

© 2024 Universidad Nacional Autónoma de México, Facultad de Estudios Superiores Zaragoza.
Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).
TIP Revista Especializada en Ciencias Químico-Biológicas, 27: 1-15, 2024.
<https://doi.org/10.22201/fesz.23958723e.2024.687>

Escherichia coli uropatógena: ejemplo de plasticidad genómica

Isabel Montserrat Cortez-de la Puente¹, Patricia Lozano-Zarain¹,
Rosa del C. Rocha-Gracia¹, Claudia Martínez-de la Peña¹,
Beatriz Eugenia Baca¹ y Margarita María de la Paz Arenas-Hernández^{1*}

¹Posgrado en Microbiología, Centro de Investigación en Ciencias Microbiológicas (CICM),
Instituto de Ciencias (ICUAP), Benemérita Universidad Autónoma de Puebla (BUAP), Puebla,
C.P 72592, Pue., México. E-mail: *margarita.arenas@correo.buap.mx

RESUMEN

Las infecciones del tracto urinario (ITU) son de las más comunes y recurrentes que afectan a hombres y mujeres a cualquier edad. Asimismo, representan un problema de salud pública a nivel mundial. El agente etiológico principal es *Escherichia coli* uropatógena (UPEC), bacteria presente en más del 80% de las ITU adquiridas en la comunidad. A la fecha, varios autores han contribuido con estudios que describen los factores de virulencia que capacitan a *Escherichia coli* uropatógena para producir ITU. Sin embargo, lo que se sabe de la genética de este patógeno se encuentra restringido a un número limitado de genes que han servido para su clasificación como son los que codifican para diversos factores de virulencia, que no son compartidos por la totalidad de las *Escherichia coli* uropatógenas. Por ello, el objetivo de esta revisión es conocer el impacto que el análisis de la secuenciación del genoma completo ha tenido para entender el papel de UPEC como patógeno y como un complemento de otros estudios.

Palabras clave: UPEC, ITU, secuenciación, genoma completo, virulencia, resistencia.

Uropathogenic *Escherichia coli*: example of genomic plasticity

ABSTRACT

Urinary tract infections (UTIs) are among the most common and recurrent bacterial infections affecting men and women at any age. They also represent a worldwide public health problem. The main etiological agent is uropathogenic *Escherichia coli* (UPEC), present in more than 80% of community-acquired UTIs. To date, several authors have contributed studies describing the virulence factors that enable uropathogenic *Escherichia coli* to produce UTI. However, what is known about the genetics of this pathogen is restricted to a limited number of genes that have served for its classification, such as those coding for various virulence factors, which are not shared by all uropathogenic *Escherichia coli*. Therefore, the objective of this review is to know the impact that whole genome sequencing analysis has had in understanding the role of UPEC as a pathogen, as a complement to other studies.

Keywords: UPEC, UTI, sequencing, whole genome, virulence, resistance.

INTRODUCCIÓN

La infección del tracto urinario (ITU) es de las infecciones bacterianas más comunes a nivel mundial. Su clasificación difiere de acuerdo con el enfoque clínico. La guía de infecciones urológicas de la Asociación Europea de Urología (EAU, por sus siglas en inglés) las agrupa en cinco tipos: ITU no complicada, complicada, recurrente, asociada a catéteres y urosepsis (Tabla I), (Bonkat *et al.*, 2021).

Por otro lado, a las ITU también se les clasifica por sus características epidemiológicas y el lugar donde se adquiere la infección que puede ser: la comunidad (CAUTI), el ambiente intrahospitalario (HAUTI) y asociadas al ambiente hospitalario con inicio de síntomas en la comunidad (CO-HAUTI). Las más comunes son las CAUTIs, afectan a más de 150 millones de personas al año (Öztürk & Murt, 2020). Sin embargo, la cantidad está subestimada debido a que la ITU no es de notificación obligatoria en los centros de atención de salud primaria a nivel mundial.

En el año 2018, la Encuesta Nacional De Atención Médica Ambulatoria (NAMCS) reportó un total de 3,577 millones de casos de ITU en los EE. UU. Además, diversos autores estiman que cada año más de 7 millones de personas suelen acudir a las instituciones de atención médica por ITU, lo que representa el 0.7% de los servicios médicos ambulatorios (Tandogdu & Wagenlehner, 2016; Taur & Smith, 2007).

En el año 2022, Yang, Chen, Zheng, Qu, Wang & Yi, realizaron una recopilación y análisis de datos globales de las ITU y el resultado fue un incremento del 60.4% en el número de casos a nivel mundial en el año 2019 con respecto al año 1990.

Ellos, reportan un aumento del 140.2% en las muertes globales asociadas con las ITU, ya que en el año 1990 fueron un total de 98,590 con respecto a las 236,790 en el año 2019. En su estudio mencionan un crecimiento de su presencia en pacientes de la tercera edad y el problema de que del 40 al 75% de los tratamientos con antibióticos son inadecuados.

En lo que respecta a Latinoamérica, existe una carencia de datos sobre las ITU. No obstante, en México, el Sistema Nacional De Vigilancia Epidemiológica (SINAVE) informa de un total de 3,231,627 casos en el año 2022 ubicándolas en el segundo lugar de las veinte principales causas de enfermedad nacional (Figura 1), (SINAVE, 2024).

Es común que a las ITU se les relacione con la aparición de cepas resistentes y a la incapacidad de erradicar completamente a los microorganismos, por esto es importante el conocimiento sobre la variación genética de los diferentes agentes etiológicos para llevar a cabo un tratamiento antimicrobiano adecuado (Abou-Heidar, Degheili, Yacoubian & Khauli, 2019). Al mismo tiempo, el conocimiento sobre la resistencia a los antibióticos en Latinoamérica, en particular en México, se limita a estudios de pequeños grupos de investigación u hospitales con un restringido número de muestras (Ballesteros-Monreal *et al.*, 2023; Barrios-Villa, Picón, Reynaga & Arenas-Hernández, 2023; de Zavaleta *et al.*, 2015).

Una ITU puede ser por microorganismos del tubo digestivo del paciente, donde abundan los bacilos gram negativos como *Escherichia coli*. A pesar de ser parte de la microbiota intestinal, algunas cepas de *E. coli* son capaces de causar una enfermedad extraintestinal como la ITU, por lo que recibe el nombre de

Tabla I. Clasificación de ITU dentro de la Guía de Infecciones Urológicas adoptada por la Asociación Europea de Urología (Traducido y modificado de Bonkat *et al.*, 2021).

Clasificación de ITU	Descripción
ITU no complicada	ITU aguda esporádica o recurrente que se presenta en vías inferiores (cistitis no complicada) y/o altas (pielonefritis no complicada). Limitado a pacientes sin anomalías anatómicas y funcionales relevantes conocidas dentro del tracto urinario o comorbilidades.
ITU complicada	ITU en pacientes con mayor probabilidad de un curso complicado, es decir, infecciones asociadas a factores que comprometen el tracto urinario o la defensa del huésped, incluyendo en éstas la obstrucción urinaria, la retención urinaria causada por enfermedades neurológicas, inmunosupresión, falla renal, embarazo y la presencia de cuerpos extraños como los cálculos renales.
ITU recurrente	Término referido indistintamente a ITU complicada y/o no complicada, cuya frecuencia es de tres episodios por año o dos episodios en los últimos seis meses.
ITU asociada a catéter	ITU en personas cuyo tracto urinario se encuentra cateterizado o ha tenido un catéter colocado en las últimas 48 hrs.
Urosepsis	Disfunción multiorgánica que compromete la vida, causada por una respuesta desregulada del huésped ante una ITU.

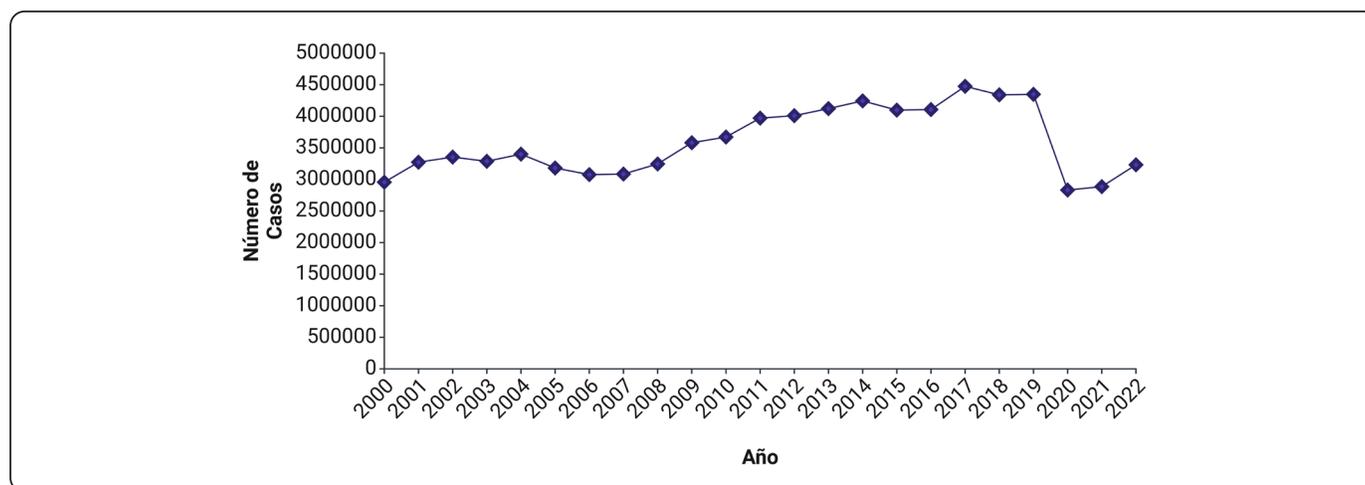


Figura 1. Reporte de casos anuales de las infecciones del tracto urinario (Reporte Anual 2000-2022. SINAVE). Creación propia realizada en BioRender.com

E. coli uropatógena (UPEC) que es reconocida como el principal agente etiológico de esta importante enfermedad (Koga *et al.*, 2014; Torres & Mattera, 2006).

Diversos estudios en Colombia, Argentina y México han encontrado a UPEC como el agente etiológico predominante en la ITU con una incidencia de más del 70% (Ambuila-González, Ramírez-López, Escobar-Bedoya & Chávez, 2015; Ballesteros-Monreal *et al.*, 2023; Córdova *et al.*, 2014; Orrego, Henao & Cardona, 2014).

