



CAPÍTULO 2

Bioinformática aplicada al análisis de variantes genéticas en poblaciones latinoamericanas: Una revisión sistemática de la literatura

Bioinformatics applied to the analysis of genetic variants in Latin American populations: A systematic review of the literature

Hernandez N., Medina M.
DOI: 10.55204/pmea.134.c211

Noelia Meliza Hernandez Aparcana, 0000-0002-9348-0642 

Universidad Nacional San Luis de Gonzaga, Facultad de Ciencias Biológicas. Av. Los maestros
S/N Ciudad Universitaria 11001, Código Postal: 11001, Ica, Ica. Perú
noelia.hernandez@unica.edu.pe

Merici Ingrid Medina Guerrero, 0000-0003-0728-9997 

Universidad Nacional San Luis de Gonzaga, Facultad de Ciencias Biológicas. Av. Los maestros
S/N Ciudad Universitaria 11001, Código Postal: 11001, Ica, Ica. Perú
ingrid.medina@unica.edu.pe

Resumen

Introducción: Las poblaciones latinoamericanas poseen una vasta diversidad genómica producto del mestizaje ancestral; sin embargo, presentan una marcada subrepresentación en los catálogos globales. Esta brecha limita la eficacia de la medicina de precisión y la interpretación de variantes genéticas en el contexto regional.

Objetivo: Analizar la aplicación actual de herramientas bioinformáticas en el estudio de variantes genéticas dentro de poblaciones de América Latina para identificar desafíos y oportunidades en la implementación de la genómica clínica local.

Métodos: Se realizó una revisión sistemática bajo la metodología PRISMA, consultando bases de datos como Scopus y Web of Science. Se seleccionaron y analizaron 61 artículos científicos publicados en el periodo comprendido entre 2021 y 2026.

Resultados: Se identificó una necesidad crítica de desarrollar flujos de trabajo bioinformáticos personalizados y fortalecer los biobancos regionales. Estas herramientas son esenciales para mejorar el filtrado de mutaciones y reducir la incidencia de variantes de significado incierto (VUS) mediante el uso de datos de ancestría local.

Conclusiones: La integración de la secuenciación de nueva generación (NGS) con contextos genómicos locales es fundamental para mitigar la desigualdad genómica. Se concluye que el desarrollo de infraestructura bioinformática propia es el camino para garantizar una medicina de precisión equitativa y soberana en la región.

Palabras Clave: Bioinformática, Variación Genética, Análisis de Secuencia de ADN, Medicina de Precisión, Grupos de Población clave.

Abstract:

Introduction: Latin American populations possess vast genomic diversity resulting from ancestral admixture; however, they remain markedly underrepresented in global genomic databases. This gap limits the effectiveness of precision medicine and the interpretation of genetic variants within the regional context.

Objective: To analyze the current application of bioinformatics tools in the study of genetic variants in Latin American populations in order to identify challenges and opportunities for the implementation of local clinical genomics.

Methods: A systematic review was conducted following the PRISMA methodology, consulting databases such as Scopus and Web of Science. A total of 61 scientific articles published between 2021 and 2026 were selected and analyzed.

Results: A critical need was identified to develop customized bioinformatics workflows and to strengthen regional biobanks. These tools are essential for improving mutation filtering and reducing the incidence of variants of uncertain significance (VUS) through the use of local ancestry data.

Conclusions: The integration of next-generation sequencing (NGS) with local genomic contexts is fundamental to mitigating genomic inequality. It is concluded that the development of proprietary bioinformatics infrastructure is key to ensuring equitable and sovereign precision medicine in the region.

Keywords: Bioinformatics, Genetic Variation, DNA Sequence Analysis, Precision Medicine, Key Population Groups.

1. INTRODUCCIÓN

En las últimas décadas, la patología molecular ha experimentado un avance sustancial gracias al crecimiento exponencial de las tecnologías de secuenciación genética, particularmente con la introducción de la secuenciación de nueva generación (NGS). La patología molecular ha avanzado sustancialmente en las últimas décadas gracias al crecimiento exponencial de las tecnologías de secuenciación genética, especialmente con la introducción de la secuenciación de nueva generación (NGS). Estas herramientas permiten la recopilación masiva de datos mediante un enfoque de alto rendimiento, interrogando múltiples genes de manera simultánea con mayor rapidez y a un costo reducido en comparación con los métodos tradicionales (1). La capacidad de identificar firmas genéticas específicas y correlacionarlas con los resultados del tratamiento ha impulsado la medicina de precisión, un proceso diagnóstico y terapéutico basado en biomarcadores genómicos aprobados. Sin embargo, la implementación clínica efectiva de estas tecnologías depende críticamente de una infraestructura bioinformática robusta capaz de procesar, alinear y analizar los archivos FASTQ para detectar variantes de nucleótido único (SNV) e inserciones y deleciones pequeñas con alta sensibilidad y especificidad (2).

Las poblaciones latinoamericanas poseen un trasfondo genético altamente diverso y mezclado, resultado de una compleja historia demográfica que involucra componentes ancestrales de origen nativo americano, europeo y africano. Los estudios de ancestría genómica han demostrado que esta composición varía significativamente entre países y regiones, con proporciones que pueden alcanzar hasta un 88% de ADN mitocondrial nativo americano en ciertas ciudades de Sudamérica (3). Esta heterogeneidad ancestral tiene implicaciones profundas en la genómica médica, ya que influye en la frecuencia de variantes genéticas asociadas a enfermedades cardiovasculares, susceptibilidad al cáncer y respuestas a tratamientos farmacológicos. Debido a esta estructura de población única, los perfiles genéticos locales no pueden asumirse a partir de datos obtenidos en otras latitudes, lo que exige un análisis específico que considere la deriva genética y los eventos de efecto fundador ocurridos durante el poblamiento de las Américas (4).

A pesar de esta rica diversidad, América Latina continúa estando severamente subrepresentada en los recursos genómicos globales, como el Atlas del Genoma del Cáncer (TCGA) y el Consorcio Internacional del Genoma del Cáncer (ICGC), donde las minorías étnicas, incluidos los hispanos, suelen estar ausentes o representadas en

proporciones mínimas. Se estima que solo el 1.3% de los estudios de descubrimiento y replicación de asociaciones genómicas (GWAS) se han realizado en estas poblaciones, lo que limita la capacidad de describir mutaciones somáticas con prevalencias bajas y superar el ruido de fondo genético en grupos específicos (1). Esta brecha de representación en las bases de datos de referencia globales, que priorizan históricamente a individuos de ascendencia europea, obstaculiza la interpretación precisa de variantes y puede exacerbar las disparidades en el diagnóstico, la eficacia del tratamiento y la evaluación de riesgos en la región.

El análisis bioinformático de variantes en pacientes latinoamericanos enfrenta desafíos técnicos adicionales debido a la ausencia de genomas de referencia que sean informativos de la variación genética específica de estas poblaciones. La falta de representación equilibrada de genomas nativos americanos en los paneles de imputación actuales puede inducir a una categorización falsa y a la sobreestimación de variantes somáticas en estudios de oncología (5). Además, la identificación de variantes de significado incierto (VUS) es un reto constante, especialmente en poblaciones menos caracterizadas donde la escasez de datos genómicos en bases públicas como ClinVar y gnomAD complica la interpretación clínica. Por tanto, es imperativo el desarrollo de flujos de trabajo bioinformáticos automatizados y personalizados que integren bases de datos de variantes germinales locales para mejorar la precisión del filtrado y la clasificación de mutaciones (6).

Para abordar la complejidad de la interpretación funcional, se han desarrollado diversos algoritmos y herramientas bioinformáticas que permiten anotar las consecuencias biológicas de las variaciones genéticas. Plataformas como Annovar, Variant Effect Predictor (VEP) y Pediatric Cancer Variant Pathogenicity Information Exchange (PeCanPIE) integran múltiples fuentes de información y guías del American College of Medical Genetics and Genomics (ACMG) para priorizar variantes candidatas (7). Estos procesos dependen de puntuaciones predictivas generadas por software como SIFT, que evalúa el impacto de las sustituciones de aminoácidos, y PolyPhen-2, que clasifica las mutaciones en categorías de daño potencial a la función proteica. Herramientas más recientes como CADD, REVEL y ClinPred combinan múltiples métricas genómicas para aumentar la robustez de las predicciones de patogenicidad, siendo fundamentales para evaluar el potencial driver de variantes raras en enfermedades complejas (8).

El impacto de estas variantes se manifiesta de manera crítica en diversas áreas de la medicina, como la farmacogenómica y la oncología de inicio temprano. Por ejemplo, variaciones moleculares en genes relacionados con la farmacocinética (ADME) determinan la variabilidad interindividual en la eficacia terapéutica y la aparición de reacciones adversas a medicamentos comúnmente prescritos, como los antineoplásicos o los tratamientos para la tuberculosis (9). En el cáncer de mama, se ha observado una susceptibilidad genética diferencial en mujeres jóvenes mexicanas, con una incidencia significativamente mayor que el promedio mundial, lo que sugiere la presencia de variantes germinales específicas aún no reportadas en poblaciones caucásicas. De manera similar, la detección de variantes accionables en genes como APOB o TTN en pacientes sudamericanos permite expandir los beneficios de las terapias dirigidas a una población más amplia, reduciendo las disparidades en el acceso a la medicina personalizada (10).

La implementación de estas tecnologías en los sistemas de salud latinoamericanos se ve limitada por barreras económicas y la escasez de profesionales capacitados en bioinformática. Los kits comerciales y los flujos de trabajo estándar suelen estar diseñados para cumplir con los estándares de países de ingresos altos, lo que a menudo no resulta costo-efectivo para la realidad regional. Sin embargo, el surgimiento de iniciativas transnacionales como el Proyecto CABANA y el desarrollo de biobancos nacionales, como el Biobanco Mexicano o el estudio CR-WGS en Costa Rica (11), están impulsando la capacidad local de generar y analizar datos genómicos a gran escala. Estos esfuerzos promueven un liderazgo científico regional y la construcción de bases de datos de referencia propias, elementos esenciales para garantizar que los beneficios de la era genómica se distribuyan de manera equitativa y beneficien a los grupos históricamente sub atendidos (4).

