
<i>Legionella</i> spp. (<i>L. pneumophila</i>)	Legionelosis Enfermedad del legionario	Alta	Puede multiplicarse	Baja	Moderada	No
<i>Francisella</i> (<i>F. tularensis</i>)	Tularemia Enfermedad ulcero-glandular Linfadenopatía Enfermedad pulmonar	Alta	Larga	Moderada	Alta	Si
Micobacterias no tuberculosas	Enfermedad pulmonar Infección de la piel	Baja	Puede multiplicarse	Alta	Baja	No
<i>Salmonella typhi</i>	Fiebre tifoidea	Alta	Moderada	Baja	Baja	No
Otras <i>Salmonelleae</i> (<i>S. enterica</i> , <i>S. bongori</i>)	Salmonellosis	Alta	Puede multiplicarse	Baja	Baja	Si
<i>Shigella</i> (<i>S. dysenteriae</i>)	Shigellosis	Alta	Corta	Baja	Alta	No
<i>Vibrio cholerae</i> (Serogrupo O1 y O139) ⁸	Cólera	Alta	Bioacumula	Baja	Baja	No

¹Adaptado de las directrices de la OMS para la calidad del agua potable (World Health Organization, 2017). ²Las especies enumeradas son comúnmente relacionadas con la transmisión por agua (otras especies también pueden causar enfermedades). ³Se relaciona con la gravedad del impacto y asociación con brotes. ⁴Período de detección para el estadio infeccioso en agua a 20° C: Corto, hasta una semana; Moderado, una semana a un mes; Largo, mayor a un mes. ⁵Dentro de las especies y grupos, es probable que haya variaciones en la resistencia (puede verse afectado por las características del suministro de agua y las condiciones de operación). La resistencia se basa en una inactivación del 99 % a 20° C donde, bajo=Ct₉₉ de <1 min.mg/L, moderado 1–30 min.mg/L y alto> 30 min.mg/L (C= la concentración de cloro libre en mg/L y t= tiempo de contacto en minutos) en las siguientes condiciones: la etapa infecciosa se suspende libremente en agua y el pH está entre 7 y 8. Los organismos que sobreviven y crecen en *biofilms*, como *Legionella* y micobacterias pueden estar protegidos contra la cloración. ⁶De acuerdo con experimentos en humanos, de evidencia epidemiológica y de estudios experimentales con animales. Alto significa que las dosis infecciosas pueden ser de 1-10² organismos o partículas, moderadas de 10² - 10⁴ y bajas > 10⁴. ⁷Incluye los grupos patogénicos: enterotoxigénica (ETEC), enteroinvasiva (EIEC), enteropatógena (EPEC), enteroagregativa (EAEC) y de adherencia difusa (DAEC). ⁸*Vibrio cholerae* puede persistir por largos períodos en asociación con copépodos y otros organismos acuáticos.

Tabla 2. Lista de bacterias patógenas para los cuales se ha propuesto la transmisión a través del agua potable, pero para los cuales la evidencia no es concluyente o no existe según la OMS¹

Microorganismo (especie/género/ grupo) ²	Evidencia de transmisión por agua (características epidemiológicas)	Presencia y comportamiento en el suministro de agua	Resistencia al cloro ³
<i>Acinetobacter</i> (Complejo <i>A. baumannii-calcoaceticus</i>)	Posible problema en los centros de salud (no gastrointestinales)	Común y puede multiplicarse	Baja
<i>Aeromonas</i> (<i>A. hydrophila</i>)	Los aislamientos clínicos no coinciden con los aislamientos ambientales	Común y puede multiplicarse	Baja
<i>Enterobacter</i> (<i>E. sakazakii</i>)	Infección asociada con fórmula infantil; no hay evidencia de transmisión por agua	Improbable	Baja
<i>Helicobacter</i> (<i>H. Pylori</i>)	Sugerido, pero sin evidencia directa; ruta primaria de transmisión familiar	Detectado, sobrevive por tiempo limitado	Baja
<i>Klebsiella</i> (<i>K. pneumoniae</i>)	Posible problema en los centros de salud (no gastrointestinales)	Puede multiplicarse	Baja
<i>Leptospira</i> (<i>L. interrogans</i>)	No hay evidencia de transmisión por ingestión de agua potable. Se propaga principalmente por contacto con aguas superficiales contaminadas; brotes asociados con inundaciones	Puede sobrevivir durante meses	Baja
<i>Pseudomonas</i> (<i>P. aeruginosa</i>)	Posible problema en los centros de salud (no gastrointestinales)	Común y puede multiplicarse	Moderada
<i>Staphylococcus</i> (<i>S. aureus</i>)	No hay evidencia de transmisión a través del agua potable; las manos son la fuente más importante	Común y puede multiplicarse	Moderada
<i>Tsukamurella</i> (<i>T. paurometabola</i>)	Posible problema en los centros de salud (no gastrointestinales)	Común y puede multiplicarse	Desconocido
<i>Yersinia</i> (<i>Y. enterocolitica</i>)	Especies detectadas en el agua probablemente no patógenas; la comida es la fuente primaria	Común y puede multiplicarse	Baja

¹Adaptado de las directrices de la OMS para la calidad del agua potable (World Health Organization, 2017). ²Las especies enumeradas son las más comúnmente relacionadas con la transmisión por agua, pero otras especies también pueden causar enfermedades. ³La resistencia se basa en una inactivación del 99 % a 20 ° C donde, bajo representa un Ct99 de <1 min.mg/L, moderado 1–30 min.mg/L y alto > 30 min.mg/L (donde C= la concentración de cloro libre en mg/L y t= tiempo de contacto en minutos) en las siguientes condiciones: la etapa infecciosa se suspende libremente en agua tratada a dosis y tiempos de contacto convencionales, y el pH está entre 7 y 8. Debe ser señalado que los organismos que sobreviven y crecen en *biofilms*, como *Pseudomonas aeruginosa*, pueden estar protegidos contra la cloración.

1.2.3. Eutrofización, Floraciones Tóxicas y Cianobacterias

El aumento de las actividades productivas y el incremento poblacional han alterado el funcionamiento y estructura de los principales ecosistemas acuáticos y terrestres a nivel mundial (Paerl, 2017). El cambio en el uso del suelo como resultado de la intensificación y expansión de las actividades agropecuarias e industriales (situación que también ocurre en Uruguay, Aubriot *et al.*, 2018) han incrementado el ingreso de nitrógeno, fósforo y diferentes compuestos contaminantes a los cursos de agua. Estos elementos son responsables de la eutrofización de los ecosistemas acuáticos (enriquecimiento excesivo de nutrientes de los cuerpos de agua) siendo una de las principales causas del aumento de la frecuencia e intensidad de las floraciones

algales (Mazzeo *et al.*, 2002; Aubriot *et al.*, 2005; Chalar, 2009; Kruk *et al.*, 2013). Como consecuencia, aparece un crecimiento explosivo de determinadas especies de cianobacterias en cortos periodos de tiempo, fenómeno denominado floración (Kruk *et al.*, 2015; Martínez de la Escalera *et al.*, 2017).

Las cianobacterias son microorganismos procariotas fotosintéticos, se encuentran formando parte del fitoplancton y son ubicuas en la naturaleza, habitando una amplia gama de hábitats acuáticos y terrestres, incluso ambientes extremos como desiertos, aguas termales y ambientes helados (Whitton y Potts, 2000; Rastogi, Madamwar y Incharoensakdi, 2015). Tienen una antigüedad datada en 3500 millones de años y su larga historia de vida es una de las principales razones de su éxito evolutivo actual (Dokulil & Teubner, 2000; Reynolds, 2006). Las floraciones son de duración variable y ocurren en aguas ricas en nutrientes, particularmente fosfatos, nitratos y amonio, cuando la temperatura es alta y/o cuando las condiciones hidrológicas son favorables. Su creciente incidencia es una gran preocupación, ya que no solamente alteran la calidad del agua, modificando el color y el sabor de esta, sino que, además, tienen la capacidad de producir toxinas (cianotoxinas) (Chorus & Bartram, 1999; Carmichael, 2001; Chorus, 2001).

Diversas especies de cianobacterias planctónicas de los tres Órdenes: Chroococcales, Oscillatoriales y Nostocales, producen floraciones potencialmente tóxicas en cuerpos de agua de todo el mundo (Whitton & Potts, 2000). Entre las especies más comunes se encuentran *Microcystis aeruginosa* y otras especies y géneros similares que comparten características fenotípicas y origen filogenético y se agrupan en el complejo *Microcystis aeruginosa* (CMA) (Otsuka *et al.*, 2000; Komárek & Komárková, 2002). El CMA es frecuente en cuerpos de agua continentales y costeros del mundo (Harke *et al.*, 2016), de la región (Andrinolo *et al.*, 2007; Dörr *et al.*, 2010; O'Farrell, Bordet & Chaparro, 2012; Sathicq, Bauer & Gómez, 2015; Bordet, Fontanarrosa & O'Farrell, 2017; Aguilera *et al.*, 2018) y de nuestro país (Ferrari & Vidal, 2006; Vidal & Britos, 2012; González-Piana *et al.*, 2017; Kruk *et al.*, 2019; Deus-Álvarez *et al.*, 2020). Estas floraciones producen en general las cianotoxinas denominadas microcistinas y son primariamente hepatotóxicas. Las microcistinas se producen en otra gran variedad de géneros (ej: *Planktothrix*, *Oscillatoria*, *Anabaena*, *Nostoc*, *Anabaenopsis*, *Aphanocapsa*), siendo la toxina más frecuente (De León & Yunes, 2001; Brena, *et al.*, 2006; Andrinolo, *et al.*, 2007; Feola, *et al.*, 2010; Pérez, *et al.*, 2013; Bonilla, *et al.*, 2015; Kruk, *et al.*, 2015).

Otra especie que ha ganado la atención científica reciente es *Cylindrospermopsis raciborskii* (Orden Nostocales). Esta especie es de origen tropical, pero ha comenzado a invadir latitudes más frías. Si bien puede producir dos potentes toxinas (cylindrospermopsina y saxitoxina), las poblaciones americanas sólo han registrado saxitoxinas (neurotoxinas). En este sentido, se han reportado saxitoxinas en cepas de *C. raciborskii* aisladas de diversos lagos de nuestro país (Piccini, *et al.*, 2011).

En Uruguay las floraciones de cianobacterias son frecuentes en cuerpos de agua superficiales y han sido documentadas por varios estudios académicos (Revisado por Bonilla *et al.*, 2015). Así, se han descrito en: lagos artificiales (Bonilla, *et al.*, 1995; Vidal & Kruk, 2008; Fabre, *et al.*, 2010; Aubriot & Bonilla, 2012), lagunas costeras (Pérez, *et al.*, 1999; Bonilla & Conde, 2000; Bonilla, *et al.*, 2006; Vidal & Kruk, 2008; Conde, *et al.*, 2009; Vidal, *et al.*, 2009; Pacheco, *et al.*, 2010), ríos y embalses (Quirós

& Luchini, 1982, Ferrari, *et al.*, 2011; O'Farrell, *et al.*, 2012; Chalar, *et al.*, 2014) y costas del Río de la Plata (Ferrari & Vidal, 2006; Sienra & Ferrari, 2006; Kruk *et al.*, 2017; Kruk *et al.*, 2019; Deus-Álvarez *et al.*, 2020). A su vez, se han detectado microcistinas en floraciones del CMA en playas del Río de la Plata (De León & Yunes, 2001; Sienra & Ferrari, 2006; Pírez, *et al.*, 2013; Martínez de la Escalera *et al.*, 2017; Kruk *et al.*, 2019; Deus-Álvarez *et al.*, 2020), en embalses del Río Negro (González-Piana, *et al.*, 2011), muestras del Río Uruguay con diversas especies de cianobacterias (Gravier, *et al.*, 2009; Martínez de la Escalera *et al.*, 2017; Deus-Álvarez *et al.*, 2020) y diversos cuerpos de agua (microcistina-LR) (Simoens, 2009).

Otro aspecto relevante es que algunos episodios de floraciones se registraron en ecosistemas acuáticos destinados a la potabilización de agua. Entre los más destacados, en el año 2013, ocurre en la cuenca del río Santa Lucía la aparición de una floración de *Dolichospermum sp.* con claros síntomas de deterioro ambiental (Aubriot *et al.*, 2018). Este río y sus afluentes son la fuente de agua potable para 1.700.000 uruguayos, siendo, además, fuente de agua de riego para una de las zonas de mayor actividad agroindustrial en el país. Esta floración fue acompañada de una alta producción de geosmina, que otorgó mal sabor y olor al agua, generando una alarma pública principalmente debido a las características que tenía el agua que llegaba a los hogares. Otros episodios de floración en fuentes de agua potable ocurrieron en otras zonas del país (Aubriot *et al.*, 2018) como la Laguna del Sauce, (Bonilla, *et al.*, 2015) y el Río Negro (Chalar, *et al.*, 2015; González-Piana, *et al.*, 2017).

Entre los eventos más recientes, en el verano del 2019 ocurrió una de las floraciones de mayor intensidad registrada en la costa estuarina y Atlántica de Uruguay (Kruk *et al.*, 2019). Las floraciones fueron extensas (Carmelo a Rocha; 500 km) y persistentes (ca.: 4 meses), Figura 2. En todos los casos fueron organismos del CMA con estructura genética similar y su presencia estuvo asociada a niveles cuantificables y elevados de Microcistina (Kruk *et al.*, 2019; Kruk *et al.*, 2021).



Figura 2. Floraciones de cianobacterias en la costa estuarina y oceánica de Uruguay en el verano de 2019 (tomado de Kruk *et al.*, 2019). Distribución espacio-temporal (día/mes): 1) Nueva Palmira: 9/1; 2) Balneario Zagarzazú: 15/1 y Colonia 16/1; 3) Montevideo 28/1; 4) Maldonado 29/1 y 5) Rocha La Paloma 30/1. A) El Ensueño, Colonia; B) costa de Montevideo 10/2; C) Rambla de Montevideo; D) vista aérea playa de Piriápolis; E) playa La Balconada zona cercana al Faro de La Paloma (30/1), y F) Laguna de Rocha (31/1).

1.2.4. Factores de Virulencia Bacterianos

La patogenicidad de un microorganismo se define como "la cualidad o el estado de ser patógeno, es el potencial o la capacidad de producir enfermedades", mientras que la virulencia se define como "el poder productor de enfermedades de un organismo, e indica el grado de patogenicidad dentro de un grupo o especie" (Shapiro-Ilan *et al.*, 2005). La virulencia es entonces una característica medible de la capacidad del microorganismo para causar enfermedades y está destinada a comparaciones dentro de un grupo o dentro de una especie particular (Casadevall & Pirofski, 2003). Cada uno de los atributos microbiológicos que contribuyen a la virulencia en general se puede vincular a elementos arquitectónicos específicos o compuestos bioquímicos dentro del organismo. Juntos, estos elementos y compuestos generalmente pueden denominarse "factores de virulencia", y los planos para ellos se incluyen en el código genético de un organismo (National Research Council, 2001). Los factores de virulencia como elementos codificados por genes (Chen *et al.*, 2005) pueden dividirse en varias categorías, por ejemplo, sobre la base del mecanismo de virulencia al que se relacionan, como la adherencia, colonización, evasión inmunitaria, sistema de secreción, invasión, producción de toxinas, absorción de hierro, etc. (Tu *et al.*, 2014).

Desde un punto de vista funcional, Wassenaar & Gaastra, (2001) clasifican a los genes de virulencia en tres clases:

La clase I consta de genes de virulencia clásicos, también denominados "verdaderos" ya que están directamente involucrados en la interacción huésped-patógeno, son responsables del daño al huésped y están ausentes en microorganismos no patógenos. Incluye genes que codifican para la producción de toxinas, factores de adherencia y de invasión.

La clase II consta de genes que codifican factores esenciales para la actividad de los factores de virulencia clásicos (es decir, sistemas de secreción y/o reguladores de la expresión génica).

La clase III consta de genes relacionados con el estilo de vida específico de un patógeno (genes que codifican mecanismos de absorción y/o factores implicados en la colonización del huésped y la evasión de la respuesta inmune).

Esta clasificación ha sido sugerida por Tournlousse *et al.*, (2007) como potencialmente útil para la evaluación de riesgos de los patógenos vinculados al agua. Estos autores sugieren que se podría adoptar, por ejemplo, para el ranqueo de los genes en base a su importancia, donde los de la clase I recibirían la puntuación más alta.

En relación con las bacterias patógenas mencionadas en este trabajo, numerosos genes de virulencia de las tres clases han sido descritos y los mecanismos asociados con su patogenicidad o toxicidad son conocidos para muchas de las especies (World Health Organization, 2017). En la Tabla 3, se presenta a modo ilustrativo, la clasificación mencionada para algunos de estos genes.

Tabla 3. Genes de virulencia de bacterias asociadas al agua (adaptado de Tourlousse *et al.*, 2007). Se presentan algunos ejemplos de genes y su clasificación funcional*

Patógeno	Clase I	Clase II	Clase III
<i>E. coli</i> <i>EHEC</i>	Hemolisina (hlyA), toxina Shiga (<i>stx1A</i> , <i>stx1B</i> , <i>stx2A</i> , <i>stx2B</i>), intimina (<i>eae</i>), receptor de intimina (<i>tir</i>)	Sistemas de secreción tipo III (<i>escC</i> , <i>espA</i>), secreción y activación de hemolisinas (<i>hlyB</i> , <i>hlyC</i> , <i>hlyD</i> , <i>tolC</i>)	Mecanismo absorción hierro (<i>chuA</i>)
<i>Salmonella</i> <i>spp.</i>	Fimbria tipo I (<i>fimA</i> , <i>fimD</i>), curli (<i>agf</i>), fimbria polar mayor (<i>lpfA</i> , <i>lpfB</i>), proteínas de secreción (<i>sipB</i> , <i>sipC</i>)	Reguladores de dos componentes (<i>phoP</i> , <i>phoQ</i>), regulador de absorción de hierro (<i>fur</i>), sistema secreción tipo III (<i>invA</i> , <i>invE</i>)	
<i>Yersinia</i> <i>enterocolitica</i>	Enterotoxina termoestable (<i>yst</i>), factor de adhesión (<i>yadA</i>), locus fijación invasión (<i>ail</i>), invasina (<i>inv</i>)	Sistemas de secreción tipo III (<i>yopB</i> , <i>yscL</i> , <i>virG</i>), reguladores (<i>ysaV</i> , <i>ysaW</i> , <i>yspB</i> , <i>yspC</i>)	Síntesis de Yersiniabactina (<i>irp3</i> , <i>irp9</i>), Yersiniabactina receptor (<i>fyuA</i>)
<i>Aeromonas</i> <i>hydrophila</i>	Aerolisina (<i>aerA</i>), hemolisina (<i>hlyA</i>), enterotoxinas citotónica (<i>act</i>), termoestable (<i>ast</i>), termolábil (<i>alt</i>), pili tipo IV (<i>tapA</i> , <i>tapB</i> , <i>tapC</i> , <i>tapD</i>)	Sistema de secreción tipo III (<i>aopB</i> , <i>axsC</i> y <i>ascL</i>), tipo II (<i>exeD</i> y <i>exel</i>)	
<i>Staphylococcus</i> <i>aureus</i>	Hemolisinas (<i>hla</i> , <i>hlyA</i> , <i>hlyB</i> , <i>hlyC</i>), toxina exfoliativa (<i>etA</i> , <i>etB</i>), leucocidinas (<i>lukF</i> , <i>lukS</i>), enterotoxinas (<i>seA</i> - <i>seR</i> sin <i>seF</i>), toxina del shock séptico (<i>tsst-1</i>), proteínas de unión al colágeno (<i>cna</i>)	Reguladores (<i>sarA</i> , <i>agr</i>), de dos componentes (<i>lytR</i> , <i>lytS</i> ; <i>arlR</i> , <i>arlS</i> , <i>saeR</i> , <i>saeS</i>)	Mecanismo absorción hierro (<i>isdE</i> , <i>isdG</i>)
<i>Helicobacter</i> <i>pylori</i>	Citotoxina vacuolante (<i>vacA</i>), genes reguladores de citotoxina (<i>cagA</i>), factor de adhesión (<i>babA</i>)	Sistema de secreción tipo IV (<i>virD4</i> , <i>virB11</i>)	
<i>Pseudomonas</i> <i>aeruginosa</i>	Exoenzimas (<i>exoA</i> , <i>exoU</i> , <i>exoT</i>), pili tipo IV (<i>pilA</i> , <i>pilB</i> , <i>pilC</i> , <i>pilD</i>)	Sistema de secreción tipo III (<i>popB</i> , <i>exsC</i> , <i>pilC</i>), secreción tipo II (<i>xcpQ</i> , <i>xcpV</i>)	
<i>Vibrio</i> <i>cholerae</i>	Toxina colérica (<i>ctxA</i> , <i>ctxB</i>), zónula de unión celular (<i>zot</i>), gen accesorio a la toxina (<i>stx</i>), pili tipo IV (<i>pilA</i> , <i>pilB</i> , <i>pilC</i> y <i>pilD</i>)	Sistema secreción tipo II (<i>epsD</i> , <i>epsI</i>), (<i>toxR</i> , <i>toxS</i>)	
<i>Listeria</i> <i>monocytogenes</i>	Listeriolisina (<i>hlyA</i>), internalinas (<i>inlA</i> , <i>inlB</i>), Proteínas asociada a la invasión (<i>iap</i>)	Regulador (<i>prfA</i>)	
<i>Legionella</i> <i>pneumophila</i>	Toxina RTX (<i>rtxA</i>), pili tipo IV (<i>pilB</i> , <i>pilC</i> , <i>pilD</i> , <i>pilE</i>)	Sistemas de secreción: tipo IVb (<i>dotA</i> , <i>icmB</i> , <i>icmR</i>), tipo II (<i>lspD</i> , <i>lspH</i>), tipo I (<i>lssB</i> , <i>lssD</i>)	Mecanismo absorción hierro (<i>iraA</i> , <i>iraB</i> , <i>feoA</i> , <i>feoB</i>)

*Los genes se clasifican según su función descrita por Wassenaar & Gaastra, (2001) y se dividen en tres categorías: genes de virulencia clásicos (clase I), genes que codifican factores esenciales para la actividad de los factores de virulencia clásicos (clase II) y genes relacionado con el estilo de vida específico de un patógeno (clase III).

1.3. Métodos Para la Detección de Patógenos en Agua

La OMS define la vigilancia y el control del agua para consumo humano como la “evaluación y examen, de forma continua y vigilante, considerando la Salud Pública, la inocuidad y la aceptabilidad de los sistemas de abastecimiento de agua” (Ríos, Agudelo & Gutiérrez, 2017; World Health Organization, 2017). Este marco incluye conocer la calidad del agua en sus fuentes y sistemas de potabilización e identificar los microorganismos presentes en ella. Esta vigilancia exige el uso de herramientas diagnósticas para identificar los microorganismos presentes; sin embargo, hay gran limitación para determinar la totalidad de agentes involucrados en su contaminación, tanto en las fuentes de abastecimiento como en los sistemas de tratamiento del agua (Ríos, Agudelo & Gutiérrez, 2017).

1.3.1. Métodos clásicos

Durante décadas, la seguridad microbiológica del agua para fines potables y recreativos se ha monitoreado mediante la enumeración y la detección de bacterias indicadoras fecales, por ejemplo, coliformes totales, *Escherichia coli* o *Enterococcus spp.* (Figueras & Borrego, 2010). En este sentido, diversas técnicas han sido empleadas, como por ejemplo, la fermentación por tubos múltiples (Número más probable), filtración por membrana y/o el cultivo bacteriano estándar en placa (Deshmukh *et al.*, 2016). Sin embargo, cultivar bacterias patógenas es una tarea laboriosa, que incluye el enriquecimiento y el crecimiento en medios selectivos, con el objetivo de separar uno o un conjunto de cepas o especies específicas del contexto bacteriano existente en la muestra (Sohier *et al.*, 2014; Deshmukh *et al.*, 2016). En algunos casos las concentraciones son muy pequeñas como para lograr el cultivo, pero suficientes para causar infección (Oliver, 2010).

Por otro lado, muchas bacterias, incluyendo importantes patógenos humanos, responden al estrés ambiental generando formas viables, pero no cultivables (VPNC) (Oliver, 2010). En algunos casos, como el de *Helicobacter pylori* y *Vibrio cholerae*, la formación de VPNC hace imposible su detección por cultivos (Oliver, 2010; Sohier *et al.*, 2014). Entre los patógenos que entran en estado VPNC se encuentran *Campylobacter spp.*, *Escherichia coli* (incluyendo cepas enterohemorrágicas, EHEC), *Francisella tularensis*, *Helicobacter pylori*, *Legionella pneumophila*, *Listeria monocytogenes*, *M. tuberculosis*, *Pseudomonas aeruginosa*, varias cepas de *Salmonella* y *Shigella spp.*, así como numerosos *Vibrio spp.* (Xu *et al.*, 1982; Lindbäck *et al.*, 2010; Oliver, 2010).

En relación con las cianobacterias, actualmente en nuestro país la detección se realiza mediante la inspección microscópica, la cual es trabajosa e insume mucho tiempo. Además, la ocurrencia de especies problema no es condición suficiente para la presencia de toxinas en el agua y la detección en base a características ambientales es difícil (Huisman & Weissing, 2001). Tradicionalmente se han utilizado estimadores globales de la comunidad del fitoplancton como indicadores de calidad de agua y para la predicción, como la concentración de pigmentos, en particular la clorofila-a (Vollenweider, 1976; Scheffer *et al.*, 2003). Estos indicadores no consideran diferentes respuestas y efectos (toxicidad, riesgo sanitario) causados por la presencia de distintas especies o grupos. Es necesario desarrollar, entonces, herramientas que contemplen los mecanismos ecológicos que controlan el funcionamiento de las

comunidades (Le Quéré *et al.*, 2005) pero a su vez sintetizan el comportamiento de los cientos de especies de fitoplancton que pueden ocurrir simultáneamente (Kruk *et al.*, 2011). Los métodos ideales para cada tipo de ecosistema o uso del agua varían, siendo más crítica la detección directa de cianotoxinas en una planta de potabilización de agua que en aguas recreacionales (Brena & Bonilla 2009; Fries *et al.* 2001).

Entre los métodos de detección y cuantificación de cianotoxinas, principalmente microcistina, se encuentran el ensayo de inhibición de fosfatasa (Chorus, 2001), ELISA (McDermott, Feola & Plude, 1995) y HPLC (Lawton, Edwards & Codd, 1994), LC-MS (Robillot *et al.*, 2000), y NMR (Harada *et al.*, 1995). Todas estas técnicas tienen sus ventajas y desventajas, siendo el ELISA y la inhibición de fosfatasa los más sensibles, con algunas desventajas como reactividad cruzada, altos costos y uso de elementos radiactivos en algunos casos.

1.3.2. Aplicación de Métodos de Secuenciación Masiva al Estudio de Patógenos

El desarrollo de métodos moleculares independientes del cultivo, para la detección y cuantificación de bacterias específicas ha sido empleado para evitar los problemas asociados con bacterias difíciles de cultivar, que requieren de un huésped o que entran en forma VPNC (Ramamurthy *et al.*, 2014).

La PCR se ha convertido en la prueba “gold-estándar” para la detección de muchos patógenos virales, fúngicos, parasitarios y bacterianos, superando en sensibilidad a los métodos clásicos de aislamiento viral y cultivo bacteriano (van Elden *et al.*, 2002; Yang & Rothman, 2004). Si bien los ensayos de PCR son muy sensibles y específicos, su costo puede elevarse especialmente cuando se realizan pruebas con múltiples patógenos (Laupland & Valiquette, 2013). De manera rutinaria, esto se hace utilizando muchas PCR simples o algunas PCR multiplex. La PCR multiplex reduce los costos, pero el método se limita a una pequeña cantidad de patógenos que pueden detectarse en cada prueba, potencialmente a expensas de la sensibilidad para cada patógeno (Wernike, Hoffmann & Beer, 2015).

A medida que aumenta la disponibilidad de nuevas tecnologías, existe el potencial para el desarrollo de mejores pruebas de diagnóstico que expanden los límites de la PCR. Una de estas tecnologías es la secuenciación masiva (también denominada como “*Next Generation Sequencing*”, NGS) (Tan *et al.*, 2015).

Desde el punto de vista del diagnóstico, los métodos derivados del uso de secuenciación masiva permiten la detección simultánea de segmentos de ADN únicos y característicos. En este sentido, la detección por identidad de secuencia es más específica e informativa que la PCR, lo que resulta muy útil para detectar diferentes especies y/o cepas de microorganismos en distintas muestras biológicas (Deurenberg *et al.*, 2017). A su vez, han demostrado excelentes rendimientos, dadas sus longitudes de lectura más largas y mayor precisión, utilizando aplicaciones que requieren menos consumibles y menos mano de obra que los métodos de secuenciación tradicionales (Caliendo *et al.*, 2013). Estos sistemas basados en secuenciación podrían reemplazar los procedimientos de identificación microbiológica tradicionales y utilizarse para la detección de elementos asociados a la virulencia y a marcadores genéticos de resistencia a los antimicrobianos (RAM), entre otras aplicaciones (Caliendo *et al.*, 2013; Deurenberg *et al.*, 2017).

Entre las técnicas derivadas de NGS con posible aplicación para el diagnóstico se encuentra la metagenómica (Kwong *et al.*, 2015) que se basa en la secuenciación de todo el ADN presente en una muestra. Esto representa una limitante, ya que la mayor parte de los recursos utilizados (99 %) terminan en la amplificación del genoma del hospedero, que es más abundante en la mayoría de las muestras biológicas, y se pierde sensibilidad de forma significativa (Prachayangprecha *et al.*, 2014; Motro & Moran-Gilad, 2017).

Una alternativa a la secuenciación metagenómica es la NGS target o dirigida (*targeted next-generation sequencing*). Ésta se refiere a la amplificación selectiva de regiones genómicas específicas de interés antes de la secuenciación masiva paralela (Khodakov, Wang & Zhang, 2016), lo que ofrece algunas ventajas en comparación por ejemplo con la metagenómica. El uso de la amplificación dirigida y el enriquecimiento/captura específicos, aseguran que las regiones más discriminantes seleccionadas en los genomas blanco sean secuenciadas independientemente del “fondo” complejo de la muestra. Este abordaje posibilita la secuenciación de regiones únicas del microorganismo y/o regiones que contienen genes funcionalmente importantes o SNPs que discriminen filogenéticamente (Gardner *et al.*, 2015). Si bien la posibilidad de la NGS target para descubrir nuevos patógenos es limitada, la secuenciación selectiva proporciona mejor sensibilidad, especificidad, facilidad de análisis posterior y tiene un menor costo al permitir el análisis de un número muy alto de muestras en una sola corrida (Dong *et al.*, 2015; Gardner *et al.*, 2015; Brinkmann *et al.*, 2017; Anis *et al.*, 2018).

En relación con el enriquecimiento, la tecnología Ampliseq™ de *Thermo Fisher Scientific*, permite la amplificación simultánea de cientos/miles de productos de PCR (multiplex de alto rendimiento). Además, presenta otras ventajas, como requisitos de ADN input bajos (1 a 10 ng) y un protocolo de trabajo relativamente rápido (3-6 h). El flujograma de trabajo de Ampliseq se puede resumir como un protocolo de tres pasos: la amplificación por PCR-multiplex específica, digestión parcial de los amplicones y el ligado de secuencias adaptadoras (Khodakov, Wang & Zhang, 2016).

Esta aproximación se ha aplicado con éxito en el diagnóstico clínico en humanos, utilizándose para la detección de genes y tipificación de SNPs, de enfermedades como el cáncer (Belardinilli *et al.*, 2015; Hovelson *et al.*, 2015; Péterfia *et al.*, 2017; Rathi *et al.*, 2017; Lee *et al.*, 2018), patologías cardíacas (Blue *et al.*, 2014; Millat, Chanavat & Rousson, 2014; Cho *et al.*, 2017), desordenes hematológicos (He *et al.*, 2017; Jang *et al.*, 2019; Steinberg-Shemer & Tamary, 2020) o con aplicaciones en medicina forense (Seo *et al.*, 2013; Børsting *et al.*, 2014; Minogue, Koehler & Norwood, 2017; Zhang *et al.*, 2017; Sobiah *et al.*, 2018) entre otros. Incluso existe una variedad de paneles comerciales disponibles para diferentes tipos de cáncer aprobados para su uso diagnóstico por la Administración de Drogas y Alimentos de los Estados Unidos (FDA) (Rebollar, Arriaga & de la Rosa, 2018).

En las áreas relacionadas al diagnóstico microbiológico varios trabajos reportan su implementación y diferentes enfoques han sido descritos. Así por ejemplo, se ha utilizado en el diagnóstico clínico viral en humanos, para la detección de especies/cepas y tipificación de sus variantes (Brinkmann *et al.*, 2017) o más recientemente en el diagnóstico y estudios epidemiológicos de SARS-CoV-2

(Alessandrini *et al.*, 2020; Goletic *et al.*, 2020; Stenhoff *et al.*, 2020) incluso para su detección en el agua de ríos y aguas residuales (Rimoldi *et al.*, 2020).

En bacterias, el enfoque se ha centrado en el análisis de comunidades microbianas a partir de la secuenciación de amplicones de marcadores específicos como el gen que codifica para el ARNr 16S (Caporaso *et al.*, 2010) y a partir de ella explorar la diversidad de posibles taxones patógenos. Esta metodología se ha aplicado en el estudio de patógenos en comunidades de muestras de ríos (Ibekwe, Leddy & Murinda, 2013), agua potable (Vierheilig *et al.*, 2015), aguas residuales (Ye & Zhang, 2011; Kumaraswamy *et al.*, 2014; Lu *et al.*, 2015) y biosólidos (Bibby, Viau & Peccia, 2010; Luo & Angelidaki, 2014). Aunque esta metodología es de utilidad para estimar la comunidad microbiana, los resultados sólo infieren un perfil semicuantitativo de los potenciales patógenos. Este hecho limita su aplicación práctica para el monitoreo de patógenos, debiéndose asociar con otras técnicas que logren la cuantificación de los patógenos con buena precisión, como por Ej. qPCR dirigida a genes de virulencia y así suplementar la información filogenética con información del potencial patogénico (Cui *et al.*, 2017).

En los últimos años, varios trabajos se han centrado en el empleo de NGS dirigida a genes de virulencia, con resultados muy alentadores. En estos trabajos, estos genes han sido empleados para el estudio de bacterias asociadas al bioterrorismo (Gardner *et al.*, 2015), enfermedades infecciosas del ganado (Anis *et al.*, 2018) o su empleo conjunto con genes de resistencia a antimicrobianos (Park *et al.*, 2018; Ko *et al.*, 2019).

Todas estas ventajas sugieren que la NGS dirigida a genes de virulencia se puede utilizar para pruebas de detección de patógenos más completas que las actuales. Dado que hasta el momento no existe un panel comercial que contemple la detección de un rango amplio de patógenos bacterianos, mucho menos aquellos asociados a sistemas acuáticos, buscamos diseñar y probar un panel piloto AmpliSeq™ que los contemple.

2. OBJETIVOS Y ESTRATEGIA DE TRABAJO

2.1. Objetivo General

Contribuir a la mejora del diagnóstico de patógenos bacterianos en sistemas acuáticos utilizando secuenciación dirigida a genes de virulencia y NGS.

2.2. Objetivos Específicos

- Diseñar un panel de *primers* AmpliSeq™ capaz de amplificar genes involucrados en la virulencia bacteriana para su implementación en el diagnóstico molecular por NGS.
- Evaluar el sistema de amplificación y la posterior NGS en controles con ADN de cepas potencialmente patogénicas.
- Valorar la utilidad diagnóstica del panel en una muestra ambiental previamente caracterizada.
- Implementar y validar la secuenciación de ARNr 16S y su posterior análisis para suplementar la información del panel.

2.3. Estrategia de Investigación

Para alcanzar el objetivo, se tomó una aproximación molecular basada en el análisis de secuencias de ADN características de especies y/o cepas y de regiones genómicas responsables de la patogenia y/o toxicidad de un grupo de organismos relevantes para la salud pública. La estrategia de investigación utilizada para el abordaje de este problema se resume en la Figura 3.

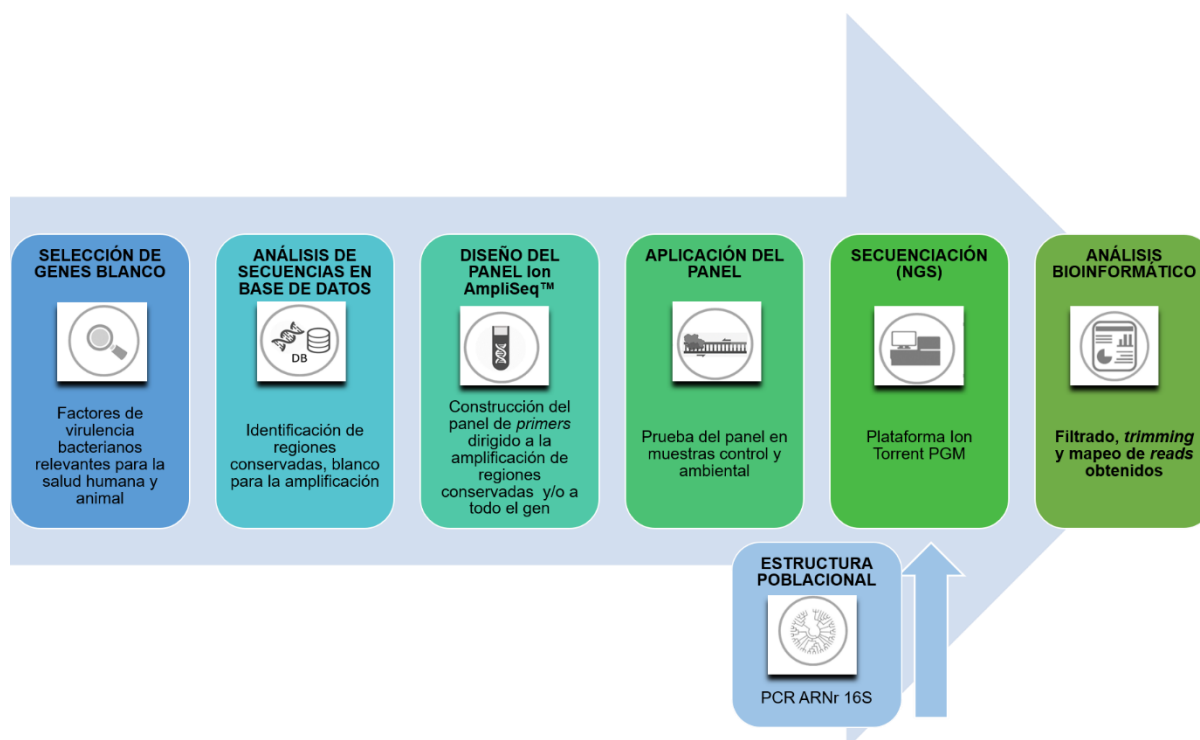


Figura 3. Flujograma general de la estrategia definida en el presente trabajo para el diseño y la aplicación del panel de *primers* AmpliSeq™.

Para ello se hizo uso del secuenciador disponible en el IIBCE, el cual al igual que otros secuenciadores de segunda generación, se caracteriza por la generación masiva de secuencias en periodos cortos de tiempo. El diseño de las regiones a secuenciar se realizó en base a información identificatoria (secuencia) de los marcadores de virulencia. El sistema que permite la amplificación simultanea de cientos de productos de PCR y su secuenciación conjunta se conoce con el nombre comercial de AmpliSeq™. Las propiedades esenciales son las siguientes: diseño de cientos de pares de *primers* compatibles para realizar una reacción de PCR multiplex para que puedan ser secuenciados asociado al sistema de secuenciación Ion Torrent (ThermoFisher Scientific). El fabricante garantiza la bioinformática del diseño y la integración al sistema del secuenciador.

3. MATERIALES Y MÉTODOS

3.1. Diseño Del Sistema de Amplificación de Genes de Virulencia Bacterianos

3.1.1. Selección de Agentes Patógenos y Genes Candidatos

Para la selección de los agentes patógenos y los genes vinculados a la patogenicidad/toxicidad bacteriana, se procedió en primer lugar a la revisión bibliográfica de la temática, realizándose la selección de los agentes patógenos en base al grado de patogenia e incidencia de la enfermedad causada. Como primera aproximación se optó por incluir únicamente patógenos bacterianos dada su incidencia como agentes causales de enfermedad, los recursos y experiencia del equipo de investigación del proyecto. Se confeccionó un catálogo de los posibles patógenos y genes definidos como blanco.

3.1.2. Evaluación de la Conservación en Los Genes de Virulencia

Con la finalidad de identificar regiones conservadas en los genes de virulencia seleccionados, blanco para el diseño de *primers* de AmpliSeq™, se realizó un análisis comparativo/evolutivo de las secuencias disponibles en la base de datos de NCBI-GenBank. Utilizando la secuencia de proteínas de referencia del factor de virulencia blanco, se realizó la búsqueda por similitud con la herramienta BLAST (tblastn). Las secuencias que cumplieron con los criterios de: identidad $\geq 70\%$ y cobertura $\geq 95\%$ fueron descargadas y se utilizaron para crear una base de datos para cada gen. Posteriormente se realizó el alineamiento y el filtrado de secuencias repetidas con el programa MUSCLE en MEGA v.7 (Edgar, 2004). La identificación de bloques conservados se realizó utilizando la herramienta Gblocks® v.0.91b, seleccionando la opción más rigurosa de no permitir posiciones contiguas no conservadas (Castresana, 2000). Finalmente la selección de una secuencia consenso representante del bloque conservado se realizó en BioEdit v.7.2.6 (Hall, 1999) (Figura 4).

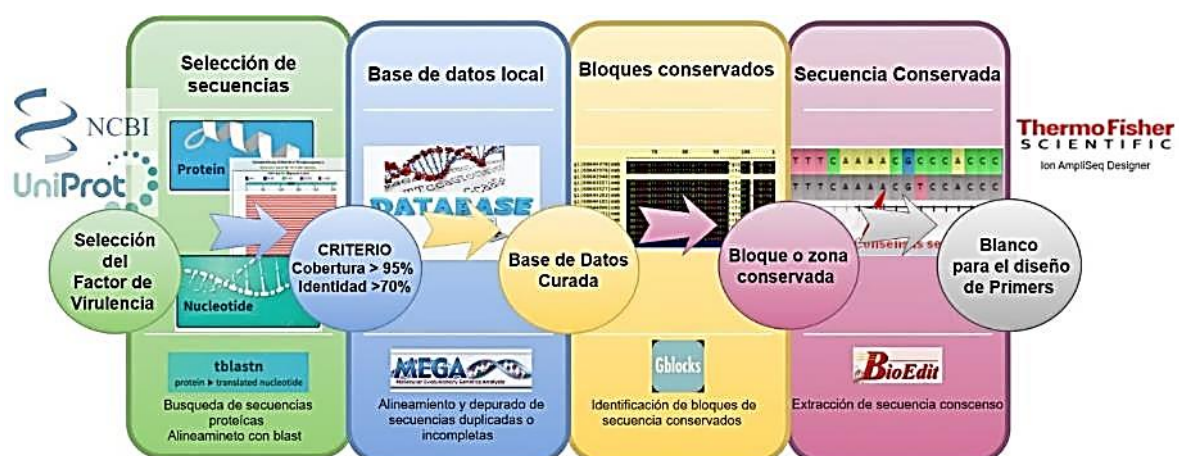


Figura 4. Pasos secuenciales para la determinación de bloques de secuencias conservadas en los genes de virulencia seleccionados.

Con la base de datos de secuencias generada, se analizó en MEGA v.7, la taxonomía correspondiente a las secuencias de cada gen (géneros/especies anotadas) y se estimó la distancia evolutiva (número de sustituciones: transiciones + transversiones)

utilizando el método de Máxima Verosimilitud Compuesta (MLC), (Nei & Kumar, 2000).

Como metodología alternativa para aquellos genes que no mostraron una región conservada suficientemente extensa para su amplificación, se utilizaron genomas de referencia mejor representados (cepas de colecciones de referencia vinculadas al gen de patogenicidad) y se optó por el diseño de múltiples juegos de *primers* capaces de amplificar todo el gen blanco.

En la Figura 5 se esquematiza de manera ilustrativa el diseño de *primers* utilizado en un gen de virulencia con una zona conservada. Una vez identificadas las zonas conservadas para cada gen o identificados los genes sin región conservada se procedió a la construcción del panel de *primers* AmpliSeq™.

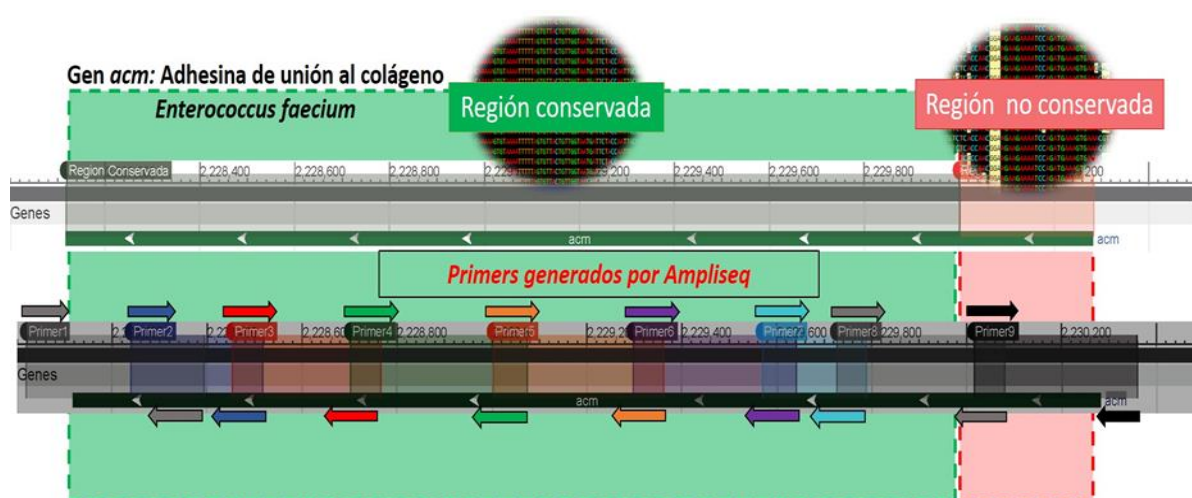


Figura 5. Esquema de la localización de *primers* en un gen seleccionado. La figura muestra la conservación en la secuencia (regiones conservada y no conservada) y localización de los juegos de *primers* generados sobre ambas regiones.