La vigilancia epidemiológica de las cepas UPEC en México, se basa principalmente en la determinación de los perfiles de virulencia y de resistencia al evaluar la presencia y/o ausencia de biomarcadores genéticos mediante pruebas fenotípicas, pruebas de susceptibilidad antimicrobiana y moleculares, como la reacción en cadena de la polimerasa (Polymerase Chain Reaction o PCR). Se han utilizado otros métodos como la electroforesis en gel de campos pulsados (Pulsed Field Gel Electrophoresis o PFGE) para la determinación de huellas dactilares del ADN Multi-*locus* y su tipificación de secuencias (Multi-*locus* Sequence Typing o MLST) para establecer una relación clonal entre los aislados. Sin embargo, el uso de la secuenciación del genoma completo (Whole Genome Sequencing o WGS) ha proporcionado un análisis más profundo y detallado de la filogenia, virulencia y resistencia de las cepas de UPEC, lo que permite la identificación y caracterización de cepas emergentes, su origen y la identificación de cepas portadoras de genes de resistencia y virulencia, y la ubicación de estos en el genoma bacteriano.

La WGS ha revolucionado los enfoques microbiológicos tradicionales para ofrecer un método más confiable para caracterizar las cepas clínicas de UPEC. Por ello, el objetivo de

esta revisión es conocer el impacto del análisis de la WGS como un complemento de otros métodos de estudio para entender el papel de UPEC como patógeno.

Escherichia coli es un bacilo Gram negativo, anaerobio facultativo, mide aproximadamente 1.1–1.5 µm por 2.0–6.0 µm. pertenece a la familia *Enterobacteriaceae*. *E. coli* se puede diferenciar serológicamente mediante tres antígenos principales de la superficie celular: el O (somático), H (flagelar) y K (capsular).

La mayor parte de las cepas de *E. coli* están en la microbiota intestinal. No obstante, algunas han evolucionado a través de la adquisición de diversos genes que codifican para factores de virulencia, que les confieren una mayor adaptación a nuevos nichos y provoca una amplia gama de enfermedades (Kaper, Nataro & Mobley, 2004).

Las cepas de *E. coli* que tienen la capacidad de generar una enfermedad se dividen en dos grupos principales: *E. coli* diarreogénica (DEC) y *E. coli* extraintestinal (ExPEC), (Figura 2).

Las cepas de ExPEC, provocan diversas enfermedades y son las ITU las más frecuentes con una alta incidencia a nivel mundial donde UPEC es el agente etiológico principal responsable de más del 80% de los casos.

Las cepas de UPEC generalmente tienen una cantidad heterogénea de factores de virulencia que lo convierten en uno de los patógenos más exitosos en el desarrollo de una ITU (Barrios-Villa *et al.*, 2023). UPEC tiene la capacidad de utilizar diversos entornos nutricionales como el intestino, la vejiga, los riñones y el torrente sanguíneo, ya que su metabolismo se

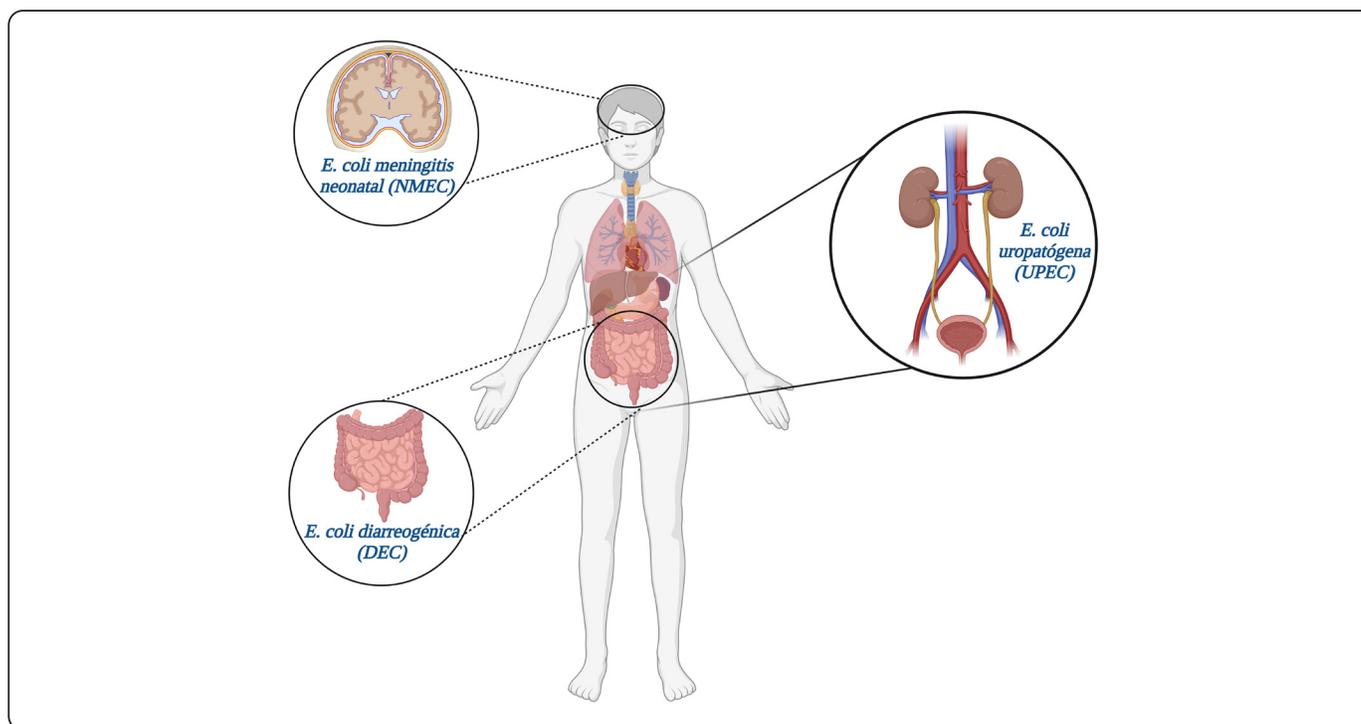


Figura 2. Patotipos de *Escherichia coli*. Creación propia realizada en BioRender.com

encuentra altamente regulado y es sensible a la disponibilidad de nutrientes, que le permiten sobrevivir en entornos competitivos y fluctuantes de nutrientes (Mann, Mediat, Duggin, Harry & Bottomley, 2017).

DIVERSIDAD GENÉTICA DE *E. coli*

E. coli es uno de los microorganismos más utilizados en el laboratorio por su versatilidad y facilidad de manejo, es el modelo microbiológico más importante dentro de la biología molecular, ya que ha sido una pieza medular para el desarrollo de la fisiología y genética bacteriana (Blount, 2015). Por la versátil y diversidad de su genomas es multifacético capaz de colonizar tanto a humanos como animales; esto se aprecia en las diferentes combinaciones de factores de virulencia que varían de acuerdo con el nicho ecológico y/o hospedero en el que se encuentre. Por ello, los diferentes patotipos de *E. coli* se encuentran asociados a: condiciones clínicas, la epidemiología, la filogenia y los factores de virulencia diferentes (García & Fox, 2021; Leimbach, Hacker & Dobrindt, 2013).

La versatilidad genómica de *E. coli* ha sido revelada gracias a la WGS. Antes de la introducción de esta herramienta, su estructura clonal se determinaba mediante la serotipificación de cepas clínicas, cuya variabilidad serológica no era aleatoria, ya que algunos serotipos están asociados a patotipos específicos (Ingle *et al.*, 2016; Ørskov & Ørskov, 1992). La electroforesis enzimática multi-locus (Multilocus Enzyme Electrophoresis o MLEE) apoyó la teoría de que la variabilidad de *E. coli* no

se debe a una selección natural, sino a su aptitud bacteriana (Tenailon, Skurnik, Picard, & Denamur, 2010). La WGS también reveló su plasticidad genómica, producto de su capacidad para cambiar, adquirir o perder genes, y de la selección natural, lo que resulta en su diversidad genética y fenotípica (Blount, 2015).

Serotipificación

La serotipificación es una herramienta clave en el análisis bacteriano y, desde una perspectiva epidemiológica, facilita la identificación de la prevalencia de una serovariedad en diferentes regiones geográficas. En el caso de *E. coli*, la serotipificación se basa en el uso de **anticuerpos de conejo**¹ generados contra los diferentes tipos de antígenos O y H. *E. coli* presenta una gran diversidad en su antígeno O, lo que lo convierte en un biomarcador importante para su clasificación serológica. Esta diversidad ha hecho de la serotipificación el estándar de oro para la tipificación de *E. coli*. La combinación de los antígenos O y H (por ejemplo, UPEC O25:H4) son fundamentales para la detección de brotes epidemiológicos, la diferenciación taxonómica de *E. coli*, la identificación de serotipos patógenos dentro de la especie, así como para estudios clonales y evolutivos (García & Fox, 2021).

¹ Los anticuerpos de conejo obtenidos al inyectar antígenos (antígeno somático o antígeno del flagelo) se usan en la serotipificación de *E. coli* a través del método de aglutinación.

Se han descrito al menos 186 antígenos O, 56 antígenos H y 80 antígenos K. El uso de los antígenos O y H es suficiente para identificar a la mayoría de las cepas de *E. coli*. Emplear el antígeno O contribuye no sólo a determinar el serogrupo, y la combinación de este con el antígeno H, sino también el serotipo (Fratamico, DebRoy, Liu, Needleman, Baranzoni & Feng, 2016; Tenaillon *et al.*, 2010).

El serotipo es considerado un marcador de tipos o linajes específicos de *E. coli*, ya que el número de posibles combinaciones de antígenos O y H conocidos, de esta bacteria, es superior a 10 000. Se han reportado un gran número de eventos de recombinación alrededor del *locus* del antígeno O, la región está sujeta a una fuerte presión de selección de los sistemas inmunitarios, de los hospederos mamíferos y a la depredación de los bacteriófagos. Por tanto, es posible que dos aislados estrechamente relacionados tengan serotipos diferentes, o que dos aislados de *E. coli* de linajes no relacionados hayan convergido en el mismo serotipo (Ingle *et al.*, 2016).

E. coli uropatógena se ha asociado a un número limitado de serogrupos (O1, O2, O4, O6, O7, O8, O16, O18, O25, O62 y O75) y serotipos (O1:H4, O1:H6, O1:H7, O1:H-, O2:H1, O2:H4, O4:H5, O6:H1, O7:H4, O7:H6, O7:H-, O18ac:H7, O18ac:H-, O22:H1, O25:H1, O75:H5 y O75:H7), (Arenas-Hernández, Navarro-Ocaña, Molina-Villa, Martínez-Alvarado, Aroche-Camarillo & Martínez-Laguna, 2011). Sin embargo, se han encontrado nuevos serotipos asociados a las cepas de *E. coli* que producen infecciones urinarias, aunque es importante mencionar que un gran porcentaje de estas son no tipificables (NT), (Ballesteros-Monreal *et al.*, 2020).