La aplicación de flujos de trabajo bioinformáticos ha permitido una caracterización sin precedentes de enfermedades monogénicas y raras en el contexto latinoamericano, revelando una arquitectura alélica distinta a la observada en poblaciones caucásicas. En el estudio de las retinopatías asociadas al gen ABCA4, como la enfermedad de Stargardt, se han identificado variantes fundadoras específicas, destacando la mutación p.Ala1773Val en México y perfiles diferenciales en Argentina y Brasil debido a sus composiciones ancestrales únicas (12). De manera similar, la detección de variantes noveles en genes como GCK e HNF1A en pacientes brasileños con diabetes monogénica subraya la necesidad de un cribado genético que contemple el

trasfondo de mezcla étnica local (13). En oncología de precisión, el desarrollo de paneles de NGS personalizados ha facilitado la identificación de biomarcadores accionables en tumores con alta prevalencia regional, como el cáncer de vesícula biliar, permitiendo el acceso a terapias dirigidas que antes no se consideraban para estos pacientes. Asimismo, la integración de análisis basados en ARN ha demostrado ser superior a los métodos tradicionales de ADN para detectar variantes críticas como el skipping del exón 14 de MET en cohortes sudamericanas de cáncer de pulmón de células no pequeñas (7). Estos hallazgos demuestran que la bioinformática no solo refina el diagnóstico, sino que también permite ajustar los algoritmos clínicos para reflejar la realidad genómica de la región.

Más allá de las enfermedades hereditarias, la bioinformática ha sido fundamental para fortalecer la vigilancia genómica y el estudio de la susceptibilidad del hospedero ante enfermedades infecciosas en América Latina. Iniciativas transnacionales como el proyecto CABANA han permitido rastrear la evolución de patógenos en tiempo real, identificando genotipos de SARS-CoV-2 que emergieron localmente, como las variantes Gamma en Brasil, Mu en Colombia y Lambda en Perú (11). El uso de paneles de NGS personalizados en poblaciones como la colombiana ha revelado factores genéticos del hospedero asociados con la severidad del COVID-19 y el desarrollo de síntomas a largo plazo, integrando datos clínicos y genómicos en modelos predictivos de riesgo. En el ámbito de las enfermedades metabólicas, el análisis de genomas completos en Costa Rica ha permitido identificar una carga significativa de variantes de riesgo para dislipidemias en genes como APOB, LPL y LDLR, muchas de las cuales presentan frecuencias alélicas distintas a las reportadas en bases globales (14). Estos estudios locales son esenciales para mejorar el rendimiento de la imputación genotípica, ya que la inclusión de genomas nativos americanos en los paneles de referencia aumenta drásticamente la precisión en la detección de variantes raras y específicas de la población. Por ende, la caracterización detallada de genes ADME mediante secuenciación de exoma completo en individuos mestizos ha evidenciado que las variantes raras son una fuente principal de la variabilidad interindividual en la respuesta farmacológica, lo que desafía el uso de microarreglos estándar diseñados para poblaciones europeas.

La justificación de esta investigación radica en la necesidad de cerrar la brecha de desigualdad genómica que afecta a las poblaciones latinoamericanas, las cuales permanecen severamente subrepresentadas en las bases de datos de referencia globales

diseñadas mayoritariamente para individuos de ascendencia europea. Debido a la compleja mezcla ancestral de la región, los algoritmos bioinformáticos estándar y los paneles de variantes comerciales suelen presentar limitaciones para identificar con precisión mutaciones locales, lo que deriva en diagnósticos erróneos o en la clasificación incorrecta de variantes de significado incierto. Al sistematizar el conocimiento sobre las herramientas bioinformáticas aplicadas a este contexto específico, este estudio busca fundamentar la creación de flujos de trabajo personalizados que reconozcan la arquitectura genética propia de la región. Esto no solo permitirá mejorar la eficacia de la medicina de precisión y la farmacogenómica en pacientes locales, sino que también fortalecerá la soberanía científica de América Latina, garantizando que los avances en secuenciación de nueva generación se traduzcan en beneficios equitativos para la salud pública regional.

El objetivo general de esta revisión es analizar el estado actual de la bioinformática aplicada al estudio de variantes genéticas en poblaciones latinoamericanas, evaluando las herramientas y flujos de trabajo existentes para identificar desafíos técnicos y brechas de representación genómica en la región. Como objetivos específicos, se busca categorizar los algoritmos y plataformas computacionales más utilizados para la anotación funcional de variantes en contextos de mezcla ancestral; examinar la eficacia de las bases de datos globales frente a los biobancos locales en la interpretación de variantes de significado incierto; y proponer estrategias bioinformáticas que mejoren la precisión del diagnóstico clínico y la medicina de precisión adaptada a la diversidad genética de los pacientes latinoamericanos.

2. METODOLOGÍA

Para asegurar la integridad de la revisión, se utilizó el modelo PRISMA, el cual permite organizar la búsqueda científica de manera minuciosa. El uso de esta estructura es fundamental para ofrecer claridad en el informe, permitiendo que los lectores confirmen la validez de los datos y la objetividad de los criterios de selección utilizados.

La búsqueda bibliográfica se llevó a cabo en Scopus y Web of Science, bases de datos seleccionadas por su prestigio y amplio catálogo de publicaciones arbitradas. Ambas plataformas garantizan un acceso integral a investigaciones de alto impacto, siendo referentes críticos para el análisis de la gestión hospitalaria y las ciencias de la salud a escala internacional. La ecuación de búsqueda para la base de datos Scopus:

(TITLE-ABS-KEY ("bioinformatics" OR "computational genomics" OR "genome analysis" OR "NGS" OR "next-generation sequencing") AND TITLE-ABS-KEY ("genetic variation" OR "genetic variants" OR SNP* OR mutation* OR polymorphism* OR "variant calling") AND TITLE-ABS-KEY ("Latin America" OR "Latinoamerica" OR "Hispanic population*" OR "admixed population*"))

En Scopus se obtuvieron 211 artículos.

Ecuación de búsqueda para la base de datos Web of Science:

TS = (("bioinformatics" OR "computational genomics" OR "genome analysis" OR "NGS" OR "next-generation sequencing") AND ("genetic variation" OR "genetic variants" OR SNP* OR mutation* OR polymorphism* OR "variant calling") AND ("Latin America" OR "Latinoamerica" OR "Hispanic population*" OR "admixed population*"))

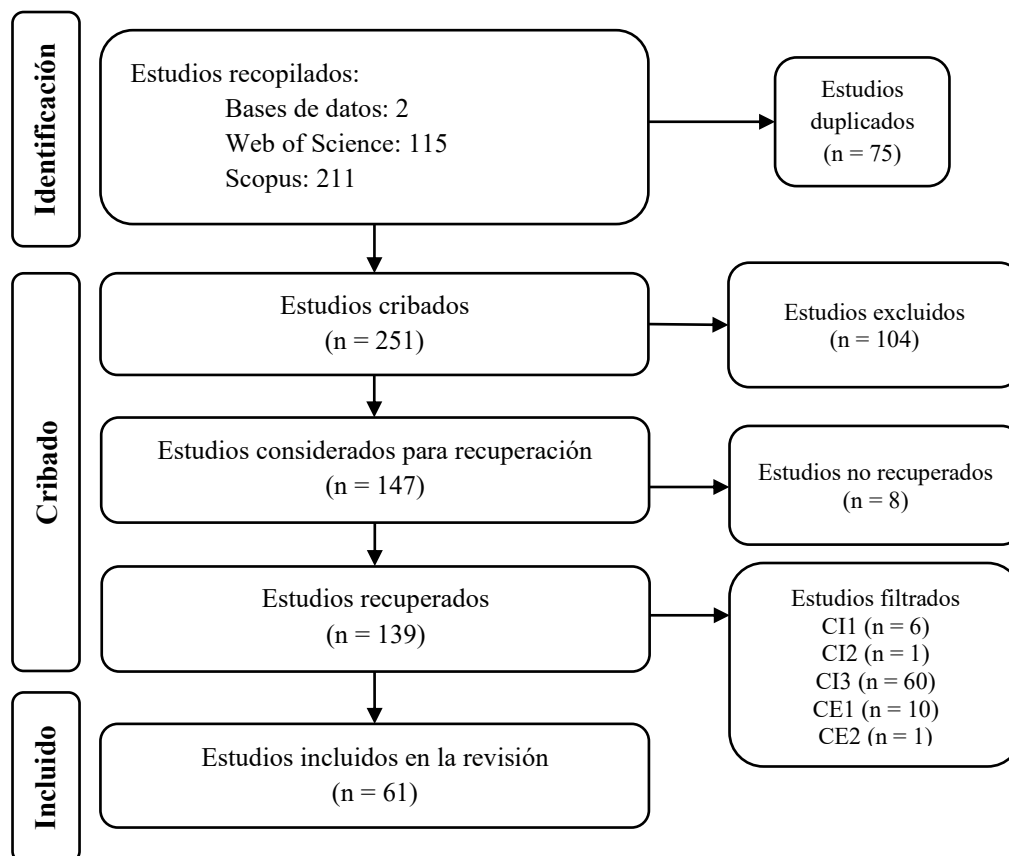
En Web of Science se obtuvieron 115 artículos.

Los artículos recopilados de ambas bases de datos fueron seleccionados mediante los siguientes criterios:

Tabla 4.
Criterio de inclusión y exclusión

Criterios de inclusión	Criterios de exclusión
CI1: Investigaciones que incluyan muestras de poblaciones de países latinoamericanos o individuos con ascendencia latina/hispana claramente definida.	CE1: Estudios que reporten variantes genéticas sin describir el uso de flujos de trabajo (pipelines) bioinformáticos o herramientas de análisis de datos masivos.
CI2: Estudios que utilicen herramientas bioinformáticas para identificar o caracterizar variantes genéticas.	CE2: Cartas al editor, editoriales, resúmenes de congresos, libros, notas técnicas y, muy importante, otras revisiones sistemáticas o bibliométricas.
CI3: Estudios publicados entre 2021 y 2026.	CE3: Artículos no publicados en español, portugués e inglés.

Figura 5.
Flujograma Prisma de selección de investigaciones para revisión



Tras aplicar los criterios de selección, se determinó la elegibilidad de 61 artículos para un análisis detallado. Con el fin de presentar la evidencia de forma transparente, los hallazgos se organizaron en la Tabla 2, la cual sistematiza variables críticas como el objetivo, el diseño metodológico, el país de origen y las conclusiones principales. Esta estructura facilita tanto la interpretación individual de los estudios como la síntesis global de los resultados.