3.1.3. Construcción del Panel de Primers AmpliSeq

Como se mencionó anteriormente el sistema de AmpliSeq™ permite la amplificación simultánea de cientos de productos de PCR (multiplex de alto rendimiento) para que puedan ser secuenciados asociado al sistema de secuenciación, en este caso Ion Torrent. Para lograr el diseño correcto de los *primers* el proveedor proporciona acceso gratuito al software *Ion AmpliSeq™ Designer* (<https://www.ampliseq.com/>) que utiliza algoritmos que logran un alto nivel de especificidad y rendimiento de los paneles generados. Esta herramienta permite la utilización de paneles de ADN y ARN prediseñados, su modificación o construcción *de novo* con diseño propio "customized". Para el diseño de los *primers* se procedió a la confección de un panel utilizando las secuencias de los genes de virulencia anteriormente analizadas y de acuerdo con el manual de uso del software (Ion AmpliSeq Designer™, 2018).

La delimitación de las regiones de interés se realizó sobre genomas de referencia en los que se sustituyeron las secuencias de los genes blanco por aquellas secuencias consenso generadas a través del estudio de conservación. Estos genomas, se utilizaron para crear un archivo FASTA que fue cargado en el software como

“referencia personalizada”. Se recomienda la utilización de un genoma de referencia completo para evitar generar *primers* con posibles falsos positivos, *hits* fuera del blanco y otros errores relacionados con el mapeo.

Una vez cargado el archivo con los genomas de referencia, se creó un archivo CSV con las coordenadas de los genes blanco y se procedió al diseño del panel por el software (Figura 6).

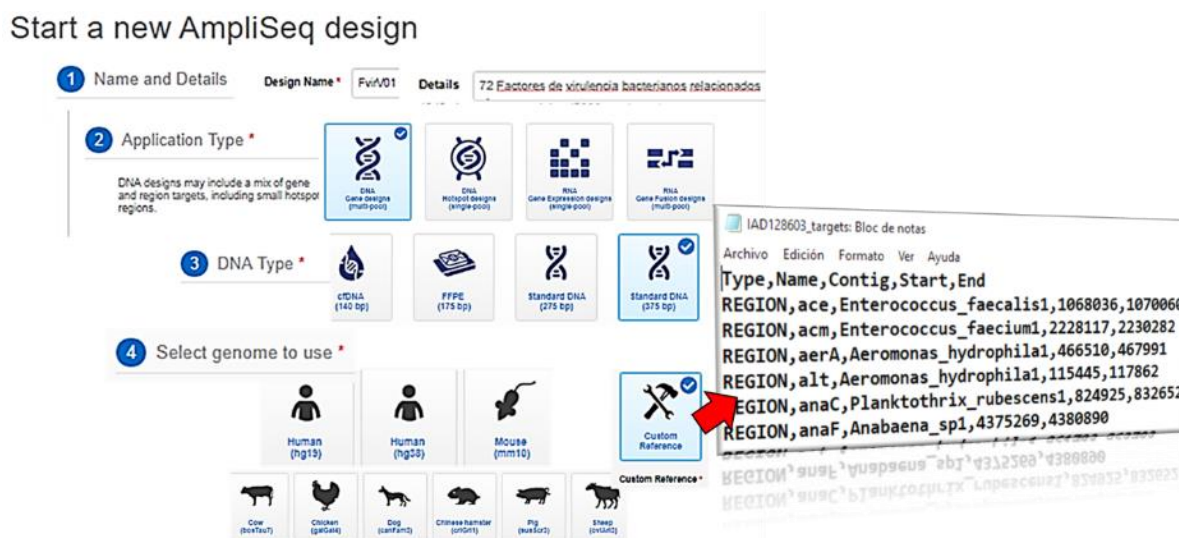


Figura 6. Diseño de *primers* mediante AmpliSeq™ designer. Pasos durante su utilización (izquierda) y personalización del panel (derecha).

3.1.4. Análisis in Silico Del Panel de Primers AmpliSeq

Para conocer detalladamente y verificar que los *primers* sintetizados estén dentro de los parámetros razonables para la PCR-multiplex se realizó la evaluación *in silico* del panel.

Se evaluaron características particulares de cada *primer* como contenido en G+C, temperatura de fusión o *melting*, formación de dímeros y horquillas, entre otros. Se determinó *in silico* la hibridación específica a los genes blanco utilizando el software Geneious Prime® (Kearse *et al.*, 2012) y la base de datos de secuencias de genomas anteriormente confeccionada. Se realizó el mapeo de cada juego de *primers* con su secuencia de referencia específica evaluando la orientación relativa, el tamaño del amplicón, la superposición, la discordancia o *mismatches* totales, así como el alineamiento fuera del gen o región conservada blanco. La hibridación inespecífica (blancos no bacterianos) de cada juego de *primers* se evaluó mediante la herramienta *BLASTn*, utilizando la base de datos de GenBank®.

3.2. Evaluación Del Panel en Muestras Controles Con Potencial Patogénico

3.2.1. Colección de Cepas y Extracción de ADN

El funcionamiento del panel se evaluó empleando muestras de ADN de bacterias patógenas y cianobacterias tóxicas portadoras de algunos de los genes blanco. La colección se conformó a partir de aislamientos o ADN obtenidos del Dpto. de Microbiología (IIBCE) y a través de diferentes grupos de investigación colaboradores (Instituto de Higiene, Instituto de Investigaciones Pesqueras, Facultad de Veterinaria) ver Tabla S1 en Anexo I. La extracción de ADN de los aislamientos obtenidos se llevó a cabo utilizando un método físico-químico puesto a punto previamente en el laboratorio de Microbiología de IIBCE, que incluye el uso de un paso de ruptura de las células con Fast-Prep (MP, *Biomedicals*), sucesivos pasos de centrifugación y precipitación del ADN con cloroformo-isopropanol (Martínez de la Escalera *et al.*, 2017) o utilizando kits comerciales de acuerdo a las recomendaciones del fabricante. El ADN obtenido fue cuantificado por espectrofotometría a 260 nm en Nanodrop (Thermo-Fisher™) y almacenado en freezer de -80 °C (backup). Previamente se tomaron alícuotas y se mantuvieron congeladas a -20 °C para su utilización.

3.2.2. Diseño Experimental

Se realizaron dos ensayos experimentales utilizando pooles de ADN de los patógenos obtenidos en la colección.

En el ensayo experimental I, control denominado “de cepas únicas”, se utilizó un pool de ADN de cuatro especies de colección portadoras de 8 genes blanco. Este ensayo tuvo como objetivo, evaluar el funcionamiento del panel, poner a punto la concentración de ADN y los ciclos de PCR iniciales. Para ello, se evaluaron cuatro tratamientos ajustando la concentración de ADN molde y los ciclos de PCR iniciales (Tabla 4). Se conformó un pool de ADN inicial con concentración igual a 30.000 copias genoma/μL de cada microorganismo. A partir del pool se realizaron 3 diluciones decimales seriadas y se ajustaron los ciclos de PCR de acuerdo con la concentración de ADN *input* y las recomendaciones de uso del panel.

Tabla 4. Descripción del ensayo experimental I, control con cepas únicas. Especies, genes blanco y tratamientos utilizados.

Nombre/Descripción	Especies/genos blanco	Tratamientos	
		Copias genoma/μL ²	Ciclos de PCR ³
Control de cepas únicas (CL) Pool de ADN utilizando 4 especies patógenas portadoras de 8 genes blanco ¹	<i>Escherichia coli</i> intimina (<i>eaeA</i>), toxina tipo Shiga 1 (<i>stx1A</i> , <i>stx1B</i>)		
	<i>Salmonella enterica</i> Typhimurium invasina A (<i>invA</i>)	CL I = 30.000	14
	<i>Listeria monocytogenes</i> listeriolisina O (<i>hyl</i>), internalina A (<i>inlA</i>)	CL II = 3.000	17
	<i>Pseudomonas aeruginosa</i> exotoxina A (<i>exotA</i>) exoenzima S (<i>exoS</i>)	CL III = 300 CL IV = 30	20 23

¹Determinado en trabajos previos, Tabla S1 Anexo I. ²El número de copias de ADN se calculó utilizando la herramienta de URI Genomics & Sequencing Center (<http://cels.uri.edu/gsc/cndna.html>) y corresponde a las copias totales en la PCR inicial. ³Los ciclos de PCR se ajustaron según el manual de uso de AmpliSeq™.

En el ensayo experimental II, “control de cepas múltiples”, se evaluó el funcionamiento del panel en muestras pooles utilizando toda la colección de cepas obtenida. Para ello se realizó un pool de ADN con 92 aislamientos de 8 géneros bacterianos posibles portadoras de algún gen blanco (Información disponible en la Tabla S1, Anexo I). Se realizaron dos tratamientos con diferente relación de copias de genomas iniciales/género para explorar si el enriquecimiento de la muestra en un microorganismo afecta la detección de otros menos representados. Uno de los tratamientos (C+) se conformó con igual copia inicial de genoma/género (10^9 copias/ μ L) y el otro (C50) enriquecido un orden logarítmico en uno de los géneros (*Listeria monocytogenes*, 10^{10} copias/ μ L) (Tabla 5).

Tabla 5. Descripción del ensayo experimental II, control con cepas múltiples. Número de aislamientos y tratamientos utilizados.

Nombre/Descripción	Nº de aislamientos utilizados ¹	Tratamientos (Copias genoma/ μ L) ²
Control Cepas Múltiples Pool de ADN utilizando 92 aislamientos de 8 géneros posiblemente patógenos La presencia de algunos genes blanco en los aislamientos era desconocida	<i>Escherichia coli</i> (7) + <i>Salmonella enterica</i> Typhimurium (1) + <i>Listeria monocytogenes</i> (2) + <i>Pseudomonas aeruginosa</i> (1)	C+ <i>Listeria</i> = 10^9 <i>Staphylococcus</i> = 10^9 <i>Salmonella</i> = 10^9 <i>E.coli</i> = 10^9 <i>Pseudomonas</i> = 10^9 <i>Cylindrospermopsis</i> = 10^9 <i>Aeromonas</i> = 10^9 <i>Enterococcus</i> = 10^9
	+ <i>Enterococcus</i> spp. (49) + <i>Staphylococcus aureus</i> (23) + <i>Aeromonas</i> spp. (8) + <i>Cylindrospermopsis raciborskii</i> (1)	C50 <i>Listeria</i> = 10^{10} <i>Staphylococcus</i> = 10^9 <i>Salmonella</i> = 10^9 <i>E.coli</i> = 10^9 <i>Pseudomonas</i> = 10^9 <i>Cylindrospermopsis</i> = 10^9 <i>Aeromonas</i> = 10^9 <i>Enterococcus</i> = 10^9

¹Los datos de los aislamientos utilizados se encuentran en la Tabla S1 del Anexo I. ²El número de copias de ADN se calculó utilizando la herramienta de URI Genomics & Sequencing Center (<http://cels.uri.edu/gsc/cndna.html>) y corresponde a las copias/ μ L en la PCR inicial.

3.2.3. Construcción de Pooles de ADN

Para la conformación de los diferentes pooles, las muestras de ADN almacenadas a -20 °C se descongelaron a temperatura ambiente y se colocaron en baño de agua termostatzado a 37 °C por 5 min (para resuspender ADN genómico). La concentración de ADN se cuantificó por fluorometría en Qubit® 2.0 (Invitrogen) utilizando el kit *Qubit dsDNA HS Assay* (ThermoFisher Scientific) de acuerdo con las instrucciones del fabricante. El número de copias de ADN se estimó a partir de la concentración de ADN molde y el tamaño del genoma de cada microorganismo utilizando la herramienta de *URI Genomics & Sequencing Center* (<http://cels.uri.edu/gsc/cndna.html>). Los pooles de ADN se conformaron mediante

dilución en agua bidestilada estéril de grado molecular, de acuerdo con lo descrito en el diseño experimental.

3.2.4. Amplificación y Construcción de Librerías

La amplificación utilizando el panel y su asociación al sistema Ion Torrent constó de los pasos secuenciales que se ilustran en la Figura 7. Los pools de ADN descritos anteriormente se utilizaron para la amplificación por PCR multiplex utilizando el panel AmpliSeq™ (desde ahora panel Fvir). La reacción de PCR y la construcción de librerías se realizaron utilizando los kits: *Ion AmpliSeq™ Library Kit 2.0*, adaptadores con secuencias identificatorias (barcodes) *Ion Xpress™ Barcode Adapter 1-96* (ThermoFisher Scientific) y el sistema de purificación *Agencourt AMPure™ XP* (Beckman Coulter™) siguiendo la guía de uso de AmpliSeq™ (Ion AmpliSeq, 2017), ver figuras 8 y 9.



Figura 7. Flujograma general de la técnica de secuenciación utilizando la plataforma de Ion Torrent PGM™, desde la extracción de ADN al análisis de datos, indicando el resultado obtenido en cada etapa.

PCR multiplex: El panel Fvir consta de dos pools de *primers* (Pool 1 y 2 respectivamente). Se preparó una reacción de amplificación con: 5 µL de 5x *Ion AmpliSeq™ HiFi Mix*, 1 µL de ADN (ver concentración en diseño) y agua ultrapura (Invitrogen) hasta volumen final de 12.5 µL. Se transfirió 5 µL de la mix a dos tubos de PCR y se le adicionaron 5 µL de cada pool de *primers* respectivo (amplificación de cada pool por separado). La reacción se incubó en un termociclador (MJ Mini®, BioRad) utilizando el siguiente programa: 1 ciclo de 99 °C x 2 min, seguido de n ciclos de 99 °C x 15 s y 60 °C x 4 min (el números de ciclos se ajustó de acuerdo con la

concentración de ADN molde, como se indicó en el diseño experimental). Una vez finalizado el programa se combinaron los 10 μL de la reacción de cada pool.

Digestión parcial de amplicones: los productos de PCR se digirieron con la enzima FuPA (proceso patentado por AmpliSeq™). Esta enzima digiere los *primers* modificados creando terminaciones que facilitan la ligación de adaptadores a ambos extremos (evita secuenciar innecesariamente los *primers* sobrantes y dímeros mejorando la eficiencia). Para ello se adicionaron 2 μL de la enzima y se incubaron con el siguiente programa en el termociclador: 50 °C x 10 min, 55 °C x 10 min y 60 °C x 20 min (Figura 8).

Ligado de adaptadores con barcodes: los adaptadores con barcodes son secuencias cortas de ADN que se unen a cada fragmento amplificado, permitiendo la identificación bioinformática de cada muestra durante el análisis posterior a la secuenciación (mediante la secuencia conocida de 8 nucleótidos del barcode). El primer paso en esta etapa fue diluir los adaptadores y preparar una mix de adaptadores con cada barcodes para cada muestra. Esta mix se preparó con: 2 μL de *Ion P1 Adapter*, 2 μL de *Ion Xpress™ Barcode* (uno diferente para cada muestra) y 4 μL de agua ultrapura.

Para la ligación la reacción fue preparada adicionando los reactivos en el siguiente orden: 4 μL de *Switch Solution*, 2 μL de la mix de adaptadores con barcodes, 2 μL de la ADN ligasa y los 22 μL de la digestión de amplicones. La reacción se incubó en el termociclador a 22 °C x 30 min, 68 °C x 5 min y 72 °C x 5 min (Figura 8).

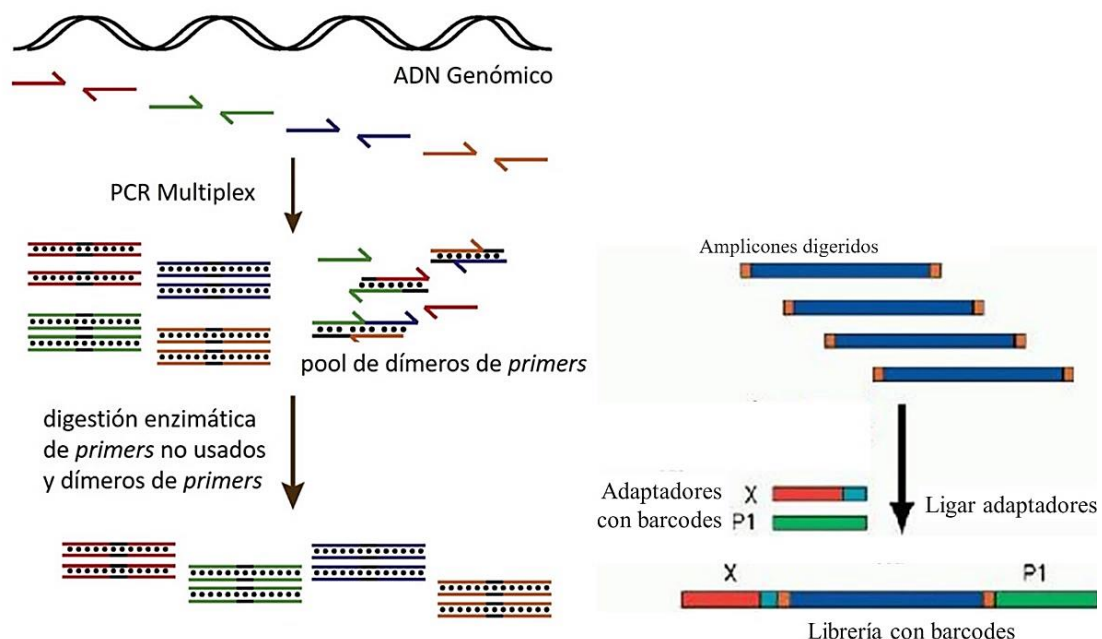


Figura 8. Amplificación y construcción de librerías con AmpliSeq™. Pasos de PCR multiplex, digestión enzimática y ligado de adaptadores con barcodes (adaptado de Khodakov, 2016 y Ion AmpliSeq™, 2017).

Purificación de librerías: Se utilizó un sistema de purificación por perlas magnéticas (permite eliminar contaminantes como dNTP's, *primers*, sales y dímeros de *primers*). Se realizó adicionando el reactivo AMPure™ XP Reagent a 1.5 X, previamente templado a ambiente. La mezcla se incubó 5 min y se colocó en el rack magnético

hasta que la solución se aclaró (separación de perlas). El sobrenadante fue descartado y las perlas lavadas con etanol 70 dos veces. El etanol se descartó y las perlas se secaron por 5 minutos (*pellet* de perlas), Figura 9.

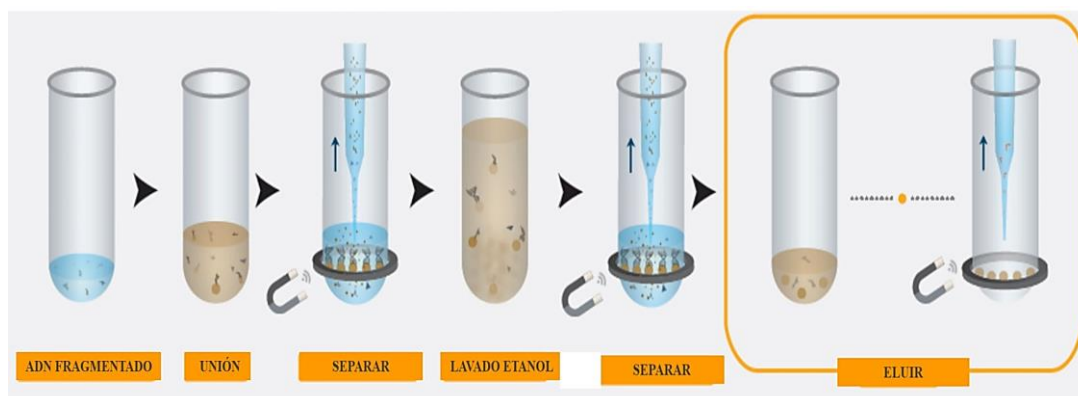


Figura 9. Pasos en la purificación de amplicones con perlas magnéticas. Unión de amplicones a las perlas, separado de perlas de los contaminantes, lavado de perlas con etanol 70 para remover contaminantes, elución de amplicones desde las perlas, transferencia de amplicones purificados a un nuevo tubo (AMPure™ XP, 2019).

Amplificación de las librerías: las librerías contenidas en los *pellets* de perlas de AMPure™ se amplificaron. Para ello se agregó 50 µL de *Platinum™ PCR SuperMix High Fidelity* y 2 µL de *Library Amplification Primer Mix*. La mezcla se incubó en el rack magnético por 2 min para separar las perlas y posteriormente se transfirieron 50 µL de sobrenadante a un nuevo tubo de PCR. La reacción se incubó en el termociclador utilizando el siguiente programa: 1 ciclo de 98 °C x 2 min, seguido de 5 ciclos de 98 °C x 15 s y 64 °C x 1 min.

Purificación de librerías amplificadas: luego de amplificadas las librerías se realizó la purificación con AMPure™ en dos etapas. La primera purificación se realizó con volumen 0.5 X de la muestra (aquí el ADN de alto peso molecular se une a las perlas; amplicones y *primers* se mantienen en la solución). En la segunda se purificó con volumen 1.2 X de la muestra (aquí los amplicones se unen a la muestra y los *primers* se mantienen en solución). Finalmente, las librerías se eluyeron con 50 µL de *buffer Low TE* y se almacenaron a 4 °C hasta su cuantificación y secuenciación.

Cuantificación por Qubit™: cada librería amplificada se cuantificó usando el fluorímetro Qubit™ 2.0 y el kit *Qubit™ dsDNA HS Assay*. Previo a la medición se preparó cada standard de concentración para generar la curva de calibración. Para la medición, se combinaron 10 µL de la librería amplificada con 190 µL de *Qubit™ dsDNA HS Reagent* (dilución de trabajo 1:200 en *dsDNA HS Buffer*). La mezcla se incubó por 2 minutos y se midió en el fluorímetro de acuerdo con el manual de uso. Como referencia se considera que las librerías amplificadas tienen entre 300 y 1500 ng/mL, aunque librerías por debajo de 300 ng/mL pueden aún proveer lecturas de buena calidad.

Normalización de la librería: se utilizó la electroforesis de ADN de alta sensibilidad mediante el sistema *Bioanalyzer DNA HS* (Agilent). Este sistema permite la cuantificación del ADN (5 pg/µL-100 ng/µL) así como la determinación de tamaños e integridad. Este paso se llevó a cabo por el técnico de la plataforma de secuenciación, utilizando los protocolos y reactivos del laboratorio. Basado en la concentración

calculada cada biblioteca de amplicones se diluyó a 26 pM y fueron agrupadas para secuenciarlas al mismo tiempo

3.2.5. Amplificación Clonal, Enriquecimiento y Secuenciación

La emulsión de PCR se llevó a cabo con el *Ion OneTouch™ System e Ion OneTouch 200 Kit v2* (Life Technologies). Después de la PCR en emulsión, las *Ion Sphere™* que contienen los moldes fueron enriquecidas con las perlas *Dynabeads™ MyOne™ Estreptovidina C1* (Life Technologies) y se lavaron con *Ion OneTouch™ Wash Solution*. Este proceso se llevó a cabo utilizando el *Ion OneTouch™ ES System* (Life Technologies). Después de la preparación de las *Ion Sphere Particles* se realizó el *Massively Parallel DNA Sequencing* con un sistema de *Ion Torrent Personal Genome Machine* (PGM) utilizando el *Ion 200 PGM™ Sequencing Kit* y el *Ion 316 chip*.

3.2.6. Análisis Bioinformático

Las secuencias de datos se procesaron con *Ion Torrent Software Suite™* y *Torrent Server* removiendo los adaptadores y barcodes, separando cada muestra en formato FASTq (de-multiplexado). Las secuencias fueron analizadas con *CLC Genomics Workbench 11* (CLC Bio, Aarhus). Se realizó el control de calidad por secuencia y por base (distribución de tamaños, contenido de G+C, contenido de bases ambiguas, cobertura y contribución de nucleótidos) así como el análisis de sobrerrepresentación en las secuencias obtenidas (niveles de duplicación, y secuencias duplicadas).

Las secuencias analizadas fueron filtradas (*trimming*) de acuerdo con criterios de calidad, tamaño y ambigüedad de bases (Remoción de secuencias: baja calidad=límite 0,05, nucleótidos ambiguos=máximo 2, tamaño=75-400 nucleótidos). Las secuencias descartadas se guardaron para su posterior mapeo.

El mapeo se realizó contra los archivos de referencia que contienen los genomas, las regiones conservadas de los genes blanco o los amplicones de cada juego de *primer*, generadas en la etapa de diseño. Para el mapeo se utilizaron los criterios por defecto del programa (costo de inserción=3, costo de eliminación=3, costo de discrepancia o *mismatch*=2, longitud de la fracción=0,5 y similitud de la fracción=0,8). Los resultados del mapeo se analizaron evaluando el número de lecturas o *reads* asignados y no asignados, distribuciones del tamaño en los mismos y comparando las secuencias obtenidas con la base de datos de NCBI-GenBank mediante *BLASTn*. Los *reads* no mapeados se analizaron de igual manera.

Los resultados del mapeo se expresaron como la proporción de *reads* (%) que mapean (en cada genoma, gen, región o amplicón) normalizados al total de *reads* mapeados de esa muestra.

$$\text{Reads mapeados (\%)} = \left(\frac{\text{Número de reads mapeados/gen o amplicón}}{\text{Número de reads mapeados de la muestra}} \right) \times 100$$

3.2.7. Análisis Estadístico

Se realizó el análisis exploratorio de datos, comparando el número de *reads* obtenidos en los distintos tratamientos y entre especies, genes y amplicones. Se evaluó si existían diferencias significativas entre estos grupos, así como entre pools de *primers*, tipo de genes y localización de los *primers* (conservados vs no conservados).

La sensibilidad y especificidad de los *primers* se calcularon *in silico* y para el ensayo I, control con cepas únicas, dado que se conocen los genes de virulencia que portan y considerando los *reads* mapeados de los cuatro tratamientos como repeticiones. Los parámetros se definieron y calcularon en base a lo reportado por Lemmon & Gardner, (2008):

$$\text{Sensibilidad}^* = 1 - \text{ratio falso negativo} = \frac{\text{falsos negativos}}{\text{verdaderos positivos} + \text{falsos negativos}}$$

$$\text{Especificidad} = 1 - \text{ratio falso positivo} = \frac{\text{falsos positivos}}{\text{verdaderos negativos} + \text{falsos positivos}}$$

*Para los cálculos de sensibilidad *in silico* se consideró como “hit de mapeo” cuando los *primers* hibridan a la secuencia específica del gen y tienen menos de 2 *mismatches* en el primer *forward* o *reverse* (Brinkmann *et al.*, 2017).

Dependiendo de las características de los datos obtenidos se aplicaron métodos paramétricos: ANOVA simple para establecer diferencias significativas ($\alpha=0,05$), seguido del test de Tukey o métodos no paramétricos: test de Kruskal-Wallis para establecer diferencias significativas ($\alpha=0,05$), seguido del test de Mann-Whitney.

3.3. Evaluación Del Panel en Una Muestra Ambiental

Para validar el panel Fvir en una muestra ambiental, se optó por el empleo de una muestra proveniente de una floración de cianobacterias del embalse de Salto Grande. Se seleccionó este sitio dado que es uno de los ecosistemas más afectados por las floraciones de cianobacterias en Uruguay. Particularmente las floraciones del complejo *Microcystis aeruginosa* (CMA) en el embalse han causado grandes problemas para la industria, recreación, turismo y para el consumo del agua. La muestra utilizada fue caracterizada previamente y cuenta con información de un gran número de variables bióticas y abióticas como por ejemplo la abundancia, expresión y diversidad de genotipos tóxicos y la concentración de microcistinas (Deus-Álvarez, 2018; Deus-Álvarez *et al.*, 2020).

3.3.1. Sitio Del Muestreo

El embalse de Salto Grande constituye un represamiento del Río Uruguay de carácter binacional (Uruguay-Argentina) que se construyó para la generación de energía hidroeléctrica en 1979. Está ubicado en el curso medio del río Uruguay, a unos 15 km al norte de las ciudades de Salto (Uruguay) y Concordia (provincia de Entre Ríos, Argentina) al norte de Uruguay y Argentina. La muestra proviene de la estación de Gualeguaycito, situada en uno de los brazos del río Uruguay, y fue tomada durante un evento de floración en diciembre de 2016 (GUAdic16). Se consideró floración como la multiplicación y acumulación de estos organismos en el cuerpo de agua de tal manera que cubrían completamente la superficie del agua y adquiría un color verdoso (Deus-Álvarez, 2018) Figura 10.

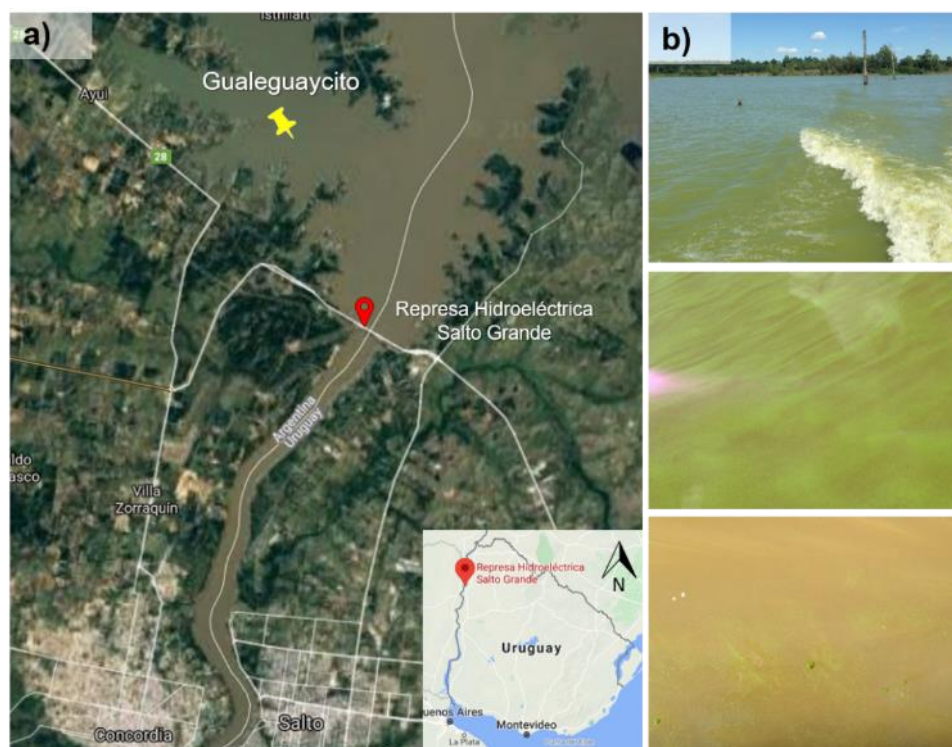


Figura 10. a) Localización del sitio de muestreo, estación Gualeguaycito. b) Evento de floración de cianobacterias en Gualeguaycito en diciembre de 2016 (Imágenes de Deus Álvarez, 2018).

3.3.2. Procesamiento de la Muestra y Extracción de ADN

Deus Álvarez, (2018) evaluó la relación entre la toxicidad y el tamaño de los organismos del CMA, por lo que las muestras fueron tamizadas utilizando mallas de diferentes tamaños, obteniendo cuatro fracciones: menores a 20 μm (células individuales, C), entre 20 y 60 μm (pequeñas, P), entre 60 y 150 μm (medianas, M) y mayores a 150 μm (grandes, G). Para la extracción de ADN genómico, 250-300 mL de agua correspondiente a cada fracción fue filtrada (membrana de celulosa de 0.2 μm). Posteriormente se utilizó la metodología puesta a punto en el Laboratorio de Microbiología (Martínez de la Escalera *et al.*, 2014). Las distintas fracciones de ADN se cuantificaron por fluorometría en Qubit® 2.0 (Invitrogen) utilizando el kit *Qubit dsDNA HS Assay* (ThermoFisher Scientific) de acuerdo con las instrucciones del fabricante.

3.3.3. Amplificación y Secuenciación de la Muestra

Se conformó un pool de ADN con concentración equimolar (50 ng/uL) a partir de las fracciones C, P y G. La fracción M (60-150 μm) se descartó debido a una baja concentración (<1 ng/ μL). Se ajustó la concentración final de ADN del pool a 10 ng/uL y se realizó la amplificación utilizando los sets de *primers* del panel y 17 ciclos de PCR (*Annealing* y extensión a 60 °C por 4 min) según los recomendados por Ampliseq (Ion Ampliseq, 2017). Los procedimientos de construcción de librerías, purificación, cuantificación, secuenciación y análisis bioinformático fueron los mismos que se utilizaron en las muestras control.

3.4. Evaluación de la Estructura Poblacional

Para complementar la información obtenida de los genes de virulencia detectados por el panel Fvir, se realizó el control de la estructura poblacional mediante una aproximación metagenómica usando la secuenciación masiva del gen que codifica para ARNr 16S.

Se puso a punto una metodología para la detección de bacterias en muestras de agua mediante la amplificación y secuenciación de la región V4 de dicho gen. Esta etapa constituyó la tesina de grado de la Lic. Eliana Nervi. En dicho trabajo se analizaron muestras de agua de estaciones costeras a lo largo del Río Uruguay y Río de la Plata. Se probaron distintos sets de *primers*, se ajustaron las condiciones de PCR y el análisis bioinformático (Nervi, 2020). Las condiciones determinadas como óptimas se utilizaron en el presente trabajo.

Debido al costo que implica la creación de librerías y secuenciación, el control poblacional mediante esta aproximación se realizó sólo en las muestras control C50 y ambiental de GUAdic2016.

3.4.1. Amplificación Por PCR

Se utilizó el ADN anteriormente extraído y cuantificado. Se amplificó por triplicado en un volumen de 50 µL utilizando los *primers* universales 520F y 802R (Claesson *et al.*, 2009; Caporaso *et al.*, 2011). La reacción de PCR consistió en: 5 µL Buffer-Taq (10 x), 3,5 µL de MgCl₂ (3,5 mM), 1 µL de dNTPs (10 mM), 2 µL de cada *primer* (10 µM), 1 µL de albúmina sérica bovina (BSA, 30 mM), 0,4 µL de Taq polimerasa (5 U/µL, Invitrogen), 2 µL de ADN molde (35 ng/reacción) y agua a completar los 50 µL.

El ciclado constó de una etapa inicial de desnaturalización a 95 °C x 5 min, seguida de 30 ciclos de 94 °C x 30 s, 50 °C x 30 s y 72 °C x 30 s. Se realizó una extensión final a 72 °C x 7 min. Los productos fueron analizados por electroforesis en gel de agarosa al 2 % en buffer TBE (Tris, Borato, EDTA), y fueron visualizados a través de la tinción con GelRed. Los productos de PCR se cuantificaron por fluorometría en Qubit® 2.0 (Invitrogen) utilizando el kit *Qubit dsDNA HS Assay* (ThermoFisher Scientific) de acuerdo con las instrucciones del fabricante. Se conformó un pool equimolar a partir del triplicado.

3.4.2. Construcción de Librerías y Secuenciación

Esta etapa se realizó utilizando los reactivos y protocolos anteriormente descritos para el uso del panel Fvir. La construcción de librerías se realizó utilizando los kits: *Ion AmpliSeq™ Library Kit 2.0*, adaptadores con barcodes *Ion Xpress™* y se purificó con el sistema de *Agencourt AMPure™*. Las librerías se cuantificaron en *Bioanalyzer DNA HS* y se normalizó su concentración. Se realizó la amplificación clonal, el enriquecimiento y por último la secuenciación en la plataforma ION Torrent PGM utilizando un chip 316.

3.4.3. Análisis Bioinformático

Para el análisis se siguió el flujograma mostrado en la Figura 11, utilizando las pipelines de programa QIIME (*Quantitative Insights Into Microbial Ecology*) (Caporaso *et al.*, 2010).

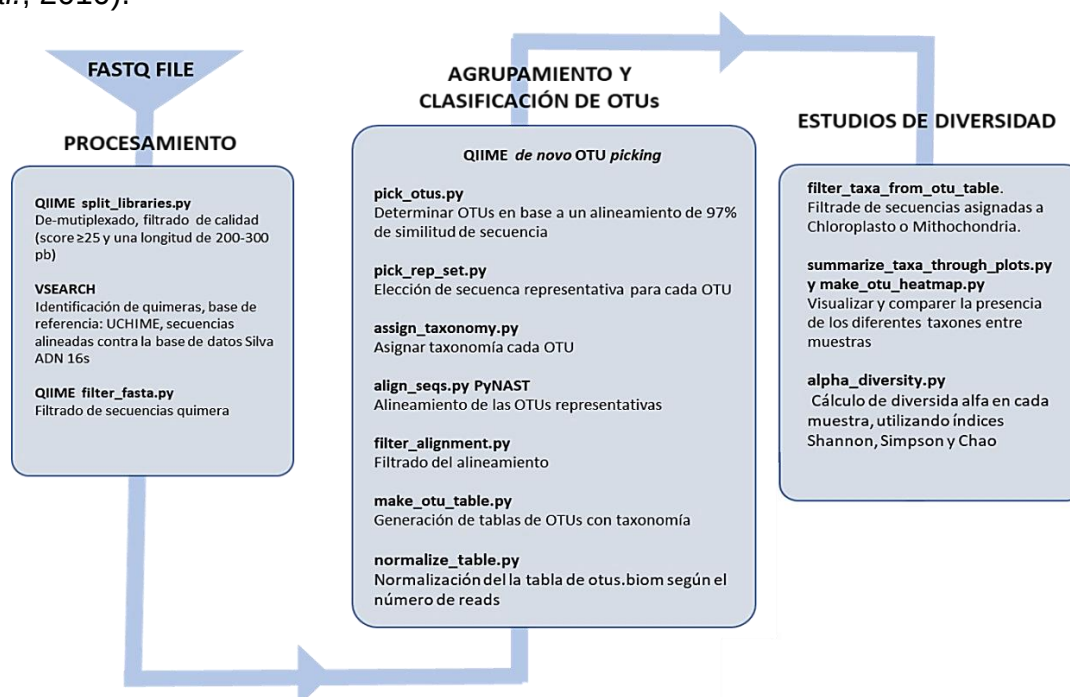


Figura 11. Flujograma del proceso de análisis con QIIME. Etapas de procesamiento, clasificación y estudio de diversidad a partir de secuencias crudas (Adaptado de Nervi, 2020).

Las lecturas obtenidas de la secuenciación contenidas en un archivo FASTA (.fna) se convirtieron en Fastq (*quality score file*). Se realizó el de-multiplexado, se eliminaron *primers*, adaptadores y *barcodes* y se filtraron secuencias de baja calidad o ambiguas. Se conservaron aquellas secuencias con un score de calidad ≥ 25 y una longitud de 200-300 pb (*quality filtering, trimming tool*). Se realizó la detección y filtración de secuencias quiméricas utilizando el programa VSEARCH y la base de datos de referencia UCHIME (Rognes *et al.*, 2016). Para la agrupación en OTUs (unidad taxonómica operativa) se utilizó la base de datos de secuencias de ARNr 16S Silva (SILVA_123.1_SSURef_Nr97_tax) con un criterio de identidad de secuencia del 97 %. Se eligió una secuencia representativa para cada OTU y se asignó la taxonomía utilizando la base de datos anteriormente mencionada. Se realizó el alineamiento de todas las OTUs representativas de manera de inferir la filogenia utilizando PyNAST. Se creó una tabla de la abundancia de cada OTU (Tabla de OTUs en formato BIOM) y se graficaron los resultados de taxonomía asignada para cada muestra a distintos niveles. Se realizó la rarefacción de la tabla de OTUs obtenida de manera de computar la riqueza de especies para cada muestra (diversidad alfa).

Los resultados obtenidos de la estructura poblacional para cada muestra se compararon con los *reads* mapeados en factores de virulencia por el panel Fvir correlacionando los grupos taxonómicos con el factor de virulencia correspondiente. Los resultados se presentarán en el análisis de cada muestra.

4. RESULTADOS

4.1. Diseño Del Sistema de Amplificación de Genes de Virulencia Bacterianos

4.1.1. Genes de Virulencia Candidatos

Se confeccionó un catálogo inicial de 72 genes de virulencia bacterianos, blanco para el diseño de *primers* (genes de virulencia de bacterias cuyo genoma y/o regiones genómicas son conocidos). Esta selección abarca genes vinculados con: el daño al hospedero (endo/exotoxinas, enzimas de la pared bacteriana), la adherencia y colonización del huésped (diversos tipos de adhesinas, fimbrias, pili y otros), los sistemas de secreción de tipo III y otros factores relacionados a la supervivencia (transportadores de urea y mecanismos de absorción de hierro entre otros). En algunos casos se seleccionaron más de un gen vinculado al mecanismo de virulencia (operones, islas de patogenicidad), así como las variantes del gen más relevantes que se han descrito. Los genes seleccionados contemplan más de 50 especies bacterianas, abarcando una amplia diversidad filogenética dentro del dominio Bacteria. En la Figura 12, se ilustra la distribución filogenética de los microorganismos seleccionados en este trabajo, así como el número de genes blanco en cada grupo.

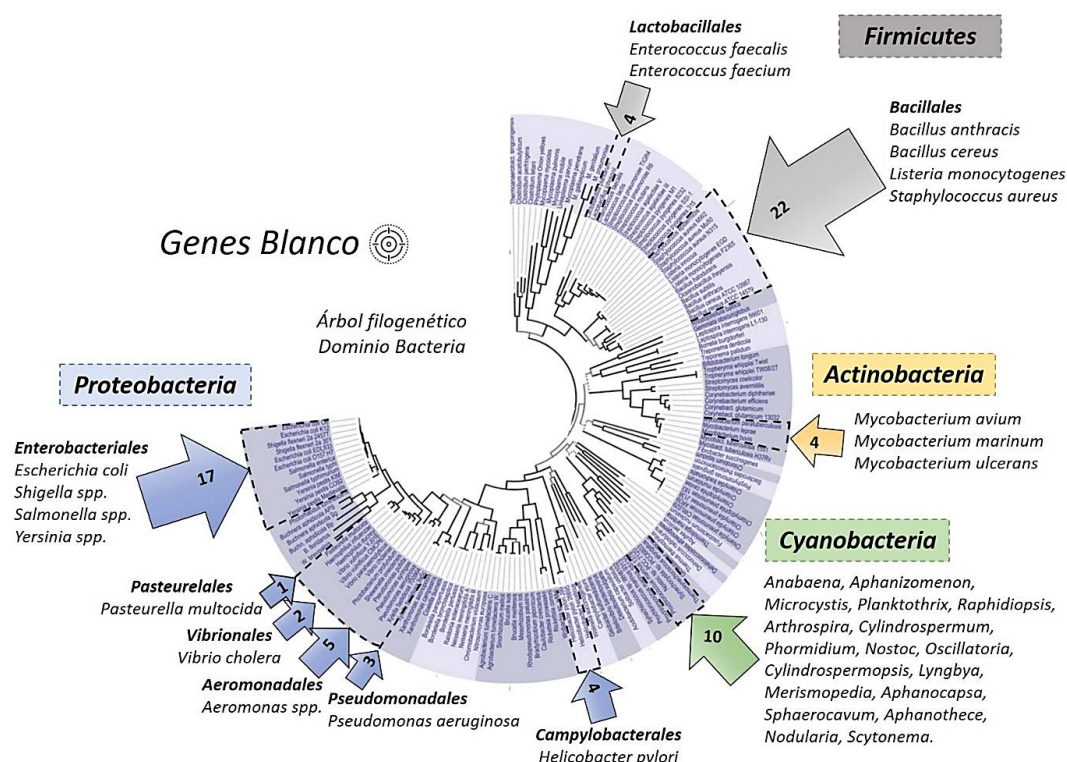


Figura 12. Distribución filogenética de los microorganismos seleccionados como blanco. El tamaño de las flechas indica el número de genes objetivos para el diseño de *primers*.

De los 72 genes seleccionados 46 se localizan en el cromosoma bacteriano y 26 en elementos extracromosomales (20 en plásmidos y 6 en fagos). En la Tabla 6 se presenta una lista detallada de los genes de virulencia seleccionados, los géneros o especies que los portan y las referencias bibliográficas consultadas para su selección.

Tabla 6. Factores de virulencia seleccionados como blanco para el diseño del panel de *primers* de AmpliSeq™. Géneros/especies vinculados al factor y referencias bibliográficas utilizadas para la selección.