La serotipificación tradicional tiene limitaciones como su elevado costo, la presencia de reactividad cruzada de los antisueros, algunos serogrupos, y la variación de anticuerpos entre lotes. Es por ello que se desarrolló una metodología molecular para la determinación del serogrupo y el serotipo. A diferencia de la serotipificación tradicional que utiliza antisueros producidos contra los diferentes tipos O y H de *E. coli*, la serotipificación molecular generalmente se refiere a ensayos basados en la genética que se dirigen a genes específicos dentro del operón que codifica al antígeno O y a genes del antígeno H que codifican los diferentes tipos de antígenos flagelares (Fratamico *et al.*, 2016; Ingle *et al.*, 2016).

Por su estructura variable, se han reportado hasta 181 antígenos O y algunos de ellos se dividen en subgrupos, como el O9 (O9 y O9a), O18 (O18ab y O18ac), O28 (O28ab y O28ac) y O112 (O112ab y O112ac), (Liu *et al.*, 2020).

La determinación genotípica del serotipo se basa en genes relacionados a la biosíntesis de los antígenos O y H. Para el

antígeno O se usan los genes *wzx* (codifica para una flipasa), *wzy* (codifica para una polimerasa), *wzm* y *wzt* (codifican para los transportadores ABC). La determinación del antígeno H se basa principalmente en los genes *fliC*, *fliA* y *fliM* que (codifican para flagelina); también se usa el gen *flkA* que (codifica para una proteína gancho), (Bessonov *et al.*, 2021; Iino *et al.*, 1988; Ratiner, 1998).

La WGS ha permitido una mayor precisión en la identificación del serotipo mediante pruebas *in silico*. El Serotype Finder, es una herramienta web del Centro de Epidemiología Genómica (Epidemiology Genomic Center o CGE que para efecto de lo anterior utiliza el genoma completo y la base de datos de los genes del antígeno O (*wzx*, *wzy*, *wzm* y *wzt*) y del antígeno H (*fliC*, *flkA*, *fliA*, *fliM* y *fliN*) que está en constante actualización (Joensen, Tetzschner, Iguchi, Aarestrup & Scheutz, 2015). Por consiguiente, la WGS da mejores resultados con respecto a la tipificación tradicional, ya que se puede determinar si existen mutaciones puntuales en los genes utilizados para su identificación o si hay discrepancias. Asimismo, se ha observado que los antígenos O y H tienen alelos que son muy similares entre sí, un ejemplo de esto son los antígenos O17, O44, O73, O77 y O106 los cuales comparten una similitud del 99% en los alelos de los genes *wzx* y *wzy*. Los antígenos H4 y H17 coinciden en un 97-99% de similitud en el gen *fliC* (Bessonov *et al.*, 2021).

Tipificación de secuencias multi-*locus* (MLST)

La tipificación de secuencias multi-*locus* (MLST, por sus siglas en inglés) utiliza secuencias completas o fragmentos internos de **genes de mantenimiento**² (usualmente seis o siete genes/fragmentos) que están sujetos a escasa presión selectiva. Estas secuencias/fragmentos son catalogados de acuerdo con su variación alélica y se les asigna un tipo de secuencia o linaje de acuerdo con la combinación alélica observada en cada uno de los genes (Maiden *et al.*, 1998; Urwin & Maiden, 2003).

Para llevar a cabo la MLST de una cepa bacteriana en el laboratorio se amplifican los siete u ocho genes mediante PCR, se someten a secuenciación y se realiza la comparación de los perfiles alélicos para determinar su secuencia tipo (ST), (Figura 3) (Uelze *et al.*, 2020). Los genes utilizados en los esquemas MLST son: *adk* (adenilato kinasa), *fumC* (fumarato hidratasa), *gyrB* (ADN girasa), *icd* (Isocitrato deshidrogenasa), *mdh* (Malato deshidrogenasa), *purA* (Adenilsuccinato sintetasa), y *recA* (Motivo de unión ATP/GTP) para el esquema de Achtman y, *dinB* (ADN polimerasa IV), *icdA* (Isocitrato deshidrogenasa), *pabB* (p-aminobenzoato sintasa), *polB* (ADN polimerasa II),

² Genes esenciales para el mantenimiento de las funciones celulares básicas. Son genes que se expresan de manera estable en todas las células independientemente de la etapa de desarrollo o del ciclo celular en la que se encuentre la célula (Joshi, Drangowska-Way, O'Rourke & Lewis, 2022).

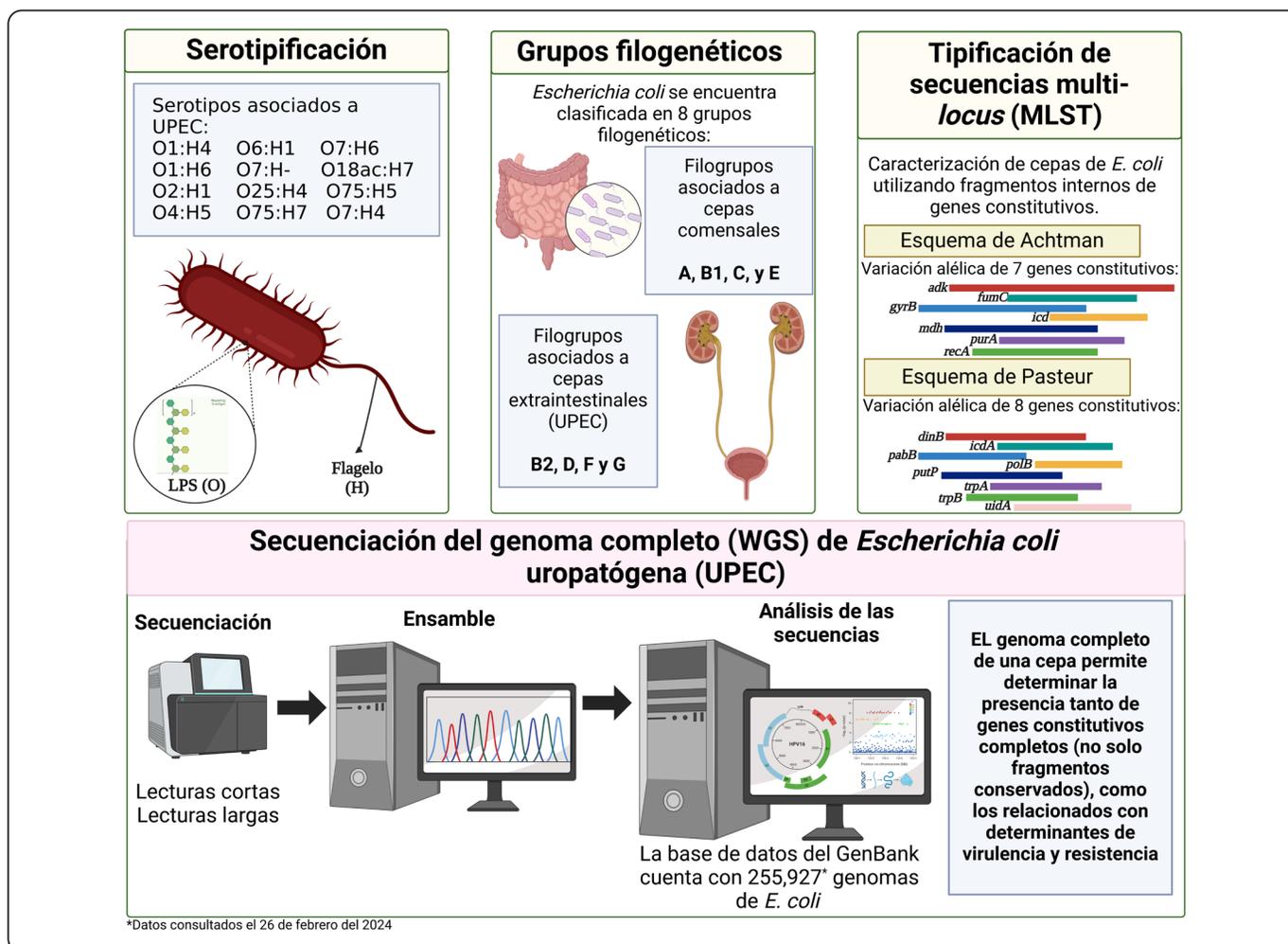


Figura 3. Análisis integral para el estudio de *E. coli*. UPEC. *Escherichia coli* uropatógena; LPS. Lipopolisacárido; O. Antígeno somático; H. Antígeno flagelar; *adk*: adenilato kinasa; *fumC*: fumarato hidratasa; *gyrB*: ADN girasa; *icd*: isocitrato deshidrogenasa; *mdh*: malato deshidrogenasa; *purA*: adenilosuccinato sintetasa; *recA*: motivo de unión ATP/GTP; *dinB*: ADN polimerasa; *icdA*: isocitrato deshidrogenasa; *pabB*: p-amniobenzoato sintasa; *polB*: polimerasa II; *putP*: prolin permeasa; *trpA*: subunidad A de la triptófano sintasa; *trpB*: subunidad B de la triptófano sintasa; *uidA*: Beta-glucuronidasa (Clermont *et al.*, 2013; Genome-NCBI-NLM, n.d.; Kauffmann, 1947). Creación propia realizada en BioRender.com

putP (Prolin permeasa), *trpA* (subunidad α de la triptófano sintasa), *trpB* (subunidad β de la triptófano sintasa), *uidA* (β -glucuronidasa) para el esquema Pasteur. Este método es capaz de proporcionar datos para la vigilancia epidemiológica y su análisis requiere la interpretación detallada de los datos de secuenciación obtenidos de los genes de mantenimiento y la información generada a partir de una serie de herramientas informáticas como las “variaciones alélicas” proporcionadas por el Centro de Epidemiología Genómica (<https://cge.food.dtu.dk/services/MLST/>) y la del Instituto Pasteur (<https://bigsd.web.pasteur.fr/ecoli/>), (Jauregui *et al.*, 2008).

