

# Vigilancia de la seguridad alimentaria con MiSeq™ i100 Series

Detección precisa de patógenos de origen alimentario



Reducción del tiempo de participación activa con una preparación de librerías simplificada.



Menor duración del experimento de secuenciación y mayor rendimiento.



Montaje fiable de genomas bacterianos de alta calidad.

## Introducción

El uso de métodos de laboratorio para clasificar las bacterias entéricas, incluida *Salmonella*, ha sido fundamental para identificar posibles brotes y vincular las enfermedades causadas por bacterias con posibles fuentes del brote. A medida que aumenta la resolución de los métodos de subclasificación del laboratorio, también aumenta la precisión en la conexión de las enfermedades con una fuente. En los Estados Unidos, la vinculación de posibles casos de brote se realiza a través de PulseNet, una red nacional de vigilancia molecular para enfermedades entéricas de transmisión alimentaria, por el agua y relacionadas con One Health que se encuentran en más de 80 laboratorios estatales, locales y federales. La red PulseNet evita más de 250 000 enfermedades causadas por bacterias entéricas cada año y ahorra medio millón de dólares en costes médicos y pérdida de productividad.<sup>1,2</sup>

PulseNet comenzó a implementar la secuenciación del genoma completo (WGS, whole-genome sequencing) con instrumentos de Illumina a principios de la década de 2010 con socios federales de la Administración de Alimentos y Medicamentos (FDA, Food and Drug Administration) de los EE. UU. y el Servicio de Inspección y Seguridad Alimentaria del Departamento de Agricultura de los EE. UU. (USDA-FSIS, US Department of Agriculture Food Safety and Inspection Service).<sup>3</sup> Mediante la WGS, los científicos de salud pública pueden identificar posibles brotes identificando cepas de casos que son muy similares genéticamente y asociados a enfermedades que se producen en el mismo momento o región geográfica.

La WGS también se utiliza para identificar tendencias en cepas bacterianas que no se caracterizaron previamente por métodos moleculares de menor resolución.<sup>4,5</sup> Estas cepas se repitieron con el tiempo y provocaron brotes, surgieron y provocaron un aumento de la enfermedad o persistieron causando enfermedades durante meses o años. Los Centros para el Control y la Prevención de Enfermedades (CDC, Centers for Disease Control and Prevention) comenzaron a describir estas cepas como cepas recurrentes, emergentes y persistentes (REP, recurring, **emerging and persisting**). Las cepas REP se pueden identificar por su relación genética y otras características definitorias, incluida la resistencia a múltiples fármacos, la virulencia u otros marcadores genéticos. Dado que solo aproximadamente el 10 % de las bacterias causantes de enfermedades en PulseNet están asociadas a un brote, el uso de la WGS para identificar cepas REP puede proporcionar una oportunidad para prevenir más enfermedades al vincular cepas adicionales que pueden compartir una fuente común.

Esta nota de aplicación demuestra la detección y caracterización precisas de una cepa persistente de *Salmonella enterica* de serotipo Infantis multirresistente (REP JFX01) mediante un flujo de trabajo de NGS eficiente que integra Illumina DNA Prep, MiSeq i100 Series y herramientas de análisis de terceros (figura 1).



## Métodos

### Muestras

Los aislados bacterianos recogidos como parte de la vigilancia rutinaria de PulseNet se cultivaron en base de agar con sangre BBL (sin añadir sangre) (BD, n.º de catálogo 211037). El ADN genómico (ADNg) se extrajo con Wizard Genomic DNA Purification Kit (Promega, n.º de catálogo A1120) siguiendo un protocolo descrito anteriormente para el aislamiento de ADNg de bacterias gramnegativas.<sup>8</sup> El ADNg purificado se cuantificó con el kit de ensayo Qubit dsDNA BR (Thermo Fisher Scientific, n.º de catálogo Q32853) antes de la preparación de librerías.

### Preparación de librerías

Las librerías listas para secuenciación se prepararon manualmente a partir de ADNg extraído de 100 ng con Illumina DNA Prep (Illumina, n.º de catálogo 20060060) con modificaciones en el protocolo para mejorar el rendimiento en MiSeq i100 Series, incluida una ronda adicional de purificación basada en bolas con una relación de bolas a muestras de 0,5× para eliminar fragmentos cortos.<sup>7</sup> La calidad y concentración de las librerías amplificadas por PCR se evaluaron con el sistema 4200 TapeStation (Agilent Technologies, n.º de catálogo G2991BA) y Qubit 1X dsDNA High Sensitivity (HS) Assay Kit (Thermo Fisher Scientific, n.º de catálogo Q33231) antes de la agrupación.