Tabla 5.
Investigaciones analizadas

N.º	Autores	Objetivos	Hallazgo
1	Rhead <i>et al.</i> (15)	Analizaron mutaciones somáticas y expresión génica en tumores colorrectales de 8,454 pacientes según ascendencia y etnia.	La ascendencia africana se asoció a más mutaciones en APC, KRAS y PIK3CA, menos en BRAF y mayor tasa del subtipo CMS3.
2	Salvo <i>et al.</i> (1)	Validaron un panel NGS de 25 genes en tejidos de parafina para detectar mutaciones accionables en tumores sólidos de Latinoamérica.	El flujo de trabajo detectó más de 50 biomarcadores en 272 muestras con un 100% de concordancia y sensibilidad del 1%.
3	Angulo-Aguado <i>et al.</i> (2)	Usaron un panel NGS personalizado para caracterizar factores genéticos de severidad y	La variante rs11385942 permitió crear modelos predictivos de gravedad (AUC 0.86) y COVID prolongado (AUC 0.83).

		COVID prolongado en una muestra colombiana.	
4	Gionda <i>et al.</i> (16)	Realizaron la secuenciación de genomas completos de los genotipos F y H del virus de la hepatitis B en Brasil y México.	Se secuenciaron 18 genomas completos, detectando variantes en genes preS/S en 7 casos y en la región precore/core en 11.
5	Molina-Mora <i>et al.</i> (11)	Secuenciaron 1,534 genomas de siete países para la vigilancia del SARS-CoV-2 en América Latina durante 2021.	La frecuencia de mutaciones en la región S1 aumentó significativamente hasta alcanzar 28 por mes a inicios de 2022.
6	Mena <i>et al.</i> (12)	Describieron el espectro molecular de la enfermedad de Stargardt en 95 probandos argentinos mediante el análisis de cuatro genes.	Se hallaron variantes causales en ABCA4 en el 77% de los probandos, identificando 173 variantes totales y 9 novedosas.
7	Sohail <i>et al.</i> (4)	Geno tipificaron 6,057 individuos de las 32 entidades de México para analizar su historia demográfica y rasgos complejos.	El Biobanco Mexicano predice mejor los rasgos poblacionales que el del Reino Unido mediante el análisis de tamaños ancestrales.
8	Cecchini <i>et al.</i> (17)	Usaron NGS para describir la resistencia a INSTI en una cohorte de 56 mujeres embarazadas con VIH en Argentina.	Con sensibilidad del 1%, la resistencia principal subió al 8.9% y la accesoria al 23.2% con cargas mayores a 1,000 c/mL.
9	Fernández-Gajardo <i>et al.</i> (3)	Investigaron impulsores moleculares de glioblastoma y astrocitomas de alto grado en 70 pacientes chilenos.	Se confirmaron marcadores IDH y TERTp, se hallaron mutaciones en TP53 y PTEN, y una baja frecuencia en NF1 ($p < 0.01$).
10	Chacon-Camacho <i>et al.</i> (18)	Analizaron el genotipo de degeneraciones retinianas vinculadas a ABCA4 en 211 casos índice no relacionados en México.	Se logró una tasa diagnóstica del 86.47% con 128 variantes en ABCA4, incluyendo 22 nuevas y la mutación fundadora p.Ala1773Val.
11	Silgado-Guzmán <i>et al.</i> (19)	Estudiaron la variabilidad de 35 genes del panel ADME mediante la secuenciación del exoma completo de 509 colombianos.	Las variantes raras de sentido erróneo fueron 2.1 veces más frecuentes, revelando 89 variaciones farmacogenéticas en 28 genes.
12	Cadena-Ullauri <i>et al.</i> (20)	Analizaron variantes de APOB y ascendencia en 60 pacientes ecuatorianos con afecciones cardíacas hereditarias.	Se identificaron 227 variantes en APOB (principalmente benignas) en una población con 58.3% de ascendencia nativa y 32.9% europea.
13	De Oliveira Cavagna <i>et al.</i> (21)	Evaluaron asociaciones clinicopatológicas en 1,131 pacientes brasileños con cáncer de pulmón mediante perfiles de ascendencia.	El 88.0% de los pacientes tuvo alteraciones oncogénicas, siendo TP53 (58.0%) la mutación asociada a menor supervivencia.
14	Silva <i>et al.</i> (22)	Desarrollaron paneles multiétnicos para la imputación de alelos HLA clásicos y no clásicos en poblaciones mezcladas.	El modelo con 3,675 sujetos superó al Michigan Imputation Server en la precisión de predicción de alelos HLA-B en brasileños.
15	Cadena-Ullauri <i>et al.</i> (10)	Compararon la concordancia de variantes entre tejido y ADN tumoral circulante en 36 mujeres ecuatorianas con cáncer de tiroides.	Se detectaron 71 variantes con una concordancia del 81.69% entre tejido y ctDNA, destacando TP53 como el gen más mutado.
16	Cancela <i>et al.</i> (23)	Monitorearon la abundancia de variantes de SARS-CoV-2 en Uruguay mediante secuenciación de la región RBD en aguas residuales.	La vigilancia detectó la variante Gamma semanas antes del reporte clínico y la Delta a frecuencias del 1% al 3% desde abril de 2021.

17	Cadena-Ullauri <i>et al.</i> (24)	Realizaron un cribado genómico para condiciones cardíacas en dos parientes ecuatorianos para optimizar su manejo clínico.	Se identificó por primera vez en Latinoamérica la mutación Arg302Gln en PRKAG2 en dos sujetos con miocardiopatía y Wolff-Parkinson-White.
18	Douillard <i>et al.</i> (25)	Evaluaron la precisión de 1000 Genomes como panel de referencia para la imputación de HLA en ancestrías mezcladas.	El set de 2,504 individuos logró una puntuación F1 de 0.66 en HLA-B, superando modelos poblacionales (0.53 frente a 0.42).
19	Montella <i>et al.</i> (26)	Analizaron retrospectivamente las mutaciones de EGFR en 7,413 pacientes brasileños con cáncer de pulmón (NSCLC).	La detección de mutaciones en EGFR fue del 24.24%, con la delección del exón 19 (12.8%) como la más prevalente.
20	Nóbrega <i>et al.</i> (27)	Informaron 9 casos de lipofuscinosis ceroides neuronal tipo 11 para profundizar en el fenotipo genético latinoamericano.	El 100% de los pacientes mostró deterioro visual y ataxia, hallando la variante p.(Gln257ProfsTer27) en 7 brasileños y una nueva en 1 colombiano.
21	Luce <i>et al.</i> (28)	Caracterizaron el gen DMD en 400 pacientes argentinos para identificar candidatos a terapias de salto de exón y lectura de codones.	Se logró una detección del 97%, hallando delecciones en el 56,6% de los casos y potencial de restauración del marco de lectura en el 87,5%.
22	Sohail <i>et al.</i> (29)	Utilizaron el Biobanco Mexicano para revelar ancestrías de escala fina y su impacto en rasgos complejos de la población.	El análisis de 6,000 muestras asoció ancestrías de las Américas con menor estatura y mayores niveles de glucosa y triglicéridos.
23	Cadena-Ullauri <i>et al.</i> (30)	Realizaron el análisis genómico de una joven ecuatoriana con trastornos de conducción cardíaca sin antecedentes familiares.	Se identificó la variante inédita p.Arg893Cys en SCN5A en una paciente con 43,9% de ancestría europea y 35,7% nativa americana.
24	Giacomazzi <i>et al.</i> (31)	Evaluaron factores de riesgo y fenotipos de cáncer hereditario en pacientes jóvenes del sur de Brasil.	El historial familiar elevó el riesgo de cáncer (OR 6,1) y el 32% de las pacientes menores de 50 años portaban variantes patogénicas germinales.
25	Brito <i>et al.</i> (32)	Caracterizaron hallazgos clínicos y moleculares en 28 pacientes chilenos con sospecha de síndrome de Rett.	Se confirmó síndrome de Rett en el 39,2% de los participantes (MECP2) y se hallaron diagnósticos alternativos en otros genes en un 14,2% adicional.
26	Romer <i>et al.</i> (33)	Establecieron un perfil de epítomos del antígeno TSSA de <i>T. cruzi</i> para optimizar la serotipificación del Chagas.	Se identificaron tres epítomos discretos, 9 variantes nuevas de TSSA y se registró el seroreconocimiento de <i>T. cruzi</i> marinkellei en un paciente.
27	Alvarado-Orosco <i>et al.</i> (34)	Descifraron la resistencia de 111 aislados de <i>Pseudomonas aeruginosa</i> en Bolivia entre 2023 y 2024.	El 60,3% de los aislados portaba carbapenemasas, destacando el clon ST277 (35%), con una sensibilidad a colistina del 98,2%.
28	Cuadros-Espinoza <i>et al.</i> (35)	Evaluaron el poder de estadísticas de neutralidad para detectar selección natural en 15 poblaciones mundiales.	Se validó que las estadísticas Fadm y LAD detectan eficazmente la mezcla adaptativa en poblaciones de África Occidental, Madagascar y el Sahel.
29	Guevara-Ramírez <i>et al.</i> (36)	Describieron variantes del gen TTN y ancestría en 60 pacientes ecuatorianos con cardiopatías.	Se hallaron 29 variantes de interés en TTN en una población con 63,1% de ancestría nativa, sugiriendo al gen como modulador de enfermedad.
30	Aguilar-Zamora <i>et al.</i> (37)	Estudiaron la epidemiología molecular y resistencia de Clostridioides en hospitales de la Ciudad de México.	Se identificaron 20 tipos de secuencia en 94 aislados, predominando el ST1 (63%) y una multiresistencia antibiótica del 65%.
31	Rivas <i>et al.</i> (7)	Analizaron 1,560 biopsias de tumores de pulmón en Chile,	Se halló una prevalencia de variantes de MET del 8%, donde el análisis de ARN