La herramienta del Centro de Epidemiología Genómica, denominada MLST 2.0, con una base de datos amplia almacena la información en cinco bases de datos online. Esta información

es recolectada por la Universidad de Oxford mediante una base de datos pública llamada PubMLST útil para la tipificación molecular y el estudio de la diversidad del genoma bacteriano (Jolley, Bray & Maiden, 2018; Larsen *et al.*, 2012). La plataforma MLST 2.0 cuenta con alrededor de 70 esquemas MLST para 66 especies bacterianas diferentes. Esta herramienta se encuentra en actualización constante mediante un **script**³ automatizado para la descarga mensual de todas las secuencias de alelos y perfiles de ST de las bases de datos del PubMLST. Para la determinación de la ST, se sube en la plataforma el genoma de interés, se especifica la bacteria que se está trabajando y la plataforma realiza un BLAST para la búsqueda en el genoma

³ Un script o guión (en español) es una secuencia de comandos que indican a un ordenador una serie de instrucciones a realizar.

de todos los alelos de los genes que conforman el MLST. En el caso de *E. coli* se cuentan con dos esquemas MLST: el esquema 1, el cual utiliza 7 genes (esquema de Achtman) y el esquema 2 que emplea 8 genes (esquema de Pasteur). Una vez identificados los alelos, se determina la ST que se basa en la combinación resultante de estos alelos (Larsen *et al.*, 2012). El uso de herramientas web de libre acceso que se encuentran en constante actualización permite un resultado más específico en la asignación de la ST.

Filogenia

E. coli es un microorganismo diverso presente en la mayoría de los mamíferos del mundo. Algunas cepas de *E. coli*, como se mencionó, tienen la capacidad de causar enfermedades intra y extraintestinales tanto en humanos como en animales (Clermont, Bonacorsi, & Bingen, 2000; Ørskov & Ørskov, 1992). Al estudiar la diversidad de las cepas patógenas de esta bacteria mediante técnicas de **electroforesis enzimática multi-locus**⁴ (MLEE, por sus siglas en inglés) y **ribotipificación**⁵, encontraron la existencia de subestructuras genéticas asociadas con su sitio de aislamiento. Los primeros ensayos filogenéticos determinaron la presencia de 4 grupos filogenéticos principales (A, B1, B2 y D) y dos accesorios (C y E), (Selander, Caugant, Ochman, Musser, Gilmour, & Whittam, 1986; Herzer, Inouye, Inouye, & Whittam, 1990; Tenaillon *et al.*, 2010). Posteriormente, con el desarrollo de herramientas moleculares como PCR, Clermont *et al.* (2000) desarrollaron una PCR triple para poder llevar a cabo la clasificación filogenética de las cepas de una manera más rápida y eficaz. La PCR triple consiste en la determinación de 3 genes, el gen *chuA* (involucrado en el transporte del grupo hemo), *yjaA* (presente en *E. coli* K-12, función desconocida) y TspE4.C2 (fragmento de ADN) (Clermont *et al.*, 2000).

En 2008, Jauregui *et al.*, al trabajar con 161 cepas de *E. coli* obtenidas de bacteriemias, e identificar los polimorfismos de los nucleótidos obtenidos de la secuenciación de 8 genes de mantenimiento (*dinB*, *icdA*, *pabB*, *polB*, *putP*, *trpA*, *trpB* y *uidA*), encuentran algunas inconsistencias entre la PCR triple propuesta por Clermont *et al.* (2000) y los filogrupos obtenidos en su trabajo. Al hacer un análisis filogenético conjunto observan que las cepas se agrupan en 6 filogrupos, los filogrupos A, B1, B2, D y E, previamente descritos por otros autores, y un filogrupo nuevo que denominan F, en donde se encuentran cepas previamente asignadas al grupo D de acuerdo con la PCR triple. Además, Clermont *et al.* (2011) trabajan con 234 cepas de *E. coli* aisladas de humanos y animales, en las que detectan la presencia de 7 grupos filogenéticos (A, B1, B2, C, D, E y F); el filogrupo C en la PCR triple se identificaba como filogrupo A.

Por ello, Clermont, Christenson, Denamur & Gordon, (2013) hacen una revisión de la PCR triple previamente propuesta y plantean el uso de una PCR cuádruple, donde modifican la secuencia de nucleótidos de los genes *chuA*, *yjaA* y TspE4.C2 y adicionan los genes *arpA* y *trpA* para confirmar los filogrupos D y C respectivamente. Esta metodología proporciona una correcta asignación en alrededor del 95% de las 234 cepas analizadas. Estos datos se validaron al comparar el filogrupo previamente establecido mediante la secuenciación de los ocho genes de mantenimiento utilizados en el esquema del MLST del Instituto Pasteur, con los resultados obtenidos por la PCR cuádruple. Para el procesamiento de los datos de la secuencia de nucleótidos y ubicar el filogrupo se utilizaron las herramientas BAPS y STRUCTURE, que emplean un algoritmo de asignación poblacional (Clermont *et al.*, 2011; Gordon, Clermont, Tolley & Denamur, 2008).

En la actualidad con el incremento de datos de WGS se han podido determinar con mayor precisión los grupos filogenéticos y con potencial de reportar nuevos filogrupos. Lu *et al.* (2016), mediante la WGS de cepas de *E. coli* aisladas de *Marmota himalayana*, identificaron dos nuevos grupos filogenéticos (G y H) que divergen antes que los otros siete filogrupos previamente reportados. De estos dos nuevos grupos, el filogrupo H solo está en una cepa, por lo que aún no hay datos suficientes para considerarlo como un filogrupo nuevo. El filogrupo G se localiza en una colección de cepas previamente reportadas de acuerdo con la PCR cuádruple, como F o B2 demostrando que el filogrupo G es un grupo intermedio entre estos dos (Clermont *et al.*, 2019).

González-Alba, Baquero, Cantón & Galán (2019), trabajan con 6,220 genomas de *E. coli* disponibles en las bases de datos del NCBI y realizan una reconstrucción estratificada de la diversificación de los filogrupos de la bacteria, a partir de la construcción de un árbol filogenético, ellos reportan que los filogrupos se encuentran en 3 linajes: el EB1A (con los filogrupos E, B1 y A), el D y el FGB2 (filogrupos F, G y B2). Además, mediante el uso de **SP-mPH**⁶, observaron que el filogrupo D es el que presenta un menor número de cambios en su genoma, las cepas de este compartían un 99% de los **genes ancestrales** de *E. coli*. Del linaje FGB2, el primero en divergir del genoma ancestral hipotético fue el B2, seguido por el G y al final el F. En el linaje EB1A, el A fue el primero en divergir y el E el último. El C, no pudo ser diferenciado del B1. Dado que solo hay una cepa que pertenece al H, no se encontró información suficiente para validarlo como un filogrupo nuevo. Del mismo modo, un grupo de cepas (0.3%) no las pudieron clasificar en alguno de

⁴ Método basado en el polimorfismo de proteínas (enzimas) determinado mediante su migración electroforética relativa.

⁵ Permite obtener una “huella genética” para la identificación de microorganismos mediante el uso de polimorfismos de longitud de fragmentos de restricción (RFLP) de secuencias genómicas codificantes de ARN ribosomal.

⁶ Combinación de las técnicas de “filogenia estratificada (SP, por sus siglas en inglés)” la cual permite una reconstrucción filogenética de los procesos de diversificación genómica de *E. coli*, y la técnica denominada “sello distintivo de polimorfismo molecular (mPH, por sus siglas en inglés)” que complementa a la SP al realizar un análisis de identificación de los nucleótidos conservados entre las ramas del árbol filogenético y las variables con respecto al ancestro hipotético.

los grupos filogenéticos. Dentro de este 0.3%, 23 genomas (el 0.2% del total de genomas) se agruparon dentro de una rama monofilética que en un futuro podrían considerarse como el filogrupo I. Sin embargo, aún se necesitan más estudios y nuevas herramientas bioinformáticas para proponer la presencia de nuevos filogramos.

Por otro lado, Abram *et al.* (2021) mediante un análisis basado en **Mash**⁷ de 10,667 genomas ensamblados y 95,525 genomas no ensamblados de *E. coli*, determinaron la presencia de 12 filogramos (A, B1, B2-1, B2-2, C, D1, D2, D3, E1, E2 (O157), F y G). Estos comparten genes centrales únicos entre cada uno de ellos que no están en el genoma central de los demás. También observaron que el genoma central de los filogramos denominados como patógenos (B2-1, B2-2, D1, D2, D3, E2 (O157), F y G) están más conservados que los identificados como cepas

comensales (B1 y A), lo que concuerda con trabajos previos donde los eventos de recombinación son menos frecuentes en cepas de *E. coli* patógenas (intra y extraintestinales) que en cepas comensales (McNally, Cheng, Harris & Corander, 2013). Se vieron discrepancias en la clasificación del 7% de los genomas entre el análisis basado en Mash y la PCR del **ClermonTyping**⁸. Otro 9% de ellos no se pudo clasificar mediante esta herramienta y 47 más del filogrupo G quedaron como F con base en ClermonTyping.

Los trabajos de González-Alba *et al.* (2019) y Abram *et al.* (2021) indican que el uso de herramientas analíticas en los genomas secuenciados permite la determinación de cambios genómicos entre los diversos filogramos, lo cual es útil para el entendimiento de la diversidad genómica de microorganismos como *E. coli* (Figura 4).



Figura 4. Línea del tiempo de la clasificación filogenética de *E. coli*. Creación propia realizada en Canva.com

⁷ Análisis basados en k-meros que permite la agrupación escalable de todos el genoma y las distancias de Mash resultantes es una aproximación de la tasa de mutación, lo que realiza aproximaciones rápidas de filogenias mediante una agrupación jerárquica (Ondov *et al.*, 2016).

⁸ Método *in silico* que reproduce los resultados de filotipificación de *E. coli* por PCR de acuerdo a lo propuesto por Clermont *et al.*, 2000 y actualizado en 2013 (Beghain, Bridier-Nahmias, Le Nagard, Denamur & Clermont, 2018).

SECUENCIACIÓN DEL GENOMA COMPLETO (WGS) DE *E. coli* UROPATÓGENA (UPEC)

La WSG ha revolucionado diversas áreas de la investigación, incluida la microbiología. Desde la aparición de los primeros genomas bacterianos completos, se han secuenciado miles incluyendo cepas de referencia y aislados de diversos orígenes. *E. coli* es uno de los microorganismos más estudiados, como también se mencionó, en el laboratorio a nivel mundial, en la actualidad en la base de datos del GenBank se encuentran un total de 255,927 genomas obtenidos de diversos nichos ecológicos. Estos se encuentran hasta febrero del año 2024, en diferentes niveles de ensamblaje, siendo los no cerrados (incompletos: scaffold y contig) los más predominantes (Figura 5) (*Genome-NCBI-NLM*, n.d.).