### Secuenciación

Las librerías preparadas se agruparon por volumen siguiendo un protocolo descrito anteriormente,<sup>8</sup> con una ligera modificación que incluía una ronda adicional de purificación con bolas mediante inmovilización reversible en fase sólida (SPRI, solid-phase reversible immobilization) a 0,5× realizada después de la cuantificación y agrupación equimolar de cada librería. Las librerías agrupadas se diluyeron a una concentración de carga de 60 pM (32 librerías/proceso). La secuenciación se realizó en MiSeq i100 Plus System con una celda de flujo de 25M con una configuración de experimento de 2 × 301 pb. En el caso de estudios más grandes, los experimentos de secuenciación se pueden ampliar a NextSeq™ 1000 System, NextSeq 2000 System, NovaSeq™ 6000 System o NovaSeq X System.

### Análisis de datos

Una vez finalizada la secuenciación, los datos se transmitieron a BaseSpace™ Sequence Hub y se analizaron con la aplicación DRAGEN™ FastQC + MultiQC para proporcionar estadísticas y métricas generales, asignación de secuencias y alineación.

El análisis se realizó en la aplicación web PulseNet 2.0. Dentro de esta aplicación, se calcularon los criterios de medición de control de calidad (CC) basados en lectura y ensamblaje,<sup>5</sup> se realizó la caracterización de secuencias multilocus del genoma central (cgMLST, core genome multilocus sequence typing) con el llamador de alelos PulseNet 2.0,<sup>6</sup> se realizó el serotipado con SeqSero2,<sup>8</sup> se estableció el perfil de resistencia a los antimicrobianos (AMR, antimicrobial resistance) con AMRFinderPlus<sup>9</sup> y se identificaron plásmidos a partir de las bases de datos de PlasmidFinder.<sup>10</sup> El árbol de cgMLST y los genotipos asociados se visualizaron con la aplicación web iTOL v7.<sup>11</sup>

## Resultados

### Criterios de medición de secuenciación

Los datos de demultiplexado se generaron después de 20 ciclos con una duración total del experimento de 14 horas y 20 minutos, lo que confirma que las librerías se cuantificaron y agruparon correctamente. MiSeq i100 Plus System generó datos de secuenciación con un promedio de 1,88 millones de lecturas por muestra con una longitud de lectura de 2 × 301 pb con ≥97,5 % de bases por encima de Q30 (figura 2A). El porcentaje observado de contenido de GC de las lecturas siguió una distribución aproximadamente normal, lo que indica que la química de preparación de librerías y la secuenciación no tuvieron tendencia a inclinarse hacia regiones de GC más altas o más bajas (figura 2B).

Las 32 colonias bacterianas aisladas secuenciadas demostraron una calidad de secuenciación sistemáticamente alta, con una mediana de puntuación Q de 37,1 (tabla 1). La profundidad de cobertura estimada osciló entre 46,1× y 109,9×, con un promedio de 71,9×. El porcentaje de locus principales del esquema de cgMLST de *Salmonella* en PulseNet 2.0 fue superior al 99 % para todas las muestras, lo que indica una integridad del genoma suficiente para la agrupación mediante cgMLST.

### Predicción del serotipo de *Salmonella*

Los archivos FASTQ se analizaron con el software SeqSero2 para la predicción del serotipo de *Salmonella*. Treinta y una muestras se predijeron correctamente como *Salmonella infantis* (I 7:r:1,5) y una se predijo correctamente como *Salmonella enteritidis* (I 9:g,m:-) (figura 3).