		Brasil y Perú para detectar la omisión del exón 14 de MET.	detectó un 27% más de casos que el de ADN.
32	Abreu <i>et al.</i> (13)	Evaluaron genes de diabetes monogénica en 57 pacientes brasileños y 34 familiares mediante secuenciación de Sanger.	Se identificaron 20 pacientes con variantes patogénicas (13 GCK-MODY y 4 HNF1A-MODY), incluyendo cuatro variantes nuevas.
33	Jiménez-Kaufmann <i>et al.</i> (5)	Exopandieron el panel de 1000 Genomes con 134 genomas de nativos americanos para mejorar la imputación de genotipos.	La inclusión de estos genomas aumentó los SNP de calidad (ej. 3,269 en Perú) y mejoró la imputación en tratos de ascendencia nativa.
34	Jaramillo-Valverde <i>et al.</i> (38)	Estimaron la prevalencia de genotipos NAT2 y CYP2E1 en peruanos para asociarlos al tratamiento de tuberculosis.	En 472 participantes, se hallaron 36 haplotipos de NAT2 en mestizos y 6 en nativos, vinculando variantes específicas con el riesgo de lesión hepática.
35	Leitão <i>et al.</i> (8)	Investigaron exomas de 18 genes de susceptibilidad a leucemia linfoblástica en 64 amerindios brasileños.	Se identificaron 125 variantes en amerindios brasileños, detectando siete nuevas y seis asociadas a la susceptibilidad de desarrollar LLA.
36	Molina-Mora <i>et al.</i> (39)	Analizó la diversidad de SARS-CoV-2 en Costa Rica mediante 185 secuencias obtenidas en 2020.	El análisis reveló 283 variantes de nucleótidos y un aumento de la variante de espiga T1117I al 29,2% en noviembre de 2020.
37	Paixão <i>et al.</i> (6)	Evaluaron mutaciones de línea germinal en 321 pacientes brasileñas con cáncer de mama usando un panel multigenético.	Se detectaron variantes patogénicas en el 25,2% de las pacientes, con mayor frecuencia en el subtipo triple negativo (31,6%).
38	Mueller <i>et al.</i> (40)	Analizaron microARN en suero de 228 pacientes de Chile, Bolivia e Italia para predecir miocardiopatía chagásica.	Se identificaron tres microARN con niveles bajos en pacientes con ECG anormal ($p < 0.0002$) que correlacionan con la gravedad de Kuschnir.
39	Zuccherato <i>et al.</i> (41)	Evaluaron el espectro de variantes de F8 en 265 brasileños con hemofilia A grave y su relación con inhibidores.	Se hallaron variantes deletéreas en el 98,1% de los pacientes y una asociación del dominio A3 con el desarrollo de inhibidores (21,0%; $p < 0,001$).
40	Uricoechea Patiño <i>et al.</i> (42)	Analizaron el ADN mitocondrial de 33 restos antiguos para delinear la genética de tres grupos prehispánicos colombianos.	Se identificaron los cuatro haplogrupos fundadores (A, B, C, D) y 21 sub-haplogrupos en restos antiguos, confirmando una alta heterogeneidad.
41	Melendez-Montañez <i>et al.</i> (43)	Investigaron la predisposición genética a obesidad e hiperfagia en 103 niños puertorriqueños mediante tamizaje de variantes raras.	Se halló una prevalencia del 73% de variantes reportadas, con un 7% de variantes patogénicas por ciliopatías y un 90% de variantes de significado incierto.
42	Barra-Matos <i>et al.</i> (44)	Caracterizaron la carga de INDEL mitocondriales en 179 mestizos de la Amazonía brasileña con enfermedad de Parkinson.	Pacientes con Parkinson mostraron una carga de INDEL significativamente mayor (OR = 9,23; FDR = 0,044), incluyendo siete nuevas mutaciones en el gen ND5.
43	Gómez-Flores-Ramos <i>et al.</i> (45)	Analizaron variantes patogénicas en 862 genes del exoma germinal de 115 mexicanas jóvenes con cáncer de mama.	Se identificaron 49 variantes patogénicas en el 34% de las pacientes, afectando a 40 genes como BRCA1/2 y otros poco reportados en la región.
44	Galán-Huerta <i>et al.</i> (46)	Describieron la prevalencia de variantes VOC de SARS-CoV-2 en infecciones posvacunación en Monterrey durante abril-agosto de 2021.	La variante Delta predominó en el 67,92% de infecciones posvacunación, con el linaje B.1.617.2 como el más común en casos ambulatorios ($p < 0.01$).
45	Gondola <i>et al.</i> (47)	Determinaron genotipos y mutaciones de resistencia de C.	Se identificaron cinco genotipos, siendo el D (47%) el más prevalente, y una triple

		trachomatis en orina de 166 adolescentes de Panamá.	mutación en rplV en el 90% de las muestras.
46	Martínez <i>et al.</i> (48)	Evaluaron el papel de genes de resistencia a metales pesados en la divergencia del linaje ST5-SCCmecI de <i>S. aureus</i> .	Se identificó el plásmido pSCL4752 en el 66% de los aislados, con una frecuencia mayor en Concepción (88%) que en Santiago (29%).
47	Valverde-Hernández <i>et al.</i> (14)	Identificaron variantes en 69 genes metabólicos en costarricenses para cuantificar la carga genética de dislipidemias.	Se detectaron 2,600 variantes polimórficas, de las cuales 18 pueden alterar la función génica y 9 tienen implicaciones farmacogenómicas o protectoras.
48	Chen <i>et al.</i> (49)	Exploraron la diversidad genómica de tres serotipos de <i>Salmonella enterica</i> en aguas de Chile, México y Brasil.	Se halló una correlación fuerte ($R > 0.90$, $P < 0.01$) entre integrones y multiresistencia en Chile y México, pero no se detectaron en Brasil.
49	Siegel <i>et al.</i> (50)	Describieron un marco de descubrimiento de microhaplotipos para capturar la identidad por descendencia en <i>Plasmodium vivax</i> .	Un panel de 100 microhaplotipos con alta diversidad ($HE = 0.70-0.81$) identificó el 89% de las infecciones policlonales de <i>P. vivax</i> .
50	Marín <i>et al.</i> (51)	Describieron las características de seis casos de <i>N. meningitidis</i> resistente a antibióticos en El Salvador.	El análisis mostró que seis aislados compartían el ST3587 y el gen ROB-1, resultando en desenlaces fatales para dos de los pacientes.
51	Spangenberg <i>et al.</i> (52)	Investigaron la ancestría indígena uruguaya mediante secuenciación del genoma completo de individuos con ancestría Charrúa.	Se estimó que el primer contacto europeo con el grupo Charrúa ocurrió hace 11 generaciones, situándolo a mediados del siglo XVII.
52	Chen <i>et al.</i> (53)	Desarrollaron un modelo de supermezcla para medir el parentesco individual en poblaciones con antecedentes mezclados.	El modelo demostró que el parentesco en poblaciones americanas cambia de forma fluida entre individuos debido a eventos históricos de mezcla.
53	Zambrano-O <i>et al.</i> (54)	Compararon la precisión de un panel multigenético frente al genoma completo para inferir ancestría en mujeres colombianas.	Se reportó una correlación de 0.96 para las ancestrías, con una media de 45.7% europea, 46.2% nativa americana y 8.11% africana.
54	Bispo <i>et al.</i> (2026) (55)	Analizaron el impacto de la ancestría y el mitogenoma en la respuesta a levodopa en pacientes brasileños con Parkinson.	La ancestría mitocondrial nativa (macrohaplogrupo C) fue más frecuente en respondedores a levodopa, con alta heteroplasmia en MT-ND5.
55	Deiana <i>et al.</i> (56)	Realizaron la caracterización genómica completa del primer aislado europeo del virus Oropouche importado desde Cuba.	El análisis determinó que el aislado es un virus reordenado que comparte alta similitud con el nuevo clúster OROV_SCDC_2024.
56	Chen <i>et al.</i> (57)	Examinaron la diversidad y persistencia de <i>Salmonella</i> en aguas agrícolas de Brasil, Chile y México.	<i>Salmonella</i> se aisló en el 45.8% de 3,291 muestras de agua, con 128 serovares y resistencia multidroga en el 47.4% de los casos.
57	Jaramillo-Valverde <i>et al.</i> (9)	Determinaron la prevalencia de genotipos metabolizadores y riesgos de lesión hepática en poblaciones nativas peruanas.	La población amazónica mostró mayor riesgo de lesión hepática por alelos en AGLB4 ($OR = 3.83$) y SLCO1B1 ($OR = 3.81$).
58	Nøhr <i>et al.</i> (58)	Desarrollaron un método de máxima verosimilitud para estimar parentesco usando secuenciación de baja profundidad (1x).	El método propuesto superó a herramientas existentes logrando estimaciones precisas de parentesco con una profundidad de solo 1x.
59	Fu <i>et al.</i> (59)	Propusieron un método de inferencia de ancestría local basado en polimorfismos de	El método ASA logró alta concordancia en mezclas recientes y superó a otros en casos

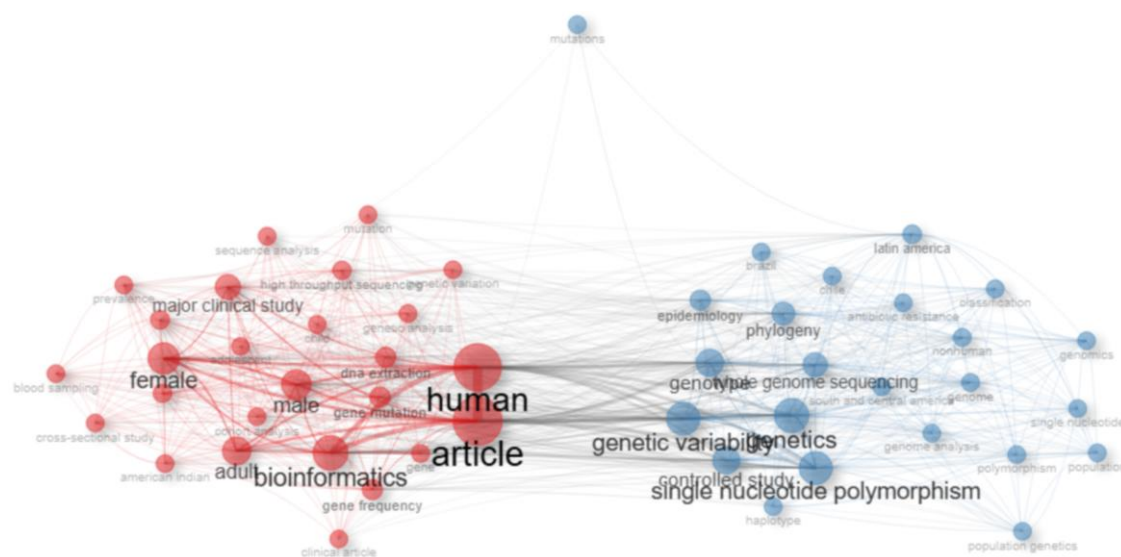
		nucleótido único (SNP) de mezclas antiguas con referencias poblacionales.	de mezclas antiguas con referencias contemporáneas deficientes.
60	Honorato-Mauer <i>et al.</i> (60)	Caracterizaron factores demográficos y de paneles de referencia que afectan la precisión de la ancestría local en mezclas.	Los fragmentos indígenas tuvieron menor precisión de inferencia (88%-94%) que los europeos y africanos (96%-99%) por errores de asignación.
61	Esteves <i>et al.</i> (61)	Caracterizaron proteínas diferenciales en cepas de <i>L. braziliensis</i> de lesiones típicas y atípicas en indígenas de Brasil.	Se identificaron 4,048 proteínas, hallando una mayor abundancia de enzimas antioxidantes en cepas de lesiones atípicas de <i>Leishmania</i> .