La disponibilidad de estos genomas ha permitido la identificación de diversos genes relacionados con la virulencia, la formación de biopelículas y con la resistencia a antibióticos incluidos los presentes en islas genómicas (GEIs, por sus siglas en inglés), (Subhadra et al., 2018; Torres-Puig et al., 2022; Welch et al., 2002). El libre acceso a genomas secuenciados ha permitido la creación de la Base de Datos De Factores De Virulencia (VFDB) y la Base de Datos Integral de Resistencia a los Antibióticos (CARD). Estas bases sirven como servidores en línea para la identificación de genes de virulencia y de resistencia respectivamente, así como de mutaciones en sitios blanco de antibióticos que confieren resistencia (Alcock et al., 2023; Liu, Zheng, Jin, Chen & Yang, 2019). Son también

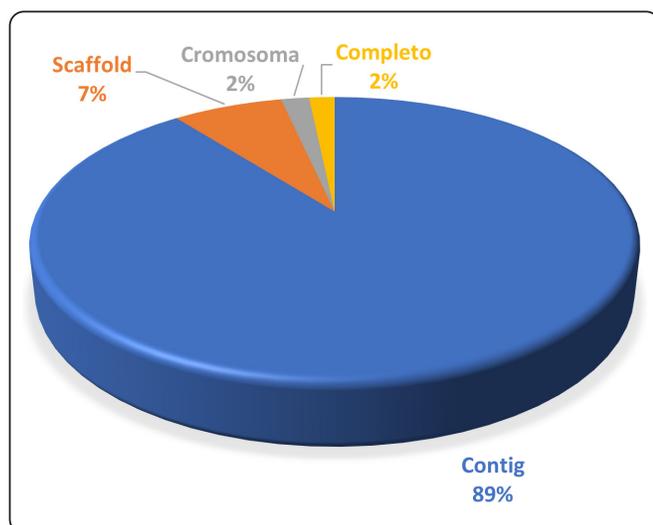


Figura 5. Nivel de ensamblaje de los 255,927 genomas depositados en la base de datos del NCBI. Desde el año 1980 a febrero de 2024 se han depositado diversos genomas de libre acceso. Contig: genomas no cerrados (incompletos); Scaffold: genomas no cerrados fragmentados compuesto por contigs y gaps (incompletos); Cromosoma: Genoma cromosómico cerrado (completo); Genoma completo: Genoma cromosómico y plasmídico (sí los contiene) (*Genome-NCBI-NLM*, n.d.). Creación propia realizada en Excel.

utilizadas como base por otras plataformas en línea como PATRIC, VirulenceFinder y VFAnalyzer para factores de virulencia y ResFinder y PointFinder para resistencia y así como ABRicate para ambos. Estas plataformas realizan una búsqueda mediante BLAST para la determinación *in silico* de los factores de virulencia y resistencia en genomas secuenciados, la mayoría, están en constante actualización (Bortolaia et al., 2020; Joensen et al., 2014; Seeman, 2016; Wattam et al., 2017).

La disponibilidad de los genomas en conjunto con las bases de datos ha permitido el desarrollo de trabajos bioinformáticos integrales sobre la diseminación de resistencia entre cepas clínicas, así como las discrepancias entre lo reportado fenotípicamente con lo reportado a nivel genómico. Por ejemplo, en el diagnóstico de susceptibilidad a la amoxicilina con ácido clavulánico, por una concentración estandarizada para pruebas de susceptibilidad antibiótica, puede dar lugar a falsas cepas sensibles (Sánchez-Osuna, Barbé, & Erill, 2023; Vanstokstraeten et al., 2023). Por ello, la WSG no solo ha proporcionado un mayor entendimiento en la diseminación de resistencia, sino que puede servir como un valioso apoyo para la predicción de la resistencia fenotípica de diversas cepas mediante el uso de más de una herramienta bioinformática.

Las cepas de UPEC pueden ser parte asintomática de la microbiota intestinal. Sin embargo, una vez que UPEC accede al tracto urinario causa ITU que puede ser recurrente, crónica o en casos graves progresar a una urosepsis (Desvaux, Dalmasso, Beyrouthy, Barnich, Delmas & Bonnet, 2020). Si bien la serotipificación, la filogenia y la MLST han servido en la detección de brotes, para la vigilancia epidemiológica, el conocimiento que se tenía sobre UPEC se encontraba restringido a genes conservados (Figura 3). Las regiones de estos genes conservados han proporcionado escasa información sobre su evolución como patógeno, ya que no hay datos sobre los procesos por los cuales los factores asociados a la patogenicidad han sido adquiridos por UPEC con el paso del tiempo (Lo, Zhang, Foxman & Zöllner, 2015).

La información sobre la evolución patogénica de UPEC ha sido proporcionada por la secuenciación del genoma completo (WGS) de cepas clínicas que han sido usadas como modelos de estudio para el entendimiento de su mecanismo de patogenicidad y el origen de los genes asociados a este evento.

El genoma de las cepas prototipo de UPEC oscila entre 5.23 (CFT073), 5.07 (UTI 89) y 4.93 (536) Mpb y es mayor al tamaño del genoma de una cepa comensal como el de *E. coli* K-12 substr. MG1655 (4.64 Mpb). Se estima que la diferencia en tamaño entre una cepa patógena y la comensal puede llegar a ser hasta de 1 Mpb. La diferencia de tamaño se asocia con la ganancia y pérdida de los genes relacionados a factores de virulencia y/o resistencia a través de elementos genéticos móviles como las secuencias de inserción, los transposones,

los plásmidos y los bacteriófagos diseminados mediante eventos de transferencia genética horizontal (TGH) como la transformación, la transducción y la conjugación bacteriana (Croxen, Law, Scholz, Keeney, Wlodarska & Finlay, 2013; Croxen & Finlay, 2009).

Se requiere una mejor comprensión sobre los genes que codifican para los factores de virulencia y de resistencia clínicamente relevantes, su entorno genético, su regulación y como fueron adquiridos para el desarrollo de nuevas estrategias terapéuticas y profilácticas que puedan contribuir en la reducción de la morbilidad asociada con las infecciones urinarias causadas por UPEC.

Los avances en WSG junto con el análisis bioinformático, para la interpretación de los datos de la secuencia genómica, han respondido a cuestiones básicas de la evolución del genoma de *E. coli* y específicamente de UPEC que la ha llevado a adaptarse mejor al tracto urinario (Figura 6). Además, los

estudios de genómica comparativa mediante el uso de diversas herramientas bioinformáticas y analíticas de las cepas de UPEC sugieren que su evolución a partir de una cepa comensal se debe a la adquisición de GEIs mediante diversos eventos de transferencia genética horizontal. Se ha demostrado que las GEIs en UPEC contribuyen a la aptitud y a la adaptación bacteriana a nuevos nichos ecológicos dado que portan genes relacionados a la virulencia (adhesinas, toxinas, formación de biofilm) y resistencia antibiótica (beta-lactamasas), (Desvaux *et al.*, 2020, Dobrindt, Hochhut, Hentschel & Hacker, 2004; Tenaillon *et al.*, 2010).

El análisis del genoma de la cepa prototipo de UPEC CFT073 aislada de una mujer con pielonefritis aguda en EE. UU., muestra que las partes conservadas de su genoma o “core genome” han sufrido cambios lentos en su estructura. Sin embargo, en el resto de su genoma o “genoma accesorio” se presentan una serie de pérdidas y ganancias de genes recién introducidos a través de numerosos eventos de TGH. Algunos de los sitios usados en

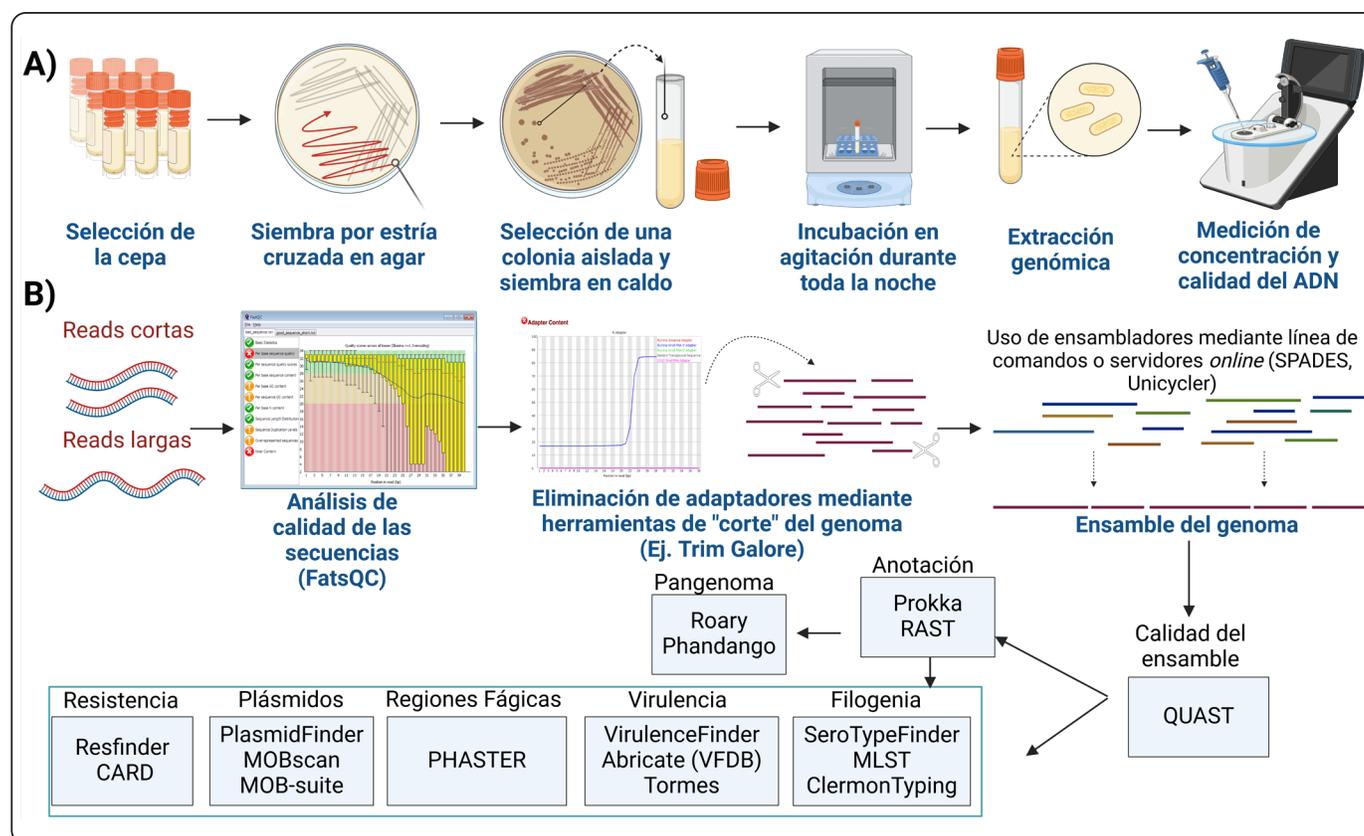


Figura 6. Guía para la secuenciación del genoma completo. **A)** Pasos previos a la secuenciación del genoma. Realizar una buena extracción de ADN de una colonia aislada de la cepa bacteriana problema. Se requiere que el ADN tenga una concentración > 3µg de ADN genómico y una relación de absorbancia de 260/280 de 1.8 a 2.0 y absorbancia de 260/230 de 2 a 2.2 **B)** Análisis de calidad, ensamble y análisis bioinformático. La calidad de las secuencias se determina usando la herramienta FastQC que da una distribución idónea de las puntuaciones de calidad de las secuencias (Por ejemplo: calidad de secuencia por base, el contenido de secuencia por base y la presencia de adaptadores), la presencia de adaptadores interfiere al realizar el ensamblaje. Continúa con la eliminación de adaptadores, ensamble, determinación de la calidad del ensamble. Finalmente, el análisis bioinformático incluye el uso de varias herramientas. Creación propia realizada en BioRender.com

la TGH sirven como objetivos de inserción universal o “sitios calientes” utilizados de forma independiente en linajes separados (Luo, Hu, & Zhu, 2009; Welch *et al.*, 2002). Además, al hacer una comparación entre cepas de UPEC con cepas de *E. coli* comensales se observa que las cepas de UPEC son oportunistas y no muestran una selección preferencial por los factores de virulencia que adquieren. Como si no fuera suficiente, las cepas de UPEC que se encuentran estrechamente relacionadas a nivel filogenético pueden tener fenotipos uropatógenos diferentes (Lo *et al.*, 2015).