Phylum/Clase	Orden	Factor de Virulencia	N° de genes seleccionados	Gen blanco	Géneros/especies	Referencias
Cyanobacteria Cyanophyceae	Chroococcales, Nostocales, Oscillatoriales	Anatoxina	2	<i>anaC</i> , <i>anaF</i>	<i>Anabaena</i> , <i>Aphanizomenon</i> , <i>Microcystis</i> , <i>Planktothrix</i> , <i>Raphidiopsis</i> , <i>Arthrospira</i> , <i>Cylindrospermum</i> , <i>Phormidium</i> , <i>Nostoc</i> y <i>Oscillatoria</i>	(Osswald <i>et al.</i> , 2007)
		Cylindrospermopsina	2	<i>cyrA</i> , <i>cyrB</i>	<i>Cylindrospermopsis raciborskii</i> , <i>Aphanizomenon ovalisporum</i> , <i>Raphidiopsis curvata</i> , <i>Aphanizomenon flos-aquae</i> , <i>Anabaena lapponica</i> , <i>Lyngbya wollei</i>	(Stüken & Jakobsen, 2010)
		Microcistina	2	<i>mcyE</i> , <i>mcyJ</i>	<i>Microcystis aeruginosa</i> , <i>Microcystis novacekii</i> , <i>Microcystis flos-aquae</i> , <i>Microcystis panniformis</i> , <i>Microcystis protocystis</i> , <i>Merismopedia tenuissima</i> , <i>Aphanocapsa delicatissima</i> , <i>Sphaerocavum brasiliense</i> , <i>Aphanothece stagnina</i>	(Tooming-Klunderud <i>et al.</i> , 2008; Bonilla, 2009; Vidal & Britos, 2012)
		Nodularina	1	<i>ndaI</i>	<i>Nodularia spumigena</i> , <i>Nodularia sphaerocarpa</i>	(Moffitt & Neilan, 2004)
		Saxitoxina	3	<i>sxtA</i> , <i>sxtI</i> , <i>sxtU</i>	<i>Anabaena</i> , <i>Aphanizomenon</i> , <i>Planktothrix</i> , <i>Cylindrospermopsis</i> , <i>Lyngbya</i> , <i>Scytonema</i>	(Kellmann <i>et al.</i> , 2008; Wiese <i>et al.</i> , 2010; Smith <i>et al.</i> , 2012)
Actinobacteria Actinobacteria	Actinomycetales	Complejo Antígeno 85	3	<i>fbpA</i> , <i>fbpB</i> , <i>fbpC</i>	<i>Mycobacterium avium</i> , <i>Mycobacterium marinum</i> , <i>Mycobacterium ulcerans</i>	(Dheenadhayalan <i>et al.</i> , 2002; Uchiya <i>et al.</i> , 2013; Stevenson, 2015)
		Antígeno P27 (Lipoproteína LprG)	1	<i>lprG</i>		
Proteobacteria Gammaproteobacteria	Aeromonadales	Complejo Enterotoxina citotóxica termolábil /Aerolisina/ hemolisina	1	<i>aerA</i> (<i>act/hlyA/aer</i>)	<i>Aeromonas hydrophila</i> / <i>Aeromonas</i> spp.	(Sen & Rodgers, 2004; Balsalobre <i>et al.</i> , 2009; Khajanchi <i>et al.</i> , 2010; Shakila <i>et al.</i> , 2013)
		Enterotoxina Citotónica Termolábil	1	<i>alt</i>		
		Enterotoxina Citotónica Termoestable	1	<i>ast</i>		
		Elastasa	1	<i>ela</i>		
		Fosfolipasa C	1	<i>plc</i>		

Continúa →

Tabla 6 (Continuación). Factores de virulencia seleccionados como blanco para el diseño del panel de *primers* de AmpliSeq™. Géneros/especies vinculados al factor y referencias bibliográficas utilizadas para la selección

Phylum/Clase	Orden	Factor de Virulencia	Nº de genes seleccionados	Gen blanco	Géneros/especies	Referencias
Proteobacteria Gammaproteobacteria	Enterobacterales	Pili tipo IV	2	<i>bfpA, bfpB</i>	<i>Escherichia coli</i>	(Croxen <i>et al.</i> , 2013)
		Intimina	1	<i>eaeA</i>		
		Enterotoxina termolábil	2	<i>eltA, eltB</i>		
		Enterotoxina termoestable	2	<i>est1, est2</i>		
		Fimbrias F17 y F41	2	<i>f17A, f41A</i>	<i>Escherichia coli, Shigella dysenteriae, Shigella sonnei y Shigella flexneri</i>	(Bertin <i>et al.</i> , 1996; Nagy & Fekete, 1999, 2005; Valat <i>et al.</i> , 2014)
		Toxina tipo Shiga I y II	4	<i>stx1A, stx1B, stx2A, stx2B</i>		(Croxen <i>et al.</i> , 2013)
		Invasina A	1	<i>invA</i>		(Turki <i>et al.</i> , 2014)
		Adhesina yad	1	<i>yadA</i>		(Drancourt, 2012)
		Invasina A	1	<i>ye-invA</i>		(Cornelis & Wolf-Watz, 1997; Black <i>et al.</i> , 2000; Drancourt, 2012)
		Tirosin-Fosfatasa	1	<i>yopH</i>	<i>Yersinia pestis, Yersinia enterocolitica y Yersinia pseudotuberculosis</i>	
	Pasteurelales	Toxina dermonecrótica	1	<i>toxA</i>	<i>Pasteurella multocida</i>	(Wilson & Ho, 2013)
	Pseudomonadales	Exotoxina A	1	<i>exotA</i>	<i>Pseudomonas aeruginosa</i>	(Gellatly & Hancock, 2013)
		Exoenzima S	1	<i>exoS</i>		
		Exoenzima U	1	<i>exoU</i>		
	Vibrionales	Toxina Colérica	2	<i>ctxA, ctxB</i>	<i>Vibrio cholerae, Vibrio spp.</i>	(Lencer <i>et al.</i> , 1992; De Haan & Hirst, 2004)

Continúa→

Tabla 6 (Continuación). Factores de virulencia seleccionados como blanco para el diseño del panel de *primers* de AmpliSeq™. Géneros/especies vinculados al factor y referencias bibliográficas utilizadas para la selección.

Phylum/Clase	Orden	Factor de Virulencia	Nº de genes seleccionados	Gen blanco	Géneros/especies	Referencias
Proteobacteria Epsilonproteobacteria	Campylobacterales	Ureasa	3	<i>ureA, ureB, ureI</i>	<i>Helicobacter pylori</i>	(de Bernard & Josenhans, 2014)
		Citotoxina vacuolizante	1	<i>vacA</i>		
Firmicutes Bacilli	Bacillales	Capsula	3	<i>capA, capB, capC</i>	<i>Bacillus anthracis</i>	(Duesbery & Woude, 1999; Little & Ivins, 1999; Zakowska <i>et al.</i> , 2012)
		Toxina antrácica: antígeno protector (AP), factor letal (FL) y factor de edema (FE)	3	<i>pag, lef,cya</i>		
		Toxina emética	4	<i>cesA, cesB, cesC, cesD</i>	<i>Bacillus cereus</i>	(McKillip, 2000; Guinebretiere <i>et al.</i> 2006; Bottone, 2010; Dommel <i>et al.</i> 2011; Økstad & Kolstø, 2011)
		Citotoxina K	1	<i>cytK</i>		
		Toxina hemolítica	3	<i>hblA, hblC,hblD</i>		
		Enterotoxina no hemolítica	3	<i>nheA, nheB, nheC</i>		
		Internalina A	1	<i>inlA</i>	<i>Listeria monocytogenes</i>	(Reddy & Lawrence, 2014)
		Listeriolisina O	1	<i>hlyA</i>		
		Coagulasa	1	<i>coa</i>	<i>Staphylococcus aureus</i>	(Rasmussen <i>et al.</i> , 2013)
		Proteína A de unión a fibronectina	1	<i>fnbA</i>		
		Toxina de shock séptico (TSST)	1	<i>tst</i>		
	Lactobacillales	Proteína de unión al colágeno (adhesina)	1	<i>ace</i>	<i>Enterococcus faecalis</i>	(Soheili <i>et al.</i> , 2014)
		Proteína de unión al colágeno (adhesina)	1	<i>acm</i>	<i>Enterococcus faecium</i>	
		Hialuronidasa	1	<i>hyl</i>		
		Pili	1	<i>pilB</i>		

4.1.2. Evaluación de la Conservación en Los Genes de Virulencia

La búsqueda de regiones conservadas en los genes de virulencia (blanco de *primers*) implicó una evaluación exhaustiva de las bases de datos de secuencias disponibles. Como resultado, se generó una base de datos local con 9544 secuencias que cumplieron los criterios establecidos de identidad y cobertura (secuencias reportadas en GenBank). Esta se conformó en su mayoría (75 %) con secuencias de genes o *contigs* y en menor grado (25 %) con secuencias de genomas completos. Solo 39 de los 72 genes presentaron al menos un genoma completo reportado al momento de su estudio (Tabla S1, Anexo II). En la base de datos generada, se exploró para cada uno de los genes la variabilidad genética (distancia evolutiva), el número de secuencias reportadas y la taxonomía de las secuencias (Figura 13). En estos resultados se observa, por ejemplo, la variación genética alta de algunos genes (cianobacterias, Complejo *Mycobacterium avium*), el bajo reporte de secuencias (cianobacterias o algunos genes de *E. coli*) o el rango de distribución particular entre los taxones (amplio o acotado).

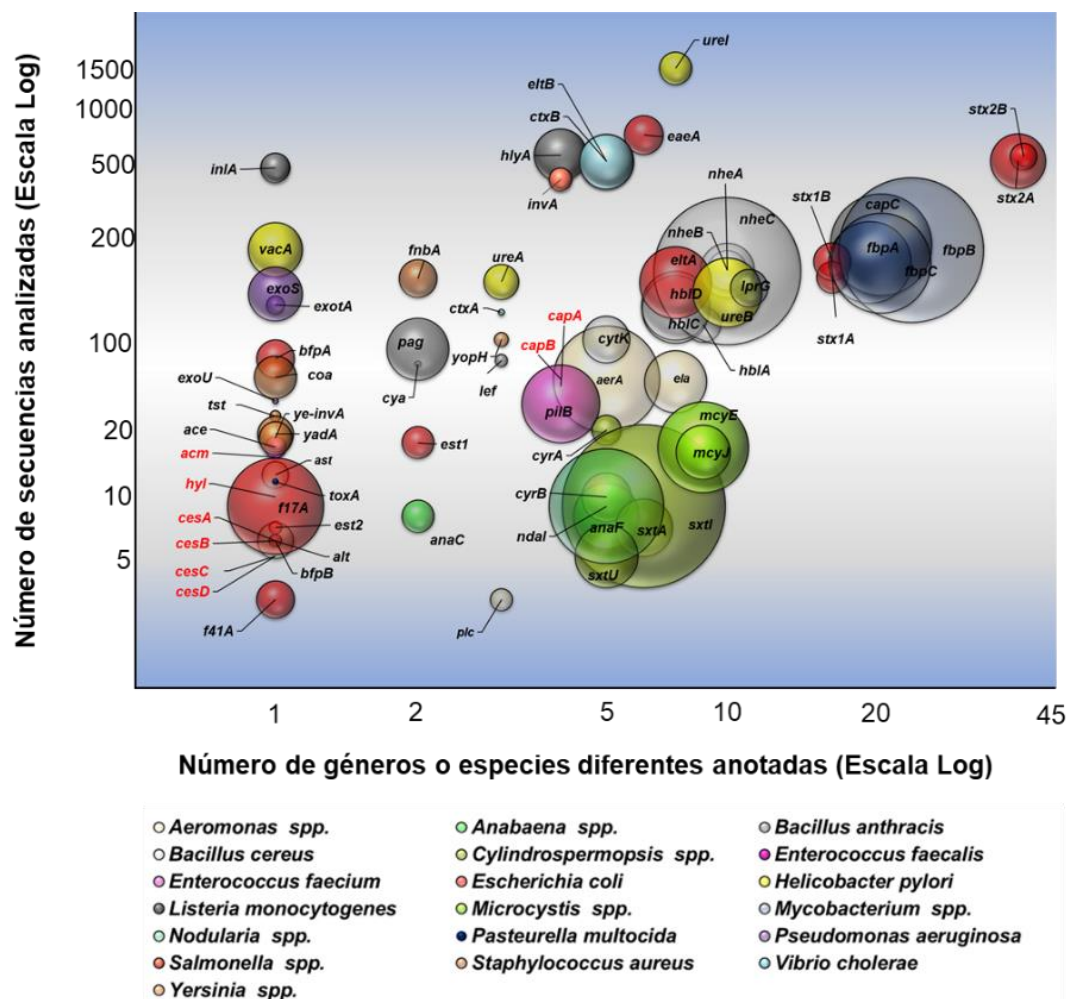


Figura 13. Análisis de la base de datos construida a partir de secuencias reportadas de GenBank. Número de secuencias analizadas por especie/género anotado. El tamaño de la burbuja = distancia evolutiva (número de sustituciones: transiciones + transversiones). El rango de variación fue de 0-62 %, donde 0 corresponde a los genes marcados en rojo y 62 al gen *sxtI*. El número de genes/especies anotadas incluye también secuencias de fagos y plásmidos.

El estudio de la variabilidad en la secuencia de los genes permitió establecer los bloques conservados más extensos y definirlos como blanco para el diseño de *primers*. Se identificaron regiones conservadas (>200 pb) en 39/72 genes (54 %), grupo denominado como “genes conservados”. En la Figura 14 se observa el tamaño y porcentaje de conservación de cada gen (tamaño de la región conservada más extensa en relación con el tamaño del gen). En estos genes el grado de conservación promedio fue alto, del 75 % de la secuencia del gen (rango 10-100 %) y si bien algunos genes presentaron valores inferiores al 20 % de conservación (4 genes) su tamaño > a 1000 pb permitió establecer una zona conservada blanco para *primers*.

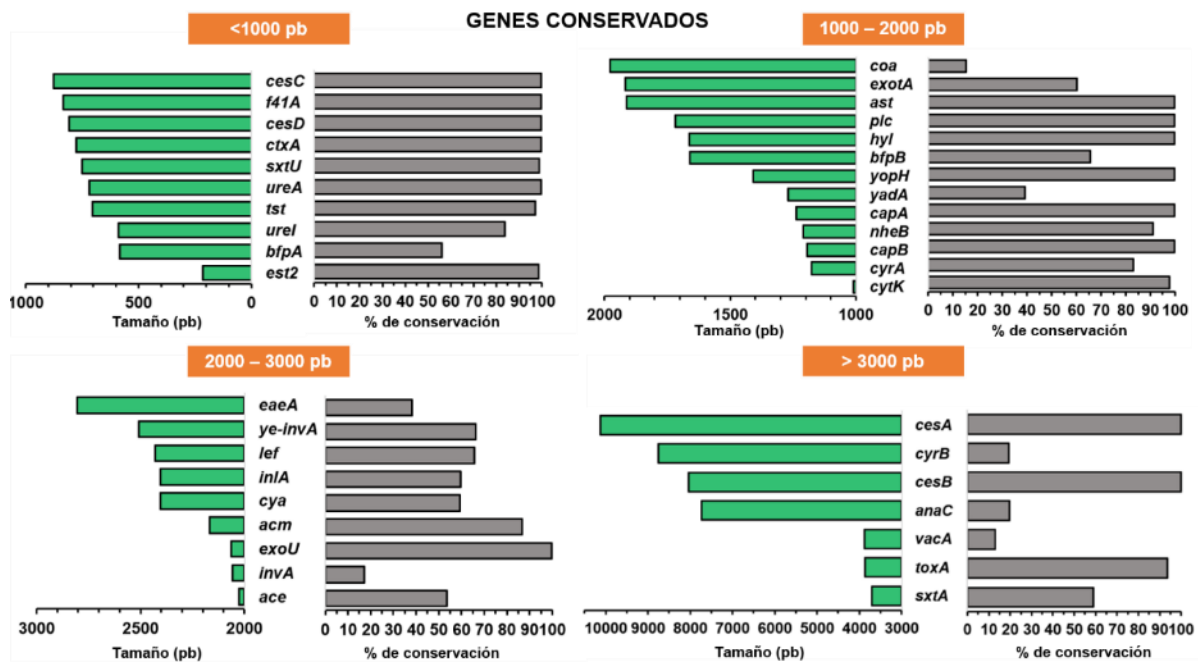


Figura 14. Genes Conservados, n=39. Tamaño del gen y porcentaje de conservación de la secuencia (longitud del bloque conservado más grande/longitud del gen). Distribución de genes por tamaño. Los genes representados pertenecen a las especies: *Aeromonas* spp. (*ast*, *plc*), *Bacillus anthracis* (*capA*, *capB*, *cya*, *lef*), *Bacillus cereus* (*cesA*, *cesB*, *cesC*, *cesD*, *cytK*, *nheB*), *Cianobacterias* (*anaC*, *cyrA*, *cyrB*, *sxtA*, *sxtU*), *Enterococcus* spp. (*ace*, *acm*, *hyl*), *Escherichia coli* (*bfpA*, *bfpB*, *eaeA*, *est2*, *f41A*), *Helicobacter pylori* (*ureA*, *ureI*, *vacA*), *Listeria monocytogenes* (*inlA*), *Pasteurella multocida* (*toxA*), *Pseudomonas aeruginosa* (*exotA*, *exoU*), *Salmonella* spp. (*invA*), *Staphylococcus aureus* (*coa*, *tst*), *Vibrio cholerae* (*ctxA*) y *Yersinia* spp. (*yadA*, *ye-invA*, *yopH*).

Los restantes 33 genes (46 %) no presentaron una región conservada suficientemente extensa (>200 pb), grupo denominado “genes no conservados”. En la Figura 15 se observa el tamaño y porcentaje de conservación de cada gen. Si bien varios genes de este grupo son de tamaño > a 1000 pb, incluso algunos alcanzan las 10.000 pb, el grado de conservación fue bajo (< 13 % de la secuencia del gen), con un promedio de 3,25 %, y la región conservada más extensa observada fue de 90 pb.

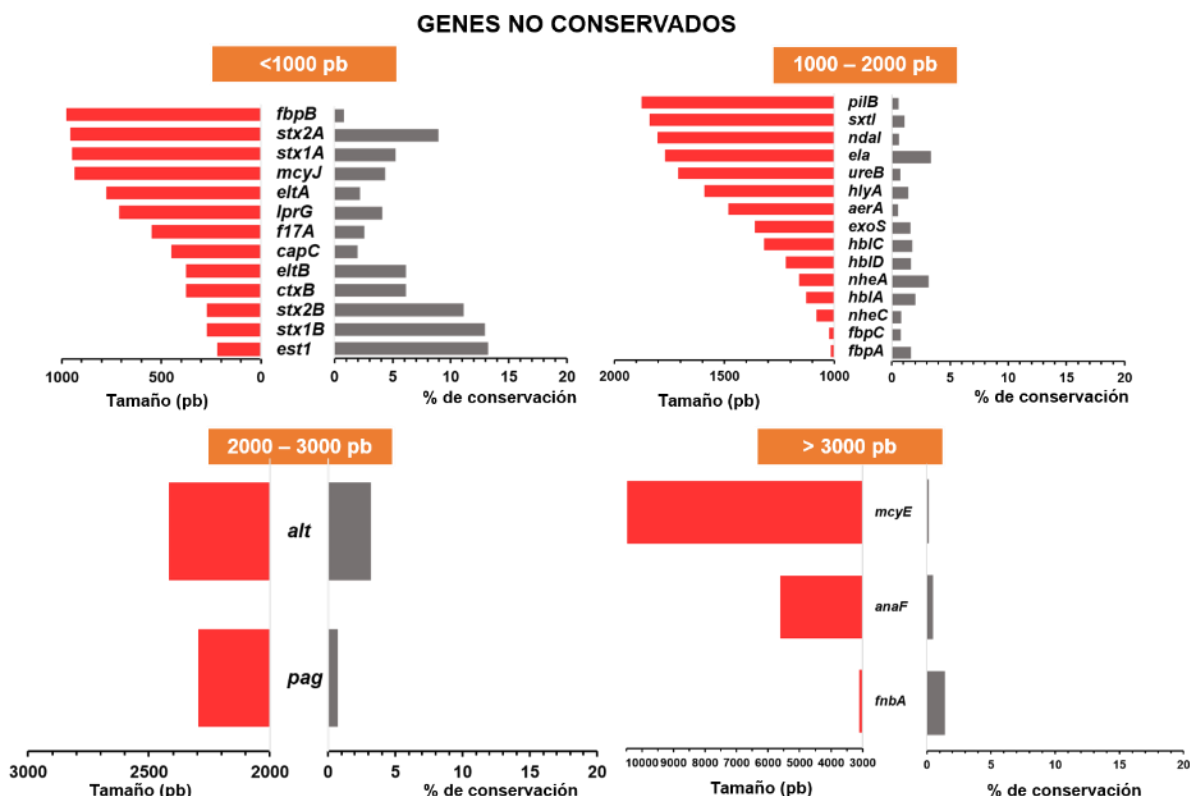


Figura 15. Genes no conservados, n=33. Tamaño del gen y porcentaje de conservación de la secuencia (longitud del bloque conservado más grande/longitud del gen). Distribución de genes por tamaño. Los genes representados pertenecen a las especies: *Aeromonas* spp. (*aerA*, *alt*, *ela*), *Bacillus anthracis* (*capC*, *pag*), *Bacillus cereus* (*hblA*, *hblC*, *hblD*, *nheA*, *nheC*), Cianobacterias (*anaF*, *mcyE*, *mcyJ*, *ndal*, *sxtI*), *Enterococcus* spp. (*pilB*), *Escherichia coli* (*eltA*, *eltB*, *est1*, *f17A*, *stx1A*, *stx1B*, *stx2A*, *stx2B*), *Helicobacter pylori* (*ureB*), *Listeria monocytogenes* (*hlyA*), *Mycobacterium* spp. (*fbpA*, *fbpB*, *fbpC*, *lprG*), *Pseudomonas aeruginosa* (*exoS*), *Staphylococcus aureus* (*fmbA*) y *Vibrio cholerae* (*ctxB*).

La variabilidad genética media, el número de géneros o especies anotadas y el número de secuencias analizadas en cada gen presentaron diferencias significativas entre los grupos de genes conservados y no conservados (Figura 16).

El número de secuencias reportadas en los genes no conservados varió entre 6 y 559 (mediana= 113) en comparación con los genes conservados 3 y 1551 (mediana= 26).

En los genes conservados la variabilidad genética es más baja (0-9 % mediana=1,1 %) y la distribución entre taxones más acotada (1-9 y mediana=1 géneros o especies), limitándose en algunos genes a una única especie o una cepa en particular (por Ej: genes que codifican toxinas como la dermonecrótica de *Pasteurella multocida* (*toxA*), la colérica (*ctxA*) o las del ántrax (*cya*, *lef*).

En contraste, Los genes no conservados presentan una variabilidad genética alta (2-62 %, mediana=9,1 %) y se distribuyen en un rango más amplio de taxones (1-38, mediana= 7). Como lo observado en los genes del complejo *Mycobacterium avium* (complejo antigénico 85: *fbpA*, *fbpB*, *fbpC*) o en algunos de los genes que codifican toxinas de cianobacterias (Microcistina: *mcyE* y *mcyJ*).

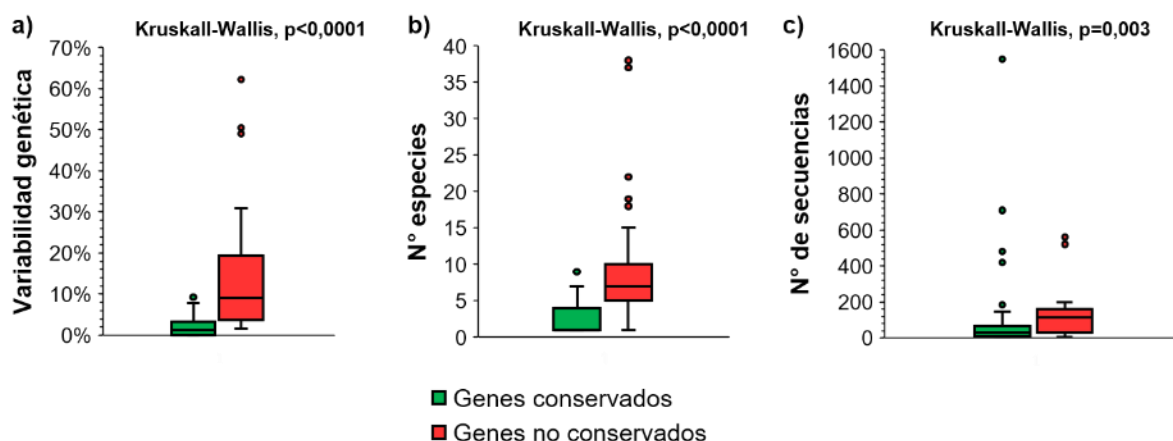


Figura 16. Evaluación de la base de datos de secuencia, genes conservados vs no conservados. Diagrama de cajas de: variabilidad genética media (a), número de géneros/especies anotadas (b) y número de secuencias reportadas (c) en los grupos de genes clasificados como conservados y no conservados. La variabilidad genética media se calculó como el número de sustituciones (transiciones + transversiones) utilizando el método de Máxima Verosimilitud Compuesta (MLC), (Nei & Kumar, 2000) y se expresó en porcentaje. El número de géneros/especies anotadas incluye las secuencias de fagos y plásmidos.

Los resultados del estudio de conservación de los genes blanco definieron la estrategia para el diseño de *primers* y en base a la clasificación en conservado o no conservado se procedió de la siguiente manera. En los genes con regiones conservadas se determinó el bloque más extenso de conservación y se definió como blanco para el diseño de los *primers*. La secuencia consenso de dicho bloque se mapeó en genomas de referencia donde se sustituyó por la original. En los genes no conservados se utilizó el genoma de una cepa de referencia con una secuencia única de cada gen. En la Tablas S2 y S3 del Anexo I se detallan los genomas de referencia utilizados, la localización de los genes y/o las zonas conservadas para ambos grupos respectivamente.

Los genomas de referencia y un archivo con las coordenadas de localización de los genes se cargaron en Ampliseq™ *designer* para el diseño del panel de *primers*.

4.1.3. Panel de Primers Fvir: Características Generales

El diseño de los *primers* se realizó mediante el software *Ampliseq Designer*™, el cual proporciona una estimación de cobertura de las regiones a secuenciar. Una vez aplicados los criterios de delimitación de las regiones de interés, se obtuvo un total de 548 juegos de *primers* divididos en 2 pooles con 284 y 264 juegos respectivamente, que permitirían amplificar el ADN para la creación de librerías mediante una PCR multiplex. Se logró una cobertura global del 97,35 % del total de bases seleccionadas como blanco, lo que implicaría la secuenciación de 148,74 kilobases por muestra analizada. La cobertura de los *primers* en la mayoría de los genes (62/72) fue del 100 % de la secuencia y sólo 2 genes resultaron con baja cobertura (<30 %, *exotA* y *exoS* de *Pseudomonas aeruginosa*), diseñándose un menor número de *primers*. En el caso del gen *est1* de *E. coli* no fue posible el diseño de *primers*, por lo que el panel se redujo a 71 genes candidatos (Figura 17).

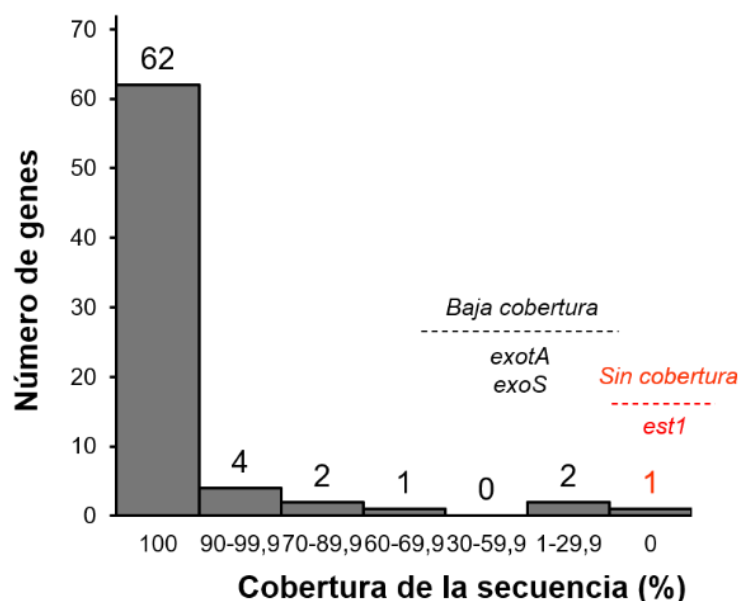


Figura 17. Cobertura del Panel de *primers* Fvir en la secuencia de los genes candidatos. Genes con baja cobertura: *exoA* y *exoS* (*Pseudomonas aeruginosa*), gen sin cobertura: *est1* (*Escherichia coli*).

En las Figuras 18,19 y 20 se ilustra el número de juegos de *primers* que conforman el Panel Fvir distribuidos por gen y taxón. El 54 % de los *primers* generados apuntan hacia los grupos de cianobacterias y *Bacillus* (*anthracis* y *cereus*) teniendo como blanco 27/71 genes candidatos. Los *primers* para el grupo *Enterobacterales* representan el 12 % del total de *primers* con más de la mitad (59 %) dirigidos a únicamente a genes de *E. coli*, dado que algunos de ellos definen distintos patotipos de la especie.

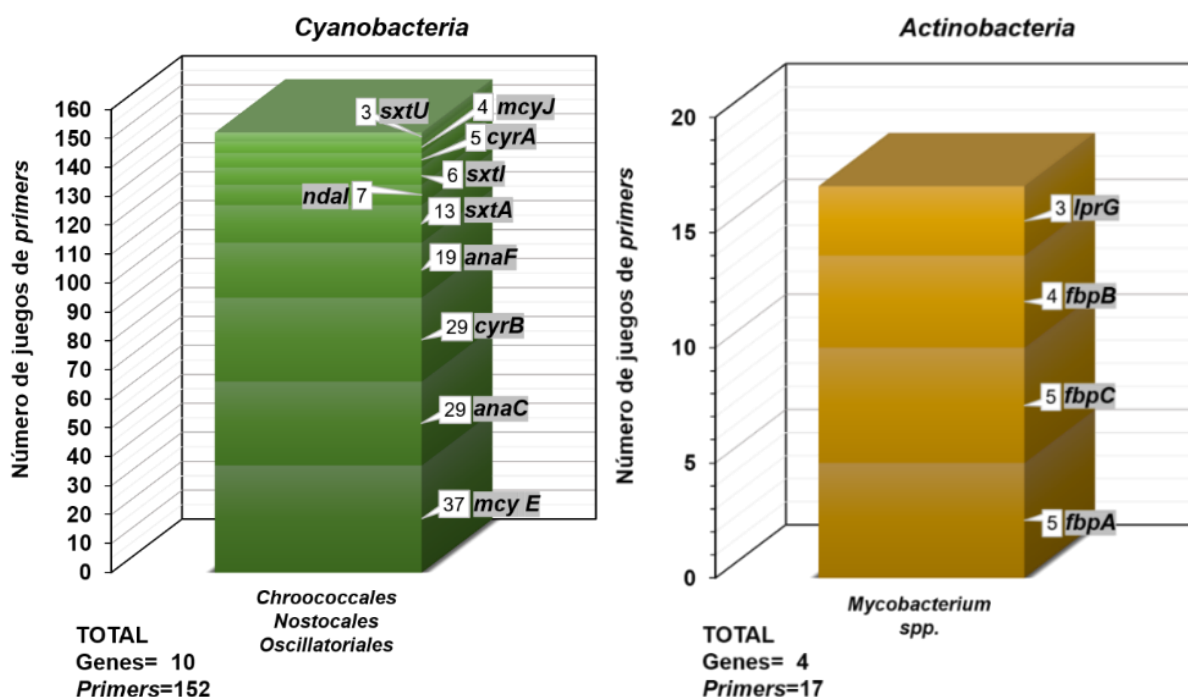


Figura 18. Número de juegos de *primers* generados en el panel Fvir para la amplificación de genes de virulencia de los filos: Cyanobacteria y Actinobacteria.

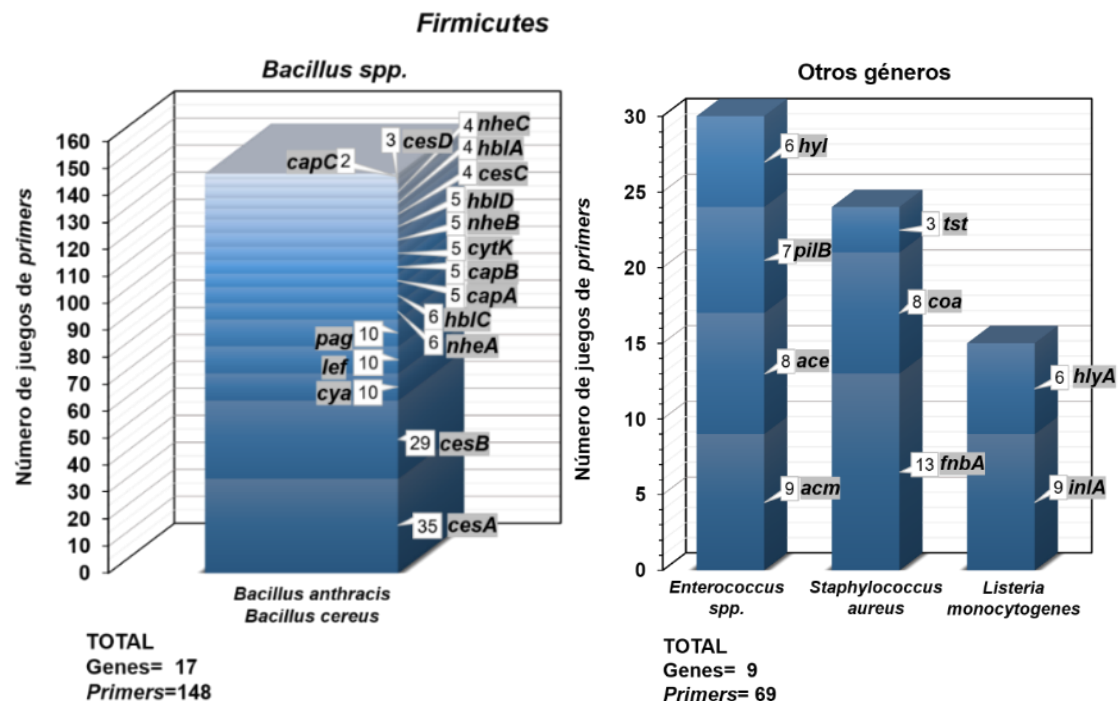


Figura 19. Número de juegos de *primers* generados en el panel Fvir para la amplificación de genes de virulencia del filo Firmicutes.

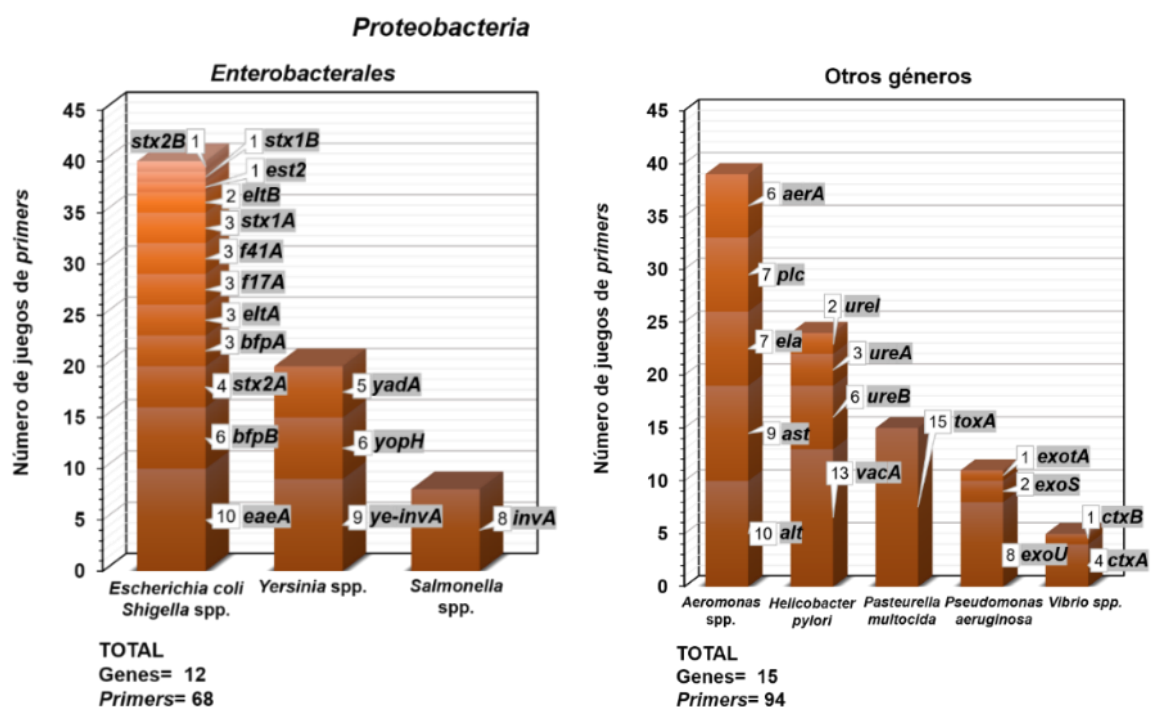


Figura 20. Número de juegos de *primers* generados en el panel Fvir para la amplificación de genes de virulencia del filo Proteobacteria.

4.1.4. Evaluación in Silico Del Panel

4.1.4.1. Características generales

Se evaluaron *in silico* algunas propiedades de importancia para el adecuado funcionamiento de los *primers* explorando los dos pools que conforman el panel, dado que se utilizan por separado en la PCR inicial. Los resultados se describen en la Figura 21. La evaluación *in silico* del panel reflejó lo eficiente del sistema *Ampliseq™ designer* en el diseño de los *primers*. No existieron diferencias significativas entre pools ($p < 0,05$) para las variables de longitud, contenido en G+C y temperatura de *melting* (T_m). La longitud de los *primers* varía entre 17-33 pb con el 75 % de ellos en 29 pb. El contenido en G+C varió en mayor medida (27,6-64,7 %) y sólo el 20 % de los *primers* contienen más de 50 % en G+C. Los valores de T_m oscilaron entre 58,8-64,6 °C con más del 95 % de los *primers* con valores superiores a 60 °C.

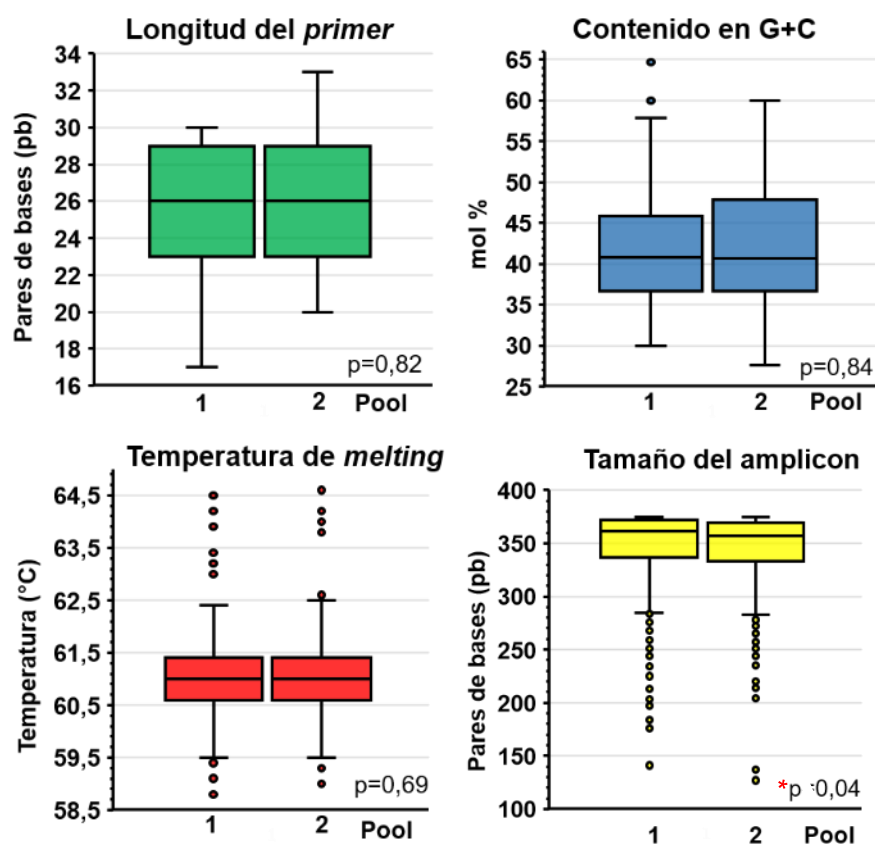


Figura 21. Características generales de los *primers* del panel Fvir. Longitud, contenido en G+C, temperatura de *melting* y tamaño del amplicón en pools 1 y 2. *Diferencias significativas test de Kruskal-Wallis ($p < 0.05$).

En relación con la estructura secundaria de los *primers* (formación de dímeros en/entre *primers* y de horquillas) no existieron diferencias significativas entre los pools ($p > 0,05$). La capacidad de formación de dímeros de *primer* y entre *primers* estuvo presente en 58 y 31 % de los *primers* respectivamente. En ningún caso se observó a temperaturas ≥ 60 °C (temperaturas de *annealing* y extensión de la PCR). La capacidad de formación de horquillas estuvo presente en el 80 % de los *primers* y sólo 2 % de ellas presentó temperaturas ≥ 60 °C (Figura 22).

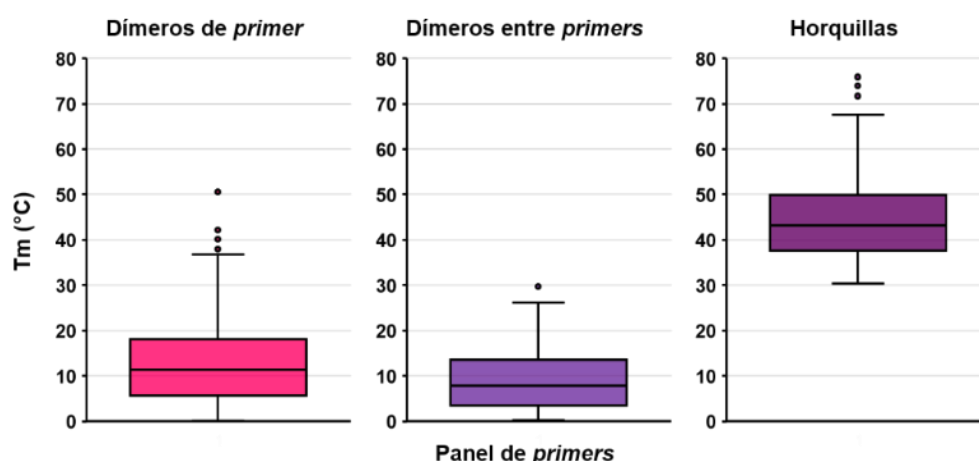


Figura 22. Análisis de la estructura interna de los *primers* del panel Fvir. Formación de dímeros de *primer* (n=641), entre los *primers* (n=171) y formación de horquillas (n=882). No existieron diferencias significativas entre pools, test de Kruskal-Wallis ($p < 0,05$).

La distribución de los tamaños de los amplicones fue levemente diferente entre ambos pools ($p = 0,04$). En la Figura 23 se observan las frecuencias por tamaño de los insertos de amplicones del panel y las distribuciones en los dos pools. Los amplicones se diseñaron en un rango entre 127-375 pb, el 75 % de ellos cercanos a 370 pb y sólo 7 tienen tamaño ≤ 200 pb.

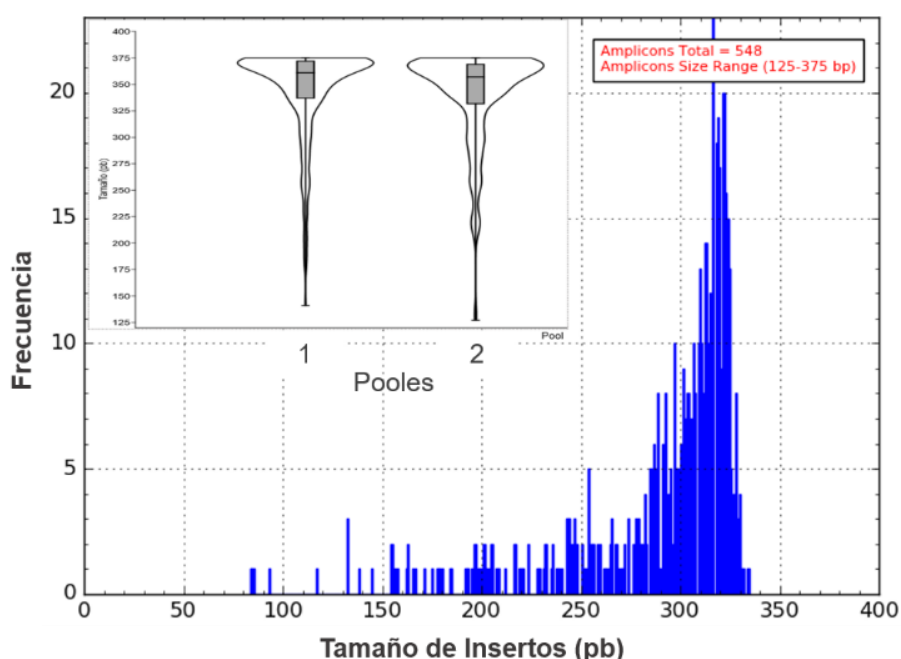


Figura 23. Histograma de frecuencias del tamaño de insertos de amplicones (pb) del Panel Fvir. Distribución de tamaños de amplicón en los dos pools (esquina superior izquierda).

4.1.4.2. Análisis del mapeo

Los juegos de *primers* se mapearon contra sus genomas de referencia considerando su orientación relativa y una distancia entre sí <1000 pb.

Los 584 juegos mapean sobre la región del gen blanco, con un bajo porcentaje de *mismatches* (< 0,5 %), observándose un máximo de 2 *mismatches* en un sólo *primer* (considerado como hit de mapeo según el criterio establecido). En base a estos resultados la sensibilidad *in silico* del panel fue de 100 %. De los 548 juegos, 3 no fueron específicos de la región de interés, mapeando además en otra región del genoma y se excluyeron del análisis ya que existían otros que amplifican la misma región. En base a este resultado la especificidad general *in silico* del panel es de 99,5 %.

No existió mapeo inespecífico a blancos no bacterianos que generaran productos <1000 pb en ninguno de los *primers*.

En relación con el mapeo en zonas conservadas fue posible generar 170 juegos de *primers* que tiene como blanco 32 de los 39 genes candidatos. En los restantes 7 genes (*anaC*, *bfpA*, *coa*, *est2*, *exotA*, *invA* y *ureI*) la localización de los *primers* resultó fuera de la región conservada, como se esquematiza en la Figura 24.