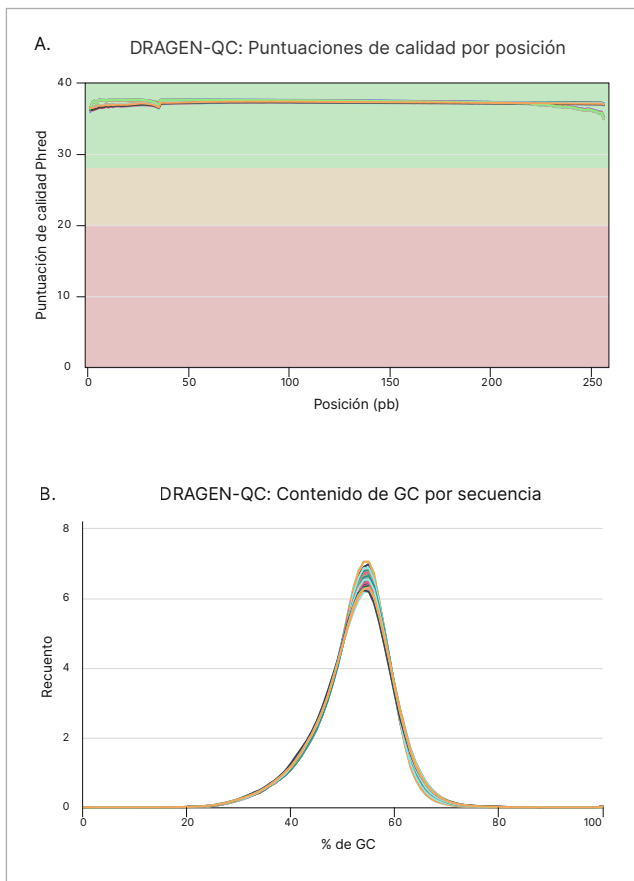
### Caracterización molecular de patógenos

El análisis de grupos de 3002 locus de cgMLST en el esquema de *Salmonella* de PulseNet 2.0 permitió deducir el grado de relación genómica entre las colonias aisladas en el conjunto de muestras.

Las muestras se agruparon como era de esperar, con la agrupación de las muestras caracterizadas como REPJFX01. Según lo previsto, muestras adicionales de *Salmonella infantis* y una muestra de *Salmonella enteritidis* que no se caracterizaron dentro del grupo REPJFX01 se agruparon por separado (figura 3). Se encontraron siete tipos de secuencias génicas multilocus con loci *aroC*, *dnaN*, *hemD*, *hisD*, *purE*, *sucA* y *thrA*. Los tipos de secuencia coincidieron con el resultado esperado para las 32 muestras. La AMR prevista coincidió con los resultados esperados para las 32 muestras y, en particular, se detectó la mutación *gyrA\_D87Y* que es característica de las cepas REPJFX01 en todas las muestras REP del conjunto de datos (figura 3). El contenido plasmídico coincidió con los resultados esperados para las 32 muestras y se detectó el plásmido IncFIB (pN55391) en todas las cepas REPJFX01, que también es una característica definitoria de esta cepa REP (figura 3).

Tabla 1: Criterios de medición de calidad específicos

Muestra	Calidad media (puntuación Q)	Profundidad de cobertura estimada	Porcentaje de locus principales llamados
1	37,1	76,7×	99,3 %
2	37,1	93,0×	99,3 %
3	37,1	68,3×	99,3 %
4	37,1	67,2×	99,2 %
5	37,1	65,5×	99,3 %
6	37,0	96,9×	99,4 %
7	37,1	77,3×	99,3 %
8	37,0	62,1×	99,3 %
9	37,1	59,9×	99,2 %
10	37,0	57,7×	99,2 %
11	37,0	63,4×	99,3 %
12	37,0	52,8×	99,3 %
13	37,1	86,2×	99,3 %
14	37,0	55,7×	99,3 %
15	37,0	66,8×	99,2 %
16	37,0	55,6×	99,4 %
17	37,1	109,9×	99,5 %
18	37,1	70,5×	99,5 %
19	37,0	54,8×	99,5 %
20	37,0	46,1×	99,5 %
21	37,1	77,0×	99,3 %
22	37,1	78,2×	99,2 %
23	37,1	76,6×	99,2 %
24	37,1	100,6×	99,2 %
25	37,1	80,1×	99,2 %
26	37,2	78,0×	99,2 %
27	37,1	63,0×	99,3 %
28	37,0	73,4×	99,3 %
29	37,1	88,8×	99,2 %
30	37,0	61,6×	99,5 %
31	37,1	71,3×	99,5 %
32	37,1	66,8×	99,4 %



#### Figura 2: Datos de alta calidad para la detección de patógenos de origen alimentario

(A) MiSeq i100 Plus System genera datos de alta calidad con  $\geq 97,5$  % de bases por encima de Q30 (área verde) para todas las muestras analizadas (líneas de color). (B) El porcentaje de distribución del contenido de GC para las lecturas secuenciadas en MiSeq i100 Plus System cumplió los niveles esperados para las 32 muestras analizadas (líneas de color). Los datos se procesaron con la aplicación DRAGEN FastQC + MultiQC.