La cepa UPEC 26-1 fue aislada de un paciente que sufría de pielonefritis aguda en Corea, con una alta virulencia y al comparar su genoma se encontró que contiene 98,325 pb adicionales al de CFT073, donde se encuentran genes que codifican para proteínas involucradas en la virulencia (por ejemplo, adhesinas), resistencia a múltiples fármacos (bomba de expulsión MFS resistente a tetraciclina) y formación de biopelículas (sistema toxina-antitoxina), (Subhadra *et al.*, 2018).

Además, la WGS ha permitido dilucidar la carga genética de las cepas previamente utilizadas como modelos en pruebas fenotípicas, como CI5 y NU14 cuyo genoma ha servido para el entendimiento del ciclo infectivo de UPEC, así como su capacidad para formar colonias bacterianas intracelulares. Estas cepas han servido también de modelo para el estudio de vacunas usadas en el tratamiento de las ITU por UPEC (Meher Shahi, Abraham & Chen, 2015; Meher Shahi & Chen, 2017).

En México aun cuando se cuenta con poca información sobre el genoma completo de las cepas de UPEC que circulan dentro de nuestra población, en el año 2019, Paniagua-Contreras *et al.* (2019) realizaron la WGS de 24 UPEC aisladas de pacientes femeninos (n=20) y de pacientes masculinos (n=4) y observaron que las secuenciadas presentan un genotipo de resistencia a múltiples fármacos de uso común en la clínica. Además tienen una gran cantidad de genes de virulencia, su pangenoma es “abierto” y con una gran familia de genes **ortólogos**⁹ (Paniagua-Contreras *et al.*, 2019).

CONCLUSIONES

Las infecciones del tracto urinario son de las enfermedades más recurrentes a nivel mundial que afectan tanto a hombres como a mujeres, siendo éstas últimas las más propensas a padecer la enfermedad, pero en los varones es más grave. *E. coli* uropatógena es el principal agente etiológico de las ITU que provoca más del 70% de estas infecciones. Las nuevas tecnologías de secuenciación del genoma completo se han convertido en una herramienta poderosa para describir la relación que existe entre la presencia de algunos genes, principalmente aquellos relacionados a su clasificación filogenética, virulencia

y resistencia bacteriana, con fenotipos que le han permitido a este importante patógeno diseminarse mundialmente.

El número de secuencias del genoma completo de *E. coli* disponibles en las bases de datos de acceso libre al crecer de forma exponencial igual que las herramientas bioinformáticas permiten realizar estudios integrales de genómica comparativa donde se emplean diversas herramientas. Estos trabajos son de gran valor, no sólo, para entender el potencial de virulencia de UPEC, sino también para clasificarla filogenéticamente con la ayuda de “biomarcadores específicos” compartidos por todas las cepas UPEC obtenidos del análisis de WGS de diversas cepas alrededor del mundo. Sin embargo, esto no descarta la posibilidad de hallar genes que sean diferentes de acuerdo con la región geográfica de donde estas procedan. Así como continuar con el estudio de los genomas de *E. coli*, en especial del patotipo uropatógeno de la población mexicana aislado de ITU, para un mejor control y tratamiento de la enfermedad.

AGRADECIMIENTOS

Este trabajo es apoyado por los proyectos: Estudio de la plataforma genética involucrada en la transferencia de la resistencia a antimicrobianos y genes de virulencia en clonas de *Escherichia coli* de interés médico y alimentario. CONAHCYT A1-S-22136. FSSEP02-C-2018-4 y Análisis del genoma de *E. coli* patógena extraintestinal y comensal para la prevención, control y tratamiento adecuados de infección de tracto urinario. Proyecto VIEP 2023.

Isabel Montserrat Cortez de la Puente es apoyada con una beca CONAHCYT para sus estudios de doctorado en el Posgrado de Microbiología del CICM. ICUAP. BUAP. No de Becario 911787.

REFERENCIAS

- Abou-Heidar, N., Degheili, J., Yacoubian, A. & Khauli, R. (2019). Management of urinary tract infection in women: A practical approach for everyday practice. *Urology Annals*, **11(4)**, 339. https://doi.org/10.4103/UA.UA_104_19
- Abram, K., Udaondo, Z., Bleker, C., Wanchai, V., Wassenaar, T. M., Robeson, M. S & Ussery, D. W. (2021). Mash-based analyses of *Escherichia coli* genomes reveal 14 distinct phylogroups. *Communications Biology*, **4(1)**, 117. <https://doi.org/10.1038/s42003-020-01626-5>
- Alcock, B. P., Huynh, W., Chalil, R., Smith, K. W., Raphenya, A. R., Wlodarski, M. A., Edalatmand, A., Petkau, A., Syed, S. A., Tsang, K. K., Baker, S. J. C., Dave, M., McCarthy, M. C., Mukiri, K. M., Nasir, J. A., Golbon, B., Imtiaz, H., Jiang, X., Kaur, K., Kwong, M., Liang, Z. C., Niu, K. C., Shan, P., Yang, J. Y. J., Gray, K. L., Hoad, G. R., Jia, B., Bhandu, T., Carfrae, L. A., Farha, M. A., French, S., Gordzevich, R., Rachwalski, Tu, K., M., Bordeleau, E., Dooley, D., Griffiths, Zubyk, E., H. L., Brown, E. D., Maguire, F., Beiko, R. G., Hsiao, W. W. L., Brinkman, F. S. L., Van Domselaar, G. & McArthur A. G. (2023). CARD

⁹ Son genes homólogos donde un gen diverge después de un evento de especiación, pero el gen y su función principal se conservan.

- 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Research*, **51(D1)**, D690–D699. <https://doi.org/10.1093/nar/gkac920>
- Ambuila-González, E., Ramírez-López, L., Escobar-Bedoya, A. & Chávez, M. (2015). Prevalencia de uropatógenos Gram negativo aislados de pacientes adultos en la ciudad de Cali. *Ciencia & Salud*, **4(13)**, 11–17.
- Arenas-Hernández, M. M. P., Navarro-Ocaña, A., Molina-Villa, T., Martínez-Alvarado, J., Aroche-Camarillo, F. & Martínez-Laguna, Y. (2011). *Escherichia coli* UROPATÓGENA. En *Modelos de la Patogénesis de las Enfermedades Infecciosas II*. Rocha-Gracia Rosa del Carmen, Lozano-Zarain, Patricia y Martínez-Laguna, Ygnacio. (Eds.). Publicación especial de la Benemérita Universidad Autónoma de Puebla. Puebla, México. ISBN. 978-607-487-476-1.
- Ballesteros-Monrreal, M. G., Arenas-Hernández, M. M. P., Enciso-Martínez, Y., Martínez-De la Peña, C. F., Rocha-Gracia, R. D. C., Lozano-Zarain, P., Navarro-Ocaña, A., Martínez-Laguna, Y. & de la Rosa-López, R. (2020). Virulence and Resistance Determinants of Uropathogenic *Escherichia coli* Strains Isolated from Pregnant and Non-Pregnant Women from Two States in Mexico. *Infection and Drug Resistance*, **13**, 295. <https://doi.org/10.2147/IDR.S226215>
- Ballesteros-Monrreal, M. G., Mendez-Pfeiffer, P., Barrios-Villa, E., Arenas-Hernández, M. M. P., Enciso-Martínez, Y., Sepúlveda-Moreno, C. O., Bolado-Martínez, E. & Valencia, D. (2023). Uropathogenic *Escherichia coli* in Mexico, an Overview of Virulence and Resistance Determinants: Systematic Review and Meta-analysis. *Archives of Medical Research*, **54(3)**, 247–260. <https://doi.org/10.1016/J.ARCMED.2023.01.001>
- Barrios-Villa, E., Picón, L. R., Reynaga, R. B., & Arenas-Hernández, M. M. de la P. (2023). An Updated Overview on the Resistance and Virulence of UPEC. In A. G. Torres (Ed.), *Trending Topics in Escherichia coli Research* (pp. 249–276). Springer International Publishing. Cham, Switzerland https://doi.org/10.1007/978-3-031-29882-0_11
- Beghain, J., Bridier-Nahmias, A., Le Nagard, H., Denamur, E. & Clermont, O. (2018). ClermonTyping: an easy-to-use and accurate in silico method for *Escherichia* genus strain phylotyping. *Microbial Genomics*, **4(7)**, e000192. <https://doi.org/10.1099/mgen.0.000192>
- Bessonov, K., Laing, C., Robertson, J., Yong, I., Ziebell, K., Gannon, V. P. J., Nichani, A., Arya, G., Nash, J. H. E. & Christianson, S. (2021). ECTyper: *in silico* *Escherichia coli* serotype and species prediction from raw and assembled whole-genome sequence data. *Microbial Genomics*, **7(12)**, 728. <https://doi.org/10.1099/MGEN.0.000728>
- Blount, Z. D. (2015). The unexhausted potential of *E. coli*. In *eLife* (Vol. 4). eLife Sciences Publications, Ltd. <https://doi.org/10.7554/eLife.05826>
- Bonkat, G., Bartoletti, R., Bruyere, F., Cai, T., Geerlings, S. E., Koves, B., Schubert, S. & Wagenlehner, F. (2021). EAU Guidelines on Urological Infections. Edn. presented at the EAU Annual Congress Milan Italy 2021, March, 18–20. ISBN 978-94-92671-13-4.
- Bortolaia, V., Kaas, R. S., Ruppe, E., Roberts, M. C., Schwarz, S., Cattoir, V., Philippon, A., Allesoe, R. L., Rebelo, A. R., Florensa, A. F., Fagelhauer, L., Chakraborty, T., Neumann, B., Werner, G., Bender, J. K., Stingl, K., Nguyen, M., Coppens, J., Xavier, B. B., Malhotra-Kumar, S., Westh, H., Pinholt, M., Anjum, M. F., Duggett, N. A., Kempf, I., Nykäsenoja, S., Olkkola, S., Wiecek, K., Amaro, A., Clemente, L., Mossong, J., Losch, S., Ragimbeau, C., Lund, O. & Aarestrup, F. M. (2020). ResFinder 4.0 for predictions of phenotypes from genotypes. *The Journal of Antimicrobial Chemotherapy*, **75(12)**, 3491–3500. <https://doi.org/10.1093/jac/dkaa345>
- Clermont, O., Bonacorsi, S. & Bingen, E. (2000). Rapid and simple determination of the *Escherichia coli* phylogenetic group. *Applied and Environmental Microbiology*, **66(10)**, 4555–4558. <https://doi.org/10.1128/AEM.66.10.4555-4558.2000>
- Clermont, O., Christenson, J. K., Denamur, E. & Gordon, D. M. (2013). The Clermont *Escherichia* and Relatives of *Escherichia coli* B, and the Derivation of B Strains REL606 and BL21(DE3). *Journal of Molecular Biology*, **394(4)**, 634–643. <https://doi.org/10.1016/J.JMB.2009.09.022>
- Clermont, O., Dixit, O. V. A., Vangchhia, B., Condamine, B., Dion, S., Bridier-Nahmias, A., Denamur, E. & Gordon, D. (2019). Characterization and rapid identification of phylogroup G in *Escherichia coli*, a lineage with high virulence and antibiotic resistance potential. *Environmental Microbiology*, **21(8)**, 3107–3117. <https://doi.org/10.1111/1462-2920.14713>
- Clermont, O., Olier, M., Hoede, C., Diancourt, L., Brisse, S., Keroudean, M., Glodt, J., Picard, B., Oswald, E. & Denamur, E. (2011). Animal and human pathogenic *Escherichia coli* strains share common genetic backgrounds. *Infection, Genetics and Evolution*, **11(3)**, 654–662. <https://doi.org/10.1016/J.MEEGID.2011.02.005>
- Córdova, E., Lespada, M., Cecchini, D., Jacob, N., Gomez, N., Gutfraind, G., Badía, M. & Rodríguez Ismael, C. (2014). Prevalencia de gérmenes multirresistentes en infecciones del tracto urinario de la comunidad y asociadas a los cuidados de la salud. En *Actual. SIDA. Infectol.*, 33–38. <https://doi.org/10.52226/revista.v22i84.268>
- Croxen, M. A. & Finlay, B. B. (2009). Molecular mechanisms of *Escherichia coli* pathogenicity. *Nature Reviews Microbiology*, **8(1)**, 26–38. <https://doi.org/10.1038/nrmicro2265>
- Croxen, M. A., Law, R. J., Scholz, R., Keeney, K. M., Wlodarska, M. & Finlay, B. B. (2013). Recent advances in understanding enteric pathogenic *Escherichia coli*. *Clinical Microbiology Reviews*, **26(4)**, 822–880. <https://doi.org/10.1128/CMR.00022-13>