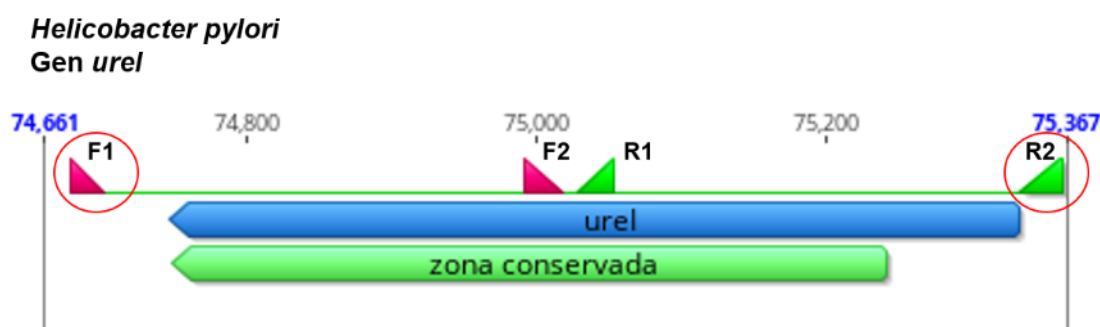


Figura 24. Ejemplo de localización de *primers* fuera de la región de conservación. Se representa un gen con región conservada (gen *ureI* e *Helicobacter pylori*) y la localización de *primers* fuera de la región conservada.

4.2. Evaluación Del Panel en Muestras Controles con Potencial Patogénico

Se realizaron dos ensayos experimentales utilizando pools de ADN de los patógenos obtenidos en la colección (ensayo con cepas únicas y con cepas múltiples). Durante estos ensayos se secuenciaron 6 muestras provenientes de las distintas configuraciones de los tratamientos. Los resultados generales relacionados a las etapas de cuantificación de librerías y controles de calidad de la secuenciación se presentan juntos para todas las muestras, mientras que el mapeo en los genes se presenta por separado para cada ensayo en particular.

4.2.1. Cuantificación de Librerías y Secuenciación

Todas las librerías construidas pudieron ser cuantificadas con concentraciones detectables en *Bioanalyzer*. Los picos detectados y cuantificados fueron del tamaño esperado de los amplicones cubiertos por el panel Fvir. Se cuantificaron los picos en el rango entre 125-400 pb. El número de picos y las concentraciones variaron de acuerdo con los ensayos experimentales. En el ensayo I (menor número de genes candidatos) se detectaron de 2-6 picos, con una concentración entre 65-130 pg/μL entre las cuatro muestras. En el ensayo II (múltiples cepas) se detectó mayor número de picos (10-13) y la concentración estuvo comprendida entre 750-830 pg/μL entre las dos muestras. No se observaron picos cuantificables dentro del rango estudiado en los controles negativos en ninguno de los ensayos. En la Figura 25 se ilustra a manera de ejemplo la cuantificación por esta técnica en una de las muestras.

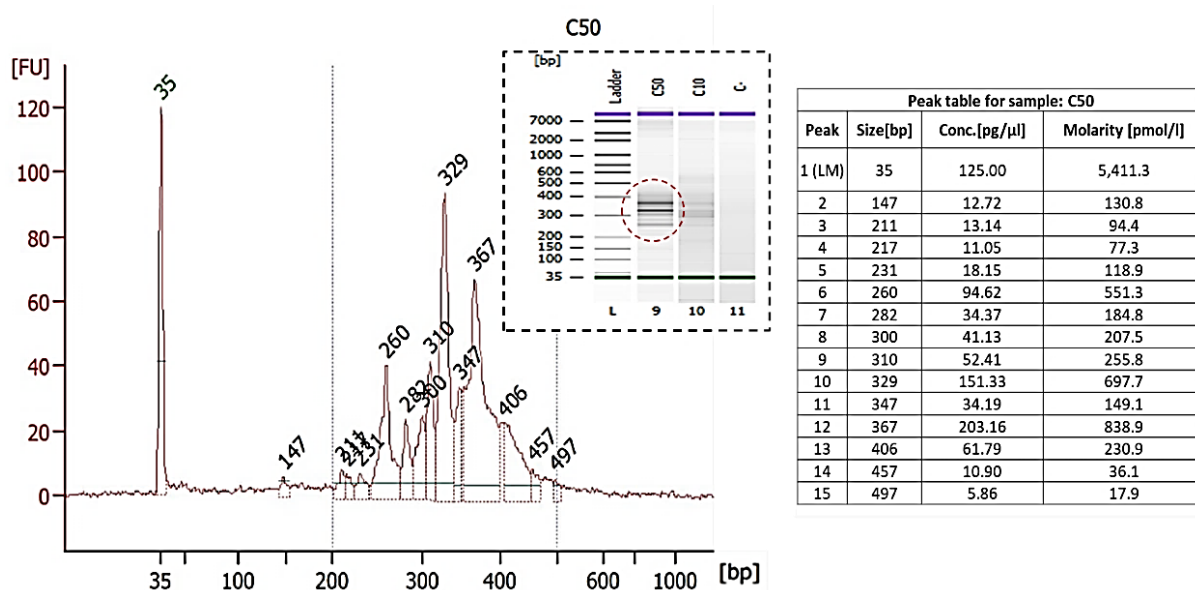


Figura 25. Evaluación cualitativa y cuantitativa de librerías mediante *Bioanalyzer 2100 HS* en una de las muestras control (C50). Electroferograma (izquierda) que representa la distribución de tamaños de los amplicones contenidos en las librerías. Cada pico correspondiente es cuantificado (derecha). Electroforesis mostrando el gel con los amplicones (centro).

Los resultados de la secuenciación en *Ion Torrent PGM™* se visualizaron en el software *Torrent Suite*. Todas las muestras secuenciadas cumplieron con los criterios establecidos de calidad: densidad de perlas o *Ion Sphere™* (ISP), ISP policlonales,

ISP de baja calidad, dímeros de adaptadores y el porcentaje de lecturas útiles (ISP con biblioteca que pasan los filtros anteriores). La densidad de carga en el chip fue alta en las dos secuenciaciones realizadas.

El número de secuencias obtenidas en los dos ensayos superó los 100.000 *reads* totales en todas las muestras y varió de acuerdo con la profundidad de cada secuenciación (total de muestras secuenciadas en el chip). En la Tabla 7 se resume el número de *reads* totales (crudos), post filtrado, así como el porcentaje de mapeo sobre los genes blanco para todas las muestras.

Tabla 7. Número de Reads obtenidos en las distintas etapas (secuenciación, *trimming* y mapeo) durante el análisis de las muestras control.

Ensayo*	Muestras	NÚMERO DE READS					Mapeo (%)**
		Totales	Post Filtrado	Descartados (%)	Mapeados	No mapeados	
I	CLI	427.569	419.309	1,97	418.459	789	99,80
	CLII	844.113	823.686	2,48	823.155	531	99,94
	CLIII	761.846	745.593	2,18	745.398	195	99,97
	CLIV	574.564	561.372	2,35	561.261	111	99,98
II	C50	133.250	130.156	2,38	130.003	153	99,88
	C+	163.858	160.143	2,32	159.701	442	99,72

*Las muestras dentro de cada ensayo se secuenciaron en la misma corrida. Los ensayos corresponden a dos corridas en diferente momento. ** Mapeo en los amplicones de referencia correspondientes a cada juego de *primers*. Ensayo I (cepas únicas): CLI 30000 copias/14 ciclos, CLII 3000 copias/17 ciclos, CLIII 300 copias/20 ciclos, CLIV 30 copias/23 ciclos. Ensayo II (cepas múltiples): C+ (igual copia inicial de genomas/género) y C50 (enriquecido en *Listeria monocytogenes*).

En todas las muestras, la proporción de *reads* descartados (por tamaño o calidad) fue baja y estuvo entorno al 2 % del total. Con relación al mapeo en los amplicones de referencia, en todas las muestras se obtuvo un alto nivel de mapeo, cercano al 100 % de los *reads*. Los *reads* descartados en el *trimming* y los no mapeados se alinearon contra la base de datos de Genbank y corresponden en su mayoría a secuencias cortas o con gaps que se alinean en algunos de los genes blanco.

4.2.2. Control Con Potencial Patogénico (Cepas Únicas)

Con el objetivo de evaluar el funcionamiento del panel, poner a punto la concentración de ADN inicial y los ciclos de PCR se construyó una mezcla con cuatro cepas de colección portadoras de 8 genes blanco (*E. coli*: *eaeA*, *stx1A*, *stx1B*; *Listeria monocytogenes*: *inlA*, *hlyA*, *Pseudomonas aeruginosa*: *exoA*, *exoS* y *Salmonella enterica*: *invA*). Se realizaron 4 tratamientos variando la cantidad inicial de copias de genomas y los ciclos de PCR (CLI: 30.000 copias/14 ciclos, CLII: 3.000 copias/17 ciclos, CLIII: 300 copias/20 ciclos, CLIV: 30 copias/23 ciclos).

4.2.2.1. Mapeo general

Las secuencias obtenidas de cada tratamiento mapearon en los ocho genes candidatos, lo que se ilustra en la Figura 26. Adicionalmente, un amplicón correspondiente al gen *exoU* de *Pseudomonas aeruginosa* (gen no presente en la cepa de referencia) se detectó en bajo número (< 0,0005 %) y con baja cobertura (fragmento de 40 pb). Este amplicón sólo se detectó en el tratamiento con más alta concentración inicial de molde (CLI), sin embargo, fue considerado para los cálculos de especificidad.

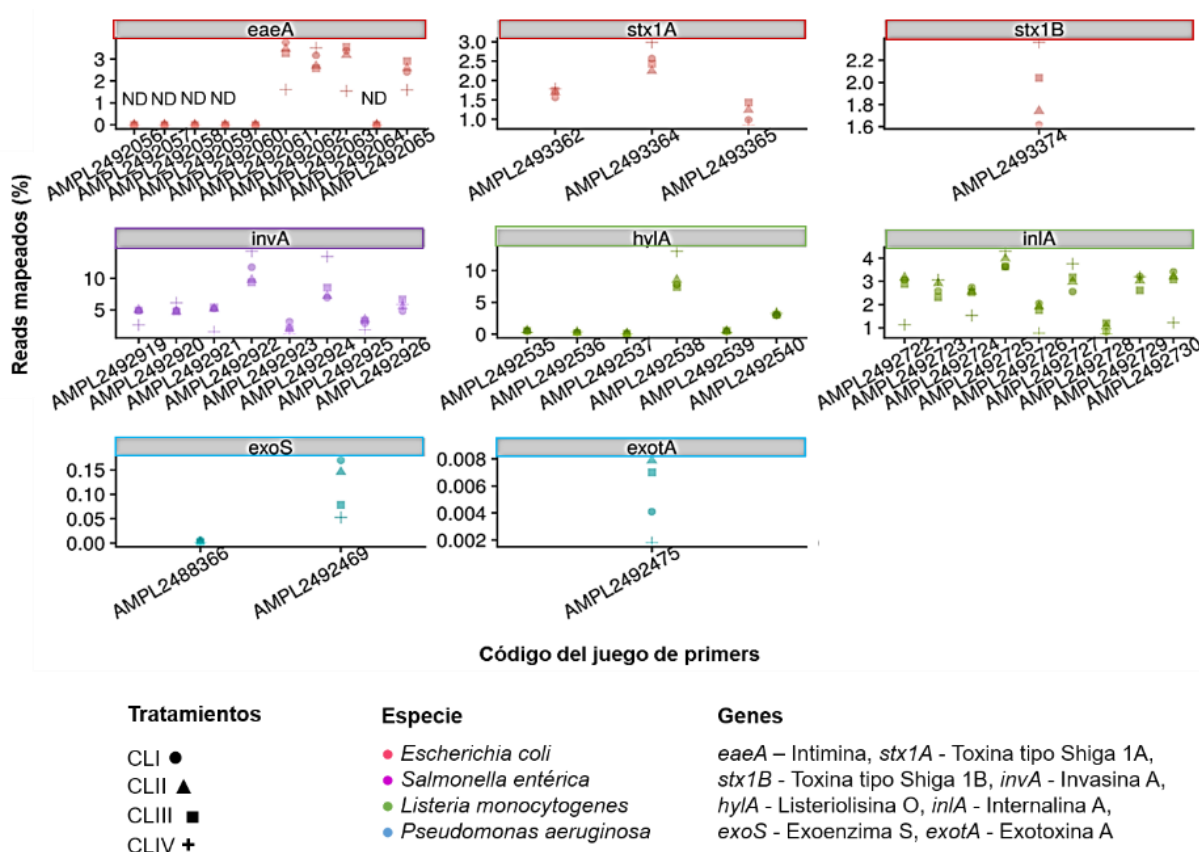


Figura 26. Mapeo general de *reads* del control con cepas únicas. Porcentaje de *reads* mapeados en cada amplicón para cada uno de los tratamientos del ensayo de cepas únicas. Distribución por especie y gen. Tratamientos: CLI 30000 copias/14 ciclos, CLII 3000 copias/17 ciclos, CLIII 300 copias/20 ciclos, CLIV 30 copias/23 ciclos. El porcentaje de *reads* mapeados se normalizó en función de los *reads* totales del tratamiento. ND= no detectado.

Para determinar si el aumento de ciclos de PCR compensaba la disminución de ADN molde dentro de los rangos recomendados por Ampliseq™, se evaluaron cuatro tratamientos variando la cantidad inicial de copias de genomas y los ciclos de PCR. No existieron diferencias significativas entre los tratamientos por lo que el aumento de ciclos de PCR compensaría la disminución de molde en los rangos estudiados (30-30.000 copias de genoma). En la Figura 27 se observan las distribuciones del mapeo en los cuatro tratamientos.

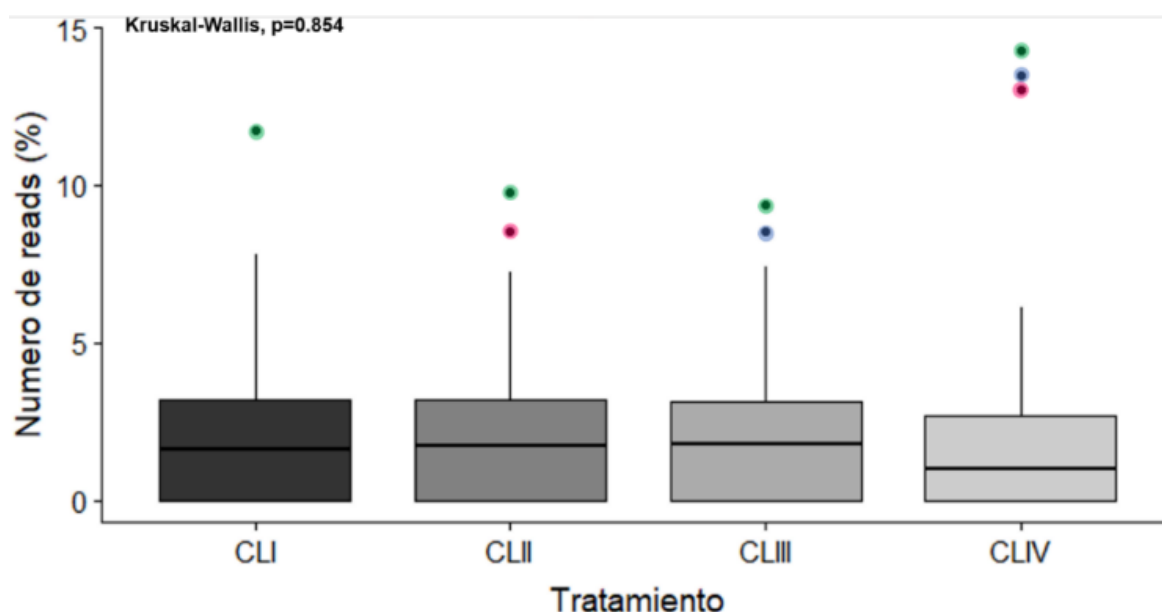


Figura 27. Reads mapeados del control de cepas únicas, comparación entre tratamientos. Diagrama de cajas del número de *reads* mapeados (%) de cada juego de *primer* por tratamiento. Tratamientos: CLI 30000 copias/14 ciclos, CLII 3000 copias/17 ciclos, CLIII 300 copias/20 ciclos, CLIV 30 copias/23 ciclos. El porcentaje de *reads* mapeados se normalizó en función de los *reads* totales del tratamiento. Los puntos en color corresponden a los mismos juegos de *primers*, gen *invA* de *Salmonella* (verde/azul) y gen *hlyA* de *Listeria monocytogenes* (rosado).

4.2.2.2. Mapeo por especie, gen y primer

El número de *reads* mapeados (%) fue diferente entre las cuatro especies estudiadas y los 8 genes candidatos (Figura 28), lo que puede deberse a diferencias en el número de copias del gen que porta cada cepa, a un desbalance en las copias iniciales de los genomas o al número y/o eficiencias de cada *primer*, entre otros. Se observó un mayor mapeo en el gen de la invasina de *Salmonella enterica* seguido por los genes de *Listeria monocytogenes* y *Escherichia coli*. El nivel de detección fue bajo en los genes de *Pseudomonas aeruginosa* en todos los tratamientos (<1 % del total de *reads*).

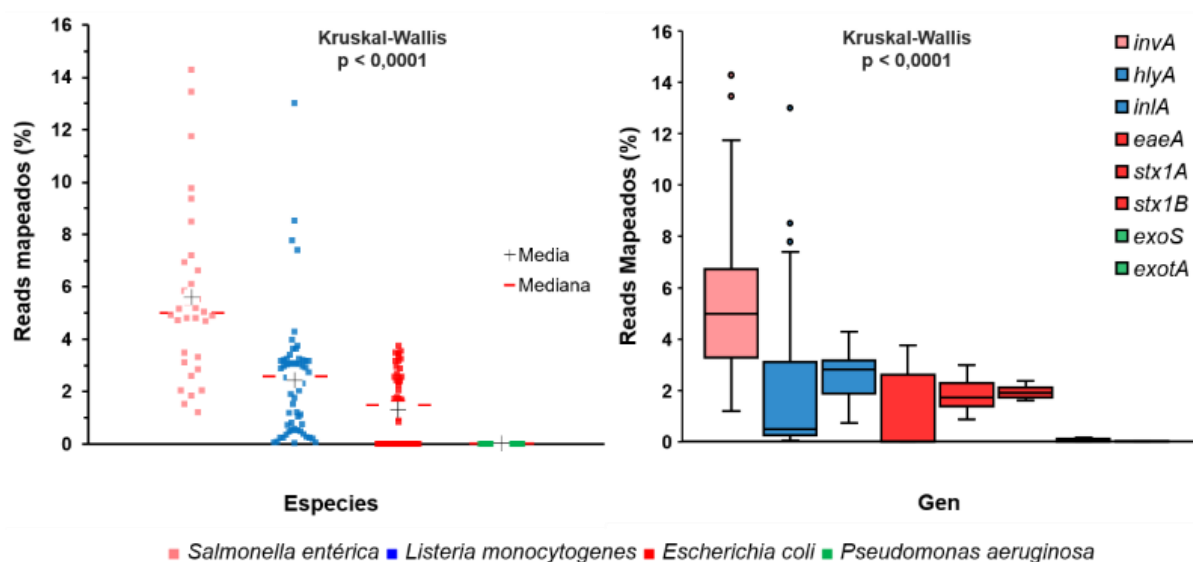


Figura 28. Reads mapeados del control de cepas únicas, comparación entre especies y genes. Diagrama de cajas del número de *reads* mapeados (%) de cada juego de *primer* por tratamiento. El porcentaje de *reads* mapeados se normalizó en función de los *reads* totales del tratamiento.

En la Figura 29, se ilustra la distribución de *reads* mapeados por juego de *primer* indicando las posibles diferencias entre regiones conservadas y no conservadas del gen (*primers* localizados en región conservada remarcados en rojo).

Con relación a los genes con zonas conservadas, por ejemplo, en el caso del gen de la intimina de *E. coli* (*eaeA*), no todos los *primers* amplificaron y los dirigidos a una región conservada presentaron valores más altos y homogéneos en comparación con los otros. En contraposición, el gen de la internalina A de *Listeria monocytogenes* (*inlA*) no mostró diferencias entre las zonas y cada juego de *primer* mapeo diferente.

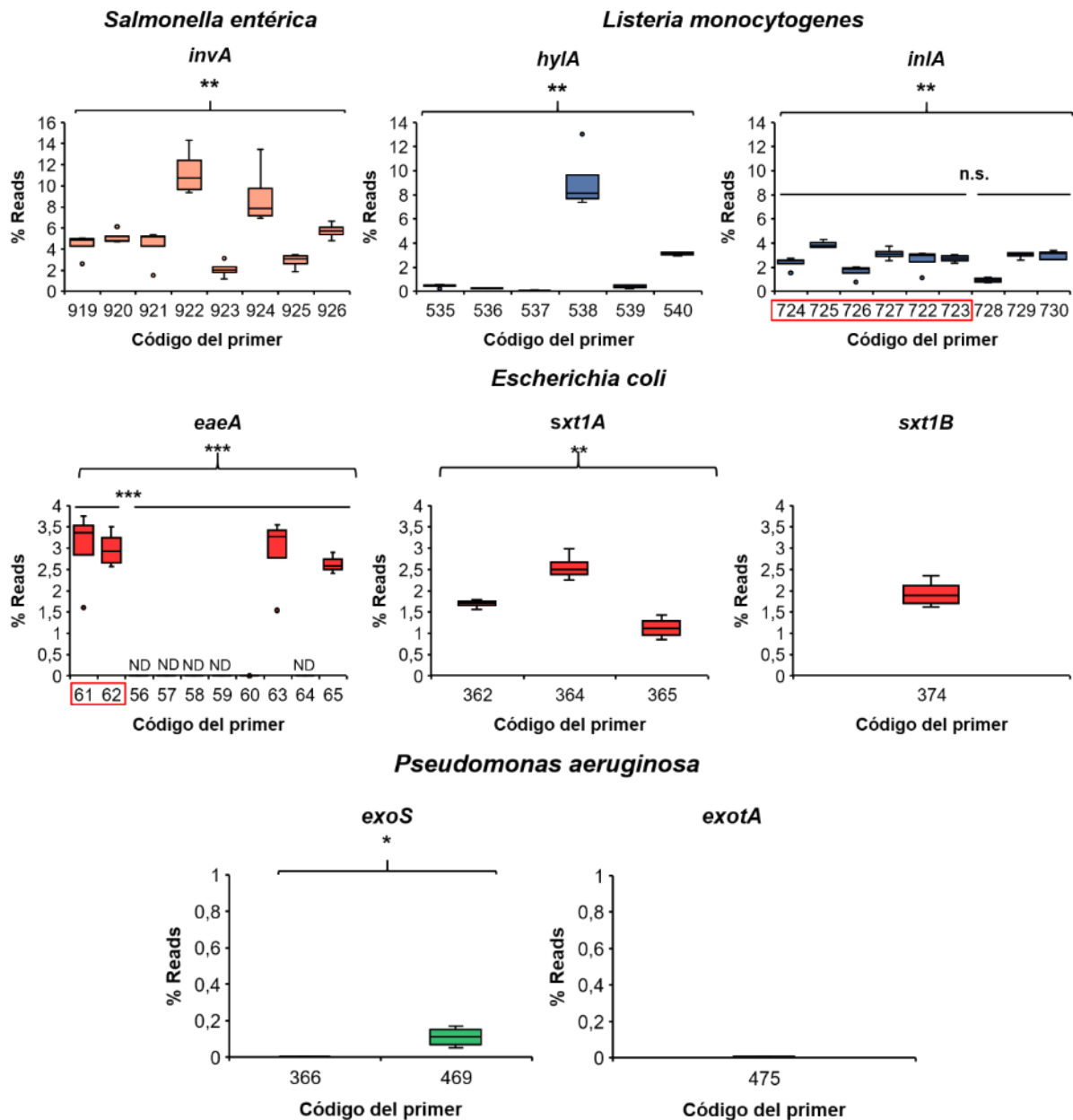


Figura 29. Reads mapeados del control de cepas únicas, comparación entre *primers* según la especie. Reads mapeados (%) en genes de virulencia de *Salmonella*, *Listeria*, *E. coli* y *Pseudomonas* para todos los tratamientos. Distribución por juego de *primers*. Asteriscos indican diferencias entre juegos de *primers* o entre grupos de *primers* con blanco en zonas conservadas (recuadro rojo) y no conservadas (sin marcar). Test de Kruskal-Wallis con p : * $<0,05$ - $0,01$; ** $<0,01$ - $0,001$; *** $<0,0001$, n.s.= no significativo. ND= no detectado.

4.2.2.3. Distribución por set de primers

Como se mencionó anteriormente los *primers* del panel Fvir se encuentran distribuidos en dos sets, de manera de maximizar el grado la compatibilidad entre *primers*. Esto implica que las PCR iniciales se realicen de manera independiente una de la otra, por lo que exploramos su *performance* por set para evaluar el posible sesgo. No existieron diferencias en el mapeo entre los dos sets, la *performance* fue similar entre las PCR, si bien los valores *outliers* observados en el ensayo se encontraron en el set de *primers* N° 2 (Figura 30).

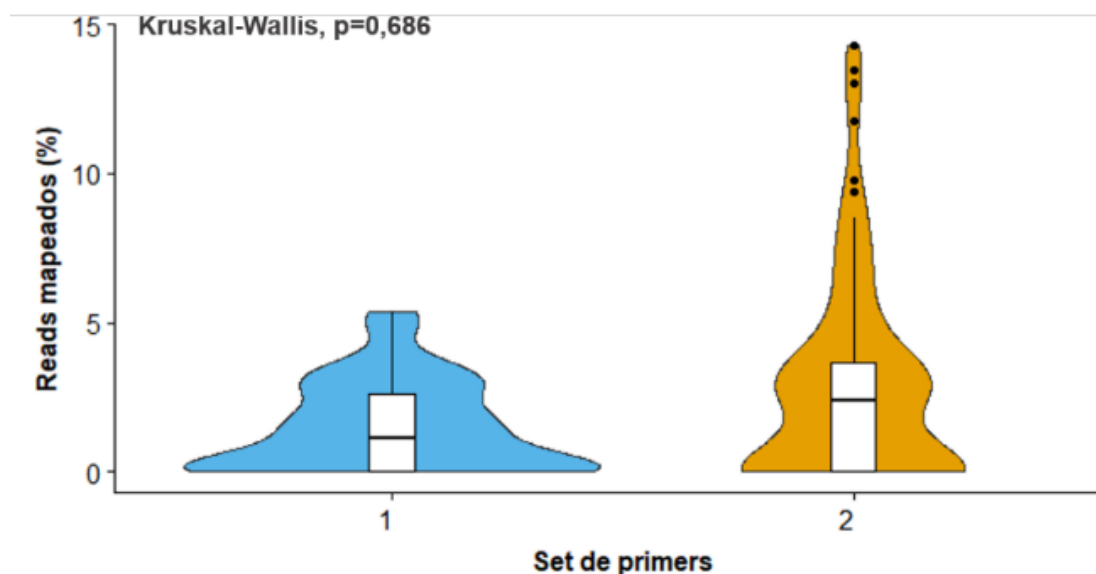


Figura 30. Reads mapeados del control de cepas únicas, comparación entre *sets* de *primers*. Diagrama de violines y cajas del número de *reads* mapeados (%) por *set* de *primers* en todos los tratamientos.

4.2.2.4. Resumen general del funcionamiento del panel

Evaluamos el funcionamiento del panel en una muestra control con 8 genes candidatos (40 juegos de *primers* dirigidos a ellos) en un rango de 30-30.000 copias de genomas, lo que se ilustra en la Figura 31. En esta se observa que se logró la detección de todos los genes candidatos debido a la amplificación generada por el 87,5 % de los *primers* diseñados. Los *primers* que no amplificaron corresponden a regiones no conservadas de un gen (*eaeA*), cuyo diseño se basó en un genoma de referencia posiblemente con variantes en la secuencia en comparación a la cepa utilizada (cepa nativa o de campo).

La sensibilidad general de los *primers* (promedio) se calculó en base a los cuatro tratamientos del ensayo y fue del 95, 2 %. Para los genes de *Salmonella*, *Listeria* y *Pseudomonas* fue del 100 % mientras que en *E. coli* alcanzó el 81 % .

En relación con la especificidad de los *primers*, sólo se observó un falso positivo en un par de *primers* (gen *exoU* de *Pseudomonas*) y en uno de los tratamientos (CLI) por lo que la especificidad promedio alcanzó el 99,5 %.

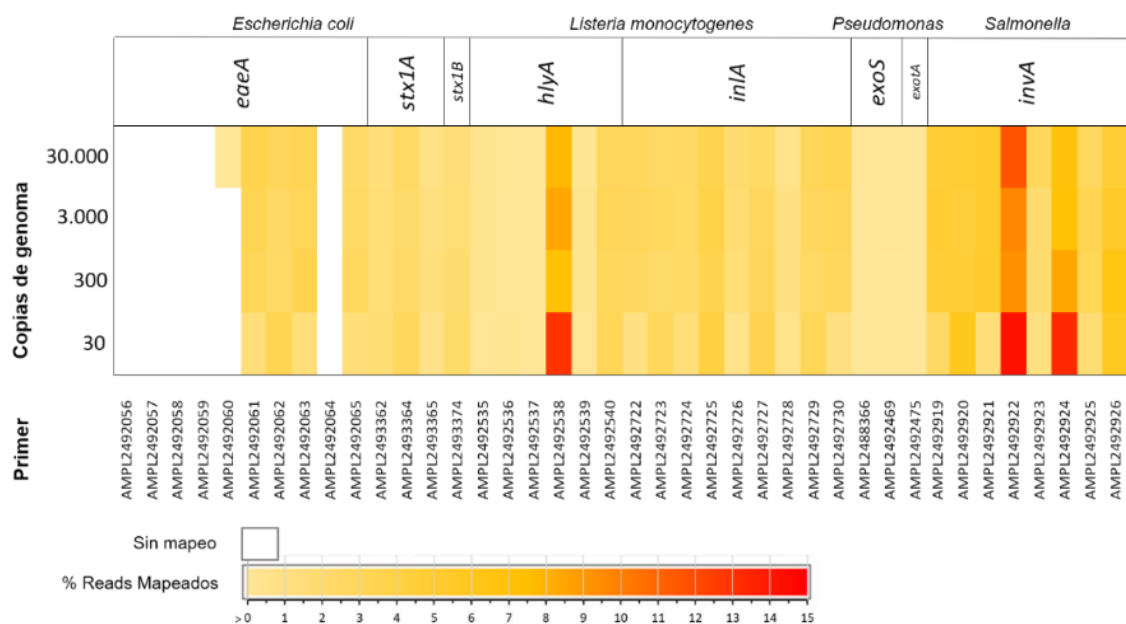


Figura 31. Resumen de la performance del panel en el control con cepas únicas. Mapeo general de *primers* por tratamiento. Tratamientos: CLI 30000 copias/14 ciclos, CLII 3000 copias/17 ciclos, CLIII 300 copias/20 ciclos, CLIV 30 copias/23 ciclos. El porcentaje de *reads* mapeados se normalizó en función de los reads totales del tratamiento.

4.2.3. Control Con Potencial Patogénico (Cepas Múltiples)

Con el objetivo de ampliar el número de *primers* a probar se realizó un ensayo experimental con una muestra más diversa utilizando un pool de ADN con el total de cepas que conforman la colección (92 aislamientos de 8 géneros bacterianos, Tabla S1, Anexo I). Se realizaron dos tratamientos con diferente relación de copias de genomas iniciales/género para explorar si el enriquecimiento de la muestra en un microorganismo afecta la detección de otros menos representados. El tratamiento C+ con igual copia inicial de genomas de cada género (10^9 copias/ μ L). El tratamiento C50 con un orden logarítmico más de *Listeria monocytogenes* en comparación con los otros géneros del pool (10^{10} copias/ μ L).

4.2.3.1. Mapeo general

En la Figura 32 se presenta de manera ilustrativa los *reads* mapeados sobre cada amplicón en ambos tratamientos, donde se detectaron 25 genes pertenecientes a los géneros que conformaron el pool (*Escherichia*, *Salmonella*, *Listeria*, *Pseudomonas*, *Enterococcus*, *Staphylococcus*, *Aeromonas* y *Cylindrospermopsis*). Los mismos genes fueron detectados en ambos tratamientos, 16 de los 25 genes detectados habían sido caracterizados anteriormente y reportados en alguna de las cepas (ver Tabla S1 en Anexo I). Adicionalmente se detectaron 9 genes de los que se desconocía su presencia en las cepas: *fnba*, *tst* (*S. aureus*), *f41A*, *stx2A*, *stx2B* (*E. coli*), *ace* (*E. faecalis*) y *acm*, *hyl* y *pilB* (*E. faecium*). En ningún caso se detectaron genes pertenecientes a otros grupos bacterianos distintos a los que se utilizaron. En ambos experimentos (C+ y C50) se utilizó la misma cepa de *Pseudomonas aeruginosa* que en el ensayo experimental I, no detectándose el gen *exoU* (ausente en la cepa). Existieron variaciones en el mapeo según el tratamiento, especie, gen y juego de *primers*. Los resultados más relevantes se presentaran a continuación.

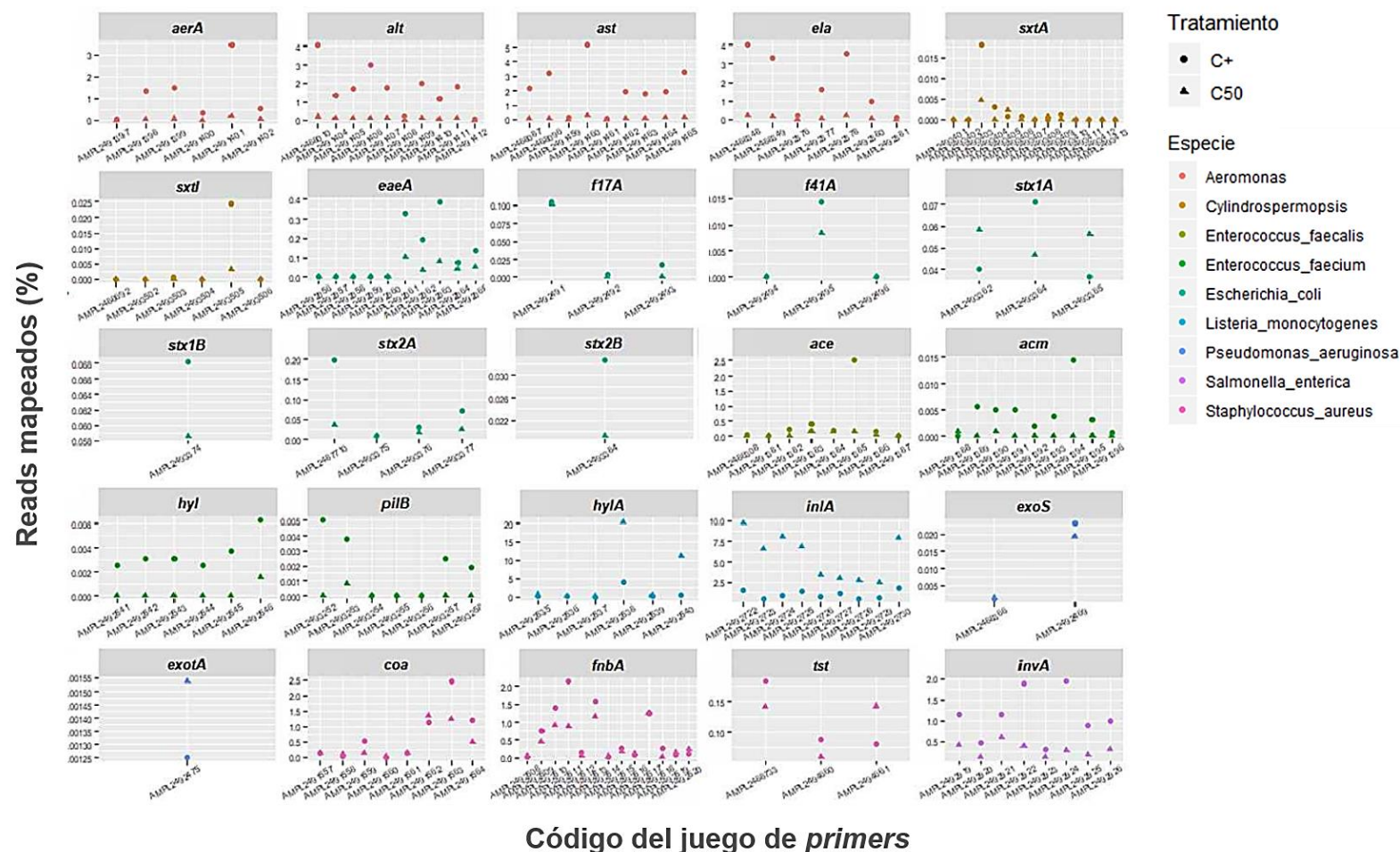


Figura 32. Mapeo general de *reads* del control con múltiples cepas. Reads mapeados (%) en cada amplicón, distribución por tratamiento, especie y gen. Tratamientos: C+ (igual copia inicial de genomas/género) y C50 (enriquecido en género *Listeria*). Genes: Aerolisina (*aerA*); Enterotoxinas citotónicas (*alt*, *ast*); Elastasa (*ela*); Saxitoxina (*sxtA*, *sxtI*); Intimina (*eaeA*); Fimbrias 17 y 41 (*f17A*, *f41A*); Toxina tipo Shiga (*stx1A*, *stx1B*, *stx2A*, *stx2B*); Adhesinas de unión al colágeno (*ace*, *acm*); Hialuronidasa (*hyl*); Pili (*pilB*); Listeriolisina O (*hylA*); Internalina A (*inlA*); Exoenzima S (*exoS*); Exotoxina A (*exotA*); Coagulasa (*coa*); Proteína de unión a la fibronectina (*fnbA*); Toxina del shock séptico (*tst*); Invasina A (*invA*). *El porcentaje de *reads* mapeados se normalizó en función de los *reads* totales del tratamiento.

4.2.3.2. Mapeo por tratamiento y género

En la Figura 33 se resume: el mapeo general en ambos tratamientos (a), la relación entre las copias iniciales de genomas y los *reads* mapeados según las especies (b) y las distribuciones del mapeo por especie para los dos tratamientos (c y d).

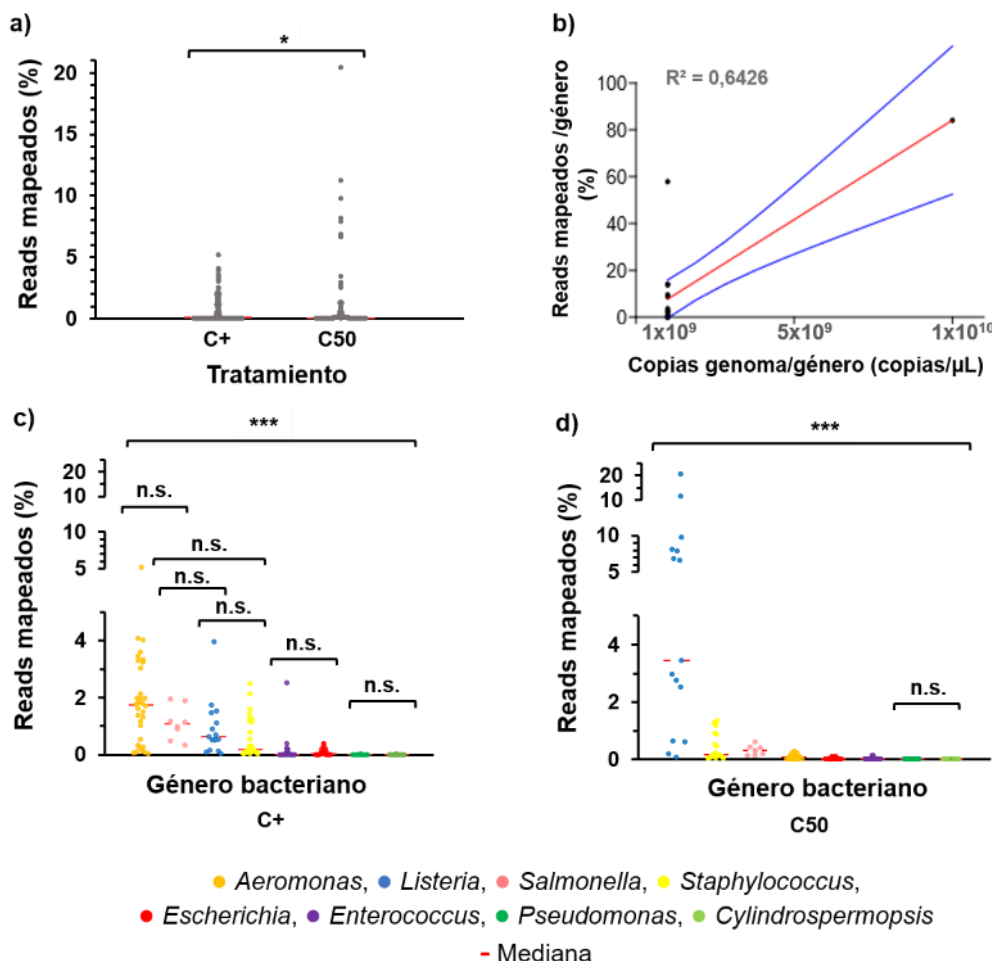


Figura 33. Reads mapeados del control de cepas múltiples, comparación de tratamientos y especies. a) Distribución de *reads* totales por tratamiento (%). b) Relación entre copias iniciales de genomas y *reads* mapeados según el microorganismo (todos los géneros bacterianos y en ambos tratamientos, las líneas azules representan el intervalo de confianza de 95 %). En c) y d) Distribución de *reads* por género bacteriano en los tratamientos C+ y C50 respectivamente. El porcentaje de *reads* mapeados se normalizó en función de los *reads* totales del tratamiento. Asteriscos indican diferencias entre grupos, test de Kruskal-Wallis con p: *<0,05-0,01; **<0,01-0,001; ***<0,0001 y n.s.= no significativo.

Si bien se detectaron los mismos genes en ambos tratamientos, existieron diferencias significativas entre ellos (33-a). Las diferencias del mapeo entre las especies pueden explicarse en parte al diseño, dada la estructura poblacional (géneros bacterianos) que conforman tratamiento (ver más adelante el control mediante el ARNr 16S). Los *reads* mapeados se correlacionaron con las copias de genomas iniciales en los tratamientos (33-b) pero con diferencias de acuerdo con el grupo microbiano, lo que podría explicarse en parte al número de cepas que porten o no el gen candidato dentro del pool de ese género. En el tratamientos C+ (igual copia de genomas/género) existieron diferencias en los *reads* mapeados entre los géneros, pero con valores

similares entre algunos de ellos, acercándose más a la homogeneidad perseguida en el diseño para este tratamiento (Figura 33-c). El enriquecimiento del género *Listeria* en el tratamiento C50 se refleja en el mapeo de *reads* para dicho género, donde se observó que estos superaron el 80 % del total de la muestra (Figura 33-d).

En la Figura 34 se presentan los resultados del mapeo por gen en ambos tratamientos.

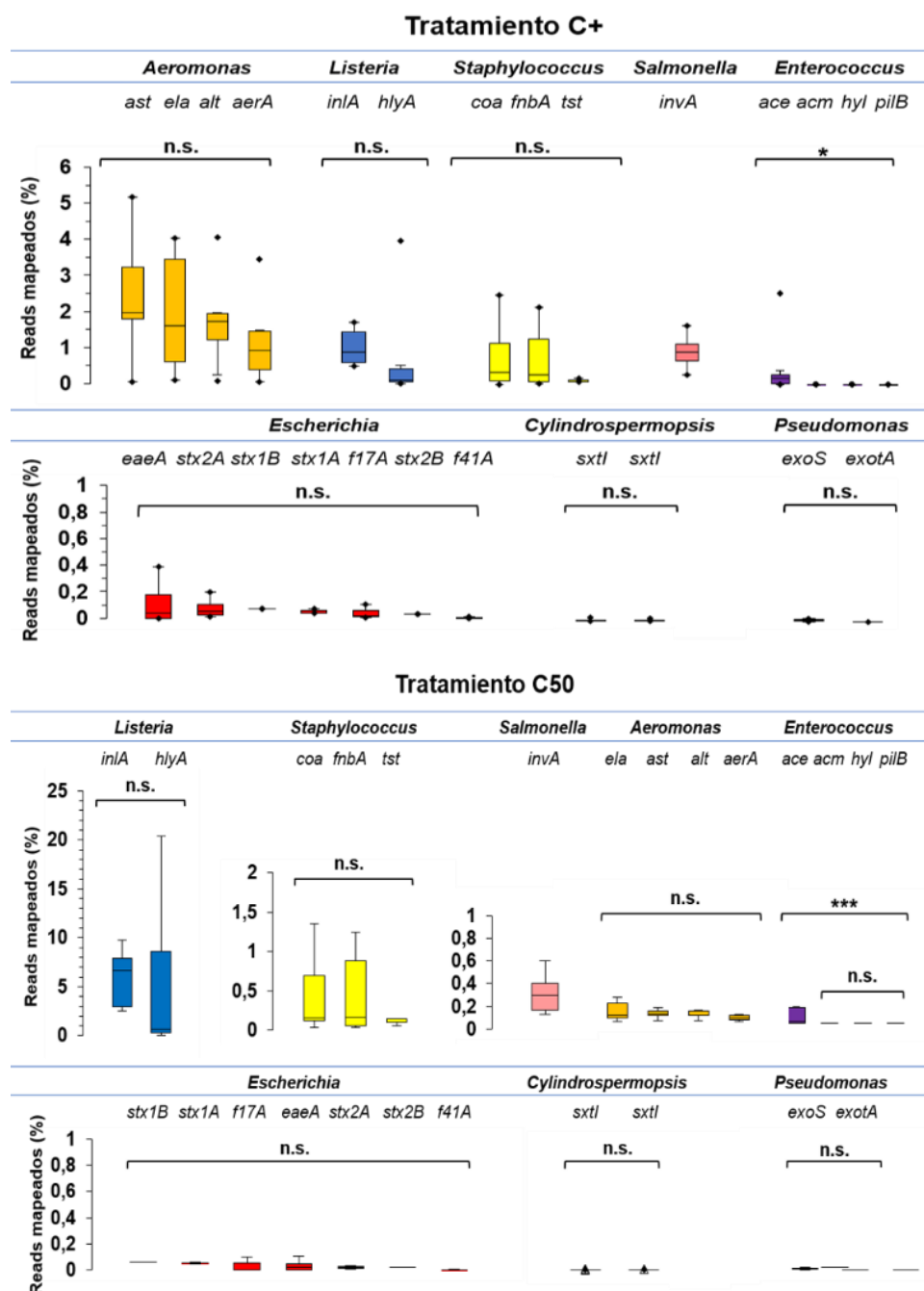


Figura 34. Reads mapeados del control de cepas múltiples, comparación de tratamientos y genes. Porcentaje de *reads* mapeados (%) en los tratamientos del control de múltiples cepas. De arriba hacia abajo tratamientos C+ (igual relación copias de genomas) y C50 (mayor copias de genomas de *L. monocytogenes*) respectivamente. El porcentaje de *reads* mapeados se normalizó en función de los *reads* totales de la muestra. Asteriscos indican diferencias entre grupos, test de Kruskal-Wallis con p: * $<0,05$ - $0,01$; ** $<0,01$ - $0,001$; *** $<0,0001$. n.s.= no significativo.