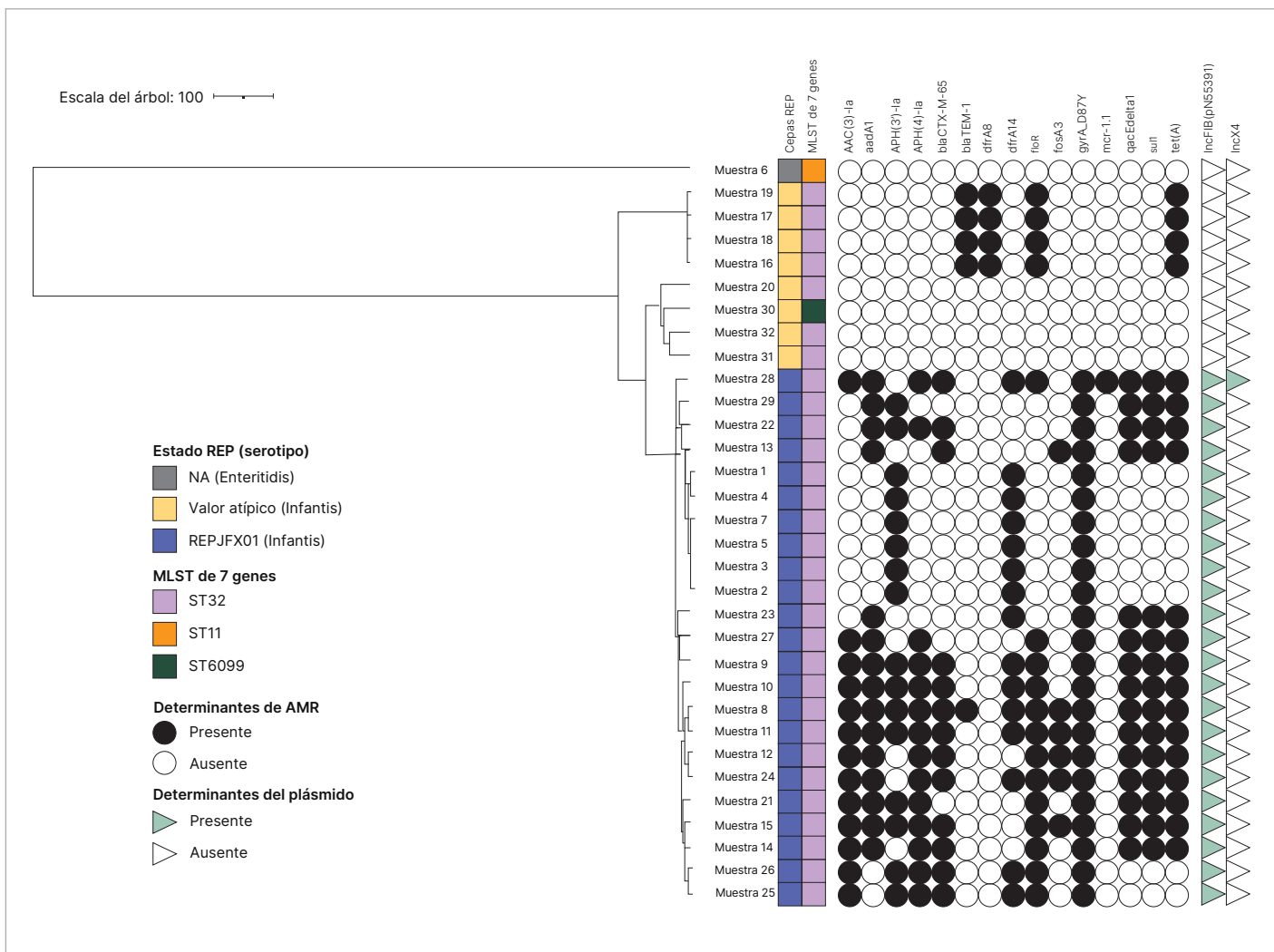


Figura 3: Agrupación de colonias aisladas del estudio basada en cgMLST y datos de genotipo asociados

Método de grupos de pares no ponderados basado en cgMLST con media aritmética (UPGMA, unweighted pair group method with arithmetic mean) de 32 muestras de *Salmonella enterica* con información clave del genotipo; a la derecha del árbol se encuentran el estado REP y la información del serotipo, el MLST de 7 genes, los determinantes de AMR y los determinantes plasmídicos.