- De Zavaleta, S. M., de León Garduño, P. A., Esquivel, G. J., Nava, R. E., Covarrubias, R. F., Ramírez, G. A., Hernández, A. J., del Barco, E. L., Alpuche, C. J., Narváez, C. J., Juárez, C. P., González, T. P., Escoto, G. P., Carrera, F. O., Grajales, S. L., Orozco, H. H., Molina, C. R., Moreno, D. LA, Sagastegui, S., Mendoza, O., Pérez, M., Anzures A., Zuno R., Rodríguez, G., Rodríguez, C., Rivera Martínez G. A. R., Almeida G. & Barbosa, A. P. (2015). Recomendaciones de expertos mexicanos en el tratamiento de las infecciones del tracto urinario en pacientes adultos, embarazadas y niños. *Rev. Mex. Urol.*, **75**(2), 1–46. www.elsevier.es/uromx
- Desvaux, M., Dalmasso, G., Beyrouthy, R., Barnich, N., Delmas, J. & Bonnet, R. (2020). Pathogenicity Factors of Genomic Islands in Intestinal and Extraintestinal *Escherichia coli*. *Frontiers in Microbiology*, **11**, 558927. <https://doi.org/10.3389/FMICB.2020.02065/BIBTEX>
- Dobrindt, U., Hochhut, B., Hentschel, U. & Hacker, J. (2004). Genomic islands in pathogenic and environmental microorganisms. *Nature Reviews Microbiology*, **2**(5), 414–424. <https://doi.org/10.1038/nrmicro884>
- Fratamico, P. M., DeRoy, C., Liu, Y., Needleman, D. S., Baranzoni, G. M. & Feng, P. (2016). Advances in Molecular Serotyping and Subtyping of *Escherichia coli*. *Frontiers in Microbiology*, **7**, 644. <https://doi.org/10.3389/FMICB.2016.00644>
- García, A. & Fox, J. G. (2021). A One Health Perspective for Defining and Deciphering *Escherichia coli* Pathogenic Potential in Multiple Hosts. *Comparative Medicine*, **71**(1), 3. <https://doi.org/10.30802/AALAS-CM-20-000054>
- Genome-NCBI-NLM. (n.d.). Retrieved February 26, 2024, from <https://www.ncbi.nlm.nih.gov/datasets/genome/?taxon=562>
- González-Alba, J. M., Baquero, F., Cantón, R. & Galán, J. C. (2019). Stratified reconstruction of ancestral *Escherichia coli* diversification. *BMC Genomics*, **20**(1), 1–15. <https://doi.org/10.1186/S12864-019-6346-1/FIGURES/6>
- Gordon, D. M., Clermont, O., Tolley, H. & Denamur, E. (2008). Assigning *Escherichia coli* strains to phylogenetic groups: multi-locus sequence typing versus the PCR triplex method. *Environmental Microbiology*, **10**(10), 2484–2496. <https://doi.org/10.1111/J.1462-2920.2008.01669.X>
- Herzer, P. J., Inouye, S., Inouye, M. & Whittam, T. S. (1990). Phylogenetic distribution of branched RNA-linked multicopy single-stranded ADN among natural isolates of *Escherichia coli*. *Journal of Bacteriology*, **172**(11), 6175–6181. <https://doi.org/10.1128/jb.172.11.6175-6181.1990>
- Iino, T., Komeda, Y., Kutsukake, K., Macnab, R. M., Matsumura, P., Parkinson, J. S., Simon, M. I. & Yamaguchi, S. (1988). New unified nomenclature for the flagellar genes of *Escherichia coli* and *Salmonella typhimurium*. *Microbiological Reviews*, **52**(4), 533. <https://doi.org/10.1128/MR.52.4.533-535.1988>
- Ingle, D. J., Valcanis, M., Kuzevski, A., Tauschek, M., Inouye, M., Stinear, T., Levine, M. M., Robins-Browne, R. M. & Holt, K. E. (2016). *In silico* serotyping of *E. coli* from short read data identifies limited novel O-loci but extensive diversity of O:H serotype combinations within and between pathogenic lineages. *Microbial Genomics*, **2**(7), e000064. <https://doi.org/10.1099/mgen.0.000064>
- Jauregui, F., Landraud, L., Passet, V., Diancourt, L., Frapy, E., Guigon, G., Carbonnelle, E., Lortholary, O., Clermont, O., Denamur, E., Picard, B., Nassif, X. & Brisse, S. (2008). Phylogenetic and genomic diversity of human bacteremic *Escherichia coli* strains. *BMC Genomics*, **9**(1), 1–14. <https://doi.org/10.1186/1471-2164-9-560/TABLES/4>
- Joensen, K. G., Scheutz, F., Lund, O., Hasman, H., Kaas, R. S., Nielsen, E. M. & Aarestrup, F. M. (2014). Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli*. *Journal of Clinical Microbiology*, **52**(5), 1501–1510. <https://doi.org/10.1128/JCM.03617-13>
- Joensen, K. G., Tetzschner, A. M. M., Iguchi, A., Aarestrup, F. M. & Scheutz, F. (2015). Rapid and Easy *in silico* Serotyping of *Escherichia coli* Isolates by Use of Whole-Genome Sequencing Data. *Journal of Clinical Microbiology*, **53**(8), 2410–2426. <https://doi.org/10.1128/JCM.00008-15>
- Jolley, K. A., Bray, J. E. & Maiden, M. C. J. (2018). Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Research*, **3**, 124. <https://doi.org/10.12688/WELLCOMEOPENRES.14826.1>
- Joshi, C. J., Ke, W., Drangowska-Way, A., O'Rourke, E. J. & Lewis, N. E. (2022). What are housekeeping genes? *PLoS Computational Biology*, **18**(7), e1010295. <https://doi.org/10.1371/JOURNAL.PCBI.1010295>
- Kaper, J. B., Nataro, J. P. & Mobley, H. L. T. (2004). Pathogenic *Escherichia coli*. *Nature Reviews. Microbiology*, **2**(2), 123–140. <https://doi.org/10.1038/NRMICRO818>
- Kauffmann, F. (1947). The serology of the coli group. *The Journal of Immunology*, **57**(1), 71–100.
- Koga, V. L., Tomazetto, G., Cyويا, P. S., Neves, M. S., Vidotto, M. C., Nakazato, G. & Kobayashi, R. K. T. (2014). Molecular screening of virulence genes in extraintestinal pathogenic *Escherichia coli* isolated from human blood culture in Brazil. *BioMed Research International*. **2014**, 465054. <https://doi.org/10.1155/2014/465054>
- Larsen, M. V., Cosentino, S., Rasmussen, S., Friis, C., Hasman, H., Marvig, R. L., Jelsbak, L., Sicheritz-Pontén, T., Ussery, D. W., Aarestrup, F. M. & Lund, O. (2012). Multilocus sequence typing of total-genome-sequenced bacteria. *Journal of Clinical Microbiology*, **50**(4), 1355–1361. <https://doi.org/10.1128/JCM.06094-11>
- Leimbach, A., Hacker, J. & Dobrindt, U. (2013). *E. coli* as an all-rounder: The thin line between commensalism and pathogenicity. *Current Topics in Microbiology and Immunology*, **358**, 3–32. https://doi.org/10.1007/82_2012_303
- Liu, B., Furevi, A., Perepelov, A. V., Guo, X., Cao, H., Wang,