A excepción del género *Enterococcus spp.*, no existieron diferencias significativas en el mapeo entre distintos genes de un mismo género bacteriano. Esto indicaría una distribución homogénea de los genes candidatos entre los aislamientos de un mismo género que se utilizaron para conformar el pool de ADN. En los enterococos, se presentó un mayor mapeo del gen *ace* (gen particular de *E. faecalis*) en comparación con los genes *acm*, *hyl* y *pilB* (genes de *E. faecium*). Este resultado puede responder al número de aislamientos utilizados en el pool de ADN (n=49) y las proporciones entre las dos especies (mayor proporción de *E. faecalis*).

4.2.3.3. Funcionamiento general del panel

El porcentaje de *primers* que amplificaron en relación al total diseñado fue alto y similar entre los tratamientos (Figura 35). En el género *Enterococcus* se observó menor proporción de amplificación en el tratamiento C50 probablemente por disminuir las copias de genoma en relación al enriquecimiento en el género *Listeria* en este tratamiento.

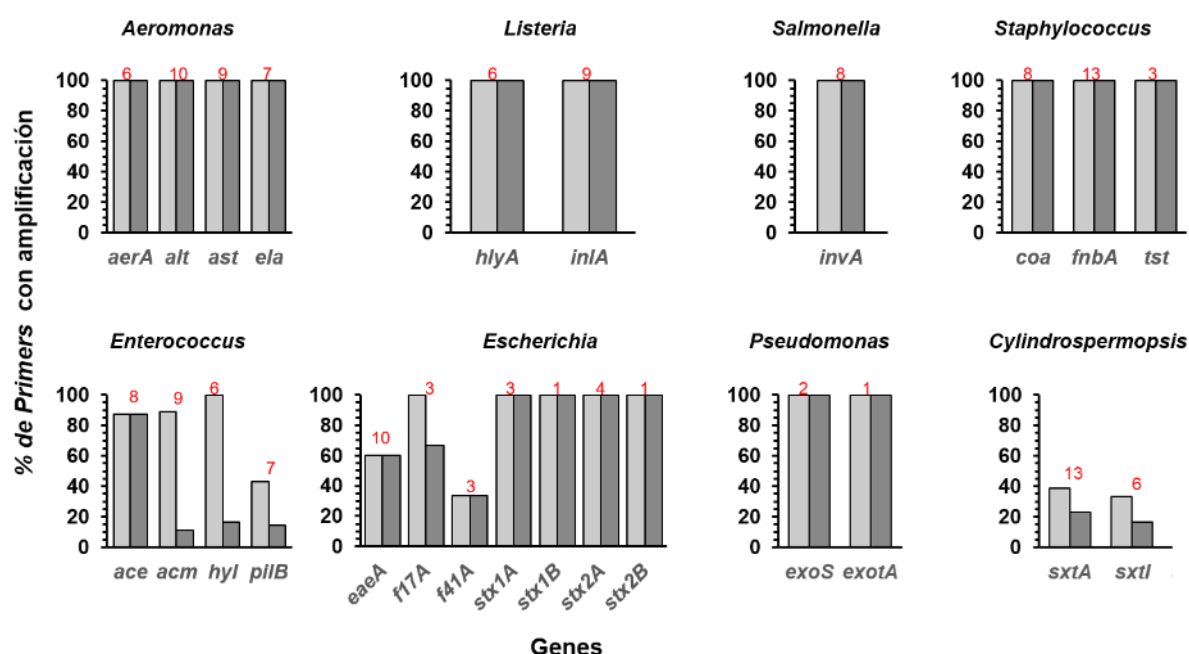


Figura 35. Porcentaje de *primers* con amplificación del control de cepas múltiples (*primers* amplificados/ total de *primers* por gen). Distribución por género y gen. Tratamientos C+ (■) y C50 (■). En rojo se indica el número de juegos de *primers* diseñados para ese gen.

En resumen en el ensayo se detectaron 25 genes de virulencia que involucran el funcionamiento de 167 juegos de *primers* (Figura 36). El 67 % de los *primers* amplificaron en ambos tratamientos (n=112) y el 18 % no amplificó en ninguno de los dos (n=31). El tratamiento C+ (igual copias de genomas iniciales) presentó mayor número de *primers* que amplifican (80 %, n=133) muestra más homogénea. El enriquecimiento con *Listeria* del tratamiento C50 posiblemente afectó el desempeño de los *primers* en algunos de los genes menos representados (69 %, n =115). Estos juegos de *primers* que no amplificaron podrían considerarse como falsos negativos en la muestra. Sin embargo, la existencia de múltiples *primers* sobre cada gen posibilitó la detección de los mismos genes en ambas muestras.

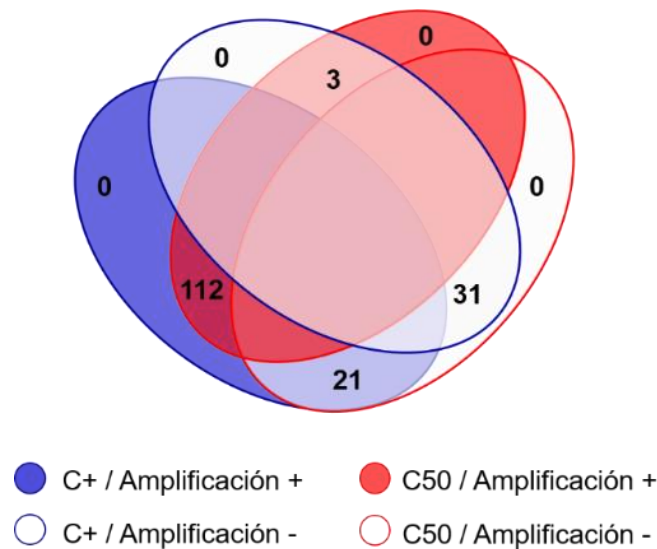


Figura 36. Resumen del funcionamiento de *primers* en el control con cepas múltiples, comparación entre tratamientos. Diagrama de Venn representando el número de juegos de *primers* con y sin amplificación en los tratamientos C+ y C50 (n=167).

Por ultimo, no existieron diferencia en el mapeo entre los dos *sets* de *primers* que conforman el panel Fvir ($p=0,29$). Tampoco existieron diferencias en el porcentaje de mapeo entre *primers* localizados en zonas conservadas y no conservadas ($p=0,12$).

4.2.3.4. Control de la estructura poblacional

El control de la estructura poblacional se realizó para el tratamiento C50 mediante la secuenciación masiva del gen ribosomal 16 S. De la secuenciación se obtuvo un total de 258.513 *reads*, de los cuales se filtraron 1981 secuencias que no cumplieron los criterios de tamaño, calidad o presentaban secuencias quiméricas (total =0,8 %). Posteriormente se eliminaron las secuencias correspondientes al ARN 16S de organelos (mitocondrias y cloroplastos), conformando la tabla de OTUs con un total de 256.513 secuencias que se agruparon en 712 OTUs. Las especies de la comunidad contenida en la muestra se estimaron mediante la evaluación de la diversidad alfa (Figura 37).

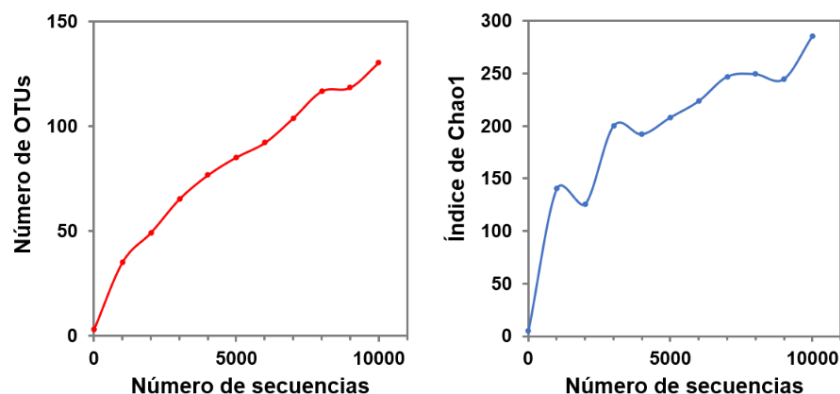


Figura 37. Estimadores de diversidad alfa en el control de la estructura poblacional del tratamiento C50. Curva de rarefacción e Índice de Chao1 calculados a un tamaño de muestra estandarizado = 10.000 secuencias.

El análisis mediante QIIME permitió la asignación taxonómica del 98,5 % de las OTUs a un nivel de género o de grupo de géneros para el caso de *Escherichia-Shigella*. La estructura observada en la comunidad concuerda con lo esperado para la muestra de acuerdo con el diseño experimental (8 géneros con un enriquecimiento de 50 % en *Listeria monocytogenes*) y con la aproximación mediante el estudio de los genes de virulencia del panel Fvir (sin diferencias significativas entre las aproximaciones), Figura 38.

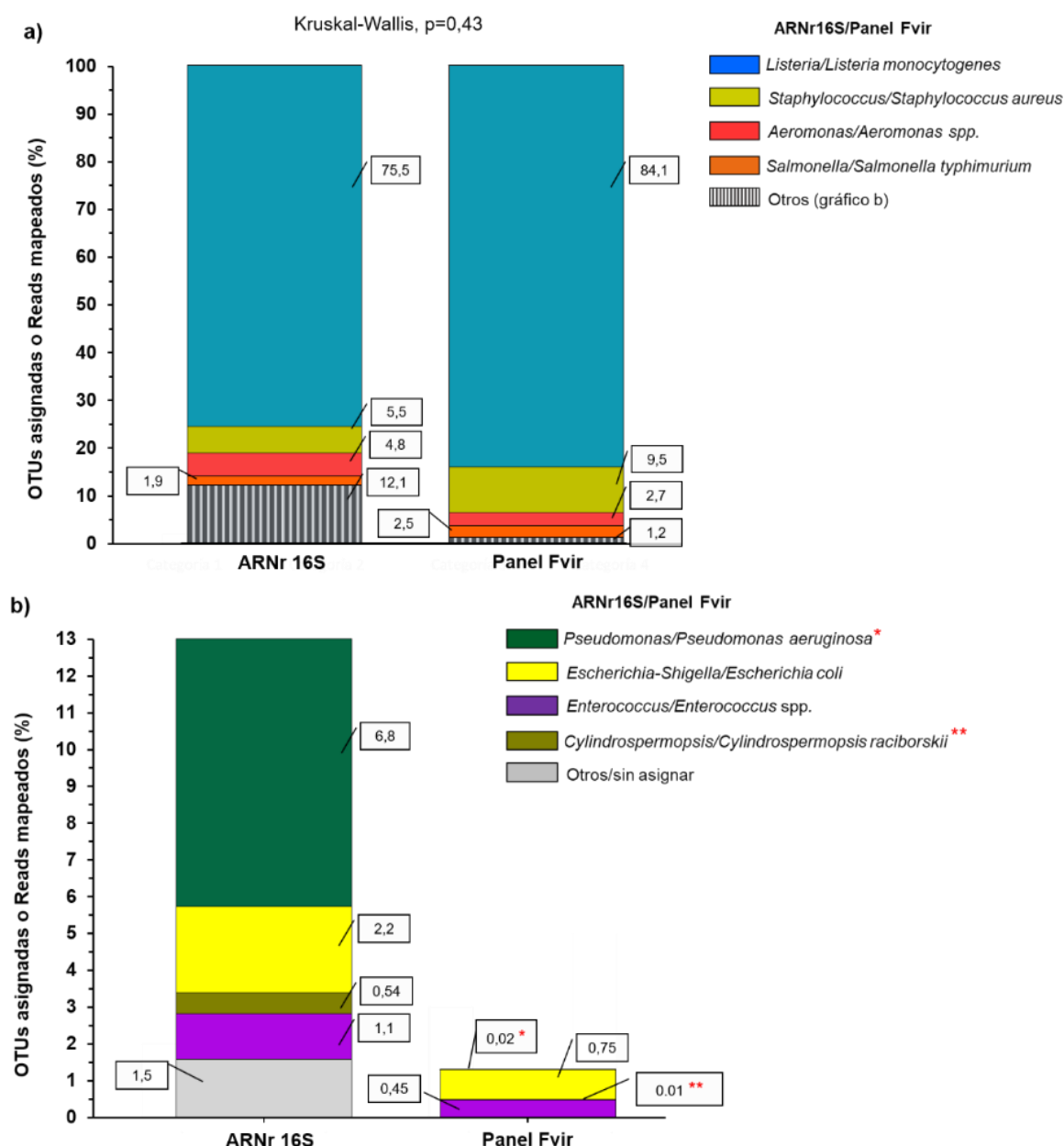


Figura 38. Distribución y comparación entre el porcentaje de OTUs asignadas/taxón (ARNr 16S) y el porcentaje de *reads* mapeados en factores de virulencia/especie (panel Fvir) en la muestra control C50. a) Géneros mayoritarios en ambas aproximaciones, b) géneros menos abundantes (<10 %).

El porcentaje de OTUs asignadas se correlacionó positivamente con el número de copias iniciales de genomas y con los *reads* mapeados en factores de virulencia del panel (Figura 39). Estos resultados validan el diseño experimental y dan sustento al funcionamiento del panel.

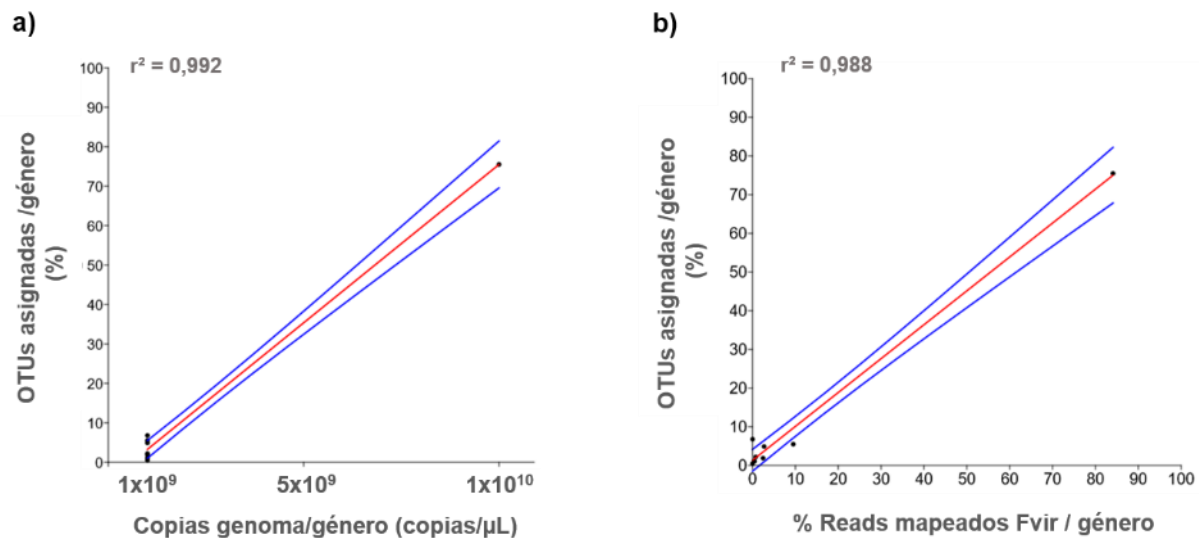


Figura 39. Correlaciones entre la aproximación metagenómica del ARNr 16S y las copias de genomas iniciales o los *reads* mapeados en genes de virulencia en el tratamiento C50. Se representa en: a) Correlación entre porcentaje de OTUs asignadas/género y copias de genomas iniciales/género y en b) Correlación entre porcentaje de OTUs asignadas/género y porcentaje de *reads* mapeados en factores de virulencia/género (Panel Fvir). Las líneas azules representan el intervalo de confianza de 95 %.

4.3. Evaluación Del Panel en Una Muestra Ambiental

La muestra fue seleccionada en base a sus antecedentes de toxicidad y corresponde a un evento de floración del complejo *Microcystis aeruginosa* en la estación de Gualeguaycito, situada en uno de los brazos del río Uruguay. La muestra fue caracterizada por Deus-Álvarez, (2018) quien evaluó la relación entre el tamaño colonial y la capacidad de producción de microcistinas, evaluando la abundancia del gen *mcyE* y la concentración de microcistinas totales. En la Tabla 8 se presenta un resumen de las características evaluadas.

Tabla 8. Información de la muestra de Gualeguaycito obtenida por Deus, (2018). Análisis de toxicidad en las fracciones obtenidas de la CMA.

Parámetros Evaluados	Fracciones*			
	G	M	P	C
Tamaño de la fracción (µm)	> 150	60-150	20-60	< 20
Copias del gen <i>mcyE</i> ** (copias/mL)	19614	8	1661	0,08
Microcistinas totales (µg/L) ***	483 ± 70	127 ± 12	72 ± 11	0,66 ± 0,13

*Las fracciones corresponden según el tamaño a células individuales (C), pequeñas (P), medianas (M) y grandes (G). **Determinado por qPCR. ***Concentración de microcistinas totales determinado por ELISA policlonal (valores promedio ± desvío estándar).

4.3.1. Amplificación y Secuenciación

Para la aplicación del panel Fvir se conformó un pool de ADN a partir de las fracciones. En base a los resultados reportados y de acuerdo con el volumen y concentración de ADN de las fracciones, el número estimado de copias del gen *mcyE* que se utilizaron en el pool fue de 4×10^5 copias/µL.

De la secuenciación de la muestra se obtuvo un total de 322.383 *reads* de los cuales se filtró el 7,3 % en base a los criterios establecidos de calidad y tamaño de las secuencias. Posteriormente las secuencias se mapearon contra el archivo de referencia de amplicones del panel, obteniendo un mapeo del 87,4 %. Las secuencias no mapeadas corresponden a aquellas descartadas en base a los criterios de *mismatches*, longitud e identidad con el archivo de referencia.

4.3.2. Genes de Virulencia Amplificados

En la Figura 40 se representa el mapeo sobre los factores de virulencia de la muestra. Más del 99,8 % de los *reads* mapearon en los genes relacionados a la producción de microcistinas (*mcyE* y *mcyJ*) observando un mayor mapeo sobre el gen *mcyE* (96,8%). Otros factores de virulencia relacionados a otros géneros de cianobacterias y a otros grupos bacterianos fueron detectados con valores menores al 1 % del total de *reads*. En total se detectaron 16 genes de virulencia que involucran a varios grupos/géneros de cianobacterias, *Aeromonas* spp., *Listeria monocytogenes*, *Staphylococcus aureus*, *Escherichia coli* y *Enterococcus faecalis*.

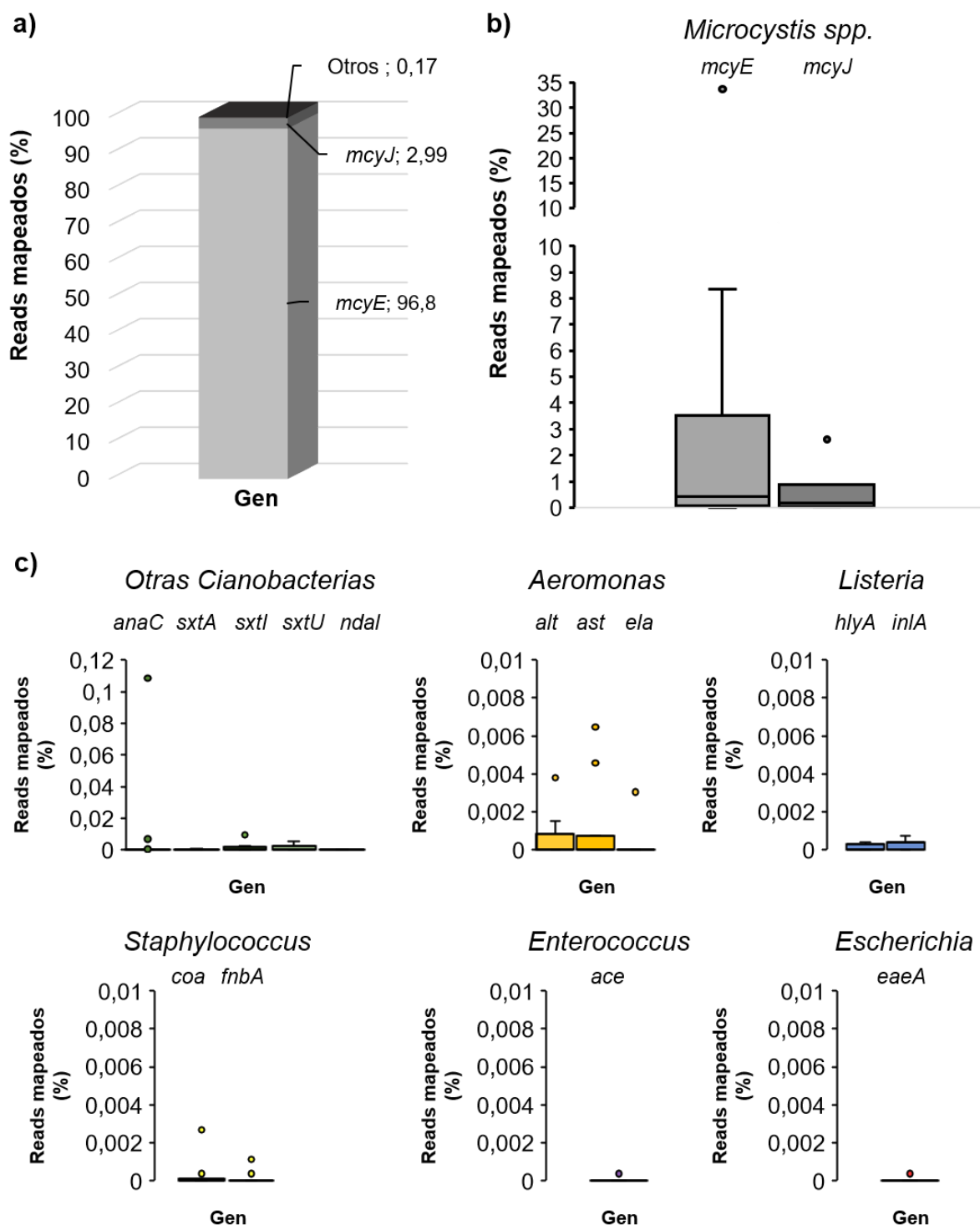


Figura 40. Reads mapeados en factores de virulencia bacterianos en la muestra ambiental de Gualaguaycito (porcentaje normalizado al total de *reads* mapeados en la muestra). Los *boxplots* corresponden a los valores de mapeo de cada juego de *primers*. a) Mapeo general. b) Genes con mayor abundancia: producción de microcistinas (*mcy E* y *mcyJ*). c) Genes menos abundantes (< 1 % de *reads*): Otras cianobacterias: anatoxina (*anaC*), saxitoxina (*sxtA*, *sxtI*, *sxtU*) y nodularina (*ndaI*); *Aeromonas* spp.: enterotoxina termoestable y termolábil (*alt*, *ast*), elastasa (*ela*); *Listeria monocytogenes*: listeriolisina O (*hlyA*), internalina A (*inlA*); *Staphylococcus aureus*: coagulasa (*coa*), fibronectina A (*fnbA*); *Enterococcus faecalis*: adhesina de unión al colágeno (*ace*); *Escherichia coli*: intimina (*eaeA*).

4.3.3. Funcionamiento Del Panel

En la Figura 41 se muestra el porcentaje de *primers* que amplificaron en cada uno de los genes. Los 16 genes detectados involucran el funcionamiento de 179 pares de *primers* de los cuales 80 mostraron amplificación (45 %). Los genes con mayor mapeo (*mcyE* y *mcyJ*) mostraron amplificación en todos los juegos de *primers* diseñados (41 juegos de *primers*) mientras que en los genes con menor mapeo el porcentaje de *primers* con amplificación varió entre 10-67 %.

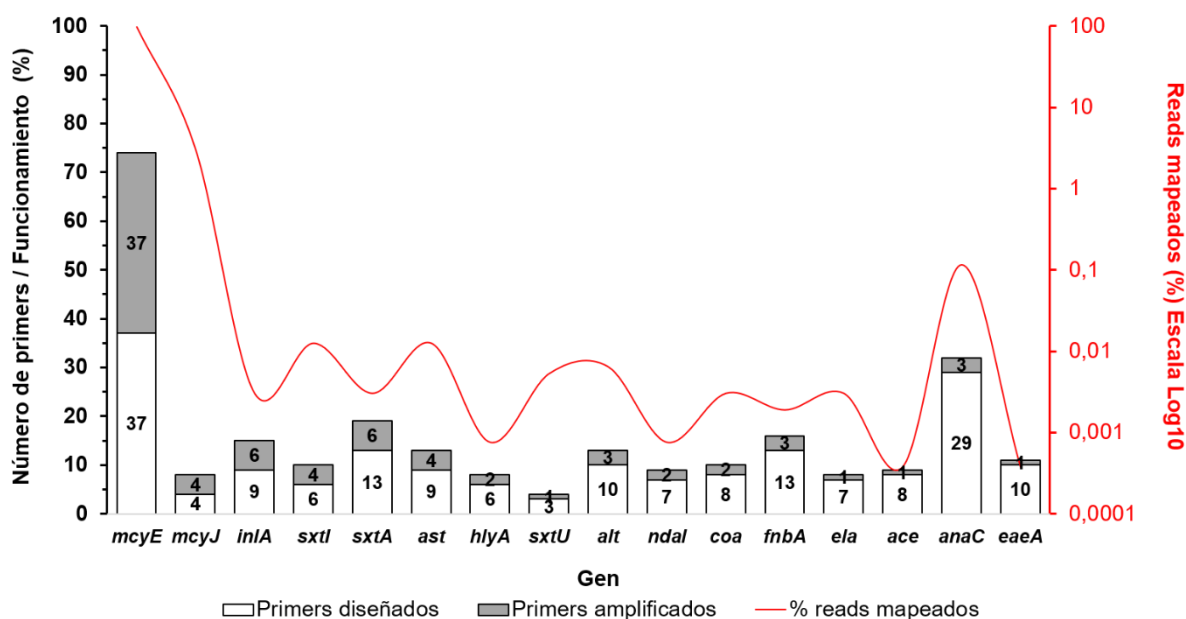


Figura 41. Genes de virulencia amplificados en la muestra de Gualguaycito. Las barras indican el número de *primers* diseñados y número de *primers* con amplificación. En el eje Y el porcentaje de funcionamiento se calculó como el cociente entre *primers* diseñados y amplificados. En el eje secundario se representa el porcentaje de *reads* mapeados para cada gen (escala Log10).

El porcentaje de mapeo de estos genes varió de acuerdo con el juego de *primers* diseñado. Sin embargo, en ambos genes se observó que uno de los juegos de *primers* presentó un mapeo mayor en comparación con el resto.

Los genes *mcy* no mostraron una región conservada durante el análisis (*primers* generados en base a *Microcystis aeruginosa* PCC 7806) (Frangeul *et al.*, 2008) por lo que resulta interesante explorar su funcionamiento en la comunidad del CMA en nuestros sistemas, así como compararlos con otros *primers* que actualmente se utilizan para su cuantificación.

Los resultados del mapeo de estos genes y la comparación con otros *primers* reportados por otros autores ilustran en la Figura 42. En el caso del gen *mcyE*, el juego de *primers* con mayor mapeo (AMPL2492956) apunta a una región diferente a la utilizada como blanco de otros *primers* reportados.

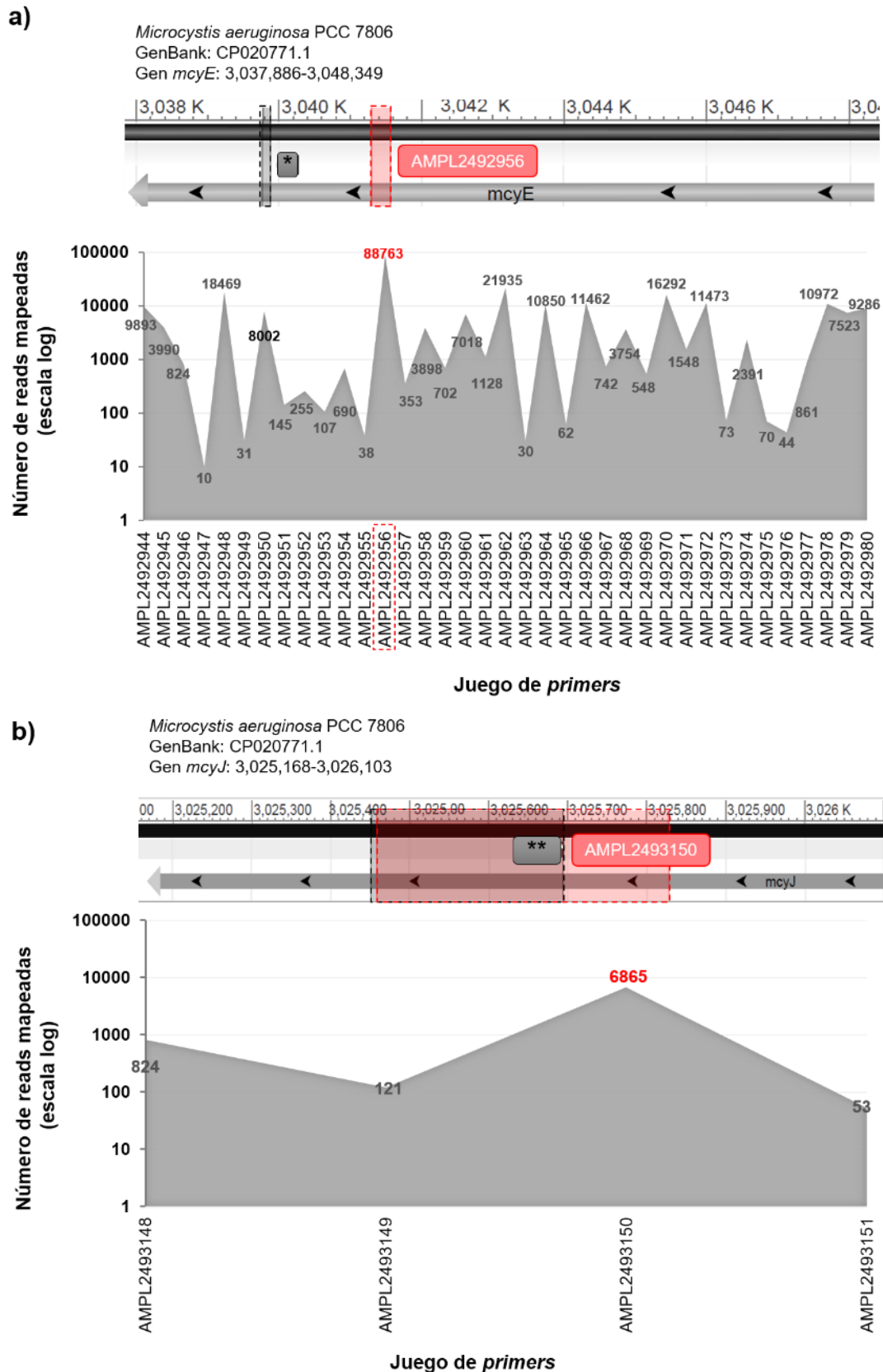


Figura 42. Cobertura de los genes *mcy* por el panel de *primers* Fvir en la muestra de Gualaguaycito. a) Gen *mcyE* y b) Gen *mcyJ*. En ambos se representa la localización de los *primers* del panel Fvir con mayor número de *reads* (rojo) y la localización de *primers* reportados por otros autores (negro). *Primers* reportados por: *(Sipari et al., 2010) y **(Kim et al., 2010).

4.3.4. Control de la Estructura Poblacional

Al igual que en las muestras control, la estructura poblacional se determinó mediante la secuenciación masiva del gen ribosomal 16S. De la secuenciación se obtuvo un total de 337.124 *reads*, de los cuales se filtraron 2086 secuencias que no cumplieron los criterios establecidos (total =0,6 %). La tabla de OTUs se conformó con un total de 328.750 secuencias. Los resultados de las curvas de rarefacción obtenidas mostraron que no fue posible secuenciar en su totalidad la comunidad microbiana presente (Figura 43).

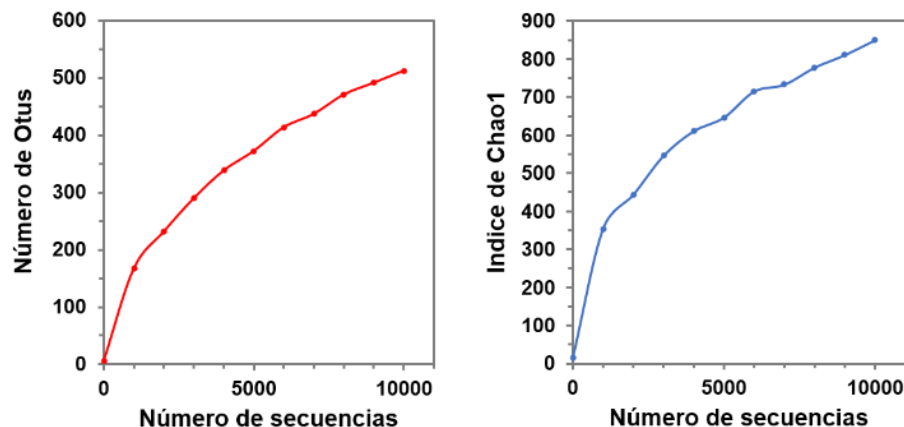


Figura 43. Estimadores de diversidad alfa. Curva de rarefacción e Índice de Chao1 calculados a un tamaño de muestra estandarizado = 10.000 secuencias.

Se obtuvo un total de 1932 OTUs donde 10 OTUs representan el 70 % de las secuencias de la muestra. En relación con la asignación taxonómica de las secuencias, el 97,54 % se asignaron al dominio Bacteria, 0,001 % a Archea y el 2,45 % no fue asignada. En el 79,5 % de las secuencias fue posible la asignación taxonómica de las OTUs a nivel de género y sólo el 0,5 % fue asignado a nivel de especie. La muestra presentó mayor abundancia del Filo *Cyanobacteria* seguido por *Proteobacteria* y *Bacteroidetes*. En relación con las cianobacterias, dos OTUs fueron mayoritarias (91 % del Filo) y fueron identificadas como *Dolichospermum uruguayense* y *Microcystis* spp. (Figura 44).

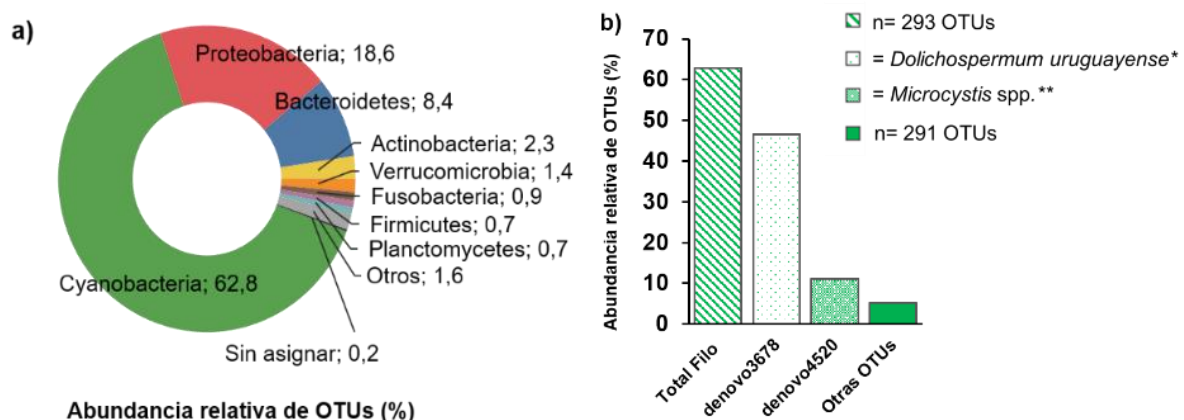


Figura 44. Estructura poblacional en la muestra de Gualaguaycito. a) Abundancia relativa de OTUs asignadas a nivel de Filo. b) OTUs más abundantes del Filo Cyanobacteria. *Determinado por BLAST como el hit más alto. ** No diferenciable a nivel de especie.

El control de la estructura poblacional tuvo como objetivo corroborar la presencia de posibles taxones portadores de los genes de virulencia detectados por el panel Fvir, Figura 45. En 15 de 16 genes de virulencia fue posible constatar a nivel de género el posible microorganismo asociado. En el caso del gen que codifica para la Nodularina (*ndaI*) no fue constatada la presencia del género *Nodularia*, pero sí de la familia *Aphanizomenonaceae*. No existieron diferencias significativas entre la abundancia relativa de OTUs asignadas a género y los factores de virulencia detectados en esos géneros. La abundancia relativa de OTUs se correlacionó positivamente con el mapeo de *reads* en los factores de virulencia detectados ($r^2= 0,945$).

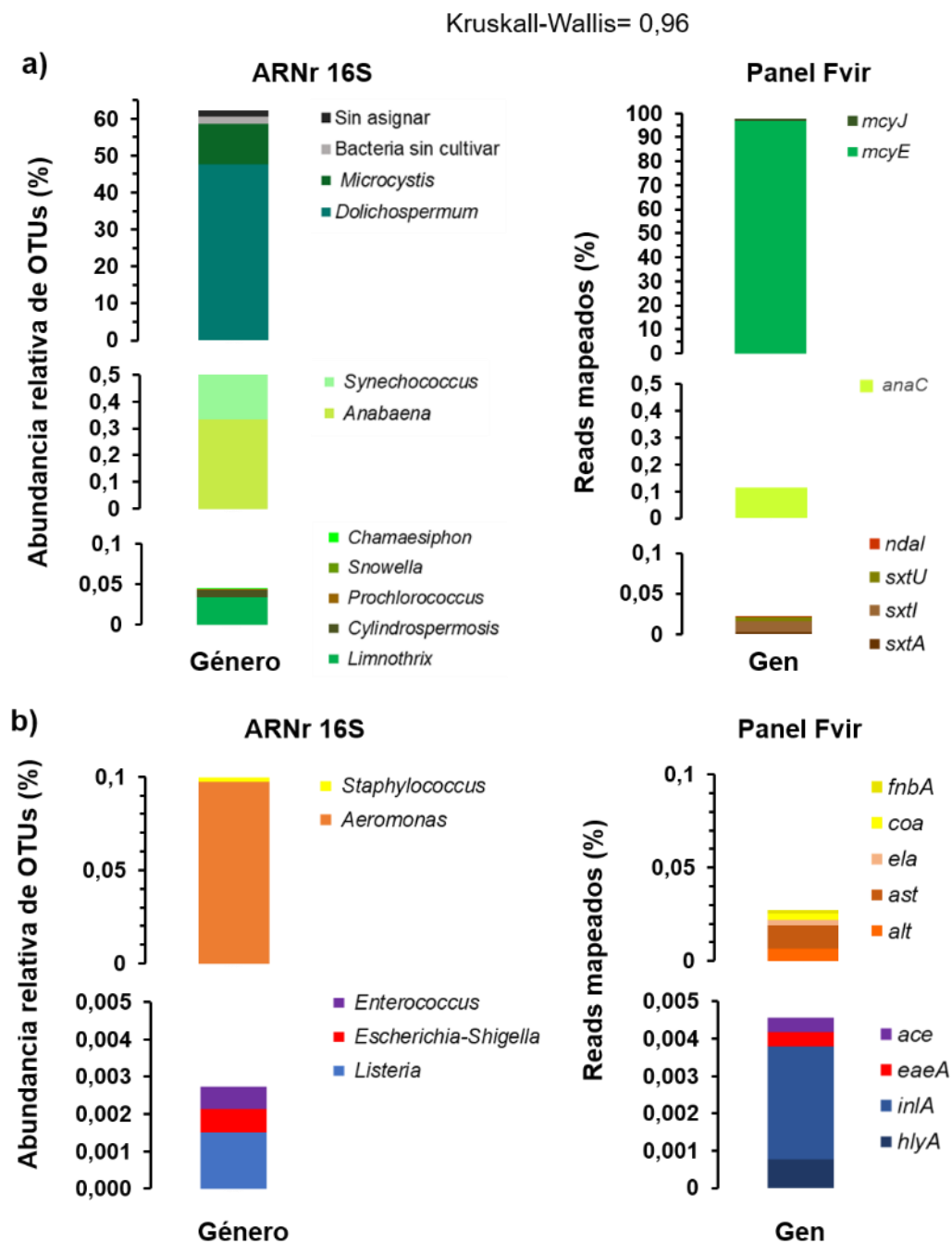


Figura 45. Comparación entre la estructura de la comunidad ARNr16S (% OTUs asignadas a nivel de género) y los posibles genes de virulencia asociados (Panel Fvir) en la muestra de Gualeguaycito. a) Cianobacterias. b) Otros géneros bacterianos.

Por otra parte, en la Figura 46 se ilustra la abundancia relativa de OTUs asignadas a otros posibles géneros patógenos mediante la aproximación ribosomal, sin embargo, no se detectaron factores de virulencia por el panel Fvir en esos géneros.

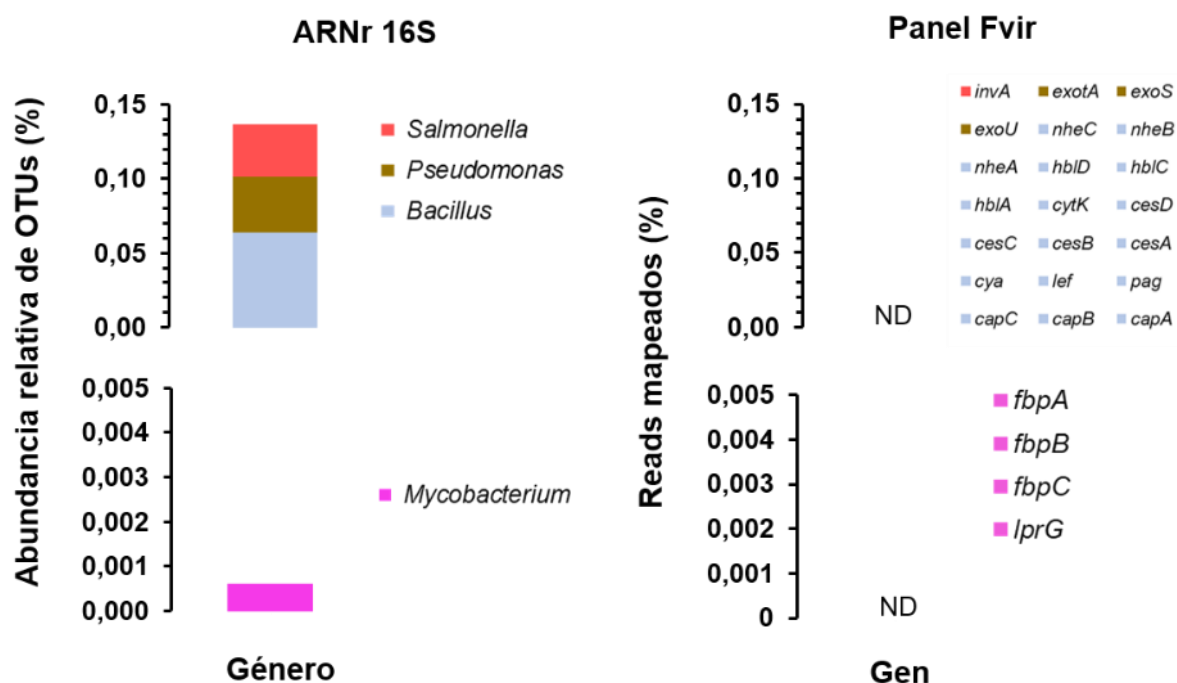


Figura 46. Abundancia relativa (% OTUs asignadas a nivel de género) de géneros bacterianos en los que no se detectaron genes de virulencia asociados (Panel Fvir) en la muestra de Gualaguaycito.

4.4. Funcionamiento General Del Panel Fvir en Las Muestras Analizadas

En la Tabla 9, se presenta un resumen general del funcionamiento de los *primers* del panel Fvir considerando todas las muestras analizadas (n=7). Durante este estudio se detectaron 30/71 genes candidatos contemplados por el panel (42 %). Mas allá de que algunos *primers* no mostraron amplificación en algunas de las muestras, fue posible la detección de todos los genes de virulencia en los casos que se contaba con dicha información. Tomando en cuenta sólo los genes amplificados, esto implica que se evaluaron 236/548 juegos de *primers* del panel (43 %). En promedio, el 85 % de estos *primers* mostró amplificación en al menos una muestra. Algunos de estos genes (n=18) fueron evaluados en 3 o más oportunidades (entre muestras controles y ambiental) mostrando un alto *porcentaje* de *primers* con amplificación (promedio > 70 %).

Tabla 9. Resumen de la amplificación de los *primers* del panel Fvir en las muestras analizadas. Se indica el número de *primers* diseñados y con amplificación, así como el tipo y número de muestras en las que se evaluó y el porcentaje promedio de *primers* con amplificación.