## Resumen

Las enfermedades de origen alimentario relacionadas con la contaminación microbiana por bacterias representan una amenaza considerable para la salud humana. MiSeq i100 Series ofrece un flujo de trabajo de NGS rápido e integral que permite una detección altamente precisa de patógenos bacterianos transmitidos por los alimentos para una vigilancia eficaz como parte de los esfuerzos de salud pública.

## Más información

[Illumina DNA Prep](#)

[MiSeq i100 Series](#)

## Bibliografía

- Scharff RL, Besser J, Sharp DJ, Jones TF, Peter GS, Hedberg CW. [An Economic Evaluation of PulseNet: A Network for Foodborne Disease Surveillance](#). *Am J Prev Med*. 2016;50(5 Suppl 1):S66-S73. doi:10.1016/j.amepre.2015.09.018
- Ribot EM, Freeman M, Hise KB, Gerner-Smidt P. [PulseNet: Entering the Age of Next-Generation Sequencing](#). *Foodborne Pathog Dis*. 2019;16(7):451-456. doi:10.1089/fpd.2019.2634
- Stevens EL, Carleton HA, Beal J, et al. [Use of Whole Genome Sequencing by the Federal Interagency Collaboration for Genomics for Food and Feed Safety in the United States](#). *J Food Prot*. 2022;85(5):755-772. doi:10.4315/JFP-21-437
- Tolar B, Joseph LA, Schroeder MN, et al. [An Overview of PulseNet USA Databases](#). *Foodborne Pathog Dis*. 2019;16(7):457-462. doi:10.1089/fpd.2019.2637
- Leeper MM, Tolar BM, Griswold T, et al. [Evaluation of whole and core genome multilocus sequence typing allele schemes for Salmonella enterica outbreak detection in a national surveillance network, PulseNet USA](#). *Front Microbiol*. 2023;14:1254777. Fecha de publicación: 21 de septiembre de 2023. doi:10.3389/fmicb.2023.1254777
- Leeper MM, Schroeder MN, Griswold T, et al. [Validation of Core and Whole-Genome Multi-Locus Sequence Typing Schemes for Shiga-Toxin-Producing E. coli \(STEC\) Outbreak Detection in a National Surveillance Network, PulseNet 2.0, USA](#). *Microorganisms*. 2025;13(6):1310. Fecha de publicación: 4 de junio de 2025. doi:10.3390/microorganisms13061310
- Illumina. Maximizing performance on the MiSeq i100 Series. [illumina.com/content/dam/illumina/gcs/assembled-assets/marketing-literature/miseq-i100-library-optimization-technote-m-gl-03322/miseq-i100-library-optimization-technote-m-gl-03322.pdf](https://www.illumina.com/content/dam/illumina/gcs/assembled-assets/marketing-literature/miseq-i100-library-optimization-technote-m-gl-03322/miseq-i100-library-optimization-technote-m-gl-03322.pdf). Año de publicación: 2024. Fecha de consulta: 5 de febrero de 2025.
- Zhang S, den Bakker HC, Li S, et al. [SeqSero2: Rapid and Improved Salmonella Serotype Determination Using Whole-Genome Sequencing Data](#). *Appl Environ Microbiol*. 2019;85(23):e01746-19. Fecha de publicación: 14 de noviembre de 2019. doi:10.1128/AEM.01746-19
- Feldgarden M, Brover V, Gonzalez-Escalona N, et al. [AMRFinderPlus and the Reference Gene Catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence](#). *Sci Rep*. 2021;11(1):12728. Fecha de publicación: 16 de junio de 2021. doi:10.1038/s41598-021-91456-0
- Carattoli A, Zankari E, García-Fernández A, et al. [In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing](#). *Antimicrob Agents Chemother*. 2014;58(7):3895-3903. doi:10.1128/AAC.02412-14
- Letunic I, Bork P. [Interactive Tree of Life \(iTOL\) v6: recent updates to the phylogenetic tree display and annotation tool](#). *Nucleic Acids Res*. 2024;52(W1):W78-W82. doi:10.1093/nar/gkae268



1 800 809 4566 (llamada gratuita, EE. UU.) | tel.: +1 858 202 4566  
techsupport@illumina.com | www.illumina.com

© 2025 Illumina, Inc. Todos los derechos reservados. Todas las marcas comerciales pertenecen a Illumina, Inc. o a sus respectivos propietarios. Si desea consultar información específica sobre las marcas comerciales, consulte [www.illumina.com/company/legal.html](http://www.illumina.com/company/legal.html).  
M-GL-02912 ESP v2.0