3. RESULTADOS Y DISCUSIÓN

3.1. Análisis bibliométrico

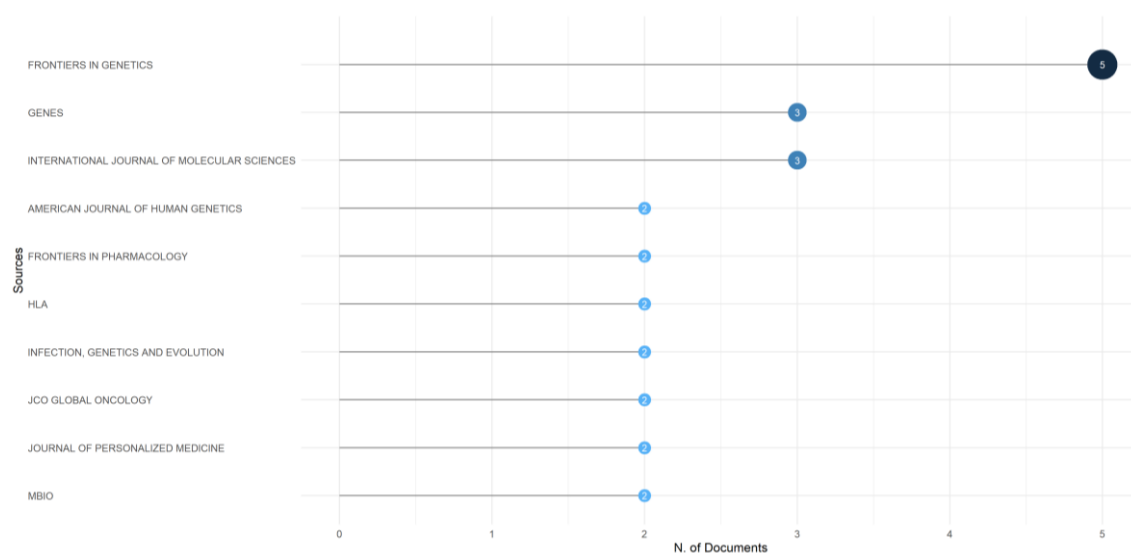
Se obtuvieron nodos temáticos críticos que revelan una transición hacia la medicina personalizada, en esta red, destaca una fuerte asociación entre la secuenciación de nueva generación (NGS) y el análisis de variantes genéticas, donde conceptos como high-throughput nucleotide sequencing y computational biology actúan como puentes metodológicos fundamentales para el estudio de la diversidad genómica. La Figura 2 evidencia que el flujo de trabajo bioinformático no solo se limita al procesamiento técnico de datos, sino que se integra con aplicaciones clínicas directas, vinculando la identificación de polimorfismos de un solo nucleótido (SNP) con la susceptibilidad a enfermedades y la farmacogenética. Asimismo, la densidad de las conexiones sugiere un interés creciente en la ancestría y la estructura poblacional, elementos que son determinantes para refinar los algoritmos de interpretación de variantes en grupos tradicionalmente subrepresentados.

Figura 2.
Mapa de co-ocurrencia de términos



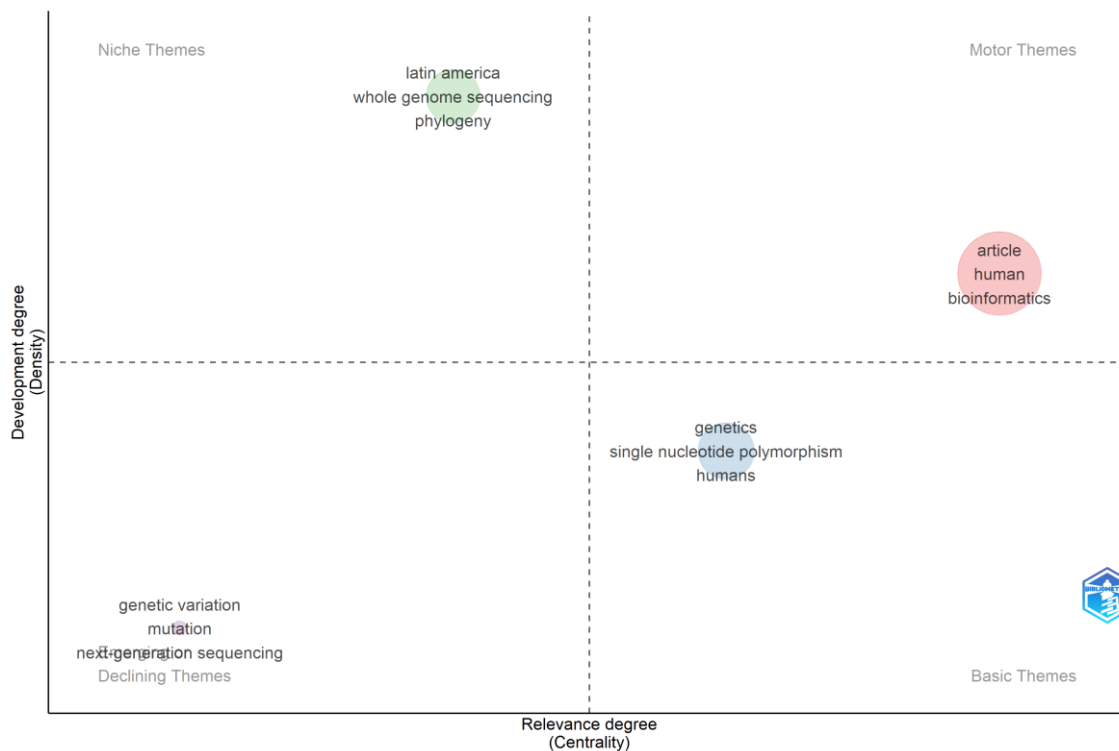
Respecto a la distribución editorial de la producción científica, se observa una concentración significativa en revistas de alto impacto con un enfoque interdisciplinario y de acceso abierto, como se detalla en la Figura 3. Revistas como *Frontiers in Genetics* y *Genes* lideran el volumen de publicaciones, esta hegemonía de fuentes especializadas en genética molecular y sistemas biológicos subraya la consolidación de la bioinformática no solo como una herramienta de soporte, sino como un eje central en la literatura científica contemporánea para la validación de hallazgos en medicina de precisión.

Figura 3.
Producción científica

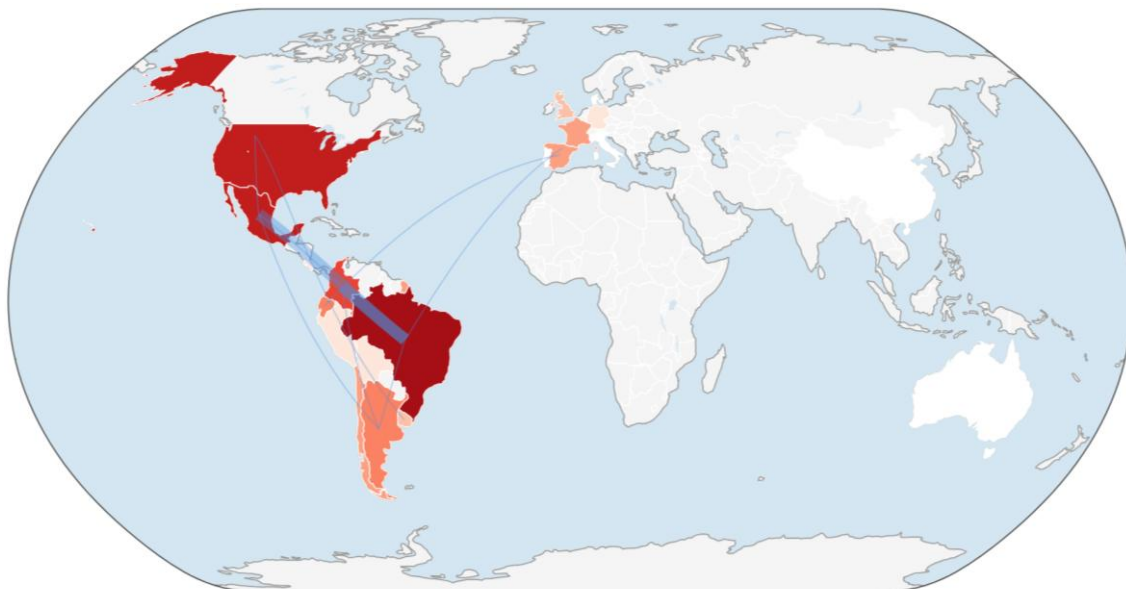


En la Figura 4, el cuadrante de temas motores, destaca la convergencia entre la secuenciación de nucleótidos de alto rendimiento y la biología computacional, consolidándose como los pilares técnicos más robustos y centrales de la disciplina. Por otro lado, el análisis de variantes genéticas y el estudio de polimorfismos de nucleótido único (SNP) se posicionan como temas básicos, lo que indica que, aunque son fundamentales y transversales para toda la literatura analizada, requieren de una integración continua con nuevas metodologías para no perder su capacidad de innovación. El surgimiento de nichos específicos relacionados con la ancestría y la estructura poblacional sugiere una especialización creciente, orientada a cerrar la brecha de datos genómicos en poblaciones diversas, permitiendo que el flujo de trabajo bioinformático transite desde la estandarización técnica hacia una aplicación clínica con mayor pertinencia local.

