

## Bioinformatics in Biomedical Analysis and its impact on Personalized Medicine

### El Papel de la Bioinformática en el Análisis Biomédico y su Impacto en la Medicina Personalizada

Marcelo D. Gamarra<sup>1,\*</sup>; Belén M. Brizuela<sup>1</sup>

1- Instituto de Genética Humana de Misiones (IGeHM). Parque de la Salud de la provincia de Misiones. Misiones, Argentina.

\* E-mail: mdgamarraok@gmail.com

Received: 17/10/2023; Accepted: 10/11/2023

#### Abstract

Biomedical analysis represents a crucial process for understanding and diagnosing various diseases, including those of genetic origin. In this context, bioinformatics plays a fundamental role in facilitating the interpretation of genomic data and providing essential information that can guide diagnosis and clinical decisions. Bioinformatics is an area that promotes the use of computers to address questions of biological nature, tackling the massive analysis of available biological data. On the other hand, one of the most revolutionary techniques in the clinical field is Next-Generation Sequencing (NGS), which allows for the rapid and precise reading of the DNA of any patient suspected of having a genetic condition. With NGS, it is possible to identify mutations, genomic variants, and other changes that may be related to medical conditions. Other diagnostic techniques, such as Multiplex Ligation-dependent Probe Amplification (MLPA) and DNA arrays, also excel in detecting genetic changes. MLPA detects changes in the number of copies of DNA segments (known as CNVs), while arrays analyze multiple markers simultaneously, identifying CNVs and point mutations. When used together, these techniques generate a significant amount of information that can be crucial for approaching and understanding a pathology. Thus, bioinformatics is present in all stages of biomedical analysis. It begins with the alignment of sequencing reads against the Human Reference Genome using various software tools like *BWA* or *Bowtie*. Reads undergo variant filtering, which identifies discrepancies between the patient's genome and the reference genome. In this phase, tools like *GATK* and *SAMtools* are essential. Beyond all intermediate steps, variant interpretation is the decisive stage of the entire process. In this phase, factors such as heritability, penetrance, clinical impact, and population frequency are evaluated, determining whether variants are pathogenic, benign, or of uncertain significance, according to various criteria like those established by the *American College of Medical Genetics and Genomics (ACMG)*. In some cases, information from the structural biology of each variant is used, through simulations and calculations of structural parameters such as aggregability, switchability, protein-protein or protein-DNA interactions, and energy changes, to complement this assessment. For all these sequence and structural analyses, there is a range of freely available software and databases, such as InterPro, dbSNP, ClinVar, Uniprot, Ensembl, and Varsome. It is evident that the integration of bioinformatics into any clinical research or healthcare group is essential to comprehend and effectively utilize these mentioned tools. This fact is reflected in clinical practice, where this integration has a significant impact on the quality of life of patients and their families, enabling more effective medical interventions and better-informed long-term care planning. Moreover, this integration fosters the development of personalized therapies, promising to revolutionize healthcare in the future.

Keywords: Genetic Counseling, Bioinformatics, Biomedical Analysis, Personalized Medicine.

#### Resumen

El análisis biomédico representa un proceso crucial para comprender y diagnosticar diversas enfermedades, incluyendo las de origen genético. En este contexto, la bioinformática desempeña un papel fundamental al facilitar la interpretación de datos genómicos y proporcionar información esencial que puede orientar el diagnóstico y las decisiones clínicas. La bioinformática es un área que fomenta el uso de la computadora para responder preguntas de naturaleza biológica, abordando el análisis masivo de datos biológicos disponibles. Por su parte, una de las técnicas más revolucionarias en el ámbito clínico es la Secuenciación de Nueva Generación (NGS, por sus siglas en

inglés), que permite la rápida y precisa lectura del ADN de cualquier paciente con sospecha de una enfermedad de base genética. Con ella, es posible identificar mutaciones, variantes genómicas y otros cambios que pueden estar relacionados con condiciones médicas. Otras técnicas diagnósticas, como el *Multiplex Ligation-dependent Probe Amplification* (MLPA) y los *Array* de ADN, también se destacan en la detección de cambios genéticos. El MLPA detecta cambios en el número de copias de segmentos de ADN (conocidas como CNVs), mientras que los *Arrays* analizan múltiples marcadores simultáneamente, identificando CNVs y mutaciones puntuales. Estas técnicas, cuando se utilizan en conjunto, generan una cantidad significativa de información que puede ser fundamental para el abordaje y comprensión de una patología. Así, la bioinformática está presente en todas las etapas del análisis biomédico. Comienza con el alineamiento de lecturas de secuenciación contra el *Genoma de Referencia Humano* utilizando diferentes softwares como *BWA* o *Bowtie*. Las lecturas se someten a un filtrado de variantes que identifica discrepancias entre el genoma del paciente y el de referencia. En esta fase, herramientas como *GATK* y *SAMtools* son esenciales. Más allá de todos los pasos intermedios, la interpretación de variante es la etapa decisiva de todo el proceso. En esta fase, se evalúan factores como heredabilidad, penetrancia, impacto clínico y frecuencia poblacional, determinando si las variantes son patogénicas, benignas o de significado incierto, según diversos criterios como los establecidos por el *American College of Medical Genetics and Genomics* (ACMG). En algunos casos, se recurre a información de la biología estructural de cada variante, mediante simulaciones y cálculos de parámetros estructurales como la agregabilidad, la *switchabilidad*, las interacciones proteína-proteína o proteína-ADN y cambios energéticos para complementar esta evaluación. Para todos estos análisis, tanto de secuencia como estructurales, existen un abanico de softwares o bases de datos disponibles de manera libre, como *InterPro*, *dbSNP*, *ClinVar*, *Uniprot*, *Ensembl* y *Varsome*. Es evidente que la integración de la bioinformática en cualquier grupo de investigación clínica o atención médica es esencial para comprender y utilizar efectivamente estas herramientas. Este hecho se refleja en la clínica, donde esta incorporación tiene un impacto significativo en la calidad de vida de los pacientes y sus familias, permitiendo intervenciones médicas más efectivas y una planificación del cuidado a largo plazo mejor informada. Además, esta integración fomenta el desarrollo de terapias personalizadas, lo que promete revolucionar la atención médica en el futuro.

Palabras clave: Genómica clínica, Bioinformática, Análisis biomédico, Medicina personalizada.