Microorganismo blanco	Factor de Virulencia	Gen blanco	Primers Evaluados		Performance Global		Tipo de muestra
			Nº de <i>primers</i> diseñados	Primers con amplificación (%) [*]	Nº de muestras evaluadas	$\bar{x}$ de <i>primers</i> con amplificación (%) ^{**}	
<i>Aeromonas spp.</i>	Aerolisina	<i>aerA</i>	6	100	2	100	C
	Enterotoxina TL	<i>elt</i>	10	100	3	77	C+A
	Enterotoxina TE	<i>ast</i>	9	100	3	81	C+A
	Elastasa	<i>ela</i>	7	100	3	71	C+A
<i>Cyanobacterias</i>	Anatoxina	<i>anaC</i>	29	10	1	10	A
	Microcistina	<i>mcxE</i>	37	100	1	100	A
		<i>mcyJ</i>	4	100	1	100	A
	Nodularina	<i>ndaI</i>	7	29	1	29	A
	Saxitoxina	<i>sxtA</i>	13	62	3	36	C+A
		<i>sxtI</i>	6	67	3	39	C+A
		<i>sxtU</i>	3	33	3	11	C+A
<i>Enterococcus spp.</i>	Adhesina	<i>ace</i>	8	100	3	63	C+A
	Adhesina	<i>acm</i>	9	100	2	56	C
	Hialuronidasa	<i>hyl</i>	6	100	2	58	C
	Pili	<i>pilB</i>	7	57	2	36	C
<i>E. coli- Shigella spp.</i>	Intimina	<i>eaeA</i>	10	60	7	57	C+A
	Fimbria 17	<i>f17A</i>	3	100	2	83	C
	Fimbria41	<i>f41A</i>	3	33	3	33	C+A
	Toxina shiga tipo I	<i>stx1A</i>	3	100	6	100	C
		<i>stx1B</i>	1	100	6	100	C
	Toxina shiga tipo II	<i>stx2A</i>	4	100	2	100	C
		<i>stx2B</i>	1	100	2	100	C
<i>Listeria monocytogenes</i>	Listeriolisina O	<i>hlyA</i>	6	100	7	90	C+A
	Internalina	<i>inlA</i>	9	100	7	95	C+A
<i>Pseudomonas aeruginosa</i>	Exoenzima S	<i>exoS</i>	2	100	6	100	C
	Exotoxina A	<i>exotA</i>	1	100	6	100	C
<i>Salmonella spp.</i>	Invasina	<i>invA</i>	8	100	6	100	C
<i>Staphylococcus aureus</i>	Coagulasa	<i>coa</i>	8	100	3	75	C+A
	Unión a la Fibronectina	<i>fnbA</i>	13	100	3	74	C+A
	Toxina Shock Séptico I	<i>tst</i>	3	100	2	100	C

Tipo de muestra en las que amplificó el *primer*: C (controles), A (ambiental), C+A (controles y ambiental). *Número de juegos de *primers* que han amplificado por lo menos en una de las muestras.

**Promedio del número de *primers* que amplifican en el total de las muestras evaluadas.

5. DISCUSIÓN

Asegurar la calidad microbiológica del agua exige el uso de herramientas diagnósticas sensibles y específicas que garanticen la identificación de patógenos y sus capacidades patogénicas o tóxicas. Hasta la fecha no existe un método de detección que integre a todas las bacterias patógenas de interés en una muestra de agua (Choudhury, 2020) y la detección se basa principalmente en cultivos, contando con claras limitaciones (Ramírez-Castillo *et al.*, 2015; Deshmukh *et al.*, 2016). Los avances en la detección de patógenos, particularmente mediante técnicas genómicas como la NGS, que permiten un análisis masivo del ADN de amplicones o ácidos nucleicos ambientales, marcan el comienzo de una nueva era de desarrollo para la evaluación de la calidad del agua (Tan *et al.*, 2015). En este contexto, el objetivo de este estudio fue contribuir a la mejora en el diagnóstico de bacterias patógenas asociadas a sistemas acuáticos mediante NGS dirigida a genes implicados en la patogenicidad y/o toxicidad bacteriana.

5.1. Diseño del Panel

5.1.1. Genes Candidatos

Seleccionamos 72 genes candidatos de más de 50 especies que cubren un amplio espectro filogenético y están relacionados a diferentes mecanismos de virulencia ya estudiados (Figura 12, Tabla 6). La diversidad de los genes candidatos es vasta e incluye más de un gen por especie, varios genes de un clúster, variantes más relevantes descritas, así como elementos codificados en islas de patogenicidad, plásmidos y fagos.

Este abordaje tan extenso indica la robustez que tendría el desarrollo de esta herramienta y aporta valiosa información para la evaluación de patógenos en los sistemas acuáticos. En este sentido, la capacidad de los patógenos bacterianos para establecer una infección y causar una enfermedad está determinada directa o indirectamente por múltiples factores de virulencia que actúan individualmente o en conjunto (Wu, Wang & Jennings, 2008; Lee *et al.*, 2013).

Los factores de virulencia como elementos codificados por genes, (genes de virulencia), (Chen *et al.*, 2005) representan uno de los principales descriptores (variables descriptivas), útiles para predecir la respuesta de los microorganismos en los sistemas acuáticos (Tourlousse *et al.*, 2007). Estos descriptores analizados en conjunto con variables de respuesta, como la potencia, la severidad/gravedad y la persistencia de los patógenos, han sido utilizados como un enfoque para clasificar y ranquear a los patógenos transmitidos por el agua y establecer directrices en términos de calidad del agua potable (NDWAC, 2004; Jenkins *et al.*, 2004; Tourlousse *et al.*, 2007; Cangelosi, 2009). En este sentido, uno de los grandes desafíos, es encontrar métodos que cuantifiquen a los descriptores genéticos de virulencia y que sean capaces de identificar y discriminar sus secuencias específicas con precisión (Tourlousse *et al.*, 2007) situación a la que el desarrollo de nuestra herramienta podría contribuir.

A su vez, la múltiple selección de genes de virulencia ha sido señalada como una estrategia más adecuada para estimar o predecir la virulencia de una cepa dado que

la detección de un gen en particular por sí sola no aporta información suficiente (Tourlousse *et al.*, 2007). Otro aspecto relevante considerado en este abordaje, son los elementos genéticos móviles asociados con la virulencia y presentes en varias de las especies/cepas blanco (por ejemplo, las toxinas de *Shigella spp.*). Estos elementos genéticos se comparten con frecuencia entre los patógenos transmitidos por el agua que tienen un nicho ecológico similar (por ejemplo, *Salmonella*, *Shigella*, *Vibrio*, *Escherichia*) y son objeto de preocupación constante para la salud pública (National Research Council, 2001).

El amplio rango de genes seleccionados planteó un desafío adicional, la búsqueda de regiones conservadas blanco para el diseño de *primers*.

5.1.2. Evaluación de la conservación en los genes de virulencia

Establecimos regiones de conservación en 39 de los 72 genes candidatos (Tabla S2, Anexo I). Estos genes se caracterizan por presentar una variabilidad genética menor (en general se distribuyen en un menor rango de géneros o especies distintas) (Figuras 13 y 14). En contraste, los 33 genes restantes, “genes no conservados” presentaron un grado de variabilidad genética alta (distribución en un rango más amplio de taxones) (Figuras 13 y 15). Es relevante mencionar que este estudio fue realizado en el periodo 2016-2017 y representa las secuencias disponibles hasta ese momento. En este sentido, los valores de la variabilidad genética deben analizarse con precaución ya que pueden subestimarse cuando el número de secuencias disponibles es limitado (Tourlousse *et al.*, 2007).

Entre los genes no conservados se encuentran, los que codifican toxinas de cianobacterias particularmente, Saxitoxina (*sxtI*), Microcistina (*mcyE*) y Anatoxina (*anaF*) (Figura 15). Esta variabilidad ya ha sido descrita en estudios previos, donde se indican como los más variables de sus clústeres respectivos (Casero *et al.*, 2014; Ballot *et al.*, 2016; Ballot, Scherer & Wood, 2018; Martínez de la Escalera *et al.*, 2019). En este sentido, estos trabajos proponen que la variabilidad del gen *sxtI* o *anaF* permitirían su utilización como marcador genético para estudios filogenéticos (Casero *et al.*, 2014; Ballot *et al.*, 2016; Ballot, Scherer & Wood, 2018). De manera similar nuestro equipo de trabajo ha señalado que la variabilidad en el gen *mcyJ* podría relacionarse a preferencias geográficas o nichos ecológicos particulares (Martínez de la Escalera *et al.*, 2019). Otros genes que mostraron una variabilidad genética alta, son los que codifican para el antígeno 85 del complejo *Mycobacterium avium* (*fbpA*, *fbpB*, *fbpC*) y el gen de la capsula de *Bacillus anthracis* (*nheC*) cuya variabilidad también ha sido descrita (Content *et al.*, 1991; Harrell, Andersen & Wilson, 1995). La heterogeneidad en la secuencia también se ha descrito en elementos genéticos móviles como por ejemplo, los fagos que codifican shiga-toxina (*stx2*) de *E. coli*, donde se sugiere que las variantes de secuencia podrían asociarse a diferentes rangos de hospederos o diferencias en la producción de toxina (Wagner, Acheson & Waldor, 1999; Gamage *et al.*, 2004).

En relación con la distribución de los genes entre géneros/especies distintas, ésta es particular en cada gen estudiado y en asociación con la variabilidad genética, pueden brindar información relevante sobre determinado patotipo, rasgo patogénico particular, o incluso la evolución del gen o del clúster de genes en estudio (Tourlousse *et al.*, 2007; D’Agostino *et al.*, 2016). La distribución de estos genes puede ser amplia

entre taxones muy distintos, como lo observado en los genes que codifican microcistina en cianobacterias (*Microcystis*, *Anabaena*, *Nostoc*, *Fischerella*, *Hapalosiphon*, *Oscillatoria/Planktothrix* y *Phormidium*) (Heck *et al.*, 2018). En cambio, en otros se distribuyen entre cepas de una única especie (por ejemplo, algunos genes de virulencia de *E. coli*) e incluso su presencia o ausencia se traduce en diferentes patotipos específicos (Bekal *et al.*, 2003). Esta variabilidad en la distribución de genes de virulencia ha sido descrita también en aislamientos de *Staphylococcus aureus* (Becker *et al.*, 2004); *Aeromonas spp.* (Chacón *et al.*, 2004) y *Pseudomonas aeruginosa* (Feltman *et al.*, 2001) entre otros. Si bien su entendimiento se ha ampliado, el origen de esta variabilidad es aún incierto para muchas de estas especies patógenas (Tourlousse *et al.*, 2007).

El estudio realizado sobre la conservación en los genes de virulencia (indispensable para el diseño de *primers*), también posibilitó el abordaje de enfoques complementarios. Como se mencionó anteriormente, las regiones variables presentes en algunos de los genes podrían distinguir diferencias de secuencias entre distintas especies o relacionarse a determinados nichos ecológicos, incluso estos genes podrían ser más específicos que el gen del ARNr 16S para diferenciarlas (Miller *et al.*, 2008). En este contexto y utilizando nuestra base de datos, comparamos la variabilidad de algunos genes de virulencia con la del ARNr 16S en su mismo genoma (Anexo II). Observamos que, en su conjunto, los genes de virulencia tienen una mayor variabilidad genética en comparación al ARNr16S y proponemos que la variabilidad en algunos genes o combinaciones de genes (por ejemplo, en genes de cianobacterias o de *E. coli*) reflejarían distintas preferencias ambientales y definirían ecotipos (poblaciones que presentan cohesión genética, pero difieren a nivel ecológico) (González-Revello *et al.*, 2017).

En suma, los resultados del estudio de conservación de los genes blanco definieron la estrategia para el diseño de *primers*.

Los genes que presentaron zonas con alto grado de conservación (representativas de los bancos de secuencias disponibles) representaron la aproximación ideal para localizar los *primers* que se diseñaron en el panel. A su vez, en los genes no conservados, como estrategia alternativa, se utilizó el genoma de una cepa de referencia (secuencia única de cada gen), por lo que la capacidad de los *primers* para cubrir variantes específicas podría verse limitada. Nuestro diseño de múltiples *primers* que apuntan a la cobertura de todo el gen candidato, podrían establecer diferencias de amplificación entre *primers* para su selección en base a los mejores resultados.

5.1.3. Panel FVIR, Características Generales y Evaluación in Silico

Diseñamos una primera versión de un panel Ampliseq™, el panel Fvir (V 1.0). Este consta de 548 juegos de *primers*, que cubren 71 genes de virulencia bacterianos (Figuras 18, 19 y 20) compatibles en dos reacciones de PCR-multiplex (284 y 264 juegos respectivamente) y asociados al sistema de secuenciación Ion Torrent (PGM). Mediante el software *Ampliseq designer*™, logramos una alta cobertura sobre la secuencia de los genes de virulencia candidatos (97,3 %) y sólo en un gen no fue posible el diseño de *primers* (gen *est1*, enterotoxina termoestable de *E. coli*) quedando por fuera del alcance del panel (Figura 17).

El número de *primers* de nuestro panel es similar a los reportados por Brinkmann *et al.*, (2017) para la detección de 46 especies virales (541 amplicones) y por Gardner *et al.*, (2015) para la detección de 6 especies bacterianas consideradas como amenazas biológicas (467 amplicones). En estos trabajos se destaca el alto grado de especialización en el diseño de *primers* alcanzado por la tecnología de AmpliSeq™. Éste es el producto insignia de la plataforma de secuenciación de Ion Torrent y ha logrado una notable capacidad de PCR-multiplex, considerado como difícilmente igualable por otros sistemas (Khodakov, Wang & Zhang, 2016). A modo de ejemplo, su panel *Ampliseq Comprehensive Cancer* utiliza 16.000 pares de *primers* compatibles en 4 tubos de reacción, lo que implica un promedio de 4.000 juegos por tubo de PCR (Li *et al.*, 2013). En este sentido, nuestro panel puede considerarse relativamente pequeño y lejos de alcanzar los límites ofrecidos por esta tecnología. A su vez, da cuenta del grado de escalabilidad al que se puede acceder, por ejemplo, mediante el agregado de nuevos *primers* para la detección de otros genes, sus variantes o según sea necesario con el descubrimiento de nuevos patógenos de interés (Anis *et al.*, 2018). Otras ventajas de este sistema, que también han sido reportadas (Khodakov, Wang & Zhang, 2016) son: los requisitos de ADN input bajos (1-10 ng), de especial relevancia en muestras de agua, el protocolo de trabajo rápido (3-6 h) y la combinación con otras plataformas de secuenciación como MiSeq Illumina o MinION (Oxford Nanopore Technologies) (Quick *et al.*, 2017).

La evaluación *in silico* del panel reflejó lo eficiente del sistema *Ampliseq™ designer* en el diseño de los *primers* (Figuras 21 y 22). En este sentido no existieron diferencias en las principales características de los *primers* que conforman los dos pools multiplex, indicando un correcto balance entre ellos. A su vez el 95 % de los *primers* del panel tienen Tm mayores a 60° C, así como una baja formación de dímeros de/entre *primers* y horquillas a esta temperatura. Estas son condiciones satisfactorias para la multiplex-PCR inicial cuyo *annealing* y extensión se realizan a esa temperatura.

El panel Fvir se diseñó en un rango de tamaños de amplicones compatibles al sistema de secuenciación (125-375 pb) con un alto poder discriminatorio de los genes candidatos, que sumado a la redundancia del diseño (múltiples amplicones / gen), se traduce en altos valores de sensibilidad (100 %) y especificidad (99,5 %) evaluados *in silico*. En relación con el mapeo en zonas conservadas, se diseñaron 170 juegos de *primers*, blanco de 32 de los 39 genes candidatos y en los restantes 7 no fue posible su diseño (Figura 24).

Resultados similares son reportados por Brinkmann *et al.*, (2017) en su panel para detección viral señalando los altos niveles alcanzados (sensibilidad= 97,9 % y especificidad= 100 %). Sin embargo, estos autores describen que la sensibilidad de su panel se vio afectada por nuevas secuencias virales no incluidas en el diseño original (agregadas a la base de datos posteriormente). A su vez, indican que el panel debe actualizarse a las nuevas secuencias disponibles y que la divergencia en las secuencias nuevas puede solucionarse simplemente con la adición de algunos *primers*. De igual manera, planteamos que nuestro panel deberá actualizarse según los avances en las bases de datos y proponemos que la adición de nuevos *primers* podría ser una opción para las zonas conservadas de aquellos genes que no se cubrieron.

5.2. Evaluación Del panel en Controles Con Potencial Patogénico

Evaluamos la capacidad de detección del panel Fvir empleando una serie de controles y ADN de cepas portadoras de algunos de los genes candidatos.

La aproximación utilizando ADN/ARN transcrito de cepas de colección bacterianas, virales, fúngicas y parasitarias ha sido utilizada como una estrategia inicial para evaluar los nuevos paneles diseñados (Gardner *et al.*, 2015; Brinkmann *et al.*, 2017; Anis *et al.*, 2018). Sin embargo, esta aproximación no es adecuada para determinar la sensibilidad analítica absoluta de los paneles (Anis *et al.*, 2018) recomendándose que los límites de detección (LOD) se evalúen con secuencias blanco clonadas en plásmidos o a partir de ARN transcrito *in vitro* (Jennings *et al.*, 2017; Anis *et al.*, 2018). A su vez, existe falta de disponibilidad de cepas de referencia para muchos de los patógenos que contempla nuestro panel o las mismas tienen un costo elevado. Teniendo en cuenta estas limitaciones nuestros ensayos se orientaron a la evaluación preliminar del panel, explorando la sensibilidad relativa en el número de cepas a las que se logró acceder.

5.2.1. Control Con Potencial Patogénico (Cepas Únicas)

En el ensayo control con un número reducido de cepas, el panel logró detectar los 8 genes candidatos referidos a las cepas utilizadas (*E. coli*, *L. monocytogenes*, *P. aeruginosa* y *S. enterica* Typhimurium) (Figura 26). Los genes se detectaron en un rango estimado entre 30-30.000 copias de genomas/ μ L, amplificando el 87,5 % de los *primers* diseñados y alcanzando niveles altos de sensibilidad (95,2 %) y especificidad (99,5 %) (Figuras 27 y 31).

Resultados similares en relación con la sensibilidad, especificidad y la capacidad de detección a bajas concentraciones son reportados en los trabajos anteriormente mencionados (Gardner *et al.*, 2015; Brinkmann *et al.*, 2017; Anis *et al.*, 2018). Nuestros resultados muestran la amplificación de los genes candidatos aún en la concentración más baja utilizada (30 genomas/uL). Sin embargo, esto se logró mediante el aumento de los ciclos de PCR inicial dentro de los límites recomendados para el tamaño de nuestro panel (14-23 ciclos) (Ion AmpliSeq, 2017). En este sentido, el empleo de amplificación dirigida proporciona un gran aumento en las lecturas de las regiones discriminatorias, lo que permite realizar una identificación precisa de las secuencias blanco a niveles muy bajos (Brinkmann *et al.*, 2017). Específicamente esto es demostrado en el trabajo de Gardner *et al.*, (2015) en una muestra con tan sólo 10 copias de genoma/ml (0,001 copias/uL) y manteniendo en 16 los ciclos de PCR iniciales. A su vez, la muestra contenía múltiples patógenos y estaba enriquecida en genoma humano, mostrando la capacidad de las NGS dirigida para detectar patógenos de interés en una matriz compleja (Anis *et al.*, 2018).

Por otra parte, si bien el panel detectó los 8 genes candidatos, el número de *reads* mapeados fue diferente entre los genes y especies estudiadas (Figura 28). Se observó un mayor mapeo en el gen de la invasina de *Salmonella enterica* (*invA*), seguido por los genes de *Listeria monocytogenes* (*hylA*, *inlA*) y *Escherichia coli* (*eaeA*, *sxt1A*, *sxt1B*). A su vez, el nivel de detección fue bajo en los genes de *Pseudomonas*

aeruginosa (*exoS*, *exotA*). Esto puede deberse a diferencias en el número de copias del gen que porta cada cepa, a un desbalance en las copias iniciales de los genomas o al número y/o eficiencias de cada *primer* en particular entre otros factores.

En referencia a que las cepas porten más de una copia del gen candidato, en oposición a lo observado en otros genes como los del ARNr (suelen encontrarse en múltiples copias) las copias de los genes de virulencia suelen ser únicas en el cromosoma bacteriano (Klappenbach, Dunbar & Schmidt, 2000; Søborg *et al.*, 2016). Esto se ha reportado en los genes candidatos evaluados de *Salmonella* (Fey *et al.*, 2004), *Listeria* (Dubail *et al.*, 2000), *Pseudomonas* (Kung, Ozer & Hauser, 2010) y para el gen de la intimina (*eaeA*) de *E. coli* (Barbau-Piednoir *et al.*, 2018). Sin embargo, se ha descrito que los genes de la toxina shiga (*stx1* y *stx2*) son genes profágicos y que un sólo huésped bacteriano puede albergar más de un profago (Fogg *et al.*, 2012). Dado que en nuestro ensayo estos genes no presentaron valores de mapeo mayores a los genes de copia única consideramos que este no sería el factor determinante.

En relación con las copias iniciales de genomas, en nuestro diseño estimamos su número a través de la concentración de ADN y el tamaño específico de un genoma de referencia. Esta metodología también fue empleada por Gardner *et al.*, (2015), sin embargo, ha sido señalado que esta aproximación puede conducir a errores (sobre todo relacionados con la calidad/pureza del ADN extraído) (Jennings *et al.*, 2017). En este sentido, para una cuantificación más precisa, podrían utilizarse técnicas como la qPCR de un gen específico, por lo menos en alguno de los genes candidatos (Brinkmann *et al.*, 2017; Anis *et al.*, 2018). Consideramos que una incorrecta estimación del número de copias iniciales, debido por ejemplo a una contaminación de ADN durante su extracción puede ser factible, ya que en muchos casos el ADN fue obtenido como tal de colecciones externas. En este sentido, podría haber generado un desbalance entre los genomas de los cuatro microorganismos, explicando en parte estas diferencias, así como el bajo mapeo en los genes de *Pseudomonas aeruginosa*. Este tema será abordado más adelante en la discusión (control de la estructura poblacional).

Con respecto a las diferencias de mapeo que se observaron entre *primers* de un mismo gen, estas podrían deberse en parte al diseño de los *primers* sobre zonas conservadas o no conservadas. En este sentido, fue posible explorarlas en los genes de la intimina (*eaeA*) de *E. coli*, y de la internalina A (*inlA*) de *L. monocytogenes* ya que se diseñaron *primers* para ambas zonas (Figura 29).

El gen de la intimina de *E. coli* presentó diferencias según su localización en una u otra región. Los 2 juegos de *primers* localizados sobre la zona conservada del gen mostraron un mapeo similar y valores altos de amplificación, mientras que en la región no conservada la amplificación fue variable y 6 de los 8 *primers* diseñados no amplificaron. La variabilidad en la secuencia de este gen es conocida y ha sido estudiada detalladamente (Yang *et al.*, 2020). Así, se reporta que el extremo N-terminal de la proteína está altamente conservado en cepas de diferentes fuentes, mientras que el extremo C-terminal, donde ocurre la actividad de unión celular, es altamente variable (Hernandes *et al.*, 2009). En base a esta variabilidad se han definido al menos 30 subtipos de intimina (Ooka *et al.*, 2012), los que han sido correlacionados a la especificidad de huésped y el tropismo tisular (Zhang *et al.*,

2002). En nuestro estudio observamos que los *primers* que no amplificaron, corresponden en su mayoría (5/6) a esta región de alta variabilidad. En este sentido, la falta de amplificación de esta región podría indicar que la cepa control utilizada en el ensayo (*E. coli* PL74.1, aislamiento de campo) presentaría diferencias de secuencia con la cepa de referencia utilizada en el diseño de los *primers* (*E. coli* O157H7 cepa Sakaï-subtipo *eae-γ1*). Si bien los *primers* dirigidos a la zona conservada amplificaron y detectaron el gen, resultaría interesante explorar la funcionalidad de los *primers* de la región variable en otras cepas.

En contraste, el gen de la internalina A de *L. monocytogenes* no mostró diferencias significativas de mapeo entre las regiones, amplificando el total de *primers* diseñados. Esto indicaría que el genoma de referencia utilizado para el diseño es similar al de la cepa control utilizada, incluso en las regiones no conservadas.

Los restantes genes evaluados (sin zonas conservadas) presentaron amplificación de todos los *primers* diseñados y el mayor mapeo observado de algunos indicaría mejor eficiencia para la cepa evaluada (por ejemplo, el juego de *primers* 538 del gen *hylA* de *L. monocytogenes* (Figura 29).

5.2.2. Control Con Potencial Patogénico (Cepas Múltiples)

Evaluamos la funcionalidad del panel Fvir en un control más complejo, utilizando 92 aislamientos pertenecientes a los géneros *Escherichia*, *Salmonella*, *Listeria*, *Pseudomonas*, *Enterococcus*, *Staphylococcus*, *Aeromonas* y *Cylindrospermopsis*. Los aislamientos utilizados eran portadores de algunos genes de virulencia cubiertos por el panel Fvir (Tabla S1, Anexo I) mientras que otros genes se desconocía su presencia.

En relación con los resultados generales de este ensayo, podemos observar que la capacidad de detección lograda fue alta. El panel Fvir detectó 25 genes asociados a los géneros que conformaron el pool (Figura 32), 16 de ellos caracterizados anteriormente (Tabla S1, Anexo I) y reportados en alguna de las cepas (De Los Santos *et al.*, 2014; Umpiérrez *et al.*, 2017; Perretta, Antúñez & Zunino, 2018). Este resultado da cuenta de la sensibilidad de los *primers* diseñados, ya que la presencia de estos genes previamente reportados puede considerarse como un control positivo. A su vez, no fueron detectados otros genes asociados a otros grupos bacterianos diferentes a los que conformaron el pool, lo que indicaría niveles adecuados de especificidad. Adicionalmente se detectaron 9 genes, cuya presencia en las cepas era desconocida pero factible dado la especie u origen: *fnba*, *tst* (*S. aureus*), *f41A*, *stx2A*, *stx2B* (*E. coli*), *ace* (*E. faecalis*) y *acm*, *hyl* y *pilB* (*E. faecium*). Esto proporcionó información valiosa para los investigadores colaboradores, como por ejemplo en las cepas de *E. coli* la detección de genes de fimbrias particulares (F41) o variantes de toxina tipo Shiga (*stx2*). Aunque esta muestra no refleja la complejidad de una verdadera matriz ambiental, mostró la capacidad del panel para detectar los genes de virulencia de interés en un pool bastante completo. Presumiendo que no todos los genes de virulencia estén presentes en cada uno de los aislamientos el panel proporcionó suficiente cobertura a pesar de la posible interferencia del ADN de las cepas que no los portan.

Por otra parte, exploramos si el enriquecimiento de la muestra en un microorganismo afectaba la detección de otros menos representados. Para ello, utilizamos dos tratamientos. En el tratamiento C+ se mantuvo la relación de copias de genoma iniciales entre los géneros (10^9 copias/ μ L). En el tratamiento C50 aumentamos las copias de genoma de *Listeria monocytogenes* un orden logarítmico (10^{10} copias/ μ L) en relación con los otros géneros (10^9 copias/ μ L). En ambos tratamientos detectamos los mismos genes (Figura 34), lo que implica que el panel fue capaz de detectar genes en muy baja representación, ya que en la muestra C50 el ADN de *Listeria* representaba más del 50 % del ADN total. En este sentido Brinkmann *et al.*, (2017), señalan la dificultad que enfrentan los métodos de detección en muestras complejas simuladas y como el empleo de amplificación dirigida permite realizar una identificación segura de las especies aún a niveles muy bajos.

El porcentaje de *primers* que amplificaron fue alto en ambos tratamientos (67 % de los *primers* diseñados) (Figura 35). Sin embargo, el tratamiento con igual copias de genomas iniciales (C+) presentó mejores niveles de amplificación (80 %), lo que sugiere que el enriquecimiento con *Listeria* del tratamiento C50 afectó el funcionamiento de los *primers* (69 %) en algunos de los genes menos representados (Figura 36).

5.2.3. Control de la estructura poblacional

En el tratamiento C50 (Mayor número de copias de genoma/ μ L de *Listeria*) realizamos el control de la estructura poblacional mediante secuenciación masiva del ARNr 16s.

La estructura observada concuerda con lo esperado para la muestra, de acuerdo con el diseño experimental, donde se observan los 8 géneros bacterianos utilizados en el pool de ADN (Figura 38). El enriquecimiento en *Listeria* también fue constatado, con un 75,5 % de OTUs asignadas a nivel de género. El porcentaje de OTUs asignadas se correlacionó positivamente con el número de copias iniciales de genomas calculado, lo que valida en parte el diseño experimental (Figura 39).

Valores de OTUs asignadas, inferiores a los esperados (entorno al 6 %) se observaron en los géneros *Escherichia* (2,2 %), *Salmonella* (1,9 %), *Enterococcus* (1,1 %) y *Cylindrospermopsis* (0,54 %).

Esto indicaría que el método de estimación no fue muy preciso y apoya lo planteado con anterioridad en cuanto a las causas que llevan a una subestimación (calidad/pureza del ADN o contaminación). En este sentido, la presencia de ADN de otros grupos fue constatada y el 1 % de las OTUs fueron asignadas a otros 95 géneros diferentes a los ensayados. El origen podría explicarlo en parte, la utilización de *Cylindrospermopsis raciborskii* en los controles, ya que si bien el cultivo de esta cianobacteria es monoespecífico, no es axénico, y no se puede descartar la presencia de bacterias heterótrofas durante la extracción del ADN (Vico *et al.*, 2020).

Por otra parte, el porcentaje de OTUs asignadas se correlacionó positivamente con los *reads* mapeados en factores de virulencia del panel ($r^2=0,98$) (Figura 39).

Estos resultados dan sustento al funcionamiento del panel, ya que la estructura poblacional se vió reflejada en la mayoría de los genes de virulencia amplificados. En

este sentido, permite dislucidar el grado de sensibilidad del panel, capaz de detectar los genes candidatos en los grupos bacterianos presentes en baja proporción. Sin embargo, observamos una menor amplificación en genes de virulencia de *Pseudomonas* (0,02 %) en relación al control poblacional (6,8 %), lo que indicaría un menor grado de funcionalidad de estos *primers*. Resulta importante mencionar que durante los controles se utilizó una única cepa (ATCC® 27853™) siendo conveniente su evaluación en un espectro más amplio, de manera de garantizar su funcionamiento.

5.3. Evaluación Del Panel en una Muestra Ambiental

5.3.1. Genes de Virulencia Detectados

En esta etapa del trabajo, se evaluó la utilidad diagnóstica del panel en una muestra del Río Uruguay durante un evento de floración del CMA. La muestra, previamente caracterizada (Deus-Álvarez, 2018) era abundante en genotipos tóxicos productores de microcistinas (gen *mcy*) (Tabla 8).

Los resultados concuerdan con lo reportado por Deus, y mostraron más del 99,8 % del total de *reads* con mapeo en los genes relacionados a la producción de microcistinas (*mcyE* y *mcyJ*) (Figura 40). Todos los *primers* del panel diseñados para estos genes presentaron amplificación (*mcyE*= 37 juegos, *mcyJ*= 4 juegos) (Figura 41). Sin embargo, un mayor mapeo fue observado en el gen *mcyE* (96,8 %) en comparación al gen *mcyJ* (3 %) independientemente del número de *primers* para cada gen. Este resultado coincide con lo reportado anteriormente por nuestro grupo de trabajo, quienes evaluaron mediante *primers* basados en la literatura la cuantificación por qPCR de los genes *mcyE*, J, D y B, indicando al gen *mcyE* con los mejores resultados para utilizar en la comunidad del CMA encontrada en Uruguay (Martínez de la Escalera *et al.*, 2017; Deus-Álvarez, 2018).

En relación con la amplificación de estos *primers*, en ambos genes (*mcyE* y *mcyJ*), se observó un mapeo mayor (10 veces más de *reads*) de un juego de *primers* en comparación con el resto (Figura 42). En el caso del gen *mcyJ* estos *primers* apuntan a la misma región reportada por los de Kim *et al.*, (2010) y que actualmente son utilizados por nuestro grupo (Martínez de la Escalera *et al.*, 2017). En el gen *mcyE* nuestros *primers* apuntan a una región diferente a los de Sipari *et al.*, (2010) lo que podría indicar que serían más adecuados para evaluar al CMA en nuestros sistemas acuáticos. Sin embargo, son necesarios más estudios que avalen estos resultados, así como otros que evalúen los *primers* actualmente utilizados, ya que pueden estar sesgando la cuantificación de células tóxicas del CMA.

Por otra parte, otros 16 genes de virulencia fueron detectados en la muestra, aunque con baja profundidad (mapeo inferior al 1 % de los *reads*) (Figura 40). Entre ellos, se detectaron genes que codifican toxinas relacionadas a varios grupos/géneros de cianobacterias: anatoxina (*anaC*), saxitoxina (*sxtA*, *sxtI*, *sxtU*) y nodularina (*ndaI*). En este sentido, existen escasos reportes de estas toxinas en los sistemas acuáticos del país, limitándose a algunos trabajos referidos a saxitoxinas producidas por cepas de *C. raciborskii* proveniente de lagos/lagunas (Vidal & Kruk, 2008; Piccini *et al.*, 2011). Los demás genes detectados en la muestra se relacionan a diferentes factores de virulencia (adhesinas/toxinas/enzimas) de los géneros: *Aeromonas* spp. (*alt*, *ast*, *ela*),

Listeria monocytogenes (*hylA*, *inlA*), *Staphylococcus aureus* (*coa*, *fnbA*), *Escherichia coli* (*eaeA*) y *Enterococcus faecalis* (*ace*). La presencia de bacterias potencialmente patógenas ha sido descrita en la comunidad heterótrofa asociada a las floraciones de cianobacterias. Por ejemplo, se ha reportado asociaciones con los géneros *Aeromonas*, *Vibrio*, *Acinetobacter* y *Pseudomonas* (Ward & Cohan, 2006; Berg *et al.*, 2009). Si bien los estudios de riesgo durante las floraciones se centran en las cianotoxinas, es importante mencionar que muchos problemas de salud humana, (efectos dermatológicos/gastrointestinales) pueden ser causados por estas bacterias heterótrofas (Berg *et al.*, 2009).

5.3.2. Control de la Estructura Poblacional

En relación con la estructura poblacional de la muestra (ARNr 16S), la comunidad logró ser capturada en parte y mostró a las cianobacterias como el Filo más abundante (62,8 %). En este sentido, dos OTUs fueron mayoritarias (91 % del Filo) y se asignaron al género *Microcystis* sp. y a la especie *Dolichospermum uruguayense* (determinado como el hit más alto por Blast) (Figura 44).

Estos resultados concuerdan con lo esperado para la muestra ya que proviene de una floración del CMA (Deus-Álvarez, 2018). A su vez, la especie de *Dolichospermum uruguayense*, fue reportada en el Río Uruguay por Kozlikova-Zapomelova, Ferrari & Perez, (2016) quienes señalan que los eventos de floración en este río son frecuentes. Además, en su trabajo indican que la mayor abundancia observada de esta especie ocurrió durante una floración de *Microcystis aeruginosa* cercana al sitio de Fray Bentos. En este sentido, la coexistencia de estas especies también ha sido documentada en el Río de la Plata por Pérez *et al.*, (2016) señaladas como especies dominantes del fitoplancton en sus muestras de verano 2014 y 2015.

Todos los genes de virulencia detectados (panel Fvir) se pudieron relacionar a posibles taxones presentes en la muestra (ARNr16 S) la mayoría a nivel de género (Figura 45). En este sentido ambas aproximaciones mostraron abundancias similares, para todos los genes detectados, correlacionándose positivamente. La comparación entre ambas no mostró diferencias significativas.

Otros grupos bacterianos potencialmente patógenos se observaron en la comunidad de la muestra, pero no fueron detectados por el panel (Figura 46). Lo que podría sugerir que se tratara de cepas no patógenas. Sin embargo, son necesarios más estudios que avalen estos resultados, dado que los *primers* para estos grupos aún no se han probado en ningún control.

En comparación con el ARNr 16S la secuenciación dirigida a genes de virulencia presentó un mayor valor diagnóstico de la muestra, ya que fue posible determinar el potencial patogénico de algunos géneros presentes. La información sobre las propiedades de virulencia de un medio ambiente, como la abundancia, la distribución y su correlación con las propiedades ambientales, es fundamental para comprender la naturaleza y el alcance de su amenaza potencial (Lee *et al.*, 2013).

6. CONCLUSIONES

En resumen, la amplificación dirigida a genes de virulencia bacterianos seguida de NGS se presenta como una estrategia novedosa e innovadora que contribuye a los actuales sistemas de diagnóstico de bacterias patógenas en sistemas acuáticos. Las conclusiones principales que se extraen de esta tesis son:

- Se diseñó un panel Ampliseq™ dirigido a 71 genes de virulencia bacterianos, de interés para su utilización diagnóstica en sistemas acuáticos. Los *primers* del panel diseñado (548 juegos) son compatibles en 2 tubos de PCR-multiplex acoplados al sistema NGS de Ion Torrent y mostraron *in silico* niveles elevados de sensibilidad y especificidad sobre los genes candidatos.
- La evaluación del panel Fvir en una colección de cepas patógenas mostró una alta capacidad de detección de los genes candidatos presentes. Los niveles de detección se mantuvieron constantes dentro de los límites de ADN molde (mínimo 30 copias de genoma) y ciclos de PCR recomendados para el uso del panel. A su vez, el panel amplió la información conocida sobre los genes portados por estas cepas.
- La utilidad diagnóstica del panel fue comprobada en una muestra ambiental, brindando información más profunda sobre los genotipos tóxicos presentes en una floración del CMA. Asimismo, complementó la información conocida detectando otros genes de virulencia asociados a este grupo y a la comunidad heterótrofa. Los resultados obtenidos del uso del panel Fvir dan cuenta del potencial patogénico de la muestra.
- Se logró implementar la secuenciación masiva del ARNr 16S, y la información obtenida de la estructura poblacional resultó un control útil para la evaluación preliminar del panel. Los patrones observados a partir de los datos de secuenciación del ARNr 16S fueron concordantes con los factores de virulencia encontrados, y la secuencia apoyó un análisis filogenético más detallado que la metagenómica del 16S, expandiendo el uso de identificación del panel.

7. PERSPECTIVAS

Dado que este trabajo constituyó un primer acercamiento al desarrollo de un panel Ampliseq™ específico y que sólo fue posible una evaluación preliminar del panel generado, considero que:

A nivel de diseño, debe tenerse en cuenta la gran plasticidad que presenta esta tecnología, donde nuevos genes candidatos o variantes de secuencia podrán incorporarse al panel (por ejemplo, a medida que aumenten las secuencias en las bases de datos). Asimismo, podrán excluirse los *primers* con problemas de especificidad o que presenten baja/nula amplificación. Mejorando la performance general del panel.

A nivel de pruebas, es necesario ampliar el número de genes candidatos a evaluar, sobre todo de aquellos genes en los que por motivos de seguridad o disponibilidad no es factible el uso de ADN/cepas de referencia. Actualmente buscamos diseñar fragmentos sintéticos de ADN que podrán ser utilizados como control para factores de virulencia poco frecuentes (genes de toxina del ántrax y otros).

Además, el panel deberá ser desafiado con muestras ambientales más diversas de manera de evaluar la capacidad de detección de aquellos genes con mayor variabilidad en su secuencia.

La estandarización de las copias de los genes candidatos con métodos más precisos, como la qPCR, permitirá un mayor conocimiento de la sensibilidad analítica de la técnica (límites de detección y límites de cuantificación) y constituye (junto a otros estudios) los primeros pasos hacia una posible validación.

Con una perspectiva más amplia consideramos que esta tecnología es factible de expandirse a otras áreas y a otros microorganismos (virus, hongos, parásitos, etc.) donde sea relevante establecer el diagnóstico patogénico de una manera más integral (Salud humana/animal, seguridad alimentaria, monitoreo de contaminaciones en distintos procesos).

8. REFERENCIAS BIBLIOGRÁFICAS

- Aguilera, A. *et al.* (2018) "Bloom-forming cyanobacteria and cyanotoxins in Argentina: A growing health and environmental concern," *Limnologica*. Elsevier, 69(October), pp. 103–114. doi: 10.1016/j.limno.2017.10.006.
- Alessandrini, F. *et al.* (2020) "Evaluation of the ion ampliseq SARS-CoV-2 research panel by massive parallel sequencing," *Genes*, 11(8), pp. 1–14. doi: 10.3390/genes11080929.
- Alhamlan, F. S., Al-Qahtani, A. A. & Al-Ahdal, M. N. (2015) "Recommended advanced techniques for waterborne pathogen detection in developing countries," *Journal of Infection in Developing Countries*, 9(2), pp. 128–135. doi: 10.3855/jidc.6101.
- Alzweghaibi, A. B. *et al.* (2018) "Rapid molecular identification and differentiation of common *Salmonella* serovars isolated from poultry, domestic animals and foodstuff using multiplex PCR assay," *Archives of Microbiology*. Springer Berlin Heidelberg, 200(7), pp. 1009–1016. doi: 10.1007/s00203-018-1501-7.
- AMPure XP (2019) "Manual or Automated Purification and Clean-up". USA. 4 p. Disponible en: <https://www.beckman.com/reagents/genomic/cleanup-and-size-selection/pcr>
- Andrinolo, D. *et al.* (2007) "Occurrence of *Microcystis aeruginosa* and microcystins in Rio de la Plata river (Argentina)," *Acta toxicol. argent*, 15(1), pp. 8–14.
- Anis, E. *et al.* (2018) "Evaluation of Targeted Next-Generation Sequencing for Detection of Bovine Pathogens in Clinical Samples," *Journal of Clinical Microbiology*, 56(7), pp. 1–11. doi: 10.1128/jcm.00399-18.
- Aubriot, L. E. *et al.* (2005) "Vulnerabilidad de una laguna costera en una Reserva de Biósfera: indicios recientes de eutrofización.," in JVPIPK, ed. Taller Internacional de Eutrofización y Embalses CYTED VXII B. Santiago de Chile: Patagonia Impresores, pp. 65–87.
- Aubriot, L. & Bonilla, S. (2012) "Rapid regulation of phosphate uptake in freshwater cyanobacterial blooms," *Aquatic Microbial Ecology*, 67(3), pp.251-263.
- Aubriot, L. E. *et al.* (2018) "Evolución de la eutrofización en el Río Santa Lucía: influencia de la intensificación productiva y perspectivas," *INNOTEC*, 14(14), pp. 7–16. doi: 10.26461/14.04.
- Ballot, A. *et al.* (2016) "Variability in the *sxt* gene clusters of PSP toxin producing *Aphanizomenon gracile* strains from Norway, Spain, Germany and North America," *PLoS ONE*, 11(12). doi: 10.1371/journal.pone.0167552.
- Ballot, A., Scherer, P. I. & Wood, S. A. (2018) "Variability in the anatoxin gene clusters of *Cuspidothrix issatschenkoi* from Germany, New Zealand, China and Japan," *PLoS ONE*, 13(7), pp. 1–13. doi: 10.1371/journal.pone.0200774.
- Balsalobre, L. C. *et al.* (2009) "Molecular detection of enterotoxins in environmental strains of *Aeromonas hydrophila* and *Aeromonas jandaei*," *Journal of Water and Health*, 7(4), pp. 685–691. doi: 10.2166/wh.2009.082.
- Barbau-Piednoir, E. *et al.* (2018) "Detection and discrimination of five *E. coli* pathotypes using a combinatory SYBR® Green qPCR screening system," *Applied Microbiology and Biotechnology*, 102(7), pp. 3267–3285. doi: 10.1007/s00253-018-8820-0.
- Bartram, J. *et al.* (2015) "Handbook of Water and Health," 1st edition Howick Place, London UK, Routledge, chapter 3, Bartram, J., Hunter, P., Bradley classification of disease transmission Routes for water related hazards: 20-3.
- Becker, K. *et al.* (2004) "Systematic survey on the prevalence of genes coding for staphylococcal enterotoxins SEIM, SEIO, and SEIN," *Molecular Nutrition and Food Research*. doi: 10.1002/mnfr.200400044.
- Bekal, S. *et al.* (2003) "Rapid identification of *Escherichia coli* pathotypes by virulence gene detection with DNA microarrays," *Journal of Clinical Microbiology*. doi: 10.1128/JCM.41.5.2113-2125.2003.

- Belardinilli, F. *et al.* (2015) "Validation of the Ion Torrent PGM sequencing for the prospective routine molecular diagnostic of colorectal cancer," *Clinical Biochemistry*. doi: 10.1016/j.clinbiochem.2015.04.003.
- Berg, K. *et al.* (2009) Heterotrophic Bacteria Associated with Cyanobacteria in Recreational and Drinking Water, *The ISME Journal*.
- de Bernard, M. & Josenhans, C. (2014) "Pathogenesis of *Helicobacter pylori* infection," *Helicobacter*, 19, pp. 11–18. doi: 10.1111/hel.12160.
- Bertin, Y. *et al.* (1996) "Rapid and specific detection of F17-related pilin and adhesin genes in diarrheic and septicemic *Escherichia coli* strains by multiplex PCR," *Journal of Clinical Microbiology*, 34(12), pp. 2921–2928. doi: 10.1128/jcm.41.12.5798-5802.2003.
- Bibby, K., Viau, E. & Peccia, J. (2010) "Pyrosequencing of the 16S rRNA gene to reveal bacterial pathogen diversity in biosolids," *Water Research*. Elsevier Ltd, 44(14), pp. 4252–4260. doi: 10.1016/j.watres.2010.05.039.
- Bitton, G. (2014) "Microbiology of Drinking Water Production and Distribution," John Wiley & Sons Inc. doi: 10.1002/9781118743942.
- Black, D. S. *et al.* (2000) "The *Yersinia* tyrosine phosphatase YopH targets a novel adhesion-regulated signalling complex in macrophages," *Cellular Microbiology*, 2(5), pp. 401–414. doi: 10.1046/j.1462-5822.2000.00061.x.
- Blue, G. M. *et al.* (2014) "Targeted next-generation sequencing identifies pathogenic variants in familial congenital heart disease," *Journal of the American College of Cardiology*, 64(23), pp. 2498–2506. doi: 10.1016/j.jacc.2014.09.048.
- Bonilla, S. *et al.* (2015) "Cianobacterias y cianotoxinas en ecosistemas límnicos de Uruguay," *Innotec*, 10(10 ene-dic), pp. 9–22. doi: 10.26461/innotec.v0i10.
- Bonilla, S. (2009) "Cianobacterias Planctónicas del Uruguay: Manual para la identificación y medidas de gestión," *Documento Técnico PHI*, vol. 16.
- Bonilla, S. *et al.* (2006) "Procesos estructuradores de las comunidades biológicas en lagunas costeras de Uruguay" En: Menafrá, R., Rodríguez-Gallego, L., Scarabino, F. y Conde, D. (eds.) *Bases para la conservación y el manejo de la costa uruguaya*. Montevideo: *Vida Silvestre*. pp.611-630. ISBN: 9974-7589-2-0.
- Bonilla, S. & Conde, D. (2000) "El fitoplancton como descriptor sensible de cambios ambientales en las lagunas costeras de la Reserva Bañados del Este," Seminario-Taller sobre monitoreo ambiental. Rocha: *PROBIDES/UNESCO. (Documentos de Trabajo, N° 31)*. pp.63-74.
- Bonilla, S., Pérez, M. & De León, L. (1995) "Cianofíceas Planctónicas del Lago Ton-Ton, Canelones, Uruguay," *Hoehnea*, 21, pp.85-192.
- Bordet, F., Fontanarrosa, M. S. & O'Farrell, I. (2017) "Influence of light and mixing regime on bloom-forming phytoplankton in a subtropical reservoir," *River Research and Applications*, 33(8), pp. 1315–1326. doi: 10.1002/rra.3189.
- Børsting, C. *et al.* (2014) "Evaluation of the Ion Torrent™ HID SNP 169-plex: A SNP typing assay developed for human identification by second generation sequencing," *Forensic Science International: Genetics*. Elsevier Ireland Ltd, 12, pp. 144–154. doi: 10.1016/j.fsigen.2014.06.004.
- Bottone, E. J. (2010) "*Bacillus cereus*, a volatile human pathogen," *Clinical Microbiology Reviews*, 23(2), pp. 382–398. doi: 10.1128/CMR.00073-09.
- Braga, V. *et al.* (2017) "Prevalence and serotype distribution of *Listeria monocytogenes* isolated from foods in Montevideo-Uruguay," *Brazilian Journal of Microbiology*. Sociedade Brasileira de Microbiologia, 48(4), pp. 689–694. doi: 10.1016/j.bjm.2017.01.010.
- Bradley, D. (2009). *The Spectrum of Water-Related Disease Transmission Processes*. Global Issues in Water, Sanitation and Health, pp.60-73. The National Academies Press, Washington DC, USA.