Figura 4.
Mapa temático

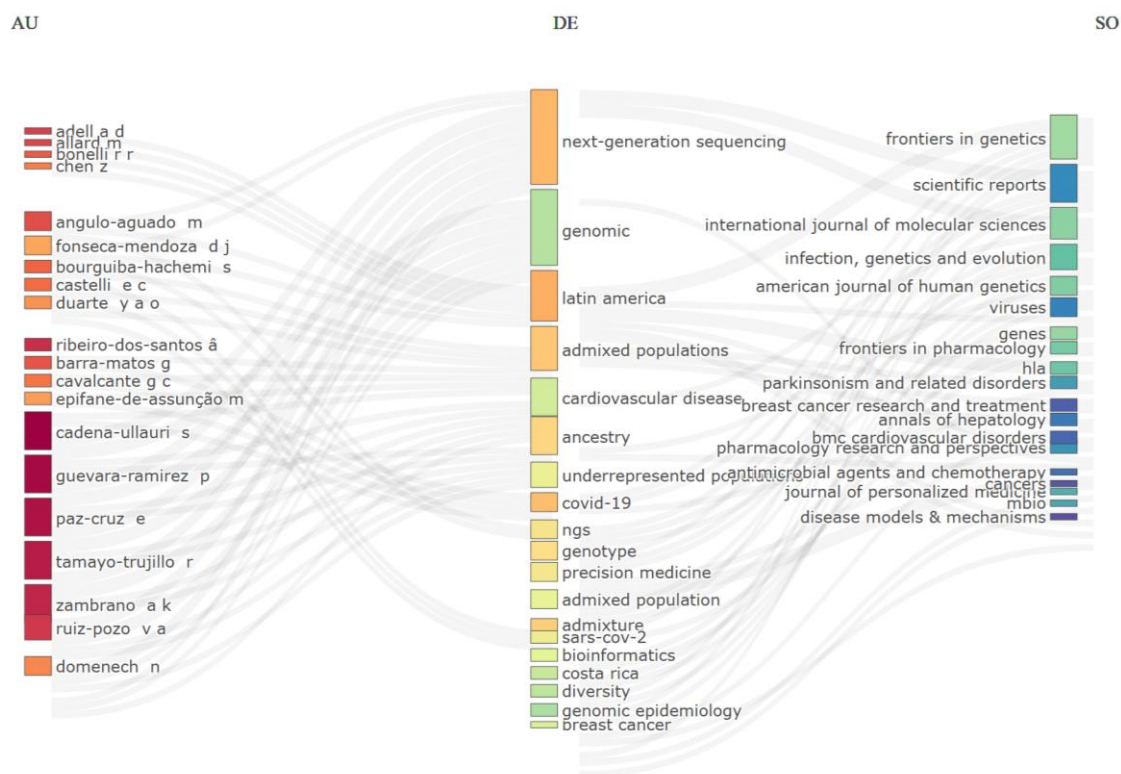


La red de cooperación internacional, representada en el mapa de colaboración de la Figura 5, revela un ecosistema científico globalizado con nodos de producción centralizados y puentes estratégicos de conocimiento. Se observa que Brasil y México actúan como los principales núcleos de colaboración, estableciendo redes densas que conectan con varios países, lo cual es indicativo de su liderazgo en infraestructura tecnológica y financiamiento para proyectos genómicos a gran escala. Esta dinámica de colaboración no solo facilita el intercambio de datos bioinformáticos y metodologías avanzadas de secuenciación, sino que también subraya la necesidad de fortalecer la autonomía científica regional para garantizar que los hallazgos en medicina de precisión respondan de manera específica a la diversidad y estructura poblacional propia de América Latina.

Figura 5.*Mapa de colaboración entre países*

La integración sistémica de los componentes de la investigación se visualiza en el diagrama de tres campos de la Figura 6, que establece el flujo relacional entre los autores principales (izquierda), las palabras clave (centro) y las fuentes editoriales (derecha). Este análisis permite identificar que las palabras clave actúan como el eje articulador de la disciplina, donde términos como admixed populations, ancestry y genetics conectan el trabajo de autores referentes como Albrechtsen A., Moltke I. y Belanger S. con las revistas de mayor impacto en el área. Se observa una convergencia masiva de estos flujos temáticos hacia publicaciones como Genes, Frontiers in Genetics y Scientific Reports, lo que demuestra que la producción científica está altamente cohesionada en torno a la genética de poblaciones y la biología computacional.

Figura 6.
Three-Field Plot



3.2. Discusión

La discusión sobre el estado actual de la bioinformática en Latinoamérica revela una transición crítica hacia la medicina de precisión, impulsada por la consolidación de infraestructuras para el análisis de secuenciación de nueva generación (NGS). Los hallazgos demuestran que, si bien se han validado flujos de trabajo y paneles personalizados que logran una alta sensibilidad diagnóstica en tumores sólidos de la región (1), el principal desafío técnico continúa siendo la profunda brecha de representación en los recursos genómicos globales. Esta subrepresentación limita la precisión en la interpretación de variantes en poblaciones con mezclas ancestrales complejas, lo que ha hecho imperativo el desarrollo de estrategias para expandir las bases de datos internacionales con genomas de referencia nativos americanos (5). Asimismo, la evaluación de las herramientas existentes subraya que el fortalecimiento de los biobancos locales es esencial, dado que estos recursos presentan una capacidad predictiva de rasgos poblacionales superior a los modelos diseñados para poblaciones europeas (4). Finalmente, la evolución de las estrategias bioinformáticas regionales, que ahora incluyen la vigilancia genómica en tiempo real y la integración de análisis de ARN para mejorar

la detección de variantes críticas (11), evidencia un esfuerzo coordinado por superar las limitaciones técnicas y garantizar una aplicación clínica equitativa adaptada a la diversidad genética latinoamericana.

La categorización de las herramientas computacionales identificadas en esta revisión revela una evolución desde el uso de plataformas de anotación estándar hacia metodologías diseñadas específicamente para abordar la complejidad de las poblaciones mezcladas. Mientras que algoritmos como Annovar, Variant Effect Predictor (VEP), SIFT y PolyPhen-2 se mantienen como los pilares fundamentales para la predicción del impacto funcional y la patogenicidad, su aplicación en Latinoamérica se ve significativamente fortalecida mediante la integración de modelos de ancestría y paneles de referencia personalizados. En este sentido, se observa que el desarrollo de paneles multiétnicos para la imputación de alelos complejos supera en precisión a servidores globales convencionales (22). Asimismo, la expansión de recursos internacionales, como el panel de 1000 Genomes, mediante la inclusión de genomas nativos americanos, ha demostrado ser una estrategia eficaz para mejorar la calidad de la imputación en tractos de ascendencia indígena, mitigando las limitaciones de los paneles de referencia que no representan adecuadamente la diversidad regional (25).

Por otro lado, la inferencia de ancestría local (LAI) se posiciona como una categoría crítica para la interpretación funcional precisa. Algoritmos como el método ASA han mostrado una alta concordancia en el análisis de mezclas recientes, superando a herramientas previas en escenarios con referencias contemporáneas limitadas (59). Sin embargo, persisten desafíos técnicos notables, ya que la precisión de la inferencia suele ser menor en fragmentos de origen indígena que en los de origen europeo o africano, debido a errores de asignación en los paneles actuales. Complementariamente, el uso de modelos de supermezcla (57) y plataformas para la estimación de parentesco en datos de baja profundidad, como NGSremix (58), junto con el empleo de estadísticas de neutralidad como Fadm y LAD para detectar mezcla adaptativa, conforman un ecosistema bioinformático robusto que permite transitar de una anotación genómica general a una interpretación funcional con pertinencia poblacional.

La evaluación de la eficacia de los recursos genómicos subraya una limitación crítica en la interpretación clínica para Latinoamérica: la dependencia de bases de datos globales que, al estar centradas en poblaciones de ascendencia europea, exacerban la clasificación de variantes de significado incierto (VUS). La evidencia analizada

demuestra que esta falta de representatividad en repositorios internacionales como ClinVar y gnomAD complica el filtrado de mutaciones, alcanzando niveles alarmantes donde hasta el 90% de las variantes detectadas en ciertas cohortes del Caribe carecen de una interpretación clínica definitiva (43). En contraste, el desarrollo de biobancos locales ha demostrado una eficacia superior para la medicina personalizada; por ejemplo, el Biobanco Mexicano exhibe una mayor capacidad para predecir rasgos poblacionales complejos que el Biobanco del Reino Unido, gracias a un análisis de ancestría de escala fina que captura la arquitectura genética propia de la región.

Esta superioridad diagnóstica de los datos locales se ve reforzada por la identificación de variantes patogénicas en genes que están poco reportados en la literatura global, pero que presentan una prevalencia significativa en pacientes latinoamericanos (45). Asimismo, se ha validado que la inclusión de genomas nativos americanos en los paneles de referencia aumenta drásticamente la precisión de la imputación y permite detectar variantes polimórficas con implicaciones farmacogenómicas que los microarreglos estándar suelen omitir. En conjunto, estos hallazgos confirman que, aunque las bases de datos globales son puntos de partida necesarios, solo la integración de datos genómicos locales puede garantizar la equidad y precisión requeridas para la aplicación clínica en la región.

La propuesta de estrategias bioinformáticas para fortalecer la medicina de precisión en Latinoamérica debe fundamentarse en la transición de flujos de trabajo genéricos hacia modelos ancestry-aware. Los hallazgos de esta revisión sugieren que una estrategia prioritaria es la implementación de grafos de referencia pangénomicos, los cuales permiten capturar variaciones estructurales y alelos específicos de poblaciones indígenas y afrodescendientes que son ignorados por el genoma de referencia lineal actual (5). Complementariamente, la integración de la vigilancia genómica mediante plataformas de análisis en tiempo real y el uso de técnicas de supermezcla para el ajuste de puntajes de riesgo poligénico (PRS) representan herramientas esenciales para mejorar la estratificación de pacientes en contextos de alta heterogeneidad.