- Q., Reeves, P. R., Knirel, Y. A., Wang, L. & Widmalm, G. (2020). Structure and genetics of *Escherichia coli* O antigens. *FEMS Microbiology Reviews*, **44**(6), 655. <https://doi.org/10.1093/FEMSRE/FUZ028>
- Liu, B., Zheng, D., Jin, Q., Chen, L. & Yang, J. (2019). VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Research*, **47**(Database issue), D687. <https://doi.org/10.1093/NAR/GKY1080>
- Lo, Y., Zhang, L., Foxman, B. & Zöllner, S. (2015). Whole-genome sequencing of uropathogenic *Escherichia coli* reveals long evolutionary history of diversity and virulence. *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases*, **34**, 244. <https://doi.org/10.1016/J.MEEGID.2015.06.023>
- Lu, S., Jin, D., Wu, S., Yang, J., Lan, R., Bai, X., Liu, S., Meng, Q., Yuan, X., Zhou, J., Pu, J., Chen, Q., Dai, H., Hu, Y., Xiong, Y., Ye, C. & Xu, J. (2016). Insights into the evolution of pathogenicity of *Escherichia coli* from genomic analysis of intestinal *E. coli* of Marmota himalayana in Qinghai-Tibet plateau of China. *Emerging Microbes & Infections*, **5**(12), e122. <https://doi.org/10.1038/emi.2016.122>
- Luo, C., Hu, G. Q. & Zhu, H. (2009). Genome reannotation of *Escherichia coli* CFT073 with new insights into virulence. *BMC Genomics*, **10**, 552. <https://doi.org/10.1186/1471-2164-10-552>
- Maiden, M. C. J., Bygraves, J. A., Feil, E., Morelli, G., Russell, J. E., Urwin, R., Zhang, Q., Zhou, J., Zurth, K., Caugant, D. A., Feavers, I. M., Achtman, M. & Spratt, B. G. (1998). Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. *Proceedings of the National Academy of Sciences of the United States of America*, **95**(6), 3140. <https://doi.org/10.1073/PNAS.95.6.3140>
- Mann, R., Mediati, D. G., Duggin, I. G., Harry, E. J. & Bottomley, A. L. (2017). Metabolic Adaptations of Uropathogenic *E. coli* in the Urinary Tract. *Frontiers in Cellular and Infection Microbiology*, **7**(241). <https://doi.org/10.3389/FCIMB.2017.00241>
- McNally, A., Cheng, L., Harris, S. R. & Corander, J. (2013). The Evolutionary Path to Extraintestinal Pathogenic, Drug-Resistant *Escherichia coli* Is Marked by Drastic Reduction in Detectable Recombination within the Core Genome. *Genome Biology and Evolution*, **5**(4), 699–710. <https://doi.org/10.1093/gbe/evt038>
- Mehershahi, K. S. & Chen, S. L. (2017). Complete Genome Sequence of the Uropathogenic *Escherichia coli* Strain NU14. *Genome Announcements*, **5**(18), e00306-17. <https://doi.org/10.1128/genomeA.00306-17>
- Mehershahi, K. S., Abraham, S. N. & Chen, S. L. (2015). Complete Genome Sequence of Uropathogenic *Escherichia coli* Strain CI5. *Genome Announcements*, **3**(3), e00558-15. <https://doi.org/10.1128/genomeA.00558-15>
- Ondov, B. D., Treangen, T. J., Melsted, P., Mallonee, A. B., Bergman, N. H., Koren, S. & Phillippy, A. M. (2016). Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biology*, **17**(1), 132. <https://doi.org/10.1186/s13059-016-0997-x>
- Orrego, C., Henao, C. & Cardona, J. (2014). Prevalence of urinary infection, uropathogens and antimicrobial susceptibility profile. *Rev Acta Medica Colombiana*, **39**(352–358), 7. <http://www.scielo.org.co/pdf/amc/v39n4/v39n4a08.pdf>
- Ørskov, F. & Ørskov, I. (1992). *Escherichia coli* serotyping and disease in man and animals. *Canadian Journal of Microbiology*, **38**(7), 699–704. <https://doi.org/10.1139/m92-115>
- Öztürk, R. & Murt, A. (2020). Epidemiology of urological infections: a global burden. *World Journal of Urology*, **38**(11), 2669–2679. <https://doi.org/10.1007/s00345-019-03071-4>
- Paniagua-Contreras, G. L., Monroy-Pérez, E., Díaz-Velásquez, C. E., Uribe-García, A., Labastida, A., Peñalosa-Figueroa, F., Domínguez-Trejo, P., García, L. R., Vaca-Paniagua, F. & Vaca, S. (2019). Whole-genome sequence analysis of multidrug-resistant uropathogenic strains of *Escherichia coli* from Mexico. *Infection and Drug Resistance*, **12**, 2363–2377. <https://doi.org/10.2147/IDR.S203661>
- Ratiner, Y. A. (1998). New Flagellin-Specifying Genes in Some *Escherichia coli* Strains. *Journal of Bacteriology*, **180**(4), 979. <https://doi.org/10.1128/JB.180.4.979-984.1998>
- Sánchez-Osuna, M., Barbé, J. & Erill, I. (2023). Systematic *in silico* Assessment of Antimicrobial Resistance Dissemination across the Global Plasmidome. *Antibiotics (Basel, Switzerland)*, **12**(2), 281. <https://doi.org/10.3390/antibiotics12020281>
- Seeman, T. (2016). ABRicate: mass screening of contigs for antibiotic resistance genes. Retrieved February 28, 2024, from <https://github.com/tseemann/abricate>
- Selander, R. K., Caugant, D. A., Ochman, H., Musser, J. M., Gilmour, M. N., & Whittam, T. S. (1986). Methods of multilocus enzyme electrophoresis for bacterial population genetics and systematics. *Applied and Environmental Microbiology*, **51**(5), 873–884. <https://doi.org/10.1128/aem.51.5.873-884.1986>
- SINAVE. (2024). Boletín Epidemiológico Sistema Nacional de Vigilancia Epidemiológica Sistema Único de Información | Secretaría de Salud | Gobierno | gob.mx. <https://www.gob.mx/salud/documentos/boletinepidemiologico-sistema-nacional-de-vigilancia-epidemiologica-sistema-unico-de-informacion-355523>
- Subhadra, B., Kim, D. H., Kim, J., Woo, K., Sohn, K. M., Kim, H. J., Han, K., Oh, M. H. & Choi, C. H. (2018). Complete genome sequence of uropathogenic *Escherichia coli* isolate UPEC 26-1. *Genes and Genomics*, **40**(6), 643–655. <https://doi.org/10.1007/s13258-018-0665-5>
- Tandogdu, Z. & Wagenlehner, F. M. E. (2016). Global epidemiology of urinary tract infections. *Current Opinion in*

- Infectious Diseases*, **29(1)**, 73–79. <https://doi.org/10.1097/QCO.0000000000000228>
- Taur, Y. & Smith, M. A. (2007). Adherence to the Infectious Diseases Society of America guidelines in the treatment of uncomplicated urinary tract infection. *Clinical Infectious Diseases*, **44(6)**, 769–774. <https://doi.org/10.1086/511866>
- Tenaillon, O., Skurnik, D., Picard, B. & Denamur, E. (2010). The population genetics of commensal *Escherichia coli*. *Nature Reviews Microbiology*, **8(3)**, 207–217. <https://doi.org/10.1038/nrmicro2298>
- Torres, M. & Mattered, A. (2006). Infección urinaria. En *Temas de Bacteriología y Virología Médica* (2da edición, pp. 189–196) Montevideo. Facultad de Medicina. Instituto de Higiene: Oficina del Libro-FEFMUR.
- Torres-Puig, S., García, V., Stærk, K., Andersen, T. E., Møller-Jensen, J., Olsen, J. E. & Herrero-Fresno, A. (2022). “Omics” Technologies - What Have They Told Us About Uropathogenic *Escherichia coli* Fitness and Virulence During Urinary Tract Infection? *Frontiers in Cellular and Infection Microbiology*, **12**, 824039. <https://doi.org/10.3389/FCIMB.2022.824039>
- Uelze, L., Grützkke, J., Borowiak, M., Hammerl, J. A., Juraschek, K., Deneke, C., Tausch, S. H. & Malorny, B. (2020). Typing methods based on whole genome sequencing data. *One Health Outlook*, **2(1)**, 1–19. <https://doi.org/10.1186/S42522-020-0010-1>
- Urwin, R. & Maiden, M. C. J. (2003). Multi-locus sequence typing: a tool for global epidemiology. *Trends in Microbiology*, **11(10)**, 479–487. <https://doi.org/10.1016/J.TIM.2003.08.006>
- Vanstokstraeten, R., Piérard, D., Crombé, F., De Geyter, D., Wybo, I., Muyltermans, A., Seyler, L., Caljon, B., Janssen, T. & Demuyser, T. (2023). Genotypic resistance determined by whole genome sequencing versus phenotypic resistance in 234 *Escherichia coli* isolates. *Scientific reports*, **13(1)**, 449. <https://doi.org/10.1038/s41598-023-27723-z>
- Wattam, A. R., Davis, J. J., Assaf, R., Boisvert, S., Brettin, T., Bun, C., Conrad, N., Dietrich, E. M., Disz, T., Gabbard, J. L., Gerdes, S., Henry, C. S., Kenyon, R. W., Machi, D., Mao, C., Nordberg, E. K., Olsen, G. J., Murphy-Olson, D. E., Olson, R., Overbeek, R., Parrello, B., Pusch, G. D., Shukla, M., Vonstein, V., Warren, A., Xia, F., Yoo, H. & Stevens, R. L. (2017). Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center. *Nucleic Acids Research*, **45(D1)**, D535–D542. <https://doi.org/10.1093/nar/gkw1017>
- Welch, R. A., Burland, V., Plunkett, G., Redford, P., Roesch, P., Rasko, D., Buckles, E. L., Liou, S. R., Boutin, A., Hackett, J., Stroud, D., Mayhew, G. F., Rose, D. J., Zhou, S., Schwartz, D. C., Perna, N. T., Mobley, H. L. T., Donnenberg, M. S. & Blattner, F. R. (2002). Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, **99(26)**, 17020. <https://doi.org/10.1073/PNAS.252529799>
- Yang, X., Chen, H., Zheng, Y., Qu, S., Wang, H. & Yi, F. (2022). Disease burden and long-term trends of urinary tract infections: A worldwide report. *Frontiers in Public Health*, **10**, 888205. <https://doi.org/10.3389/fpubh.2022.888205>