Brena, B. *et al.* (2006) "ITREOH building of regional capacity to monitor recreational water: development of a non- commercial microcystin ELISA and its impact on public health policy," *International Journal of Occupational and Environmental Health*, 12 (4), e377-e385.

Brena, B. & S. Bonilla, 2009. Producción de toxinas y otros metabolitos. In Bonilla, S. (ed) *Cianobacterias planctónicas del Uruguay: Manual para la identificación y medidas de gestión* Documento Técnico PHI N° 16. UNESCO, Montevideo, 16-18.

Brinkmann, A. *et al.* (2017) "Development and preliminary evaluation of a multiplexed amplification and next generation sequencing method for viral hemorrhagic fever diagnostics," *PLoS Neglected Tropical Diseases*, 11(11). doi: 10.1371/journal.pntd.0006075.

Caliendo, A. M. *et al.* (2013) "Better tests, better care: improved diagnostics for infectious diseases.," *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America*, 57 Suppl 3(Suppl 3), pp. 139–170. doi: 10.1093/cid/cit578.

Cangelosi, G. A. (2009) "Prospects for applying virulence factor-activity relationships (VFAR) to emerging waterborne pathogens," *Journal of Water and Health*, 7(SUPPL. 1), pp. 64–74. doi: 10.2166/wh.2009.045.

Caporaso, J. G. *et al.* (2010) "QIIME allows analysis of high-throughput community sequencing data.," *Nature methods*, 7(5), pp. 335–6. doi: 10.1038/nmeth.f.303.

Caporaso, J. G. *et al.* (2011) "Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample," *Proceedings of the National Academy of Sciences of the United States of America*, 108(SUPPL. 1), pp. 4516–4522. doi: 10.1073/pnas.1000080107.

Carmichael, W. W. (2001) "Health Effects of Toxin-Producing Cyanobacteria: 'The CyanoHABs,'" *Human and Ecological Risk Assessment: An International Journal*, 7(5), pp. 1393–1407. doi: 10.1080/20018091095087.

Casadevall, A. & Pirofski, L. A. (2003) "The damage-response framework of microbial pathogenesis," *Nature Reviews Microbiology*. doi: 10.1038/nrmicro732.

Casero, M. C. *et al.* (2014) "Characterization of saxitoxin production and release and phylogeny of sxt genes in paralytic shellfish poisoning toxin-producing *Aphanizomenon gracile*," *Harmful Algae*, 37, pp. 28–37. doi: 10.1016/j.hal.2014.05.006.

Castresana, J. (2000) "Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis.," *Molecular biology and evolution*, 17(4), pp. 540–552. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10742046>.

Chacón, M. R. *et al.* (2004) "Type III Secretion System Genes in Clinical *Aeromonas* Isolates," *Journal of Clinical Microbiology*. doi: 10.1128/JCM.42.3.1285-1287.2004.

Chalar, G. (2009) "The use of phytoplankton patterns of diversity for algal bloom management," *Limnologia*, 39(3), pp. 200–208. doi: 10.1016/j.limno.2008.04.001.

Chalar, G., Gerhard, M., González-Piana, M. & Fabián, D. (2014) "Hidroquímica y eutrofización en tres embalses subtropicales en cadena,". En: Marcovecchio, J. E., Botté, S. E. y Freije, R. H., *Procesos geoquímicos superficiales en Sudamérica*. Salamanca: Nueva Graficesa. pp.121-148. ISBN: 978-84-937437-6-5.

Chalar, G., Fabián, D., González-Piana, M. & Piccardo, A. (2015) "Estado y evolución de la calidad de agua de los tres embalses del Río Negro," Montevideo: Facultad de Ciencias, UdelaR, UTE.

Chen, L. *et al.* (2005) "VFDB: A reference database for bacterial virulence factors," *Nucleic Acids Research*. doi: 10.1093/nar/gki008.

Cho, S. *et al.* (2017) "SNP-Based Fetal DNA Detection in Maternal Serum Using the HID-Ion AmpliSeq™ Identity Panel," *Korean Journal of Legal Medicine*. doi: 10.7580/kjlm.2017.41.2.41.

Chorus, I. (2001) "Cyanotoxin Occurrence in Freshwaters," in *Cyanotoxins*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 5–101. doi: 10.1007/978-3-642-59514-1_2.

- Chorus, I. & Bartram, J. (1999) "A guide to their public health consequences, monitoring and management," in *Toxic Cyanobacteria in Water*: doi: 10.1046/j.1365-2427.2003.01107.x.
- Choudhury, S. (2020) Molecular tools for the detection of waterborne pathogens, *Waterborne Pathogens*. Elsevier. doi: 10.1016/b978-0-12-818783-8.00012-8.
- Claesson, M. J. *et al.* (2009) "Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structures in the human distal intestine," *PLoS ONE*, 4(8). doi: 10.1371/journal.pone.0006669.
- Collier, S. A. *et al.* (2012) "Direct healthcare costs of selected diseases primarily or partially transmitted by water," *Epidemiology and Infection*, 140(11), pp. 2003–2013. doi: 10.1017/S0950268811002858.
- Conde, D., Hein, V. & Bonilla, S. (2009). "Floraciones de cianobacterias en lagunas costeras". En: Bonilla, S. (ed). *Cianobacterias planctónicas del Uruguay. Manual para la identificación y medidas de gestión*. Montevideo: UNESCO. (Documento técnico, 16). pp.79-80.
- Content, J. *et al.* (1991) "The genes coding for the antigen 85 complexes of *Mycobacterium tuberculosis* and *Mycobacterium bovis* BCG are members of a gene family: Cloning, sequence determination, and genomic organization of the gene coding for antigen 85-C of *M. tuberculosis*," *Infection and Immunity*, 59(9), pp. 3205–3212. doi: 10.1128/iai.59.9.3205-3212.1991.
- Cornelis, G. R. & Wolf-Watz, H. (1997) "The *Yersinia* Yop virulon: A bacterial system for subverting eukaryotic cells," *Molecular Microbiology*, 23(5), pp. 861–867. doi: 10.1046/j.1365-2958.1997.2731623.x.
- Croxen, M. A. *et al.* (2013) "Recent advances in understanding enteric pathogenic *Escherichia coli*," *Clinical Microbiology Reviews*, 26(4), pp. 822–880. doi: 10.1128/CMR.00022-13.
- Cui, Q. *et al.* (2017) "Evaluation of bacterial pathogen diversity, abundance and health risks in urban recreational water by amplicon next-generation sequencing and quantitative PCR," *Journal of Environmental Sciences (China)*. Elsevier B.V., 57, pp. 137–149. doi: 10.1016/j.jes.2016.11.008.
- D'Agostino, P. M. *et al.* (2016) "Advances in genomics, transcriptomics and proteomics of toxin-producing cyanobacteria," *Environmental Microbiology Reports*, 8(1), pp. 3–13. doi: 10.1111/1758-2229.12366.
- De León, L. & Yunes, J. S. (2001) "First report of a microcystin- containing bloom of the cyanobacterium *Microcystis aeruginosa* in the La Plata River, South America," *Environmental Toxicology*, 16(1), pp.110-112.
- Deshmukh, R. A. *et al.* (2016) "Recent developments in detection and enumeration of waterborne bacteria: a retrospective minireview," *MicrobiologyOpen*, 5(6), pp. 901–922. doi: 10.1002/mbo3.383.
- Deurenberg, R. H. *et al.* (2017) "Application of next generation sequencing in clinical microbiology and infection prevention," *Journal of Biotechnology*. Elsevier B.V., 243, pp. 16–24. doi: 10.1016/j.jbiotec.2016.12.022.
- Deus-Álvarez, S. (2018) "Caracterización morfológica, genética y óptica de floraciones tóxicas y no tóxicas del Complejo *Microcystis aeruginosa*." Udelar. FC.
- Deus-Álvarez, S. *et al.* (2020) "Morphology captures toxicity in *Microcystis aeruginosa* complex: Evidence from a wide environmental gradient☆," *Harmful Algae*. Elsevier, 97(January), p. 101854. doi: 10.1016/j.hal.2020.101854.
- Dheenadhayalan, V. *et al.* (2002) "Cloning and characterization of the genes coding for antigen 85A, 85B and 85C of *Mycobacterium avium* subsp. *paratuberculosis*," *Mitochondrial DNA*, 13(5), pp. 287–294. doi: 10.1080/1042517021000019269.
- Dokulil, M. T. & Teubner, K. (2000) "Cyanobacterial dominance in lakes," *Hydrobiologia*, 438, pp. 1–12. doi: 10.1023/A:1004155810302.
- Dommel, M. K. *et al.* (2011) "Transcriptional kinetic analyses of cereulide synthetase genes with respect to growth, sporulation and emetic toxin production in *Bacillus cereus*," *Food Microbiology*. Elsevier Ltd, 28(2), pp. 284–290. doi: 10.1016/j.fm.2010.07.001.

- Dong, L. et al. (2015) "Clinical Next Generation Sequencing for Precision Medicine in Cancer," *Current Genomics*, 16(4), pp. 253–263. doi: 10.2174/1389202915666150511205313.
- Dörr, F. A. et al. (2010) "Microcystins in South American aquatic ecosystems: Occurrence, toxicity and toxicological assays," *Toxicon*. Elsevier Ltd, 56(7), pp. 1247–1256. doi: 10.1016/j.toxicon.2010.03.018.
- Drancourt, M. (2012) "Plague in the genomic area," *Clinical Microbiology and Infection. European Society of Clinical Infectious Diseases*, 18(3), pp. 224–230. doi: 10.1111/j.1469-0691.2012.03774.x.
- Dubail, I. et al. (2000) "Listeriolysin O as a reporter to identify constitutive and in vivo- inducible promoters in the pathogen *Listeria monocytogenes*," *Infection and Immunity*, 68(6), pp. 3242–3250. doi: 10.1128/IAI.68.6.3242-3250.2000.
- Duesbery, N. S. & Woude, G. F. Van de (1999) "Review Anthrax toxins," *Cellular and Molecular Life Sciences CMLS*, 55, pp. 1599–1609.
- Edgar, R. C. (2004) "MUSCLE: Multiple sequence alignment with high accuracy and high throughput," *Nucleic Acids Research*, 32(5), pp. 1792–1797. doi: 10.1093/nar/gkh340.
- van Elden, L. J. R. et al. (2002) "Polymerase chain reaction is more sensitive than viral culture and antigen testing for the detection of respiratory viruses in adults with hematological cancer and pneumonia," *Clinical Infectious Diseases*, 34(2), pp. 177–183. doi: 10.1086/338238.
- Fabre, A. et al. (2010) "El nitrógeno y la relación zona eufótica/zona de mezcla explican la presencia de cianobacterias en pequeños lagos subtropicales, artificiales de Uruguay," *Pan-American Journal of Aquatic Sciences*, 5(1), pp.112-125.
- Feltman, H. et al. (2001) "Prevalence of type III secretion genes in clinical and environmental isolates of *Pseudomonas aeruginosa*," *Microbiology*. doi: 10.1099/00221287-147-10-2659.
- Feola, G. et al. (2010) "Programa de monitoreo de agua de playas y de costa de Montevideo," Informe 2009-2010. Montevideo: Intendencia de Montevideo. (Informe técnico).
- Ferrari, G. et al. (2011) "Planktic Cyanobacteria in the lower Uruguay River, South America," *Fottea*, 11(1), pp.225-234.
- Ferrari, G. & Vidal, L. (2006) "Fitoplancton de la zona costera uruguaya: Río de la Plata y Océano Atlántico.," En: Menafrá, R., Rodríguez-Gallago, L., Scarabino, F. y Conde, D., eds. Bases para la conservación y el manejo de la costa uruguaya. Montevideo: Vida Silvestre., pp. 45–56.
- Fey, A. et al. (2004) "Establishment of a real-time PCR-based approach for accurate quantification of bacterial RNA targets in water, using *Salmonella* as a model organism," *Applied and Environmental Microbiology*, 70(6), pp. 3618–3623. doi: 10.1128/AEM.70.6.3618-3623.2004.
- Figueras, J. M. & Borrego, J. J. (2010) "New perspectives in monitoring drinking water microbial quality," *International Journal of Environmental Research and Public Health*. doi: 10.3390/ijerph7124179.
- Fogg, P. C. M. et al. (2012) "Cumulative effect of prophage burden on Shiga toxin production in *Escherichia coli*," *Microbiology*. doi: 10.1099/mic.0.054981-0.
- Forstinus, N. et al. (2016) "Water and Waterborne Diseases: A Review," *International Journal Of Tropical Disease & Health*, 12(4), pp. 1–14. doi: 10.9734/ijtdh/2016/21895.
- Frangeul, L. et al. (2008) "Highly plastic genome of *Microcystis aeruginosa* PCC 7806, a ubiquitous toxic freshwater cyanobacterium," *BMC Genomics*, 9, pp. 1–20. doi: 10.1186/1471-2164-9-274.
- Fries, D., Carmichael, W., State, R., Scholin, C. & Doucette, G. (2001) "Cyanotoxin Detection and Quantification and Instrumentation," Workshop (CDQIW). Florida, USA.
- Gamage, S. D. et al. (2004) "Diversity and host range of Shiga toxin-encoding phage," *Infection and Immunity*, 72(12), pp. 7131–7139. doi: 10.1128/IAI.72.12.7131-7139.2004.
- Gardner, S. N. et al. (2015) "Targeted amplification for enhanced detection of biothreat agents by next-generation sequencing," *BMC Research Notes. BioMed Central*, 8(1), pp. 1–10. doi: 10.1186/s13104-015-1530-0.

- Gellatly, S. L. & Hancock, R. E. W. (2013) "Pseudomonas aeruginosa: New insights into pathogenesis and host defenses," *Pathogens and Disease*, 67(3), pp. 159–173. doi: 10.1111/2049-632X.12033.
- Gerba, C. P. & Pepper, I. L. (2019) *Microbial Contaminants*. 3rd ed, Environmental and Pollution Science. 3rd ed. Elsevier Inc. doi: 10.1016/b978-0-12-814719-1.00013-6.
- Goletic, T. *et al.* (2020) "Phylogenetic pattern of SARS-CoV-2 from COVID-19 patients from Bosnia and Herzegovina: lessons learned to optimize future molecular and epidemiological approaches," *bioRxiv*.
- González-Piana, M., Fabian, D., Delbene, L. & Chalar, G. (2011) "Toxics blooms of *Microcystis aeruginosa* in three Rio Negro reservoirs, Uruguay," *Harmful Algae News*, 43, pp.16-17.
- González-Piana, M. *et al.* (2017) "Dynamics of Total Microcystin LR Concentration in Three Subtropical Hydroelectric Generation Reservoirs in Uruguay, South America," *Bulletin of Environmental Contamination and Toxicology*. Springer US, 99(4), pp. 488–492. doi: 10.1007/s00128-017-2158-7.
- González-Revello. *et al.* (2017) "Potencial Patogénico Del Agua: Diversidad Filogenética Vs. Diversidad De Ecotipos," En: 1er Taller de la Red Colaborativa en Ecología Microbiana Acuática en América Latina (MicroSudAqua), Rocha, Uruguay.
- Gravier, A., Vidal, L. & Britos, A. (2009). *Cianobacterias como organismos interferentes en el servicio de agua potable de Bella Unión (Uruguay)*. En: Resúmenes del 1º Encuentro Uruguayo. *Cianobacterias: del conocimiento a la gestión*. Montevideo: LATU; Facultad de Ciencias.
- Guinebreiere, M. H. *et al.* (2006) "Rapid discrimination of *cytK-1* and *cytK-2* genes in *Bacillus cereus* strains by a novel duplex PCR system," *FEMS Microbiology Letters*, 259(1), pp. 74–80. doi: 10.1111/j.1574-6968.2006.00247.x.
- De Haan, L. & Hirst, T. R. (2004) "Cholera toxin: A paradigm for multi-functional engagement of cellular mechanisms (Review)," *Molecular Membrane Biology*, 21(2), pp. 77–92. doi: 10.1080/09687680410001663267.
- Hall, T. A. (1999) "BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.," *Nucl. Acids. Symp. Ser.* doi: citeulike-article-id:691774.
- Harada, K. *et al.* (1995) "Two cyclic peptides, anabaenopeptins, a third group of bioactive compounds from the cyanobacterium *Anabaena flos-aquae* NRC 525-17," *Tetrahedron Letters*, 36(9), pp. 1511–1514. doi: 10.1016/0040-4039(95)00073-L.
- Harke, M. J. *et al.* (2016) "A review of the global ecology, genomics, and biogeography of the toxic cyanobacterium, *Microcystis* spp.," *Harmful Algae*. Elsevier B.V., 54, pp. 4–20. doi: 10.1016/j.hal.2015.12.007.
- Harrell, L. J., Andersen, G. L. & Wilson, K. H. (1995) "Genetic variability of *Bacillus anthracis* and related species," *Journal of Clinical Microbiology*, 33(7), pp. 1847–1850. doi: 10.1128/jcm.33.7.1847-1850.1995.
- He, Y. *et al.* (2017) "Novel mutations in patients with hereditary red blood cell membrane disorders using next-generation sequencing," *Gene*. doi: 10.1016/j.gene.2017.07.009.
- Heck, K. *et al.* (2018) "Biosynthesis of microcystin hepatotoxins in the cyanobacterial genus *Fischerella*," *Toxicon*, 141, pp. 43–50. doi: 10.1016/j.toxicon.2017.10.021.
- Hernandes, R. T. *et al.* (2009) "An overview of atypical enteropathogenic *Escherichia coli*," *FEMS Microbiology Letters*. doi: 10.1111/j.1574-6968.2009.01664.x.
- Hirigoyen, D. *et al.* (2018) Proyecto MGAP. Ministerio de Ganadería Agricultura y Pesca, Programa: +Tecnologías Para La Producción. Caracterización molecular y conservación de suero fermento de quesos artesanales (03/2015 - 07/2017).
- Hovelson, D. H. *et al.* (2015) "Development and Validation of a Scalable Next-Generation Sequencing System for Assessing Relevant Somatic Variants in Solid Tumors," *Neoplasia*. doi: 10.1016/j.neo.2015.03.004.
- Huisman, J. & Weissing, F. J. (2001) "Biological conditions for oscillations and chaos generated by multispecies competition," *Ecology*. doi: 10.1890/0012-9658(2001)082[2682:BCFOAC]2.0.CO;2.

Ibekwe, A. M., Leddy, M. & Murinda, S. E. (2013) "Potential Human Pathogenic Bacteria in a Mixed Urban Watershed as Revealed by Pyrosequencing," *PLoS ONE*. Edited by R. K. Aziz, 8(11), p. e79490. doi: 10.1371/journal.pone.0079490.

Ion AmpliSeq™ library preparation (2017). Ion AmpliSeq™ Library Kit 2.0 User Guide: Targeted DNA and RNA Library Preparation. Pub. No. MAN0006735.Revisión. E.O. 91 p. Disponible en: https://assets.thermofisher.com/TFS-assets/LSG/manuals/MAN0006735_AmpliSeq_DNA_RNA_Lib-Prep_UG.pdf

Ion AmpliSeq™ Designer (2018). Ion AmpliSeq™ Designer: Getting Started. USER GUIDE. Publication Number MAN0010907. Revisión F.O. 89 p. Disponible en: https://s3.amazonaws.com/ampliseq-ui-public-assets/MAN0010907_AmpliSeq_Designer_UG_RevF_12Jun2018.pdf

Ingerson-Mahar, M. & Reid, A. (2012) "Microbes in Pipes: The Microbiology Of Water Distribution Systems," A Report On An American Academy Of Microbiology Colloquium.

Jang, W. *et al.* (2019) "Hereditary spherocytosis caused by copy number variation in SPTB gene identified through targeted next-generation sequencing," *International Journal of Hematology*. doi: 10.1007/s12185-019-02630-0.

Jenkins, T. M. *et al.* (2004) "Assessment of virulence-factor activity relationships (VFARs) for waterborne diseases," *Water Science and Technology*, 50(1), pp. 309–314. doi: 10.2166/wst.2004.0072.

Jennings, L. J. *et al.* (2017) "Guidelines for Validation of Next-Generation Sequencing–Based Oncology Panels: A Joint Consensus Recommendation of the Association for Molecular Pathology and College of American Pathologists," *Journal of Molecular Diagnostics*. American Society for Investigative Pathology and the Association for Molecular Pathology, 19(3), pp. 341–365. doi: 10.1016/j.jmoldx.2017.01.011.

Jin, L. & Nei, M. (1990) "Limitations of the evolutionary parsimony method of phylogenetic analysis," *Molecular Biology and Evolution*. doi: 10.1093/oxfordjournals.molbev.a040596.

Kearse, M. *et al.* (2012) "Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data," *Bioinformatics*, 28(12), pp. 1647–1649. doi: 10.1093/bioinformatics/bts199.

Kellmann, R. *et al.* (2008) "Biosynthetic intermediate analysis and functional homology reveal a saxitoxin gene cluster in cyanobacteria," *Applied and Environmental Microbiology*, 74(13), pp. 4044–4053. doi: 10.1128/AEM.00353-08.

Khajanchi, B. K. *et al.* (2010) "Distribution of virulence factors and molecular fingerprinting of *Aeromonas* species isolates from water and clinical samples: Suggestive evidence of water-to-human transmission," *Applied and Environmental Microbiology*, 76(7), pp. 2313–2325. doi: 10.1128/AEM.02535-09.

Khodakov, D., Wang, C. & Zhang, D. Y. (2016) "Diagnostics based on nucleic acid sequence variant profiling: PCR, hybridization, and NGS approaches," *Advanced Drug Delivery Reviews*. The Authors, 105, pp. 3–19. doi: 10.1016/j.addr.2016.04.005.

Kim, S. G. *et al.* (2010) "Annual variation of *Microcystis* genotypes and their potential toxicity in water and sediment from a eutrophic reservoir," *FEMS Microbiology Ecology*, 74(1), pp. 93–102. doi: 10.1111/j.1574-6941.2010.00947.x.

Klappenbach, J. A., Dunbar, J. M. & Schmidt, T. M. (2000) "rRNA operon copy number reflects ecological strategies of bacteria," *Applied and Environmental Microbiology*, 66(4), pp. 1328–1333. doi: 10.1128/AEM.66.4.1328-1333.2000.

Ko, D.-H. *et al.* (2019) "Application of next-generation sequencing to detect variants of drug-resistant *Mycobacterium tuberculosis*: genotype-phenotype correlation.," *Annals of clinical microbiology and antimicrobials*. *BioMed Central*, 18(1), p. 2. doi: 10.1186/s12941-018-0300-y.

Komárek, J. & Komárková, J. (2002) "Review of the European *Microcystis* morphospecies (Cyanoprokaryotes) from nature.," *Fottea*, 2(1), pp. 1–24.

- Kozlikova-Zapomelova, E., Ferrari, G. & Perez, M. del C. (2016) "Dolichospermum uruguayense sp. nov., a planktic nostocacean cyanobacterium from the Lower Uruguay River, South America," *Fottea*, 16(2), pp. 189–200. doi: 10.5507/fot.2016.009.
- Kruk, C. et al. (2011) "Phytoplankton community composition can be predicted best in terms of morphological groups," *Limnology and Oceanography*, 56(1), pp. 110–118. doi: 10.4319/lo.2011.56.1.0110.
- Kruk, C. et al. (2013) "Análisis calidad de agua en Uruguay," Asesoramiento Ambiental Estratégico, Vida Silvestre, IUCN. Montevideo, Uruguay.
- Kruk, C. et al. (2015) "Herramientas para el monitoreo y sistema de alerta de floraciones de cianobacterias nocivas: Río Uruguay y Río de la Plata," *INNOTECH*, 10, pp. 23–39.
- Kruk, C. et al. (2017) "A multilevel trait-based approach to the ecological performance of *Microcystis aeruginosa* complex from headwaters to the ocean," *Harmful Algae*. Elsevier B.V., 70, pp. 23–36. doi: 10.1016/j.hal.2017.10.004.
- Kruk, C. et al. (2021) "Rapid freshwater discharge on the coastal ocean as a mean of long distance spreading of an unprecedented toxic cyanobacteria bloom," *Science of the Total Environment*. Elsevier B.V., 754, p. 142362. doi: 10.1016/j.scitotenv.2020.142362.
- Kruk, C. et al. (2019) "Floración excepcional de cianobacterias tóxicas en la costa de Uruguay, verano 2019," *INNOTECH*, 18(August), pp. 36–68. doi: 10.26461/18.06.
- Kulinkina, B. A. V. et al. (2016) "The situation of water-related infectious diseases in the pan-European region," *Protocol on Water and Health*.
- Kumaraswamy, R. et al. (2014) "Molecular analysis for screening human bacterial pathogens in municipal wastewater treatment and reuse," *Environmental Science and Technology*, 48(19), pp. 11610–11619. doi: 10.1021/es502546t.
- Kung, V. L., Ozer, E. A. & Hauser, A. R. (2010) "The Accessory Genome of *Pseudomonas aeruginosa*," *Microbiology and Molecular Biology Reviews*, 74(4), pp. 621–641. doi: 10.1128/mmbr.00027-10.
- Kwong, J. C. et al. (2015) "Whole genome sequencing in clinical and public health microbiology," *Pathology*. doi: 10.1097/PAT.0000000000000235.
- Laupland, K. B. & Valiquette, L. (2013) "The changing culture of the microbiology laboratory," *Canadian Journal of Infectious Diseases and Medical Microbiology*, 24(3), pp. 125–128. doi: 10.1155/2013/101630.
- Lawton, L., Edwards, C. & Codd, G. (1994) "Extraction and high-performance liquid chromatographic method for the determination of microcystins in raw and treated waters.," *The Analyst*, 119(7), pp. 1525–1530. doi: 10.1039/AN9941901525.
- Lee, A. et al. (2018) "Use of the Ion AmpliSeq Cancer Hotspot Panel in clinical molecular pathology laboratories for analysis of solid tumours: With emphasis on validation with relevant single molecular pathology tests and the OncoPrint Focus Assay," *Pathology Research and Practice*. doi: 10.1016/j.prp.2018.03.009.
- Lee, Y. J. et al. (2013) "The PathoChip, a functional gene array for assessing pathogenic properties of diverse microbial communities," *ISME Journal*. doi: 10.1038/ismej.2013.88.
- Leimeister-Wachter, M. et al. (1990) "Identification of a gene that positively regulates expression of listeriolysin, the major virulence factor of *listeria monocytogenes*," *Proceedings of the National Academy of Sciences*, 87(21), 8336–8340. doi:10.1073/pnas.87.21.8336
- Lemmon, G. H. & Gardner, S. N. (2008) "Predicting the sensitivity and specificity of published real-time PCR assays," *Annals of Clinical Microbiology and Antimicrobials*, 7. doi: 10.1186/1476-0711-7-18.
- Lencer, W. I. et al. (1992) "Mechanism of cholera toxin action on a polarized human intestinal epithelial cell line: Role of vesicular traffic," *Journal of Cell Biology*, 117(6), pp. 1197–1209. doi: 10.1083/jcb.117.6.1197.

- Li, K. *et al.* (2013) "Abstract 4218: Validation of the Ion AmpliSeq™ Comprehensive Cancer Panel (CCP) using cast PCR™ technologies.," in. doi: 10.1158/1538-7445.am2013-4218.
- Lindbäck, T. *et al.* (2010) "The ability to enter into an avirulent viable but non-culturable (VBNC) form is widespread among *Listeria monocytogenes* isolates from salmon, patients and environment," *Veterinary Research*. doi: 10.1051/vetres/2009056.
- Lipp, E. K., Farrah, S. A. & Rose, J. B. (2001) "Assessment and impact of microbial fecal pollution and human enteric pathogens in a coastal community," *Marine Pollution Bulletin*. doi: 10.1016/S0025-326X(00)00152-1.
- Little, S. F. & Ivins, B. E. (1999) "Molecular pathogenesis of *Bacillus anthracis* infection," *Microbes and Infection*. doi: 10.1016/S1286-4579(99)80004-5.
- Logan, N. A. (2012) "Bacillus and relatives in foodborne illness," *Journal of Applied Microbiology*, 112(3), pp. 417-29.
- De Los Santos, R. *et al.* (2014) "Characterization of *Staphylococcus aureus* isolated from cases of bovine subclinical mastitis in two Uruguayan dairy farms," *Archivos de Medicina Veterinaria*, 46(2), pp. 315–320. doi: 10.4067/S0301-732X2014000200018.
- Lu, X. *et al.* (2015) "Bacterial Pathogens and Community Composition in Advanced Sewage Treatment Systems Revealed by Metagenomics Analysis Based on High-Throughput Sequencing," *PLOS ONE*. Edited by H. Bach, 10(5), p. e0125549 1-15. doi: 10.1371/journal.pone.0125549.
- Luo, G. & Angelidaki, I. (2014) "Analysis of bacterial communities and bacterial pathogens in a biogas plant by the combination of ethidium monoazide, PCR and Ion Torrent sequencing," *Water Research*. Elsevier Ltd, 60, pp. 156–163. doi: 10.1016/j.watres.2014.04.047.
- Magana-Arachchi, D. N. & Wanigatunge, R. P. (2020) Ubiquitous waterborne pathogens, *Waterborne Pathogens*. Elsevier. doi: 10.1016/b978-0-12-818783-8.00002-5.
- Martínez de la Escalera, G. *et al.* (2017) "Dynamics of toxic genotypes of *Microcystis aeruginosa* complex (MAC) through a wide freshwater to marine environmental gradient," *Harmful Algae*, 62, pp. 73–83. doi: 10.1016/j.hal.2016.11.012.
- Martínez de la Escalera, G. *et al.* (2019) "Genotyping and functional regression trees reveals environmental preferences of toxic cyanobacteria (*Microcystis aeruginosa* complex) along a wide spatial gradient," *bioRxiv*. doi: 10.1101/2019.12.20.885111.
- Mazzeo, N. *et al.* (2002) "Eutrofización: causas, consecuencias y manejo," *Perfil Ambiental del Uruguay*. Nordan-Comunidad. Montevideo, Uruguay, pp. 39–55.
- McDermott, C. M., Feola, R. & Plude, J. (1995) "Detection of cyanobacterial toxins (microcystins) in waters of northeastern wisconsin by a new immunoassay technique," *Toxicon*, 33(11), pp. 1433–1442. doi: 10.1016/0041-0101(95)00095-4.
- McKillip, J. L. (2000) "Prevalence and expression of enterotoxins in *Bacillus cereus* and other *Bacillus* spp., a literature review," *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*. doi: 10.1023/A:1002706906154.
- Millat, G., Chanavat, V. & Rousson, R. (2014) "Evaluation of a new NGS method based on a custom AmpliSeq library and Ion Torrent PGM sequencing for the fast detection of genetic variations in cardiomyopathies," *Clinica Chimica Acta*. Elsevier B.V., 433, pp. 266–271. doi: 10.1016/j.cca.2014.03.032.
- Miller, S. M. *et al.* (2008) "In situ-synthesized virulence and marker gene biochip for detection of bacterial pathogens in water," *Applied and Environmental Microbiology*, 74(7), pp. 2200–2209. doi: 10.1128/AEM.01962-07.
- Minogue, T. D., Koehler, J. W. & Norwood, D. A. (2017) "Targeted next-generation sequencing for diagnostics and forensics," *Clinical Chemistry*, 63(2), pp. 450–452. doi: 10.1373/clinchem.2016.256065.

- Moffitt, M. C. & Neilan, B. A. (2004) "Characterization of the nodularin synthetase gene cluster and proposed theory of the evolution of cyanobacterial hepatotoxins," *Applied and Environmental Microbiology*, 70(11), pp. 6353–6362. doi: 10.1128/AEM.70.11.6353-6362.2004.
- Motro, Y. & Moran-Gilad, J. (2017) "Next-generation sequencing applications in clinical bacteriology," *Biomolecular Detection and Quantification*. Elsevier, 14(February), pp. 1–6. doi: 10.1016/j.bdq.2017.10.002.
- Nagy, B. & Fekete, P. Z. (1999) "Enterotoxigenic Escherichia coli (ETEC) in farm animals.," *Veterinary research*.
- Nagy, B. & Fekete, P. Z. (2005) "Enterotoxigenic Escherichia coli in veterinary medicine," *International Journal of Medical Microbiology*, 295(6–7), pp. 443–454. doi: 10.1016/j.ijmm.2005.07.003.
- National Research Council (2001) *Classifying Drinking Water Contaminants for Regulatory Consideration*, Classifying Drinking Water Contaminants for Regulatory Consideration. doi: 10.17226/10080.
- NDWAC (2004) National Drinking Water Advisory Council (2004) Report of the CCL Classification Process Work Group to the National Drinking Water Advisory Council; National Drinking Water Advisory Council: Washington, D.C., <http://www.awwa.org/Advocacy/govtaff/documents/CCLRatificationDraft.pdf>
- Nei, M. & Kumar, S. (2000) "Molecular evolution and phylogenetics". Oxford: Oxford University Press.
- Nervi, E. (2020) "Análisis de comunidades bacterianas en muestras de agua: puesta a punto de un protocolo in-house de secuenciación masiva de amplicones del gen ribosomal 16S". Facultad de Ciencias, Udelar. 74p.
- Nocker, A., Burr, M. & Camper, A. (2013) "Pathogens in Water and Biofilms," in *Microbiology of Waterborne Diseases: Microbiological Aspects and Risks: Second Edition*, pp. 3–32. doi: 10.1016/B978-0-12-415846-7.00001-9.
- O'Farrell, I., Bordet, F. & Chaparro, G. (2012) "Bloom forming cyanobacterial complexes co-occurring in a subtropical large reservoir, validation of dominant eco-strategies", *Hydrobiologia*, 698(1), pp.175-190. doi: 10.1007/s10750-012-1102-4.
- Økstad, O. A. & Kolstø, A.-B. (2011) "Genomics of Bacillus Species," in *Genomics of Foodborne Bacterial Pathogens*. doi: 10.1007/978-1-4419-7686-4_2.
- Oleiwi, S. R. (2020) "Water and Waterborne Diseases: A Review," *Plant Archives*, 20(1), pp. 547–550. Available at: http://plantarchives.org/20-1/547-550__6238_.pdf.
- Oliver, J. D. (2010) "Recent findings on the viable but nonculturable state in pathogenic bacteria," *FEMS Microbiology Reviews*, 34(4), pp. 415–425. doi: 10.1111/j.1574-6976.2009.00200.x.
- Ooka, T. et al. (2012) "Clinical significance of Escherichia albertii," *Emerging Infectious Diseases*. doi: 10.3201/eid1803.111401.
- Osswald, J. et al. (2007) "Toxicology and detection methods of the alkaloid neurotoxin produced by cyanobacteria, anatoxin-a," *Environment International*, 33(8), pp. 1070–1089. doi: 10.1016/j.envint.2007.06.003.
- Otsuka, S. et al. (2000) "Morphological variability of colonies of Microcystis morphospecies in culture," *Journal of General and Applied Microbiology*, 46(1), pp. 39–50. doi: 10.2323/jgam.46.39.
- Pacheco, J. P. et al. (2010) "Phytoplankton community structure in five subtropical shallow lakes with different trophic status (Uruguay): a morphology-based approach," *Hydrobiologia*, 646(1), pp.187-197.
- Paerl, H. W. (2017) "Controlling harmful cyanobacterial blooms in a climatically more extreme world: management options and research needs," *Journal of Plankton Research*, 39(5), pp. 763–771. doi: 10.1093/plankt/fbx042.
- Park, J. et al. (2018) "Determining genotypic drug resistance by ion semiconductor sequencing with the ion ampliSeq™ TB panel in multidrug-resistant mycobacterium tuberculosis isolates," *Annals of Laboratory Medicine*, 38(4), pp. 316–323. doi: 10.3343/alm.2018.38.4.316.

- Pérez, M. C., Bonilla, S., De León, L., Smarda, J. & Komárek, J. (1999) "A bloom of *Nodularia baltica*-*spumigena* group (Cyanobacteria) in a shallow coastal lagoon of Uruguay, South America," *Algological Studies*, 128, pp.91-101.
- Pérez, M. C., Bonilla S., Haakonsson S. & Arocena R. (2016) "Dynamics of Chroococcales and Nostocales in a turbid subtropical estuary: Río de la Plata, South America". En: 20th IAC Cyanophyte Cyanobacteria Research Symposium. Austria. 65 p. Disponible en: <https://www.uibk.ac.at/congress/iac-symposium2016/abstract-volume.pdf>
- Perretta, A., Antúnez, K. & Zunino, P. (2018) "Phenotypic, molecular and pathological characterization of motile aeromonads isolated from diseased fishes cultured in Uruguay," *Journal of Fish Diseases*, 41(10), pp. 1559–1569. doi: 10.1111/jfd.12864.
- Péterfia, B. *et al.* (2017) "Construction of a multiplex mutation hot spot PCR panel: The first step towards colorectal cancer genotyping on the GS Junior platform," *Journal of Cancer*. doi: 10.7150/jca.16037.
- Piccini, C. *et al.* (2011) "Genetic and eco-physiological differences of South American *Cylindrospermopsis raciborskii* isolates support the hypothesis of multiple ecotypes," *Harmful Algae*, 10(6), pp. 644–653. doi: 10.1016/j.hal.2011.04.016.
- Piccini, C. *et al.* (2013) "Revealing Toxin Signatures in Cyanobacteria: Report of Genes Involved in Cylindrospermopsin Synthesis from Saxitoxin-Producing *Cylindrospermopsis raciborskii*" *Advances in Microbiology*, 03(03), pp. 289–296. doi: 10.4236/aim.2013.33041.
- Pérez, M. *et al.* (2013) "Limited analytical capacity for cyanotoxins in developing countries may hide serious environmental health problems: simple and affordable methods may be the answer," *Journal of Environmental Management*, 114, pp.63-71.
- Prachayangprecha, S. *et al.* (2014) "Exploring the potential of next-generation sequencing in detection of respiratory Viruses," *Journal of Clinical Microbiology*. doi: 10.1128/JCM.01641-14.
- Le Quéré, C. *et al.* (2005) "Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models," *Global Change Biology*, 11(11), pp. 2016–2040. doi: 10.1111/j.1365-2486.2005.1004.x.
- Quick, J. *et al.* (2017) "Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples," *bioRxiv*, p. 098913. doi: 10.1101/098913.
- Quirós, R. & Luchini, L. (1982) "Características limnológicas del embalse Salto Grande. III. Fitoplancton y su relación con parámetros ambientales," *Revista Asociación Ciencias Naturales del Litoral*, 13, pp.40-66.
- Rahn, K. *et al.* (1992) "Amplification of an *invA* gene sequence of *Salmonella* Typhimurium by polymerase chain reaction as a specific method of detection of *Salmonella*," *Molecular and Cellular Probes*. doi: 10.1016/0890-8508(92)90002-F.
- Ramamurthy, T. *et al.* (2014) "Current perspectives on viable but non-culturable (VBNC) pathogenic bacteria," *Frontiers in Public Health*. doi: 10.3389/fpubh.2014.00103.
- Ramírez-Castillo, F. *et al.* (2015) "Waterborne Pathogens: Detection Methods and Challenges," *Pathogens*, 4(2), pp. 307–334. doi: 10.3390/pathogens4020307.
- Rasmussen, G. *et al.* (2013) "Prevalence of Clonal Complexes and Virulence Genes among Commensal and Invasive *Staphylococcus aureus* Isolates in Sweden," *PLoS ONE*, 8(10), pp. 1–10. doi: 10.1371/journal.pone.0077477.
- Rastogi, R. P., Madamwar, D. & Incharoensakdi, A. (2015) "Bloom dynamics of cyanobacteria and their toxins: Environmental health impacts and mitigation strategies," *Frontiers in Microbiology*, 6(NOV), pp. 1–22. doi: 10.3389/fmicb.2015.01254.
- Rathi, V. *et al.* (2017) "Clinical validation of the 50 gene AmpliSeq Cancer Panel V2 for use on a next generation sequencing platform using formalin fixed, paraffin embedded and fine needle aspiration tumour specimens," *Pathology*. doi: 10.1016/j.pathol.2016.08.016.

- Rebollar, R. G., Arriaga, C. & de la Rosa, I. A. (2018) "Clinical Applications of Next-Generation Sequencing," *Revista de investigación Clínica*, 70(4), pp. 153–157. doi: 10.24875/RIC.18002544.
- Reddy, S. & Lawrence, M. L. (2014) "Virulence characterization of *Listeria monocytogenes*," *Listeria monocytogenes*, 1157, pp. 157–165. doi: 10.1007/978-1-4939-0703-8.
- Reynolds, C. S. (2006) The ecology of phytoplankton, *The Ecology of Phytoplankton*. doi: 10.1017/CBO9780511542145.
- Rimoldi, S. G. et al. (2020) "Presence and infectivity of SARS-CoV-2 virus in wastewaters and rivers," *Science of the Total Environment*. Elsevier B.V., 744, p. 140911. doi: 10.1016/j.scitotenv.2020.140911.
- Ríos, S., Agudelo, R. M. & Gutiérrez, L. A. (2017) "Patógenos e indicadores microbiológicos de calidad del agua para consumo humano," *Revista Facultad Nacional de Salud Pública*, 35(2), pp. 236–247. doi: 10.17533/udea.rfnsp.v35n2a08.
- Robillot, C. et al. (2000) "Hepatotoxin production kinetics of the cyanobacterium *Microcystis aeruginosa* PCC 7820, as determined by HPLC - mass spectrometry and protein phosphatase bioassay," *Environmental Science and Technology*. doi: 10.1021/es991294v.
- Rognes, T. et al. (2016) "VSEARCH: A versatile open source tool for metagenomics," *PeerJ*, 2016(10), pp. 1–22. doi: 10.7717/peerj.2584.
- Sathicq, M. B., Bauer, D. E. & Gómez, N. (2015) "Influence of El Niño Southern Oscillation phenomenon on coastal phytoplankton in a mixohaline ecosystem on the southeastern of South America: Río de la Plata estuary," *Marine Pollution Bulletin*, 98(1–2), pp. 26–33. doi: 10.1016/j.marpolbul.2015.07.017.
- Scheffer, M. et al. (2003) "Why plankton communities have no equilibrium: Solutions to the paradox," in *Hydrobiologia*. doi: 10.1023/A:1024404804748.
- Scheut, F. (2014) "Taxonomy Meets Public Health: The Case of Shiga Toxin-Producing *Escherichia coli*," *Microbiology Spectrum*. doi: 10.1128/microbiolspec.ehec-0019-2013.
- Semenza, J. C. (2020) "Cascading risks of waterborne diseases from climate change," *Nature Immunology*. Springer US, 21(5), pp. 484–487. doi: 10.1038/s41590-020-0631-7.
- Sen, K. & Rodgers, M. (2004) "Distribution of six virulence factors in *Aeromonas* species isolated from US drinking water utilities: A PCR identification," *Journal of Applied Microbiology*, 97(5), pp. 1077–1086. doi: 10.1111/j.1365-2672.2004.02398.x.
- Seo, S. B. et al. (2013) "Single nucleotide polymorphism typing with massively parallel sequencing for human identification," *International Journal of Legal Medicine*, 127(6), pp. 1079–1086. doi: 10.1007/s00414-013-0879-7.
- Shakila, R. J. et al. (2013) "Prevalence of Hemolytic and Enterotoxigenic *Aeromonas* spp. in Healthy and Diseased Freshwater Food Fishes as Assessed by Multiplex PCR," (January). doi: 10.7726/ajafst.2013.1007.
- Shapiro-Ilan, D. I. et al. (2005) "Definitions of pathogenicity and virulence in invertebrate pathology," *Journal of Invertebrate Pathology*. doi: 10.1016/j.jip.2004.10.003.
- Sienra, D & Ferrari, G. (2006) "Monitoreo de cianobacterias en la costa de Montevideo (Uruguay)". En: Menafrá, R., Rodríguez-Gallego, L., Scarabino, F. y Conde, D. (eds.). Bases para la conservación y el manejo de la costa uruguaya. *Montevideo: Vida Silvestre*. pp.413-420. ISBN: 9974-7589-2-0.
- Simoens, M. (2009) "Microcistina – LR en el Uruguay". En: Resúmenes del 1º Encuentro Uruguayo. Cianobacterias: del conocimiento a la gestión. Montevideo: LATU; Sección Limnología, Facultad de Ciencias, Universidad de la República.
- Sipari, H. et al. (2010) "Development of a chip assay and quantitative PCR for detecting microcystin synthetase e gene expressions," *Applied and Environmental Microbiology*, 76(12), pp. 3797–3805. doi: 10.1128/AEM.00452-10.
- Smith, F. M. J. et al. (2012) "Survey of *Scytonema* (Cyanobacteria) and associated saxitoxins in the littoral zone of recreational lakes in Canterbury, New Zealand," *Phycologia*, 51(5), pp. 542–551. doi: 10.2216/11-84.1.