Asimismo, el éxito del diagnóstico clínico depende de la creación de consorcios regionales que estandaricen el intercambio de datos a través de biobancos locales interconectados. Esta infraestructura permitiría reducir la tasa de variantes de significado incierto (VUS) al contrastarlas con frecuencias alélicas propias de la región, superando la eficacia de los repositorios globales (43). Finalmente, se propone que la precisión

diagnóstica se potencie mediante la combinación de datos genómicos con análisis transcriptómicos (RNA-seq) para validar el impacto funcional de variantes raras, asegurando así que las intervenciones terapéuticas estén verdaderamente alineadas con la arquitectura genética única de los pacientes latinoamericanos (7).

4. CONCLUSIONES

La presente revisión permite concluir que la bioinformática aplicada a poblaciones latinoamericanas se encuentra en un punto de inflexión, donde la transición hacia una medicina de precisión efectiva depende de la superación de la brecha de representación genómica regional. Se ha determinado que, si bien existen algoritmos robustos para la anotación funcional, su eficacia en contextos de mezcla ancestral está condicionada a la integración de modelos de ancestría local y paneles de referencia que incluyan componentes nativos americanos y afrodescendientes. Asimismo, se evidencia que las bases de datos globales resultan insuficientes para una interpretación clínica precisa, siendo los biobancos locales las herramientas con mayor capacidad predictiva para reducir la incidencia de variantes de significado incierto. En consecuencia, la implementación de estrategias basadas en pangenomas de referencia, la vigilancia genómica en tiempo real y la colaboración académica regional se posicionan como los pilares fundamentales para garantizar que los avances en biotecnología se traduzcan en diagnósticos equitativos y tratamientos personalizados adaptados a la singular diversidad genética de Latinoamérica.

FINANCIACIÓN

Los autores no recibieron financiación para el desarrollo de la presente investigación.

CONFLICTO DE INTERESES

Los Autores declaran que no existe conflicto de intereses

CONTRIBUCIÓN DE AUTORÍA

En concordancia con la taxonomía establecida internacionalmente para la asignación de créditos a autores de artículos científicos (<https://credit.niso.org/>). Los autores declaran sus contribuciones en la siguiente matriz:

	Autor 1	Autor 2	
Participar activamente en:			
Conceptualización	X	X	
Análisis formal	X	X	
Adquisición de fondos	X	X	
Investigación	X	X	
Metodología	X	X	

Administración del proyecto	X	X	
Recursos	X	X	
Redacción –borrador original	X	X	
Redacción –revisión y edición	X	X	
La discusión de los resultados	X	X	
Revisión y aprobación de la versión final del trabajo.	X	X	

REFERENCIAS

1. Salvo M, González-Feliú E, Toro J, Gallegos I, Maureira I, Miranda-González N, et al. Validation of an NGS Panel Designed for Detection of Actionable Mutations in Tumors Common in Latin America. *JPM*. 2021; 11: p. 899.
2. Angulo-Aguado M, Carrillo-Martinez JC, Contreras-Bravo NC, Morel A, Parra-Abaunza K, Usaquén W, et al. Next-generation sequencing of host genetics risk factors associated with COVID-19 severity and long-COVID in Colombian population. *Sci Rep*. 2024; 14: p. 8497.
3. Fernández-Gajardo R, Urra H, Sáez M, Pihán P, Fonseca B, Sanchez-Doñas C, et al. Characterization of the mutational status of glioblastoma and high-grade astrocytomas in a Latin American cohort. *Sci Rep*. 2025; 15: p. 34485.
4. Sohail M, Palma-Martínez MJ, Chong AY, Quinto-Cortés CD, Barberena-Jonas C, Medina-Muñoz SG, et al. Mexican Biobank advances population and medical genomics of diverse ancestries. *Nature*. 2023; 622: p. 775–783.
5. Jiménez-Kaufmann A, Chong AY, Cortés A, Quinto-Cortés CD, Fernandez-Valverde SL, Ferreyra-Reyes L, et al. Imputation Performance in Latin American Populations: Improving Rare Variants Representation With the Inclusion of Native American Genomes. *Front. Genet*. 2022; 12: p. 719791.
6. Paixão D, Torrezan GT, Santiago KM, Formiga MN, Ahuno ST, Dias-Neto E, et al. Characterization of genetic predisposition to molecular subtypes of breast cancer in Brazilian patients. *Front. Oncol*. 2022; 12: p. 976959.
7. Rivas S, Sepúlveda RV, Tapia I, Estay C, Soto V, Blanco A, et al. MET Exon 14 Skipping and Novel Actionable Variants: Diagnostic and Therapeutic Implications in Latin American Non-Small-Cell Lung Cancer Patients. *IJMS*. 2024; 25: p. 13715.
8. Leitão LPC, De Carvalho DC, Rodrigues JCG, Fernandes MR, Wanderley AV, Vinagre LWMS, et al. Identification of Genomic Variants Associated with the Risk of Acute Lymphoblastic Leukemia in Native Americans from Brazilian Amazonia. *JPM*. 2022; 12: p. 856.

9. Jaramillo-Valverde L, Horton MK, Poterico JA, Lanata CM, Guio H. Pharmacogenetic Study of Anti- TB Drugs in the Native Ancestry Peruvian Population. *Pharmacology Res & Perspec.* 2025; 13: p. e70135.
10. Cadena-Ullauri S, Ruiz-Pozo VA, Paz-Cruz E, Tamayo-Trujillo R, Guevara-Ramírez P, Jaramillo-Calvas O, et al. Evaluating Liquid Biopsy for Circulating Tumor DNA (ctDNA) Detection as a Complementary Diagnostic Tool in Thyroid Cancer Among Ecuadorian Women. *IJMS.* 2025; 26: p. 6987.
11. Molina-Mora JA, Reales-González J, Camacho E, Duarte-Martínez F, Tsukayama P, Soto-Garita C, et al. Overview of the SARS-CoV-2 genotypes circulating in Latin America during 2021. *Front. Public Health.* 2023; 11: p. 1095202.
12. Mena MD, Moresco AA, Vidal SH, Aguilar-Cortes D, Obregon MG, Fandiño AC, et al. Clinical and Genetic Spectrum of Stargardt Disease in Argentinean Patients. *Front. Genet.* 2021; 12: p. 646058.
13. Abreu GDM, Tarantino RM, Da Fonseca ACP, Andrade JRFDO, De Souza RB, Soares CDAPD, et al. Identification of Variants Responsible for Monogenic Forms of Diabetes in Brazil. *Front. Endocrinol.* 2022; 13: p. 827325.
14. Valverde-Hernández JC, Flores-Cruz A, Chavarría-Soley G, Silva De La Fuente S, Campos-Sánchez R. Frequencies of variants in genes associated with dyslipidemias identified in Costa Rican genomes. *Front. Genet.* 2023; 14: p. 1114774.
15. Rhead B, Hein DM, Pouliot Y, Guinney J, De La Vega FM, Sanford NN. Association of genetic ancestry with molecular tumor profiles in colorectal cancer. *Genome Med.* 2024; 16: p. 99.
16. Gionda PO, Gomes-Gouvea M, Malta FDM, Sebe P, Salles APM, Francisco RDS, et al. Analysis of the complete genome of HBV genotypes F and H found in Brazil and Mexico using the next generation sequencing method. *Annals of Hepatology.* 2022; 27: p. 100569.
17. Cecchini D, Sfalcin J, Zapiola I, Gomez A, Fernandez-Giuliano S, Rodriguez C, et al. Integrase strand transfer inhibitors resistance-associated mutations in HIV-infected pregnant women. *Rev Esp Quimioter.* 2024; 37: p. 479–485.
18. Chacon-Camacho OF, Xilotl-de Jesús N, Calderón-Martínez E, Ordoñez-Labastida V, Neria-Gonzalez MI, Villafuerte-de La Cruz R, et al. Genotypic spectrum of ABCA4-associated retinal degenerations in 211 unrelated Mexican patients: identification of 22 novel disease-causing variants. *Mol Genet Genomics.* 2024; 299: p. 79.

19. Silgado-Guzmán DF, Angulo-Aguado M, Morel A, Niño-Orrego MJ, Ruiz-Torres DA, Contreras Bravo NC, et al. Characterization of ADME Gene Variation in Colombian Population by Exome Sequencing. *Front. Pharmacol.* 2022; 13: p. 931531.
20. Cadena-Ullauri S, Guevara-Ramírez P, Ruiz-Pozo VA, Tamayo-Trujillo R, Paz-Cruz E, Becerra-Fernández M, et al. Population Admixture and APOB Variant Landscape in Ecuadorian Mestizo Patients with Cardiac Diseases: Potential Implications for Familial Hypercholesterolemia Genetics. *JCDD.* 2026; 13: p. 36.
21. De Oliveira Cavagna R, Escremim De Paula F, Berardinelli GN, Bonatelli M, Santana I, Reis MT, et al. Clinical and molecular characterization of a large Brazilian lung cancer cohort: a real-world observational study. *The Lancet Regional Health - Americas.* 2026; 56: p. 101429.
22. Silva NSB, Bourguiba-Hachemi S, Ciriaco VAO, Knorst SHY, Carmo RT, Masotti C, et al. A multi-ethnic reference panel to impute classical and non-classical class I alleles in admixed samples: Testing imputation accuracy in an admixed sample from Brazil. *HLA.* 2024; 103: p. e15543.
23. Cancela F, Ramos N, Smyth DS, Etchebehere C, Berois M, Rodríguez J, et al. Wastewater surveillance of SARS-CoV-2 genomic populations on a country-wide scale through targeted sequencing. *PLoS ONE.* 2023; 18: p. e0284483.
24. Cadena-Ullauri S, Guevara-Ramirez P, Ruiz-Pozo V, Tamayo-Trujillo R, Paz-Cruz E, Sánchez Insuasty T, et al. Case report: Genomic screening for inherited cardiac conditions in Ecuadorian mestizo relatives: Improving familial diagnose. *Front. Cardiovasc. Med.* 2022; 9: p. 1037370.
25. Douillard V, Dos Santos Brito Silva N, Bourguiba-Hachemi S, Naslavsky MS, Seliar MO, Duarte YAO, et al. Optimal population-specific HLA imputation with dimension reduction. *HLA.* 2024; 103: p. e15282.
26. Montella T, Zalis M, Zukin M, Cordeiro De Lima VC, Baldotto C, De Marchi P, et al. EGFR Mutation Detection in Brazilian Patients With Non-Small-Cell Lung Cancer: Lessons From Real-World Data Scenario of Molecular Testing. *JCO Global Oncology.* 2023;: p. e2200426.
27. Nóbrega PR, Paiva ARB, Amorim Junior AD, Lima PLGSB, Cabral KSS, Barcelos IP, et al. Further description of the phenotypic spectrum of neuronal ceroid lipofuscinosis type 11. *Genetics in Medicine.* 2025; 27: p. 101291.

28. Luce L, Carcione M, Mazzanti C, Buonfiglio PI, Dalamón V, Mesa L, et al. Theragnosis for Duchenne Muscular Dystrophy. *Front. Pharmacol.* 2021; 12: p. 648390.
29. Sohail M, Moreno-Estrada A. The Mexican Biobank Project promotes genetic discovery, inclusive science and local capacity building. *Disease Models & Mechanisms.* 2024; 17: p. dmm050522.
30. Cadena-Ullauri S, Guevara-Ramírez P, Ruiz-Pozo VA, Tamayo-Trujillo R, Paz-Cruz E, Simancas-Racines D, et al. Genomic analysis of an Ecuadorian individual carrying an SCN5A rare variant. *BMC Cardiovasc Disord.* 2024; 24: p. 388.
31. Giacomazzi J, Ziegelmann PK, Da Costa S, Bittar CM, Obst FM, Rosset C, et al. Cancer Risk Factors in Southern Brazil: Report of a Comprehensive, Matched Case-Control Study. *JCO Glob Oncol.* 2023;: p. e2300006.
32. Brito F, Lagos C, Cubillos J, Orellana J, Gajardo M, Böhme D, et al. Genomic analysis in Chilean patients with suspected Rett syndrome: keep a broad differential diagnosis. *Front. Genet.* 2024; 15: p. 1278198.
33. Romer G, Bracco LA, Ricci AD, Balouz V, Berná L, Villar JC, et al. Deep serological profiling of the *Trypanosoma cruzi* TSSA antigen reveals different epitopes and modes of recognition by Chagas disease patients. *PLoS Negl Trop Dis.* 2023; 17: p. e0011542.
34. Alvarado-Orosco N, Gomis-Font MA, Sastre-Femenia MA, Ribera Ortiz S, Miranda Arce M, Araúz Barba ME, et al. First nationwide survey on in Bolivia: susceptibility profiles, resistome, and genomic epidemiology. *Antimicrob Agents Chemother.* 2025; 69: p. e01163–25.
35. Cuadros-Espinoza S, Laval G, Quintana-Murci L, Patin E. The genomic signatures of natural selection in admixed human populations. *The American Journal of Human Genetics.* 2022; 109: p. 710–726.
36. Guevara-Ramírez P, Cadena-Ullauri S, Tamayo-Trujillo R, Ruiz-Pozo VA, Paz-Cruz E, Ibarra-Castillo R, et al. Spectrum and Clinical Interpretation of TTN Variants in Ecuadorian Patients with Heart Disease: Insights into VUS and Likely Pathogenic Variants. *IJMS.* 2025; 26: p. 11896.
37. Aguilar-Zamora E, Weimer BC, Torres RC, Gómez-Delgado A, Ortiz-Olvera N, Aparicio-Ozores G, et al. Molecular Epidemiology and Antimicrobial Resistance of *Clostridioides difficile* in Hospitalized Patients From Mexico. *Front. Microbiol.* 2022; 12: p. 787451.

38. Jaramillo-Valverde L, Levano KS, Tarazona DD, Capristano S, Sanchez C, Poterico JA, et al. Pharmacogenetic variability of tuberculosis biomarkers in native and mestizo Peruvian populations. *Pharmacology Res & Perspec.* 2024; 12: p. e1179.
39. Molina-Mora JA, Cordero-Laurent E, Godínez A, Calderón-Osorno M, Brenes H, Soto-Garita C, et al. SARS-CoV-2 genomic surveillance in Costa Rica: Evidence of a divergent population and an increased detection of a spike T1117I mutation. *Infection, Genetics and Evolution.* 2021; 92: p. 104872.
40. Mueller M, Blandino A, Scherer D, Zulantay I, Apt W, Varela NM, et al. Small-RNA sequencing identifies serum microRNAs associated with abnormal electrocardiography findings in patients with Chagas disease. *Journal of Infection.* 2025; 91: p. 106613.
41. Zuccherato LW, Souza RP, Camelo RM, Santana MAP, Dias MM, Jardim LL, et al. Deleterious variants cluster in the A3 domain of factor VIII in people with severe hemophilia A and inhibitors. *Research and Practice in Thrombosis and Haemostasis.* 2025; 9: p. 103006.
42. Uricoechea Patiño D, Collins A, García OJR, Santos Vecino G, Cuenca JVR, Bernal JE, et al. High Mitochondrial Haplotype Diversity Found in Three Pre-Hispanic Groups from Colombia. *Genes.* 2023; 14: p. 1853.
43. Melendez-Montañez JM, De Jesus-Rojas W. The Tip of the Iceberg: Genotype of Puerto Rican Pediatric Obesity. *Genes.* 2024; 15: p. 394.
44. Barra-Matos G, Gouvea De Souza F, Santos Silva C, Epifane-de-Assunção M, Belém Souza MV, Cavaleiro De Macêdo LC, et al. Mitochondrial insertions/deletions (INDELs) burden in Parkinson's disease: an analysis from a Brazilian cohort. *Parkinsonism & Related Disorders.* 2026; 145: p. 108250.
45. Gómez-Flores-Ramos L, Barraza-Arellano AL, Mohar A, Trujillo-Martínez M, Grimaldo L, Ortiz-Lopez R, et al. Germline Variants in Cancer Genes from Young Breast Cancer Mexican Patients. *Cancers.* 2022; 14: p. 1647.
46. Galán-Huerta KA, Flores-Treviño S, Salas-Treviño D, Bocanegra-Ibarias P, Rivas-Estilla AM, Pérez-Alba E, et al. Prevalence of SARS-CoV-2 Variants of Concern and Variants of Interest in COVID-19 Breakthrough Infections in a Hospital in Monterrey, Mexico. *Viruses.* 2022; 14: p. 154.
47. Gondola J, Aguilar C, Castillo J, Gonzalez C, Moreno A, Chavarria O, et al. Genotypes and Mutations Associated with Macrolide Resistance of *Chlamydia trachomatis* in Urine Samples in School-Going, Urban Adolescents 14–19 Years

- from Panama: A Cross-Sectional Descriptive Study. *Microbiology Research*. 2024; 15: p. 2007–2014.
48. Martínez JRW, Alcalde-Rico M, Jara-Videla E, Reyes J, Carvajal LP, Rincon S, et al. Geographic divergence of methicillin-resistant *h* ST5-SCC I in the aftermath of a major earthquake and tsunami: impact of a plasmid harboring heavy metal resistance genes. *mBio*. 2025; 16: p. e03669–24.
49. Chen Z, Toro M, Moreno-Switt AI, Adell AD, Delgado-Suárez EJ, Bonelli RR, et al. Unveiling the genomic landscape of serotypes Typhimurium, Newport, and Infantis in Latin American surface waters: a comparative analysis. *Microbiol Spectr*. 2024; 12: p. e00047–24.
50. Siegel SV, Trimarsanto H, Amato R, Murie K, Taylor AR, Sutanto E, et al. Lineage-informative microhaplotypes for recurrence classification and spatio-temporal surveillance of *Plasmodium vivax* malaria parasites. *Nat Commun*. 2024; 15: p. 6757.
51. Marín JEO, Villatoro E, Luna MJ, Barrientos AM, Mendoza E, Lemos APS, et al. Emergence of MDR invasive in El Salvador, 2017–19. *Journal of Antimicrobial Chemotherapy*. 2021; 76: p. 1155–1159.
52. Spangenberg L, Fariello MI, Arce D, Illanes G, Greif G, Shin JY, et al. Indigenous Ancestry and Admixture in the Uruguayan Population. *Front. Genet*. 2021; 12: p. 733195.
53. Chen D, Storey JD. Coancestry superposed on admixed populations yields measures of relatedness at individual-level resolution. *PLoS Comput Biol*. 2025; 21: p. e1013848.
54. Zambrano-O YT, Mejía-García A, Morales PD, Tsao HM, Rey-Vargas L, Montero-Ovalle W, et al. Inference of genetic ancestry from a multi-gene cancer panel in Colombian women with cancer. *Breast Cancer Res Treat*. 2025; 210: p. 251–259.
55. Bispo AG, Souza FGD, Epifane-de-Assunção M, Sena-dos-Santos C, Barra-Matos G, Moura DD, et al. Molecular analysis of mitochondrial complex I in the levodopa short duration response in Parkinson's disease. *Parkinsonism & Related Disorders*. 2026; 145: p. 108252.
56. Deiana M, Malagò S, Mori A, Accordini S, Matucci A, Passarelli Mantovani R, et al. Full Genome Characterization of the First Oropouche Virus Isolate Imported in Europe from Cuba. *Viruses*. 2024; 16: p. 1586.
57. Chen Z, Moreno-Switt AI, Reyes-Jara A, Delgado Suarez E, Adell AD, Oliveira CJB, et al. A multicenter genomic epidemiological investigation in Brazil, Chile, and

Mexico reveals the diversity and persistence of populations in surface waters. *mBio*. 2024; 15: p. e00777–24.

58. Nøhr AK, Hanghøj K, Garcia-Erill G, Li Z, Moltke I, Albrechtsen A. NGSremix: a software tool for estimating pairwise relatedness between admixed individuals from next-generation sequencing data. *G3 Genes|Genomes|Genetics*. 2021; 11: p. jkab174.
59. Fu H, Shi G. Local Ancestry Inference Based on Population-Specific Single-Nucleotide Polymorphisms—A Study of Admixed Populations in the 1000 Genomes Project. *Genes*. 2024; 15: p. 1099.
60. Honorato-Mauer J, Shah NN, Maihofer AX, Zai CC, Belangero S, Nievergelt CM, et al. Characterizing features affecting local ancestry inference performance in admixed populations. *The American Journal of Human Genetics*. 2025; 112: p. 224–234.
61. Esteves B, Melo-Braga M, Gorshkov V, Verano-Braga T, Larsen M, Gontijo C, et al. Characterization of Differentially Abundant Proteins Among *Leishmania* (*Viannia*) *braziliensis* Strains Isolated From Atypical or Typical Lesions. *Frontiers in Cellular and Infection Microbiology*. 2022; 12: p. 824968.