- Sobiah, R. *et al.* (2018) "Implications of Targeted Next Generation Sequencing in Forensic Science," *Journal of Forensic Research*, 09(02), pp. 1–8. doi: 10.4172/2157-7145.1000416.
- Søborg, D. A. *et al.* (2016) "Bacterial human virulence genes across diverse habitats as assessed by in silico analysis of environmental metagenomes," *Frontiers in Microbiology*, 7(NOV). doi: 10.3389/fmicb.2016.01712.
- Soheili, S. *et al.* (2014) "Wide Distribution of Virulence Genes among *Enterococcus faecium* and *Enterococcus faecalis* Clinical Isolates," *The Scientific World Journal*, 2014, pp. 1–6. doi: 10.1155/2014/623174.
- Sohier, D. *et al.* (2014) "Evolution of microbiological analytical methods for dairy industry needs," *Frontiers in Microbiology*. doi: 10.3389/fmicb.2014.00016.
- Steinberg-Shemer, O. & Tamary, H. (2020) "Impact of Next-Generation Sequencing on the Diagnosis and Treatment of Congenital Anemias," *Molecular Diagnosis and Therapy*. Springer International Publishing, 24(4), pp. 397–407. doi: 10.1007/s40291-020-00478-3.
- Stenhoff, A. *et al.* (2020) "Quadrangle Jo ur na l P re," *Journal of Contextual Behavioral Science*. Elsevier Inc., p. 105398. doi: 10.1016/j.celrep.2020.108352.
- Stepińska, M. & Trafny, E. A. (2008) "Diverse type III secretion phenotypes among *Pseudomonas aeruginosa* strains upon infection of murine macrophage-like and endothelial cell lines," *Microbial Pathogenesis*, 44(5), pp. 448–458. doi: 10.1016/j.micpath.2007.11.008.
- Stevenson, K. (2015) "Genetic diversity of *Mycobacterium avium* subspecies paratuberculosis and the influence of strain type on infection and pathogenesis: A review Modeling Johne's disease: From the inside out Dr Ad Koets and Prof Yrjo Grohn," *Veterinary Research*, 46(1), pp. 1–13. doi: 10.1186/s13567-015-0203-2.
- Stewart, J. R. *et al.* (2008) "The coastal environment and human health: Microbial indicators, pathogens, sentinels and reservoirs," in *Environmental Health: A Global Access Science Source*. doi: 10.1186/1476-069X-7-S2-S3.
- Stüken, A. & Jakobsen, K. S. (2010) "The cylindrospermopsin gene cluster of *Aphanizomenon* sp. strain 10E6: Organization and recombination," *Microbiology*, 156(8), pp. 2438–2451. doi: 10.1099/mic.0.036988-0.
- Symonds, E. M., Griffin, D. W. & Breitbart, M. (2009) "Eukaryotic viruses in wastewater samples from the United States," *Applied and Environmental Microbiology*. doi: 10.1128/AEM.01899-08.
- Tan, B. *et al.* (2015) "Next-generation sequencing (NGS) for assessment of microbial water quality: Current progress, challenges, and future opportunities," *Frontiers in Microbiology*, 6(SEP). doi: 10.3389/fmicb.2015.01027.
- Tooming-Klunderud, A. *et al.* (2008) "The mosaic structure of the *mcyABC* operon in *Microcystis*," *Microbiology*, 154(7), pp. 1886–1899. doi: 10.1099/mic.0.2007/015875-0.
- Tourlousse, D. M. *et al.* (2007) "Virulence Factor Activity Relationships: Challenges and Development Approaches," *Water Environment Research*, 79(3), pp. 246–259. doi: 10.2175/106143007x156826.
- Tu, Q. *et al.* (2014) "GeoChip 4: A functional gene-array-based high-throughput environmental technology for microbial community analysis," *Molecular Ecology Resources*. doi: 10.1111/1755-0998.12239.
- Turki, Y. *et al.* (2014) "Molecular typing, antibiotic resistance, virulence gene and biofilm formation of different *Salmonella enterica* serotypes," *The Journal of General and Applied Microbiology*, 60(4), pp. 123–130. doi: 10.2323/jgam.60.123.
- Uchiya, K. *et al.* (2013) "Comparative Genome Analysis of *Mycobacterium avium* Revealed Genetic Diversity in Strains that Cause Pulmonary and Disseminated Disease," *PLoS ONE*, 8(8). doi: 10.1371/journal.pone.0071831.
- Umpiérrez, A. *et al.* (2017) "Zoonotic potential and antibiotic resistance of *Escherichia coli* in neonatal calves in Uruguay," *Microbes and Environments*, 32(3), pp. 275–282. doi: 10.1264/jsme2.ME17046.

UNESCO-Programa Hidrológico Internacional. (2016) Participation of the International Hydrological Programme in the Implementation and Monitoring of Water Related Sustainable Development Goals, p. www.unesco.org/water/ihp/index_es.shtml.

United Nations General Assembly (UNGA), (2010) The Human Right to Water and Sanitation: Resolution/Adopted by the General Assembly.UN Doc.64/292.

United Nations Sustainable Development (UNSD), (2015) Transforming Our World: The 2030 Agenda for Sustainable Development. United Nations General Assembly Resolution. A/RES/70/1.

U.S. Department of Agriculture/Food Safety & Inspection Service (USDA/FSIS) and U.S. Environmental Protection Agency (EPA) (2012) "Microbial Risk Assessment Guideline: Pathogenic Microorganisms with Focus on Food and Water, " USDA/FSIS and EPA Ed.; FSIS Publication; EPA Publication: USA, 128 p.

Valat, C. *et al.* (2014) "Assessment of adhesins as an indicator of pathovar-associated virulence factors in bovine *Escherichia coli*," *Applied and Environmental Microbiology*, 80(23), pp. 7230–7234. doi: 10.1128/AEM.02365-14.

Vico, P. *et al.* (2020) "Biogeography of the cyanobacterium *Raphidiopsis* (*Cylindrospermopsis*) *raciborskii*: Integrating genomics, phylogenetic and toxicity data," *Molecular Phylogenetics and Evolution*. Elsevier, 148(April), p. 106824. doi: 10.1016/j.ympev.2020.106824.

Vidal, L., Britos, A. & Gravier, A. (2009) "Cianobacterias planctónicas en aguas superficiales destinadas para consumo humano en Uruguay: taxonomía, distribución y toxicidad," En: Resúmenes del 1º Encuentro Uruguayo. Cianobacterias: del conocimiento a la gestión. Montevideo: LATU; Sección Limnología, Facultad de Ciencias, Universidad de la República.

Vidal, L. & Britos, A. (2012) "Uruguay: occurrence, toxicity and regulation of Cyanobacteria," Current approaches to Cyanotoxin risk assessment, risk management and regulations in different countries, p. 130.

Vidal, L. & Kruk, C. (2008) "*Cylindrospermopsis raciborskii* (Cyanobacteria) extends its distribution to Latitude 34°53'S: Taxonomical and ecological features in Uruguayan eutrophic lakes," *Pan-American Journal of Aquatic Sciences*.

Vierheilig, J. *et al.* (2015) "Potential applications of next generation DNA sequencing of 16S rRNA gene amplicons in microbial water quality monitoring," *Water Science and Technology*, 72(11), pp. 1962–1972. doi: 10.2166/wst.2015.407.

Vollenweider, R. A. (1976) "Advances in defining critical loading levels for phosphorus in lake eutrophication," *Memorie dell'Istituto italiano di idrobiologia dott. Marco De Marchi*.

Wagner, P. L., Acheson, D. W. K. & Waldor, M. K. (1999) "Isogenic lysogens of diverse Shiga toxin 2-encoding bacteriophages produce markedly different amounts of Shiga toxin," *Infection and Immunity*. doi: 10.1128/iai.67.12.6710-6714.1999.

Ward, A. D. & Trimble, S. W. (2003) "Environmental Hydrology". (Lewis,P., Ed.). CRC Press. 695p.

Ward, D. & Cohan, F. (2006) "Microbial Diversity in Hot Spring Cyanobacterial Mats: Pattern and Prediction," *Geothermal Biology and Geochemistry in Yellowstone National Park*.

Wassenaar, T. M. & Gaastra, W. (2001) "Bacterial virulence: Can we draw the line?," *FEMS Microbiology Letters*, 201(1), pp. 1–7. doi: 10.1016/S0378-1097(01)00241-5.

Wernike, K., Hoffmann, B. & Beer, M. (2015) "Simultaneous detection of five notifiable viral diseases of cattle by single-tube multiplex real-time RT-PCR," *Journal of Virological Methods*. Elsevier B.V., 217, pp. 28–35. doi: 10.1016/j.jviromet.2015.02.023.

White, G., D. Bradley & A. White (1972). *Drawers of Water*. University of Chicago Press, Chicago, IL, USA.

Whitton, B.A. & Potts, M. (2000) "The biology and ecology of cyanobacteria,". 669p.

Wiese, M. *et al.* (2010) "Neurotoxic alkaloids: Saxitoxin and its analogs," *Marine Drugs*, 8(7), pp. 2185–2211. doi: 10.3390/md8072185.

- Wilson, B. A. & Ho, M. (2013) "Pasteurella multocida: From Zoonosis to cellular microbiology," *Clinical Microbiology Reviews*, 26(3), pp. 631–655. doi: 10.1128/CMR.00024-13.
- Wingender, J. & Flemming, H. C. (2011) "Biofilms in drinking water and their role as reservoir for pathogens," *International Journal of Hygiene and Environmental Health*. Elsevier GmbH., 214(6), pp. 417–423. doi: 10.1016/j.ijheh.2011.05.009.
- Woolhouse, M. E. J. (2006) "Where do emerging pathogens come from?," *Microbe*. doi: 10.1128/microbe.1.511.1.
- World Health Organization (2017) Guidelines for Drinking-water Quality. 4th ed. Geneva: World Health Organization; Licence: CC BY-NC-SA 3.0 IGO. Available at: <https://apps.who.int/iris/bitstream/handle/10665/254636/9789241550017-eng.pdf>.
- World Health Organization and the United Nations Children's Fund (WHO/UNICEF) (2017). Progress on Drinking Water, Sanitation and Hygiene: Update and Sustainable Development Goal Baselines. World Health Organization (WHO) and the United Nations Children's Fund (UNICEF), Geneva. License: CC BY-NC-SA 3.0 IGO.
- Wu, H. J., Wang, A. H. J. & Jennings, M. P. (2008) "Discovery of virulence factors of pathogenic bacteria," *Current Opinion in Chemical Biology*. doi: 10.1016/j.cbpa.2008.01.023.
- Xu, H. S. et al. (1982) "Survival and viability of nonculturable Escherichia coli and Vibrio cholerae in the estuarine and marine environment," *Microbial Ecology*. doi: 10.1007/BF02010671.
- Yang, S. & Rothman, R. E. (2004) "PCR-based diagnostics for infectious diseases: uses, limitations, and future applications in acute-care settings," *The Lancet Infectious Diseases*, 4(6), pp. 337–348. doi: 10.1016/S1473-3099(04)01044-8.
- Yang, X. et al. (2020) "Genetic diversity of the intimin gene (eae) in non-O157 Shiga toxin-producing Escherichia coli strains in China," *Scientific Reports*. Springer US, 10(1), pp. 1–9. doi: 10.1038/s41598-020-60225-w.
- Ye, L. & Zhang, T. (2011) "Pathogenic bacteria in sewage treatment plants as revealed by 454 pyrosequencing," *Environmental Science and Technology*. doi: 10.1021/es201045e.
- Yoder, J. S. et al. (2011) Surveillance for waterborne disease and outbreaks associated with recreational water use and other aquatic facility- associated health events - United States, 2007- 2008. *The Morbidity and Mortality Weekly Report (CDC)*. 60:1–38.
- Yousefi-Avarvand, A. et al. (2015) "The Frequency of Exotoxin A and Exoenzymes S and U Genes Among Clinical Isolates of Pseudomonas aeruginosa in Shiraz, Iran," (1).
- Zakowska, D. et al. (2012) "New aspects of the infection mechanisms of bacillus anthracis," *Annals of Agricultural and Environmental Medicine*, 19(4), pp. 613–618.
- Zhang, S. et al. (2017) "Developmental validation of a custom panel including 273 SNPs for forensic application using Ion Torrent PGM," *Forensic Science International: Genetics*. Elsevier Ireland Ltd, 27, pp. 50–57. doi: 10.1016/j.fsigen.2016.12.003.
- Zhang, W. L. et al. (2002) "Genetic diversity of intimin genes of attaching and effacing Escherichia coli strains," *Journal of Clinical Microbiology*. doi: 10.1128/JCM.40.12.4486-4492.2002.

9. ANEXO I. MATERIAL ADICIONAL

Tabla S1. Aislamientos bacterianos utilizados en los ensayos experimentales. Procedencia, origen del aislamiento, información genética y referencias de las cepas.

	Número/Código del aislamiento	Colección utilizada	Origen	Genes/Referencia
Ensayo I	<i>Escherichia coli</i> PL74.1	Dpto. de Microbiología, IIBCE.	Materia Fecal Terneros	Cepa STEC, <i>stx1+</i> / <i>ea+</i> (Umpiérrez et al., 2017)
	<i>Listeria monocytogenes</i> 488	Dpto. de Bacteriología, Instituto de Higiene, Facultad de Medicina, Udelar.	Alimentos	<i>hlyA+</i> / <i>inlA+</i> Aislamiento utilizados para el control rutinario de la especie (Braga et al., 2017)
	<i>Pseudomonas aeruginosa</i> ATCC® 27853™	Dpto. de Ciencia y Tecnología de alimentos, Facultad de Veterinaria, Udelar	American Type Culture Collection hemocultivo	<i>exotA+</i> / <i>exoS+</i> / <i>exoU-</i> (Stepińska & Trafny, 2008; Yousefi-Avarvand et al., 2015)
	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar. Typhimurium ATCC® 14028™	Dpto. de Microbiología, IIBCE.	American Type Culture Collection Tejido animal (pollo)	<i>invA</i> + (Rahn et al., 1992; Alzwghaibi et al., 2018)
Ensayo II	<i>Aeromonas</i> sp. 8 Aislamientos (B1, B2, H1, H2, H3, H6, H8, H9)	Instituto de Investigaciones Pesqueras, Facultad de Veterinaria, Udelar.	Infección de Peces	<i>aerA+</i> / <i>alt+</i> / <i>ast+</i> / <i>ela+</i> / <i>lip+</i> Especies <i>A. hydrophila</i> y <i>A. bestiarium</i> (Perretta, Antúnez & Zunino, 2018)
	<i>Escherichia coli</i> 7 Aislamientos (4.1, 8.1, 42.3, 74.1, P.AG2, AG3, AF26)	Dpto. de Microbiología, IIBCE.	Materia Fecal Terneros	<i>stx1+</i> / <i>ea+</i> / <i>fl17+</i> / <i>est+</i> / <i>elt+</i> otros genes sin determinar (Umpiérrez et al., 2017)
	<i>Enterococcus</i> spp. 49 Aislamientos numeración consecutiva (En2 al En50)	Dpto. de Ciencia y Tecnología de alimentos, Facultad de Veterinaria, Udelar	Productos lácteos artesanales	Sin caracterizar factores de virulencia Especies <i>E. faecalis</i> , <i>E. faecium</i> , <i>E. durans</i> (Hirigoyen et al., 2018)
	<i>Staphylococcus aureus</i> 23 Aislamientos numeración consecutiva (SA1 al SA23)	Dpto. de Ciencia y Tecnología de alimentos, Facultad de Veterinaria, Udelar	Leche, mastitis subclínica bovina	<i>coa+</i> otros genes sin determinar (De Los Santos et al., 2014)
	<i>Cylindrospermopsis raciborskii</i> 1 aislamiento, MVCC19	Dpto. de Microbiología, IIBCE.	Lago Javier, Uruguay	<i>sxt+</i> (Piccini et al., 2011, 2013; Vico et al., 2020)
	<i>Pseudomonas aeruginosa</i> 1 aislamiento, ATCC® 27853™	Dpto. de Ciencia y Tecnología de alimentos, Facultad de Veterinaria, Udelar	American Type Culture Collection hemocultivo	<i>exotA+</i> / <i>exoS+</i> / <i>exoU-</i> (Stepińska & Trafny, 2008; Yousefi-Avarvand et al., 2015)
	<i>Listeria monocytogenes</i> 2 Aislamientos (888, H43)	Dpto. de Bacteriología, Instituto de Higiene, Facultad de Medicina, Udelar	Alimentos, Humano (hemocultivo/LCR*)	<i>hlyA+</i> / <i>inlA+</i> Aislamientos utilizados para el control rutinario de la especie (Braga et al., 2017)
	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar. Typhimurium 1 aislamiento, ATCC® 14028™	Dpto. de Microbiología, IIBCE.	American Type Culture Collection Tejido animal (pollo)	<i>invA</i> + (Rahn et al., 1992; Alzwghaibi et al., 2018)

* Líquido Cefalorraquídeo

Tabla S 2. Genes de virulencia con región conservada (n=39). Genomas de referencia utilizados, código de referencia de NCBI, localización, sentido y tamaño de la región conservada.

Gen	Genoma de referencia	Referencia NCBI	Coordenada genómica de la región conservada	Sentido	Región conservada (pb)
ace	<i>Enterococcus faecalis</i> ATCC 700802	NC_004668.1	1068036-1069118	+	2024
acm	<i>Enterococcus faecium</i> ATCC BAA-472	NC_017960.1	2228120-2230000	-	2165
anaC	<i>Planktothrix rubescens</i> NIVA-CYA	NZ_KE734717.1	829419-830942	-	7723
ast	<i>Aeromonas hydrophila</i> subsp. <i>hydrophila</i> ATCC 7966	NC_008570.1	861796-863703	-	1910
bfpA	<i>Escherichia coli</i> 0127:H6 E2348/69 plasmid pMAR2	NC_011603.1	2646-2971	+	581
bfpB	<i>Escherichia coli</i> 0127:H6 E2348/69 plasmid pMAR2	NC_011603.1	4248-5337	+	1658
capA	<i>Bacillus anthracis</i> strain Larissa	CP012521.1	53882-55114	-	1235
capB	<i>Bacillus anthracis</i> strain Larissa	CP012521.1	55593-56783	-	1193
cesA	<i>Bacillus cereus</i> strain F1-15scaffold1 whole genome shotgun	NZ_KN049962.1	616611-626783	+	10124
cesB	<i>Bacillus cereus</i> strain F1-15scaffold1 whole genome shotgun	NZ_KN049962.1	626800-634842	+	8045
cesC	<i>Bacillus cereus</i> strain F1-15scaffold1 whole genome shotgun	NZ_KN049962.1	635011-635883	+	875
cesD	<i>Bacillus cereus</i> strain F1-15scaffold1 whole genome shotgun	NZ_KN049962.1	635903-636706	+	806
coa	<i>Staphylococcus aureus</i> subsp. <i>aureus</i> Rosenbach ATCC 700698	AP009324.1	265857-266157	+	1976
ctxA	<i>Vibrio cholerae</i> ATCC 39315	NC_002505.1	1567341-1568114	-	776
cya	<i>Bacillus cereus</i> strain 03BB87	CP009941.1	1288774-1290198	+	2402
cyrA	<i>Cylindrospermopsis raciborskii</i> CS-505	NZ_LYXA01000001.1	1402336-1403313	-	1175
cyrB	<i>Cylindrospermopsis raciborskii</i> CS-505	NZ_LYXA01000001.1	1398531-1400235	-	8753
cytK	<i>Bacillus cereus</i> ATCC 14579	NC_004722.1	1091473-1092459	-	1010
eaeA	<i>Escherichia coli</i> O157:H7 str. Sakai	NC_002695.1	4597751-4598825	-	2804
est2	<i>Escherichia coli</i> F18+ strain EC2173 plasmid pTC1	CP000913.1	5057-5269	-	213
exotA	<i>Pseudomonas aeruginosa</i> ATCC 15692	NC_002516.2	1241346-1242500	-	1916
exoU	<i>Pseudomonas aeruginosa</i> NCGM 1984	NZ_AP014646.1	4738700-4740760	-	2063

Continúa→

Tabla S2 (continuación). Genes de virulencia con región conservada (n=39). Genomas de referencia utilizados, código de referencia de NCBI, localización, sentido y tamaño de la región conservada.

Gen	Genoma de referencia	Referencia NCBI	Coordenada genómica de la región conservada	Sentido	Región conservada (pb)
f41A	<i>Escherichia coli</i> strain 08-00022 plasmid pCFSAN004179G	CP012501.1	102222-103052	+	833
hyl	<i>Enterococcus faecium</i> ATCC BAA-472	NC_017960.1	145590-147248	-	1661
inlA	<i>Listeria monocytogenes</i> ATCC BAA-679	NC_003210.1	454912-456345	+	2402
invA	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Typhi str. CT18	NC_003198.1	2891939-2892294	-	2057
lef	<i>Bacillus cereus</i> strain 03BB87	CP009941.1	1314,460-1316058	-	2429
nheB	<i>Bacillus cereus</i> ATCC 10987, complete genome	AE017194.1	1903949-1905049	+	1101
plc	<i>Aeromonas veronii</i> strain TH0426	CP012504.1	540439-542154	-	1718
sxtA	<i>Aphanizomenon gracile</i> NIVA-CYA 676	LT549449.1	5561-7745	-	3704
sxtU	<i>Aphanizomenon gracile</i> NIVA-CYA 676	LT549449.1	18108-18849	+	749
toxA	<i>Pasteurella multocida</i> subsp. <i>multocida</i> str. HN06	NC_017027.1	2223481-2227095	+	3857
tst	<i>Staphylococcus aureus</i> subsp. <i>aureus</i> Rosenbach ATCC 700698	AP009324.1	2139151-2139834	+	704
ureA	<i>Helicobacter pylori</i> ATCC 700392	NC_000915.1	77243-77956	-	716
ureI	<i>Helicobacter pylori</i> ATCC 700392	NC_000915.1	74750-75241	-	587
vacA	<i>Helicobacter pylori</i> ATCC 700392	NC_000915.1	941019-941523	+	3872
yadA	<i>Yersinia enterocolitica</i> subsp. <i>enterocolitica</i> strain 8081	NC_008791.1	47639-48134	-	1268
ye-invA	<i>Yersinia enterocolitica</i> subsp. <i>enterocolitica</i> strain 8081	NC_008800.1	2756136-2757799	-	2507
yopH	<i>Yersinia enterocolitica</i> subsp. <i>enterocolitica</i> strain 8081	NC_008791.1	36049-37452	+	1406

Tabla S 3. Genes de virulencia sin región conservada (n=33). Genomas de referencia utilizados, código de referencia de NCBI, localización, sentido y tamaño del gen.

Gen	Genoma Referencia	Referencia NCBI	Coordenada genómica del gen	Sentido	Tamaño del gen (pb)
aerA	<i>Aeromonas hydrophila</i> subsp. <i>hydrophila</i> ATCC 7966	NC_008570.1	466510-467991	+	1482
alt	<i>Aeromonas hydrophila</i> subsp. <i>hydrophila</i> ATCC 7966	NC_008570.1	115445-117862	+	2418
anaF	<i>Anabaena</i> sp. WA102	NZ_CP011456.1	4375269-4380890	-	5622
capC	<i>Bacillus anthracis</i> strain Larissa	CP012521.1	55126-55575	-	450
ctxB	<i>Vibrio cholerae</i> ATCC 39315	NC_002505.1	1566967-1567341	-	375
ela	<i>Aeromonas hydrophila</i> subsp. <i>hydrophila</i> ATCC 7966	CP000462.1	913092-914861	+	1770
eltA	<i>Escherichia coli</i> ETEC H10407 plasmid pEntH10407	AP010910.1	21719-22495	-	777
eltB	<i>Escherichia coli</i> ETEC H10407 plasmid pEntH10407	AP010910.1	21348-21722	-	375
est1	<i>Escherichia coli</i> F18+ strain EC2173 plasmid pTC1	CP000913.1	14840-15058	+	219
exoS	<i>Pseudomonas aeruginosa</i> strain X78812	NZ_CP008872.2	1199811-1201172	+	1362
f17A	<i>Escherichia coli</i> Vir68 plasmid pVir68	CP001162.1	9749-10297	+	549
fbpA	<i>Mycobacterium tuberculosis</i> ATCC 25618	NC_000962.3	4265642-4266658	-	1011
fbpB	<i>Mycobacterium tuberculosis</i> ATCC 25618	NC_000962.3	2134890-2135867	-	978
fbpC	<i>Mycobacterium tuberculosis</i> ATCC 25618	NC_000962.3	156578-157600	-	1023
fnbA	<i>Staphylococcus aureus</i> subsp. <i>aureus</i> Rosenbach ATCC 700698	AP009324.1	2643954-2647070	-	3117
hblA	<i>Bacillus cereus</i> ATCC 14579	NC_004722.1	3062634-3063761	-	1128
hblC	<i>Bacillus cereus</i> ATCC 14579	NC_004722.1	3065080-3066399	-	1320
hblD	<i>Bacillus cereus</i> ATCC 14579	NC_004722.1	3063798-3065018	-	1221
hlyA	<i>Listeria monocytogenes</i> ATCC BAA-679	NC_003210.1	205819-207408	+	1590
lprG	<i>Mycobacterium tuberculosis</i> ATCC 25618	NC_000962.3	1587772-1588482	-	711
mcyE	<i>Microcystis aeruginosa</i> PCC 7806SL chromosome	NZ_CP020771.1	3037886-3048349	-	10464
mcyJ	<i>Microcystis aeruginosa</i> PCC 7806SL chromosome	NZ_CP020771.1	3025168-3026103	-	936
ndaI	<i>Nodularia spumigena</i> CCY9414	NZ_CP007203.1	4293700-4295505	+	1806

Continúa→

Tabla S3 (continuación). Genes de virulencia sin región conservada (n=33). Genomas de referencia utilizados, código de referencia de NCBI, localización, sentido y tamaño del gen.

Gen	Genoma Referencia	Referencia NCBI	Coordenada genómica del gen	Sentido	Tamaño del gen (pb)
<i>ndal</i>	<i>Nodularia spumigena</i> CCY9414	NZ_CP007203.1	4293700-4295505	+	1806
<i>nheA</i>	<i>Bacillus cereus</i> ATCC 10987	NC_003909.8	1902733-1903893	+	1161
<i>nheC</i>	<i>Bacillus cereus</i> ATCC 10987	NC_003909.8	1905394-1906473	+	1080
<i>pag</i>	<i>Bacillus anthracis</i> str. Ames 0581	NC_007322.2	143779-146073	+	2295
<i>pilB</i>	<i>Enterococcus faecium</i> ATCC BAA-472	NC_017960.1	1376322-1378199	-	1878
<i>stx1A</i>	<i>Escherichia coli</i> O157:H7 str. Sakai	NC_002695.1	2924769-2925716	-	948
<i>stx1B</i>	<i>Escherichia coli</i> O157:H7 str. Sakai	NC_002695.1	2924490-2924759	-	270
<i>stx2A</i>	<i>Escherichia coli</i> O157:H7 str. Sakai	NC_002695.1	1266965-1267924	+	960
<i>stx2B</i>	<i>Escherichia coli</i> O157:H7 str. Sakai	NC_002695.1	1267936-1268205	+	270
<i>sxtI</i>	<i>Aphanizomenon gracile</i> NIVA-CYA 676	LT549449.1	23584-25422	+	1839
<i>ureB</i>	<i>Helicobacter pylori</i> ATCC 700392	NC_000915.1	75527-77236	-	1710

10. ANEXO II. TRABAJOS SUPLEMENTARIOS

El desarrollo del presente trabajo permitió el análisis desde otros abordajes que lo suplementan y de los cuales algunos resultados se presentan a continuación.

POTENCIAL PATOGENICO DEL AGUA: DIVERSIDAD FILOGENÉTICA VS. DIVERSIDAD DE ECOTIPOS

Álvaro González Revello^{1,3}, Claudia Piccini¹, Eliana Nervi¹, Andrés Iriarte², Pablo Zunino¹, y José Sotelo-Silveira³.

¹Departamento de Microbiología, Instituto de Investigaciones Biológicas Clemente Estable (IIBCE).

²Departamento de Desarrollo Biotecnológico, Instituto de Higiene, Facultad de Medicina, Udelar.

³Departamento de Genómica, Instituto de Investigaciones Biológicas Clemente Estable (IIBCE).

Objetivo

Utilizando la base de datos de secuencias generadas para la construcción del panel Fvir, nos planteamos explorar la variabilidad de cada gen de virulencia seleccionado entre las cepas/especies que los portan y compararla con la de genes marcadores tales como el gen del ARNr 16S.

Métodos

La base de datos previamente creada fue filtrada seleccionando aquellas secuencias provenientes de genomas completos (39 genes, Tabla S4). A partir de ella se dividieron las secuencias provenientes de factores de virulencia y del gen del ARNr 16S del mismo genoma. Las secuencias se alinearon (MUSCLE/SSU-ALIGN) y se creó una matriz de distancia (distmat/Emboss) calculando la distancia genética media corregida según el método de Jin-Nei Gamma (Jin & Nei, 1990).

Tabla S4. Genes seleccionados para el estudio con al menos un genoma de referencia disponible (n=39). Especie bacteriana, factor de virulencia, localización del gen y numero de genomas analizados.

Especie Bacteriana	Factor de Virulencia	Gen	Localización	N° de secuencias (Genomas)
<i>Aeromonas spp</i>	Aerolysin	<i>aerA</i>	cromosoma	16
	Enterotoxin	<i>ast</i>	cromosoma	12
	Elastase	<i>ela</i>	cromosoma	21
	Lipase A1	<i>pla</i>	cromosoma	20
<i>Bacillus anthracis</i>	Capsule biosynthesis protein	<i>capC</i>	plasmido	130
	Cytotoxin K	<i>cytK</i>	cromosoma	57
	Hemolysin bl-binding component	<i>hblA</i>	cromosoma	64
	Hemolysin BL lytic component L2	<i>hblC</i>	cromosoma	63
<i>Bacillus cereus</i>	Hemolysin BL lytic component L1	<i>hblD</i>	cromosoma	64
	Non-hemolytic enterotoxin A	<i>nheA</i>	cromosoma	133
	Non-hemolytic enterotoxin A	<i>nheB</i>	cromosoma	133
	Non-hemolytic enterotoxin A	<i>nheC</i>	cromosoma	131
	Anatoxin-a synthetase F	<i>anaF</i>	cromosoma	3
<i>Cyanobacteria</i> (<i>Anabaena sp.</i> , <i>C. raciborskii</i> , <i>Microcystis spp.</i> , <i>Plankthotrix sp.</i> , <i>Nostoc sp.</i> , <i>Nodularia spp.</i>)	Saxitoxin	<i>sxtI</i>	cromosoma	2
	Microcystin	<i>mcvE</i>	cromosoma	4
	Microcystin	<i>mcvJ</i>	cromosoma	4
	Nodularin ABC transporter ndai	<i>ndal</i>	cromosoma	2
<i>Enterococcus faecalis</i>	Collagen adhesin precursor	<i>ace</i>	cromosoma	10
<i>Enterococcus faecium</i>	Collagen-binding MSCRAMM	<i>acm</i>	cromosoma	14
	Pil B	<i>pilB</i>	cromosoma	29

Tabla S4 (Continuación). Genes seleccionados para el estudio con al menos un genoma de referencia disponible (n=39). Especie bacteriana, factor de virulencia, localización del gen y numero de genomas analizados.

Especie Bacteriana	Factor de Virulencia	Gen	Localización	Nº de secuencias (Genomas)
<i>Escherichia coli</i>	Intimin	<i>eaeA</i>	cromosoma	58
	Heat-labile (LT) enterotoxin	<i>eltA</i>	plasmido	11
	Heat-labile (LT) enterotoxin	<i>eltB</i>	plasmido	11
	Shiga toxin Stx1 subunit A	<i>stx1A</i>	fago	27
	Shiga toxin Stx1 subunit B	<i>stx1B</i>	fago	27
	Shiga toxin Stx2 subunit A	<i>stx2A</i>	fago	55
	Shiga toxin Stx2 subunit B	<i>stx2B</i>	fago	55
<i>Mycobacterium sp.</i>	Diacylglycerol acyltransferase/mycolyltransferase	<i>fbpA</i>	cromosoma	135
	Diacylglycerol acyltransferase/mycolyltransferase	<i>fbpB</i>	cromosoma	142
	Diacylglycerol acyltransferase/mycolyltransferase	<i>fbpC</i>	cromosoma	144
	Lipoarabinomannan carrier protein LprG	<i>lprG</i>	cromosoma	113
<i>Pseudomonas aeruginosa</i>	Exoenzyme S	<i>exoS</i>	cromosoma	93
	Exotoxin A	<i>exxA</i>	cromosoma	92
	ExoU protein	<i>exoU</i>	cromosoma	25
<i>Salmonella spp</i>	Type III secretion system protein InvA	<i>invA</i>	cromosoma	369
<i>Staphylococcus aureus</i>	Coagulase	<i>coa</i>	cromosoma	21
	fibronectin-binding protein A	<i>fnbA</i>	cromosoma	128
<i>Vibrio cholerae</i>	Heat-labile enterotoxin alpha chain	<i>ctxA</i>	fago	10
	Heat-labile enterotoxin betha chain	<i>ctxB</i>	fago	11

Resultados

Existió una correlación significativa entre la tasa de evolución filogenética y la de los genes de virulencia, lo que se indican en la Figura S1.

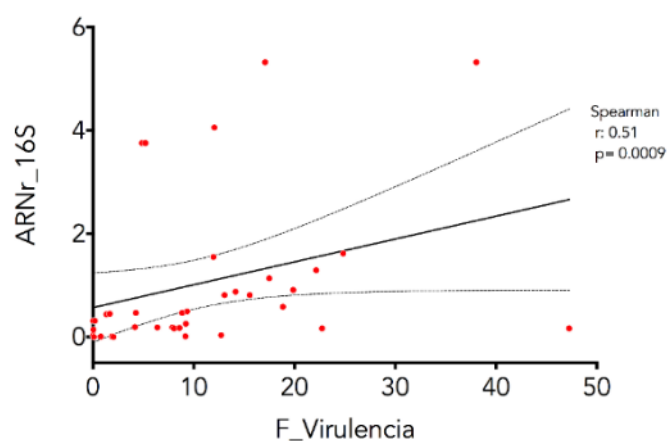


Figura S1. Relación de las distancias genéticas medias (número de sustituciones cada 100 bases) entre las secuencias del gen codificante del ARNr 16S y entre los factores de virulencia evaluados.

Existe una correlación significativa entre la tasa de evolución filogenética y la de los genes de virulencia, si bien la variabilidad de los factores de virulencia en su conjunto es mayor que la del gen para el ARNr 16S (Figura S2).

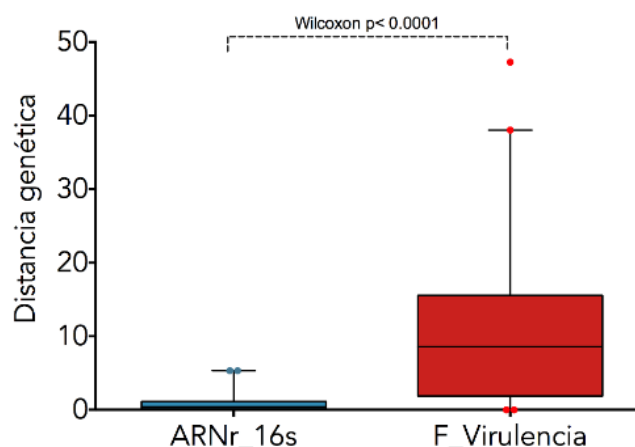


Figura S2. Variabilidad en las secuencias del gen para el ARNr 16S y de los genes de virulencia para todas las especies evaluadas.

Entre los genes de virulencia que mostraron mayor variabilidad genética se encuentran aquellos pertenecientes a cianobacterias, los que también presentan la mayor variabilidad en el gen del ARNr 16S en relación con el resto de los organismos estudiados (Figura S3 A). Entre ellos, los que codifican para enzimas involucradas en la síntesis de microcistinas como *mcyJ* mostraron mayor variación, concordando con estudios de nuestro grupo que sugieren que las diferencias en este gen reflejarían distintas preferencias ambientales y definirían ecotipos (Martínez de la Escalera *et al.*, 2019) (Figura S3 B).

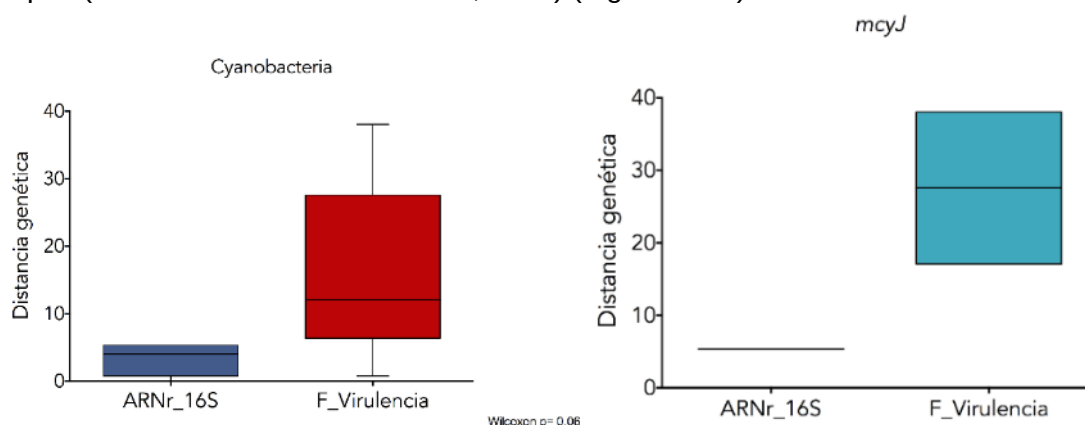


Figura S3. A (izquierda) Variabilidad genética del ARNr 16S y de los genes de toxinas en las especies de cianobacterias (todas las especies de cianobacterias evaluadas). B (derecha) Variabilidad genética del gen del ARNr 16S de las especies productoras de microcistina y entre los genes *mcyJ* de dichas especies.

En el caso de *E. coli*, los factores de virulencia estudiados presentaron una menor variabilidad genética que la observada para las cianobacterias, mientras que la variabilidad en el marcador filogenético fue similar (Figura S4).

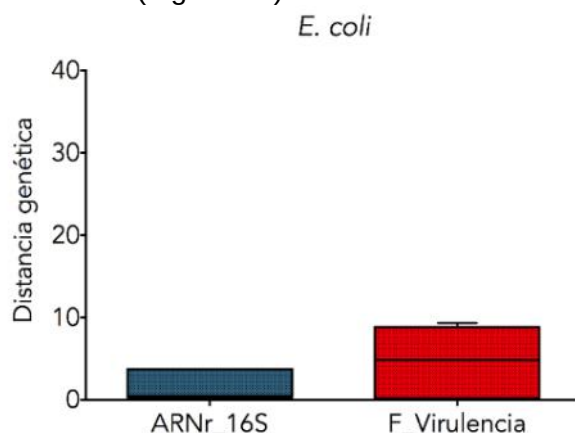
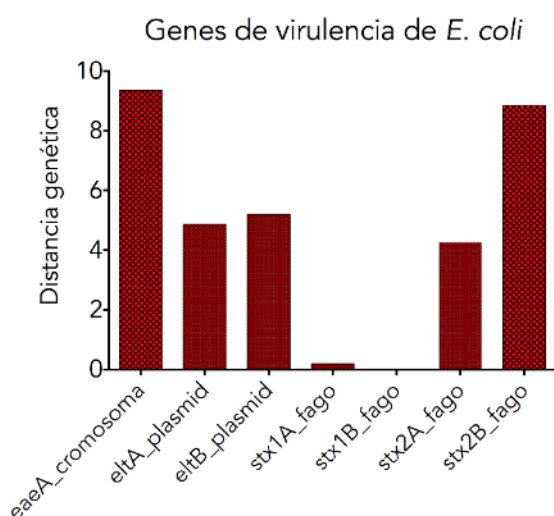


Figura S4. Variabilidad del gen para el ARNr 16S y de todos los genes de virulencia evaluados en *E. coli*.

La variabilidad de los factores de virulencia se asoció en algunos casos a la localización del gen (cromosomal o extracromosomal), Figura S5A, presentando mayor variabilidad los genes que codifican para una intimina (*eaeA*/cromosomal) y la toxina tipo Shiga 2 (*stx2b*/bacteriófago). Esto se refleja en la abundancia de subtipos (21 hasta el momento) que se han descrito para la toxina *Stx2* (Scheutz, 2014) (Figura S5B).



Subtipos y variantes de toxina tipo shiga descritos. Parsimony tree Scheutz, 2014.

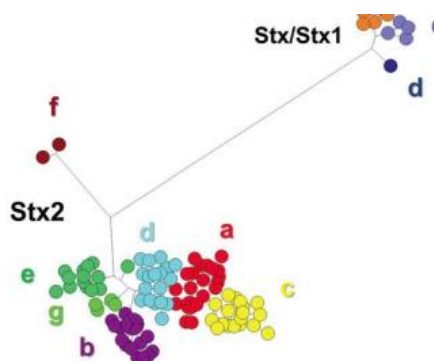


Figura S5. A (izquierda). Variabilidad de los factores de virulencia de *E. coli* según su localización (cromosomal/extracromosomal). Figura S5B (derecha). Subtipos y variantes de toxina tipo shiga descritos (árbol de parsimonia, Scheutz, 2014).

Las variaciones de las secuencias en los genes *eaeA* y *stx2B* se estudiaron en 55/58 genomas respectivamente. El árbol filogenético para el gen *eaeA* se ilustra en la Figura S6. Proponemos que la combinación de variantes de *eaeA* y *stx2* podrían definir ecotipos de virulencia en *E. coli*.

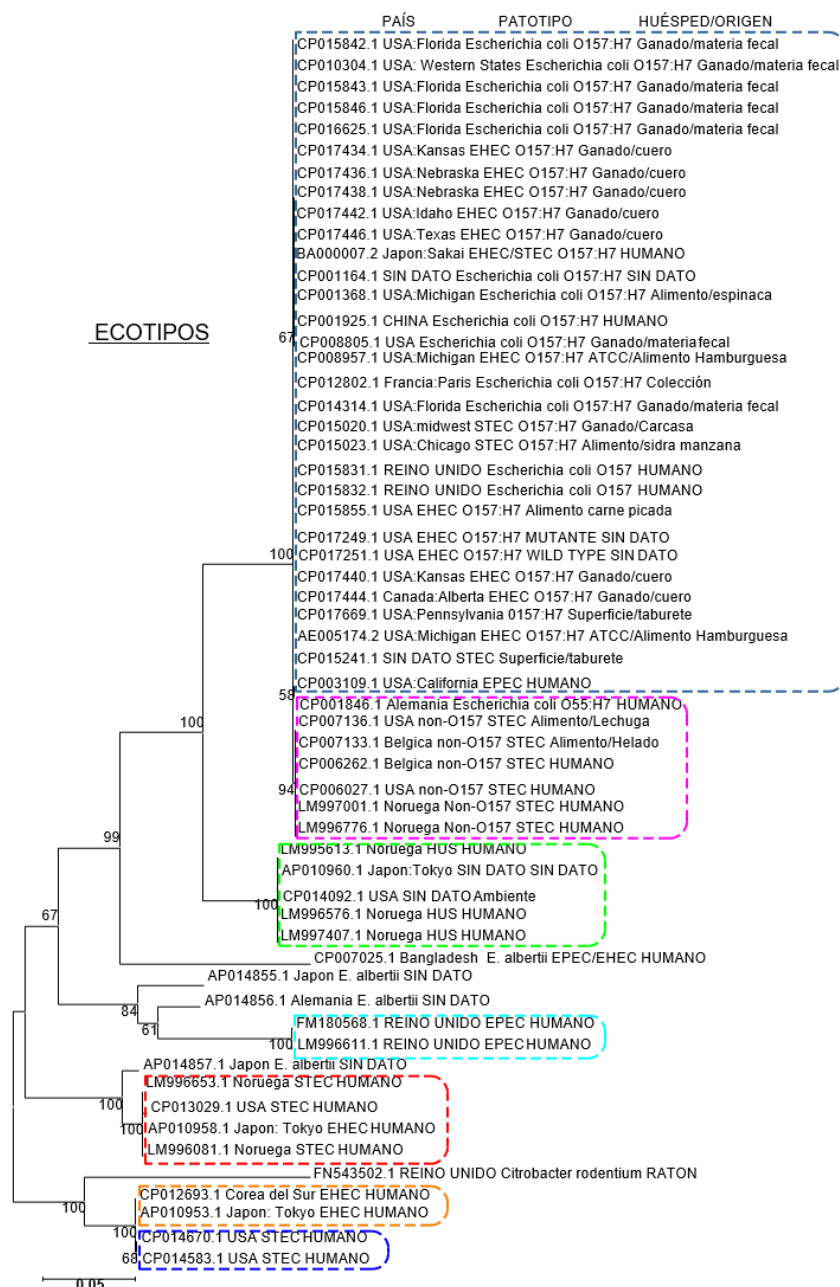


Figura S6. Árbol filogenético construido a partir de las secuencias alineadas del gen *eaeA* para todos los genomas analizados. Información del país de origen, patotipo y huésped del aislamiento. *Shiga toxin producing* (STEC), *Enterohemorrhagic* (EHEC), *Enteropathogenic* (EPEC), *Hemolytic Uremic Syndrome* (HUS).